Fusion, fission, and scrambling of the bilaterian genome in Bryozoa

Thomas D. Lewin,[1] Isabel Jiah-Yih Liao,[1] Mu-En Chen,[1] John D. D. Bishop,[2] Peter W. H. Holland,[3] and Yi-Jyun Luo[1]*

[1]Biodiversity Research Center, Academia Sinica, Taipei, Taiwan
[2]Marine Biological Association, Plymouth, UK
[3]Department of Biology, University of Oxford, Oxford, UK

*Corresponding author: yjluo@gate.sinica.edu.tw

Running title: Genome scrambling in bryozoans

Abstract

Groups of orthologous genes are commonly found together on the same chromosome over vast evolutionary distances. This extensive physical gene linkage, known as macrosynteny, is seen between bilaterian phyla as divergent as Chordata, Echinodermata, Mollusca, and Nemertea. Here, we report a unique pattern of genome evolution in Bryozoa, an understudied phylum of colonial invertebrates. Using comparative genomics, we reconstruct the chromosomal evolutionary history of five bryozoans. Multiple ancient chromosome fusions followed by gene mixing led to the near-complete loss of bilaterian linkage groups in the ancestor of extant bryozoans. A second wave of rearrangements, including chromosome fission, then occurred independently in two bryozoan classes, further scrambling bryozoan genomes. We also discover at least five derived chromosomal fusion events shared between bryozoans and brachiopods, supporting the traditional but highly debated Lophophorata hypothesis. Finally, we show that chromosome fusion and fission processes led to the partitioning of genes from bryozoan Hox clusters onto multiple chromosomes. Our findings demonstrate that the canonical bilaterian genome structure has been lost across all studied representatives of an entire phylum; reveal that linkage group fission can occur very frequently in specific lineages; and provide a powerful source of phylogenetic information.

Introduction

Advances in genome sequencing and assembly have in recent years made it possible to compare the organization of entire genomes across diverse branches of the animal kingdom. The emerging picture is one of a common set of metazoan genes, superimposed with extensive gene duplication, gene loss, and gene novelty in each lineage, plus a surprising degree of large-scale organizational similarity. Many groups of orthologous genes have been retained on the same chromosome for hundreds of millions of years and remain together in diverse and highly divergent lineages (Putnam et al. 2007, 2008; Simakov et al. 2013, 2022; Schultz et al. 2023; Lin et al. 2024). This conservation of chromosome-scale gene linkages, known as macrosynteny, is intriguing, as the selective pressures maintaining it and its significance for genome function are still not fully understood.

The persistence of gene linkages enables the identification of orthologous chromosomes between distantly related phyla. Furthermore, it permits the reconstruction of ancestral linkage groups (ALGs)—hypothesized sets of orthologous genes located on the same chromosome or chromosome arm in ancestral species (Simakov et al. 2022; Mackintosh et al. 2023a). Lineage-specific ALGs have been identified in several animal groups, including vertebrates (Sacerdot et al. 2018; Simakov et al. 2020) and Lepidoptera (butterflies and

48 moths) (Wright et al. 2024). Most strikingly, the discovery of 24 ALGs shared between
49 spiralians and deuterostomes (Simakov et al. 2022; Marlétaz et al. 2023) revealed the genome
50 structure of the most recent bilaterian common ancestor. These highly-conserved bilaterian
51 blocks of orthologous genes are retained in species from phyla as divergent as chordates,
52 echinoderms, molluscs, and annelids (Wang et al. 2017; Simakov et al. 2022; Marlétaz et al.
53 2023; Martín-Zamora et al. 2023; Lin et al. 2024), suggesting that there are substantial
54 functional constraints to genome organization. There are lineages within phyla that have
55 experienced genomic scrambling, including the breaking up or mixing of genes from different
56 ALGs, such as planarians (Ivankovic et al. 2023), tunicates (Plessy et al. 2024) and clitellate
57 annelids (Schultz et al. 2024; Vargas-Chávez et al. 2024; Lewin et al. 2024a). However,
58 conservation of ALGs seems to be the norm at the phylum level. Despite this general
59 conservation, fluctuations in chromosome number—often leading to ALG rearrangement—are
60 a relatively frequent evolutionary process (White 1954; King 1995; Rice et al. 2015; Román-
61 Palacios et al. 2021). In animals, most of these changes are attributed to inter-chromosomal
62 rearrangements such as translocations and fusions, though they may also arise through whole
63 genome duplication (Orr 1990; Gregory and Mable 2005; Moriyama and Koshiba-Takeuchi
64 2018; Muffato et al. 2023). Following chromosome fusion, orthologous genes from different
65 ALGs may either be maintained in separate chromosome sections or shuffled and
66 interspersed along the chromosome in a process known as 'fusion-with-mixing' (Simakov et
67 al. 2022). Conversely, chromosomes can undergo fission, splitting into separate parts, but the
68 division of ALGs through this process is rare (Simakov et al. 2022; Wright et al. 2024).
69 The interplay between the stability of orthologous gene sets and the dynamic nature
70 of inter-chromosomal rearrangements shapes the genomic landscape. These
71 rearrangements, while disrupting gene function and expression (Harewood and Fraser 2014)
72 and altering recombination rates (Yoshida et al. 2023), also play a role in driving adaptation
73 (Coyle and Kroll 2007; Dunham et al. 2002; Cheng et al. 2012) and contributing to
74 macroevolutionary processes like speciation (Rieseberg 2001; de Vos et al. 2020; Berdan et
75 al. 2023; Mackintosh et al. 2023b). This dynamic quality of chromosomal changes has made
76 macrosynteny a powerful tool for phylogenetic studies (Steenwyk and King 2024). In particular,
77 fusion-with-mixing events can be particularly powerful phylogenetic markers as the intermixing
78 of genes from different ALGs is considered rare and irreversible. In light of this, macrosynteny
79 was recently used to investigate early branching metazoan lineages (Schultz et al. 2023) and
80 teleost evolution (Parey et al. 2023).
81 Even in the age of phylogenomics, when the amino acid sequences of thousands of
82 deduced proteins are aligned and analyzed, there remains a small group of 'Problematica':
83 taxa for which a robust phylogenetic placement remains tantalizingly out of reach (Jenner and
84 Littlewood 2008; Jenner et al. 2009). Many of these taxa are rare or have few species. Yet

one phylum—the Bryozoa—stands out among Problematica as highly abundant, species-rich, and ecologically important (Gordon and Costello 2016; Lombardi et al. 2020), increasingly as alien invasive species (Loxton et al. 2017; Weaver et al. 2018; Pratt et al. 2022). Bryozoans are benthic, suspension-feeding, aquatic invertebrates that are important ecosystem engineers (Wood et al. 2012; Bock and Gordon 2013). They form colonies of individual zooids, generally with calcified skeletons (Zhang et al. 2021), and each zooid can be morphologically differentiated for specialized roles in feeding, brooding, or defense. Despite their unique biological traits, bryozoans and bryozoan genomes remain relatively understudied (Orr et al. 2021; Liao et al. 2023). Several chromosome-scale assemblies have recently been made available (Hoencamp et al. 2021; Bishop et al. 2023a, 2023b; Wood et al. 2023; Bishop et al. 2024), but these are without comparative analysis, and beyond an accepted placement within Spiralia, the phylum's phylogenetic position remains unresolved (Jenner and Littlewood 2008; Jenner et al. 2009; Khalturin et al. 2022; Liao et al. 2023) (Supplemental Fig. S1).

Here, we present a comparative genomic analysis of five bryozoan genomes, focusing on genome rearrangements and chromosome evolution. Our study aims to (i) characterize the inter-chromosomal rearrangements that have occurred with the phylum Bryozoa; (ii) determine how the 24 bilaterian ALGs are organized in bryozoans; (iii) reconstruct the genome organization of the last common ancestor of extant bryozoans; and (iv) use macrosynteny to elucidate the phylogenetic position of Bryozoa.

Results

Chromosome-scale genomes of five bryozoans

To reconstruct the evolutionary history of bryozoan chromosomes, we first assembled a dataset of five chromosome-scale genomes. Four assemblies, *Bugulina stolonifera* (BST) (Wood et al. 2023), *Cryptosula pallasiana* (CPA) (Bishop et al. 2023b), *Membranipora membranacea* (MME) (Bishop et al. 2023a), and *Watersipora subatra* (WSU) (Bishop et al. 2024), are marine bryozoans in the class Gymnolaemata, which accounts for approximately 90% of described bryozoan species (Bock and Gordon 2013) (Fig. 1A). The fifth, *Cristatella mucedo* (CMU) (Hoencamp et al. 2021), is an outgroup belonging to the freshwater class Phylactolaemata (Taylor and Waeschenbach 2015). Four of the five assemblies were generated and made openly accessible by the Darwin Tree of Life (DToL) project (The Darwin Tree of Life Project Consortium et al. 2022), sequenced to chromosome-level using PacBio HiFi long reads and high-throughput chromosome conformation capture (Hi-C) scaffolding (Wood et al. 2023; Bishop et al. 2023a, 2023b, 2024).

