Brachiopods possess a split Hox cluster with signs of spatial, but not temporal collinearity

Sabrina M. Schiemann[1], José M. Martín-Durán[1], Aina Børve[1], Bruno C. Vellutini[1], Yale J. Passamaneck[2], Andreas Hejnol[1*]

[1] Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen 5008, Norway

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

*Corresponding author: Andreas Hejnol (andreas.hejnol@uib.no)

Abstract

Hox genes are often clustered in animal genomes and exhibit spatial and/or temporal collinearity. It is generally believed that temporal collinearity is the major force preserving Hox clusters. However, studies combining genomic and gene expression analyses of Hox genes are scarce, particularly within Spiralia and Lophotrochozoa (e.g. mollusks, segmented worms, and flatworms). Here, we use two brachiopod species –*Terebratalia transversa* and *Novocrania anomala*– that respectively belong to the two major brachiopod lineages to characterize their Hox complement, the presence of a Hox cluster, and the temporal and spatial expression of their Hox genes. We demonstrate that the Hox complement consists of ten Hox genes in *T. transversa* (*lab*, *pb*, *Hox3*, *dfd*, *scr*, *lox5*, *antp*, *lox4*, *post2* and *post1*) and nine in *N. anomala* (missing *post1*). Additionally, *T. transversa* has an ordered, split Hox cluster. Expression analyses reveal that Hox genes are neither temporally nor spatially collinear, and only the genes *pb* (in *T. transversa*), *Hox3* and *dfd* (in both brachiopods) show staggered expression in the mesoderm. Remarkably, *lab*, *scr*, *antp* and *post1* are associated with the development of the chaetae and shell-forming epithelium, as also observed in annelid chaetae and mollusk shell fields. This, together with the expression of *Arx* homeobox, supports the deep conservation of the molecular basis for chaetae formation and shell patterning in Lophotrochozoa. Our findings challenge the current evolutionary scenario that (temporal) collinearity is the major mechanism preserving Hox clusters, and suggest that Hox genes were involved in the evolution of lophotrochozoan novelties.

Introduction

Hox genes are transcription factors that bind to regulatory regions via a helix-turn-helix domain to enhance or suppress gene transcription (McGinnis and Krumlauf 1992; Pearson, et al. 2005). Hox genes were initially described in the fruit fly *Drosophila melanogaster* (Lewis 1978; McGinnis, Levine, et al. 1984) and later on in vertebrates (Carrasco, et al. 1984; McGinnis, Garber, et al. 1984; McGinnis, Hart, et al. 1984) and the nematode *Caenorhabditis elegans* (Costa, et al. 1988). In all these organisms, Hox genes were shown to provide a spatial coordinate system for cells along the anterior-posterior axis (Akam 1989). Remarkably, the Hox genes of these organisms are clustered in their genomes and exhibit a staggered spatial (Lewis 1978) and temporal (Dollé, et al. 1989; Izpisúa-Belmonte, et al. 1991) expression during embryogenesis that corresponds to their genomic arrangement (Lewis 1978; Duboule and Morata 1994; Lemons and McGinnis 2006). These features were used to classify Hox genes in four major orthologous groups –anterior, Hox3, central and posterior Hox genes– and were proposed to be ancestral attributes to all bilaterally symmetrical animals (McGinnis and Krumlauf 1992; Garcia-Fernàndez 2005; Lemons and McGinnis 2006).

However, the study of the genomic arrangements and expression patterns of Hox genes in a broader phylogenetic context has revealed multiple deviations from that evolutionary scenario. Hox genes are prone to gains (de Rosa, et al. 1999; Simakov, et al. 2013; Zwarycz, et al. 2016) and losses (Aboobaker and Blaxter 2003a; Aboobaker and Blaxter 2003b; Tsai, et al. 2013; Smith, et al. 2016), and their arrangement in a cluster can be interrupted, or even completely disintegrated (Seo, et al. 2004; Duboule 2007; Albertin, et al. 2015; Serano, et al. 2016). Furthermore, the collinear character

of the Hox gene expression can fade temporally (Lowe and Wray 1997; Irvine and Martindale 2000; Seo, et al. 2004) and/or spatially (Lee, et al. 2003). Hox genes have also diversified their roles during development, extending beyond providing spatial information. In many bilaterian embryos, Hox genes are expressed during early development, well before the primary body axis is patterned (Wada, et al. 1999; Irvine and Martindale 2000; Aronowicz and Lowe 2006; Hejnol and Martindale 2009). They are also involved in patterning different tissues (Chauvet, et al. 2000) and have been often recruited for the evolution and development of novel morphological traits, such as vertebrate limbs (Zakany and Duboule 2007; Woltering and Duboule 2015), cephalopod funnels and arms (Lee, et al. 2003), and beetle horns (Wasik, et al. 2010).

It is thus not surprising that Hox genes show diverse arrangements regarding their genomic organization and expression profiles in the Spiralia (Barucca, et al. 2016), a major animal clade that includes a high disparity of developmental strategies and body organizations (Hejnol 2010; Dunn, et al. 2014; Struck, et al. 2014; Laumer, et al. 2015). An example is the bdelloid rotifer *Adineta vaga,* which belongs to the Gnathifera, the possible sister group to all remaining Spiralia (Struck, et al. 2014; Laumer, et al. 2015). As a result of their reduced tetraploidy, its Hox complement includes 24 genes, albeit it lacks posterior Hox genes and a Hox cluster (Flot, et al. 2013). The freshwater flatworms *Macrostomum lignano* and *Schmidtea mediterranea* also lack a Hox cluster (Wasik, et al. 2015; Currie, et al. 2016) and parasitic flatworms have undergone extensive Hox gene losses, likely associated with their particular life style (Tsai, et al. 2013). Interestingly, the limpet mollusk *Lottia gigantea* (Simakov, et al. 2013) shows a well-organized Hox cluster. Other mollusks (e.g. the pacific oyster *Crassostrea gigas*) and the segmented annelid *Capitella teleta*

exhibit organized split Hox clusters (Fröbius, et al. 2008; Zhang, et al. 2012). On the other hand, the cephalopod mollusk *Octopus bimaculiodes* has lost several Hox genes and lacks a Hox cluster (Albertin, et al. 2015); and the clitellate annelids *Helobdella robusta* and *Eisenia fetida* do not show a Hox cluster and have greatly expanded some of the Hox classes (Simakov, et al. 2013; Zwarycz, et al. 2016).

Although Hox gene expression is known for a handful of spiralian species (Kourakis, et al. 1997a, b; Irvine and Martindale 2000; Irvine and Martindale 2001; Kourakis and Martindale 2001; Hinman, et al. 2003; Fröbius, et al. 2008; Samadi and Steiner 2009, 2010; Fritsch, et al. 2015; Hiebert and Maslakova 2015a, b; Currie, et al. 2016; Fritsch, et al. 2016), the relationship between genomic organization and expression domains is known for only three of them, namely the annelids *C. teleta* and *H. robusta*, and the planarian *S. mediterranea*. Consistent with the lack of a Hox cluster, *H. robusta* and *S. mediterranea* show neither temporal nor spatial collinearity (Kourakis, et al. 1997a, b; Kourakis and Martindale 2001; Currie, et al. 2016). Conversely, *C. teleta*, which has an organized, broken cluster, does exhibit these features (Fröbius, et al. 2008). These observations support that the presence of collinearity –in particular, temporal collinearity– is associated with the retention of a more or less intact Hox cluster (Duboule 1994; Ferrier and Minguillon 2003; Garcia-Fernàndez 2005; Duboule 2007). However, more studies combining genomic and expression information, and including the vast spiralian morphological diversity, are essential to draw robust conclusions about Hox gene evolution and regulation in Spiralia and Metazoa (Monteiro and Ferrier 2006) and to test hypotheses about the correlation between collinearity and cluster organization (Duboule 2007).

Here, we present a comprehensive study of the genomic arrangement and expression of Hox genes in Brachiopoda, a lineage of the Spiralia whose origins date back to the Lower Cambrian (Rudwick 1970). Brachiopods are marine, sessile, filter-feeding animals. They are protected by two dorsoventral mineralized shells and reproduce by external fertilization, often developing through an intermediate, free-living larval stage (Brusca, et al. 2016). In this study, we use two brachiopod species –the 'articulate' *Terebratalia transversa* and the 'inarticulate' *Novocrania anomala*– that respectively belong to the two major brachiopod lineages, thus allowing the reconstruction of putative ancestral characters for Brachiopoda as a whole. By transcriptomic and genomic sequencing we demonstrate that the Hox complement consists of ten Hox genes in *T. transversa* and nine in *N. anomala*. In addition, the ten *Hox* genes of *T. transversa* are ordered in a split Hox cluster that differs from the genomic arrangement reported for the brachiopod *Lingula anatina* (Luo, et al. 2015). We show that Hox genes are restricted to the 'trunk' region of the larva, and are overall neither temporally nor spatially collinear. However, the genes *pb* (only in *T. transversa*), *Hox3* and *dfd* show spatially collinear expression in the mesoderm of both brachiopod species. Additionally, the Hox genes *lab*, *scr*, *antp* and *post1* appear to be associated with the development of two brachiopod features: the chaetae and the shell-forming epithelium. Altogether, our findings demonstrate that the presence of a split Hox cluster in the Brachiopoda is not associated with a temporally collinear expression of Hox genes, which challenges the hypothesized correlation between temporal collinearity and the retention of a Hox cluster (Duboule 1994; Ferrier and Minguillon 2003; Garcia-Fernàndez 2005; Duboule 2007) and suggests that alternative/additional genomic forces might shape Hox clusters during animal evolution.

Results

The Hox *gene complement of* T. transversa *and* N. anomala

Transcriptomic and genomic searches resulted in the identification of ten Hox genes in *T. transversa*. In the brachiopod *N. anomala*, we identified seven Hox genes in the transcriptome and two additional fragments corresponding to a Hox homeodomain in the draft genome assembly. Attempts to amplify and extend these two genomic sequences in the embryonic and larval transcriptome of *N. anomala* failed, suggesting that these two Hox genes might be expressed only during metamorphosis and/or in the adult brachiopod.

Maximum likelihood orthology analyses resolved the identity of the retrieved Hox genes (Figure 1; Supplementary Figure S1). The ten Hox genes of *T. transversa* were orthologous to *labial* (*lab*), *proboscipedia* (*pb*), *Hox3*, *deformed* (*dfd*), *sex combs reduced* (*scr*), *lox5*, *antennapedia* (*antp*), *lox4*, *post2* and *post1*. The nine Hox genes identified in *N. anomala* corresponded to *lab*, *pb*, *Hox3*, *dfd*, *scr*, *lox5*, *antp*, *lox4*, and *post2*. Therefore, *T. transversa* has a Hox complement similar to the one described in the brachiopod *L. anatina* (Luo, et al. 2015), while *N. anomala* lacks the *post1* Hox gene.

Genomic organization of Hox *genes in* T. transversa *and* N. anomala

We used the draft assemblies of *T. transversa* and *N. anomala* genomes to investigate the genomic arrangement of their Hox genes. In *T. transversa*, we identified three scaffolds containing Hox genes (Figure 2A). Scaffold A spanned 81.7 kb and contained *lab* and *pb* in a genomic region of 15.4 kb, flanked by other genes with no

known linkage to the Hox cluster in other animals. Scaffold B was the longest (284.8 kb) and included *Hox3*, *dfd*, *scr*, *lox5*, *antp*, *lox4* and *post2*, in this order (Figure 2A) including the micro RNA *mir-10* between *dfd* and *scr*. As in scaffold A, other genes flanked the Hox genes, which occupied a genomic region of 76.2 kb. Finally, *post1* aligned to various short scaffolds. We could not recover any genomic linkage between the identified Hox genes in *N. anomala* due to the low contiguity (N50 of 3.5 kb) of the draft genome assembly. Altogether, these data demonstrate that *T. transversa* has a split Hox cluster broken into three sub-clusters, each of them with an organized arrangement. Importantly, the potential genomic disposition of these three sub-clusters is similar to that observed in other spiralians, such as *C. teleta* and *L. gigantea* (Figure 2B), which suggests that the lineage leading to the brachiopod *L. anatina* experienced genomic rearrangements that modified the ordered and linkage of the Hox genes.

Hox *gene expression in* T. transversa

To investigate the presence of temporal and/or spatial collinearity in the expression of the clustered Hox genes in *T. transversa*, we first performed whole-mount *in situ* hybridizations in embryos from blastula to late, competent larval stages (Figure 3).

Anterior Hox genes

The anterior Hox gene *lab* was first detected in the mid gastrula stage in two faint bilaterally symmetrical dorsal ectodermal domains (Figure 3Ad, Ae). In late gastrulae, *lab* expression consisted of four dorsal ectodermal clusters that corresponded to the position where the chaetae sacs form (Figure 3Af, Ag). In early larva, the expression was strong and broad in the mantle lobe (Figure 3Ah, Ai), and in late larvae it became

restricted to a few mantle cells adjacent to the chaetae sacs (Figure 3Ij, Ik). These cells do not co-localize with tropomyosin, which labels the muscular mesoderm of the larva (Figure 4A). This suggests that *lab* expressing cells are likely ectodermal, although we cannot exclude localization in non-muscular mesodermal derivates.

