Genomic resources and toolkits for developmental study of whip spiders (Amblypygi) provide insights into arachnid genome evolution and antenniform leg patterning

Guilherme Gainett[1], Prashant P. Sharma[1]

[1]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, USA 53706

Correspondence: guilherme.gainett@wisc.edu

Abstract

Background

The resurgence of interest in the comparative developmental study of chelicerates has led to important insights, such as the discovery of a genome duplication shared by spiders and scorpions, inferred to have occurred at the most recent common ancestor of Arachnopulmonata (a clade comprised of the five book lung-bearing arachnid orders). Nonetheless, several arachnid groups remain understudied in the context of development and genomics, such as the order Amblypygi (whip spiders). The phylogenetic position of Amblypygi in Arachnopulmonata posits them as an interesting group to test the incidence of the proposed genome duplication in the common ancestor of Arachnopulmonata, as well as the degree of retention of duplicates over 450 Myr. Moreover, whip spiders have their first pair of walking legs elongated and modified into sensory appendages (a convergence with the antenna of mandibulates), but the genetic patterning of these antenniform legs has never been investigated.

Results

We established genomic resources and protocols for cultivation of embryos and gene expression assays by in situ hybridization to study the development of the whip spider *Phrynus marginemaculatus*. Using embryonic transcriptomes from three species of Amblypygi, we show that the ancestral whip spider exhibited duplications of all ten Hox genes. We deploy these resources to show that paralogs of the leg gap genes *dachshund* and *homothorax* retain arachnopulmonate-specific expression patterns in *P. marginemaculatus.* We characterize the expression of leg gap genes *Distal-less*, *dachshund-1/2* and *homothorax-1/2* in the embryonic antenniform leg and other appendages, and provide evidence that allometry, and by extension the antenniform leg fate, is specified early in embryogenesis.

Conclusion

This study is the first step in establishing *P. marginemaculatus* as a model for modern chelicerate evolutionary developmental study, and provides the first resources sampling whip spiders for comparative genomics. Our results suggest that Amblypygi share a genome duplication with spiders and scorpions, and set up a framework to study the genetic specification of antenniform legs. Future efforts to study whip spider development must emphasize the development of tools for functional experiments in *P. marginemaculatus*.

Keywords: Arachnida | Arachnopulmonata | gene duplication | paralogs | leg gap genes | sensory biology

Background

In the past 25 years, comparative developmental study of chelicerates has resurged with the integration of molecular biology and genomics, spearheaded by developmental genetic

investigations of the spiders *Cupiennius salei* and *Parasteatoda tepidariorum* [1-6]. Such works have revealed several important aspects of chelicerate development, such as the molecular mechanism of dorso-ventral axis patterning [3,7-9], segmentation of the prosoma [1,10-13] and opisthosoma [2,13], and specification of prosomal versus opisthosomal fate [14,15]. A few non-spider chelicerate models have also emerged in recent years and provided new perspectives on chelicerate development, such as the horseshoe crab *Limulus polyphemus* (Xiphosura) [16,17] , the mites *Archegozetes longisetosus* [18-21] and *Tetranychus urticae* (Acariformes) [22-24], the tick *Rhipicephalus microplus* (Parasitiformes) [25], the Arizona bark scorpion *Centruroides sculpturatus* (Scorpiones) [26] and the harvestmen *Phalangium opilio* (Opiliones) [27-30]. Studies in *A. longisetosus* and *C. sculpturatus* have suggested that changes in Hox gene number and expression domains are responsible for such phenomena as the reduced segmentation of mites and the supernumerary posterior appendage identities of scorpions [18,26].

In addition to the insights into chelicerate development, one of the major recent outcomes of increasing availability of genomic resources in in the group was the discovery of a whole or partial genome duplication in spiders and scorpions. This genome duplication is inferred to trace back to the most recent common ancestor of the recently proposed clade Arachnopulmonata, which includes spiders, scorpions, and three other arachnid orders with book lungs [6,26,31-33]. This duplication event is also inferred to be independent from the multiple rounds of whole genome duplication undergone by horseshoe crabs (Xiphosura) [34,35]. The duplication in spiders and scorpions encompasses several important developmental genes, such as the homeobox family [6,32]. Excitingly, there is increasing evidence of divergent expression and function of paralogs, such as in the case of Hox genes, leg gap genes and Retinal Determination Gene Network homologs [6,26,36-40].

Nevertheless, knowledge on the development of several arachnid orders still remains scarce or entirely unexplored in the contexts of embryological study and genomic architecture. One particularly interesting arachnid order whose genomic evolution and developmental biology remains largely unexplored is Amblypygi. Commonly known as whip spiders, Amblypygi comprise approximately 220 described species of nocturnal predators that are distributed primarily in the tropics and subtropics worldwide [41,42]. Amblypygi is part of the clade Pedipalpi (together with the orders Thelyphonida and Schizomida), which in turn is sister group to spiders (Araneae) [31,43-45]. Considering the phylogenetic position of Amblypygi within Arachnopulmonata, a better understanding of Amblypygi genomics is anticipated to facilitate exploration of the evolutionary outcomes of the genome duplication in Arachnopulmonata, and better characterizing the extent to which arachnopulmonate orders retained and specialized the ensuing paralogous genes. Unfortunately, the embryology of the order Amblypygi is known only from the works of Pereyaslawzewa [46], a brief mention in Strubell [47] ("Phrynidae"), Gough [48], and the comprehensive study of Weygoldt [49] in *Phrynus marginemaculatus* (formerly *Tarantula marginemaculata*).

Beyond the interest in their genomic architecture, whip spiders possess a fascinating biology and natural history. Whip spiders are emerging as model organisms in behavioral ecology, learning, and neurophysiology in arachnids [50,51]. Their pedipalps, the second pair of appendages, are robust raptorial devices used for striking prey, including mostly arthropods, and even small vertebrates [41,50,52]. However, their most conspicuous characteristic and namesake is their "whip", a modified antenniform first walking leg that is not used for locomotion, but instead as a sensory appendage. The antenniform legs are extremely elongated relative to the other leg pairs, and the tibia and tarsus are pseudo-segmented into

hundreds of articles that harbor an array of chemo-, thermo-, hygro- and mechanoreceptive sensilla [53-55]. The peripheral circuitry of the antenniform legs is complex, exhibiting peripheral synapses and giant neurons which are possibly involved in fast sensory responses [55-57]. The antenniform legs also have been shown to be important for foraging, mating, and intrasexual contests [58,59]. Notably, concentration of sensory structures, elongation and pseudo-segmentation of these legs constitute a striking convergence with the antenna of mandibulate arthropods (Pancrustacea and Myriapoda). Evidence from Hox gene expression and functional experiments support the view that the antenna of mandibulates is positionally homologous to the chelicera of chelicerates (deutocerebral appendage) [21,30]. Despite their different positions along the antero-posterior axis of the body, the serial homology of antennae (mandibulates) and antenniform legs (whip spiders) to walking legs constitutes a potentially useful comparison to address whether a striking morphological convergence is associated with convergence in mechanisms of genetic patterning.

Most of the knowledge about the genetics of antenna versus leg fate specification in arthropods was discovered in the fruit fly *Drosophila melanogaster* and involves the regulation by leg gap genes and Hox genes. Broadly overlapping domains of leg gap genes *homothorax* (*hth*), *dachshund* (*dac*) and *Distal-less* (*Dll*) in the antenna promote antennal fate through activation of downstream antennal determinants such as *spineless* [60-63] (reviewed by [64]); leg identity in thoracic appendages is mediated by mutually antagonistic *hth* (proximal), *dac* (median) and *Dll* (distal) domains, and repression of *hth* by the Hox gene *Antennapedia* (*Antp*) [62,65,66]. Ectopic expression of *hth* on the leg discs results in leg-to-antenna transformation, whereas ectopic *Antp* in the antennal disc results in antenna-to-leg transformation [62,65,66]. In chelicerates, similar to *D. melanogaster*, the chelicera (antennal positional homolog) has broadly overlapping domains of *Dll* and *hth*; legs and pedipalps lack

a distal *hth*, and a proximal *Dll* expression domain, indicating a conservation of expression of positional homologs [28,67-69]. In line with this evidence, knockdown of *hth* in the harvestmen *Phalangium opilio* results in chelicera-to-leg transformation (and also partial pedipalp-to-leg transformation) [30]. In contrast to *D. melanogaster* and other insects, arachnid *Antp* is not expressed in the leg-bearing segments, and knockdown of this gene in the spider *Parasteatoda tepidariorum* has revealed a role in repressing limbs in the opisthosoma [14]. Expression of *dac* in a mid-domain of chelicerate pedipalps and legs suggests a conserved role in mid segment patterning, and, accordingly, knockdown of *dac* in the harvestmen *P. opilio* and the spider *P. tepidariorum* results in defects on the mid-leg segments [27,38]. Additionally, loss of a *dac* domain of the chelicera of spiders in comparison to harvestman has been implicated in the transition from three- to two-segmented chelicera [27,28].

Towards revitalizing embryological studies in Amblypygi, we generated herein protocols and comprehensive genomic resources for developmental genetic study of the whip spider *Phrynus marginemaculatus* (Fig. 1A–B), resuming the seminal studies begun by Peter Weygoldt with this promising model species. Leveraging a comprehensive developmental transcriptome for this species, we first investigated the presence and retention of systemic gene duplications in Amblypygi predicted from their phylogenetic position in Arachnopulmonata; to mitigate biases stemming from survey of a data point from a single exemplar, we included as well the embryonic transcriptomes we recently generated for two species of the distantly related family Charinidae (*Charinus ioanniticus* and *Charinus israelensis*) [40,41,70]. As a first step towards elucidating the patterning of antenniform legs, we investigated a possible role of the leg gap genes *homothorax*, *dachshund* and *Distal-less* in specifying antenniform leg fate, by characterizing their expression pattern in developing

embryonic stages of *P. marginemaculatus* using newly established procedures for colorimetric whole mount in situ hybridization.

Results

Overview of embryogenesis in Amblypygi

A characterization of the postembryonic development of *P. marginemaculatus* is available in the works of Weygoldt [71,72]. The following summary is after Weygoldt [49], complemented with observations from the present study. We note that the timing of embryonic stages is approximate, as the temperature in a previous work [49] was kept in a range of 24–27˚C. Details on the culture and collection of embryos are described in *Methods*.

In our colony, a sample of 18 clutches from different females had an average of 19 eggs (min: 8 eggs; max: 39 eggs). Each egg is ovoid and measures 1.2–1.4 mm at the major axis. The yolk granules are light green, and fixation turns them light yellow. Upon secreting the brood sac and eggs into the ventral side of the opisthosoma, the brood sac is transparent, but gradually hardens and becomes dark brown in the course of two days. Around 6 dAEL, a blastoderm is formed. At 7–8 dAEL, the yolk is more condensed, a perivitelline space is formed, cells move into the yolk presumably in the region of the future blastopore. At 8–9 dAEL a blastopore is recognizable on the dorsal surface of the embryo, and a germ disc is formed. Towards 10 dAEL, the germ band starts to take form, the blastopore region is displaced posteriorly, and the first signs of segmentation of the prosoma are visible. A presumable cumulus, a group of migrating cells with dorso-ventral organizing properties, was reported by Weygoldt (1975) at this stage, but was not investigated by us. At 11-12 dAEL, the nascent prosomal segments elongate anteriorly and by 13-14 dAEL the segmented germ

band occupies the ventral hemisphere of the egg. The limb buds of prosomal appendages begin to appear (Fig. 2A). The process of reversion begins, with a sagittal split of the germ band (Fig. 2A). From this point, until 19 dAEL, the ventral sulcus progressively enlarges with the dorsal displacement of both halves, and then begin to ventrally fuse again starting anteriorly (Fig. 2A–E). The head lobes on 15 dAEL have a circular outline and uniform ectoderm (Fig. 2F). From 15–19 dAEL, ectodermal cells invaginate along the ventral midline and head lobes (neural precursor cells), which can be seen as evenly spaced dark dots (Fig. 2F–H). Concomitantly, the anterior rim of the head lobes forms semilunar groves (i.e., folds over itself) (Fig. 2F–H). Opisthosomal segments 1–12 sequentially appear from the posterior growth zone (Fig. 2I–K). At 19 dAEL, the yolk moves into the caudal papilla (growth zone) and the opisthosoma flexes ventrally (Fig. 2K, N). The first pair of legs (future antenniform legs) is already longer than the other appendages in an early limb bud stage (15 dAEL), and more clearly in subsequent stages (Fig. 2L–N). A protuberance on the proximal end of the second walking leg pair (L2), the nascent lateral organ (osmoregulatory embryonic organ), is discernible at 16 dAEL and later stages, and becomes reniform (Fig. 2L–N). From 20–21 dAEL, the embryo begins to secrete a cuticle, which is darker on the fully formed lateral organ. The head lobes have completely folded onto themselves, and the distal segment of the chelicera is sclerotized. The embryo then ruptures the egg shell, and is now termed deutembryo (Fig. 2O–Q). Deutembryos remain in the brood sac for further 70 days without discernible external changes. The primordia of the eyes are formed towards the end of deutembryo development (70–106 dAEL), and eye spots of the median and lateral eyes become visible. From 90-106 dAEL the deutembryos molt into praenymphs, leave the brood sac and climb onto the adult female's back (Fig. 1B). For a histological description of gastrulation, of the mesoderm, brain and nervous system, digestive system and gonad formation, we refer the reader to the exquisite original descriptions of Weygoldt [49].

Phrynus marginemaculatus embryonic transcriptome

In order to provide a comprehensive resource for Amblypygi to facilitate developmental genetic study and comparative genomics, we assembled an embryonic transcriptome spanning different embryonic stages. The assembly of *P. marginemaculatus* transcriptome resulted in 544.372 transcripts composed of 277.432.101 bp and N50 of 440 bp. BUSCO scores, which reflect the presence and number of copies of widely conserved single-copy genes (BUSCO-Arthropoda; n=1066), indicated 94.2% completeness with 5.9% of genes duplicated. From the total number of transcripts (including isoforms), we predicted 37,540 peptide sequences with the Trinotate pipeline with a top blastp hit in Swiss-Prot database. A total of 801 of these peptides (2.1%) had a top hit to non-metazoan proteins, and presumed microbial contaminants. The total number of predicted proteins is higher in comparison with the predicted protein set from the genome of the spider *Parasteatoda tepidariorum* (27,900) and the scorpion *Centruroides sculpturatus* (30,456) [6], which is attributable to the typical fragmentation of genes in such transcriptomic assemblies.

To investigate the possibility of systemic gene duplication in Amblypygi and further test the transcriptome completeness, we searched for all ten Hox genes, emphasizing discovery of paralogs known from the genomes of the spider *Parasteatoda tepidariorum* (except for *fushi tarazu*) and the scorpion *Centruroides sculpturatus* (except for *Hox3*) [6,26,32]. One of the *Hox3* copies of *P. tepidariorum* has a highly divergent homeodomain sequence, constituting an unusual long branch in our phylogenetic analysis (Additional file 1, Fig. S1). The same is true for the *Hox3* ortholog *bicoid* of *Drosophila melanogaster* (Additional file 1, Fig. S1). We therefore removed both paralogs from our main analysis. We found orthologs of all ten Hox genes in *Phrynus marginemaculatus*, namely *labial*,

proboscipedia, *Hox3*, *Deformed*, *Sex combs reduced*, *fushi tarazu*, *Antennapedia*, *Ultrabithorax*, *abdominal A* and *Abdominal B* (Fig. 3A–B). We found evidence that all Hox genes, except *proboscipedia* are assembled in two genes, and the predicted proteins sequences overlap and have amino acid differences in *P. marginemaculatus* transcriptome (Fig. 3B). A *Pmar-Dfd* and *Pmar-abdA* were assembled in a third gene, but their non-overlapping sequences on the alignment and nearly identical amino acid sequence to the respective paralog 2 indicate that they are fragmented assemblies of the same copy. We additionally analyzed two embryonic transcriptomes of *Charinus* whip spiders recently generated by us [40]. For *C. israelensis*, we discovered orthologs of all ten Hox genes, with two genes for all cases except *ftz* and *Antp* (Fig. 3A–B). A third gene was annotated as *Cisr-Dfd*, but we interpret it as an assembly error, given its largely non-overlapping sequence with respect to the other *Cisr-Dfd* sequences. For *C. ioanniticus*, we discovered one gene each for *ftz*, *Antp*, *Ubx*, *AbdB*, and two genes each for *Hox3*, *Dfd*, and *Scr*. We did not detect orthologs of *lab*, *pb* and *AbdB* in *C. ioanniticus* transcriptome (Fig. 3A–B).

These absences almost certainly reflect incompleteness of transcriptomic assemblies, and cannot be equated with true losses. We therefore consider the union of the Hox gene counts across the three species as the reconstruction of the whip spider ancestral Hox complement. Our results suggest that the ancestral Hox complement in Amblypygi consists of 20 genes (two copies of every Hox gene), a trait they share with spiders and scorpions. While a genome is not available for whip spiders, the phylogenetic placement of this order, as well as the pattern of Hox duplication, strongly suggests that whip spiders share the condition of two Hox clusters observed in spider and scorpion genomes.

Gene expression analyses in *Phrynus marginemaculatus*

All three methods we trialled for fixing embryos of *P. marginemaculatus* yielded embryos which are suitable for in situ hybridization (ISH). However, fixation of intact embryos in formaldehyde/heptane phase sometimes results in the vitelline membrane adhering to the embryo. Upon dissection of the vitelline membrane for ISH, parts of the embryo are commonly lost. Piercing a hole in the embryo for fixation in formaldehyde achieves a similar result, and is therefore not recommended. Dissecting the vitelline membrane under PBS and immediately fixing the embryos in formaldehyde yielded the best results. We did not detect any difference in the signal-to-background ratio of colorimetric ISH between embryos fixed for 2 h or 24 h. To test gene expression assays and investigate the genetic patterning of antenniform legs, we selected the leg gap genes *Distal-less*, *dachshund* and *homothorax*, which constitute the subset of PD axis patterning genes associated with functional datasets in other arachnid model species [27,30,73].

