Wnt and FGF Signaling in *C. elegans* Vulval Cell Lineage Polarity

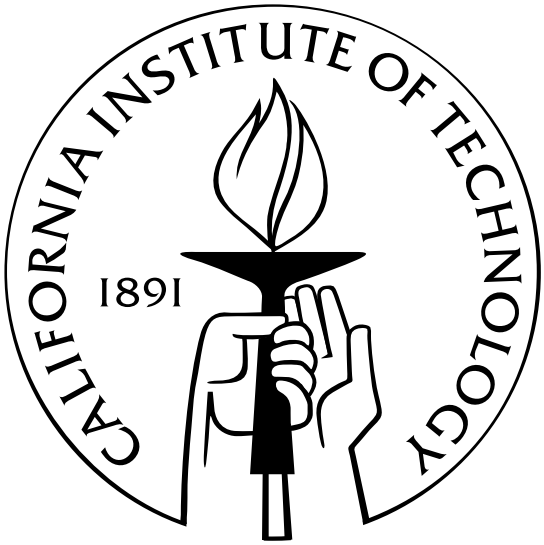
Thesis by

Paul Joseph Minor

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2014

Defended August 1, 2013

To my parents, Greg and Rita Minor, for instilling in me a lifelong love of learning and inciting an incessant curiosity for the world around me.

"An investment in knowledge pays the best interest."

-Benjamin Franklin

"Science is the poetry of reality."

-Richard Dawkins

ACKNOWLEDGMENTS

Attempting to list all of the individuals that have influenced me in my first 30 years of life is an impossible task. With that in mind, I can honestly say the last seven years I have spent at Caltech have been some of the most exciting, intellectually stimulating, humbling, and unforgettable years of my life. I have been fortunate enough to work with some of the best and brightest scientists as well as meet some of my closest friends. Without either of these two groups of people, I would have never finished my doctorate. The completion of this thesis can be attributed as much to my own work as to their assistance and I will forever be thankful for their support.

First and foremost I would like to thank my adviser, Paul Sternberg. There are no words I could write that would accurately convey the amount of gratitude and appreciation I have for Paul. I cannot imagine working for anyone else during graduate school. I came to Caltech with an engineering background, but I knew I wanted to follow my heart and become a biologist. Paul's enthusiasm for biology and learning about the world around us is contagious. It is impossible to not be excited about science after talking with him. His office door was always open to me regardless of what I wanted to talk to him about. I would not be the scientist I am today without his guidance. It is a rare opportunity to be able to work for someone who is one of the best and brightest minds in the field. It is even more rare when that person happens to also be one of the nicest, most understanding people you have ever met. I have been lucky enough to have a boss that fits that description. As I move forward in life and take the next step in the path to become a successful biologist, I will never forget my time in the Sternberg lab. The

experience has shaped me both as a scientist and as a person. Under the guidance of Paul I have become the biologist I hoped I would when I enrolled at Caltech.

I would also like to thank all of my friends and colleagues in the Sternberg lab. Adler Dillman has been a great friend and scientific inspiration. Without Adler I would not be the person I am today. Adler's curiosity for the world around him is infectious. So much of my love of biology and evolution come from the discussions that took place over the many lunches and afternoon coffee and soda breaks we took together. I will forever be grateful that our paths happened to cross in the Sternberg lab. I must also thank all of the people in Kerckhoff 217 that I have been lucky enough to work with over the years, including Mihoko Kato, Adeline Seah, Hillel Schwartz, Elly Chow, and Cheryl van Buskirk. I would like to thank each of them for helpful discussions and advice. It has been great working in the same room as them. Having seen each of them almost every day over the last few years, I do not just see them as colleagues, but also as great friends. I would also like to thank the following individuals from the Sternberg lab: Jennifer Green and Takao Inoue for laying the groundwork to my project, helping me out when I first joined the lab, and for their continued support and assistance; Amir Sapir for always being there to answer my questions and to listen to my ideas no matter how out of left field they may be; John DeModena for his countless hours of entertainment and stories; Jagan Srinivasin for helpful scientific advice, friendship, and always making my days a little more pleasant; Chris Cronin for every amusing story and joke that helped me get through the day; Andrea Choe for making lab much more exciting and enjoyable; Oren Schaedel for being a great labmate and neighbor; Ted Ririe for helping me get acclimated to the Sternberg lab; and Chris Grove who has become a good friend over the last few

years and has opened my eyes to how I think about science and how to convey those ideas to others. There are many others from lab that I would like to thank, including Margaret Ho, James Lee, Srimoyee Ghosh, Pei Shih, Gladys Medina, Barbara Perry, Ryoji Shinya, Daniel Leighton, Alli Akagi, Ping Hseuh, Ryan Baugh, Elissa Hallem, and Vivian Chiu.

Along the way I have interacted with so many graduate students and postdocs during my time at Caltech as well as others from outside campus and would like to thank them all for their friendship, encouragement, and advice. I would like to thank the following individuals: Jade Takahashi for all of the love and support, for always being there for me, and for being the first person I can go to no matter what the circumstances are; I could not have done it without her; Peter Weir for being a great friend and roommate, for always being there through the ups and downs of graduate school, and for countless hours spent discussing science, baseball, or any other random aspect of our lives; Faisal Amlani for his friendship, introducing me to a plethora of new music, and being there to help me keep my sanity throughout the years at Caltech; Evans and Susanna Boney for the many hours we spent sitting around talking about any and all topics under the sun and for being an endless source of entertainment; Trevor Johnson and Marie Suver for being amazing friends while in Pasadena and even after moving a thousand miles away; Andrew English for being one of the best friends anyone could ask for; Scott Kelber for all the lunches and conversations that could make me laugh no matter how down I was; Chris Rogan for showing me it is possible to be a great scientist while living the good life; John Ngo for always being able to cheer me up no matter how stressed I was about science and lab work; I will always appreciate his friendship. I would also like to thank the countless

others that at one time or another have befriended me, offered guidance and a listening ear, or have in some way helped me get to where I am today, including: Jon Valencia, Brian Duistermars, Eric Erkenbrack, Ronnie Bryan, Michael Rome, Stephen Chapman, Geoffrey Smith, Geoffrey Lovely, Alex Romero, Lauren LeBon, Allan Wong, and Anand Asthagiri.

I also owe thanks to the members of my thesis committee, Marianne Bronner, Long Cai, Bruce Hay, and Paul Sternberg for taking time out of their busy schedules to meet with me, read my work and offer constructive criticism. I also offer thanks and appreciation to the Howard Hughes Medical Institute and the National Institutes of Health for funding the majority of the research in the Sternberg lab. Thanks is also due to Ting-Fang He, Chang Ho Sohn, Andrew Davenport, and André Hoelz. Collaborating with them was an excellent experience and helped open my eyes to other branches of biology.

Finally, my family has also been extremely supportive since day one. Without them this work would have never been possible. For as long as I can remember my family has encouraged my education and higher learning even when that meant moving 2000 miles away from them. I am very grateful to always have them in my corner and thankful for their love, support, and encouragement. I would not be the person I am today without them.

PREFACE

For as long as I can remember I have been fascinated by the complexity of life. As a child I would obsess over the diversity found within nature, an obsession that has remained strong to this day. Growing up in Alabama, surrounded by forests, mountains, rivers, and wildlife, fostered my curiosity, and this childlike wonder soon turned into questions of "how" and "why." One of the aspects of nature that has always amazed me was the beautiful patterns and symmetry found across all Metazoa. As a child I observed these patterns with amazement; as an adult I have the opportunity to study the genes and underlying mechanisms that control the patterns found not only within nematodes but the entire animal kingdom. As I finish my graduate school career and prepare for postdoctoral studies, I realize I am still very much interested in the "how" and "why" questions that perplexed me as a child, only now I also have the capabilities of answering them.

Marie Curie once said, "A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him as though they were fairy tales." I could not agree more.

ABSTRACT

The interpretation of extracellular cues leading to the polarization of intracellular components and asymmetric cell divisions is a fundamental part of metazoan organogenesis. The *C. elegans* vulva, with its invariant cell lineage and interaction of multiple cell signaling pathways, provides an excellent model for the study of cell polarity within an organized epithelial tissue. Herein I discuss the interaction of Wnt and FGF signaling in controlling vulval cell lineage polarity with emphasis on the posterior-most cell that forms the vulva, P7.p.

The mirror symmetry of the *C. elegans* vulva is achieved by the opposite division orientation of the vulval precursor cells (VPCs) flanking the axis of symmetry. Opposing Wnt signals control the division patterns of the VPCs by controlling the localization of SYS-1/ β-catenin toward the direction of the Wnt gradient. Multiple Wnt signals, expressed at the axis of symmetry, promote the wild-type, anterior-facing, P7.p orientation, whereas Wnts EGL-20 and CWN-1 from the tail and posterior body wall muscle, respectively, promote the daughter cells of P7.p to face the posterior. EGL-20 acts through a member of the LDL receptor superfamily, LRP-2, along with Ror/CAM-1 and Van Gogh/VANG-1. All three transmembrane proteins control orientation through the localization of the SYS-1.

The Fibroblast Growth Factor (FGF) pathway acts in concert with LIN-17/Frizzled to regulate the localization of SYS-1. The source of the FGF ligand is the 1° VPC, P6.p, which controls the polarity of the neighboring 2° VPC, P7.p, by signaling through the sex myoblasts (SMs), activating the FGF pathway. The Wnt, *cwn-1*, is expressed in the posterior body wall muscle of the worm as well as the SMs, making it the only Wnt

expressed on the posterior and anterior sides of P7.p at the time of the polarity decision. Both sources of *cwn-1* act instructively to influence P7.p polarity in the direction of the Wnt gradient. The FGF pathway leads to the regulation of *cwn-1* transcripts in the SMs. These results illustrate the first evidence of the interaction between FGF and Wnt in *C. elegans* development and vulval cell lineage polarity as well as highlight the promiscuous nature of Wnt signaling within *C. elegans*.

TABLE OF CONTENTS

Chapter 1

Introduction

Thesis Overview

Proper tissue architecture and organogenesis are fundamental aspects of Metazoan development. The arrangement of cells into functional structures is achieved through cell division patterning and orientation, resulting from the polarization of intracellular components. Loss of cell polarity and asymmetry is a major factor in tumor formation, and growing evidence illustrates its importance in understanding human cancer (Wodarz and Nathke, 2007). Proper orientation is often achieved through communication between cells in the form of intracellular signaling pathways. The action of these pathways is initiated when an extracellular ligand binds with its transmembrane receptor, triggering a biochemical response within the cell and relaying this message through the downstream components of the pathway leading to the polarization of cellular components or the transcriptional regulation of genes. Two such pathways involved in Metazoan development are the Wnt and Fibroblast Growth Facter (FGF) pathways.

Wnt signaling is implicated in many aspects of development, including cell proliferation, migration, polarity, terminal differentiation, and the self-renewal of stem cells (Boutros and Mlodzik, 1999; Gao and Chen, 2010; Sugimura et al., 2012; Tauriello et al., 2012), and deregulation of pathway components is associated with multiple human diseases (Luo et al., 2007). Wnt signaling has evolved to function in multiple pathways, broadly divided into the canonical/β-catenin-dependent pathway and the noncanonical/β-catenin-independent pathway, of which the planar cell polarity (PCP) pathway is the most studied (Gao and Chen, 2010). In the canonical pathway, Wnt ligands are received by a Frizzled and LRP5/6 (Arrow) coreceptor complex. The binding of the Wnt ligand recruits Dishevelled and Axin to the cell membrane and results in the inactivation of the β-catenin

destruction complex. With the inactivation and dissolution of this destruction complex, β-catenin is able to translocate within the nucleus and interact with the transcription factor TCF/Lef to regulate gene expression (Wharton, 2003). Noncanonical Wnt signaling encompasses several different signaling pathways, of which the most studied and, therefore, most understood is the PCP pathway. PCP signaling leads to the polarization of cells along an epithelial sheet. A wide range of components exist within PCP signaling, including the transmembrane proteins Frizzled, Flamingo, and Strabismus (Van Gogh), as well as the cytoplasmic proteins Diego, Prickle, and Dishevelled (Seifert and Mlodzik, 2007). Interactions between these core components leads to the asymmetric enrichment and distribution within the cell, which in turn enables the polarization of cells along the epithelial sheet.

Much like Wnt signaling, FGF signaling is also involved in a wide range of developmental functions, including mesodermal patterning in the early embryo, cellular proliferation, survival, migration, and differentiation (Turner and Grose, 2010). FGF ligands are secreted glycoproteins that bind with a receptor tyrosine kinase leading to ligand-dependent dimerization and activation of the kinase domain on the receptor. Once tyrosine residues on the receptor are phosphorylated, they serve as docking sites for adaptor proteins, which may also be phosphorylated by the FGF receptor. Activation of the FGF pathway leads to the activation of Raf and Ras proteins resulting in the downstream activation of the Mitogen-activated protein (MAP) kinase pathway. Both Wnt and FGF pathways are conserved in *C. elegans*, and play a role in the patterning of the vulva, a classic model system used to study cell-signaling pathways.

In chapter 2, I discuss the role of the low-density lipoprotein receptor, *lrp-2*, and its

role in vulval patterning. Nematodes, including *C. elegans,* do not have a copy of LRP5/6 (Arrow) in their genome. LRP5/6 plays an important role in canonical Wnt signaling in higher order organisms, including *Drosophila*, *Xenopus*, mouse, and humans. How then does Wnt signaling compensate for the absence of LRP5/6 in *C. elegans*? Furthermore, it is often hypothesized that LRP5/6 evolved from larger low-density lipoprotein receptors, LRP1 and LRP2 (megalin), due to the presence of all LRP5/6 protein domains within LRP1 and LRP2 and the presence of these receptors in lower order organisms that lack LRP5/6. Interestingly, *C. elegans* does have orthologs of these larger low-density lipoprotein receptors; *lrp-1* is an ortholog of megalin while *lrp-2* appears to be a recent duplication of *lrp-1* that has diverged to take on specialized function and different expression patterns from its paralog. The impetus to study *lrp-1* and *lrp-2* was to better understand these questions of Wnt signaling evolution and decipher a potential role for these genes within *C. elegans* Wnt signaling, a role that could possibly enhance our knowledge of the evolution of the function of low-density lipoprotein receptors. Within this thesis I show that *lrp-2* is downstream of the Wnt, *egl-20*, and functions with the *cam-1*/Ror and *vang-1*/Van Gogh transmembrane proteins to direct the localization of SYS-1/β-catenin during anaphase of the first vulval cell division, a role that shows conservation between the early low-density lipoprotein and their potential evolutionary derivative, LRP5/6.

Chapter 3 looks at the origin and evolution of dishevelled across Metazoa, with emphasis on nematodes. As the hub of multiple Wnt signaling pathways, dishevelled is a highly studied and medically relevant protein. Most work on dishevelled takes for granted the protein architecture and assumes that the conservation of protein domains and

motifs holds true across all of Metazoa. With this in mind, we found it interesting that most animals have multiple orthologs of dishevelled, and there are multiple studies that imply a functional specialization between different orthologs. To us, the functional specialization of orthologs indicated an underlying difference in protein architecture, something we wished to explore and better understand. We find evidence of dynamic evolution of dishevelled. We identify a new domain specific to some nematode lineages, the DEP-like fragment, and find an unexpected nuclear localization signal conserved in many dishevelled orthologs, presenting the potential of dishevelled acting as a transcription factor in some animal lineages. Our findings raise questions of protein evolution in general and provide clues as to how animals have dealt with the complex intricacies of having a protein, such as dishevelled, act as a central messenger hub connected to many different and vitally important pathways. Furthermore our work highlights the fact that the classic domain architecture of dishevelled does not hold true across Metazoa. Despite the majority of literature illustrating dishevelled as a highly conserved protein, we find that it is dynamically evolving across multiple animal lineages, especially within Nematoda. Our work also highlights many future experiments that could potentially be performed in different model systems to test how divergence in protein architecture leads to diverse function of paralagous proteins.

In chapter 4, I look at the interaction of Wnt and FGF signaling in vulval cell lineage polarity. How cells are capable of interpreting instructions from complicated cell signaling networks that often involve crosstalk between multiple signaling pathways and the relaying of this instructional signal between multiple cells of varying types is a highly studied and important topic in developmental biology. Two such pathways that have a

large role in animal development and also medical implications in humans are the Wnt and FGF pathways. Recent work has shown crosstalk between these two pathways in vertebrate model systems, a connection that has proven essential to proper animal development (Stulberg et al., 2012; Yardley and Garcia-Castro, 2012). Prior to this work, no interaction between Wnt and FGF signaling in *C. elegans* had been presented. With its invariable cell lineage, short generation time, ease of genetics and transgenics, and conservation of both Wnt and FGF pathways, *C. elegans* is a phenomenal system to better understand the relationship between Wnt and FGF signaling. Furthermore, the *C. elegans* vulva, which has been shown to express both signaling cascades, albeit independent of each other, during its development, is the ideal system to study the interaction. Here, I present evidence that FGF signaling is necessary for the regulation of *cwn-1*/Wnt in the migrating sex myoblasts. I show that despite receptor specificity, all Wnt ligands controlling vulval orientation have the same molecular output, directing the localization of SYS-1/β-catinin toward the predominant Wnt gradient at the time of anaphase during the first vulval cell division.

These findings present new insights into Wnt and FGF signaling both within *C. elegans* as well as other model systems. My work has added a new component, the migrating sex myoblasts, into the beautiful network of cells controlling vulval cell lineage polarity, shown the interaction between Wnt and FGF signaling within *C. elegans*, and brought us one step closer to understanding how Wnt signaling determines the orientation of the 22 cells that make up the vulva.

REFERENCES

Boutros, M. and Mlodzik, M. (1999). Dishevelled: at the crossroads of divergent intracellular signaling pathways. *Mechanisms of development* **83**, 27-37.
Gao, C. and Chen, Y. G. (2010). Dishevelled: The hub of Wnt signaling. *Cellular signalling* **22**, 717-727.
Luo, J., Chen, J., Chen, Z., Deng, I., Luo, X., Song, W. X., et al. (2007). Wnt signaling and human diseases: what are the therapeutic implications? *Lab. Invest.* **87**: 97–103.
Seifert, J.R. and Mlodzik, M. (2007). Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. *Nat. Rev. Genet.* **8**, 126-138.
Sugimura, R., He, X. C., Venkatraman, A., Arai, F., Box, A., Semerad, C., Haug, J. S., Peng, L., Zhong, X. B., Suda, T. et al. (2012). Noncanonical Wnt signaling maintains hematopoietic stem cells in the niche. *Cell* **150**, 351-365.
Stulberg, M. J., Lin, A., Zhao, H. and Holley, S. A. (2012). Crosstalk between Fgf and Wnt signaling in the zebrafish tailbud. *Developmental biology* **369**, 298-307.
Tauriello, D. V., Jordens, I., Kirchner, K., Slootstra, J. W., Kruitwagen, T., Bouwman, B. A., Noutsou, M., Rudiger, S. G., Schwamborn, K., Schambony, A. et al. (2012). Wnt/beta-catenin signaling requires interaction of the Dishevelled DEP domain and C terminus with a discontinuous motif in Frizzled. *Proc Natl Acad Sci U S A* **109**, E812-820.
Turner, N., Grose, R. (2010) Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer* **2**, 116-129.
Wharton, K. A., Jr. (2003). Runnin' with the Dvl: proteins that associate with Dsh/Dvl and their significance to Wnt signal transduction. *Dev Biol* **253**, 1-17.
Wodarz, A. and Nathke, I. (2007). Cell polarity in development and cancer. *Nature cell biology* **9**, 1016-1024.
Yardley, N. and Garcia-Castro, M. I. (2012). FGF signaling transforms non-neural ectoderm into neural crest. *Developmental biology* **372**, 166-177.

Chapter 2

The role of *lrp-2* in vulval cell lineage polarity

ABSTRACT

During *C.elegans* vulval development the anchor cell induces 3 of 6 multipotent vulval precursor cells (VPCs). The closest VPC generates a 1° lineage; the flanking VPCs, P5.p and P6.p, each generate a 2°, mirror symmetric, lineage. Two Wnt signals from the anchor cell promote the wild-type, anterior-facing, P7.p orientation, whereas the EGL-20/Wnt signal from the tail promotes the daughter cells of P7.p to face the posterior, antagonizing two parallel Wnt pathways consisting of Frizzled and Ryk and receptors necessary to direct the wild-type vulval orientation. We show EGL-20/Wnt acts through a member of the LDL receptor superfamily, LRP-2, along with CAM-1/Ror and VANG-1/Van Gogh. Using a promoter fusion construct, we examined the expression pattern of *lrp-2* and found expression beginning in the 2-cell stage of the vulva and lasting through the fourth larval stage (L4). All three transmembrane proteins control orientation through the localization of the β-catenin-like transcriptional coactivator, SYS-1. This work lays a solid foundation for the role of LRP-2 in *C. elegans* vulval formation and brings several testable hypotheses to light.

INTRODUCTION

The orientation of asymmetric cell divisions is an essential part of Metazoan development (Strutt, 2005). Loss of cell polarity and asymmetry is a major factor in tumor formation, and growing evidence illustrates its importance in understanding human cancer (Wodarz and Nathke, 2007). To ensure proper cell divisions and organ formation, high amounts of signaling redundancy and cell-cell interactions involving crosstalk between multiple signaling pathways are often incorporated to tightly regulate these processes. The *Caenorhabditis elegans* vulva provides a simple model to study this phenomenon due to the small number of cells, invariant cell lineage and developmental timing, and cell signaling mechanisms involved within vulval formation (reviewed by Sternberg, 2005; reviewed by Gupta et al., 2012). Here we examine the antagonism of competing Wnt pathways in the development of the *C. elegans* vulva.

The *C. elegans* vulva is formed from divisions of three VPCs: P5.p, P6.p, and P7.p – arranged along the anterior-posterior axis in the ventral epithelium (Sulston and Horvitz, 1977). During the L3 (third larval) stage, a combination of EGF, Notch, and Wnt signals instructs the VPCs to adopt fates corresponding to particular lineage patterns. P6.p adopts a 1° fate and undergoes three rounds of symmetric divisions that lead to eight cells that form the vulval lumen. P5.p and P7.p adopt the 2° fate, which leads to three rounds of asymmetric cell divisions forming seven cells that create the anterior and posterior sides of the vulva (Figure 1). The outermost progeny of P5.p and P7.p adhere to the epidermis whereas the innermost progeny join the descendants of P6.p in forming the vulval lumen. The descendants of P5.p and P7.p display mirror symmetry about the center of the vulva.

Previous analyses show the orientation of P5.p and P7.p descendants is determined by the interaction of multiple Wnt signals. In the absence of all Wnts, the VPCs display a randomized orientation, which is likely the default (Green et al., 2008; Figure 1). Two separate Wnts from the anchor cell, LIN-44 and MOM-2 acting through receptors LIN-17/Frizzled and LIN-18/Ryk, respectively, regulate P7.p orientation (Ferguson et al., 1987; Sternberg and Horvitz, 1988; Sawa et al., 1996; Inoue et al., 2004; Gleason et al., 2006). In the absence of these signals the orientation of the progeny of P7.p mimic those of P5.p and face toward the posterior of the worm, a phenotype referred to as posterior-reversed vulval lineage (P-Rvl; Figure 2). This posterior orientation is dependent on the instructive signal of EGL-20, a Wnt expressed in the tail acting through CAM-1/ROR and VANG-1/Van Gogh, and is referred to as "ground polarity." In response to the Wnt signals from the anchor cell, LIN-17 and LIN-18 orient P7.p to face the center. This reorientation is described as "refined polarity" and is the wild-type orientation (Green et al., 2008; Figure 1).

Here we examine the role of a low-density lipoprotein receptor, *lrp-2*, and its role in controlling the orientation of P7.p daughter cells. Our genetic data and expression analysis indicate *lrp-2* functions with *cam-1* and *vang-1* to antagonize the *lin-17* pathway. We find that *lrp-2* works downstream of *egl-20* and controls the localization of SYS-1 at the time of the first cell division of P7.p

MATERIALS AND METHODS

Strains and Genetics

C. elegans was handled as described previously (Brenner, 1974). All strains used are derivatives of *C. elegans* N2 Bristol strain. The alleles used are as follows. LGI: *lin-17(n671)*, *lrp-2(gk272)*, *lrp-2(gk292)*. LGII: *cam-1(gm122)*. LGIII: *qIs95[pSYS-1::VNS::SYS-1* with *pttx-3::dsRed]*. LGX: *lin-18(e620)*, *vang-1(ok1142)*.

Scoring Vulval Phenotypes

To determine the vulval phenotype as wild type or P-Rvl, animals were scored in the mid-L4 stage. Animals were classified as P-Rvl if the 1° and 2° VPCs were induced but separated by adherent cells (Katz et al., 1995). Only fully induced vulvae were scored.

Transgenics

To make the *Plrp-2* GFP construct 2.5 kb of the promoter region of *lrp-2* was amplified and then fused to GFP using the PCR fusion technique previously described (Hobert, 2002). The *Plrp-2::*GFP extrachromosomal array was generated by creating an injection mix consisting of 15 ng/µl *Plrp-2::*GFP, 40 ng/µl *unc-119(+)*, and 95 ng/µl DNA ladder and injecting the mix into *unc-119(ed4)* animals as described (Mello et al., 1991).

Anterior expression of *egl-20* was achieved by fusing the promoter region of *fos-1a* to the cDNA of *egl-20* and injecting the mix, that was created by Jennifer Green, into *lin-17(n671) lrp-2(gk272)* animals as described (Mello et al., 1991).

RESULTS

LRP-2 Functions in Ground Polarity

The regulation of vulval cell lineage polarity is controlled by Wnt signaling (Figure 1). Previously known components involved in the regulation of vulval cell lineage polarity include LIN-17, LIN-18, CAM-1, and VANG-1 (Inoue et al., 2004; Gleason et al., 2006; Green et al., 2008). A directed screen of known Wnt pathway components was performed to find additional genes involved in directing vulval orientation. A BLAST was run using other known Wnt receptors and it was determined that *C. elegans* does not contain a true ortholog of *Drosophila* LRP5/6 (Arrow) (He et al., 2004; Eisenmann, 2005), but does have multiple low-density lipoprotein receptors, including LRP-1 and LRP-2 (Figure 3). Like other low-density lipoprotein receptors, both LRP-1 and LRP-2 contain many LDLR Domain Class A and Class B repeats, EGF-like domains, and a transmembrane domain. However, having approximately three times as many amino acids, LRP-1 and LRP-2 are more similar to megalin than LRP5/6 (Yochem et al., 1999). The absence of LRP5/6 within *C. elegans* but presence in flies and all other higher order organisms suggests that the gene encoding LRP5/6 arose after nematodes, potentially from either LRP1 or LRP2/megalin, as both receptors contain the entire extracellular portion of LRP5/6 in a single contiguous sequence block (Figure 3).

Examining the protein sequence of LRP-1 and LRP-2 we find that most nematodes have at least two copies of LRP-like proteins with *C. elegans* LRP-1 and LRP-2 being highly similar possibly due to a recent duplication and divergence (Figure 4). Comparing the sequences across *Caenorhabditis* we find that LRP-1 proteins cluster together and LRP-2 proteins also form their own cluster. Based on location in the genome and

sequence similarity from protein alignment, we believe that *Caenorhabditis lrp-2* is a recent duplication and divergence of *lrp-1* (Figure 4).

To examine the potential role of *lrp-1* and *lrp-2* vulval cell lineage polarity we examined null mutants of each gene. Previous work has shown a role for *lrp-1* and *lrp-2* in *dab-1*/Disabled signaling and lipid transport (Kamikura and Cooper, 2003; Holmes et al., 2007; Branicky et al., 2009), but no previous role in vulval formation has been investigated. *l*rp-*1* null worms arrest early in development and due to this reason we were not able to examine its role in vulval polarity. At the time of this investigation, two potential null mutations of *lrp-2* existed, allele *gk272* with a 253 bp deletion and allele *gk292* with an approximately 1800 bp deletion. Both mutants are remarkably wild-type in most aspects of development including vulval formation (Table 1).

lrp-2 expresses in the vulval cells beginning at the two-cell stage and lasting through L4, similar to the components known to be involved in vulval cell lineage polarity (Figure 5). To investigate this interaction double mutants were constructed with both alleles of *lrp-2* and *lin-17(n671)* (Table 1). Much like *cam-1(gm122)* and *vang-1(ok1142)*, both alleles of *lrp-2* suppress the *lin-17(n671)* phenotype from 74 to approximately 50% P-Rvl, leading us to hypothesize that *lrp-2* functions in the same pathway as *cam-1* and *vang-1*. To ensure that this phenotype was a result of the null function of *lrp-2* we injected a fosmid containing the full-length sequence of *lrp-2* and found that it does rescue the double mutant phenotype of *lin-17(n671) lrp-2(gk272)* as well as *lin-17(n671) lrp-2(gk292)*. In order to better test the hypothesis that *lrp-2* functions with *cam-1*, a triple mutant was constructed between *lin-17(n671), lrp-2(gk272)*, and *cam-1(gm122)* (Table 1). This triple mutant displays the same P-Rvl

penetrance as both the *lin-17(n671) lrp-2(gk272)* and *lin-17(n671); cam-1(gm122)* double mutants confirming that *lrp-2* functions in the same pathway as *cam-1*.

LRP-2 is Downstream of EGL-20/Wnt

egl-20 is expressed in the tail (Whangbo and Kenyon, 2000) and forms a posterior-to-anterior concentration gradient (Coudreuse et al., 2006). It has previously been shown that EGL-20 acts instructively in the vulva by imparting directional information opposed to being permissive, where it would only be required for polarization (Green et al., 2008). By moving the source of *egl-20* expression from the posterior of the worm to the anchor cell, the axis of symmetry of the developing vulva, we can reorient the daughter cells of P5.p and P7.p toward the center in a wild-type configuration.

Expressing *egl-20* from the axis of symmetry suppresses the *lin-17(n671)* phenotype (Green et al., 2008; see also Table 2). To test whether LRP-2 is downstream of EGL-20 we ectopically expressed *egl-20* from the anchor cell in a *lin-17(n671) lrp-2(gk272)* background. If LRP-2 is not downstream of EGL-20 we would expect further suppression of the double mutant phenotype. However, if LRP-2 is downstream of EGL-20 we would expect no further suppression since EGL-20 would be lacking a pathway component to work through in the *lrp-2* null mutation. Because we do not see further suppression of the *lin-17(n671) lrp-2(gk272)* when expressing *egl-20* from the anchor cell we can conclude that like the transmembrane proteins CAM-1 and VANG-1, LRP-2 is downstream of EGL-20 (Table 2).

LRP-2 Controls the Localization of SYS-1/β-catenin

The polarity of the P7.p cell divisions is controlled by the Wnt/β-catenin asymmetry pathway (Green et al., 2008). This pathway includes the β-catenin-like proteins SYS-1 and WRM-1, POP-1/TCF, and the Nemo-like-kinase, LIT-1 (reviewed by Mizumoto and Sawa, 2007). The Wnt/β-catenin asymmetry pathway ensures different ratios of SYS-1 to POP-1, controlling the differential transcription of Wnt target genes between daughters of an asymmetric cell division. Because or genetic data indicate an antagonism between LRP-2 and LIN-17, similar to that between CAM-1 and VANG-1 and LIN-17, we wanted to determine if LRP-2 can control the asymmetric localization of SYS-1 between the daughter cells of P7.p during anaphase of the first cell division. The initial establishment of vulval polarity can be observed through the localization of VENUS::SYS-1 (VNS::SYS-1), localized in a high (P7.pa)/low (P7.pp) pattern in the wild-type worm, reciprocal to the localization of POP-1/TCF (Phillips et al., 2007; Green et al., 2008).

As previously reported, VNS::SYS-1 asymmetry in P7.p daughter cells is often lost in *lin-17(n671)* and *lin-18(e620)* mutants (Figure 6). These mutants display two aberrant patterns of VNS::SYS-1 localization as well as the wild-type pattern, though less frequently. The two deviant localization patterns include one in which both P7.pa and P7.pp express equal amounts of VNS::SYS-1 and a reversed VNS::SYS-1 pattern in which P7.pp is enriched with VNS::SYS-1. By observing VNS::SYS-1 localization in a *lin-17(n671) lrp-2(gk272)* background we see that the aberrant localization of SYS-1 is suppressed in a similar degree to that of *lin-17(n671); cam-1(gm122)* and *lin-17(n671); vang-1(ok1142)*. This observation confirms LRP-2 controls vulval cell polarity by

antagonizing LIN-17 in a similar fashion to CAM-1 and VANG-1 and that the effect of LRP-2 is at the level of P7.p rather than its progeny.

DISCUSSION

We have investigated the role of *lrp-2* in *C. elegans* vulval cell lineage polarity. We find that despite the high conservation of the Wnt signaling component, LRP5/6, in higher order organisms it appears to have evolved after the split of Nematoda due to its lack of presence in all nematode genomes examined. *C. elegans* contains multiple low-density lipoprotein receptors within its genome, two of which are *lrp-1* and *lrp-2*. Due to the position in the genome and high sequence similarity we believe that *lrp-2* is the product of a recent duplication of *lrp-1* within the *Caenorhabditis* lineage.

lrp-1 mutants are sick and arrest during an early larval stage. For this reason we were not able to examine the role of *lrp-1* in vulval formation. Despite high sequence similarity and proposed functional redundancy with *lrp-1*, *lrp-2* is remarkably wild-type as a single mutant. *lrp-2* expresses in the developing vulva at the same time as both *cam-1* and *vang-1*. By examining double and triple mutant strains we find that *lrp-2* functions downstream of *egl-20* along with transmembrane proteins *cam-1* and *vang-1* (Figure 7). All three genes antagonize the *lin-17*/Frizzled pathway by directing the aberrant localization of SYS-1 to the posterior daughter cell of P7.p leading to the posterior orientation of the P7.p lineage.

This work provides evidence that despite lacking a true LRP5/6 ortholog, the formation of the *C. elegans* vulva is controlled by another member of the low-density lipoprotein superfamily, *lrp-2*. This data could potentially lead to insight into the evolution of both structure and function of the highly important Wnt pathway component, LRP5/6. Despite strong genetic evidence, this work does not describe the physical interaction between LRP-2 and CAM-1, VANG-1, EGL-20. Can LRP-2 bind with the

other transmembrane proteins, CAM-1 and VANG-1, involved in this pathway? Can LRP-2 physically bind the Wnt ligand, EGL-20? Future work should focus on the biochemistry of this pathway. Answers to these questions could provide interesting insights into the evolution of low-density lipoprotein receptors, including LRP5/6, as well as how Wnt signaling has evolved within nematodes without the presence of one of the most important and highly conserved transmembrane proteins.

ACKNOWLEDGMENTS

We would like to thank Jennifer Green and Takao Inoue for insightful comments. Thank you to Adler Dillman for assistance with the phylogenetics and evolutionary trees. Thanks is also due to Barbara Perry and Gladys Medina for strains and reagent preparation. Strains were provided by the CGC. This work was supported by a National Institutes of Health (NIH) United States Public Health Service Training Grant (T32GM07616) to P.J.M., and by the Howard Hughes Medical Institute (with which P.W.S. is an investigator).

REFERENCES

Branicky, R., Desjardins, D., Liu, J., and Hekimi, S. (2010). Lipid Transport and Signaling in *Caenorhabditis elegans*. *Developmental Dyanmics* **239,** 1365-1377.
Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
Coudreuse, D. Y., Roel, G., Betist, M. C., Destree, O. and Korswagen, H. C. (2006). Wnt gradient formation requires retromer function in Wnt-producing cells. *Science* **312**, 921-924.
Eisenmann, D.M., (2005). Wnt signaling. *Wormbook : the online review of C. elegans biology*, 1-17.
Ferguson, E. L., Sternberg, P. W. and Horvitz, H. R. (1987). A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*. *Nature* **326**, 259-267.
Gleason, J. E., Szyleyko, E. A. and Eisenmann, D. M. (2006). Multiple redundant Wnt signaling components function in two processes during *C. elegans* vulval development. *Developmental biology* **298**, 442-457.
Green, J. L., Inoue, T. and Sternberg, P. W. (2008). Opposing Wnt pathways orient cell polarity during organogenesis. *Cell* **134**, 646-656.
Gupta, B. P., Hanna-Rose, W. and Sternberg, P. W. (2012). Morphogenesis of the vulva and the vulval-uterine connection. *WormBook : the online review of C. elegans biology*, 1-20.
He, X., Semenov, N. Kelko, T., and Zeng, X. (2004). LDL receptor-related proteins 5 and 6 in Wnt/b-catenin signaling:
Arrows point the way. *Development* **131**, 1663-1677.
Hobert, O. (2002). PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *BioTechniques* **32**, 728-730.
Holmes, A., Flett, A., Coudreuse, D., Korswagen, H., and Pettitt, J. (2007). C. elegans Disabled is required for cell-type specific endocytosis and is essential in animals lacking the AP-3 adaptor complex. *Journal of Cell Science* **120**, 2741-2751.
Inoue, T., Oz, H. S., Wiland, D., Gharib, S., Deshpande, R., Hill, R. J., Katz, W. S. and Sternberg, P. W. (2004). *C. elegans* LIN-18 is a Ryk ortholog and functions in parallel to LIN-17/Frizzled in Wnt signaling. *Cell* **118**, 795-806.
Kamikura, D. and Cooper, J. (2003). Lipoprotein receptors and a Disabled family cytoplasmic adaptor protein regulate EGL-17/FGF export in *C. elegans*. *GENES & DEVELOPMENT* **17**, 2798–2811.
Katz, W. S., Hill, R. J., Clandinin, T. R. and Sternberg, P. W. (1995). Different levels of the *C. elegans* growth factor LIN-3 promote distinct vulval precursor fates. *Cell* **82**, 297-307.
Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal* **10**, 3959-3970.
Mizumoto, K. and Sawa, H. (2007). Two betas or not two betas: regulation of asymmetric division by beta-catenin. *Trends in cell biology* **17**, 465-473.
Phillips, B. T., Kidd, A. R., 3rd, King, R., Hardin, J. and Kimble, J. (2007). Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric

divisions in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3231-3236.
Sawa, H., Lobel, L. and Horvitz, H. R. (1996). The *Caenorhabditis elegans* gene *lin-17*, which is required for certain asymmetric cell divisions, encodes a putative seven-transmembrane protein similar to the Drosophila frizzled protein. *Genes & development* **10**, 2189-2197.
Sternberg, P. W. and Horvitz, H. R. (1988). *lin-17* mutations of *Caenorhabditis elegans* disrupt certain asymmetric cell divisions. *Developmental biology* **130**, 67-73.
Sternberg, P. W. (2005). Vulval development. *WormBook : the online review of C. elegans biology*, 1-28.
Strutt, D. (2005). Organ shape: controlling oriented cell division. *Current biology : CB* **15**, R758-759.
Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Developmental biology* **56**, 110-156.
Wodarz, A. and Nathke, I. (2007). Cell polarity in development and cancer. *Nature cell biology* **9**, 1016-1024.
Yochem, J., Tuck, S., Greenwald, I., and Han, M. (1999). A gp330/megalin-related protein is required in the major epidermis of Caenorhabditis elegans for completion of molting. *Development* **126**, 597-606.
Whangbo, J., Harris, J., Kenyon, C. (2000). Multiple levels of regulation specify the polarity of an asymmetric cell division in *C. elegans*. *Development* **127**, 4587-4598.

Table 1

Relevant Genotype	% P-Rvl	n	p value
N2	0	100	
lin-17(n671)	74	100	
lin-18(e620)	31	100	
lin-17(n671); lin-18(e620)	100	40	
egl-20(hu120)	0	66	
lin-17(n671); egl-20(hu120)	6	52	<0.0001
cam-1(gm122)	0	54	
vang-1(ok1142)	0	58	
lrp-2(gk272)	0	40	
lrp-2(gk292)	0	40	
lin-17(n671); cam-1(gm122)	50	54	0.0042
lin-17(n671); vang-1(ok1142)	48	60	0.0013
lin-17(n671) lrp-2(gk272)	55	40	0.0430
lin-17(n671) lrp-2(gk292)	54	55	0.0198
lin-17(n671) lrp-2(gk272); cam-1(gm122)	49	39	

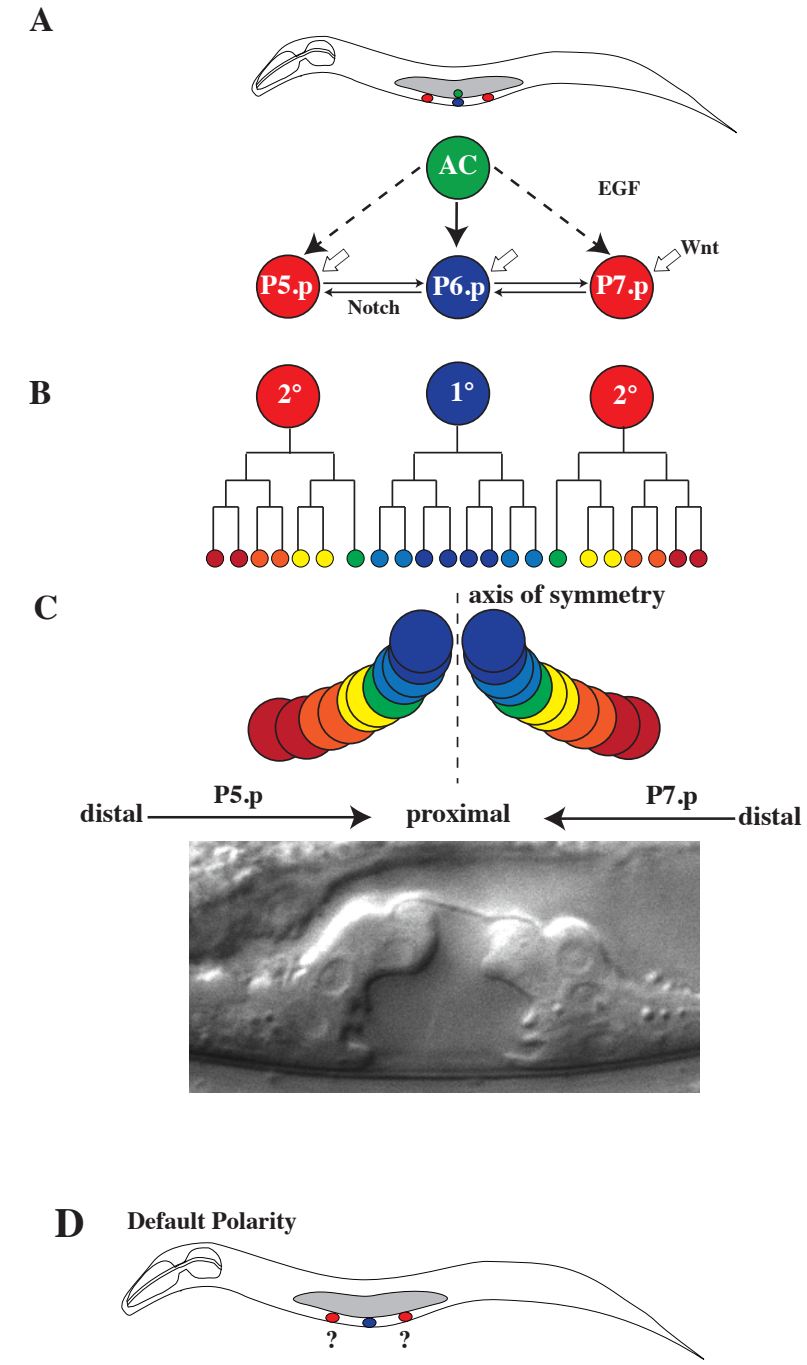
Table 1. Genetic analysis of *lrp-2*
lrp-2 suppresses the phenotype of *lin-17(n671)* in roughly the same manner as *cam-1(gm122)* and *vang-1(ok1142)*. The triple mutant of *lin-17(n671) lrp-2(gk272); cam-1(gm122)* does not show further suppression of either of the two double mutants.