The Hox gene *pb* was first detected asymmetrically on one lateral of the ectoderm of the early gastrula (Figure 3Bb, Bc). In the mid gastrula, the ectodermal domain located dorsally and extended as a transversal stripe (Figure 3Bd, Be). Remarkably, this domain disappeared in late gastrula embryos, where *pb* was detected in the anterior mantle mesoderm (Figure 3Bf, Bg). This expression was kept in the early and late larva (Figure 3Bh–Bk; Figure 4B)

Hox3

The gene *Hox3* was detected already in blastula embryos in a circle of asymmetric intensity around the gastral plate (Figure 3Ca). In early gastrulae, Hox3 is restricted to one half of the vegetal one, which is the prospective posterior side (Figure 3Cb, Cc). With axial elongation, *Hox3* becomes expressed in the anterior mantle mesoderm and in the ventral ectoderm limiting the apical and mantle lobe (Figure 3Cd, Ce). This expression is maintained in late gastrula stages and in the early larva (Figure 3Cf–Ci). In the late larva, *Hox3* is detected in part of ventral, internal mantle ectoderm and in the most anterior part of the pedicle mesoderm (Figure 3Cj, Ck; Figure 4C)

Central Hox genes

The Hox gene *dfd* was asymmetrically expressed on one side of the vegetal pole of the early gastrula of *T. transversa* (Figure 3Db, Dc). This expression was maintained

in the mid gastrula, and corresponded to the most posterior region of the embryo (Figure 3Dd, De). In the late gastrula, *dfd* becomes strongly expressed in the posterior mesoderm (Figure 3Df, Dg). In the early larva, the expression remained in the pedicle mesoderm, but new domains in the posterior ectoderm and in the anterior, ventral pedicle ectoderm appear (Figure 3Dh, Di). These expression domains are also observed in the late larva (Figure 3Dj, Dk; Figure 4D).

The central Hox gene *scr* was first expressed in the medial dorsal ectoderm of the mid gastrula (Figure 3Ed, Ee). In late gastrula stages, the expression expanded towards the ventral side, forming a ring (Figure 3Ef, Eg). In the early larva, *scr* was detected in a ring encircling the most anterior ectoderm of the pedicle lobe and extending anteriorly on its dorsal side (Figure 3Eh, Ei). With the outgrowth of the mantle lobe in the late larva, the expression became restricted to the periostracum, the internal ectoderm of the mantle lobe that forms the shell (Figure 3Ej, Ek; Figure 4E).

The Hox gene *Lox5* is expressed on one side of the early gastrula (Figure 3Fb, Fc). During axial elongation, the expression became restricted to the most posterior ectoderm of the embryo (Figure 3Fd–Fg). This domain remained constant in larval stages, where it was expressed in the whole posterior ectoderm of the pedicle lobe (Figure 3Fh–Fk).

The *antp* gene is weakly detected at the mid gastrula stage, in one posterior ectodermal domain and one dorsal ectodermal patch (Figure 3Gd, Ge). In the late gastrula, the posterior expression is maintained and the dorsal domain extends ventrally, encircling the embryo (Figure 3Gf, Gg). These two domains remained in

the larvae: the ectodermal anterior-most, ring-like domain localized to the periostracum, and the posterior domain limited to the most posterior tip of the larva (Figure 3Gh–Gk).

The Hox gene *Lox4* is first detected in the dorsal, posterior end of the late gastrula and early larva (Figure 3Hf–Hi). In the late larva, *Lox4* is expressed dorsally and posteriorly, although it is absent from the most posterior end (Figure 3Hj, Hk).

Posterior Hox genes

The posterior Hox gene *post2* was first detected in mid gastrula stages at the posterior tip of the embryo (Figure 3Id, Ie). This expression was maintained in late gastrulae (Figure 3If, Ig). In early larva, *post2* expression extended anteriorly and occupied the dorso-posterior midline of the pedicle lobe (Figure 3Ih, Ii). In late, competent larvae, *post2* was detected in a T-domain in the dorsal side of the pedicle ectoderm (Figure 3Ij, Ik).

The Hox gene *post1* was transiently detected in late gastrula stages in the four mesodermal chaetae sacs (Figure 3Jf, Jg).

We verified the absence of temporal collinearity in the expression of the Hox genes in *T. transversa* by quantitative real-time PCR and comparative stage-specific RNA-seq data (Supplementary Figure S2).

Hox *gene expression in* N. anomala

In order to infer potential ancestral Hox expression domains for the Brachiopoda, we

investigated the expression of the nine Hox genes of *N. anomala* during embryogenesis and larval stages (Figure 5).

Anterior Hox genes

The Hox gene *lab* was first detected at the mid gastrula stage in three bilaterally symmetrical ectodermal cell clusters that appear to correlate with the presumptive site of chaetae sac formation (Figure 5Ad, Ae). The expression in the most posterior pair was stronger than in the two most anterior ones. This expression was maintained in the late gastrula (Figure 5Af, Ag). In larval stages, *lab* was detected in the two most anterior chaetae sacs of the mantle lobe (Figure 5Ah, Ai), expression that fainted in late larvae (Figure 5Aj, Ak).

The Hox gene *pb* was asymmetrically expressed already at blastula stages, in the region that putatively will rise to the most posterior body regions (Figure 5Ba). With the onset of gastrulation, the expression of *pb* extended around the vegetal pole, almost encircling the whole blastoporal rim (Figure 5Bb, Bc). During axial elongation, *pb* was first broadly expressed in the region that forms the mantle lobe (Figure 5Bd, Be) and later on the ventral mantle ectoderm of the late gastrula (Figure 5Bf, Bg). In early larvae, *pb* was detected in the anterior ventral mantle ectoderm (Figure 5Bh, Bi). This domain was not detected in late, competent larvae (Figure 5Bj, Bk).

Hox3

The Hox gene *Hox3* was asymmetrically detected around half of the vegetal pole of the early gastrulae (Figure 5Cb, Cc). In mid gastrulae, the expression almost encircled

the whole posterior area and the blastoporal rim (Figure 5Cd). In addition, a domain in the mid-posterior mesoderm became evident (Figure 5Ce). By the end of the axial elongation, *Hox3* was strongly expressed in the posterior mesoderm and weakly in the ventral posterior mantle ectoderm (Figure 5Cf, Cg). Noticeably, the posterior most ectoderm did not show expression of *Hox3*. This expression pattern was maintained in early and late larval stages (Figure 5Ch–Ck).

Central Hox genes

The central Hox gene *dfd* was first detected in the posterior ectodermal tip of mid gastrulae (Figure 5Dd, De). In late gastrula stages, *dfd* was expressed in the posterior ectodermal end (Figure 5Df) and in the posterior mesoderm (Figure 5Dg). Early larvae showed expression of *dfd* in the posterior mesoderm and posterior mantle ectoderm (Figure 5Dh, Di). This expression remained in late larvae, although the most posterior ectodermal end was devoid of expression (Figure 5Dj, Dk).

The Hox gene *scr* was only detected in late larval stages, in a strong dorsal ectodermal domain (Figure 5Ej, Ek).

The gene *Lox5* was detected asymmetrically around half of the blastoporal rim in early gastrula stages (Figure 5Fb, Fc). During axial elongation, the expression progressively expanded around the blastoporal rim (Figure 5Fd, Fe) and limited to the ventral midline (Figure 5Ff, Fg). In the larvae, *Lox5* was expressed in the ventral, posterior-most midline (Figure 5Fh–Fk).

The Hox gene *antp* was first expressed asymmetrically in one lateral side of the early

gastrula (Figure 5Gj, Gk). In the mid gastrula, *antp* was detected in the dorsal ectodermal mantle in a cross configuration: dorsal midline and the mantle cells closer to the apical-mantle lobe boundary (Figure 5Gd, Ge). In late gastrulae, *antp* was only expressed in a mid-dorsal ectodermal region (Figure 5Gf, Gg). This expression pattern was also observed in early larval stages, although the size of the domain reduced (Figure 5Gh, Gi). In late larvae, antp was detected in a small mid-dorsal patch and a weak ventro-posterior ectodermal domain (Figure 5Gj, Gk).

We could neither identify nor amplify *Lox4* in a transcriptome and cDNA obtained from mixed embryonic and larval stages, suggesting that either it is very transiently and weakly expressed during embryogenesis or it is only expressed in later stages (metamorphosis and adulthood).

Posterior Hox genes

The only posterior Hox gene present in *N. anomala*, *post2*, could not be amplified in cDNA obtained from mixed embryonic and larval stages, suggesting that it is not expressed –or at least expressed at really low levels– during these stages of the life cycle.

Discussion

The brachiopod Hox complement and the evolution of Hox genes in Spiralia

Our findings on *T. transversa* and *N. anomala* reveal an ancestral brachiopod Hox gene complement consistent with what has been hypothesized to be ancestral for Spiralia and Lophotrochozoa on the basis of degenerate PCR surveys (de Rosa, et al. 1999; Halanych and Passamaneck 2001; Balavoine, et al. 2002; Passamaneck and

Halanych 2004). This ancient complement comprises eight Hox genes – *lab*, *pb*, *Hox3*, *Dfd*, *Scr*, *Lox5*, *Lox4* and *Post2* – and has been confirmed by genomic sequencing of representative annelids and mollusks (Zhang, et al. 2012; Simakov, et al. 2013; Albertin, et al. 2015), rotifers and platyhelminthes (Flot, et al. 2013; Tsai, et al. 2013; Wasik, et al. 2015; Currie, et al. 2016) and the linguliform brachiopod *L. anatina* (Luo, et al. 2015). While *T. transversa* has retained this ancestral Hox complement, independent losses have occurred in the brachiopods *N. anomala* (*Post1*; this study) and *L. anatina* (*Lox4*) (Luo, et al. 2015) (Figure 2).

The draft genomes and available deep transcriptomes of platyhelminthes, rotifers, nemerteans, bryozoans and entoprocts did not reveal a *Lox2* ortholog (Figure 7). Similarly, genomic sequencing (Luo, et al. 2015) did not confirm the presence of a *Lox2* gene in *L. anatina* obtained by degenerate PCR (de Rosa, et al. 1999). Considering the Hox complement of chaetognaths (i.e. arrow worms) as outgroup (Matus, et al. 2007), the diversification of Hox genes in the studied spiralians indicates that the presence of a *Lox2* ortholog is a unique trait of mollusks and annelids. Altogether, the available data suggest that *Lox2* arose possibly by duplication of the ancestral *Lox4*/*Hox8*/*AbdA* gene in the lineage to Annelida + Mollusca, which is more parsimonious than considering *Lox2* ancestral to Lophotrochozoa and subsequent multiple losses of this gene in brachiopods, bryozoans, entoprocts, brachiopods, nemerteans and phoronids. Similarly, the emergence of two posterior Hox genes – *Post1* and *Post2* – in Lophotrochozoa is likely a result of a duplication event of a *Hox9* ortholog. However, more sampling of different spiralian taxa is needed to identify the exact timings of these events.

Our genomic information shows that the Hox cluster of *T. transversa* is split in three parts, with *lab* and *pb* separate from the major cluster and *Post1* also on a separate scaffold (Figure 2A). Overall, the cluster extends over 100 kb, which is significantly shorter than those of other lophotrochozoans, such as *C. teleta* (~345kb) (Fröbius, et al. 2008) and *L. gigantea* (~455 kb) (Simakov, et al. 2013). Its compact size is related to short intergenic regions and introns, comparable to the situation observed in vertebrate Hox clusters (Duboule 2007). The order and orientation of the Hox genes in *T. transversa* is preserved and more organized than in the Hox cluster reported for the brachiopod *L. anatina*, which misses *Lox4* and exhibits genomic rearrangements that placed the *Antp* gene upstream *lab* (Luo, et al. 2015). Interestingly, the Hox cluster of *L. anatina* is also split, broken into two pieces between *Lox5* and *Post2*, suggesting that the evolution of a split cluster in *T. transversa* and *N. anatina* occurred independently. Indeed, the split Hox clusters reported so far in lophotrochozoan taxa exhibit all different conformations, indicating that lineage-specific genomic events have shaped Hox gene clusters in Spiralia.

Signs of spatial, but not temporal, collinearity in T. transversa despite a split cluster

The analysis of Hox clustering in different animal species together with the temporal and spatial expression patterns of their Hox genes grounded the hypotheses that the regulatory elements required for their collinearity –mostly temporal– maintain the clustered organization of Hox genes (Duboule 1994; Ferrier and Holland 2002; Ferrier and Minguillon 2003; Patel 2004; Lemons and McGinnis 2006; Monteiro and Ferrier 2006; Duboule 2007). Although there are cases in which spatial collinearity is displayed in the absence of a cluster, as in the appendicularian chordate *O. dioica* (Seo, et al. 2004), all investigated clustered Hox genes show at least one type of

collinearity that could account for their genomic organization (Monteiro and Ferrier 2006; Duboule 2007) (Figure 7). Within Spiralia, this evolutionary scenario appears to be supported by the staggered temporal and spatial expression of the Hox genes in the split cluster of the annelid *C. teleta* (Fröbius, et al. 2008). In the other investigated spiralians, there is only either genomic information (e.g. the mollusks *L. gigantea* and *C. gigas*) or expression analysis (e.g. the mollusks *G. varia*, *Haliotis asinina*) (Hinman, et al. 2003; Samadi and Steiner 2010; Zhang, et al. 2012; Simakov, et al. 2013). Most of these gene expression studies have demonstrated coordinated spatial or temporal expression of Hox genes along the anteroposterior axis of the animal (Kulakova, et al. 2007; Fritsch, et al. 2015; Fritsch, et al. 2016) or in organ systems, such the nervous system (Hinman, et al. 2003; Samadi and Steiner 2010). However, the absence of a correlation between the expression of Hox genes and their genomic organization in these animals hampers the reconstruction of the putative mechanisms that preserve Hox clusters in Lophotrochozoa.

