Title
Wnt signaling and divergent expression of duplicate genes in the leech Helobdella austinensis

Author
Flores, Vanessa Rae

Publication Date
2017

Peer reviewed|Thesis/dissertation

Wnt signaling and divergent expression of duplicate genes in the leech *Helobdella austinensis*

By

Vanessa R. Flores

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David A. Weisblat, Chair
Professor Gian Garriga
Professor Craig Miller
Professor Lew Feldman

Spring 2017

Abstract

Wnt signaling and divergent expression of duplicate genes in the leech *Helobdella austinensis*

by

Vanessa R. Flores

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor David Weisblat, Chair

Evolutionary developmental biology (Evo-Devo) studies the relationships between changes in the genome, development and body plan throughout evolutionary history. In Evo-Devo, the development of different species is compared to determine similarities and differences, which are interpreted in light of phylogenetic relationships. This helps us reconstruct ancestral states and understand the level of conservation and plasticity of developmental signaling pathways.

Molecular phylogenies have revealed three bilaterian superphyla: Lophotrochozoa, Ecdysozoa, and Deuterostomia. All three superphyla contain both segmented and non-segmented taxa, and the question of when and/or how often this trait evolved remains open. To understand this further we must compare species from all three superphyla, however, most developmental models are drawn from Ecdysozoa and Deuterostomia. *Helobdella austinensis* is a useful model for studying segmentation in Lophotrochozoa. They have large, identifiable blastomeres and develop via largely stereotyped cell lineages. All the segmental mesodermal and ectodermal tissues of the leech arise from a posterior growth zone (PGZ) made of five bilateral pairs of lineage restricted segmentation stem cells (teloblasts) and their mitotic progeny. Cell lineage analyses reveal that leeches undergo a distinct, lineage-driven process of segmentation, as opposed to a boundary-driven process operating in vertebrate and arthropod models.

The goal of this work is to contribute to the understanding of the molecular process regulating the PGZ of *Helobdella austinensis*, focusing on the Wnt signaling pathway. With this knowledge, we can make comparisons with established model systems and draw inferences about the ancestry of this trait. This signaling pathway is conserved throughout bilaterian animals as a regulator of anterior-posterior patterning, segmentation, and stem cell regulation.

It has previously been found that 10/13 of the Wnt ligands encoded in the *Helobdella* genome are expressed in the PGZ. I found activating the Wnt pathway using the small molecule lithium chloride (LiCl), causes the primary neurogenic stem cells (the N teloblasts) of the PGZ to divide symmetrically in a subset of embryos, whereas normally this cell only undergoes highly asymmetric divisions. In some cases, both resulting teloblast-like cells carry out stem cell-like divisions. I also tested for changes in the expression patterns of genes associated with this lineage, and obtained evidence suggesting that β-catenin, a key intermediate in the transduction of Wnt signaling, was upregulated in these embryos. This suggests the Wnt pathway plays a role in the formation of the stereotyped set of leech lineage-restricted stem cells.

Gene duplication is a major contributing process to the diversification of genomes. I carried out two investigations aimed at exploring the consequences of this process in *Helobdella*. One dealt with duplication of an ancestral *wnt16* gene, that appears to have occurred at some point in the lineage leading to the leech from a polychaete ancestor. Previous work showed that *wnt16a* is expressed in ventral ectoderm (N teloblast lineage) between stages 8-10. I show that *wnt16b* exists in at least two isoforms, and that these are expressed in lateral ectoderm (the O and P lineages). Moreover, the expression of at least one of these isoforms switches from ventrolateral ectoderm (O lineage) to dorsolateral ectoderm (P lineage) during embryogenesis. These results provide evidence that the *wnt16* genes have undergone neofunctionalization or subfunctionalization since their duplication. These two genes may contribute to the segmental pattern in the leech by specifying fate and/or division patterns of primary blast cells.

I also investigated the consequences of gene duplication within a more rapidly evolving gene family, namely, the *innexin*s, which encode invertebrate gap junctions. In particular, I compared expression patterns in *Helobdella* to those in the medicinal leech *Hirudo verbana*, to determine levels of conservation and divergence in the expression of *innexin* genes that have undergone duplication in the lineages leading to these modern leeches since their last common ancestor.

In summary, the work presented here suggests that Wnt signaling plays a role in cell division and/or fate specification in multiple stages of embryogenesis in the leech, and thus supports the conclusion that the function of the Wnt signaling pathway is conserved across distantly related taxa. It also provides examples of divergent gene expression between duplicated genes in two different gene families.

Dedication

I dedicate this work to my Mom and Dad, because I could not have accomplished it without their encouragement and belief in my ability to succeed.

Table of Contents

List of Figures …… iii

Chapter 1. Introduction to Evo-Devo, the leech, and Wnt signaling …… 1

Chapter 2. A potential role for Wnt signaling in stem cell division patterns in the posterior growth zone of the leech *Helobdella austinensis* …… 12

Chapter 3. Divergent expression of duplicate *wnt* genes in the leech segmental ectoderm …….. 34

Chapter 4: Characterization of the rapidly evolving *innexin* gene family in the leech *Helobdella* …… 50

Chapter 5: Conclusions and Future Directions……. 67

References ……..... 73

Appendix A: β-catenin sequences and epitopes……... 80

Appendix B: *Fz1/2/7b* sequence with primers and sgRNAs……..... 83

Appendix C: *Wnt16* sequences…… 85

List of Figures

Chapter 1:

Figure 1.1. Phylogenetic tree of Bilateria

Figure 1.2. Stages of Development in *Helobdella austinensis*

Figure 1.3. Boundary Driven vs. lineage-driven segmentation processes

Figure 1.4. Canonical Wnt pathway

Figure 1.5. Components of the *wnt* pathway are highly expressed in the leech PGZ

Chapter 2:

Figure 2.1. Parental vs Grandparental stem cell lineages

Figure 2.2. LiCl appears to affect teloblast division pattern

Table 2.1. Cell number in N lineage after 24 hour LiCl treatment

Figure 2.3. N lineage tracing of LiCl treated embryos reveals two teloblast-like cells in a subset of embryos

Table 2.2. Cell number in OPQ lineage after 24 hour LiCl treatment

Figure 2.4. OPQ lineage tracing of LiCl treated embryos

Figure 2.5. Average number of cells after 24 hour LiCl treatment

Figure 2.6. LiCl-treated embryos with 2 N tlcs make significantly more cells

Figure 2.7. A subset of LiCl treated embryos with two N tlcs have two N bandlets

Figure 2.8. LiCl treatment causes an increase in *sfrp1/2/5c expression*

Figure 2.9. LiCl treatment causes an increase in *sfrp1/2/5c expression*

Figure 2.10. *Pax6a expression* is diminished in LiCl treated embryos

Figure 2.11. β-catenin1 expression increases in the micromere cap after LiCl treatment

Figure 2.12. β-catenin2 expression increases in the micromere cap after LiCl treatment

Figure 2.13. CRISPR-STAT genotyping results

Chapter 3:

Table 3.1. List of primers used to amplify *wnt16* sequences

Figure 3.1. Distinct division pattern of s and f blast cells

Figure 3.2. Genomic architecture of *wnt16a* and *wnt16b*

Figure 3.3. *wnt16a* and *wnt16b-i* expression patterns through stage 8

Figure 3.4. ISH for *wnt16bi* in early stage 8 embryos

Figure 3.5. ISH for *wnt16bi* in mid-stage 8 embryos

Figure 3.6. ISH for *wnt16bi* in late stage 8 embryos

Figure 3.7. ISH for *wnt16bi* in stage 9 embryos

Figure 3.8. ISH for *wnt16bi* in stage 10 embryos

Figure 3.9. *wnt16b-i* is expressed normally in embryos with a knockdown of *bmp5/8*

Chapter 4:

Figure 4.1. Expression of *inx* genes in cleavage stages of *Helobdella*

Figure 4.2. Expression of *inx* genes in stage mid 8

Figure 4.3. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with conserved expression patterns during late development

Figure 4.4. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with potentially conserved expression patterns during late development

Figure 4.5. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with no evidence for conserved expression patterns during late development

Figure 4.6. Late expression of the *innexins* present as multi-copy genes in either *Helobdella* or *Hirudo*

Chapter 5:

Figure 5.1. Expression of β-catenin1 throughout leech development

Figure 5.2. Expression of β-catenin2 throughout leech development

Acknowledgements

I thank all the members of the Weisblat lab, especially my advisor Dr. David Weisblat, for all the helpful, as well as entertaining conversations, and for supporting a positive work environment. I also thank the members of my thesis committee, Dr. Gian Garriga, Dr. Craig Miller, and Dr. Lew Feldman, for all the helpful comments.

Chapter 1. Introduction to Evo-Devo, the leech, and Wnt signaling

Evolutionary developmental biology uses a comparative approach to understand diversity

The animal kingdom exhibits a great diversity of body plans, and a fundamental question in biology is how these differences arose. The field of evolutionary developmental biology (Evo-Devo) aims to address this question by comparing developmental processes between species in light of their phylogenetic relationships. By mapping phenotypic and genotypic states to phylogenies we can reconstruct the ancestral state and better understand evolutionary progression. A major goal of Evo-Devo is to understand the molecular changes that have occurred throughout evolutionary history, what drives these changes, and what kinds of changes are possible. Specific questions addressed by Evo-Devo include: How highly conserved are the fundamental developmental pathways? When differences arise, what molecular changes have taken place? Do certain traits arise once or multiple times? When convergent traits arise, is it due to changes in the same molecular pathway? To address these questions, Evo-Devo studies draw comparisons between animals representing different clades.

By reconstructing animal phylogenies then comparing traits and how they develop and are regulated, we can understand past evolutionary trajectories, and gain insight into the level of plasticity and conservation that exists in developmental pathways. Specifically, if multiple facets of a trait are found to be similar when comparing two species, it suggests that they are homologous traits derived from a common ancestor. Indeed, we can use this methodology to gain insight into what the last common ancestor of all animals looked like, by determining what commonalities exist in many different clades across vast evolutionary distances. For example, a common feature of all bilaterian animals is axial polarity (Petersen and Reddien 2009). This refers to the distinct regions of different body axes, including the head and tail of the anterior-posterior (A-P) axis and the back and belly of the dorsal-ventral (D-V) axis. This brings up many interesting questions, such as, what are the fundamental defining factors of animals, i.e., what makes us distinct from plants, bacteria, archaea, and fungi? and, what is the fundamental toolkit necessary for driving animal development? On the other hand, when we draw comparisons and find differences between species, it can give us clues as to what changes contribute to evolutionary differentiation. In order to understand these questions, it is necessary to study a diversity of animals.

Three bilaterian superphyla

Molecular phylogenies have revolutionized our understanding of animal relationships and evolutionary plasticity, revealing three bilaterian superphyla: Lophotrochozoa, Ecdysozoa, and Deuterostomia (Aguinaldo *et al.* 1997, Figure 1.1). The superphylum Lophotrochozoa includes annelids, mollusks, and platyhelminth flatworms, Ecdysozoa includes arthropods and nematodes, and Deuterostomia includes vertebrates and echinoderms. As discussed, in order to understand the evolutionary history of animals and the vast diversity that is represented in extant animals, it is necessary to compare representatives from all three of these groups. However, most major developmental model organisms are drawn from only the Ecdysozoa and Deuterostomia, including *Drosophila melanogaster, Caenorhabditis elegans*, the mouse, and sea urchins. In contrast, the lophotrochozoan superphylum is not represented by major developmental model

organisms. The work in this thesis aims to contribute to knowledge of an understudied group of organisms, which will further our ability to compare many organisms and further understand evolutionary relationships.

An interesting observation made when comparing these superphyla is the fact that all three contain both segmented and non-segmented clades. Segmentation is a fundamentally interesting trait, as segmented taxa are considered to be some of the most successful, in that they are some of the most speciose and complex (Hannibal and Patel 2013). These taxa include arthropods, annelids, and vertebrates. Although even the definition of what a true segment is has been debated, it is generally defined as repeating body units along the anterior-posterior (A-P) axis of an animal (Hannibal and Patel 2013, Blair 2008). Many questions have been posed regarding this process and its evolutionary history. Because all three superphyla have both segmented and non-segmented taxa, there are two possible scenarios: The trait evolved once at the base of Bilateria, and was lost multiple times, or the trait evolved two or more times independently (Figure 1.1). Some have argued that because similar sets of genes are expressed during segmentation in different animals, that the ancestral bilaterian must have been segmented (Balavoine 2014, Couso 2009, De Robertis 2008). However, others draw attention to the fact that there are many different segmental systems that clearly have little in common, even within the same group of organisms. For example, vertebrates are segmented in three different areas of the body: somites, rhombomeres, and pharyngeal arches. A 2014 review by Graham *et al.* suggests that these are all very distinct processes with different evolutionary histories, and the fact that they all occur in a single group of organisms shows that "segmentation" is just how we label a process, and it does not necessarily reflect a common evolutionary history. To pursue this issue further, we must elucidate segmentation mechanisms in phylogenetically diverse taxa from each superphylum.

Helobdella as a lophotrochozoan model

For this purpose, the Weisblat lab studies the leech *Helobdella austinensis*, a segmented representative of the Lophotrochozoa. Leeches of the genus *Helobdella* are useful for developmental studies. Their embryos are large (~400µm), easy to collect, readily survive in easy to prepare medium, and have large, identifiable cells that can be micro-injected with tracers, plasmids, and/or RNA. Their development is stereotyped, enabling us to manipulate the embryos and identify aberrant cell divisions and cell fates. There are also several genes that have been found to be expressed in specific lineages, and can therefore be used as lineage markers. The genome for a sister species, *Helobdella robusta*, was sequenced in 2007 (Simakov *et al.* 2013), and the high similarity between the two species allows us to utilize this information for studying *H. austinensis*.

Leech segmental tissues arise from a posterior growth zone (PGZ), which forms during cleavage. Leeches of the genus *Helobdella* undergo spiralian cleavage (as do all annelids) and go through 11 stages of development (Figure 1.2). In early stages, unequal cleavage gives rise to an 8-cell stage consisting of four macromeres (A', B', C' and D') and four micromeres (a', b', c', and d''). The D' macromere inherits yolk-free, RNA-enriched cytoplasm called teloplasm. This cytoplasmic inheritance leads D' into a unique series of further divisions, giving rise to five bilateral pairs of lineage-restricted stem cells called teloblasts; there is one pair of mesodermal stem cells (M teloblasts), and four pairs of ectodermal stem cells (N, O, P, and Q teloblasts, from ventral to dorsal). Teloblasts undergo highly asymmetric divisions to give rise to bandlets of

segmental founder cells called blast cells. The teloblasts and their immediate progeny are referred to as the PGZ. Later in development left and right germinal bands migrate anteriorly where they coalesce to form the germinal plate. Segments differentiate in anteroposterior progression in stages 9-11, after which the juvenile emerges (Weisblat and Kuo 2014).

Different modes of segmentation

Comparisons between leeches and model organisms suggest there are at least two cellular modes of segmentation. I define the first as boundary-driven segmentation. During this process, variants of which are seen in vertebrate and arthropod models, boundaries are imposed on an initially equipotent field of cells. As patterning occurs subsequently within each segmental unit, clones descended from individual cells are often variable in spatial distribution and cell type composition but fail to cross the boundaries of segmental or sub-segmental divisions. In contrast, what is observed in leech is referred to as lineage-driven segmentation. In this case, segmental founder cells of fixed lineages undergo stereotyped cell divisions. Progeny interdigitate, and patterned units arise due to stereotyped division patterns, even though the clones of several types of individual founder cells are distributed across two or more segments (Weisblat 1985, Weisblat and Shankland 1985, Seaver and Shankland 2001, Figure 1.3).

It is also important to note the difference between one of the most well-studied segmented invertebrates, *Drosophila*, and most other segmented invertebrates and vertebrates. *Drosophila* undergoes long germ band development, where all segments are formed at once. However, most annelids, and indeed many arthropods, undergo short germ band development. In short germ band development, segments are formed sequentially, where new segments are formed posteriorly in a growth zone (McGregor *et al.* 2009, Bolognesi *et al.* 2008). This is thought to be the ancestral form of segmentation, which is a good reason to understand the PGZ of diverse animals, including the leech.

The Wnt pathway

One of the highly conserved developmental signaling pathways that has been shown to be important in the segmentation process of many species is the Wnt pathway. This pathway is highly conserved across the Metazoa and functions not only in segmentation, but also in axial polarity, cell fate specification, and stem cell maintenance and/or differentiation (Nusse 2012, Buechling and Boutros 2011). Components of this pathway have been shown to be involved in a signal transduction cascade which results in gene transcription, as well as in non-transcriptional events. However, the most well described (and the first discovered) Wnt pathway is what we refer to as the canonical Wnt signaling pathway (Figure 1.4). In the canonical pathway, the Wnt ligand binds the seven-pass transmembrane receptor Frizzled (Fz), which transduces the signal to Disheveled (Dsh), which is thought to directly interact with what is called the destruction complex (Fiedler *et al.* 2011). The destruction complex is composed of GSK3β, Adenomatous Polyposis Coli (APC), and Axin proteins, and is so named because, in the absence of Wnt signaling, they form a functional unit that actively phosphorylates β-catenin, thereby targeting it for degradation. However, when the Wnt ligand is present, Dsh functions to dissociate the complex, thereby allowing β-catenin to accumulate and enter the nucleus. There it acts as a transcriptional modulator by interacting with transcription factors (Nusse 2012, Buechling and Boutros 2011).

The Wnt pathway is very important in animal development and it is important to understand the molecular details and level of conservation across species for many reasons. Not only is it involved in fundamental cellular processes, but it has also been shown that the dysregulation of this pathway can result in cancer in many tissues (Sedgwick and D'Souza-Schorey 2016). Therefore, by studying the role of this pathway in the leech PGZ, we can also increase our general understanding of conserved features of this pathway and how they may be dysregulated. Below I discuss some of the conserved features of Wnt signaling across bilaterian animals.

The Wnt pathway is conserved throughout Bilateria as a regulator of A-P patterning

The leech PGZ gives rise to cells in anterior to posterior progression. The Wnt pathway has been found to regulate A-P patterning across the Metazoa, thus making this an excellent candidate pathway to investigate in the leech PGZ. One of the most striking similarities found across many different species, including deuterostomes and protostomes, is that the Wnt signal is located in the posterior of the developing embryo, whereas Wnt inhibitors are expressed at the anterior (Janssen *et al.* 2010, Petersen and Reddien 2009). Wnt signaling has even been shown to play a role in axial patterning in cnidarians, which predate bilaterians, suggesting that this is an ancestral feature (Lengfeld *et al.* 2009). Interestingly, this A-P pattern occurs not only in the patterning of the main body axis, but specific tissues as well. For example, in *Xenopus* embryos, an A-P gradient of Wnt signaling is required to form the neural plate (Kiecker and Niehrs 2001). This pathway is also required for correct A-P patterning in developing tissue of the blastema in regenerating planaria (De Robertis 2010, Petersen and Reddien 2008). It has also been found in many organisms that Wnt signals provide axonal guidance cues along the A-P axis (Zou 2006, Lyuksyutova 2003).

Wnt pathway in segmentation

It has long been recognized that the *Drosophila wingless (wg)* gene, homologous to *wnt1* in other species, plays a crucial role in segmentation in this species as a segment polarity gene. In this long germ band developing species, a succession of genes expressed in transverse stripes delineate narrower bands of expression with tighter boundaries until the last set, the segment polarity genes, define the distinct segmental boundaries. *Wg* is a segment polarity gene and upregulates the *engrailed* gene in the directly posterior cells. This turns on expression of *hedgehog*, which functions to maintain *wg* expression in the anterior cells. This imposes a border defining the boundary of a parasegment (Swarup and Verheyen 2012).

Wnt signaling plays a role in segmentation processes of other protostomes as well, such as the short-germ developing arthropod *Tribolium castaneum* (the flour beetle). In this insect, it was found that multiple *wnt* genes are expressed segmentally and in the posterior growth zone, and that knocking down gene function results in failure to form abdominal segments (Bolognesi *et al.* 2008). A 2010 study of multiple protostomes by Janssen *et al.* also found multiple *wnt* genes to be expressed in a segmental pattern as well as the PGZ of these animals. Thus, it may be that the involvement of a single *wnt* gene (*wg*) in *Drosophila* segmentation represents a secondary simplification.

The Wnt pathway has also been found to be important in segmentation of vertebrate paraxial mesoderm. This tissue forms segmental units called somites. This process involves

what is known as a segmentation clock. Genes that are involved in this process are cyclically expressed (Geetha-Loganathan *et al.* 2008). Nuclear β-catenin is expressed in a posterior to anterior gradient in this tissue (Graham *et al.* 2014). It is thought to regulate the interval at which the clock genes are expressed (Gibb *et al.* 2009).

Wnt and the leech posterior growth zone

Given the stereotyped nature, cellular processes of leech development are fairly well understood. The overall goal of my thesis is to contribute to our understanding of the molecular mechanisms regulating the PGZ and segmentation process of the leech. With this knowledge, we can make comparisons with established model systems and make inferences about the ancestry of this trait. As discussed, the Wnt pathway is an excellent candidate for PGZ formation and axial patterning because it is a conserved pathway seen throughout Bilateria as a regulator of anterior-posterior patterning and segmentation. The Weisblat lab has been interested in studying the Wnt pathway for some time. A former postdoctoral researcher in the lab, Sung-jin Cho, led a study surveying expression patterns of the *wnt* genes and other components of the pathway (Cho *et al.* 2010). He found that the *Helobdella* genome contains 13 *wnts* representing 9 of the ancestral subgroups. Ten of these genes are highly expressed in at least part of the PGZ (Figure 1.5). This is consistent with expression data observed in other animals and supports the hypothesis that a conserved function of this pathway is in axial patterning. The leech PGZ is composed of the teloblasts and nascent blast cells. Therefore, the specific function of the Wnt signaling pathway in regulating the leech PGZ may be in teloblast formation/function and/or blast cell division pattern and/or fate specification, which would contribute to the segmentation pattern, due to the lineage-driven segmentation process in the leech. The work I present here further supports this hypothesis. I show that small molecule activation of the pathway results in an aberrant teloblast division. I also show that the two copies of a duplicated *wnt* gene show divergent expression patterns, although both are segmentally iterated. This suggests that Wnt signaling may be involved in division patterns and/or cell fate specification at multiple stages of development in the leech *Helobdella austinensis*. In addition, I explore conservation and divergence of duplicate genes in a different gene family, the *innexins*.

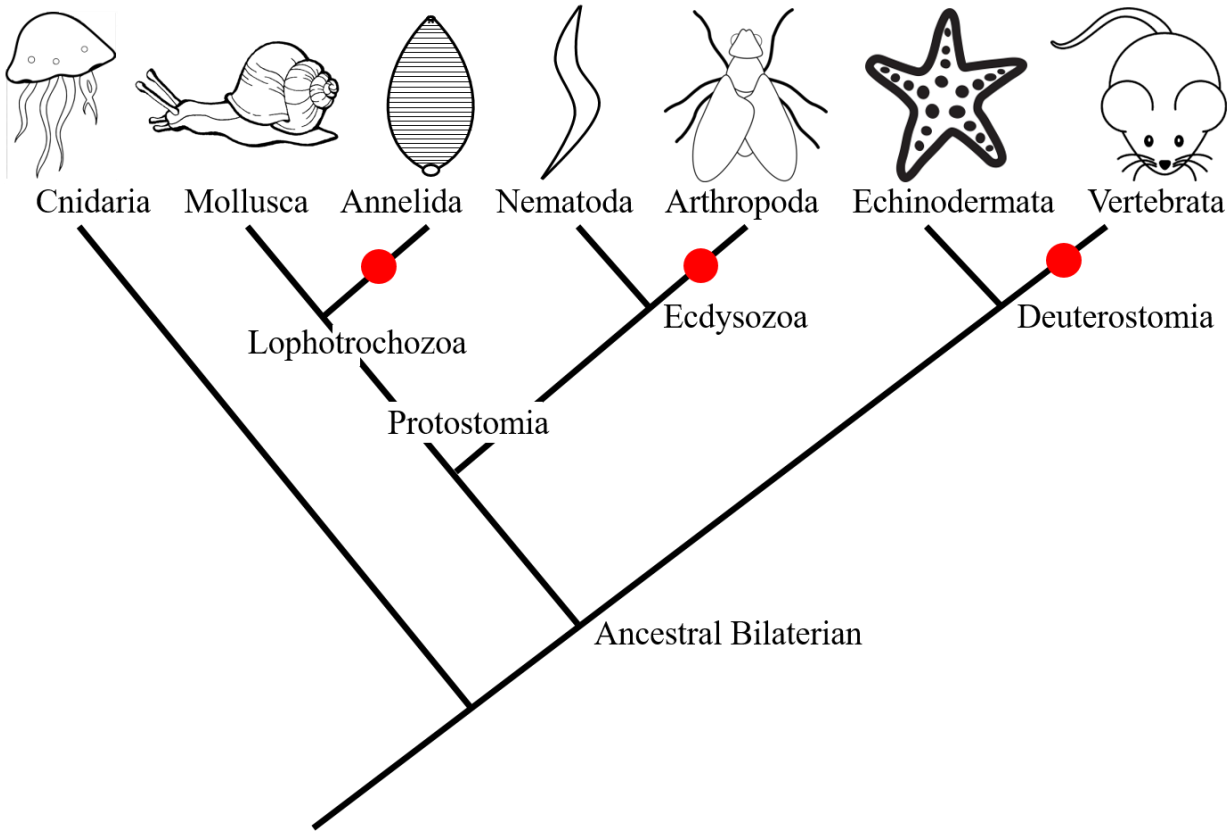


Figure 1.1. Phylogenetic tree of Bilateria. The tree depicts the three superphyla of bilaterian animals, Lophotrochozoa, Ecdysozoa, and Deuterostomia. Red dots indicate groups of segmented animals, illustrating that each superphylum contains both segmented and unsegmented clades. The Cnidaria are an outgroup to the Bilateria, and are unsegmented.

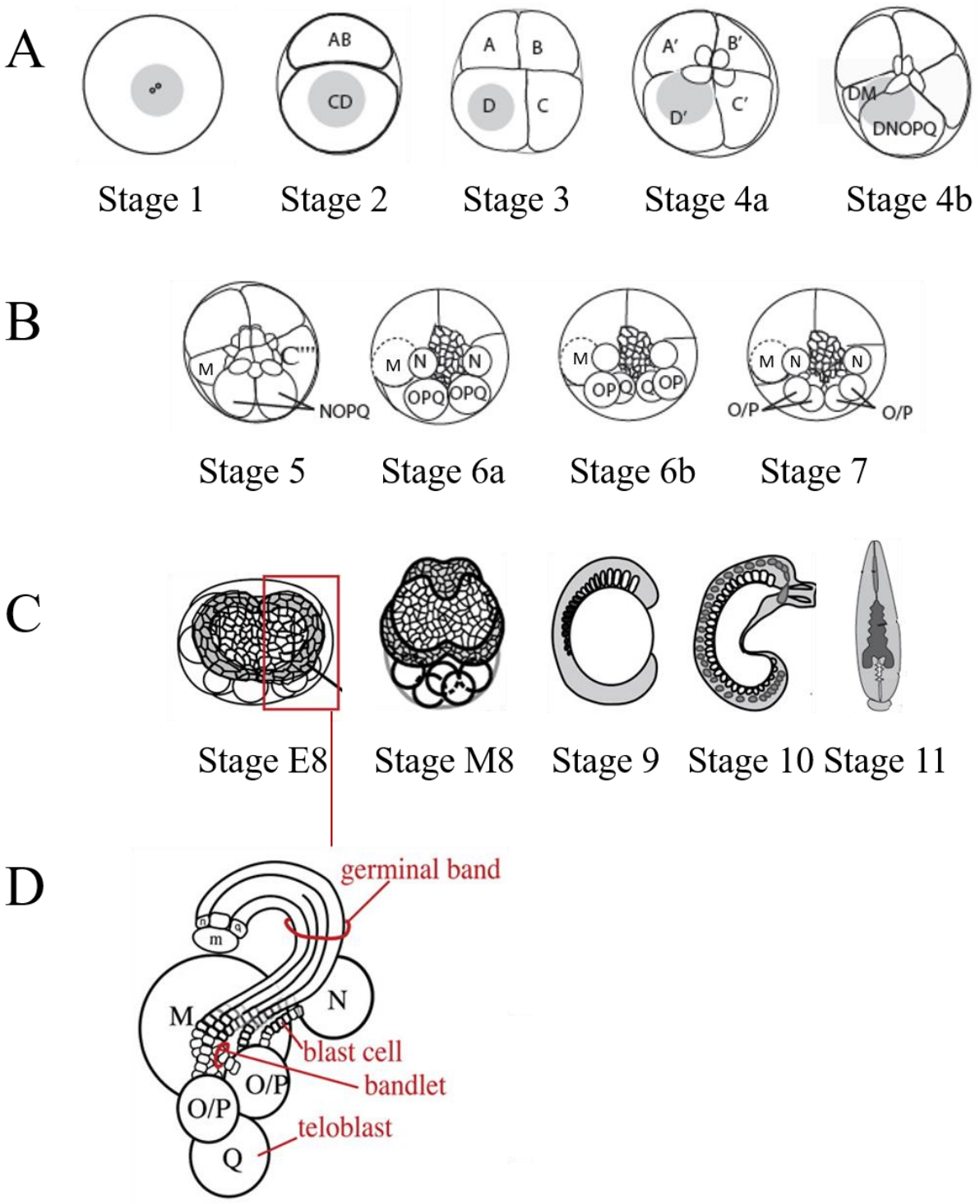


Figure 1.2. Stages of Development in *Helobdella austinensis*. Adapted from Weisblat and Kuo 2014. Leech development has been subdivided into 11 stages. Panel A depicts early cleavage stages. During this time, the embryo undergoes unequal spiral cleavage to produce macromeres and micromeres, and segregates RNA-rich cytoplasm called teloplasm (gray circle) into the D macromere. Panel B depicts teloblastogenesis, a stereotyped set of divisions of the

mesodermal proteloblast DM and ectodermal proteloblast DNOPQ, to produce the five bilateral sets of teloblasts: M, N, O/P, O/P, and Q. As illustrated in Panel D, these teloblasts undergo highly asymmetric divisions to produce a column of blast cells called a bandlet which progress anteriorly into the germinal band. The germinal bands coalesce to form the germinal plate at the anterior of the embryo, and zipper together along the ventral midline. This is illustrated by the stage 8 embryos in Panel C. Panel C also depicts the late stages of development, stages 9-11, when segmentation, dorsal closure, and yolk exhaustion occur to complete development. Panels A and B depict embryos in animal view. Stage 8 and 11 embryos are shown as dorsal views. Stage 9 and 10 are lateral views.

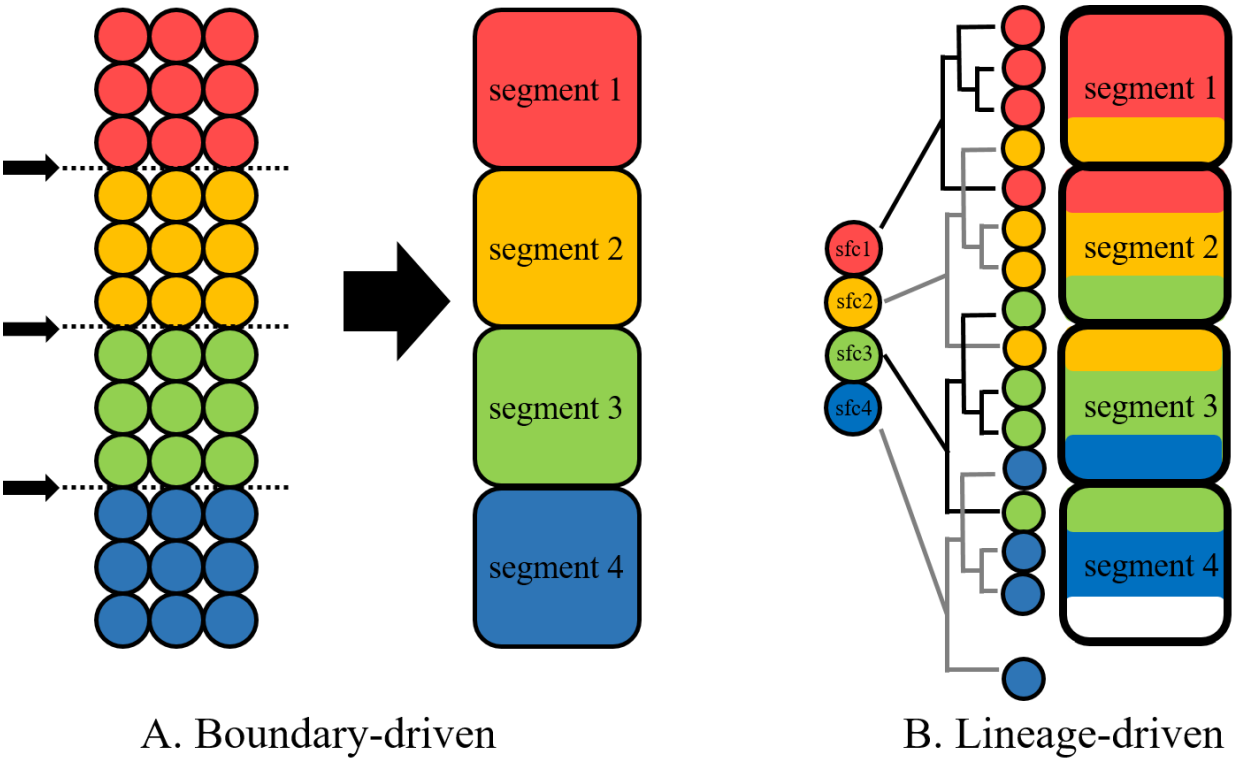


Figure 1.3. Boundary Driven vs. lineage-driven segmentation processes. From Weisblat and Kuo 2014. The segmentation process in model organisms such as *Drosophila melanogaster* are boundary-driven, where boundaries are imposed on a field of initially equipotent cells by signaling events, and patterning occurs within each unit. On the other hand, the leech undergoes lineage-driven segmentation, where there are bandlets of cells of fixed lineages. These segmental founder cells undergo a stereotyped division pattern. Clones interdigitate across segmental boundaries, but due to the stereotyped process, patterned units arise. In panel B, sfc stands for segmental founder cell.