Table 2

Relevant Genotype	Anterior EGL-20 Source	% P-Rvl	n	p value
lin-17(n671)	-	74	100	
lin-17(n671)	+	25	44	<0.0001
lin-17(n671) lrp-2(gk272)	-	54	40	
lin-17(n671) lrp-2(gk272)	+	52	21	

Table 2. LRP-2 is downstream of EGL-20
EGL-20 acts instructively. By ectopically expressing EGL-20 from the anterior side of P7.p using the *Pfos-1a* promoter, P5.p and P7.p reorient to face the anterior gradient. The anterior source of EGL-20 suppresses the *lin-17(n671)* phenotype. Because we do not see further suppression of the *lin-17(n671) lrp-2(gk272)* phenotype when EGL-20 is ectopically expressed from the anterior side of P7.p we conclude that LRP-2 is downstream of EGL-20.

Figure 1


A
AC
EGF
Wnt
P5.p
P6.p
P7.p
Notch
B
2°
1°
2°
C
axis of symmetry
distal
P5.p
proximal
P7.p
distal
D
Default Polarity
?
?

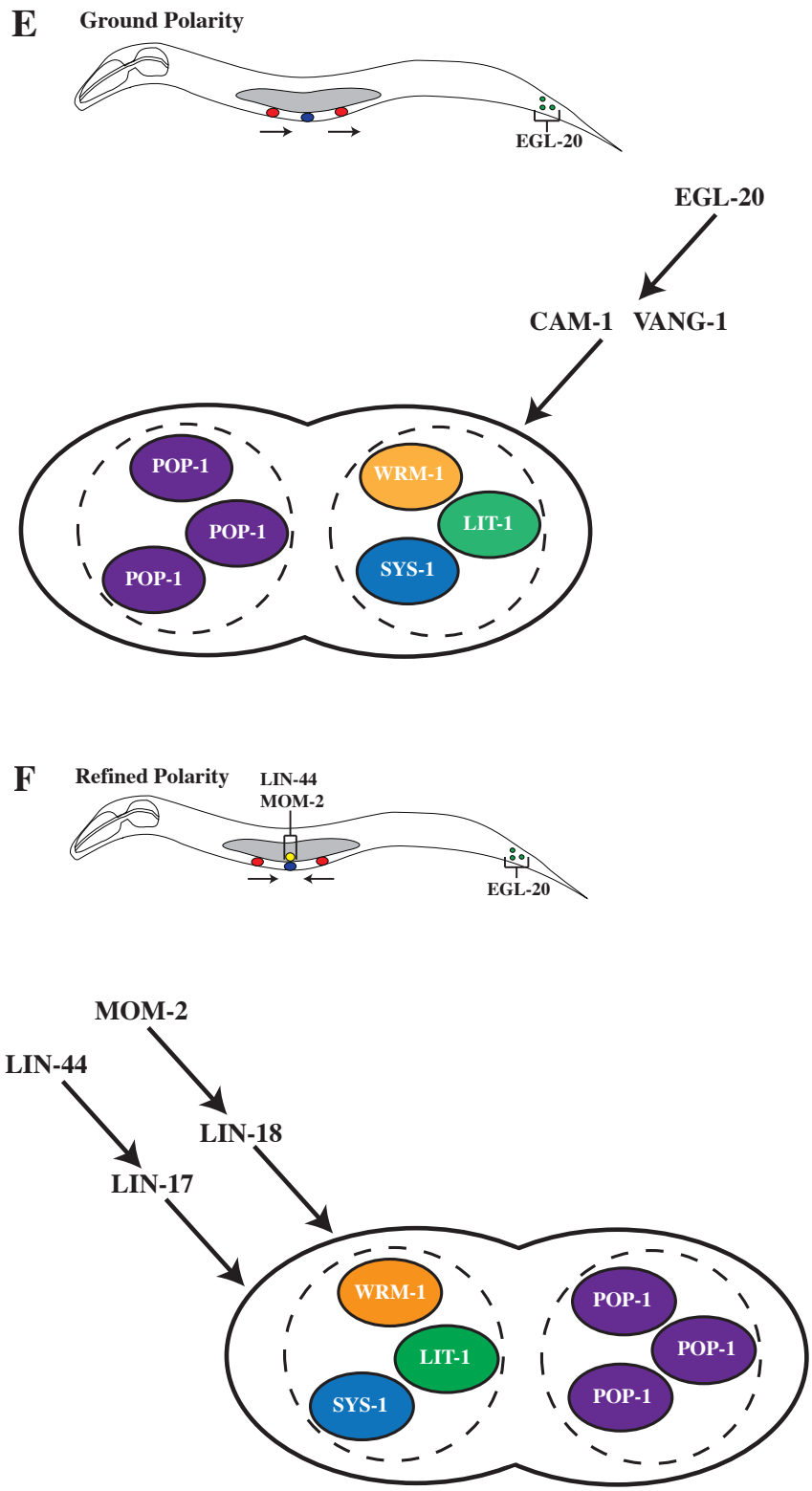


Fig. 1. *C. elegans* vulval development
(A) Schematic of vulval induction illustrating sources of EGF, Notch, and Wnt. (B) Lineage trees of VPC progeny: P5.p, 2° fate, on the left, P6.p, 1° fate, in center, and P7.p, 2° fate, on left. The progeny of each cell is color coded: A cells – red, B cells (B1 and

B2) – orange, C cells – yellow, D cells – green, E cells light blue, and F cells dark blue. (C) Final conformation of vulval cells shown as a cartoon and Nomarski image in mid-L4 stage. Mirror symmetry is noted about the vulval center. Proximal daughter cells of P5.p and P7.p join the daughters of P6.p in forming the vulval lumen whereas the distal most daughters of P5.p and P7.p adhere to the ventral epidermis. (D) The default polarity of P5.p and P7.p is random in the absence of all Wnts. (E) *egl-20* is expressed in the tail (green circles) and establishes ground polarity in which both P5.p and P7.p face the posterior as a result of asymmetric localization of SYS-1, LIT-1, WRM-1 to the posterior daughter of P7.p and POP-1 to the anterior daughter. (F) *lin-44* and *mom-2* are expressed in the anchor cell (yellow circle) resulting in refined polarity where both P5.p and P7.p both face towards the center as a result of asymmetric localization of SYS-1, LIT-1, and WRM-1 to the anterior daughter cell of P7.p and POP-1 to the posterior daughter cell.

Figure 2

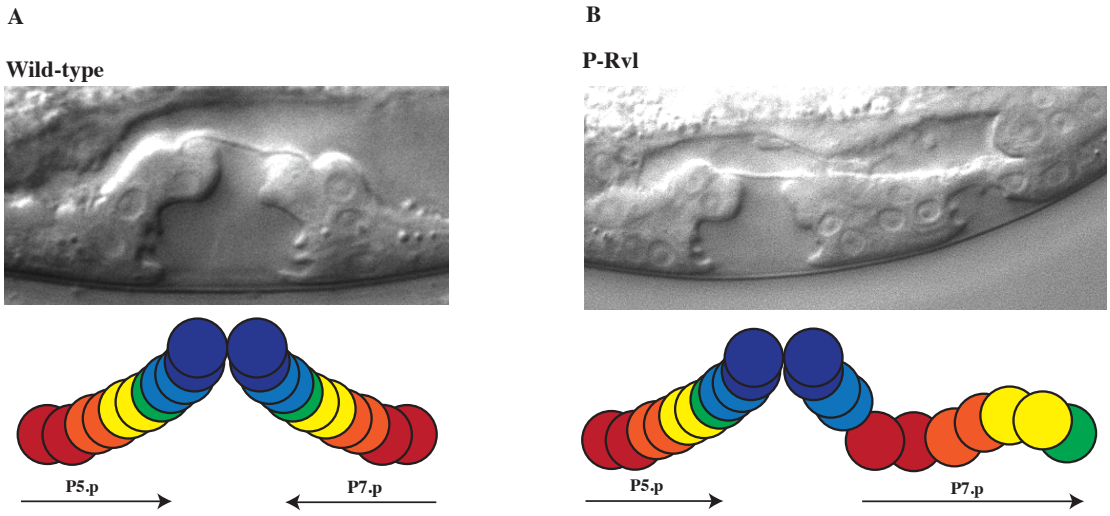


Fig. 2. Wild-type vulva vs. Posterior-reversed vulval lineage vulva
(A) Wild-type vulva formed from 22 progeny of 3 VPCs: P5.p, P6.p, and P7.p. The progeny of P5.p and P7.p form mirror symmetry about the vulval center. (B) Posterior-reversed vulval lineage: the daughter cells of P7.p mimic those of P5.p. Both images taken with *sem-5(n1779)* background.

Figure 3

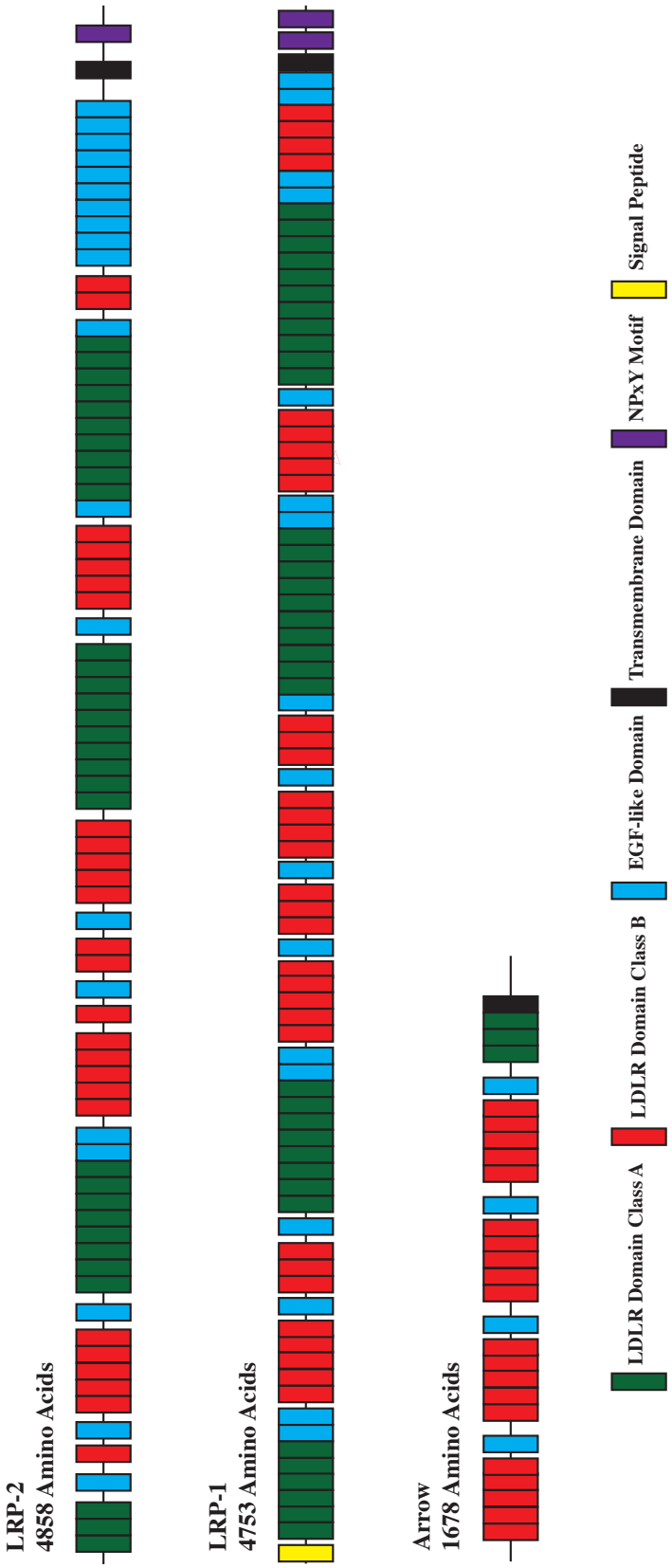

LRP-2
4858 Amino Acids
LRP-1
4753 Amino Acids
Arrow
1678 Amino Acids
LDLR Domain Class A
LDLR Domain Class B
EGF-like Domain
Transmembrane Domain
NPxY Motif
Signal Peptide

Figure 3. Protein domains of LRP-2, LRP-1, and *Drosophila* Arrow
C. elegans does not possess a true ortholog of Arrow (LRP5/6); however, it does possess multiple megalin-like proteins that contain LDLR Class A repeats, LRDR Class B repeats, and EGF-like domains that are found in varieties of low density like lipoprotein receptors such as megalin and Arrow. All domains are color-coded and drawn to approximate scale according to the SMART database.

Figure 4

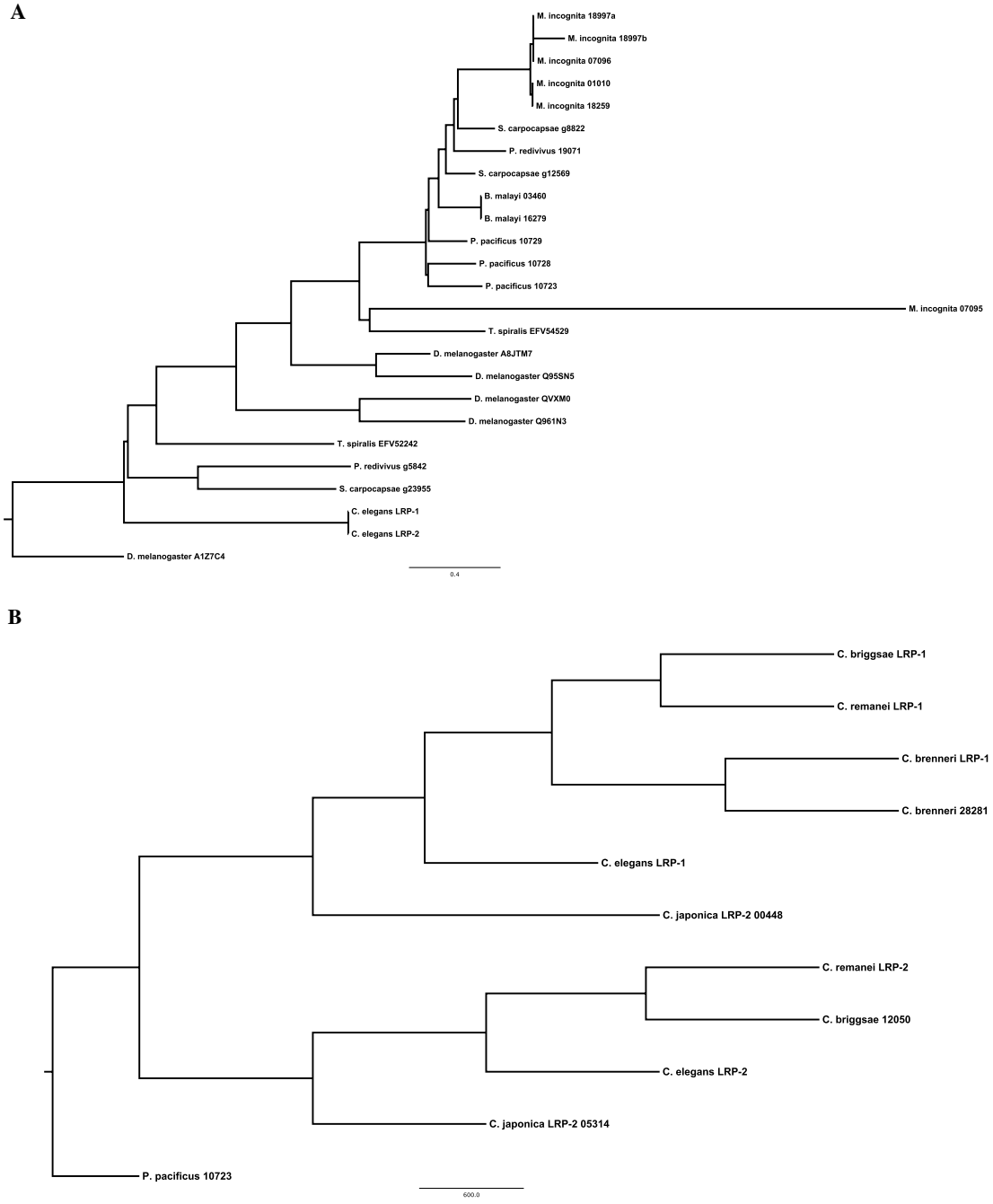


Figure 4. Evolutionary trees of *lrp-2*

(A) A tree based on the protein sequence of LRP-1 and LRP-2 in nematodes and megalin in *Dropsophila*. Based on sequence similarity, position in the genome, and clustering, it appears that LRP-2 is the result of a recent duplication in *Caenorhabditis*. (B) Within *Caenorhabditis*, LRP-1 orthologs cluster together and LRP-2 orthlogs cluster. *Pristionchus pacificus* is used as the outgroup.

Figure 5

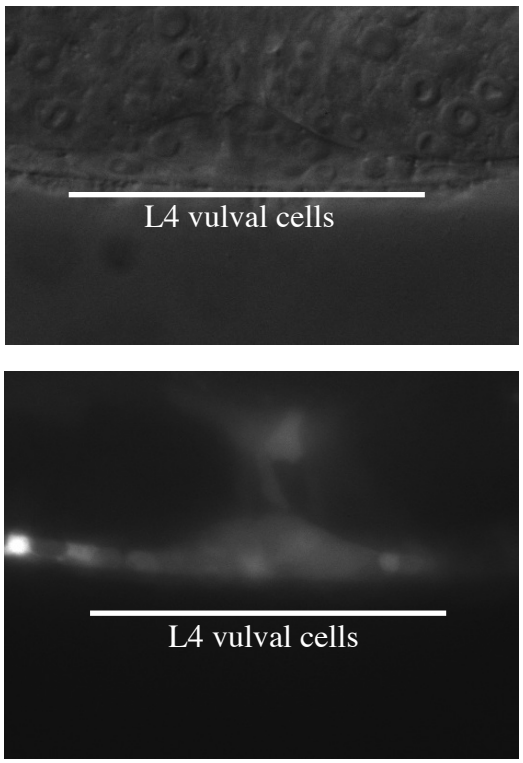
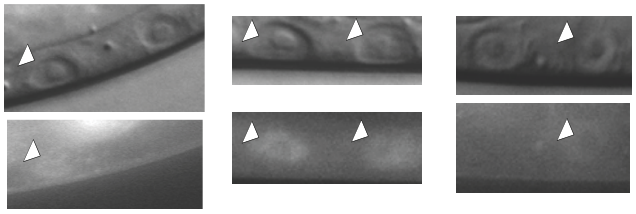


Figure 5. *lrp-2* expresses in the vulva
lrp-2 expresses in the vulva beginning as early as the two-cell stage. The L4 vulva is pictured using a *lrp-2* promoter gfp fusion. The fluorescent picture is shown below the Nomarski image.

Figure 6



	Number of Worms		
Relevant Genotype	P7.pa > P7.pp	P7.pa = P7.pp	P7.pa < P7.pp
+	20	0	0
lin-17(n671)	3	8	9
lin-17(n671); cam-1(gm122)	8	8	4
lin-17(n671); vang-1(ok1142)	6	12	2
lin-17(n671) lrp-2(gk272)	7	10	3

Figure 6. Subcellular localization of VNS::SYS-1
The localization pattern of VNS::SYS-1 in P7.p daughter cells. The resulting pattern was classified by eye into three categories: SYS-1 enriched in the anterior daughter (P7.pa > P7.pp), SYS-1 present at similar levels in both daughters (P7.pa = P7.pp), and SYS-1 enriched in the anterior daughter (P7.pa < P7.pp). A representative image of each scenario is shown.

Figure 7

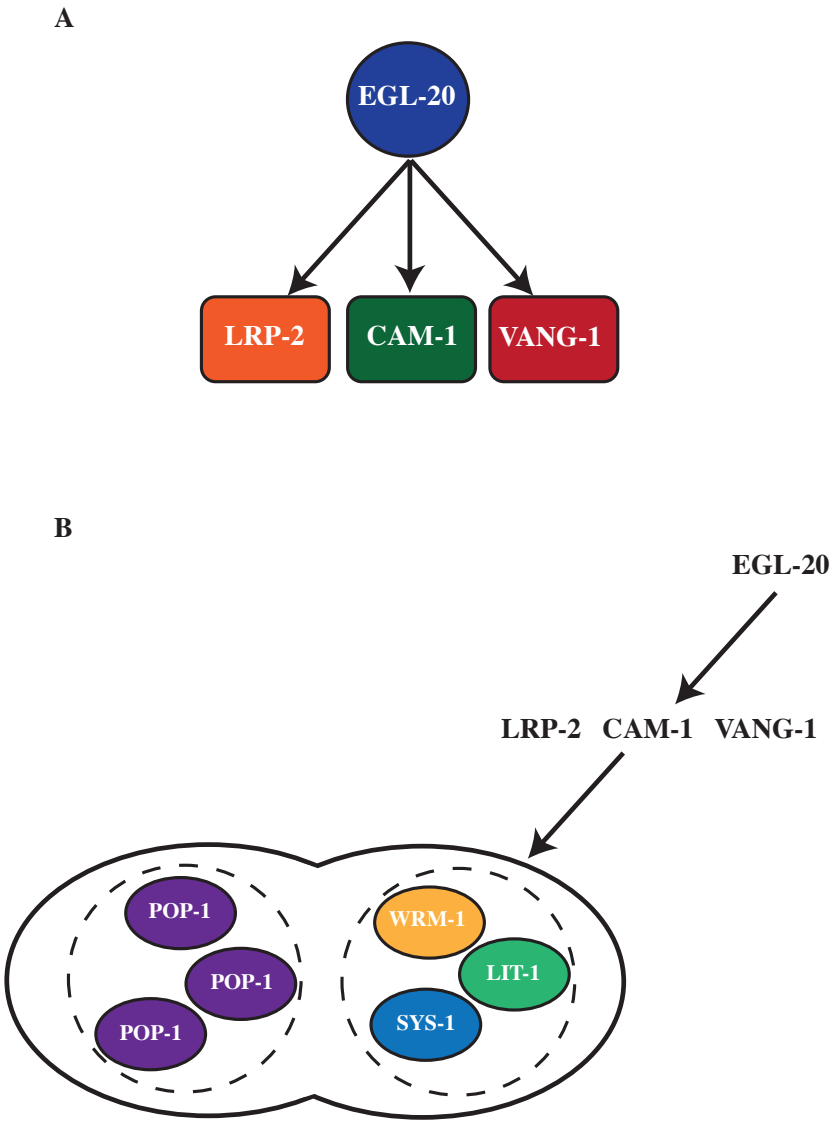


Figure 7. The role of LRP-2 in vulval lineage polarity
(A) EGL-20 acts upstream of LRP-2, CAM-1, and VANG-1. (B) EGL-20 is expressed on the posterior side of P7.p and acts through LRP-2, CAM-1, and VANG-1 to drive the localization of SYS-1 to the posterior daughter cell of P7.p. This localization results in posterior-facing orientation of the P7.p daughter cells.

Chapter 3

The origin and evolution of dishevelled

This work was done in collaboration with Adler R. Dillman

ABSTRACT

Dishevelled (Dsh or Dvl) is an important signaling protein, playing a key role in Wnt signaling and relaying cellular information for several developmental pathways. Dsh is highly conserved among metazoans and has expanded into a multigene family in most bilaterian lineages, including vertebrates, planarians, and nematodes. These orthologs, where explored, are known to have considerable overlap in function, but evidence for functional specialization continues to mount. We performed a comparative analysis of Dsh across animals to explore protein architecture and identify conserved and divergent features that could provide insight into functional specialization with an emphasis on invertebrates, especially nematodes. We find evidence of dynamic evolution of Dsh, particularly among nematodes, with taxa varying in ortholog number from one to three. We identify a new domain specific to some nematode lineages and find an unexpected nuclear localization signal conserved in many Dsh orthologs. Our findings raise questions of protein evolution in general and provide clues as to how animals have dealt with the complex intricacies of having a protein, such as Dsh, act as a central messenger hub connected to many different and vitally important pathways. We discuss our findings in the context of functional specialization and bring many testable hypotheses to light.

INTRODUCTION

Dishevelled (Dsh or Dvl) is a multifunctional phosphoprotein originally discovered in *Drosophila* and named for its disruptions in hair and bristle polarity (Fahmy and Fahmy, 1959; Klingensmith et al., 1994). Dsh plays a key role in Wnt signaling, thus affecting cell proliferation, migration, polarity, terminal differentiation, and the self-renewal of stem cells (Boutros and Mlodzik, 1999; Gao and Chen, 2010; Sugimura et al., 2012; Tauriello et al., 2012). Deregulation of pathway components is associated with multiple human diseases. Wnt signaling has evolved to act in multiple pathways, broadly divided into the canonical/β-catenin-dependent pathway and the non-canonical/β-catenin-independent pathway, with Dsh acting in a key role, relaying signals from receptors to downstream effectors (Gao and Chen, 2010). Several components of the Wnt signaling pathway, including Frizzled, GSK3, and β-catenin, can be found in protozoans, but it is not until the emergence of Metazoa that we see a complete Wnt pathway (Holstein, 2012). The early branching metazoan lineage Porifera has only the major components of the canonical pathway, with critical non-canonical pathway components arising subsequently in eumetazoan lineages (Kusserow et al., 2005; Adamska et al., 2007; Adamska et al., 2010). Thus, although Wnt signaling is conserved across Metazoa from sponges to humans, it seems that this pathway's origin and the original role of Dsh lies in the canonical/β-catenin-dependent pathway, with non-canonical signaling developing later. Figure 1 summarizes the evolution of Dsh across animals including the loss and gain of orthologs, paralogs, and protein domains.

In the current model of the canonical pathway, Wnt signals are received by a Frizzled (Fz) and LRP5/6 co-receptor complex, leading to the recruitment of Dsh and Axin to the

cell membrane. This recruitment results in the inactivation and dissolution of the β-catenin destruction complex allowing for the nuclear translocation of β-catenin where it interacts with members of the TCF/Lef transcription family to regulate gene expression (Wharton, 2003). Non-canonical signaling encompasses several different pathways that do not necessarily lead to the activation of β-catenin, of which the best understood is the planar cell polarity (PCP) pathway. PCP signaling is responsible for the polarization of cells along an epithelial sheet. The core components of this pathway include the transmembrane proteins Fz, Flaming (Fmi), and Strabismus (Stbm), as well as the cytoplasmic proteins Diego (Dgo), Prickle (Pk), and Dsh (Seifert and Mlodzik, 2007). In general PCP signaling relies on complex interactions between these core components that lead to their asymmetric enrichment and distribution within a cell. For example, during polarization in the *Drosophila* wing two distinct protein complexes antagonize each other and localize to opposite ends of the cell: a Fz-Dsh-Dgo complex becomes enriched at the distal end of each cell, whereas a Stbm-Pk complex concentrates proximally (Simons and Mlodzik, 2008).

The literature establishes the archetypal Dsh protein to contain three conserved domains: an N-terminal DIX (Dishevelled and Axin) domain, a central PDZ (Post Synaptic Density-95, Discs Large, and Zonula occludens-1) domain, and a C-terminal DEP (Dishevelled, Egl-10, Pleckstrin) domain (Penton et al., 2002; Gao and Chen, 2010) (Figure 2). In addition to these three conserved domains, Dsh is known to contain a basic region that precedes the N-terminal of the PDZ domain as well as a proline-rich region that includes an SH3 binding motif located between the PDZ and DEP domains (Figure 2). Both of these regions are thought to be conserved and have functional significance.

There is a fourth domain, reported to be conserved, in Dsh, the DSV or Dishevelled domain, although its functional significance is rarely discussed (Figure 2). The Dsh protein contains approximately 15% serine and threonine residues many of which are phosphorylated; however, the functional significance of these residues has been questioned (Yanfeng et al., 2011). Expansions in Dsh among metazoan lineages seem common and variable, with most vertebrates containing three Dsh homologs (although the chicken *Gallus gallus* has two), while insects have only one (Klingensmith et al., 1994; Sweetman et al., 2008; Gray et al., 2009; Gao and Chen, 2010). The nematode *Caenorhabditis elegans* has three Dsh homologs and the planarian *Schmidtea mediterranea* has two (Ruvkun and Hobert, 1998; Gurley et al., 2008; Figure 1).

With the ever-increasing amount of genomic data available for analysis, we leveraged the currently available data to study the evolution of Dsh across animals with an emphasis on nematodes. In addition to exploring the potential conservation of the three *C. elegans* Dsh homologs among nematodes, we were interested in identifying conserved or divergent protein features that correlate with the known functional divergence between Dsh orthologs observed in several animal taxa, and that could provide hypotheses about the evolution Dsh. For example, the planarian Dsh paralogs, *Smed-dvl-1* and *Smed-dvl-2*, appear to be functionally specialized such that only *Smed-dvl-2* is thought to be involved in β-catenin-dependent signal transduction, suggesting underlying physical differences in these proteins that have not yet been linked to their divergent function (Almuedo-Castillo et al., 2011). Similarly, the function of Dsh orthologs seems to have diverged among vertebrates, where *Dvl1* and *Dvl2*, but not *Dvl3*, are necessary to mediate the Wnt-dependent signals that control neural crest specification in *Xenopus*, but in murines it is

thought that *Dvl2* and *Dvl3* function in neural crest development whereas *Dvl1* apparently does not (Lijam et al., 1997; Hamblet et al., 2002; Monsoro-Burq et al., 2005; Etheridge et al., 2008; Gray et al., 2009). It is still not known whether Dsh's role in neural crest development is through the canonical or non-canonical pathways, or both (Etheridge et al., 2008). In the our study we found that Dishevelled is a highly conserved protein that has undergone dynamic evolution across metazoans and variation in protein architecture provides clues about its functional roles in β-catenin-dependent and -independent pathways.

MATERIALS AND METHODS

Orthology Analyses

To study the evolution of Dishevelled, we used the available predicted protein datasets from WormBase release WS225 (www.wormbase.org) for the following species – *Brugia malayi*, *Caenorhabditis elegans*, *C. angaria*, *C. japonica*, *C. brenneri*, *C. remanei*, *C. briggsae*, *Meloidogyne hapla*, *Pristionchus pacificus*, and *Trichinella spiralis*. We also included the *Ascaris suum, Bursaphelenchus xylophilus*, and *Meloidogyne incognita* predicted proteome data sets from WormBase release WS229. For outgroup and comparative analysis we used the predicted protein datasets of *Arabidopsis thaliana* (vGNOMON 7/9/07), *Drosophila melanogaster* (v10/30/11), *Homo sapiens* (v9/7/11), *Mus musculus* (v3/4/11), *Nasonia vitripennis* (v1.2), *Saccharomyces cerevisiae* (v2/3/11), and *Tribolium castaneum* (vTcas 3.0) genome projects, obtained from the NCBI/NIH repository (ftp://ftp.ncbi.nih.gov/genomes). Pre-released proteomes for *Panagrellus redivivus* and *Steinernema carpocapsae* were also used, from manuscripts in preparation (Dillman et al., 2012). Dsh orthologs from the jellyfish, *Clytia hemisphaerica* (AFI99114.1), the planarian *Schmidtea mediterranea* (Smed-DVL-1 ADZ58511.1 and Smed-DVL-2 ADZ58512.1), the frog *Xenopus tropicalis* (DVL1 NP_001116886.1, DVL2 NP_001072660.1, and DVL3 NP_01116929.1), the sponge *Amphimedon queenslandica* (XP_003384321), and the tunicate *Ciona intestinalis* (NP_001027754.1) were acquired from GenBank (http://blast.ncbi.nlm.nih.gov/).

Version 1.4 of the OrthoMCL pipeline was used to cluster proteins from the proteomes into families of orthologous genes (http://www.orthomcl.org) (Li et al., 2003). To identify orthologs of Dsh across animals, we ran OrthoMCL using the full proteomes

of *C. elegans*, *P. redivivus*, *T. spiralis*, *N. vitripennis*, *D. melanogaster*, *T. castaneum*, *M. musculus*, *H. sapiens*, *S. cerevisiae*, and *A. thaliana*. To identify orthologs across Nematoda, we ran OrthoMCL using the full proteomes of *B. malayi*, *A. suum*, *P. pacificus*, *C. elegans*, *B. xylophilus*, *M. hapla*, *M. incognita*, *P. redivivus*, *S. carpocapsae*, *T. spiralis*, with *N. vitripennis* as an outgroup. To identify orthologs within *Caenorhabditis* we ran OrthoMCL using the full proteomes of *C. angaria*, *C. briggsae*, *C. brenneri*, *C. japonica*, *C. remanei*, and *C. elegans*. All orthology analyses were run using OrthoMCL version 1.4 with default settings and the BLAST parameters recommended in the OrthoMCL documentation (Li et al., 2003). Identified orthology clusters are available in the supporting material online (Supplementary Material online).

Domain Analysis

Each identified Dsh ortholog was analyzed for protein domains using the SMART protein domain analysis website (http://smart.embl-heidelberg.de), used in normal mode (Letunic et al., 2012). All additional options (outlier homologues, PFAM domain, signal peptides, internal repeats, and intrinsic protein disorder) were turned on for the analysis. The full protein sequences and identified domains are available in the supplementary material.

Sequence Alignment, Phylogenetics, and Selection Detection

Sequence alignments were made using all of the amino acid sequence from the beginning of the PDZ domain to the end of the DEP domain, since these were the only identified domains conserved across all of the animal taxa we evaluated. Protein

sequence alignments of this region were made using the online MUSCLE service (http://www.ebi.ac.uk/Tools/msa/muscle) (Edgar, 2004). These protein alignments were then replaced with the appropriate nucleotide sequences using the RevTrans server (http://www.cbs.dtu.dk/services/RevTrans), which preserves the alignment obtained with the amino acid sequences, but replaces each amino acid with the user-supplied protein coding nucleotides (Wernersson and Pedersen, 2003). Coding sequence for the proteomes was downloaded along with the proteomes from the sites listed above, though the sequences for most of the nematodes we used could also be acquired from WormBase (www.wormbase.org). Two separate alignments were made using this method, one that included the Dsh orthologs across animals, including *T. castaneum*, *N. vitripennis*, *D. melanogaster*, *M. musculus*, *H. sapiens*, *T. spiralis*, *A. suum*, *B. malayi*, *P. pacificus*, *C. elegans*, *S. carpocapsae*, *B. xylophilus*, *P. redivivus*, *M. hapla*, *M. incognita*, and the jellyfish *C. hemisphaerica*. The other alignment focused on Dsh orthologs within caenorhabditid nematodes, utilizing genes from *C. elegans*, *C. angaria*, *C. japonica*, *C. brenneri*, *C. remanei*, *C. briggsae*, with the intracellular parasite *T. spiralis* and the parasitoid wasp *N. vitripennis* as outgroups. Alignments were then shaded to reflect sequence conservation using GeneDoc (http://www.nrbsc.org/gfx/genedoc) (Nicholas et al., 1997).

The nucleotide alignments were then evaluated for the best-fit model of evolution using jModelTest2 (http://code.google.com/p/jmodeltest2) (Guindon and Gascuel, 2003; Darriba et al., 2012). For the alignment across animals, the corrected Akaike information criterion, the Bayesian inference criterion, and the decision theory criterion all selected the GTR+I+G model of evolution, with a p-invar=0.0640 and a gamma shape parameter

of 0.9840. The analysis of the *Caenorhabditis* alignment resulted in the GTR+G model being chosen by all criteria, with a gamma shape parameter of 0.5140.

Following model selection, maximum likelihood (ML) analyses with 1,000 bootstraps were done using the RAxML BlackBox server (http://phylobench.vital-it.ch/raxml-bb) (Stamatakis et al., 2008). New technology parsimony analyses were done using TNT (http://www.cladistics.com/aboutTNT.html) (Goloboff, 1999; Nixon, 1999). Maxtrees was set to 10,000. A new technology, random driven search was performed using ratchet, drift, and tree fusing options. A bootstrap analysis of 1,000 was performed by resampling.

Selection was detected using two methods. First, the alignment files of the protein-coding nucleotide sequences were uploaded into MEGA 5.05 (http://www.megasoftware.net) (Tamura et al., 2011). The selection analysis option in MEGA, which estimates selection for each codon using HyPhy, was used. Our ML analysis served as the guide tree, and the ML statistical method was chosen using the GTR model, as selected by jModelTest2. All sites were used in the analysis. Following the MEGA analysis of selection, we used the HyPhy package as implemented by the Datamonkey adaptive evolutionary server (http://www.datamonkey.org). Alignment files with the ML phylogenetic analysis written into them were uploaded using the codon data type and the universal genetic code. We used the recommended meme method in our analyses, setting the options to estimate the global dN/dS value and to average encountered ambiguities in the consensus sequence (Murrell et al., 2012). We chose to set the level of significance at p=0.1.

RESULTS

Dishevelled Conservation and Expansions Among Animals

We evaluated the conservation and potential expansion of Dsh using cluster analysis of seventeen whole proteomes, including vertebrates, insects, nematodes, and a fungal and plant proteome as outgroups (see Materials and Methods). We found no evidence of Dsh or Dsh-like genes outside Metazoa. It was previously known that *D. melanogaster* and potentially all insects have one Dsh (*Dmel-dsh*), the model nematode *C. elegans* has three Dsh homologs (*Cele-dsh-1*, *Cele-dsh-2*, and *Cele-mig-5*), the planarian *Schmidtea mediterranea* has two (*Smed-dvl-1* and *Smed-dvl-2*), and most vertebrates have three (*Dvl1*, *Dvl2*, and *Dvl3*; Figure 1). We found three distinct clusters of Dsh genes, the largest included all of the insect orthologs (*Dmel-dsh*, *Nvit-dsh*, *Tcas-dsh*; one copy in each insect proteome), all nematode *dsh-1* orthologs, and the vertebrate orthologs *Dvl1* and *Dvl3* (*Mmus-Dvl1*, *Mmus-Dvl-3*, *Hsap-Dvl1*, and *Hsap-Dvl3*). A second cluster included exclusively nematode *mig-5* genes, while the vertebrate *Dvl2* orthologs (*Mmus-Dvl-2* and *Hsap-Dvl2*) formed their own cluster, apparently having no orthologs outside vertebrates. The *C. elegans dsh-2* remained an unclustered orphan in this broad analysis.

Using the three *C. elegans* Dsh homologs (*Cele-dsh-1*, *Cele-dsh-2*, and *Cele-mig-5*) as queries, we found that only *Cele-dsh-1* has orthologs outside of Nematoda, being highly conserved across metazoans, with all insects and nematodes having only one strict ortholog, and vertebrates having two, *Dvl1* and *Dvl3*. Among the nematode genera in this analysis, *C. elegans* is unique in having three Dsh homologs. In addition to having an ortholog of *dsh-1*, most nematodes also have an ortholog of *mig-5*, but none of the nematodes in this analysis have orthologs of *Cele-dsh-2*. Unlike the rest of the nematodes

we studied, *T. spiralis*, which is in the basal clade 2 of Nematoda, retains only a single Dsh ortholog (*Tspi-dsh-1*)(Holterman et al., 2006; Figure 3). This result shows that Dsh has experienced dynamic evolution within Nematoda, with extant taxa possessing one, two, or three Dsh homologs.

In evaluating the relationships among Dsh genes across animals, we included known homologs in organisms for which we did not perform whole genome searches (*C. hemisphaerica*, *S. mediterranea*, and *X. tropicalis*). We found that each of the three vertebrate Dsh orthologs shares ancestry, and that the planarian *S. mediterranea* and the nematode *T. spiralis* Dshs (*Smed-dvl-1*, *Smed-dvl-2*, and *Tspi-dsh-1*) are not very similar to the other nematode Dsh homologs (Figure 4). The rest of the nematode Dsh orthologs in this analysis formed two distinct clades, the *mig-5* orthologs forming one clade, and the *dsh-1* orthologs forming the other (Figure 4). Although we did not find any proteins with significant similarity to *Cele-dsh-2* in our broad clustering analysis, the phylogenetic analyses place it in the clade containing all nematode *dsh-1* orthologs (Figure 4). These *C. elegans* paralogs, *dsh-1* and *dsh-2*, are approximately 130 kb apart on chromosome II, have the same orientation, and form a conserved gene cluster with a recombination frequency of 0.61%. Together with the phylogenetic analyses, this suggests that *Cele-dsh-2* is a diverging duplication of *Cele-dsh-1* and is specific to *C. elegans* or perhaps the *Caenorhabditis* lineage.

To evaluate the apparent expansion of Dsh among caenorhabditids and determine the origin of *dsh-2*, we performed a cluster analysis of whole caenorhabditid proteomes, including *C. angaria*, *C. japonica*, *C. elegans*, *C. brenneri*, *C. remanei*, and *C. briggsae*. The relationships among *Caenorhabditis* nematodes are becoming increasingly refined as

more species are described (Kiontke et al., 2011). Our analysis includes members of both the *Elegans* and *Drosophilae* supergroups within the *Caenorhabditis* genus, and resulted in three unsurprising clusters of Dsh genes: *dsh-1*, *dsh-2*, and *mig-5*. We found that all caenorhabditids have *dsh-1* and *mig-5* orthologs, but that *C. angaria* lacks a *dsh-2* ortholog, suggesting that only members of the *Elegans* supergroup (*C. japonica*, *C. elegans*, *C. brenneri*, *C. remanei*, and *C. briggsae*) have *dsh-2* orthologs. Furthermore, this clustering analysis revealed the possibility of species-specific expansions. For example, *C. angaria* appeared to have two potential *dsh-1* orthologs, *C. japonica* appeared to have two *dsh-1* orthologs and two *dsh-2* orthologs, and *C. brenneri* appeared to have two orthologs each of *dsh-1*, *dsh-2*, and *mig-5*. Detailed protein analyses of this kind rely on the quality of the assemblies and gene predictions of the proteomes used, and the results can often improve the annotations. Despite valiant efforts to inbreed these nematodes prior to genomic sequencing, the current assemblies of *C. brenneri*, *C. remanei*, and *C. japonica* (WormBase release WS225) are known to have considerable heterozygosity, with some genes being represented by allelic variants (Barrière et al., 2009). Furthermore, the genome assembly for *C. angaria* is still quite fragmented (Mortazavi et al., 2010), although additional sequencing is ongoing. We explore these genes in more detail in the following section.

The gene relationships among the *Caenorhabditis* Dsh paralogs is consistent with those recovered using a broader sampling of animal taxa (Figure 4 and Supporting Information): we recapitulate three clades, one for each of the three paralogs with *dsh-1* and *dsh-2* being more closely related, supporting the notion that *dsh-2* is the result of a fairly recent duplication event and has subsequently diverged from *dsh-1*. This

duplication could have occurred after the split between the *Elegans* and *Drosophilae* supergroups, or may have occurred earlier and been subsequently lost in *C. angaria*. More could be inferred about the evolution of Dsh among caenorhabditids from sequencing additional taxa from this genus.

Conservation and Diversification of Dsh Domain Architecture

Next we wanted to assess the protein domains in Dsh and evaluate the conservation of domain structure across animal evolution among orthologs and paralogs. The SMART database recognizes the DIX, DSV, PDZ, and DEP domains as being around 80, 72, 80, and 75 amino acids respectively, with some variation between species, particularly in the DIX and DSV domains (Figures 2, 5, and 6). The basic region located between DIX and PDZ, and the proline-rich region containing an SH3 binding domain located between PDZ and DEP are not recognized by SMART, but they were identifiable by sequence alignment similarity with known sequences (Penton et al., 2002).