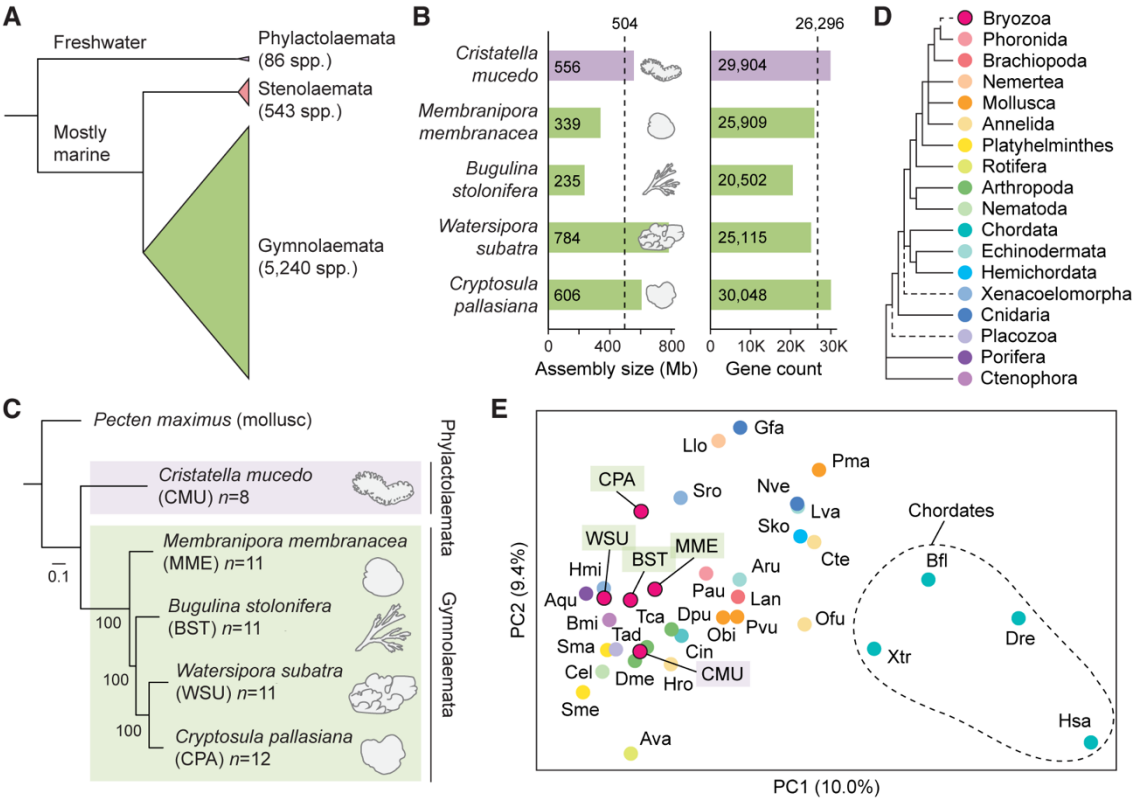

Figure 1. Comparative genomics of five chromosome-scale bryozoan genomes. (*A*) Relationships between the three classes of bryozoans. Class Phylactolaemata is the most basal, containing just 86 species, all of which inhabit freshwater environments (Bock and Gordon 2013). Stenolaemata comprises 543 species, all of which are marine. Gymnolaemata is the most species-rich class, comprising 5,240 species. (*B*) Genome size and gene model count of bryozoan genome assemblies used in this study. Dashed lines mark mean values. The clade containing *W. subatra* and *Cry. pallasiana* has larger genomes. (*C*) Tree of bryozoan species constructed with chromosome-scale genomes with the scallop *Pecten maximus* as the outgroup. The tree was constructed using the maximum likelihood method (LG+F+R6 model) with 1,000 bootstrap replicates in IQ-TREE, utilizing 994 orthologous genes. The scale bar represents the number of substitutions per site. (*D*) Phylogenetic relationships among animal phyla, with dashed lines indicating areas of uncertainty in their placement. Phylogeny based on (Telford et al. 2015) and (Liao et al. 2023). (*E*) Principal component analysis (PCA) conducted using orthologous gene group counts as determined by OrthoFinder. Points are colored by their phylum, as indicated in panel *D*. PCs 1 and 2 explain 10.0% and 9.4% of the variance in the dataset, respectively. Variance ratios for the top 30 principal components are available as Supplemental Fig. S2. Species abbreviations: Ava, *Adineta vaga*; Aqu, *Amphimedon queenslandica*; Aru, *Asterias rubens*; Bfl, *Branchiostoma floridae*; Bmi, *Bolinopsis microptera*; Cel, *Caenorhabditis elegans*; Cin, *Ciona intestinalis*; Cte, *Capitella teleta*; Dme, *Drosophila melanogaster*; Dpu, *Daphnia pulex*; Dre, *Danio rerio*; Gfa, *Galaxea fascicularis*; Hmi, *Hofstenia miamia*; Hro, *Helobdella robusta*; Hsa, *Homo sapiens*, Lan, *Lingula anatina*; Llo, *Lineus longissimus*; Lva, *Lytechinus variegatus*; Nve, *Nematostella vectensis*; Obi, *Octopus bimaculoides*; Ofu, *Owenia fusiformis*; Pau, *Phoronis australis*; Pma, *Pecten maximus*; Pvu, *Patella vulgata*; Sko, *Saccoglossus kowalevskii*; Sma, *Schistosoma mansoni*; Sme, *Schmidtea mediterranea*; Sro, *Symsagittifera roscoffensis*; Tad, *Trichoplax adhaerens*; Tca, *Tribolium castaneum*; Xtr, *Xenopus tropicalis*.

We first performed comprehensive gene prediction using the BRAKER3 pipeline, utilizing transcriptomic data as supporting evidence (Fig. 1B; Supplemental Tables S1–S3). Benchmarking Using Single-Copy Orthologs (BUSCO) (Simão et al. 2015; Manni et al. 2021) reveals a high degree of completeness in gene predictions while also indicating some losses in bryozoans, as evidenced by completeness scores of 82.7% to 88.6% from the Metazoa database (odb10) (Supplemental Table S1). Given accurate phylogenetic relationships are key to inferring chromosome rearrangement events, we built a maximum likelihood phylogenetic tree of the five bryozoan species using gene models obtained from these chromosome-scale genomes (Fig. 1C); this tree is consistent with recent phylogenetic analyses using mitochondrial and ribosomal data (Orr et al. 2022). Principal component analysis of orthologous gene counts finds bryozoans to have a gene set broadly similar to that of other lophotrochozoans (Fig. 1D,E; Supplemental Tables S4,S5).

Massive chromosome rearrangements within Bryozoa

Haploid chromosome number in our bryozoan dataset ranges from eight in *Cri. mucedo* to 12 in *Cry. pallasiana*. To investigate how the chromosomes of each species are related and understand whether large genome restructuring events have occurred within the phylum Bryozoa, we used OrthoFinder (Emms and Kelly 2019) to identify 4,242 single-copy orthologs and mapped the position of these orthologs within each genome. We then constructed synteny maps which we displayed as ribbon plots (Fig. 2) and Oxford dot plots (Supplemental Fig. S3). Throughout our analysis of chromosome rearrangements, we apply the algebraic symbol conventions established by Simakov et al., wherever applicable (Simakov et al. 2022). Specifically, we use triple bar for one-to-one relationships, also known as syntenic equivalences (≡); tensor for fusion with mixing (⊗); dot product for end-to-end fusion (●); down arrow for centric insertion (↘); and addition for fission (+) (Fig. 2A). However, although these descriptions are effective for taxa with relatively conserved genomes, we found that the highly derived nature of bryozoan genomes resulted in many chromosomes not fitting clearly into these predefined categories. Furthermore, the notation may become excessively complex and confusing when addressing chromosomes that have undergone numerous ALG fusions and have intricate evolutionary histories. We therefore also utilize narrative descriptions, especially for highly derived chromosomes.

