Genomic and transcriptomic survey of bryozoan Hox and ParaHox genes with emphasis on phylactolaemate bryozoans

Ahmed J. Saadi[1*], André Luiz de Oliveira[2], Kevin M. Kocot[3] and Thomas Schwaha[1]

Abstract

Background Bryozoans are mostly sessile aquatic colonial invertebrates belonging to the clade Lophotrochozoa, which unites many protostome bilaterian phyla such as molluscs, annelids and brachiopods. While Hox and ParaHox genes have been extensively studied in various lophotrochozoan lineages, investigations on Hox and ParaHox gene complements in bryozoans are scarce.

Results Herein, we present the most comprehensive survey of Hox and ParaHox gene complements in bryozoans using four genomes and 35 transcriptomes representing all bryozoan clades: Cheilostomata, Ctenostomata, Cyclostomata and Phylactolaemata. Using similarity searches, phylogenetic analyses and detailed manual curation, we have identified five Hox genes in bryozoans (*pb*, *Dfd, Lox5, Lox4* and *Post2*) and one ParaHox gene (*Cdx*). Interestingly, we observed lineage-specific duplication of certain Hox and ParaHox genes (*Dfd, Lox5* and *Cdx*) in some bryozoan lineages.

Conclusions The bryozoan Hox cluster does not retain the ancestral lophotrochozoan condition but appears relatively simple (includes only five genes) and broken into two genomic regions, characterized by the loss and duplication of serval genes. Importantly, bryozoans share the lack of two Hox genes (*Post1* and *Scr*) with their proposed sister-taxon, Phoronida, which suggests that those genes were missing in the most common ancestor of bryozoans and phoronids.

Keywords Bryozoans, Phylactolaemate, Hox and ParaHox genes, Comparative genomic and transcriptomic, Gene duplication

*Correspondence:
Ahmed J. Saadi
ahmed.saadi@univie.ac.at
[1]Department of Evolutionary Biology, Unit for Integrative Zoology, University of Vienna, Schlachthausgasse 43, Vienna A-1030, Austria
[2]Department of Symbiosis, Max-Planck-Institute for Marine Microbiology, Celsiustraße,1, D-28359 Bremen, Germany
[3]Department of Biological Sciences and Alabama Museum of Natural History, University of Alabama, Tuscaloosa, Alabama 35487, USA

Background

The Hox and ParaHox gene complements are a set of genes that encode transcription factors with a highly conserved sequence region (60 amino acid homeodomain) [1]. They belong to a class of homeobox genes (Antennapedia-class) which are responsible for the regulation of early embryonic development and are involved in patterning the anterior-posterior body axis in Bilateria [2, 3]. Hox genes were originally discovered in the fruit fly *Drosophila melanogaster* [4, 5] and later reported in vertebrates [6, 7], while the ParaHox cluster (the proposed evolutionary sister of the Hox cluster) was initially described in the invertebrate chordate amphioxus *Branchiostoma floridae* [8]. Genomic and transcriptomic data suggest that Hox and ParaHox genes have arisen by duplication and divergence of an ancestral ProtoHox cluster early in metazoan evolution [9]. Based on sequence similarity, the Hox gene complement can be classified into four major classes: the Anterior class, *Hox3* class, Central class and Posterior class [10]. The genetic architecture of these classes varies across taxa due to duplication, inversion, or gene loss events that took place during evolution [9, 10]. Variation in the expression patterns, number, and sequence of Hox and ParaHox genes correlate with body plan evolution during the radiation of several clades such as Arthropoda, Annelida, Chordata and Mollusca [9, 11–13]. Therefore, reconstructing the evolutionary history and determining the expression patterns of these genes are crucial for better understanding animal evolution and the relationships between genetics and different levels of morphological complexity [14].

One of the major clades of protostome animals is Lophotrochozoa, which is mainly characterized by the trochophore, a free-swimming ciliated larva (Trochozoa) and/or the lophophore feeding organ composed of ciliated tentacles surrounding the mouth (Lophophorata) [15]. The ancestral Hox gene complement of Lophotrochozoa is hypothesized to have been composed of 11 genes (*Hox1/lab, Hox2/pb, Hox3, Hox4/Dfd, Hox5/Scr, Lox5, Antp, Lox4, Lox2, Post1 and Post2*) whilst the ParaHox cluster generally comprises three genes (*Gsx, Xlox* and *Cdx*) [16]. The Hox and ParaHox gene complements of lophotrochozoans have been examined in a variety of taxa based on genomic or transcriptomic data (for details see [16]). In the lophophorate clades Brachiopoda and Phoronida, it seems that the Hox and ParaHox gene complements retain the ancestral lophotrochozoan condition [17–19]. For example, the Hox cluster in the phoronids (*Phoronis australis* and *P. harmeri*) consists of eight Hox genes, with *Antp, Post1* and *Scr* being absent [17, 19]. Likewise, the Hox cluster in brachiopods (*Lingula anatina* and *Terebratalia transversa*) comprises almost a complete set of genes, with only *Post1* missing *in L. anatine and Scr missing in T. transversa* [17, 20]. Additionally,

Both Brachiopoda and Phoronida possess a complete set of ParaHox genes [17, 19, 20].

Within Lophotrochozoa, Bryozoa represents a rather large aquatic phylum with more than 6,000 extant species of almost entirely colonial suspension feeders [21, 22]. The evolution of bryozoan diversity might be linked to the role of developmental genes, particularly Hox genes [23]. It has been shown that Hox genes can be recruited and coopted into lineage-specific morphological structures, in addition, to their ancestral role in head-to-tail body plan specification [16, 24]). Although the Hox and ParaHox gene complements of lophotrochozoans have been extensively investigated, in the lophophorate group Bryozoa, the presence or absence of Hox genes has been only examined in a single bryozoan species, *Bugula turrita*, based on a targeted search using polymerase chain reaction (PCR) [25] while the ParaHox genes have never been investigated. The examination of Hox genes in Bryozoa revealed a possible loss of six genes including *Hox1*, *Hox5*, *Antp*, *Lox4*, *Lox2*, and *Post1* [25]. Genomic and transcriptomic studies of bryozoans are limited [26–28] and have not focused on the evolution or the organization of Hox and ParaHox genes.