Our findings robustly demonstrate that split Hox cluster of *T. transversa* overall show neither spatial nor temporal collinearity (Figures 3, 4), and not even quantitative collinearity (Monteiro and Ferrier 2006), as it has been shown in mouse (Spitz, et al. 2003). These observations are also supported by the absence of a coordinated spatial and temporal expression of the Hox genes in *N. anomala* (Figure 5). In *T. transversa*, the early expression of *Hox3* breaks temporal collinearity, while it is *pb* that becomes first expressed in *N. anomala*. In both species, the gene *Lox5* is also expressed before *Scr*, as it is also the case in the annelid *N. virens* (Kulakova, et al. 2007). Ectodermal spatial collinearity is absent in the two brachiopods even when considering the future location of the larval tissues after metamorphosis (Nielsen 1991; Freeman 1993a).

The most anterior class gene *lab* is exclusively expressed in the chaetae of *T. transversa* and *N. anomala*, and thus is not affiliated with anterior neural or foregut tissues as in other lophotrochozoans, such as annelids (Fröbius, et al. 2008; Steinmetz, et al. 2011). Similarly, the most posterior Hox gene, *Post1*, is very transiently expressed in the chaetae sacs, which occupy a mid-position in the larval body. We only detected spatial collinearity in the staggered expression of the Hox genes *pb*, *Hox3* and *Dfd* along the anterior-posterior axis of the developing larval mesoderm in both *T. transversa* and *N. anomala* (Figure 6).

Altogether, the absence of a global, temporal and spatial collinearity in the brachiopod *T. transversa*, albeit the presence of a split Hox cluster, challenges the hypothesis that temporal collinearity is the underlying factor keeping Hox genes clustered (Duboule 1994; Ferrier and Minguillon 2003; Garcia-Fernàndez 2005; Monteiro and Ferrier 2006; Duboule 2007). Therefore, alternative mechanisms might need to be considered. In this regard, why do Hox clusters split in different positions between related species, as seen for instance in brachiopods (this study) and drosophilids (Negre and Ruiz 2007), but still display similar expression profiles? This might indicate that the control of expression in large split Hox clusters relies more on gene-specific short-range transcriptional control than on a global, coordinated cluster regulation, as seen in the small Hox vertebrate clusters (Spitz, et al. 2003; Duboule 2007; Acemel, et al. 2016). The conservation of Hox clusters could then be a consequence of the general conservation of syntenic relationships of a given genome. Our findings thus highlight the necessity of further detailed structure-function analyses of spiralian Hox clusters to better understand the intricate evolution of the genomic organization and regulation of Hox genes in metazoans.

Recruitment of Hox genes into morphological novelties

The bristle-like chaetae (or setae) of annelids and brachiopods, and shell valves in mollusks and brachiopods are the most prominent hard tissues found in lophotrochozoan spiralians (Brusca, et al. 2016). The ultrastructural morphology of the brachiopod and annelid chaetae is nearly identical (Lüter 2000) and with the placement of brachiopods as close relatives of annelids and mollusks (Halanych, et al. 1995), the homology of these structures appeared more likely (Lüter and Bartolomaeus 1997). In this context, the anterior hox gene *lab* is expressed in the chaetae of *Chaetopterus* sp. (Irvine and Martindale 2000) and *Post1* is expressed in the chaetae of *C. teleta*, *P. dumerilii* and *N. virens* (Kulakova, et al. 2007; Fröbius, et al. 2008). Similarly, *lab* and *Post1* are expressed in the chaetae of the brachiopods *T. transversa* and *N. anomala* (Figures 3, 5). Further evidence of a common molecular profile comes from the expression of the homeodomain gene *Aristaless-like* (*Arx*) and the zinc finger *Zic*. These genes are expressed at each chaetae sac territory in the *Platynereis* larva (Fischer 2010), in *Capitella teleta* (Layden, et al. 2010), and also in the region of the forming chaetae sac territories in *T. transversa* (Figure S3). Therefore, the expression of the Hox genes *lab* and *Post1* and the homeodomain gene *Arx* indicate that similar molecular signature underlays the development of chaetae in annelids and brachiopods. This, together with the evident morphological similarities shared by brachiopod and annelid chaetae, support considering these two structures homologous, and thus, common lophotrochozoan novelties. This would be consistent with placing the fossil *Wiwaxia*, which contains chaetae, as a stem group lophotrochozoan (Smith 2014).

The shell is a mineralized tissue present in brachiopods and mollusks. In the gastropod mollusk *G. varia*, the Hox genes *lab*, *Post1* and *Post2* are first expressed in the shell field, and later is *Dfd* (Samadi and Steiner 2009). In *H. asinina* also *lab* and *Post2* are related to shell formation (Hinman, et al. 2003). In brachiopods, *Dfd* is associated to the adult shell in *L. anatina* (Luo, et al. 2015). During embryogenesis of *T. transversa* and *N. anomala*, however, only *Scr* and *Antp* are expressed in the shell fields, but not *lab* or *Post1*, which are expressed in the chaetae sacs. The different deployment of Hox genes in the shell fields of brachiopods and mollusks might indicate that these genes do not have an ancient role in the specification of the shell-forming epithelium. However, their consistent deployment during shell development might reflect a more general, conserved role in shaping the shell fields according to their position along the anterior posterior axis.

Conclusions

In this study, we characterize the Hox gene complement of the brachiopods *T. transversa* and *N. anomala*, and demonstrate the last common ancestor to all brachiopods likely had ten Hox genes (*lab*, *pb*, *Hox3*, *dfd*, *scr*, *Lox5*, *antp*, *Lox4*, *post2*, *post1*). Noticeably, brachiopod Hox genes do not show global temporal and spatial collinearity, albeit *T. transversa* exhibits an ordered, split Hox cluster. Only the genes *pb* (in *T. transversa*), *Hox3* and *dfd* (in both brachiopods) show spatial collinearity in the 'trunk' mesoderm. In addition, the Hox genes *lab* and *post1*, as well as the homeobox *Arx*, are expressed in the developing chaetae, as also described for other annelid species (Irvine and Martindale 2001; Kulakova, et al. 2007; Fröbius, et al. 2008). These molecular similarities, together with evident morphological resemblances (Lüter 2000), support considering brachiopod and annelid chaetae

homologous structures and reinforce considering the fossil *Wiwaxia* as a stem group lophotrochozoan (Smith 2014). Altogether, our findings challenge the current scenario that temporal collinearity is the major force preserving Hox clusters (Duboule and Morata 1994; Ferrier and Minguillon 2003; Garcia-Fernàndez 2005; Monteiro and Ferrier 2006; Duboule 2007), and indicate that alternative/additional genomic mechanisms might account for the great diversity of Hox gene arrangements observed in extant animals.

Material and Methods

Animal cultures

Gravid adults of *Terebratalia transversa* (Sowerby, 1846) were collected around San Juan Island, Washington, USA and *Novocrania anomala* (Müller, 1776) around Bergen, Norway. Animal husbandry, fertilization and larval culture were conducted following previously published protocols (Reed 1987; Freeman 1993b, 2000).

Hox *cluster reconstruction in* T. transversa *and* N. anomala

Male gonads of *T. transvesa* and *N. anomala* were preserved in RNAlater (Life Technologies) for further genomic DNA (gDNA) isolation. Paired end and mate pair libraries of 2 kb and 5 kb insert sizes of *T. transversa* gDNA were sequenced using an Illumina HiSeq2000 platform. First we trimmed Illumina adapters with Cutadapt 1.4.2 (Martin 2011). Then, we assembled the paired end reads into contigs, scaffolded the assembly with the mate pair reads, and closed the gaps using Platanus 1.21 (Kajitani, et al. 2014). The genomic scaffolds of *T. transversa* including *Hox* genes are published on GenBank with the accession numbers KX372775 and KX372776. Paired end libraries of *N. anomala* gDNA were sequenced using an Illumina HiSeq2000

platform. We removed Illumina adapters as above and assembled the paired end reads with MaSuRCA 2.2.1 (Zimin, et al. 2013).

Gene isolation

Pooled samples of *T. transversa* and *N. anomala* embryos at different developmental stages (cleavage, blastula, gastrula, mid gastrula, late gastrula, early larva, and late/competent larva) were used for RNA isolation and Illumina sequencing (NCBI SRA; *T. transversa* accession SRX1307070, *N. anomala* accession SRX1343816). We trimmed adapters and low quality reads from the raw data with Trimmomatic 0.32 (Bolger, et al. 2014) and assembled the reads with Trinity 2.0.6 (Grabherr, et al. 2011). *Hox* genes were identified by BLAST searches on these transcriptomes and their respective draft genomes (see above). First-strand cDNA template (SuperScript^{TM}, Life Technologies) of mixed embryonic stages was used for gene-specific PCR. RACE cDNA of mixed embryonic stages was constructed with SMARTer RACE cDNA Amplification Kit (Clontech) and used to amplify gene ends when necessary. All fragments were cloned into the pGEM-T-Easy vector (Promega) and sequenced at the University of Bergen sequencing facility. *T. transversa* and *N. anomala Hox* gene sequences were uploaded to GenBank (accession numbers KX372756–KX372774).

Orthology analyses

Hox gene sequences of a representative selection of bilaterian lineages (Supplementary Table S1) were aligned with MAFFT v.7 (Katoh and Standley 2013). The multiple sequence alignment, which is available upon request, was trimmed to include the 60 amino acids of the homeodomain. ProtTest v.3 (Darriba, et al. 2011)

was used to determine the best fitting evolutionary model (LG+G+I). Orthology analyses were conducted with RAxML v.8.2.6 (Stamatakis 2014) using the autoMRE option. The resulting trees were edited with FigTree and Illustrator CS6 (Adobe).

Gene expression analyses

T. transversa and *N. anomala* embryos at different embryonic and larval stages were fixed in 4% paraformaldehyde in sea water for 1 h at room temperature. All larval stages were relaxed in 7.4% magnesium chloride for 10 min before fixation. Fixed samples were washed several times in phosphate buffer saline (PBS) with 0.1% tween-20 before dehydration through a graded methanol series and storage in 100% methanol at -20 ºC. Single colorimetric whole mount *in situ* hybridization were carried out following an established protocol (detailed protocol available in Protocol Exchange: doi:10.1038/nprot.2008.201) (Hejnol and Martindale 2008; Santagata, et al. 2012). Double fluorescent *in situ* hybridizations were conducted as described elsewhere (Grande, et al. 2014). Representative stained specimens were imaged with bright field Nomarski optics using an Axiocam HRc connected to an Axioscope Ax10 (Zeiss). Fluorescently labeled embryos were mounted in Murray's clearing reagent (benzyl alcohol: benzyl benzoate, 1:2) and imaged under a SP5 confocal laser-scanning microscope (Leica). Images and confocal z-stacks were processed with Fiji and Photoshop CS6 (Adobe) and figure panels assembled with Illustrator CS6 (Adobe). Contrast and brightness were always adjusted to the whole image, and not to parts of it.

***Quantitative* Hox *gene expression in* T. transversa**

Thousands of synchronous *T. transversa* embryos collected at 14 specific stages

(oocytes, 8h mid blastula, 19h late blastula, 24h moving late blastula, 26h early gastrula, 37h asymmetric gastrula, 51h bilateral gastrula, 59h bilobed, 68h trilobed, 82h early larva (first chaetae visible), 98h late larva (long chaetae, eye spots), 131h competent larva, 1d juvenile, 2d juvenile) were pooled together and preserved in RNAlater (Life Technologies). Total RNA was isolated with Trizol Reagent (Life Technologies). For quantitative real time PCR, total RNA was DNAse treated and preserved at -80 ºC. Gene specific primers bordering an intron splice-site and defining an amplicon of 80-150 bp sizes were designed for each gene (Supplementary Table S2). Expression levels of two technical replicates performed in two biological replicates were calculated based on absolute quantification units. For comparative stage-specific transcriptomic analyses, total RNA was used for constructing Illumina single end libraries and sequenced in four lanes of a HiSeq 2000 platform. Samples were randomized between the lanes. To estimate the abundance of transcripts per stage, we mapped the single end reads to the transcriptome of *T. transversa* with Bowtie, calculated expression levels with RSEM, and generated a matrix with TMM normalization across samples by running Trinity's utility scripts. Expression levels obtained after quantitative real-time PCR and comparative stage-specific transcriptomics were plotted with R.