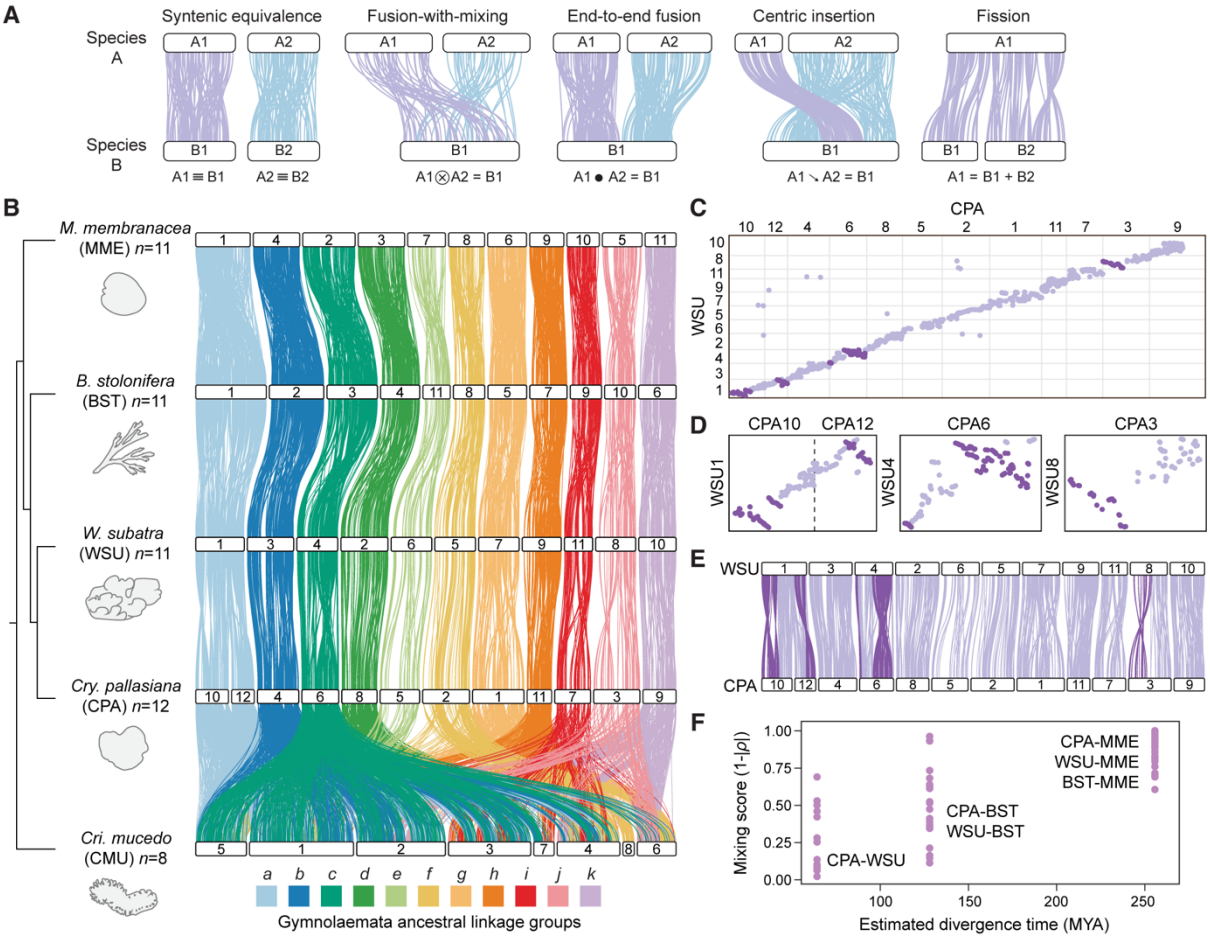

Figure 2. Chromosome rearrangement in gymnolaemate and phylactolaemate bryozoans. (*A*) Schematic representations of possible inter-chromosomal rearrangements. White bars represent chromosomes; colored lines connect orthologous genes; the same color line represents genes in the same ancestral linkage group. (*B*) Chromosome-scale gene linkage between five bryozoan species revealed by ribbon plots. Horizontal bars represent chromosomes. Vertical lines connect the genomic position of orthologous genes in each genome. Chromosomes connected by many orthologs are inferred to be homologous (i.e., orthologous chromosomes). In gymnolaemate species, 11 ALGs *a* to *k* are highly conserved, though the chromosome containing ALG *a* underwent a fission event to form chromosomes CPA10 and CPA12 in *Cry. pallasiana*. These ALGs are not conserved in the phylactolaemate species *Cri. mucedo*. Links between chromosomes sharing five or fewer genes between chromosomes are removed for clarity but can be seen in the Oxford dot plots. (*C*) Oxford dot plot revealing chromosome-scale gene linkage between the two most closely related species in our dataset, *Cry. pallasiana* and *W. subatra*. Each axis represents the entire length of the genome of one species. Gray lines separate chromosomes. Each point represents a pair of orthologs placed by their ordinal position in each genome. Strings of orthologs perpendicular to the main diagonal axis reveal clear intra-chromosomal inversions (dark purple). (*D*) Single-chromosome Oxford dot plots for *Cry. pallasiana* and *W. subatra* chromosomes with clear intra-chromosomal inversions (dark purple). (*E*) Ribbon plot for *Cry. pallasiana* and *W. subatra* highlighting intra-chromosomal inversions (dark purple) (*F*) Microsynteny mixing score versus divergence time for gymnolaemate bryozoan chromosomes. Each point represents a chromosome. Microsynteny mixing score measures how genes are shuffled within a chromosome by intra-chromosomal inversions (see

Materials and Methods): a higher score indicates more mixing. Mixing score increases with phylogenetic distance. Abbreviation: MYA, million years ago.

We used the chromosomal position of orthologous genes to identify orthologous chromosomes among the four gymnolaemate species. There is complete syntenic equivalence among three gymnolaemate bryozoans (*B. stolonifera*, *M. membranacea,* and *W. subatra*), which all have a haploid chromosome number of 11, suggesting there have been no instances of chromosome fusion or fission (Fig. 2B). Based on gene position in these species, we assigned 11 gymnolaemate ALGs, labeled ALGs *a* to *k* (Fig. 2B). The final gymnolaemate species, *Cry. pallasiana*, has undergone one fission event to diverge from this state, with the chromosome containing gymnolaemate ALG *a* (*M. membranacea* chr1 ≡ *B. stolonifera* chr1 ≡ *W. subatra* chr1) splitting to form two smaller chromosomes (*Cry. pallasiana* chr10 + chr12) (Fig. 2B). Given the reported rarity of chromosome fission events in animal evolution, we verified that this is not an assembly error by generating a Hi-C contact map, which confirmed that the fission is genuine (Supplemental Fig. S4). This aside, chromosome structure is highly conserved between these species. In contrast, a striking lack of conservation is observed when the phylactolaemate bryozoan *Cri. mucedo* is compared. None of the 11 gymnolaemate ALGs are conserved, and most are completely dispersed and scrambled across the entire *Cri. mucedo* genome (Fig. 2B). For instance, we identified four gymnolaemate ALGs on seven of eight *Cri. mucedo* chromosomes (gymnolaemate ALGs *a*, *b*, *c*, and *g*) (Supplemental Fig. S5). This represents an almost complete loss of chromosome relationships between the two groups. Overall, genome structure is highly conserved within gymnolaemate bryozoans but large inter-chromosomal rearrangements have resulted in highly divergent genome structures between members of Gymnolaemata and Phylactolaemata.

Macrosynteny studies over long evolutionary distances often overlook the precise order of genes, known as microsynteny (Renwick 1971; Passarge et al. 1999). This is because intra-chromosomal rearrangements like inversions rapidly alter gene order, leading to weak microsyntenic signals which are difficult to trace between distantly related species. Within our dataset, there is minimal conservation of gene order between bryozoans with the exception of the two most closely related species, *W. subatra* and *Cry. pallasiana* (Supplemental Fig. S3), which are estimated to have diverged 60–90 million years ago (Orr et al. 2022). Conservation of microsynteny between these species permitted us to investigate the process by which intra-chromosomal rearrangements occur. By generating Oxford dot plots for whole genomes (Fig. 2C) and individual chromosomes (Fig. 2D), we identified six clear cases of large intra-chromosomal inversions between these two species (e.g., *W. subatra* chr1 vs. *Cry. pallasiana* chr12). We noticed many inversions with a complete lack of intermixing, suggesting that these inversions involved entire chromosome arms (e.g., *W. subatra* chr8 vs. *Cry. pallasiana* chr3)

241 (Fig. 2D,E). We developed a microsynteny mixing score (see Methods) to quantify the extent
242 of intra-chromosomal gene shuffling. Between the closely related species *W. subatra* and *Cry.*
243 *pallasiana*, mixing score varies greatly between chromosomes, suggesting genes are not
244 shuffled at equal rates on different chromosomes (Fig. 2F). In addition, the mixing score
245 increases with estimated divergence time in our dataset (Fig. 2F; Supplemental Tables
246 S6,S7). Overall, the mixing of genes within a chromosome increases with the divergence time
247 between species, and occurs at least in part due to large inversions.

