Molecular Biology and Evolution

Evolutionary Dynamics of the *wnt* Gene Family: A Lophotrochozoan Perspective

Sung-Jin Cho,†,[1] Yvonne Vallès,*†,[1] Vincent C. Giani Jr,[2] Elaine C. Seaver,[2] and David A. Weisblat*,[1]

[1]Department of Molecular and Cell Biology, University of California

[2]Kewalo Marine Laboratory, Pacific Biosciences Research Center, University of Hawaii

†These authors contributed equally to this work.

***Corresponding author:** E-mail: weisblat@berkeley.edu; kyvovas@gmail.com.

Associate editor: Billie Swalla

Abstract

The *wnt* gene family encodes a set of secreted glycoproteins involved in key developmental processes, including cell fate specification and regulation of posterior growth (Cadigan KM, Nusse R. 1997. Wnt signaling: a common theme in animal development. Genes Dev. 11:3286–3305.; Martin BL, Kimelman D. 2009. Wnt signaling and the evolution of embryonic posterior development. Curr Biol. 19:R215–R219.). As for many other gene families, evidence for expansion and/or contraction of the *wnt* family is available from deuterostomes (e.g., echinoderms and vertebrates [Nusse R, Varmus HE. 1992. Wnt genes. Cell. 69:1073–1087.; Schubert M, Holland LZ, Holland ND, Jacobs DK. 2000. A phylogenetic tree of the Wnt genes based on all available full-length sequences, including five from the cephalochordate amphioxus. Mol Biol Evol. 17:1896–1903.; Croce JC, Wu SY, Byrum C, Xu R, Duloquin L, Wikramanayake AH, Gache C, McClay DR. 2006. A genome-wide survey of the evolutionarily conserved Wnt pathways in the sea urchin Strongylocentrotus purpuratus. Dev Biol. 300:121–131.]) and ecdysozoans (e.g., arthropods and nematodes [Eisenmann DM. 2005. Wnt signaling. WormBook. 1–17.; Bolognesi R, Farzana L, Fischer TD, Brown SJ. 2008. Multiple Wnt genes are required for segmentation in the short-germ embryo of Tribolium castaneum. Curr Biol. 18:1624–1629.]), but little is known from the third major bilaterian group, the lophotrochozoans (e.g., mollusks and annelids [Prud'homme B, Lartillot N, Balavoine G, Adoutte A, Vervoort M. 2002. Phylogenetic analysis of the Wnt gene family. Insights from lophotrochozoan members. Curr Biol. 12:1395.]). To obtain a more comprehensive scenario of the evolutionary dynamics of this gene family, we exhaustively mined *wnt* gene sequences from the whole genome assemblies of a mollusk (*Lottia gigantea*) and two annelids (*Capitella teleta* and *Helobdella robusta*) and examined them by phylogenetic, genetic linkage, intron–exon structure, and embryonic expression analyses. The 36 *wnt* genes obtained represent 11, 12, and 9 distinct *wnt* subfamilies in *Lottia*, *Capitella*, and *Helobdella*, respectively. Thus, two of the three analyzed lophotrochozoan genomes retained an almost complete ancestral complement of wnt genes emphasizing the importance and complexity of this gene family across metazoans. The genome of the leech *Helobdella* reflects significantly more dynamism than those of *Lottia* and *Capitella*, as judged by gene duplications and losses, branch length, and changes in genetic linkage. Finally, we performed a detailed expression analysis for all the *Helobdella wnt* genes during embryonic development. We find that, although the patterns show substantial overlap during early cleavage stages, each *wnt* gene has a unique expression pattern in the germinal plate and during tissue morphogenesis. Comparisons of the embryonic expression patterns of the duplicated *wnt* genes in *Helobdella* with their orthologs in *Capitella* reveal extensive regulatory diversification of the duplicated leech *wnt* genes.

Key words: *wnt* family genes, lophotrochozoan genomes, gene duplication and diversification, annelid, leech, polychaete.

Research article

Introduction

One current view of animal evolution posits that differential expression of conserved sets of regulatory genes, including those encoding signaling pathway proteins, accounts for body plan diversification, with little contribution from new genes that might arise in different groups of animals (Carroll et al. 2001). In particular, the expansion and contraction of conserved gene families, by gene duplication and loss, respectively, play key roles in developmental diversification (Li et al. 2005); for example, gene duplications permit the relaxation of selective pressures, allowing paralogs to undergo changes in their regulatory and/or coding sequences that can alter developmental outcomes (Ohno 1970; Li et al. 2005).

The family of Wnt signaling ligands provides a good example of this ongoing dynamic of gene duplication, loss, and modification. Wnt ligands are secreted glycoproteins 350–400 amino acids in length, whose domain structure is defined by the presence of 23–24 cysteines positioned at conserved sites throughout the protein's length (Nusse and Varmus 1992; Cadigan and Nusse 1997). Comparisons of available whole-genome sequences indicate that 13 *wnt* subfamilies were present in the common ancestor of cnidarians and bilaterally symmetric animals (Kusserow et al. 2005; Croce et al. 2006; Garriock et al. 2007; Bolognesi et al. 2008; Lengfeld et al. 2009). Of these, 12 have been retained in *Nematostella vectensis*, a modern cnidarian, which shows one case of gene duplication, resulting in a total of 12 *wnt*

genes (Kusserow et al. 2005; note that originally two gene duplications were reported, however, our close examination indicates that *wnt7a* and *wnt7b* are splice variants of a single gene). The human genome (super-phylum Deuterostomia) contains 19 *wnt* genes, representing 12 subfamilies with seven duplications (Garriock et al. 2007). In contrast, the beetle *Tribolium* (super-phylum Ecdysozoa) retains only 9 subfamilies, with no duplications (Bolognesi et al. 2008), and two other ecdysozoans, Drosophila and *Caenorhabditis*, have just seven and five *wnt* genes, respectively (Eisenmann 2005; Bolognesi et al. 2008).

These differences reflect diverse genome dynamics throughout evolution and lead naturally to further questions. Does the relative paucity of *wnt* genes in the three ecdysozoan species reflect a general protostome trend? To what extent are the *wnt* subfamilies specialized for specific developmental functions, as opposed to being essentially interchangeable with one another? What happens to *wnt* gene expression and function in response to gene duplication or loss?

To begin addressing these questions and to further investigate the evolvability of the *wnt* family genes, we have taken an inventory of the *wnt* genes from three members of Lophotrochozoa, the third super phylum of bilaterally symmetric animals. The whole-genome sequences were examined from the mollusk *Lottia gigantea* and two annelids, the polychaete *Capitella teleta* previously known as *Capitella* sp. I (Blake et al. 2009) and the leech *Helobdella robusta*. We find 11, 12, and 13 *wnt* genes, representing 11, 12, and 9 subfamilies in these species, respectively. Phylogenetic analyses reveal that multiple duplications and losses of *wnt* subfamily genes have occurred in the leech lineage.

To begin elucidating *wnt* gene function in lophotrochozoan development, we examined the expression patterns of the complete set of *wnt* genes during embryogenesis in the leech *H. robusta* with particular attention to those subfamilies (*wnt5*, *wnt11*, and *wnt16*) that have undergone duplications relative to *Capitella*, the closest relative of *Helobdella* for which the whole-genome sequence is available. Comparing the expression patterns of these genes with their unduplicated orthologs in *Capitella* suggests that the duplicated genes in leech have undergone regulatory subfunctionalization with respect to the single-copy gene in an annelid ancestor. These findings provide new lophotrochozoan models for elucidating the evolutionary constraints operating on *wnt* genes in metazoan evolution while permitting tremendous body plan diversification.

Materials and Methods

Cloning, Sequencing, and In Vitro Transcription

Specific primers (supplementary table S1, Supplementary Material online) for the genes of interest were designed against the genomic sequences of *H. robusta* (http://genome.jgi-psf.org/Helro1/Helro1.home.html) and *C. teleta* (http://genome.jgi-psf.org/Capca1/Capca1.home.html) and amplified from *H. robusta* var. *austin* cDNA and *C. teleta* cDNA, respectively. A fragment of the *Capitella wnt5* gene was isolated by degenerate polymerase chain reaction (PCR), and additional sequence was recovered using the SMART RACE amplification kit. PCR products were gel extracted with a Quiagen Gel Extraction Kit and ligated into pGEM-T easy vector according to supplier's instructions. These clones were introduced into *Escherichia coli* DH105 competent cells by heat shock, plated onto 1% agar plates, and grown overnight at 37 °C. Colonies were selected and grown overnight in LB Broth with 1% ampicillin. Plasmids were isolated using the Quiagen Miniprep Kit following manufacture's instructions and sequenced on an Applied Biosystems platform. Riboprobes labeled with digoxygenin or biotin were made using the MEGAscript (Ambion) kit, according to the manufacturer's instructions. A 1.4 kb cDNA clone representing the *Capitella wnt16* gene was identified from searches of a *Capitella* expressed sequence tag library, recovered from a glycerol stock and sequenced for verification.

Alignment and Phylogenetic Analyses

All 90 amino acid sequences used in the analyses (supplementary table S2, Supplementary Material online) were prealigned using ClustalW2 (Larkin et al. 2007) followed by manual editing in Se-Al v2.0a11 (http://tree.bio.ed.ac.uk/software/seal/). Two data sets were used for the analyses, one including all the amino acids from the Wnt domain and the other using Gblocks to exclude poorly aligned amino acid positions under the least stringent parameters as implemented in the online version of the program (Castresana 2000). The resulting alignments comprised 584 and 215 amino acid positions, respectively (TREEBASE M5070). The Whelan and Goldman model of protein evolution (Whelan and Goldman 2001) with gamma distribution + invariant sites was implemented in all analyses as selected by ProtTest v1.4 as the model that best fits the data (Abascal et al. 2005). Bayesian inference (BI) analyses were executed in MrBayes v3.0 (Huelsenbeck and Ronquist 2001) for 15,000,000 generations, and maximum likelihood (ML) analyses were executed in RAxML-v7.0.4 (Stamatakis et al. 2005) for which support was obtained from 1,000 bootstrap replicates. Trees were viewed and edited in Dendroscope v2.2.2 (Huson et al. 2007). NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/) was used to determine all potential N-glycosylation sites of the wnt genes (supplementary table S3 and fig. S1, Supplementary Material online).