Distal-less identification and expression

Similar to other arachnopulmonates and arthropods in general, a single *Distal-less* gene, *Pmar-Dll*, was found by similarity searches and its orthology to other metazoan *Distal-less* genes was confirmed with a phylogenetic analysis (Additional file 2, Fig. S2). *Pmar-Dll* is strongly expressed in all developing appendages, from early limb bud stage (15 dAEL) and is localized to the distal portion of the chelicera, pedipalp and legs throughout leg elongation (Fig. 4A–I; Fig. 5A–L). In 17-18 dAEL embryos, the expression is heterogeneous and forms bands of stronger expression on the forming podomeres of the pedipalps and legs (Fig. 5H–L). We did not detect expression on ventral midline (*contra* [28] and[16]), or in the opisthosoma, and observed only diffuse expression on the head lobes of the stages surveyed. Expression on L1 of 16-17 dAEL embryos is subtly stronger on the proximal part of the

expression domain (Fig. 5C–C'), but we did not detect clear differences in younger or older embryonic stages. The proximal and distal boundaries of strong *Distal-less* expression do not differ between L1 and the other legs, spanning two-thirds of the legs from the tip of the tarsus and excluding the most proximal third (Fig. 5C, 5I).

In contrast to other arachnid embryos (mygalomorph spiders [69]; araneomorph spiders [67,73]; Opiliones, Sharma et al. 2012; and mites [20,74]), *Dll* expression was not detected as a strong domain in the endite of pedipalp, which is very inconspicuous at these embryonic stages. We were not able to assess *Dll* expression in the pedipalpal endite in deutembryo stages due to deposition of cuticle.

dachshund identification and expression

dachshund is typically present as a single copy in Arthropoda, but is found independently duplicated in Arachnopulmonata and Xiphosura [38,75]. We found two genes with high similarity to *Dmel-dac* and confirmed their identity as dachshund paralogs with a phylogenetic analysis (Additional file 3, Fig. S3). *Pmar-dac1* and *Pmar-dac2* are each nested in a clade with paralogs 1 and 2 of other arachnopulmonates, which appears as independent from horseshoe crab duplications (Additional file 3, Fig. S3).

Pmar-dac1 is strongly expressed as broad median band in the pedipalp and the four pairs of legs (Fig. 6A–F). This sharp mid band excludes the most proximal and the distal domains of these appendages (Fig. 7). No staining was detected on the chelicera (Fig. 6A–F; 7A, E). Additionally, expression occurs in three domains in the head lobes and strongly on the ventral

midline of the prosoma. The continuous expression on the ventral midline stops before the opisthosoma (Fig. 6A–F).

Pmar-dac2 is expressed as ring at the proximal region of chelicera, pedipalp and legs (Fig. 6G–I). On 16-17 dAEL embryos, a second thinner ring of expression is also present on the pedipalps and legs (Fig. 6G–I; 7I–L). Expression domains also occur on the neuromere of L1 and pedipalp, being larger on the pedipalpal neuromere (Fig. 6G). Three spot domains also occur on O1, O2 and O3 segments (Fig. 6H–I).

homothorax identification and expression

Similar to the case with arachnid *dachshund*, *homothorax* is also found duplicated in arachnopulmonates [32,39]. We annotated two genes in *P. marginemaculatus* transcriptome and the phylogenetic analysis shows that each copy is nested in two clades with the arachnopulmonate paralogs (Additional file 4, Fig. S4).

Pmar-hth1 is conspicuously expressed on the developing appendages throughout the stages studied (15-18 dAEL; Fig. 8–9). On 16-17 dAEL embryos, *Pmar-hth1* is expressed more strongly on the proximal parts of the chelicera, pedipalp, and legs, and is absent in a small distal domain (Fig. 9A–D). On 17-18 dAEL, expression in all elongated appendages is similar, and spans most of the appendage length except for a small distal territory (Fig. 9E-H). Additionally, three stronger stripe domains occur distal to the coxa of pedipalp and legs. Expression in L1 and other legs is similar in the studied stages, save for a slightly expanded distal domain on the proximal tarsus of L1 (Fig. 9G–H). Expression on the ventral ectoderm

and head lobes is weak and homogenous in early stages, and becomes stronger in 17-18 dAEL (not shown)

Pmar-hth2 is expressed in the developing appendages, but is reduced in intensity and different in pattern in comparison to *Pmar-hth1* (Fig. 8B–B'). *Pmar-hth2* spans the whole chelicera (Fig. 9I). Expression on the pedipalp spans most of the appendage, excluding only a distal territory (Fig. 9J). In addition, there is a proximal and a sub terminal strong domain (Fig. 9J). Expression on L1 and other legs is identical in all stages studies. *Pmar-hth2* expression on the legs is consistently weaker in comparison to the pedipalp, but likewise possess a proximal and a distal sub-terminal domain (Fig. 9K–L). Expression on the ventral ectoderm, opisthosoma and head lobes is ubiquitous (Fig. 8B).

Discussion

Genomic resources for whip spiders corroborate widespread retention of duplicates across Arachnopulmonata

Changes in Hox gene expression, regulation and targets are linked to changes in morphology and Hox gene duplications have also been implicated in morphological innovations [76-79]. Duplication in the Hox cluster is common in the vertebrates, which may have up to eight Hox clusters (teleost fishes) due to successive rounds of genome duplications [76]. On the other hand, genome and Hox cluster duplications are considered rare in Arthropoda [80]. Complete Hox cluster duplication in the phylum has only been reported in horseshoe crabs (Xiphosura) [34,35,81,82], and in spiders and scorpions (presumed ancestral in Arachnopulmonata). The phylogenetic placement of Amblypygi

makes this group an opportune system to test the downstream prediction that *P. marginemaculatus* retains and expresses Hox paralogs shared by spiders and scorpions. In accordance with this hypothesis, we discovered that the whip spider *P. marginemaculatus* presents homologs of all ten Hox genes, and all excepting *proboscipedia* appear to be duplicated, as supported by our alignment and phylogenetic analysis. Our survey in the two *Charinus* transcriptomes confirmed the duplication of several Hox genes, and additionally revealed that two copies of *proboscipedia* (absent in *P. marginemaculatus*) are present at least in one species. Our results suggest that Amblypygi retain a complete duplicated Hox complement with 20 genes. Similarly, a recent transcriptomic survey in *Charinus acosta* and *Euphrynichus bacillifer* reported duplications of nearly all Hox genes (excepting *proboscipedia*), as well as Wnt and Frizzled genes [83]. Taken together with the recent discovery of duplicated RDGN genes in *Charinus*, our results suggest that Amblypygi retain ohnologs of numerous developmental patterning genes, as a function of an ancient, shared whole genome duplication.

A second line of evidence of a shared genome duplication event comes from conserved expression patterns of leg gap gene paralogs. It has recently been shown that duplicates of leg patterning genes *extradenticle*, *optomotor blind*, *dachshund* (*dac*) and *homothorax* (*hth*) present one copy with an expression pattern resembling a stereotypical panarthropod domain, and a second copy with a divergent expression, with these two sets of expression patterns shared by a spider and a scorpion [75]. For instance, *dac1* of surveyed arachnopulmonates retains the plesiomorphic expression pattern of a broad median band in the appendages, the condition observed in the single-copy *dac* of all non-arachnopulmonate arachnids surveyed to date [20,28], whereas *dac2* is expressed as a proximal and a medial domain at the patella-tibia boundary [38,67,69,75]. Similarly, *hth1* retains a plesiomorphic expression pattern on

the proximal-median appendage excluding the tarsus, whereas *hth2* presents a strong short proximal domain and stripes of expression at podomere boundaries [39,67,69,75]. In the whip spider *P. marginemaculatus*, we indeed recovered two copies of *dachshund* and *homothorax*, with paralogs nested in two clades of arachnopulmonate copies with high nodal support (Additional file 3, Fig. S3; Additional file 4, Fig. S4). Moreover, *Pmar-dac* and *Pmar-hth* have paralog-specific expression patterns shared with spiders and a scorpion: *Pmar-dac1/2* and *Pmar-hth1/2* conserve the plesiomorphic expression pattern of arachnids, while copy 2 retains the arachnopulmonate-specific pattern. These data, together with the evidence from Hox gene duplications and the topologies of the *dac* and *hth* gene trees, provide compelling support for genome duplication as a synapomorphy of Arachnopulmonata.

A *dachshund2* domain in the two-segmented clasp-knife chelicera of *P. marginemaculatus*

Expression of *Pmar-Dll*, *Pmar-dac1/2* and *Pmar-hth1/2* on the appendages is largely similar to spiders and scorpions (see above) [38,39,67-69,75]. A notable difference in the chelicera of the whip spider is the retention of proximal expression of one of the *dachshund* paralogs. In arachnids with three segmented chelicerae, such as the daddy-long-legs *Phalangium opilio*, *dac* is expressed in the proximal segment of the appendage [28]. The same is true for the three-segmented chelicera of scorpions, in which both *dac1* and *dac2* are expressed in the proximal segment [75]. In the two-segmented chelicera of spiders, *dac1* is not expressed in the chelicera save for a non-ectodermal faint staining in late embryonic stages [38,67-69,84]. *dac2* of spiders is also largely not expressed in the chelicera, save for a small ventral domain at the base of the chelicera of *Parasteatoda tepidariorum* and *Pholcus phalangioides* [38]. It has been proposed that the loss of a proximal *dac* domain is associated with the loss of the

proximal segment in the chelicera of spider and other arachnids. This hypothesis is supported by functional data in the daddy-long-leg *P. opilio*: knockdown of the single-copy *dac* resulted in the loss of the proximal segment of the chelicera (as well as mid-segment truncations in pedipalps and legs) [27]. Similar to spiders, the two-segmented chelicera of the whip spider *P. marginemaculatus* does not express *Pmar-dac1*. Nonetheless, a clear complete ring of proximal expression was detected for *Pmar-dac2*.

In contrast to the classic leg gap phenotype incurred by knockdown of the ancestral *dac* copy in the harvestman, knockdown of *dac2* in the spider *P. tepidari*orum resulted in the loss of the patella-tibia segment boundary in the pedipalp and walking legs; effects of this knockdown on the chelicera were not reported by the authors, nor is the function of *dac1* known in *P. tepidariorum* [38]. Given the interpretation that *dac2* has a derived role in patterning a segment boundary in the spider, the incidence of the *Pmar-dac2* stripe in the middle of the whip spider chelicera (and the absence of *Pmar-dac1* in the chelicera) is tentatively hypothesized here to constitute evidence of a segment boundary patterning activity that separates the basal and distal cheliceral segment of Amblypygi (rather than evidence of a proximal segment identity in whip spiders that is homologous to the proximal cheliceral segment of groups like harvestmen and scorpions). Exploring this hypothesis in future will require the establishment of RNAinterference tools in *P. marginemaculatus.*

Antenniform leg distal allometry is specified early in development

To investigate a possible role of changes in leg gap genes expression in the patterning of the antenniform leg pair (L1) of whip spiders, we specifically searched for changes in expression domain between embryonic antenniform leg (L1) and the posterior legs (L2–4).

The expression patterns of *Pmar-Dll*, *Pmar-dac1/2* and *Pmar-hth1/2* are mostly identical between the antenniform leg L1 and the posterior legs. We found only a slightly expanded *Pmar-hth1* domain in the distal leg (compare Fig. 9G–H). It is unclear if this difference bears any functional significance in the differential fate of L1. Similar distal expression dynamics appear to distinguish the identity of the pedipalp and walking legs of the harvestman, as knockdown of the single copy of *hth* in *P. opilio* results in homeotic transformation of pedipalps to legs [30].

Besides the difference in sensory equipment, the antenniform leg pair of adults differ from the posterior walking legs in being much longer, and presenting numerous articles (pseudo-segments) on the tibia, metatarsus and especially the tarsus, whereas the walking legs have pseudo-segments only in the tarsus [41,53,55]. Antenniform legs may be more than 2.5 times longer than the walking legs, and this allometry in adults is mostly accounted for by the length added by the distal leg segments, which may attain more than 100 distal articles (tibiomeres + tarsomeres) [53]. Understanding the timing of the ontogenetic divergence of L1 relative to the walking legs is crucial for circumscribing target stages for investigation of the genetic patterning of this phenomenon.

Current evidence indicates that this allometry on the antenniform legs takes place both embryonically and postembryonically. The praenymph, the stage that molts from the deutembryo and climbs to the mother's back, has the full number of tibiomeres, but not of tarsomeres. The final number of tarsomeres is added in the next molt to the protonymph [53,85], and from this moment onwards the subsequent growth is isometric [53,85]. In embryos of *P. marginemaculatus*, Weygoldt [49] observed that the primordia of L1 are the first to appear in the germ band and are already longer than the remaining prosomal

appendages, an observation paralleling the allometric growth of the L2 limb bud of the harvestman *P. opilio* (wherein L2 is the longest leg and also bears the most tarsomeres) [29]. It is presently unclear at what point in development the distal podomeres of antenniform legs attain their allometry with respect to the proximal podomeres in the whip spider. In other words, are embryonic legs simply growing longer overall with respect to walking legs, or are they longer due to distal growth? The expression pattern *dachshund* in *P. marginemaculatus* early embryos reveals a further interesting aspect of this allometry. The expression of *dac* (including duplicates) in the legs of arachnids surveyed to date occurs on podomeres proximal to the tibia, so that the tibia-metatarsus-tarsus territory may be identified as the a *dac*-free domain [20,27,38,67,69,75,84]. The distal boundary of *Pmar-dac1* and *Pmar-dac2* in stages long before podomere formation reveals that the distal domain (*Pmar-dac1*-free cells) of L1 is already longer with respect to the proximal leg and the other legs (compare Fig. 7C, D). These data support the inference that the distal allometry, and by extension the antenniform leg fate, are already specified from the onset of limb bud development.

Conclusion

The study of the embryology of Amblypygi has faced 45 years of hiatus since the seminal work describing the development of *Phrynus marginemaculatus* [49]. Here, we take the first steps in establishing *P. marginemaculatus* as a model for modern chelicerate evolutionary developmental study and for comparative genomics. We provide evidence for a shared genome duplication in Amblypygi; show that the distal allometry of antenniform legs is specified early in appendage embryogenesis; and observe that expression patterns of the leg gap genes *Distal-less*, *dachshund* and *homothorax* do not substantian their role in fate specification of the sensory legs of whip spiders. Future efforts to uncover the genetic

underpinnings of this fascinating convergence with the mandibulate antenna should prioritize discovery of leg fate specification mechanisms in early whip spider development. Despite the establishment of a few promising arachnid models for studying chelicerate embryology, a pervasive challenge has been the establishment of toolkits for testing gene function [3,22,27,73]. Therefore, future efforts to study whip spider development must emphasize the development of tools for functional experiments in *P. marginemaculatus*.

Methods

Establishment of amblypygid colony in the laboratory

Adult individuals of *Phrynus marginemaculatus* were purchased from a commercial supplier collects animals in the Florida Keys (USA). The maintenance of a colony of *Phrynus marginemaculatus* (Fig. 1A–B) in the laboratory followed guidelines provided is the original developmental works of Peter Weygoldt [49,71,72] and notes on behavioral studies on this species [86-88].While most species of Amblypygi have been reported as aggressive and better kept in separate containers [41], individuals of *P. marginemaculatus* can be kept in a stable laboratory colony in large community containers [86]. A maximum of 20 animals may be allowed to interact and mate freely in plastic boxes (67.3 x 41.9 x 32.4 cm) with a 2 cm layer of damp coconut fiber saturated with distilled water. The substrate is continuously dampened with approximately 500 mL of dH_2O on the substrate every week. All animals are kept in an acclimatized room at 26-28˚C in a controlled reversed 16L:8D light cycle to simulate summer months in Florida. To prevent flies from reproducing in this humid environment, the lid of the container is provided with a 10 x 20 cm fine mesh cloth, which also allows for air circulation. Each container is furnished with pieces of tree bark, egg

carton, and retreats made from polystyrene foam or florist's foam, to provide enough space for animals to hide and suspend themselves during molting and egg-laying. After about two weeks of setting up a terrarium, the substrate hosted a stable colony of collembolans (possibly brought by the tree bark). The co-occurrence of collembolans does not appear to be disadvantageous in any way. Collembolans are invariably observed feasting on carcasses and likely contribute to preventing mite infestations (see also [41]). Animals are fed at least twice a week with nymphs of *Acheta domestica ad libitum.* Carcasses are removed as soon as detected to prevent the spread of mites. Freshly hatched praenymphs are kept with the female, and aggregate on the back of the brooding female (Fig. 1B). The next instar, the protonymph, disperse from the female upon molting. Protonymphs are transferred to community terraria as described above. Protonymphs and subsequent instars are fed pinheads of *A. domestica* and *Gryllus bimaculatus ad libitum* three times per week until adulthood. The first adult individuals from our laboratory-bred animals were detected 8 months after hatching, which accords to what has been described by Weygoldt [49].