248 **Near-complete disruption of bilaterian ancestral linkage groups in Bryozoa**

249 To facilitate synteny analysis using bilaterian ancestral linkage groups, we created a new
250 pipeline, SyntenyFinder (see Methods). SyntenyFinder integrates OrthoFinder and RIdeogram
251 and is capable of conducting large-scale comparisons. We first validated our synteny analysis
252 methods using bilaterians with established chromosome relationships. We applied the
253 bilaterian ALG chromosome assignment to the amphioxus *Branchiostoma floridae* and found
254 it to be consistent with previous work (Simakov et al. 2022), identifying bilaterian ALG fusions
255 on chr1 (A1⊗A2), chr2 (C1⊗J2), chr3 (C2●Q), and chr4 (O1●I) . This result confirms the
256 accuracy of our ortholog identification and macrosynteny methodologies.
257 We next investigated how bryozoan chromosomes relate to the conserved bilaterian
258 genome organization. Using only the bryozoan data, it remains uncertain whether the
259 Gymnolaemata, Phylactolaemata, or neither is similar to the ancestral state of bilaterians. To
260 resolve this, in addition to the amphioxus *B. floridae* (Simakov et al. 2020), we compared the
261 macrosynteny of bryozoans with two other bilaterians with conserved genome organization:
262 the nemertean *Lineus longissimus* (Kwiatkowski et al. 2021) and the mollusc *Pecten maximus*
263 (Kenny et al. 2020). Both *L. longissimus* and *P. maximus* belong to the Lophotrochozoa, the
264 same major protostome clade as bryozoans, while the deuterostome *B. floridae* serves as an
265 outgroup. We generated a dataset of 995 single-copy orthologs and associated each ortholog
266 with one of the 24 bilaterian ALGs described by Simakov et al. (Simakov et al. 2022)
267 (Supplemental Data S1). Synteny maps show that, despite their divergence at the base of the
268 Bilateria, the genome organizations of *B. floridae*, *P. maximus*, and *L. longissimus* exhibit high
269 levels of similarity, having largely preserved the ancestral bilaterian ALGs (Fig. 3;
270 Supplemental Fig. S6; Supplemental Table S8). Our analysis revealed that *P. maximus* and
271 *L. longissimus* share four fusion events that have also been recorded in annelids and
272 brachiopods (J2⊗L, O1⊗R, O2⊗K, and H⊗Q) (Simakov et al. 2022; Martín-Zamora et al.
273 2023), suggesting that they are ancestral to lophotrochozoans. In addition, we detected
274 several lineage-specific fusion events. Specifically, we identified a fusion on chr2 (B2⊗M) in
275 *P. maximus*; and a fusion on chr2 (C1⊗G) in *L. longissimus*. Overall, genome structures in

the deuterostome *B. floridae* and lophotrochozoans *P. maximus* and *L. longissimus* retain most bilaterian ALGs alone on single chromosomes while each have a few lineage-specific fusion events. These changes are modest, considering that these species have been diverging for over half a billion years since the last common ancestor of bilaterians.

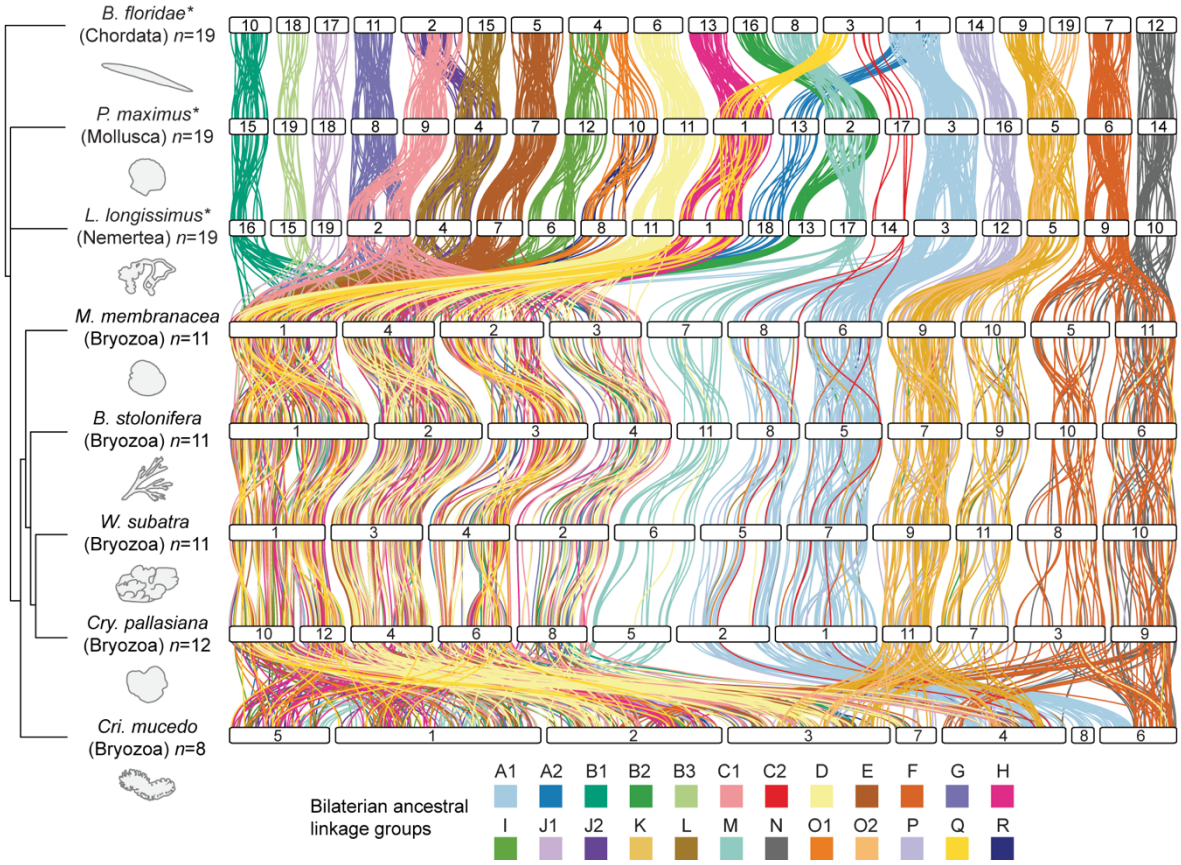

Figure 3. Extensive fusion-with-mixing of bilaterian ancestral linkage groups in bryozoans. Chromosome-scale gene linkage between five bryozoan species and three outgroup bilaterians: a nemertean, a mollusc, and a chordate. Horizontal bars represent chromosomes. Vertical lines connect the genomic position of orthologous genes in each genome. Lines are colored by the bilaterian ALGs defined by (Simakov et al. 2022). These bilaterian ALGs are highly conserved in species ranging from the amphioxus *Branchiostoma floridae* to the scallop *Pecten maximus* and the nemertean *Lineus longissimus*, as indicated by asterisks, but are extensively scrambled in bryozoans.

Like *B. floridae*, *P. maximus*, and *L. longissimus*, bryozoan genomes contain genes from all 24 bilaterian ALGs. However, in contrast to their highly conserved state in other lophotrochozoans and deuterostomes, these bilaterian ALGs have been extensively combined and mixed in the bryozoan lineage (Fig. 3; Supplemental Fig. S6; Supplemental Table S9). We found that 18 out of the 24 bilaterian ALGs (75%) have been combined and mixed across four gymnolaemate chromosomes, indicating that gymnolaemate bryozoans underwent substantial irreversible fusion-with-mixing events followed by fragmentation of a single,

298 ancestrally fused chromosome into four distinct ones. We also detect other bryozoan-specific
299 fusions: ALGs A1 and C2 initially fused (A1⊗C2) and then divided into two chromosomes
300 (e.g., *M. membranacea* chr6 and chr8); ALG P joined with the lophotrochozoan fusion K⊗O2
301 (P⊗K⊗O2) and later separated into two chromosomes (e.g., *M. membranacea* chr9 and
302 chr10); ALGs F and N fused (F⊗N) and subsequently split into two chromosomes (e.g., *M.*
303 *membranacea* chr5 and chr11). One chromosome contains only ALG M genes (e.g., *M.*
304 *membranacea* chr7), although various genes from ALG M were translocated to other
305 chromosomes. The identification of the same chromosome structure across *M. membranacea*,
306 *B. stolonifera*, *W. subatra*, and *Cry. pallasiana* confirms that the observed genome scrambling
307 is not due to assembly errors but rather represents genuine evolutionary divergence. We also
308 found that the *Cri. mucedo* genome arrangement, while unique from other bryozoans, similarly
309 does not preserve the bilaterian ALGs on separate chromosomes. Massive rearrangements
310 have independently taken place within the lineage leading to *Cri. mucedo*, resulting in the
311 dispersion of bilaterian ALGs throughout its genome (Fig. 3). Overall, the genome
312 organizations of the two classes of bryozoans studied here are distinct from each other and
313 highly divergent from the bilaterian ancestral state.