Our understanding of Hox and ParaHox gene evolution in Bryozoa is currently limited and it is largely based on targeted search of Hox genes using PCR where it is difficult to recover the full Hox cluster. Therefore, a comprehensive investigation of Hox and ParaHox clusters using genomic and transcriptomic data is required to ascertain whether the alleged missing Hox genes are indeed absent from the genomes of bryozoans. In this study, we provide the first examination of Hox and ParaHox genes in bryozoans by analysing four genomes and 35 transcriptomes from representative of all main bryozoan linages: Gymnolaemata, Phylactolaemata and Stenolaemata. Using similarity searches, phylogenetic inferences and a detailed manual curation (i.e., identification of conserved residues and protein motifs), we have identified and confirmed the orthology of five Hox and one ParaHox genes in 35 bryozoan species, providing clustering evidence for Hox genes.

Results

Hox and ParaHox genes in bryozoans

Assessment of the 35 transcriptomes and the four genomes with BUSCO showed that the majority of the investigated transcriptomes (28) and two genomes have BUSCO values for complete and fragmented sequences above 94%, indicating high completeness scores (Supplementary Figure S1). Of the 11 ancestral Hox genes in Lophotrochozoa, only five candidate orthologues of Hox genes were identified in the bryozoan genomes and transcriptomes. Those genes represent orthologues of the Hox genes *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Lox5*,