Terraria with adults are inspected daily for freshly molten individuals or ovigerous females, which are immediately separated from the main colony to avoid aggressive interactions. Ovigerous females carry the eggs outside the body in a brood sac attached to the ventral side of the opisthosoma [41]. Eggs are allowed to develop with the female until the desired stage for developmental work. Approximate stages may be inferred from days after egg laying (dAEL) at 26˚C following the staging in [49]. For obtaining the eggs, females are anesthetized with CO_2 and immobilized in customized plastic petri dishes (35 x 11 mm) with a central whole and foam (Fig. 10A–B). With a blunt forceps pair, the pleura of the opisthosoma embracing the eggs may be gently displaced from the brood sac, and the brood sac removed as whole (Fig. 10C). The female may be returned to the colony unharmed. The

brood sac easily dissolves when immersed in a phosphate buffer solution (PBS), and eggs may be cleaned with forceps under PBS. We obtained high mortality in keeping isolated eggs developing in petri dishes with humid (wet filter paper) or dry in 26˚C incubators, as mold frequently develops. However, we found out that embryos will survive and develop immersed in PBS supplemented with kanamycin (1:5000) at 26˚C incubator and even hatch into deutembryos (~20 dAEL).

Phrynus marginemaculatus transcriptome assembly

We collected embryos of different embryonic stages as described above, and conducted total RNA extractions of each stage separately: hatched deutembryo (female #5, n=3; undetermined age, before eyespots), unhatched deutembryo (female #2 n=3; approx. 21 dAEL), and leg elongation stage (female #3, n=2; approx. 16-17 dAEL), late limb bud stage (female 6, n=1; 16 dAEL) and early limb bud stage (female 6, n=1; 14 dAEL) (voucher images are available upon request). Embryos were homogenized in Trizol TRIreagent (Invitrogen) and stored at –80˚C until extraction. The combined total RNA of the five extractions was submitted to cDNA library preparation and sequencing in an Illumina High-Seq platform using paired-end short read sequencing strategy (2×100 bp) at the Center for Gene Expression, Biotechnology Center of University of Wisconsin-Madison (USA). The quality of the raw reads was assessed with FastQC (Babraham Bioinformatics). De novo assembly of compiled paired-end reads was conducted with the software Trinity v. 2.6.6 [89], using Trimmomatic v. 0.36 [90] to remove low quality reads and adaptors. Completeness of the transcriptome was assessed with the Trinity script '*TrinityStats.pl*', and BUSCO v. 3 [91] on the longest isoforms using the database 'Arthropoda' and analysis of Hox genes. Whole transcriptome annotation was conducted following the Trinotate pipeline, by predicting

proteins with TransDecoder v. 5.2.0 [92] and searching best hits with blastx and blastp against Swiss-Prot and Pfam databases [93].

Gene identification

Hox genes (*labial*, *proboscipedia*, *Hox3*, *Deformed*, *Sex combs reduced*, *fushi tarazu*, *Antennapedia*, *Ultrabithorax*, *abdominal A* and *Abdominal B*), *Distal-less*, *dachshund*, and *homothorax* orthologs were identified from the embryonic transcriptome of *Phrynus marginemaculatus* using BLAST searches (tblastn) with peptide queries of the homologs of *Drosophila melanogaster* and *Parasteatoda tepidariorum*. The best hits were reciprocally blasted to NCBI database and further assessed with a phylogenetic analysis.

For the phylogenetic analysis of Hox genes, we selected arthropod terminals with available genomes and well annotated references: the fruit fly *Drosophila melanogaster* (FlyBase), the beetle *Tribolium castaneum* (GCA_000002335.3) (Insecta), the centipede *Strigamia maritima* (GCA_000239455.1) (Myriapoda), the tick *Ixodes scapularis* (GCA_002892825.2), the mite *Tetranychus urticae* (GCA_000239435.1), the spider *Parastetoda tepidarioum* (GCA_000365465.2) and the scorpion *Centruroides sculpturatus* (GCA_000671375.2) (Chelicerata) [6,94-97] (Additional file 5, Table S1). For the analysis of *Distal-less*, *dachshund*, and *homothorax* we selected terminals used in Nolan et al. [75] and included the peptide sequence for *P. marginemaculatus* candidates (Additional file 5, Table S1). Peptides were predicted with ORF Finder (NCBI), TransDecoder v. 5.5.0 [92], or obtained as CDS when possible from publicly available NCBI database. Peptide sequences were aligned using Clustal Omega [98] in the software SeaView v. 4 [99]. Gene trees were inferred under maximum likelihood using IQTREE [100] with automatic model selection (-m TEST) and

1000 ultrafast bootstrap resampling (-bb 1000). Alignments for the Hox gene analysis are available in Additional files 6–7. Genes referred throughout text refer to the longest isoform assembled by Trinity in a contig, and annotated by the procedure detailed above. Paralogs were discerned from gene fragments using as criteria the presence of largely overlapping sequences with amino acid differences.

Fixation and in situ hybridization

Collecting freshly laid eggs and keeping them in controlled temperature with the female or incubator (26˚C) allows to select desirable stages for fixation. We fixed embryos for in situ hybridization with three methods: (1) intact eggs in a phase of 4% formaldehyde/PBS and heptane overnight, after an established spider fixation protocol [3]; (2) piercing a hole in the vitelline membrane, and immersing eggs in 4% formaldehyde/PBS overnight; (3) dissecting the embryo out of the vitelline membrane and fixing in 4% formaldehyde/PBS for periods varying from 2–24h. All three procedures were conducted in a shaking platform at room temperature (~22 ˚C). Following the fixation period, embryos were rinsed several times in 1x PBS/0.1% Tween-20 (Sigma Aldrich) (PBS-T), and gradually dehydrated into pure methanol. Embryos were then stored in 20 ˚C freezer until next procedures.

cDNA was synthetized with SuperScriptIII kit (Thermo Fisher) following the manufacturer's instructions, using the same total RNA used to generate the embryonic transcriptome. Riboprobe templates for the genes *Pmar-dac1*, *Pmar-dac2*, *Pmar-hth1*, and *Pmar hth-2* in situ hybridizations were generated as follows: genes were PCR amplified from cDNA using gene-specific primers designed with Primer3 v. 4.1.0 [101] and complemented with T7 linker sequences (5′ -ggccgcgg-3′ ; 5′ - cccggggc-3′), and a second PCR using universal T7 primer

to generate T7 polymerase templates specific for sense (control) and anti-sense (signal) probes (Additional file 8, Table S2). Sense and antisense riboprobes for *Pmar-Dll* were generated with T7 and T3 polymerases using a plasmid template. For cloning of *Pmar-Dll*, gene-specific primers were used to amplify an 848 bp fragment (Additional file 8, Table S2), which was incorporated to vectors and *E. coli* using TOPO TA cloning kit (Thermo Fisher) following the manufacturer's instructions, and Sanger sequenced to confirm their identities.

For colorimetric in situ hybridization, we adapted a protocol for the spider *Parasteatoda tepidariorum* after Akiyama-Oda and Oda (2003). The staining reactions lasted between 2h and 8h at room temperature using nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-3′-indolyphosphate (BCIP) staining reactions (Sigma Aldrich). Embryos were then rinsed in PBS-T (0.05% Tween-20), counter-stained with 10µg/mL Hoechst 33342 (Sigma-Aldrich, St. Louis, MO, USA), post-fixed in 4% formaldehyde/PBS-T and stored at 4˚C. Images were taken in a petri dish with agar under PBS-T, using a Nikon SMZ25 fluorescence stereomicroscope mounted with either a DS-Fi2 digital color camera (Nikon Elements software). Appendage mounts were imaged in 70% glycerol/PBS-T in a Zeiss LCM 710 confocal microscope.

Declarations

Ethics approval and consent to participate

Studies with arachnids do not required ethics approval or consent to participate.

Consent for publication

Not applicable.

Availability of Data and Materials

The data generated and analyzed in this study is included in the main article, supplementary information and available upon request. Raw reads are deposited in NCBI Sequence Read Archive under accession SRR12232018.

Competing interests

The authors declare having no conflict of interest.

Funding

GG was supported by a Wisconsin Alumni Research Foundation Fall Research Competition award. Materials for study were supported by an Oscar Franke student competition award to GG from the International Society of Arachnology. This material is based on work supported by the National Science Foundation under grant IOS-1552610 to PPS.

Author's contributions

Both authors designed the study, analyzed data, developed protocols and contributed to writing. GG conducted gene searches, orthology assessment, phylogenetic analysis, in situ

hybridization, assembled figures and drafted the manuscript. Both authors read and approved the final version of the manuscript.

Acknowledgments

This work is dedicated to Peter Weygoldt, whose pioneer work in the development of *Phrynus marginemaculatus* and contributions to the study of Amblypygi biology continue to inspire the authors of this paper. Comments from two anonymous reviewers improved an early draft of this manuscript. Sequencing was performed at the University of Wisconsin-Madison BioTechnology Center. Microscopy was performed at the Newcomb Imaging Center, Department of Botany, University of Wisconsin-Madison.

Author information

Affiliation

Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI, USA 53706

Guilherme Gainett; Prashant P. Sharma

Corresponding author

Correspondence to Guilherme Gainett (guilherme.gainett@wisc.edu)

References

1. Damen WGM, Hausdorf M, Seyfarth E-A, Tautz D. A conserved mode of head

segmentation in arthropods revealed by the expression pattern of Hox genes in a spider. PNAS. National Academy of Sciences; 1998;95:10665–70.

2. Stollewerk A, Schoppmeier M, Damen WGM. Involvement of *Notch* and *Delta* genes in spider segmentation. Nature. Nature Publishing Group; 2003;423:863–5.

3. Akiyama-Oda Y, Oda H. Early patterning of the spider embryo: a cluster of mesenchymal cells at the cumulus produces Dpp signals received by germ disc epithelial cells. Development. 2003;130:1735–47.

4. Akiyama-Oda Y, Oda H. Cell migration that orients the dorsoventral axis is coordinated with anteroposterior patterning mediated by Hedgehog signaling in the early spider embryo. Development. 2010;137:1263–73.

5. Schwager EE, Schoppmeier M, Pechmann M, Damen WG. Duplicated Hox genes in the spider *Cupiennius salei.* Front Zool. 2007;4:10–1.

6. Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, et al. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. BMC Biol. 2017;15:62.

7. Akiyama-Oda Y, Oda H. Axis specification in the spider embryo: *dpp* is required for radial-to-axial symmetry transformation and *sog* for ventral patterning. Development. 2006;133:2347–57.

8. Oda H, Akiyama-Oda Y. Differing strategies for forming the arthropod body plan: lessons from Dpp, Sog and Delta in the fly *Drosophila* and spider *Achaearanea*. Develop. Growth Differ. 2008;50:203–14.

9. Pechmann M, Benton MA, Kenny NJ, Posnien N, Roth S. A novel role for *Ets4* in axis specification and cell migration in the spider *Parasteatoda tepidariorum*. eLife. 2017;6.

10. Schwager EE, Pechmann M, Feitosa NM, McGregor AP, Damen WGM. *hunchback* functions as a segmentation gene in the spider *Achaearanea tepidariorum.* Curr. Biol. 2009;19:1333–40.

11. Pechmann M, Khadjeh S, Turetzek N, McGregor AP, Damen WGM, Prpic N-M. Novel function of *Distal-less* as a gap gene during spider segmentation. PLoS Genet. 2011;7:e1002342.

12. Setton EVW, Sharma PP. Cooption of an appendage-patterning gene cassette in the head segmentation of arachnids. PNAS. 2018;128:201720193–10.

13. Paese CLB, Schoenauer A, Leite DJ, Russell S, McGregor AP. A SoxB gene acts as an anterior gap gene and regulates posterior segment addition in a spider. eLife. 2018;7:1735.

14. Khadjeh S, Turetzek N, Pechmann M, Schwager EE, Wimmer EA, Damen WGM, et al. Divergent role of the Hox gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression. PNAS. National Academy of Sciences; 2012;109:4921–6.

15. Pechmann M, Schwager EE, Turetzek N, Prpic N-M. Regressive evolution of the

arthropod tritocerebral segment linked to functional divergence of the Hox gene *labial*. Proceedings of the Royal Society B: Biological Sciences. 2015;282:20151162–6.

16. Mittmann B, Scholtz G. *Distal-less* expression in embryos of *Limulus polyphemus* (Chelicerata, Xiphosura) and *Lepisma saccharina* (Insecta, Zygentoma) suggests a role in the development of mechanoreceptors, chemoreceptors, and the CNS. Dev. Genes Evol. 2001;211:232–43.

17. Blackburn DC, Conley KW, Plachetzki DC, Kempler K, Battelle B-A, Brown NL. Isolation and expression of *Pax6* and *atonal* homologues in the American horseshoe crab, *Limulus polyphemus*. Satir P, Fallon JF, editors. Dev. Dyn. 2008;237:2209–19.

18. Barnett AA, Thomas RH. Posterior Hox gene reduction in an arthropod: *Ultrabithorax* and *Abdominal-B* are expressed in a single segment in the mite *Archegozetes longisetosus*. EvoDevo. 2013;4:23.

19. Barnett AA, Thomas RH. The delineation of the fourth walking leg segment is temporally linked to posterior segmentation in the mite *Archegozetes longisetosus* (Acari: Oribatida, Trhypochthoniidae). Evolution & Development. 2012;14:383–92.

20. Barnett AA, Thomas RH. The expression of limb gap genes in the mite *Archegozetes longisetosus* reveals differential patterning mechanisms in chelicerates. Evolution & Development. 2013;15:280–92.

21. Telford MJ, Thomas RH. Expression of homeobox genes shows chelicerate arthropods retain their deutocerebral segment. PNAS. 1998;95:10671–5.

22. Khila A, Grbić M. Gene silencing in the spider mite *Tetranychus urticae*: dsRNA and siRNA parental silencing of the *Distal-less* gene. Dev. Genes Evol. 2007;217:241–51.

23. Grbić M, Khila A, Lee K-Z, Bjelica A, Grbić V, Whistlecraft J, et al. Mity model: *Tetranychus urticae*, a candidate for chelicerate model organism. Bioessays. 2007;29:489–96.

24. Dearden PK. Expression of pair-rule gene homologues in a chelicerate: early patterning of the two-spotted spider mite *Tetranychus urticae*. Development. 2002;129:5461–72.

25. Santos VT, Ribeiro L, Fraga A, de Barros CM, Campos E, Moraes J, et al. The embryogenesis of the Tick *Rhipicephalus* (*Boophilus*) *microplus*: The establishment of a new chelicerate model system. genesis. 2013;51:803–18.

26. Sharma PP, Schwager EE, Extavour CG, Wheeler WC. Hox gene duplications correlate with posterior heteronomy in scorpions. Proc. Biol. Sci. 2014;281:20140661–1.

27. Sharma PP, Schwager EE, Giribet G, Jockusch EL, Extavour CG. *Distal-less* and *dachshund* pattern both plesiomorphic and apomorphic structures in chelicerates: RNA interference in the harvestman *Phalangium opilio* (Opiliones). Evolution & Development. 2013;15:228–42.

28. Sharma PP, Schwager EE, Extavour CG, Giribet G. Evolution of the chelicera: a *dachshund* domain is retained in the deutocerebral appendage of Opiliones (Arthropoda, Chelicerata). Evolution & Development. 2012;14:522–33.

29. Sharma PP, Schwager EE, Extavour CG, Giribet G. Hox gene expression in the harvestman *Phalangium opilio* reveals divergent patterning of the chelicerate opisthosoma. Evolution & Development. 2012;14:450–63.

30. Sharma PP, Tarazona OA, Lopez DH, Schwager EE, Cohn MJ, Wheeler WC, et al. A conserved genetic mechanism specifies deutocerebral appendage identity in insects and arachnids. Proceedings of the Royal Society B: Biological Sciences. 2015;282:20150698–8.

31. Sharma PP, Kaluziak ST, Pérez-Porro AR, González VL, Hormiga G, Wheeler WC, et al. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. Mol. Biol. Evol. 2014;31:2963–84.

32. Leite DJ, Baudouin-Gonzalez L, Iwasaki-Yokozawa S, Lozano-Fernandez J, Turetzek N, Akiyama-Oda Y, et al. Homeobox gene duplication and divergence in arachnids. Mol. Biol. Evol. 2018;35:2240–53.

33. Sharma PP, Santiago MA, González-Santillán E, Monod L, Wheeler WC. Evidence of duplicated Hox genes in the most recent common ancestor of extant scorpions. Evolution & Development. 2015;17:347–55.

34. Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, et al. Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. Heredity. Nature Publishing Group; 2015;116:190–9.

35. Shingate P, Ravi V, Prasad A, Tay B-H, Garg KM, Chattopadhyay B, et al. Chromosome-level assembly of the horseshoe crab genome provides insights into its genome evolution. Nature Communications. Springer US; 2020;:1–13.

36. Schomburg C, Turetzek N, Prpic N-M. Candidate gene screen for potential interaction partners and regulatory targets of the Hox gene labial in the spider *Parasteatoda tepidariorum*. Development Genes and Evolution; 2020;:1–16.

37. Samadi L, Schmid A, Eriksson BJ. Differential expression of retinal determination genes in the principal and secondary eyes of *Cupiennius salei* Keyserling (1877). EvoDevo. BioMed Central; 2015;6:16–7.

38. Turetzek N, Pechmann M, Schomburg C, Schneider J, Prpic N-M. Neofunctionalization of a Duplicate *dachshund* Gene Underlies the Evolution of a Novel Leg Segment in Arachnids. Mol. Biol. Evol. 2015;33:109–21.

39. Turetzek N, Khadjeh S, Schomburg C, Prpic N-M. Rapid diversification of *homothorax* expression patterns after gene duplication in spiders. BMC Evolutionary Biology; 2017;:1–12.

40. Gainett G, Ballesteros JA, Kanzler CR, Zehms JT, Zern JM, Aharon S, et al. How spiders make their eyes: Systemic paralogy and function of retinal determination network homologs in arachnids. bioRxiv. Cold Spring Harbor Laboratory; 2020;129:1143–34.