Acknowledgements

We thank the crew of the "Centennial" boat and office stuff at Friday Harbor Laboratories (USA) and the crew of the "Hans Brattström" and "Aurelia" boats at the Espeland Marine Station (Norway) for their invaluable help during animal collections. We also thank Daniel Thiel and Anlaug Boddington for their help with animal collections and spawnings, Daniel Chourrout for his valuable comments on early

versions of this manuscript, and Kevin Kocot for the access to entoproct transcriptomes. The trip to Friday Harbor Laboratories was funded by a Meltzer Fond grant. The research conducted in this study was funded by the Sars Centre core budget.

Author contributions

A.H. designed the study. A.H., S.M.S. and J.M.M.D. conducted the gene isolation and *in situ* hybridization studies. J.M.M.D. performed the gene orthology analyses. A.H., J.M.M.D., Y.P. and B.V. collected the stage-specific samples of *T. transversa* embryos. A.B. and J.M.M.D. isolated the genomic DNA of *T. transversa* and *N. anomala*. J.M.M.D. and B.V. did the draft genome assemblies and S.M.S. analyzed the Hox genomic organization. J.M.M.D. performed the stage-specific RNA isolations; A.B. did the quantitative real time PCR experiments, and B.V. conducted the analysis of the stage-specific transcriptomes. A.H. and J.M.M.D. wrote the manuscript. All authors discussed the data and edited the text.

References

Aboobaker A, Blaxter M. 2003a. Hox gene evolution in nematodes: novelty conserved. Curr Opin Genet Dev 13:593-598.

Aboobaker AA, Blaxter ML. 2003b. Hox Gene Loss during Dynamic Evolution of the Nematode Cluster. Curr Biol 13:37-40.

Acemel RD, Tena JJ, Irastorza-Azcarate I, Marletaz F, Gómez-Marín C, de la Calle-Mustienes E, Bertrand S, Diaz SG, Aldea D, Aury JM, et al. 2016. A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. Nat Genet 48:336-341.

Akam M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. Cell 57:347-349.

Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524:220-224.

Aronowicz J, Lowe CJ. 2006. Hox gene expression in the hemichordate *Saccoglossus kowalevskii* and the evolution of deuterostome nervous systems. Integ Comp Biol 46:890-901.

Balavoine G, de Rosa R, Adoutte A. 2002. Hox clusters and bilaterian phylogeny. Mol Phylogenet Evol 24:366-373.

Barucca M, Canapa A, Biscotti MA. 2016. An Overview of Hox Genes in Lophotrochozoa: Evolution and Functionality. J Dev Biol 4:12.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Brusca RC, Moore W, Shuster SM. 2016. Invertebrates. Sunderland, MA: Sinauer Associates, Inc.

Carrasco AE, McGinnis W, Gehring WJ, De Robertis EM. 1984. Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. Cell 37:409-414.

Chauvet S, Merabet S, Bilder D, Scott MP, Pradel J, Graba Y. 2000. Distinct hox protein sequences determine specificity in different tissues. Proc Natl Acad Sci U S A 97:4064-4069.

Costa M, Weir M, Coulson A, Sulston J, Kenyon C. 1988. Posterior pattern formation in *C. elegans* involves position-specific expression of a gene containing a homeobox. Cell 55:747-756.

Currie KW, Brown DD, Zhu S, Xu C, Voisin V, Bader GD, Pearson BJ. 2016. HOX gene complement and expression in the planarian Schmidtea mediterranea. Evodevo 7:7.

Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1164-1165.

de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB, Balavoine G. 1999. Hox genes in brachiopods and priapulids and protostome evolution. Nature 399:772-776.

Dollé P, Izpisúa-Belmonte JC, Falkenstein H, Renucci A, Duboule D. 1989. Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation. Nature 342:767-772.

Duboule D. 2007. The rise and fall of Hox gene clusters. Development 134:2549-2560.

Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. Dev Suppl:135-142.

Duboule D, Morata G. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. Trends Genet 10:358-364.

Dunn CW, Giribet G, Edgecombe GD, Hejnol A. 2014. Animal phylogeny and its evolutionary implications. Ann Rev Ecol Evol Syst 45:371-395.

Ferrier DEK, Holland PWH. 2002. *Ciona intestinalis* ParaHox genes: evolution of Hox/ParaHox cluster integrity, developmental mode, and temporal colinearity. Mol Phylogenet Evol 24:412-417.

Ferrier DEK, Minguillon C. 2003. Evolution of the Hox/ParaHox gene clusters. Int J Dev Biol 47:605-611.

Fischer A. 2010. Mesoderm formation and muscle development of *Platynereis dumerilii* (Nereididae, Annelida). [Dissertation]. [Berlin]: Freie Universität Berlin.

Flot JF, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EG, Hejnol A, Henrissat B, Koszul R, Aury JM, et al. 2013. Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. Nature 500:453-457.

Freeman G. 1993a. Metamorphosis in the Brachiopod *Terebratalia*: Evidence for a Role of Calcium Channel Function and the Dissociation of Shell Formation from Settlement. Biol Bull 184:15-24.

Freeman G. 1993b. Regional specification during embryogenesis in the articulate brachiopod *Terebratalia*. Dev Biol 160:196-213.

Freeman G. 2000. Regional Specification during Embryogenesis in the Craniiform Brachiopod *Crania anomala*. Dev Biol 227:219-238.

Fritsch M, Wollesen T, de Oliveira AL, Wanninger A. 2015. Unexpected co-linearity of Hox gene expression in an aculiferan mollusk. BMC Evol Biol 15:151.

Fritsch M, Wollesen T, Wanninger A. 2016. Hox and ParaHox gene expression in early body plan patterning of polyplacophoran mollusks. J Exp Zool B Mol Dev Evol 326:89-104.

Fröbius AC, Matus DQ, Seaver EC. 2008. Genomic organization and expression demonstrate spatial and temporal Hox gene colinearity in the lophotrochozoan *Capitella* sp. I. PLoS One 3:e4004.

Garcia-Fernàndez J. 2005. The genesis and evolution of homeobox gene clusters. Nature Reviews Genetics 6:881-892.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29:644-652.

Grande C, Martín-Durán JM, Kenny NJ, Truchado-García M, Hejnol A. 2014. Evolution, divergence and loss of the Nodal signalling pathway: new data and a synthesis across the Bilateria. Int J Dev Biol 58:521-532.

Halanych KM, Bacheller JD, Aguinaldo AM, Liva SM, Hillis DM, Lake JA. 1995. Evidence from 18S ribosomal DNA that the lophophorates are protostome animals. Science 267:1641-1643.

Halanych KM, Passamaneck Y. 2001. A Brief Review of Metazoan Phylogeny and Future Prospects in Hox-Research. Amer. Zool. 41:629-639.

Hejnol A. 2010. A Twist in Time-The Evolution of Spiral Cleavage in the Light of Animal Phylogeny. Integrative and Comparative Biology 50:695-706.

Hejnol A, Martindale MQ. 2008. Acoel development indicates the independent evolution of the bilaterian mouth and anus. Nature 456:382-386.

Hejnol A, Martindale MQ. 2009. Coordinated spatial and temporal expression of Hox genes during embryogenesis in the acoel Convolutriloba longifissura. BMC Biol 7:65.

Hiebert LS, Maslakova SA. 2015a. Expression of *Hox*, *Cdx*, and *Six3/6* genes in the hoplonemertean *Pantinonemertes californiensis* offers insight into the evolution of maximally indirect development in the phylum Nemertea. Evodevo 6:26.

Hiebert LS, Maslakova SA. 2015b. *Hox* genes pattern the anterior-posterior axis of the juvenile but not the larva in a maximally indirect developing invertebrate, *Micrura alaskensis* (Nemertea). BMC Biol 13:23.

Hinman VF, O'Brien EK, Richards GS, Degnan BM. 2003. Expression of anterior Hox genes during larval development of the gastropod *Haliotis asinina*. Evol Dev 5:508-521.

Irvine SQ, Martindale MQ. 2001. Comparative analysis of Hox gene expression in the polychaete Chaetopterus: Implications for the evolution of body plan regionalization. American Zoologist 41:640-651.

Irvine SQ, Martindale MQ. 2000. Expression Patterns of Anterior Hox Genes in the Polychaete Chaetopterus: Correlation with Morphological Boundaries. Dev Biol 217:333-351.

Izpisúa-Belmonte J, Falkenstein H, Dollé P, Renucci A, Duboule D. 1991. Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. The EMBO journal 10:2279.

Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, et al. 2014. Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Research 24:1384-1395.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780.

Kourakis MJ, Martindale MQ. 2001. Hox gene duplication and deployment in the annelid leech Helobdella. Evol Dev 3:145-153.

Kourakis MJ, Master VA, Lokhorst DK, Nardelli-Haefliger D, Wedeen CJ, Martindale MQ, Shankland M. 1997a. Conserved anterior boundaries of Hox gene expression in the central nervous system of the leech *Helobdella*. Dev Biol 190:284-300.

Kourakis MJ, Master VA, Lokhorst DK, Nardelli-Haefliger D, Wedeen CJ, Martindale MQ, Shankland M. 1997b. Evolutionary conservation of Hox gene expression in the CNS of the leech *Helobdella*. Dev Biol 186:B98-B98.

Kulakova M, Bakalenko N, Novikova E, Cook CE, Eliseeva E, Steinmetz PRH, Kostyuchenko RP, Dondua A, Arendt D, Akam M, et al. 2007. Hox gene expression in larval development of the polychaetes *Nereis virens* and *Platynereis dumerilii* (Annelida, Lophotrochozoa). Dev Genes Evol 217:39-54.

Laumer CE, Bekkouche N, Kerbl A, Goetz F, Neves RC, Sørensen MV, Kristensen RM, Hejnol A, Dunn CW, Giribet G, et al. 2015. Spiralian Phylogeny Informs the Evolution of Microscopic Lineages. Curr Biol 25:2000-2006.

Layden MJ, Meyer NP, Pang K, Seaver EC, Martindale MQ. 2010. Expression and phylogenetic analysis of the *zic* gene family in the evolution and development of metazoans. Evodevo 1:12.

Lee PN, Callaerts P, de Couet HG, Martindale MQ. 2003. Cephalopod Hox genes and the origin of morphological novelties. Nature 424:1061-1065.

Lemons D, McGinnis W. 2006. Genomic Evolution of Hox Gene Clusters. Science 313:1918-1922.

Lewis EB. 1978. A gene complex controlling segmentation in Drosophila. Nature 276:565-570.

Lowe CJ, Wray GA. 1997. Radical alterations in the roles of homeobox genes during echinoderm evolution. Nature 389:718-721.

Luo YJ, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, Fujie M, Yamasaki SI, Endo K, Satoh N. 2015. The *Lingula* genome provides insights into brachiopod evolution and the origin of phosphate biomineralization. Nat Commun 6:8301.

Lüter C. 2000. Ultrastructure of Larval and Adult Setae of Brachiopoda. Zool Anz 239:75-90.

Lüter C, Bartolomaeus T. 1997. The phylogenetic position of Brachiopoda - a comparison of morphological and molecular data. Zool Scripta 26:245-253.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17.

Matus DQ, Halanych KM, Martindale MQ. 2007. The Hox gene complement of a pelagic chaetognath, Flaccisagitta enflata. Integrative and Comparative Biology 47:854-864.

McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ. 1984. A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. Cell 37:403-408.

McGinnis W, Hart CP, Gehring WJ, Ruddle FH. 1984. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. Cell 38:675-680.

McGinnis W, Krumlauf R. 1992. Homeobox genes and axial patterning. Cell 68:283-302.

McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ. 1984. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. Nature 308:428-433.

Monteiro AS, Ferrier DEK. 2006. Hox genes are not always Colinear. Int J Biol Sci 2:95-103.

Negre B, Ruiz A. 2007. HOM-C evolution in Drosophila: is there a need for Hox gene clustering? Trends Genet 23:55-59.

Nielsen C. 1991. The development of the brachiopod *Crania (Neocrania) anomala* (O. F. Müller) and its phylogenetic significance. Acta Zoologica 72:7-28.

Passamaneck YJ, Halanych KM. 2004. Evidence from Hox genes that bryozoans are lophotrochozoans. Evol Dev 6:275-281.

Patel NH. 2004. Evolutionary biology: time, space and genomes. Nature 431:28-29.

Pearson JC, Lemons D, McGinnis W. 2005. Modulating Hox gene functions during animal body patterning. Nature Reviews Genetics 6:893-904.

Reed C. 1987. Phylum Brachiopoda. In: Strathmann MF, edi- tor. Reproduction and Development of the Marine Inverte- brates of the Northern Pacific Coast. Seattle. University of Washington Press.:pp 486–493.

Rudwick MJS. 1970. Living and fossil brachiopods: Hutchinson.

Samadi L, Steiner G. 2010. Expression of Hox genes during the larval development of the snail, *Gibbula varia* (L.)-further evidence of non-colinearity in molluscs. Dev Genes Evol 220:161-172.