Lox4 and *Post2* (Figs. 1, 2 and 3 and Supplementary Figures S2-S8). Of the three genes belonging to the ParaHox complement in Lophotrochozoa, only one candidate orthologue was found (*Cdx*) in our analyses (Figs. 1, 2 and 3, and Supplementary Figures S2 and S8). The recovered Hox and ParaHox genes of bryozoans (except *Lox4*) form distinct clades in our phylogenetic trees with bootstrap support values ≥70 (Fig. 1) and posterior probabilities ≥0.90 (Supplementary Figure S2). Exact orthology assessment of the *Lox4* gene in bryozoans was not possible based on our phylogenetic analyses (Supplementary Figures S6) as *Lox4* orthologues in bryozoans showed a sister-group relationship to a clade comprising both *Lox2* and *Lox4* genes of non-bryozoan taxa. Therefore, we have
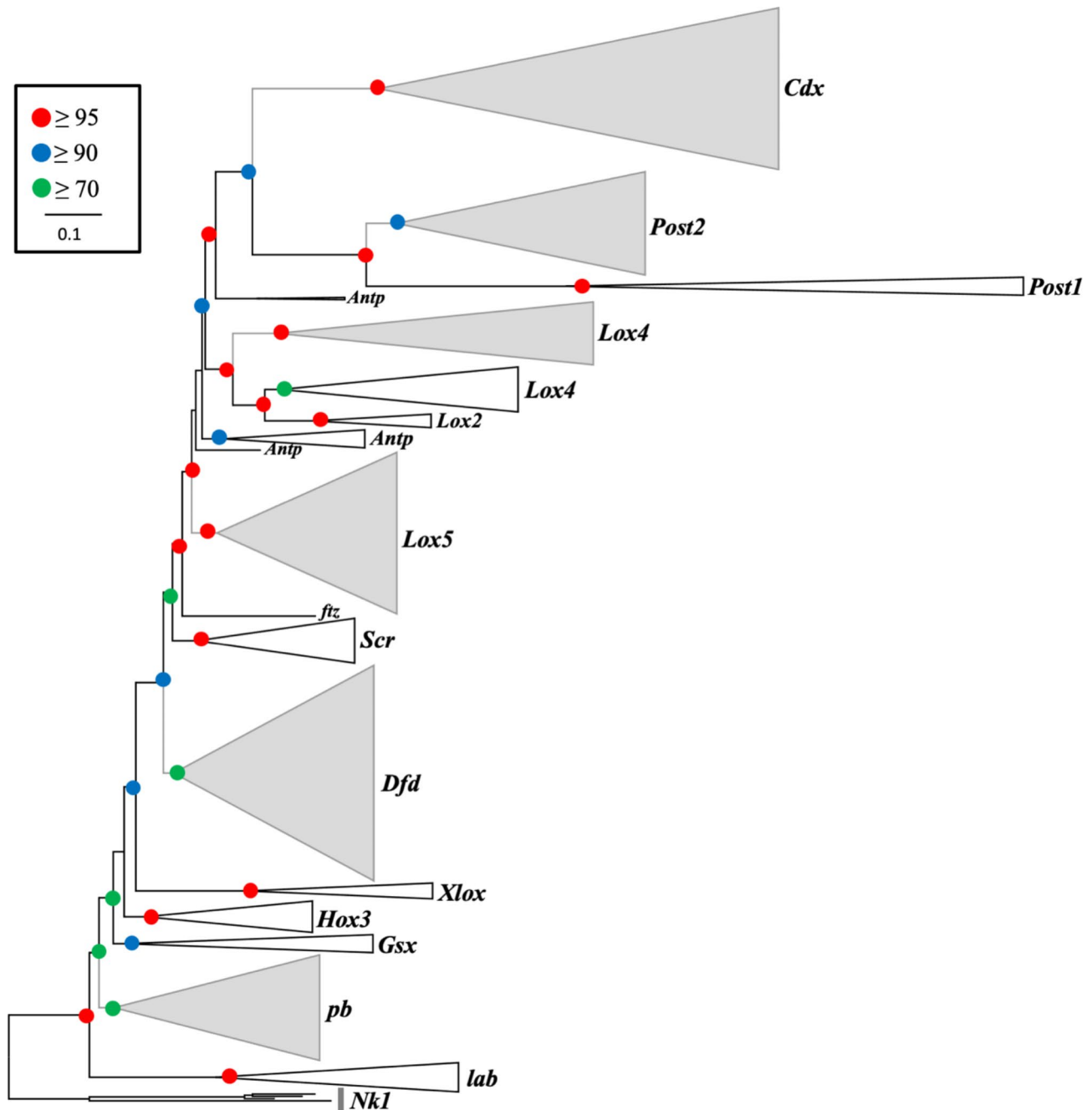
Fig. 1 Maximum likelihood phylogeny of Hox and ParaHox genes based on amino acid sequences containing homeodomain and flanking regions of bryozoans and a selection of lophotrochozoan species. ML bootstrap support values are represented by the coloured circles on tree nodes. Clades were collapsed to allow better visibility. Recovered Hox and ParaHox genes in bryozoans are highlighted in grey and their expanded trees are provided in Supplementary Figure S3 -S8. The scale bar indicates amino acid substitutions per site. The homeobox gene Nk1 was used as the outgroup
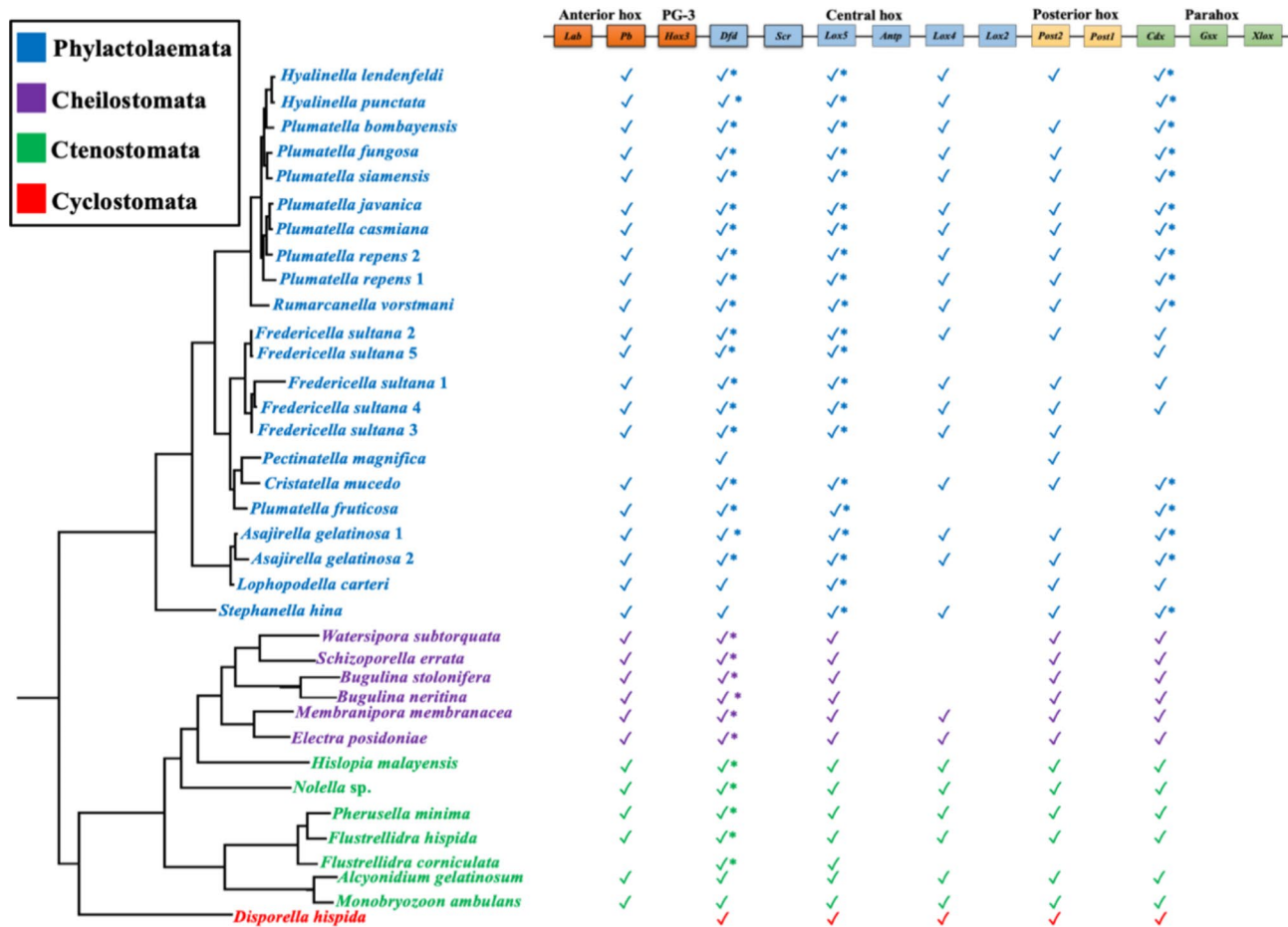
Fig. 2 Summary of the Hox and ParaHox genes identified in the 35 bryozoan species studied herein. For comparison, the putative ancestral lophotrochozoan Hox/ParaHox complement is provided. Tree topology follows Saadi et al. [30]. Check signs represent the presences of genes and asterisks indicate duplicated genes

relied on the sequence alignment to confirm the identity of *Lox4* in bryozoans and found that the sequence of this gene has most of the diagnostic signature residues of *Lox4* [29] (Supplementary Figure S9) and lacks typical signatures of *Lox2* [29]. For the Hox *Dfd* gene, two gene copies *Dfda* and *Dfdb* were identified in bryozoans, both copies form distinct clades in Phylactolaemata with strong support, whereas *Dfda* is distributed on two separate clades in Myolaemata (Supplementary Figures S4). Likewise, two copies were recovered for the Hox gene *Lox5* (*Lox5a*, *Lox5b*, Supplementary Figures S5) and the ParaHox gene *Cdx* (*Cdxa* and *Cdxb*, Fig. 9) in phylactolaemate bryozoans. These gene copies form well-supported clades in the phylogenetic trees. Our genomic analysis has demonstrated that each copy of the *Dfd*, *Lox5* and *Cdx* genes exists in distinct genomic environments, confirming they are authentic duplicates, and not merely a result of alternative splicing or assembly artifacts.