We found the PDZ and DEP domains to be the most highly conserved structural components of Dsh across taxa, being present in all taxa from the sponge *A. queenslandica* to mammals (Figure 5 and Supplemental Figure 1). The basic region, just anterior to the PDZ domain, is also highly conserved and only absent from the *Bxyl-dsh-1* ortholog, which is truncated. The proline-rich region extends over an approximately 20 amino acid window and contains a class I core SH3 binding motif RxEPV/IR/QP (where x stands for any amino acid), with ligand preference varying around the PxxP core (Penton et al., 2002). Although the proline-rich region is not always conserved, the SH3 binding domain is conserved in the *dsh-1* orthologs of all taxa, but is absent in in

nematode *mig-5* orthologs (Figures 5 and 6 and Supplemental Figures 2 and 3). We refer to this region as the SH3 binding motif rather than the proline-rich region due to the conservation of the motif across taxa, although the area surrounding the motif is not necessarily proline-rich in some nematode taxa. The DIX domain is conserved in all Dsh orthologs but is conspicuously missing from two nematode *dsh-1* orthologs; *Ppac-dsh-1* and *Asuu-dsh-1* (Figure 5). The understudied DSV domain appears to have experienced dynamic evolution, being absent from both sponge and jellyfish taxa and arising in bilaterian taxa (Figures 1 and 5). The DSV domain is conserved in planaria, vertebrates, and two of the three insect taxa we investigated (*Dmel-dsh* and *Nvit-dsh*) but is missing from *Tcas-dsh* and is absent from all nematode Dsh homologs except *Tspi-dsh-1*, the only Dsh homolog in the most basal nematode lineage included in our analysis (Holterman et al., 2006; Figure 3). The Dsh-C domain is vertebrate specific, but appears to be truncated in *Xtrop-Dvl1* (Figure 5). We found a previously unreported DEP-like fragment (DLF) domain, recognized by the SMART database, and is present and conserved in several nematode species from clades 8, 9, and at least one species, *S. carpocapsae*, from nematode clade 10 (Figures 3 and 5). The amino acid sequence conservation and codon variation that we detect both in the DSV and DLF domains suggest that these are functionally relevant, despite the current lack of functional data (Supplemental Figures 2 and 3). The absence of a recognizable DSV domain in early branching lineages (i.e. *A. queenslandica* and *C. hemisphaerica*), and its apparent loss in *T. castaneum* and all evaluated nematode lineages branching after clade 2 suggest that its conservation among some insects, planarians, and vertebrates has functional significance and should be tested.

Similarly, the conservation of DLF among clades 8, 9, and at least one clade 10 nematode (*S. carpocapsae*), suggest that it too has functional significance.

The domain architecture of Dsh orthologs within *Caenorhabditis* genus is more dynamic than that observed across a broader sampling of animals, likely facilitated by the presence of three Dsh orthologs (Figures 5 and 6). The PDZ and DEP domains are the most highly conserved across caenorhabditid orthologs, with both only being absent from *Cjap- dsh-1a* and *Cang-dsh-1b*. The DEP domain is missing from *Cang-mig-5*, and PDZ and DEP are separated between the *Cjap-dsh-2a* and *Cjap-dsh-2b* (Figure 6 and Supplemental Figure 3). The basic region is also highly conserved and is present in all orthologs with protein sequence N-terminal to the PDZ domain (Figure 6). The SH3 binding motif is conserved in *dsh-1* and *dsh-2* orthologs that contain protein sequence C-terminal to the PDZ and/or N-terminal to the DEP domain, but is entirely absent from all *mig-5* orthologs (Figure 6 and Supplemental Figure 3). The newly discovered DLF domain, where present, is between the PDZ and DEP domains, just C-terminal to the SH3 binding motif. The DLF domain is only present in nematode *dsh-1* orthologs that have PDZ and DEP domains except *Cbre-dsh-1a*, where it is conspicuously missing.

We investigated the splice isoforms of all three Dsh paralogs within *Caenorhabditis* and used *B. malayi* and *P. pacificus* for outgroup comparison (Figure 7). There is conserved isoform architecture among these species for all three paralogs, although no species has been as thoroughly studied as *C. elegans*, which has 3 isoforms of *Cele-dsh-1* as well as *Cele-mig-5* (Figure 7). For example, all *dsh-1* isoforms that have a DIX domain have it split across three exons, whereas the PDZ domain appears to be split across two exons in all caenorhabditid taxa except *C. japonica*, where it might be split

across three. In addition to partitioning domains among proteins, as *P. pacificus* seems to have done with the DIX domain being present in *Ppac-dsh-1* and absent from *Ppac-mig-5*, other taxa can produce isoforms with and without certain domains (e.g. *Cele-dsh-1* and *Bmal-dsh-1*; Figure 7). Too little is known about splice isoforms in the other species to draw strong conclusions from these data, but interesting features of conservation and divergence are apparent. Additionally, this analysis sheds light on the potential paralogs identified within *C. japonica* and *C. brenneri*. *Cjap-dsh-1a* and *Cjap-dsh-1b* are tandem in the same orientation, with *Cjap-dsh-1a* being <3 kb upstream from *Cjap-dsh-1b*, suggesting that these are fragments of the same gene (Figures 6 and 7). However, *Cjap-dsh-2a* and *Cjap-dsh-2b*, although still in the same orientation, are >10 kb apart with *Cjap-dsh-2b*, which has the DEP domain, being upstream of *Cjap-dsh-2b*, inverting the traditional order of DIX, PDZ, and then DEP, suggesting that these might actually be separate genes, representing a physical partitioning of DIX, PDZ, and the basic region on one protein and the SH3 binding motif and DEP on the other (Figures 6 and 7). The potential paralogs within *C. brenneri* (*Cbre-dsh-1a*, *Cbre-dsh-1b*, *Cbre-dsh-2a*, *Cbre-dsh-2b*, *Cbre-mig-5a*, and *Cbre-mig-5b*) were each on separate contigs, offering no potential clarification. However, a nucleotide alignment of the PDZ and DEP domains revealed that each paralogous pair has identical nucleotide sequence, suggesting that these are likely allelic variants or splice isoforms (Barrière et al., 2009). Splice isoforms seem particularly likely in cases in which a paralagous pair differ in domain content (*e.g. Cbre-dsh-1a*, *Cbre-dsh-1b*, and *Cbre-mig-5a*, and *Cbre-mig-5b*).

The amino acid sequence alignments of Dsh across animals and across caenorhabditids show clear regions of high conservation and other regions with

considerable divergence. Across animals, we detect at least two codons that are experience strong negative selection (codons 318 and 335) and at least 10 codons that are experiencing diversifying selection (codons 6, 221, 222, 234, 235, 238, 239, 252, 284, and 354; Supplemental Figure 2). Focusing on caenorhabditids, we detect at least three codons experiencing negative selection (codons 113, 252, and 261) and at least 4 codons that are experiencing diversifying selection (codons 63, 85, 155, and 193; Supplemental Figure 3). It is not surprising that areas of functional significance are highly conserved across species, whereas those regions that show considerable divergence or are experiencing diversifying selection may play important roles in the acquisition of novel functions but remain to be functionally tested.

Nuclear Transport

In addition to the conserved elements of Dsh shown in Figure 2, there are other motifs, structural components, and phosphorylation sites that affect the function of Dsh. For example, the presence of a nuclear export signal (NES) and a nuclear localization signal (NLS) affect the subcellular distribution of Dsh. A conserved NES has been identified as M/LxxLxL, where mutations in the leucines lead to nuclear localization of Dsh in *Xenopus* (Itoh et al., 2005). We found this NES to have patchy conservation, being present in *Aque-Dvl*, *Cint-Dvl,* and all vertebrate Dsh orthologs, except *Xtrop-Dvl1* (Supplemental Figure 2). It was not present in *Chem-Dvl* or any insect or nematode Dsh orthologs (Supplemental Figure 2); however, it was present in *Smed-dvl-2* but absent from *Smed-dvl-1*. Previous studies indicate that Dsh translocates to the nucleus and is actively exported into the cytoplasm, presumably via NLS and NES signals, and that

blocking the nuclear export by mutating the NES or chemically inhibiting nuclear export leads to nuclear localization of Dsh in vertebrates (Torres and Nelson, 2000; Itoh et al., 2005). A NLS sequence was previously identified in vertebrates, flies, and *Hydra*, and identified as IxLT/VAK (Itoh et al., 2005). We found this NLS to be highly conserved across the taxa in our analyses, being present in the Dsh orthologs of most taxa we examined, but absent in all nematode *mig-5* orthologs, and identifiable yet slightly altered in *dsh-2* orthologs (Supplemental Figures 2 and 3). This NLS has been shown to be necessary and sufficient for nuclear translocation of Dsh in vertebrates, although this has not been pursued in invertebrate taxa (Torres and Nelson, 2000; Itoh et al., 2005).

Phosphorylation of Tyrosine473

The phosphorylation of tyrosine 473 (Y473), located in the DEP domain, is essential for PCP signaling (Yanfeng et al., 2011). The substitution of *Dmel-dsh* Y473 to phenylalanine (Dsh^{Y473F}) leads to strong PCP specific defects in *Drosophila*, but has no effect on canonical Wnt signaling. It is believed that this site in the DEP domain is phosphorylated by an Abelson family tyrosine kinase (Abl), which is also required for PCP signaling, but not canonical Wnt signaling (Singh et al., 2010). We find that Y473 is conserved across all evaluated Dsh orthologs except *Mhap-mig-5* and *Minc-mig-5* as well as *Smed-dvl-1*. All evaluated organisms have at least one Dsh with a conserved Y473 implying an ancient and essential function across Metazoa, and suggesting another potential mechanism for partitioning the function of Dsh paralogs and/or splice isoforms.

DISCUSSION

The origin of Dsh lies in the common ancestor of Metazoa and likely had the three major functional domains DIX, PDZ, and DEP (Figure 8A). Dsh has experienced dynamic evolution across animal evolution, acquiring new domains and experiencing duplications in several animal lineages. The DSV domain seems to have evolved prior to the bilaterian split and been subsequently lost in some nematode and insect taxa. In no phylum where multiple taxa were examined did we find complete conservation of both domain architecture and number of Dsh orthologs. We have identified many structural features that are conserved and divergent and that suggests potential mechanisms for partitioning the various functions of Dsh among isoforms and/or paralogs. We discuss these findings in the context of known functional specializations among invertebrates.

Dishevelled Across Nematodes

Nematoda is an ancient lineage, originating during the Precambrian or Cambrian explosion over 500 million years ago (Ayala et al., 1998; Rodriguez-Trelles et al., 2002). With this abundance of evolutionary time, nematodes have evolved to inhabit virtually every habitat known and nearly every ecological niche. The model nematode *C. elegans* was the first metazoan to have its genome sequenced and is among the most studied and best understood animal on earth (Consortium, 1998). Often what is learned about *C. elegans* is assumed to be conserved among nematodes, and although this may be largely true for some features, *e.g.*, neuroanatomy and CO_2 detection and response (Bumbarger et al., 2007; Hallem and Sternberg, 2008; Bumbarger et al., 2009; Ragsdale et al., 2009; Hallem, E.A. et al., 2011; Hallem, E. A. et al., 2011), *C. elegans* is a derived nematode

with many unique features (Blaxter, 1998; Blaxter, 2011). We have shown that the number of Dsh homologs varies across nematodes, at least from one to three, but many taxa remain unstudied, especially within the basal clades of the phylum (Figure 8B). Most genera in our study have two Dsh homologs, *dsh-1* and *mig-5*. The acquisition of *mig-5* is ancient, occurring sometime after the split of clade 2 and before the split of clade 8, although additional taxon sampling would improve this estimate (Figure 8B). The *C. elegans* genome encodes three Dsh genes, *Cele-dsh-1*, *Cele-dsh-2*, and *Cele-mig-5*. We have shown that *dsh-2* is likely a paralog of *dsh-1* and a derived character among *Caenorhabditis* species, perhaps only among members of the *Elegans* supergroup (Figure 8C). We identify only one Dsh ortholog in *T. spiralis*, *Tspi-dsh-1*, and find that among nematodes, it has unique similarity to insect Dsh as it is the only nematode Dsh known to have a DSV domain (Figures 5 and 6).

The domain architecture among nematode Dshs is variable and suggests potential mechanisms of functional divergence. We have discovered a novel Dep-like fragment domain that is present and highly conserved in half of the 10 nematode taxa we examined (Figures 5, 6, and 8B). The domain architecture of *mig-5* is conserved, having the same structural features (DIX, PDZ, DEP, and the basic region) in all taxa (except *Cang-mig-5*, which is missing DEP), while *dsh-1* orthologs are more diverse (Figures 5 and 6). *Asuu-dsh-1* and *Ppac-dsh-1* lack the DIX domain, *Bxyl-dsh-1* lacks the DIX domain and the basic region while *Scar-dsh-1* seems to have acquired a signal peptide and a coiled domain that is unknown in any other Dsh homologs. Finally, we detected a conserved NLS in all nematode *dsh-1* orthologs (and both *Cbre-dsh-2* orthologs), suggesting that these proteins may be translocated to the nucleus, as has been shown in vertebrates. It is

worth noting that the presence of an NLS and a basic region, features that are broadly conserved in Dsh orthologs across animals, are hallmarks of transcription factors, although this possibility has not been experimentally explored (Grove et al., 2009).

Although there are many examples in *C. elegans* of the functional overlap of Dsh paralogs, there are also known specializations for each. For example, B cell polarity in males is controlled by Wnt signaling, where *Cele-mig-5* males have altered B cell daughter size (Herman et al., 1995; Sawa et al., 1996; Wu and Herman, 2006). Neither *Cele-dsh-1* nor *Cele-dsh-2* affects the polarity of the B cell as single mutants and neither enhances the phenotype of the *Cele-mig-5* mutant, showing specialization of *Cele-mig-5* in this pathway (Wu and Herman, 2006). The divergence of Dsh function in *C.* elegans can also be seen in the outgrowth of neurites from RME head motor neurons. In this pathway, Cele-DSH-1 physically interacts with Ror/CAM-1 to transmit the Wnt/CWN-2 signal to downstream components enabling neurite outgrowth (Song et al., 2010). The binding activity of Cele-DSH-1 to Ror/CAM-1 lies in its PDZ and DEP domains, whereas the DIX domain is not required for binding. Furthermore, only *Cele-dsh-1b,* the isoform that lacks the DIX domain (Figure 7; *Cele-dsh-1b* is the *Cele-dsh-1* 'b' isoform from WormBase), was shown to express in the RME cells and Cele-DSH-1b is sufficient to rescue the *dsh-1* null phenotype, suggesting that alternative splicing of Dsh can lead to functional specialization within *C. elegans* (Song et al., 2010). An example of domain specialization within a Dsh homolog can be seen in the asymmetric cell division of the ABpl/rpppa neuroblast via a β-catenin independent pathway (Hingwing et al., 2009). Domain analysis has shown that the DIX domain is not required for ABpl/rpppa asymmetric division, but the DEP domain is essential. Hingwing *et al.* (Hingwing et al.,

2009) go on to show that *Cele-dsh-2* is involved in the asymmetric divisions of SGP cells along the proximal-distal axis of the developing gonad, which leads to the formation of distal tip cells from distal daughters and an anchor or ventral uterine cell from the proximal daughters. Loss of *Cele-dsh-2* results in two proximal daughters. Unlike the asymmetric division of the ABpl/rpppa neuroblast, both the DIX and DEP domains are essential for proper SGP cell division, thus demonstrating the divergent functional roles of domains in a single Dsh ortholog (Hingwing et al., 2009).

We have identified and shown the conservation of a Dep-like fragment domain across all *Caenorhabiditd dsh-1* orthologs along with *Asuu-dsh-1, Bmal-dsh-1*, *Ppac-dsh-1*, and *Scar-dsh-1*. Furthermore we have shown that the basic region, DIX, PDZ, DEP, SH3 binding motif, and the NLS are conserved across nematode *dsh-1* orthologs (with a few exceptions lacking the DIX domain and the absence of the basic region in *Bxyl-dsh-1*). We have shown the extreme conservation of structure across all *mig-5* orthologs, and that these uniformly lack the SH3 binding motif as well as the NLS. The functional relevance of these features and what role, if any, they play in the partitioning of Dsh function would be interesting to explore. These results suggest, for example, that *Ppac-dsh-1* and *Ppac-mig-5* might have evolved to function in separate pathways and perform at least some non-overlapping functions (Figure 5). The apparent lack of an NES in any nematode or insect Dsh is also striking, especially considering the presence of an NLS among most *dsh-1* orthologs. Perhaps nematodes and insects have an alternative and as yet unidentified NES, since these proteins are not reported to be nuclear-specific.

Dishevelled in Other Invertebrates

S. mediterranea has two orthologs of Dsh, *Smed-dvl-1* and *Smed-dvl-2* (Gurley et al., 2008). Initial studies in this flatworm have investigated the functional specialization of these paralogs. Only *Smed-dvl-2* appears to be involved in β-catenin-dependent signaling: phenotypes described after silencing canonical Wnt ligands are reproduced upon the silencing of *Smed-dvl-2* (Almuedo-Castillo et al., 2011). Conversely, both *Smed-dvl-1 and Smed-dvl-2* transduce the noncanonical signals that control neural connectivity as well as mediolateral patterning of the central nervous system, neither of which involves components of the PCP pathway. Components of the PCP pathway, including Van Gogh and Diversin, have been implicated in the apical positioning of the basal body in epithelial cells. Interestingly, only *Smed-dvl-2* has been shown to function alongside these core PCP components. Our domain analysis of planarian Dsh supports these experimental results. Only *Smed-dvl-2* contains the NLS and NES sequences, implying its role in β-catenin-dependent signaling. Because *Smed-dvl-1* lacks both sequences, we would suggest that it cannot function in a β-catenin dependent pathway, and this hypothesis is supported experimentally (Almuedo-Castillo et al., 2011). Furthermore, it has been shown that tyrosine473 is essential for PCP signaling. This amino acid is present in *Smed-dvl-2*, but not *Smed-dvl-1*, supporting the experimental finding that only *Smed-dvl-2* can function in the PCP pathway (Almuedo-Castillo et al., 2011).

Insects have only one copy of Dsh, at least the taxa that have been investigated so far. Significant effort has gone into understanding how *Drosophila*, with one Dsh ortholog, channels a Wnt signal into distinct pathways. It has been shown that specificity is achieved by the presence or absence of binding partners as well as the subcellular

localization of Dsh (Wallingford and Habas, 2005). Other work has shown that qualitatively different Fz-Dsh interactions underlie PCP and canonical Wnt signaling (Strutt et al., 2012).

The insect proteins, *Dmel-dsh*, *Tcas-dsh*, and *Nvit-dsh* are very similar in architecture. All have a DIX, PDZ, and DEP domain as well as the basic region and SH3 binding motif. *T. castaneum* is the only one that lacks a DSV domain, but this suggests that additional taxon sampling could reveal a broader trend. Interestingly, all insect Dsh proteins have a NLS, but none have the known NES that has been shown in *Xenopus*. It is currently not known whether invertebrate Dsh translocates to the nucleus, but if Dsh does, it must employ a different export signal than the one found in *Xenopus*. More work must be done to better understand the localization and transport in invertebrates.

CONCLUSIONS

We have discussed the origin and evolution of Dsh in a variety of metazoan lineages, emphasizing a recurring theme of Dsh duplication and expansion in many phyla. The data we have evaluated suggest that Dsh arose in the most recent common ancestor of Metazoa and possessed many of the structural features that have come to characterize Dsh (Figure 2). Most basal lineages within explored phyla appear to have only a single Dsh ortholog, leading us to conclude that the ancestral state of Wnt signaling pathways was built using a single Dsh protein acting as the hub, and has then experienced lineage-specific expansions in many phyla. The deuterostome taxa wherein Dsh has been explored reveal that early branching deuterostome phyla (*e.g.*, Echinodermata and Hemichordata) have only one Dsh ortholog, which is also true of basal chordate lineages like lancelets and sea squirts (Cephalochordata and Urochordata, respectively) (Gray et al., 2009). It is noteworthy that as more taxa in a particular phylum are explored, the derived lineages seem to have convergently evolved multiple Dsh orthologs, although there may be exceptions such as insects, where even the more recent lineages seem to us the ancestral strategy of partitioning Dsh function in ways other than protein duplication and subsequent divergence.

As the hub of Wnt signaling, Dsh plays an essential role in animal development and homeostasis. We have shown that Dsh has experienced dynamic evolution across Metazoa, including the acquisition and loss of domains as well as gene duplication in many lineages. Our findings on the divergent and varied architecture of Dsh across taxa provide testable hypotheses about the means of these specializations. The dynamic evolution of Dsh among nematodes both by paralogous duplication and the formation of

lineage-specific splice isoforms raises questions of protein evolution and provides clues as to how these organisms have dealt with the complex intricacies of having a protein, like Dsh, act as a central messenger hub connected to so many different and vitally important pathways.

ACKNOWLEDGMENTS

We thank Byron J. Adams, Geoffrey T. Smith, and Christian A. Grove for their critical reading of the manuscript and Art Vandelay for helpful insights into the details of importing and exporting. This work was supported by a National Institutes of Health (NIH) United States Public Health Service Training Grant (T32GM07616) to A.R.D. and P.J.M., and by the Howard Hughes Medical Institute (with which P.W.S. is an investigator).

SUPPORTING INFORMATION

All Dsh proteins identified and used in this analysis

The supporting information included all of the protein sequences used in this analysis and has the various protein domains, as identified using the SMART database (http://smart.embl-heidelberg.de/) highlighted, according the the key below.

Domain Highlight Key:
Green – DIX
Yellow – PDZ
Pink – DEP
Blue – Basic Region
Red lettering – SH3 binding domain in the proline-rich region
Dark Green – DEP-like fragment
Red – Dishevelled specific domain
Teal – Dishevelled C domain (mammalian specific)
Black & White – Signal sequence
Maroon – Coil domain
Grey – internal repeat

>Cele-DSH-1
MAESPPPVDSSLNAPNVGSPTTMMERLRLRDQTEENGKEDDFDNKSVSSAQYSQ
TSEATTAVKQQPFLHTMTKVYCHIDDETDPYMLEVHVPPDLITLGDLKRVLMRT
NFKYYRKALDPDSGYEVKAEIRDDSQRLTPSPNNLFELFLLTIEGSTHSDGSSGK
MRKYPSVPGPAPSNRNGPPMNYQHAAYQFDNSMMSTDSESMISAAIPGYLKSAA
YNRRFPQHYLGHRRHLEESTIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKMA
KKKKNKRNFRKPSRASSFSSITESSMSLDVITVNLNMDTVNFLGISIVGQTSNCGD
NGIYVANIMKGGAVALDGRIEAGDMILQVNETSFENFTNDQAVDVLREAVSRRG
PIKLTVAKSFENGQSCFTIPRNSREEPVRPIDTQAWIQHTNAMRGMPSIVEESAPTP
IPGEWPHGRPPSSSTVTSNGSNGQNTVVGGGAHIILDIHTDKKKVVEIMAMPGSG
LDIKNRTWLKIPIPMSFLGSDLVEWLLDHVEGLRERKTARNFAADLLKLKYIAHV
VNKVTFTEQCYYVLGDECSDYARFRNEDGGPKYQWTIGMNGMSAGNGSSVML
PPPHLPGGMAGPPGAFKGMAPSMVSGYASMPASPFPPAQLQQQRREGSTTSGSS
GGGIRKQRVVVLPRKPSSSANVPFDDSSTIYEESNNSFLMATGQRYEY

>Cele-DSH-2
MTDSPSPIDSSFDASDVATPCTVIAAKISLRNRNGLEEDQENLDSFDAFTETHETQ
ESKNIAHGEHEEDVSNIYVDDFSKEFGDTVSSVMEPLPKPLTFARTITKVYCHLD
DQEHPYMVEVHVPPDCITLRDVKRKLMRTNFKYYCIALDPDTGLEVKAEVRDD
SRRLYPLKNGRFELYLLTVEGSVHSDTSSGRHRRKQDGSSKGSSGSREYLRAAH
HYDNPTPFSDDESQASSLPTYVKKAHAYNRKHAPQAYERHLPHMKHNNRHNHH
RQNHYEESTFDVTTESDDHYRDGVTYYDEDEDDSRSINTDLTSVSQVHLKQKW

RQQQKEMRNKWKRMPSISTASSSFSSITESSMGLEVITVRLNLETIPLGMTPSGHT
NARGDAGLYVGDIQDRGAVALDGRIDIGDMIVGINEISLGNYSNKEAVQLLREA
VQRQYLTLTIAKTGDPKQNAFPRNPRAEPIRPIDPNEWVKHATNAMKAMPSISEE
SSSTPIPDDWPTNSSASGTPFGGPPPANCLNVMTDKKYVVEVMAAPGSGLDIKDR
YWFKIPIPMSFLGTDLVEWLVKHVQGLETKKKAREFAEEMLKLGYIRPGVGKQS
FTKECYYVMGDECADYTQLRGPDGGYKYPQSHASSASGHSSNNLIFPPSMYPPQ
PPTAGAVQSSKFGHSTFFNDWRIRITTGVSISSRKGICQ

>Cele-MIG-5
MEPPCTSDCSQIKVFYYLDDETTPYVSVIEAREGVATLGNFKNSFTKRGYKYYA
KELDPDIQREVKVELTTDSDRLRKSQNGFYEIFLVSTPGYGTLPRNSGTMTRPQR
TALDKRRRRSADFDATPYSDASLAPSTIVSRRAGEHLAELYTSNSEDPYQYDEHT
RRTGDDSSLYEPLAARDMNKIYDDDRRRKKQKKERFRRPYVPSTISSATESSVNS
GLPRILEIYLPMKNVPYLGLSVCTIDGHIFVSEIAPEGAVEKDGRVNVGDQILQVN
RVSFEELSGPQAVRSLREAASSKRPITLYISKFARGAPSEYDDPLASMASETMPLD
VGVWVETAVQNTEKMKALGLDPQEQTATTIDDGTLPFTSTASDDEERMLYDQR
RNGIPRALIEEAERKRENEQNEKIEQLTEMIDPIIVVRSMARPDSGLAVKNRKWL
KILVPMSFIGRDLVDWLVDHMADIHNRKKARIYAARLLAAGLIRHVVSKLTFTE
KCYYVFGDGILGNDRNSTDTTGTSGTTMRVEATTEVTYVGSPAPHALAARVGR
NIPHRLETTTLSPVAHDQTWLRRRRDCESPMTNDYASMVGESQIGMNPVGNYH
VFGTKNNHRQVPAPSQVTSSSLTNGSGGLGGPPPTPLSSTMVLAASPIQSQNAVN
HDFDGENSSNSRTRILRT

>Bmal-DSH-1a
MDVSDTLDTSNTTDTIIEKTNKLKISSAVESKDESSKNSSCTSKAPLAKGTSCLEH
GIGVPSQQTSTKVYYHIDDEMVPYCTDVMVPPDKITLGDFKRVLTRSNFKYYCK
APAPDSGVFPEVKVEIRDDNECLHRSANGQFELFLLTSEGSSHSDGSSGLPLKSAR
LMFSKIIFRPIIVHCEKNFLVASTDSDSFISDMRALPMQVKGISRRPFPQQYISQGH
RGGRRFEDSTLGSESDARLFSDDDDRSRVSTSTDITSVSRQHHAPAYRKRRNRRR
FRQPSRASSFSSITESSMSLDVITVTLNMDTVNFLGISIVGQSSSRGDNGIYVANIM
KGGAVALDGRIEPGDMILQVNDISFENFTNDQAVDVLRESVARRGPIKLTVAKM
WDSGPRSAFTVPRHRDEPVRPIDTQAWIQHTNAMRGMPSILEGSEGAPTPIPGQY
GRPASSSTATSNGSVPNTIVGGAHFRLDAMTDKKKVVQMMVMPNSGLDIKNRT
WLKIPIPMSFLGSDLVDWLMEHVDGLRDRKDGRKFAGELLKEKLISHVVNKITF
TEQCYYILGEECADYARLRQNPGGDDPGVRSEVGSVLPPPPPGLVAAAAAQQSG
RAWPQPTMIPQSAPSMVSGIENPANADGTTRFYLTNL

>Bmal-DSH-1b [SPLICE ISOFORM]
MLFSVPGPAPTMYPVGSMAYRQAVQQFDQSMASTDSDSFISDMRALPMQVKGI
SRRPFPQQYISQGHRGGRRFEDSTLGSESDARLFSDDDDRSRVSTSTDITSVSRQH
HAPAYRKRRNRRRFRQPSRASSFSSITESSMSLDVITVTLNMDTVNFLGISIVGQSS
SRGDNGIYVANIMKGGAVALDGRIEPGDMILQVNDISFENFTNDQAVDVLRESV
ARRGPIKLTVAKMWDSGPRSAFTVPRHRDEPVRPIDTQAWIQHTNAMRGMPSIL
EGSEGAPTPIPGQYGRPASSSTATSNGSVPNTIVGGAHFRLDAMTDKKKVVQMM
VMPNSGLDIKNRTWLKIPIPMSFLGSDLVDWLMEHVDGLRDRKDGRKFAGELL

KEKLISHVVNKITFTEQCYYILGEECADYARLRQNPGGDDPGVRSEVGSVLPPPPP
GLVAAAAAQQSGRAWPQPTMIPQSAPSMVSGIENPANADGTTRFYLTNL

>Bmal-MIG-5
MREEAAAISTKVYYYLDDSTPYLSVVPVADDAITLGDFKKVFNKKGYKYFCKQL
DEAVGCEVKVEIRDDSTKLVKSANGLIELVLLSSSDNVHCSGTLPRVSNKTNKGQ
LTGAKLKDFNLRKRRSLHDLASENRDLLRIHRNKSDENSITASSLSTVISKRAGEG
LAELYASNSEDPYHLEDVNSRYFKHPGACXPVFPASPLASPSGVLPCPRRQRRPR
KERYRKAYVPSTISSVTESSMTSLSLPRIDVITLPMKNGVFLGISVLSHDGGIFVSD
VHSGGIVDLDGRIEVGDQIVQVNRSSFENLSDVEAVDLLRKAAASRKPITLYVAK
RTCNNSDKRADILSGIASETMPIDISLWVESTKHNIVRPPKGLEEMVSVNDGDAT
LVAEEAETDLEGAYAERRNGHIPSIQNCVKLKQLNPPDLNTSLNIEDIARRRENEE
NEQQLDNLNVDMDPVIILKYMALPSSGLQIKNRKWLKIPVPMSFIGCDLVDWLM
EHVHGITDRKAARIYASKLLAEGHIRHVVNKLTFTEKCYYIFEDSILSVRNKNKS
DSSLGKAGAEVTTEVTYVGSPAPAHLSRTSARNTLGGKAIFDQSWPHLTITSSEQ
RKSFCGSSTNDYASVMGPDMIDSTLLTEAPTLKLSHRTLPNRLDMEQRINGCEVA
QPPNTPNSLLHEQRNADSETEFETVEDNKFLVIQK

>Bxyl-DSH-1
MSSMTETSMSLQVMHVRLNMDTVKFLGITVVGQSSARGDNGIYVAHVMPGGA
VALDGRIEIGDMILEVNDVSLEKMTNDEAVEFLREAVTTKGPIKLTVAKCVDSNR
ANFLVSSREPVRPIDTRAWVQHTNAMIGMPMNSIPESAEEAPTPIPGQYPQNNSH
PVFCRPQSSSTATSNGSGGPKNTVVGIPGAYFALPPRLDLSTDKKIVAKAMAMPN
SGLEVRNRTWLKIPVPMSFLGSALLDWIHEHVEGIRDRKEARKYASELLKDRLIA
HVVNKSSFTEQCYYVFGEDCQEILKLRNEDGTPRTDLMPQPPMHPPKPIPGHSAF
GWTQARSTGDYASMPVSPYPGNGPFIPTNSLNQPLNKHGDIHSQASGNSNDGSSS
SDQRRKPLLPAAPPLPMGLNPLYQQGPQPPPSQPPPLPPFEARDPSSLKLDDLGSN
QQLQALVSQGFTVDHL

>Bxyl-MIG-5
MASTKVYSYLDDDNTPFLTSVPVSASQITLGDFKKALPKRNLTYSQKVFDENVK
KHVKRTIHDDSEPLQLNENGVVELFLTSNGPIGTLKTNGRHYKNLPPHLANYGV
GDLDVCGGQRVSMVTAEPSISGAGSYLSKRAGEQLASIDNTSASEDPYVFEQSSI
ANQSSDYLRVMNTTDYSRRRKKRERYRRPYVPSTISSASGMSSVPQVTEIMLDLR
EHALGIDIGMCDGAVLITSIAKNSAADRCGALGVGDQIVQVDNTSFEELSDEQVV
SLLRKISSQKRVVRIVVARYKRDSDQRSDALSALCETAEIDVSLWVEGTKQANPD
APFDDFHQNPNTSAHINEAAEETSDEERAAYDDRRNGIGAHFVPQIKLMRAIHQR
TEANGHIPNDENDRHSRLSATDAMEIVVRQMARLDSGLKIKDRKWLKIPIPMSFI
GHELVNWLLENVDGLDNRKQARNYAKQLLEKGFIKHAVNMNSFSEKCYYKFQ
EKICEERLLMEQQAKLNTVDSPTEITYMSGPSSPHQPHRQIPNNYHTGPLDVLQQ
RQINYSGTQPPPPRINPQMTSKSQPALPNTLPMNSTQVQRPMNSTAVDLSKWEFS
PIPARNYRDCESTSGIYDVVKQ

>Tspi-DSH
MEIRKDSSRTSVIGKGNYNGENDKKVVKEEISDDAVKLPTFNGCVESWLVTSEG
STHSDANGPTTQHGHSSPRHPPTRSSGLGESRPPSFHGGRADSRDNLVPSGCSSET

ESTLSGLPKFPRYGTKLDRRSFHPQFHERVRYPKGMRARCFDVGHGLTSDLDSTS
FMDSEDDQASRISSSTDITSVSRQYLKRKKRRRLPRTLLSRASSVSSITESSMSLNII
TVTLNMDTVNFLGISIVGHSNQLGGDGGIYVGSIMKGSARGAVALDGRIEPGDMI
LQVNDISFESMSNDDAVRVLREAVQKPGPIKLVVAKCWDPNPKGYFTIPRTEPVR
PIDPGAWVAHTNALRAEMPLDYPGPLSVNTGSFTPVPEAEKFLEECNLDVNVHDI
LTIVRAMAKPESGLEIRDRTWLKITIPNAFLGSDVVEWLYTNVQGFYDRRDARK
YAARMLKEGYIKHTVNKITFAEQCYYVFGDICDNFAGLRLEPHREEPPQLEHDSI
SALPPPPMTPSTLWSSNVGHPYYPHQACPSFVQTSSVASGYVPIPYQYNNEAGSF
PSLVVGGAPPTSRSDANSAGSAGSNGSNESERVRKGVVGGLLPMTSTPGVPLDP
NIQLKGYGPPVGAMTSKGMIGPLSPTSRNVNDLQSDISGSRQSFRMAMGNPCEFF
VDVM

>Asuu-DSH-1
MGEGIGMSLLLDCPISQAACYLNFVKHPNVSMYMKLFVEVTSSLQMPTDDLVD
VDENRRVVKLSQSFELKGRKSGRSRSLPHTDKKGCRCHRAGRRFEDSTIGSESDA
RLFSDDDDRSRVSTSTDITSVSRQHQANAYRKRRNRRKFRQPSRASSFSSITESSM
SLDVITVTLNMDTVNFLGISIVGQSSSRGDNGIYVANIMKGGAVALDGRIEPGDM
ILQVNDISFENFTNDQAVDVLRESVARRGPIKLTVAKMWDGGPRSAFTVPRHRD
EPVRPIDTQAWIQHTNAMRGMPSILEGSEGAPTPVPGQYGRPPSSSTMTSNGSAP
NTVVGGTHIRLDTTTDKKKIVHMMVLPNSGLDIKNRTWLKIPIPMSFLGSDLVD
WLMEHVDGLRDRKDGRKFAGELLKEKLISHVVNKITFTEQCYYILGDECAEFAR
LRQNPSAGDEQGVRSDVGSVLPPPPPGLMAAAAAQQTATRAWPQTTLLQNAPPS
MVSGYASMPISPYPGQPPTQPFAPSIMHGVGNPVSIYKGGCAPDVHSQTSSNDGS
SGSEHIRASQLAHSIIEKGTACPHLRLRDTDVVIYQDNVVTQSHGTRARSSLLEIR
FGKSEQAAHVRIKCGLYFPTCCGLCFPFYRQKYAGAAFTRERKCEFL

>Asuu-MIG-5
MPLDENAAVASTKVYYYLDDNTPYLSVIPVPENKVTLGDFKKIFTRKGYKYFCK
QLDKAIGCEVKVEIRDDSSRLEKSANGLIELVLLSTVTPSGTLPRAITNKTEHNGV
GNRDPVVNFESDLKVRKRRSLHELADNVVVNADPLANGCRASMRNSNEDSITA
SSLSTVISSYSILERAGEGLAELYTSNSEDPYRFDGSNSRFTLNSEACSSAYGGIRM
AASSAASCLKAHRQRRPRKERYRKAYMPSTISSITESSMASLSLPRIEIVKLLMTN
GAFLGISVLSNDGGIFVSDIIKGGAVALDGRIEVGDQIVQVNKNSFENLTDAQAV
QLLRQAAVSRRPITLYVVKRPCNTDSRSDVLSGLASETLPIDISLWIESAKQNSVK
PLKPFAIDETNSIMVENTLGEEEHETDMEGAYAERHDMIQIPPTSNKGRTTQVVT
TKRGVMTTEDVARRRENEENEQLVDNLNVNMDPRIILKFMARPDSGLQIKNRK
WLKIPVPMSFIGRELVDWLLDHVHGLHDRKAARSFASKLLADGHIRHVVNKLTF
TEKCYYVFDDSILSVRSINHSDGSNGKTGAEATTEVTYVGSPAPQAAARLAVRN
CNDSGMPPPPINNKMPAEIDQTWPVSPITIYGPTQRRKDCDSPVTNDYASVIGPDV
VTSTMLGTSLTEAPTLKLPRHGFIGRGNEMRRRMDDEIVVAQPPNTPSSLSAAPNI
GVDLGADFDSLEQDRRNMLKENR

>Ppac-DSH-1
MSGYQVDTDDVSSTSGSGGGYYGYGGSTSGGSVYYGMSGATVRNRGTMTIREA
DEDGGERDDGEGSMSTDLTSVSRQHEKMLARRRKAQRTYHRRPSRASSFSSITES
SMSLHVETVTLNMDTVNFLGISIVGQSSARGDNGIYVANVMKGGAVALDGRIEP

GDMILQVNEVAFENFTNDQAVDVLRDAVERKGTIRLTVAKSIDSNQRSDEPVRPI
DTGAWIQHTNAMRGMPSILEGSEGAPTPLPGERLHHHQQQMQQRPATSSSATST
GSNGQNTVVGSAGLPGTMPPLPRLDVHTDKRRVVEVMTRPGSGLDIKNRTWLKI
SIPMSFLGSDLVEWLLEHVSGLRERKEARKYATELLRQRLIAHVVNKISFTEQCY
YVLGEQCADFARYRTVQVEEGGVNGGGGTQPTWQWGGGGPRPTGITPGGVGG
SMSTLPAESHLFFLVKCNFINCASQVSGYASMPMSPHPPLFGGASAPGVRTFGGE
GSQISSTDGSGSSDAHKRGSI

>Ppac-MIG-5
MQVHDIASILKKKEMGVESELMTNQLPLRIYESEGLVFEWMWLRSVPVMGLDS
NTYGKVFEDFDLHELKEMLTVNYTSWRKYVWMVLDGMHVLSFPLPIPPQELIIY
RLEDMKMKYTSIQSESPNNLLRYFELRSFDFRPADEELKGFIESELRNGSMAIDGV
DWREWLCESDVLNGMSDTVAEDRQLCWIEDVNGKEKKRLMAATEAMRVDTP
YLWKDVMELAMEKLQLMALPTMKFNVMKYGETKRQRILMKEEKSGEIPPRYF
KEIHCMSRENIVPPGVNETTKVFYYLNDEPTPYVEIIHVGAESVTLGHFKREVKV
ELIGDETKLTRNGENGLFELFLLSTGTQNGGGGGTLQRKTNGTLTRRVGKERND
RDIYNTHRVHSDDYDSSSRQDTLMSRRAGEILAESITSASENLYEYDDSRQYNAV
MGDEASRIDGPRRRKARKNRVRKSYVPSTVGSQPESRYGGIPGTSLSLPRILEVNL
QIGPNDLLGISVVSVEGSILISDVFPVGVVARDGRIDVGDQIVQVNTRSFENLSDQ
QAIMILRKVAAAKKPLTLYVAKRTMSTAESDPLCTLASETLPLDISLWVENAVHC
TERQRFGVDGSVDGTILSEGVGRAASICTEDEEEERMLYVQRRNGMGIRERGLE
QPPIHLHSAPPPRGNYSEGSGYTERLSTRINPHSLINIISQPNSGLTVKNRKWLKIPV
PQSFIGVELVDWLVQNVEDLGERKEARKYATHLLEKGLIKHVVNKRDFTEKCYY
VFNGE

>Scar-DSH-1
MYFSLAIWPTVATRGFVLVLVMFDLLSFHRGLCNVRVSSPVLLPLIVKPNMSDSA
VSNSTLEPIVAVTARMTIKDDSALVGPSSDVSDRQSDVSGSQSLQPIGGPESDQCV
ANANVPSSKQTTKVYYNIDDETTPYCTEVPVPPDRITLRDFKNAFNRQQNFKYY
CKRIDKELGKEVKSELRDDSQPIERSQNDIFELFLLTAEGSTHSDGSSGAHTRPSL
RGTHMVPAPAPSGAHFDAYNAQYMHGRNFDPSCKITSLQEKTDLYHFSGEHRLR
IDVRKRTPVLPAGQHEPKAAVLPASPSADSSCLRSFQSNEGYGKRPPCRLSLFAN
TLLLLNPNLAGLGCCSSPSSFYDFQSWLPLLIYCAAFDNLSFIPRSIIPHRKKIHMS
GADRNSAERIDRLRERAKRLQELITDRRRSCRYVSEKCERFDVSAEPETSLKANK
PTTAPEPMVKMREPKPYVKSRTRELVESLINILPREPRDVRRSESLAVIVPSSTSTT
APMVIRHSSVSASNSKRQKDSDAPPGKQEQYAVIKKKSKRLSLVEKLPFLSAYSE
RQRPKSTSVAFERDTESLLPNRSEPPEAESESKAPPKAESESKAPPNAEELVEPAE
KIHRRSLPNPVRYRPKSIYQQSQMLPSGSTPKTLTRAMTMVDSPRRLPPLIEVTAP
EPSEAGESAAQSQSEAAAEKTSRKPRRYTGSHTTSTYQGGRDAAPDASPRRHLIG
FLHRTSHLSLTSELSADASFYLIGHRGRGRFEESTIASESDARIFSDDDDQTRVSTS
TDITSVSRQQHHNIYRKRKNRRKYRQPSRASSFTSITESSMSLDVITVTLNMDTVN
FLGISIVGQSSNRGDNGIYVANIMKGGAVALDGRIEAGDMILQVNDVSFENFTND
QAVDVLRDAVARRGPIKLTVAKCWESAQKNCFPVPRPRDEPVRPIDTQAWIQHT
NAMRGMPSIMEGSEGAPTPVPGQYSQHQRPHSSSTVTSGGSGPNTVIGNGPNHIQ
LDSKMDKRKVVEAMAMPNSGLDIKDRTWLKILIPMCFLGTDLVNWLIEHVHGL
KDSKEAKKYAMELLQAKLIKHVVKKVTFTEQCYYELGEEFARHRLMTGECSSG