Samadi L, Steiner G. 2009. Involvement of Hox genes in shell morphogenesis in the encapsulated development of a top shell gastropod (*Gibbula varia* L.). Dev Genes Evol 219:523-530.

Santagata S, Resh C, Hejnol A, Martindale MQ, Passamaneck YJ. 2012. Development of the larval anterior neurogenic domains of *Terebratalia transversa* (Brachiopoda) provides insights into the diversification of larval apical organs and the spiralian nervous system. Evodevo 3.

Seo HC, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, Flaat M, Weissenbach J, Lehrach H, Wincker P, et al. 2004. Hox cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica. Nature 431:67-71.

Serano JM, Martin A, Liubicich DM, Jarvis E, Bruce HS, La K, Browne WE, Grimwood J, Patel NH. 2016. Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. Dev Biol 409:297-309.

Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, et al. 2013. Insights into bilaterian evolution from three spiralian genomes. Nature 493:526-531.

Smith FW, Boothby TC, Giovannini I, Rebecchi L, Jockusch EL, Goldstein B. 2016. The Compact Body Plan of Tardigrades Evolved by the Loss of a Large Body Region. Curr Biol 26:224-229.

Smith MR. 2014. Ontogeny, morphology and taxonomy of the soft-bodied cambrian 'mollusc' *Wiwaxia*. Palaeontology 57:215-229.

Spitz F, Gonzalez F, Duboule D. 2003. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell 113:405-417.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.

Steinmetz PRH, Kostyuchenko RP, Fischer A, Arendt D. 2011. The segmental pattern of otx, gbx, and Hox genes in the annelid *Platynereis dumerilii*. Evol Dev 13:72-79.

Struck TH, Wey-Fabrizius AR, Golombek A, Hering L, Weigert A, Bleidorn C, Klebow S, Iakovenko N, Hausdorf B, Petersen M, et al. 2014. Platyzoan paraphyly based on phylogenomic data supports a noncoelomate ancestry of spiralia. Mol Biol Evol 31:1833-1849.

Tsai IJ, Zarowiecki M, Holroyd N, Garciarrubio A, Sanchez-Flores A, Brooks KL, Tracey A, Bobes RJ, Fragoso G, Sciutto E, et al. 2013. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496:57-63.

Wada H, Garcia-Fernandez J, Holland PW. 1999. Colinear and segmental expression of amphioxus Hox genes. Dev Biol 213:131-141.

Wasik BR, Rose DJ, Moczek AP. 2010. Beetle horns are regulated by the Hox gene, Sex combs reduced, in a species- and sex-specific manner. Evol Dev 12:353-362.

Wasik K, Gurtowski J, Zhou X, Ramos OM, Delas MJ, Battistoni G, El Demerdash O, Falciatori I, Vizoso DB, Smith AD, et al. 2015. Genome and transcriptome of the regeneration-competent flatworm, *Macrostomum lignano*. Proc Natl Acad Sci U S A 112:12462-12467.

Woltering JM, Duboule D. 2015. Tetrapod axial evolution and developmental constraints; Empirical underpinning by a mouse model. Mech Dev 138 Pt 2:64-72.

Zakany J, Duboule D. 2007. The role of Hox genes during vertebrate limb development. Curr Opin Genet Dev 17:359-366.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Zimin AV, Marcais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. Bioinformatics 29:2669-2677.

Zwarycz AS, Nossa CW, Putnam NH, Ryan JF. 2016. Timing and Scope of Genomic Expansion within Annelida: Evidence from Homeoboxes in the Genome of the Earthworm *Eisenia fetida*. Genome Biol Evol 8:271-281.

Figures

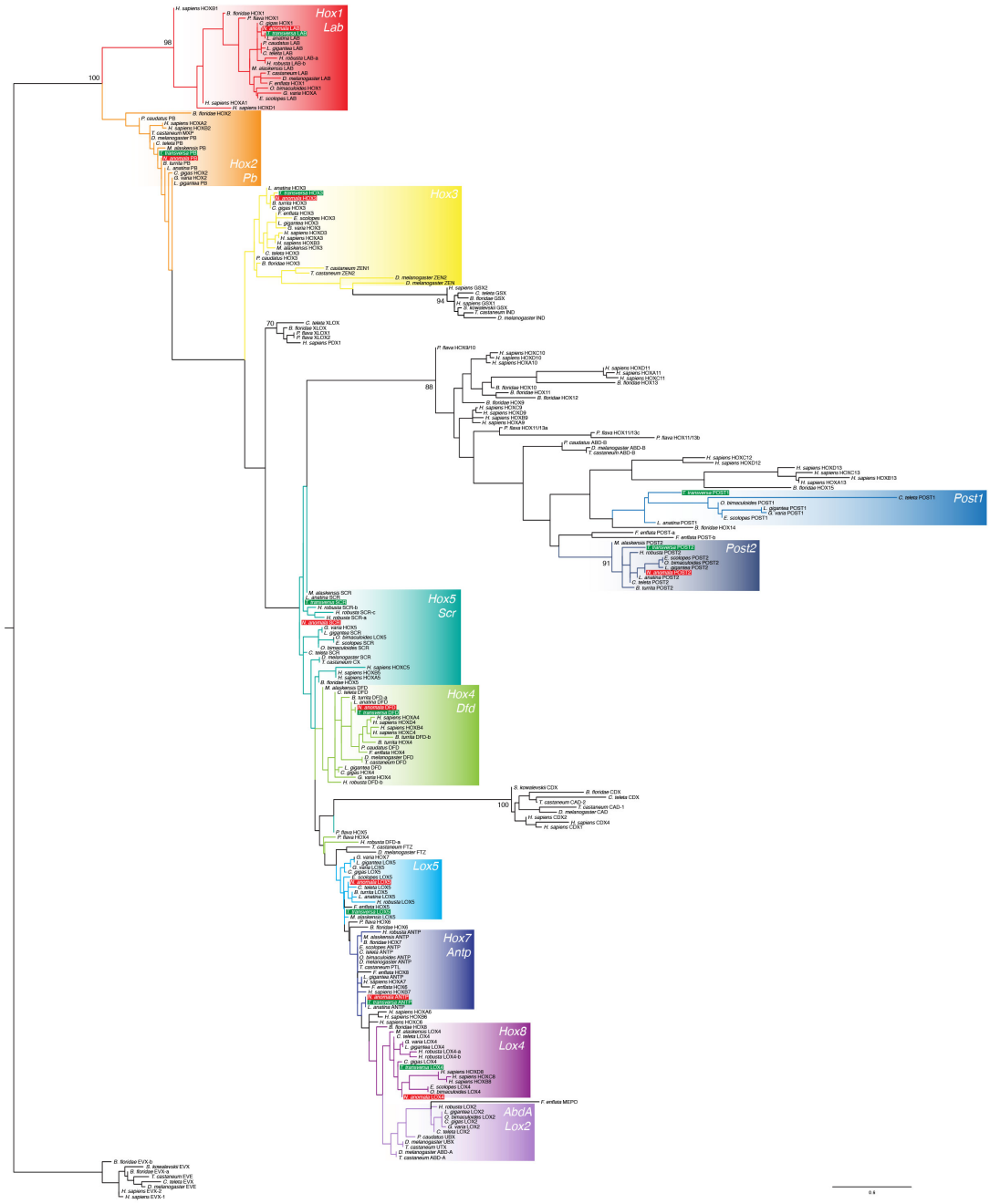

Figure 1. Orthology analysis of *T. transversa* and *N. anomala* Hox genes.

Maximum likelihood phylogenetic analysis of bilaterian Hox and ParaHox genes, using as outgroup the *even-skipped* (EVX) subfamily. Colored boxes indicate Hox ortholog groups present in spiralian representatives. *T. transversa* sequences are

highlighted by green boxes and *N. anomala* sequences by red boxes. Only high bootstrap values are shown.

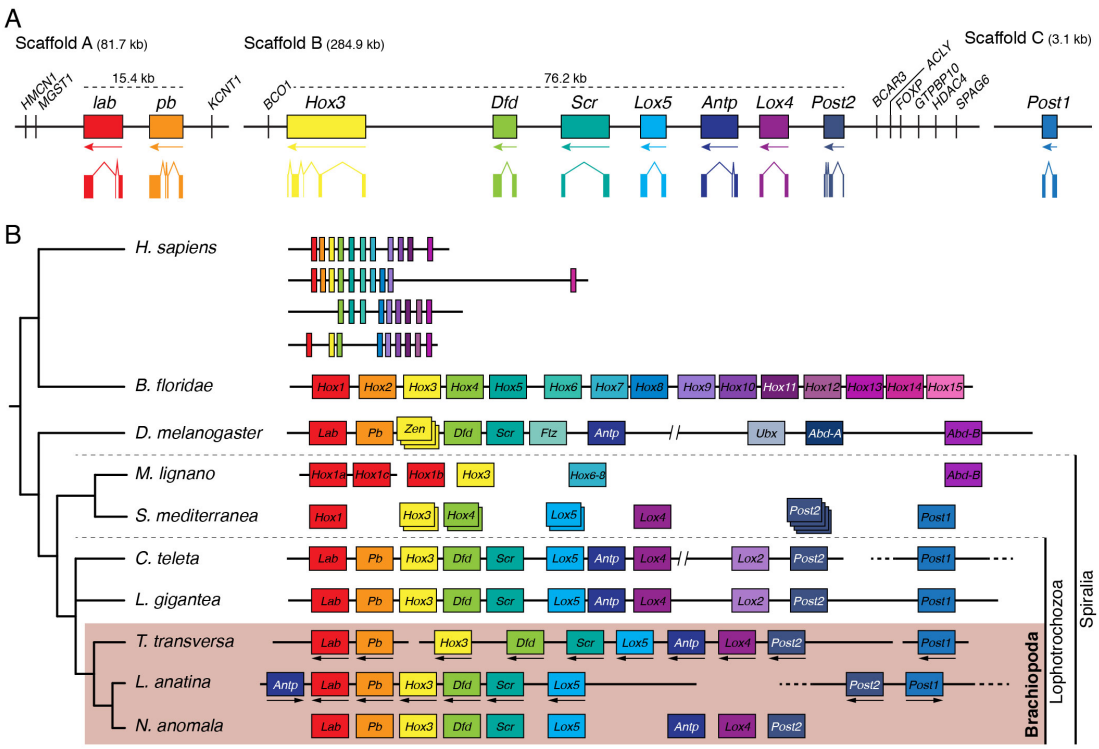

Figure 2. Genomic organization of Hox genes in *T. transversa*. (**A**) The ten Hox genes of *T. transversa* are ordered along three genomic scaffolds and are flanked by external genes (vertical lines; gene orthology is based on best blast hit). Thus, *T. transversa* has a split Hox cluster composed of three sub-clusters. No predicted ORFs were identified between the Hox genes in scaffold A and B. A colored box represents each Hox gene, and below each box there is the direction of transcription and the exon-intron composition. The genomic regions containing Hox genes are represented in scale. (**B**) The genomic organization of brachiopod Hox genes in a phylogenetic context (adapted from (Albertin, et al. 2015; Luo, et al. 2015)). The genomic order of Hox genes in *T. transversa* is similar to that observed in other spiralians (e.g. *Capitella teleta* and *Lottia gigantea*), which suggests that the translocation of the Hox gene *Antp* to the most upstream region of the Hox cluster in the brachiopod *Lingula anatina* is a lineage-specific feature (in *T. transversa* and *L. anatina* the arrows below the genes show the direction of transcription). The low contiguity of the draft genome

assembly of *N. anomala* hampered recovering genomic linkages between the identified Hox genes. Each ortholog group is represented by a particular color.