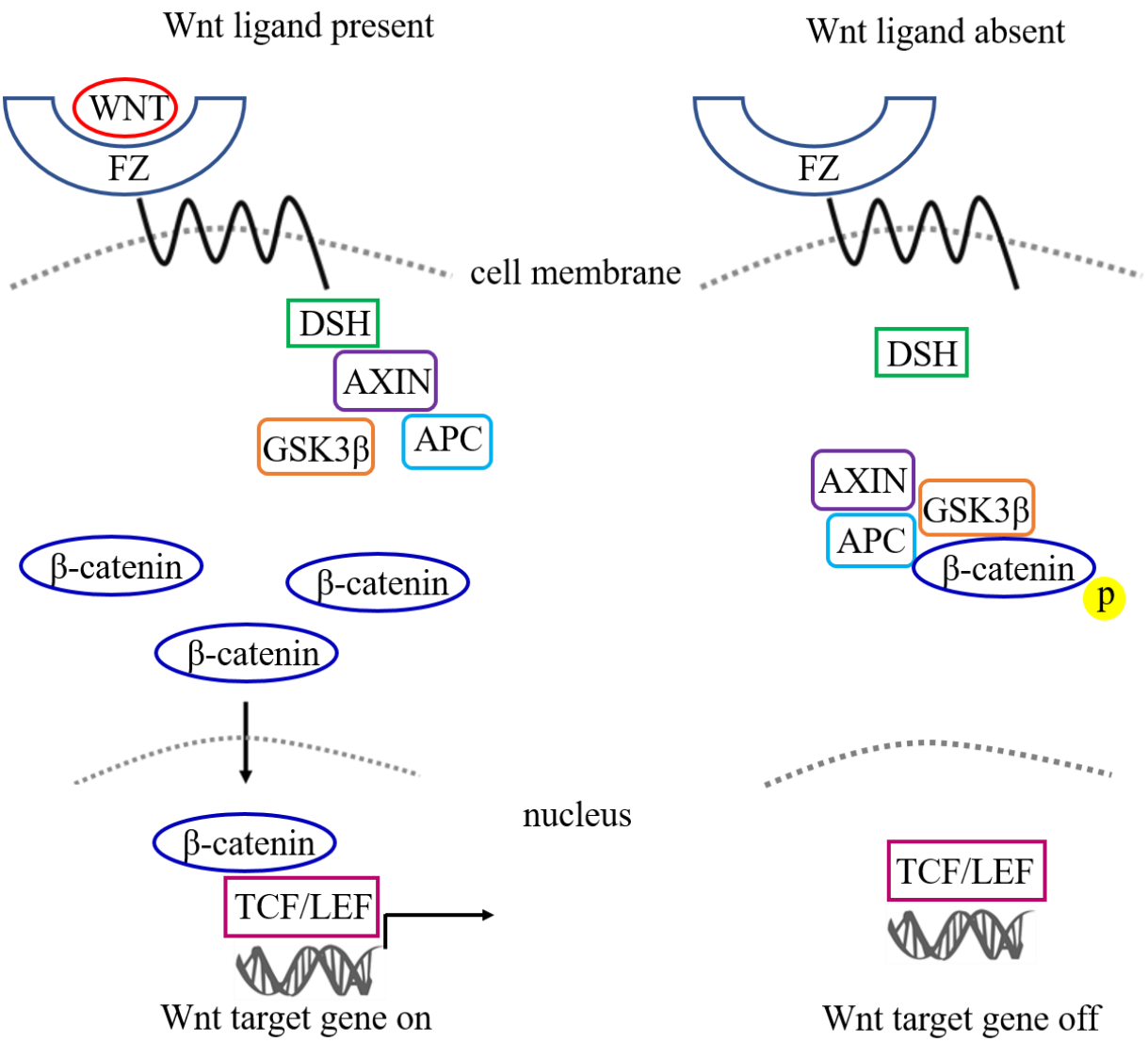


Figure 1.4. Canonical Wnt pathway. This shows the main players in the Wnt pathway and their interactions. When the Wnt ligand is present it binds to the seven-pass transmembrane receptor Frizzled, which transduces the signal to Disheveled. Disheveled inactivates the destruction complex, composed in part of Axin, APC, and GSK3β. This allows β-catenin to accumulate then enter the nucleus and interact with the TCF/LEF transcription factor. With no Wnt present the destruction complex phosphorylates β-catenin and targets it for degradation.

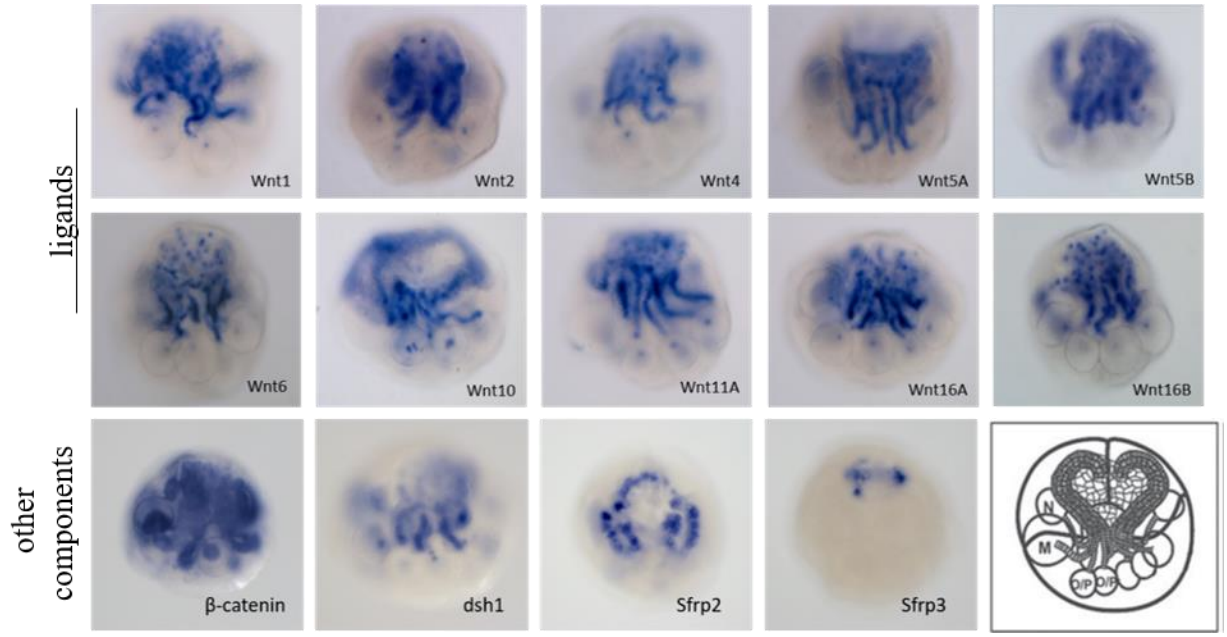


Figure 1.5. Components of the *wnt* pathway are highly expressed in the leech PGZ. Adapted from Cho *et al.* 2010. Ten of the thirteen leech *wnt* genes are highly expressed in the posterior growth zone, which includes the teloblasts and their immediate progeny. Other components shown illustrate how other genes in this pathway are also either highly expressed in the PGZ, or show a specific expression pattern.

Chapter 2. A potential role for Wnt signaling in stem cell division patterns in the posterior growth zone of the leech *Helobdella austinensis*

Introduction

It is important to study embryonic stem cells across many species

Research in embryonic stem cells (ESCs) has produced great advances in our understanding of what genes regulate cell fate decisions during animal development. The advent of induced pluripotent stem cells (iPSCs) has revolutionized the study of medicine (Takahashi and Yamanaka 2006). Stem cell biology also gives us insight into mechanisms of cancer. It is therefore extremely important to understand the biology of stem cells. Much work has been done in mammalian stem cells; in the past several years we have gained great insight into what genes are expressed in ESCs to maintain self-renewal, as well as what genes promote differentiation into specific cell types (Nusse *et al.* 2008, Okita and Yamanaka 2006). In addition, many adult stem cell types, such as intestinal stem cells, are well studied. However, ESCs in groups of animals besides mammals and the major model organisms are not as well studied. If we can understand what genes are involved in ESC division pattern, self-renewal, and fate specification across a vast array of species, we can understand more about the level of conservation of developmental signaling pathways. It is important to understand the fundamental nature of stem cells and how deeply engrained certain genetic pathways are in associated processes. This could give us greater insight into the best pathways to target in medicine, as well as what the rudimentary necessities are in animal development. On the other hand, studying stem cells in diverse systems may also reveal a diversity of mechanisms by which the operationally defined phenomena characteristic of stem cell biology can be achieved.

Leeches are an excellent model in which to study stem cells. Embryos of clitellate annelids, as exemplified by those of glossiphoniid leeches, have five pairs of lineage-restricted stem cells that undergo a stereotyped division pattern. Within this set of stem cells, there are several interesting features to study which could give us more insight into stem cell function. For example, of the five lineages, three (M, and the two O/P) undergo a parental stem cell division pattern, whereas two (N and Q) undergo a grandparental stem cell division pattern (Figure 2.1). In a parental stem cell lineage, each division produces the same daughter cells: the stem cell and the same kind of blast cell each time. On the other hand, in a grandparental stem cell lineage, besides the stem cell, two different kinds of blast cells are produced in exact alternation (Zackson 1984). It is unknown whether the two kinds of blast cells differ at birth or if some later event results in their differentiation. Another interesting aspect of leech stem cells is the opportunities they provide for studying different mechanisms of lineage specification. For example, it has been shown that the two O/P teloblasts are equipotent at birth in species of the leech genus *Helobdella*. In early stage 8, a *bmp5-8* signal from the Q lineage directs the adjacent lineage towards to the P fate, whereas the other lineage takes on the O fate, which is the default fate (Kuo and Weisblat 2011). In contrast, teloblasts that generate the other three lineages (M, N, and Q), are committed to their fates at the time of their birth, as judged by their failure to change fates in response to various combinations of teloblast ablations.

Leech stem cells could also give us insight into regeneration processes. Leeches undergo determinate growth, making exactly 32 segments during embryogenesis and no more during their

remaining life; They are also unable to regenerate segments (Bely 2006). By contrast, most oligochaetes, representing the group from within which leeches evolved, undergo indeterminate growth, meaning they can continue to add segments to their posterior end throughout their adult life, and many exhibit robust ability to regenerate segments. Understanding the differences between the stem cells and the rest of the posterior growth zone (PGZ) between leeches and oligochaetes could give insight into the field of regeneration.

The cells of the leech PGZ are produced through a stereotyped set of cell divisions. In the initial cleavage stages the D quadrant of the four-cell embryo inherits yolk-free, RNA-rich cytoplasm (teloplasm). The D' macromere (after contributing one micromere) undergoes a stereotyped pattern of cell divisions to generate the five pairs of bilateral teloblasts. First it divides to form the separate mesodermal and ectodermal precursors, DM and DNOPQ. After additional rounds of asymmetric divisions, the granddaughter (DM") and great-granddaughter of (DNOPQ"') of these precursors divide symmetrically to form the two mesodermal teloblasts M_L and M_R, plus the left and right NOPQ ectodermal precursors. Each NOPQ cell then divides to give rise to the N teloblast, plus OPQ proteloblast. OPQ gives rise to the Q teloblast and the OP proteloblast. The OP proteloblast gives rise to four blast cells before the division giving rise to the two O/P teloblasts (additional micromeres are also produced during these divisions). The five pairs of teloblasts are the ten lineage-restricted stem cells within the leech PGZ (Weisblat and Kuo 2014, Figure 1.2).

Wnt signaling in stem cells

The Wnt signaling pathway has been shown to be involved in various stem cell processes, depending on the type of stem cell. It is required for self-renewal and stem cell maintenance (Sokol 2011, Nusse *et al.* 2008). In a 2012 study by Blauwkamp *et al.*, it was shown that different levels of Wnt signaling result in differentiation into distinct lineages in cultures of human ESCs (hESCs). They showed that cells treated with a Wnt inhibitor (IWP) maintained pluripotency, and were more likely to differentiate into neural cells when they also inhibited signaling by the transforming growth factor-β (TGF-β) pathway. On the other hand, hESCs treated with Wnt protein were more likely to be directed towards mesodermal or endodermal fates.

Multiple studies have shown that a directed Wnt signal can influence the symmetry of division in different cell types. Goldstein *et al.* in 2006 showed that presenting a directed Wnt signal to blastomeres of *C. elegans* orients the plane of division, directs the localization of the Fz receptor, and polarizes cell fates. Similar experiments were later done using Wnt3a protein-coated beads to present directed signals to ESCs. A 2013 study by Habib *et al.* showed that a directed Wnt signal can influence the symmetry of division in ESCs. By presenting the Wnt3a protein to one side of the ESC, they showed that expression of other signaling components of the pathway was induced in the proximal daughter. In addition, the majority of cells that divided did so on an axis oriented in line with the signal. They found that cells with the Wnt3a protein presented on either side divided symmetrically.

In the leech, it has been observed that *Fz1/2/7b* is the only *Fz* expressed at stage 6 and 7, and that it is specifically expressed in the N teloblast (Cho and Wang unpublished). This suggests a potential role of Wnt signaling in N teloblast fate and/or division pattern. I found further support for this idea when preliminary experiments with LiCl, which activates the Wnt pathway by inhibiting GSK3β (a component of the β-catenin destruction complex), appeared to

cause a symmetric division of the N teloblast (Figure 2.2). Normally, once the N teloblast is born it only undergoes highly asymmetric divisions which give rise to primary blast cells and a micromere. Here, I test the hypothesis that Wnt signaling plays a role in stem cell maintenance, specification, and/or division pattern of the N teloblast. I show that treatment of embryos with 2mM LiCl at the onset of teloblastogenesis results in a symmetric division of the N teloblast in a subset of embryos, suggesting that this highly conserved signaling pathway does play a role in the stereotyped division patterns which give rise to the stereotypical set of five bilateral pairs of teloblasts.

Materials and Methods

Small molecule treatment

At the onset of stage 5, just after the bilaterally symmetric division of ectodermal precursor DNOPQ, embryos were placed in a bath solution of Htr containing 2mM LiCl. Control embryos were placed in a bath solution of Htr containing an additional 2mM of NaCl.

Lineage tracing

While embryos were being continuously treated (with either LiCl or NaCl), just after NOPQ completes cleavage, either the N teloblast or OPQ proteloblast was injected with rhodamine dextran amine (RDA) and mRNA of green fluorescent protein conjugated to *histone2b* (*h2b:gfp)*. To determine the number of cells and cell types, embryos were fixed after 24 hours and imaged using an LSM 710 confocal microscope.

ISH

Embryos were treated for 24 hours, then washed with Htr medium (Weisblat and Kuo 2009). They were allowed to develop until the NaCl-treated control embryos reached mid-late stage 8, then fixed in 4% paraformaldehyde (PFA) for one hour at room temperature, rinsed in PBS, and dehydrated in a methanol series to 100% and stored at -20°C. Embryos were devitellinized in 50% methanol using insect pins. Riboprobes labeled with digoxygenin were made using the kit. The initial ISH steps were performed at room temperature. Before ISH, embryos were rehydrated into PBS. Embryos were permeabilized with Proteinase K (concentration, time, temperature), followed by three five minute washes of glycine solution to stop the reaction. Embryos were post-fixed for one hour, washed in PBT (PBS with 0.1% Tween-20), and transferred into pre-hybridization buffer (prehyb) for 2-16 hours at 64.7°C. Following prehyb, embryos are incubated in gene-specific anti-sense riboprobe for 16-48 hours at 64.7°C. Probes were first denatured at 80°C for 10 minutes. Following probe incubation embryos were washed in a series of SSC buffer, first in combination with prehyb, then from 2X SSC to 0.1X SSC. Embryos were then moved to room temperature and transferred to PBS in a series of 0.1X SSC/PBS solutions. Embryos were then blocked for 2 hours with shaking, using Roche Western blocking reagent. They were then incubated overnight at 4°C in the primary antibody anti-digoxigenin conjugated to alkaline phosphatase. The remaining steps are at room temperature. The next day embryos were washed for three hours in PBT, with buffer changes

every hour. They were then incubated for 30 minutes in coloration buffer, followed by the staining reaction. Embryos were stained in a solution of NBT/BCIP, NaCl, and Tris-HCl Ph 9.5 in water. Embryos were stained until a distinct pattern can be discerned, without the entirety of the embryo, including the yolk, turning purple, which indicates over-staining. Embryos were then washed with PBT and in some cases stained with 1µg/ml DAPI in PBT for one hour. They were then washed in an ethanol series (80, 95, and 100%) for 10 minutes each, washed in PPO for 10 minutes, then a 1/1 solution of PPO/EPON overnight. Embryos were then transferred to fresh EPON for imaging.

Immunohistochemistry (IHC)

To determine if the expression of β-catenin increased, as expected with activation of the canonical Wnt pathway, custom peptide antibodies against both leech β-catenins were tested in LiCl-treated and control embryos. Two antibodies against each protein were produced in rabbits using the PolyExpress Silver Package by GenScript (Piscataway, NJ). The sequence of each protein, as well as the epitopes selected for antibody production are shown in Appendix A.

Embryos were incubated in blocking solution (PBS with 0.1% Triton-100 (PBTr) with 2.5mg/ml BSA) overnight, incubated in primary antibody diluted in blocking solution for 3-5 days (BC1 1:1000, BC2 1:2000), overnight wash in blocking solution, then overnight in secondary antibody (monoclonal mouse anti-rabbit conjugated to HRP) diluted 1:1000 in blocking solution. Embryos were then washed in PBS with 0.1% Tween-20 (PBT) for six hours. They were then stained with DAB for five minutes, washed with PBS, then incubated in DAPI for one hour. Embryos were once again washed with PBS, then treated with an ethanol series, 80%, 95%, 100% for five minutes each, then cleared in BBBA.

CRISPR/Cas9

I attempted to knockout the function of *fz1/2/7b* using CRISPR/Cas9 technology. I selected three distinct 20bp sequences at the 5' region of the gene that were upstream of a PAM motif (nucleotide sequence NGG, Carroll 2012, Doench *et al.* 2014). Selected sequences and their complements were ordered as oligonucleotides from Elim Biopharmaceuticals, Inc. (Hayward, CA). These were annealed, phosphorylated, then ligated into the pX330 plasmid (Addgene, Cambridge, MA), containing tracrRNA. This enabled amplification of a full single guide RNA (sgRNA) containing the gene-specific target sequence. I then used the MegaShortScript kit from Ambion to in vitro transcribe the sgRNA. The three sgRNAs were co-injected with Cas9 protein and Cas9 buffer into early blastomeres of *Helobdella* embryos.

To evaluate the efficacy of the sgRNAs, I genotyped several embryos to determine if an indel was produced in the expected location, using the CRISPR-STAT method (Carrington *et al.* 2015). I extracted genomic DNA from single embryos by digestion in a 50mM Tris-HCl (pH 8.0) buffer containing 1mM EDTA, 0.5% Tween-20, and 200µg/ml proteinase K at 55°C for four hours. I amplified the target region using primers flanking all sgRNAs. Genotyping was performed at the UCB DNA sequencing facility. Results were analyzed using PeakStudio v2.2 from the Fodor Lab at the University of North Carolina, Charlotte (Figure 2.13). Sequences for *fz1/2/7b* including sgRNAs and primers used are shown in Appendix B.

Results

LiCl treatment resulted in fewer progeny and a symmetric division of the N teloblast in a subset of embryos

In this experiment, embryos were treated at the onset of stage 5 with either 2mM LiCl, or 2mM NaCl as a control. Then, after division of the NOPQ proteloblasts, one of the daughter cells (N or OPQ) was injected with two lineage tracers: RDA and *h2b:gfp*. RDA is a cytoplasmic lineage tracer used to assess cell shape and identity, and *h2b:gfp* is a nuclear marker used to assess cell number. Embryos were then fixed 24 hours post injection, and cell number was counted. The results are summarized in Table 2.1 and Table 2.2. In control embryos, I found an average of 11.06 ± 1.73 progeny of the N lineage, and 16 ± 2.99 progeny of the OPQ lineage 24 hours after injection. There were significantly fewer cells produced when embryos were treated with LiCl. After 24 hours, LiCl-treated embryos had 7.26 ± 1.29 progeny of the N lineage (t-test, $p<0.0001$), and 8.28 ± 1.49 progeny of the OPQ lineage (t-test, $p<0.0001$). Morphologically, there were no discernible differences in cell size and shape between control and treated embryos. The most striking difference was in the number of teloblast-like cells (tlcs) that were produced. LiCl-treated embryos whose nominal N teloblasts were injected had two N tlcs 29.6% of the time. This is in contrast to NaCl-treated control embryos which produced only one tlc 100% of the time (Figures 2.3 and 2.5). LiCl-treated embryos whose nominal OPQ proteloblasts were injected have ≤2 tlcs 100% of the time (Figures 2.4 and 2.5). However, control embryos did appear to have ≤2 tlcs 21.4% of the time, instead of three, indicating that 24 hours post-injection may have been insufficient time to allow the normal cleavage divisions to reach completion in the OPQ lineage. Therefore, all the following experiments focused on the effect of LiCl in the N lineage.

Some aberrant teloblasts produce blast cells

Given that two N tlcs are produced in some embryos, it brings up the question of whether both produce blast cells. Indeed, comparisons between the number of cells produced in LiCl-treated embryos with one versus two N tlcs show that a significantly greater number of cells are produced in embryos where the teloblast has made a symmetric division (t-test, t-value=3.4909, $p=0.0018$). In order to address whether this was due to both tlcs producing blast cells, I allowed embryos to develop longer than 24 hours after injection with lineage tracer. However, in order to complete these experiments, it was necessary to remove the embryos from LiCl after 24 hours, because embryos maintained in 2mM LiCl were not viable beyond stage 8. Embryos maintained in LiCl for over 24 hours did produce bandlets of cells, however, they did not go on to form recognizable germinal bands. It is unknown whether germinal band formation failed due to aberrant cell divisions or migration or both.

Embryos were assessed using light and confocal microscopy. NaCl-treated control embryos had only one bandlet arise from the N teloblast injection 100% of the time (n=12). However, a subset of LiCl-treated embryos did have two bandlets after the nominal N teloblast was injected with RDA (Figure 2.7). Six of the 26 embryos had two N tlcs. Of these six, it was clear that three had distinct bands that were produced from those tlcs. The other three either had a single bandlet, or the blast cells were so distorted that it was impossible to tell if there were multiple bandlets.

Expression of N-lineage markers after treatment with LiCl

If LiCl treatment causes a duplication of the N-lineage, we expect to see an increase in the expression of N-lineage markers. To test this, I treated embryos at the onset of stage 5, washed with Htr after 24 hours, then fixed and processed embryos for ISH after control embryos reached mid-late stage 8, and performed ISH for two genes known to be expressed in the N lineage of the leech: *sfrp1/2/5c*, and *pax6a*. One thing to note is that LiCl treated embryos do generate germinal bands, but they are disorganized and do not follow the normal patterns of movement over the embryo (Figures 2.2 and 2.7).

Sfrp1/2/5c is normally expressed in a segmentally iterated pattern in the N and Q lineages during stage 8. This gene does appear to be upregulated after LiCl treatment, but as the bands are disorganized it is yet to be determined if this is a general upregulation, possibly in response to the hypothesized increase in *wnt* signaling itself, or if it is a result of lineage duplication (Figures 2.8 and 2.9). It is also possible that other lineages were transformed by LiCl treatment.

Pax6a is normally expressed continuously along the N-lineage, as well as in discrete spots along the adjacent O lineage – the patterns are easily distinguishable between lineages. Although control embryos show the expected pattern of expression, it is difficult to analyze the effect in LiCl-treated embryos. There is some staining, although it is no longer in a distinct pattern (Figure 2.10). The staining is diffuse throughout the germinal bands. It is therefore not possible to make any conclusions about the effect of LiCl on this N lineage marker.

Is β-catenin upregulated in these cells?

LiCl may have off target effects, so in order to confirm that it is truly upregulating the Wnt pathway, we must visualize the β-catenin protein to confirm its expression level has increased. To test for an increase in β-catenin expression, particularly nuclear expression, I treated with 2mM LiCl or 2mM NaCl at the onset of stage 5, fixed within 24 hours and performed immunostaining for either β-catenin1 or β-catenin2.

I tested the β-catenin1 antibody on embryos that were treated for 16 hours and for 20 hours. After 16 hours, there was not much discernible difference between control and treated embryos. In both cases, all embryos had staining in the micromere cap, and a subset appeared to have some level of staining in at least one N teloblast (30% for control embryos vs. 55% for treated embryos). There was a clearer difference in embryos treated for 20 hours. After this length of treatment, all embryos showed β-catenin1 expression specifically in the micromere cap. However, in control embryos this was only in a subset of micromeres. In nearly all treated embryos, there was clearly an expansion of this expression pattern within the micromere cap. More cells were stained, and the morphology of the cells appears to be affected as well. LiCl-treated embryos appear to have larger micromeres (Figure 2.11).

The β-catenin2 antibody was tested in embryos treated for 19 hours and 24 hours. Similar to β-catenin1, embryos at the earlier time point looked fairly similar when comparing the control to experimental conditions. In both cases the micromere cap is stained throughout. It is possible that the micromere cap in LiCl-treated embryos covers more of the surface of the embryo. However, in embryos treated for 24 hours, a discernible difference becomes apparent. By this point β-catenin2 staining is largely diminished from the micromere cap in most embryos. However, 50% of treated embryos do have staining in this region (Figure 2.12).

Does inhibition of the wnt signal show a converse phenotype?

I attempted to knock down the *wnt* signaling pathway by targeting the *fz1/2/7b* gene using CRISPR/Cas9 gene knock out technology. I hypothesized that this would eliminate all downstream *wnt* signaling in stages 6-7, as this is the only *fz* receptor expressed in the embryo at this stage. I tested five different gRNAs targeted to the 5' end of the gene. I used traditional sequencing techniques as well as CRISPR-STAT genotyping, however, I never found evidence that an indel was produced (Figure 2.13).

Discussion and Conclusion

The N lineage gives rise to much of the leech CNS. Wnt signaling has been shown to function in neural stem cell divisions in other systems as well, suggesting a conserved function for this pathway in this process. For example, one study tested the effects of LiCl in the mouse neural stem/progenitor cells after irradiation. They found that activating the Wnt pathway led to increased proliferation of these cells (Zanni *et al.* 2015). In addition, a study by Kalani *et al.* in 2008 found a population of Wnt-responsive cells in the developing mouse brain. When put into culture, it was found that this population was enriched for self-renewing stem cells. They suggest that Wnt signaling is necessary for expansion of this clonogenic population of cells, which generates new stem cells, neurons, astrocytes, and oligodendrocytes.

It is also notable that LiCl treatment resulted in a symmetric division of a cell that normally undergoes only highly asymmetric divisions, given that a directed Wnt signal has been found to contribute to asymmetric stem cell divisions. If there is a directed Wnt signal, then only specific daughter cells should upregulate β-catenin. This kind of asymmetry has been shown to polarize the plane of mitotic division in some cases (Habib *et al.* 2013, Goldstein *et al.* 2006). However, if β-catenin is upregulated all around the embryo, this could result in the division being symmetrized. Based on results of β-catenin antibody staining, I cannot conclude that this protein was upregulated in all cells of the embryo. However, given the large amount of yolk in early stage embryos, it is possible that the antibodies do not fully penetrate these cells, and it is therefore possible that some expression was not able to be visualized.

Overall what I have shown is that treatment with 2mM LiCl at the onset of stage 5 resulted in a symmetric division of the N teloblast approximately one third of the time. Taken together with the specific expression of the *wnt* receptor *fz1/2/7b* in the N teloblast, this suggests that *wnt* signaling is involved in some function of this cell: it may impart fate or control the division pattern. This division results in two N "teloblast-like cells," and in some cases, both give rise to blast cells. The expression of at least one lineage marker is upregulated after LiCl treatment, suggesting that the effect may be to duplicate the N lineage or transform other lineages.

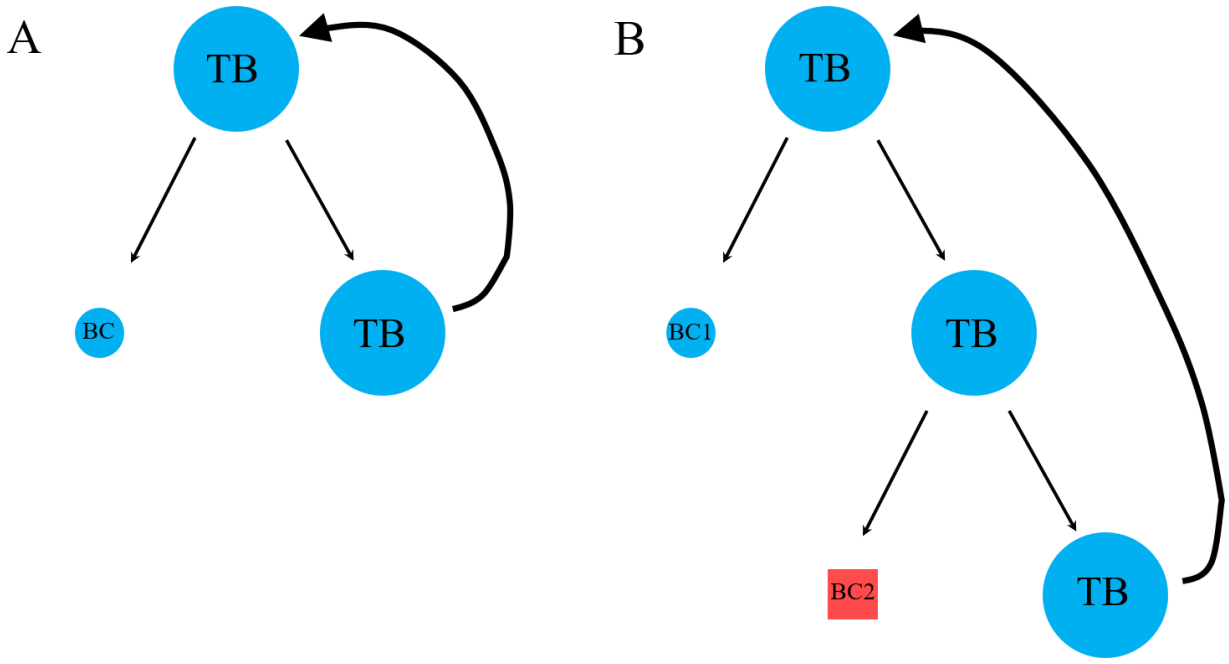


Figure 2.1. Parental vs Grandparental stem cell lineages. In the leech the M and O/P lineages are parental, whereas N and Q are grandparental stem cell lineages. Panel A demonstrates a parental stem cell lineage. Each division gives rise to a teloblast (TB) and blast cell (BC), and the same cell types are produced with each division. In the grandparental stem cell lineages, two distinct blast cell types are produced in exact alternation.

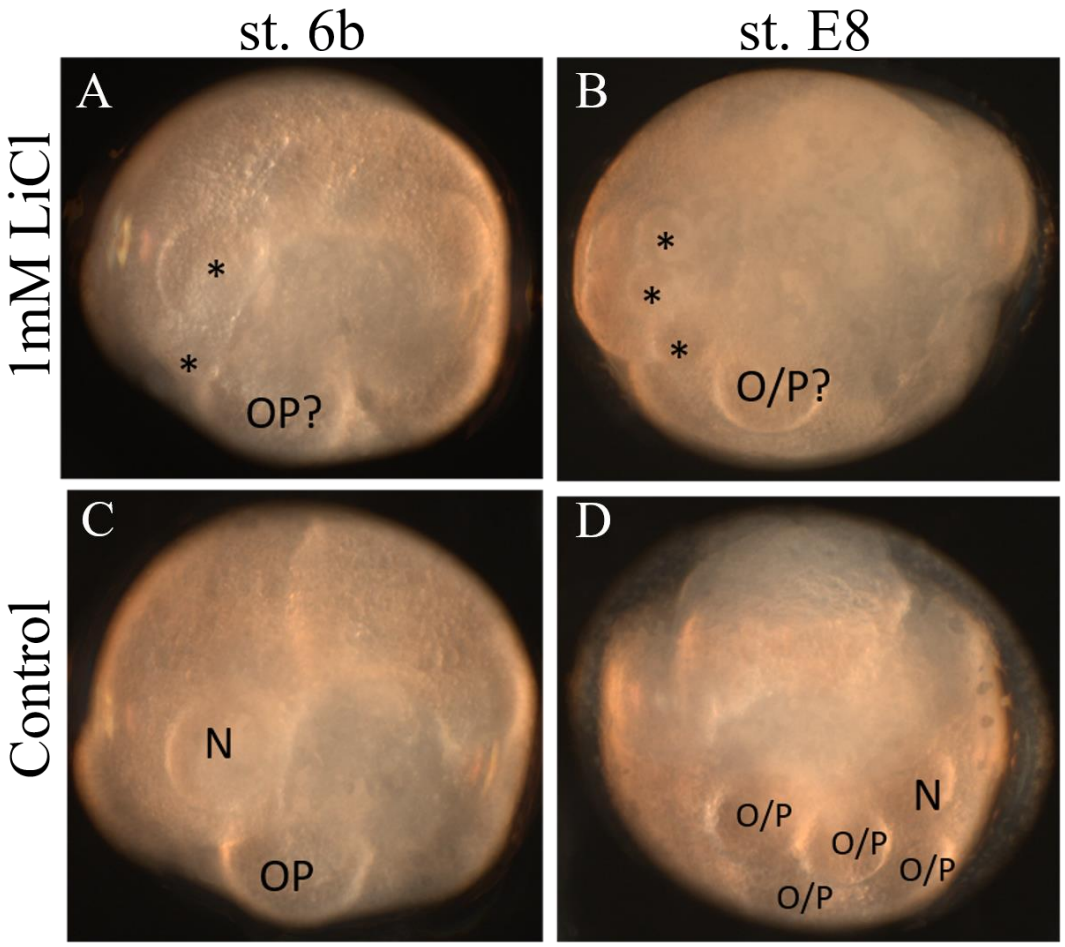


Figure 2.2. LiCl appears to affect teloblast division pattern. Initial trials of LiCl treatment of embryos in a bath solution appeared to affect a teloblast division. Embryos were treated at stage 5. Panel A shows what a representative treated embryo looks like when a sibling control embryo has reached stage 6b (Panel C). Panel B shows the morphology of a representative treated embryo when a sibling control embryo has reached mid stage 8 (Panel D). No germinal bands have formed on the treated embryo. Asterisks denote aberrant cells in a position where one large teloblast would normal exist. Panels A-C are animal views. Panel D is a dorsal view. Note that the trial experiment was performed in 1mM LiCl, whereas all future experiments were performed in 2mM LiCl.

Table 2.1. Cell number in N lineage after 24 hour LiCl treatment. Number of teloblast-like cells (tlc) and total cells after N lineage tracing. This was determined by counting the number of GFP-labeled nuclei. The error shown is standard deviation. Control embryos made significantly more cells.