314 The presence of several ALGs on two or four chromosomes in gymnolaemate
315 bryozoans led us to question whether the splitting of ALGs was due to whole genome
316 duplication (WGD) or chromosome duplications rather than chromosome fission. We tested
317 this using two complementary methods, each using *Metaphire vulgaris*, an annelid with a
318 known WGD (Jin et al. 2020), as a positive control. We first tested for whole genome
319 duplication using synonymous substitutions per site (K_s) distributions for paralogs within each
320 genome. The large peak in the distribution for *M. vulgaris* is caused by its WGD, and the
321 absence of similar large peaks within the K_s distributions of gymnolaemate bryozoans does
322 not support the hypothesis that a recent WGD occurred in this group (Supplemental Fig. S7).
323 There may be very shallow peaks at higher K_s values (K_s = 3 to 4) in the gymnolaemate
324 bryozoans which could indicate ancient WGD in this lineage, but the signal is very weak
325 (Supplemental Fig. S7). Given this uncertainty, we then used MCScanX (Wang et al. 2012) to
326 identify and align collinear blocks of genes within each genome. The presence of duplicated
327 gene blocks on chromosomes with genes from the same ALG would suggest that they were
328 formed by duplication rather than fission. Accordingly, strong pairwise relationships were
329 identified between duplicated chromosomes in *M. vulgaris.* However, in gymnolaemate
330 bryozoans, this analysis revealed zero or few such homologous blocks of genes, which
331 provides no evidence that they are derived from WGD or whole chromosome duplication
332 events (Supplemental Fig. S8). Finally, we used a Duplicate Genes Classifier (Wang et al.
333 2012) to determine the origin of duplicates found in the genomes. While 23.5% of *M. vulgaris*
334 genes are derived from WGD/segmental duplication events, only 0.4 to 3.6% of genes in the

335 four gymnolaemate species fall into this category (Supplemental Table S10). This again lends
336 no support to the idea of a WGD in gymnolaemate bryozoans. In the phylactolaemate *Cri.*
337 *mucedo,* we identified several sections of homology between *Cri. mucedo* chromosomes (1
338 with 2; 1 with 5; 4 with 6) (Supplemental Fig. S8) and several very small very small K_s peaks
339 (Supplemental Fig. S7), suggesting possible segmental duplications in this lineage. Overall,
340 we failed to find evidence that the presence of ALGs on two or four chromosomes in
341 gymnolaemate bryozoans is the product of WGD or large-scale chromosomal duplications,
342 though this cannot be conclusively ruled out.
343

344 **The last common ancestor of extant bryozoans had a highly rearranged genome**

345 Since the Gymnolaemata and Phylactolaemata have such distinct genome arrangements, we
346 sought to uncover the genomic structure of their ancestor, the last common ancestor of extant
347 bryozoan lineages. Did it maintain the general bilaterian organization or had rearrangements
348 already occurred? We leveraged the conserved gene linkages found in modern genomes to
349 reconstruct the ancestral state by searching for fusion and fission events that are shared by
350 both gymnolaemate and phylactolaemate species. Since these two classes diverged at the
351 base of the extant Bryozoa, inter-chromosomal rearrangements found in both sets of species
352 can be inferred to have been present in their ancestor. This implicitly assumes that they did
353 not arise multiple times independently, but the complex nature of the shared rearrangements
354 makes that scenario highly improbable. This analysis revealed the likely presence of six
355 linkage groups in the ancestor of living bryozoans, suggesting an ancestral karyotype of six
356 chromosomes (designated here as α, β, γ, δ, ε, and ζ) (Fig. 4A,B). Notably, bryozoan ALGs α
357 and β comprise 10 and eight bilaterian ALGs, respectively, indicating numerous events of
358 sequential fusion-with-mixing (Fig. 4C; Supplemental Fig. S9; Supplemental Table S11). Thus,
359 chromosomal rearrangements had fused and shuffled bilaterian ALGs extensively, even
360 before the extant clades of bryozoans diverged.
361

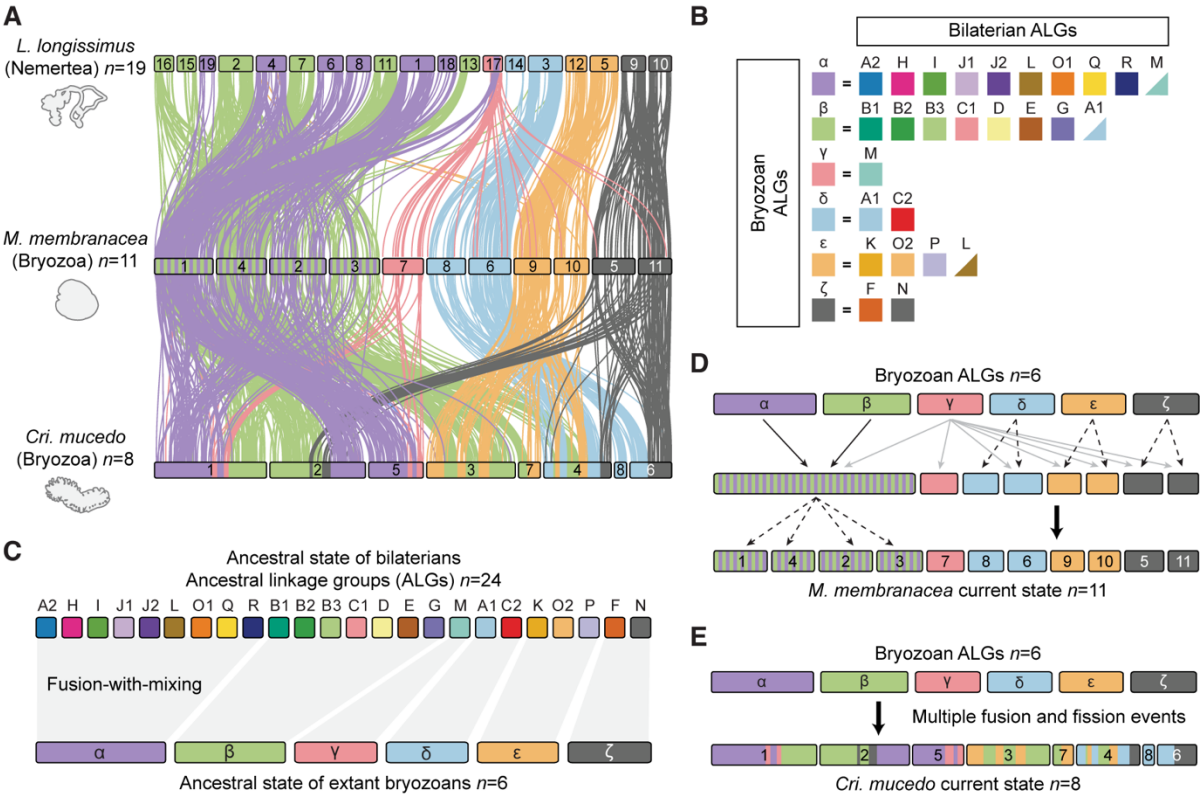

Figure 4. Reconstruction of ancestral genome architecture of extant bryozoans. (*A*) Chromosome-scale gene linkage between *L. longissimus*, *M. membranacea,* and *Cri. mucedo*. Links between orthologs are colored by their 6 bryozoan ALGs. Bryozoan ALGs are inferred from sets of genes that are co-located on the same chromosome (region) in both gymnolaemate and phylactolaemate bryozoans. *L. longissimus* is used as an outgroup because it largely preserves the 24 bryozoan ALGs on separate chromosomes but possess the four fusion events that are shared by lophotrochozoans (J2⊗L, O1⊗R, O2⊗K, and H⊗Q). (*B*) Composition of bryozoan ALGs by bilaterian ALGs, alongside the reconstructed bryozoan ancestral state (six chromosomes: α, β, γ, δ, ε, and ζ). Triangles indicate the inclusion of several genes from the bilaterian ALG but the majority of genes from that ALG are elsewhere. (*C*) Reconstruction of the genome rearrangements that formed the bryozoan ancestral state. The top row shows bilaterian ALGs, and the bottom row shows bryozoan ALGs. (*D*) Reconstruction of the genome rearrangements that led from the bryozoan ancestral state to the extant gymnolaemate genomes. Solid black arrows indicate fusion-with-mixing. Solid gray arrows indicate dispersal around the genome. Dashed arrows indicate chromosome fission. (*E*) Reconstruction of the genome rearrangements that led from the bryozoan ancestral state to the extant phylactolaemate genomes.

Based on this ancestral state of extant bryozoans, we reconstructed the events leading to the observed genome structures in contemporary species (Fig. 4D). In the gymnolaemates, the chromosomes containing ALGs α and β underwent fusion-with-mixing (α⊗β) to create one enormous chromosome containing 18 bilaterian ALGs. This chromosome then underwent sequential fission events to form four daughter chromosomes containing these 18 bilaterian ALGs (e.g., *M. membranacea* chr1 to chr4). Meanwhile, bryozoan ALGs δ, ε, and ζ each

underwent a single fission event to form two separate chromosomes. ALG γ was maintained alone on one chromosome (chr7), but genes from this ALG also dispersed to all other chromosomes (Fig. 4D). The *Cri. mucedo* genome exhibits a structure more derived from the ancestral state, characterized by numerous fission and fusion events with limited mixing, resulting in chromosomes consisting of distinct segments derived from different bryozoan ALGs (Fig. 4E). Overall, the bryozoan common ancestor had a highly derived genome, and lineage-specific fusion and fission events have further modified the genomes of both gymnolaemate and phylactolaemate bryozoans.

Shared derived chromosomal rearrangements between bryozoans and brachiopods

Recent studies have proposed the application of macrosynteny analysis to phylogeny inference, arguing that the infrequency and irreversibility of fusion-with-mixing allows these events to serve as phylogenetic markers (Schultz et al. 2023; Parey et al. 2023; Steenwyk and King 2024). In light of the contentious debate surrounding the phylogenetic position of bryozoans, we sought to determine if macrosynteny could offer insights in this regard. Two contrasting hypotheses exist: (a) the Lophophorata hypothesis, which proposes a close relationship to phoronids and brachiopods (Nesnidal et al. 2013; Luo et al. 2017; Marlétaz et al. 2019; Laumer et al. 2019), and (b) the Polyzoa hypothesis, which proposes closer affiliation to Entoprocta and Cycliophora (Hejnol et al. 2009; Kocot et al. 2017; Khalturin et al. 2022) (Supplemental Fig. S1). Although no chromosome-scale genomes are available for Cycliophora, Entoprocta, or Phoronida species, we recently sequenced a chromosome-level genome for the brachiopod *Lingula anatina* (Lewin et al. 2024b). We first validated the quality of this assembly using genome statistics and found it to be highly contiguous (Supplemental Tables S12,S13), then verified that the chromosome number was supported by cytological karyotype studies (n = 10) (Nishizawa et al. 2010). Next, we checked the assembly by comparing chromosome structure to two species with highly conserved genomes: *P. maximus* and *L. longissimus* (Supplemental Fig. S10). The recovery of many ancestral bilaterian ALGs on singular chromosomes further supports that the genome is correctly assembled (Supplemental Fig. S10).