Finally, orthologues of six Hox genes (*lab*/*Hox1*, *Hox3*, *Scr*/*Hox5*, *Antp*, *Lox2* and *Post1*) and two ParaHox genes (*Gbx* and *Xlox*) were not identified in the investigated

bryozoan genomes or transcriptomes. Additionally, we found that the gene identified by YL Passamaneck and KM Halanych [25] as *Hox3* in *Bugula turrita* is most likely contaminated or misidentified (Supplementary Figure S10).

Hox cluster organization in bryozoan genomes

We examined the organization of the Hox cluster in four bryozoan species whose genomes are publicly available, three of them belonging to Cheilostomata (*Bugula neritina*, *Bugulina stolonifera* and *Membranipora membranacea*) and one to Phylactolaemata (*Cristatella mucedo*). In the three cheilostome genomes (*B. neritina*, *B. stolonifera* and *M. membranacea*), Hox genes are located in two different genomic regions (chromosomes or scaffolds), one containing *pb*, *Dfda*, *Lox5* and *Post2* genes (in *M. membranacea* it also includes *Lox4*) and the other one containing the *Dfdb* gene. Hox genes have a consistent order among these three genomes and they share the same transcriptional orientation except *dfdb* in *B. neritina* which is transcribed in a different direction (Fig. 3).
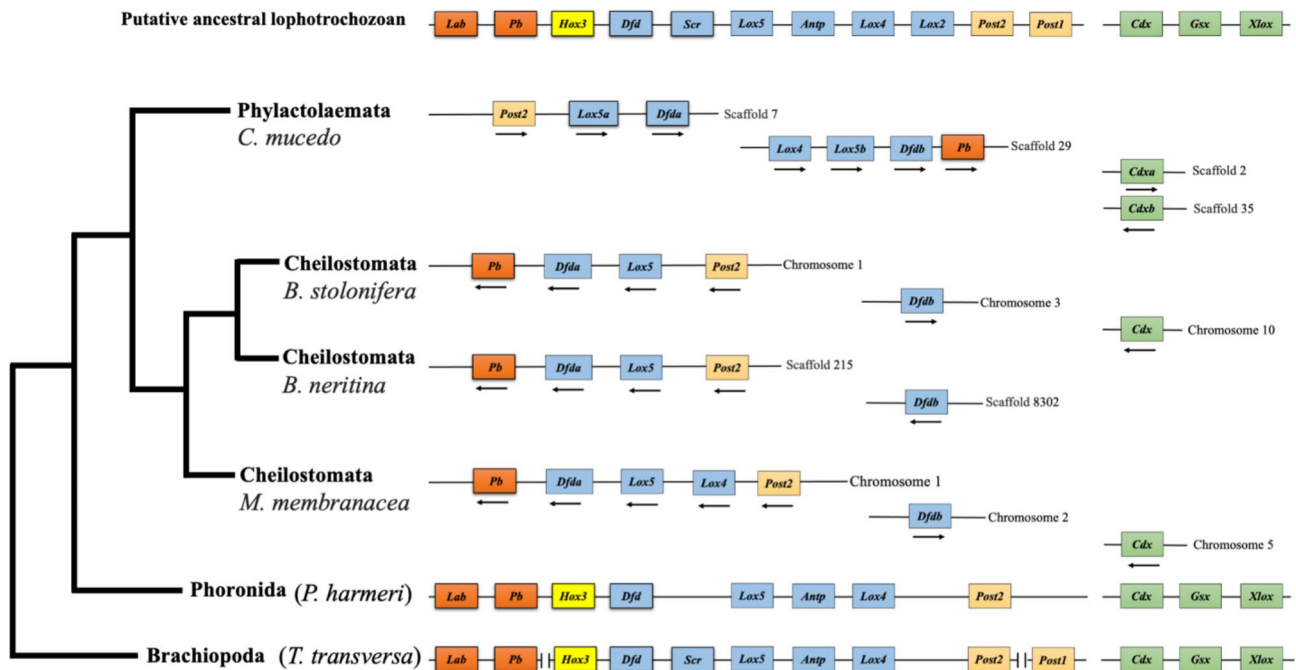
Fig. 3 A comparison of Hox and ParaHox clusters among lophophorates. On the left, bryozoan phylogeny with Phoronida and Brachiopoda follows Saadi et al. [30]. On the right, schematic representation of Hox and ParaHox gene complements of Lophophorata (Hox gene clusters of brachiopods and phoronids follows [20]). For comparison, the putative ancestral lophotrochozoan gene toolkit is provided. The coloured boxes indicate the presences of Hox and ParaHox genes. Arrows indicate transcript directions. Hox genes that are hosted within the same chromosome or scaffold are connected by a gray line. The intergenic spaces are not scaled

Table 1 Number of introns (first value) and protein length (second value) of the identified Hox and ParaHox genes in four bryozoan genomes (absent genes = "-")

Genes Species	*pb*	*dfda*	*dfdb*	*lox5a*	*lox5b*	*lox4*	*post2*	*cdxa*	*cdxb*
B. neritina	2/256	2/233	3/231	2/255	-	-	1/265	-	-
B. stolonifera	2/264	2/236	3/231	1/259	-	-	1/270	2/254	-
M. membranacea	2/251	1/183	3/234	1/242	-	1/77	1/249	2/342	-
C. mucedo	2/309	2/258	1/261	1/219	1/259	3/222	1/219	2/213	2/299

In both of *B. stolonifera and M. membranacea* genomes, Hox genes occur on the same chromosome (chromosome two) except *Dfdb* which is located on chromosome three in *B. stolonifera* and on chromosome one in *M. membranacea* (Fig. 3). In the phylactolaemate species *Cristatella mucedo*, Hox genes also fall into two different scaffolds, one harbouring *pb*, *Dfdb*, *Lox5b* and *Lox4* genes and the other containing all the remaining Hox genes (*Dfda*, *Lox5a* and *Post2*), (Fig. 3). The Hox gene order in *C. mucedo* diverge from the other three cheilostome genomes. The first part of the Hox cluster in *C. mucedo* includes three genes in the following order: *Post2*, *Lox5a*, and *Dfda* while the second part contains four genes with the order *Lox4*, *Lox5b*, *Dfdb* and *pb* (Fig. 3). Furthermore, in *C. mucedo*, the ParaHox cluster exhibits distinctions compared to the three other cheilostome genomes due to the presence of two copies of the *Cdx* gene, each located on separate genomic regions. It is worth noting

that the intron numbers of *Dfdb* and *Lox4* genes in *C. mucedo* also diverge from those in the other cheilostome genomes (Table 1).