Treatment	Embryos with 1tlc	Embryos with ≥2tlc	Avg #tlc ± 1 s.d.	Avg # cells ± 1 s.d.	n	t-value	p-value
2mM NaCl	16 100%	0 (0%)	1 ± 0	11.06 ±1.73	16	8.2226	<0.0001
2mM LiCl	19 (70.4%)	8 (29.6%)	1.33±0.55	7.26±1.29	27		

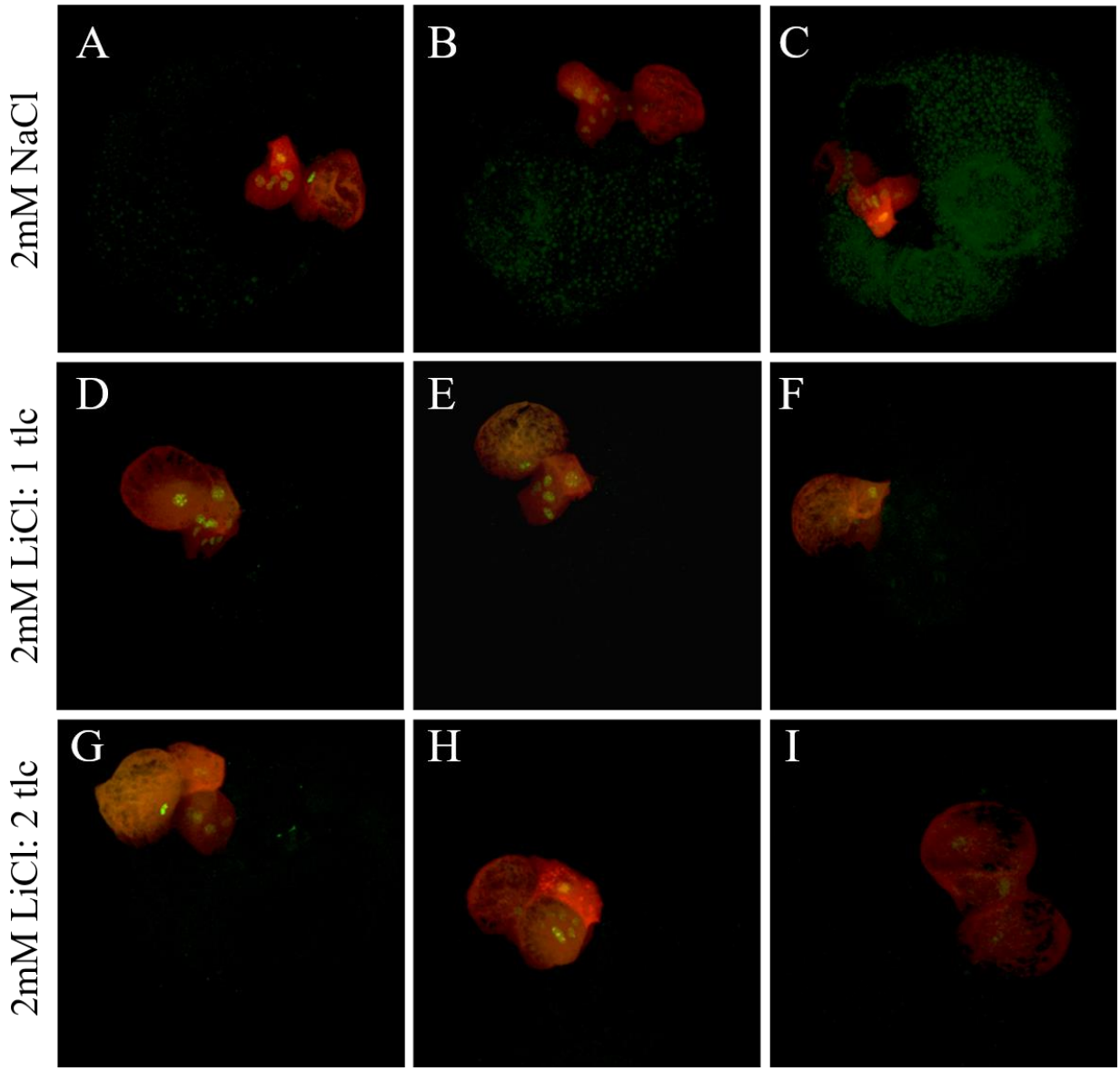


Figure 2.3. N lineage tracing of LiCl treated embryos reveals two teloblast-like cells in a subset of embryos. Panels A-C are three separate control embryos, which shows the stereotypical single large yolky teloblast, with a bandlet of small disc-shaped blast cells which normally loops around to join the anteriorly-progressing germinal band. Panels D-F show three separate LiCl-treated embryos that only produced one teloblast. Panels G-I show three separate LiCl treated embryos that have two cells that have the morphology of a teloblast, which is large, spherical, and yolky. Images are Z-stack projections of confocal slices.

Table 2.2. Cell number in OPQ lineage after 24 hour LiCl treatment. Number of teloblast-like cells (tlc) and total cells after OPQ lineage tracing. This was determined by counting the number of GFP-labeled nuclei. The error shown is standard deviation. Control embryos made significantly more cells.

Treatment	Embryos with 3 tlc	Embryos with ≤2 tlc	Avg # tlc ± 1 s.d.	Avg # cells ± 1 s.d.	n	t-value	p-value
2mM NaCl	11(78.6%)	3 (21.4%)	2.71 ± 0.61	16 ± 2.99	14	10.8051	<0.0001
2mM LiCl	0 (0%)	25 (100%)	1.96 ± 0.2	8.28 ± 1.49	25		

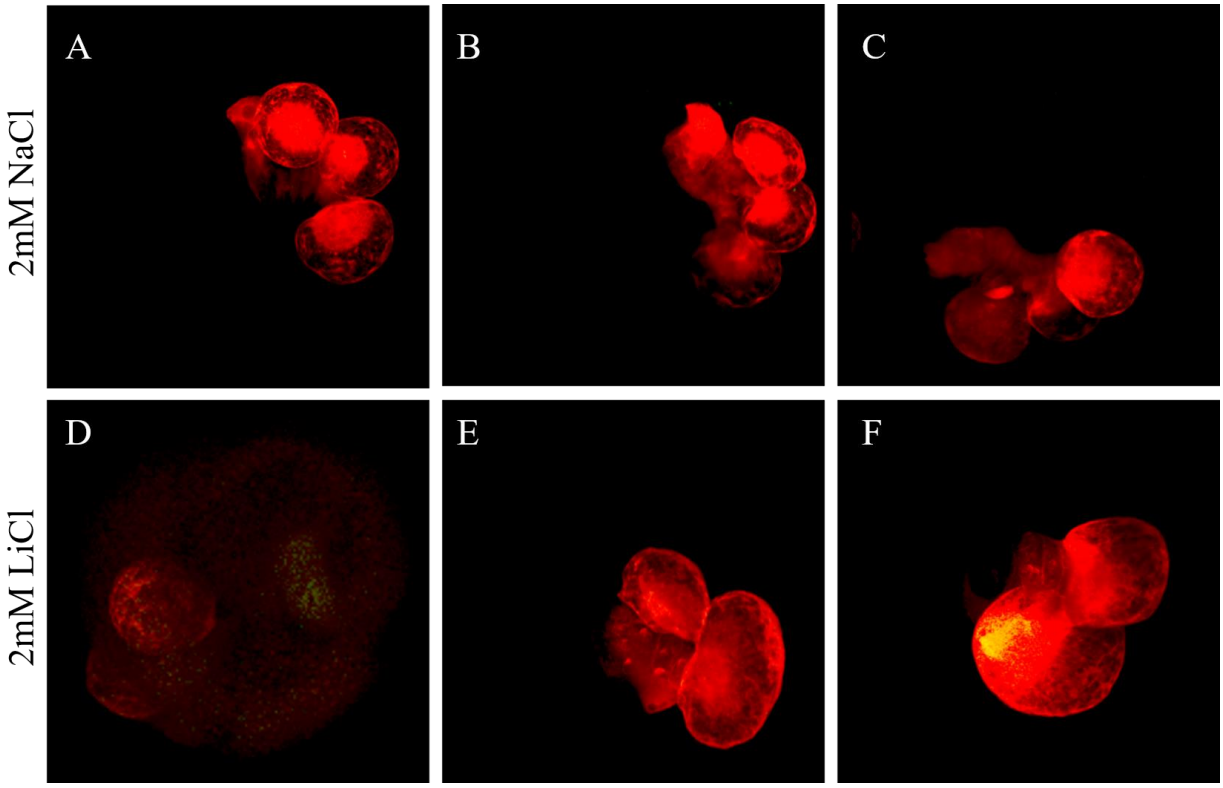


Figure 2.4. OPQ lineage tracing of LiCl treated embryos. Panels A-C are three separate control embryos, which shows the stereotypical O/P and Q teloblasts. There are three total, and they are approximately equal sized. Panels D-F show three separate LiCl-treated embryos, each of which only produced two teloblasts. In most cases, as in Panels E and F, one was much larger than the other. Images are Z-stack projections of confocal slices.

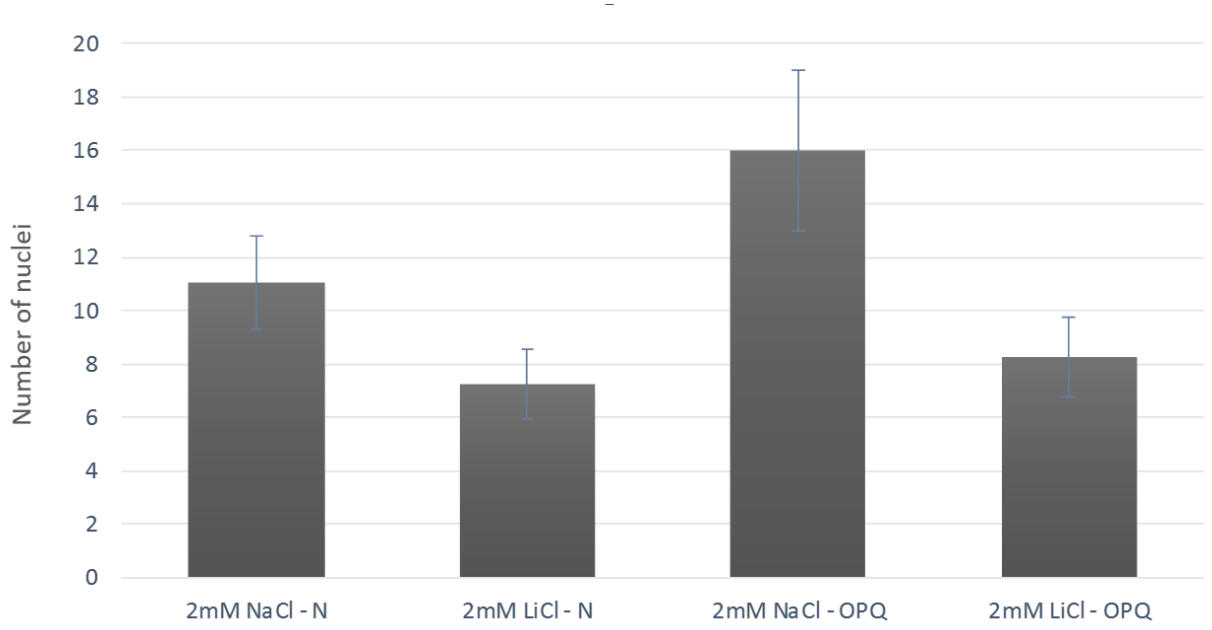


Figure 2.5. Average number of cells after 24 hour LiCl treatment. Bars represent the average number of nuclei counted for each treatment in the N and OPQ lineages. Error bars show one standard deviation.

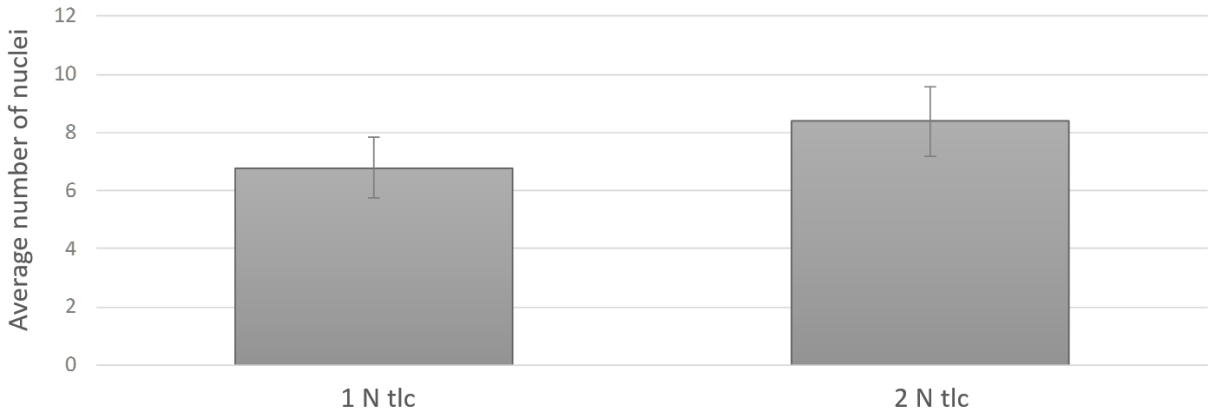


Figure 2.6. LiCl-treated embryos with 2 N tlcs make significantly more cells. After 24 hours of LiCl treatment, embryos with N-lineage tracing had either one or two N tlcs, shown on the X-axis. On the Y-axis is the average number of nuclei. Error bars show one standard deviation. Embryos with 2 N tlcs have a significantly greater number of cells (t-test, t-value=3.4909, p=0.0018).

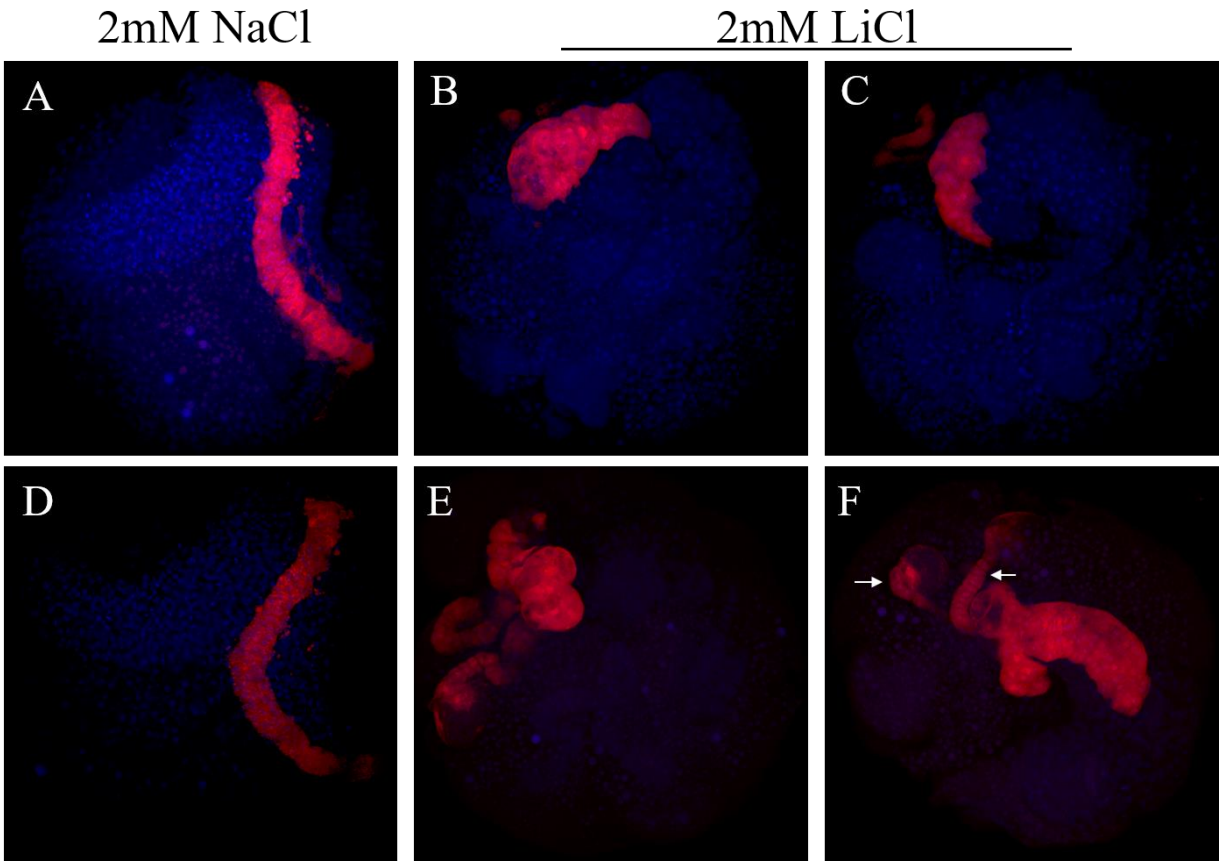


Figure 2.7. A subset of LiCl treated embryos with two N tlcs have two N bandlets. Panels A and D show ventral views of control embryos. In these embryos, there is clearly one bandlet labeled an it is in the stereotypic position of the N lineage. Panels B, C, E, and F show different examples of embryos treated with LiCl for 24 hours, then washed and allowed to develop for five days. Panels B, C, and E illustrate typical defects which make it difficult to determine the number of bandlets present. Panel F clearly shows two N tlcs, each producing a bandlet of blast cells. Images are Z stacks of confocal slices. Nuclei are stained with DAPI. Red fluorescence is the lineage tracer RDA.

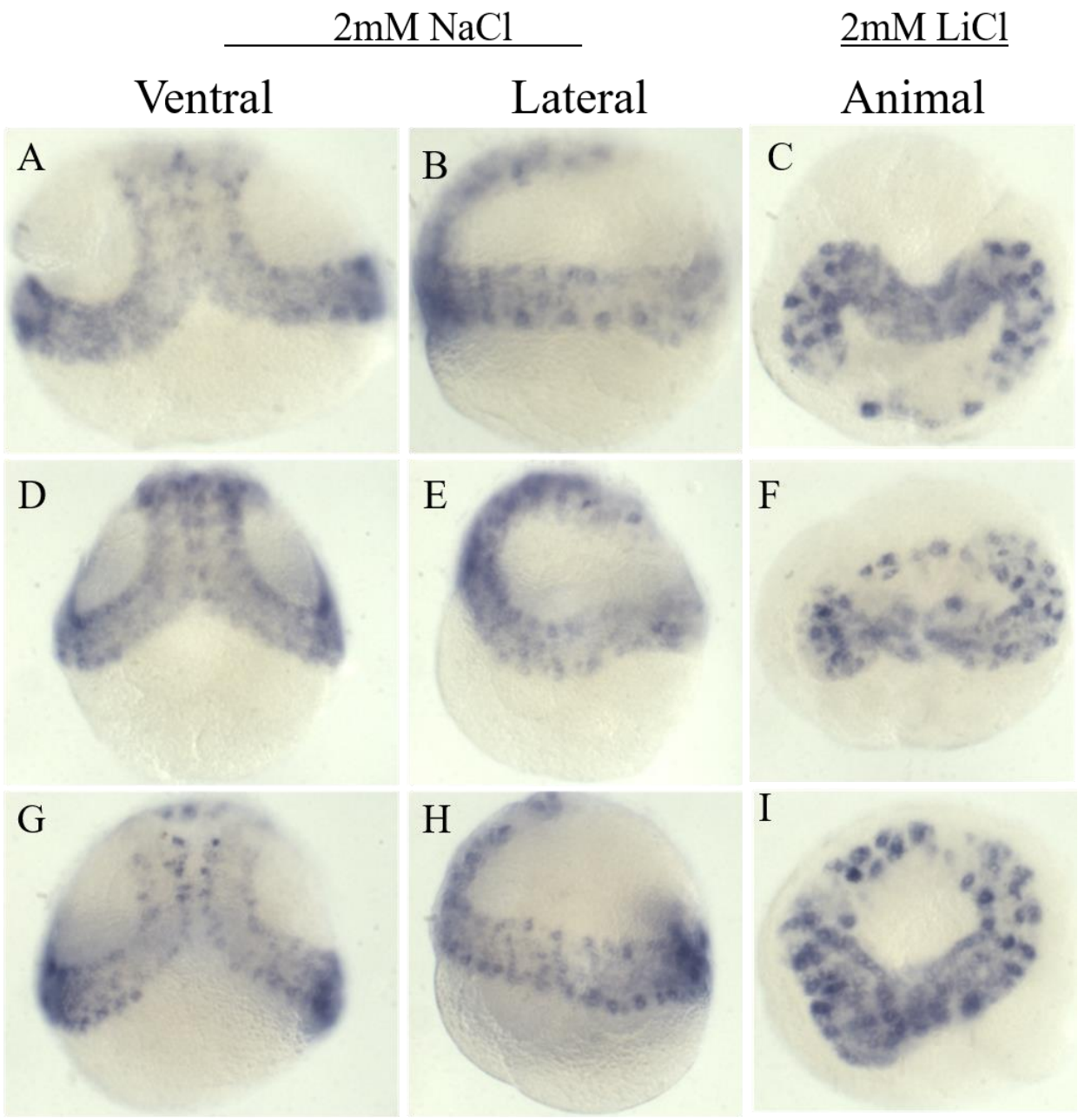


Figure 2.8. LiCl treatment causes an increase in *sfrp1/2/5c expression*. Panels A, D, and G show ventral views of three different control embryos, and Panels B, E, and H show lateral views of the corresponding embryos. *Sfrp1/2/5c* is normally expressed in the N and Q lineages. Panels C, F, and I show animal views of three different sibling embryos treated with LiCl. Expression no longer appears to be in two distinct bandlets, and appears to have increased. These embryos were treated with LiCl for 24 hours, washed, then fixed after 142 hours for ISH against *sfrp1/2/5c*.

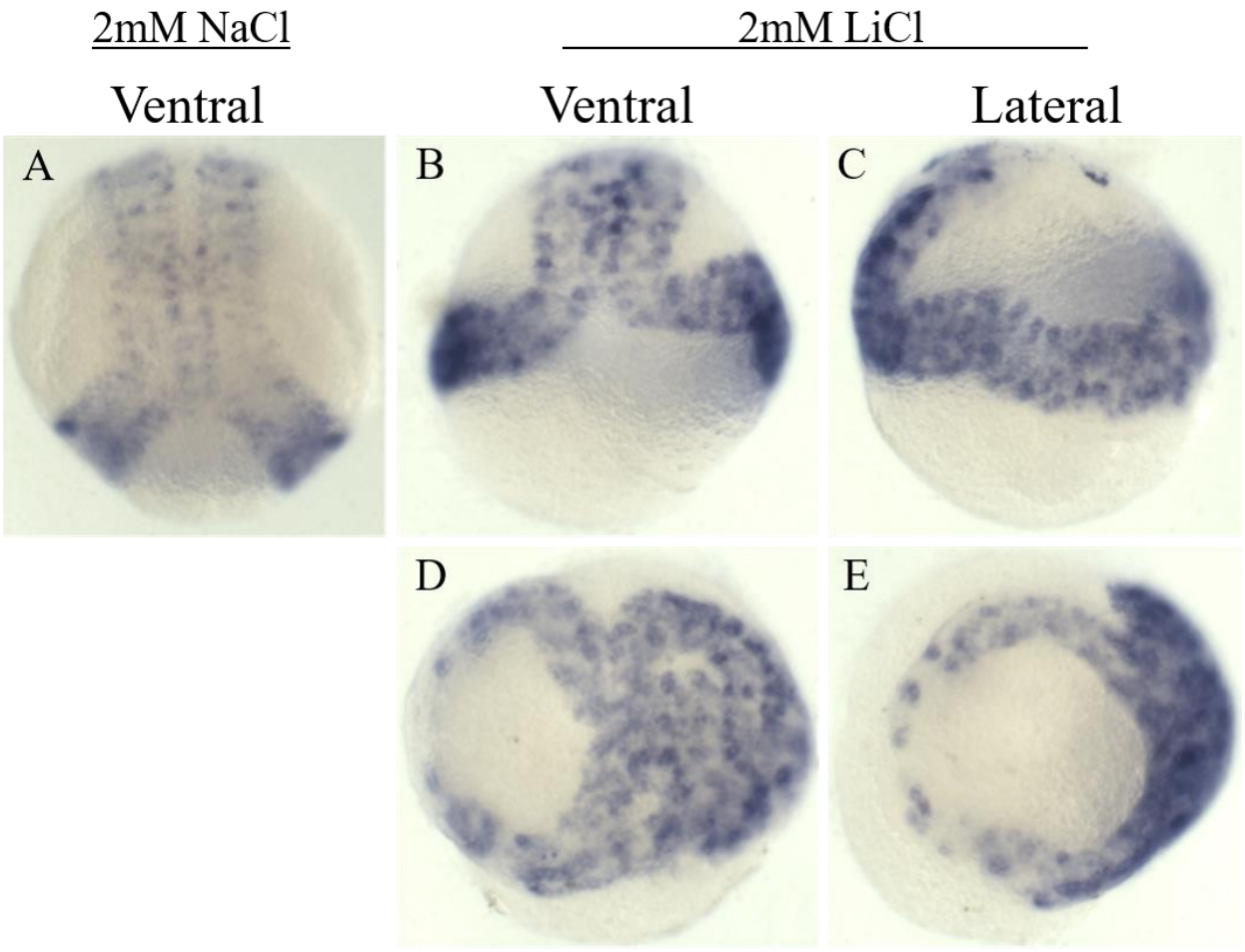


Figure 2.9. LiCl treatment causes an increase in *sfrp1/2/5c expression*. Panels A shows a ventral view of a control embryos. Panels B and D show ventral views of LiCl-treated embryos, and Panels C and E show the lateral views of the corresponding embryos. These embryos were treated with LiCl for 24 hours, washed, then fixed after 167 hours for ISH against *sfrp1/2/5c*.

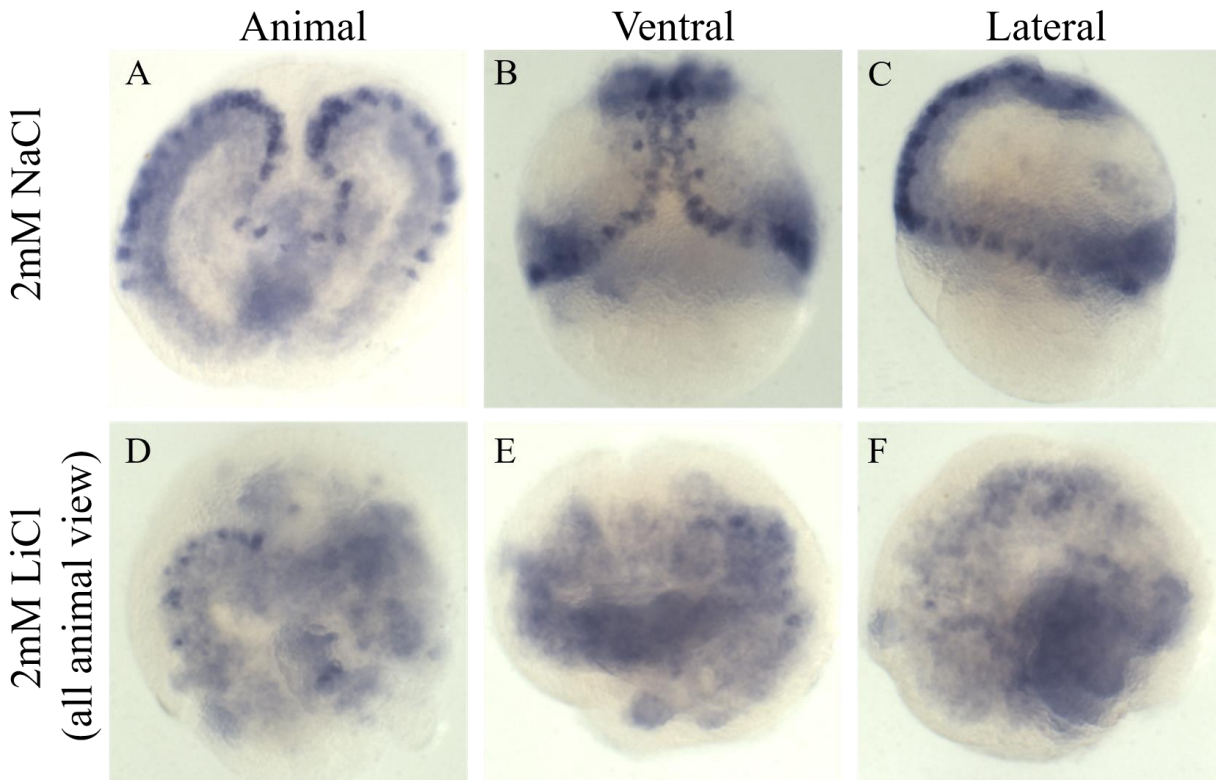


Figure 2.10. *Pax6a* expression is diminished in LiCl treated embryos. Panels A-C show different views of the normal mid stage 8 expression pattern in control embryos. *Pax6a* is normally expressed in the N and O lineages. In the N lineage, it is continuously expressed throughout the bandlet. In the O lineage, it is expressed in distinct dots. Panels D-F show three different LiCl-treated embryos. The staining is diffuse throughout the germinal bands in no apparent pattern. Embryos were treated with LiCl for 24 hours, washed, then fixed between 120 and 160 hours for ISH against *pax6a*.

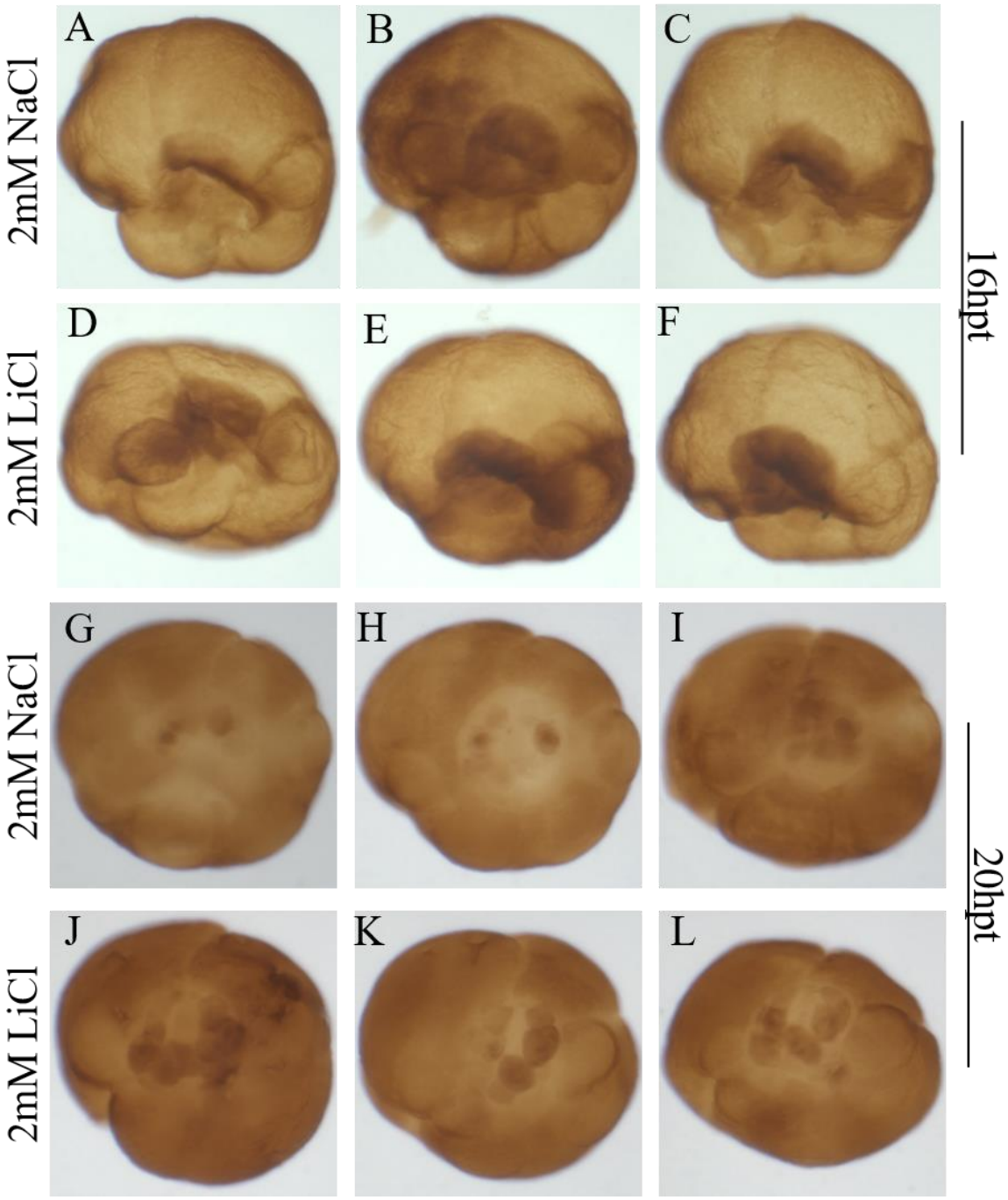


Figure 2.11. β-catenin1 expression increases in the micromere cap after LiCl treatment. Panels A-C and G-I are control embryos, and Panels D-F and J-L are treated embryos. Embryos shown in Panels A-F were fixed 16 hours post treatment (hpt), and embryos in Panels G-L were treated for 20 hours. After 16 hours, the staining appears to be in the same area of the micromere cap, although it appears darker in treated embryos. After 20 hours staining in control embryos is in a subset of cells in the micromere cap. Staining is in an expanded number of cells in treated

embryos. These cells also appear to be larger than in control embryos. Embryos were treated with primary antibody against leech β-catenin1, secondary antibody conjugated to HRP, then stained with DAB for five minutes.