We then conducted a macrosynteny comparison between the bryozoan *M. membranacea*, representative of Gymnolaemata, and the brachiopod *L. anatina* to test for the existence of shared derived rare genomic changes in the form of chromosome fusion-with-mixing events. We identified nine fusion-with-mixing events shared between *L. anatina* and *M. membranacea* (Fig. 5A,B). Four of these events (J2⊗L, O1⊗R, O2⊗K, and H⊗Q) are also shared with annelids, nemerteans, and molluscs (Fig. 3) (Simakov et al. 2022; Martín-Zamora

et al. 2023). The remaining five fusion-with-mixing events are exclusive to brachiopods and bryozoans, making them shared derived characteristics unique to these groups (Fig. 5C, marked by ⊗ occurring outside of parentheses). This suggests that bryozoans and brachiopods are closely related, having inherited these changes from a recent common ancestor, supporting the Lophophorata hypothesis (Fig. 5D, scenario Type A). The changes could also be attributed to convergent evolution (Fig. 5E,F, Type B) or a mixture of inheritance from a common ancestor and convergent evolution (Fig. 5G, Type C), but inheritance from a recent common ancestor is the most parsimonious explanation, requiring the fewest fusion-with-mixing events.

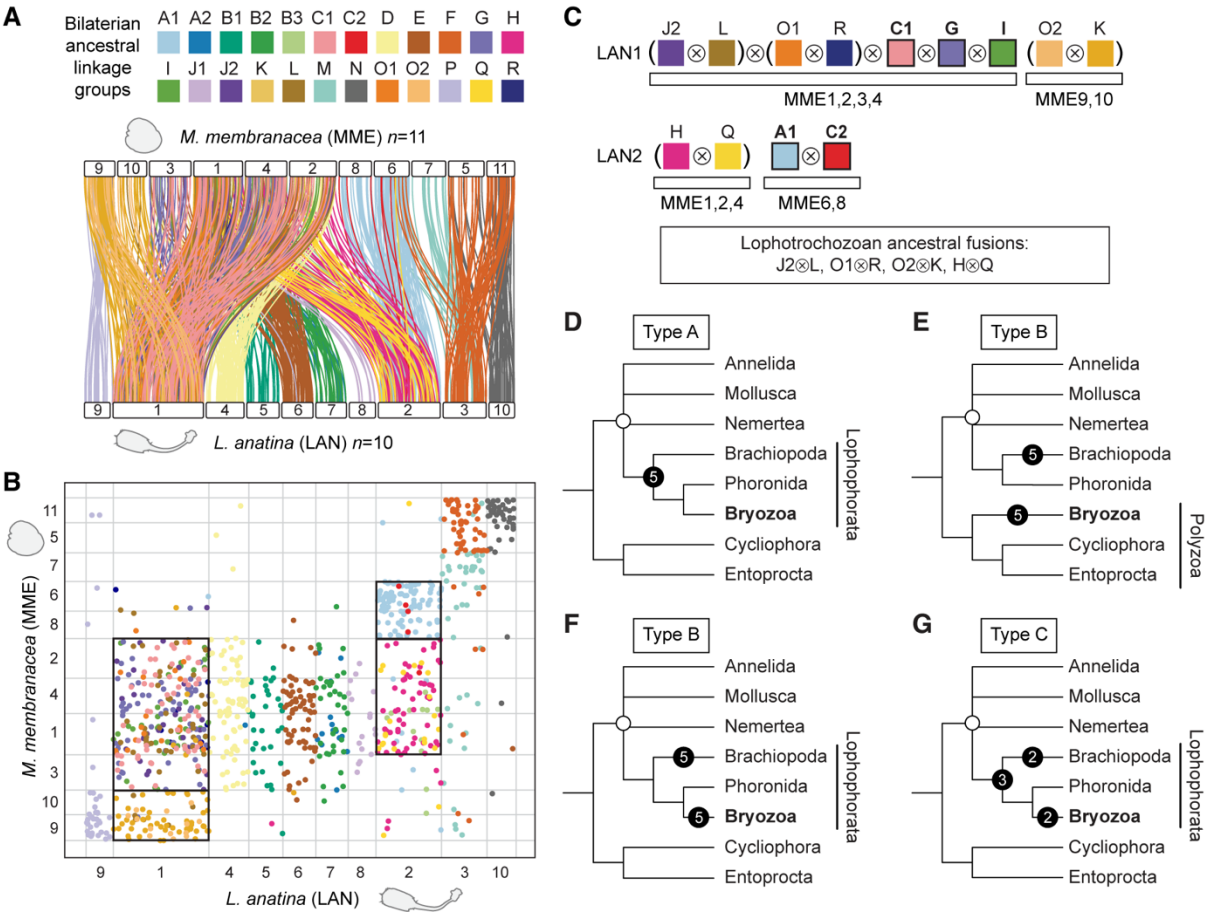

Figure 5. Shared derived chromosome features between bryozoans and brachiopods support the Lophophorata hypothesis. (*A*) Chromosome-scale gene linkage between the bryozoan *M. membranacea* and the brachiopod *L. anatina*. Horizontal bars represent chromosomes. Vertical lines connect the genomic position of orthologous genes in each genome. Lines are colored by bilaterian ALGs. (*B*) Oxford dot plot showing shared ALG fusion events between *L. anatina* and *M. membranacea*. Each point represents a pair of orthologs placed by their ordinal position in each genome, colored by their bilaterian ALG. Black boxes highlight shared fusion-with-mixing events, where two ALGs appear together on the same chromosome in both species. (*C*) There are nine shared fusion-with-mixing events between *L. anatina* and *M. membranacea*. Several spiralian phyla share four of these events ('lophotrochozoan ancestral fusions'), marked by ⊗ within parentheses, but their timing is

uncertain. These events may, therefore, not be useful phylogenetic markers for the position of bryozoans. Confirmed derived fusion-with-mixing events shared between bryozoans and brachiopods but not present in molluscs and annelids are marked by ⊗ occurring outside of parentheses. (*D*–*G*) Visual representation of four possible evolutionary scenarios that could give rise to the five derived fusion-with-mixing events shared between bryozoans and brachiopods (Types A, B, and C). An open circle marks the base of the Lophotrochozoa. Black circles mark chromosome fusion-with-mixing events. (*D*) Scenario Type A: Brachiopods, phoronids, and bryozoans are closely related, forming the Lophophorata. The five fusion-with-mixing events shared between bryozoans and *L. anatina* occurred in the ancestor of the Lophophorata. Total number of fusions required: 5. (*E*) Scenario Type B-1: Bryozoans, Cycliophora, and Entoprocta are closely related, forming the Polyzoa. Bryozoans are distantly related to brachiopods. The five fusion-with-mixing events shared between bryozoans and *L. anatina* occurred independently in bryozoans and brachiopods. Total number of fusions required: 10. (*F*) Scenario Type B-2: Brachiopods, phoronids, and bryozoans are closely related, forming the Lophophorata. The five fusion-with-mixing events shared between bryozoans and *L. anatina* occurred independently in bryozoans and brachiopods. Total number of fusions required: 10. (*G*) Scenario Type C: Brachiopods, phoronids, and bryozoans are closely related, forming the Lophophorata. Some of the five fusion-with-mixing events shared between bryozoans and *L. anatina* occurred in the ancestor of Lophophorata, while some occurred independently in bryozoans and brachiopods. Total number of fusions required: 6–9.

Under the assumption that fusion-with-mixing events are irreversible, the Type A scenario would be contradicted if any bryozoan, phoronid, or brachiopod were discovered without the described fusion-with-mixing events. To test this, we examined whether the same fusion events are present in *Cri. mucedo*, the most divergent bryozoan in our dataset. Despite *Cri. mucedo* having a highly divergent chromosome organization, all five derived fusion-with-mixing events shared by *M. membranacea* and *L. anatina* also occur together on the same chromosomes in *Cri. mucedo* (Supplemental Fig. S11). Altogether, we observed nine shared fusion chromosome events between bryozoans and brachiopods, with at least five being derived fusion-with-mixing events and thus considered potential phylogenetic markers. These events are notably absent in molluscs, annelids, and nemerteans, which we interpret as indicating a close relationship between bryozoans and brachiopods, supporting the Lophophorata hypothesis. In the future, sequencing a phoronid genome to chromosome level will be a key test of this hypothesis, revealing whether the same fusions are found in this phylum.