Discussion

In this study, we examined the Hox and ParaHox genes in 35 bryozoan transcriptomes and four genomes representing all bryozoan clades: Cheilostomata, Ctenostomata, Cyclostomata and Phylactolaemata. Five Hox genes (*pb*, *Dfd*, *Lox5, Lox4* and *Post2*) and one ParaHox gene (*Cdx*) were identified in bryozoans. Our results revealed that bryozoans have experienced significant gene losses in the Hox and ParaHox complements including loss of six Hox genes (*lap*, *Hox3*, S*cr*, *Antp*, *Lox2* and *Post1*) and two ParaHox (*Gsx* and *Xlox*) genes. Prior to our study, Hox genes had only been investigated in a single bryozoan species [25] in which five Hox genes (*pb*, *Hox3*, *Dfd*, *Lox5*, and *Post2*) were also identified. However, the *Hox3*

gene was not recovered from any bryozoan in our analyses indicating that this gene is evidently missing in bryozoans and that the *Hox3* sequence previously reported by YL Passamaneck and KM Halanych [25] may have been the result of contamination. This is also suggested by our phylogenetic analysis as the previously published *Hox3* gene with non-bryozoan *Hox3* orthologues with strong support.

The *Hox3* gene has been reported in most of the lophotrochozoan linages [16] including the lophophorate groups Brachiopoda [18, 31] and Phoronida [17, 19]. However, specific role of *Hox3* varies between different lophotrochozoan taxa and it has not been fully characterized in most phyla. For example, in the phoronid *Phoronopsis harmeri*, the *Hox3* gene is exclusively expressed in ectodermal cells [17, 19] while in the brachiopod species, *Novocrania anomala* and *Terebratalia transversa*, *Hox3* is expressed in ectodermal and mesodermal cells [31]. Similarly, the *Antp* gene is also missing in bryozoans, but present in many other lophotrochozoan groups [16] including brachiopods and phoronids where it is only ectodermally expressed [17, 19]. Since both *Hox3* and *Antp* genes are missing in bryozoans and present in brachiopods and phoronids, we suggest that those two genes were lost after bryozoans split from their last common ancestor.

Although it has been suggested that *Post1* and *Scr* genes are involved in biomineralization of brachiopods [20, 31], the lack of these two genes in bryozoans is not surprising as the ancestor of bryozoans was soft-bodied [30, 32]. The mineralized skeleton of cheilostomes and stenolaemates has most likely evolved independently in these two clades [33, 34]. Furthermore, the lack of *Post1* and *Scr* genes in bryozoans is shared with phoronids which might indicate that these genes were lost in their last common ancestor. A sister-group relationship between phoronids and bryozoans has been proposed by several phylogenomic studies [35–38] and our finding regarding the shared missing Ho*x* genes in bryozoans and phoronids are supportive of this hypothesis. However, a recent phylogenomic analyses recovered a sister group relationship between Kamptozoa (=Entoprocta) and Bryozoa, which contradicts the monophyly of Lophophorata [39]. In comparison to bryozoans, almost complete sets of Hox genes (10 of 11 Hox genes commonly found in Lophotrochozoans) were identified in kamptozoan transcriptomes and genomes [40]. Thus, the data obtained in our study provides evidence for a closer relationship of bryozoans to phoronids than to kamptozoans.

Interestingly, two of the identified Hox gene orthologues in bryozoans (*Dfd* and *Lox5*) have two copies in most bryozoans. The phylogenetic analyses provide strong support for each copy of those orthologue groups. The duplication of *Dfd* and *Lox5* genes is also confirmed by the genomic analysis since each copy is located on a different genomic region. In the *Dfd* gene, the duplication occurs in three bryozoan lineages: Cheilostomata, Ctenostomata and Phylactolaemata while the duplication of *Lox5* gene occurs only in phylactolaemate bryozoans. Two copies of the *Dfd* gene were previously reported in *Bugula turrita* by YL Passamaneck and KM Halanych [25] where they suggested that duplication of *Dfd* gene happened when bryozoans formed an independent lineage. Duplication events in the *Dfd* gene have also been reported in other spiralian lineages including Rotifera (*Adineta vaga*) [13], Annelida (*Perionyx excavatus*) [41] and in the nemertean *Notospermus geniculatus* [17]. Meanwhile, the duplication of *Lox5* gene is less common in lophotrochozoans and has only been shown in the platyhelminths *Dugesia japonica* and *Girardia tigrine* [16].

Overall, since the duplication of the *Dfd* gene occurs in both phylactolaemates and myolaemates, it is probable that this gene has undergone duplication independently after the divergence of bryozoans from their last common ancestor while the duplication of *Lox5* gene seems to be linage specific and occurred in last common ancestor of all phylactolaemates.