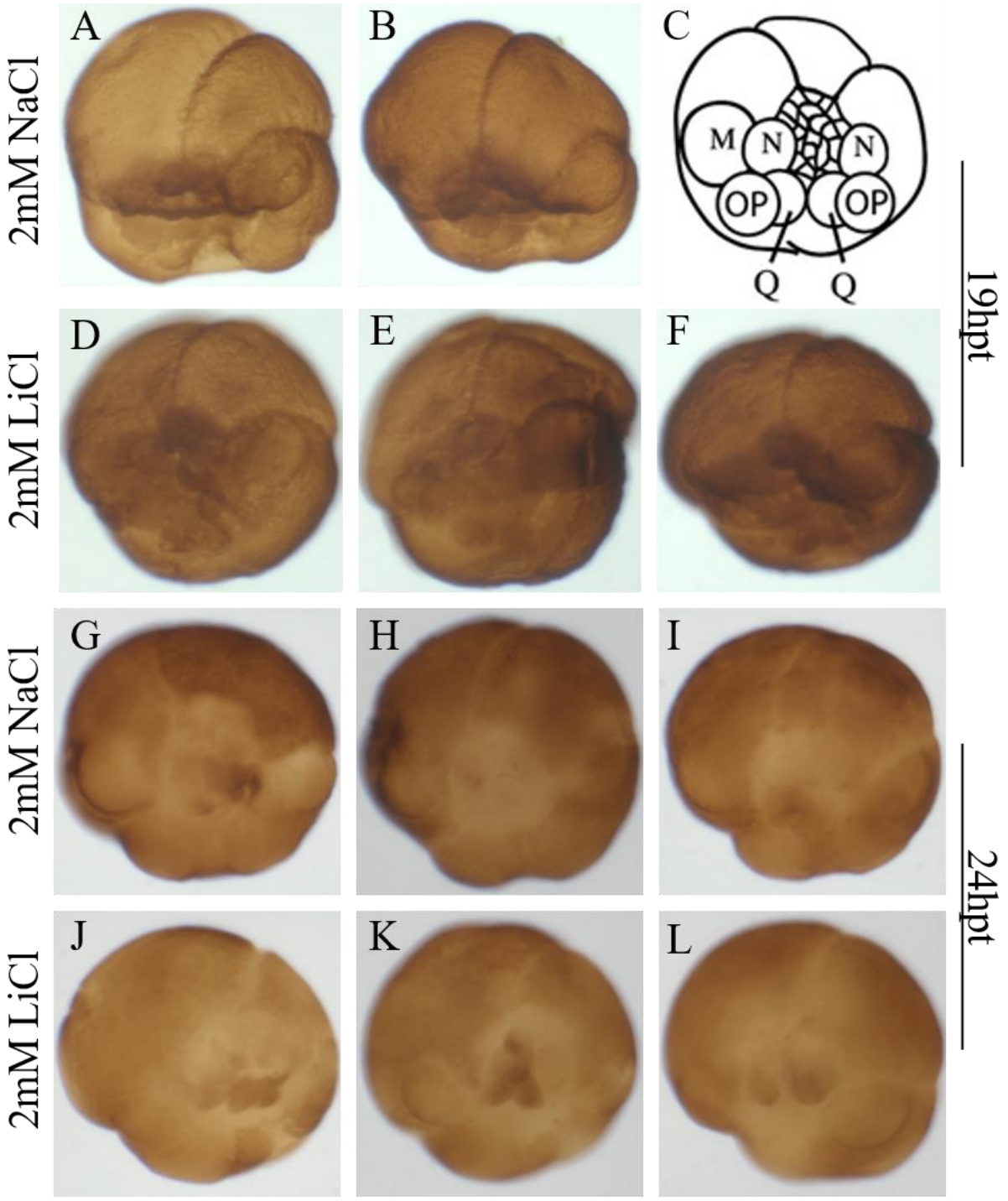
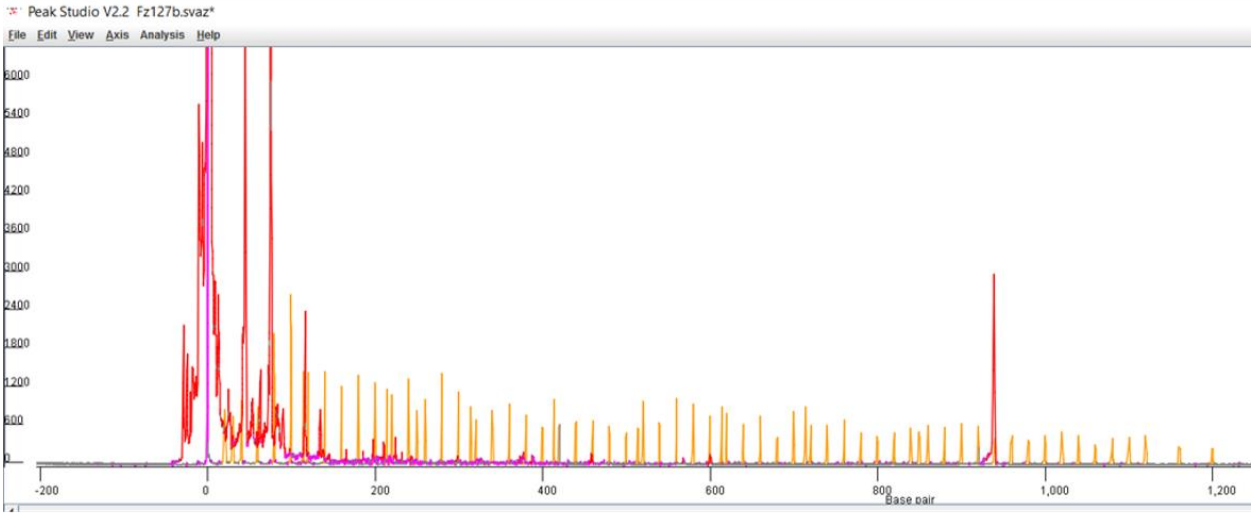
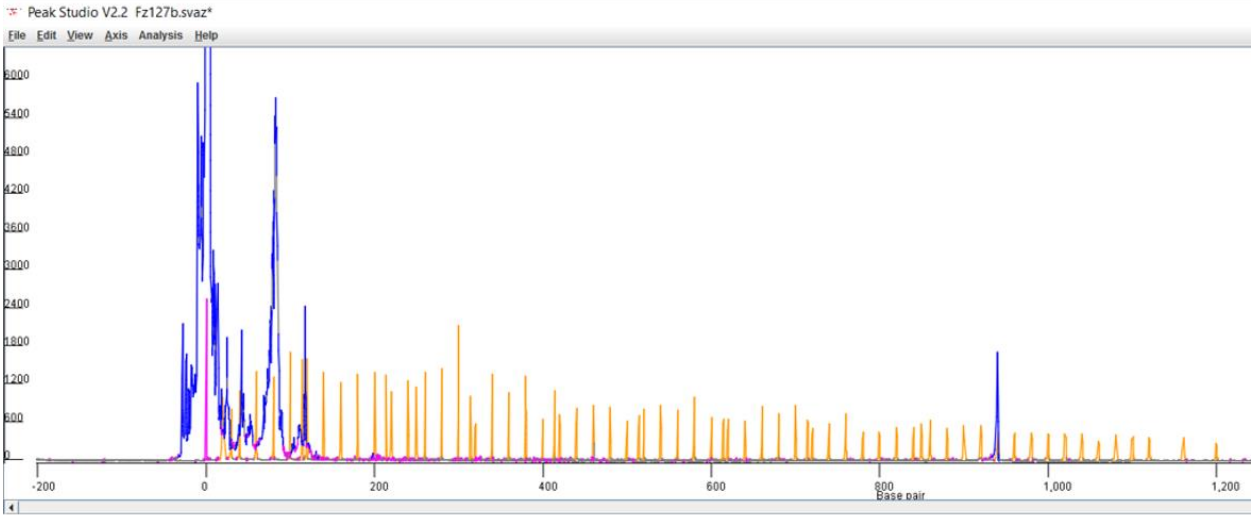


Figure 2.12. β-catenin2 expression increases in the micromere cap after LiCl treatment. Panels A-B and G-I are control embryos, and Panels D-F and J-L are treated embryos. Embryos shown in Panels A-F were fixed 19 hours post treatment (hpt), and embryos in Panels G-L were treated for 24 hours. After 19 hours, the staining appeared in the micromere cap of control embryos. Staining appeared to be in an expanded area in treated embryos. After 24 hours

staining was largely diminished in control embryos, but persisted in a subset of the micromere cap in LiCl-treated embryos. Embryos were treated with primary antibody against leech β-catenin2, secondary antibody conjugated to HRP, then stained with DAB for five minutes. Panel C depicts a cartoon of an embryo at a similar stage as the embryos shown.



CRISPRstat control embryo



CRISPR injected embryo

Figure 2.13. CRISPR-STAT genotyping results. A cocktail of guide RNAs against the 5' end of *fz1/2/7b* were injected into blastomeres of early stage embryos. The expected product size for the control region was 920bp. The X-axis represents the product length. There is a single peak at 920bp for both control and gRNA injected embryos, indicating that they did not produce indels in the target gene. If indels had been produced, peaks of different sizes would be apparent in the lower graph. The peaks between 0 and 200bp are noise.

Chapter 3. Divergent expression of duplicate *wnt* genes in the leech segmental ectoderm

Introduction

It has long been recognized that gene duplication plays an important role in evolution. Gene or whole genome duplication is a source for new genes with novel functions. Gene duplication can occur via unequal crossing over, transposition, or duplication of whole chromosomes (Magadum *et al.* 2013). After a duplication event, there are three possible outcomes. One copy of the gene may undergo nonfunctionalization, in which a buildup of degenerative mutations leads to the inactivation of that copy of the gene. They can also undergo neofunctionalization, in which one or both copies of the genes assume new functions (by virtue of changes in the expression domain or sequence of the encoded protein), or subfunctionalization, where the two duplicated copies subdivide the function or domain of expression of the original gene (Lynch and Conery 2000). In reality, subfunctionalization and neofunctionalization are overlapping categories.

Helobdella austinensis has 13 Wnt genes in its genome, however these genes only represent nine of the ancestral 13 subfamilies that are inferred to have been present in the ancestral bilaterian (Cho *et al.* 2010). It appears that there were independent duplications in the lineage leading up to leeches in three of these genes. Specifically, the leech has three copies of *wnt11*, and two copies each of *wnt5* and *wnt16*. Two other lophotrochozoan species whose whole genomes have been sequenced, the polychaete *Capitella teleta* and the mollusk *Lottia gigantea*, each only have one copy of these genes. Thus, it seems that the duplications seen in leech are of relatively recent origin.

This work by Sung-Jin Cho showed that since the sedentarian ancestor of *Capitella teleta* and *Helobdella* there was a duplication of *wnt16* in the lineage leading to the leech. Using in situ hybridization (ISH) combined with lineage tracing, he showed that *wnt16a* and *wnt16b* have non-overlapping expression patterns in stage 10 embryos in transverse, segmentally repeating bands of adjacent cells within the ventral nerve cord (which are progeny of the N lineage) and also in non-overlapping sets of cells within the proboscis at stage 10. Comparisons with the expression pattern for the single *wnt16* orthologue in *Capitella* suggest the duplicate may have undergone neofunctionalization or subfunctionalization. Cho has also found that *wnt16a* is expressed in the N lineage in a segmentally iterated subset of N lineage cells starting during early stage 8. This leaves open the question of whether *wnt16b* is expressed in adjacent N lineage cells at this stage as well.

As described in the Introduction, there are five stem cell lineages within each of the leech germinal bands, four ectodermal and one mesodermal. The N lineage gives rise to ventral ectoderm, the two O/P lineages give rise to lateral ectoderm, and the Q lineage gives rise to dorsal ectoderm. Beyond the BMP signaling that breaks the initial equipotency of the O/P lineages on each side (Kuo and Weisblat 2011), it is not well understood what signaling events contribute to the fate specification of each lineage, or what contributes to the distinct blast cell division pattern of each lineage. Another interesting difference between these lineages is that the M and O/P lineages only produce one blast cell per segment, whereas the N and Q lineages produce two blast cells per segment, each with a distinct timing and pattern of division (Zackson 1984, Figure 3.1). It is unclear what differentiates the alternate cells in these two lineages. The expression of a *wnt16a* in alternate cells of the N lineage suggests that Wnt signaling may play a

role in this process (Figure 3.3). In general, lineage and cell-specific expression in the germinal band is of interest because it may indicate a role for Wnt signaling in fate specification and/or cell division patterns during the segmentation process.

Evidence suggests that Wnt signaling is involved in cell division. By activating cyclin D1 it plays a role in cell cycle progression from G1 to S phase. Components of the Wnt pathway are also involved in regulation of the mitotic spindle. It has been shown that interfering with this signaling pathway can interfere with microtubule assembly and affect the mitotic spindle (Niehrs and Acebron 2012). As mentioned in Chapter 2, it has also been shown in both *C. elegans* and mammalian ESCs that a directed Wnt signal orients the plane and symmetry of division (Goldstein *et al.* 2006 and Habib *et al.* 2013).

Wnt signaling has also been found to function in cell fate specification in many systems. One example is neuroblast fate specification in *Drosophila melanogaster*. These cells arise from the ventral ectoderm. The *Drosophila wnt* homolog *wingless* (*wg*) is expressed in transverse stripes along the A-P axis. This correlates with specific clusters of neuroblasts, and is necessary to determine the fate of these cells (Skeath 1999). It has long been known that Wnt signaling is necessary for fate specification of early blastomeres in *Caenorhabditis elegans*. In this case, the *wnt* expressing cell, called P2, is at the posterior of the embryo and it presents a polarized signal to the endomesodermal precursor EMS. This polarized signal downregulates nuclear expression of a downstream gene in the posterior daughter, which specifies it as endoderm (Thorpe *et al.* 1997).

Current evidence suggests that Wnt signaling may also be involved in cell fate specification and/or cell division patterns in the leech embryo, given the expression pattern of *wnt16a* described above. Given the stage 10 expression pattern described by Cho *et al.*, it seems that the duplicate gene may have diverged in function after segments have formed. Here I investigate the expression pattern of *wnt16b* at an earlier stage, before lineage-specific division patterns and segmentation take place. I find multiple isoforms of *wnt16b*, and show that they are not only expressed in different cells than *wnt16a*, but different lineages.

Materials and Methods

Gene cloning

Primers were designed based on genomic DNA sequence of sibling species *Helobdella robusta* (http://genome.jgi.doe.gov/Helro1/Helro1.home.html). The primers used are listed in Table 3.1. Additional 5' and 3' sequence was obtained using the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA). PCR products were extracted using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and ligated into pGEM-T Easy Vector (Promega, Fitchburg, WI) according to manufacturer's instructions. These clones were introduced into competent *E. coli* cells by heat shock, plated onto 1% agar plates with 100µg/ml ampicillin, and grown overnight at 37°C. Colonies were selected and grown overnight in LB with 100µg/ml ampicillin. Plasmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA) following manufacturer's instructions and sequenced at the UCB DNA Sequencing Facility. Sequences were aligned using NCBI BLAST.

Lineage tracing combined with in situ hybridization

Lineage tracing and ISH were performed as described in chapter 2. In this case embryos with lineage tracing were processed for ISH, then imaged on a Leica fluorescent dissecting microscope.

Results

Genomic architecture of wnt16a vs wnt16b

Annotation of the leech genome predicts that *wnt16a* (protein ID 121846) and *wnt16b* (protein ID 79030) have a very similar architecture, each with 6 exons and 5 introns, with the fourth introns being the largest (Figure 3.2), although *wnt16a* is predicted to be approximately twice the length of *wnt16b* (6829bp vs. 3476bp). However, the genome model for *wnt16b* is incomplete, as it lacks a start codon. The full genome sequences were unable to be cloned to confirm these models. Partial cDNA sequences were reported in Cho *et al.* 2010. I expanded on these sequences using RACE PCR, although I was still unable to obtain full length transcript sequences.

wnt16b exists as multiple splice variants

By conducting 3' RACE PCR, I found that *wnt16b* exists in at least two isoforms (Figure 3.2). I refer to these as *wnt16b-i* and *wnt16b-ii*. All the experiments described below were performed using the *wnt16b-i* sequence. I was unable to clone a sequence that matched the one found by Cho, although *wnt16b-ii* does share overlapping sequence. To investigate the expression pattern of the new isoform, I performed ISH combined with lineage tracing. The cytoplasmic lineage tracer RDA was injected into individual teloblasts.

Comparison of wnt16a and wnt16b-i throughout stage 8

To initially survey the expression pattern of these duplicate genes, I performed ISH for *wnt16a* and *wnt16b-i* on embryos at timepoints throughout stage 8 (Figure 3.3). As previously described, *wnt16a* does appear to be expressed in alternating cells or groups of cells of the N lineage. I can deduce that the expression occurs in this lineage because the staining is at the ventralmost edge of the germinal bands, and when these coalesce to form the germinal plate, the dots of expression are at the midline. However, this did not appear to be the case for *wnt16b-i*. This gene was detected in a segmentally iterated pattern, however the staining appeared to be more towards the center of the germinal bands, in the region of the lateral ectodermal blast cells. I therefore carried out lineage experiments in which various lineages were marked by the prior injection of the standard cell lineage tracer RDA to determine which lineage(s) expresses *wnt16b-i*.

Expression of wnt16b-i in stage 8-10 embryos

At the earliest point of stage 8, *wnt16b-i* was first seen to be expressed in cells at the anterior of the germinal band that do not overlap with N lineage tracer (Figure 3.4). These may be the four op blast cells that the OP proteloblast gives rise to before dividing symmetrically to form the two O/P teloblasts. At a timepoint later in early stage 8, but before germinal band coalescence, the expression expanded to more cells, appearing to be lineage-specific, and in a segmentally iterated pattern (Figure 3.4). The ISH stain overlapped with O lineage tracer. During mid stage 8, the ISH staining continued to overlap with O lineage tracer, and the pattern was segmentally iterated (Figure 3.5). At this point it is interesting to note that while the expression pattern extends all the way to the anterior end of the germinal bands, no *wnt16b-i* expression was detected in the most posterior cells (which are the youngest).

I began to see a change in the expression pattern in late stage 8 embryos. In these embryos, expression persisted in the O lineage in roughly the posterior half of the germinal band. However, in the anterior half of the germinal band there appeared to be a shift in the expression pattern to the P lineage (Figure 3.6). There were often a few transitional segments where the expression appears to be in both lineages. Also, by late stage 8, the expression extended completely from the most anterior to most posterior cells of these lineages. By stage 9, most of the expression pattern overlapped with P lineage tracer (Figure 3.7). However, a small number of *wnt16b-i*-expressing cells persisted in the most posterior cells of the O lineage. In stage 10 embryos, the expression of *wn16b-i* was clearly in the lateral ectoderm, although no lineage tracing was performed to this stage (Figure 3.8). This is contrary to the expression pattern published by Cho *et al.*, who found that the *wnt16b* variant he isolated was expressed in the N lineage at stage 10.

Are the two wnt16b isoforms expressed in distinct lineages?

Although I did not perform lineage tracing experiments, *wnt16b-ii* appeared to show the same expression pattern as *wnt16b-i*. It was detected in lateral ectoderm throughout stage 8-10 (not pictured).

Expression of wnt16b-i in primary o blast cells reveals intermediates between the O and P blast cell fates

It has been shown that a translation-blocking antisense morpholino oligomer (ASMO) against *bmp5/8* can give rise to an ectopic O lineage in the wildtype P position, when injected into the Q lineage. I therefore hypothesized that this ectopic O lineage would express *wnt16b-i* in early stage 8 embryos. To test this prediction, I did the ASMO injection combined with ISH. ASMO-injected embryos did undergo the P to O transformation, as confirmed by ISH showing a loss of *six1/2a* expression in the nominal p bandlet. However, the ectopic O lineage did not express *wnt16b*. Moreover, ASMO injected embryos allowed to develop to late stage 8 showed normal *wnt16b* expression; this is after expression transitions to the P lineage. However, it has been previously suggested that the ectopic O lineage can transition back to P fate after a certain amount of time, perhaps due to ASMO lifetime (Kuo and Weisblat 2011). Indeed, by mid stage 8 ASMO-injected embryos regained normal expression of *six1/2a* (Figure 3.9).

Conclusion and Discussion

Duplicated wnt16 genes show divergent expression patterns in the leech

Leech *wnt16a* and the two *wnt16b* isoforms I cloned have divergent expression patterns between stages 8-10. *Wnt16b-i* is expressed in different lineages throughout this timeframe, and the switch may be associated with blast cell clonal age. It appears that recently born primary blast cells do not express *wnt16b-i*, primary o blast cells of a specific clonal age begin to express *wnt16b-i*, which persists through stage 8. In late stage 8 or early stage 9, possibly at a specific clonal age, the expression shifts from the O lineage to the P lineage. *Wnt16b-i* continues to be expressed in the lateral ectoderm in stage 10. Previous work has shown that a second *wnt16b* isoform is expressed in the N lineage in stage 10 embryos, although I was unable to confirm this result. I was also unable to show that *wnt16b-i* is expressed in an ectopic O lineage. This may be due to an incomplete transformation of the lineage by the ASMO.

Wnt16b-i shifts from the O to P lineage in late stage 8

It is particularly interesting that *wnt16b-i* appears to switch lineages at a fairly distinct point. During mid stage 8, when expression is in the O lineage, the staining pattern appears in more tightly clustered dots of cells. However, in very late stage 8 and stage 9, the staining pattern in the anterior of the embryo overlaps with P lineage tracer. The embryo develops in an anterior to posterior progression, so anterior cells are of an older clonal age than the posterior cells. The staining pattern in the anterior half of the embryo at this stage looks more like a transverse stripe. This may indicate that the shift is associated with a specific morphological event in the embryo, such as the formation of a segment after cells of the different lineages align. Because the N and Q lineages make twice as many blast cells per segment, they are producing blast cells for twice as long as the O/P lineages, and therefore these bandlets must slide past the lateral bandlets for the segments to come into register (Weisblat and Shankland 1985).

How are the divergent expression patterns regulated?

Automated JGI annotation suggests that intron 4 of *wnt16a* shows conservation with corresponding introns of *Capitella* and *Lottia.* Upon further review, it became clear that the region shown for *Lottia* was not actually *wnt16*. However, in the *Capitella* genome browser, the track given to represent the *Lottia* sequence does indeed correspond to *wnt16*, and there are conserved regions in the same intron when comparing those two species. The annotation shows a subset of this region to be conserved in *Helobdella wnt16a*. As the duplication of *wnt16* only exists in *Helobdella*, it was unclear if this was true for *wnt16b* as well. However, a BLAST search of the *Helobdella* sequence shows no significant similarity to any region within the corresponding *wnt16b* intron. This brings up the question of whether these conserved sites represent regulatory regions. There have been several cases where an enhancer was shown to exist within an intron (Rose 2008). If this region is found to be regulatory, it is possible that the differences in the sequence of this region explain the divergent expression pattern.

Significance

These findings are significant because they suggest that the recent independent duplication of *wnt16* in the leech appears to have resulted in subfunctionalization or neofunctionalization. These two genes may contribute to the segmental pattern in the leech by specifying fate and/or division patterns of primary blast cells.

Table 3.1. List of primers used to amplify *wnt16* sequences

Name	Primer Sequence
wnt16a-5'RACE	5' – CTGCTGCGGGTGATTGCGAAGATTCCAGCG – 3'
wnt16b-5'RACE	5' – CGCAAGAGCATTCTAGAAGGAGGCCAGCGC – 3'
wnt16b-5'RACE-nested	5' – GAGCGAGTGAGGCAGACGTGATGGCGTGG – 3'
wnt16b-3'RACE	5' – CCACGCCATCACGTCTGCCTCACTCGCTC – 3'

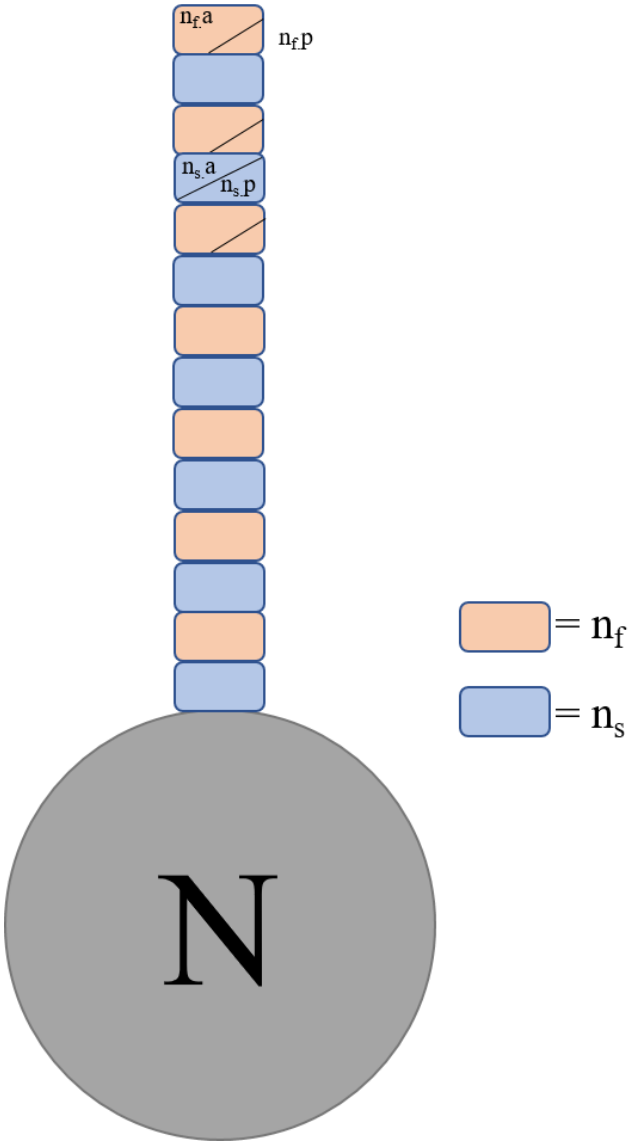


Figure 3.1. Distinct division pattern of f and s blast cells. The N and Q lineages are grandparental stem cell lineages. This means the teloblast produces two types of blast cells in exact alternation. In the case of the N lineage, these are referred to as the n_f and n_s blast cells. The f and s blast cells have unique division patterns, as illustrated.

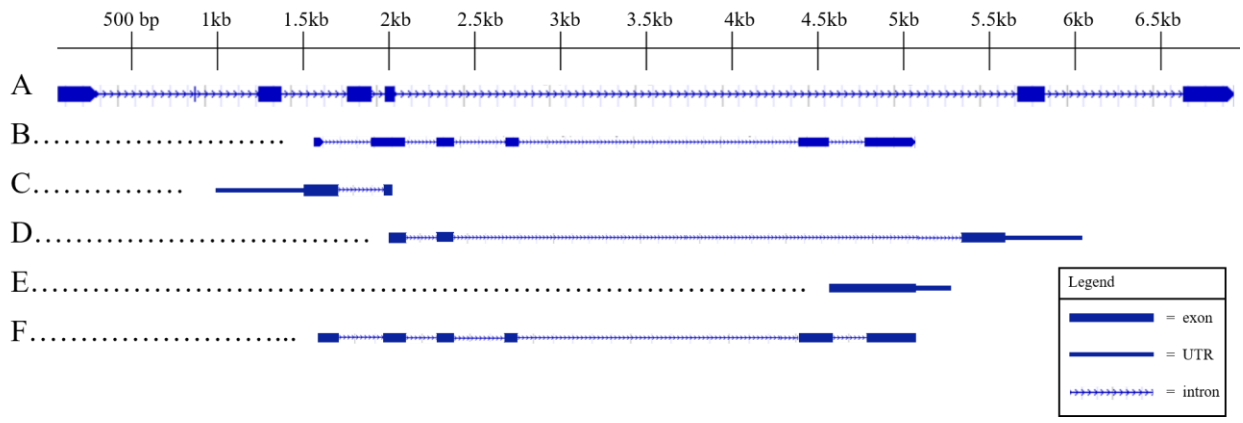


Figure 3.2. Genomic architecture of *wnt16a* and *wnt16b*. This shows the predicted architecture of the duplicated *Helobdella wnt16* genes based on the genome models as well as cloned cDNA sequences. Panel A is the predicted genome model (protein ID 121846) for *wnt16a*. The predicted length is 6829bp. Panel B is the predicted genome model (protein ID 79030) for *wnt16b*, for which the predicted length is 3476bp. Panels C-F show the predicted genomic structure for the *wnt16b* cDNA sequences I cloned. I only verified the exonic sequences in these cases (depicted by the thick blue bars). Intronic sequences (thin arrowed lines) were extrapolated by aligning my cDNA products to the genome models. Cloned regions that correspond to the genome model are aligned in the models. Panel C is the 5'RACE sequence I isolated for *wnt16b*. Panel D depicts *wnt16b-i.* Panel E depicts *wnt16b-ii.* Panel F shows the sequence isolated by Cho *et al.* in 2010. Black dotted lines are present only to clearly designate panels.

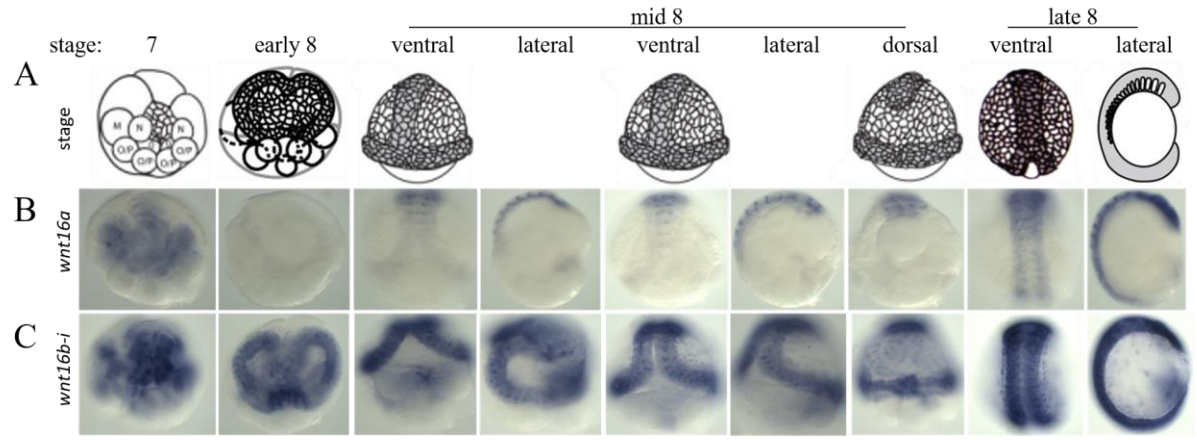


Figure 3.3. *wnt16a* and *wnt16b-i* expression patterns through stage 8. A comparison of the duplicate *wnt16* genes between sibling embryos in a timeline through stage 8. Panel A shows cartoons of the stages depicted. Panel B shows the expression pattern for *wnt16a* at these stages. The expression is in a segmentally iterated pattern in the N lineage, which is along the ventral midline. Panel C shows the expression pattern for *wnt16b-i* at the same time points. In these embryos, the dots of expression do not appear to be along the ventral midline.

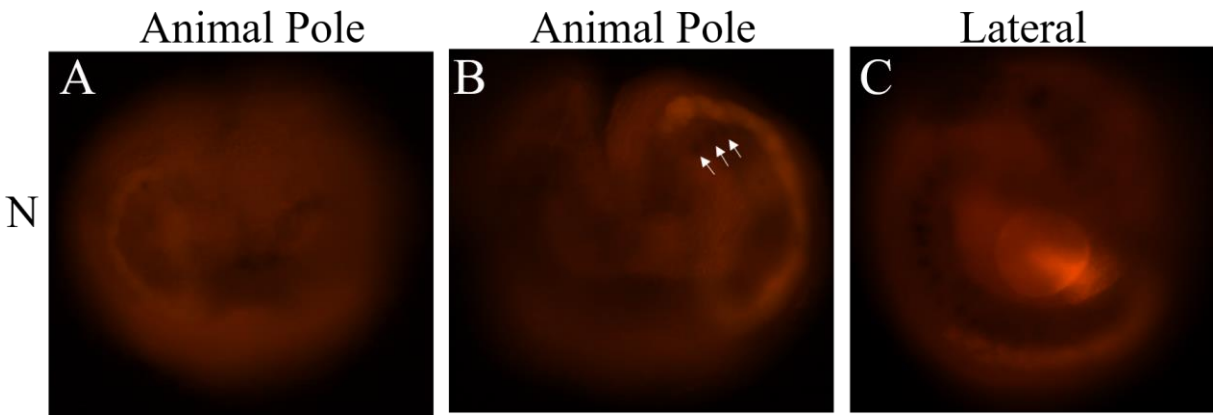


Figure 3.4. ISH for *wnt16bi* in early stage 8 embryos. The N teloblast was injected with RDA as lineage tracer at early stage 7 before being processed for ISH. Arrows indicate three dots of expression at the anterior of the germinal band at very early stage 8. Panel C shows a slightly more developed embryo with expanded expression that appears to be adjacent to the N lineage tracer.

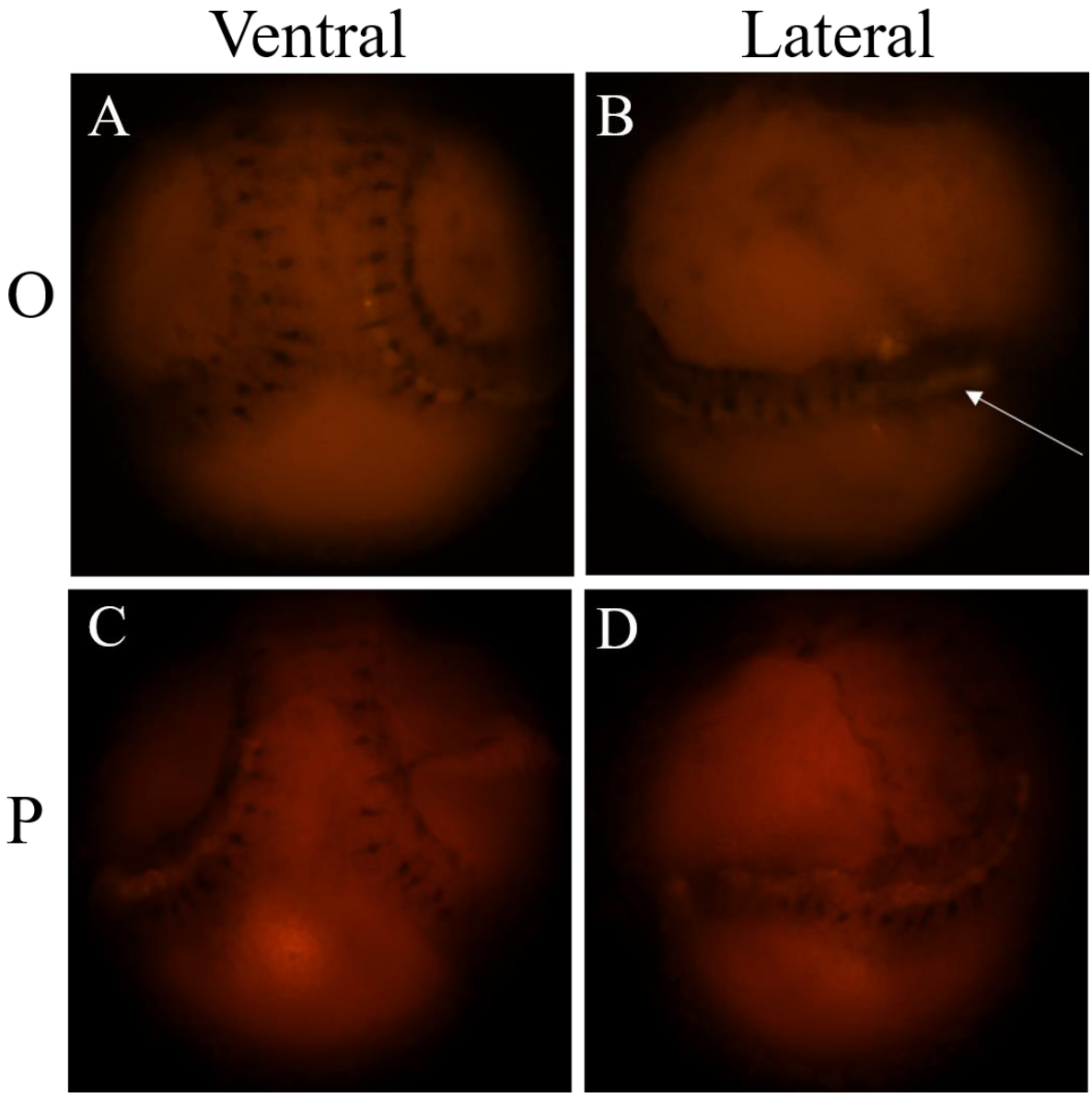


Figure 3.5. ISH for *wnt16bi* in mid-stage 8 embryos. In panels A-B the O teloblast was injected with RDA as lineage tracer at early stage 7 before being processed for ISH, and appears to overlap with the ISH signal. Arrow indicates the most posterior and youngest cells of the germinal band, where it appears the signal has ceased. In panels C-D the P teloblast was injected with RDA as lineage tracer at early stage 7, which appears to be adjacent to the ISH signal, towards the dorsal edge of the germinal band.