41. Weygoldt P. Whip Spiders (Chelicerata: Amblypygi). Their Biology, Morphology and Systematics. Apollo Books; 2000.

42. de Miranda GS, Giupponi APL, Prendini L, Scharff N. Weygoldtia, a new genus of

Charinidae Quintero, 1986 (Arachnida, Amblypygi) with a reappraisal of the genera in the family. Zoologischer Anzeiger - A Journal of Comparative Zoology. Elsevier GmbH; 2018;273:23–32.

43. Giribet G. Current views on chelicerate phylogeny—A tribute to Peter Weygoldt. Zoologischer Anzeiger - A Journal of Comparative Zoology. Elsevier GmbH; 2018;273:7–13.

44. Ballesteros JA, Sharma PP. A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. Halanych K, editor. Syst Biol. 2019;33:440–22.

45. Lozano-Fernandez J, Tanner AR, Giacomelli M, Carton R, Vinther J, Edgecombe GD, et al. Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. Nature Communications. Springer US; 2019;:1–8.

46. Pereyaslawzewa S. Développement embryonnaire des phrynes. Annales des sciences naturelles, Zoologie. 1901;13:104–17.

47. Strubell A. Zur Entwicklungsgeschichte der Pedipalpen. Zoologischer Anzeiger. 1892;15:87–93.

48. Gough LH. The development of *Admetus pumilio* Koch; a contribution to the embryology of the Pedipalps. Quarterly Journal of Microscopical Science New Series. 1902;45:595–630.

49. Weygoldt P. Untersuchungen zur Embryologie und Morphologie der Geißelspinne *Tarantula marginemaculata* C. L. Koch (Arachnida, Amblypygi, Tarantulidae). Zoomorphologie. 1975;82:137–99.

50. Chapin KJ, Hebets EA. The behavioral ecology of amblypygids. Journal of Arachnology. American Arachnological Society; 2016;44:1–14.

51. Wiegmann DD, Hebets EA, Gronenberg W, Graving JM, Bingman VP. Amblypygids: Model Organisms for the Study of Arthropod Navigation Mechanisms in Complex Environments? Front. Behav. Neurosci. 2016;10:1–8.

52. Seiter M, Lemell P, Gredler R, Wolff JO. Strike kinematics in the whip spider *Charon* sp. (Amblypygi: Charontidae). The Journal of Arachnology. 2019;47:260–7.

53. Igelmund P. Morphology, sense organs, and regeneration of the forelegs (whips) of the whip spider *Heterophrynus elaphus* (Arachnida, Amblypygi). Journal of Morphology. Wiley-Blackwell; 1987;193:75–89.

54. Foelix RF, Chu-Wang IW, Beck L. Fine structure of tarsal sensory organs in the whip spider *Admetus pumilio* (Amblypygi, Arachnida). Tissue and Cell. 1975;7:331–46.

55. Santer RD, Hebets EA. The Sensory and Behavioural Biology of Whip Spiders (Arachnida, Amblypygi). 1st ed. Spider Physiology and Behaviour. Elsevier Ltd; 2011. pp. 1–64.

56. Foelix R, Troyer D, Igelmund P. Peripheral synapses and giant neurons in whip spiders. Microsc. Res. Tech. 2002;58:272–82.

57. Foelix RF. Occurrence of synapses in peripheral sensory nerves of arachnids. Nature. 1975;254:146–8.

58. Santer RD, Hebets EA. Agonistic signals received by an arthropod filiform hair allude to the prevalence of near-field sound communication. Proceedings of the Royal Society B: Biological Sciences. 2008;275:363–8.

59. Santer RD, Hebets EA. Evidence for Air Movement Signals in the Agonistic Behaviour of a Nocturnal Arachnid (Order Amblypygi). Gronenberg W, editor. PLoS ONE. 2011;6:e22473–6.

60. Dong PDS, Dicks JS, Panganiban G. *Distal-less* and *homothorax* regulate multiple targets to pattern the *Drosophila* antenna. Development. 2002;129:1967–74.

61. Dong PD, Chu J, Panganiban G. Proximodistal domain specification and interactions in developing *Drosophila* appendages. Development. 2001;128:2365–72.

62. Casares F, Mann RS. Control of antennal versus leg development in *Drosophila*. Nature. Nature Publishing Group; 1998;392:723–6.

63. Duncan DM, Burgess EA, Duncan I. Control of distal antennal identity and tarsal development in *Drosophila* by *spineless-aristapedia*, a homolog of the mammalian dioxin receptor. Genes Dev. Cold Spring Harbor Lab; 1998;12:1290–303.

64. Setton EVW, March LE, Nolan ED, Jones TE, Cho H, Wheeler WC, et al. Expression and function of *spineless* orthologs correlate with distal deutocerebral appendage morphology across Arthropoda. Developmental Biology. Elsevier Inc; 2017;430:224–36.

65. Struhl G. A homoeotic mutation transforming leg to antenna in *Drosophila*. Nature. 1981;292:635–8.

66. Struhl G. Genes controlling segmental specification in the *Drosophila* thorax. PNAS. National Academy of Sciences; 1982;79:7380–4.

67. Prpic N-M, Damen WM. Expression patterns of leg genes in the mouthparts of the spider *Cupiennius salei* (Chelicerata: Arachnida). Dev. Genes Evol. Springer-Verlag; 2004;214:1–7.

68. Prpic N-M, Janssen R, Wigand B, Klingler M, Damen WGM. Gene expression in spider appendages reveals reversal of *exd*/*hth* spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling. Developmental Biology. 2003;264:119–40.

69. Pechmann M, Prpic N-M. Appendage patterning in the South American bird spider *Acanthoscurria geniculata* (Araneae: Mygalomorphae). Dev. Genes Evol. 2nd ed. Springer-Verlag; 2009;219:189–98.

70. Garwood RJ, Dunlop JA, Knecht BJ, Hegna TA. The phylogeny of fossil whip spiders. BMC Evolutionary Biology; 2017;:1–14.

71. Weygoldt P. Beobachtungen zur fortpflanzungsbiologie und zum verhalten der geißelspinne *Tarantula marginemaculata* C. L. Koch (Chelicerata, Amblypygi). Z. Morph. Tiere. Springer-Verlag; 1969;64:338–60.

72. Weygoldt P. Lebenszyklus und postembryonale entwicklung der geibelspinne *Tarantula marginemaculata* C. L. Koch (Chelicerata, Amblypygi) im laboratorium. Z. Morph. Tiere. Springer-Verlag; 1970;67:58–85.

73. Schoppmeier M, Damen WGM. Double-stranded RNA interference in the spider *Cupiennius salei*: the role of *Distal-less* is evolutionarily conserved in arthropod appendage formation. Dev. Genes Evol. 2001;211:76–82.

74. Thomas RH, Telford MJ. Appendage development in embryos of the oribatid mite *Archegozetes longisetosus* (Acari, Oribatei, Trhypochthoniidae). Acta Zoologica. John Wiley & Sons, Ltd; 1999;80:193–200.

75. Nolan ED, Santibáñez López CE, Sharma PP. Developmental gene expression as a phylogenetic data class: support for the monophyly of Arachnopulmonata. Dev. Genes Evol. Springer Berlin Heidelberg; 2020;230:137–53.

76. Wagner GP, Amemiya C, Ruddle F. Hox cluster duplications and the opportunity for evolutionary novelties. PNAS. National Academy of Sciences; 2003;100:14603–6.

77. Hughes CL, Kaufman TC. Hox genes and the evolution of the arthropod body plan. Evolution & Development. 2002;4:459–99.

78. Ronshaugen M, McGinnis N, McGinnis W. Hox protein mutation and macroevolution of the insect body plan. Nature. Nature Publishing Group; 2002;415:914–7.

79. Kaufman CLHATC. RNAi analysis of *Deformed*, *proboscipedia* and *Sex combs reduced* in the milkweed bug *Oncopeltus fasciatus*: novel roles for Hox genes in the Hemipteran head. 2000;:1–12.

80. Pace RM, Grbić M, Nagy LM. Composition and genomic organization of arthropod Hox clusters. EvoDevo. BioMed Central; 2016;7:11–1.

81. Nossa CW, Havlak P, Yue J-X, Lv J, Vincent KY, Brockmann HJ, et al. Joint assembly and genetic mapping of the Atlantic horseshoe crab genome reveals ancient whole genome duplication. GigaSci. 2014;3:708–21.

82. Zhou Y, Liang Y, Yan Q, Zhang L, Chen D, Ruan L, et al. The draft genome of horseshoe crab *Tachypleus tridentatus* reveals its evolutionary scenario and well-developed innate immunity. BMC Genomics. 4 ed. BioMed Central; 2020;21:137–15.

83. Harper A, Baudouin-Gonzalez L, Schönauer A, Seiter M, Holzem M, Arif S, et al. Widespread retention of ohnologs in key developmental gene families following whole genome duplication in arachnopulmonates. bioRxiv. Cold Spring Harbor Laboratory; 2020;160:1067–40.

84. Abzhanov A, Kaufman TC. Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. Developmental Biology. 2000;227:673–89.

85. Beck L, Foelix R, Gödeke E, Kaiser R. Morphology, larval development, and hair sensilla of the antenniform legs of the whip spider *Heterophrynus longicornis* Butler (Arach., Amblypygi). Zoomorphologie. Springer-Verlag; 1977;88:259–76.

86. Rayor LS, Taylor LA. Social behavior in amblypygids, and a reassessment of arachnid social patterns. Journal of Arachnology. 2006;34:399–421.

87. Santer RD, Hebets EA. Tactile learning by a whip spider, *Phrynus marginemaculatus* C.L. Koch (Arachnida, Amblypygi). J Comp Physiol A. 4 ed. 2009;195:393–9.

88. Fowler-Finn KD, Hebets EA. An Examination of agonistic interactions in the whip spider *Phrynus marginemaculatus* (Arachnida, Amblypygi). Journal of Arachnology. 2006;34:62–76.

89. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. Nature Publishing Group; 2011;29:644–52.

90. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

91. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO Applications from quality assessments to gene prediction and phylogenomics. Mol. Biol. Evol. 2017;35:543–8.

92. Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, et al. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. CellReports. 2017;18:762–76.

93. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research. Oxford University Press; 2018;47:D506–15.

94. Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. Nature Communications. Nature Publishing Group; 2016;7:10507–13.

95. Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, Grbić V, et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature. Nature Publishing Group; 2011;479:487–92.

96. Herndon N, Shelton J, Gerischer L, Ioannidis P, Ninova M, Dönitz J, et al. Enhanced genome assembly and a new official gene set for *Tribolium castaneum*. BMC Genomics; 2020;:1–13.

97. Chipman AD, Ferrier DEK, Brena C, Qu J, Hughes DST, Schröder R, et al. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. Tyler-Smith C, editor. PLoS Biol. Public Library of Science; 2014;12:e1002005–24.

98. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. Nature Publishing Group; 2011;7:1–6.

99. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Mol. Biol. Evol. 2010;27:221–4.

100. Nguyen L-T, Schmidt HA, Haeseler von A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 2015;32:268–74.

101. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007;23:1289–91.

102. Nong W, Qu Z, Li Y, Barton-Owen T, Wong AYP, Yip HY, et al. Horseshoe crab genomes reveal the evolutionary fates of genes and microRNAs after three rounds (3R) of whole genome duplication. bioRxiv. Cold Spring Harbor Laboratory; 2020;6:2020.04.16.045815.

Figure legends

Figure 1: *Phrynus marginemaculatus* as a model to study Amblypygi development. A: adult male *P. marginemaculatus*, dorsal view. B: adult female P. *marginemaculatus*, carrying praenymphs on the opisthosoma. The yellow ink is used to mark the females in the colony.

Figure 2: Overview of the embryonic development of *Phrynus marginemaculatus*. Z-stack automontage pictures of embryos imaged with nuclear staining (Hoechst) fluorescence. A–E: five embryonic stages, in ventral view, spanning the elongation of appendages, and showing germ band inversion and development of the ventral midline. F–H: head development, frontal view. F: undifferentiated head lobes. G: head lobe with neuron precursor cell invaginations. H: head lobes with anterior rim folding (semi-lunar grooves). I–K: opisthosomal development, posterior view. I: embryo with one opisthosomal segment. J: embryo with five opisthosomal segments. K: embryo ventral curved opisthosoma and most opisthosomal segments formed (number unclear). L–N: lateral organ formation and opisthosomal curving, lateral view. L: lateral organ as an undifferentiated bud proximal in leg II. M: lateral organ with a discreet “C” shape, and opisthosomal segments curved up. N: lateral organ in its final “kidney” shape, and opisthosoma curved ventrally. Insets in L–N show the lateral organ (Lo),

which is outlined with dashed lines in the respective main panel. Dashed lines on the posterior regions of the embryos in L–N mark the curvature of the opisthosoma. O–Q: deutembryo, in frontal, ventral and lateral view respectively. Ch: chelicera; Gz: growth zone; H: head lobe; Iv: neural precursor cell invagination; St: stomodeum. Scale bars: 500 µm.

Figure 3: Hox gene duplications in Amblypygi. A: Schematic representation of the number of copies of each of the ten Hox genes in the selected arthropod terminals. Dotted squares in the amblypygids *Phrynus marginemaculatus*, *Charinus israelensis*, and *C. ioanniticus* indicate absence from the transcriptome. Empty squares in *Parasteatoda* and *Tetranychus* indicate absence from the genome. Number of Hox genes in *Carcinoscorpius rotundicauda* after [35] (but see [102] for evidence of additional copies). B: Tree topology inferred from maximum likelihood analysis of a conserved region (71 amino acid characters) using the same terminals of the schematics (ln L = -3442.810). Numbers on the notes are ultrafast bootstrap resampling frequencies (only >80 shown). Species: *Phrynus marginemaculatus* (Pmar); *Charinus israelensis* (Cisr); *C. ioanniticus* (Cioa); *Parasteatoda tepidariorum* (Ptep); *Centruroides sculpturatus* (Cscu); *Ixodes scapularis* (Isca); *Tetranychus urticae* (Ttur); *Strigamia maritima* (Smar); *Drosophila melanogaster* (Dmel); *Tribolium castaneum* (Tcas). *lab*: *labial*; *pb*: *proboscipedia*/*maxillopedia*; *Hox3*: *Hox3*/*zerknullt*/*z2*; *Dfd*: *Deformed*; *Scr*: *Sex combs reduced*; *ftz*: *fushi tarazu*; *Antp*: *Antennapedia*/*prothorax-less*; *Ubx*: *Ultrabithorax*; *abdA*: *abdominal A*; *AbdB*: *Abdominal B.*

Figure 4: *Pmar-Distal-less* (*Pmar-Dll*) colorimetric in situ hybridization. Whole mount bright field images (Z-stack automontage) overlaid with nuclear staining (Hoechst). Ch: chelicera;

Pp: pedipalp; L1–4: leg 1–4; A/P: anterior/posterior; dAEL: days after egg laying. Scale bars: 500 µm.

Figure 5: *Pmar-Distal-less* (*Pmar-Dll*) colorimetric in situ hybridization, appendage dissections in lateral view. Distal is to the left. A–F, G–L: bright field. A'–F', G'-L': nuclear staining (Hoechst), confocal Z-stack maximum projection. A–F: 16–17 dAEL embryo. G–L: 17-18 dAEL embryo. J and J' have been vertically reflected to adjust orientation. Distal is to the left. Arrowheads in A–F indicate proximal boundary of expression. Arrowheads in G–L indicate stronger domains of expression. Ch: chelicera; Pp: pedipalp; L1–4: leg 1–4. Scale bar: 250µm.

Figure 6: *Pmar-dachshund 1* (*Pmar-dac1*) (A–F) and *Pmar-dachshund 2* (*Pmar-dac2*) G–I) colorimetric in situ hybridization. Whole mount bright field images (Z-stack automontage) overlaid with nuclear staining (Hoechst). Ch: chelicera; Pp: pedipalp; L1–4: leg 1–4; Lo: lateral organ; A/P: anterior/posterior; arrowhead: expression domain; dAEL: days after egg laying. Scale bars: 500 µm.

Figure 7: *Pmar-dachshund 1* (*Pmar-dac1*) (A–H) and *Pmar-dachshund 2* (*Pmar-dac2*) (I–L) colorimetric in situ hybridization, appendage dissections in lateral view. Distal is to the left. White arrowhead indicates discreet domains of expression. Asterisk in "E" marks expression from the lateral margin of the headlobe. Ch: chelicera; Pp: pedipalp; L1, 4: leg 1, 4. dAEL: days after egg laying. Scale bar: 200 µm.

Figure 8: *Pmar-homothorax 1* (*Pmar-hth1*) (A–A') and *Pmar-homothorax 2* (*Pmar-hth2*) (B–B') colorimetric in situ hybridization, ventral view. A, B: whole mount bright field images

(Z-stack automontage). A’, B’: nuclear staining (Hoechst) (Z-stack automontage). Ch: chelicera; Pp: pedipalp; L1–4: leg 1–4. A/P: anterior/posterior; arrowhead: expression domain; dAEL: days after egg laying. Scale bar: 500 µm.

Figure 9: *Pmar-homothorax 1* (*Pmar-hth1*) (A–H) and *Pmar-homothorax 2* (*Pmar-hth2*) (I–L) colorimetric in situ hybridization, appendage dissections in lateral view. Distal is to the left. White arrowhead indicates discreet domains of expression. Dotted line in “G” marks expanded distal expression domain. Ch: chelicera; Pp: pedipalp; L1, 4: leg 1, 4. dAEL: days after egg laying. Scale bar: 200 µm.

Figure 10: Tool for removing the brood sac from whip spiders. A: dissecting dish made of plastic petri dish and foam. B: detail of an immobilized female ready to have the brood sac dissected. C: intact brood sac dissected from the female. Scale bars: A: 1 cm (approximate); B: 5 mm (approximate); C: 1 mm.