RSDVGSLPPPPPGFLQSQPQQAARWQVMPGSNGVPSAPSMVGNHAPSSHMVPPR
HGGVPSSSEYGIVSGGHPCARVTIENGTTSHFILDI

>Scar-MIG-5
MEVKVFYYLNEGDVPYVSVLNPASGAAPTLGDFKKVFNRAGFKYYCKELDTDI
GREVKVELTDDKAPLLRAGSGLIELFLVQQQQQNFNSGTLPRTGKQNHEAAAGA
AFEPGGRLKKRRSLYGLSSVDGFDIGGGQRVSMVTNEQSLAESSQGTVLSRRAG
EHLADMYATNSEDPYNMEDPSASSFSAASSAYGGIPIGNNGRLPQNRHRKPRKE
RYRKAYVPSTISSAAESSLTSQSLPRIDVIKVSMKNAITLGIKVVGHDGGIFVSLIL
PDGAASQDGRLEVGDQIVQINEESFENLNDQQAVSILKKASKSKRSVTLYVSKRP
RAHDDGSSDVLTGMTANETMPLNISSWVKSTMHRKVEKHVPFQSVVGESTLDP
SESMTIADETSDEEQAAYLDRRGGVGPRFVPALGMRNQIAEDVFMRQKENDEN
DVMIDTLSVNMDPRIILKVMAKPDSGLQIKNRKWLKILVPMSFIGSGLVDWLLM
HVQGLHDRKSAREYASQLLQEGLIRHVVNKMTFTEKCYYVFDESIMQMANRGS
RGGAHSSGGEVTTEVTYVGSPAPGTDKTQPRPLDSNLTVQNANVNATWPISPITL
YGNQTARRCESPAVTNDYASMIGTEFVQQQHQMPVMMPSEAPTLKIGTQSPSTR
SPFFPPPPPSLQRSMSPPPNTPNTLLSSVLLSGATGSSRLTHQQSSHM

>Pred-DSH-1a [SPLICE ISOFORM]
ARKKRAPRRYIHSHTTSTYDAAANEKVPDASERRKLIGFLHRTSHLSPLVPEGIDT
VTTVASVTNVVGNRRRYLEDSTVGTESDARVFSDDESRVSTSTDNITSVSRQHN
NYRVRQRRMRRHLRQPSRASSMSSMTETSLALEVMTVTLNMDSVNFLGISIVGQ
SSTGGDNGIYVANIMKGGAVALDGRIQPGDMILQVNDTSFENFSNDQAVEVLKE
AVVHRGPIKLTVAKSDTGRFDAFNVPSEPVRPIDMRAWIQHTNAANNLPHIPEGS
EGAPTPIPGQYPHNYGRAPSSSTATSNGSNGHTILGPNGQVFILPKKLDLSTDRKR
VVSVMALPNSGLDIKDRTWLKIPIPMSFLGSDLVEWLIDHVEGLCDRKEAKKYA
TDLLKEGYITHMVNGNKFSEQSYYMIGRECSDAVRLRLSEEGASRSDVASSLPAP
PPHLAPQPWTANSAMTSDYVPISPFPQQHQMFYRGHDMFSQVSGQSNGDSSGSE
ARRARPILPPAPSSVTNVPYNAGANFFRQQTARLGDFPKNLAASRQSFRIACGNQ
TSEEFFIDHL

>Pred-DSH-1b
TFCVAVFIGQPMLSHLMSESRDTSATTAHSSSVSVDDAVTAIDKLALTSNGCNGA
LPGPSSSDADAKSDAASSEARSVPRTNGSGSTVSANVTKIVYHVDDESMPYVTE
VPVAPDRVTLLDFKKMLNKLNYKYYCKSNDPEVGGEVKAEIRDDNQQLFRSLN
GQFELFLLTTDGSNNSDGGASSGFSRNVTQSVPGPAPPSAYPPFGLPQHMRQYSG
YDNGNRRRYLEDSTVGTESDARVFSDDESRVSTSTDNITSVSRQHNNYRVRQRR
MRRHLRQPSRASSMSSMTETSLALEVMTVTLNMDSVNFLGISIVGQSSTGGDNGI
YVANIMKGGAVALDGRIQPGDMILQVNDTSFENFSNDQAVEVLKEAVVHRGPIK
LTVAKSDTGRFDAFNVPSEPVRPIDMRAWIQHTNAANNLPHIPEGSEGAPTPIPGQ
YPHNYGRAPSSSTATSNGSNGHTILGPNGQVFILPKKLDLSTDRKRVVSVMALPN
SGLDIKDRTWLKIPIPMSFLGSDLVEWLIDHVEGLCDRKEAKKYATDLLKEGYIT
HMVNGNKFSEQSYYMIGRECSDAVRLRLSEEGASRSDVASSLPAPPPHLAPQPW
TANSAMVGHPHHPPQMPQQGGRLLVATPNSTRKSPVLNGVPNGGGNGAYPTSY
AADNSSYGVIGGVNGVAAPPPTWRLTEESLGSLRHMADV

>Pred-MIG-5
MSDAGSSSFTTVYYYLDGATPFKSEVPVPPEKITLGDFKRVFDRQNYKYFCKEFD
PMLKKDVKVELTKDSQKLNKSESGVIQLVLLPVYPASARASDSSGGTLPRASRSD
TGRNPPRRGFAQPKRALSITDETDYANNRYSLATDPASTAYSRRAGEQMAETLSS
SDFDDDNEREDTSDEEANRESGSSTSCRRPHGNRRNRQPRAYVPSSRGSKSATSQ
SDSLPGIAEICIQIYPHQNLGFNVADHDGGIFISDIFDDTAAGNCPDLSVGDQILEV
NSVCFEHLTFEQALAQIKKATKTAKSEATETKPGKIKMHVARLRVSDQHSESGLS
AGLGDTIPFEVSEWVMATTAENVDRFDDPLNNSRFDDGGVTSDEERAAYIDRRN
GVGARLVPALHNFRSNNGFDVPSSPHNRNHPALLMPPPPPRSRENDENEFLNAPL
SVDTDPITILKRMVHPASGLEITTRKWLKIPFPDSFIGNEMLAWLMEHVEGLKNK
KAARKYAASLLTKGLIRHVLDGKQFDKKRYYFFADNIISYRRQIEHARATASASR
LPTSATAATEVTFLGSPSYPPPAGPTANGGLLKSIPASCGALPSSSNNFFSSQNTYL
TGSRAPTRSMHSQQRLPQATRSPYYPAPIGPAMYMPQPQASPGLPQWPISPLIGSN
ERRPSCTSPVTTNEYASMINAEVNSNYMSVPTLPRGYGNARLMNNGTPPPNTPN
TALQGTASRSQVPIQHR

>Minc-DSH-1
METTKLDNDESNEVRMLEKETVQLDINKAMNTNNDLNSQAEKIDNVSDASSHL
THSSMSICSSAAGRGATKVYYHIDDEQTPYCTELPVPNNRVTLGDFKRVLNRTNF
KYYCKAIDSEVGGEVKAEIRDDAQQLTRSSNGHFELFLLTAENSNSDGNSSGVSK
LASMKKVPLGPTTIYPFTSSQHYRPRNRCDSNYVSGNHTRQPFDDSTLYTNETDH
RLYSDDESRLSTSTDNITSVSRQRYNLYRRRRRQAPRNRRKPSRASSLSSMTETS
MALEVITVTLNMDSAVNFLGISIVGQSSSRGDNGIYVANVIKGGAVALDGRIEPG
DMILQVNDVSFENFKNDKAVEVLKQAVNRQGPIKLTVAKSFDSGRANYFSVPVR
EPVRPIGSEGAPTPIPAHNQYHPREQMVNMLNIPPSITQHQQRSACSSNTTATSTN
GSGGVAPQIVLGQGGVFLAVQPRLDINSDKRIIIRAMVAIGSGLEIRDRTWLKIPIP
MSFLGSSLVDWLIQNVDGLKTRKEARKYASDLLKERFIAHVVNKQVFTEQCYY
VFGDNCSDLLLLRNGIEHGSALVRNPQNQAVLANQRLFHDRDCGTQVPYMLHT
NFGGSEYAMPSVSPYPPPVIPMQTTLEYSNQQRQLHKLVPQNPVHTQTRFDQFG
ANSQASNNSNEESSGSEHRRKAMLPPAPSLTSIGQPQYFQPHQFFPGQIHRTSQPT
GPPPTTPDDVNSMMLNRFCGINALLAMDVYANYMIQQNGIRQEQQPTLSQMKQ
EFAPYSGSSSQATGTGKFVAMPLLQQTGKVFVGPGQHREAQLIGLGLRKLRTFE
REDIERAKRYALDQSVKFVMLKQREAHQQQVNFIWDFLCFKLLFSRIRRMTILFC
RGHMYSID

>Minc-MIG-5a
MSEQPKGSTKVYYHLDDSTPYMSELMTSPDRLTLGDFKRAFNRKGYSYFCKEW
DPNLKREVKVEIVNDRQLLRKSVNGLFELFLLSQQNSTQPFSNTTSNMGADQESD
LQTLISKRAGENLAEMYNSTSEDPYNRTSVSSSSIFQRRPLPGAPMPSAESTAESA
AERTLLVNTSAARQKRLRKQRYVPSTISSESESTYSLPRIEEVKLRLQDAPLGISV
ASQCGSIFIYHIQHGSAAERCCRLEVGDQIVQIDETRFEDLNEKQALEVLKKLNSV
KKTITMYVAKRARTNGGESSEDHKSDPLSLLCETQQLDISQWVESTTNKNCVEQ
VRPFAEIPPVAKVERGTLPIDGCKVKNQTVDETSDEEKAAYLDRRNGVGARLVPI
IHQVRFHQQQQQQQKSEEMHKMQMINLENNPTTSSCPVMLGTLEESSPHLDCHP
PLIQLPLHAAMDPKIILFRMVEFDSGLEIRNRKWLKIPVPMSFIGDEMIDWLISNV
QGFRDRRHARSFASNLLAQGLIKHDEIIADRLRIEEQQKQQHLLKQHKPDADFQT

TSIKKPPALAPESNTEITYMSNPPVTSTCLGSSGAPQNVSSHTHPTLTDQQQRTQY
VIANNPRYAPGPMSNNFCGQKINKNTAALITDAANLNKIDKNAPKNAVATNIPLR
TLNPPKFGLKLPVWPISPILSFYRPQNQEVSQQQRGRCDSPEASKTNDDYASMIQ
GEINCGAQFDDGLNTFQNQQQ

>Minc-MIG-5b
MSELSITPDRLTLGDFKRAFNRKGYSYFCKEWDPNLKREVKVEIVNDRQLLRKS
RGQAKIWQKCTIQHRKILTTELQSSSSIFQRRPLPGASMPSAESTAESAAERTLLV
NTSAARQKRLRKQRYVPSTISSESESTYSLPRIEEVKLRLQDAPLGISVASQCGSIFI
YHIQHGSAAERCCRLEVGDQIVQIDETRFEDLNEKQALEVLKKLNSVKKTITMY
VAKRARTNGGESSEDHKSDPLSLLCETQQLDISQWVESTTNKNCVEQVRPFAEIP
PVAKVERGTLPVDGSKVKNQTVEETSDEEKAAYLDRRNGVGARLVPIIHQVRFH
QQQQQQKSEEMHKMQMINLENNPTVSSCPVMLGTLEESSPHLDCHPPLIQLPLH
AAMDPKIILSRMVEFDSGLEIRNRKWLKIPVPMSFIGDEMIDWLISNVQGFRDRR
HARSFASNLLAQGWIKHVVNISSFNEKCYYVFNDEIIADRLRVEEQQKQQHLLD
QKPDADFQTASIKKPPAPAPESNTEITYMSNPPVTSTCLGSSGAPQNVSSHLHPTL
TDQQQRTQYVIANNPRYAPGPMSNNFCGQRINKNTAALITDAAHLNKIDKNAPK
NAAATNIPLRTLNPPKFGLKLPVWPISPILSFYRPQNREVSQQQQRGRCDSPEASK
TNDDYASMIQGEINCGAQFDDGLTFQNQHQQQHLLKSKVTSPLGTGKYSKDLK
NDVNMSNKLSQSLATTNTTKPPPLPKHQFPSEFLTLQ

>Minc-MIG-5c
MSEQPKASIKVYYHLDDSTPYMSELPIPPDRLTLGDFKHAFNRKGYSYFCKEWD
PNLKREVKVEIVNDRQLLRKSVNGLIELFLLSQQNSTQPFSNTTTNLGADQESDL
QTLISKGRRKFGRNVQFNIRGSLQQNFNPPLAFFKEGLYLVLLCLQLKAQQRTLL
VNTSAARQKRLRKQRYVPSTISSESESTYSLPRIEEVKLRLQDAPLGISVASQCGSI
FIYHIQHGSAAERCCRLEVGDQIVQIDETRFEDLNEKQALEVLKKLNSVKKTITM
YVAKRARTNGGESSEDHKSDPLSLLCETQQLDISQWVESTTNKNCVEQVRPFAEI
PPVAKVERGTLPVDGSKVKNQTVEETSDEEKAAYLDRRNGVGARLVPIIHQVRF
HQQQQQQKSEEMNKMQMINLENNPTTSSCPVMLGTLEESSPHLDCHPPLIQLPL
HAAMDPKIILSRMVEFDSGLEIRNRKWLKIPVPMSFIGDEMIDWLISNVQGFRDR
RHARSFASNLLAQGLIKHVVNISSFNEKCYYVFNDEIIADRLRIEEEQKQQHLLKQ
HKPDADFQTASIKKPPAPAPESNTEITYMSNPPVTSTCLGSSGAPQNVTSSHPHPT
LTDQQQRTQYVIANNPRYAPGPMSNNFCGQKINKNTAALITDAANLNKIDKNAP
KNAAATNIPLRTLNPPKFGLKLPVWPISPILSFYRPQNQEVSQQQQRGRCDSPDAS
KTNDDYASMIQGEINCGAQFDDGLTFQNQHQQQHLLKSKVVS

>Mhap-DSH-1
MEATKLEFNEMQVLEKETVQLNINEANDTSKKDLNSKAENVDRTSDVSSHLTDS
SMSVCSSAAARGATKVYYHIDDEQTPYCTELPVPNDIVTLGFLFANQGHQFEIIG
DFKRVLNRTNFKYYCKAIDSEVGGEVKAEMRDDAQQLIRSSNGHFELFLLTAES
SNSDGNSSGVSKLASMKKRQTFEDSTIYTTETDHRLYSDDESRLSTSTDITSVSRQ
RYNLYRRRRRQEPRNRRQPSRASSLSSMTETSMALEVITVTLNMDSTVNFLGISIV
GQSSSRGDNGIYVANVIKGGAVALDGRIEPGDMILQVNDVSFENFKNDKAVEVL
KQAVSRRGPIKLTVAKSFDSGRANYFSVPAREPVRPIGSEGAPTPIPAQYPQGHMI
NILNAPPSISQRQQRSACSSNTTATSTNGSGGVAPQTVIGPGVFLTLQSQLDLNTD

KQIIMRVMAEINSGLEIRDRTWLKIPIPMSFLGSSLVDWLLQNVDGLKSRKEARK
YASDLLKERFIAHVVNKQASYL

>Mhap-MIG-1
MSEVQIPLDRLTLGDFKRVFNRKGYSYFCKEWDPNLKRSSSSIFRRGPLPGAPMP
SAESMTESAAERTLLVNTSTARQKRLRKQRYVPSTISSESESTYSLPRIEEVKLRL
QDAPLGISVASQCGSIFIYHIQNGSAAERCCRLEVGDQIVQIDETRFEDLNEKQAL
EVLKKLSYVKNKNCVEPFAEIPPVESRIGINQITGALPIDEFKVKNQTIDETSDEEK
AAYLDRRNGVGARLVPIIHQVRFHQQQKADVKQRIKMVDLENNPTTSSFPITLGI
VEECSPHPDCHPPLIQLPLHAAMDPIVILRRMVELDSGLEIRNRKWLKIPVPMSFI
GDEMIDWLIANVQGFRDRRHARSFASNLLAQGLIKHDEIIASRLHFEEEQKQLLL
KHQQNIDCQNTSIKKAPAPPPESNTEIT

>Dmel-DSH
MDADRGGGQETKVIYHIDDETTPYLVKIPIPSAQVTLRDFKLVLNKQNNNYKYFF
KSMDADFGVVKEEIADDSTILPCFNGRVVSWLVSADGTNQSDNCSELPTSECELG
MGLTNRKLQQQQQQHQQQQQQQQQQHQQQQQQQQQVQVPVQLAQQQQQQV
LHHQKMMGNPLLQPPPLTYQSASVLSSDLDSTSLFGTESELTLDRDMTDYSSVQ
RLQVRKKPQRRKKRAPSMSRTSSYSSITDSTMSLNIITVSINMEAVNFLGISIVGQS
NRGGDGGIYVGSIMKGGAVALDGRIEPGDMILQVNDVNFENMTNDEAVRVLRE
VVQKPGPIKLVVAKCWDPNPKGYFTIPRTEPVRPIDPGAWVAHTQALTSHDSIIA
DIAEPIKERLDQNNLEEIVKAMTKPDSGLEIRDRMWLKITIPNAFIGADAVNWVL
ENVEDVQDRREARRIVSAMLRSNYIKHTVNKLTFSEQCYYVVNEERNPNLLGRG
HLHPHQLPHGHGGHALSHADTESITSDIGPLPNPPIYMPYSATYNPSHGYQPIQYG
IAERHISSGSSSSDVLTSKDISASQSDITSVIHQANQLTIAAHGSNKSSGSSNRGGG
GGGGGGGNNTNDQDVSVFNYVL

>Hsap-Dvl1
MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFFKSMDQ
DFGVVKEEIFDDNAKLPCFNGRVVSWLVLAEGAHSDAGSQGTDSHTDLPPPLER
TGGIGDSRPPSFHPNVASSRDGMDNETGTESMVSHRRERARRRNREEAARTNGH
PRGDRRRDVGLPPDSASTALSSELESSSFVDSDEDGSTSRLSSSTEQSTSSRLIRKH
KRRRRKQRLRQADRASSFSSITDSTMSLNIVTVTLNMERHHFLGISIVGQSNDRG
DGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSQT
GPISLTVAKCWDPTPRSYFTVPRADPVRPIDPAAWLSHTAALTGALPRYELEEAP
LTVKSDMSAVVRVMQLPDSGLEIRDRMWLKITIANAVIGADVVDWLYTHVEGF
KERREARKYASSLLKHGFLRHTVNKITFSEQCYYVFGDLCSNLATLNLNSGSSGT
SDQDTLAPLPHPAAPWPLGQGYPYQYPGPPPCFPPAYQDPGFSYGSGSTGSQQSE
GSKSSGSTRSSRRAPGREKERRAAGAGGSGSESDHTAPSGVGSSWRERPAGQLS
RGSSPRSQASATAPGLPPPHPTTKAYTVVGGPPGGPPVRELAAVPPELTGSRQSFQ
KAMGNPCEFFVDIM

>Hsap-Dvl2(>hsap_g0044131)
MAGSSTGGGGVGETKVIYHLDEEETPYLVKIPVPAERITLGDFKSVLQRPAGAKY
FFKSMDQDFGVVKEEISDDNARLPCFNGRVVSWLVSSDNPQPEMAPPVHEPRAE
LAPPAPPLPPLPPERTSGIGDSRPPSFHPNVSSSHENLEPETETESVVSLRRERPRRR

DSSEHGAGGHRTGGPSRLERHLAGYESSSTLMTSELESTSLGDSDEEDTMSRFSS
STEQSSASRLLKRHRRRRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFL
GISIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDMNFENMSNDD
AVRVLRDIVHKPGPIVLTVAKCWDPSPQAYFTLPRNEPIQPIDPAAWVSHSAALT
GTFPAYPGSSSMSTITSGSSLPDGCEGRGLSVHTDMASVTKAMAAPESGLEVRDR
MWLKITIPNAFLGSDVVDWLYHHVEGFPERREARKYASGLLKAGLIRHTVNKIT
FSEQCYYVFGDLSGGCESYLVNLSLNDNDGSSGASDQDTLAPLPGATPWPLLPTF
SYQYPAPHPYSPQPPPYHELSSYTYGGGSASSQHSEGSRSSGSTRSDGGAGRTGR
PEERAPESKSGSGSESEPSSRGGSLRRGGEASGTSDGGPPPSRGSTGGAPNLRAHP
GLHPYGPPPGMALPYNPMMVVMMPPPPPPVPPAVQPPGAPPVRDLGSVPPELTA
SRQSFHMAMGNPSEFFVDVM

>Hsap-Dvl3
MGETKIIYHLDGQETPYLVKLPLPAERVTLADFKGVLQRPSYKFFFKSMDDDFG
VVKEEISDDNAKLPCFNGRVVSWLVSAEGSHPDPAPFCADNPSELPPPMERTGGI
GDSRPPSFHPHAGGGSQENLDNDTETDSLVSAQRERPRRRDGPEHATRLNGTAK
GERRREPGGYDSSSTLMSSELETTSFFDSDEDDSTSRFSSSTEQSSASRLMRRHKR
RRRKQKVSRIERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGI
YIGSIMKGGAVAADGRIEPGDMLLQVNEINFENMSNDDAVRVLREIVHKPGPITL
TVAKCWDPSPRGCFTLPRSEPIRPIDPAAWVSHTAAMTGTFPAYGMSPSLSTITST
SSSITSSIPDTERLDDFHLSIHSDMAAIVKAMASPESGLEVRDRMWLKITIPNAFIG
SDVVDWLYHNVEGFTDRREARKYASNLLKAGFIRHTVNKITFSEQCYYIFGDLC
GNMANLSLHDHDGSSGASDQDTLAPLPHPGAAPWPMAFPYQYPPPPHPYNPHP
GFPELGYSYGGGSASSQHSEGSRSSGSNRSGSDRRKEKDPKAGDSKSGGSGSESD
HTTRSSLRGPRERAPSERSGPAASEHSHRSHHSLASSLRSHHTHPSYGPPGVPPLY
GPPMLMMPPPPAAMGPPGAPPGRDLASVPPELTASRQSFRMAMGNPSEFFVDV
M

>Mmus-Dvl3
MGETKIIYHLDGQETPYLVKLPLPAERVTLADFKGVLQRPSYKFFFKSMDDDFG
VVKEEISDDNAKLPCFNGRVVSWLVSAEGSHPEPAPFCADNPSELPPSMERTGGI
GDSRPPSFHPHASGGSQENLDNDTETDSLVSAQRERPRRRDGPEHAARLNGTTK
GERRREPGGYDSSSTLMSSELETTSFFDSDEDDSTSRFSSSTEQSSASRLMRRHKR
RRRKQKVSRIERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGI
YIGSIMKGGAVAADGRIEPGDMLLQVNEINFENMSNDDAVRVLREIVHKPGPITL
TVAKCWDPSPRGCFTLPRSEPIRPIDPAAWVSHTAAMTGTFPAYGMSPSLSTITST
SSSITSSIPDTERLDDFHLSIHSDMAAIVKAMASPESGLEVRDRMWLKITIPNAFIG
SDVVDWLYHNVEGFTDRREARKYASNLLKAGFIRHTVNKITFSEQCYYIFGDLC
GNMANLSLHDHDGSSGASDQDTLAPLPHPGAAPWPMAFPYQYPPPPHPYNPHP
GFPELGYSYGGGSASSQHSEGSRSSGSNRSGSDRRKEKDPKAGDSKSGGSGSESD
HTTRSSLRGPRERAPSERSGPAASEHSHRSHHSLTSSLRSHHTHPSYGPPGVPPLY
GPPMLMMTPPPAAMGPPGAPPGRDLASVPPELTASRQSFRMAMGNPSEFFVDV
M

>Mmus-Dvl2
MAGSSAGGGGVGETKVIYHLDEEETPYLVKIPVPAERITLGDFKSVLQRPAGAKY
FFKSMDQDFGVVKEEISDDNARLPCFNGRVVSWLVSSDTPQPEVAPPAHESRTEL
VPPPPPLPPLPPERTSGIGDSRPPSFHPNVSSSHENLEPETETESVVSLRRDRPRRRD
SSEHGAGGHRPGGPSRLERHLAGYESSSTLMTSELESTSLGDSDEDDTMSRFSSST
EQSSASRLLKRHRRRRKQRPPRMERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGI
SIVGQSNERGDGGIYIGSIMKGGAVAADGRIEPGDMLLQVNDMNFENMSNDDA
VRVLRDIVHKPGPIVLTVAKCWDPSPQAYFTLPRNEPIQPIDPAAWVSHSAALTG
AFPAYPGSSSMSTITSGSSLPDGCEGRGLSVHMDMASVTKAMAAPESGLEVRDR
MWLKITIPNAFLGSDVVDWLYHHVEGFPERREARKYASGLLKAGLIRHTVNKIT
FSEQCYYVFGDLSGGCESYLVNLSLNDNDGSSGASDQDTLAPLPGATPWPLLPTF
SYQYPAPHPYSPQPPPYHELSSYTYGGGSASSQHSEGSRSSGSTRSDGGAGRTGR
PEERAPESKSGSGSESELSSRGGSLRRGGEPGGTGDGGPPPSRGSTGAPPNLRALP
GLHPYGAPSGMALPYNPMMVVMMPPPPPPVSTAVQPPGAPPVRDLGSVPPELTA
SRQSFHMAMGNPSEFFVDVM

>Mmus-Dvl1
MAETKIIYHMDEEETPYLVKLPVAPERVTLADFKNVLSNRPVHAYKFFFKSMDQ
DFGVVKEEIFDDNAKLPCFNGRVVSWLVLAEGAHSDAGSQGTDSHTDLPPPLER
TGGIGDSRPPSFHPNVASSRDGMDNETGTESMVSHRRERARRRNRDEAARTNGH
PRGDRRRDLGLPPDSASTVLSSELESSSFIDSDEEDNTSRLSSSTEQSTSSRLVRKH
KCRRRKQRLRQTDRASSFSSITDSTMSLNIITVTLNMERHHFLGISIVGQSNDRGD
GGIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSQTG
PISLTVAKCWDPTPRSYFTIPRADPVRPIDPAAWLSHTAALTGALPRYGTSPCSSAI
TRTSSSSLTSSVPGAPQLEEAPLTVKSDMSAIVRVMQLPDSGLEIRDRMWLKITIA
NAVIGADVVDWLYTHVEGFKERREARKYASSMLKHGFLRHTVNKITFSEQCYY
VFGDLCSNLASLNLNSGSSGASDQDTLAPLPHPSVPWPLGQGYPYQYPGPPPCFP
PAYQDPGFSCGSGSAGSQQSEGSKSSGSTRSSHRTPGREERRATGAGGSGSESDH
TVPSGSGSTGWWERPVSQLSRGSSPRSQASAVAPGLPPLHPLTKAYAVVGGPPG
GPPVRELAAVPPELTGSRQSFQKAMGNPCEFFVDIM

>Nvit-DSH
MEETKIIYHIDDEETPYLVKLNISPERVTLADFKNVLNRPNYKYFFKSMDDDFGV
VKEEIVDDDAHLPCFNGRVVSWLVSAEGSNVSDGASQCTDTVPHQDPKHDRVD
HVTPGHTNRAQLSLSHEDTLTETESIISSRQGHHLHKSSRHHTDKYEKYNKYNGL
RINGHSKHRSGMGYESASILSSDLETTTFLESDDDASSRITSTTGRHTNMSSAVDR
ATLDRRRPQRRRRHRLPPMSRTSSFSSITDSTMSLNIITVSLNMDTVNFLGISIVGQ
SNKGGDGGIYVGSIMKGGAVALDGRIEPGDMILQVNDINFENMSNDEAVRVLRE
VVQKPGPIKLVVAKCWDPNPKGYFTIPRTEPVRPIDPGAWVAHTAAIRGEGFPPR
PPSATTLTSTSSSLASTLPDTERPFDELDLSVNTDMPTIVRAMARPDSGLEIRDRM
WLKITIPNAFIGADVVDWLHNHVKGFIDRRDARKYASLMLKAGFIRHTVNKITFS
EQCYYIFGDLCSAMSNMKLDCDTVGPLPPPSAWDMPYSGTYAPHSATGYSPMPF
NFTNEPTVYGYHREESLHSGSGGSSAGSELMFKAPMHDVKSCCSASESELQMPV
VPAMPKSTATTVTTGGSGNGNGSGSNGKRSNGSRSRSSGSEQSVQTGTGSGGGG
GSSVGPQQQQQQQDLSGSRQSFRIAMGNPCEFFVDVM

>Tcas-DSH
MDETKVIYHIDDEETPYLVKIPISPEKVTLSDFKNVLNRPNYKFFFKSMDDDFGV
VKEEIIDDSAHLPCFNGRVVSWLVSADGSNQSDGGSQCTDSVTHQSERVPPQAPE
SICTDTESIISSRQGRRHHKYSSRINGHLPRVYETASIVSSDLETTSFQSETTGRHTA
LSECSSVSRLHVAGRKRPQRRRKQRMQAMSRTSSYSSITDSTMSLNIITVTLNMD
TVNFLGISIVGQSNKGGDGGIYVGSIMKGGAVALDGRIEPGDMILQVNDVNFEN
MSNDEAVRVLREVVQKPGPIKLVVAKCWDPNPKGYFTIPRTEPVRPIDPGAWVA
HTAAVRGDPVARPPSSSTVSSTSITSTIPANERFPDLEEPLTVNTPMATVVQAMQR
PDSGLEIRDRMWLKITIPNAFIGTDMIDWLLTHVDGFQERRDARKYASHLLKAGF
IRHTVNKITFSEQCYYIFGDLCSAMNNLKIQGDTDSVGPLPNVPNYMPYSGTYNP
LEYMPMPFYTASENTVYGYNREESVLSGSGGSSNGSDHLKDAAAGHSSASDSDL
TSLGPRSALPMATGNGNGSSNGSDQSSGTQVAAQSKDIAGSRQSFKIAMGNPCE
MFVDVM

>Chem-Dvl
MAEKETKIIYHVDDEETPYLVKIPKPPDQVTLGDFKSVINRPNFKFFFKSMDDDF
GVVKEEIIDDDAPLPCFNGRVVSWVVPPEDGSCDGQSQHSGDGIFVPVQSSNSNI
SRSSTMRSKERVQDSDAESIVSRRSSRSRSSRKYESDADRRSRRSHREHRNYDQY
DSASMMSSDLETTSFVDSEEESQMSSATESSRYVGGNKRRRRRRQRMPRVERCS
SFSTITESTMSLNIITVTLNMDKINFLGISIVGQASKKGDGGIYVGSVMKGGAVDA
DGRVEPGDMILAVGDVNFENMSNDDAVRVLRECVHKPGPIQLTVAKCWDPNPK
GYFTVPKDDVTRPIDPAAWVQHSEAMRAGGGLMGRGSPSMSTMTSTSDSFSSSI
PEADRYLEHDLGLTLTIDTDMLTIVKVMNQENSGLTVRDRMWLKITIPNAFIGSD
LVDWLFANVEGFQDRREARKYASKLLKANLIRHTVNKVTFSEQCYYVFGDLSD
RRRGMHMQDIEGSEEDTLAPLASTSAQAQMGGYIQQPYGMAMIPGGMMVTGA
PPAYQQVVQPGMYAQYPYVAGSIAPSHLSGGSGGSGSQKSQKIHDDQATVRSSG
SSDRSHSSKGSRSRKPMLDDRSSLSSFHSDTISLRSEIIAPDNRSISSHHSIPPMMVM
PPGVAYSAAPVGLTPEQQQHIQVQMHQQQQMHHQIQQQLQQLQLQSPAQTPQP
PPAGSPQPPQDGTRTPAEQHRELGRLDTLPASLSASRQSFRMAMGNSSNEFFVDV
M

>Aque-Dvl
MEETKVIYYVDDEETPYMTKIPLSPSKVKLSDLKEQLSRPGPFKKFFFKSIDDDIG
VVKEEIIEDSALLPTAKGRIVCWVVSEGSASGSDVASKDLETIKEKDDDEASIVSG
VSKSSRHSTSSRRHHHHHHHRHHRRRHDRHKQPGSVTDYETATELTATDLEASC
YETEDTASRLSVETTSTVTSHKHPRKKVLTKKKVPRSMSMSTMTTSTMSDASMD
ILTVTLNMDAYNFLGISIVGHANDDGVGGIYVGTVMKGGAVAADGRIETGDML
LQVNDISFENMSNDDAVRTLREIVQQPGPIILTVAKCLEPEAYAPMFEPRLEPIRPL
DPSAWVMHTNAQRAGDYGRPFTSSPTMSTMTSNSSPSLASSIPESERDLVKLNLT
SPLYRVAKAMAAPDSGLEVKDRMWLKMPIPKSFIGSDMVEWLHNNVEGFIDRR
HARKYAAIMLKQGFVKHAVNKYTFSEQCYYIFGDFKKAAANNLSTELTQLHLS
DNSGGETDTLGPLPSDRNNGQGWSSSDDQISLPTIEGSFYPHLMHSHSNSPEGQM
QHHTGSSASSSSAAGSGYAHYPPRQHSGSPSDPRQYPSPRPSHSHMELPPQMRMR
HSNSISSASEFTSVSQQQYHHPGNGGAMGGASLPPDRMSLRSLNIGGQVPPGVSA
SRASFQQALDNPCEYFIDVM

>Cint-Dvl
MSDETKIVYYLGDEQTPYVSKINLPPDSITLGDFKAAIKKINYKFFFKSTDADFGV
VKEEVTNDKSILPLCDNRIVAWLKAPDMDSVSQCSLNSNTLPMQDKNEDMQLR
PNSTNEGELSIENKRVSPHYPLYNRYGSQSTLMSSDIETTSWDSRDDMSEYSTCT
DMTASSRRKPRPKRQRYKKKRPRPGSDTTSTYSSSITDSSMSLNILTVTLNMEKY
NFLGISIVGQTNDKGDGGIYIGSIMKGGAVAADNRIEPGDMLLQVNEVNFENMS
NEDAVRVLRNIVHKPGPITLTVAKCWDPNPDNYFTIPKDEPVRPIDPAAWANHIM
TVKGDMHGPSPPYSLNPESSSASSLPESERYDMPLSTCTDMGAVVKSLKMPDSG
LDVKTRMWLKITIPNAFIGSDLVDWLQQKVHGLTERRDARKYASSLLKAGYIRH
TVNKITFSEQCYYVFGDYTGTLDKDMSNLSLVESNSDRDSDTLGPLHGHHNTQP
WMHQPSTSNHGYASYSQYPYTQSGQLSVGPPIPLNNGISSMVDYSSPPPTYPGLA
SHHSDGDNLSVHSAHSLHRAASEASTYRQHLATFGANGNNSAGSESESRYSSRS
GKSGSGGSAKPRSRSGSEKSSDVLHQQMAYGGSTSYHNSLQRSNHPNQRTNQHS
RPLSAIPPNLSNSQQSFQQAMGNPCDYFVDVM

>Smed-DVL-1
MEETRIIYYVDDEETPYLIKFHSPPEQITLGDFKNALNRPNYKFFFKSLDDDFGVV
KEEITDDDAKLPYVNGRVVSWLVVSEGSTQSDNHSSSGKEVLLVDSDKSKDKGT
VSDSSDPKSPSFRNYNKIPTKHSSSSKKQEESNKIHRQNHKFTGPEKITLDETDDA
FDEIDSIYNEDKVPPLRKFSDFKHSVKLKKLRNAQGSHGNSNSSSNNNNSNSNNA
SNNAPAKQQPIYESSSSMMSSDLDTTSFFDSEDDSSRFSSATETTMSSKYGKQRR
QLRRRRKMPHLSRASSFSSMTDSTVSLNIITVTLNMDTVPFLGISIVGQTNGNQEN
GDGGIYVGSIMKGGAVALDGRIEPGDMILEVNGISFENVSNEEAVRTLREQVQKL
GPVTLVVAKSWDPNPTGYMLPQQDPVRPIDPRAWVLHTQAMGNMAPPNQPPV
ASGDQFIQSGKYLPYAGAMSTVASTITTTSSSLKSRTDETIPSPLTTNHEPSVIIRA
MAQADSGLPIRDRLWLKITIYNAFIGSDLVDWLYSHVQGFTDRKDARKFATNLL
KMGFIRHTVNKSSFSEHQRLTRFRLLPANSNPKRDWCIKILAVRSRQFSHRDSSLL
AVTNGQVTISSRILVQKCIHKHKS

>SmedDVL-2
MTNCATSGNVISDETRIIYHIDEEETPYLIKLSISPDKVTLGDLKNTLNRPHYKYFF
KSMDDDFGVVKEEITDDEAKLPCFKGRVISWLVTAEGSTVSDNVDSNGILDKNE
SRMLPFQESHFPLINNIKASGGTTTNESDTICDTCTDTDSVYSAAQDRVGPPRSFH
DYKQAGRVAAHANRVNTNTPNGQNPIYETNSSMMSSDLESTSFFDSEDESSRFST
TTCTTMSSRYGRQKQQRRRRRPPAISRASSFSSITDSTMSLNIVTVRLNMDTVKFL
GISIVGQSNKGGDGGIYVGSIMKGGAVAQDGRIEPGDMILEVNDISFEDMSNDDA
VRTLREQVQKPGPINLVVAKCWDPNPKGYFTIPRQEPVRPIDPRAWVLHTNAMT
AGASEPPSSVNGVHPQVSNLVAPSMQSLLSGGTMLAGTSAATFNAAAFGYMPQP
QNINQNTASVSTVGGPPGASVGFFGYPMGMPGQFSQGAGSIVTTSSSLPESERYQ
EELHLTKNTDVGTILRVLSQPDSGLDIRDRLWLKITLPNAFIGSNLVDWLYRHIEG
FSDRKEARKYAANLLKFGYIKHTVNKVTFSEQCYYVLGNTTLNMSRLSLDQVES
VSEVGVNGPHHLAALPPPNFSSNKQPISSCINQPPLNINPQLTATSEPLPSNNANVA
TATASSNSQYSVVGPLPCSQPSQHASSNASASAIKKSGSCNSLSGSSSSTSSSSSSN
RSNTRINGNASSVSNMISKNPPPKIPPRTIASVSTNSTNPIISGFQNRGQSSVSQ

>Xtro-Dvl1
MAETRIIYHIDEEETPYLVKLPVPPEKVTLADFKNVLSNRPVHHYKFFFKSMDQD
FGVVKEEISDDNAKLPCFNGRVVSWLVLAESSHSDGGSQSTESRTDLPLPIERTG
GIGDSRPPSFHPNASSSRDGLDNETGTDSVVSHRRDRHRRKNRETHDDVPRINGH
PKLDRIRDPGGYDSASTVMSSELESSSFVDSDEDENTSRLSSSTEQSTSSRLIRKHK
RRRRKQKMRQIDRSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNDRGDG
GIYIGSIMKGGAVAADGRIEPGDMLLQVNDVNFENMSNDDAVRVLREIVSKPGPI
SLTVAKCWDPTPRSYFTIPRAEPVRPIDPAAWITHTSALTGAYPRYEQEDSPLSVK
SDMATIVKVMQLPDSGLEIRDRMWLKITISNAVIGADVVDWLYTHVEGFKERRE
ARKYASSMLKHGYLRHTVNKITFSEQCYYVFGDLCGNVAALNLNEGSSGTSDQ
DTLAPLPHPAAPWPLGQGYSYQYPLAPPCFPPTYQEPGFSYGSGSAGSQHSEGKS
TYSGVFLP

>Xtro-Dvl2
MAETKVIYHLDEEETPYLVKVPVPANEIRLRDFKAALGRGHAKYFFKAMDQDF
GVVKEEISDDNAKLPCFNGRVVSWLVSSETSQTDSAPPAAEVRPDPPPVPPPVPPP
PAERTSGIGDSRPPSFHPNVSGSTEQLDQDNESVISMRRDRVRRRDSTEQGVARG
VNGRAERHLSGYESSSTLLTSEIETSICDSEEDDAMSRFSSSTEQSSASRLLKRHRR
RRKQRPPRLERTSSFSSVTDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIY
IGSIMKGGAVAADGRIEPGDMLLQVNDINFENMSNDDAVRVLRDIVHKPGPIILT
VAKCWDPSPQGYFTLPRNEPIQPIDPAAWVSHSAALSGSFPVYPGSASMSSMTSS
TSVTETELSHALPPVSLFSLSVHTDLASVAKVMASPESGLEVRDRMWLKITIPNA
FLGSDMVDWLYHHVEGFQDRREARKFASNLLKAGLIRHTVNKITFSEQCYYIFG
DLTGCENYMANLSLNDNDGSSGASDQDTLAPLPLPGASPWPLLPTFSYQYPAPH
PYSTQPPAYHELSSYSYGMGSAGSQHSEGSRSSGSNRSDGGRGTQKDERSGVVG
VGGGESKSGSGSESEYSTRSSIRRIGGGEAGPPSERSTSSRPPLHHPPSVHSYAAPG
VPLSYNPMMLMMMPPPPLPPPGACPPSSSVPPGAPPLVRDLASVPPELTASRQSF
HMAMGNPSEFFVDVM

>Xtro-Dvl3
MGETKVIYHLDEQETPYLVKLPVPAEKVTLGDFKNVLNKPNYKFFFKSMDDDFG
VVKEEISDDNAKLPCFNGRVVCWLVSADGSQSDAGSVCADNQSDLPPPIERTGGI
GDSRPPSFHPNTRGSQENLDNETETDSVVSAHRERPRRKETPEHATRLNGTSKME
RRRDTGGYESSSTLMSSELDSTSFFDSDEDDSTSRFSNSTEQSSASRLMRRHKRRR
RKPKAPRIERSSSFSSITDSTMSLNIITVTLNMEKYNFLGISIVGQSNERGDGGIYIG
SIMKGGAVAADGRIEPGDMLLQVNDTNFENMSNDDAVRVLREIVHKPGPITLTV
AKCWDPSPRNCFTLPRSEPIRPIDPAAWVSHTAAMTGSYPAYGMSPSMSTITSTSS
SITSSIPETERFDDFQLSIHSDMVTIVKAMRSPESGLEVRDRMWLKITIPNAFIGSD
VVDWLYHHVEGFTDRREARKYASNLLKAGYIRHTVNKITFSEQCYYIFGDLCGN
MANLSLNDHDGSSGTSDQDTLAPLPHPGAAPWPIAFQYQYPLPHPYSPHPGFPDP
AYSYGGGSAGSQHSEGSRSSGSNRSSTEKRKEREAKGGDTKSGGSGSESDHTTRS
SVRRERAASERSVPASEHSHRSHHSIAHSIRSHHTHHSFGPPGIPPLYGAPMMMM
PAPASVIGPPGAPPSRDLASVPPELTASRQSFRMAMGNPTKNSGVFDFL

>Cang-DSH-1a(CAN07318)
MDTVNFLGISIVGQTSTRGDNGIYVANIMKGGAVALDGRIEAGDMILQVNDISFE
NFTNDQAVDVLRDAVSRRGPIKLTVAKSFENGPQSCFTIPRNSREEPVRPIDTQA
WIQHTNAMRGMPSIVEGVEGAPTPLPGDWPPHANGRPQSSSTVTSNGSNGQNTV
VGGGNGTQIKLDLATDKKKIVEVMAMPNSGLDIKNRTWLKIPIPMSFLGSDLVE
WLLDHVEGLRERKDARKFAAELLKLKYIAHVVNKITFTEQCYYVLGDECSALL
MLNWNFCWKIHEN

>Cang-DSH-1b
MPERKIGGGGSLVTTIVSRFESKNLSGGPATKKSPQPPIAKLDDDFDAAPPPKSKV
LHQLKRTYTNLPTEEILSKKQLKKRQSESALQHLGGNKNKEGGAGQMFIGTLMR
LASLKSLVASVEDIQIGEGKKKDQKKKRKRKRADAKEALKNRLLNWNSVEEPA
PSWVNRSPSPSKIHIPTPVFKRSPEVEEEPEGLKKSSSRSLFSFSRSESKESLLSQKT
SPRSLSPARNLAENIKKKIWGRSDSLGSQASLGSAKSPRSLSPARNLVNRFLGKSA
SQTSIRSQETLKERSSSLKSKQEGSEVEGRSGSLRSLSKKILGSNLKLEARPSEPNIS
SPVATSAPTSAPAATPAPKRRLVRSDATIGARLKKKLSVEEDEPTEPEPLPKSPLR
SPLPKLVIEHLPDRPQVTQRSLSEIVSTDRLRPLSLPAGRRKLPPMARTMSLIPPPV
LPPLYEEESLAETSHNSLGTTTHTKRRYGGSNTTSTYQGVKDAAPDASPRRHLIG
FLHRTSHLSLTSELSGNASFYLIDLGTELEPEVGSETFGAQHQLSRRDLSSYFCSSG
HPGSKKKTCPIGRHDRSQAQEEAVGHRRQHLEDSTIASESDARIFSDEEDDR