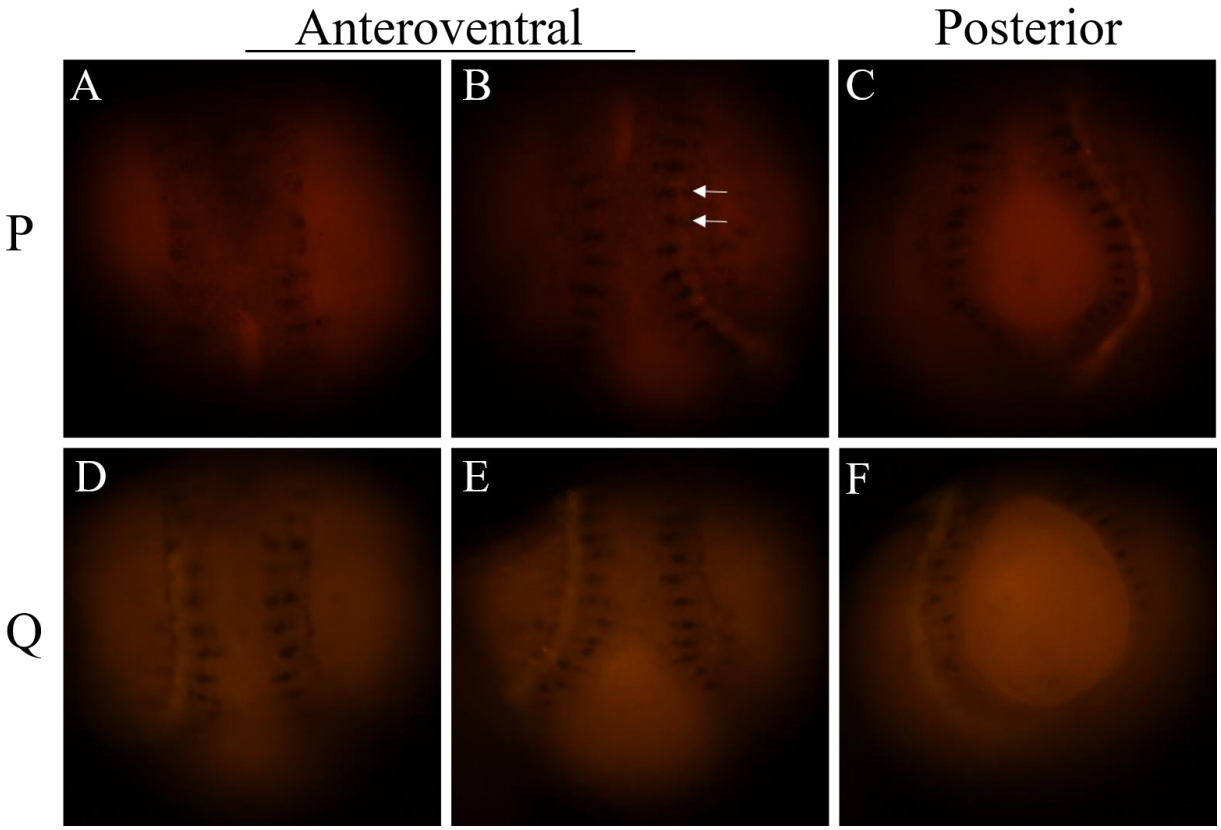


Figure 3.6. ISH for *wnt16bi* in late stage 8 embryos. In panels A-C, the P teloblast was injected with RDA as lineage tracer at early stage 7 before being processed for ISH, and appears to overlap with the ISH signal at the anterior reaches of the lineage tracer, where the arrows point in panel B. In the posterior region of the band the ISH signal and lineage tracer do not overlap, and the signal is towards the ventral edge of the band and adjacent to the tracer, suggesting that in these younger cells the signal is in the O lineage. Panels D-F show embryos with Q lineage tracing, and show that at no point does the ISH signal overlap.

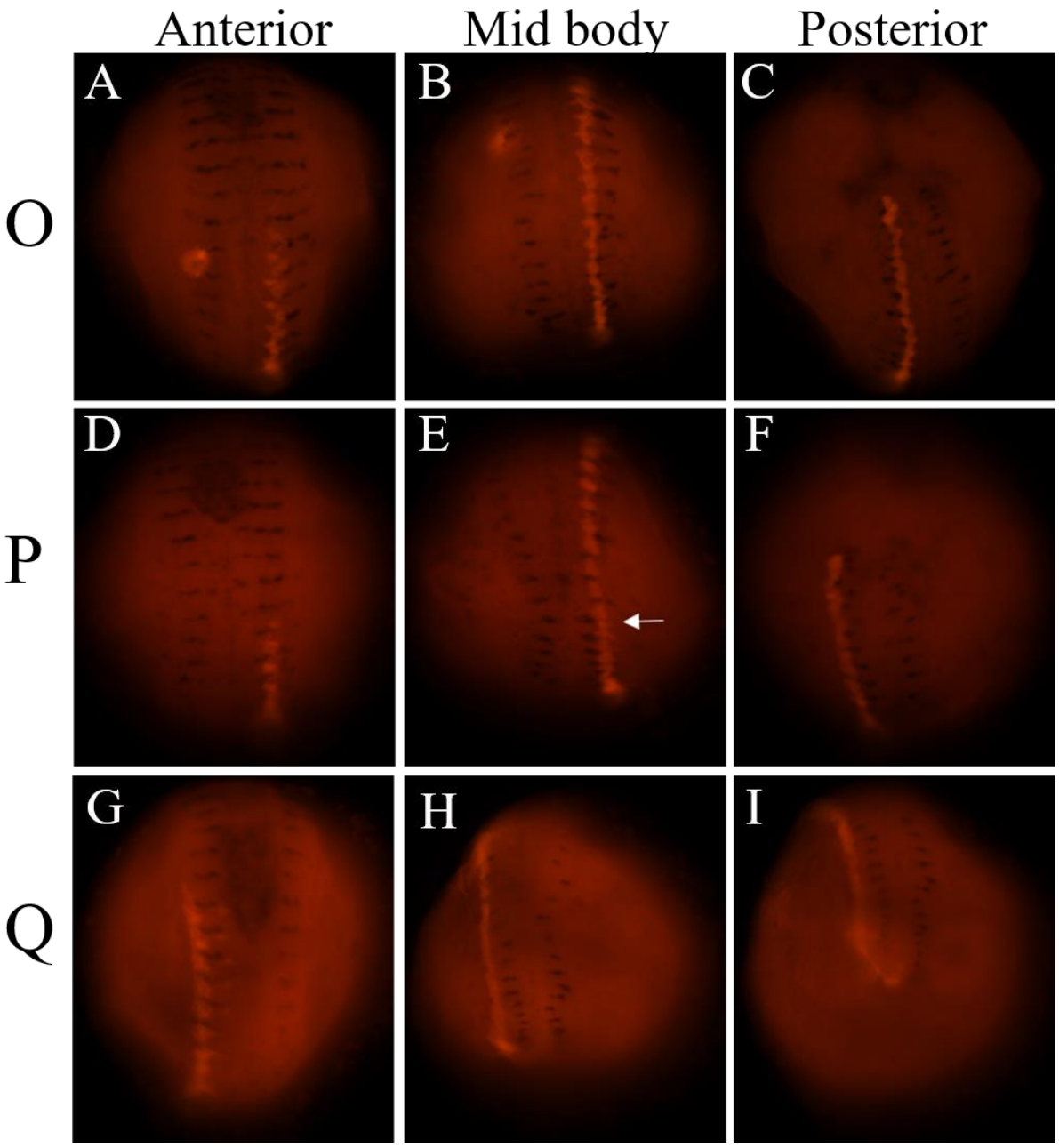


Figure 3.7. ISH for *wnt16bi* in stage 9 embryos. In panels A-C, the O teloblast was injected with RDA as lineage tracer at early stage 7 before being processed for ISH, and appears to overlap with the ISH signal at only the few most posterior cells. In panels D-F the P teloblast was injected with lineage tracer and appears to overlap with the ISH signal throughout the majority of the germinal band. The arrow in panel E shows the point where the ISH signal switches from the O to P lineage. In panels G-I the Q teloblast was injected with RDA, and does not overlap with any ISH signal.

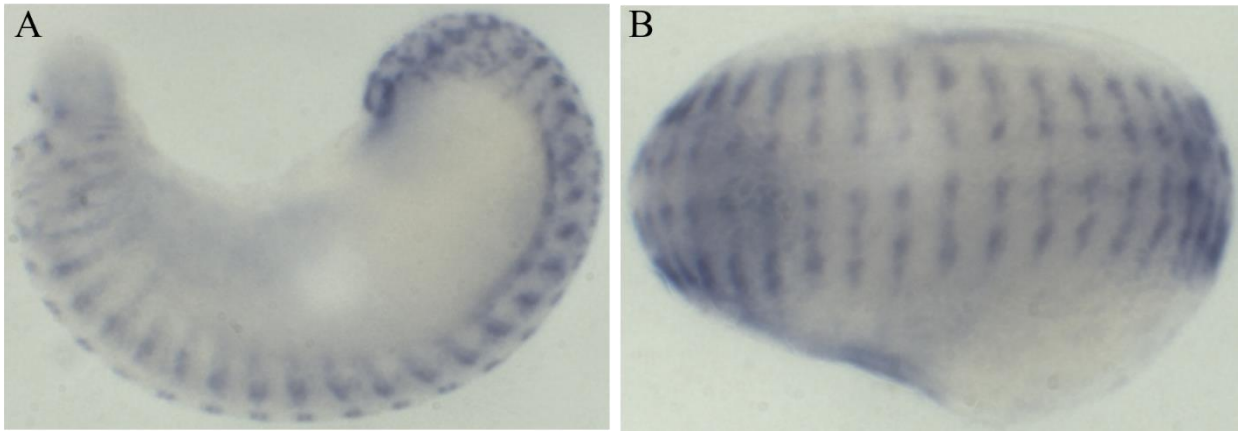


Figure 3.8. ISH for *wnt16bi* in stage 10 embryos. Panel A shows the embryo in a lateral view and Panel B shows the embryo in a ventral view. Staining is presumably in the lateral ectoderm, as the ventral midline is clear of any staining.

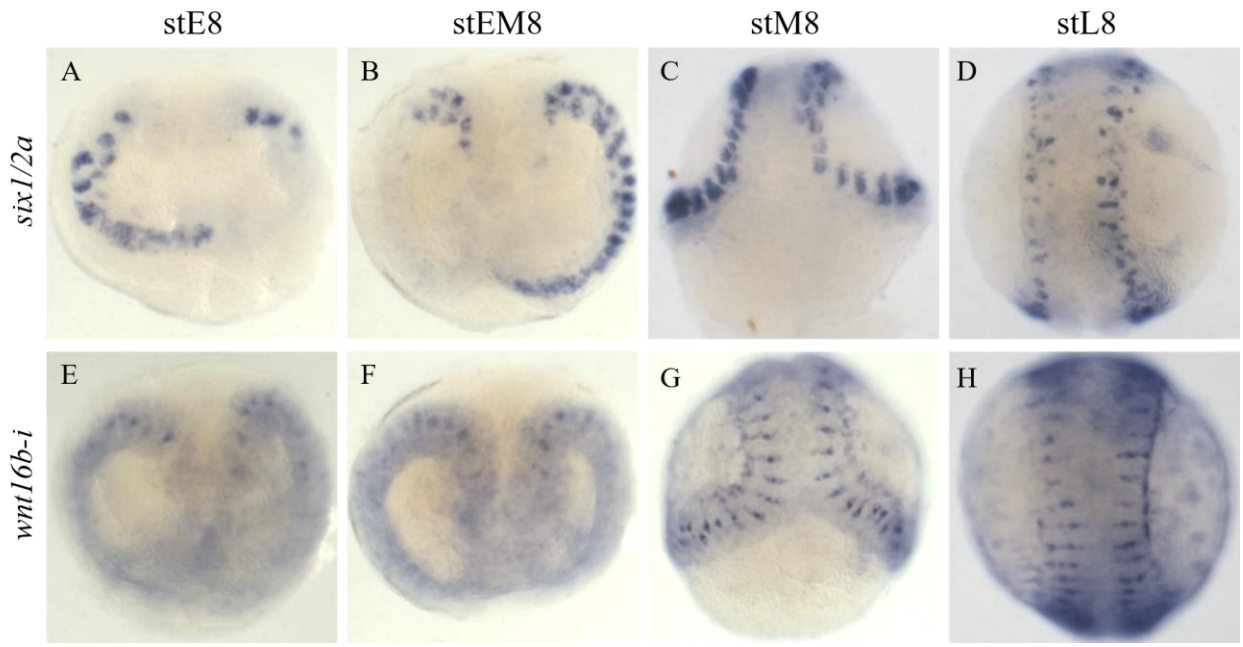


Figure 3.9. *wnt16b-i* is expressed normally in embryos with a knockdown of *bmp5/8*. ASMO to *bmp5-8* was injected into the Q lineage at stage 6. Embryos were grown to stage 8 then processed for ISH. Some embryos from each batch were used to test for expression of *six1/2a*, whose P lineage expression was shown to be abolished after *bmp5-8* KD (Kuo and Weisblat 2011). Panels A-D show the control embryos. As expected, *six1/2a* expression is abolished in early stage 8 embryos, as shown in Panels A and B. By mid to late stage 8, the normal expression pattern has returned (Panels C and D). Panels E-H show injected embryos tested for *wnt16b-i* expression. Early to mid-stage 8 embryos were expected to have two bandlets expressing *wnt16b*, as it is normally expressed in the O lineage at this stage. However, all embryos showed normal expression patterns.

Chapter 4: Characterization of the rapidly evolving *innexin* gene family in the leech *Helobdella*

Introduction

Innexin genes code for the subunits of invertebrate gap junctions

Direct cytoplasmic connections among cells, through which ions and small molecules can pass, are a common feature of both multicellular plants (plasmodesmata) and animals (gap junctions). In animals, gap junctions form by the coupling of hexameric hemichannels in the apposed membranes of adjacent cells, and the hemichannel subunits are four-pass transmembrane proteins. Despite these structural similarities, however, the gap junction proteins in chordates and protostomes are coded by non-homologous families of genes, called *connexins* and *innexins*, respectively (Phelan 2005). Thus, the vertebrate hemichannel is called a connexon, and the invertebrate analog is called an innexon. Vertebrate *inx* homologs exist (*pannexins*) but they are not responsible for most gap junctions (Dahl and Muller 2014).

Gap junction function is most well characterized in the nervous system, where they form electrical synapses (Phelan 2005, Todd *et al.* 2010). They are also essential in nervous system development (Richard *et al.* 2017). There is also evidence that indicates gap junctions are important in other aspects of development. *Drosophila inx2* is required for wing epithelial morphogenesis (Bauer *et al.* 2004). In *C. elegans*, mutants for *inx3* exhibit various developmental defects, including loss of anterior cells (Starich 2003). Surveys of transcript localization in these models, as well as studies in *Chaetopterus variopedatus*, grasshopper, planarian, and silkworm, show that *innexins* are expressed throughout development in many cell types, and paralogs are found in both overlapping and distinct regions (Stebbings *et al.* 2002, Altun *et al.* 2009, Potenza *et al.* 2003, Ganfornina *et al.* 1999, Oviedo and Levin 2007, and Hong *et al.* 2008). Intercellular communication via gap junction may also play a part in regulating differentiation and cell growth in ESCs (Wong *et al.* 2008).

The metazoan *inx* gene family has undergone significant and variable expansions (Kandarian *et al.* 2012). Among cnidarians for example, *Hydra* has 17 *inx* genes, while *Nematostella* has two (Steele *et al.* 2011). Among ecdysozoans, *Drosophila* has eight, and *Caenorhabditis* has 25 (Stebbings *et al.* 2002, Altun *et al.* 2009). Among lophotrochozoans, the genomes of a mollusk (*Lottia*), a polychaete (*Capitella*) and a leech (*Helobdella*), encode 11, 27 and 21 *inx* genes, respectively. For the most part, the expansions of the *inx* gene family appear to have occurred by independent duplication events within different evolutionary lineages, starting with only one or two copies at the base of the Bilateria (Phelan 2005, Kandarian *et al.* 2012, Simakov *et al.* 2013). This indicates a high degree of evolutionary dynamism relative to many other gene families.

Intriguingly, the expansion of this gene family appears to be an ongoing process, as differences in *inx* gene complement are observed even at lower taxonomic levels. Genome surveys of the hirudinid leech *Hirudo verbana* and of the glossiphoniid leech *Helobdella robusta* each found 21 *innexin* genes, but molecular phylogenies suggest that these 21 genes arose from just 19 genes in the ancestral leech (Kandarian *et al.* 2012). Thus, at least two independent *inx* duplications have occurred just within leeches.

Kandarian and co-workers also characterized the expression of the *Hirudo inx* genes in adults and late stages of development. As expected, the majority of the *inx* genes were expressed in subsets of neural cells. For three genes, however, no expression was observed in either the

Hirudo late stage embryo or adult. A challenge of using *Hirudo* in embryology is that, unlike *Helobdella*, their embryos are not readily accessible until later stages of development, well after cleavage and gastrulation. Thus, in the work reported here, we used the leech *Helobdella austinensis* to compare expression of orthologous genes between the two leech species, to test for functional divergences in the recently duplicated paralogs, and to study the expression of leech *inx* genes in early development. This also provides an opportunity to examine the level of divergence in more recent duplication events within a gene family that has greatly expanded throughout evolution.

We found homologous expression patterns for many genes between leeches, although there are some differences. Duplicate genes within *Helobdella* may be diverging in function. Most *innexins* are broadly expressed in early *Helobdella*.

Materials and Methods

Cloning, Sequencing, and In Vitro Transcription

Specific primers (Table 4.1) were designed against *Helobdella robusta* genomic sequences and used for amplification from *Helobdella austinensis* cDNA. Procedures for cloning, sequencing, and in vitro transcription of antisense riboprobes were described in Chapter 2.

ISH in Helobdella

In situ hybridizations were performed as described in Chapter 2. Approximate probe lengths were as follows: *Hau-inx1* (170896): 900bp; *Hau-inx2A* (73579): 1100bp; *Hau-inx2B* (185033): 900bp; *Hau-inx3* (193752): 1100bp; *Hau-inx4* (113191): 1200bp; *Hau-inx5* (100634): 1100bp; *Hau-inx6* (132793): 850bp; *Hau-inx7* (122048): 800bp; *Hau-inx8* (185762): 950bp; *Hau-inx9* (114074): 700bp; *Hau-inx10* (154179): 1100bp; *Hau-inx11* (98251): 950bp; *Hau-inx12* (179748): 1250bp; *Hau-inx13* (106068): 1000bp; *Hau-inx14* (70423): 1000bp; *Hau-inx15A* (82276): 1000bp; *Hau-inx15B* (95018): 1000bp; *Hau-inx16* (182016): 1000bp; *Hau-inx17* (96495): 1200bp; *Hau-inx18* (176569): 1200bp; *Hau-inx19* (132794): 850bp.

Results

Helobdella development

Leeches of the genus *Helobdella* are well-suited for studies of early development due to the large size (400µm diameter), and accessibility of their embryos. Fertilization occurs internally and the zygotes are then deposited prior to completion of the first maternal meiosis into cocoons on the ventral surface of the adult, from which they can be removed and cultured to adulthood (Weisblat and Kuo 2009). Their development has been divided into 11 stages, starting with a version of unequal spiral cleavage that segregates yolk-deficient, RNA-enriched, fate-determining cytoplasm called teloplasm to macromere D' of the 8-cell embryo (Weisblat and Kuo 2014, Figure 1.2). At fourth cleavage, macromere D' cleaves in an obliquely equatorial manner; the vegetal daughter, DM, is the precursor of segmental mesoderm and the animal

daughter is the precursor of segmental ectoderm. Subsequent cleavages of these two cells generate five bilateral pairs of lineage-restricted stem cells called teloblasts: one pair of mesoteloblasts (M), and four pairs of ectoteloblasts (N, O/P, O/P, and Q, from ventral to dorsal), along with additional small cells. Each teloblast undergoes highly asymmetric divisions to give rise to a column (bandlet) of segmental founder cells (blast cells). Ipsilateral bandlets coalesce into parallel arrays (left and right germinal bands), which migrate ventrovegetally over the surface of the embryos, gradually zippering together along the prospective ventral midline into a bilaterally symmetric sheet of cells (germinal plate) from which segmental ectoderm and mesoderm arise in anteroposterior progression.

To gain an overview of changing patterns of *innexin* gene expression throughout development, we selected four "early" stages, and three "late" stages of development (depicted by line drawing in Figure 1.2) for examination by *in situ* hybridization (ISH). The early stages are cleavage stages 2 and 5, stage 7 when primary blast cells begin to be produced, and mid stage 8, during epiboly and formation of the germinal plate. The late stages are 9 and 10, during which segmental differentiation becomes increasingly evident and the final embryonic stage, 11, by which segmental differentiation is largely complete.

inx expression in early Helobdella development: cleavage stages

The differential regulation of gap junctional coupling (and presumably, by proxy, of *innexin* gene expression) in early stages of spiralian embryos is suggested by the specific coupling of cells within the "molluscan cross" of *Patella* (a characteristic composition of embryonic cells; Serras and van del Biggelaar 1987), although the developmental significance of this phenomenon remains to be determined. *Helobdella* blastomeres are also extensively coupled via gap junctions from the 2-cell stage through stage 7, as judged by the ready diffusion throughout the embryo of small molecules (<~1200 Da) injected into any given cell (Weisblat *et al.* 1978, 1980; Bissen and Weisblat 1989). The tractability of the early *Helobdella* embryos allowed us to discover a dynamic, gene-specific, but extensively overlapping patterns of *inx* expression at stages of development that are not readily available in *Hirudo*.

For six of the 21 *Helobdella inx* genes, we detected no expression (*inx3*, *5*, and *6*) or very light expression (*inx10*, *13* and *16*) *during* stages 2-7 (Figures 4.1 and 4.2). In contrast, six other *Helobdella inx* genes (*inx2b 7*, *12*, *15A*, *15B* and *19*) appeared to be strongly expressed in one or more of the three cleavage stages examined as judged by the intensity of the ISH signal. For most of these genes transcripts appeared to be concentrated in the yolk-free cytoplasm (teloplasm), a D lineage determinant factor (Astrow and Weisblat 1989) that is a prominent component of cell CD of at the 2-cell embryo (stage 2) and of cells $NOPQ_L$ and $NOPQ_R$ (the left and right ectodermal precursors) in stage 5 embryos (Figures 4.1 and 4.2). Exceptions to this generalization include *inx7*, which was more broadly distributed in the cytoplasm of the cells at stage 2, and *inx12*, for which expression declined sharply in stage 5 relative to stage 2. Several of these transcripts exhibited a marked decline in abundance by the end of cleavage (stage 7), including *inx2B*, *12*, *15B* and *19*.

The remaining 9 *inx* genes were detected at intermediate levels during one or more cleavage stages. Several of these (*inx2A*, *8*, *9*, *14* and *18*) appeared to be localized primarily in the teloplasm during early cleavage and then largely disappear by stage 7. In contrast, the expression of *inx17* appeared stronger at stage 7 than at stages 2 and 5. During these cleavage stages, *inx2A*, *8*, *9*, *11*, *12*, *14*, *15B* and *19* all exhibit expression in the macromeres, which are

the main endodermal precursors (Figures 4.1 and 4.2), while *inx1, 7, 15A* and *19* exhibited enhanced expression in micromeres, a heterogeneous population of small cells that contribute to nonsegmental structures, including the dorsal anterior ganglion, proboscis, and the epithelial layer of the provisional integument that envelops the embryo during gastrulation (Weisblat and Kuo 2014).

inx expression in early Helobdella development: germinal bands and germinal plate

By the end of cleavage, the D quadrant of the early *Helobdella* embryo has generated five bilateral pairs of lineage-restricted segmentation stem cells called teloblasts. The teloblasts' iterated stem cell divisions generate columns of much smaller segment founder cells, which merge into parallel arrays (left and right germinal bands) during stage 7. During stage 8, these gradually coalesce in an anteroposterior progression along the ventral midline during to form a sheet of cells (the germinal plate) from which segmental ectoderm and mesoderm arise (for review see Weisblat and Kuo 2014, Figure 1.2). The derived state of development in *Hirudo* relative to *Helobdella* and other clitellate annelids makes it hard to homologize their embryonic stages, but the fully coalesced germinal plate of *Helobdella* (early stage 9) roughly corresponds to the earliest stage at which *inx* expression has been examined in *Hirudo* (embryonic day 6; Fernandez and Stent 1982; Dykes and Macagno 2006).

Remarkably, all but 3-4 of the 21 *Helobdella inx* genes were expressed during stage 8 within the germinal bands and/or germinal plate (Figure 4.2; *inx5, 6* and *9* were not detected at all in these tissues and *inx13* was barely detected). A hypothesis for future investigation is that this apparently overlapping expression represents emerging cellular heterogeneity as tissues differentiate during segmentation. The observation that the expression of several *inx* genes, mostly prominently *inx3,* is up-regulated in anteroposterior progression within the differentiating germinal plate is consistent with this hypothesis. Unfortunately, because of technical difficulties arising from the small size of its embryo, nothing is known about the expression of *innexin* genes in *Hirudo* at stages prior to germinal plate formation.

Expression of innexins present as single copy genes in Helobdella and Hirudo during late development

Molecular phylogenies reveal that 15 of the 21 *inx* genes exist as single copy genes in both *Hirudo* and *Helobdella* (Kandarian *et al..*, 2012). The simplest interpretation of such a one-to-one correspondence is that these 15 pairs of genes are mutual orthologs (although this is not necessarily the case--see Discussion), and should therefore show similar patterns of expression. As indicated above, nothing is known about the expression of *innexin* genes in *Hirudo* at stages prior to germinal plate formation. Where comparable structures and stages of development could be examined, however, most of the *inx* gene pairs showed conserved patterns of expression between the two species.

Unambiguously homologous patterns of expression were observed for *Hau-inx4* and *Hau-inx10*. These genes were both expressed selectively in nephridia at stages 10-11 (Figure 4.3), as are the *Hirudo* homologs at comparable stages of segment differentiation (Dykes and Macagno, 2006). Each of these genes is more closely related to other *inx* that are not selectively expressed in nephridia than to each other (Kandarian *et al..*, 2012).

The central nervous system provided other opportunities to test for the homology of expression among *inx* genes. In *Hirudo*, *inx1* and *inx14* appeared to be expressed in all neurons of the segmental ganglia, and were thus classified as pan-neuronal (Dykes and Macagno, 2006; Kandarian *et al*.., 2012). We observed similarly strong and seemingly uniform expression of their *Helobdella* homologs (Figure 4.3).

Six other *inx* genes (*5*, *6*, *8*, *16*, *17*, and *19*) were reported as being differentially expressed by subsets of neurons in the segmental ganglia of adult *Hirudo* (Dykes and Macagno, 2006; Kandarian *et al*.., 2012). Of these, *inx16* and *inx17* are clearly expressed in the segmental ganglia of *Helobdella* as well (Figure 4.4); *inx19* was also detected in segmental ganglia, though at lower levels, at stage 11 (Figure 4.4) Expression of these three genes appeared to be pan-neuronal, but this does not preclude the possibility of relatively subtle differences in expression among cells that we could not detect. In contrast, we observed no in situ hybridization signal above background levels for *Hau-inx5*, *Hau-inx6*, or *Hau-inx8* in the late stage embryos, except for transient expression of *Hau-inx8* at the mouth during stage 9 (Figure 4.5). We note, however, that the levels of expression of these genes in *Hirudo* was quite low overall and showed little contrast between background and the subsets of *inx*-positive ganglionic neurons. Thus, we cannot exclude the possibility that we were simply unable to pick up homologous patterns of expression in the much smaller *Helobdella* embryos.

Five other *inx* genes are present as single copy genes in both *Hirudo* and *Helobdella*. For two of these, *inx12* and *inx18*, no clear expression was reported in *Hirudo*. In *Helobdella*, we also failed to observe reproducible patterns of expression for *inx18*, but observed *inx12* expression in the proboscis and foregut during stages 10-11 (Figure 4.5). *Hirudo inx7* is one of several *inx* genes that is expressed in the male and female reproductive primordia and is also expressed prominently in a pair of lateral stripes. In *Helobdella*, *inx7* is expressed throughout the germinal plate including the segmental ganglia, and at higher levels in some lateral tissues (Fig 4.4), but whether these are homologous with the lateral stripes of *inx7* in *Hirudo* remains to be determined, because the fates of the *inx7*-positive cells in either species remain to be determined. The fourth gene, *inx13* is expressed in cells associated with the lumenal openings of nephridia and other organs in *Hirudo*, but was not detected in comparable stages of *Helobdella* development (Figure 4.5). Finally, *inx3* is expressed in macroglia in *Hirudo*, whereas in *Helobdella*, *inx3* is expressed strongly but transiently during stage 10 in the developing proboscis (Figure 4.5), and at lower levels during stage 9 within the germinal plate.

Thus, of the 15 *inx* genes present as single copy genes in both *Hirudo* and *Helobdella*, at least four and arguably as many as eight or nine showed homologous patterns of expression. For the remaining seven genes, there was no strong evidence either for or against homologous patterns of expression.

Late expression of the innexins present as multi-copy genes in either Helobdella or Hirudo: genes duplicated in Helobdella relative to Hirudo

In addition to the 15 single copy *inx* genes described above, the *Hirudo* genome encodes two paralogs each for *inx9* and *inx11*, and that the *Helobdella* genome encodes two paralogs each for *inx2* and *inx15*. One interpretation of these observations is that each of the four genes underwent independent duplication events in the parallel lineages from the common ancestor to *Hirudo* or *Helobdella*, respectively, but other scenarios are possible, as will be discussed below.

In the *Helobdella* genome, the paralogous *inx2* genes occur on separate scaffolds and are flanked by non-homologous genes. Moreover, the gene models for *Hau-inx2A* and *Hve-inx2* both contain one intron, whereas that for *Hau-inx2B* is intronless. These observations suggest that duplication of *inx2* occurred by a retrotransposition, with *Hau-inx2A* representing the original gene and *Hau-inx2B* the duplicate, which is presumed to have been inserted randomly into the genome and thus to have lost association with its original regulatory elements. This scenario predicts that *Hau-inx2A,* and not *Hau-inx2B* will exhibit a pattern of expression similar to that of *Hve-inx2*, Consistent with this prediction both *Hve-inx2* (Kandarian *et al.* 2012) and *Hau-inx2A* (Figure 4.6) are expressed along the midline in a pattern associated with the giant glia of the segmental ganglia and interganglionic connective nerves, while *Hau-inx2B* appears to be expressed only in the early embryo, as described above.

For *inx15*, the two *Helobdella* paralogs are also on separate scaffolds flanked by non-homologous genes. In this case, *Hau-15B* contains 4/5 introns, while *Hau-15A* appears as an intronless gene model, which could again represent an instance of retrotransposon-mediated gene duplication. In this case, however, the *Hve-inx15* is also intronless. One explanation for this result is that duplication of an intronless *inx15* in the lineage leading to *Helobdella* was followed by the rapid acquisition of multiple introns; a more parsimonious scenario is that duplication of a multi-intron *inx15* by retrotransposition occurred prior to the divergence of the two leech species and that the original gene has been lost from *Hirudo.* Both these scenarios predict that the expression patterns of *Hau-inx15A* and *Hve-inx15* should resemble one another more closely. Unfortunately, in this case, there is no clear evidence either way. As described above, both *Hau-inx15A* and *Hau-inx15B* are strongly expressed in stages of early development for which no data is available from *Hirudo*. We observe expression of *Hau-inx15B* in the midgut and adjacent ganglia, while *Hau-inx15A* appears to be broadly expressed during stages 9-10 and much less or not at all by stage 11 (Figure 4.6). *Hve-inx15* was detected in embryonic (and adult nervous) system by qRT-PCR but no in situ signal was obtained for the embryonic nervous system (Kandarian *et al.* 2012).

Late expression of the innexins present as multi-copy genes in either Helobdella or Hirudo: genes duplicated in Hirudo relative to Helobdella

In contrast to the situation described above for duplicated *inx* genes in *Helobdella,* paralogous *inx* genes in *Hirudo* have retained similar numbers of exons: 7 in both *Hve-inx9A* and *Hve-inx9B;* 5 in *Hve-in11A* and 4 in *Hve-inx11A* (Kandarian *et al.* 2012). Both *Hau-inx9* and *Hau-in11* also occur as multi-exon gene models (7 and 8 exons, respectively, though these have yet to be validated experimentally). These observations argue against the possibility of duplication by retrotransposition for the two *Hirudo* genes. In this case, then, we might predict that comparing the expression patterns of the duplicated *Hirudo* paralogs to the single copy *Helobdella* gene might give evidence for how expression of the ancestral gene has been modified following duplication. For the *inx9* paralogs, *Hve-inx9A* is expressed in the mesoderm of the early germinal plate and, in what appears to be microglia and/or ganglion sheath cells of the late embryo and adult (Dykes and Macagno 2006). Consistent with this, we observe weak expression of *Hau-inx9* in presumptive mesoderm of the stage 9 embryo and in association with segmental ganglia at stages 10-11 (Figure 4.6). By comparison, *Hve-inx9B* appeared to be more weakly expressed, though it was detected by qRT-PCR in the adult nervous system (Kandarian *et al.* 2012), possibly indicating a loss of function for the *Hve-inx9B* paralog.