ISH in *Helobdella*

All in situ hybridizations (ISHs) of embryos from the zygote to stage 8 were performed under the same conditions as specified in Rivera et al. (2005). Stage 9–11 embryos ISH protocol is the same as the aforementioned except that it was preceded by the following treatments: embryos were collected and relaxed for 10 min in a relaxant solution (10 mM $MgCl_2$, 5 mM NaCl, 1 mM KCl in 8% ethanol in water), then fixed in 4% paraformaldehyde (PFA) for 1 h. Fixed embryos were washed 5 × 5 min in 0.1% PBTw (phosphate-buffered saline [PBS] and 0.1 % Tween 20), and then permeabilized by treatment with 0.5 mg/ml pronase E (Sigma) in PBS for 12 min at 37 °C. The pronase treatment was followed by 2 × 5 min rinses in a solution containing

2 mg/ml glycine in PBS and afterward 2 × 5 min washes in PBS at room temperature. Then, 2 × 5 min washes in 1 ml 0.1 triethanolamine buffer (TEA, pH 8.0) was followed by 1 × 5 min wash in 1 ml of TEA to which was added 6 µl of acetic anhydride. The embryos were then rinsed 3 × 5 min in PBS and postfixed for 20 min in 4% PFA. From this point on the protocol conditions were identical to those described in Rivera et al. (2005). Probe lengths were as follows: *wnt1* (GU289716): 600 bp; *wnt2* (GU289717): 906 bp; *wnt4* (GU289718): 927 bp; *wnt5a* (GU289719): 894 bp; *wnt5b* (GU289720): 936 bp; *wnt6* (GU289721): 714 bp; *wnt7* (GU289722): 1155 bp; *wnt10* (GU289723): 609 bp; *wnt11a* (GU289724): 951 bp; *wnt11b* (GU289725): 1131 bp; *wnt11c* (GU289726): 1020 bp; *wnt16a* (GU289727): 624 bp; *wnt16b* (GU289728): 870 bp.

For double fluorescent in situ hybridization (DFISH), we used the NEN tyramide signal amplification (TSA) Plus kit (PerkinElmer, Wellesley, MA). The protocol was identical to that in Rivera et al. (2005) except for the blocking and color reactions steps, where embryos were rinsed once in x1 PBTw and blocked for 2 h at room temperature in a solution of 0.1 M Tris-HCl pH 7.5, 0.15 M NaCl, 0.1% Tween 20 (TNT) containing 1% NEN TSA blocking reagent (collectively designated TNB). Embryos were incubated overnight at 4 °C with either peroxidase-conjugated anti-digoxygenin or anti-biotin antibody (Roche Applied Science) at a 1:1000 dilution in TNB. Subsequent washes (3 × 20 min) in TNT at room temperature were followed by a single 1 × 30 min rinse in NEN TSA Plus amplification solution. The color reaction was initiated by adding a 1:50 dilution of reconstituted cyanine-3 or fluorescent tyramide in NEN amplification solution. Embryos were imaged on a Leica SMRE microscope equipped with a TCS SL scanning head. DFISH was visualized using a spinning-disk confocal microscope (Leica Microsystems).

Lineage Tracing and Cross Sectioning

To determine the embryonic origin of the cells expressing specific *wnts*, teloblasts or proteloblasts of embryos in stages 5 and 6 were injected with one or more fluorescently labeled, fixable dextran lineage tracers: fluorescein-conjugated dextran amine (FDA); tetramethylrhodamine-conjugated dextran amine (RDA); or Far Red-conjugated dextran amine. Injected embryos were cultured in Helobdella triserialis saline at 23 °C to the desired embryonic stage, then fixed and processed by ISH for the different *wnts* as described above. After ISH, the embryos were dehydrated in ethanol and propylene oxide, followed by infiltration with plastic embedding medium (Poly/Bed 812; Polysciences). Embryos were sectioned by using a glass knife on a microtome (MT-2B; Sorvall, Newtown, CT) or handcut by razor blade into 0.1-mm sections. Sections were imaged on a Leica compound microscope or spinning-disk confocal microscope (Leica Microsystems).

ISH in *Capitella*

Whole-mount ISH was performed according to previously published protocols for *Capitella* (Seaver et al. 2001; Seaver and Kaneshige 2006). Larvae were hybridized for 72 h at 65 °C with a probe concentration of 1 ng/µl for all three genes. Probe lengths are as follows: *wnt5* (GU323406), 1.3 kb; *wnt11* (GU323408), 850 bp; *wnt16* (GU323407), 1.4 kb. Specimens were imaged using an Axioskop 2 mot plus (Zeiss) with a SPOT FLEX digital camera (Diagnostic Instruments, Inc.).

Results

Characterization of Lophotrochozoan wnt Protein Sequences

Wnt proteins differ from many other well-conserved developmental regulators in that they lack compact, highly conserved structural domains, as defined by clearly bounded stretches of highly conserved amino acids. Instead they are characterized by small groups of well-conserved amino acid residues (usually on the order of 100 residues total) scattered throughout the protein sequence, of which approximately 23–24 are cysteines (Nusse and Varmus 1992). A total of 11, 12, and 13 Wnt genes were identified from the whole-genome sequences of the mollusk, polychaete, and leech, respectively (table 1 and supplementary table S2, Supplementary Material online; for details, see Materials and Methods). Although obviously similar to their deuterostome and ecdysozoan orthologs, the conceptually translated lophotrochozoan Wnt proteins reveal greater structural variability (supplementary table S3, Supplementary Material online). Most of the outliers are from *Helobdella*, such as Wnt2 with only 13 cysteines or Wnt7 with 27. Interestingly, although *Capitella* Wnts range between 21 and 26 cysteines, the *Lottia* Wnts all have exactly 24 conserved cysteines, except for Wnt9 (although these values could change as the genome is further annotated and the gene models validated).

As secreted glycoproteins, Wnts also feature N-linked glycosylation sites, characterized in part by the consensus sequence Asn-Xaa-Ser/Thr, where Xaa can be any amino acid except for a proline (http://www.cbs.dtu.dk/services/NetNGlyc/). Scanning the metazoan *wnt* genes used in this study for candidate N-linked glycosylation sites, we found between 0 and 5 sites per protein (supplementary table S3, Supplementary Material online). Two sites, located at the N and C termini of the protein, respectively (supplementary fig. S1, Supplementary Material online), are conserved across metazoans in almost all Wnts examined, suggesting that they might be essential for normal function of the Wnt proteins.

Diversity, Phylogeny, and Genetic Linkage of Lophotrochozoan *wnt* Genes

To assign the *wnt* genes identified from whole-genome sequences to specific *wnt* subfamilies, four separate phylogenetic analyses were carried out, applying both ML and BI to each of two data matrices, one comprising the full-length proteins (584 aa) and the other comprising the most conserved portions of the alignment (215 aa sites; for details, see Materials and Methods). For these analyses, a total of 90 *wnt* sequences were used representing all major metazoan groups. All four analyses yielded similar topologies and recovered the 13 proposed *wnt* subfamilies as monophyletic with high posterior probabilities (pp) and

Table 1. Distribution of the *wnt* Genes in Metazoans.

	Cnidaria	Vertebrata	Ecdysozoa	Lophotrochozoa		
Gene	*Nematostella vectensis*	*Homo sapiens*	*Tribolium castaneum*° other *	Mollusk, *Lottia gigantea*	Annelid, *Helobdella robusta*	Annelid, *Capitella teleta*
wntA	1 (24)		1*° (29)	1 (24)		1 (22)
wnt1	1 (23)	1 (23)	1*° (25)	1 (24)	1 (20)	1 (24)
wnt2	1 (24)	2 (24, 26)	1* (26)	1 (24)	1 (13)	1 (24)
wnt3	1 (23)	2 (25, 24)				
wnt4	1 (24)	1 (25)	1* (26)	1 (24)	1 (24)	1 (26)
wnt5	1 (25)	2 (24, 24)	1*° (24)	1 (24)	2 (23, 24)	1 (21)
wnt6	1 (17)	1 (25)	1*° (24)	1 (24)	1 (20)	1 (26)
wnt7	1 (24)	2 (25, 25)	1*° (25)	1 (24)	1 (27)	1 (24)
wnt8	2 (21, 23)	2 (25)	1*° (23)			1 (24)
wnt9		2 (24, 25)	1*° (18)	1 (25)		1 (24)
wnt10	1 (24)	2 (24, 25)	1*° (24)	1 (24)	1 (24)	1 (24)
wnt11	1 (25)	1 (26)	1*° (22)	1 (24)	3 (24, 25, 24)	1 (24)
wnt16	1 (20)	1 (25)	1* (24)	1 (24)	2 (24, 21)	1 (25)

NOTE.—Numbers in parenthesis indicate the number of cysteines in that gene. In the Ecdysozoa column, (°) indicates genes present in *T. castaneum*; (*) indicates genes present in other species.

moderate bootstrap support (bs), except for the bs of *wnt2* and *wntA* subfamilies (fig. 1).

In addition to supporting the recognized subfamilies, our Bayesian analysis provides good support (pp values > 98) for pairing particular subfamilies (i.e., *wnt10* and *wnt9*; *wnt1* and *wnt6*; *wnt4* and *wnt11*) that have been linked in previous phylogenies (Kusserow et al. 2005; Croce et al. 2006), consistent with the hypothesis that these pairs are the products of specific gene duplication events in the metazoan ancestor. Strong support in the Bayesian analyses was also obtained for placing the *wnt8* subfamily as a sister group to the *wnt9*/wnt10 ancestor (pp $= 100$) (fig. 1).