>Cang-MIG-5
MEPQGSSDQIKVFYFLDDESTPYVSVIESRGGIVTLDDFKNSFTKRGYKYFCKEL
DPDIQCEVKVELTEGTDRLRKSQNGFYEIFLVSTPGYGTLPRNSGTLTRKNVKGT
LDRRRRRSADFDAVPYSDASLAPSTIVSRRAGEHLAELYTSNSEDPYQYDDSTRL
TGESSLYEPLATRDMNKMYEDDRKQKRKIKKERHHRRPYVPSTISSATESSAGSG
LPRILEVFLPMKNVPYLGLNVCTIDGHIFVSEIAQNGAVERDGRIDVGDQILQVNR
ISFEELSVPQAVRALREAAASRKPMTLYVSKFSKAAGSEXXXXSSYIANIAVRAL
REASRKPMTLYTYLL

>Cbri-DSH-1
MAESPPRVDSSAPDVGDPTSMMERLRLGDDGGKDDEFDNKSVSSAQYSQASEA
TTAVKQQPFLHTMTKVYCHIDDETDPYMLEVHVPPDLITLGDLKRVLMRTNFKY
YRKALDPDSGYEVKAEIRDDSQRLTPSSNNLFELFLLTIEGSTHSDGSSGKMRKY
PSVPGPAPSNRSGPPMNYQHAAYQFDNSMMSTDSESMISAAIPGYLKSAAYNRR
FPQHYLAMSIKLGVARERAIPASSNTPSTSSMTTSSDVLRSLLFTAKERKVLSTYT
APPSHRNHSPGVVSQLINKIEQKASPPTTSSRTKKKKKAPPLAQIDDLDLIEDEKT
VVEVAKSPRRKKYTGNPLGNLKILNGFLPSSKTTDDKENDSVPSTGKKLKKSKS
KSRVEPETDPPFRLKRTLTDLTNNHTPSGSTWRQMVSNALGGPLKKRLSEGALFF
PGTSKAADDCVDEEGGNTKTKTKSPFLQRRKQNEKSAGGANRKESGGSASSSFF
GALIRLSHSAASLTSLTSLGGSRSNSASPSSSRSNTKEFKEELKPPQPPPPDLISPAA
ALSKPIFSLAPPSSPVSLTPPSRELRKSKTCQITTDRPPIIPSITISESRSLNRIDRCRPV
TVDGSNLTPDRRPLVSRRSTMSRTMSLIPTSPSLPPLYEEEMASTAAMVDEEREE
KTQKRRMRRYGGSNTTSTYQGRKDAAPDASPRRHLIGFLHRTSHLSLTSELSAD
ASFYLIGHRRHLEESTIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKMAKKKKN
KRNFRKPSRASSFSSITESSMSLDVITVNLNMDTVNFLGISIVGQTSNCGDNGIYV

ANIMKGGAVALDGRIEAGDMILQVNETSFENFTNDQAVDVLREAVSRRGPIKLT
VAKSFENGQSCFTIPRNSREEPVRPIDTQAWIQHTNAMRGMPSIVEESAPTPIPGE
WPHGRPPSSSTVTSNGSNGQNTVVGNGAHIHLDIHTDKKKVVEIMAMPGSGLDI
KNRTWLKIPIPMSFLGSDLVEWLLDRIEGLRERKTARNYAADLLKLKYIAHVVN
KVTFTEQCYYVLGDECADYARFRNEDGGPKYQWTMGMNGMSAGNGSSVMLP
PPQFPGAVPFKGMAPSMVSDGESRMYVMHI

>Cbri-DSH-2
MTDSPSPIDSSFDASDVATPCTVIAAKSSIRNFRDLEEGDDFDDDDDGIEDQTDYT
ESFPQPEDRALDGEDPSNIYVDDLSKEFSGCASSVMEPLPKPLTFARTMTKVFYH
LDNDPVPYTIDVHVPPDCMTLRDVKRKLPRTNYQYYCIALDPDSGKEVRAEIRD
DSQRLYPLRSGEFRLYMLTIEGSVHSDTSSGRHRRKNKGGSSNGSSSSRDYLRAA
RDYDNQSESHAVLSRYFTFILSAPFTDDESQVSETPVYVKKANAFNRRQASQAY
DQYQPRHMLHARHHQNHYEDSTFDVTSESESQFRSGALYDEDVDDAQSINTDLT
SVSQVQLKKRWRQQQKEMRNKWKRMPSMSTASSSLSSITESSMGLEVITVRLNL
QSMPLGMIPYGLKTARGGDAGLYVGDILGRGAVALDGRIEVGDMISEINEIDLSN
YSNEAAAQLLKDAVAPRQFVTLTIAKSLDSRKAAAAASARNTRNEPIRPIDTNEW
IKHANAMKGMPSISEESCSTPIPDDWPTNSSASGTPFGGPHPAIAHMTTETNKKRI
MEILAAPGSGLEIKDREWMKLPLKMCFLGKDLVNWLLDHVEGLRKHKEAKQFA
KEMWKLGYIVDALGQNVFSENCYYKMGEECADYTQLRAPDGGFKYAQSHTSS
ASGHSSNNNIFPPSMYPPQMKPMPSTAGLNAHHRNSAVLNSMVSGYASMPSSPF
PNVKPSAVDCGRTRDDVRSQTSGSSQGSSRRYVELPRKPSSLGSGSGISDQMNID
RVASRSSFRAAMSGSLRHFSID

>Cbri-MIG-5
MEPPCSSDSNQIKSKQDGFQVFYYLDDETTPYVSVIDAREGVATLGNFKNSFTKR
GYKYYAKELDPDIQREVKVELISDTDRLRRSQNGFYEVFLVSTPGYGTLPRNTGT
MTRPQRAAIDKRRRRSADFDATPYSDASLAPSTIVSRRAGEHLAEMYTSNSEDPY
QYDESTRRTTDDSSMYEPLTARDMNRYHDDDRRKKKQKKDRFRRPYVPSTISSA
TESSVNSGLPRILEIYLSMKNVPYLGLSVCTMDGHIFVSEIAPEGAVEKDGRVNV
GDQILQVNRVSFEDLTGPQAVRALREAAGSKRPITLYISKYRRAAPSEYDDPLAS
MASETMPLDVGVWVETAVQATEKMKALGIDPQEQTMTSVDDGTMPFTSTASD
DEERILYDQRRNGIPRALLEEAERKKENERNEKAEQLTELIDPIIVVRAMARPDSG
LVIKNRKWLKISVPMSFIGQDLIDWLVDHMTDIHNRKRAKLYAARLLAAGLIRH
VVSKLTFTEKCYYVFGDGILPTDRTSADNSGTSGTTTTRVEATTEVTYVGSPAPH
ALANRMGRTIPPHRLETTTLSPVAHDQTWLRRRRDCESPMTNDYASMVGESQIG
MNPAGNYYGAKNSRQMMVPASSQVTSSSLTNGSGGIGGPPPTPLSSTMVLAASPI
QS

>Cbre-DSH-1a
MAEPPPQVDSTNAPDIGNPTSMMERLRLRDQTEDGGENDDFDNKSVSSAQYSQA
SEATTAVKQQPFLQTMTKVYCHIDDEPDPYMLEVHVPPDLITLGDLKRVLMRTN
FKYYRKALDPDSGYEVKAEIRDDSQRLTPSSNNLFELFLLTIEGSTHSDGSSGKM
RKYPSVPGPAPSNRNGPPMNYQHAAYQFDNSMMSTDSESMISAAIPGYLKSAAY
NRRFPPQQHYLGHRRHLEESTIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKMT
KKKKNKRNFRKPSRASSFSSITESSMSLDVITVNLNMDTVNFLGISIVGQTSNCGD

NGIYVANIMKGGAVALDGRIEAGDMILQVNETSFENFTNDQAVDVLREAVSRRG
PIKLTVAKAFENGQSCFTIPRNSREEPVRPIDTQAWIQHTNAMRGMPSIVEGSDLV
EWLLDRIEGLRERKTARNYAAELLKLKYIAHVVNKVTFTEQCYYVLGDECADY
ARFRNEDGGPKYQWTIGMNGMSAGNGSSIMLPPPHLPGPPGAFKGMAPSMVSG
YASMPSSPFPAQLQQQMQNHRREGSTTSGSSGGGIRKQRVVLPRQPPPSANGNV
HYDDSSTIYEGSNNSFLAATGQRYEFN

>Cbre-DSH-1b
MSIKLEARKTRALPTDSSSPSTSMSTSSDVLRSLLFTAKERKVLSTYTAPPSNRHH
SPGVVSQLINKIEQKTSPPSTSSRPKKKKKAPPLAQIDDLDIIEDEKVEEVVKSPRR
KKYTGNPLGNLKILNNILPKATSSEDKIDKEDKENSSHPTTGKKLKKSKSKSRVEP
EIEPPFRLKRTLTDLTNNHTPAGSTWRQMVSNALGGPLKKRLSEGALFLPGPSNC
GDEEGEKTTTTKQRRFLQRRKQSEKSAGDGKESSGGRGGSASSSFFGALIRLSHS
AASLTSLTSLGGSKSRSTSASPSSSRSNTKEFKEEIKPPQPPPPDLISPAAALSKPIFS
LASPPSPVELAPPSRELRKSKTCQITTDRPPLIPSITISESRSLNRIDRCRPVTVDGSG
LTPDRRPLVSRRSTMSRTMSLIPTSPSLPPLYEEETASTAAMVEEERLDKVQKRR
MRRYGGSHTTSTYQGRKDAAPDASPRRHLIGFLHRTSHLSLTSELSADASFYLIG
HRRHLEESTIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKMTKKKKNKRNFRK
PSRASSFSSITESSMSLDVITVNLNMDTVNFLGISIVGQTSNCGDNGIYVANIMKG
GAVALDGRIEAGDMILQVNETSFENFTNDQAVDVLREAVSRRGPIKLTVAKAFE
NGQSCFTIPRNSREEPVRPIDTQAWIQHTNAMRGMPSIVEESAPTPIPGEWPHGRP
PSSSTVTSNGSNGQNTVVGGGAHIHLDIHTDKKKVVEIMAMPGSGLDIKNRTWL
KIPIPMSFLGKVFMIDKPFSINFFLGSDLVEWLLDRIEGLRERKTARNYAAELLKL
KYIAHVVNKVTFTEQCYYVLGDECADYARFRNEDGGPKYQWTIGMNGMSAGN
GSSIMLPPPHLPGPPGAFKGMAPSMVSDGESRMYVMHI

>Cbre-DSH-2a
MTDSPSPIDSSFDISDVGTPATVIHKSIFRKEAEEEDFDLETDEDYTEQYREHDGV
DQSELSSSFLVDDYSKDCDSSISAPIPKPSFFRTITKVYYHVDDENIPYTADIHVPP
DCITLGDVKRKLPRTNFKYYCIALDPESGLEVKAEVRDDSQRLYPLRDGRFVLYL
LTIEGSVHSDTSSGRHRKNKLSSKGSNSSREYLKAAHHFDNPASYSDSESQASSIP
AYFKKAKAFNKRQAFQAHDRHHHHQLPRHKPHGRHHHNHYDEESTFDITTESD
DHYRDGITYYDEDEDDSRSINTDLTSVSQVALKAKWRQQQREMRNKYKRMPST
ASSTLSSITESSMGVEVITVRLNIQEFPIGMVPSILTTARGDDGGLYVGQVNPRGA
VALDGRIVVGDMISEINNIDLSNYSGKEAVNILKQAVTNQPYITLTVVKTGENKK
AAPAVLRNPRAEPIRPIDTNEWLKHATNAMKAMPSISEESCSTPIPDDWPTNSSAS
GTPFGGPPPNIHCLTVTTDKKDLVQAMMAPGSGLEIKNHEWLKILIPMSFLGKDL
VDWLLDHVQGLKNRDDSCKYAGKMLKEHYIVQPNGKKKFSENCYYVVGEKCG
DYTSLRGNDGEYKYAQSQTSSASGHSSNNNVFPPSMYPPPLPPSALGAHHRNSA
VLNSIGSGYASMTSSPLPSEKPSNCGRTRDDQRSQTSGSSRGSSRRYVELPRKPSS
LGSGSGVSDQINLDRVASRSSFRAAMSGSLRQFNIDS

>Cbre-DSH-2b
MTDSPSPIDSSFDISDVGTPATVIHKSIFRKEAEEEDFDLETDEDYTEQYREHDGV
DQSELSSSFLVDDYSKDCDSSISAPIPKPSFFRTITKVIIFVKVEFFNWVLGCITKSD
ENTAKTHLSAKEPTFFCVELGPFLTELFSWSIETSKKKVVTENSKSSNKTFFPRIEL

FYLNFQVYYHVDDENIPYTADIHVPPDCITLGDVKRKLPRTNFKYYCIALDPESG
LEVKAEVRDDSQRLYPLRDGRFVLYLLTIEGSVHSDTSSGRHRKNKLSSKGSNSS
REYLKAAHHFDNPASYSDSESQASSIPAYFKKAKAFNKRQAFQAHGTQPLSLFSS
FVSFSDRHHHHQLPRHKPHGRHHHNHYDEESTFDITTESDDHYRDGITYYDEDE
DDSRSINTDLTSVSQVALKAKWRQQQREMRNKYKRMPSTASSTLSSITESSMGV
EVITVRLNIQEFPIGMVPSILTTARGDDGGLYVGQVNPRGAVALDGRIVVGDMIS
EINNIDLSNYSGKEAVNILKQAVTNQPYITLTVVKTGENKKAAPAVLRNPRAEPI
RPIDTNEWLKHATNAMKAMPSISEESCSTPIPDDWPTNSSASGTPFGGPPPNIHCL
TVTTDKKDLVQAMMAPGSGLEIKNHEWLKILIPMSFLGKDLVDWLLDHVQGLK
NRDDSCKYAGKMLKEHYIVQPNGKKKFSENCYYVVGEKCGDYTSLRGNDGDY
KYAQSQTSSASGHSSNNNVFPPSMYPPPLPPSALGAHHRNSAVLNSIGSGYASMT
SSPLPTVLNSIGSGYASMTSSPLPSEKPSNCGRTRDDQRSQTSGSSRGSSRRYVELP
RKPSSLGSGSGISDQINLDRVASRSSFRAAMSGSLRQFNIDS

>Cbre-MIG-5a
MDAPCTSDTHQIKVFYYLDDETTPYVSVIDTREGVATLGNFKNSFTKRGYKYYG
KELDPDIQREVKVELISDSDRLRKSQNGFYEVFLVSTPGYGTLPRNTGTMTRPQR
AALDKRRRRSADFDATPYSDASLAPSTIVSRRAGEHLAELYTSNSEDPYQYDEHT
RRTTTDDSSMYEPLAARDMNRYHEEERRKKKQKKDHRFRRPYVPSTISSATESS
VNSGLPRILEIYLPMKNVPYLGLSVCTIDSHIFVSEIAPEGAVEKDGRVSCGDQILQ
VNRVSFEDLTATAAVKALRDAAASKRPITLYISKFVRGAPSEYDDPLASIASETM
PLDVGVWVETAVQNTEKMKALGLDPQEQTMTSVDDGTLPFTSTASDDEERILY
DQRRNGIPRALMEEAERKRENEQNEKIEQLTELIDPIIVVRAMARPDSGLVVKNR
KWLKILVPNSFIGRDLVYWLVDHMTDIHSRKHARLYAARLLAAGLIRHVVSKLT
FTEKCYYVFGDGILPPTATVNDRNSTDTSGTSATTMRVEATTEVTYVGSPAPHAL
ATRIGRNIPPHRLETTTLSPVAHDQTWLRRRRDCESPMTNDYASMVGESQLGMG
MNTGNYHGYVAKNPRVVPAPSQVTSSSLTNGSGGIGGPPPTPLSSTMVLAAVPSP
IQSSPNVALLMHDFDAENNSGNSKSSRILRA

>Cbre-MIG-5b
MTRPQRAALDKRRRRSADFDATPYSDASLAPSTIVSRRAGEHLAELYTSNSEDPY
QYDEHTRRTTTDDSSMYEPLAARDMNRYHEEERRKKKQKKDHRFRRPYVPSTIS
SATESSVNSGLPRILEIYLPMKNVPYLGLSVCTIDSHIFVSEIAPEGAVEKDGRVSC
GDQILQVNRVSFEDLTATAAVKALRDAAASKRPITLYISKFVRGAPSEYDDPLAS
MASETMPLDVGVWVETAVQNTEKMKALGLDPQEQTMTSVDDGTLPFTSTASD
DEERILYDQRRNGIPRALMEEAERKRENEQNEKIEQLTELIDPIIVVRAMARPDSG
LVVKNRKWLKILVPNSFIGRDLVYWLVDHMTDIHSRKHARLYAARLLAAGLIRH
VVSKLTFTEKCYYVFGDGILPPTATINDRNSTDTSGTSATTMRVEATTEVTYVGS
PAPHGLATRIGRNIPPHRLETTTLSPVAHDQTWLRRRRDCESPMTNDYASMVGES
QLGMGMNTGNYHGYVAKNPRVVPAPSQVTSSSLTNGSGGIGGPPPTPLSSTMVL
AAVPSPIQSSPNVALLMRDFDAENNSGNSKSSRILR

>Cjap-DSH-1a
MADPPVDELEKLENLRIADDPDSKEDDFDTKSGSSAQYSQASEATTAVKQQPFL
QQTMTKVYCHIDDETDPYMLEVHVPPDMITLGDLKRVLMRTNFKYYRKALDPD
SGYEVKAEIRDDSQRLAPSPNNLFELFLLTIEGSTHSDGSSGKLRKYPSVPGPAPS

NRNGPPMNYQHAAYQFDNSMMSTDSESMISAAVPGYLKNAYNRRFPAQYLELK
VQFFDLITKYRTTRQNRVISRVSQRCSL

>Cjap-DSH-1b1
MSNHGSTRHSDATLRALKFTAKERKVLSNYTAPPPNRHHSPGVVSQLVNKIEKK
SSPPLASSKRKKIPPLAEIDDLEMAVEDVVEVVKSPRRKKYTGNPLGNLKILNHL
LPGTSKDPKDEVPIKKLPENKENTATQTPTPSTGKRLKKSKSKGRVEPEDERFFRL
KRTMTDNTNNHSKSSPTGSGGWRQIVSSALGGPLKKRLSEGALLPSTSSGSMYP
GGEKKSKRGRLQRMKSAEGSASSSFFGALIRLSSSAVSLTSLTSLGSLGSNKSKPH
SPSNTKEFVESLRRSPPPPTLVSESDSVQLPTQPPPPLRKSKTCQNTPDRPPIIPSITIT
ESRSLNRIDRCRPITVDGSGLTPRSSRRSAMSRTMSLIPTSPSLPPLYEEETASTAA
MVDEERIAEKKHRMRRYGGSNTTSTYQGRKDVAPDASPRRHLIGFLHRTSHLSL
TSELSGNASFYLIGHRRHLEESTIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKM
AKKKKRNIKNFRKPSRASSFSSITESSMSLDVITVNLNMDTVNFLGISIVGQTSTG
GDNGIYVANIMKGGAVALDGRIEAGDMLLQVNDVSFENFTNDQAVDVLRGSTT
TDFTSVSRQHEKMAKKKKRNIKNFRKPSRASSFSSITESSMSLDVITVNLNMDTV
NFLGISIVGQTSTGGDNGIYVANIMKGGAVALDGRIEAGDMLLQVNDVSFENFT
NDQAVDVLREAVSRRGPIKLTVAKSFENGAQSCFTIPRNSREEPVRPIDTQAWIQ
HTNAMRGMPSIVEESAPTPIPGEWPHGRPPSSSTVTSNGSNGQNTVVGGGAHIHL
DVHTDKKKVVEIMAMPGSGLDIKNRTWLKIPIPMSFLGSDLVEWLLDHIDGLRE
RKTARNYAAELLKLKYIAHVVNKVTFTEQCYYVLGEECADYARFRNEDGGPKY
QWTIGMNNGVAASNGGSVMLPPPHLSGPGAHFKGMAPSMVSDGESRMYVMHI

>Cjap-DSH-2a
MTDSPSPIDSSFDASDIATPCTVIAAKTRNLRDLKIEEEDGEDSSHGDEEEVSAIYV
DDFSKTDFSEGESSVMEPLPRPPSFARTITKVYCHMDNEEVPYMVEVHVPPDCIT
LRDVKRKLTRTNYKFFCIALDPDSGLEVKAEIRDDSNKLYPLKDGRFQLFLLTIEG
SVHSDTSSGRHRKQKTSSKGSSSSREFRAGYEHASVMSEVSSDASSLPTYVKKAH
AYNRRHGAPPQYGDLRQHLLQKQRHRNIPYQQQQNPYEESSFDVTESDVYGGG
HHHHHQHRDGETFYDEDDDSRSINTDLTSVSQQHLKKMYREQQARAQNKWKS
AMSTTSSSFTDITESSMGVEIITVRLNLETLPLGMIPCGDTDSRGDSGLFVGSITDR
GAVALDGRIDIGDMILEINGVSLQNHTNQQAANLLKLRPLPCYFLL

>Cjap-DSH-2b
MHLESAVQRQFLTLTIAKTDKKKTAFLRNTRNEPVRPIDTNEWIKHATQHMKAM
PSISEESSSTPIPDEWPSHSSASGTPFGGPTPTINQLSVITDKKYVVEVMAAPGSGL
EIKDREWLKIPIPMSFLGRDLVDWLLDHIKGLTKREEACNFAGEMLKMGYIQHV
VNKKHFSEKCYYVMGEECADYTQLRAPDGGFKYPQSRESSTSNSTTNNNNNNV
FPAHMYPPNQNEVQSANSSQNHQRNSVVLPNGIPMQQSVSGYASMPSSPFPPKN
GGDCGRTRDDQRSQTSGSSRGSSRRYVELPRKPSSQGSGSAHENSLMDRVASRS
SFRAAMSGSLRQFNIDG

>Cjap-MIG-5
MEQPCTSESSQIKVFYYLDDETTPYVSVIDTQDGVATLGNFKNSFTKRGYKYYG
KELDPDIQREVKVELTLDSDRLRRSQNGFYEVFLVSTPGYGTLPYSTMTRTQRTA
LDKRRRRSADFDAQPYSDASLAPSTIVSKEIMIIYLLSARENDQTRSAELSAXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXNSEDPYDDSAHRTGDTMYEPLATRDMNQIYEDDRR
QKRKPKKERFRRPYVPSTISSATESSVGSGLPRILEIYLPMKNVPYLGLSVCTIDGH
IFVSEIAAEGAVEKDGRMNVGDQILQVNRISFEELGGPQAVRALRDAAASGRPIT
LYISKYTKGAASEYDDPLASMASETMPLDVGVWVETAVQNTEKMKALGLDPQE
QTITTIDDGTLPFTSTASDDEEKLLYDQRRNGIPRVVHEDVERKKKESENEHLTEL
IDPMIVVRNMARPDSGLVVKNRKWLKILVPMSFIGRDLICWLMEHMTDIHSRKQ
ARVYAARLLAAGLIRHVVSKLTFTEKCYYVFADGILPEGTVDRNSTGTTSTRAD
ATTEVTYVGSPAPRGLALGARRGIPPHRLETTTLSPVAHDQTWLRRRRDCESPMT
NDYASMVGESQLGLNPTHIFGDQKKNHQRPTTTTMTAASQVTSSSLTNGSGGIG
GPPPTPLSSTMVLASPTLSHTCIADSEAGDGAKSRSKIIRS

>Crem-DSH-1
MSIKLEARERAIPTGSNTPSTSTSSDVLRSLLFTSKERKVLSTYTAPPPNRHHSPGV
VSQLINKIEQKTSPPTTSSRPKKKKKAPPLAQIDDLDLIEDDKTAVEVVKSPRRKK
YTGNPLGNLKILNNILPLSKTSEDKENSNQPTVGKKLKKSKSKSRVEPESEPPFRL
KRTLTDLTNNHSPAGSTWRQMVSNALGGPLKKRLSEGALFFPGTSKTADDCDDE
EGDNTKTTTKRRFLQRRKQSEKSAGGANGKESGGSASSSFFGALIRLSHSAASLT
SLTSLGGSRSNSASPSSSRSNTKEFKEELKPPQPPPPDLISPVAALSKPIFTLDSSHP
VSLTPPSRELRKSKTCQITTDRPPIIPSITISESRSLNRIDRCRPVTVDGSGLTPDRRP
LVSRRSTMSRTMSLIPTSPSLPPLYEEETASTAAMVEEEREDKAQKRRMRRYGGS
NTTSTYQGRKDVAPDASPRRHLIGFLHRTSHLSLTSELSADASFYLIGHRRHLEES
TIGSESDARVFSDDDDRGSTTTDFTSVSRQHEKMAKKKKNKRNFRKPSRASSFSS
ITESSMSLDVITVNLNMDTVNFLGISIVGQTSNCGDNGIYVANIMKGGAVALDGR
IEAGDMILQVNETSFENFTNDQAVDVLREAVSRRGPIKLTVAKSFENGQSCFTIPR
NSREEPVRPIDTQAWIQHTNAMRGMPSIVEESAPTPIPGEWPHGRPPSSSTVTSNG
SNGQNTVVGNGTHIHLDIHTDKKKVVEIMAMPGSGLDIKNRTWLKIPIPMSFLGK
YLFPVHIIIFFYSGSDLVEWLLDRIEGLRERKSARNYAADLLKLKYIAHVVNKVTF
TEQCYYVLGDECSGKRNEELTFIQSKYSDYARFRNEDGGPKYQWTMGMNGMS
AGNGSSVMLPPPHLPGGVPPGAFKGMAPSMVSDGESRMYVMHI

>Crem-DSH-2
MTDSPSPIDSSLDYSDVATPCTVIAAKCSIRNQKDLENEEDDLENQEDYTESFQQ
QEGDSVEQEDLSHIYVDDLSKDFSDAASSVMEPLPKPHTFARTITKVYYHLDDET
VPYMVDVHVPPDCITLRDVKRKLPRTNFKYYCIALDPESGREVKAEIRDDSQRLY
PLRCGKFELYLLTVEGSVHSDTSSGRHRKKHQTSSKGSSSSREYQRAAHHYDNP
TPYSDNESQASSIPTYVKKAHAFNRRQASQAYDRHQPRHRLHERHHQNHYDDST
FDVTTESDDHYRDGVTYYDEDEDDSRSINTDLTSVSQIHLKQRWKQQQQREARN
KWKRMPSMSTASSSLSSITESSMGLELLTVRLNLQTMPLGMVPYGLKTARGGDA
GLYVGDILDRGAVALDGRIDVGDMISEINNIDLSNYSNEAAAQLLRDAVAPRQF
VTLTIAKSIDSRKAVAAAFTKNTRAEPTRPIDTNEWLKHATNAMKAMPSISEESC
STPIPDEWPTNSSASGTPFGGPPPPSIACMNTSTNKKFVVEVMAAPGSGLEIKDRE
WLKIPIPMSFLGKDLVDWLLDHIQGLRKRGEAGKFAGEMLKLGYIQHVLNKNKF
SENCYYIMGEECADYTQLRAPDGGFKYAQSQTSSASAHSSNNNIFPPSMYPSQTQ
PSSAAGVNAHHRNSAILNSMVSGYASMPSSPFPNSKPAVGDCGRTRDDQRSQTS
GSSQGSSRR

>Crem-MIG-5
MEPPCTSDSNQIKVFYYLDDETTPYVSVIDTREGVATLGNFKNSFTKRGYKYYG
KELDPDIQREVKVELTSDSDRLRKSQNGFFEVFLVSTPGYGTLPRNTGTMTRTQR
TALDKRRRRSADFDATPYSDASLAPSTIVSRRAGEHLAELYTSNSEDPYQYDEHT
RRTIDDSSIYEPLGTRDMNKFHDDDRRKRKQKKERFRRPYVPSTISSATESSVNSG
LPRILEIYLPMKNVPYLGLSVCTMDGHIFVSEIAPEGAVEKDGRVNVGDQILQVN
RVSFEDLSGPQAVRALRDAAASKRPITLYISKFARGAPSEYDDPLASMASETMPL
DVGVWVETAVQNTEKMKALGLDPQEQTMTSVDDGTLPFTSTASDDEERILYDQ
RRNGIPRALLEEAERKKENERNEKAEQLTELIDPIIVVRAMARPDSGLVVKNRKW
LKILVPMSFIGCDLIDWLVEHMTDIHSRKHARLYAARLLAAGLIRHVVSKLTFTE
KCYYVFGDGILSTDRNSTDTSGTSGTTMRVEATTEVTYVGSPAPHAVATRIGRNI
PPHRLETTTLSPVAHDQTWLRRRRDCESPMTNDYASMVGESQIGMNPAGHYNP
YATKNNRQVPAPSQVTTSSLTNGEKLNPLAVLFSTNSVTNSRKRWYWRAPTDTS
VQYYGSSSVSDPITEHHQPRFRRGE

REFERENCES

Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., Craigie, A., Larroux, C. and Degnan, B. M. (2007). Wnt and TGF-beta expression in the sponge Amphimedon queenslandica and the origin of metazoan embryonic patterning. *Plos One* **2**, e1031.
Adamska, M., Larroux, C., Adamski, M., Green, K., Lovas, E., Koop, D., Richards, G. S., Zwafink, C. and Degnan, B. M. (2010). Structure and expression of conserved Wnt pathway components in the demosponge Amphimedon queenslandica. *Evolution & development* **12**, 494-518.
Almuedo-Castillo, M., Salo, E. and Adell, T. (2011). Dishevelled is essential for neural connectivity and planar cell polarity in planarians. *Proc Natl Acad Sci U S A* **108**, 2813-2818.
Ayala, F. J., Rzhetsky, A. and Ayala, F. J. (1998). Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc Natl Acad Sci U S A* **95**, 606-611.
Barrière, A., Yang, S.-P., Pekarek, E., Thomas, C. G., Haag, E. S. and Ruvinsky, I. (2009). Detecting heterozygosity in shotgun genome assemblies: Lessons from obligately outcrossing nematodes. *Genome Res*. **19**, 470-480.
Blaxter, M. (1998). *Caenorhabditis elegans* is a nematode. *Science* **282**, 2041-2046.
Blaxter, M. (2011). Nematodes: The worm and its relatives. *PLoS Biology* **9**, 1-9.
Blaxter, M. L., P. De Ley, J. R. Garey, L. X. Liu, P. Scheldeman et al. (1998) A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**: 71–75.
Boutros, M. and Mlodzik, M. (1999). Dishevelled: at the crossroads of divergent intracellular signaling pathways. *Mechanisms of development* **83**, 27-37.
Bumbarger, D. J., Crum, J., Ellisman, M. H. and Baldwin, J. G. (2007). Three-dimensional fine structural reconstruction of the nose sensory structures of Acrobeles complexus compared to Caenorhabditis elegans (Nematoda : Rhabditida). *J. Morphol.* **268**, 649-663.
Bumbarger, D. J., Wijeratne, S., Carter, C., Crum, J., Ellisman, M. H. and Baldwin, J. G. (2009). Three-dimensional reconstruction of the amphid sensilla in the microbial feeding nematode, Acrobeles complexus (Nematoda: Rhabditida). *J Comp Neurol* **512**, 271-281.
***C. elegans* Sequencing Consortium.** (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012-2018.
Darriba, D., Taboada, G. L., Doallo, R. and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* **9**, 772.
Dillman, A. R., Mortazavi, A. and Sternberg, P. W. (2012). Incorporating genomics into the toolkit of nematology. *J. Nematol.* **44**, 191-205.
Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics* **5**, 113.
Etheridge, S. L., Ray, S., Li, S., Hamblet, N. S., Lijam, N., Tsang, M., Greer, J., Kardos, N., Wang, J., Sussman, D. J. et al. (2008). Murine dishevelled 3 functions in redundant pathways with dishevelled 1 and 2 in normal cardiac outflow tract, cochlea, and neural tube development. *Plos Genet* **4**, e1000259.
Fahmy, O.G. and Fahmy, M. (1959). New mutants report. *Dros. Inf. Serv.* **33**, 82-94.

Gao, C. and Chen, Y. G. (2010). Dishevelled: The hub of Wnt signaling. *Cellular signalling* **22**, 717-727.
Goloboff, P. A. (1999). Analyzing large data sets in reasonable times: Solutions for composite optima. *Cladistics-the International Journal of the Willi Hennig Society* **15**, 415-428.
Gray, R. S., Bayly, R. D., Green, S. A., Agarwala, S., Lowe, C. J. and Wallingford, J. B. (2009). Diversification of the expression patterns and developmental functions of the dishevelled gene family during chordate evolution. *Developmental dynamics : an official publication of the American Association of Anatomists* **238**, 2044-2057.
Grove, C. A., De Masi, F., Barrasa, M. I., Newburger, D. E., Alkema, M. J., Bulyk, M. L. and Walhout, A. J. (2009). A multiparameter network reveals extensive divergence between C. elegans bHLH transcription factors. *Cell* **138**, 314-327.
Guindon, S. and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704.
Gurley, K. A., Rink, J. C. and Sanchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science* **319**, 323-327.
Hallem, E. A. and Sternberg, P. W. (2008). Acute carbon dioxide avoidance in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **105**, 8038-8043.
Hallem, E. A., Dillman, A. R., Hong, A. V., Zhang, Y., Yano, J. M., DeMarco, S. F. and Sternberg, P. W. (2011). A sensory code for host seeking in parasitic nematodes. *Current Biology* **21**, 377-383.
Hallem, E. A., Spencer, W. C., McWhirter, R. D., Zeller, G., Henz, S. R., Ratsch, G., Miller, D. M., 3rd, Horvitz, H. R., Sternberg, P. W. and Ringstad, N. (2011). Receptor-type guanylate cyclase is required for carbon dioxide sensation by Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **108**, 254-259.
Hamblet, N. S., Lijam, N., Ruiz-Lozano, P., Wang, J. B., Yang, Y. S., Luo, Z. G., Mei, L., Chien, K. R., Sussman, D. J. and Wynshaw-Boris, A. (2002). Dishevelled 2 is essential for cardiac outflow tract development, somite segmentation and neural tube closure. *Development* **129**, 5827-5838.
Herman, M. A., Vassilieva, L. L., Horvitz, H. R., Shaw, J. E. and Herman, R. K. (1995). The C. elegans gene lin-44, which controls the polarity of certain asymmetric cell divisions, encodes a Wnt protein and acts cell nonautonomously. *Cell* **83**, 101-110.
Hingwing, K., Lee, S., Nykilchuk, L., Walston, T., Hardin, J. and Hawkins, N. (2009). CWN-1 functions with DSH-2 to regulate C. elegans asymmetric neuroblast division in a beta-catenin independent Wnt pathway. *Dev Biol* **328**, 245-256.
Holstein, T. W. (2012). The evolution of the Wnt pathway. *Cold Spring Harbor perspectives in biology* **4**, a007922.
Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J. and Helder, J. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution* **23**, 1792-1800.
Itoh, K., Brott, B. K., Bae, G. U., Ratcliffe, M. J. and Sokol, S. Y. (2005). Nuclear localization is required for Dishevelled function in Wnt/beta-catenin signaling. *Journal of biology* **4**, 3.

Kiontke, K. C., Felix, M. A., Ailion, M., Rockman, M. V., Braendle, C., Penigault, J. B. and Fitch, D. H. (2011). A phylogeny and molecular barcodes for Caenorhabditis, with numerous new species from rotting fruits. *BMC evolutionary biology* **11**, 339.
Klingensmith, J., Nusse, R. and Perrimon, N. (1994). The Drosophila segment polarity gene dishevelled encodes a novel protein required for response to the wingless signal. *Genes Dev* **8**, 118-130.
Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H. A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M. Q. et al. (2005). Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* **433**, 156-160.
Letunic, I., Doerks, T. and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**, D302-305.
Li, L., Stoeckert, C. J. and Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res*. **13**, 2178-2189.
Lijam, N., Paylor, R., McDonald, M. P., Crawley, J. N., Deng, C. X., Herrup, K., Stevens, K. E., Maccaferri, G., McBain, C. J., Sussman, D. J. et al. (1997). Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. *Cell* **90**, 895-905.
Luo, J., Chen, J., Chen, Z., Deng, I., Luo, X., Song, W. X., et al. (2007). Wnt signaling and human diseases: what are the therapeutic implications? *Lab*. *Invest*. **87**: 97–103.
Monsoro-Burq, A. H., Wang, E. and Harland, R. (2005). Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction. *Dev Cell* **8**, 167-178.
Mortazavi, A., Schwarz, E. M., Williams, B. A., Schaeffer, L., Antoshechkin, I., Wold, B. and Sternberg, P. W. (2010). Scaffolding a *Caenorhabditis* nematode genome with RNA-seq. *Genome Res*. **20**, 1740-1747.
Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K. and Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *Plos Genet* **8**, e1002764.
Nicholas, K. B., Nicholas, H. B. and Deerfield, D. W. (1997). GeneDoc: Analysis and visualization of genetic variation. *EMBnet.news* **4**, 1-4.
Nixon, K. C. (1999). The Parsimony Ratchet, a new method for rapid parsimony analysis. *Cladistics-the International Journal of the Willi Hennig Society* **15**, 407-414.
Penton, A., Wodarz, A. and Nusse, R. (2002). A mutational analysis of dishevelled in Drosophila defines novel domains in the dishevelled protein as well as novel suppressing alleles of axin. *Genetics* **161**, 747-762.
Ragsdale, E. J., Ngo, P. T., Crum, J., Ellisman, M. H. and Baldwin, J. G. (2009). Comparative, three-dimensional anterior sensory reconstruction of Aphelenchus avenae (nematoda: Tylenchomorpha). *J Comp Neurol* **517**, 616-632.
Rodriguez-Trelles, F., Tarrio, R. and Ayala, F. J. (2002). A methodological bias toward overestimation of molecular evolutionary time scales. *Proc Natl Acad Sci U S A* **99**, 8112-8115.
Ruvkun, G. and Hobert, O. (1998). The taxonomy of developmental control in Caenorhabditis elegans. *Science* **282**, 2033-2041.
Sawa, H., Lobel, L. and Horvitz, H. R. (1996). The Caenorhabditis elegans gene lin-17, which is required for certain asymmetric cell divisions, encodes a putative seven-

transmembrane protein similar to the Drosophila frizzled protein. *Genes Dev* **10**, 2189-2197.
Seifert, J.R. and Mlodzik, M. (2007). Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. *Nat. Rev. Genet.* **8**, 126-138.
Simons, M. and Mlodzik, M. (2008). Planar cell polarity signaling: from fly development to human disease. *Annu. Rev. Genet.* **42**, 517-540.
Singh, J., Yanfeng, W. A., Grumolato, L., Aaronson, S. A. and Mlodzik, M. (2010). Abelson family kinases regulate Frizzled planar cell polarity signaling via Dsh phosphorylation. *Genes Dev* **24**, 2157-2168.
Song, S., Zhang, B., Sun, H., Li, X., Xiang, Y., Liu, Z., Huang, X. and Ding, M. (2010). A Wnt-Frz/Ror-Dsh pathway regulates neurite outgrowth in Caenorhabditis elegans. *Plos Genet* **6**.
Stamatakis, A., Hoover, P. and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* **57**, 758-771.
Strutt, D., Madder, D., Chaudhary, V. and Artymiuk, P. J. (2012). Structure-function Dissection of the Frizzled Receptor in Drosophila melanogaster Suggests Different Mechanisms of Action in Planar Polarity and Canonical Wnt Signaling. *Genetics*.
Sugimura, R., He, X. C., Venkatraman, A., Arai, F., Box, A., Semerad, C., Haug, J. S., Peng, L., Zhong, X. B., Suda, T. et al. (2012). Noncanonical Wnt signaling maintains hematopoietic stem cells in the niche. *Cell* **150**, 351-365.
Sweetman, D., Wagstaff, L., Cooper, O., Weijer, C. and Munsterberg, A. (2008). The migration of paraxial and lateral plate mesoderm cells emerging from the late primitive streak is controlled by different Wnt signals. *BMC developmental biology* **8**, 63.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731-2739.
Tauriello, D. V., Jordens, I., Kirchner, K., Slootstra, J. W., Kruitwagen, T., Bouwman, B. A., Noutsou, M., Rudiger, S. G., Schwamborn, K., Schambony, A. et al. (2012). Wnt/beta-catenin signaling requires interaction of the Dishevelled DEP domain and C terminus with a discontinuous motif in Frizzled. *Proc Natl Acad Sci U S A* **109**, E812-820.
Torres, M. A. and Nelson, W. J. (2000). Colocalization and redistribution of dishevelled and actin during Wnt-induced mesenchymal morphogenesis. *J Cell Biol* **149**, 1433-1442.
Wallingford, J. B. and Habas, R. (2005). The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity. *Development* **132**, 4421-4436.
Wernersson, R. and Pedersen, A. G. (2003). RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res* **31**, 3537-3539.
Wharton, K. A., Jr. (2003). Runnin' with the Dvl: proteins that associate with Dsh/Dvl and their significance to Wnt signal transduction. *Dev Biol* **253**, 1-17.
Wu, M. and Herman, M. A. (2006). A novel noncanonical Wnt pathway is involved in the regulation of the asymmetric B cell division in C. elegans. *Dev Biol* **293**, 316-329.
Yanfeng, W. A., Berhane, H., Mola, M., Singh, J., Jenny, A. and Mlodzik, M. (2011). Functional dissection of phosphorylation of Disheveled in Drosophila. *Dev Biol* **360**, 132-142.

Table 1

Table 1 Dsh abbreviations

Species	Abbreviation(s)
Nematoda	
Meloidogyne incognita	*Minc-dsh-1, Minc-mig-5*
Meloidogyne hapla	*Minc-dsh-1, Minc-mig-5*
Panagrellus redivivus	*Pred-dsh-1, Pred-mig-5*
Bursaphelenchus xylophilus	*Bxyl-dsh-1, Bxyl-mig-5*
Steinernema carpocapsae	*Scar-dsh-1, Scar-mig-5*
Caenorhabditis briggsae	*Cbri-dsh-1, Cbri-dsh-2, Cbri-mig-5*
Caenorhabditis remanei	*Crem-dsh-1, Crem-dsh-1, Crem-mig-5*
Caenorhabditis brenneri	*Cbre-dsh-1, Cbre-dsh-2, Cbre-mig-5*
Caenorhabditis elegans	*Cele-dsh-1, Cele-dsh-2, Cele-mig-5*
Caenorhabditis japonica	*Cjap-dsh-1, Cjap-dsh-2, Cjap-mig-5*
Caenorhabditis angaria	*Cang-dsh-1, Cang-mig-5*
Pristionchus pacificus	*Ppac-dsh-1, Ppac-mig-5*
Brugia malayi	*Bmal-dsh-1, Bmal-mig-5*
Ascaris suum	*Asuu-dsh-1, Asuu-mig-5*
Trichinella spiralis	*Tspi-dsh-1*
Arthropoda	
Nasonia vitripennis (parasitoid wasp)	*Nvit-dsh*
Tribolium castaneum (red flour beetle)	*Tcas-dsh*
Drosophila melanogaster (fruit fly)	*Dmel-dsh*
Platyhelminthes	
Schmidtea mediterranea† (planaria)	*Smed-dvl-1, Smed-dvl-2*
Chordata	
Xenopus tropicalis	*Xtro-Dvl1, Xtro-Dvl2, Xtro-Dvl3*
Mus musculus	*Mmus-Dvl1, Mmus-Dvl2, Mmus-Dvl3*
Homo sapiens	*Hsap-Dvl1, Hsap-Dvl2, Hsap-Dvl3*
Ciona intestinalis† (sea squirt)	*Cint-Dvl*
Cnidaria	
Clytia hemisphaerica† (jellyfish)	*Chem-Dvl*
Porifera	
Amphimedon queenslandica† (sponge)	*Aque-Dvl*

† Dsh orthologs for these taxa were taken from Genbank rather than identified through whole genome searches.