Genome structure and Hox gene evolution

In bilaterians, Hox genes are essential for body patterning and cell type specification (Hubert and Wellik 2023), a role facilitated by their organization into tight clusters (Akam 1989; Graham

et al. 1989; Duboule and Dollé 1989). We explored whether extensive genome scrambling has impacted Hox cluster organization in bryozoans. We annotated the Hox genes of the five bryozoan species (Supplemental Fig. S12; Supplemental Data S2) and identified extensive gene loss in the bryozoan Hox gene set, which lacks *Hox1*, *Hox3*, *Scr*, *Antp*, *Lox2*, and *Post1* in all five species. Strikingly, the remaining Hox genes are distributed across two separate chromosomes in every assembly (Fig. 6A). In bilaterians such as *P. maximus* and *L. longissimus*, the Hox cluster is associated with the bilaterian ALG B2. This ALG is fragmented across four chromosomes (five in *Cry. pallasiana*) in each bryozoan genome, as a result of successive fusion and fission events (Fig. 6B). The main Hox cluster is located on orthologous chromosomes containing ALG B2 genes in each bryozoan genome assembly, while Hox genes separated from the cluster are also on a different orthologous chromosome that contains ALG B2 genes (Fig. 6C). This evidence suggests that the presence of Hox genes across two chromosomes in bryozoans is a unique case of partial cluster tandem duplication (*Hox4*, *Lox5*, *Lox4*, and *Post2* genes) followed by separation due to chromosome fission (Fig. 6D).

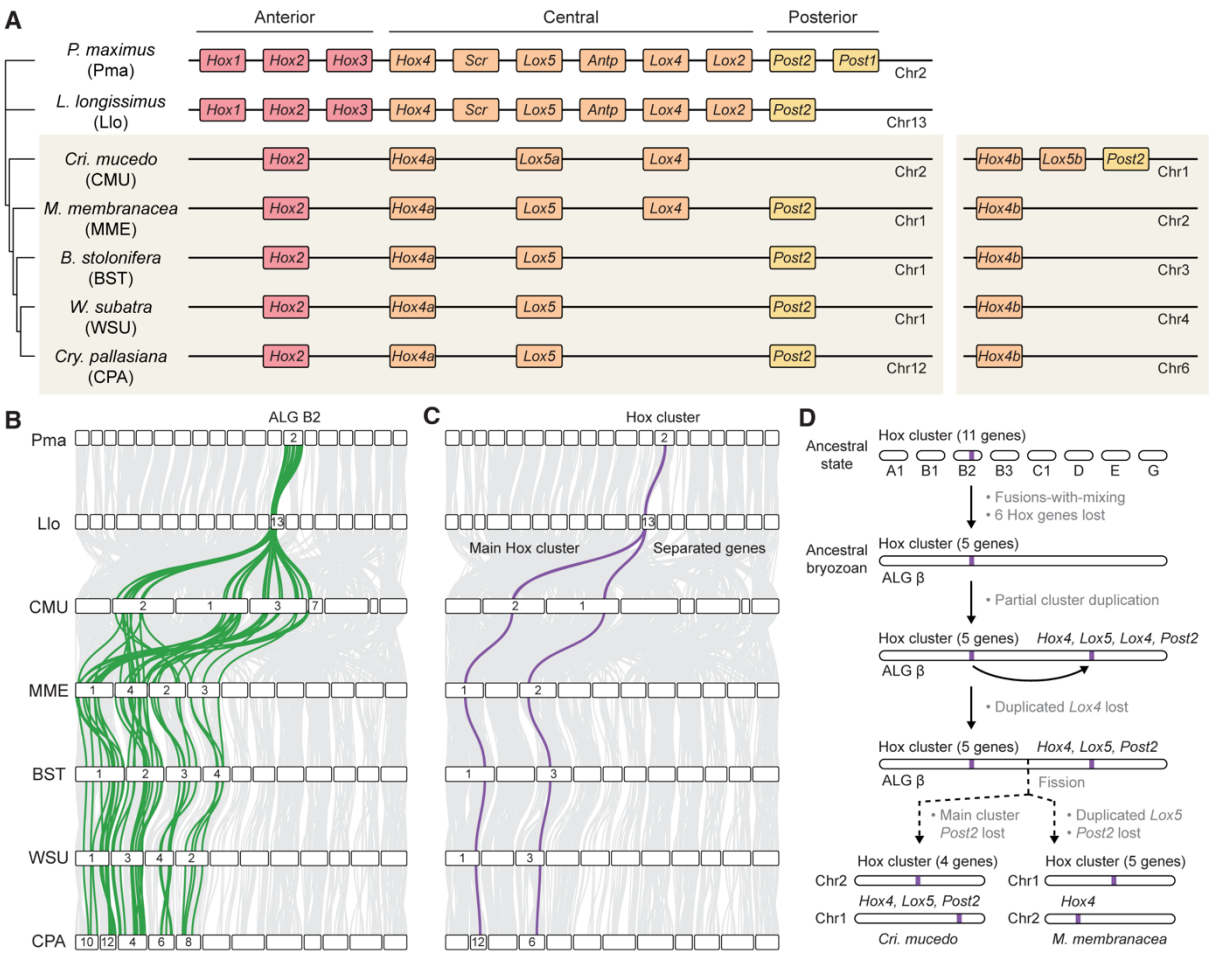

Figure 6. Evolution of the bryozoan Hox cluster. (*A*) Left: Representation of the main Hox cluster of bryozoans and two outgroups, the nemertean *L. longissimus* and the mollusc *P.*

503 *maximus*. Colored rectangles represent Hox genes. Horizontal black lines represent
504 chromosomes. Right: Separate Hox genes located in different chromosomes. (*B*) Positions
505 of ALG B2 genes (green) in bryozoan genomes. (*C*) Positions of Hox genes (purple) in
506 bryozoan genomes. (*D*) Representation of the evolutionary events that resulted in observed
507 bryozoan Hox gene arrangements. In the bryozoan ancestral lineage, chromosomes
508 containing bilaterian ALGs A1, B1, B2 (including the Hox cluster), B3, C1, D, E, and G
509 underwent sequential fusion events to form bryozoan ALG β; the Hox cluster underwent a
510 partial duplication, copying genes *Hox4*, *Lox5*, *Lox4*, and *Post2*; and the duplicated *Lox4*
511 gene was subsequently lost. In the lineage leading to *Cri. mucedo* (Phylactolaemata), the
512 main cluster *Post2* was lost, and the chromosome containing ALG β underwent fission,
513 leaving the main cluster on chr2 and duplicated genes on chr1. In the lineage leading to *M.*
514 *membranacea* (Gymnolaemata), the duplicated *Lox5* and *Post2* were lost, and the
515 chromosome containing ALG β underwent fission, leaving the main cluster on chr1 and
516 duplicated genes on chr2.
517

518 Given the critical dependence of Hox gene regulation on genomic organization, we
519 hypothesized that physical separation might alter gene expression. To investigate this, we
520 analyzed the expression of bryozoan Hox paralogs located on different chromosomes using
521 published RNA-sequencing data (Supplemental Fig. S13; Supplemental Table S14). In *Cri.*
522 *mucedo*, both the *Hox4b* and *Lox5b* genes separated from the main cluster are expressed
523 more highly than their main cluster counterparts, *Hox4a* and *Lox5a* (paired *t*-tests $p = 0.0009$
524 and $p = 0.0001$, respectively) (Supplemental Fig. S13). Therefore, the physical separation of
525 Hox cluster genes by genome rearrangements is associated with differences in expression
526 levels.
527

528 **Discussion**

529 Our study reveals dramatic rearrangements to bryozoan chromosomes that have extensively
530 fragmented the 24 ALGs inherited by all bilaterians. This presents a stark contrast to reports
531 of ALG conservation from other spiralians (Wang et al. 2017; Martín-Zamora et al. 2023),
532 deuterostomes (Lin et al. 2024), and even non-bilaterians like cnidarians and ctenophores
533 (Simakov et al. 2022; Schultz et al. 2023). We distinguish two distinct periods in which the
534 bryozoan genome was drastically remodelled: the first generating the derived bryozoan
535 ancestral state of just six ALGs from the 24 bilaterian ALGs, largely by chromosome fusion
536 events, and the second independently modifying this state further in different bryozoan
537 classes, including ALG fission events.