The 36 lophotrochozoan *wnt* sequences represent orthologs to most of the 13 known *wnt* subfamilies: 11 in *Lottia*, 12 in *Capitella*, and 9 in *Helobdella*. Curiously, none of these three genomes yielded a *wnt3* ortholog nor have we been able to identify bona fide *wnt3* orthologs in any other lophotrochozoan or ecdysozoan genomes, suggesting that this gene was lost in the ancestral protostome. (The NCBI database lists three nominal *wnt3* genes from ecdysozoan taxa but our analyses cluster all three within the *wnt7* subfamily [data not shown].) In contrast, both mollusk and leech (but not polychaete) lack *wnt8*, indicative of more recent, independent losses of this subfamily member within annelids and mollusks. Furthermore, these analyses reveal that within the annelids, in the branch leading to leeches, there have been multiple losses (*wnt8*, *wnt9* and *wntA*) along with independent duplications (*wnt5*, *wnt11* and *wnt16*).

In light of the genetic linkage between several *wnt* genes in other metazoans (Garriock et al. 2007; Bolognesi et al. 2008; Putnam et al. 2008), we also looked for scaffolds containing multiple *wnt* genes in the three lophotrochozoan species (fig. 2). In *Lottia*, 7 of the 11 *wnt* genes are found on just two scaffolds. Orthologs of *wnt1*, *wnt6*, *wnt9* and *wnt10* lie within a ~0.08 Mb stretch on one scaffold, whereas *wnt4*, *wnt5* and *wnt7* orthologs lie within ~1.08 Mb on another. Interestingly, the *Capitella* orthologs of *wnt1*, *wnt9* and *wnt10* are also found on a single scaffold, within a ~0.02 Mb region; in the more distantly related genome of the cnidarian *N. vectensis*, the *wnt10* and *wnt6* genes lie on one scaffold ~0.045 Mb apart and the *wnt5 and wnt7* genes lie on another ~0.015 Mb apart. A parsimonious explanation of these observations is that these neighbor relationships (*wnt1 and wnt6* with *wnt10* and perhaps *wnt9; wnt5* with *wnt7*) reflect the genomic organization of the metazoan ancestor of cnidarians and bilaterians. Interestingly, neither of these neighbor relationships are recovered in the *Helobdella* genome, despite the larger average scaffold size in its assembly; half of the *Helobdella* genome is contained in 21 scaffolds greater than or equal to 3.1 Mb in length, compared with 51 scaffolds greater than or equal to 1.87 Mb, and 454 scaffolds greater than or equal to 188 Kb for comparable fractions of the *Lottia* and *Capitella* genomes, respectively. In *Helobdella*, we found two scaffolds that each contained two *wnt* genes; *wnt1* and *wnt7* lie ~7.2 Mb apart on one scaffold, whereas *wnt4* and *wnt5* lie ~1.5 Mb apart on another. The failure to observe the conserved neighbor relationships of *wnt1*, *wnt6*, and *wnt10* in leech and also the observation of distant *wnt* pairings not seen in other organisms suggests that more genome rearrangements have occurred along the branch leading to *Helobdella* than along those leading to its lophotrochozoan counterparts.

The gain and loss of introns in Metazoa are far less common than point mutations and thus can be useful measures of evolution among both distantly related genes and distantly related taxa (Irimia and Roy 2008; Li et al. 2009). We have examined the exon–intron boundaries for all 85 (of 90 total) *wnt* genes used in this study for which information was available. In addition to numerous taxon-specific, sub-family–specific introns, we found that two introns are conserved in all *wnt* subfamilies across all metazoans examined, corresponding to introns #2 and #3 in most genes (supplementary fig. S1, Supplementary Material online). In addition, the *wnt1*, *wnt2*, and *wnt11* subfamilies have each acquired an additional intron that characterize these respective subfamilies across multiple clades; in the case of *wnt1*, this additional intron is present only in the lophotrochozoan taxa, whereas the additional *wnt2* and *wnt11* introns are present only in the bilaterians.

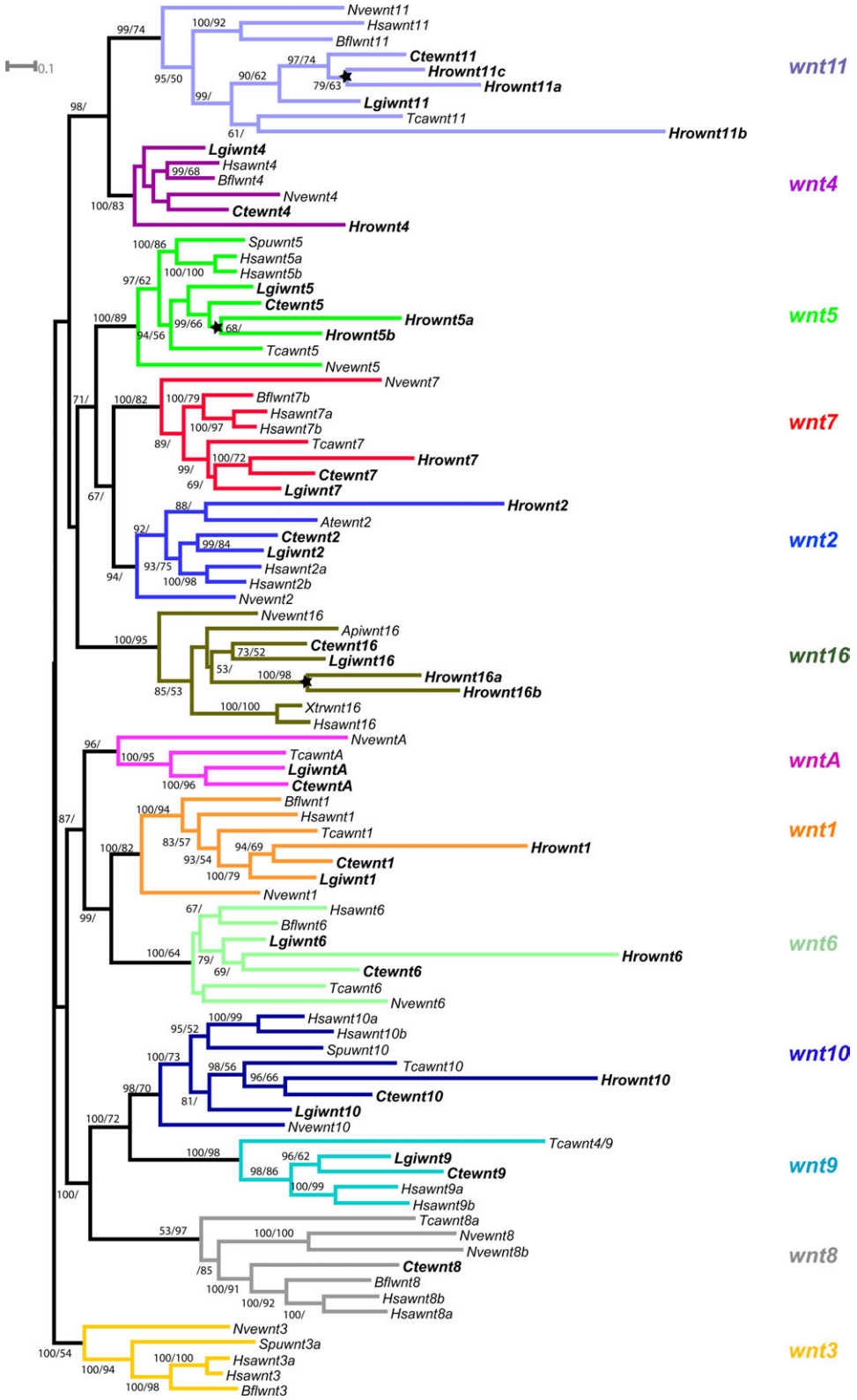

Fig. 1. ML tree showing the distribution of the *wnt* genes. Numbers on branches are pp/bs values. Species abbreviations: Api, *Acyrthosiphon pisum*; Ate, *Achaearanea tepidariorum*; Bfl, *Branchiostoma floridae*; Cte, *Capitella teleta*; Hro, *Helobdella robusta*; Hsa, *Homo sapiens*; Lgi, *Lottia gigantea*; Nve, *Nematostella vectensis*; Spu, *Strongylocentrotus purpuratus*; Tca, *Tribolium castaneum*; Xtr, *Xenopus tropicalis*. Lophotrochozoan genes are in bold. Black stars indicate the lineages tested for positive selection (for details, see Supplementary Material, Supplementary Material online).

One mechanism by which gene duplication may contribute to diversity is for the paralogs to come under different functional and/or structural constraints, in which case their diverging sequences might show signs of positive selection (Wray et al. 2003; Gu et al. 2004). We therefore tested for evidence of positive selection in the

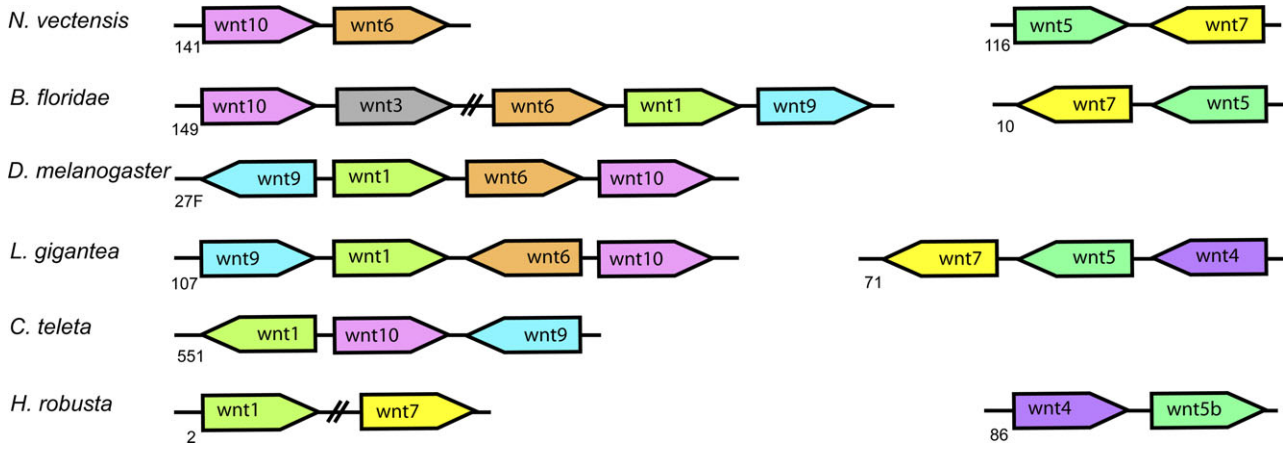

Fig. 2. *wnt* genetic linkage shared among metazoans. Cartoon showing the relative position and orientation of the *wnt* genes that share scaffolds/chromosome in *Nematostella vectensis*, *Branchiostoma floridae*, *Drosophila melanogaster*, *Lottia gigantea*, *Capitella teleta*, and *Helobdella robusta*. The distances between genes and the size of the *wnt* genes themselves are not to scale, but slanting lines indicate very large intergenic regions. Numbers indicate the specific scaffold or chromosome (in the case of *D. melanogaster*).