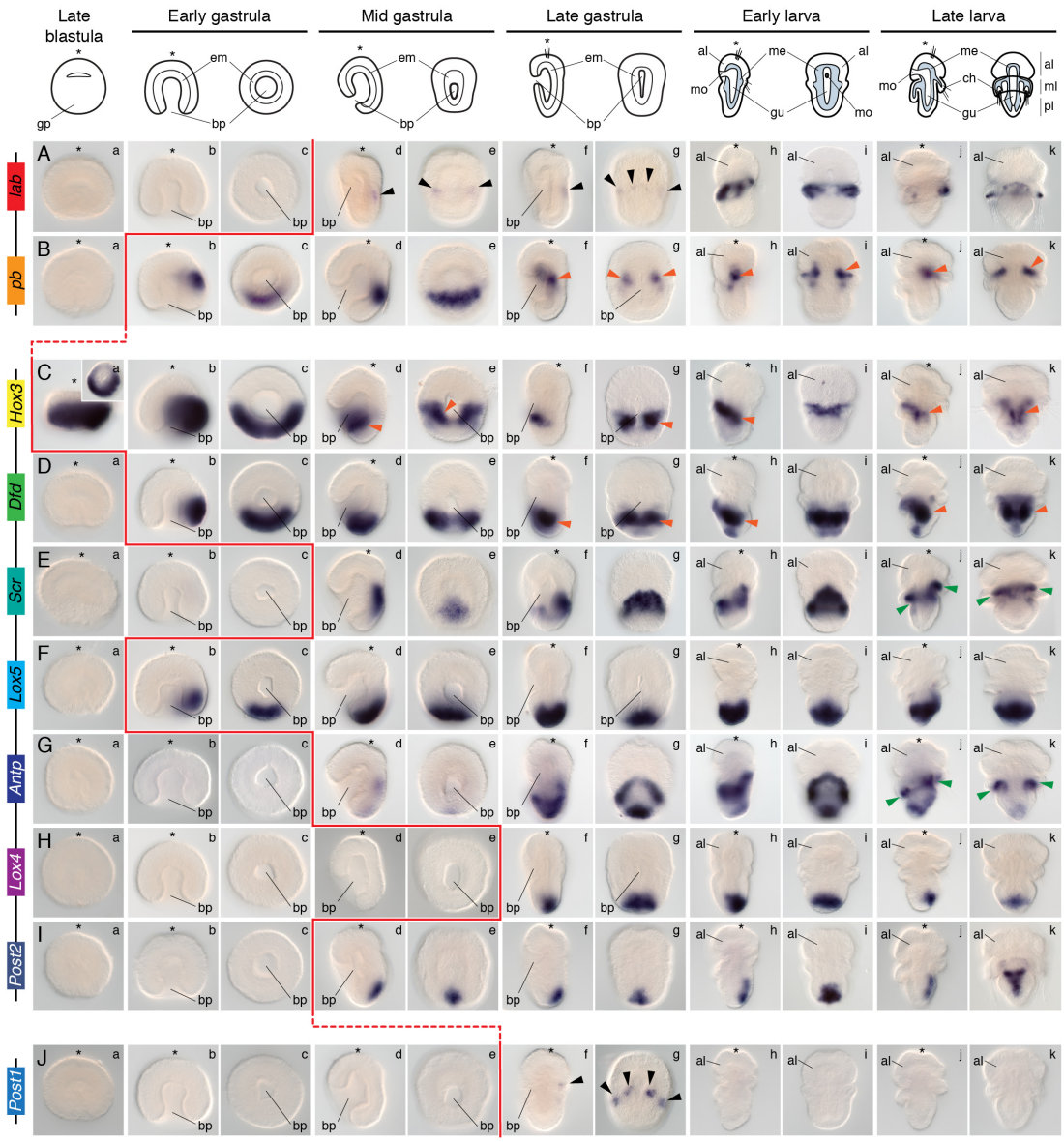

Figure 3. Expression of Hox genes in *T. transversa*. (**A–J**) Whole mount *in situ* hybridization of each Hox gene during embryonic and larval stages in *T. transversa*. The Hox genes *lab* and *post1* are expressed during chaetae formation. The genes *pb*, *Hox3* and *dfd* are collinearly expressed along the mantle and pedicle mesoderm. The Hox genes *scr* and *antp* are expressed in the periostracum, the shell-forming epithelium. *Lox5*, *Lox4* and *post2* are expressed in the posterior ectoderm of the pedicle lobe. See main text for a detailed description of each expression pattern. Black arrowheads indicate expression in the chaetae sacs. Orange arrowheads highlight mesodermal expression. Green arrowheads indicate expression in the periostracum.

The genomic organization of the Hox genes is shown on the left. On top, schematic representations of each analyzed developmental stage on its respective perspective. In these schemes, the blue area represents the mesoderm. Drawings are not to scale. The red line indicates the onset of expression of each Hox gene based on *in situ* hybridization data. The blastula stage is a lateral view (inset is a vegetal view). The other stages are in a lateral view (left column) and dorsoventral view (right column). The asterisk demarcates the animal/anterior pole. al, apical lobe; bp, blastopore; ch, chaetae; em, endomesoderm; gp, gastral plate; gu, gut; me, mesoderm; ml, mantle lobe; mo, mouth; pl, pedicle lobe.

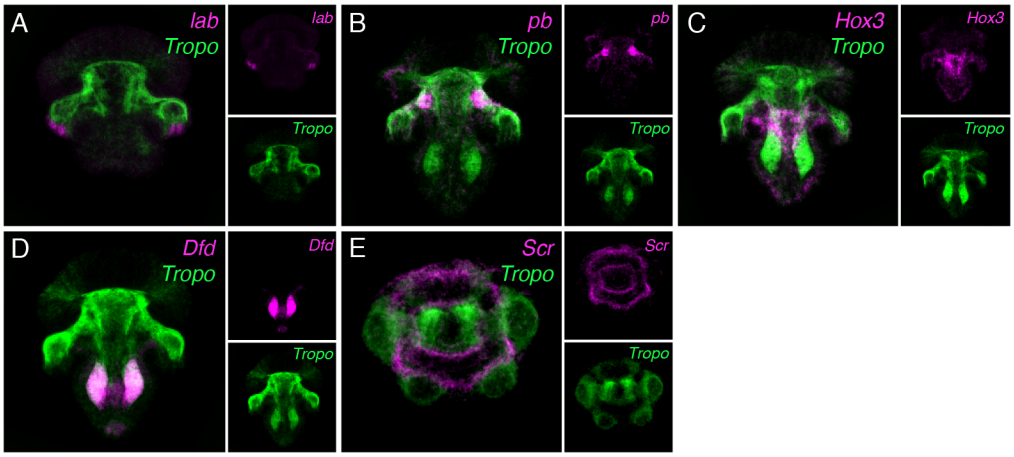

Figure 4. Hox expression in mesoderm and periostracum of *T. transversa*. (**A**–**E**) Double fluorescent *in situ* hybridization of *lab*, *pb*, *Hox3*, *dfd* and *scr* with tropomyosin (Tropo, in green) in late larval stages of *T. transversa*. (**A**) The gene *lab* is expressed in relation to the chaetae sacs, but does not overlap with the tropomyosin-expressing mesoderm. (**B**–**D**) The Hox genes *pb*, *Hox3* and *Dfd* show spatial collinearity along the mantle and pedicle mesoderm. (**E**) The gene *scr* is expressed in the periostracum, which is the epithelium that forms the shell.

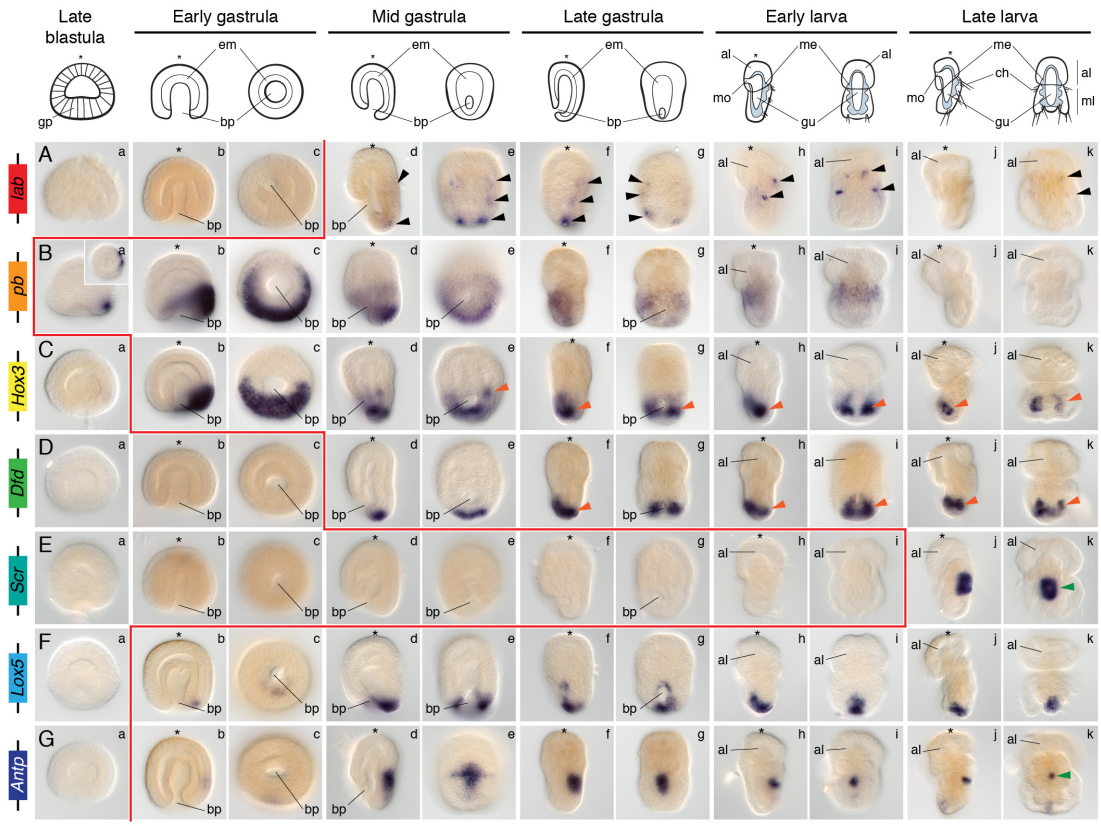

Figure 5. Expression of Hox genes in *N. anomala*. (**A–G**) Whole mount *in situ* hybridization of the Hox genes during embryonic and larval stages in *N. anomala*. The gene *lab* is expressed in the chaetae. The Hox genes *Hox3* and *dfd* are collinearly expressed in the mantle mesoderm. The genes *scr* and *antp* are expressed in the prospective shell-forming epithelium. The genes *pb* and *Lox5* are detected in the ectoderm of the mantle lobe. The genes *Lox4* and *post2* were not detected in transcriptomes and cDNA during embryonic stages. See main text for a detailed description of each expression pattern. Black arrowheads indicate expression in the chaetae sacs. Orange arrowheads highlight mesodermal expression. Green arrowheads indicate expression in the periostracum. On top, schematic representations of each analyzed developmental stage on its respective perspective. In these schemes, the blue area represents the mesoderm. Drawings are not to scale. The red line indicates the onset of expression of each Hox gene based on *in situ* hybridization data. The blastula

stage is a lateral view (inset is a vegetal view). The other stages are in a lateral view (left column) and dorsoventral view (right column). The asterisk demarcates the animal/anterior pole. al, apical lobe; bp, blastopore; ch, chaetae; em, endomesoderm; gp, gastral plate; gu, gut; me, mesoderm; ml, mantle lobe; mo, mouth.

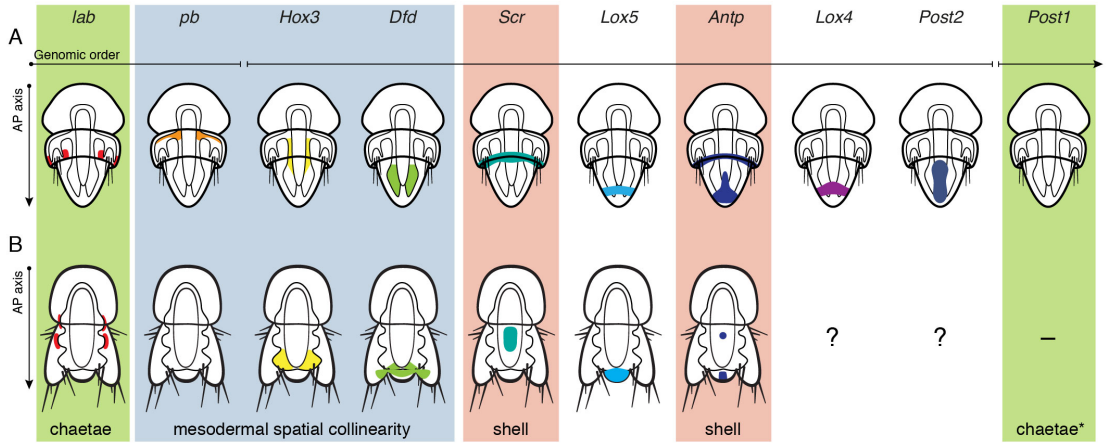

Figure 6. Summary of Hox gene expression in *T. transversa* and *N. anomala*. (**A**, **B**) Schematic drawings of late larvae of *T. transversa* and *N. anomala* depicting the expression of each Hox gene. The Hox genes *pb* (not in *N. anomala*), *Hox3* and *dfd* show staggered expression, at least in one of their domains, associated with the mesoderm (light blue box). In both brachiopods, the genes *scr* and *antp* are expressed in the periostracum, or the shell-forming epithelium (red boxes) and lab and post1 are associated to the developing chaetae (green boxes; asterisk in *post1*: *post1* is expressed in the chaetae only during late embryonic stages, not in the mature larva, and only in *T. transversa*). The expression of *Lox4* and *post2* in *N. anomala* could not be determined in this study. The gene *post1* is missing in *N. anomala*. Drawings are not to scale.