ParaHox genes have so far not been characterized in any bryozoan species, and similar to the Hox cluster, the ParaHox complement of bryozoans shows gene duplications and losses. For instance, out of the three ancestral lophotrochozoan ParaHox genes (*Gsx*, *Xlox* and *Cdx*), only *Cdx* gene (Caudal) was recovered in bryozoans. However, *Cdx* has two copies in most of the investigated phylactolaemates with each copy located on a separate scaffold of the *C. mucedo* genome. The duplication of *Cdx* gene in phylactolaemates is probably lineage-specific and could be related to the life cycle of this clade. Many studies have suggested that Caudal in lophotrochozoans is generally expressed in ectodermal and endodermal cells and less likely in mesodermal ones [42–45]. However, the functional characterization of Caudal has been less investigated [46] and recently was studied in one lophotrochozoan species (in the embryo of the mollusc *Tritia*), showing that Caudal is required for development of the hindgut in a mollusc [47]. Functional studies of other ParaHox genes are lacking for the vast majority of lophotrochozoan lineages. In contrast to bryozoans, a complete set of ParaHox genes were identified in the phoronids *Phoronis australis* [17] and *Phoronopsis harmeri* [19]. Losses of ParaHox genes is common in lophotrochozoans (e.g., *Cdx* is missing in the kamptozoan *Loxosomella murmanica*, *Gsx* is absent in the kamptozoans *Loxosomella vivipara* and *Pedicellina cernua* [40], and *Xlox* is missing in the nemertean *Notospermus geniculatus* [17]).

Our genomic information shows that the Hox cluster of the cheilostomes *B. neritina*, *B. stolonifera* and *M. membranacea* is split into two different genomic loci, with *Dfda* being separated from the major cluster. Similarly, the Hox cluster of the phylactolaemate *Cristatella mucedo* is distributed over two scaffolds, one comprises *pb*, *Dfdb*, *Lox5b* and *Lox4* and the other one includes *Dfda*, *Lox5a* and *Post2*. Dissociation of the Hox cluster is common in lophotrochozoans, for example, the annelids *Capitella teleta* has a split Hox cluster and the leech *Helobdella robusta*, shows a highly fragmented Hox complex [48]. The lack of Hox cluster is also reported in the cephalopod *Octopus bimaculoides* [49], the pacific oyster *Crassostrea gigas* [50], the brachiopod *Terebratalia transversa* [31], and the nemertean *Notospermus geniculatus* [17]. Importantly, the phoronid *Phoronis australis* has one Hox cluster [17]. Although, phoronids preserved the Hox synteny, they also lack some of the Hox genes, including *Lox2*, *Post1* and *Scr*. The secondary loss of Hox genes seems pervasive in bryozoans and phoronids. However, in bryozoans, the synteny appears to be broken as well, suggesting that there are two layers of complexity in bryozoans (Loss and shift of the genomic loci). This might indicate that bryozoan genomes are more dynamic, though further syntenic analyses are necessary to investigate these two lineages.

Conclusions

Bryozoans do not retrain the ancestral lophotrochozoan Hox and ParaHox clusters, but instead they show rather simple (including only six genes) and broken clusters with significant gene losses but also duplications. Only five Hox genes and one ParaHox gene were identified in bryozoans, which is much fewer than the number identified in their closely relative lophophorate taxa brachiopods and phoronids or even in kamptozoans, which have also been hypothesized to the sister-taxon of bryozoans based on a recent phylogenomic study. In addition, Hox gene duplications have neither been reported in other lophophorates nor in kamptozoans, which might indicate that Hox and ParaHox duplications in certain bryozoan genes are lineage specific. Still, exact details and the extent of these duplications are still unclear as most of our data were obtained from transcriptomes. Bryozoans and phoronids lack *Post1* and *Scr* genes which could be related to soft-body status of their common ancestor as those genes are generally expressed in biomineralizing tissues. Further genomic and transcriptomic studies with more taxon sampling (especially cyclostomes) are needed to determine the presence, expression patterns, and functional significance of Hox and ParaHox genes in bryozoans.

Materials and methods

Transcriptome assembly and quality assessment

A total of 35 bryozoan transcriptomes (as either raw sequence reads or assembled transcriptomes) and four genome assemblies were obtained from publicly available data (details of the specimens, GenBank Bioproject accession numbers and sources of specimens are provided in Supplementary Table S1). In order to assemble the raw Illumina reads from transcriptomes, adapters and low-quality reads were first removed from raw sequence reads using Trimmomatic v0.39 [51] with default parameters. The clean reads were *de novo* assembled in Trinity v2.8.4 [52], under default settings. The assembled transcriptomes were then run through Transdecoder. LongOrfs v5.02 to extract all possible coding sequence regions from the transcripts and through Transdecoder. Predict v5.02 with the --single_best_only to select only the single best open reading frames (ORFs) (the longest coding sequence regions within the transcript) (https://github.com/TransDecoder/TransDecoder/; last accessed July 29, 2022). Only ORFs that were at least 80 amino acids long were retained. To reduce redundancy in protein gene sets, CD-HIT v4.8.1 [53] was used with a threshold of 95% global similarity. Finally, the gene content and the completeness of the transcriptomes were assessed with BUSCO v4.1.4 [54] using the pre-defined metazoan Benchmarking set of Universal Single-Copy Orthologs with 954 evolutionary conserved orthologous groups (metazoan_odb10).

Hox and ParaHox gene sequences identification and orthology assessment

Using Blastp v2.12.0+ [55], bryozoan protein sequences from all transcriptomes were searched against well-curated publicly available metazoan Hox and ParaHox sequences including bryozoan Hox gene candidates. The top three blast hits of each similarity search were analysed and blasted back against GenBank non-redundant protein database to reconfirm the homology. Furthermore, to make sure that the full Hox and ParaHox genes in bryozoans were recovered, Exonerate v2.4.0 [56] with protein2genome model and maximum intron length set to 40 kb was used to scan the whole genome assemblies of four bryozoan species (*Cristatella mucedo*, *Bugula neritina*, *Bugulina stolonifera* and *Membranipora membranacea*), whose genome annotations were not publicly available at the time of these analyses. The longest CDS for each Hox and ParaHox gene was then selected and translated to protein using EMBOSS Transeq online server (https://www.ebi.ac.uk/Tools/st/emboss_transeq/) [57]. The putative bryozoan Hox and ParaHox genes from transcriptomes and genomes were aligned together using MAFFT 7.310 [58] combined with their representative homologs from different metazoan phyla also retrieved