leech-duplicated genes; however, none was detected (details in Supplementary Material, Supplementary Material online). This outcome was not completely surprising, both because our analysis was limited to coding sequences and because large divergences among distantly related taxa usually result in the saturation of nonsynonymous and synonymous substitution sites and thus conceal any signal of positive selection (Emes and Yang 2008).

wnt Gene Expression Patterns during Development in the Leech *Helobdella*

To explore the functional significance of *wnt* gene diversity, duplication and loss in Lophotrochozoa, we carried out ISH for all 13 *Helobdella wnt* genes throughout development. Leech development from zygote deposition (at metaphase of meiosis I) to juvenile has been divided into 11 stages. For greater precision, embryos are also staged in terms of time elapsed after zygote deposition (AZD) (Weisblat and Huang 2001). Broadly speaking, these developmental stages encompass three partially overlapping phases. First, during stages 1–7, unequal spiral cleavage segregates cytoplasmic determinants (teloplasm) to the D quadrant of the four-cell embryo and thence to five bilateral pairs of segmentation stem cells called teloblasts. One pair of teloblasts (M_R and M_L) will give rise to the mesodermal progeny, whereas the four remaining pairs—N_R, O_R, P_R, Q_R and N_L, O_L, P_L, Q_L—will give rise to the ectodermal progeny. Second, during stages 6–8, the teloblasts produce columns of segmental founder cells (blast cells) in parallel arrays that coalesce along the ventral midline to form a germinal plate, from which segmental mesoderm and ectoderm arise. Finally, during stages 7–11, individual blast cells produce stereotyped clones that interdigitate and differentiate to produce segmental mesodermal and ectodermal descendents of the juvenile leech. Gut and nonsegmental tissues of the prostomium develop from stereotyped sets of macromeres and micromeres that arise in parallel with the teloblasts during cleavage (Weisblat and Huang 2001; Huang et al. 2002).

ISH reveals that *Helobdella wnt* genes are expressed throughout development. Here, for simplicity, we only show results from stages 2 and 4b to represent cleavage (figs. 3 and 4, columns I–III), from stage 7 to illustrate the stem cell divisions of the teloblasts (figs. 3 and 4, column IV), and from stages 9 and 10 to illustrate segmental differentiation (figs. 3 and 4, columns V–VII).

Overlapping Expression Patterns in Early Leech Development

In early development (stages 2–7), *wnt* gene expression patterns overlap extensively but not entirely. For example, at stage 2, most *wnt* transcripts (e.g., *wnt2* and *wnt16a*) appear abundant in the teloplasm, which is segregated to the prospective D quadrant during the first two divisions (figs. 3 and 4, columns I–III). In contrast, *wnt1*, *wnt7*, *and wnt11a* mRNAs exhibit dynamic expression and diffuse distributions during stage 2, accumulating first in cell CD and then in cell AB (figs. 3 and 4, columns I–II); this dynamic and complementary expression pattern resembles previous results for Wnt7 immunostaining and *notch* mRNA accumulation (Huang et al. 2001; Gonsalves and Weisblat 2007).

At stage 4b, *wnt4* mRNA is observed exclusively over a subset of micromere nuclei (figs. 3 and 4, column III). In contrast, transcripts of all the other *wnt* genes are associated with the teloplasm derived from the D quadrant and (especially for *wnt1*, *wnt10*, and *wnt11a*) with the cytoplasm of micromeres located at the animal pole of the embryo (figs. 3 and 4, column III). Transcripts of *wnt5a*, *wnt7*, and *wnt16a* also exhibit macromere nuclear signals at this stage (figs. 3 and 4, column III).

Overlapping expression of *wnt* genes is also evident at stage 7. Except for *wnt7* and *wnt11b*, all *wnt* genes are strongly expressed by teloblasts and blast cells (figs. 3 and 4, column IV). In addition, *wnt6*, *wnt11a*, and *wnt16a* also show pronounced expression in micromere derivatives

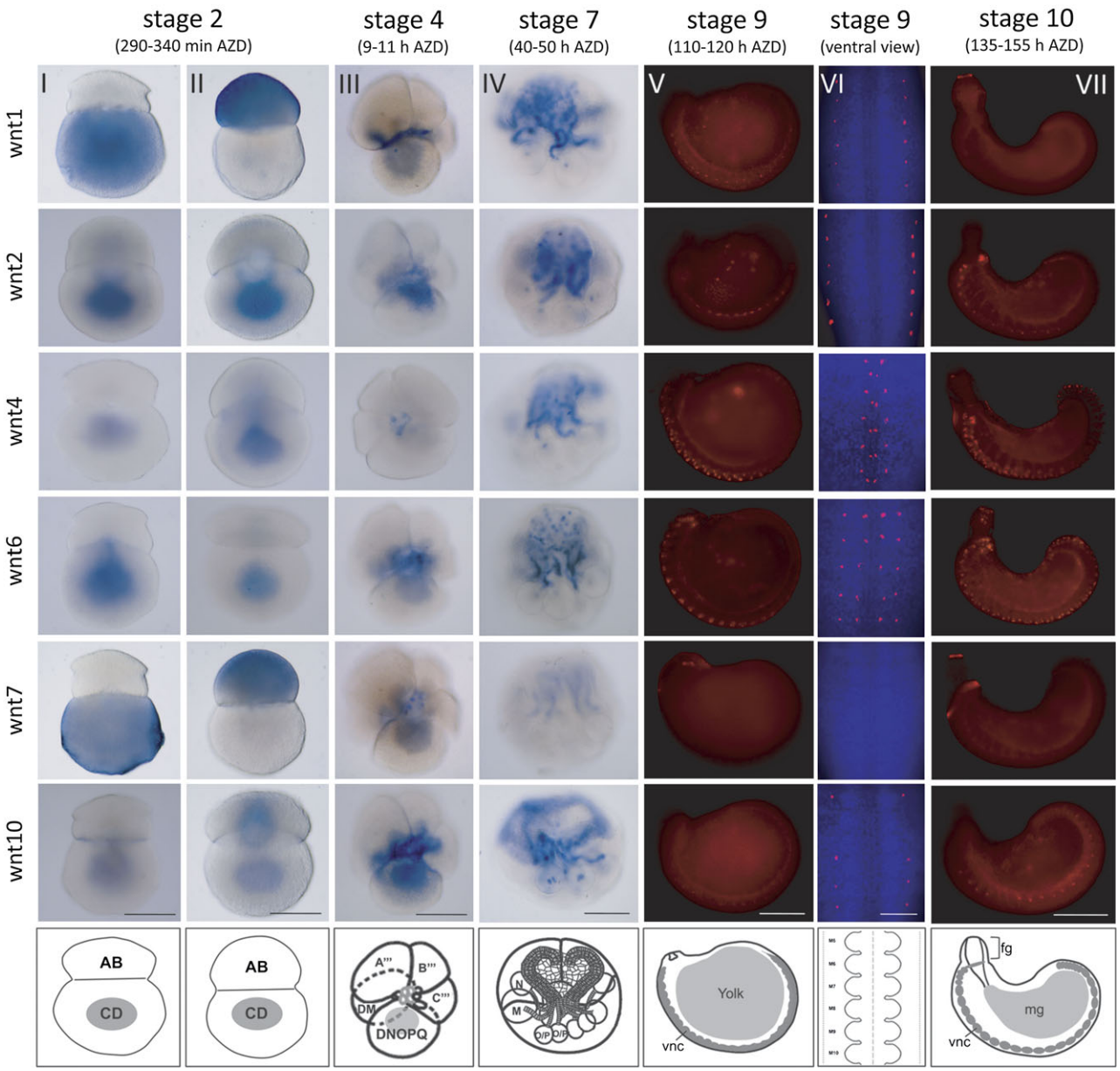

Fig. 3. Expression patterns of the nonduplicated leech *wnt* genes during embryogenesis. Leech development from zygote to juvenile has been divided into 11 stages; schematic views of selected stages are shown in the bottom row (vnc, ventral nerve cord; fg, foregut; mg, midgut). Columns I–IV show nitro blue tetrazolium chloride/5-bromo-4-chloro-3-indolyl phosphate coloration reactions and columns V–VII show FISH. Columns I and II illustrate animal views at two-cell stage at earlier (I; 290–310 min AZD) and later (II; 320–340 min AZD) time points. Two main patterns are observed, either a constant and stable association of the mRNA with the teloplasm in cell CD throughout the two-cell stage (i.e., *wnt2* and *wnt6*) or diffuse and dynamic expression, with mRNA accumulating first in CD (295–310 min AZD) and then in AB (320–340 min AZD; i.e., *wnt1* and *wnt7*). Column III depicts animal views of embryos at stage 4. At this stage, there are four macromeres and 4–15 micromeres, depending on the exact development time. In all cases except for *wnt4*, the mRNA is present in the nucleoplasm (i.e., *wnt5a* and *wnt16a*) of the four macromeres and in most cases in teloplasm (i.e., *wnt7* and *wnt10*). *wnt4*, however, is found exclusively in the nuclei of the micromeres. Column IV shows animal views of stage 7 embryos. At this point, the embryo contains ten teloblasts, each of which gives rise to a column (bandlet) of blast cells, whose progeny will generate the segmented ectoderm and mesoderm; the progeny of the micromeres, which will form nonsegmental body parts (i.e., proboscis); and the A′′′, B′′′, and C′′′ macromeres, which will contribute to the midgut. Most *wnt* genes at this stage are strongly expressed in both teloblasts and bandlets (except for *wnt11b* which is absent and *wnt7* which shows weak expression in the bandlets). In addition, *wnt6*, *wnt11a*, and *wnt16a* are expressed in the micromere-derived epithelium. Columns V (lateral view) and VI (ventral view showing details of the expression patterns in the germinal plate) illustrate stage 9 embryos. Column VI, FISH where red shows the *wnt* mRNAs. All *wnt* genes except *wnt7* and *wnt11b* are expressed in segmentally iterated patterns in the germinal plate. Additionally, *wnt7* and *wnt16a* are expressed in the earliest stages of proboscis development. Column VII shows lateral views of stage 10 embryos, except for *wnt1*, *wnt5a*, *wnt7*, and *wnt11b*, which now appear in the developing proboscis, every other *wnt* gene is expressed in a pattern similar to that observed in stage 9. Scale bar, columns I through V and VII, 100 μm; column VI, 50 μm.