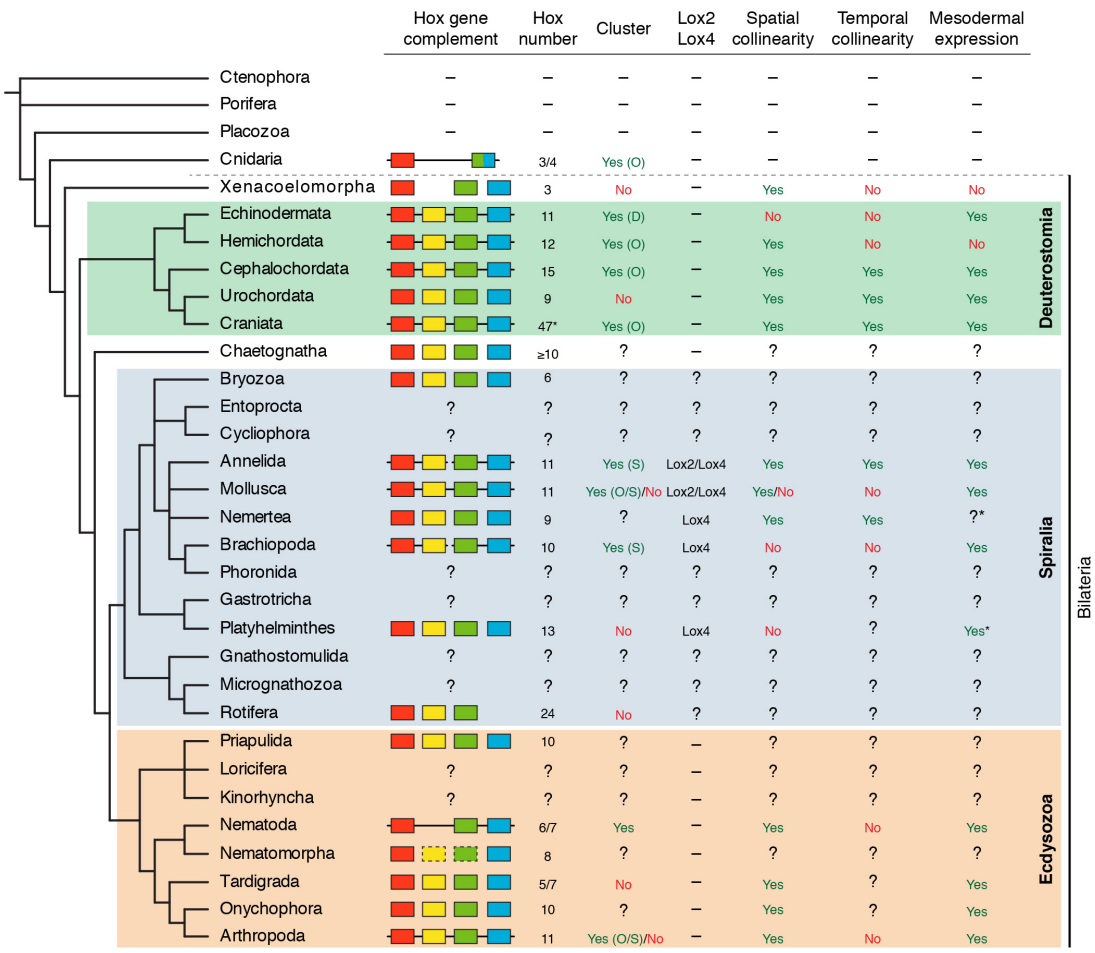

Figure 7. Evolution of Hox organization and expression across Metazoa. Table depicting the features of the Hox gene complement of each animal lineage in a phylogenetic framework. The Hox complement is summarized by the presence of at least one representative of the anterior, Hox3, central and posterior ortholog groups. The Hox number indicates the possible ancestral number, but can vary between species (in Craniata, the number corresponds to the human Hox complement, which consists of four clusters; asterisk). The cluster organization can be of three types: organized (O), disorganized (D), split (S). When there are species with an atomized cluster we write that the cluster is absent (No presence). Question marks indicate unknown data and dashes indicate absences. See main text for references.

Supplementary Material

Brachiopods possess a split Hox cluster with signs of spatial, but not temporal collinearity

Sabrina M. Schiemann, José M. Martín-Durán, Aina Børve, Bruno C. Vellutini, Yale J. Passamaneck, Andreas Hejnol

Index

- Figures S1–S3
- Table S1 and S2

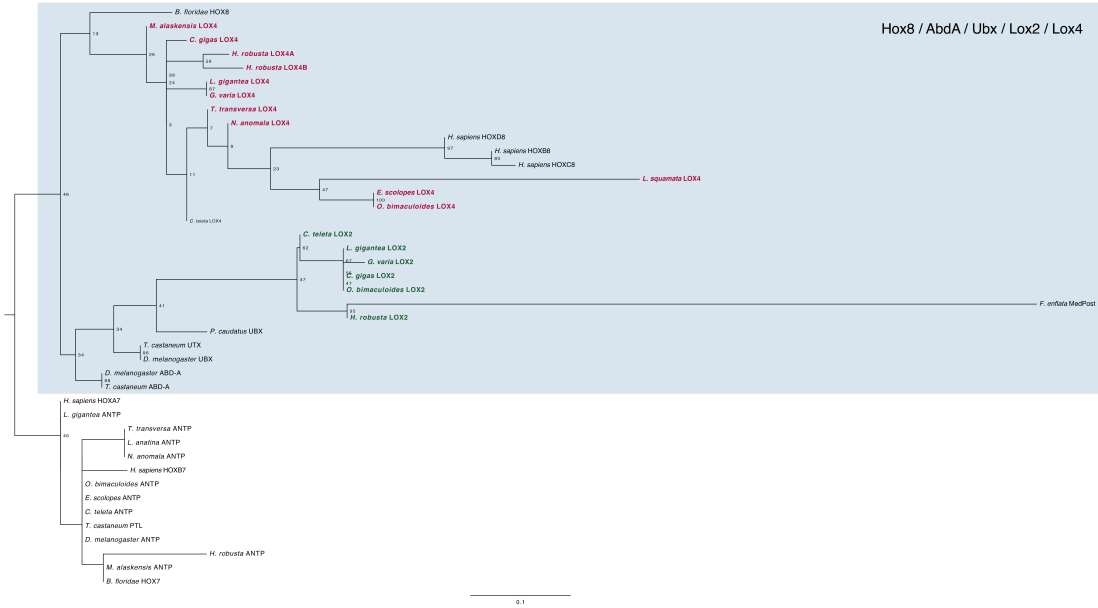

Figure S1. Phylogenetic relationships of Lox2 and Lox4 Hox genes. Maximum likelihood orthology analysis of Lox2/Lox4 genes using as outgroup the Hox7/Antp class and the LG+G+I model of protein evolution. The group Hox8/AbdA/Utx/Lox2/Lox4 includes deuterostomian (Hox8), ecdysozoan (AbdA, Utx) and spiralian representatives (Lox2/Lox4; in bold green and red, respectively). The affiliation of Lox2 sequences to the ecdysozoan AbdA and Utx and of Lox4 to deuterostomian Hox8 sequences suggests that the Lox2/Lox4 duplication might be ancestral to Spiralia, and that Lox2 got repeatedly lost in multiple spiralian and lophotrochozoan lineages. Alternatively, there was a single Lox2/4 gene in the last common spiralian ancestor and the duplication into Lox2 and Lox4 occurred at the last common ancestor to Annelida and Mollusca, although this paralogous relationship cannot be completely resolved due to the low phylogenetic signal of the homeobox domain. The Lox4 sequence from the gastrotrich *Lepidodermella squamata* is included in Table S1.

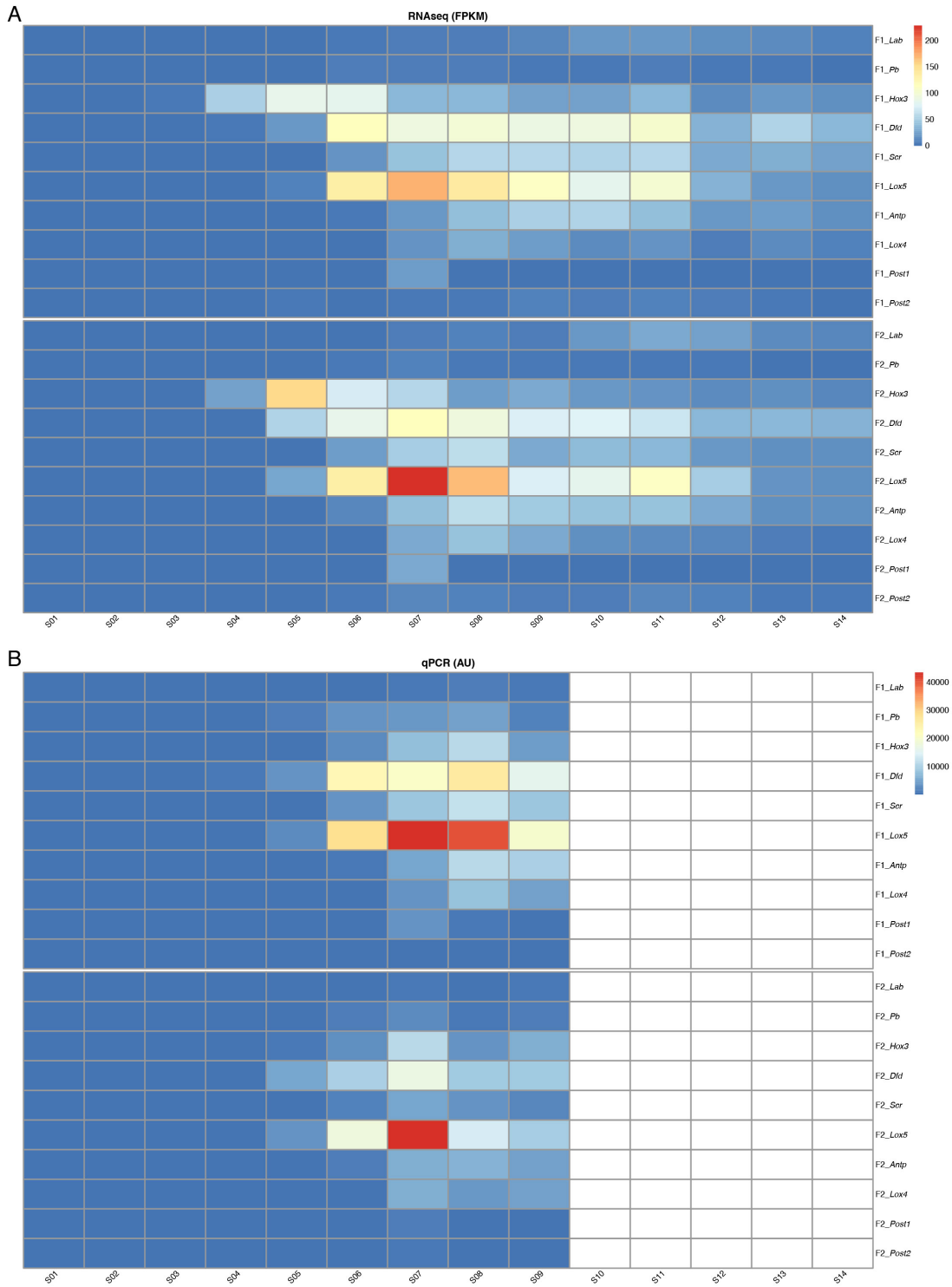

Figure S2. Quantitative expression of Hox genes in *T. transversa* developmental stages. (**A**) RNAseq expression levels calculated by fragments per kilobase of exon per million reads mapped (FPKM). As observed by whole-mount *in situ* hybridization, *Hox3* is the first gene up-regulated in the two biological replicates

(female 1, F1; female 2, F2). (**B**) quantitative real-time PCR (qPCR) expression levels based on absolute quantification units (AU). PCR was not performed for stages 10–14 (white cells). qPCR confirms the absence of temporal collinearity, although we do not detect higher levels of *Hox3* at late blastula (S04), as observed by RNAseq and *in situ* hybridization. Stages: S01, oocytes; S02, 8h mid blastula; S03, 19h late blastula; S04, 24h moving late blastula; S05, 26h early gastrula; S06, 37h mid gastrula; S07, 51h late gastrula; S08, 59h bilobed late gastrula; S09, 68h trilobed late gastrula; S10, 82h early larva; S11, 98h late larva; S12, competent larva; S13, 1 day juvenile; S14, 2 days juvenile.