from GenBank non-redundant protein database under the following options: --localpair and --maxiterate 1000. All sequences were carefully inspected then manually edited with Aliview v.2022 [59] and trimmed with tri-mAl (-gt 0.4) [60]. Phylogenetic analyses were performed using Maximum likelihood (ML) and Bayesian inference (BI) methods. The ML analysis was conducted with IQ-TREE2 v2.1.2 [61] using ModelFinder tree search with 1,000 ultrafast bootstraps and SH-aLRT test replicates [62, 63]. For BI analysis, the best substitution model was first selected (JTT+I+G4) using maximum AIC as determined using ModelTest-NG [64]. The BI phylogenetic analysis was performed using MrBayes 3.2.7a [65] with JTT+I+G4 model of amino-acid substitution as determined using maximum AIC as implemented in ModelTest-NG [64]. Two independent runs with four chains of Markov Chain Monte Carlo (MCMC) algorithm were used to explore the tree space. BI analysis was conducted for 10 million generations sampled every 100 generations. The first 25% of samples were discarded as burn-in and the remaining trees were used to calculate posterior probability values and to build the consensus tree. The final ML and BI trees were visualized in Figtree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). In case of gene orthology was not confirmed based on phylogenetic analyses (i.e., gene did not form a monophyletic group), the multiple sequence alignment was searched for the presence of diagnostic residues/motifs in the homeodomain and in the flanking regions based on de Rosa et al. (1999).

Finally, to determine Hox and ParaHox gene locations, transcriptional orientations, intron number and lengths in four bryozoan species whose genomes are publicly available (without annotation), we aligned the identified Hox and ParaHox genes of those four species back to their respective genomes using Exonerate with protein2genome as described above.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-023-09826-z.

```
Supplementary Material 1
```

Authors' contributions
A.J.S. and T.S. designed the project. A.J.S. Designed and executed the bioinformatics pipelines, performed the data analysis, and drafted the manuscript. A.L.D and K.M.K. contributed to the bioinformatics analysis. A.J.S. and T.S. finalised the manuscript. All authors read, commented on, and provided the final version of the manuscript.

Data availability
The data presented in this study including the alignment of the metazoan Hox and ParaHox sequences, the alignment of the identified Hox and ParaHox genes in bryozoans before and after trimming and the full trees are available from figshare: https://figshare.com/s/0ecf2eee0d06fd32c458.

Declarations

Ethics approval and consent to participate
Ethical approval and consent to participate were not required for this work. Specimens were collected in accordance to local guidelines.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell. 1984;37(2):403–8.
2. Pearson JC, Lemons D, McGinnis W. Modulating hox gene functions during animal body patterning. Nat Rev Genet. 2005;6(12):893–904.
3. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell. 1992;68(2):283–302.
4. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature. 1978;276(5688):565–70.
5. McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ. A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. Nature. 1984;308(5958):428–33.
6. Carrasco AE, McGinnis W, Gehring WJ, De Robertis EM. Cloning of an X. Laevis gene expressed during early embryogenesis coding for a peptide region homologous to Drosophila homeotic genes. Cell. 1984;37(2):409–14.
7. McGinnis W, Hart CP, Gehring WJ, Ruddle FH. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of Drosophila. Cell. 1984;38(3):675–80.
8. Brooke NM, Garcia-Fernàndez J, Holland PW. The ParaHox gene cluster is an evolutionary sister of the hox gene cluster. Nature. 1998;392(6679):920–2.
9. Holland PW. Evolution of homeobox genes. Wiley Interdiscip Rev Dev Biol. 2013;2(1):31–45.
10. Garcia-Fernàndez J. Hox, ParaHox, ProtoHox: facts and guesses. Heredity (Edinb). 2005;94(2):145–52.
11. Kourakis MJ, Martindale MQ. Hox gene duplication and deployment in the annelid leech Helobdella. Evol Dev. 2001;3(3):145–53.
12. Biscotti MA, Canapa A, Forconi M, Barucca M. Hox and ParaHox genes: a review on molluscs. Genesis. 2014;52(12):935–45.
13. Fröbius AC, Funch P. Rotiferan hox genes give new insights into the evolution of metazoan bodyplans. Nat Commun. 2017;8(1):9.
14. Wanninger A. Hox, homology, and parsimony: an organismal perspective. Semin Cell Dev Biol 2023.
15. Kocot KM. On 20 years of Lophotrochozoa. Organisms Divers Evol. 2016;16(2):329–43.
16. Barucca M, Canapa A, Biscotti MA. An overview of hox genes in Lophotrochozoa: evolution and functionality. J Dev Biol 2016, 4(1).
17. Luo YJ, Kanda M, Koyanagi R, Hisata K, Akiyama T, Sakamoto H, Sakamoto T, Satoh N. Nemertean and phoronid genomes reveal lophotrochozoan evolution and the origin of bilaterian heads. Nat Ecol Evol. 2018;2(1):141–51.
18. Luo YJ, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, Fujie M, Yamasaki SI, Endo K, Satoh N. The Lingula genome provides insights into

brachiopod evolution and the origin of phosphate biomineralization. Nat Commun. 2015;6:8301.

19. Gąsiorowski L, Hejnol A. Hox gene expression during development of the phoronid Phoronopsis harmeri. Evodevo. 2020;11:2.

20. Gąsiorowski L, Hejnol A. Hox gene expression in postmetamorphic juveniles of the brachiopod Terebratalia transversa. Evodevo. 2019;10:1.

21. Bock P, Gordon DP. Phylum Bryozoa Ehrenberg, 1831. Zootaxa. 2013;3703:67–74.

22. Schwaha T. Morphology of bryozoans. In: *Handbook of Zoology: Bryozoa* Edited by Schwaha T. Berlin: DeGruyter; 2020: 57–100.

23. Fuchs J, Martindale MQ, Hejnol A. Gene expression in bryozoan larvae suggest a fundamental importance of pre-patterned blastemic cells in the bryozoan life-cycle. Evodevo. 2011;2(1):13.

24. Lee PN, Callaerts P, De Couet HG, Martindale MQ. Cephalopod hox genes and the origin of morphological novelties. Nature. 2003;424(6952):1061–5.