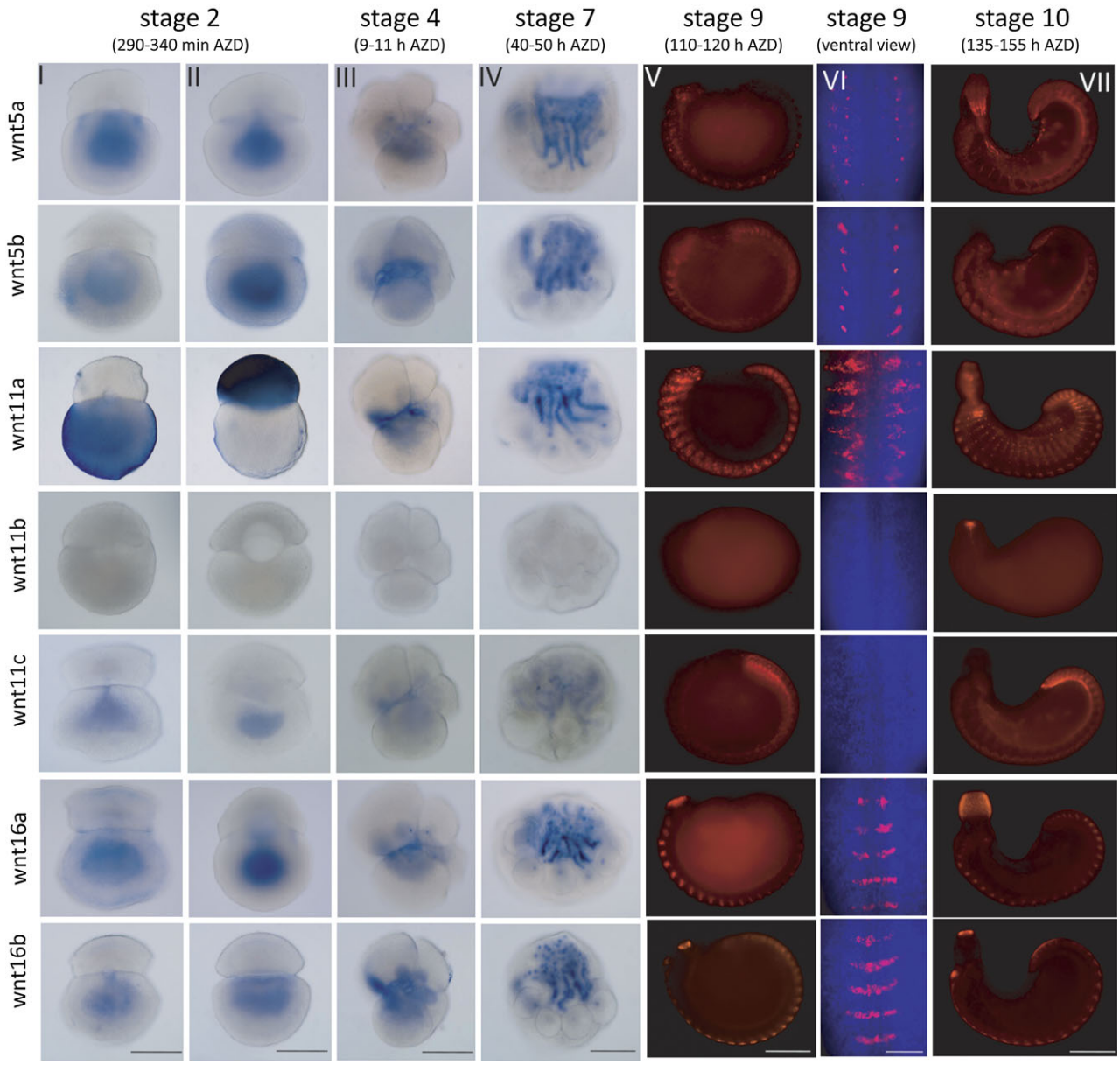

Fig. 4. Expression patterns of the duplicated leech *wnt* genes during embryogenesis. Details as described for figure 3.

lying between the nascent germinal bands (figs. 3 and 4, column IV).

Nonoverlapping Patterns in Late Leech Development

In contrast to the results obtained for stages 2–7, each of the 13 leech *wnt* genes shows a distinct pattern of expression as segments differentiate during stages 9–10, with very little overlap among the various patterns (figs. 3 and 4, columns V–VII). Except for *wnt7* and *wnt11b*, each gene is expressed in a subset of segmentally iterated progeny within the germinal plate.

To determine if the germinal plate expression of individual *wnt* genes is restricted to particular teloblast lineages or to particular germ layers, we carried out further in situ analyses on embryos in which different lineages had been labeled by injecting teloblasts or proteloblasts at earlier stages with fluorescent lineage tracers (for details, see Materials and Methods). Specifically, mesoderm was labeled by injecting the M teloblast, and ectoderm was labeled by injecting either the N teloblast, which contributes primarily to the ventral nerve cord, or the OPQ proteloblast, which contributes primarily to lateral and dorsal ectoderm (Weisblat and Shankland 1985) (fig. 5*A* and *B*). These experiments show that within the germinal plate during stages 9–10, *wnt1*, *wnt4*, *wnt5a*, *wnt5b*, *wnt6*, *wnt10*, *wnt16a*, and *wnt16b* are expressed exclusively in ectodermal derivatives (fig. 5*B′*, *C*, *E–I*, and *K–L*), and *wnt2* and *wnt11c* are expressed exclusively in mesodermal derivatives (fig. 5*B′* and *D*; data not shown for *wnt11c*), whereas *wnt11a* is expressed in both segmental ectoderm and mesoderm (fig. 5*B′* and *J*). In the case of *wnt4*, *wnt6*, *wnt16a*,

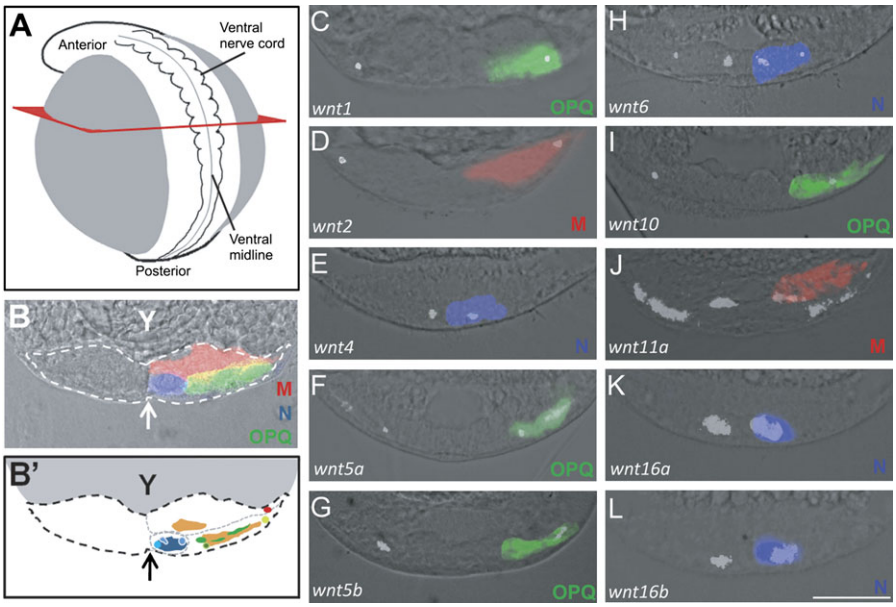

Fig. 5. Germ layer localization of *wnt*-expressing cells in *Helobdella* germinal plate. (*A*) Drawing of a stage 9 embryo showing the germinal plate and ventral nerve cord (white) and underlying yolk (gray). The red square shows where the cross sections were done. (*B*) Pseudocolored image (combined brightfield and fluorescence imaging) showing yolk (Y) and germinal plate (dotted contour) in a cross section of a stage 9 embryo in which the ventral (N) and dorsolateral (OPQ) ectodermal lineages on the right side of the germinal plate are marked with FRDA (blue) and FDA (green), respectively, and the mesodermal (M) lineage is marked with RDA (red), by lineage tracer injection into the various precursor cells at stage 6a (for details, see Materials and Methods). Arrow indicates ventral midline. (*B'*) Schematic representation of (*B*), summarizing the expression of the data shown in (*C*–*L*); *wnt1*, olive green; *wnt2*, red; *wnt4*, light blue; *wnt5a* and *wnt5b*, light green; *wnt6*, medium blue; *wnt10*, dark green; *wnt11a*, orange; *wnt16a* and *wnt16b*, dark blue. The apparent overlap in this transverse view between *wnt5a* and *wnt5b* and between *wnt16a* and *wnt16b* is resolved in ventral views shown in figure 6. (*C*–*L*) Cross sections of stage 9 embryos similar to that shown in panel B, but in which only one lineage was labeled, and which had been processed by FISH for the indicated *wnt* genes (light gray). Scale bar, 100 μm.

and *wnt16b*, the segmental expression pattern includes presumptive neurons in the ventral nerve cord.