None of the *inx11* paralogs exhibited a robust and interpretable expression pattern in the segmental tissues of either species (Figure 4.6). This could indicate that *inx11* functions primarily in early development, where expression is seen for *Helobdella*, but since these stages are not available for *Hirudo* we cannot draw any conclusions about divergence in expression between the *Hirudo* paralogs.

Discussion

Molecular phylogenies suggest that the ancestral bilaterian *inx* gene family comprised only one or two copies (Phelan 2005, Kandarian *et al.* 2012, Simakov *et al.* 2013), and that there have been multiple independent lineage-specific expansions from that point (Kandarian *et al.* 2012). Intriguingly, this process is still continuing, as judged by information available from two leech species. The genomes of both *Hirudo* and *Helobdella* encode 21 *innexin* genes, but molecular phylogenies suggest that these 21 genes arose from just 19 genes in their last common ancestor, with two independent duplications having occurred in the lineage leading to the extant species from that common ancestor (Kandarian *et al.*., 2012).

Here, we have surveyed the expression of all known members of this evolutionarily dynamic gene family in the glossiphoniid leech *Helobdella austinensis*, in stages throughout development. Our work extends previous studies of leech *innexins* by including stages of early development that are not readily accessible in the medicinal leech *Hirudo verbana*.

Apparent differences between the two species stemming from failure to observe expression in *Helobdella* are difficult to interpret in some cases because the smaller size of the *Helobdella* embryo prevents us from observing faint or spatially restricted staining patterns with the same level of resolution as in *Hirudo*, and in others because of uncertainties as to the identities of cells or tissues expressing certain *inx* genes. For example, *Hirudo* undergoes an evolutionarily derived version of indirect development compared to glossophoniid leeches and oligochaete outgroups, such as a prominent cryptolarval mouth and protonephridia for which there no obviously homologous structure exist in the *Helobdella* embryo.

innexin expression in early development

One motivation for our study was to test the hypothesis that *inx* genes for which little or no robust expression was observed in the late stage embryo and adult of *Hirudo* might function instead or as well in the early leech embryo. Previous characterization of the *inx* gene family in *Hirudo* revealed a few members (e.g., *inx8, inx11A, inx11B, inx12, inx15, inx18*) that showed only highly limited, weak or no evident expression in the stages and tissues examined (Dykes and Macagno, 2006; Kandarian *et al.*., 2012). Consistent with our expectations, and with the extensive gap junction coupling of cells in the early *Helobdella* embryo, we found that these *inx* genes and most others are expressed in dynamic, yet extensively overlapping patterns throughout the stages of early development. Expression of multiple *inx* genes was observed both in the lineages leading to the germinal plate, from which segmental mesoderm and ectoderm differentiate, as well as in the micromeres and non-D quadrant macromeres, which contribute to gut and non-segmental tissues. This broad expression of multiple *inx* genes throughout development is consistent with the plethora of possible functions that have been proposed for innexins (Altun *et al.* 2009, Oviedo and Levin 2007, Phelan 2005, Stebbings *et al.* 2002, Landesman *et al.* 2000, and Lo 1999).

innexins represented by single copy genes in both species

We presume that most or all of the 15 *inx* genes present as single copy genes in both *Helobdella* and *Hirudo* are orthologs of one another. Exceptions to this assumption are possible, however--for example, if an ancestral duplication was followed by the loss of one paralog in the lineage leading to *Hirudo* and the other paralog in the lineage leading to *Helobdella* (Catchen *et al.* 2008). Distinguishing these and other possibilities is beyond the scope of the present work but should be possible by validating the *inx* gene models in *Hirudo* and *Helobdella*, combined with similar catalogs of *inx* genes in other clitellate annelids.

In any event, we would predict to see orthologous *inx* genes expressed in similar patterns or tissues between the two species, and this is largely the case for the *inx* genes existing as one-to-one pairs between the two species. Genes expressed strongly and specifically in nephridia (*inx4* and *inx10*) and in the segmental ganglia (*inx1* and *inx14*) provide the clearest examples of homologous expression. For other *inx* genes, we cannot exclude the possibility that apparent differences in expression reflect some combination of technical issues. For example, *inx5, 6, 8, 16, 17,* and *19* are described as labeling specific subsets of neurons in *Hirudo*, while in *Helobdella* we observe either no ganglionic expression (*inx5, 6, 8*) or pan-neuronal expression (*inx16, 17, 19*). While these differences could reflect evolutionary divergence in *inx* function, we cannot exclude a more conservative hypothesis, namely that they result from differences in probe quality and/or staining conditions between the different experiments, and/or the poorer cellular resolution in the smaller *Helobdella* ganglia. As another example, for the *inx1, 5, 6, 16* genes, we cannot assess whether the gut expression observed in *Helobdella* is conserved in *Hirudo* because that tissue is dissected away in the *Hirudo* preparation.

It was also anticipated that expansion and divergence of the *inx* gene family might be reflected in morphological and developmental differences between the two species. Apart from early development, for which no information is available for *Hirudo*, the most obvious such differences between the two species are the cryptolarval structures present in *Hirudo* but not in direct developing *Helobdella*, or the dramatic differences in feeding structures--*Helobdella* features a structurally elaborate proboscis, for which there no clear homolog in *Hirudo*. The larval mouth of *Hirudo* shows strong expression of *inx1, 2, 3, 9, 10*; intriguingly, the *Hirudo* protonephridia express *inx2* most strongly, while the definitive nephridia express *inx4* and *inx10* in both species. In *Helobdella, inx3, inx12* and *inx17* are all expressed in the proboscis.

innexins present as two copies in either Helobdella or Hirudo

A priori, gene duplication events during evolution may have various consequences, grouped broadly as follows (see Lynch and Conery 2000 for a more detailed review): retention of both paralogs with their original function and patterns of expression (redundancy); inactivation and the eventual disappearance of one paralog (non-functionalization); divergence of one or both paralogs to assume different domains of expression and/or biochemical roles (variously neo-functionalization, sub-functionalization or syn-functionalization).

Which of these scenarios ensues following any given duplication event depends on a combination of chance (e.g., subsequent mutational and selection events), the original structure and function of the duplicated gene, and also by details of the duplication event, such as whether or not regulatory sequences are duplicated along with the coding sequence of the gene.

We have previously examined the evolutionary dynamics of the *wnt* gene family within the super-phylum Lophotrochozoa (Cho *et al.*., 2010). That work included comparing the expression of genes duplicated in a leech (*H. austinensis*) with respect to a polychaete (*Capitella teleta*), two rather distantly related annelids. In the work presented here, the close evolutionary relationship between *Helobdella* and *Hirudo*, combined with the rapidly evolving *inx* gene family provides an opportunity for examining relatively near term regulatory and functional consequences of gene duplications in one species relative to another.

Previous studies revealed that *inx2* and *inx15* are represented as paralog pairs in *Helobdella*, while *inx9* and *inx11* are represented by paralogs pairs in *Hirudo* (Kandarian *et al.* 2012). Without high quality genome assemblies for both species, it is impossible to be sure, but the available information available leads us to speculate that these duplicates might have arisen by different evolutionary scenarios.

For *inx2*, the most parsimonious scenario is that duplication occurred by retrotransposition after divergence of the lineages leading to *Hirudo* and *Helobdella.* Evidence in support of this hypothesis is that *Hau-inx2B* has no intron (compared with one intron in both *Hau-inx2A* and *Hve-inx2*), and that *Hau-inx2A* and *Hve-inx2* show similar patterns of expression in macroglia, while expression of *Hau-inx2B* is strong in the early embryo and appears to turn off after the germinal plate forms.

For *inx15*, we speculate that the observed duplication in *Helobdella* again reflects retrotransposition, but in this case prior to the divergence of the two leech species, followed by loss of the original gene in the *Hirudo* lineage, so that *Hau-inx15A* and *Hve-inx15* both lack introns. Expression data is again consistent for this case; the two *Helobdella* paralogs are expressed in different patterns as expected, while both *Hau-in15A* and *Hve-inx15* are detected in segmental ganglia.

For *inx9* and *inx11*, the apparent conservation of intron-exon architecture between the paralogs indicates that the duplication observed in *Hirudo* did not occur by retrotransposition *Hve-inx9A* appears to be expressed in microglia or sheath cells associated with the segmental ganglia and the rather expression we observe for *Hau-inx9* is consistent with this. Little expression was observed for *Hve-inx9B*, which could reflect either non-functionalization or expression in tissues or stages of development that were not examined. Similarly, no clear expression was observed for either paralog of *inx11* in *Hirudo*, while that of the *Hau-inx11* was restricted to stages of early development that were not available in *Hirudo.*

Conclusion

A major realization from comparative genomic analyses is the extent to which the remarkable diversification of body plans among bilaterally symmetric animals is supported by genomes that for the most part contain similar numbers and kinds of genes. For example, one recent estimate is that there were 8822 genes present as single copy progenitors in the last bilaterian ancestor (Albertin *et al.* 2015). It is conservatively estimated that that 47-85% of genes in various extant bilaterians have arisen by the expansion of these ancestral genes (Simakov *et al.*., 2013).

Thus, lineage-specific expansion of gene families is a hallmark of bilaterian genome evolution. The *innexin* gene family, which encodes proteins responsible for forming gap junctions in invertebrate species, provides an interesting example of this phenomenon, because it

exhibits extensive lineage-specific expansions among bilaterian taxa, which are still underway as evidenced by the comparisons of two extant leech species.

Table 4.1. List of primers used to amplify *innexin* genes

Name	Primer Sequence	
Hau-inx1-F	5’ GAAACGCAGAAGAACCGGCAGCTGATC	3’
Hau-inx1-R	5’ CGTACCTGTCCTTCCTGTCTTGCTGAT	3’
Hau-inx2A-F	5’ ATTAGTGCTGTTGGCAAGATACCAGAT	3’
Hau-inx2A-R	5’ TTCTTTTTTAAATCGTTCCGTTAACGG	3’
Hau-inx2B-F	5’ GAGCAGATTCTTGGAGTGTTGGGTAAG	3’
Hau-inx2B-R	5’ GTAACGGTGCTTGTCAGATTTCATTAC	3’
Hau-inx3-F	5’ CTAGTGGCAACAGCCGTCAAAATGGCC	3’
Hau-inx3-R	5’ ACCAGAGAGGAGGAGAAGGTTCTTGAC	3’
Hau-inx4-F	5’ GCGTCATCAGTTCGAAGCTCGAACGAT	3’
Hau-inx4-R	5’ CTCCTTCCCCTTCAACTCATCTGAATA	3’
Hau-inx5-F	5’ ATGTCAAAGTTGATAACCTCGAGGCGA	3’
Hau-inx5-R	5’ CGTGTCGGCGTTTATGGCTAACAATTT	3’
Hau-inx6-F	5’ CGAGATATCAAGCTAAGATCTTTCGTC	3’
Hau-inx6-R	5’ ATCCTGATCGTTGGCGTTCAGTGAAAT	3’
Hau-inx7-F	5’ TTATTACAAAGCGCCATCTGCTTCCTG	3’
Hau-inx7-R	5’ GATGTCAGATTCGACGAGTCGAGTCGA	3’
Hau-inx8-F	5’ AAGAACGACGAATCTCACTATGAGCCC	3’
Hau-inx8-R	5’ AAAGTCGTCCATCTCCATTTCCGGGTA	3’
Hau-inx9-F	5’ AGGTTCTTGGTGGAGTTCAACAAGTCC	3’
Hau-inx9-R	5’ CTCTTCCTGTTGCATCCTCCTGTTGTG	3’
Hau-inx10-F	5’ TCAGTTTCTGTTGCCAAAGGAAGAAGT	3’
Hau-inx10-R	5’ AACTTTTTTTTCAGGCGAAGTATCTTG	3’
Hau-inx11-F	5’ ACAAGTACGTACTATCTCTCCGAAGAA	3’
Hau-inx11-R	5’ AGCATATCTATCGCCAACAAAGTTGAG	3’
Hau-inx12-F	5’ AGCGGTCGGGCCGATAGCTTTAGTGAT	3’
Hau-inx12-R	5’ CATTTTCAAGCAAACCTCCTCGAGAGA	3’
Hau-inx13-F	5’ ACGTGGGATACAATACTGGAGCAGGCT	3’
Hau-inx13-R	5’ CCGGCAAAACCTTTTATGTTGCGGTT	3’
Hau-inx14-F	5’ AAAACCTTATCAGACGCCGCCAACCAC	3’
Hau-inx14-R	5’ TCGTCCGCCGTTCGCTCGATTTTTAAC	3’
Hau-inx15A-F	5’ TCTCAGTACGTCGGCGAACCTATACAC	3’
Hau-inx15A-R	5’ GGCTCCGATCATCGGTTTGTTCTTATA	3’
Hau-inx15B-F	5’ GTTATCAGCACCACGCAGTACGTCGGC	3’
Hau-inx15B-R	5’ CGTCTTGAAGTGGTCCCAAAGAGCCGC	3’
Hau-inx16-F	5’ GAGAGGCTGGACATGACGACGCGTGAC	3’
Hau-inx16-R	5’ TGGCCGTGAGCCACTACTGCTGAACTG	3’
Hau-inx17-F	5’ AGACTCTTCAAGTCGGTCCTCAGCGTT	3’
Hau-inx17-R	5’ CGAGGAACCGGAGGAAAGAAAATAGAC	3’
Hau-inx18-F	5’ GTTGAGGGTCTATTGAACACAGCGTCT	3’
Hau-inx18-R	5’ CCTCCTCCTCACTTCCGTTTTGCAGCA	3’
Hau-inx19-F	5’ CACCCAGAGGAAGTTAGACAGAGAAAG	3’
Hau-inx19-R	5’ TACTAACTCACCAGAATTTTTCGCCAC	3’

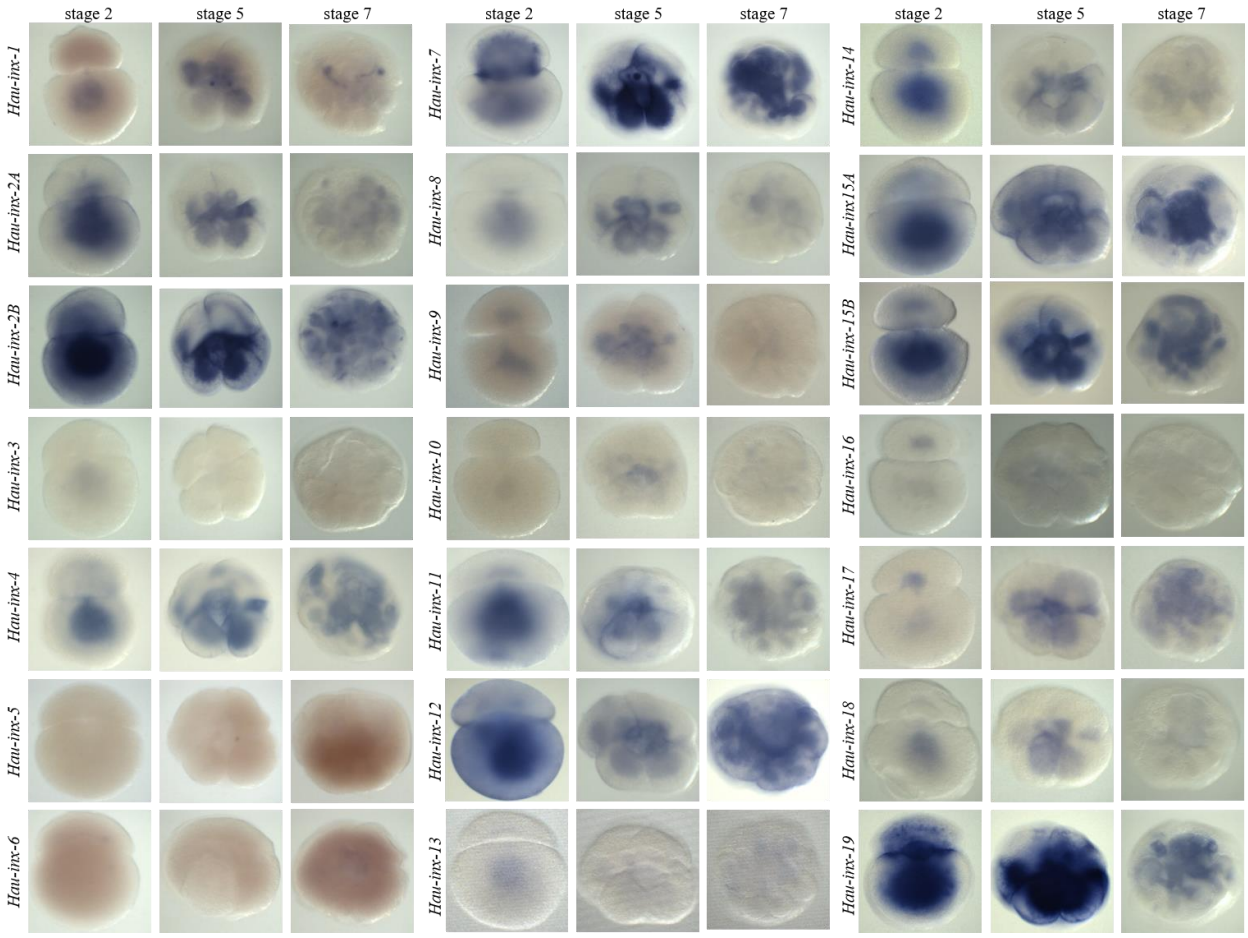


Figure 4.1. Expression of *inx* genes in early cleavage stages of *Helobdella*. The early cleavage stages are represented by stages 2, 5, and 7. The ISH staining pattern is shown for each gene in each of these stages, with the gene name labeled on the left side of each set of images.

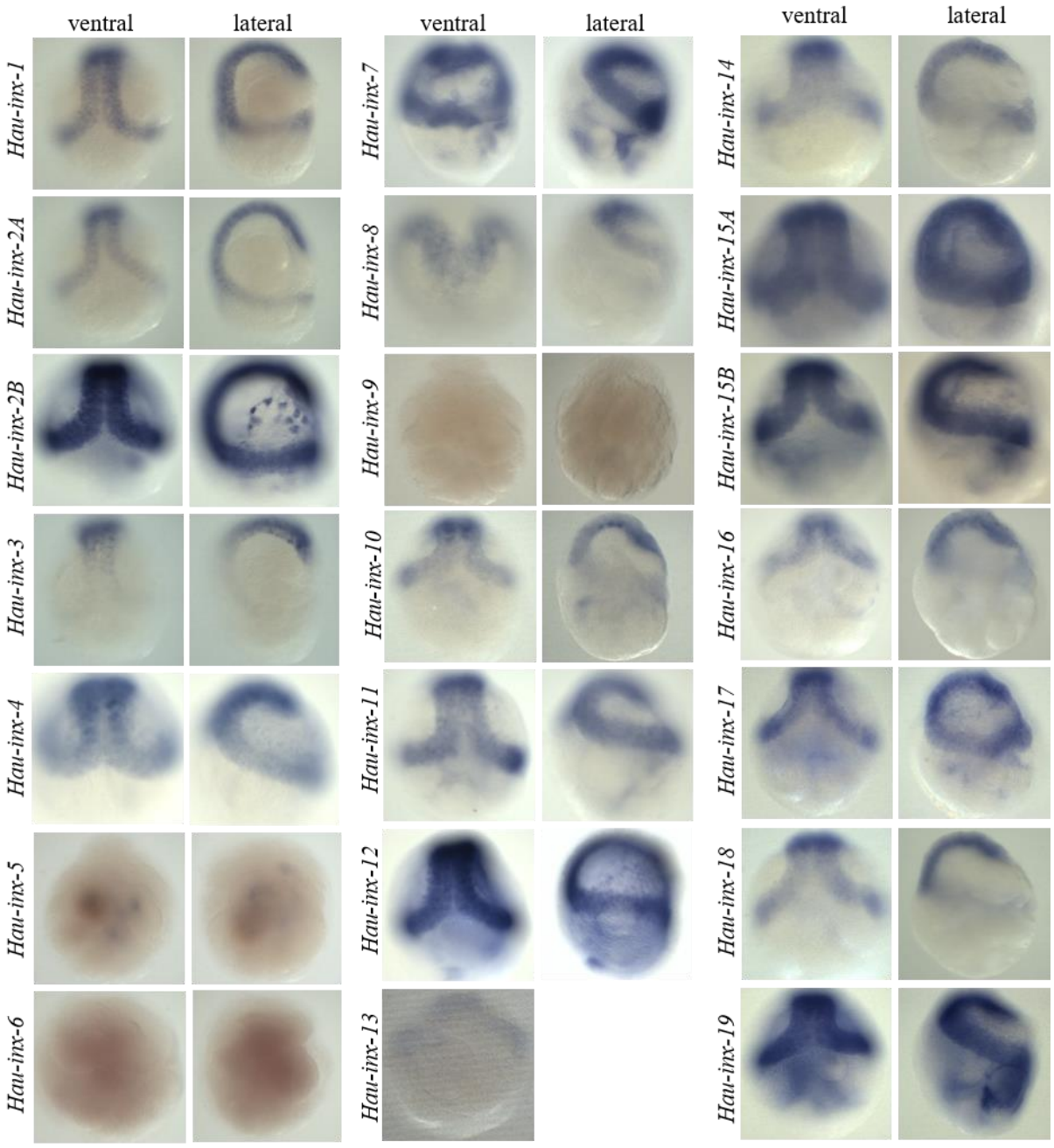


Figure 4.2. Expression of *inx* genes in stage mid 8. The expression of *innexin* genes during germinal band coalescence was tested at mid stage 8. The ISH staining pattern is shown for each gene in both ventral and lateral views, with the gene name labeled on the left side of each set of images.

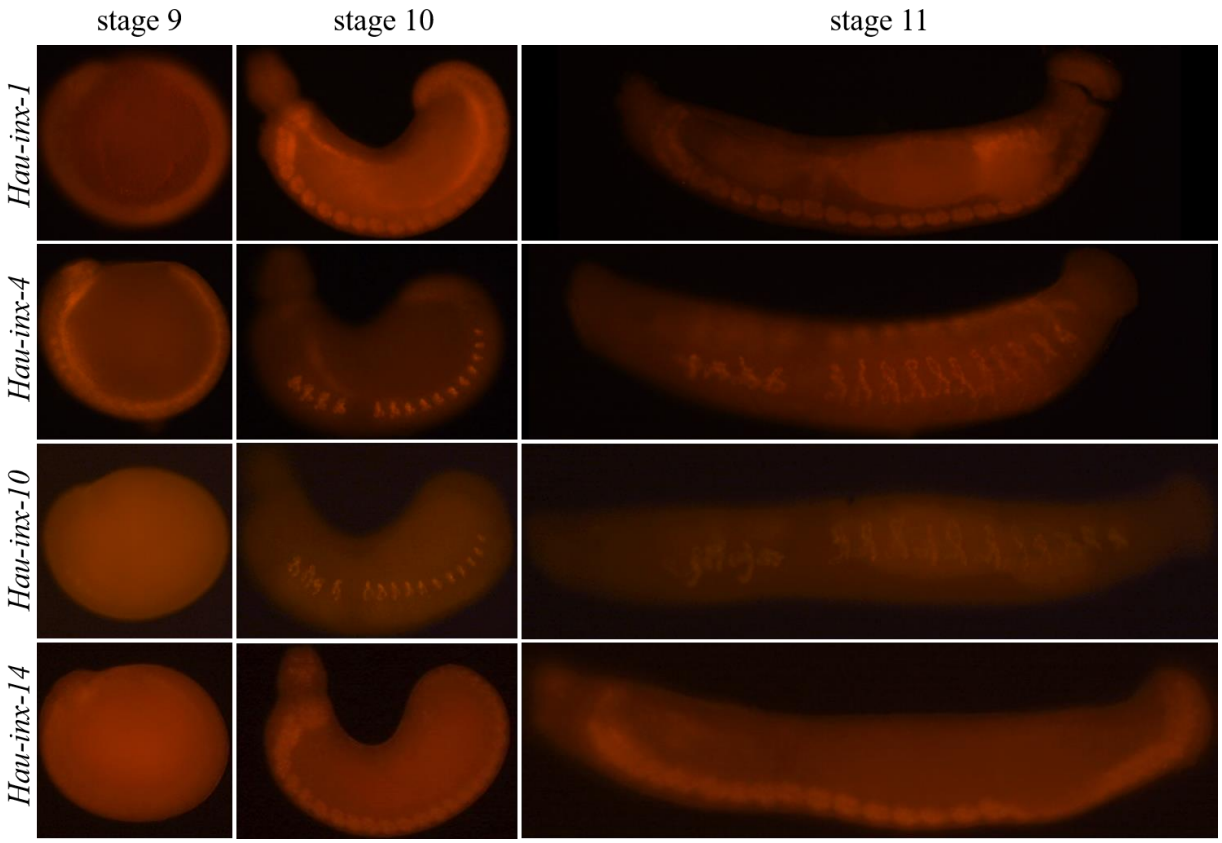


Figure 4.3. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with conserved expression patterns during late development. *Innexins 1, 4, 10,* and *14* are all present as a single copy in both *Helobdella* and *Hirudo*. The expression patterns appear to be highly conserved. *Hau-inx4, Hau-inx-10, Hve-inx-4,* and *Hve-inx10* are all highly expressed in the nephridia. *Hau-inx-1, Hau-inx-14, Hve-inx-1,* and *Hve-inx14* are all expressed pan-neuronally.

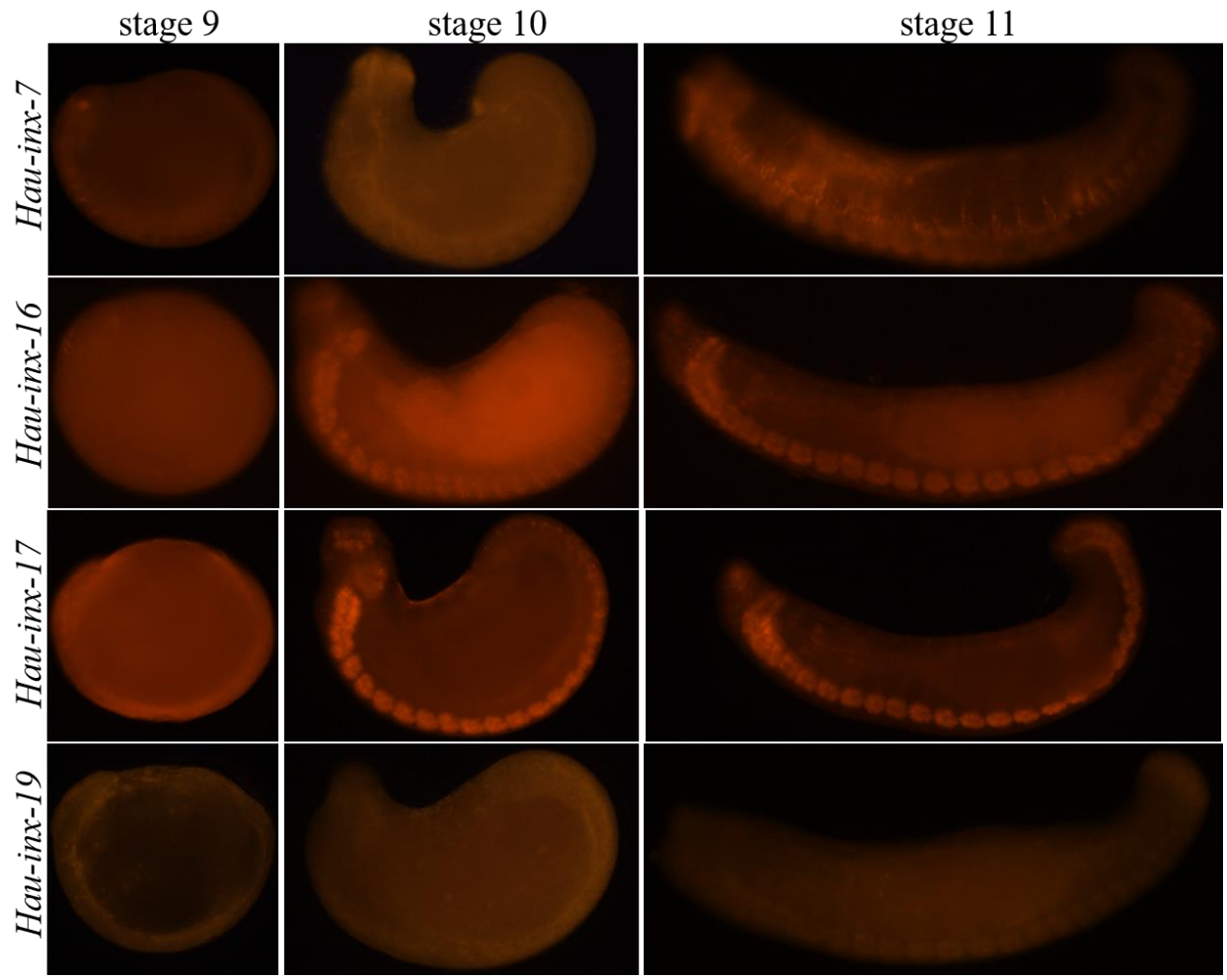


Figure 4.4. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with potentially conserved expression patterns during late development. *Innexins 7, 16, 17,* and *19* are all present as a single copy in both *Helobdella* and *Hirudo*. The expression patterns are potentially conserved, although there is not enough evidence to make that determination. *Hve-inx-16, Hve-inx17,* and *Hve-inx19* were detected in specific subsets of neurons, although the *Helobdella* homologues of these genes appeared to be expressed more broadly in the segmental ganglia. *Hve-inx-7* was detected in primordial germ cells and a lateral stripe. *Hau-inx-7* is expressed in lateral tissue, although we cannot determine if this is a homologous pattern.

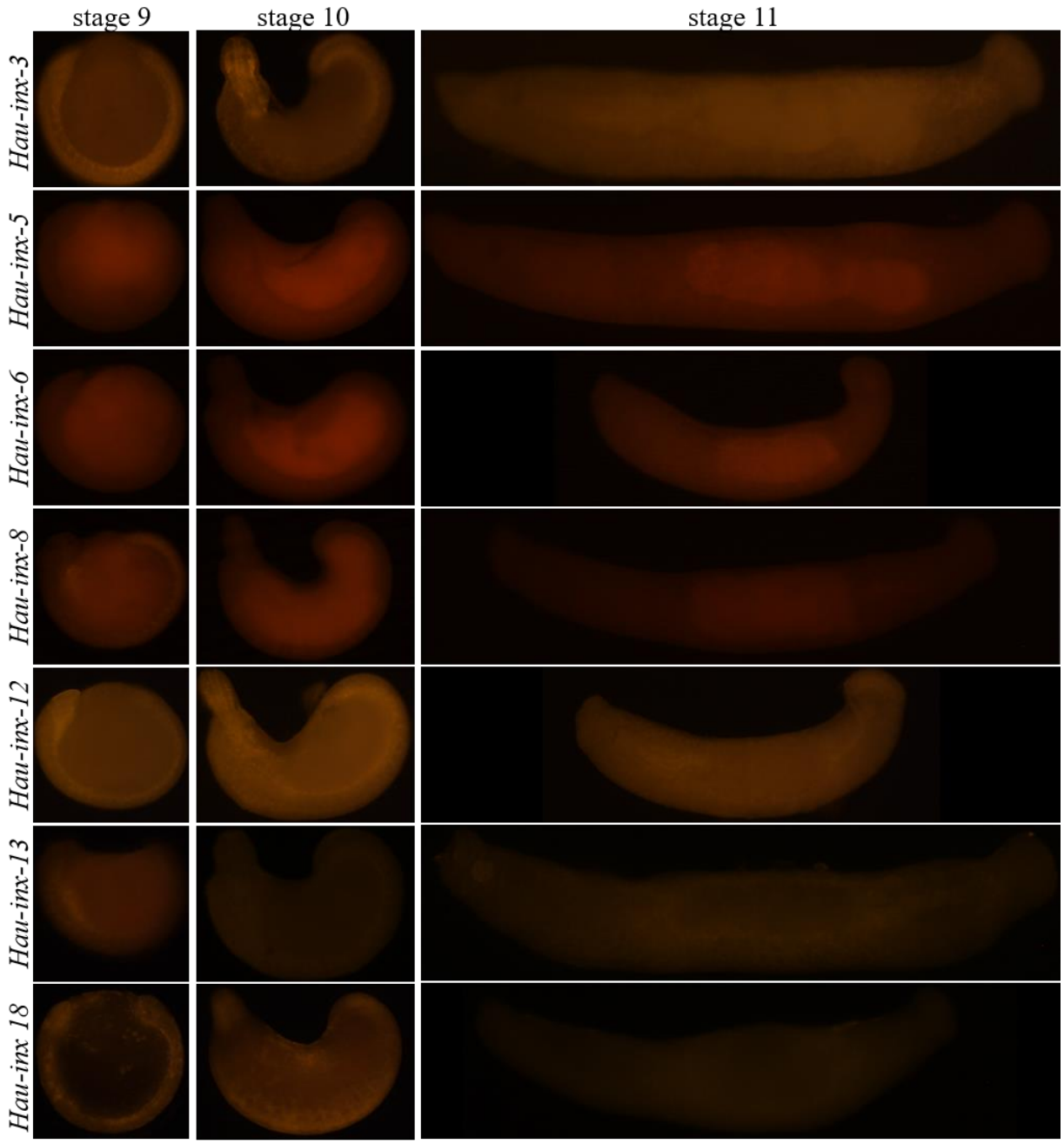


Figure 4.5. Expression of *innexins* present as single copy genes in *Helobdella* and *Hirudo* with no evidence for conserved expression patterns during late development. *Innexins 3, 5, 6, 8, 12, 13,* and *18* are all present as a single copy in both *Helobdella* and *Hirudo*. The expression patterns do not appear to be conserved between these two species. The majority of these genes were found to either not be expressed, or expressed at extremely low levels in the *Helobdella* embryo.