Table 1. Dishevelled nomenclature

Comprehensive list of all dishevelled abbreviations used throughout.

Figure 1

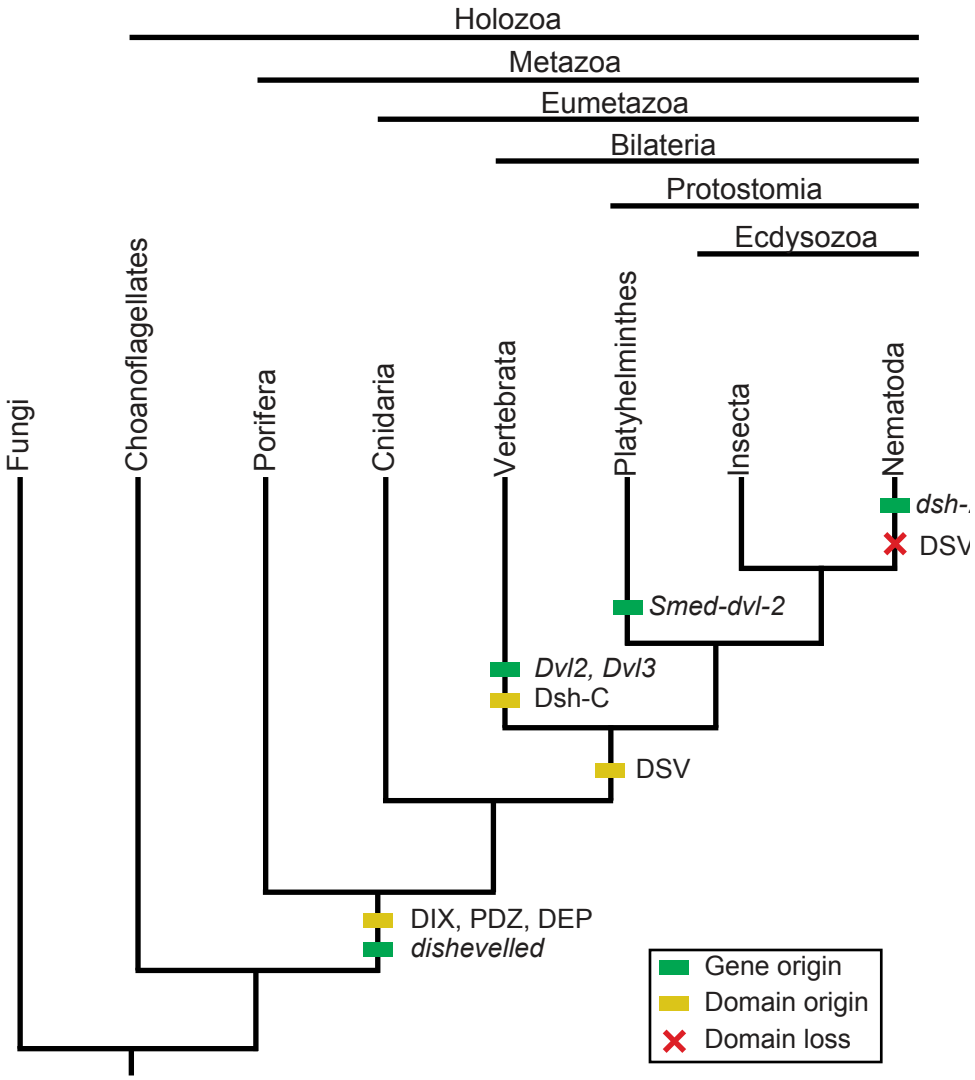


Fig. 1. Dishevelled cladogram
A cladogram showing key features during the evolution of Dsh among animals including the origin of Dsh and lineage-specific paralogs as well as the gain or loss of protein domains.

Figure 2

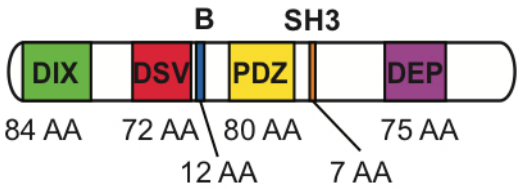


Fig. 2. Dishevelled architecture
Diagram showing the archetypal Dsh protein with conserved domains and motifs. From left to right: the DIX domain, the DSV or dishevelled domain, the basic region, the PDZ domain, the SH3 binding motif, which is often referred to as the proline-rich region, and the DEP domain.

Figure 3

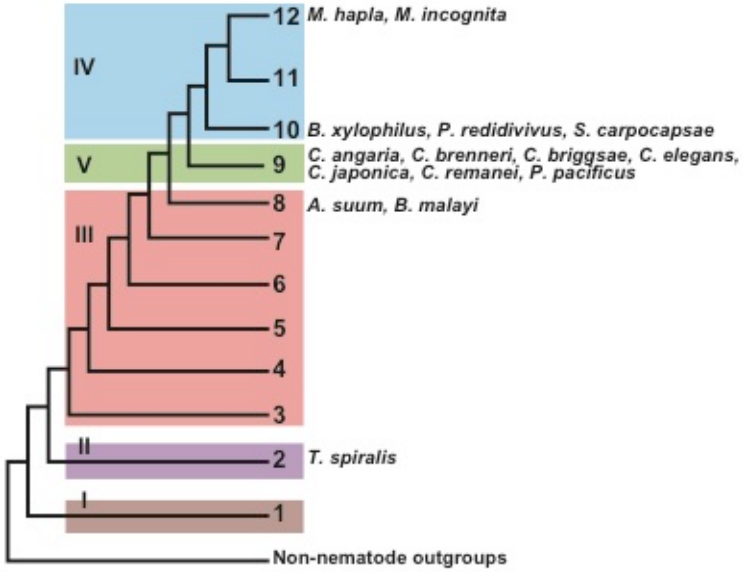


Fig. 3. Cladogram of Nematoda

A schematic representation of the division of the phylum Nematoda into clades, with the 12-clade designation after Holterman et al. 2006 (Holterman et al., 2006) and the five-clade designation after Blaxter et al. 1998 (Blaxter et al., 1998) in Roman numerals. Blaxter clades are encompassed in colored boxes.

Figure 4

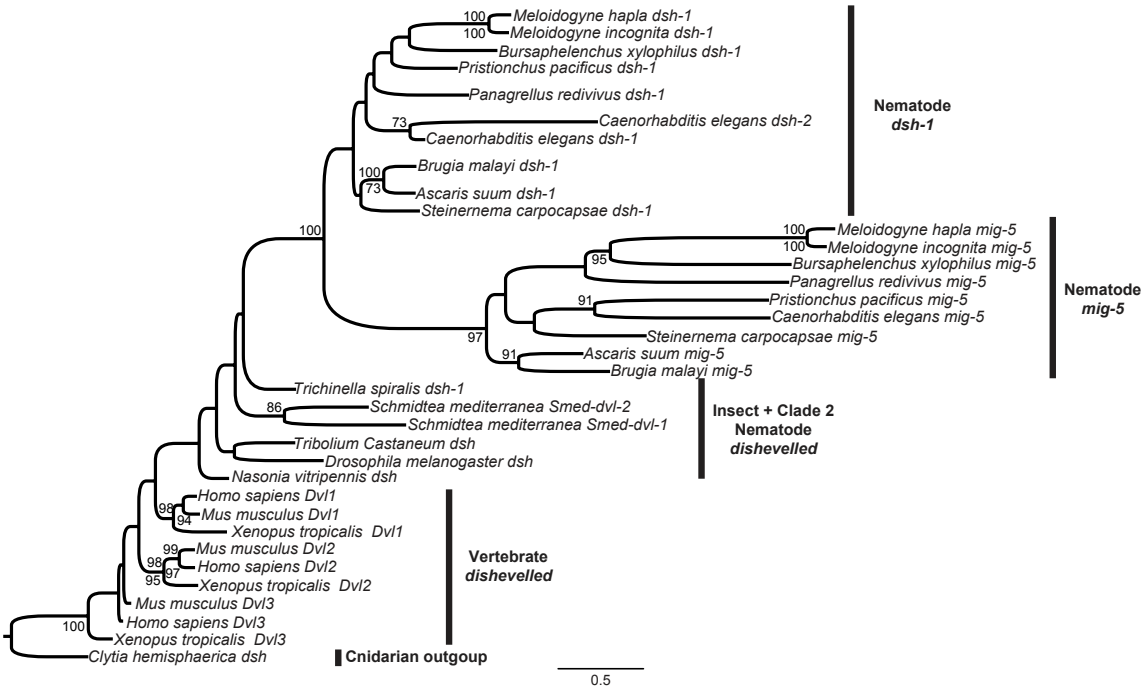


Fig. 4. Analysis of dishevelled phylogeny

Phylogenetic analysis of Dsh orthologs across animals based on the protein coding nucleotide alignment from the N-terminus of the PDZ domain through the C-terminus of the DEP domain. The ML tree (rooted with the outgroup *C. hemisphaerica*) is shown. For each node, ML bootstrap support values (1000 replicates) are above the nodes whereas parsimony bootstrap values (1000 replicates) are written below. Support values #70 are not shown.

Figure 5

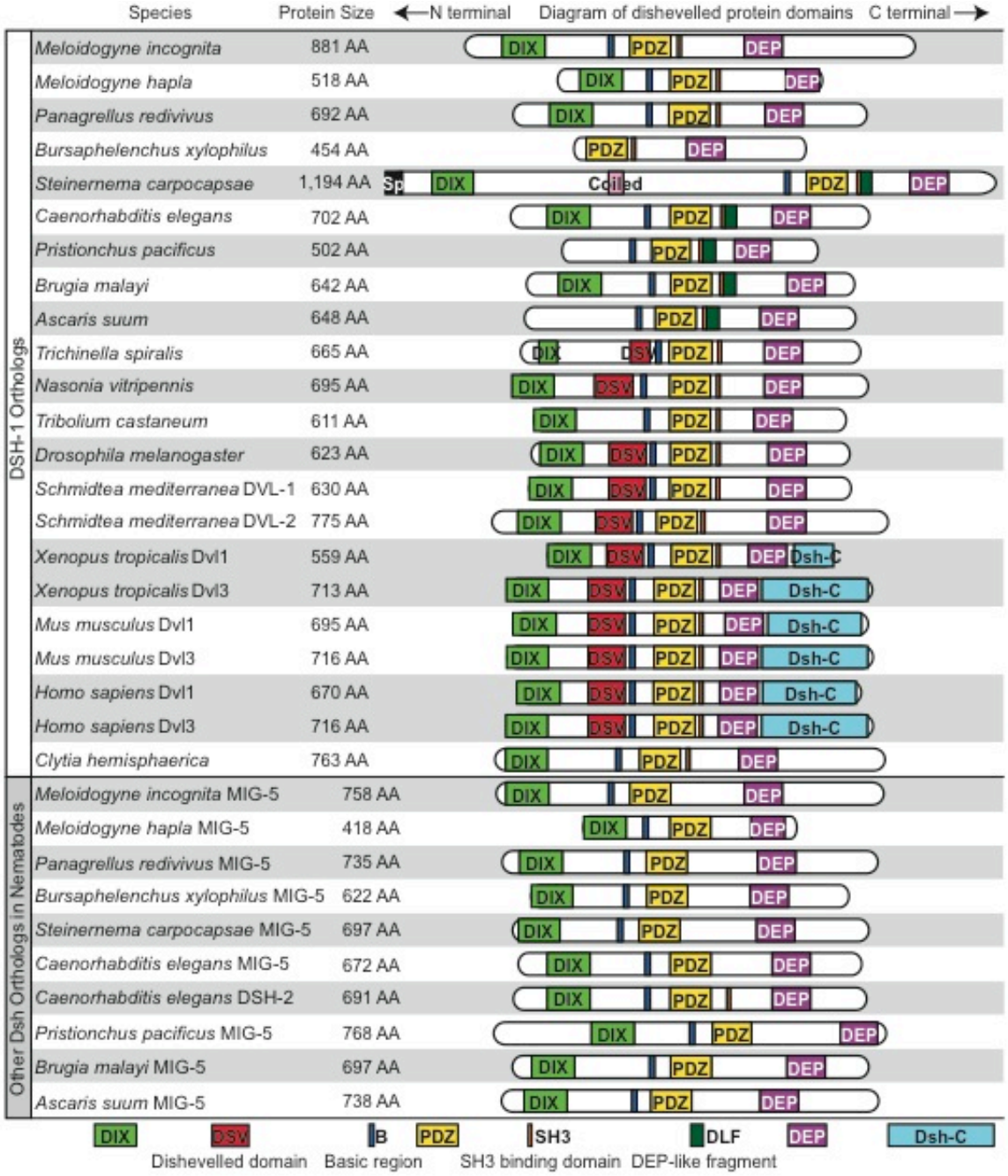


Fig. 5. Dishevelled schematic
Schematic diagram of Dsh orthologs in selected animal species. Proteins and their domains are drawn in proportion to the number of amino acids they contain.

Figure 6

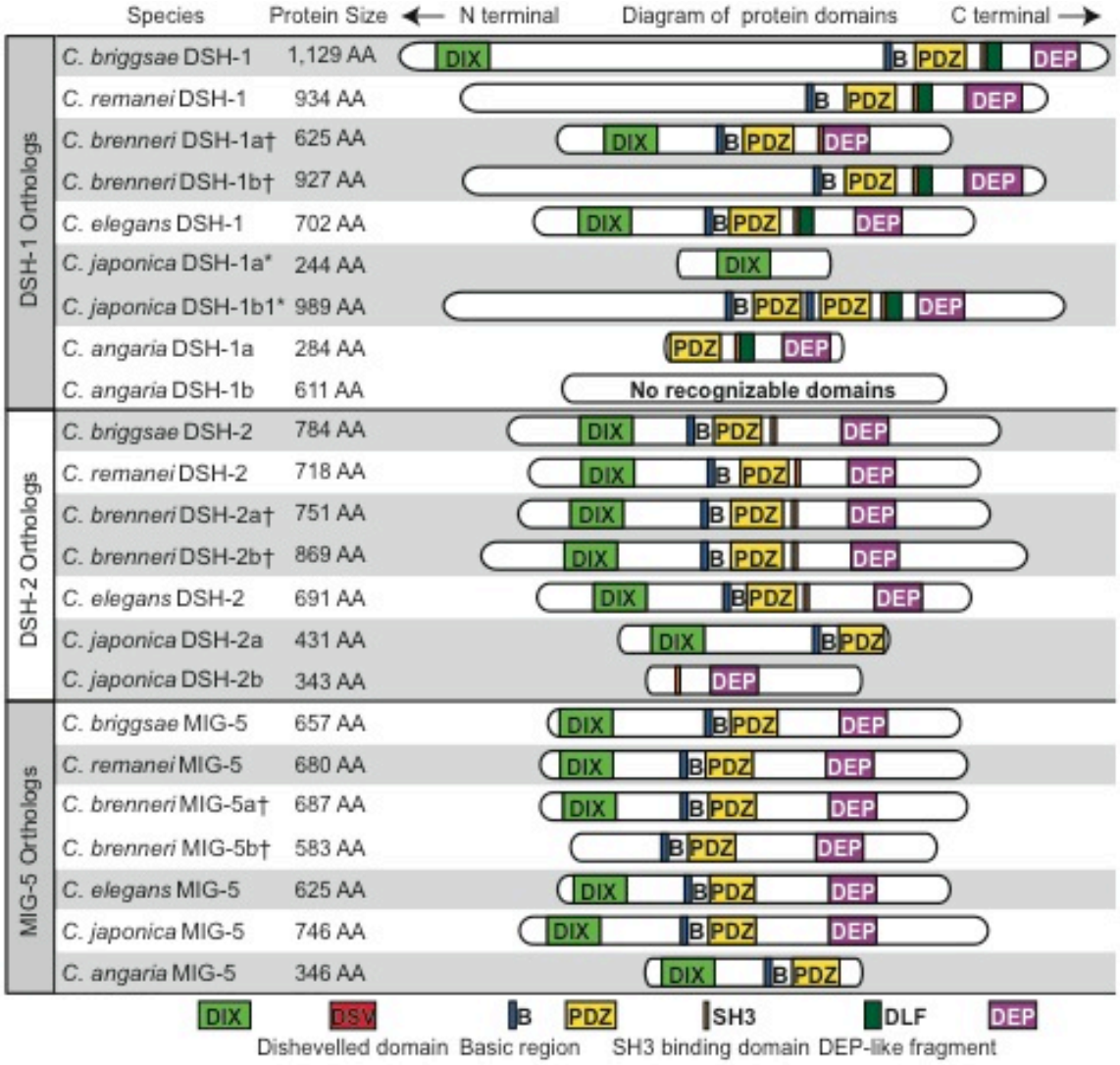


Fig. 6. Dishevelled schematic across *Caenorhabditis*
Schematic diagram of Dsh orthologs in sequenced *Caenorhabditis* species. Proteins and their domains are drawn in proportion to the number of amino acids they contain. These *C. brenneri* proteins are thought to be splice isoforms or per- haps allelic variants and not paralogous duplicates. Cjap-DSH-1a and Cjap-DSH-1b although presently annotated as separate genes, we suggest they are fragments of the same protein rather than two different orthologs of DSH-1. This is not the case with Cjap-DSH-2a and Cjap-DSH-2b, which likely are separate proteins.

Figure 7

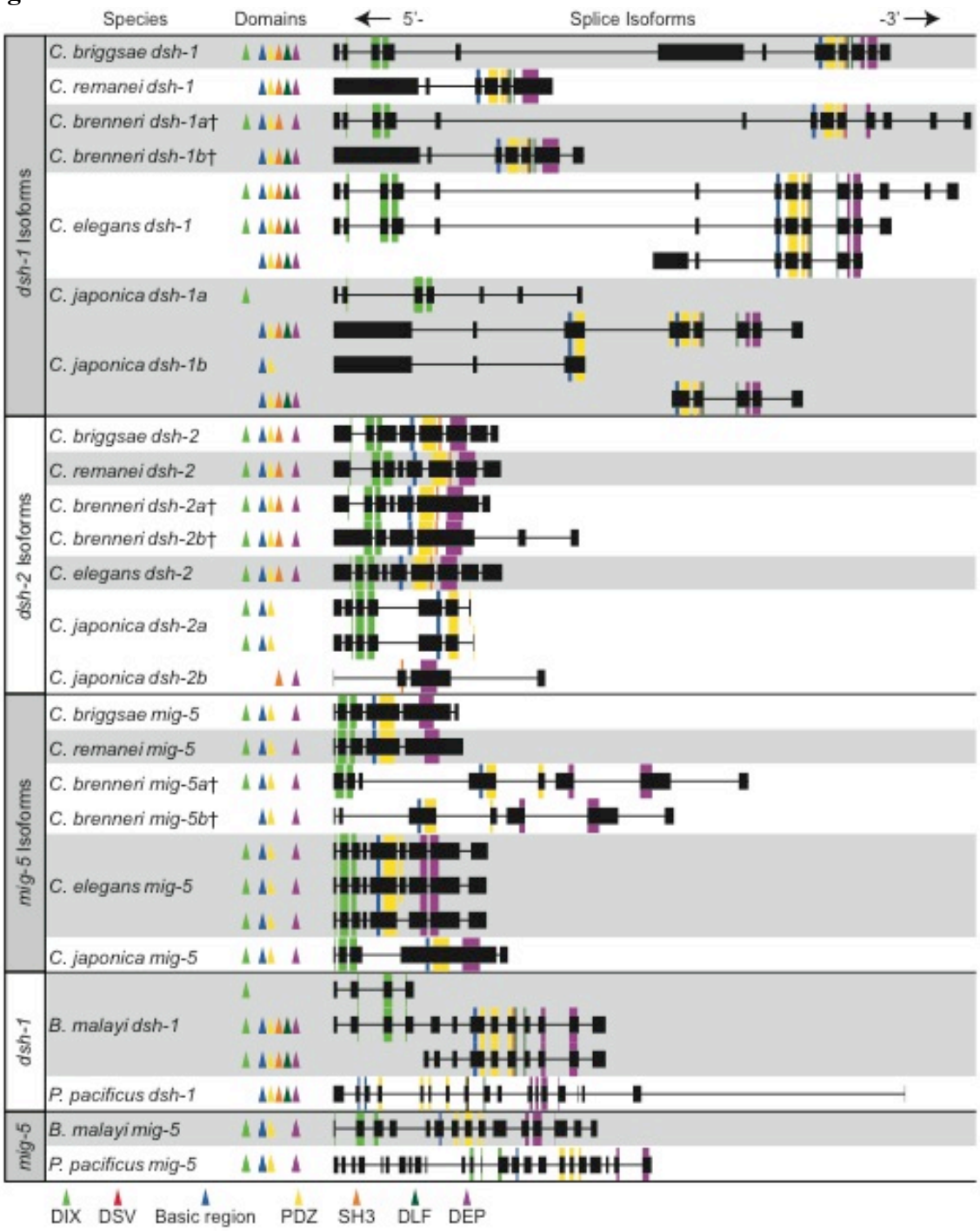


Fig. 7. Dishevelled splice isoforms

The known Dsh splice isoform architecture for *dsh-1*, *dsh-2*, and *mig-5* among caenorhabditids. All features (exons, introns, and domains) are drawn in proportion to the number of nucleotides they contain. All isoforms are shown in the same orientation, regardless of their actual orientation in their respective genomes. The known isoforms of *Bmal-dsh-1*, *Bmal-mig-5*, *Ppac-dsh-1*, and *Ppac-mig-5* are included at the bottom for outgroup comparison.

Figure 8

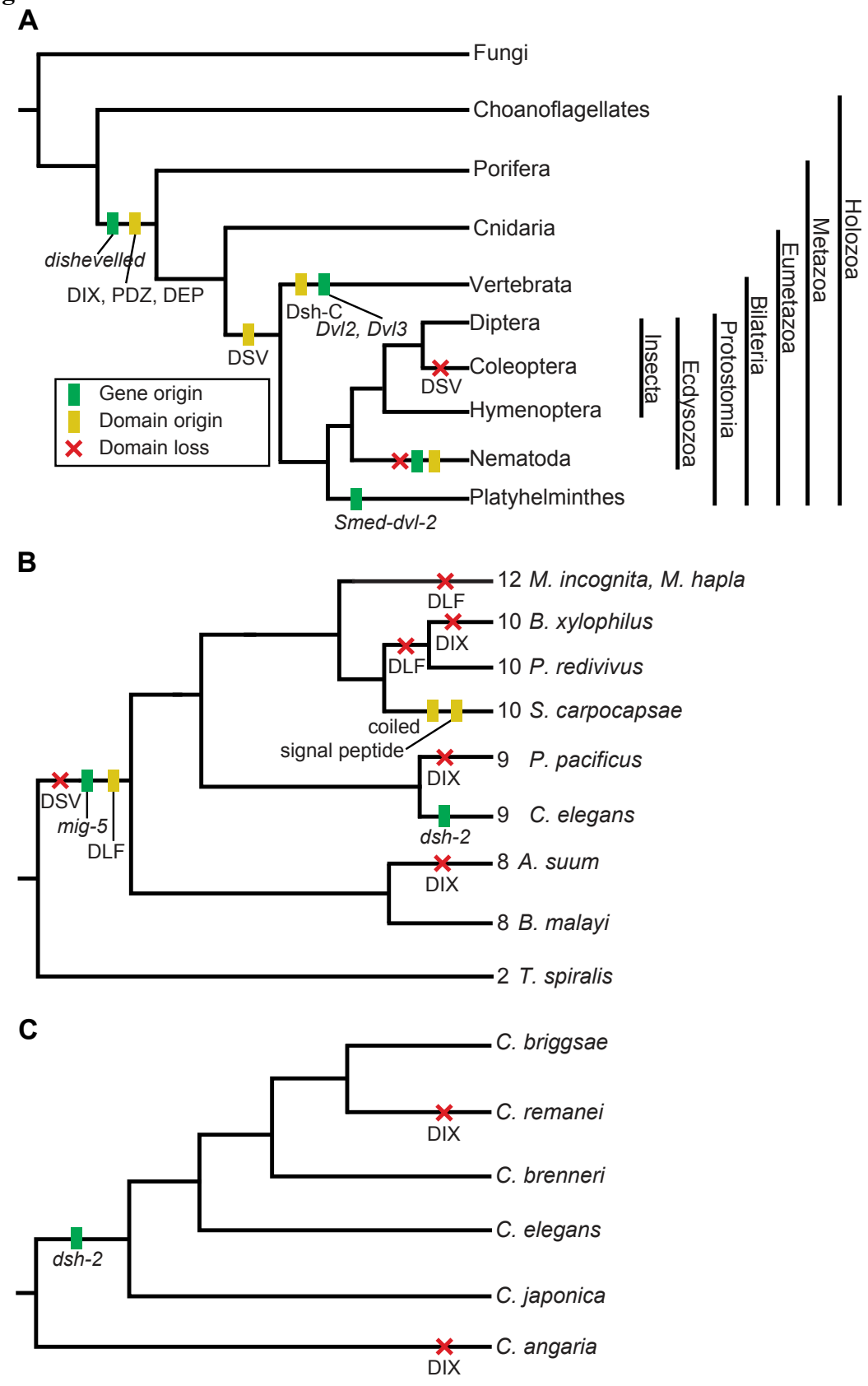
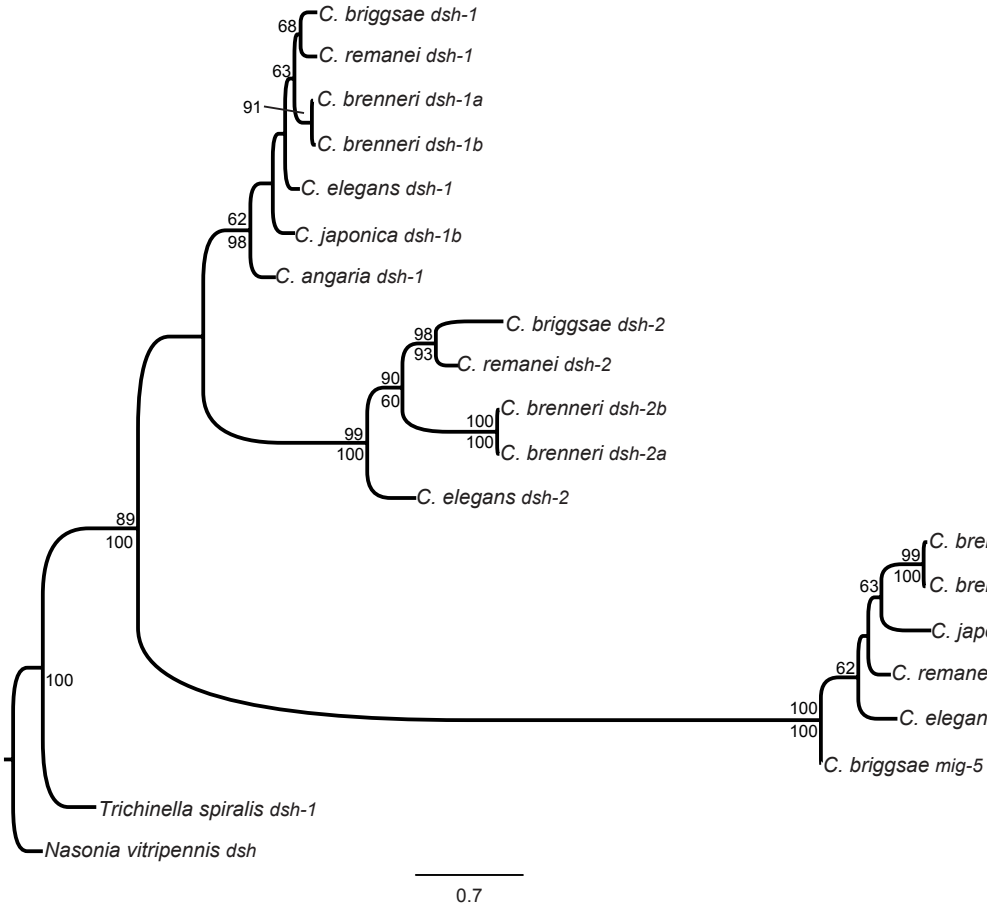

A
Fungi
Choanoflagellates
Porifera
Cnidaria
Vertebrata
Diptera
Coleoptera
Hymenoptera
Nematoda
Platyhelminthes
dishevelled
DIX, PDZ, DEP
DSV
Dsh-C
Dvl2, Dvl3
DSV
Smed-dvl-2
Gene origin
Domain origin
Domain loss
Holozoa
Metazoa
Eumetazoa
Bilateria
Protostomia
Ecdysozoa
Insecta
B
12 M. incognita, M. hapla
10 B. xylophilus
10 P. redivivus
10 S. carpocapsae
9 P. pacificus
9 C. elegans
8 A. suum
8 B. malayi
2 T. spiralis
DLF
DIX
DLF
coiled
signal peptide
DIX
dsh-2
DIX
DSV
mig-5
DLF
C
C. briggsae
C. remanei
C. brenneri
C. elegans
C. japonica
C. angaria
DIX
dsh-2
DIX

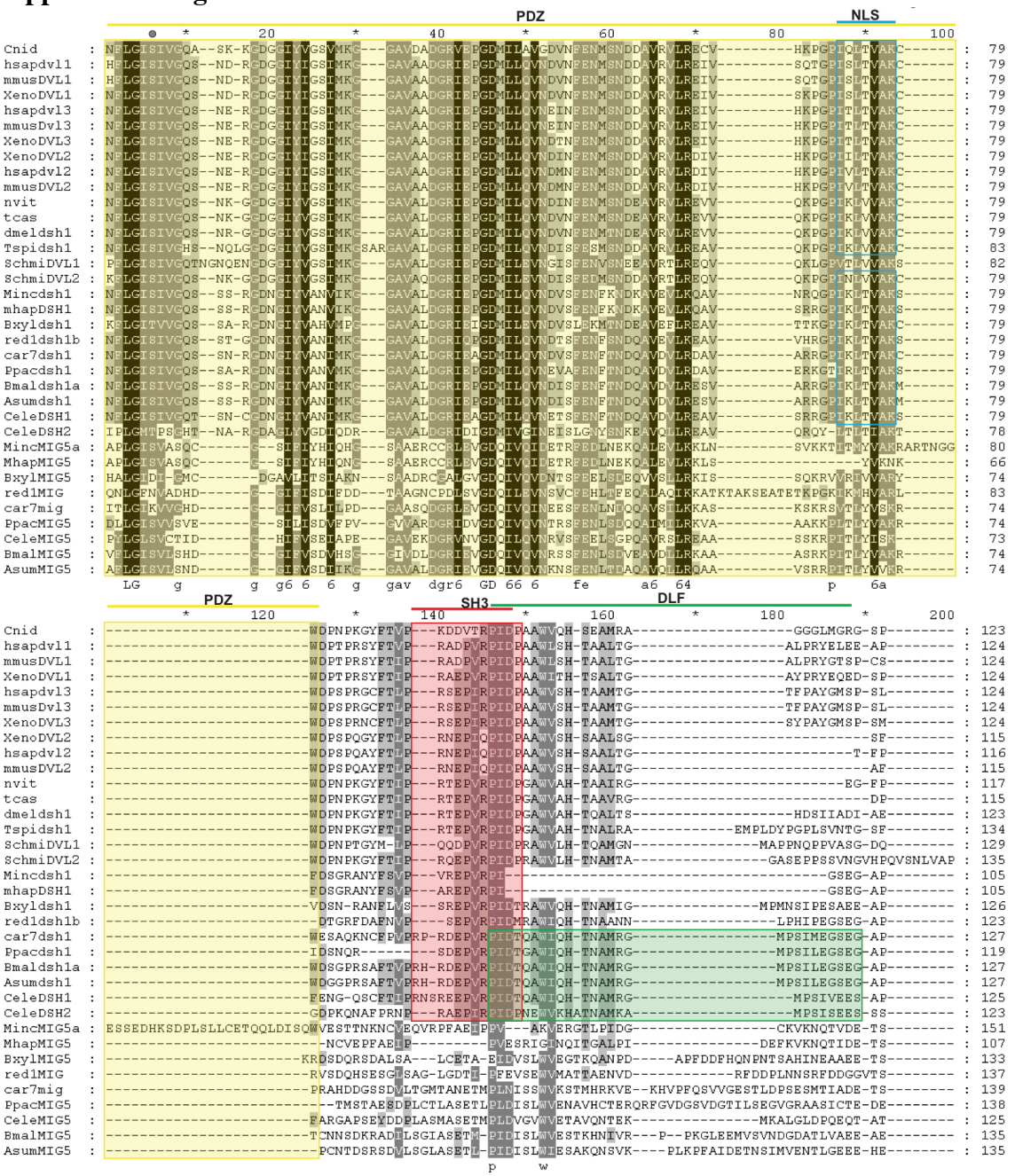
Fig. 8. Dishevelled cladograms
Graphical summary of events during the evolution of Dsh mapped onto cladograms. (A) A cladogram of animal evolution with important features of Dsh evolution mapped onto it. (B) Cladogram of nematodes with identified features of Dsh evolution mapped onto it. (C) Cladogram of caenorhabditids with identified features of Dsh evolution mapped onto it.

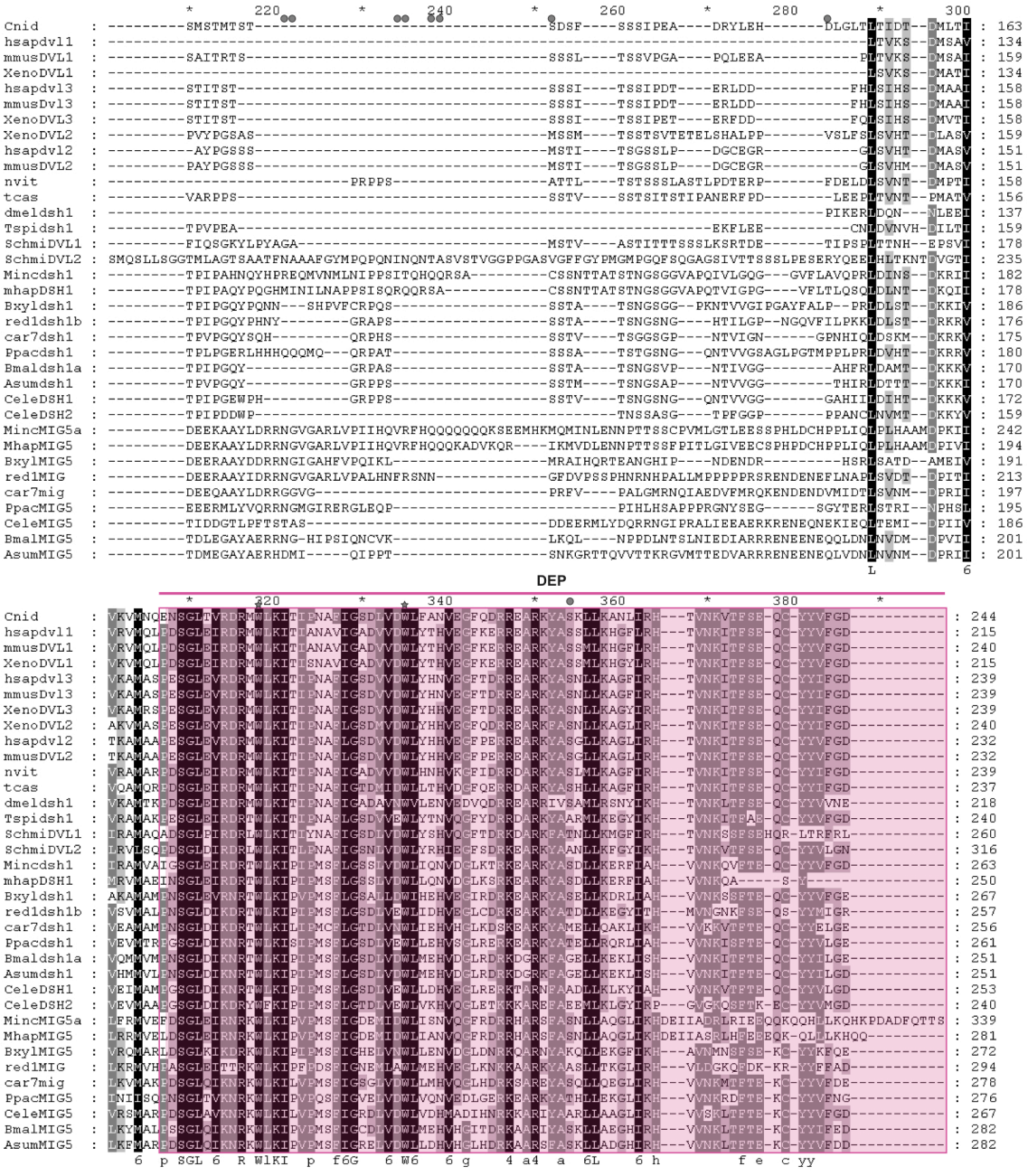
Supplemental Figure 1



Supplemental Figure 1. Dishevelled across caenorhabditids
Phylogenetic analysis of Dsh orthologs across caenorhabditids based on the protein coding nucleotide alignment from the N-terminus of the PDZ domain through the C-terminus of the DEP domain. The ML tree (rooted with the outgroup taxon *N. vitripennis*) is shown. For each node, ML bootstrap support values (1,000 replicates) are above the nodes while parsimony bootstrap values (1,000 replicates) are written below. Support values <70 are not shown.

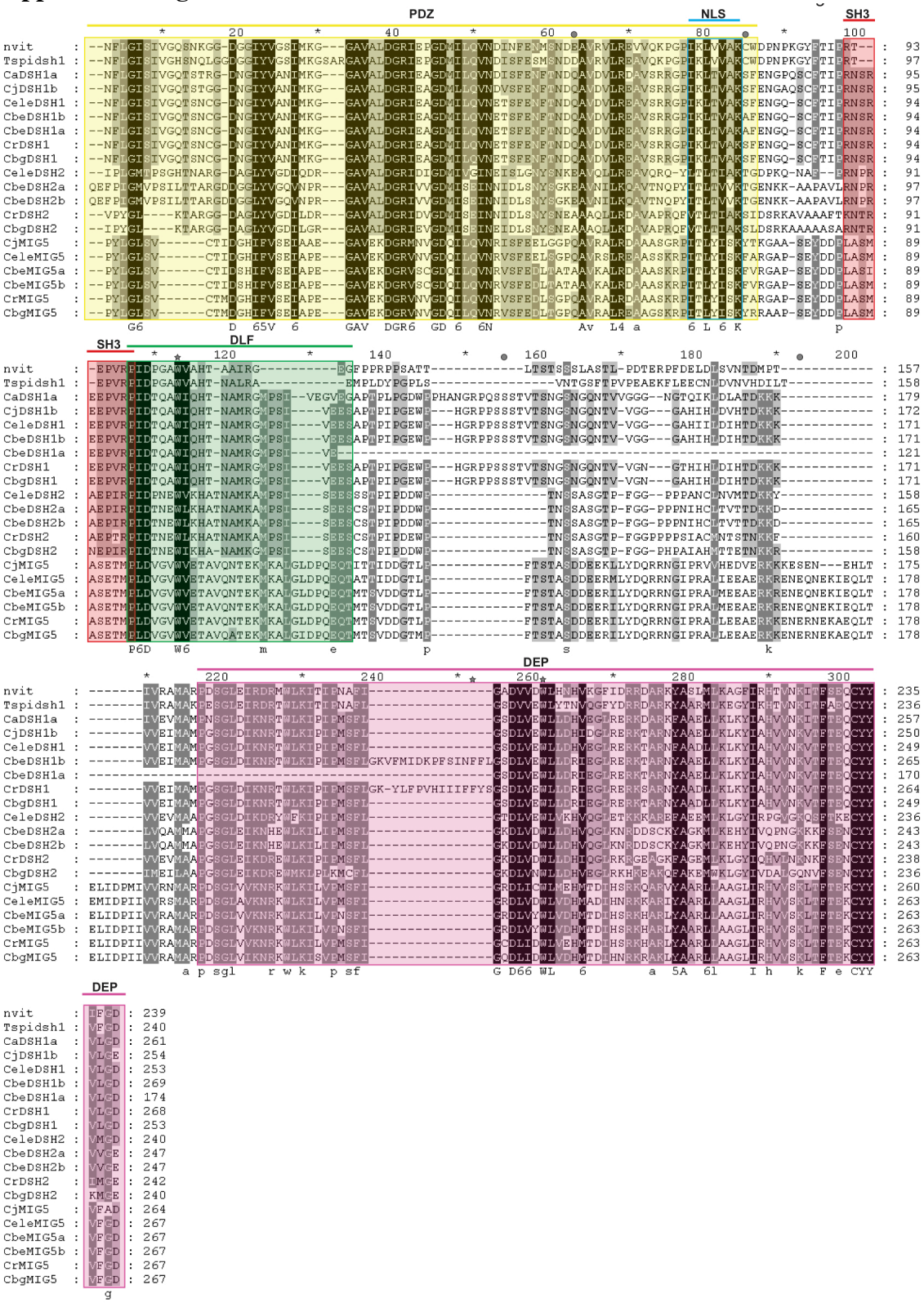
Supplemental Figure 2





Supplemental Figure 2. Dishevelled Protein Alignment
Protein alignment from the beginning of the PDZ domain through the end of the DEP domain across animals. A nucleotide version of this alignment was used to generate the phylogenetic tree from Figure 3. Domain features are highlighted in color and labeled, including PDZ, NLS, SH3, DLF, and DEP. The conserved tyrosine 473 (Y473) is labeled with a polygon. Codons identified to be under negative selection are labeled with a triangle while codons identified as experiencing diversifying selection are labeled with a circle.

Supplemental Figure 3



Supplemental Figure 3. Additional dishevelled protein alignment
Protein alignment from the beginning of the PDZ domain through the end of the DEP

domain from all caenorhabditids in this analysis, plus *N. vitripennis* and *T. spiralis* as outgoups. A nucleotide version of this alignment was used to generate the phylogenetic tree from Figure S1. Domain features are highlighted in color and labeled, including PDZ, NLS, SH3, DLF, and DEP. The conserved tyrosine 473 (Y473) is labeled with a polygon. Codons identified to be under negative selection are labeled with a triangle while codons identified as experiencing diversifying selection are labeled with a circle.

Chapter 4

FGF signaling regulates Wnt ligand expression to control vulval cell lineage polarity in *C. elegans*

ABSTRACT

The interpretation of extracellular cues leading to the polarization of intracellular components and asymmetric cell divisions is a fundamental part of metazoan organogenesis. The *C. elegans* vulva, with its invariant cell lineage and interaction of multiple cell signaling pathways, provides an excellent model for the study of cell polarity within an organized epithelial tissue. Here we show that the Fibroblast Growth Factor (FGF) pathway acts in concert with LIN-17/Frizzled to influence the localization of SYS-1/β-catenin, a component of the Wnt/ β-catenin asymmetry pathway, indirectly through the regulation of Wnt, *cwn-1*. The source of the FGF ligand is the 1º vulval precursor cell (VPC), P6.p, which controls the orientation of the neighboring 2º VPC, P7.p, by signaling through the sex myoblasts (SMs), activating the FGF pathway. The Wnt, CWN-1, is expressed in the posterior body wall muscle of the worm as well as the SMs, making it the only Wnt expressed on the posterior and anterior sides of P7.p at the time of the polarity decision. Both sources of *cwn-1* act instructively to influence P7.p polarity in the direction of the highest Wnt signal. Using single molecule FISH (smFISH), we show the FGF pathway regulates the expression of *cwn-1* in the SMs. These results demonstrate an interaction between FGF and Wnt in *C. elegans* development and vulval cell lineage polarity, as well as highlight the promiscuous nature of Wnts and the importance Wnt gradient directionality within *C. elegans*.