In addition to *wnt7* and *wnt11b*, six other *wnt* genes (*wnt1*, *wnt5a*, *wnt6*, *wnt11a*, *wnt16a*, and *wnt16b*) are also expressed in nonsegmental patterns, including subsets of cells within the developing proboscis (figs. 3 and 4; column VII), an elaborately patterned, eversible, muscular feeding apparatus (Kang et al. 2003). The proboscis arises for the most part from several micromere lineages and includes both mesodermal and ectodermal cell types (Huang et al. 2002). *Wnt11b* is exclusively expressed at stages 10 and 11 in the tip of the proboscis, *wnt7* is expressed in both the anterior part of the proboscis and in cells of the nonsegmental, dorsal anterior ganglion.

Regulatory Consequences of *wnt* Gene Duplications in Annelid Evolution

To look for evidence of redundancy, subfunctionalization and or neofunctionalization within *wnt* gene subfamilies that have experienced relatively recent duplications in annelid evolution, we compared the expression of the paralogs of the duplicated *wnt5*, *wnt11*, and *wnt16* genes in *Helobdella* with each other and with their orthologs in the polychaete *Capitella*. In the present work, we have focused on the expression patterns in postgastrulation development.

We first used whole-mount double fluorescent in situ hybridization to determine the extent of overlap in the expression among the *Helobdella* paralogs during stages 9–11 (fig. 6). The most similar expression patterns were observed for the *wnt5* and *wnt16* paralogs. *wnt5a* and *wnt5b* are expressed in slightly overlapping sets of cells in the segmental ectoderm (fig. 6*A*); *wnt5a*, but not *wnt5b*, is also expressed along the length of the inner proboscis (fig. 6*B*). Intriguingly, although the *wnt16* paralogs were both expressed in segmental ganglia and proboscis, they exhibited minimal overlap; in the germinal plate, the *wnt16* paralogs are expressed in transverse stripes within the segmental ganglia ectodermal cells, but the anterior (*wnt16b*) and posterior (*wnt16a*) stripes show little overlap (fig. 6*E*). Both genes were also expressed in the anterior portion of the proboscis but *wnt16a* was observed in the outer sheath and *wnt16b* in the inner ring (fig. 6*F*), with only a small region of overlap.

The most dissimilar expression patterns were observed for the *wnt11* paralogs; whereas *wnt11a* is expressed in a segmentally iterated pattern in both mesoderm and ectoderm of the germinal plate and in the proboscis (fig. 6*C* and *D*), *wnt11b* is expressed exclusively at the tip of the proboscis during stages 10 and 11 (fig. 6*D*), and *wnt11c* is expressed only in the posterior mesoderm of the germinal plate (fig. 6*C*). Although *wnt11a* and *wnt11b* are both expressed in the proboscis, no overlap occurs (fig. 6*D*); however, *wnt11a* and *wnt11c* expression does overlap in the posterior mesoderm (fig. 6*C*).

To compare the expression of the *wnt5*, *wnt11*, and *wnt16* genes between *Helobdella* and *Capitella*, ISH analyses

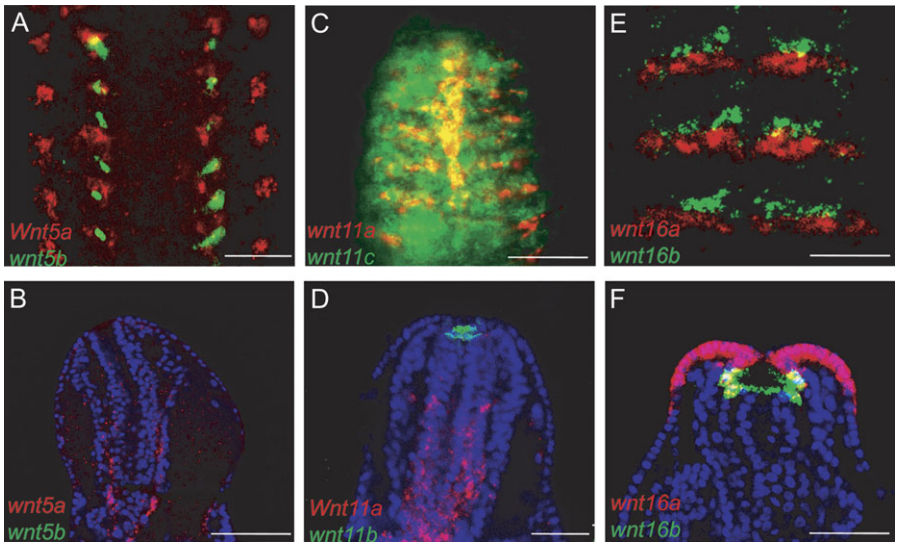

FIG. 6. Expression patterns of the leech-duplicated *wnt* genes in late development stages. Pseudocolored confocal stacks of DFISH-processed embryos showing ventral views of segmental tissues at stage 9 (*A*, *C*, *E*) and lateral views of the proboscis, counterstained with 4′,6-diamidino-2-phenylindole to reveal nuclei (blue) at stage 10 (*B*, *D*, *F*). (*A* and *B*) *wnt5a* (red) and *wnt5b* (green) are both expressed in midbody segments, with little or no overlap; only the *wnt5a* paralog is expressed in the proboscis. (*C*) *wnt11a* (red) and *wnt11c* (green) are expressed in partially overlapping patterns within the caudal segments. (*D*) Within the proboscis, *wnt11a* (red) and *wnt11b* (green) are expressed in discrete sets of posterior and anterior cells, respectively. (*E* and F) *wnt16a* (red) and *wnt16b* (green) are both expressed in neighboring sets of cells with little or no overlap in both midbody segments and at the anterior tip of the proboscis. Scale bars: *A*, *C*, and *E* 100 μm; *B*, *D*, and *F* 50 μm.

of the single subfamily members in *Capitella* were performed (fig. 7). *Capitella* is a marine polychaete annelid that passes through a nonfeeding larval phase following embryogenesis. During larval life, 13 segments are generated and morphogenesis of many of the adult tissues occurs. Embryonic and larval development is categorized by nine stages (Seaver et al. 2005), and we describe expression for the *Capitella wnt5*, *wnt11*, and *wnt16* genes at mid-larval stages (stage 7). Each gene has a unique expression pattern and includes multiple tissues. The *Capitella wnt5* transcript is present in four distinct tissues: the central nervous system, foregut, trunk mesoderm, and posterior growth zone (fig. 7*A*–*D*). In the nervous system, *wnt5* is expressed in a small ventrolateral domain in the brain (fig. 7*A* and *D*) and in a subset of segmentally iterated neurons along the anterior edge of the ventral nerve cord (fig. 7*A* and *B*).

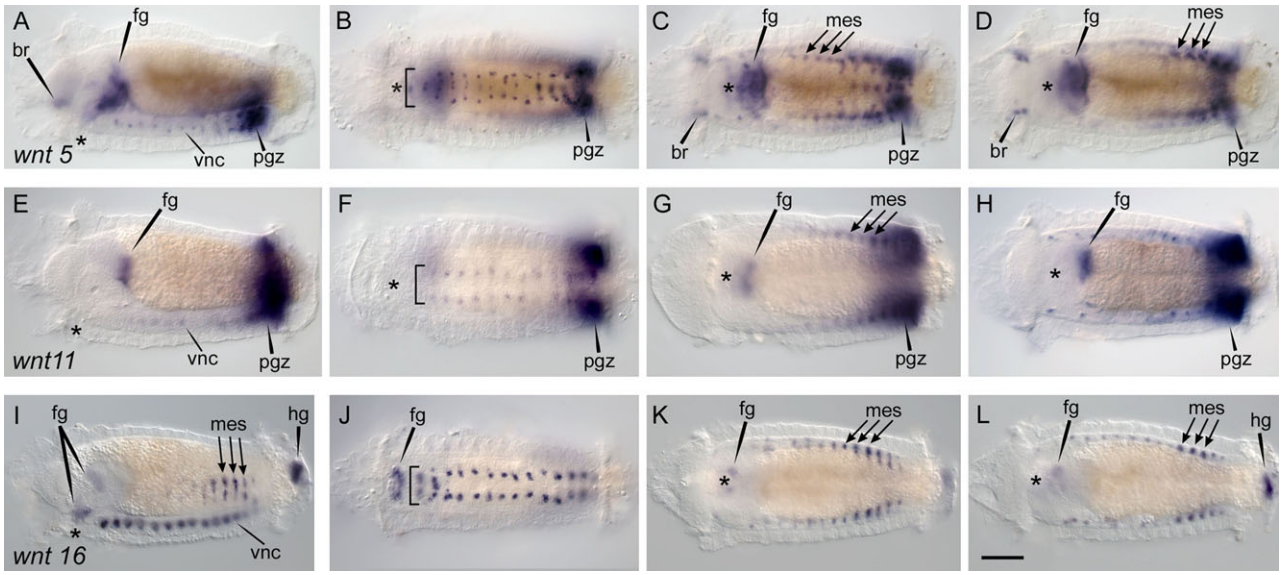

FIG. 7. Expression of *wnt5*, *wnt11* and *wnt16* genes in *Capitella*. Whole-mount ISH of stage 7 *Capitella* larvae showing expression domains of *wnt5* (*A*–*D*), *wnt11* (*E*–*H*), and *wnt16* (*I*–*L*). All panels are anterior to the left. *A*, *E*, and *I* are lateral views, with ventral down; all other panels are ventral views. *B*–*D*, *F*–*H*, and *J*–*L* are series of distinct focal planes, progressing dorsally from the ventral-most plane (*B*, *F*, *J*). *B*–*D* are the same animal and *K*–*L* are the same animal. *I* is a merge of multiple focal planes. Brackets mark the ventral nerve cord, and the mouth is marked with an asterisk, except in *J*, so as not to obscure the in situ signal. Br, brain; fg, foregut; hg, hindgut; mes, mesoderm; pgz, posterior growth zone; vnc, ventral nerve cord. Scale bar, 50 μm.