Additional file 1, Figure S1: Tree topology inferred from maximum likelihood analysis of a conserved region (75 amino acid characters) (ln L = -4780.167) using the same terminals of Figure 3 and including *Parasteatoda tepidariorum Hox3-A* paralog and *Drosophila melanogaster bicoid* (*Hox3*) (red arrow). Numbers on the notes are ultrafast bootstrap resampling frequencies (only >80 shown). For abbreviations, see Additional file 5, Table S1. File format: jpeg.

Additional file 2, Figure S2: Tree topology of *Distal-less* (*Dll*) inferred from maximum likelihood analysis amino acid sequences (ln L = -15272.780). Numbers on the notes are

ultrafast bootstrap resampling frequencies (only >80 shown). Accession numbers are available in Additional file 5, Table S1. The terminal for the whip spider *Phrynus marginemaculatus Dll* ortholog is in boldface. File format: jpeg.

Additional file 3, Figure S3: Tree topology of *dachshund* (*dac*) inferred from maximum likelihood analysis amino acid sequences (ln L = -21699.907). Numbers on the notes are ultrafast bootstrap resampling frequencies (only >80 shown). The terminals for the whip spider *Phrynus marginemaculatus Dll* orthologs are in boldface. The original protein alignment of all terminals (except *P. marginemaculatus*) is from Noland et al. 2020. File format: jpeg.

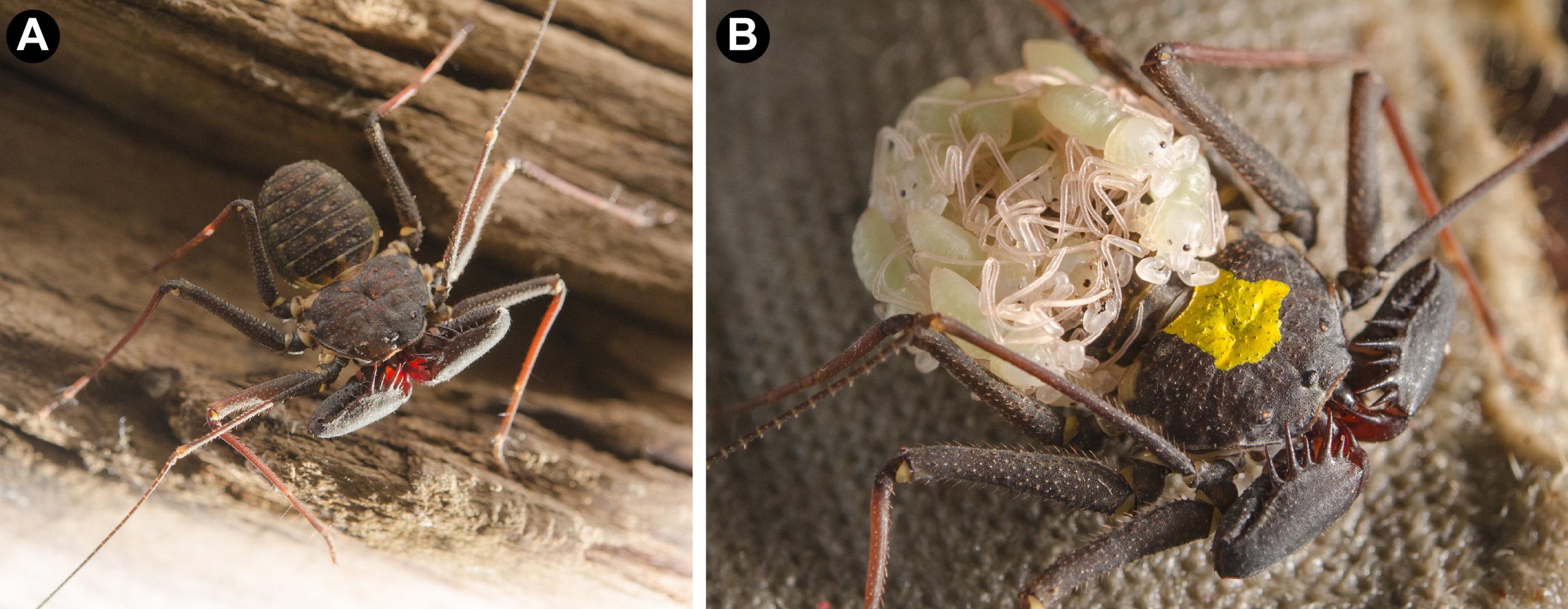
Additional file 4, Figure S4: Tree topology of *homothorax* (*hth*) inferred from maximum likelihood analysis amino acid sequences (ln L = -10459.647). Numbers on the notes are ultrafast bootstrap resampling frequencies (only >80 shown). The terminals for the whip spider *Phrynus marginemaculatus hth* orthologs are in boldface. The original protein alignment of all terminals (except *P. marginemaculatus*) is from Noland et al. 2020. File format: jpeg.

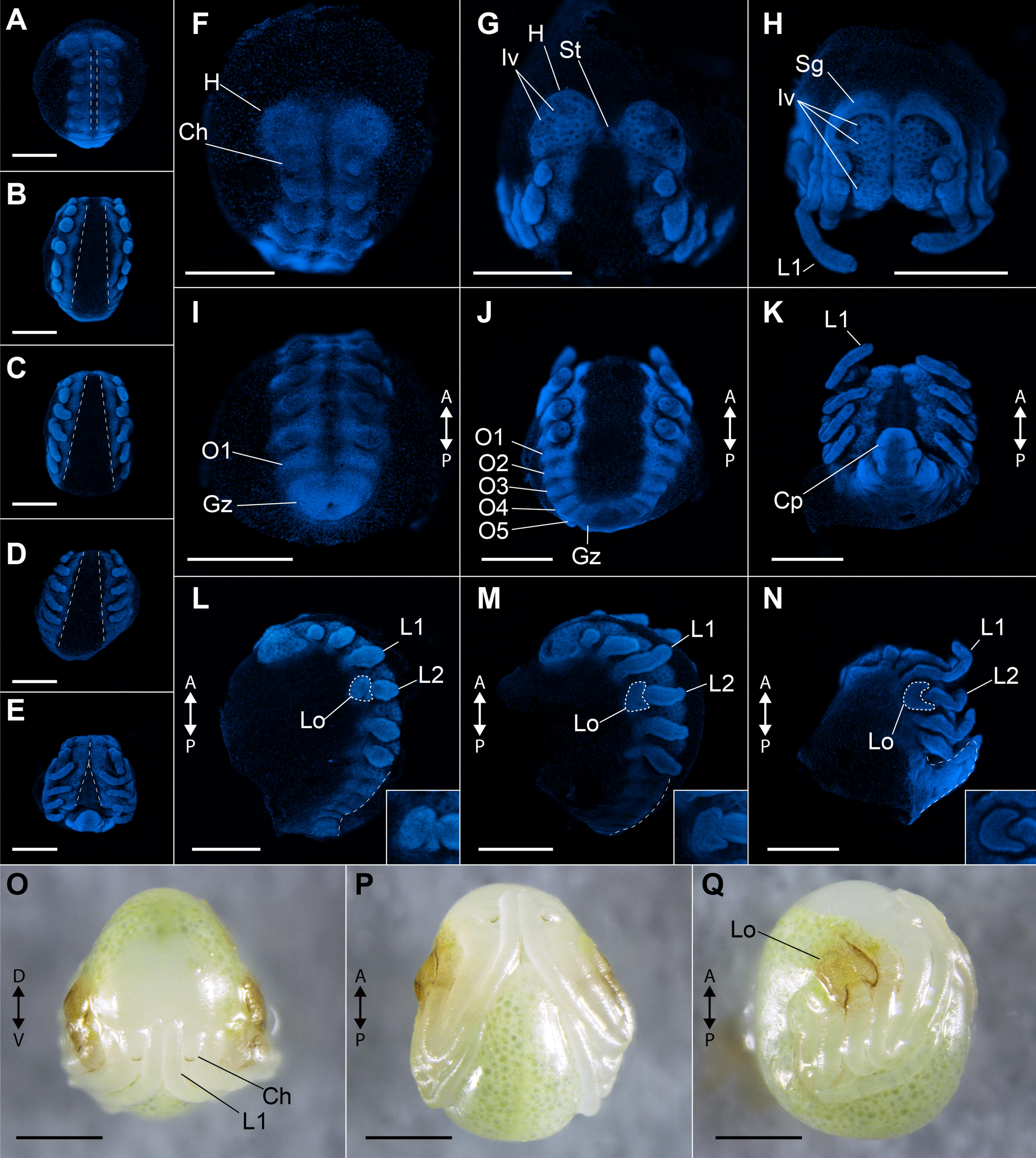
Additional file 5, Table S1: List of species, accession numbers and abbreviations of terminals used in the phylogenetic analysis of Hox genes, *Distal-less*, *dachshund* and *homothorax* homologs. File format: xlsx

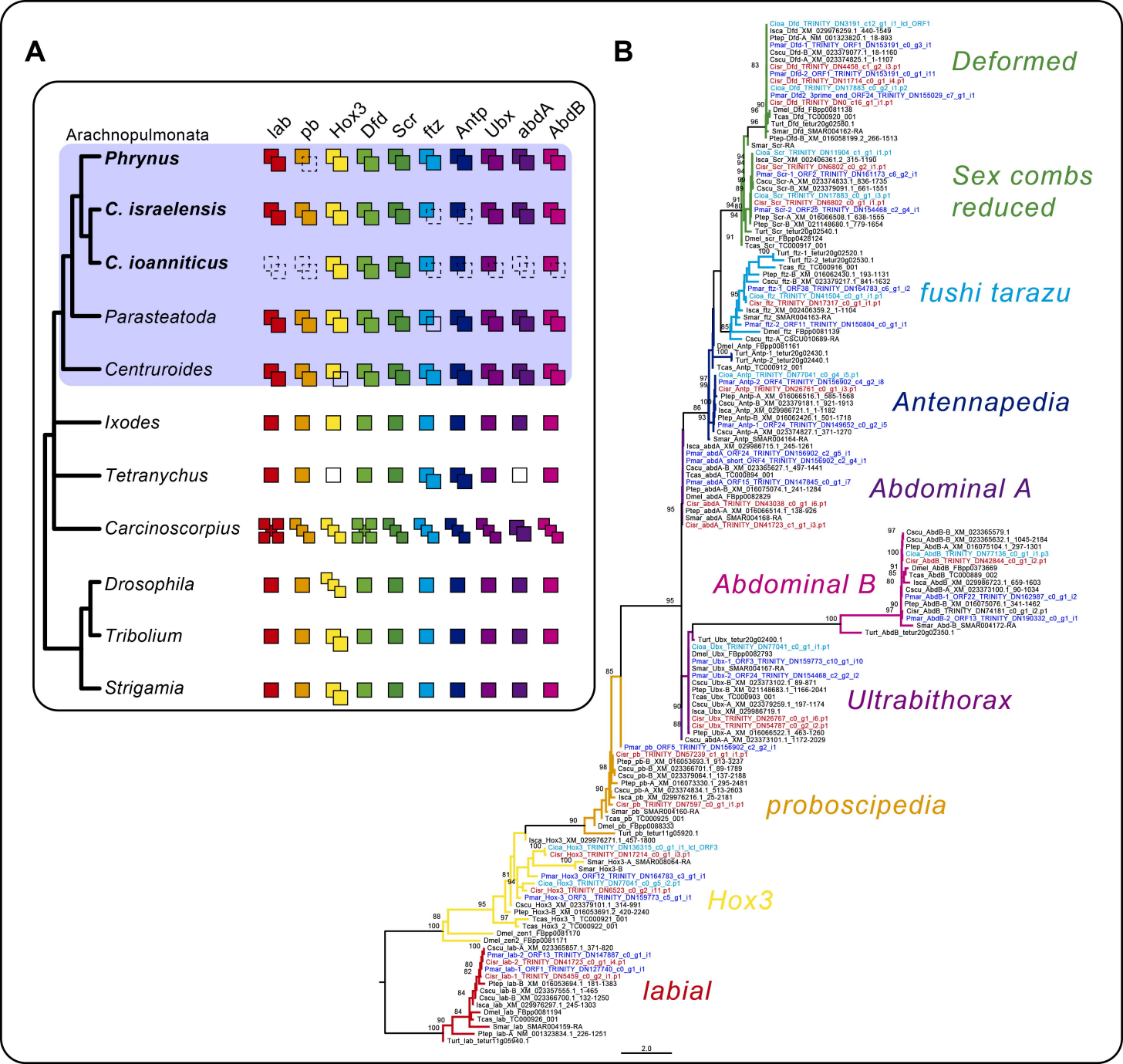
Additional file 6: Protein alignment used in the phylogenetic analysis depicted in Figure 3. File format: fasta.

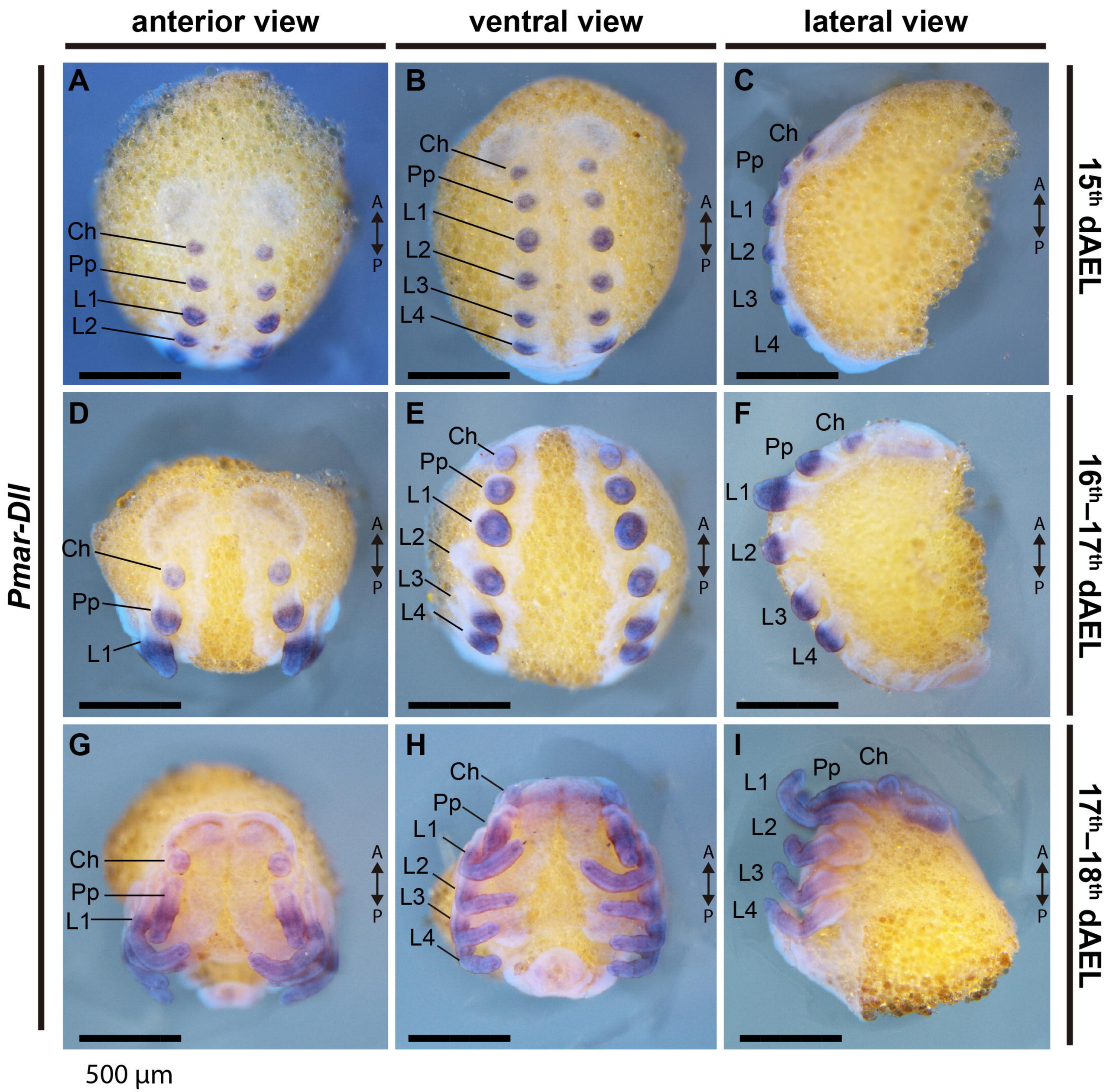
Additional file 7: Protein alignment used in the phylogenetic analysis depicted in Additional file 1, Figure S1. File format: fasta.