25. Passamaneck YL, Halanych KM. Evidence from hox genes that bryozoans are lophotrochozoans. Evol Develop 2004, 6.

26. Santagata S. Genes with evidence of positive selection as potentially related to coloniality and the evolution of morphological features among the lophophorates and entoprocts. J Exp Zool B Mol Dev Evol. 2021;336(3):267–80.

27. Rayko M, Komissarov A, Kwan JC, Lim-Fong G, Rhodes AC, Kliver S, Kuchur P, O'Brien SJ, Lopez JV. Draft genome of Bugula neritina, a colonial animal packing powerful symbionts and potential medicines. Sci Data. 2020;7(1):356.

28. Kumar G, Ertl R, Bartholomew JL, El-Matbouli M. First transcriptome analysis of bryozoan Fredericella sultana, the primary host of myxozoan parasite tetracapsuloides bryosalmonae. PeerJ. 2020;8:e9027–7.

29. de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB, Balavoine G. Hox genes in brachiopods and priapulids and protostome evolution. Nature. 1999;399(6738):772–6.

30. Saadi AJ, Bibermair J, Kocot KM, Roberts NG, Hirose M, Calcino A, Baranyi C, Chaichana R, Wood TS, Schwaha T. Phylogenomics reveals deep relationships and diversification within phylactolaemate bryozoans. Proc Biol Sci. 2022;289(1986):20221504.

31. Schiemann SM, Martín-Durán JM, Børve A, Vellutini BC, Passamaneck YJ, Hejnol A. Clustered brachiopod hox genes are not expressed collinearly and are associated with lophotrochozoan novelties. Proc Natl Acad Sci U S A. 2017;114(10):E1913–e1922.

32. Waeschenbach A, Taylor PD, Littlewood DTJ. A molecular phylogeny of bryozoans. Mol Phylogenet Evol. 2012;62(2):718–35.

33. Ernst A, Schäfer P. Palaeozoic vs. post-palaeozoic Stenolaemata: phylogenetic relationship or morphological convergence? Courier Forschungsinstitut Senckenberg. 2006;257:49–64.

34. Taylor PD, Waeschenbach A. Phylogeny and diversification of bryozoans. Palaeontology. 2015;58(4):585–99.

35. Nesnidal M, Helmkampf M, Meyer A, Witek A, Bruchhaus I, Ebersberger I, Hankeln T, Lieb B, Struck T, Hausdorf B. New phylogenomic data support the monophyly of Lophophorata and an ectoproct-phoronid clade and indicate that Polyzoa and Kryptrochozoa are caused by systematic bias. BMC Evol Biol. 2013;13:1–13.

36. Laumer CE, Bekkouche N, Kerbl A, Goetz F, Neves RC, Sørensen MV, Kristensen RM, Hejnol A, Dunn CW, Giribet G, et al. Spiralian phylogeny informs the evolution of microscopic lineages. Curr Biol. 2015;25(15):2000–6.

37. Laumer CE, Fernández R, Lemer S, Combosch D, Kocot KM, Riesgo A, Andrade SCS, Sterrer W, Sørensen MV, Giribet G. Revisiting metazoan phylogeny with genomic sampling of all phyla. Proc Biol Sci. 2019;286(1906):20190831.

38. Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. A New Spiralian Phylogeny places the Enigmatic Arrow worms among Gnathiferans. Curr Biol. 2019;29(2):312–318e313.

39. Khalturin K, Shunatova N, Shchenkov S, Sasakura Y, Kawamitsu M, Satoh N. Polyzoa is back: the effect of complete gene sets on the placement of Ectoprocta and Entoprocta. Sci Adv. 2022;8(26):eabo4400.

40. Merkel J, Wanninger A, Lieb B. Novel and conserved features of the hox cluster of Entoprocta (Kamptozoa). Ournal of Phylogenetics & Evolutionary Biology. 2018;6(149):1–8.

41. Cho S-J, Vallès Y, Kim KM, Ji SC, Han SJ, Park SC. Additional duplicated hox genes in the earthworm: Perionyx excavatus hox genes consist of eleven paralog groups. Gene. 2012;493(2):260–6.

42. Fritsch M, Wollesen T, Wanninger A. Hox and ParaHox gene expression in early body plan patterning of polyplacophoran mollusks. J Exp Zool B Mol Dev Evol. 2016;326(2):89–104.

43. Kulakova MA, Cook CE, Andreeva TF. ParaHox gene expression in larval and postlarval development of the polychaete Nereis virens (Annelida, Lophotrochozoa). BMC Dev Biol. 2008;8:61.

44. Hiebert LS, Maslakova SA. 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. 2015;6:26.

45. Le Gouar M, Lartillot N, Adoutte A, Vervoort M. The expression of a caudal homologue in a mollusc, Patella vulgata. Gene Expr Patterns. 2003;3(1):35–7.

46. Altenburger A, Martinez P, Wanninger A. Homeobox gene expression in Brachiopoda: the role of not and Cdx in bodyplan patterning, neurogenesis, and germ layer specification. Gene Expr Patterns. 2011;11(7):427–36.

47. Johnson AB, Lambert JD. The Caudal ParaHox gene is required for hindgut development in the mollusc Tritia (a.k.a. Ilyanassa). Dev Biol. 2021;470:1–9.

48. Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, et al. Insights into bilaterian evolution from three spiralian genomes. Nature. 2013;493(7433):526–31.

49. Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature. 2015;524(7564):220–4.

50. Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature. 2012;490(7418):49–54.

51. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

52. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494–512.

53. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28(23):3150–2.

54. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2.

55. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.

56. Slater GS, Birney E. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics. 2005;6:31.

57. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, Lopez R. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 2022;50(W1):W276–279.

58. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.

59. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8.

60. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3.

61. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530–4.

62. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9.

63. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the Ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518–22.

64. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-NG: a New and Scalable Tool for the selection of DNA and protein evolutionary models. Mol Biol Evol. 2020;37(1):291–4.

65. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.

Publisher's Note