Expression associated with the foregut is localized to the pharynx and esophagus and is likely to be visceral mesoderm surrounding the foregut (fig. 7*A*, *C*, and *D*). *wnt5* is expressed in a complex pattern of segmentally iterated mesodermal clusters in the trunk (fig. 7*C* and *D*) and mesoderm of the posterior growth zone. For *Capitella wnt11*, the most prominent expression domain is in the ectoderm of the posterior growth zone. Although absent from the ventral midline in the posterior growth zone, expression extends from the dorsal to the ventral side of the body. In the nervous system, *wnt11* expression is limited to bilateral pairs of segmentally repeated clusters of two to three cells each at the posterior of each ganglion in the ventral nerve cord (fig. 7*E* and *F*). Expression in the foregut is restricted to a small domain just to the left of the midline at the foregut–midgut boundary (fig. 7*E* and *H*). *wnt11* mesoderm expression is initially present as a lateral band that resolves into discrete domains positioned at segment boundaries (fig. 7*G*). *Capitella wnt16* is expressed in two distinct domains in the foregut, just subsurface to the mouth opening in the buccal cavity, and in a small region closely associated with the esophagus (fig. 7*I* and *J*). *wnt16* is also prominently expressed in the hindgut (fig. 7*I*). In the ganglia of the ventral nerve cord, expression is localized to small bilateral clusters of 2–3 cells each/segment along the posterior edge of the ganglia. Bands of *wnt16* mesodermal expression are localized to the anterior side of the segment (fig. 7*K* and *L*), most prominent in posterior segments and gradually reduced in more anterior segments.

Discussion

Lophotrochozoan *Wnt* Subfamily Diversity and Genetic Linkage

In the work presented here, we have inventoried the *wnt* genes of the three lophotrochozoan species for which whole genome information is available. Our results reveal wnt subfamily relationships suggestive of duplication events leading to the expansion of the wnt gene family early in metazoan evolution, a conservation of an extensive ancestral wnt gene complement in an extant mollusk and an annelid lophotrochozoan genome and also of an evolutionary dynamic within annelids leading to extensive wnt gene loss and duplication in the leech (fig. 1). Specifically, 12 of the 13 *wnt* subfamilies postulated for the last common ancestor of cnidarians and bilaterian metazoans (Kusserow et al. 2005; Lengfeld et al. 2009) are present in modern lophotrochozoans, whereas the 13th subfamily, *wnt3*, is lacking from all lophotrochozoan and ecdysozoan genomes sequenced so far (Bolognesi et al. 2008). Because *wnt3* is present in the basal cnidarian *Nematostella* and in the deuterostomes (i.e., echinoderms and vertebrates), its absence in lophotrochozoan and ecdysozoan lineages suggests that it was lost early during protostome evolution. In contrast, both mollusk and leech (but not polychaete) lack *wnt8* orthologs, suggesting that this gene was lost independently in the former two lineages given their phylogenetic relationships (Philippe et al. 2009). Similarly, the absence of *wnt9* in *Nematostella* and *Helobdella* suggests that it too was lost independently in these lineages because it is present in the cnidarian *Hydra*, deuterostomes, and other lophotrochozoans.

Three additional aspects of our results suggest that the annelid lineage leading to *Helobdella* has been subject to more rapid genome evolution than those leading to the other two lophotrochozoans examined here. First, the different branch lengths in the molecular phylogenies indicate that *wnt* genes in the *Helobdella* lineage have been evolving faster than those in the *Capitella* and *Lottia* lineages; the greater range in the number of cysteine residues in *Helobdella* relative to the other two species is one example of this trend. Second, the leech lineage has undergone more *wnt* gene duplications and deletions. Third, although our results support the presence in the ancestral bilaterian and lophotrochozoan of two previously postulated *wnt* gene clusters, one comprising *wnt1*, *wnt6*, *wnt9*, and *wnt10* and the other comprising *wnt5* and *wnt7* (fig. 2), these genetic linkages are absent in *Helobdella*, suggesting that the lineage leading to *Helobdella* has undergone more translocations than the other lineages.

Intron gain, loss, and migration are relatively rare events (Irimia and Roy 2008). Thus, depending on the timing of intron gain or loss relative to the expansion of the ancestral genes, we could expect more closely related genes to share specific intron–exon organizations. Alternatively, the presence or absence of introns within one lineage relative to others provides support for certain gene clades and provides information as to when introns were gained or lost. Our analyses of intron–exon boundaries (supplementary fig. S1, Supplementary Material online) does not provide any information about the *wnt* subfamily relationships, suggesting that the ur-*wnt* already contained at least two introns and that the expansion of the *wnt* family from one to 13 genes was relatively rapid, occurring without intron gain or loss in any of the intermediates. Subsequent gains or losses of introns occurred within specific lineages and thus support specific metazoan clades, that is, the Bilateria (by the presence of two introns, one in *wnt2* and the other in *wnt11* that are shared by all bilaterian lineages used in this analysis) and the Lophotrochozoa (by the intron in *wnt1* that is seen only in this group).

Conservation and Diversification of *wnt* Paralogs in Annelida

Within the diverse phylum Annelida, *Capitella* and *Helobdella* are distantly related genera (Struck et al. 2007; Zrzavy et al. 2009) and differ extensively in terms of morphology and development (Sawyer 1986). *Capitella*, for example, undergoes semidirect development, producing a nonfeeding larva that metamorphoses into a juvenile worm, whereas leech undergoes direct development, skipping the larval stage altogether (Weisblat and Huang 2001; Seaver et al. 2005). Nonetheless, their development and anatomy can be directly compared at several points.

First, both undergo spiral cleavage in early development. Second, both form segmental mesoderm and ectoderm in

anteroposterior progression due to the proliferation of cells in the posterior growth zone (Weisblat and Huang 2001; Seaver et al. 2005). Third, the central nervous system of both annelids includes a dorsal anterior ganglion and a ventral nerve cord consisting of one discrete ganglion per segment (Shain et al. 1998; Meyer and Seaver 2009). Fourth, both form the midgut endoderm largely from macromeres internalized during gastrulation (Eisig 1899; Weisblat and Huang 2001) and the foregut from micromere derivatives (Huang et al. 2002). These developmental similarities between *Helobdella* and *Capitella* allow us to investigate the regulatory consequences of *wnt* gene duplication events by comparing the expression patterns of the *wnt5*, *wnt11*, and *wnt16* subfamily members in the gut, mesoderm, posterior growth zone, and ventral nerve cord in these two species.

In general, evolutionary gene duplication is associated with one of the following consequences: 1) nonfunctionalization, that is, the loss of one paralog, which if gradual may be evidenced by a nonfunctional gene; 2) subfunctionalization, in which the function of the ancestral gene becomes divided between the two paralogs; 3) neofunctionalization, in which one paralog acquires new functions, whereas the other retains the ancestral function; and 4) synfunctionalization, in which one paralog acquires the function of a third, less closely related gene, so that gene functions are transferred from one gene to another without changing the processes in which they participate (Gitelman 2007).

Investigating these possibilities by comparing gene expression patterns in two (or more) modern species is subject to the important caveat that the modern species have all evolved since their last common ancestor, so the ancestral expression pattern of the gene in question cannot be known, even for the species in which no duplication has occurred. Moreover, expression patterns can only hint at the functions of the gene(s) in question. Finally, with the obvious exception of nonfunctionalization, these processes are not mutually exclusive, for example, subfunctionalization can be a first step toward neofunctionalization (Ohno 1970). Within these limitations, however, comparison of the expression patterns of orthologous *wnt* genes between *Helobdella* and *Capitella* are consistent with the possibilities outlined above.

Specifically, in *Capitella*, *wnt11* is expressed in distinct anterior and posterior sites within the foregut, in both mesoderm and ventral nerve cord within the trunk, and in the posterior growth zone. These expression patterns correlate well with the superposition of the largely nonoverlapping expression patterns exhibited by the three *wnt11* paralogs in *Helobdella*. Thus, comparison of the expression patterns of the *wnt11* genes between *Helobdella* and *Capitella* suggests that spatial and temporal subfunctionalization of an ancestral *wnt11* gene expression has occurred and that such expression is now accomplished by the ensemble of the three *wnt11* paralogs in *Helobdella*.

In contrast to the situation with *wnt11*, the expression patterns of the *wnt16* paralogs in *Helobdella* are strikingly similar to each other and to the *Capitella* ortholog; both genes are expressed in segmental ectoderm and at the anterior end of the foregut. Within these two tissues, *Helobdella wnt16a* and *wnt16b* exhibit minimal overlap, which could be interpreted as reflecting either spatial subfunctionalization or neofunctionalization of the duplicated *wnt16* paralogs. The differences in *wnt16* expression between *Capitella* and *Helobdella* (mesodermal expression in the trunk and hindgut of *Capitella* but not *Helobdella*; expression in the posterior growth zone of *Helobdella* but not *Capitella*) are more difficult to interpret because they could represent a variety of gains and/or losses in either species with respect to their last common ancestor.

The parallel bands of *wnt16* positive cells in the segmental ganglia of *Helobdella*, with *wnt16b*-expressing cells lying just anterior to the *wnt16a*-expressing cells are of particular interest because this pattern resembles the distribution of the respective ns and nf blast cell clones (Zhang and Weisblat 2005). But the gene expression patterns do not coincide with the clone distributions. There are some N-derived cells that do not express either *wnt16* gene and some non–N-derived ganglionic cells that do express *wnt16a* or *wnt16b*.