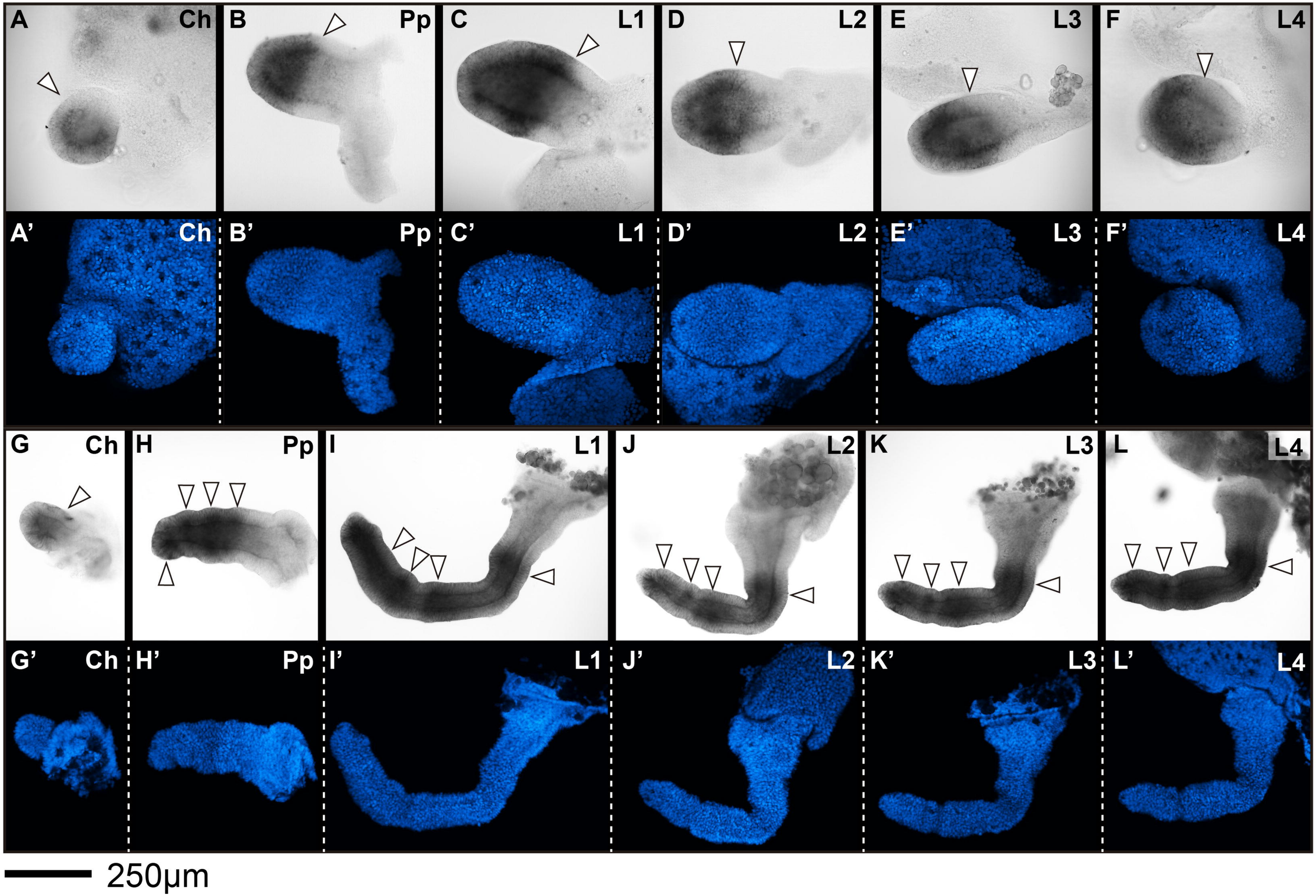
Additional file 8, Table S2: List of primers for *Pmar-Dll*, *Pmar-dac1/2*, *Pmar hth1/2*, and cloned fragments for *Pmar-Dll*. File format: xlsx

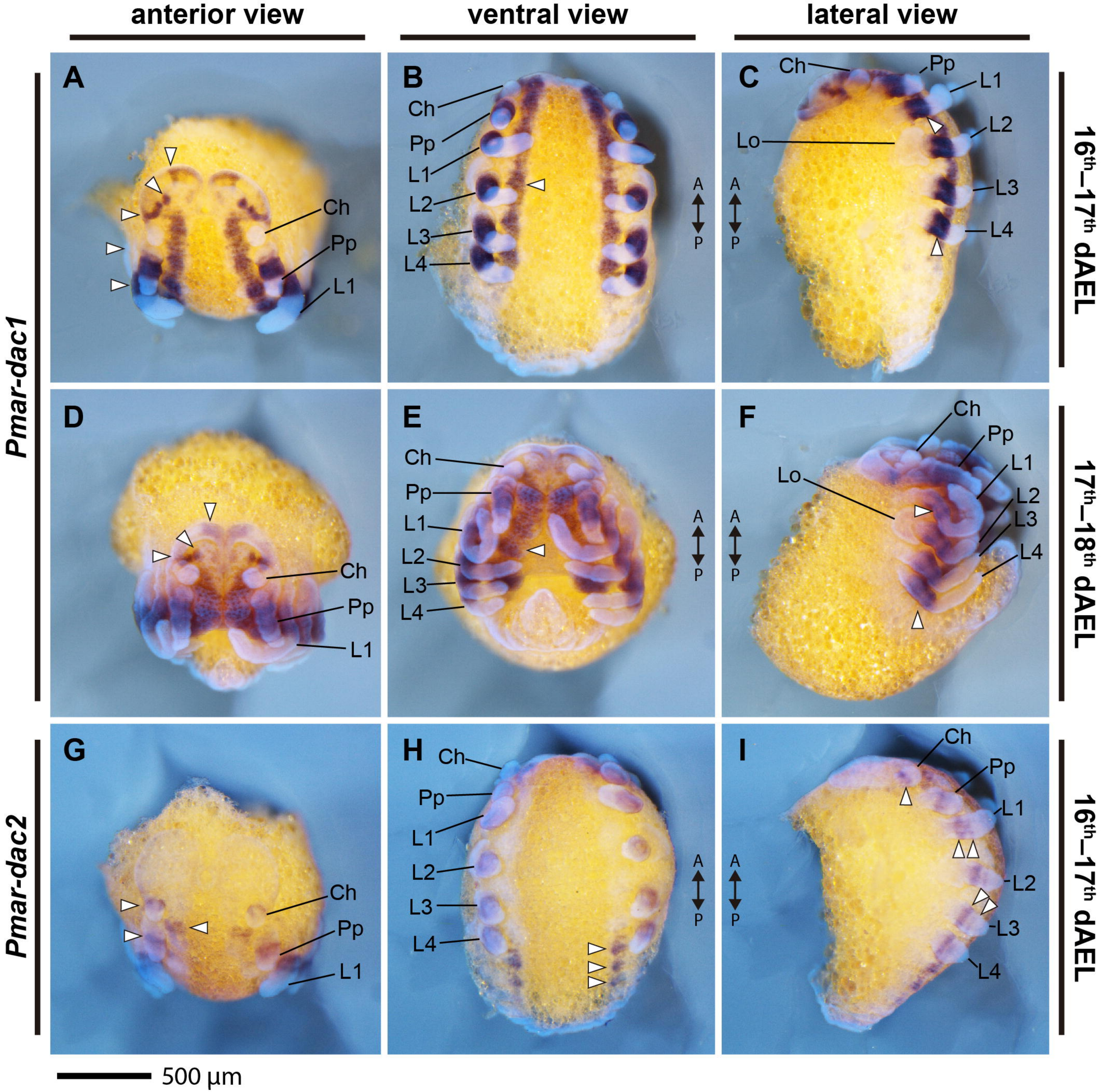

A
B

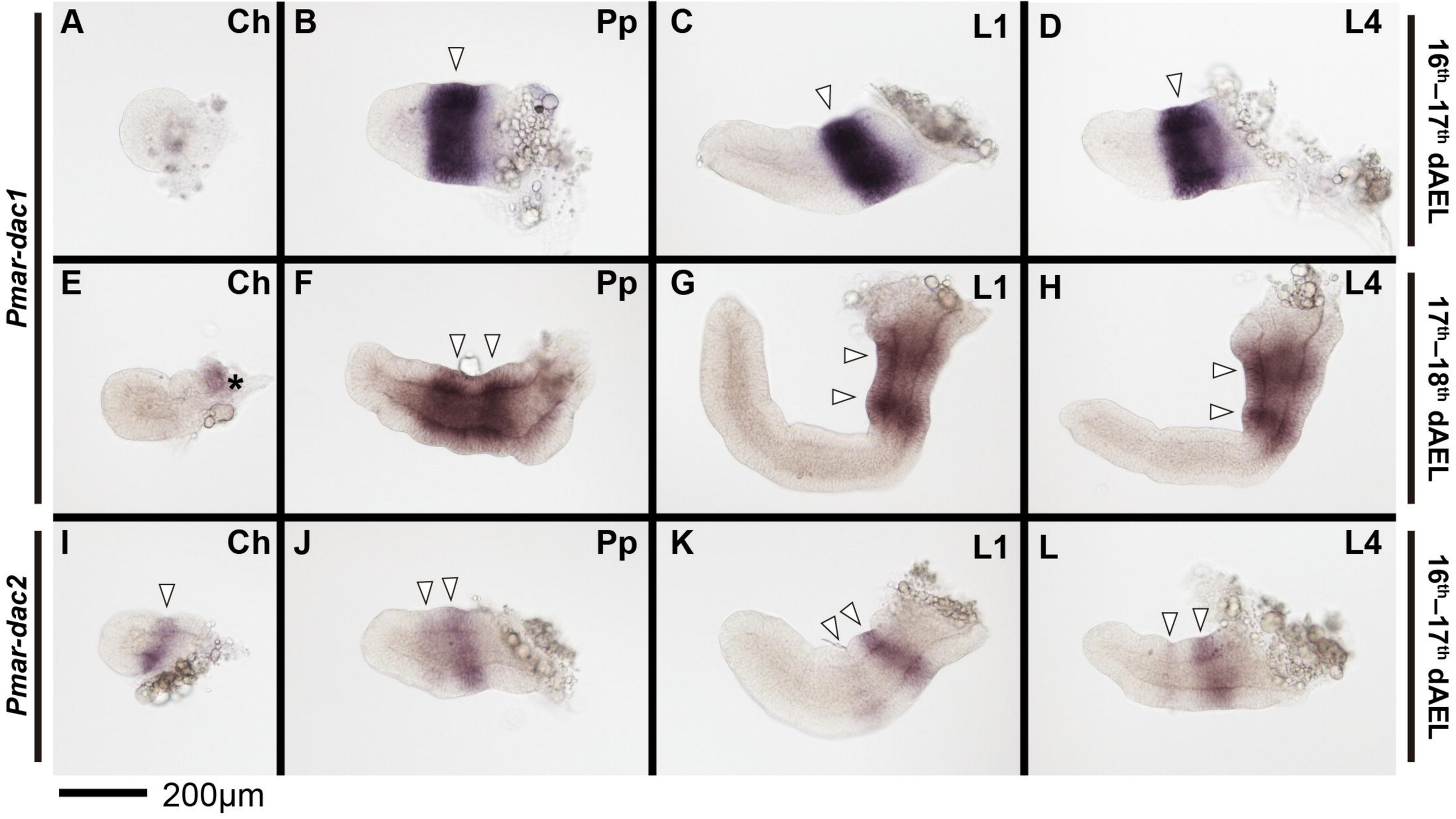

A
B
C
D
E
F
H
Ch
G
H
Iv
St
H
Sg
Iv
L1
I
O1
Gz
A
P
J
O1
O2
O3
O4
O5
Gz
A
P
K
L1
Cp
A
P
L
L1
L2
Lo
A
P
M
L1
L2
Lo
A
P
N
L1
L2
Lo
A
P
O
D
V
Ch
L1
P
A
P
Q
Lo
A
P

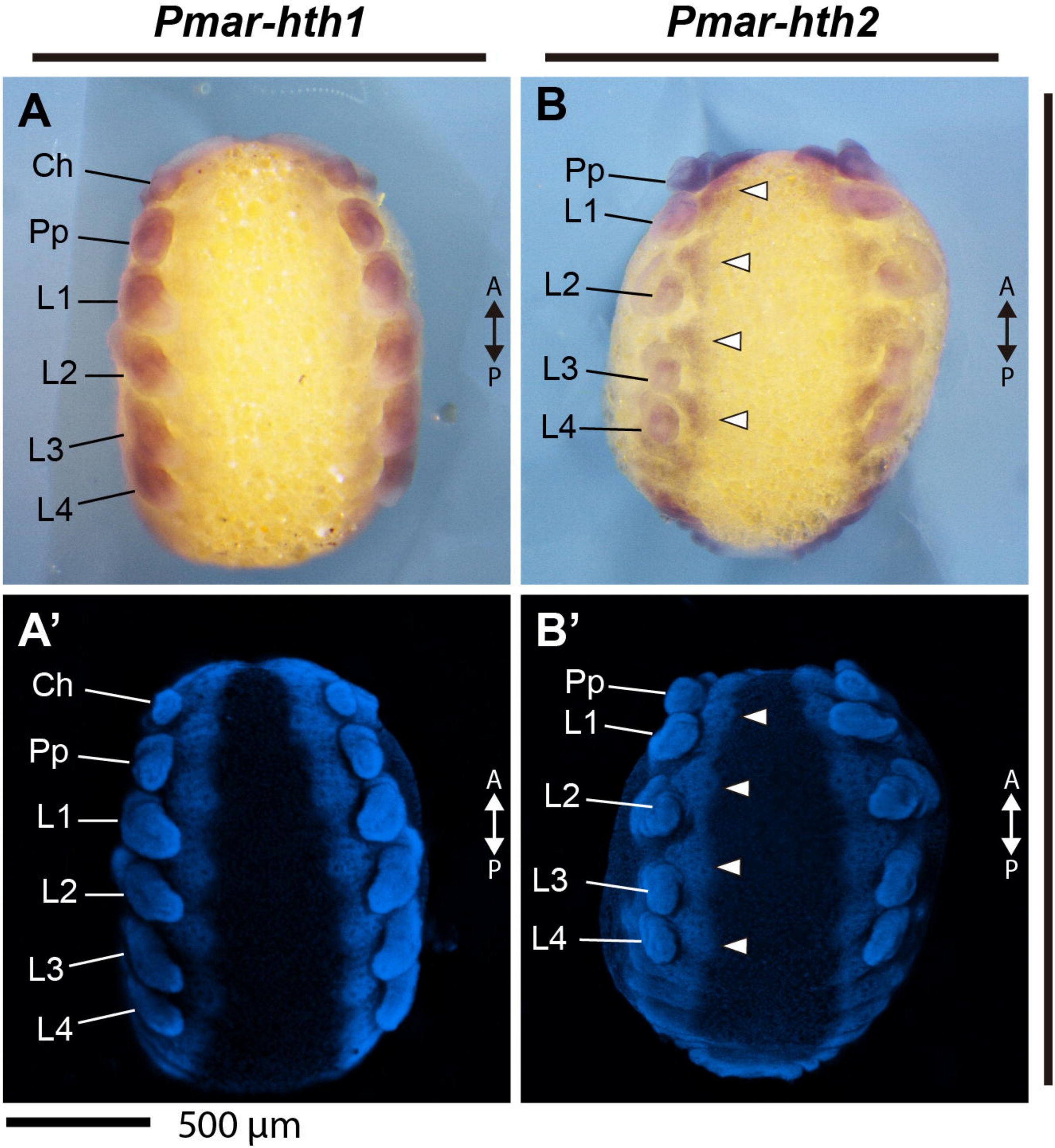

A
B
Arachnopulmonata
lab
pb
Hox3
Dfd
Scr
ftz
Antp
Ubx
abdA
AbdB
Phrynus
C. israelensis
C. ioanniticus
Parasteatoda
Centruroides
Ixodes
Tetranychus
Carcinoscorpius
Drosophila
Tribolium
Strigamia
Deformed
Sex combs reduced
fushi tarazu
Antennapedia
Abdominal A
Abdominal B
Ultrabithorax
proboscipedia
Hox3
labial
2.0

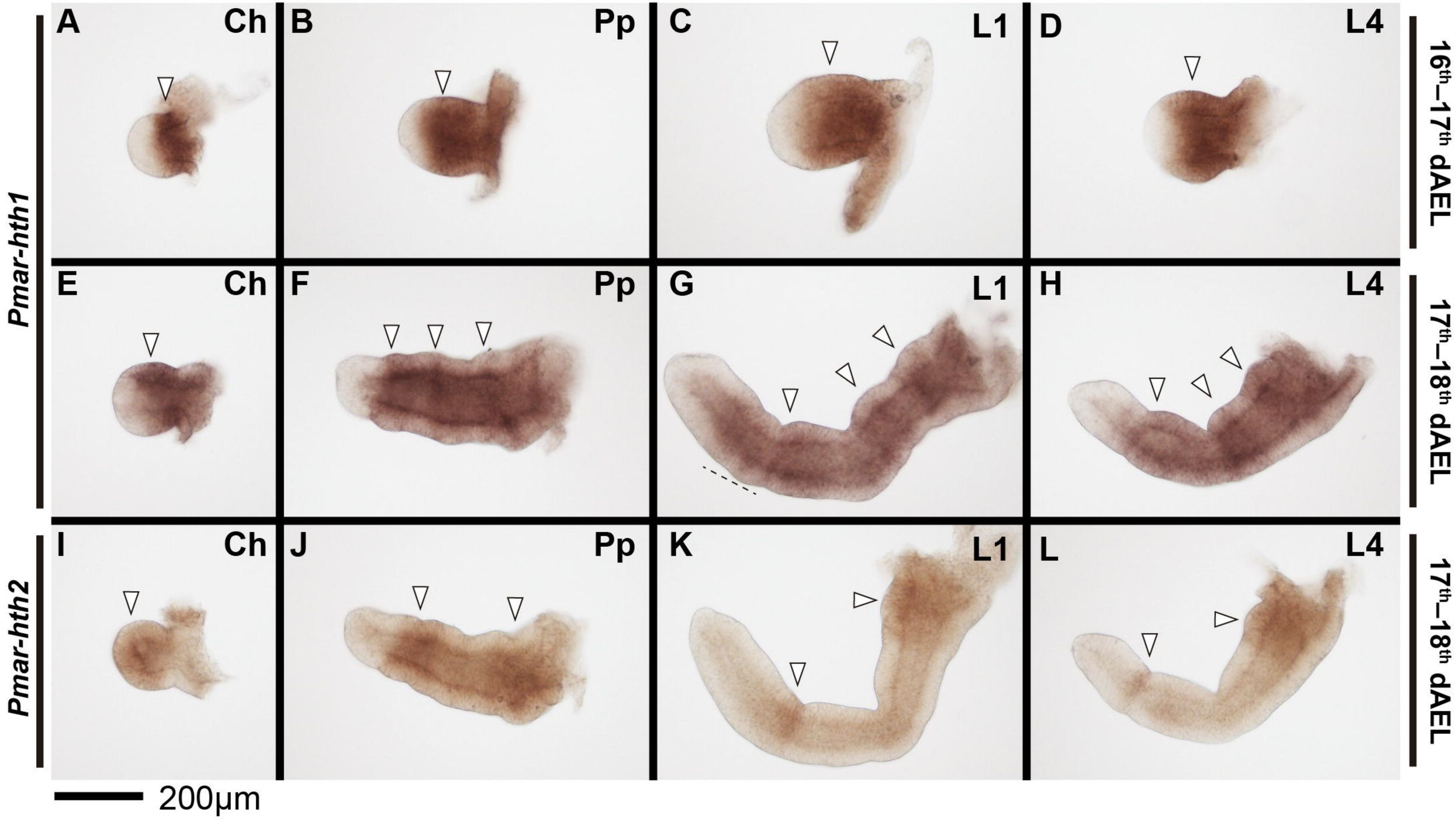

anterior view
ventral view
lateral view
Pmar-DII
15th dAEL
16th–17th dAEL
17th–18th dAEL
A
B
C
D
E
F
G
H
I
Ch
Pp
L1
L2
L3
L4
A
P
500 µm