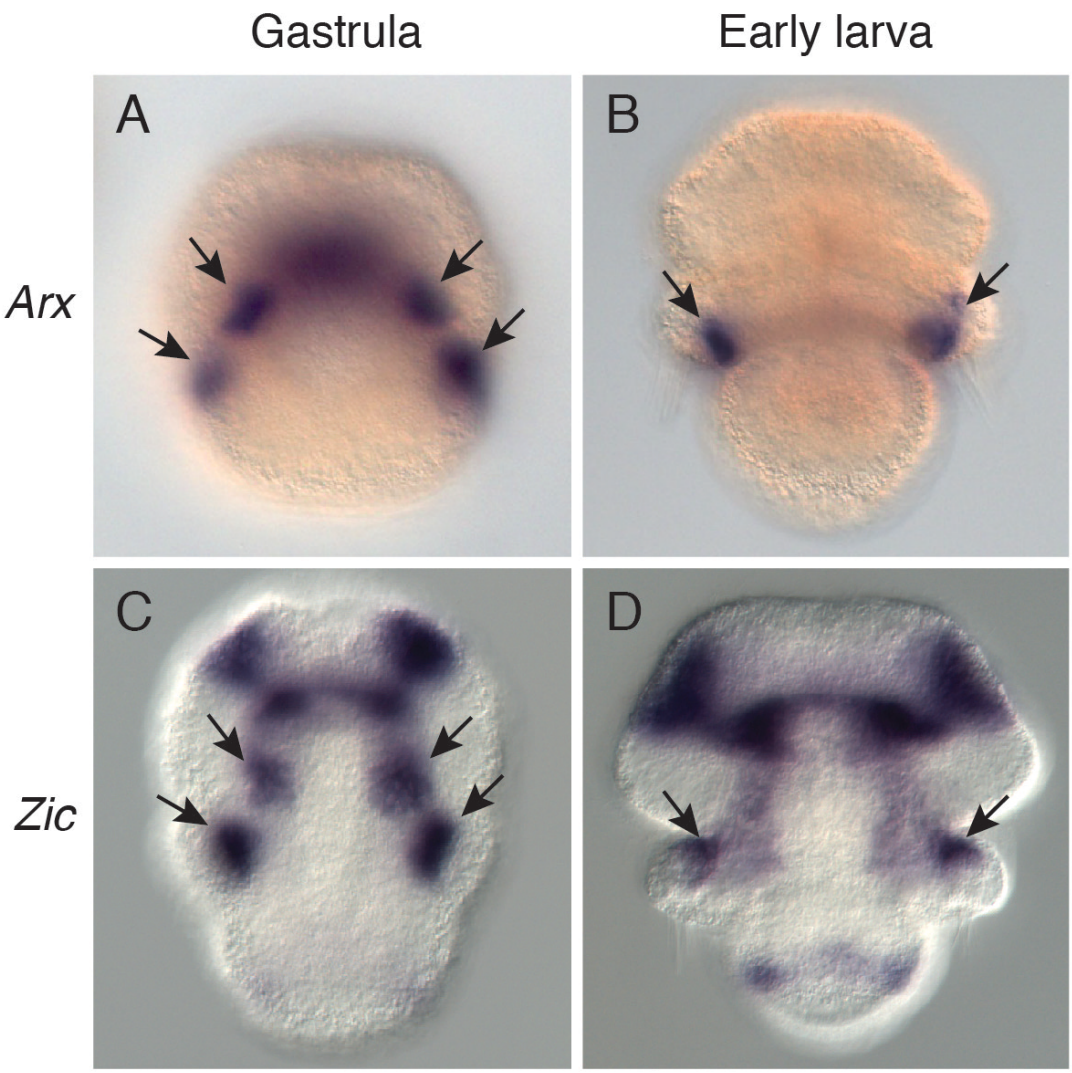

Figure S3. Expression of *Arx* and *Zic* during *T. transversa* embryogenesis. (**A–D**) Whole mount *in-situ* hybridization of *Arx* and *Zic* in gastrula embryos and early larvae of *T. transversa*. (**A**) In mid gastrulae, *Arx* is expressed in the ectoderm of the prospective chaetae sac territories (black arrows) and in a ventral domain. (**B**) In early larvae, *Arx* is expressed in the chaetae sacs (black arrows). (**C**) In late gastrulae, *Zic* is expressed in the mesoderm of the chaetae sacs (black arrows), apical lobe mesoderm and anterior ectoderm. (**D**) In early larvae, *Zic* is detected in the chaetae sacs (black arrows), in a domain in the pedicle lobe, and in the anterior mesoderm and anterior ectoderm. In all panels, the images are dorsal views, with the anterior pole to the top.

Table S1. Sequences and accession numbers used for Hox orthology assignment

Organism	Gene	Database	Accession number
H. sapiens	HoxA1	GenBank	AAB35423.2
	HoxB1		AAH99633.1
	HoxD1		AAG44444.1
	HoxA2		NP_006726.1
	HoxB2		NP_002136.1
	HoxA3		NP_705895.1
	HoxB3		AAD10852.1
	HoxD3		CAA71102.1
	HoxA4		NP_002132.3
	HoxB4		AAG45052.1
	HoxC4		AAG42145.1
	HoxD4		NP_055436.2
	HoxA5		CAG47052.1
	HoxB5		NP_002138.1
	HoxC5		EAW96748.1
	HoxA6		NP_076919.1
	HoxB6		NP_061825.2
	HoxC6		CAG33235.1
	HoxA7		CAA06713.1
	HoxB7		NP_004493.3
	HoxB8		AAG42143.1
	HoxC8		AAG42146.1
	HoxD8		AAG42152.1
	HoxA9		NP_689952.1
	HoxB9		AAG42144.1
	HoxC9		AAG42151.1
	HoxD9		NP_055028.3
	HoxA10		AAH07600.1
	HoxC10		NP_059105.2
	HoxD10		NP_002139.2
	HoxA11		NP_005514.1
	HoxC11		NP_055027.1
	HoxD11		AAF79045.1
	HoxC12		AAK16717.1
	HoxD12		AAF79044.1
	HoxA13		AAC50993.1
	HoxB13		AAH70233.1
	HoxC13		AAF73439.1
	HoxD13		AAC51635.1
	Gsx1		NP_663632.1
	Gsx2		NP_573574.1
	Pdx1		NP_000200.1
	Cdx-1		NP_001795.2
	Cdx-2		NP_001256.3
	Cdx-4		NP_005184.1
	Evx-1		NP_001291448.1

	Evx-2		NP_001073927.1
B. floridae	Hox1	GenBank	BAA78620
	Hox2		BAA78621
	Hox3		X68045
	Hox4		BAA78622
	Hox5		CAA84517
	Hox6		CAA84518
	Hox7		CAA84519
	Hox8		CAA84520
	Hox9		CAA84521
	Hox10		CAA84522
	Hox11		AAF81909
	Hox12		AAF81903
	Hox13		AAF81904
	Hox14		AAF81905
	Hox15		ACJ74394.1
	Gsx		AAC39015.1
	Xlox		AAC39016.1
	Cdx		AAC39017
	Evx-a		AAK58953.1
	Evx-b		AAK58954.1
P. flava	Hox1	GenBank	AAR07634.1
	Hox4		AAR07635.1
	Hox5		AAR07636.1
	Hox6		AAR07637.1
	Hox9/10		AAR07638.1
	Hox11/13a		AAR07639.1
	Hox11/13b		AAR07640.1
	Hox11/13c		AAR07641.1
	Xlox1		AAR07643.1
	Xlox2		AAR07644.1
S. kowalevskii	Gsx	Uniprot	A0A0U2UDE9
	Cdx	GenBank	NP_001158415.1
	Evx		NP_001164694.1
F. enflata	Hox1	GenBank	ABS18809.1
	Hox3		ABS18810.1
	Hox4		ABS18811.1
	Hox5		ABS18812.1
	Hox6		ABS18813.1
	Hox8		ABS18814.1
	MedPost		ABS18817.1
	Post-a		ABS18815.1
	Post-b		ABS18816.1
P. caudatus	Lab	GenBank	AAD40640.1
	Pb		AAD40641.1
	Hox3		AAD40642.1
	Dfd		AAD40643.1
	Ubx		AAD40647.1
	Abd-B		AAD40649.1
D. melanogaster	Lab	GenBank	CAB57787

	Pb		CAA45271
	Zen		AAF54087.1
	Zen2		P09090.2
	Dfd		P07548
	Scr		NP_524248
	Ftz		NP_477498
	Antp		CAA27417
	Ubx		CAA29194
	Abd-A		P29555
	Abd-B		CAB57859
	Ind		NP_996087.2
	Cad		AAA28409.1
	Eve		NP_523670.2
T. castaneum	Lab	GenBank	EEZ99257.1
	Mxp		NP_001107807.1
	Zen1		NP_001036813
	Zen2		AAK16425.1
	Dfd		AAK16423.1
	Cx		NP_001034523.1
	Ftz		AAK16421.1
	Ptl		NP_001034505.1
	Utx		EEZ99249.1
	Abd-A		EEZ99248.1
	Abd-B		EEZ99247.1
	Ind		AAW21974.1
	Cad-1		NP_001034498.1
	Cad-2		XP_008191732.1
	Eve		NP_001034538.1
C. teleta	Lab	GenBank	ABY67952
	Pb		ABY67953
	Hox3		ABY67954
	Dfd		ABY67955
	Scr		ABY67956
	Lox5		ABY67957
	Antp		ABY67962
	Lox4		ABY67958
	Lox2		ABY67959
	Post1		ABY67961
	Post2		ABY67960
	Gsx		AAZ23124.1
	Cdx		AAZ95508
	Xlox		AAZ95509.1
	Evx		ABG82164
H. robusta	Lab-a	Simakov *et al.* 2013	
	Lab-b	Simakov *et al.* 2013	
	Scr-a	Simakov *et al.* 2013	
	Scr-b	Simakov *et al.* 2013	
	Scr-c	Simakov *et al.* 2013	
	Dfd-a	Simakov *et al.* 2013	
	Dfd-b	Simakov *et al.* 2013	

	Lox5	Simakov *et al.* 2013	
	Antp	Simakov *et al.* 2013	
	Lox4-a	Simakov *et al.* 2013	
	Lox4-b	Simakov *et al.* 2013	
	Lox2	Simakov *et al.* 2013	
	Post2	Simakov *et al.* 2013	
L. anatina	Lab	ENSEMBL	g10891
	Pb		g10890
	Hox3		g10889
	Dfd		g10888
	Scr		g10887
	Lox5		g10886
	Antp		g10892
	Post1		g12396
	Post2		g12399
C. gigas	Hox1	ENSEMBL	CGI_10024083
	Hox2		CGI_10024086
	Hox3		CGI_10024087
	Hox4		CGI_10024091
	Lox5		CGI_10026565
	Lox2		CGI_10018592
	Lox4		CGI_10026562
L. gigantea	Lab	Simakov *et al.* 2013	
	Pb	Simakov *et al.* 2013	
	Hox3	Simakov *et al.* 2013	
	Dfd	Simakov *et al.* 2013	
	Scr	Simakov *et al.* 2013	
	Lox5	Simakov *et al.* 2013	
	Antp	Simakov *et al.* 2013	
	Lox2	Simakov *et al.* 2013	
	Lox4	Simakov *et al.* 2013	
	Post1	Simakov *et al.* 2013	
	Post2	Simakov *et al.* 2013	
O. bimaculoides	Hox1	ENSEMBL	Ocbimv22030263
	Scr		Ocbimv22018468
	Lox5		Ocbimv22010205
	Antp		Ocbimv22036189
	Lox2		Ocbimv22033340
	Lox4		Ocbimv22009726
	Post1		Ocbimv22015181
	Post2		Ocbimv22031197
G. varia	HoxA	GenBank	ACX84671.1
	Hox2		ADJ18233.1
	Hox3		ADJ18232.1
	Hox4		ACX84672.1
	Hox5		ADJ18234.1
	Lox5		ADJ18235.1
	Hox7		ADJ18236.1
	Lox2		ADJ18238.1
	Lox4		ADJ18237.1

	Post1		ACX84673.1
E. scolopes	Lab	GenBank	AY330184
	Hox3		AY330185
	Scr		AY330186
	Lox5		AY330187
	Antp		AY330188
	Lox4		AY330189
	Post1		AY330190
	Post2		AY330191
M. alaskensis	Lab	GenBank	KP762174
	Pb		KP762176
	Hox3		KP762173
	Dfd		KP762180
	Scr		KP762177
	Lox5		KP762179
	Antp		KP762171
	Lox4		KP762175
	Post2		KP762178
B. turrita	Pb	GenBank	AAS77225
	Hox3		AAS77226
	Dfd-a		AAS77227
	Dfd-b		AAS77228
	Lox5		AAS77229
	Post2		AAS77230
L. squamata	Lox4[1]	This study	

[1]`>Lepidodermella_squamata_Lox4`
`IITNAVTGANNGSSGKLMGAAHRTAPMYAWMAVVGPNSSQKRRGRQTYTRHQTIELEKEFAFCHYLARK`
`RRIELAAALSLSERQVKIWFQNRRMKLKKEKQQIADMNHISTSTTSTSNSSHSKSNRHDDYNDVNDASS`
`SDEDHLD`

Table S2. Primers used for qPCR experiments

	Forward	Reverse
Lab	CAAAGCTCCGTAGCCACTTA	TCGAGCTCTGTCAATTGCTT
Pb	AACAAATCGGATGGCTCTG	TTCATGGTCTGCTTCCTCTG
Hox3	ACTTCGCGTTAGCCAATCA	TGCAGGAACCCTTCAGAAA
Dfd	ATGCCGAGTATAAGCCGTTC	TATACCCGTGGATGAAACGA
Scr	ACGTCTGATGCCTGGTGTAG	ATAGCCATGAACAAATGCCA
Lox5	GTGTACGTTTGCCTGGTACG	GCATGTCGCAAGCGTATAGT
Antp	TCTCAAGCTCGAGTGTTTGG	GGAGACGCAGATAACGACAG
Lox4	GTTTGTCGACCGCGTCTT	AAATGGATACGGGTCTGCTC
Post2	GCTCCTGTGGCATTGTGTAG	AGCAAGCAAGCCCTGTAGAT
Post1	AACGTTGTCCCATTCTCTCC	CGATATACTATGCGGACCCA