Comparison of the *wnt5* expression patterns is difficult to interpret. Expression in the posterior growth zone and foregut region is common to *Capitella wnt5* and *Helobdella wnt5a*, which could be suggestive of a *wnt5* expression pattern ancestral to annelids. If so, the largely nonoverlapping expression pattern of *Helobdella wnt5a* and *wnt5b* could be indicative of neofunctionalization for *wnt5b*. As in the case of the *wnt16* paralogs, the noncongruent *wnt5* pattern elements (*Capitella wnt5* in the brain, ventral nerve cord, and mesodermal patches in the ventral trunk; *Helobdella wnt5a* and *wnt5b* in ectoderm lateral to the nerve cord) could reflect various scenarios of gained or lost traits. In any event, the minimally overlapping patterns of *wnt5* paralogs in late stage of *Helobdella* development must reflect regulatory changes following *wnt* gene duplication.

In conclusion, although bioinformatic analyses failed to detect positive selection, nonoverlapping expression patterns of the duplicated *wnt* genes in stages 9–11 of *Helobdella* indicate the occurrence of significant regulatory diversification between these duplicates. Combined with the extensively overlapping expression patterns of many of the *Helobdella wnt* genes during early development, these results also suggest that the *wnt* genes might provide an interesting example of modular enhancers described in other systems (Burch 2005), in which shared, conserved modules regulate early expression and more divergent gene-specific enhancers regulate later expression in specific tissues. Further studies are required to determine the extent to which this regulatory diversification has been accompanied by functional diversification and to assess the significance of overlapping in early expression patterns. By analyzing two annelid species, we identified expression domains for *wnt5*, *wnt11*, and *wnt16* that are likely to have been present in the annelid ancestor. These results also

significantly add to the data available for comparative analyses of the regulatory and functional plasticity of *wnt* genes in metazoan evolution.

Supplementary Material

Supplementary tables S1–S3 and figure S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to P. Francino and R. Nielsen for insightful discussions, E. Begovic, D. H. Kuo, and S. E. Gline for manuscript comments and the Levine Lab and T. W. Kim for access to the confocal microscopes. We thank Emi Yamaguchi, Michael Boyle, and Gemma Richards for assistance in cloning and performing ISH for the *Capitella wnt5* gene. This work was supported by grants from National Institutes of Health (GM 074619) to D.A.W. and National Science Foundation (IOB05-44869) to E.C.S. The authors declare no conflict of interest.

References

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21:2104–2105.

Blake JA, Grassle JP, Eckelbarger KJ. 2009. Capitella teleta, a new species designation for the opportunistic and experimental capitellid, Capitella sp. I, with a review of the literature for confirmed records. *Zoosymposia* 2:25–53.

Bolognesi R, Farzana L, Fischer TD, Brown SJ. 2008. Multiple Wnt genes are required for segmentation in the short-germ embryo of Tribolium castaneum. *Curr Biol.* 18:1624–1629.

Burch JB. 2005. Regulation of GATA gene expression during vertebrate development. *Semin Cell Dev Biol.* 16:71–81.

Cadigan KM, Nusse R. 1997. Wnt signaling: a common theme in animal development. *Genes Dev.* 11:3286–3305.

Carroll SB, Grenier JK, Weatherbee SD. 2001. From DNA to diversity: molecular genetics and the evolution of animal design. Malden (MA): Blackwell Science.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17:540–552.

Croce JC, Wu SY, Byrum C, Xu R, Duloquin L, Wikramanayake AH, Gache C, McClay DR. 2006. A genome-wide survey of the evolutionarily conserved Wnt pathways in the sea urchin Strongylocentrotus purpuratus. *Dev Biol.* 300:121–131.

Eisenmann DM. 2005. Signal transduction: Wnt signaling. In: WormBook, editor. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.7.1. Available from: http://www.wormbook.org.

Eisig H. 1899. Zur Entwicklungsgeschichte der Capitelliden. *Mitth Zool Stat Zu Neapel.* 13:1–292.

Emes RD, Yang Z. 2008. Duplicated paralogous genes subject to positive selection in the genome of Trypanosoma brucei. *PLoS One.* 3:e2295.

Garriock RJ, Warkman AS, Meadows SM, D'Agostino S, Krieg PA. 2007. Census of vertebrate Wnt genes: isolation and developmental expression of Xenopus Wnt2, Wnt3, Wnt9a, Wnt9b, Wnt10a, and Wnt16. *Dev Dyn.* 236:1249–1258.

Gitelman I. 2007. Evolution of the vertebrate twist family and synfunctionalization: a mechanism for differential gene loss through merging of expression domains. *Mol Biol Evol.* 24: 1912–1925.

Gonsalves FC, Weisblat DA. 2007. MAPK regulation of maternal and zygotic Notch transcript stability in early development. *Proc Natl Acad Sci U S A.* 104:531–536.

Gu Z, Rifkin SA, White KP, Li WH. 2004. Duplicate genes increase gene expression diversity within and between species. *Nat Genet.* 36:577–579.

Huang FZ, Bely AE, Weisblat DA. 2001. Stochastic WNT signaling between nonequivalent cells regulates adhesion but not fate in the two-cell leech embryo. *Curr Biol.* 11:1–7.

Huang FZ, Kang D, Ramirez-Weber FA, Bissen ST, Weisblat DA. 2002. Micromere lineages in the glossiphoniid leech Helobdella. *Development* 129:719–732.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R. 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics*. 8:460.

Irimia M, Roy SW. 2008. Spliceosomal introns as tools for genomic and evolutionary analysis. *Nucleic Acids Res.* 36:1703–1712.

Kang D, Huang F, Li D, Shankland M, Gaffield W, Weisblat DA. 2003. A hedgehog homolog regulates gut formation in leech (Helobdella). *Development* 130:1645–1657.

Kusserow A, Pang K, Sturm C, et al. (11 co-authors). 2005. Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* 433:156–160.

Larkin MA, Blackshields G, Brown NP, et al. (13 co-authors). 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948.

Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, Schmidt HA, Ozbek S, Bode H, Holstein TW. 2009. Multiple Wnts are involved in Hydra organizer formation and regeneration. *Dev Biol.* 330:186–199.

Li W, Tucker AE, Sung W, Thomas WK, Lynch M. 2009. Extensive, recent intron gains in Daphnia populations. *Science* 326:1260–1262.

Li WH, Yang J, Gu X. 2005. Expression divergence between duplicate genes. *Trends Genet.* 21:602–607.

Martin BL, Kimelman D. 2009. Wnt signaling and the evolution of embryonic posterior development. *Curr Biol.* 19:R215–R219.

Meyer NP, Seaver EC. 2009. Neurogenesis in an annelid: characterization of brain neural precursors in the polychaete Capitella sp. I. *Dev Biol.* 335:237–252.

Nusse R, Varmus HE. 1992. Wnt genes. *Cell* 69:1073–1087.

Ohno S. 1970. Evolution by gene duplication. New York: Springer-Verlag.

Philippe H, Derelle R, Lopez P, et al. (20 co-authors). 2009. Phylogenomics revives traditional views on deep animal relationships. *Curr Biol.* 19:706–712.

Prud'homme B, Lartillot N, Balavoine G, Adoutte A, Vervoort M. 2002. Phylogenetic analysis of the Wnt gene family. Insights from lophotrochozoan members. *Curr Biol.* 12:1395.

Putnam NH, Butts T, Ferrier DE, et al. (37 co-authors). 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453:1064–1071.

Rivera AS, Gonsalves FC, Song MH, Norris BJ, Weisblat DA. 2005. Characterization of Notch-class gene expression in segmentation stem cells and segment founder cells in Helobdella robusta (Lophotrochozoa; Annelida; Clitellata; Hirudinida; Glossiphoniidae). *Evol Dev.* 7:588–599.

Sawyer RT. 1986. Leech biology and behaviour. Oxford: Clarendon Press.

Schubert M, Holland LZ, Holland ND, Jacobs DK. 2000. A phylogenetic tree of the Wnt genes based on all available full-length sequences, including five from the cephalochordate amphioxus. *Mol Biol Evol.* 17:1896–1903.

Seaver EC, Kaneshige LM. 2006. Expression of 'segmentation' genes during larval and juvenile development in the polychaetes Capitella sp. I and H. elegans. *Dev Biol.* 289:179–194.

Seaver EC, Paulson DA, Irvine SQ, Martindale MQ. 2001. The spatial and temporal expression of Ch-en, the engrailed gene in the polychaete Chaetopterus, does not support a role in body axis segmentation. *Dev Biol.* 236:195–209.

Seaver EC, Thamm K, Hill SD. 2005. Growth patterns during segmentation in the two polychaete annelids, Capitella sp. I and Hydroides elegans: comparisons at distinct life history stages. *Evol Dev.* 7:312–326.

Shain DH, Ramirez-Weber FA, Hsu J, Weisblat DA. 1998. Gangliogenesis in leech: morphogenetic processes leading to segmentation in the central nervous system. *Dev Genes Evol.* 208:28–36.

Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 21:456–463.

Struck TH, Schult N, Kusen T, Hickman E, Bleidorn C, McHugh D, Halanych KM. 2007. Annelid phylogeny and the status of Sipuncula and Echiura. *BMC Evol Biol.* 7:57.

Weisblat D, Huang F. 2001. An overview of glossiphoniid leech development. *Can J Zool.* 79:218–232.

Weisblat DA, Shankland M. 1985. Cell lineage and segmentation in the leech. *Philos Trans R Soc Lond B Biol Sci.* 312:39–56.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol.* 18:691–699.

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol.* 20:1377–1419.

Zhang SO, Weisblat DA. 2005. Applications of mRNA injections for analyzing cell lineage and asymmetric cell divisions during segmentation in the leech Helobdella robusta. *Development* 132:2103–2113.

Zrzavy J, Riha P, Pialek L, Janouskovec J. 2009. Phylogeny of Annelida (Lophotrochozoa): total-evidence analysis of morphology and six genes. *BMC Evol Biol.* 9:189.