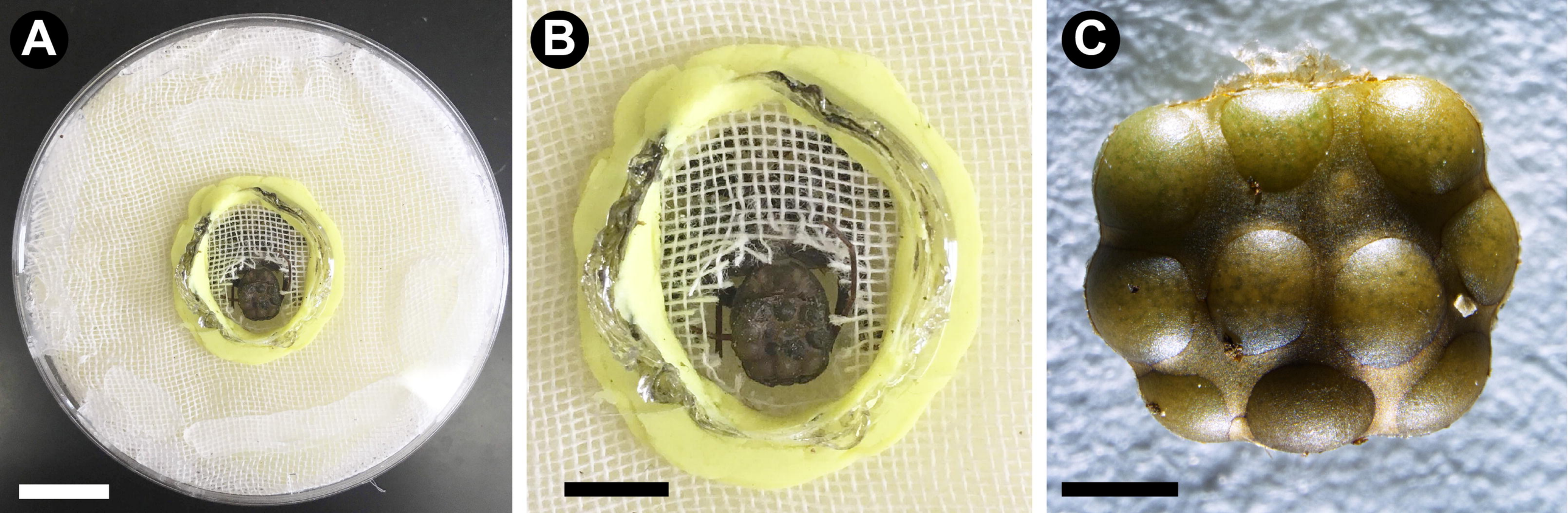

A
Ch
B
Pp
C
L1
D
L2
E
L3
F
L4
A'
Ch
B'
Pp
C'
L1
D'
L2
E'
L3
F'
L4
G
Ch
H
Pp
I
L1
J
L2
K
L3
L
L4
G'
Ch
H'
Pp
I'
L1
J'
L2
K'
L3
L'
L4
250µm

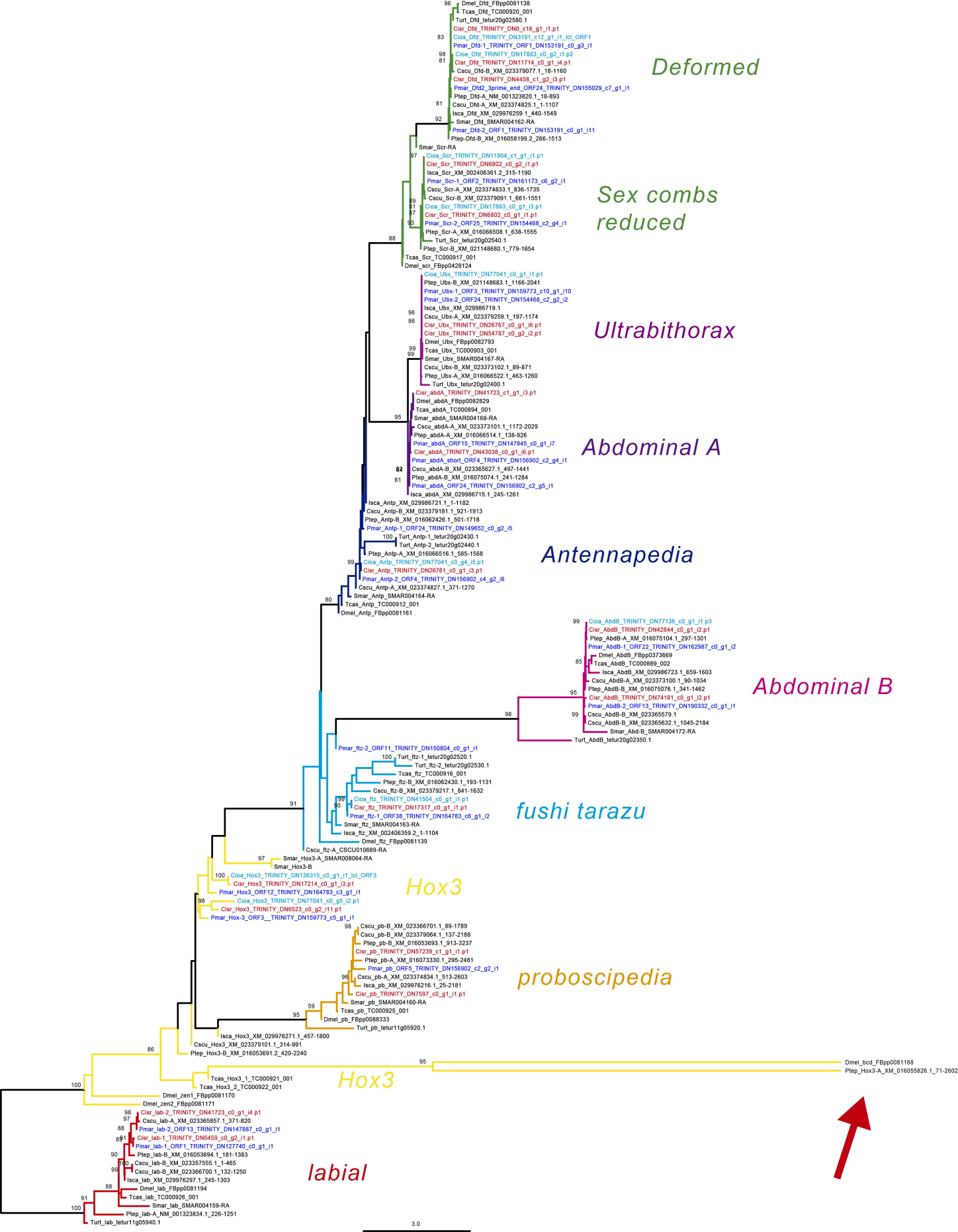

anterior view
ventral view
lateral view
A
B
C
D
E
F
G
H
I
Ch
Pp
L1
L2
L3
L4
Lo
A
P
Pmar-dac1
Pmar-dac2
16th–17th dAEL
17th–18th dAEL
16th–17th dAEL
500 µm

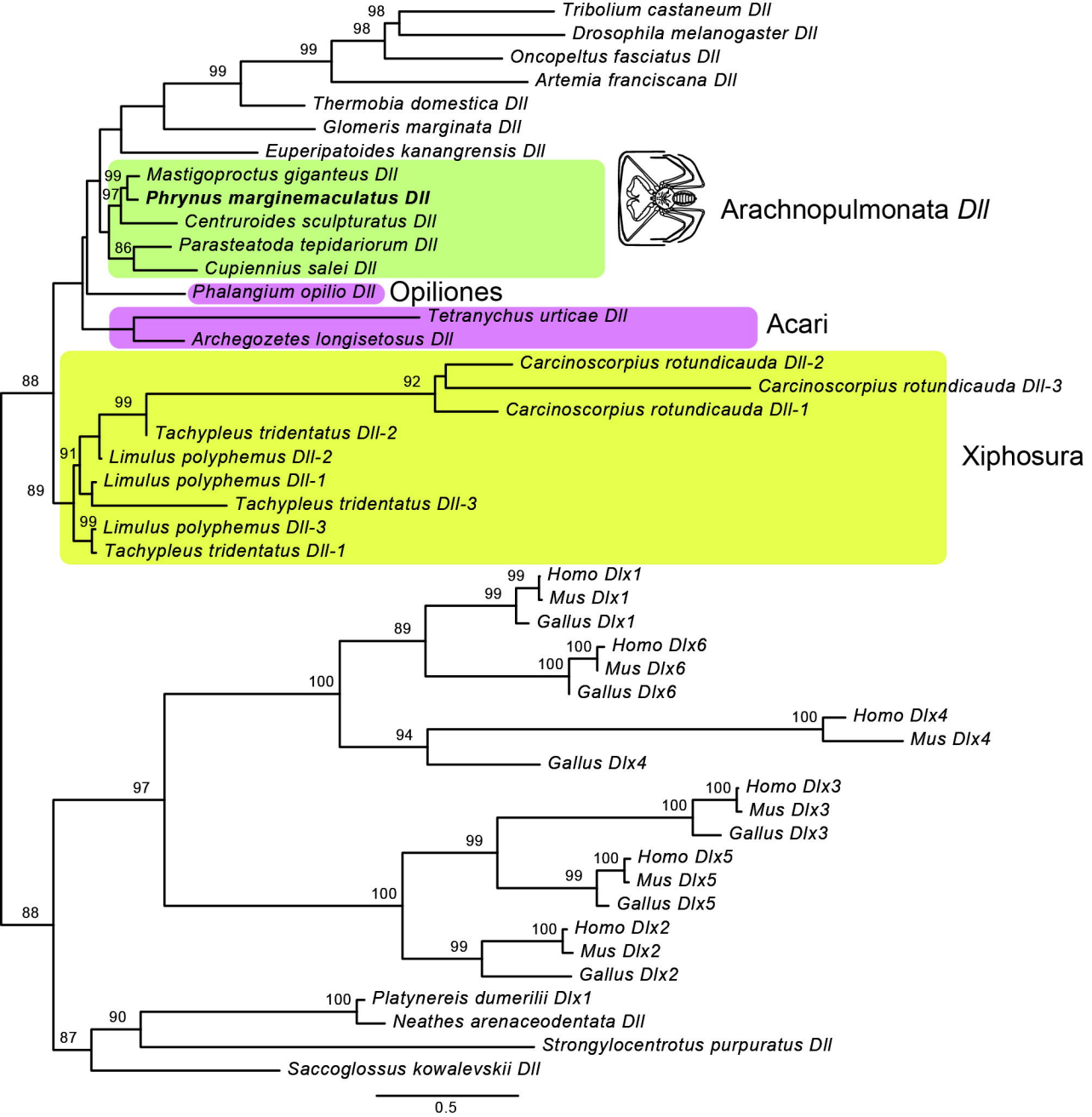

A
Ch
B
Pp
C
L1
D
L4
16th–17th dAEL
Pmar-dac1
E
Ch
F
Pp
G
L1
H
L4
17th–18th dAEL
I
Ch
J
Pp
K
L1
L
L4
16th–17th dAEL
Pmar-dac2
200μm

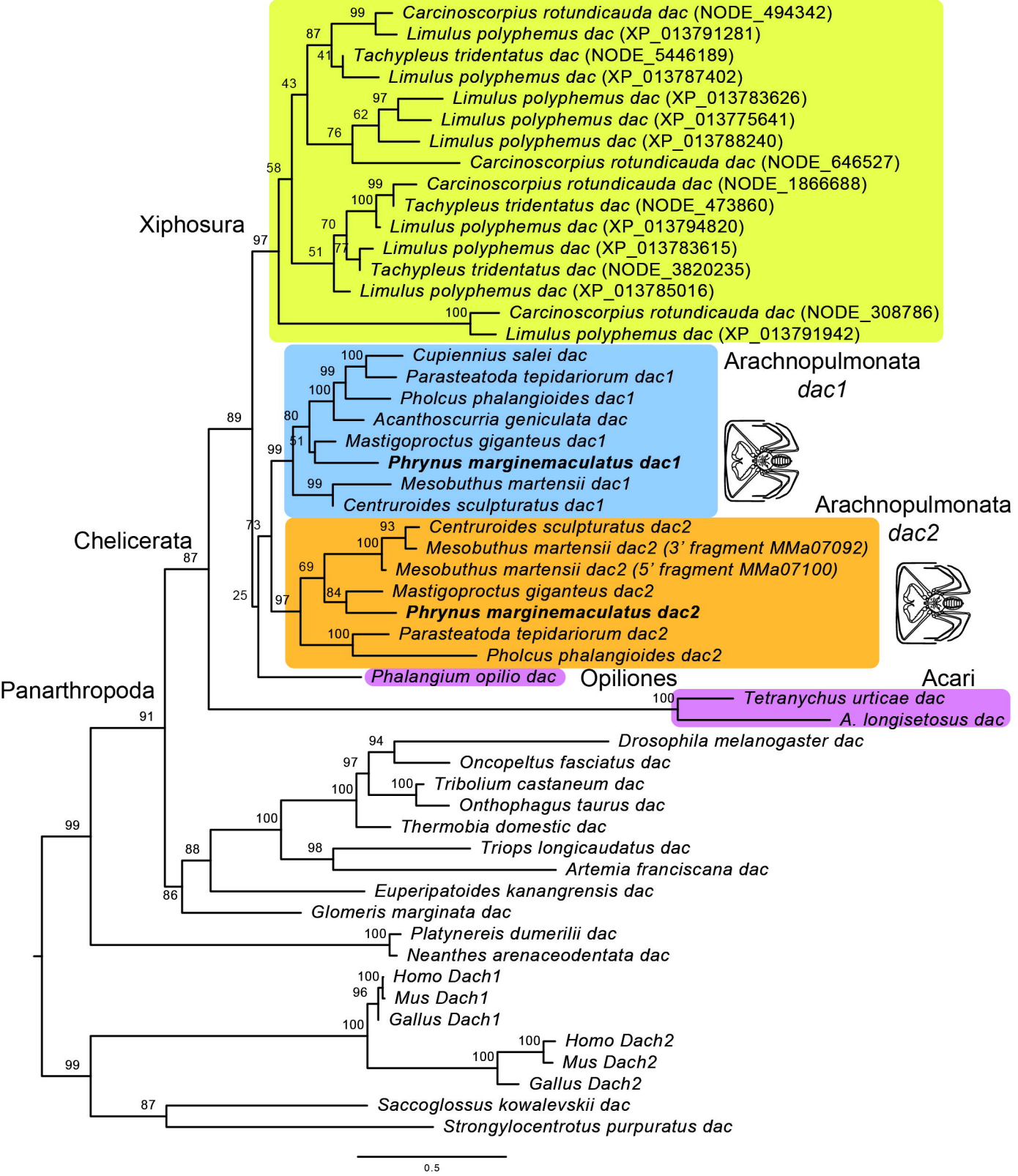

Pmar-hth1
Pmar-hth2
A
Ch
Pp
L1
L2
L3
L4
A
P
B
Pp
L1
L2
L3
L4
A
P
A'
Ch
Pp
L1
L2
L3
L4
A
P
B'
Pp
L1
L2
L3
L4
A
P
16th–17th dAEL
500 µm