INTRODUCTION

The orientation of asymmetric cell divisions is essential to proper tissue architecture and organogenesis (Strutt, 2005). Loss of cell polarity and asymmetry is a major factor in tumor formation, and growing evidence illustrates its importance in understanding human cancer (Wodarz and Nathke, 2007). Because polarity and asymmetry are such vital components of proper organ formation, cell-cell interactions involving crosstalk between multiple signaling pathways are often incorporated to tightly regulate these processes. The *Caenorhabditis elegans* vulva provides a simple model to study this phenomenon due to the small number of cells, invariant cell lineage and developmental timing, and cell signaling mechanisms involved within vulval formation (reviewed by Sternberg, 2005; reviewed by Gupta et al., 2012). Here we examine the interaction of FGF and Wnt signaling in controlling vulval cell lineage orientation.

The *C. elegans* vulva is formed from divisions of three VPCs: P5.p, P6.p, and P7.p – arranged along the anterior-posterior axis in the ventral epithelium (Sulston and Horvitz, 1977). During the L3 (third larval) stage, a combination of EGF, Notch, and Wnt signals instructs the VPCs to adopt fates corresponding to particular lineage patterns. P6.p adopts a 1º fate and undergoes three rounds of symmetric divisions that lead to eight cells that form the vulval lumen. P5.p and P7.p adopt the 2º fate, which leads to three rounds of asymmetric cell divisions forming seven cells that create the anterior and posterior sides of the vulva (Figure 1). The outermost progeny of P5.p and P7.p adhere to the epidermis whereas the innermost progeny join the descendants of P6.p in forming the vulval lumen. The descendants of P5.p and P7.p display mirror symmetry about the center of the vulva.

Previous analyses show the orientation of P5.p and P7.p descendants is determined by the interaction of multiple Wnt signals. In the absence of all Wnts, the VPCs display a randomized orientation, which is likely the default (Green et al., 2008; Figure 1). Two separate Wnts from the anchor cell, LIN-44 and MOM-2 acting through receptors LIN-17/Frizzled and LIN-18/Ryk, respectively, regulate P7.p orientation (Ferguson et al., 1987; Sternberg and Horvitz, 1988; Sawa et al., 1996; Inoue et al., 2004; Gleason et al., 2006). In the absence of these signals the orientation of the progeny of P7.p mimic those of P5.p and face toward the posterior of the worm, a phenotype referred to as posterior-reversed vulval lineage (P-Rvl; Figure 2). This posterior orientation is dependent on the instructive signal of EGL-20, a Wnt expressed in the tail acting through CAM-1/ROR and VANG-1/Van Gogh, and is referred to as "ground polarity." In response to the Wnt signals from the anchor cell, LIN-17 and LIN-18 orient P7.p to face the center. This reorientation is described as "refined polarity" and is the wild-type orientation (Green et al., 2008; Figure 1).

The adult vulva is essential for egg-laying and mating. The sex muscles, consisting of uterine and vulval muscles, are required for egg-laying. The vulval muscles are formed from the migrating SMs (Thomas et al., 1990). Both gonad-independent and dependent pathways control the anterior migration of SMs in the *C. elegans* hermaphrodite (Burdine et al., 1998; Branda and Stern, 2000). EGL-17/FGF is the gonad-dependent attractant and acts via FGF receptor EGL-15. The dorsal uterus, ventral uterus, anchor cell and P6.p produce the gonad-dependent attractant (Branda and Stern, 2000). The function of EGL-17 in SM migration requires other components of the FGF pathway: genetic mutations of each component affect the migration and final location of the SMs (Sundaram et al.,

1996). Because *egl-17* expression in P6.p is not necessary, but is sufficient, for proper SM migration, it is believed that this expression is used to fine-tune the gonadal attraction (Burdine et al., 1998). *egl-17* expression in P6.p is activated by the inductive signal from the anchor cell that occurs in early L3 at which time the SMs have reached the center of the gonad (Figure 3).

Interactions between Wnt and other signaling pathways during vulval orientation have not been explored. Here we present evidence that FGF signaling promotes the wild-type orientation of P7.p. We show that FGF signaling interacts genetically with LIN-17 and indirectly controls the localization of SYS-1/β-catenin, a key component of Wnt/β-catenin asymmetry pathway. The 1º cell, P6.p, is the source of the EGL-17 signal that controls polarity and acts through EGL-15 and the remainder of the FGF pathway in the migrating SMs. The effect of FGF signaling on vulval orientation is two-sided. First, the SMs must reach their final position, around the gonad center, then EGL-17 must activate the remainder of the FGF pathway in the SMs. Using smFISH we discovered that the FGF pathway is necessary for the regulation of a Wnt, *cwn-1*, in the left and right SMs as they flank the center of the gonad during the polarity decision of P7.p. *cwn-1* is also expressed strongly from the posterior body wall muscle, making it the only Wnt with sources of expression on both the anterior and posterior sides of P7.p. We demonstrate that these two sources act instructively and add to the overall Wnt gradient in both the anterior and posterior directing pathways.

MATERIALS AND METHODS

Strains and Genetics

C. elegans was handled as described previously (Brenner, 1974). All strains used (listed in Table S1) are derivatives of *C. elegans* N2 Bristol strain. The alleles used are as follows. LGI: *lin-17(n671)*, *sem-2(n1343)*. LGII: *cwn-1(ok546), ayIs4[egl-17::gfp, dpy-20(+)]*. LGIII: *qIs95[pSYS-1::VNS::SYS-1* with *pttx-3::dsRed]*. LGIV: *lin-45(sy96), dpy-20(e1282)*. LGX: *lin-18(e620)*, *egl-17(e1313)*, *egl-17(n1377)*, *egl-15(n484)*, *sem-5(n1779, cs15, n2109, n2195)*, *ksr-1(ku68)*. The strain *ayIs4[egl-17::GFP, dpy-20(+)]; dpy-20(e1282); lin-18(e620)* was constructed by crossing strains NH2466 with CB620 (Ferguson and Horvitz, 1985; Burdine et al., 1998). For RNAi experiments, gravid hermaphrodites were fed RNAi-expressing bacteria and their L4 progeny were scored.

Scoring Vulval Phenotypes

To determine the vulval phenotype as wild type or P-Rvl, animals were scored in the mid-L4 stage. Animals were classified as P-Rvl if the 1° and 2° VPCs were induced, but separated by adherent cells (Katz et al., 1995). Only fully induced vulvae were scored.

Transgenics

To make the *CWN-1::GFP* construct backbone, *cwn-1* was amplified from genomic DNA (forward primer, ATGTGATGTCGACAAAAATGCTGAAATCTACACAAGTGATCC; reverse primer, GCAGCTTCTAGATAAGCATAAATACTTCTCAATTCG) and inserted into Fire vector pPD95.75 using restriction sites SalI and XbaI. To create *Pegl-17::CWN-1::GFP*,

first the promoter region of *egl-17* was amplified from genomic DNA (forward primer, GCCTATGCAGCATTGGAGGATG; reverse primer, GGATCACTTGTGTAGATTTCAGCATAGCTCACATTTCGGGCACCTG). The promoter region of *egl-17* was then fused to *CWN-1::GFP* (forward primer, GCCTATGCAGCATTGGAGGATG; reverse primer, AAGGGCCCGTACGGCCGACTA) (Hobert, 2002). The *Pegl-17::CWN-1::GFP* extrachromosomal array was generated by creating an injection mix consisting of 1 ng/µl *Pegl-17::CWN-1::GFP*, 7 ng/µl *Pmyo-2::dsRed*, and 142 ng/µl DNA ladder and injecting the mix into *cwn-1(ok546); lin-18(e620)* as well as *lin-18(e620) egl-15(n484)* animals as described (Mello et al., 1991).

Ablations

Cell ablation experiments were performed as described (Bargmann and Avery, 1995). P6.p was ablated post induction, but before the first division of the VPCs. Strain NH2466 was crossed into *lin-18(e620)* in order to accurately time the experiments by monitoring *egl-17* expression in P6.p. The M cell was ablated in the early L1 stage in both a *lin-18(e620)* as well as *cwn-1(ok546); lin-18(e620)* background. After ablations, the animals were recovered from slides and grown at 20°C until the mid-L4 stage when the vulval phenotype could be scored. Mock ablations were performed by placing appropriately staged worms on a slide for approximately 10 minutes, recovering them, and then scoring their vulval phenotype in mid-L4.

Single Molecule mRNA FISH

Probes for *cwn-1* detection were provided as a gift from Dong Hyun Kim (Harterink et al., 2011). Preparation and hybridization steps were performed as previously described (Raj et al., 2008). Both strains, *N2* and *egl-15(n484)*, were prepared and imaged in an identical manner. Multiple plates were grown until full of gravid hermaphrodites and then bleached. The eggs from these bleachings were placed on fresh plates and grown at 20°C to enable an approximate synchronization of animals. After the animals had reached vulval induction they were washed from the plates using ddH_20 and fixed in 3.7% formaldehyde in 1x PBS for 1 hour. Fixed animals were then permeabilized in 70% ethanol for 48 hours. Animals were washed and the *cwn-1* probes coupled with Cy5 were added and left overnight at 37°C. The next day animals were washed and DAPI stained. Images were taken in z-stacks using an Olympus IX2-UCB microscope, Andor iKon-M 934 camera, and appropriate optical filters for Cy5 and DAPI. Z-stacks were flattened into single images using Fiji. Quantification of single mRNA transcripts within the SMs was performed using a MATLAB script and manually corrected for further accuracy.

RESULTS

FGF Signaling Defects Enhance the *lin-18*/Ryk Phenotype

The *C. elegans* Grb2 ortholog, *sem-5,* acts in both vulval induction, controlled by the Epidermal Growth Factor (EGF) pathway, and SM migration, controlled by the FGF pathway (Clark et al., 1992; reviewed by Sundaram, 2006). SEM-5 is an adaptor protein whose SH2 domain likely binds to the phospho-tyrosine residues of LET-23/EGFR and EGL-15/FGFR and recruits the RAS exchange factor SOS-1/Son of sevenless via its SH3 domains. Expression of the FGF ligand in P6.p is dependent upon vulval induction (Burdine et al., 1998).

Different alleles of *sem-5* have varying degrees of effect on vulval induction as well as SM migration, but a role in vulval orientation has not previously been reported. We scored the vulval lineage of P7.p in four different alleles of *sem-5*. Two alleles, *n2019* and *cs15,* which cause a Glycine to Alanine substitution in the first SH3 domain and an opal stop in the second SH3 domain, respectively, cause polarity and induction defects, whereas *n2195,* which causes a Glycine to Arginine substitution in the second SH3 domain, yields neither polarity nor induction defects. The fourth allele, *n1779*, causes a Glutamate to Lysine substitution in the SH2 domain, results in a 13% P-Rvl phenotype, affecting polarity, but not induction. We thus used *sem-5(n1779)* as the canonical allele. Previously known components involved in the regulation of vulval cell lineage polarity include LIN-17, LIN-18, CAM-1, and VANG-1, all Wnt signaling components (Inoue et al., 2004; Gleason et al., 2006; Green et al., 2008). SEM-5 is the first non-Wnt signaling component found to be involved in vulval orientation.

We next looked at the involvement of each component in the FGF pathway. No allele of *egl-17*, *egl-15*, or any other downstream FGF component other than *sem-5* had any effect on orienation as single mutants (Table 1), which is likely due to the involvement of *sem-5* in one of the other pathways controlling vulval orientation as well as its role in the FGF pathway. No null mutations of the downstream components of the FGF pathway are available due to their lack of viability. There are conflicting reports on whether *egl-17(n1377)* is a null or reduced-function allele, though due to the severity of its phenotype as well as the frequency with which *egl-17* mutations arise in EMS screens, it is most often considered null (Burdine et al., 1997; Chateau et al., 2010).

To understand the genetic relationship between FGF signaling and the previously known Wnt polarity pathway components required for the wild-type vulval orientation, we constructed double mutants of *egl-15(n484)* with *lin-17(n671)* and *lin-18(e620)*, the canonical null alleles (Table 1). Because *egl-15(n484)* enhances the *lin-18(e620)* P-Rvl phenotype from 31 to 63% and has no effect on *lin-17(n671)* we believe the FGF pathway is working with the LIN-17 pathway to control vulval orientation. To test this hypothesis we constructed double mutants of all known FGF pathway components with *lin-18(e620)* or used RNAi in a *lin-18(e620)* background (Table 1). Alleles of *egl-17* enhanced *lin-18(e620)* to approximately 55% P-Rvl, similar to the effect of *sem-5(n1779),* which enhanced *lin-18(e620)* to 57% P-Rvl. The double-mutant with the Son of sevenless ortholog, *sos-1,* had a P-Rvl of 63% whereas the double mutant with the Ras ortholog, *let-60*, enhanced the *lin-18(e620)* phenotype to 68% P-Rvl. Finally, the MAP kinase cascade consisting of *lin-45*, *mek-2*, *mpk-1*, and the scaffold, *ksr-1*, also enhanced the vulval phenotype to 60, 67, 68, and 66% P-Rvl, respectively. Each component of the

pathway enhanced the P-Rvl phenotype of *lin-18(e620)* to roughly the same degree implying the entire FGF pathway functions together. This pathway likely acts with LIN-17 as the mutations enhance *lin-18(lf)* but not *lin-17(lf)* alleles. If FGF signaling was working separately from the LIN-17 pathway we would expect FGF to enhance the *lin-17(lf)* phenotype as it does *lin-18(lf)*; however, because there is no effect on *lin-17(lf)* we assume FGF acts in concert with, not separately from, LIN-17.

FGF Regulates the Localization of SYS-1/ β -catenin

The polarity of the P7.p cell divisions is controlled by the Wnt/β-catenin asymmetry pathway (Green et al., 2008), which includes the β-catenin-like proteins SYS-1 and WRM-1, POP-1/TCF, and the Nemo-like-kinase, LIT-1 (reviewed by Mizumoto and Sawa, 2007). The Wnt/β-catenin asymmetry pathway ensures different ratios of SYS-1 to POP-1, controlling the differential transcription of Wnt target genes between daughters of an asymmetric cell division. Because our genetic data show an interaction between FGF and LIN-17 we wanted to determine if the FGF pathway, like LIN-17, can control the asymmetric localization of proteins between daughter cells of P7.p. The initial establishment of vulval polarity can be observed through the localization of VENUS::SYS-1 (VNS::SYS-1), localized in a high (P7.pa)/low (P7.pp) pattern in the wild-type worm, reciprocal to the localization of POP-1/TCF (Phillips et al., 2007; Green et al., 2008).

As previously reported, VNS::SYS-1 asymmetry in P7.p daughter cells is often lost in *lin-17(n671)* and *lin-18(e620)* mutants (Figure 4). These mutants display two aberrant patterns of VNS::SYS-1 localization as well as the wild-type pattern, though less

frequently. The two deviant localization patterns include one in which both P7.pa and P7.pp express equal amounts of VNS::SYS-1 and a reversed VNS::SYS-1 pattern in which P7.pp is enriched with VNS::SYS-1. By observing VNS::SYS-1 localization in *sem-5(n1779)* mutants we found 2 of 20 worms having an atypical localization of VNS::SYS-1, which reflects the small percentage of worms that have P-Rvl phenotype (13% P-Rvl). Since in wild-type worms VNS::SYS-1 invariably localized to the anterior daughter of P7.p, this result is physiologically relevant. In agreement with our model, no other VPCs show defective VNS::SYS-1 localization in a *sem-5(n1779)* background. The reversal of VNS::SYS-1 localization in *lin-18(e620) sem-5(n1779)* double mutants is slightly enhanced to a degree greater than *lin-18(e620)* alone (Figure 4). This observation confirms FGF pathway controls vulval cell polarity by interacting with LIN-17 and thus the Wnt/β-catenin asymmetry pathway and indicates the FGF effect is at the level of P7.p rather than its progeny.

P6.p is the Source of EGL-17/FGF and Controls P7.p Polarity

Once it was confirmed that FGF regulates P7.p polarity we wanted to find the source of FGF. Since the FGF ligand, EGL-17, is expressed in the 1° VPC, P6.p, after EGF has activated vulval induction (Burdine et al., 1998; Figure 3), we hypothesized that P6.p could be the source of the polarity cue. To date, only the anchor cell and the tail of the worm have been shown to be sources of polarity cues; there has been no evidence of the 1° cell regulating the polarity of its 2° neighbors despite their crosstalk during vulval induction (Sternberg and Horvitz, 1989; Levitan and Greenwald, 1998). We ablated P6.p after it received its induction cue, but prior to any polarity choice of P7.p. We used a

Pegl-17::gfp construct to time induction, and ablated the primary cell in both a wild-type background as well as a *lin-18(e620)* background to sensitize the animals to defects in FGF signaling. Worms were monitored until the *Pegl-17::gfp* construct expressed in P6.p and then P6.p was ablated using a laser microbeam (Figure 5). Similarly to the single mutants of the FGF pathway, ablating P6.p in a wild-type background does not lead to any instances of the P-Rvl phenotype. However, the ablation of P6.p in a *lin-18(e620)* background showed a strong enhancement of the *lin-18(e620)* P-Rvl phenotype, similar to that of every FGF pathway component mutant: the mock ablated animals had a 30% P-Rvl phenotype whereas the ablated animals had a 68% P-Rvl phenotype. These data suggest that P6.p produces the EGL-17 ligand cue that directs the polarity of P7.p, and the 1° vulval cell influences polarity of the neighboring 2° vulval cells.

Ablation of the Sex Myoblasts Enhances the *lin-18*/Ryk Phenotype

After verifying the location of the EGL-17 source we wanted to confirm the location of the receptor and remainder of the FGF signaling cascade that influences cell orientation. EGL-15 is expressed in the SMs and is necessary for proper SM migration (DeVore et al., 1995; Sundaram et al., 1996; Branda and Stern, 2000; Lo et al., 2008). To determine whether the polarity cue is acting through the SMs or possibly through the VPCs, we examined the expression pattern of *egl-15* using a gfp translational fusion and found no expression in P7.p or any other VPC; however, expression was seen in the M cell lineage, consistent with previous observations (Lo et al., 2008).

The SMs are born from the M cell approximately 13 hours post-hatching (Sulston and Horvitz, 1977), begin migrating approximately 2 hours after they form, and reach their

final position, flanking the gonad center, 4 hours after beginning migration (Branda and Stern, 2000; Figure 6). If the SMs are the source of the FGF polarity pathway we should see an enhancement of the *lin-18(e620)* P-Rvl phenotype; however, if the source is in another location, such as the vulval precursor cells, we would expect to see no enhancement. We ablated the M cell, the precursor to both the left and right SMs, in 29 worms, approximately 10 hours post-hatching, in a *lin-18(e620)* background. Ablation of the M cell resulted in a strong enhancement of the *lin-18(e620)* phenotype in the same manner as all FGF mutants as well as in the ablation of P6.p: specifically, the M cell ablated worms showed a 66% P-Rvl phenotype as compared to 30% in the non-ablated controls (Figure 6).

Because the M cell descendants also contribute to the posterior body wall muscle and coelomocytes, we sought a cleaner way to eliminate the SMs before the polarity cue. The SoxC ortholog, *sem-2(n1343)*, alters the M cell lineage and prevents the formation of the SMs by driving the cells initially destined to become SMs to become posterior body wall muscle (Tian et al., 2011; Figure 6). Constructing *sem-2(n1343); lin-18(e620)* double mutants results in a 68% P-Rvl phenotype, confirming that the SMs influence the polarity choice of P7.p.

We wanted to observe the effect on vulval orientation in a mutant that inhibits SM migration independently of FGF signaling and does not eliminate FGF signaling within the SMs. *mig-2* encodes a member of the Rho family of GTP-binding proteins, is expressed in the SMs, and prevents the SMs from wild-type migration in approximately half of the animals (Forrester and Garriga, 1997; Zipkin et al., 1997; Kishore and Sundaram, 2002). Because half the SMs do not migrate to their final wild-type position,

we hypothesized these SMs would not be capable of giving the polarity cue to P7.p since they do not migrate to the anterior of the cell. *mig-2* RNAi treated *lin-18(e620)* animals have a 56% P-Rvl phenotype, a significant increase from the *lin-18(e620)* single mutant, confirming the SMs must migrate to their wild-type position to transmit the polarity cue to P7.p (Figure 6).

These results, along with the expression pattern of *egl-15*, indicate the FGF polarity signal comes from P6.p, and requires the SMs. Because the polarity decision of the vulval precursor cells is made prior to anaphase of the first cell division, we believe the FGF polarity cue acts once the SMs have reached their final position flanking the center of the gonad. Mutations of each component of the FGF pathway have varying degrees of penetrance on the migration of the SMs (Sundaram et al., 1996). By contrast the effects of these mutants on vulval lineage orientation are strikingly similar. We believe the effect of FGF signaling on P7.p orientation is two-sided. First, the SMs must migrate to the anterior side of P7.p via an uncompromised FGF signal. Once the SMs have migrated to the anterior side of P7.p, the FGF signal from P6.p activates the downstream components of the pathway, activating the transcription of a gene or set of genes necessary for proper VPC orientation. If either of these two events is compromised, the FGF pathway cannot direct the anterior orientation of P7.p.

FGF Signaling Regulates Expression of *cwn-1* in the Sex Myoblasts

The *C. elegans* genome encodes five different Wnt proteins, expressed in partially overlapping patterns across the anteroposterior axis, but only one, *cwn-1*, is expressed in the SMs (reviewed by Eisenmann, 2005; Harterink et al., 2011). Work in other animals

has shown crosstalk between FGF and Wnt pathways, often leading to the regulation of Wnt by FGF (Hong et al., 2008; Stulberg et al., 2012; Yardley and Garcia-Castro, 2012). We hypothesized that FGF signaling regulates a Wnt signal produced in the SMs that controls P7.p polarity. To test this idea directly we used smFISH to quantify the number of mRNA transcripts of *cwn-1* found within the left and right SMs just prior to the polarity decision of P7.p in a wild-type and reduced FGF signaling backgrounds (Figure 7).

On average, the wild-type SMs each express 50 transcripts of *cwn-1* prior to the polarity choice of P7.p. In an *egl-15(n484)* background, the expression of *cwn-1* transcripts is reduced two-fold on average with 23% of the SMs having 1/3 the number of wild-type transcript and 10% having as little as 1/5 of the number of wild-type transcripts. There is no overlap in SM transcript count between the wild-type and mutant backgrounds. The lowest wild-type SM transcript count is still greater than the highest SM transcript count in the mutant background: 40 transcripts per SM is the lowest wild-type count compared to 37 transcripts per SM for the highest *egl-15(n484)* count (Figure 7 and Table S2). Therefore, FGF signaling regulates the expression of the Wnt ligand, *cwn-1*. It cannot be determined just how much *cwn-1* transcript is needed to produce a wild-type vulval orientation, although previous work has examined how a change in transcript count affects phenotype (Raj et al., 2010). Examining the transcript count of *egl-15(n484),* we hypothesize that the SMs with a higher *cwn-1* transcript count, similar to that of the wild-type, produce a P7.p lineage with an anterior orientation. It is the SMs with a greatly reduced *cwn-1* transcript count that likely fall below the necessary threshold to orient P7.p to the anterior and, therefore, produce a P-Rvl phenotype.

cwn-1 Acts Instructively from Both the Anterior and Posterior Sides of P7.p

cwn-1 is expressed in the posterior body wall muscle and M cell descendants, making it the only Wnt ligand expressed from the anterior and posterior sides of P7.p during the polarity decision (Harterink et al., 2011; also see Figure 7). Previous work suggested that Wnt ligands instruct P7.p to orient toward the direction of the Wnt gradient: LIN-44 and MOM-2 toward the anterior and EGL-20 toward the posterior (Figure 1). Genetic evidence indicates *cwn-1* acts upstream of *lin-17,* a receptor necessary for the anterior signal (Gleason et al., 2006), and has been shown to bind to CAM-1, a receptor necessary for the posterior signal (Green et al., 2007). Because *cwn-1* is expressed on both sides of P7.p and has been shown to interact with receptors associated with the anterior and posterior pathways, we hypothesized that each gradient might instruct P7.p to orient toward the direction of the respective gradient. A *cwn-1* mutation would, therefore, abolish the anterior and posterior signals making the overall effect of CWN-1 on P7.p minimal. We analyzed a *cwn-1(ok546)*; *lin-18(e620)* double mutant and found it had little effect on *lin-18(e620)* vulval orientation, 31 versus 26% P-Rvl.

All Wnts directing VPC polarity instruct the localization of SYS-1 to the P7.p daughter cell toward the gradient (Green et al., 2008). Despite being different Wnts, LIN-44 and MOM-2, acting through LIN-17 and LIN-18, respectively, both have the same molecular output of anterior SYS-1 localization. EGL-20, from the posterior, drives the posterior localization of SYS-1. We assume that each Wnt imparts a directional cue instructing SYS-1 to localize to the direction of the Wnt source. Therefore, CWN-1 from the SMs joins LIN-44 and MOM-2 in driving anterior localization, through an overall anterior Wnt gradient, and CWN-1 from the posterior body wall muscle joins EGL-20 in

driving posterior localization, through an overall posterior Wnt gradient (Figure 8). This assumption makes physical sense when considering mutations in FGF pathway components. The single mutants do not affect orientation because only one anterior Wnt is removed, leaving LIN-44 and MOM-2 to direct the localization of SYS-1. However in a *lin-18(e620)* double mutant, the animal has lost two anterior sources of Wnt, CWN-1 and MOM-2, being the ligand of LIN-18, and therefore the overall anterior Wnt gradient is greatly reduced allowing the posterior gradient to dominate. Likewise, if the posterior CWN-1 signal is compromised the overall posterior Wnt gradient is reduced and SYS-1 is instructed to localize to the anterior daughter cell.

To test this hypothesis we designed a construct that would provide an anterior gradient of CWN-1, namely, *Pegl-17*::CWN-1::GFP, and therefore reinforce the anterior gradient. The *egl-17* promoter activates the expression of *cwn-1* in P6.p upon vulval induction (Figure S1). By expressing this construct in a *cwn-1(ok546); lin-18(e620)* background, the only source of CWN-1 comes from the anterior side of P7.p. Anterior expressed CWN-1 suppresses the *cwn-1(ok546); lin-18(e620)* phenotype from 26 to 13% (p-value 0.1288). We hypothesized the P-Rvl phenotype of *cwn-1(ok546); lin-18(e620)* could be too mild at 26% to see the full suppression resulting from driving CWN-1 from the anterior, so we used a sensitized background that gives a higher initial P-Rvl phenotype. Treating *cwn-1(ok546); lin-18(e620)* worms with *lin-44* RNAi increases the percentage of P-Rvl to 52% due to the role of LIN-44 acting upstream of LIN-17. Expressing the *Pegl-17*::CWN-1::GFP construct in *cwn-1(ok546); lin-18(e620)* worms treated with *lin-44* RNAi results in significant suppression of the P-Rvl phenotype to 30% (p-value of 0.0210) (Table 2).

We next tested whether anterior CWN-1 could rescue the phenotype of *lin-18(e620) egl-15(n484)*, and found that it does rescue the phenotype from 63 to 38% (Table 2). We believe the construct does not rescue fully back to 30% because in a *lin-18(e620) egl-15(n484)* animal the SMs are still producing a reduced CWN-1 signal from the posterior side of P7.p.

These data illustrate that CWN-1 provides an instructive anterior gradient sufficient to suppress the posterior gradient in the wild-type nematode (Table 2). If this cue were permissive we would not expect to see a sole anterior source of CWN-1 suppress either *cwn-1(ok546); lin-18(e620), cwn-1(ok546); lin-18(e620)* grown in *lin-44* RNAi, or rescue the phenotype of *lin-18(e620) egl-15(n484)*. CWN-1, therefore, acts instructively from the anterior and posterior of P7.p. In the absence of a posterior signal, the anterior signal reinforces the progeny of P7.p to face the center and can suppress the P-Rvl phenotype. Likewise, in the absence of the anterior CWN-1 signal, through defects in the FGF pathway, removal of P6.p, or the SMs, the posterior signal instructs the progeny of P7.p to orient posteriorly when the anterior Wnt gradient has been compromised (Figure 8).

DISCUSSION

Our results describe an interaction between FGF and Wnt signaling in vulval cell lineage polarity. Through genetic analysis we have shown that each component of the FGF pathway enhances the P-Rvl phenotype LIN-18/Ryk mutants, but does not affect those of LIN-17/Fz, indicating a specific interaction between FGF and LIN-17, likely CWN-1 acting on LIN-17 but not LIN-18. The underlying mechanisms of the P-Rvl phenotype can be seen on the molecular level through the localization of the β-catenin ortholog, SYS-1. FGF signaling indirectly controls the localization of SYS-1 to the anterior daughter cell of P7.p, which leads to the wild-type vulval orientation. FGF signaling does not directly influence the vulval lineage orientation, but instead is required for the regulation of CWN-1 expression, which acts instructively from both sides of P7.p (Figure 8). CWN-1 is the only Wnt ligand expressed on the anterior and posterior of P7.p at the time of its polarity decision and acts upstream of receptors involved in directing P7.p to face the anterior and posterior: LIN-17 and CAM-1, respectively.

How does P7.p always orient toward the anterior in the wild-type worm? Genetic data suggest MOM-2 and LIN-44 have a greater ability to direct the anterior orientation of P7.p, with CWN-1 acting as a minor player. Both posterior expressed CWN-1 and EGL-20 act over a distance and form a poster-anterior gradient that has the ability to direct the orientation of P7.p toward the posterior, though the concentration of posterior Wnts might be much lower compared to anterior expressed Wnts by the time they reach the VPCs (Coudreuse et al., 2006). Expressing either CWN-1 or EGL-20 from the anterior of P7.p (from the anchor cell or P6.p) is sufficient to redirect the orientation of P7.p toward the anterior. All four Wnts involved in vulval orientation direct the localization of SYS-1

despite acting through three different receptors, all of which are present in the same cell, P7.p. There is receptor specificity, but all Wnts seem to have the same effect: P7.p orients in the direction of the highest Wnt gradient. P7.p always faces the anterior in a wild-type worm because of the three anterior sources of Wnts in close proximity to P7.p. Only by removing these sources can we begin to see the effects of the posterior Wnts; the same posterior Wnts that can impart an anterior directing cue when repositioned. The two posterior Wnts, EGL-20 and CWN-1, both activate competence to respond to LIN-3/EGF in the anterior VPCs and may have the same molecular activity (Penigault and Felix, 2011). A possible hallmark of Wnt-mediated patterning within *C. elegans* could be similar molecular outputs from genes that are not truly redundant.

How similar is Wnt driven VPC patterning to other systems? A major difference between *C. elegans* and *Drosophila* is no Wnts have been implicated in *Drosophila* planar cell polarity whereas Wnts play a major roll in patterning the VPCs. On the other hand, the receptor CAM-1/Ror and transmembrane protein, VANG-1/Van Gogh, antagonize LIN-17 and LIN-18 by directing the localization of SYS-1 to the posterior daughter of P7.p. The antagonism between Fz and Van Gogh is a hallmark of planar cell polarity in the Drosophila wing (Seifert and Mlodzik, 2007; Gao, 2012; Singh and Mlodzik, 2012), but much less is understood about the interaction between Ror and Van Gogh (Gao et al., 2011).

Other comparisons can be drawn between *C. elegans* and vertebrate Wnt signaling. Wnts LIN-44 and CWN-1 act through LIN-17/Fz and MOM-2 acts through LIN-18/Ryk to direct SYS-1 to localize to the anterior daughter of P7.p. Although the possibility of a Fz-Ryk coreceptor complex exists in the mammalian systems (Lu et al., 2004), LIN-17

and LIN-18 function in parallel pathways despite both directing the localization of SYS-1. Recent work in vertebrates has shown FGF regulates the expression of Wnt similar to our work in *C. elegans* vulval patterning. FGF regulates the expression of Wnt in the non-neural ectoderm of the chick (Yardley and Garcia-Castro, 2012). FGF also elevates Wnt expression, through inhibition of Wnt antagonists, in the zebrafish tailbud (Stulberg et al., 2012). Furthermore, our results illustrate a network of signals, relayed back and forth between different tissues: the gonadal anchor cell expresses an EGF signal that induces the ectodermal vulval cells, activating an FGF signal that is sent to the mesodermal sex myoblasts, which enables the regulation of a Wnt that directs the patterning of the ectodermal vulval cells. This relay between different tissues bears resemblance to *Xenopus* where it has been shown that Fgf8a induces neural crest indirectly through the activation of Wnt8 in the paraxial mesoderm, which then directs neural crest formation in the overlying ectoderm (Hong et al., 2008).

Using the *C. elegans* vulva as a model, we have shown a network of Wnt signals, with distinct receptor specificity directs the orientation of the vulval precursor cells through the localization of β-catenin. One of these Wnts, CWN-1, is regulated through the activity of the FGF pathway in a crosstalk between multiple tissues that enables the efficacy of its directional cue.

ACKNOWLEDGMENTS

We thank Takao Inoue, Jennifer Green, Wendy Katz, Adeline Seah, and Michael Stern for insightful comments and laying the groundwork for this project. We also thank Gladys Medina and Barbara Perry for technical assistance and members of the Sternberg laboratory, especially Mihoko Kato, Amir Sapir, James Lee, and Hillel Schwartz, for helpful discussions and critically reading the manuscript. We thank WormBase and the *Caenorhabditis* Genetics Center. P.J.M. was supported by a National Institutes of Health (NIH) United States Public Health Service Training Grant (T32GM07616), and the Howard Hughes Medical Institute, with which P.W.S. is an investigator.

REFERENCES

Bargmann, C. I. and Avery, L. (1995). Laser killing of cells in *Caenorhabditis elegans*. *Methods in cell biology* **48**, 225-250.
Branda, C. S. and Stern, M. J. (2000). Mechanisms controlling sex myoblast migration in *Caenorhabditis elegans* hermaphrodites. *Developmental biology* **226**, 137-151.
Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
Burdine, R. D., Branda, C. S. and Stern, M. J. (1998). EGL-17(FGF) expression coordinates the attraction of the migrating sex myoblasts with vulval induction in *C. elegans*. *Development* **125**, 1083-1093.
Burdine, R. D., Chen, E. B., Kwok, S. F. and Stern, M. J. (1997). *egl-17* encodes an invertebrate fibroblast growth factor family member required specifically for sex myoblast migration in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2433-2437.
Chateau, M. T., Araiz, C., Descamps, S. and Galas, S. (2010). Klotho interferes with a novel FGF-signalling pathway and insulin/Igf-like signalling to improve longevity and stress resistance in *Caenorhabditis elegans*. *Aging* **2**, 567-581.
Clark, S. G., Stern, M. J. and Horvitz, H. R. (1992). *C. elegans* cell-signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* **356**, 340-344.
Coudreuse, D. Y., Roel, G., Betist, M. C., Destree, O. and Korswagen, H. C. (2006). Wnt gradient formation requires retromer function in Wnt-producing cells. *Science* **312**, 921-924.
DeVore, D. L., Horvitz, H. R. and Stern, M. J. (1995). An FGF receptor signaling pathway is required for the normal cell migrations of the sex myoblasts in *C. elegans* hermaphrodites. *Cell* **83**, 611-620.
Eisenmann, D. M. (2005). Wnt signaling. *WormBook : the online review of C. elegans biology*, 1-17.
Ferguson, E. L. and Horvitz, H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. *Genetics* **110**, 17-72.
Ferguson, E. L., Sternberg, P. W. and Horvitz, H. R. (1987). A genetic pathway for the specification of the vulval cell lineages of *Caenorhabditis elegans*. *Nature* **326**, 259-267.
Forrester, W. C. and Garriga, G. (1997). Genes necessary for C. elegans cell and growth cone migrations. *Development* **124**, 1831-1843.
Gao, B. (2012). Wnt regulation of planar cell polarity (PCP). *Current topics in developmental biology* **101**, 263-295.
Gao, B., Song, H., Bishop, K., Elliot, G., Garrett, L., English, M. A., Andre, P., Robinson, J., Sood, R., Minami, Y. et al. (2011). Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2. *Developmental cell* **20**, 163-176.
Gleason, J. E., Szyleyko, E. A. and Eisenmann, D. M. (2006). Multiple redundant Wnt signaling components function in two processes during *C. elegans* vulval development. *Developmental biology* **298**, 442-457.

Green, J. L., Inoue, T. and Sternberg, P. W. (2007). The *C. elegans* ROR receptor tyrosine kinase, CAM-1, non-autonomously inhibits the Wnt pathway. *Development* **134**, 4053-4062.
Green, J. L., Inoue, T. and Sternberg, P. W. (2008). Opposing Wnt pathways orient cell polarity during organogenesis. *Cell* **134**, 646-656.
Gupta, B. P., Hanna-Rose, W. and Sternberg, P. W. (2012). Morphogenesis of the vulva and the vulval-uterine connection. *WormBook : the online review of C. elegans biology*, 1-20.
Harterink, M., Kim, D. H., Middelkoop, T. C., Doan, T. D., van Oudenaarden, A. and Korswagen, H. C. (2011). Neuroblast migration along the anteroposterior axis of *C. elegans* is controlled by opposing gradients of Wnts and a secreted Frizzled-related protein. *Development* **138**, 2915-2924.
Hobert, O. (2002). PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *BioTechniques* **32**, 728-730.
Hong, C. S., Park, B. Y. and Saint-Jeannet, J. P. (2008). Fgf8a induces neural crest indirectly through the activation of Wnt8 in the paraxial mesoderm. *Development* **135**, 3903-3910.
Inoue, T., Oz, H. S., Wiland, D., Gharib, S., Deshpande, R., Hill, R. J., Katz, W. S. and Sternberg, P. W. (2004). *C. elegans* LIN-18 is a Ryk ortholog and functions in parallel to LIN-17/Frizzled in Wnt signaling. *Cell* **118**, 795-806.
Katz, W. S., Hill, R. J., Clandinin, T. R. and Sternberg, P. W. (1995). Different levels of the *C. elegans* growth factor LIN-3 promote distinct vulval precursor fates. *Cell* **82**, 297-307.
Kishore, R. S. and Sundaram, M. V. (2002). ced-10 Rac and mig-2 function redundantly and act with unc-73 trio to control the orientation of vulval cell divisions and migrations in Caenorhabditis elegans. *Developmental biology* **241**, 339-348.
Levitan, D. and Greenwald, I. (1998). LIN-12 protein expression and localization during vulval development in *C. elegans*. *Development* **125**, 3101-3109.
Lo, T. W., Branda, C. S., Huang, P., Sasson, I. E., Goodman, S. J. and Stern, M. J. (2008). Different isoforms of the *C. elegans* FGF receptor are required for attraction and repulsion of the migrating sex myoblasts. *Developmental biology* **318**, 268-275.
Lu, W., Yamamoto, V., Ortega, B. and Baltimore, D. (2004). Mammalian Ryk is a Wnt coreceptor required for stimulation of neurite outgrowth. *Cell* **119**, 97-108.
Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal* **10**, 3959-3970.
Mizumoto, K. and Sawa, H. (2007). Two betas or not two betas: regulation of asymmetric division by beta-catenin. *Trends in cell biology* **17**, 465-473.
Penigault, J. B. and Felix, M. A. (2011). High sensitivity of C. elegans vulval precursor cells to the dose of posterior Wnts. *Developmental biology* **357**, 428-438.
Phillips, B. T., Kidd, A. R., 3rd, King, R., Hardin, J. and Kimble, J. (2007). Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3231-3236.
Raj, A., Rifkin, S. A., Andersen, E. and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918.

Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature methods* **5**, 877-879.
Sawa, H., Lobel, L. and Horvitz, H. R. (1996). The *Caenorhabditis elegans* gene *lin-17*, which is required for certain asymmetric cell divisions, encodes a putative seven-transmembrane protein similar to the Drosophila frizzled protein. *Genes & development* **10**, 2189-2197.
Seifert, J. R. and Mlodzik, M. (2007). Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. *Nature reviews. Genetics* **8**, 126-138.
Singh, J. and Mlodzik, M. (2012). Planar Cell Polarity Signaling: Coordination of cellular orientation across tissues. *Wiley interdisciplinary reviews. Developmental biology* **1**, 479-499.
Sternberg, P. W. (2005). Vulval development. *WormBook : the online review of C. elegans biology*, 1-28.
Sternberg, P. W. and Horvitz, H. R. (1988). *lin-17* mutations of *Caenorhabditis elegans* disrupt certain asymmetric cell divisions. *Developmental biology* **130**, 67-73.
Sternberg, P. W. and Horvitz, H. R. (1989). The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. *Cell* **58**, 679-693.
Strutt, D. (2005). Organ shape: controlling oriented cell division. *Current biology : CB* **15**, R758-759.
Stulberg, M. J., Lin, A., Zhao, H. and Holley, S. A. (2012). Crosstalk between Fgf and Wnt signaling in the zebrafish tailbud. *Developmental biology* **369**, 298-307.
Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Developmental biology* **56**, 110-156.
Sundaram, M., Yochem, J. and Han, M. (1996). A Ras-mediated signal transduction pathway is involved in the control of sex myoblast migration in *Caenorhabditis elegans*. *Development* **122**, 2823-2833.
Sundaram, M. V. (2006). RTK/Ras/MAPK signaling. *WormBook : the online review of C. elegans biology*, 1-19.
Thomas, J. H., Stern, M. J. and Horvitz, H. R. (1990). Cell interactions coordinate the development of the *C. elegans* egg-laying system. *Cell* **62**, 1041-1052.
Tian, C., Shi, H., Colledge, C., Stern, M., Waterston, R. and Liu, J. (2011). The C. elegans SoxC protein SEM-2 opposes differentiation factors to promote a proliferative blast cell fate in the postembryonic mesoderm. *Development* **138**, 1033-1043.
Wodarz, A. and Nathke, I. (2007). Cell polarity in development and cancer. *Nature cell biology* **9**, 1016-1024.
Yardley, N. and Garcia-Castro, M. I. (2012). FGF signaling transforms non-neural ectoderm into neural crest. *Developmental biology* **372**, 166-177.
Zipkin, I. D., Kindt, R. M. and Kenyon, C. J. (1997). Role of a new Rho family member in cell migration and axon guidance in C. elegans. *Cell* **90**, 883-894.

Table 1

Relevant Genotype	% P-Rvl	n
N2	0	100
lin-17(n671)	74	100
lin-18(e620)	31	100
lin-17(n671); lin-18(e620)	100	40
egl-17(n1377)	0	40
egl-17(e1313)	0	36
egl-15(n484)	0	100
sem-5(n1779)	13	45
sem-5(n2019)	33	18
sem-5(cs15)	7	30
sem-5(n2195)	0	30
sos-1(RNAi)	0	30
let-60(RNAi)	0	30
lin-45(sy96)	0	35
mek-2(RNAi)	0	30

Relevant Genotype	% P-Rvl	n	p value
mpk-1(RNAi)	0	30	
ksr-1(ku68)	0	30	
lin-17(n671); egl-15(n484)	71	31	
lin-18(e620) egl-15(n484)	63	52	0.0001
lin-18(e620) egl-15(RNAi)	63	32	0.0031
egl-17(e1313) lin-18(e620)	54	48	0.0111
egl-17(n1377) lin-18(e620)	57	30	0.0168
lin-18(e620) sem-5(n1779)	57	30	0.0168
lin-18(e620) sos-1(RNAi)	63	30	0.0024
lin-18(e620) let-60(RNAi)	68	31	0.0006
lin-18(e620) lin-45 (sy96)	60	30	0.0054
lin-18(e620) mek-2(RNAi)	67	30	0.0006
lin-18(e620) mpk-1(RNAi)	68	34	0.0002
ksr-1(ku68) lin-18(e620)	66	35	0.0005

Table 1. FGF signaling enhances the P-Rvl phenotype of *lin-18(e620)*
Double mutants were constructed between *lin-18(e620)* and each known component of the FGF pathway. Vulval phenotypes were scored during mid-L4. The p-values were calculated in comparison with *lin-18(e620)* using Fisher's exact test.