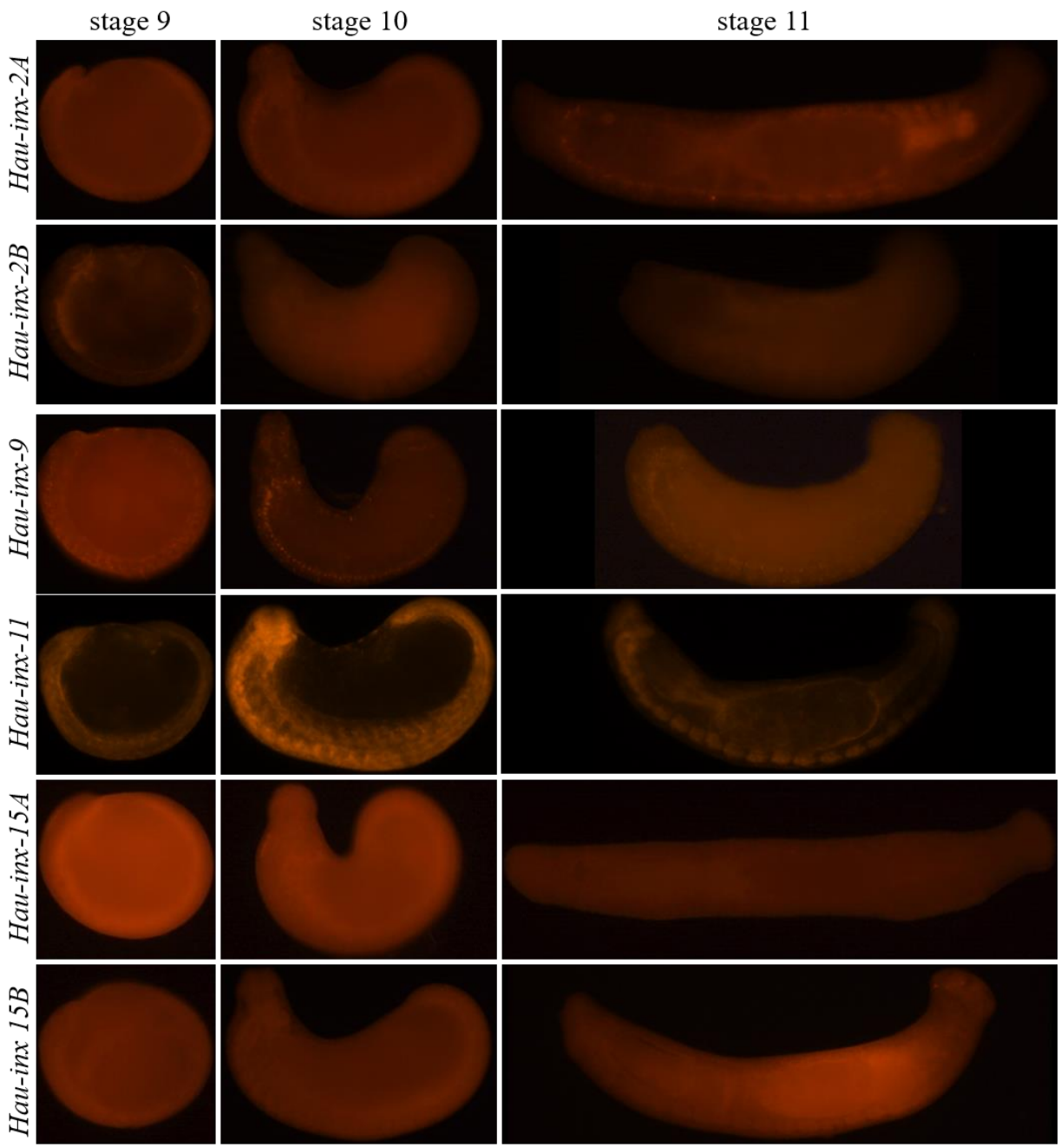


Figure 4.6. Late expression of the innexins present as multi-copy genes in either *Helobdella* or *Hirudo.* *Innexins 2* and *15* are duplicated in *Helobdella*, and *innexins 9* and *11* are duplicated in *Hirudo.* For most of these genes, some level of conservation can be observed between the expression pattern of at least one duplicate and the corresponding single copy in the other species. This excludes *innexin 15*, for which no conclusions could be made.

Chapter 5: Conclusions and Future Directions

This work has set the table for several future experiments that can shed more light on the topic of evolution of development. An important outcome of this work was the generation of leech-specific β-catenin antibodies. These can now be used to survey the subcellular expression of these proteins across leech development. I completed an initial survey of β-catenin expression throughout leech development using whole embryos and light microscopy. This survey showed that both β-catenins are expressed in the micromere cap during cleavage stages, and broadly throughout the germinal bands during stage 8, and at cell-cell contacts in general. B-catenin1 is expressed in the nerve cord ganglia in late stage embryos, whereas β-catenin2 may be expressed broadly in the epithelium at similar stages (Figures 5.1 and 5.2). However, in order to discern specifically in which lineages, cells, and cellular compartments these proteins are expressed in, more detailed studies are necessary. This includes lineage tracing, sectioning of embryos, and high-powered microscopy. This will help us make significant progress in understanding canonical Wnt signaling in the leech, as nuclear β-catenin is indicative of active canonical Wnt signaling. In order to complete these experiments, Western blots will be necessary to confirm a protein of the correct size is being targeted, and to assess differential protein levels. The efficacy of these antibodies can also be confirmed with heterologous expression of the proteins in mammalian cell culture. Because the epitopes are leech specific, they should not cross react with the mammalian protein, so this will demonstrate the specificity and functionality of the antibodies.

I would like to further explore the consequences of LiCl treatment on the N lineage. The results of the N-lineage marker analyses were unclear, which leaves many questions. It is possible that one or more lineages is being transformed. Given that the nominal N teloblast divides symmetrically, this may indicate that it is acting like the OPQ lineage (this proteloblast continues to make divisions that are more symmetric after the N teloblast is born). On the other hand, given that expression of *sfrp1/2/5c* is upregulated, other lineages may transform to N, or it is possible that after the symmetric division of the N teloblast, multiple N bandlets arise more frequently than I was able to assess. To address this I will combine lineage tracing with the *in situs* for the lineage markers. For example, if OPQ lineage tracer overlaps with an N lineage marker, this would suggest a transformation of O/P, Q, or both, to N. I believe I can address this question more easily by removing embryos from LiCl earlier than 24 hours. If I remove them as soon as the symmetric division has occurred, it is possible that there will be less germinal band defects, and the lineages will therefore be easier to analyze.

Given the divergent expression pattern of the duplicate *wnt16* genes, an interesting question to address is what regulatory changes have occurred. I noted that a conserved region in intron four is a potential enhancer region. To test this, one could clone this region into a vector with a basal promoter, upstream of GFP or some other fluorescent marker. If this region is indeed an enhancer, I expect it to drive expression of GFP in a pattern similar to what I have detected by ISH. Given the very specific expression of this gene, I believe it is plausible to expect the protein expression to be close to the RNA expression. If I can find the enhancer regions in both *wnt16a* and *wnt16b*, I can compare the sequences and arrangement in the genome and make hypotheses about what changes contributed to the regulatory divergence.

Taken together, this work contributes to our knowledge about conservation of developmental signaling pathways and other gene families. Features of the Wnt signaling

pathway that are highly conserved across Metazoa, including the posterior expression of *wnts*, do appear to hold true in the leech. However, a different family that seems to be rapidly evolving, the *innexin* gene family, appears to have several genes that do not share a homologous expression pattern even with a closely related species. The leech is an emerging developmental model organism, with more tools becoming available, and studies such as this one will contribute more to our understanding of diversity and the evolution of developmental mechanisms.

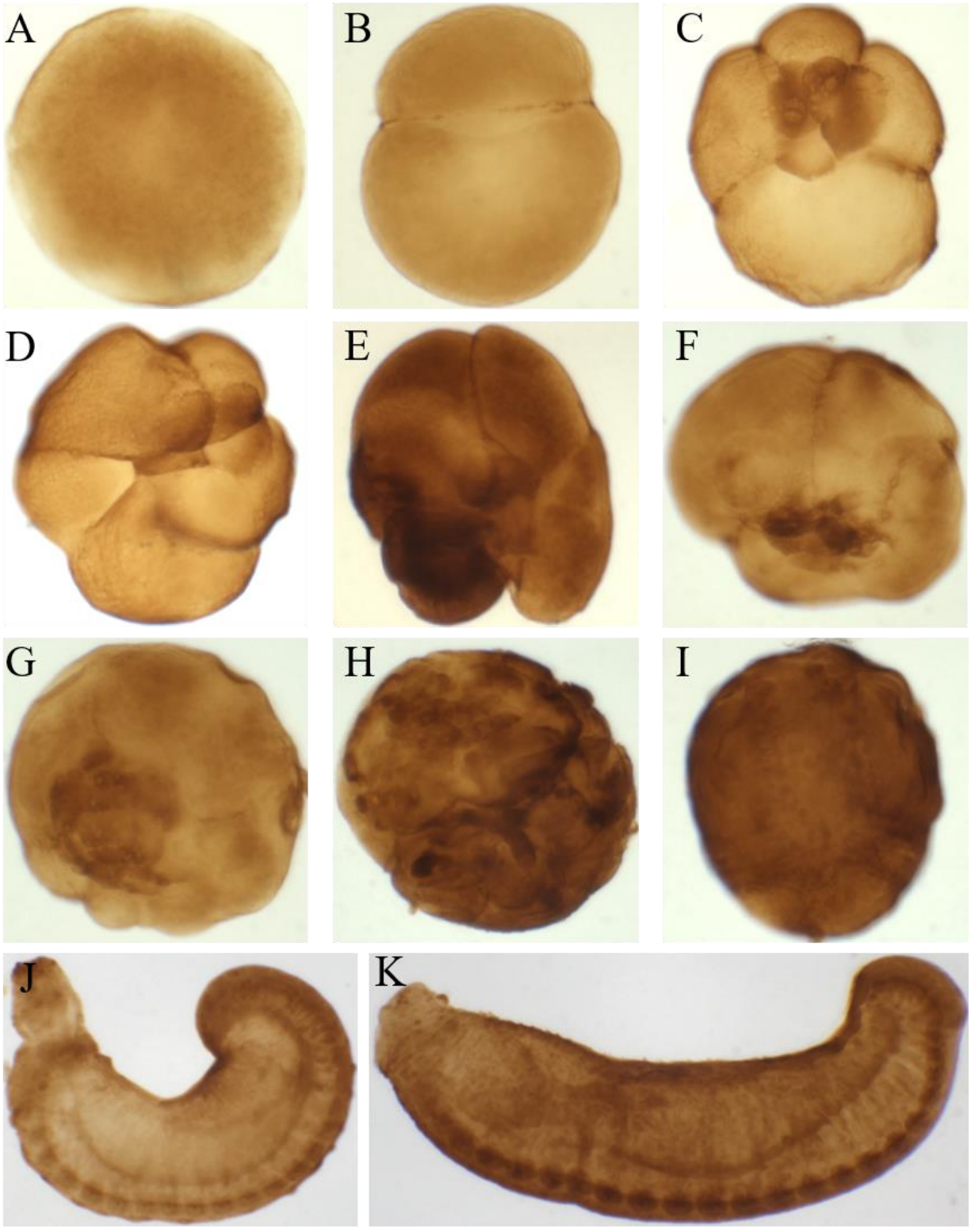

A
B
C
D
E
F
G
H
I
J
K

Figure 5.1. Expression of β-catenin1 throughout leech development. Nine stages of leech development are shown. Panel A is stage 1, B is stage 2, C is stage 4a, D is stage 4b, E is stage 5, F is stage 6b, G is stage 7, H is stage early 8, I is stage late 8, J is stage 10, and K is stage 11. Embryos were treated with primary antibody against leech β-catenin1, followed by an HRP-conjugated secondary antibody. They were stained with DAB for five minutes. Expression is primarily in the micromere cap in cleavage stages (C-G). It is expressed broadly throughout the embryo during stage 8 (H-I), and in the nerve cord ganglia in late stages (J-K). Panels A-H are animal pole views, panel I is a ventral view, and panels J-K are lateral views.

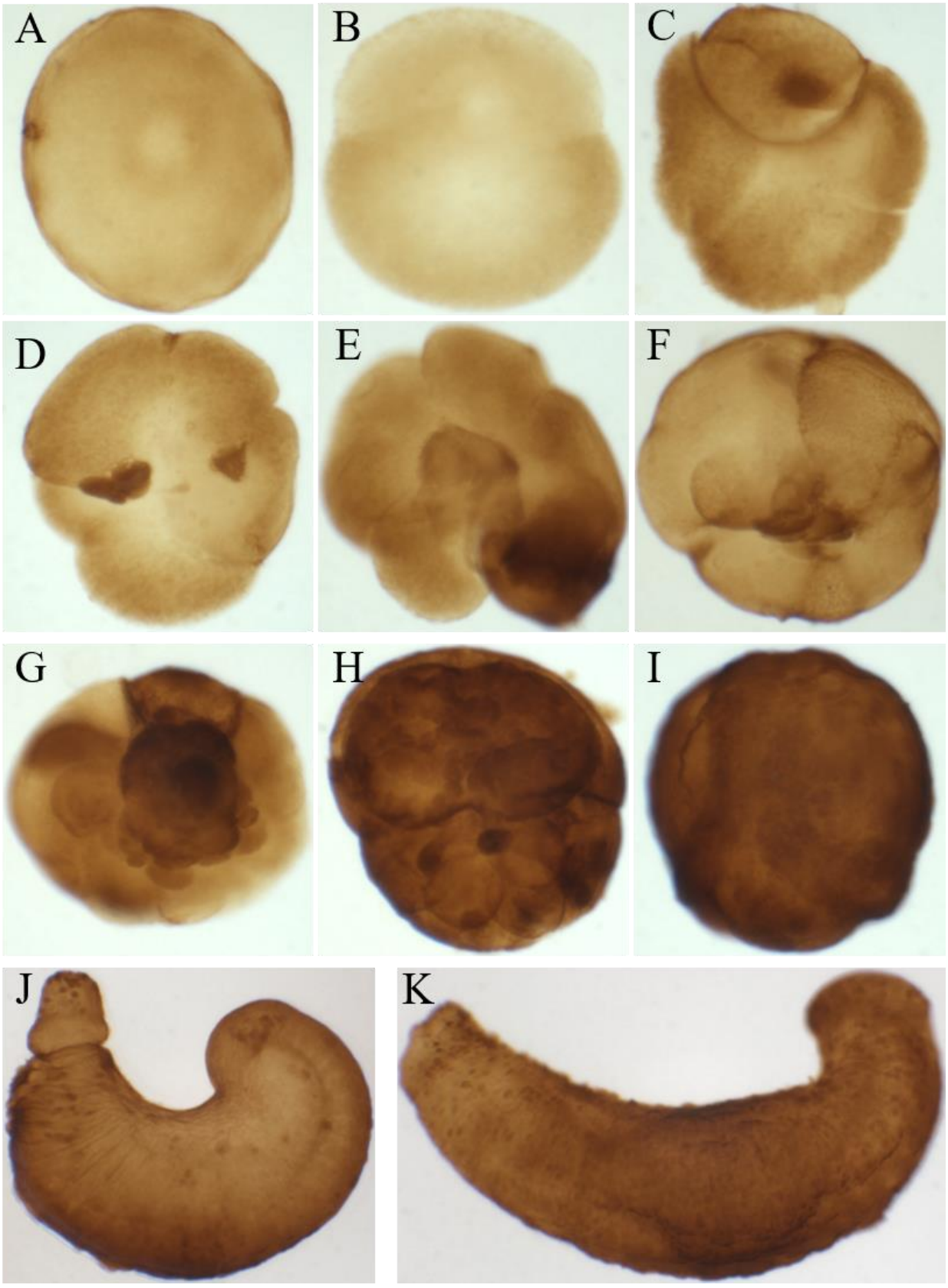

A
B
C
D
E
F
G
H
I
J
K

Figure 5.2. Expression of β-catenin2 throughout leech development. Nine stages of leech development are shown. Panel A is stage 1, B is stage 2, C is stage 4a, D is stage 4b, E is stage 5, F is stage 6b, G is stage 7, H is stage early 8, I is stage late 8, J is stage 10, and K is stage 11. Embryos were treated with primary antibody against leech β-catenin2, followed by an HRP-conjugated secondary antibody. They were stained with DAB for five minutes. Expression is in a macromere at stage 4a (Panel C). During cleavage stages the protein is primarily found in the micromere cap (D-G). It is expressed broadly throughout the embryo during stage 8 (H-I), and broadly throughout the epithelium in late stages (J-K). Panels A-H are animal pole views, panel I is a ventral view, and panels J-K are lateral views.

References

Aguinaldo, A.M.A, J.M. Turbeville, L.S. Linford, M.C. Rivera, J.R. Garey, R.A. Raff, and J.A. Lake. 1997. Evidence for a clade of nematodes, arthropods, and other moulting animals. *Nature*. 387:489-493.

Albertin, C.B., O. Simakov, T. Mitros, Z.Y. Wang, J.R. Pungor, E. Edsinger-Gonzales, S. Brenner, C.W. Ragsdale, and D.S. Rokhsar. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*. 524:220–224. doi:10.1038/nature14668.

Altun, Z.F., B. Chen, Z.-W. Wang, and D.H. Hall. 2009. High Resolution Map of *Caenorhabditis elegans* Gap Junction Proteins. *Dev. Dyn.* 238:1936–1950. doi:10.1002/dvdy.22025.

Astrow, S.H., and D.A. Weisblat. 1989. Factors specifying cell lineages in the leech. *Ciba Found. Symp.* 144:113-130, 150-155.

Balavoine, G. 2014. Segment formation in annelids: Patterns, processes and evolution. *Int. J. Dev. Biol.* 58:469-483. doi:10.1387/ijdb.140148gb.

Bauer, R., C. Lehmann, J. Martini, F. Eckardt, and M. Hoch. 2004. Gap junction channel protein Innexin 2 is essential for epithelial morphogenesis in the *Drosophila* embryo. *Mol. Biol. Cell.* 15:2992-3004.

Bauer, R., B. Löer, K. Ostrowski, J. Martini, A. Weimbs, H. Lechner, and M. Hoch. 2005. Intercellular communication: The *Drosophila* innexin multiprotein family of gap junction proteins. *Chem. Biol*. 12:515–526. doi:10.1016/j.chembiol.2005.02.013.

Bely, A.E. 2006. Distribution of segment regeneration ability in the Annelida. *Integr. Comp. Biol.* 46:508–518. doi:10.1093/icb/icj051.

Bissen, S.T., and D. A. Weisblat. 1989. The durations and compositions of cell cycles in embryos of the leech, *Helobdella triserialis*. *Development*. 106:105–118.

Blair, S.S. 2008. Segmentation in animals. *Curr. Biol.* 18:R991-R995. doi:10.1016/j.cub.2008.08.029.

Blauwkamp, T.A., S. Nigam, R. Ardehali, I.L. Weissman, and R. Nusse. 2012. Endogenous Wnt signaling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. *Nat. Commun.* 3:1070. doi:10.1038/ncomms2064.

Bolognesi, R., L. Farzana, T.D. Fischer, and S.J. Brown. 2008. Multiple Wnt Genes Are Required for Segmentation in the Short-Germ Embryo of Tribolium castaneum. *Curr. Biol.* 18:1624-1629. doi:10.1016/j.cub.2008.09.057.

Buechling, T., and M. Boutros. 2011. Wnt Signaling: Signaling at and Above the Receptor Level. *Curr. Top. Dev. Biol.* 97:21-53. doi:10.1016/B978-0-12-385975-4.00008-5.

Carrington, B., G.K. Varshney, S.M. Burgess, and R. Sood. 2015. CRISPR-STAT: an easy and reliable PCR-based method to evaluate target-specific sgRNA activity. *Nucleic Acids Res.* 43:e157. doi: 10.1093/nar/gkv802.

Carroll, D. 2012. A CRISPR approach to gene targeting. *Mol. Ther.* 20:1658-1660. doi:10.1038/mt.2012.171.

Catchen, J.M., J.S. Conery, and J.H. Postlethwait. 2008. Inferring ancestral gene order. *Methods Mol. Biol.* 452:365-383. doi: 10.1007/978-1-60327-159-2_17.

Cho, S.J., Y. Vallès, V.C. Giani, E.C. Seaver, and D.A. Weisblat. 2010. Evolutionary dynamics of the wnt gene family: A lophotrochozoan perspective. *Mol. Biol. Evol.* 27:1645–1658. doi:10.1093/molbev/msq052.

Couso, J.P. 2009. Segmentation, metamerism and the Cambrian explosion. *Int. J. Dev. Biol.* 53:1305-1316. doi:10.1387/ijdb.072425jc.

Dahl, G., and K.J. Muller. 2014. Innexin and pannexin channels and their signaling. *FEBS Lett*. 588:1396–1402. doi:10.1016/j.febslet.2014.03.007.

De Robertis, E.M. 2008. The molecular ancestry of segmentation mechanisms. *PNAS*. 105:16411-16412.

De Robertis, E.M. 2010. Wnt signaling in axial patterning and regeneration: Lessons from Planaria. *Sci. Signal*. 3. doi: 10.1126/scisignal.3127pe21.

Doench, J.G., E. Hartenian, D.B. Graham, Z. Tothova, M. Hedge, I. Smith, M. Sullender, B.L. Ebert, R.J. Xavier, and D.E. Root. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 32:1262-1267. doi:10.1038/nbt.3026.

Dykes, I.M., and E.R. Macagno. 2006. Molecular characterization and embryonic expression of innexins in the leech *Hirudo medicinalis*. *Dev. Genes Evol*. 216:185–197. doi:10.1007/s00427-005-0048-1.

Fernández, J., and G.S. Stent. 1982. Embryonic development of the hirudinid leech *Hirudo medicinalis*: structure, development and segmentation of the germinal plate. *J. Embryol. Exp. Morphol.* 72:71–96.

Fiedler, M., C. Mendoza-Topaz, T.J. Rutherford, J. Mieszczanek, and M. Bienz. 2011. Disheveled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating β-catenin. *Proc. Natl. Acad. Sci. U. S. A.* 108:1937–42. doi:10.1073/pnas.1017063108.

Ganfornina, M.D., D. Sánchez, M. Herrera, and M.J. Bastiani. 1999. Developmental expression and molecular characterization of two gap junction channel proteins expressed during embryogenesis in the grasshopper *Schistocerca americana*. *Dev. Genet*. 24:137–50.

Geetha-Loganathan, P., S. Nimmagadda, M. Scaal, R. Huang, and B. Christ. 2008. Wnt signaling in somite development. *Ann. Anat.* 190:208-222. doi:10.1016/j.aanat.2007.12.003.

Gibb, S., A. Zagorska, K. Melton, G. Tenin, I. Vacca, P. Trainor, M. Maroto, and J.K. Dale. 2009. Interfering with Wnt signaling alters the periodicity of the segmentation clock. *Dev. Biol.* 330:21-31. doi:10.1016/j.ydbio.2009.02.035.

Goldstein, B., H. Takeshita, K. Mizumoto, and H. Sawa. 2006. Wnt signals can function as positional cues in establishing cell polarity. *Dev. Cell*. 10:391-396. doi:10.1016/j.devcel.2005.12.016.

Graham, A., T. Butts, A. Lumsden, and C. Kiecker. 2014. What can vertebrates tell us about segmentation? *EvoDevo*. 5:24. doi:10.1186/2041-9139-5-24.

Habib, S.J., B-C. Chen, F-C. Tsai, K. Anastassiadis, T. Meyer, E. Betzig, and R. Nusse. 2013. A Localized Wnt Signal Orients Asymmetric Stem Cell Division in Vitro. *Science*. 339:1445-1448. doi: 10.1126/science.1231077.

Hannibal, R.L., and N.H. Patel. 2013. What is a segment? *EvoDevo*. 4:35. doi: 10.1186/2041-9139-4-35.

Hong, S.M., S.W. Kang, T.W. Goo, N.S. Kim, J.S. Lee, K.A. Kim, and S.K. Nho. 2008. Two gap junction channel (*innexin*) genes of the *Bombyx mori* and their expression. *J. Insect Physiol.* 54:180–191. doi:10.1016/j.jinsphys.2007.09.002.

Janssen, R., M. Le Gouar, M. Pechmann, F. Poulin, R. Bolognesi, E.E. Schwager, C. Hopfen, J.K. Colbourne, G.E. Budd, S.J. Brown, N.-M. Prpic, C. Kosiol, M. Vervoort, W. Gm Damen, G. Balavoine, and A.P. Mcgregor. 2010. Conservation, loss, and redeployment of Wnt ligands in protostomes: implications for understanding the evolution of segment formation. *BMC Evol. Biol.* 10:374. doi:10.1186/1471-2148-10-374.

Kalani, M.Y.S., S.H. Cheshier, B.J. Cord, S.R. Bababeygy, H. Vogel, I.L. Weissman, T.D. Palmer, and R. Nusse. 2008. Wnt-mediated self-renewal of neural stem/progenitor cells. *PNAS*. 105:16970-16975.

Kandarian, B., J. Sethi, A. Wu, M. Baker, N. Yazdani, E. Kym, A. Sanchez, L. Edsall, T. Gaasterland, and E. Macagno. 2012. The medicinal leech genome encodes 21 innexin genes: Different combinations are expressed by identified central neurons. *Dev. Genes Evol*. 222:29-44. doi:10.1007/s00427-011-0387-z.

Kiecker, C., and C. Niehrs. 2001. A morphogen gradient of Wnt/β-catenin signaling regulates anteroposterior neural patterning in *Xenopus*. *Development*. 128:4189-4201.

Kuo, D.H., and D.A. Weisblat. 2011. A new molecular logic for BMP-mediated dorsoventral patterning in the leech Helobdella. *Curr. Biol.* 21:1282–1288. doi:10.1016/j.cub.2011.06.024.

Landesman, Y., D.A. Goodenough, and D.L. Paul. 2000. Gap Junctional Communication in the Early *Xenopus* Embryo. *J. Cell. Biol*. 150:929–936.

Lo, C.W. 1999. Genes, gene knockouts, and mutations in the analysis of gap junctions. *Dev Genet.* 24:1–4. doi:10.1002/(sici)1520-6408(1999)24:1/2<1::aid-dvg1>3.0.co;2-u.

Lengfeld, T., H. Watanabe, O. Simakov, D. Lindgens, L. Gee, L. Law, H.A. Schmidt, S. Özbek, H. Bode, and T.W. Holstein. 2009. Multiple Wnts are involved in Hydra organizer formation and regeneration. *Dev. Biol.* 330:186–199. doi:10.1016/j.ydbio.2009.02.004.

Lynch, M., and J.S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. *Science*. 290:1151–5. doi:10.1126/science.290.5494.1151.

Lyuksyutova, A.I., C-C. Lu, N. Milanesio, L.A. King, N. Guo, Y. Wang, J. Nathans, M. Tessier-Lavigne, and Y. Zou. 2003. Anterior-Posterior Guidance of Commissural Axons by Wnt-Frizzled Signaling. *Science*. 302:1984-1988. doi: 10.1126/science.1089610.

Magadum, S., U. Banerjee, P. Murugan, D. Gangapur, and R. Ravikesavan. 2013. Gene duplication as a major force in evolution. *J. Genet.* 92:155–161. doi:10.1007/s12041-013-0212-8.

McGregor, A.P., M. Pechmann, E.E. Schwager, and W.G.M. Damen. 2009. An ancestral regulatory network for posterior development in arthropods. *Commun. Integr. Biol.* 2:174-176. doi:10.4161/cib.7710.

Niehrs, C., and S.P. Acebron. 2012. Focus Review Mitotic and mitogenic Wnt signaling. *EMBO J.* 31:2705–2713. doi:10.1038/emboj.2012.124.

Nusse, R., C. Fuerer, W. Ching, K. Harnish, C. Logan, A. Zeng, D. Ten Berge, and Y. Kalani. 2008. Wnt signaling and stem cell control. *Cold Spring Harb Symp Quant Biol*. 73:59-66. doi: 10.1101/sqb.2008.73.035.

Nusse, R. 2012. Wnt signaling. *Cold Spring Harb. Perspect. Biol.* 4:a011163. doi:10.1101/cshperspect.a011163.

Okita, K., and S. Yamanaka. 2006. Intracellular signaling pathways regulating pluripotency of embryonic stem cells. *Curr. Stem Cell Res. Ther.* 1:103–111. doi:10.2174/157488806775269061.

Oviedo, N.J., and M. Levin. 2007. *Smedinx-11* Is a Planarian Stem Cell Gap Junction Gene Required for Regeneration and Homeostasis. *Development*. 134:3121–3131. doi:10.1242/dev.006635.

Oviedo, N.J., and M. Levin. 2007. Gap Junctions Provide New Links in Left-Right Patterning. *Cell*. 129:645–647. doi:10.1016/j.cell.2007.05.005.

Petersen, C.P., and P.W. Reddien. 2009. Wnt Signaling and the Polarity of the Primary Body Axis. *Cell*. 139:1056–1068. doi:10.1016/j.cell.2009.11.035.

Phelan, P. 2005. Innexins: Members of an evolutionarily conserved family of gap-junction proteins. *Biochim. Biophys. Acta - Biomembr*. 1711:225–245. doi:10.1016/j.bbamem.2004.10.004.

Potenza, N., R. Del Gaudio, M.L. Chiusano, G.M.R. Russo, and G. Geraci. 2003. Specificity of Cellular Expression of *C. variopedatus* Polychaete Innexin in the Developing Embryo: Evolutionary Aspects of Innexins' Heterogeneous Gene Structures. *J. Mol. Evol*. 57:165–173. doi:10.1007/s00239-003-0023-2.

Richard, M., R. Bauer, G. Tavosanis, and M. Hoch. 2017. The gap junction protein Innexin3 is required for eye disc growth in Drosophila. *Dev. Biol*. 1–17. doi:10.1016/j.ydbio.2017.04.001.

Rose, A.B. 2008. Intron-mediated regulation of gene expression. *Curr. Top. Microbiol. Immunol.* 326:277-290.

Sedgwick, A.E., and C. D. Souza-Schorey. 2016. Wnt signaling in cell motility and invasion: Drawing parallels between development and cancer. *Cancers (Basel).* 8:1–15. doi:10.3390/cancers8090080.

Serras, F., and J.A.M. van den Biggelaar. 1987. Is a mosaic embryo also a mosaic of communication compartments? *Dev. Biol*. 120:132–138. doi:10.1016/0012-1606(87)90111-4.

Simakov, O., F. Marletaz, S.-J. Cho, E. Edsinger-Gonzales, P. Havlak, U. Hellsten, D.-H. Kuo, T. Larsson, J. Lv, D. Arendt, R. Savage, K. Osoegawa, P. de Jong, J. Grimwood, J.A. Chapman, H. Shapiro, A. Aerts, R.P. Otillar, A.Y. Terry, J.L. Boore, I. V. Grigoriev, D.R. Lindberg, E.C. Seaver, D.A. Weisblat, N.H. Putnam, and D.S. Rokhsar. 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature*. 493:526–531. doi:10.1038/nature11696.

Skeath, J.B. 1999. At the nexus between pattern formation and cell-type specification: The generation of individual neuroblast fates in the drosophila embryonic central nervous system. *BioEssays*. 21:922-931.

Sokol, S.Y. 2011. Maintaining embryonic stem cell pluripotency with Wnt signaling. *Development*. 138:4341–4350. doi:10.1242/dev.066209.

Stebbings, L.A., M.G. Todman, R. Phillips, C.E. Greer, J. Tam, P. Phelan, K. Jacobs, J.P. Bacon, and J.A. Davies. 2002. Gap junctions in *Drosophila*: Developmental expression of the entire *innexin* gene family. *Mech. Dev*. 113:197–205. doi:10.1016/S0925-4773(02)00025-4.

Steele, R.E., C.N. David, and U. Technau. 2011. A genomic view of 500 million years of cnidarian evolution. *Trends Genet.* 27:7-13. doi: 10.1016/j.tig.2010.10.002.

Swarup, S., and E.M. Verheyen. 2012. Wnt/wingless signaling in drosophila. *Cold Spring Harb. Perspect. Biol.* 4:a007930. doi:10.1101/cshperspect.a007930.

Takahashi, K., and S. Yamanaka. 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 126:663–676. doi:10.1016/j.cell.2006.07.024.

Thorpe, C.J., A. Schlesinger, J. Clayton Carter, and B. Bowerman. 1997. Wnt signaling polarizes an early C. elegans blastomere to distinguish endoderm from mesoderm. *Cell*. 90:695–705. doi:10.1016/S0092-8674(00)80530-9.

Todd, K.L., W.B. Kristan, and K.A. French. 2010. Gap junction expression is required for normal chemical synapse formation. *J. Neurosci*. 30:15277–85. doi:10.1523/JNEUROSCI.2331-10.2010.

Weisblat, D.A. 1985. Segmentation and Commitment in the Leech Embryo. *Cell*. 42:701–702. doi:0092-8674(85)90264-8 [pii].

Weisblat, D.A., and D.H. Kuo. 2014. Developmental biology of the leech *Helobdella*. *Int. J. Dev. Biol.* 58:429–443. doi:10.1387/ijdb.140132dw.

Weisblat, D.A., and D.H. Kuo. 2009. *Helobdella* (Leech): A model for developmental studies. *Cold Spring Harb. Protoc*. 4:1–6. doi:10.1101/pdb.emo121.

Weisblat, D.A., and M. Shankland. 1985. Cell Lineage and Segmentation in the Leech. *Philos. Trans. R. Soc. B Biol. Sci.* 312:39–56. doi:10.1098/rstb.1985.0176.

Weisblat, D.A., R.T. Sawyer, and G.S. Stent. 1978. Cell lineage analysis by intracellular injection of a tracer enzyme. *Science*. 202:1295-1298.

Weisblat, D.A., S.L. Zackson, S.S. Blair, and J.D. Young. 1980. Cell lineage analysis by intracellular injection of fluorescent tracers. *Science*. 209: 1538-1541.

Wong, R.C.B., M.F. Pera, and A. Pébay. 2008. Role of gap junctions in embryonic and somatic stem cells. *Stem Cell Rev*. 4:283–292. doi:10.1007/s12015-008-9038-9.

Zackson, S.L. 1984. Cell lineage, cell-cell interaction, and segment formation in the ectoderm of a glossiphoniid leech embryo. *Dev. Biol.* 104:143–160. doi:10.1016/0012-1606(84)90044-7.

Zanni, G., E. Di Martino, A. Omelyanenko, M. Andäng, U. Delle, K. Elmroth, and K. Blomgren. 2015. Lithium increases proliferation of hippocampal neural stem/ progenitor cells and rescues irradiation-induced cell cycle arrest in vitro. *Oncotarget*. 6:37083-37097.