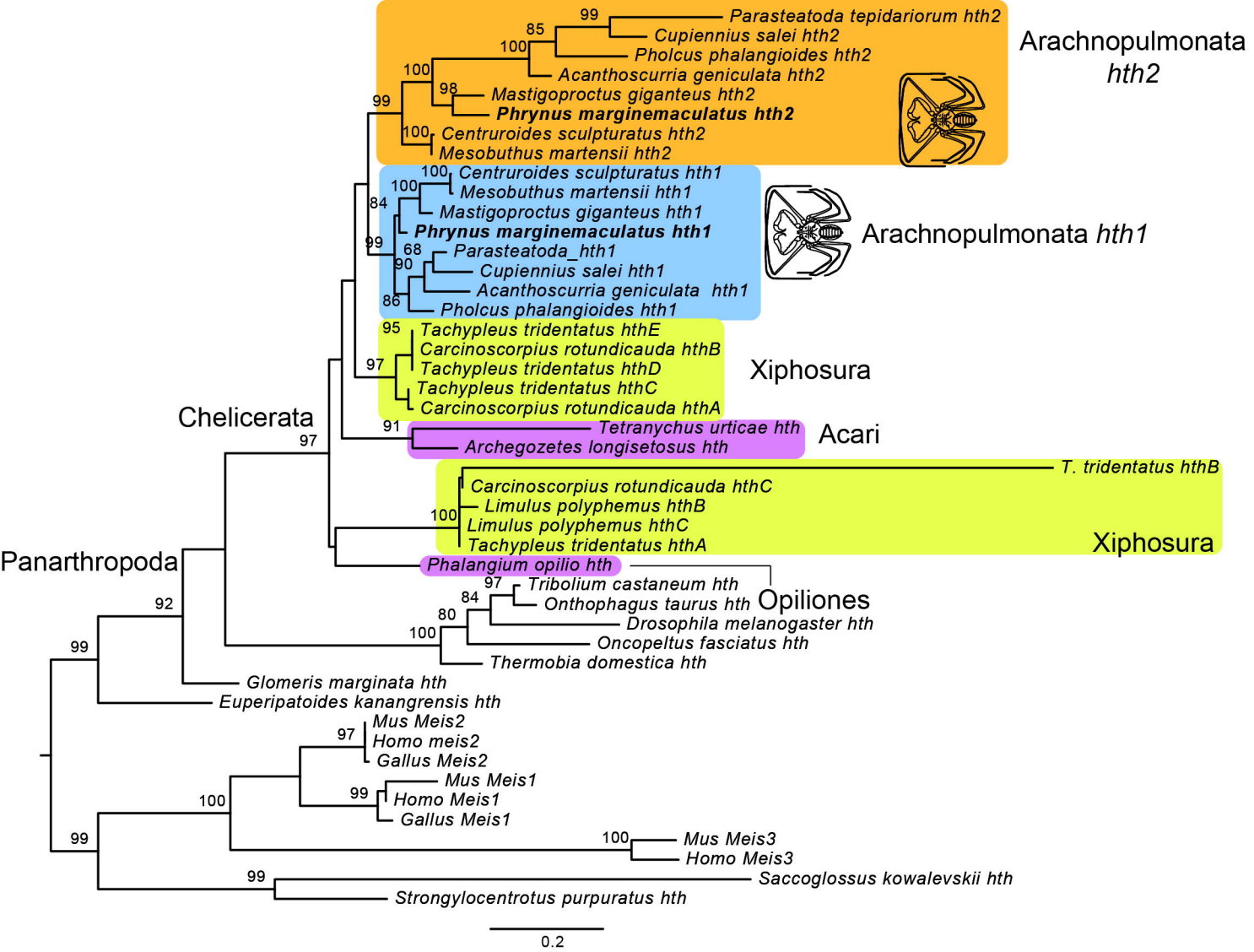

A
Ch
B
Pp
C
L1
D
L4
E
Ch
F
Pp
G
L1
H
L4
I
Ch
J
Pp
K
L1
L
L4
Pmar-hth1
Pmar-hth2
16th–17th dAEL
17th–18th dAEL
17th–18th dAEL
200μm


A
B
C


Deformed
Sex combs reduced
Ultrabithorax
Abdominal A
Antennapedia
Abdominal B
fushi tarazu
Hox3
proboscipedia
Hox3
labial
Dmel_Dfd_FBpp0081138
Tcas_Dfd_TC000920_001
Turt_Dfd_tetur20g02580.1
Cisr_Dfd_TRINITY_DN0_c16_g1_i1.p1
Cioa_Dfd_TRINITY_DN3191_c12_g1_i1_lcl_ORF1
Pmar_Dfd-1_TRINITY_ORF1_DN153191_c0_g3_i1
Cioa_Dfd_TRINITY_DN17883_c0_g2_i1.p2
Cisr_Dfd_TRINITY_DN11714_c0_g1_i4.p1
Cscu_Dfd-B_XM_023379077.1_18-1160
Cisr_Dfd_TRINITY_DN4458_c1_g2_i3.p1
Pmar_Dfd2_3prime_end_ORF24_TRINITY_DN155029_c7_g1_i1
Ptep_Dfd-A_NM_001323820.1_18-893
Cscu_Dfd-A_XM_023374825.1_1-1107
Isca_Dfd_XM_029976259.1_440-1549
Smar_Dfd_SMAR004162-RA
Pmar_Dfd-2_ORF1_TRINITY_DN153191_c0_g1_i11
Ptep-Dfd-B_XM_016058199.2_266-1513
Smar_Scr-RA
Cioa_Scr_TRINITY_DN11904_c1_g1_i1.p1
Cisr_Scr_TRINITY_DN6802_c0_g2_i1.p1
Isca_Scr_XM_002406361.2_315-1190
Pmar_Scr-1_ORF2_TRINITY_DN161173_c6_g2_i1
Cscu_Scr-A_XM_023374833.1_836-1735
Cscu_Scr-B_XM_023379091.1_661-1551
Cioa_Scr_TRINITY_DN17883_c0_g1_i3.p1
Cisr_Scr_TRINITY_DN6802_c0_g1_i1.p1
Pmar_Scr-2_ORF25_TRINITY_DN154468_c2_g4_i1
Ptep_Scr-A_XM_016066508.1_638-1555
Turt_Scr_tetur20g02540.1
Ptep_Scr-B_XM_021148680.1_779-1654
Tcas_Scr_TC000917_001
Dmel_scr_FBpp0428124
Cioa_Ubx_TRINITY_DN77041_c0_g1_i1.p1
Ptep_Ubx-B_XM_021148683.1_1166-2041
Pmar_Ubx-1_ORF3_TRINITY_DN159773_c10_g1_i10
Pmar_Ubx-2_ORF24_TRINITY_DN154468_c2_g2_i2
Isca_Ubx_XM_029986719.1
Cscu_Ubx-A_XM_023379259.1_197-1174
Cisr_Ubx_TRINITY_DN26767_c0_g1_i6.p1
Cisr_Ubx_TRINITY_DN54787_c0_g2_i2.p1
Dmel_Ubx_FBpp0082793
Tcas_Ubx_TC000903_001
Smar_Ubx_SMAR004167-RA
Cscu_Ubx-B_XM_023373102.1_89-871
Ptep_Ubx-A_XM_016066522.1_463-1260
Turt_Ubx_tetur20g02400.1
Cisr_abdA_TRINITY_DN41723_c1_g1_i3.p1
Dmel_abdA_FBpp0082829
Tcas_abdA_TC000894_001
Smar_abdA_SMAR004168-RA
Cscu_abdA-A_XM_023373101.1_1172-2029
Ptep_abdA-A_XM_016066514.1_138-926
Pmar_abdA_ORF15_TRINITY_DN147845_c0_g1_i7
Cisr_abdA_TRINITY_DN43038_c0_g1_i6.p1
Pmar_abdA_short_ORF4_TRINITY_DN156902_c2_g4_i1
Cscu_abdA-B_XM_023365627.1_497-1441
Ptep_abdA-B_XM_016075074.1_241-1284
Pmar_abdA_ORF24_TRINITY_DN156902_c2_g5_i1
Isca_abdA_XM_029986715.1_245-1261
Isca_Antp_XM_029986721.1_1-1182
Cscu_Antp-B_XM_023379181.1_921-1913
Ptep_Antp-B_XM_016062426.1_501-1718
Pmar_Antp-1_ORF24_TRINITY_DN149652_c0_g2_i5
Turt_Antp-1_tetur20g02430.1
Turt_Antp-2_tetur20g02440.1
Ptep_Antp-A_XM_016066516.1_585-1568
Cioa_Antp_TRINITY_DN77041_c0_g4_i5.p1
Cisr_Antp_TRINITY_DN26761_c0_g1_i3.p1
Pmar_Antp-2_ORF4_TRINITY_DN156902_c4_g2_i8
Cscu_Antp-A_XM_023374827.1_371-1270
Smar_Antp_SMAR004164-RA
Tcas_Antp_TC000912_001
Dmel_Antp_FBpp0081161
Cioa_AbdB_TRINITY_DN77136_c0_g1_i1.p3
Cisr_AbdB_TRINITY_DN42844_c0_g1_i2.p1
Ptep_AbdB-A_XM_016075104.1_297-1301
Pmar_AbdB-1_ORF22_TRINITY_DN162987_c0_g1_i2
Dmel_AbdB_FBpp0373669
Tcas_AbdB_TC000889_002
Isca_AbdB_XM_029986723.1_659-1603
Cscu_AbdB-A_XM_023373100.1_90-1034
Ptep_AbdB-B_XM_016075076.1_341-1462
Cisr_AbdB_TRINITY_DN74181_c0_g1_i2.p1
Pmar_AbdB-2_ORF13_TRINITY_DN190332_c0_g1_i1
Cscu_AbdB-B_XM_023365579.1
Cscu_AbdB-B_XM_023365632.1_1045-2184
Smar_Abd-B_SMAR004172-RA
Turt_AbdB_tetur20g02350.1
Pmar_ftz-2_ORF11_TRINITY_DN150804_c0_g1_i1
Turt_ftz-1_tetur20g02520.1
Turt_ftz-2_tetur20g02530.1
Tcas_ftz_TC000916_001
Ptep_ftz-B_XM_016062430.1_193-1131
Cscu_ftz-B_XM_023379217.1_841-1632
Cioa_ftz_TRINITY_DN41504_c0_g1_i1.p1
Cisr_ftz_TRINITY_DN17317_c0_g1_i1.p1
Pmar_ftz-1_ORF38_TRINITY_DN164783_c6_g1_i2
Smar_ftz_SMAR004163-RA
Isca_ftz_XM_002406359.2_1-1104
Dmel_ftz_FBpp0081139
Cscu_ftz-A_CSCU010689-RA
Smar_Hox3-A_SMAR008064-RA
Smar_Hox3-B
Cioa_Hox3_TRINITY_DN136315_c0_g1_i1_lcl_ORF3
Cisr_Hox3_TRINITY_DN17214_c0_g1_i3.p1
Pmar_Hox3_ORF12_TRINITY_DN164783_c3_g1_i1
Cioa_Hox3_TRINITY_DN77041_c0_g5_i2.p1
Cisr_Hox3_TRINITY_DN6523_c0_g2_i11.p1
Pmar_Hox-3_ORF3__TRINITY_DN159773_c5_g1_i1
Cscu_pb-B_XM_023366701.1_89-1789
Cscu_pb-B_XM_023379064.1_137-2188
Ptep_pb-B_XM_016053693.1_913-3237
Cisr_pb_TRINITY_DN57239_c1_g1_i1.p1
Ptep_pb-A_XM_016073330.1_295-2481
Pmar_pb_ORF5_TRINITY_DN156902_c2_g2_i1
Cscu_pb-A_XM_023374834.1_513-2603
Isca_pb_XM_029976216.1_25-2181
Cisr_pb_TRINITY_DN7597_c0_g1_i1.p1
Smar_pb_SMAR004160-RA
Tcas_pb_TC000925_001
Dmel_pb_FBpp0088333
Turt_pb_tetur11g05920.1
Isca_Hox3_XM_029976271.1_457-1800
Cscu_Hox3_XM_023379101.1_314-991
Ptep_Hox3-B_XM_016053691.2_420-2240
Tcas_Hox3_1_TC000921_001
Tcas_Hox3_2_TC000922_001
Dmel_zen1_FBpp0081170
Dmel_zen2_FBpp0081171
Dmel_bcd_FBpp0081168
Ptep_Hox3-A_XM_016055826.1_71-2602
Cisr_lab-2_TRINITY_DN41723_c0_g1_i4.p1
Cscu_lab-A_XM_023365857.1_371-820
Pmar_lab-2_ORF13_TRINITY_DN147887_c0_g1_i1
Cisr_lab-1_TRINITY_DN5459_c0_g2_i1.p1
Pmar_lab-1_ORF1_TRINITY_DN127740_c0_g1_i1
Ptep_lab-B_XM_016053694.1_181-1383
Cscu_lab-B_XM_023357555.1_1-465
Cscu_lab-B_XM_023366700.1_132-1250
Isca_lab_XM_029976297.1_245-1303
Dmel_lab_FBpp0081194
Tcas_lab_TC000926_001
Smar_lab_SMAR004159-RA
Ptep_lab-A_NM_001323834.1_226-1251
Turt_lab_tetur11g05940.1
3.0


Tribolium castaneum Dll
Drosophila melanogaster Dll
Oncopeltus fasciatus Dll
Artemia franciscana Dll
Thermobia domestica Dll
Glomeris marginata Dll
Euperipatoides kanangrensis Dll
Mastigoproctus giganteus Dll
Phrynus marginemaculatus Dll
Centruroides sculpturatus Dll
Parasteatoda tepidariorum Dll
Cupiennius salei Dll
Arachnopulmonata Dll
Phalangium opilio Dll
Opiliones
Tetranychus urticae Dll
Archegozetes longisetosus Dll
Acari
Carcinoscorpius rotundicauda Dll-2
Carcinoscorpius rotundicauda Dll-3
Carcinoscorpius rotundicauda Dll-1
Tachypleus tridentatus Dll-2
Limulus polyphemus Dll-2
Limulus polyphemus Dll-1
Tachypleus tridentatus Dll-3
Limulus polyphemus Dll-3
Tachypleus tridentatus Dll-1
Xiphosura
Homo Dlx1
Mus Dlx1
Gallus Dlx1
Homo Dlx6
Mus Dlx6
Gallus Dlx6
Homo Dlx4
Mus Dlx4
Gallus Dlx4
Homo Dlx3
Mus Dlx3
Gallus Dlx3
Homo Dlx5
Mus Dlx5
Gallus Dlx5
Homo Dlx2
Mus Dlx2
Gallus Dlx2
Platynereis dumerilii Dlx1
Neathes arenaceodentata Dll
Strongylocentrotus purpuratus Dll
Saccoglossus kowalevskii Dll
0.5


Carcinoscorpius rotundicauda dac (NODE_494342)
Limulus polyphemus dac (XP_013791281)
Tachypleus tridentatus dac (NODE_5446189)
Limulus polyphemus dac (XP_013787402)
Limulus polyphemus dac (XP_013783626)
Limulus polyphemus dac (XP_013775641)
Limulus polyphemus dac (XP_013788240)
Carcinoscorpius rotundicauda dac (NODE_646527)
Carcinoscorpius rotundicauda dac (NODE_1866688)
Tachypleus tridentatus dac (NODE_473860)
Limulus polyphemus dac (XP_013794820)
Limulus polyphemus dac (XP_013783615)
Tachypleus tridentatus dac (NODE_3820235)
Limulus polyphemus dac (XP_013785016)
Carcinoscorpius rotundicauda dac (NODE_308786)
Limulus polyphemus dac (XP_013791942)
Xiphosura
Cupiennius salei dac
Parasteatoda tepidariorum dac1
Pholcus phalangioides dac1
Acanthoscurria geniculata dac
Mastigoproctus giganteus dac1
Phrynus marginemaculatus dac1
Mesobuthus martensii dac1
Centruroides sculpturatus dac1
Arachnopulmonata
dac1
Centruroides sculpturatus dac2
Mesobuthus martensii dac2 (3' fragment MMa07092)
Mesobuthus martensii dac2 (5' fragment MMa07100)
Mastigoproctus giganteus dac2
Phrynus marginemaculatus dac2
Parasteatoda tepidariorum dac2
Pholcus phalangioides dac2
Arachnopulmonata
dac2
Chelicerata
Phalangium opilio dac
Opiliones
Acari
Tetranychus urticae dac
A. longisetosus dac
Panarthropoda
Drosophila melanogaster dac
Oncopeltus fasciatus dac
Tribolium castaneum dac
Onthophagus taurus dac
Thermobia domestic dac
Triops longicaudatus dac
Artemia franciscana dac
Euperipatoides kanangrensis dac
Glomeris marginata dac
Platynereis dumerilii dac
Neanthes arenaceodentata dac
Homo Dach1
Mus Dach1
Gallus Dach1
Homo Dach2
Mus Dach2
Gallus Dach2
Saccoglossus kowalevskii dac
Strongylocentrotus purpuratus dac
0.5


Parasteatoda tepidariorum hth2
Cupiennius salei hth2
Pholcus phalangioides hth2
Acanthoscurria geniculata hth2
Mastigoproctus giganteus hth2
Phrynus marginemaculatus hth2
Centruroides sculpturatus hth2
Mesobuthus martensii hth2
Arachnopulmonata hth2
Centruroides sculpturatus hth1
Mesobuthus martensii hth1
Mastigoproctus giganteus hth1
Phrynus marginemaculatus hth1
Parasteatoda_hth1
Cupiennius salei hth1
Acanthoscurria geniculata hth1
Pholcus phalangioides hth1
Arachnopulmonata hth1
Tachypleus tridentatus hthE
Carcinoscorpius rotundicauda hthB
Tachypleus tridentatus hthD
Tachypleus tridentatus hthC
Carcinoscorpius rotundicauda hthA
Xiphosura
Chelicerata
Tetranychus urticae hth
Archegozetes longisetosus hth
Acari
T. tridentatus hthB
Carcinoscorpius rotundicauda hthC
Limulus polyphemus hthB
Limulus polyphemus hthC
Tachypleus tridentatus hthA
Xiphosura
Panarthropoda
Phalangium opilio hth
Tribolium castaneum hth
Onthophagus taurus hth
Opiliones
Drosophila melanogaster hth
Oncopeltus fasciatus hth
Thermobia domestica hth
Glomeris marginata hth
Euperipatoides kanangrensis hth
Mus Meis2
Homo meis2
Gallus Meis2
Mus Meis1
Homo Meis1
Gallus Meis1
Mus Meis3
Homo Meis3
Saccoglossus kowalevskii hth
Strongylocentrotus purpuratus hth
0.2