Table 2

Relevant Genotype	Anterior CWN-1 Source	% P-Rvl	n	p value
cwn-1(ok546); lin-18(e620)	-	26	100	
cwn-1(ok546); lin-18(e620)	+	13	45	0.1288
cwn-1(ok546); lin-18(e620) lin-44 RNAi	-	52	61	
cwn-1(ok546); lin-18(e620) lin-44 RNAi	+	30	50	0.0210
egl-15(n484) lin-18(e620)	-	63	52	
egl-15(n484) lin-18(e620)	+	38	40	0.0202

Table 2. *cwn-1* acts instructively from the anterior and posterior sides of P7.p. Driving CWN-1 from the anterior side of P7.p suppresses the P-Rvl phenotype of cwn-1(ok546); *lin-18(e620)* mildly and significantly suppresses the phenotype of *cwn-1(ok546); lin-18(e620)* grown in *lin-44* RNAi. Anterior expression rescues the phenotype of *lin-18(e620) egl-15(n484)*.

Supplemental Table 1

Relevant Genotype	Strain Name	Strain Origin
lin-17(n671)	MT1306	Ferguson and Horvitz, 1985
lin-18(e620)	CB620	Ferguson and Horvitz, 1985
egl-17(n1377)	MT3188	Akerib and Meyer, 1994
egl-17(n1377) lin-18(e620)	PS6574	This Work
egl-17(e1313)	CB1313	Trent et al., 1983
egl-17(e1313) lin-18(e620)	PS6575	This Work
egl-15(n484)	PS39	Trent et al., 1983
lin-17(n671); egl-15(n484)	PS5099	This Work
lin-18(e620) egl-15(n484)	PS4937	This Work
sem-5(n1779)	MT4185	Stern and Horvitz, 1988
lin-18(e620) sem-5(n1779)	PS6576	This Work
sem-5(n2019)	MT4755	Clark et al., 1992
sem-5(cs15)	UP148	Rocheleau et al., 1992
sem-5(n2195)	MT998	Clark et al., 1992
lin-45(sy96)	PS427	Han et al., 1993
lin-18(e620) lin-45(sy96)	PS6577	This Work
ksr-1(ku68)	MH734	Sundaram and Han, 1995
ksr-1(ku68) lin-18(e620)	PS6578	This Work
ayIs4	NH2466	Burdine et al., 1998
ayIs4; lin-18(e620)	PS6349	This Work
cwn-1(ok546)	RB763	Zinovyeva and Forrester, 2005
cwn-1(ok546); lin-18(e620)	PS4970	Green et al., 2008
*Pegl-17::*CWN-1::GFP	PS6579	This Work
qIs95[pSYS-1::VNS::SYS-1]	JK3930	Phillips et al., 2007
lin-17(n671); qIs95	JK4062	Phillips et al., 2007
qIs95; lin-18(e620)	PS5593	Green et al., 2008
qIs95; sem-5(n1779)	PS6580	This Work
qIs95; lin-18(e620) sem-5(n1779)	PS6581	This Work
lin-17(n671); lin-18(e620)	PS3976	Inoue et al., 2004
sem-2(n1343)	MT3214	Garriga et al., 1993
sem-2(n1343); lin-18(e620)	PS6582	This Work

Supplemental Table 1. Original source of worm strains
A comprehensive table of the original work where each strain used can be found.

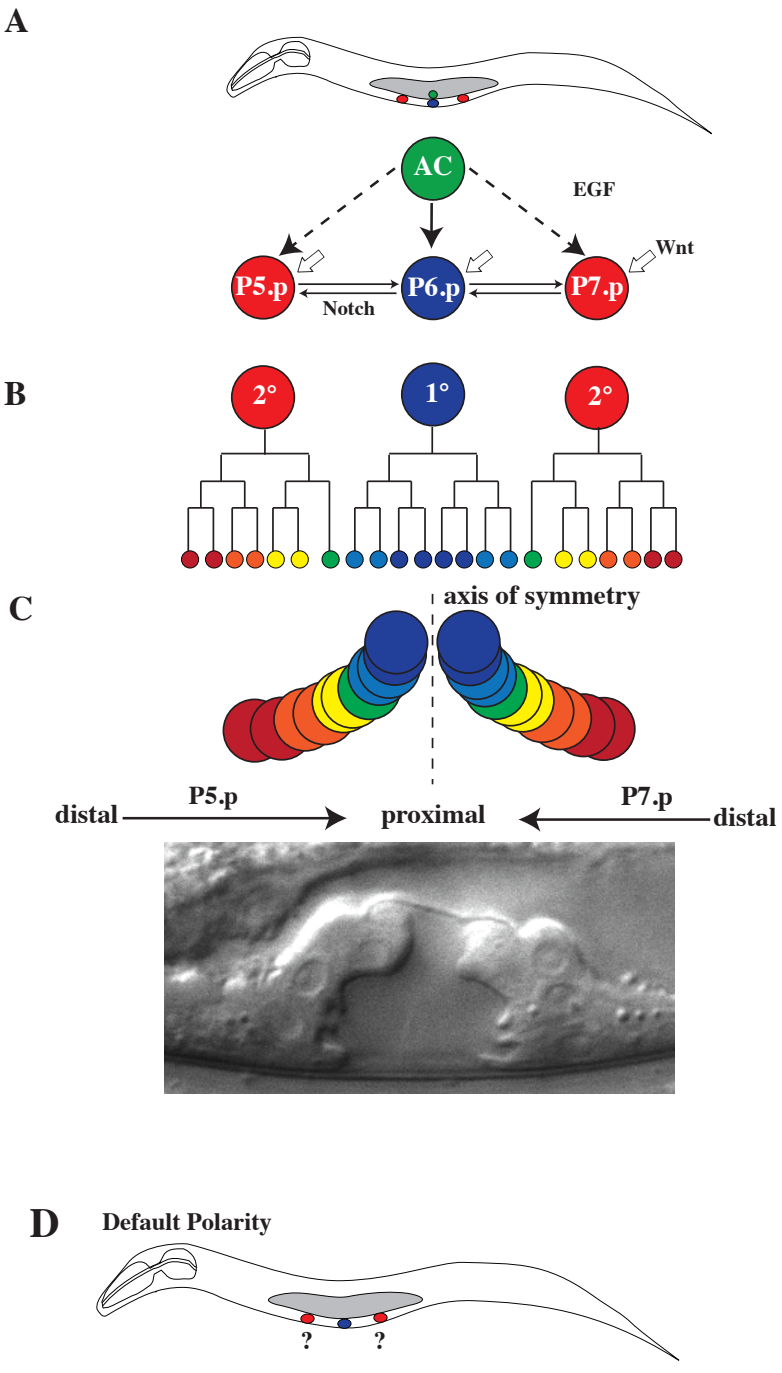
Supplemental Table 2

+	SM1	SM2	Sum
Worm 1	58	62	120
Worm 2	60	45	105
Worm 3	61	40	101
Worm 4	55	53	108
Worm 5	72	60	132
Worm 6	40	48	88
Worm 7	43	40	83
Worm 8	48	45	93
Worm 9	45	52	97
Worm 10	41	47	88
Worm 11	45	40	85

egl-15(n484)	SM1	SM2	Sum
Worm 1	27	25	52
Worm 2	29	13	42
Worm 3	21	26	47
Worm 4	36	13	49
Worm 5	26	7	33
Worm 6	28	15	43
Worm 7	35	30	65
Worm 8	31	9	40
Worm 9	36	37	73
Worm 10	36	23	59
Worm 11	34	34	68

Supplemental Table 2. Raw data from smFISH analysis.
Transcript counts from both SMs in a wild-type and *egl-15(n484)* background. 11 worms were analyzed, representing 22 SMs.

Figure 1


A
AC
EGF
Wnt
P5.p
P6.p
P7.p
Notch
B
2°
1°
2°
C
axis of symmetry
distal
P5.p
proximal
P7.p
distal
D
Default Polarity
?
?

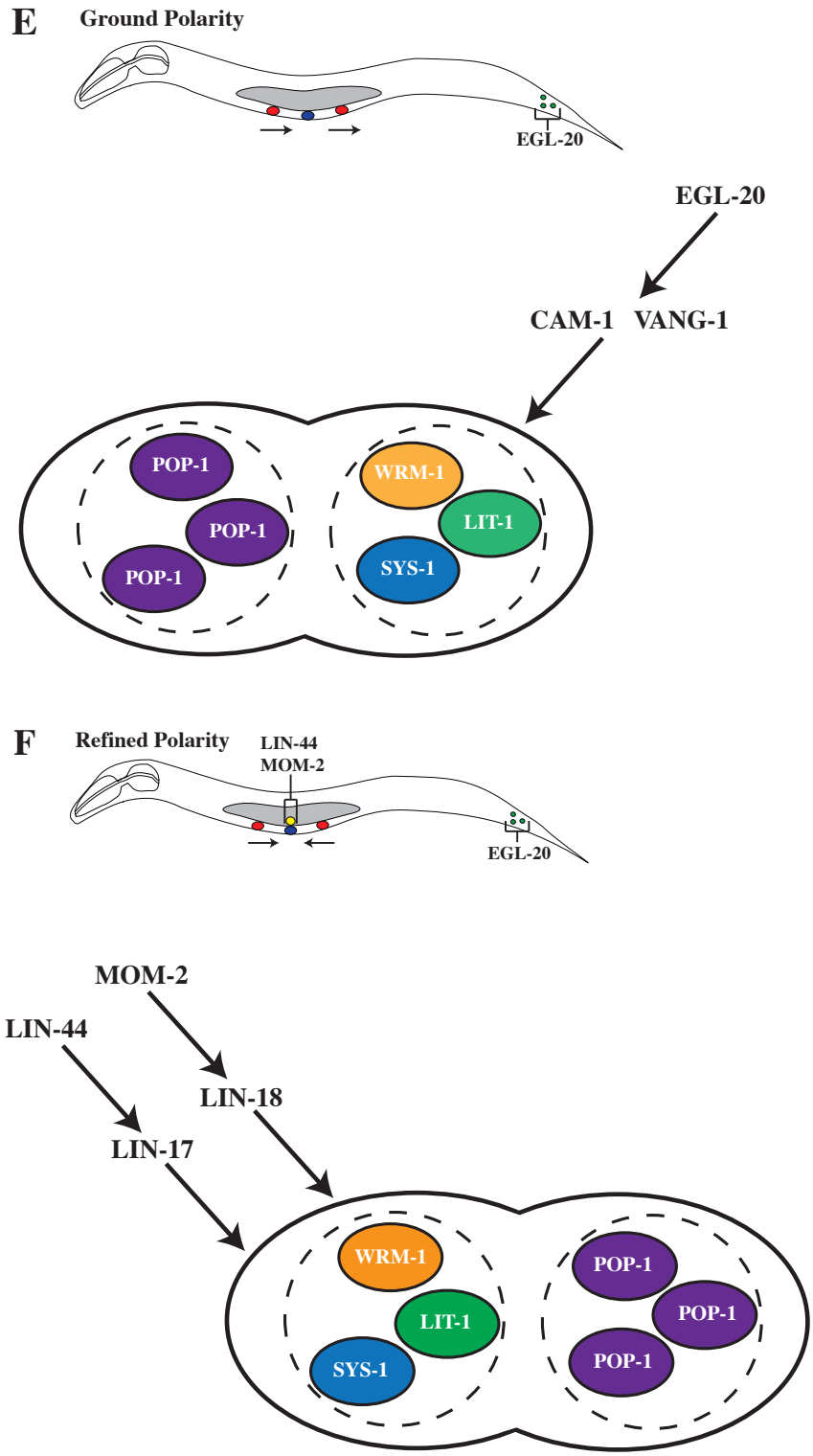


Fig. 1. *C. elegans* vulval development
(A) Schematic of vulval induction illustrating sources of EGF, Notch, and Wnt. (B) Lineage trees of VPC progeny: P5.p, 2° fate, on the left, P6.p, 1° fate, in center, and P7.p, 2° fate, on left. The progeny of each cell is color coded: A cells – red, B cells (B1 and B2) – orange, C cells – yellow, D cells – green, E cells light blue, and F cells dark blue.

(C) Final conformation of vulval cells shown as a cartoon and Nomarski image in mid-L4 stage. Mirror symmetry is noted about the vulval center. Proximal daughter cells of P5.p and P7.p join the daughters of P6.p in forming the vulval lumen whereas the distal most daughters of P5.p and P7.p adhere to the ventral epidermis. (D) The default polarity of P5.p and P7.p is random in the absence of all Wnts. (E) *egl-20* is expressed in the tail (green circles) and establishes ground polarity in which both P5.p and P7.p face the posterior as a result of asymmetric localization of SYS-1, LIT-1, WRM-1 to the posterior daughter of P7.p and POP-1 to the anterior daughter. (F) *lin-44* and *mom-2* are expressed in the anchor cell (yellow circle) resulting in refined polarity where both P5.p and P7.p both face towards the center as a result of asymmetric localization of SYS-1, LIT-1, and WRM-1 to the anterior daughter cell of P7.p and POP-1 to the posterior daughter cell.

Figure 2

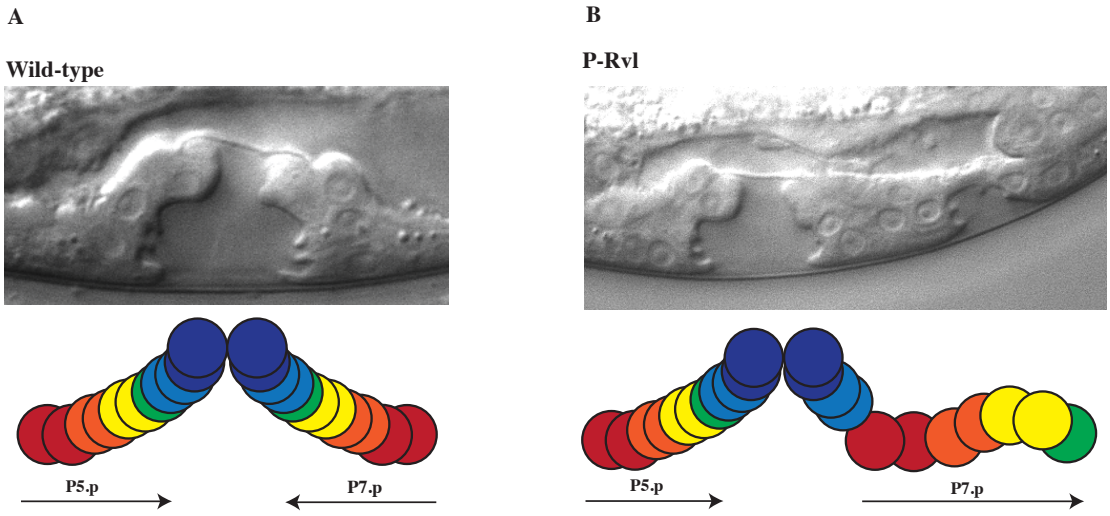


Fig. 2. Wild-type vulva vs. Posterior-reversed vulval lineage vulva
(A) Wild-type vulva formed from 22 progeny of 3 VPCs: P5.p, P6.p, and P7.p. The progeny of P5.p and P7.p form mirror symmetry about the vulval center. (B) Posterior-reversed vulval lineage: the daughter cells of P7.p mimic those of P5.p. Both images taken with *sem-5(n1779)* background.

Figure 3

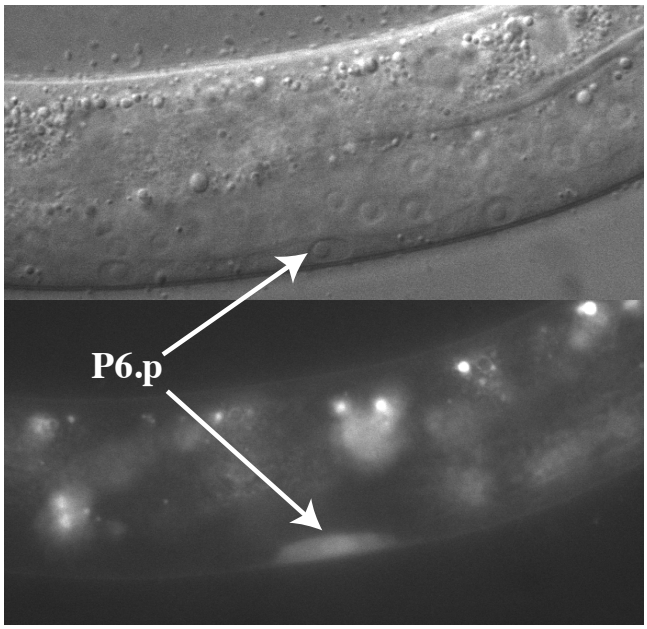
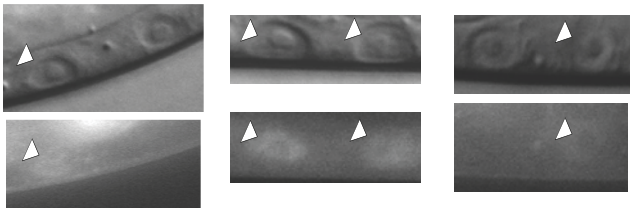


Fig. 3. *egl-17::gfp* expression in P6.p
egl-17 is activated by EGF signaling and is expressed in P6.p. Expression of *egl-17* is used as a marker for vulval induction.

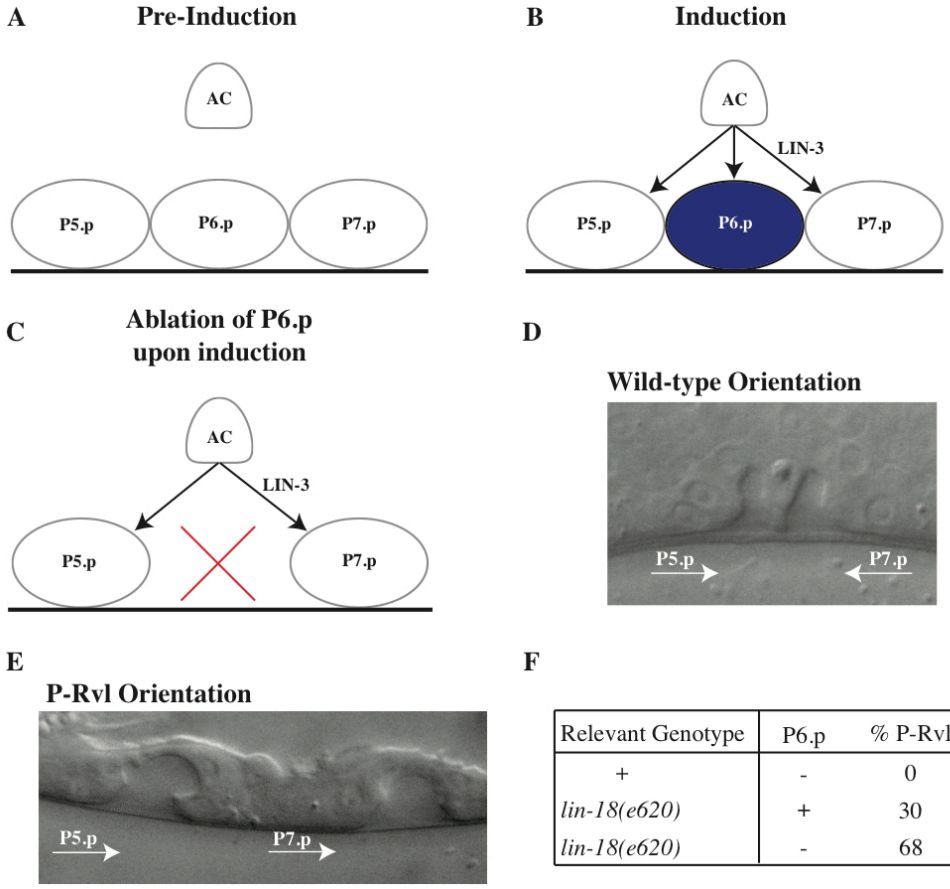
Figure 4



	Number of Worms		
Relevant Genotype	P7.pa > P7.pp	P7.pa = P7.pp	P7.pa < P7.pp
+	20	0	0
lin-17(n671)	3	8	9
lin-18(e620)	9	3	7
sem-5(n1779)	19	1	1
lin-18(e620) sem-5(n1779)	5	6	10

Fig 4. Subcellular localization of VNS::SYS-1
The localization pattern of VNS::SYS-1 in P7.p daughter cells. The resulting pattern was classified by eye into three categories: SYS-1 enriched in the anterior daughter (P7.pa > P7.pp), SYS-1 present at similar levels in both daughters (P7.pa = P7.pp), and SYS-1 enriched in the anterior daughter (P7.pa < P7.pp). A representative image of each scenario is shown.

Figure 5



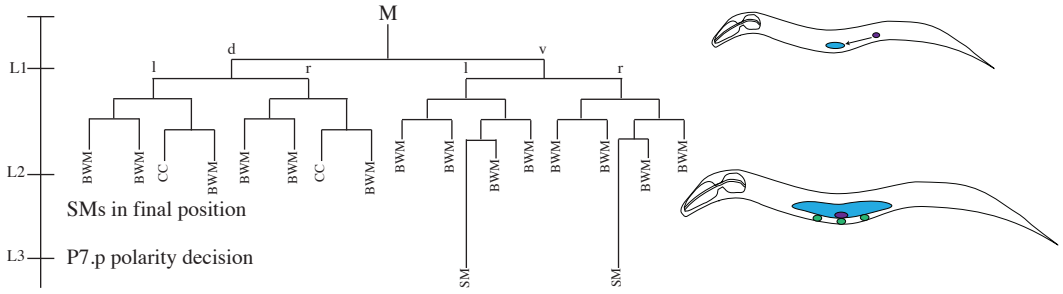
F

Relevant Genotype	P6.p	% P-Rvl	n	p value
+	-	0	25	
lin-18(e620)	+	30	40	
lin-18(e620)	-	68	25	0.004

Fig. 5. P6.p influences the polarity of P7.p
(A) Prior to induction, the anchor cell is directly dorsal to P6.p. (B) During induction the anchor cell produces LIN-3/EGF, which is supplied to P5-7.p. Induction activates *egl-17*, illustrated in blue, within P6.p. (C) P6.p is ablated at the start of induction, but prior to the VPC polarity choice, leaving only P5.p and P7.p, the 2° VPCs. (D) Wild-type orientation of a worm with P6.p ablated (E). P-Rvl orientation of a worm with P6.p ablated. (F) Ablating P6.p enhances the phenotype of *lin-18(e620)*.

Figure 6

A



B

Relevant Genotype	M Cell	% P-Rvl	n	p value
lin-18(e620)	+	30	40	
lin-18(e620)	-	66	29	0.0065
sem-2(n1343); lin-18(e620)	-	68	40	0.0016
lin-18(e620); mig-2(RNAi)	Mig	56	50	0.0188
cwn-1(ok546); lin-18(e620)	+	26	100	
cwn-1(ok546); lin-18(e620)	-	25	20	

C

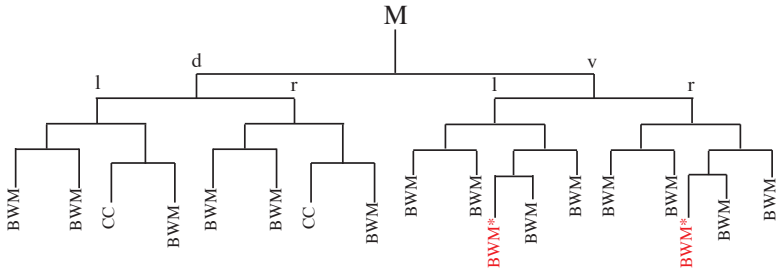


Fig. 6. The sex myoblasts influence the polarity of P7.p.
(A) The sex myoblasts arise from the M cell as shown in the cell lineage. The sex myoblasts are born 13 hours post hatch and migrate anteriorly until they reach their final position, flanking the center of the gonad. The polarity decision of P7.p is made after the sex myoblasts have reached their final position and prior to their first division. The M cell/sex myoblasts are shown in purple, the gonad blue and the VPCs green. (B) Ablating the M cell enhances the phenotype of the *lin-18(e620)* single mutant, but does not enhance the *cwn-1(ok546); lin-18(e620)*, suggesting the SMs, which arise from the M cell, regulate vulval cell lineage polarity and CWN-1 is the necessary cue expressed in the SMs. *sem-2(n1343)* genetically ablates the SMs and *mig-2(RNAi)* causes a migratory defect in the SMs, denoted by Mig. (C) The M cell lineage shown in a *sem-2(n1343)* background. The SMs do not form in this background, but instead become posterior body wall muscle, marked in red with an asterisk.

Figure 7

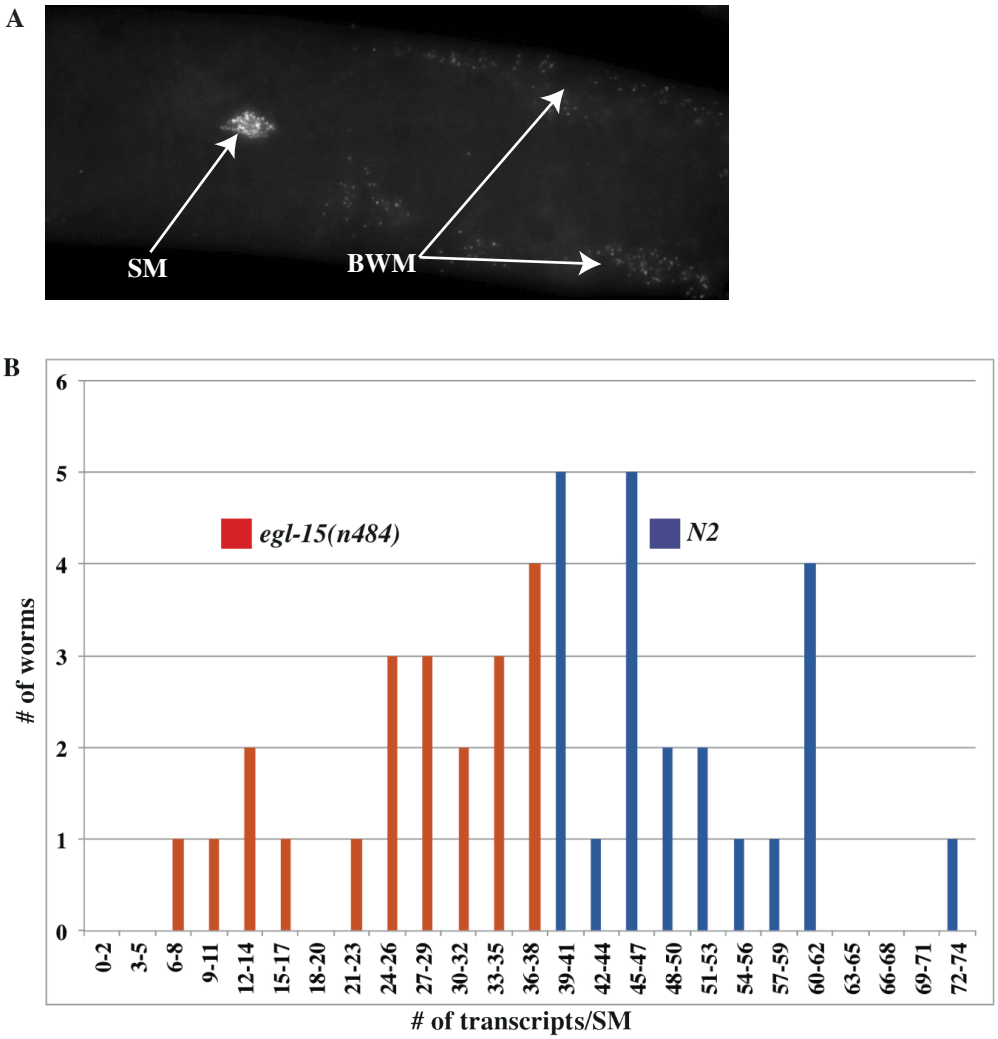
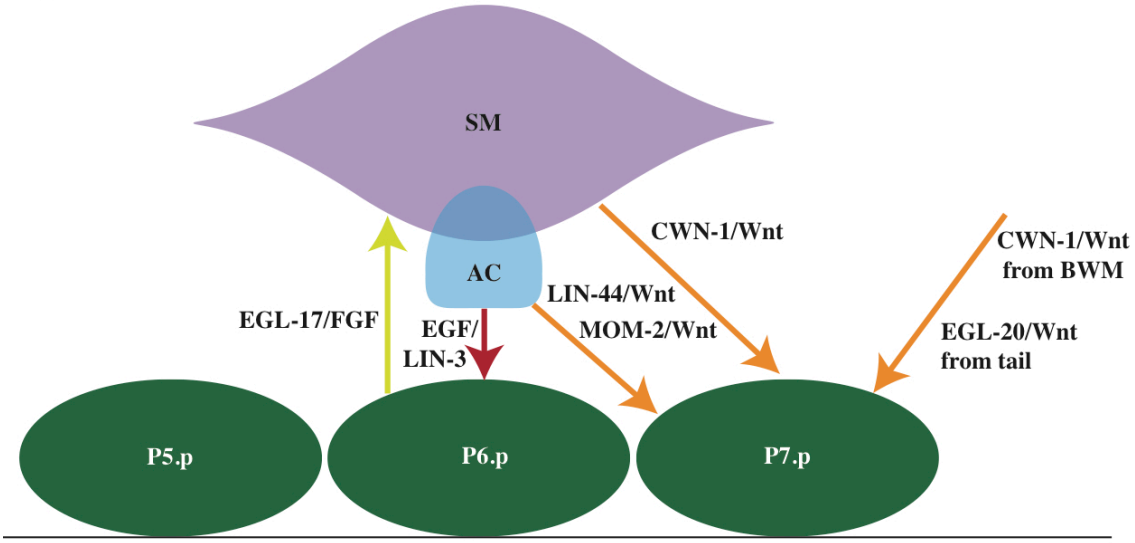


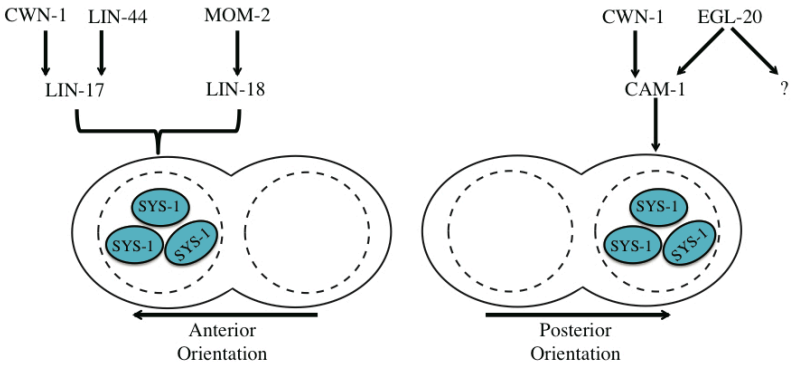
Fig. 7. FGF signaling regulates *cwn-1* expression in the SMs
(A) smFISH analysis of the *cwn-1* transcript in a wild-type worm. *cwn-1* is predominantly expressed in the posterior body wall muscle (BWM) and the M cell/SM lineage. (B) A wild-type sex myoblast at the time of the polarity decision. (C) A histogram quantifying the number of *cwn-1* transcript/SM in both a wild-type and *egl-15(n484)* background.

Figure 8

A



B

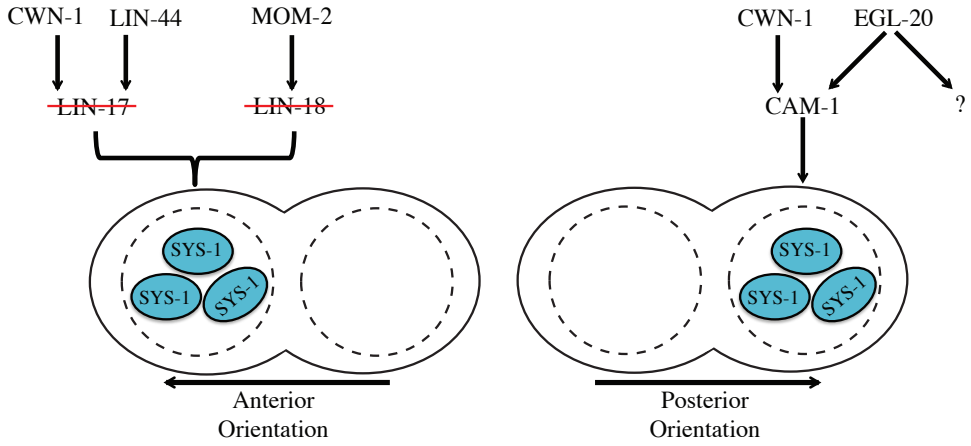


C

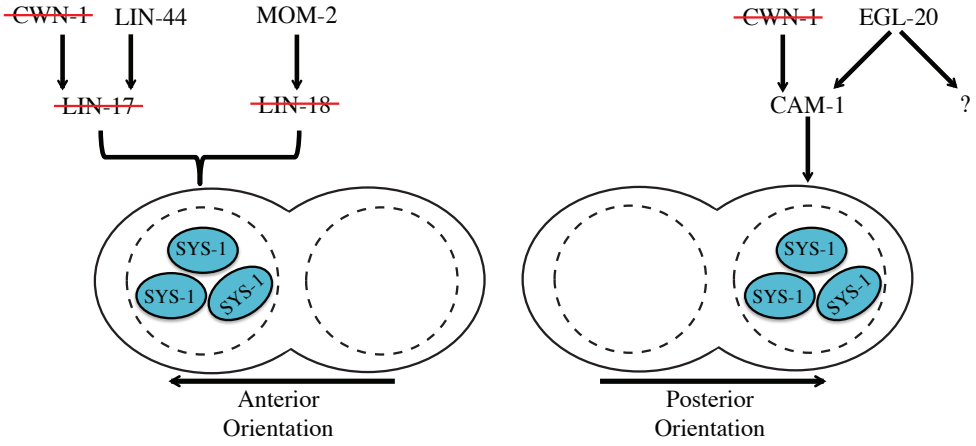
Anterior Components	% P-Rvl	n
lin-17(n671)	74	100
lin-18(e620)	31	100
lin-17(n671); lin-18(e620)	100	40
lin-18(e620); lin-44(RNAi)	70	56
lin-17(n671); mom-2(or42)	100	103
lin-18(e620) egl-15(n484)	63	52
sem-2(n1343); lin-18(e620)	68	40
sem-2(n1343); lin-18(e620); lin-44(RNAi)	82	45

Posterior Components	% P-Rvl	n
lin-17(n671); cam-1(gm122)	46	54
lin-17(n671); vang-1(ok1142)	48	60
lin-17(n671); egl-20(hu120)	6	52
egl-20(hu120); lin-18(e620)	8	51
lin-17(n671); egl-20(hu120); lin-18(e620)	50	52
lin-17(n671); cwn-1(ok546)	53	40
lin-17(n671); cwn-1(ok546); lin-18(e620)	92	47

D *lin-17(n671); lin-18(e620)*



E *lin-17(n671); cwn-1(ok546); lin-18(e620)*



F *sem-2(n1343); lin-18(e620); lin-44(RNAi)*

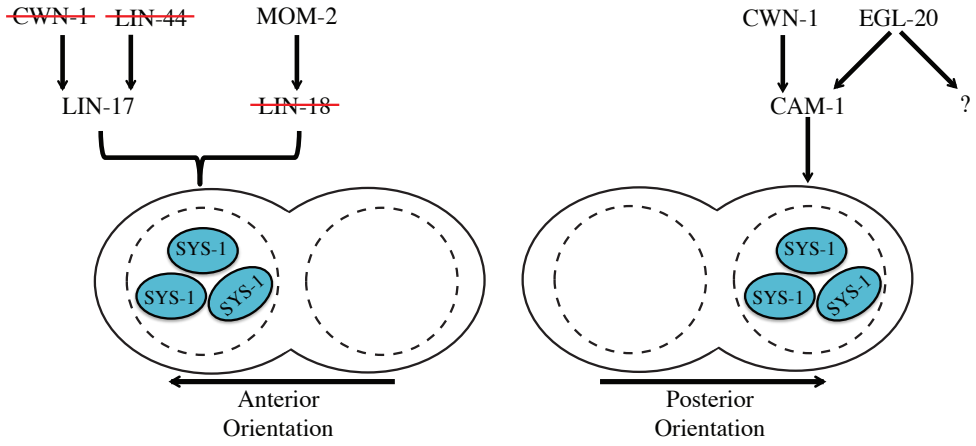
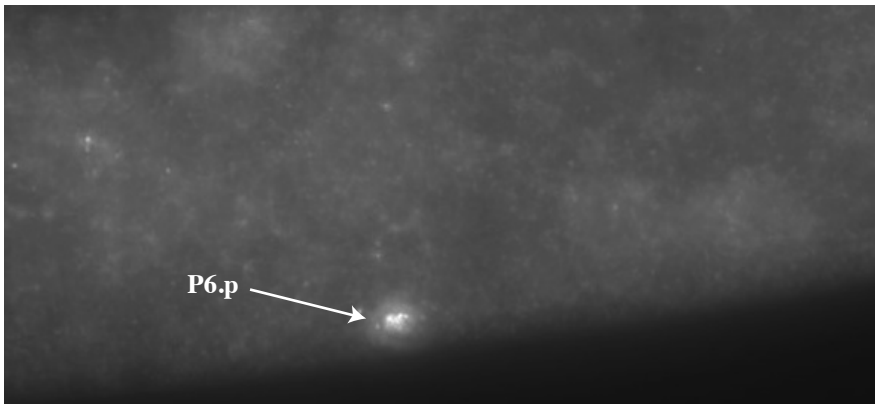


Fig. 8. The role of Wnt signaling on P7.p
(A) The anchor cell (shown in blue) releases LIN-3 (red arrow), inducing the VPCs (shown in green). Induction triggers the expression of the FGF ligand, *egl-17* (yellow arrow), in P6.p, which activates the FGF pathway located in the SMs (shown in purple). The FGF pathway regulates the expression of *cwn-1* in the SMs. The SMs are the anterior source of CWN-1 for P7.p whereas the posterior body wall muscle serves as the posterior source. All Wnt signals are depicted with an orange arrow. *lin-44* and *mom-2*, both expressed anterior to P7.p, express in the anchor cell whereas *egl-20* expresses in the tail, the posterior side of P7.p. (B) CWN-1 from the SMs and LIN-44 act through LIN-17 and MOM-2 acts through LIN-18. All three ligands act to drive SYS-1 localization to the anterior daughter of P7.p. Posterior expressed CWN-1 and EGL-20 act through CAM-1 to drive SYS-1 localization to the posterior daughter of P7.p. Genetic data indicate EGL-20 possibly acts through another, unknown receptor. (C) Genetic data for components of the anterior and posterior acting pathways. Combinations of mutations for anterior pathway components increase the P-Rvl penetrance whereas mutations in posterior components suppress the P-Rvl phenotype. *cwn-1* is the only component found in both pathways. (D) – (F) Examples of how mutations drive phenotypic output. In *lin-17(n671)*; *lin-18(e620)* all anterior receptors are eliminated resulting in a 100% P-Rvl phenotype. Mutations in *cwn-1* eliminate it from both sides of the pathway whereas *sem-2(n1343),* due to a genetic ablation of the SMs, eliminates only the anterior source of *cwn-1*.

Supplemental Figure 1



Supplemental Figure 1. *Pegl-17*::CWN-1::GFP expression
smFISH picture of *Pegl-17*::CWN-1::GFP expressing in P6.p immediately after induction.

Chapter 5

Conclusion

Concluding Remarks

The work presented in this thesis highlights the role of Wnt and FGF signaling in vulval cell lineage polarity. In chapter 2, I described how the low-density lipoprotein receptor, *lrp-2*, functions downstream of *egl-20*/Wnt with *cam-1*/Ror and *vang-1*/Van Gogh to direct the localization of SYS-1/β-catenin. The foundation that this chapter lays for future studies on low-density lipoprotein receptors and their function in Wnt signaling within *C. elegans* is extremely exciting. The first step I would like to see taken is to better understand the biochemistry of the system. Does LRP-2 have the ability to bind with CAM-1 or VANG-1 and form a true coreceptor complex? Furthermore, can LRP-2 physically bind the posterior Wnts, EGL-20 and possibly CWN-1? If there is a physical interaction between LRP-2 and the other transmembrane proteins in this signaling cascade or between LRP-2 and the posterior expressed Wnts, what protein domains are necessary? A structure function analysis would be essential. By better understanding the structure and function of LRP-2 within the context of vulval formation we, perhaps, could learn more about the evolution of other low-density lipoprotein receptors, mainly LRP5/6 (Arrow) and how it has evolved to be an essential component of Wnt signaling in higher order organisms. I would also like to see more work done investigating the role *lrp-1*. I did not pursue *lrp-1* in this work because mutations in this gene cause the worm to arrest at an early larval stage, prior to when one observes the orientation of the L4 vulva. However, this does not mean that *lrp-1* is not involved in vulval lineage patterning or other asymmetric cell divisions within *C. elegans*. I would like to see more work done with this gene to determine any possible function in Wnt signaling within the worm.

In chapter 3, I discussed the origin and evolution of dishevelled. As the hub of

multiple Wnt signaling cascades, dishevelled plays an important role in many developmental pathways across Metazoa. Our findings show that dishevelled is a protein that has undergone dynamic evolution, particularly in nematodes. We found a novel protein domain, the DEP-like fragment, and also present evidence that a nuclear localization sequence has been highly conserved across many dishevelled orthologs. We also hypothesize that much of the dynamic evolution dishevelled has undergone has readied specific orthlogs for functional specialization. Our work raises questions of protein evolution in general and provides clues as to how animals have dealt with the complex intricacies of having a protein, such as dishevelled, act as a central messenger hub connected to many different and vitally important pathways.

As a work that focuses on bioinformatics and theory, there are many future directions to take our findings. I would like to see the functional significance of the DEP-like fragment investigated. With its high level of conservation, one would assume that it does have a functional role in Wnt signaling. I would also like to see the likelihood of dishevelled acting as a transcription factor investigated. Dishevelled possesses many hallmarks of transcription factors, and with the high conservation of a nuclear localization sequence, I think it would be worth pursuing this idea further. Finally, our work has shown that many predicted functional specializations based on domain architecture hold true for the planarian, *Schmidtea mediterranea*, but not enough experimental work has been done in *C. elegans* to fully determine the role that domain architecture plays in functional specialization within nematodes. Wnt signaling through dishevelled has been implicated in many processes within *C. elegans*. I would like to see how well our bioinformatics predict which ortholog of dishevelled is involved within

these Wnt signaling cascades.

In chapter 4, I discussed the interaction between Wnt and FGF signaling in vulval cell lineage polarity. My data illustrate crosstalk between multiple cells in order to direct the orientation of the vulva. I also show that FGF is required to regulate the expression of *cwn-1*/Wnt in the sex myoblasts. The interaction between the sex myoblasts and vulval formation was not previously known, and looking back I am proud to be able to add an additional player to the already impressive list of components involved in vulval signaling. Finally, I also show that despite receptor specificity, all Wnts acting on the vulva have the same molecular output, directing the localization of SYS-1 in the direction of the overall Wnt gradient. This work brings us one step closer to fully understanding how Wnt signaling regulates the orientation of vulval cells.

Looking back at this project I feel there are several questions that should be pursued to further our understanding of the interaction between Wnt and FGF signaling. First, my data shows that FGF regulates the expression of *cwn-1* within the sex myoblasts; however, the level of this regulation is not known. It would be interesting to find the downstream effectors of the FGF pathway and see if they are necessary for upregulating *cwn-1* expression. Currently we do not know the promoter sequence necessary for *cwn-1* expression, but determining this sequence and the effector necessary for expression would be a great leap in our understanding of the Wnt and FGF interaction. On the other hand, perhaps FGF does not directly regulate the expression of *cwn-1*, but rather indirectly regulates expression level through the suppression of a Wnt inhibitor. Regardless of how FGF controls Wnt expression in *C. elegans*, determining the mechanisms would not only be a great discovery in the worm field, but also for Wnt

signaling in general.