Zou, Y. 2006. Navigating the anterior-posterior axis with Wnts. *Neuron*. 49: 787-789. doi: 10.1016/j.neuron.2006.03.004.

Appendix A: β-catenin sequences and epitopes

> β-catenin1, protein ID 108064, Cloned DNA sequence

ATGAATATCCCACCATCACCAGGCCACCAGAATTTCATAAACCTGGGAGAGATGCC
CCCAGCCATGGATAGCACTCAGCAGACATTATTGTGGCAGCAAAACCAATATATGG
CTGACTCCGGCATTCACTCAGCTGCAGGTACTCAGACCCCATCTCTCAGCAGTAAAT
TAGGGATCGACGGAGCGACTATTGGTGATAACATGTCTTCTCAGCAGTCGCAGCTAT
ACACGTTCGATGATTTCAGTGTTCCTCCTGTTTATGGACAGCCTCAGGTTGATGAAAT
TCAAGGCATCCAGTGTGGCATGCCTCTGTCGCAGGCCAACCCTCCTCCGGGTTATTT
TGATCAAATGGGTGCAGAGGGAATCTGCCCACCTAGCAACATGATGCACCCGGACC
CCACAATGGATGGCGAATGGCTTTCTGAGCCATCTCAGATGTTGAGAGAGGCCGTTG
CCAATCTGGTCAACTATCAAGACGTCTCCGACATGGCGCAGACTGCCATCCCTGAAC
TTGCGAATCTCCTGAACGGCGAGGATCATGTGGTTGTCGGCCAGGCAGCCATGATGG
TGCATCAGCTGTCGAAGAAGGAAGCCAGTAGACATGCCATAATCAACTCACCACTG
ATGGTCAAATCCCTAATTAATGGCATAAACACGACGGCCGATACTGAGACAATGAG
GTACCTCGCTGGCACCCTTCATAATCTTTCACACCACAGACAGGGACTCCTGGCTAT
TTTTCAGTCTGGAGGAATTCCTTCTCTTGTCAAGTTACTCAGCTCTAATGTAGAGTCG
GTACTTTTCTATGCGATAACCACTCTGCACAACCTGCTGCTGCATCAGGAGCGATCC
AAGATGGCGGTTCGTATGGCCGGGGGCCTCCAGAAGATGATTGCCCTACTCCAGAG
TACCAACATCAAATTCCTGGCCACCACTGCTGACTGCTTGCAACTGCTTGCTTACGG
AAATCAAGAAAGCAAATTAATCATCTTGTCGAGTAATGGTCCGACAGAGTTAGTGA
AGATCATGCACTCATACTCTTACGAGAAACTTCTCTGGACTACGTCACGGGTTCTTA
AAGTTCTGTCAGTCTGCCCACAGAACAAACCTGCTATCATTGCTGCTGGTGGTATGT
CAGCCTTGGCCATGCATTTGTGTCATCCTAGTCAGAGGTTGGTTCAAAATTGCCTGT
GGACACTCAGGAATCTGTCAGATACTGCTGCTCATGCTGAGAACCTGGAACCTCTCC
TGCAAGTCTTGGTGCAGATGCTGGCAGCCAACGACGCAAACATTGTGACCTGTGTGG
CGGGCATCCTGTCCAACTTAACTTGCAACAACCAGACGAACAAGATGGTCGTCTGTC
GGGCCGGGGGTATCGAGACGCTCATCAGGACCCTCGTGCAGGCTGGGGATAGAGAG
GACATAACTGAACCCACGATATGTGCCCTCCGTCACTTGACAAGTCGACACCCAGA
GGCTGAGATGGCCCAGAACTCCGTGCGCATGTTCAACGGTCTGCCACTTCTTATCAA
ACTCCTTCAGCCGCCGAGCCGCTGGCCTCTCATGAAGGCCATCATGGGGCTCATTAG
AAATCTCGCCCTGGCACCGGCCAATCAGGCCCCTCTCAGAGAACACGGGGCCATTC
CGAAAATTGCGCAGATTCTGGCCAGGGCTCATCAGGATATACAAAGGGCCGGAGGA
ATGGGAGTGCAGGGAAGCAACAGTAGTAGCAGCAGCAGCAGCGGCAGTGCTGTCA
GCCACCCACCAGTGTACGTGGATGGTGTGAAGATGGAGGAGGTGCTGGAGGGGGCC
ATTGGGACGCTACACATCATGTCCAGGGAGGTCAACAACAGGACTGTCATTCGATC
GATCGAAAACATCATGCCCATCCTCACACAGCTCCTCTACTACCCAAATGAGAACAT
TCAACGTATGGCTCATGGCGTGCTCTGCGAATTGAATGTCCAGCCGCAACCGTCGCC
GCTACAACAGCAGCAGCCACACATCCACGTCCAACAACCACAACAGCCACACATGC
ATCTCCAACAGCAGCAGCAGCAGCATACAATCGTTCCACAGCAACAGCAACAACAA
CAACAACATATTCATTCACATCATCAACATCAATTACAGCAGCATCAGCACAGGATG
CATAGCATCAATAGCGGCAATAATGGCATTAATAGTGGCGGTAACAGCAGCAATGG
AAGTAATAATACTAATTGGAGTAGTAATTGGTCTGGCAATGGAAACAGTGGCAACA
ATCCAATGACGATGATGATGAATCCATCTACTCCTGGTAAGATCAACAACAGCAATA
ACATCCTCAACGGCAGCAGCAGCGGTAGTAGTGGTGGTAACGCTTTCGAGTCGCAG

CAGCTGCAGCTACAGCAGCATCAACAACAGCAGCATCATAGGATGCTGGATATGAA
CACCTTCAACTATCTACTAA

> β-catenin1, protein ID 108064, inferred amino acid sequence
MNIPPSPGHQNFINLGEMPPAMDSTQQTLLWQQNQYMADSGIHSAAGTQTPSLSSKLGI
DGATIGDNMSSQQSQLYTFDDFSVPPVYGQPQVDEIQGIQCGMPLSQANPPPGYFDQMG
AEGICPPSNMMHPDPTMDGEWLSEPSQMLREAVANLVNYQDVSDMAQTAIPELANLLN
GEDHVVVGQAAMMVHQLSKKEASRHAIINSPLMVKSLINGINTTADTETMRYLAGTLH
NLSHHRQGLLAIFQSGGIPSLVKLLSSNVESVLFYAITTLHNLLLHQERSKMAVRMAGGL
QKMIALLQSTNIKFLATTADCLQLLAYGNQESKLIILSSNGPTELVKIMHSYSYEKLLWT
TSRVLKVLSVCPQNKPAIIAAGGMSALAMHLCHPSQRLVQNCLWTLRNLSDTAAHAEN
LEPLLQVLVQMLAANDANIVTCVAGILSNLTCNNQTNKMVVCRAGGIETLIRTLVQAGD
REDITEPTICALRHLTSRHPEAEMAQNSVRMFNGLPLLIKLLQPPSRWPLMKAIMGLIRN
LALAPANQAPLREHGAIPKIAQILARAHQDIQRAGGMGVQGSNSSSSSSSGSAVSHPPVY
VDGVKMEEVLEGAIGTLHIMSREVNNRTVIRSIENIMPILTQLLYYPNENIQRMAHGVLC
ELNVQPQPSPLQQQQPHIHVQQPQQPHMHLQQQQQQHTIVPQQQQQQQQHIHSHHQHQ
LQQHQHRMHSINSGNNGINSGGNSSNGSNNTNWSSNWSGNGNSGNNPMTMMMNPSTP
GKINNSNNILNGSSSGSSGGNAFESQQLQLQQHQQQQHHRMLDMNTFNYLL

> β-catenin2, protein ID 189420, cloned DNA sequence
ATGGAAAGCTACCCAAGTCACACAGCTCCTTATTCAATGGACCACAATGCAGTTTAC
ATGCAGGAACAAATTAATGACATGAACCAGCAACTGGCGCAGACTCGCTCACAGAG
GGTTCGTGCGGCCATGTTCCCCGAGACATTCAACGACGAGGAACTTCAAATTCCTTC
AACGCAGAAACACTTCGGCCAGATGACGACTGTGCAGAGACTGGCTGAGCCATCTC
GGATGCTGAAGCATGCTGTCGTCAATCTCATCAACTATCAGGAGGATGCAGATGTGG
CGGTGAGAGCGATTCCTGAGTTGATATCACTGCTGAATGATGAGGATCAGATTGTAG
TTGGCCACGCTGCTCTCATGAGTCACCAACTTTCGAAGAAGGAGGCCAGCAGACAC
GCCCTCCTCAACTCTCCACAGATCATCGCTGCCCTCATCAATGCACTCAACAAAGCC
AAAGACCCAGAAACCATTAGGTACTTGACGGGGGCATTGCACAACCTGTCCCACCA
CCAGCAAGGACTGTTGTACATCTTCAAGTCTGGTGGCATACCTGCTCTCATCAGGCT
GCTTAGTTCCCCCATTGAATCGATTCTCTTCTATGCGATGACTACCCTACACAATTTG
TTGCTACACCAGGAAGGAGCTAAGATGGCGGTGCACACCGCAGGGGGGCTGCAAAA
AATGGTGGCCCTCTTGACGTATGATGACCCAAAGTTCCTGGCTATAACCACTGACTG
CCTGCAACTTCTTGCTTATGGAAATCAGGAATGCAAATTGATCATCCTGGCGAGTGG
AGGTCCAGACAATCTGGTGCGCATCATGGAGAGGTACAACTACGAGAAGTTGCTCT
GGACCACATCCAGGGTTCTCAAGGTTCTCTCCGTCTGTCCCAGCAACAAGATGGCCC
TTGTTGAAGCCGGAGGCATACACGTCCTCTCCAACCTGCTCGTGAACCCAAGTACTC
GACTCGTTCACAACTGCATGTGGACCATCCGGAACCTCTCTGACTCTGCCACCAAAC
TGGAAGGAATGGAGCCTGTTCTGGAAGTATGCGTGCGAATGTTGCAAATGGACGAC
TTGGATATGATCATTTGCTCCTGCGGCATCCTGTCCAACCTCACCTGCAACAACCAC
GCCAACAAGTCGTACGTGTATCAAATATCTGGAGTGGAGGCACTTGTGGCCACAGTC
GTCAAGGCTGGAGACAGGGAGGACATCACCGAACCATCCGGGTGTGCACTGCGTCA
CCTGACCAGCCGACACGCGCATGCTGAATTATCTCAAAACCTGATTCGAGAAGAAG
GAGGTCTCATGCCAATCATCAACCTCCTCCACCCGCCATCCCACTGGCCTCTCCTCA
AGGCAACCGTCAGCTTGGTCAGGAATTTGGCACTTTGTGAGAAGAACAATCCAGAA

CTGCGGGGATTGGGAGCAGTCCCGAAGATCATCCAGCTGCTTGTGAGGTCCCATCAG
GAGGTTCAGAAGGCTGTACTAAGCAACTCATCCGATGGAATAATGATGGATGGAGT
GAGAATGGATGAGATGGTGGAGACGTGTGTTGGCGCTCTGCACATTCTGTCACGCG
ACGTTCACAACAGGACAATCATCAAGGAACACAACTGCATTCCTCTCTTTGTGCAGT
TACTCTACTCCCCCCACGAGAACACCCAGCGGGTGACGACGGGAGTGCTGTGCGAG
TTGGCAGCGGAGAAGGATGCGGTTGAGATGACTGAGCAGGAAGGTGCCACGGCCCC
CCTCACCGAGTTGCTGCACAGCCACAATGAAGCCATCGCGACGTACGCAGCGGCAA
TATTGTATCGCATGAGTGAAGATAAGTCGGCAGATTACAAGAAGAGACTGTCCATT
GAACTGACCAACTCTCTGGTCAGAGGCGATCCAAACACCTGGAACAACCAGCTGAC
TCTGGTTGAAGAGATGCCAAACGATGAAATGTACCAACCAAACAACAACCAGGGCA
ACAACATTTTTCCTTCGTCTAATACTCCAGATTTTAATTCTCTACAACCACTGCACCA
AGGCCTCTACCTTCACCCCAACCAAGCCAACCAATCAAATCAGTTCTATGCGGGCTC
AGAATCAGGCAGCCGACACAACCAGCCAATGAGATTTCCTGATCTAACAGAAAATG
TTTCCGATCCCTATGCGGTGAGGTATTCGAACAACACTGGCTCGCTCGTCAGCCAAT
CAGGGAGAGCTTTGCAACTGGATGGTGGACCATGGTTCAATACTGATGTGTGA

> β-catenin2, protein ID 189420, inferred amino acid sequence
MESYPSHTAPYSMDHNAVYMQEQINDMNQQLAQTRSQRVRAAMFPETFNDEELQIPST
QKHFGQMTTVQRLAEPSRMLKHAVVNLINYQEDADVAVRAIPELISLLNDEDQIVVGH
AALMSHQLSKKEASRHALLNSPQIIAALINALNKAKDPETIRYLTGALHNLSHHQQGLLY
IFKSGGIPALIRLLSSPIESILFYAMTTLHNLLLHQEGAKMAVHTAGGLQKMVALLTYDD
PKFLAITTDCLQLLAYGNQECKLIILASGGPDNLVRIMERYNYEKLLWTTSRVLKVLSVC
PSNKMALVEAGGIHVLSNLLVNPSTRLVHNCMWTIRNLSDSATKLEGMEPVLEVCVRM
LQMDDLDMIICSCGILSNLTCNNHANKSYVYQISGVEALVATVVKAGDREDITEPSGCA
LRHLTSRHAHAELSQNLIREEGGLMPIINLLHPPSHWPLLKATVSLVRNLALCEKNNPEL
RGLGAVPKIIQLLVRSHQEVQKAVLSNSSDGIMMDGVRMDEMVETCVGALHILSRDVH
NRTIIKEHNCIPLFVQLLYSPHENTQRVTTGVLCELAAEKDAVEMTEQEGATAPLTELLH
SHNEAIATYAAAILYRMSEDKSADYKKRLSIELTNSLVRGDPNTWNNQLTLVEEMPNDE
MYQPNNNQGNNIFPSSNTPDFNSLQPLHQGLYLHPNQANQSNQFYAGSESGSRHNQPM
RFPDLTENVSDPYAVRYSNNTGSLVSQSGRALQLDGGPWFNTDV

Epitopes: Highlighted regions are the sequences selected to generate peptide antibodies. Sequences were chosen so antibodies would not cross-react to other protein.

Appendix B: ***Fz1/2/7b*** **sequence with primers and sgRNAs**

>*fz1/2/7b* nucleotide sequence from cDNA, JGI protein ID 191092
ATGGCACTTTTTAAGCTTTTTTCGTCTCAGAAACTGATATTTTTAATATTTGTCGTCG
GAGATGTCACGAATTTCGTTACATCAGTTTCGCCGATGGACACCCATCATTATCATC
ATTCAAAATTGTCCGACACTCATCTGCCCACTCATGGCAAGTGCGAACCAATCACAA
TAGAGCAGTGCAAAAATTTAGAATACAACGAAACAATAATGCCAAACGTTTTGAAT
CAAATGAAGCAGTCAGATGCCAAAGAATCCATAACACAGTACAGCAGACTGATTCA
GAGCAGATGCAGTCAGTACATTCAAATTTTCCTTTGTTCAGTTTTTGTACCTGTGTGT
ACACAGTTAGAAACTGCACTTCCTCCGTGCAGATCGCTTTGTTTGAATTCAAAATTA
GATTGCGAATCAGTAATGAAAGCTGCTGGTTATGGGTGGCCAGAGGAGTTAGATTG
TAAAAAATTTCCTGACGATAACGTCAATGTACTCTGTGTTAGTGTGAATGGATCAGA
AAAGGATTCTAGTAAAGAAGGGCGTAAAAATAAAAATTGGCCAAATAAGGGATCA
AGAGGGGGTAATGATGGGGCTAGACAAGGTCCAAAGGAATTGTACTGTCCTACTTT
TATGAGAGCAGATTCAAAATCTGACTACAGATTGTTCATTGGTGATTCAGTGATGAA
GAGTTGTGGTATGCCATGCGGTGTCAAGGGAGATATGTTCTCACTTTCTGATAAAAG
AAAACTTGTCAGGTTGATAATTTGCATTGTGGCTTGCATCTGTCTGCTCAGTTCTCTT
TTTACATTTTTAACTTTCCTCGTTGACACGAGAAGGTTTCGTTATCCAGAAAGGCCGA
TAGTGTTCATTTCTGCTTGCTACTGTGTTATAGCCATTGCATATATTGTAGCATATGC
ACTGGACAATAAAATATCATGTTCAGAATTTGTTGAGCATTTTGATGATGACATTTC
ATTTCAGTCTGAACTGATCACACAGGGTACAAAGAGAGAAGGATGTACAATCATAT
TCGTCATGTTATACTTCTCTACTTTGGCAAGTTGCATCTGGTGGGTCATATTGACATT
AACCTGGTTTTTATCTGCAGGTTTGAAATGGGGACAAGAAGCTGTAGAAAGCAGAT
CACAATATTTTCATCTCTTGGCCTGGGCAGTCCCTGGAGTGATGACCATTGTGCTGCT
GGCCATGGGTCAAGTTGATGGTGATGCACTGACTGGTATTTGCAACACAGGCCTCAC
CAATAAAGAAATTGCAACAGCATTTGTTGCAGGTCCTCTCATTATTTTGCTGTCCGTT
GGAATATTCTTCCTCCTTGCTGGATTCATCTCTCTTTGCAGAGTTAGAAATGCTATGA
AGTTCGATGGTAACAGGACAGACAAGTTGGAGAAACTGATGGTGAGAATCGGTGTA
TACGCCCTACTTTATGTTGTACCAGTTGCATTCGTCGTTGGTTGTGTCTTCTACGAGC
ATGCCTGGAGGACCGAGTGGGTGAAAGGCTGGTACCATCAGACATGTTCAGACCTC
CTCATGTCCGGGGAGGGCCATTTTGGAAATGTTTTTGAACTCGACTTTGAGCCAAAC
TACCTGCGAAGTAATGTAATCTTGCAGCAACAGCAGTGGGCCTGTGGTAACTACCCA
TCCCACTTTGCTAATTATGCTGGTAGTAAACCCGATTTGACCATCTTTGTACTGAAAC
ATCTCATGACACTGATAGTAGGAGTGGCCTGTGGATTCTGGGTGTGTTCTGGAAAAA
CTGCAGAATCATGGTCAAAGTTTTTCTGTCGCTCGAATGAGGCCAGCAAGAAAAATC
ATCAGAAAGGTCAGCCAACACTCTTGTAA

>*fz1/2/7b* inferred amino acid sequence
MALFKLFSSQKLIFLIFVVGDVTNFVTSVSPMDTHHYHHSKLSDTHLPTHGKCEPITIEQ
CKNLEYNETIMPNVLNQMKQSDAKESITQYSRLIQSRCSQYIQIFLCSVFVPVCTQLETA
LPPCRSLCLNSKLDCESVMKAAGYGWPEELDCKKFPDDNVNVLCVSVNGSEKDSSKEG
RKNKNWPNKGSRGGNDGARQGPKELYCPTFMRADSKSDYRLFIGDSVMKSCGMPCGV
KGDMFSLSDKRKLVRLIICIVACICLLSSLFTFLTFLVDTRRFRYPERPIVFISACYCVIAIA
YIVAYALDNKISCSEFVEHFDDDISFQSELITQGTKREGCTIIFVMLYFSTLASCIWWVILT
LTWFLSAGLKWGQEAVESRSQYFHLLAWAVPGVMTIVLLAMGQVDGDALTGICNTGL
TNKEIATAFVAGPLIILLSVGIFFLLAGFISLCRVRNAMKFDGNRTDKLEKLMVRIGVYAL

LYVVPVAFVVGCVFYEHAWRTEWVKGWYHQTCSDLLMSGEGHFGNVFELDFEPNYLR
SNVILQQQQWACGNYPSHFANYAGSKPDLTIFVLKHLMTLIVGVACGFWVCSGKTAES
WSKFFCRSNEASKKNHQKGQPTLL-

Primers used for genotyping

Target sequences for sgRNAs

Appendix C: *Wnt16* sequences

>*wnt16a* JGI genome transcript model, protein ID 121846
CATAATTCACAATCGCCACAGCCACTACAACTCAACCAGACGACTACATGCCAATTA
CTACAACAACATGGCTACATTTCGTCGGAGCAACTTTCGCTATGCCTCGGCGACCAT
TTTTCAGTTCTGGCAGCAGTGTCTGGGCGACGACTCGGCCTAAGCGAGTGCTTTCAT
CAATTTCGTTACGAACGCTGGAACTGTTCGGCAACGGTTGCTCGCAACATCGAATCA
GGAGGCAAGCAGCGGCGCACAAAAGAGCAAGCATTTATGCAGGCTGTTATGAGCGC
TGGAGTTGTCTTCACCGTCACTGAAGCTTGCAGTTCTGGAAAACTTCTCAGATGTTCC
TGTCTTTCTATCCCATCATCATCCTCATCATCCTCATCATCCAACAACAACAACAATC
GCTTCAACGACGACACATGGAAGTGGGGTGGTTGTAGCGACGACATCGACTACGGA
TTAAGTTACGCCAAACTTTTCACCGACAAACCAATCAAAAAACAACTGCTGAAAAA
TGGAATTTTTAGAATGTCGGATTTGAAAGGATTAGTTGATTTACATAATAACGAAGT
TGGCAGACAGATTCTCTCTTCCTTAATGAAAATAAAATGCCGGTGCCACGGCGTTTC
TGGCCTCTGTGGCGTCCGGACGTGTTGGAAGTCGCTGCCGACGTTTCGCGAAGTCGG
CGATGCGTTGAAGAACAAGTACGAGACTTCAATCGAAATTTCGAGACCGTCTCAAC
ATCTTCTAAAGCGAGAAAAACGCCGACGCCGTCGCGAGCCCATCTCCTCGGCCGAC
TTGATTTTCCTAAAAAAATCTCCGGACTACTGCAAGCAAAACCTGAAGAAGGGAAT
CCCAGGAACTAGGGGACGACTCTGCAACAAAAACTCCACGGGCCCTGATGGGTGTG
ACTATCTTTGCTGTGGGCGGGGCTTCAACCTTGCTGAAACGCGATTGGTCGAGAGGT
GCCACTGCAAATTTATCTGGTGCTGCTCTGTTCAGTGTAAGATGTGTGAACGTGTTG
AGATTAAATACACGTGTAAG

>*wnt16a* cloned cDNA sequence for ISH
TGGAAGTGGGGTGGTTGTAGCGACGACATCGACTACGGATTAAGTTACGCCAAACT
TTTCACCGACAAACCAATCAAAAAACAACTGCTGAAAAATGGAATTTTTAGAATGTC
GGATTTGAAAGGATTAGTTGATTTACATAATAACGAAGTTGGCAGACAGATTCTCTC
TTCCTTAATGAAAATAAAATGCCGGTGCCACGGCGTTTCTGGCCTCTGTGGCGTCCG
GACGTGTTGGAAGTCGCTGCCGACGTTTCGCGAAGTCGGCGATGCGTTGAAGAACA
AGTACGAGACTTCAATCGAAATTTCGAGACCGTCTCAACATCTTCTAAAGCGAGAAA
AACGCCGACGCCGTCGCGAGCCCATCTCCTCGGCCGACTTGATTTTCCTAAAAAAAT
CTCCGGACTACTGCAAGCAAAACCTGAAGAAGGGAATCCCAGGAACTAGGGGACG
ACTCTGCAACAAAAACTCCACGGGCCCTGATGGGTGTGACTATCTTTGCTGTGGGCG
GGGCTTCAACCTTGCTGAAACGCGATTGGTCGAGAGGTGCCACTGCAAATTTATCTG
GTGCTGCTCTGTTCAGTGTAAGATGTGTGAACGTGTTGAGATTAAATACACGTGTAA
G

>*wnt16a* 5'RACE sequence
TCTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGATACATAGAGGAATAT
ATACAAGGAATATTTGGTGTTTTAATTAATTACTTATTCGCGTTAAATCTGATTACAA
TCCCCATGGCTAGATGAATATTGAGATGTATTGATTGATAGTCAGTCTGTGATCGTT
CACTCGGATAGATATATATCATAACCAATGGATGATACCACCATGATATCATGATAC
CTATCTGAAATCATGTAAAAAAAAGGATGACCCGACCAAGATACAACAACAACAAC
AACAACAACAACAATAACAACAACAACAACAACAACAACAGCATCATCACCAACG
CCTTCAACATCATCGTCGTCGTCGTTGTCGTCACTTCGTTATTACTGCTTGGCGGAAG
TTTCTGCTCAAATGAAAGAAACATTCCGACACGTAGCTTGACACGTAAAGAAGCCAT

ATACAAGAGCAGCAGCAGTAGCAGAAGTAGCGACAATAACTACAGCAACAAAGAT
GAAAACATCAAACACTACAAATACAACAAGAAGTTTAATAACATCAACAACAACAA
CAGCAAACAACTCAAAACAAACCAACTCCAATACAACTACTATTATTACGATGATTA
TGAATTCAACAACAACAAGCAGCAACAACTGCAACAACCACAACAACTGCAACAAC
GTTGGCAACAACAACTAGAGCAACCGCGGCAGCAACAACAACAGTTGCTGCAACAA
CAGCAACATCAAAATGATCAGCCTCACCATGACGAACGCATGTTTCTCTATCTAGAG
ACGAAANGAATTTTTGATGAGTATCAAAANCTTCAACAAATTGAAAGATTAAAGCA
ACAACAACAACAGCTGCTGCAACAACAACAGCAACAACAGCAGCCACAACGGCAG
CAGCCACAACGGCAGCAGCCACAACGGCAGCAACCACAGCAGCAGCAGCAGCAAA
ATAACCAGAAATTTTTAAGTTTGTCTCAAGAAAACAAGCTTAAACATCAAAAACAA
CAACCCNAACAACAACNTCAACAACAACACCAACAACAGCAACTTCAGCAACAAC

>*wnt16b* JGI genome transcript sequence, protein ID 79030
GGATCGGTGTTGGGACTGTCGGAATGCAGGCAGCAGTTTGAAGAGGAGAGGTGGAA
TTGTCCCATCAAAAAGCAACCATTCATCATACCCGCCCCACAAAAACCATCCAATCA
ATTGATACCAATCATCAAACTAACCGTCCAATCAGGAACGAAGGAGATTGCTTTCAT
CCACGCCATCACGTCTGCCTCACTCGCTCACTCCATCACTTCATCATGTAGCGCTGGC
CTCCTTCTAGAATGCTCTTGCGACAGGTCCCTGCAGTCGATCGTGAGTACGGACAGC
AGCTGGCGCTGGGGTGGCTGCAGCGATAACGTCCAATACGGCATCAAGTATTCGAA
GATCATCACAGATGGTGACGGAAAAAAGGCAGCTCTGATGGAACGGGTTCGATCGC
TCGTTCATCTGCATAATAATGTGTCGGAAGGAAGACCTTGCACTCCCTGATGACCCA
CAAATGTCGTTGCCATGGTGTCTCGGGGTCGTGCGCGGTTCGATCCTGTTGGAGGTC
CCTGCCCAGCTTCCGGCAGGTGGGGGACCAATTGAAATTCAAGTACCTGGACAGCG
TCGAAATTTCTCCCTCTTTGCAATTGGCTGATAGCAACTCACATCGCGACAAACGCA
ACAGCATACCCCAACCCGCCTCTGACACCGACTTGATCTTCCTCGATAAATCTCCCA
ACTACTGCCGGCCAGACCGCAAGAGAGGGGTCCAAGGAACTCGGGACCGAAACTGC
CAACCCGATACCGACAAACCGAACAATTGTAAGCATCTCTGCTGCGGGAGAGGGTA
TCGAACCCGGGTGATCGAGGTCGACGAAGCTTGTGAATGCCAATTATGTGGTGCTGT
AGTATTCAGTGCAAGATTTGCAAGAAGGTTCAAGTTGTTCATACGTGCCTGTGA

>*wnt16b* cDNA sequence cloned by SJC
CAGTTTGAAGAGGAGAGGTGGAATTGTCGTCTGCTGGACGGGGCGACCAAAGATGA
AAATATCCTGGAGACTATTTTGCAGCAAGGAACGAAGGAGATTGCTTTCATCCACGC
CATCACGTCTGCCTCACTCGCTCACTCCATCACTTCATCATGTAGCGCTGGCCTCCTT
CTAGAATGCTCTTGCGACAGGTCGGTGCAGTCGATCGTGAGTACGGACAGCAGCTG
GCGCTGGGGTGGCTGCAGCGATAACGTCCAATACGCCATCAAGTATTCGAAGATCA
TCACAGATGGTGGAAATAAACAAAAAGACGGAAAAAAGGCAGCTCTGATGGAACG
GGTTCGATCGGTCGTTCATCTGCATAATAATGATGTCGGAAGGAAGACCTTCCACTC
GCTGATGACCCACAAATGTCGTTGCCATGGTGTCTCGGGGTCGTGCGCCGTTCGATC
CTGTTGGAGGTCCCTGCCCAGGTTCCGGCAGGTGGGGGACCAATTGAAATTCAAGTA
CCTGGACAGCGTCGAAATTTCTCCCTCTTTGCAATTGGCTGATAGCAACTCAGTGAG
TGTGGCCTCTGGTCAAAAGAGCCAATGGAGA

>*wnt16b* 5’ RACE sequence
ACATGGGCACCGTGAAAAGTTTCTTCAACAACTTCCACAACAGCAACAACTACTTCA
ACAACAACAACTTCCACAACAACAACAACTTCTTCAACAACAACAACTTCCACAAC

AACAACAACTTCTTCAACAACAACAACTTCCACAACAGGAACAACTTCCACAACAA
CAACATATTCTTCAACAACAACAACTTCTTCAACAACAACAACTTCCACAACAGCAA
CAACTTCTTCAACAACAACAACATTCGCAACAACATTCACGACATACGAAACAACA
TGCTAAAGCACATCATCAGCTTAATAAAAAACAATATCTTCAGTCTACACCCCAGAA
CCAAGAACATCAATACTACTATCAACAACAACAACTTCAACAACAACAACAATTTC
AACAACAACAACTTCAACAACAACAACTTCATCAGCAACAACAGCAACAACAACAT
CAACAACAACATATCAACAACTGCACCTATCAGACAGAAATGTTGCAGCACAAAAT
TGACGCATGTTCCCCTGAGAACCACGCGATCATCAACAGCATCCAACAAGGATCGG
TGTTGGGACTGTCGGAATGCAGGCAGCAGTTTGAAGAAGAGAGGTGGAATTGTCGT
CTGCTGGACGGGACGACCAAAGATGGAAATATCCTGGAGACTATTTTGCAACAAGG
AACGAAGGAGATTGCTTTCATCCACGCCATCACGTCTGCCTCACTCGCTC

>*wnt16b* 3’ RACE product #1 *(wnt16b-i)*
TCCACGCCATCACGTCTGCCTCACTCGCTCACTCCATCACTTCATCATGTAGCGCTGG
CCTCCTTCTCGAATGCTCTTGTGACAGGTCCCTGCAGTCGATCGTGAGTACGGACAG
CAGCTGGCGCTGGGGTGGTTGCAGCGATAACGTCCAATACGGCATCAAGTATTCGA
AGATCATCTTCCTCATATCCCTCACCATCCTTCATATAATAATTATCACCATTTTGTT
GGCCACTGAAATTTATCCGGCAATTAAATTCGATAACTCTCTTGATGATTTTTGTCAA
CTTCCTTACCCACACACACACAAACACAGACAAACATACACACACGCCTACCTACA
CGTACGCATACACACACACACACACACACACAGTTGTACAATATTTATTGGTCTTAT
TTATTTACTGATACACGATACCGCTAGTACTATATAGACTACCTACATACATTTCAAT
ATATACACATAAAATTTCATTCTTTTATCCGTCCGTCTGTTTTTCAGTTCATCCATTCA
TTCATCTAATCATTCAATCGTTTGTTCGTTCATCCACTAATTCACTCATTCATTCATTC
ATTCACTTATGATCAAACCACCGACTAGTCATTTTTATTCTTTCTCGATGAACACTTC
ATTTTCACCATGAATTGAAACGTCGTCGGCTTCATTAGCCTCCTCGCAGGTTCGAAT
AGTTTGATGGCCTTATAAAATTTTCCCTGATTTATCTTTAATATTTGAATGTCTTCCAT
GGGATGGTTGATCCAATCCACTATTTCAATGTAAACTATTCGATGCAATGTACCAAT
GGGTGCTCAAGAATGTACAACCAGCCAGCCAACCAACCAGCAAGTCGTCAATTTCA
AACAGGATTTCAATTCTACCATCATATATACTGGCATTTTGGCATTGTAAATAAATT
ATTAAATGTGTGTGTAAATAAAAAAAAAAAAAAAAAAA

> *wnt16b* 3’ RACE product #2 (*wnt16b-ii)*
CTCTGGTCACAAGAGCCAATGGAGAAGGTCCGTTCATAGAAGAAAGAACGTGATTG
GTGGAAAAGTTACTGATGAAAATGAAAANAGTTTTANNAAANAAGATAAAAAAAT
GGTTGTTCGCCCAAGACAACAACATCAACAACAACATCAACAACATCAACAACAAC
ATCAACAACAACATCAACAACGACAACAACTAACCCTCGTAATGAAGCACCGCGAC
AAACGCAACAGCATACCCCAACCCGCCTCTGACACCGACTTGATCTTCCTCGATAAA
TCTCCCAACTACTGCCGGCCAGACCACAAGAGAGGGGTCCAAGGAACTCGGGACCG
AAACTGCCAACCCGATACCGACAAACCGAACAATTGTAANCATCTCTGCTGCGGGA
GAGGGTATCGAACCCGGGTGATCGAGGTCGACGAANCTTGTGAATGTCAATTTTTGT
GGTGCTGTAGTATTCANTGCAAGATTTGCAAGAAGGTTCAAGTTGTCCATACGTGCC
TGTGAATGTGTGTGTGTGTGTTTGCGTGTGTGTGTGTGTTTGCGTGTGTGTGTTTGCG
TGTGTGTTCTAGTTTGAAAAAGATTCTTCGATAGAAAATAAAATAATGTAACGAGTT
GCATGCCGTTTTTCGTTGTTATTTTTGATGTTTCTTTCGATATTTTGNTGTAACTTCNA
TTAACTTTTTCTTGTTTTCNTCACCCCCTCACCATCGTCNTCTTCTTC