538 Three aspects of bryozoan genomes stand out as particularly remarkable. Firstly, the
539 magnitude of chromosome fusions is notable, with 18 out of 24 (75%) bilaterian ALGs on the
540 same chromosomes. Secondly, the occurrence of these massive fusion events across a
541 phylum is exceptionally rare. Unlike dipterans and clitellate annelids, where such extensive

542 fusion events are specific to certain lineages (Zdobnov et al. 2005; Simakov et al. 2022;
543 Schultz et al. 2024; Lewin et al. 2024a; Vargas-Chávez et al. 2024), our data indicate that
544 these rearrangements are present across the entire extant bryozoan phylum. This suggests
545 that massive rearrangements had already occurred by the time of the last common ancestor
546 of living lineages. However, we note that much of bryozoan diversity is represented only in the
547 fossil record. Six of the seven orders of Stenolaemata are long extinct and less than 40% of
548 bryozoan genera have living representatives (Taylor and Waeschenbach 2015), while
549 approximately 70% of described bryozoan species are fossil (Gordon and Costello 2016).
550 Thirdly, many ALGs in bryozoans have been separated by ALG fissions. Major studies on
551 animal chromosome evolution suggested that a general rule in metazoans is that ALGs are
552 combined but not broken up (Simakov et al. 2022; Wright et al. 2024). In contrast, we identified
553 numerous fissions occurring in the common ancestor of the four gymnolaemate species
554 around 350 million years ago (Orr et al. 2022), with subsequent, lineage-specific fission in *Cry.*
555 *pallasiana* less than 90 million years ago (Orr et al. 2022). This indicates a prolonged period
556 during which fissions were evolutionarily tolerated, challenging the notion that they occur in
557 brief, time-limited periods and are followed by evolutionary stasis (Simakov et al. 2022).
558 Why have bryozoans specifically undergone radical ALG fissions while many animal
559 phyla retained the conserved ancestral state? Although our study doesn't offer a definitive
560 answer, we can apply Mayr's classical framework (Mayr 1961) to outline the elements of a
561 possible explanation. Firstly, we must identify the mechanism that generated these
562 rearrangements (the proximate cause). Secondly, we must understand why these
563 rearrangements conferred a fitness advantage and were thus favored by selection (the
564 ultimate cause). The high level of rearrangement in bryozoans could be due to either (i) an
565 increase in the rate of rearrangement events (the proximate cause), (ii) an increase in the rate
566 of fixation of rearrangement events (the ultimate cause), or both. Loss of DNA repair
567 machinery (Vargas-Chávez et al. 2024), transposable element invasions (Ahola et al. 2014;
568 Höök et al. 2023), or environmental mutagenic agents (Schultz et al. 2024) have been
569 proposed as potential proximate causes other systems, though additional studies are needed
570 to test whether these were contributors in bryozoans.
571 The ultimate cause of large-scale genome restructuring events is particularly intriguing
572 from an evolutionary standpoint. One possibility is that rearrangements are selectively neutral
573 or weakly deleterious but are fixed at an increased rate due to genetic drift (Wright 1941;
574 Mackintosh et al. 2023c). Notably, bryozoans are colonial and have an asexual stage in which
575 they reproduce by budding (Schwaha 2020), and genetic drift and mutational load are elevated
576 in asexual populations (Haag and Roze 2007). Alternatively, rearrangements may be adaptive
577 and favored by positive selection, as has been suggested for loss of genome structure during
578 the marine–freshwater transition in clitellate annelids (Sun et al. 2021; Schultz et al. 2024;

579 Vargas-Chávez et al. 2024; Lewin et al. 2024a). The transition to obligate colonialism is a
580 significant evolutionary innovation in bryozoans (Schwaha et al. 2020) that correlates with the
581 timing of genome restructuring, but further work is required to understand whether this, or
582 other bryozoan adaptations such as biomineralization (Taylor et al. 2015), were facilitated by
583 chromosomal rearrangement events.

584 A third possibility is that increased rate of fixation of rearrangements results from the
585 relaxation in bryozoans of strong evolutionary pressure that has been maintained in other
586 lineages. The role of conserved macrosyntenic gene linkages in genome functioning remains
587 largely unknown, but their high-fidelity maintenance in diverse lineages for over half a billion
588 years (Simakov et al. 2022) implies at least some importance. One idea is that this is due to
589 long-range interactions between distant elements on the same chromosome, which means
590 that gene linkages are significant to gene expression (Tolhuis et al. 2002; Miele and Dekker
591 2008; Dean 2011). A change in the relative importance of mechanisms of gene regulation may
592 therefore reduce the emphasis on long-range interactions and render conserved genome
593 organization obsolete. Our work highlights the need for studies relating gene regulation to
594 genome organization, which may be key to understanding the ultimate causes of the evolution
595 of highly derived genomic architectures.

596 Having elucidated the organization of bryozoan genomes, we questioned whether
597 macrosynteny could resolve their phylogenetic position. Substantial evidence has been
598 presented supporting both their association with Phoronida and Brachiopoda (Lophophorata)
599 (Nesnidal et al. 2013; Luo et al. 2017; Marlétaz et al. 2019; Laumer et al. 2019) and with
600 Entoprocta and Cycliophora (Polyzoa) (Hejnol et al. 2009; Kocot et al. 2017; Khalturin et al.
601 2022). Our study identifies nine fusion-with-mixing events shared between bryozoans and
602 brachiopods, of which five are confirmed as derived characters, supporting the Lophophorata
603 topology. Fusion-with-mixing events are essentially irreversible in evolution, but convergent
604 evolution cannot be ruled out (Steenwyk and King 2024). If these rare genomic changes are
605 specific to Lophotrochozoans (excluding Entoprocta and Cycliophora), then the presence of
606 at least nine shared derived fusion events in brachiopods and bryozoans would almost rule
607 out convergence, becoming a conclusive piece of evidence in this ongoing debate. The future
608 availability of chromosome-scale genomes from these groups is expected to provide clarity.
609 Looking further, evidence from macrosynteny is likely to become increasingly prominent in
610 discussions concerning phylogenetically challenging clades.

611 Overall, though macrosyntenic gene linkages are conserved across diverse bilaterian
612 lineages, members of the phylum Bryozoa possess highly derived genomes with scrambling
613 of these gene sets. Divergence has continued within the phylum, leading to distinct
614 organizational structures in two bryozoan classes, and ALG fission, previously
615 underappreciated in its importance, has occurred with high frequency. The addition of

616 information on the third bryozoan class (the Stenolaemata, represented by present-day
617 cyclostomes) when chromosome-level genomes become available will add considerably to
618 this picture. Furthermore, although we present comparative genomics data on Hox gene
619 duplications, additional studies on the spatial expression patterns of bryozoan Hox genes are
620 necessary to examine how genome rearrangements have affected their functions (Fuchs et
621 al. 2011). We expect the bryozoan genomes annotated in this study to be crucial for the pursuit
622 of such future research avenues.
623

624 **Methods**

625 **Assembly acquisition and gene prediction**

626 Chromosome-scale assemblies of *B. stolonifera* (GenBank accession no. GCA_935421135)
627 (Wood et al. 2023), *Cry. pallasiana* (GCA_945261195) (Bishop et al. 2023b), *M.*
628 *membranacea* (GCA_914767715) (Bishop et al. 2023a), and *W. subatra* (GCA_963576615)
629 (Bishop et al. 2024) were sequenced as part of the Darwin Tree of Life project (The Darwin
630 Tree of Life Project Consortium et al. 2022) and retrieved from the NCBI Datasets resource.
631 A chromosome-scale assembly of *Cri. mucedo* (Hoencamp et al. 2021) was obtained from the
632 NCBI Gene Expression Omnibus (GSE169088). None of the bryozoan assemblies used have
633 publicly available gene models, so gene prediction was performed in-house. Repetitive
634 elements were first annotated with RepeatModeler (v2.0.4) (Flynn et al. 2020) and masked
635 using RepeatMasker (v4.1.5) in sensitive mode (Smit et al. 2015). Gene prediction and
636 annotation were performed using the BRAKER3 pipeline (v3.0.3), incorporating hints from
637 RNA sequencing (RNA-seq) data as aligned reads in bam format (Stanke et al. 2006, 2008;
638 Li et al. 2009; Barnett et al. 2011; Lomsadze et al. 2014; Buchfink et al. 2015; Hoff et al. 2016,
639 2019; Gabriel et al. 2024).

640 BRAKER3 uses RNA-seq and genome data to train the gene prediction tools
641 AUGUSTUS and GeneMark, producing highly supported gene transcripts. Raw RNA-seq
642 reads for gene prediction (Supplemental Table S2) were sourced from the NCBI Sequence
643 Read Archive (SRA) using the SRA Toolkit (v3.0.0) (Leinonen et al. 2011) with GNU parallel
644 (Tange 2021). Transcriptomic data for *B. stolonifera*, *Cry. pallasiana*, *M. membranacea*, and
645 *W. subatra* were generated from adult colonies collected by the Marine Biological Association,
646 Plymouth, UK, in association with the Darwin Tree of Life project. For these four species, RNA
647 extraction was performed on the same colony used for DNA sequencing to generate the
648 reference assembly. For *B. stolonifera* and *M. membranacea,* this dataset was supplemented
649 with independent datasets from developmental stages (Supplemental Table S2). RNA-seq
650 data for *Cri. mucedo* was derived from mature adult colonies from a population different to that

651 from which the genome assembly was produced (Supplemental Table S2). Quality control for
652 RNA-seq data was performed with FastQC (Andrews and Others 2010). Reads were trimmed
653 with Trimmomatic (v0.39) (Bolger et al. 2014) and aligned to the genome with STAR (v2.7.10b)
654 (Dobin et al. 2013). BUSCO (v5.4.7) (Simão et al. 2015; Manni et al. 2021) was used to
655 evaluate the quality of gene prediction. Hox genes were identified with reciprocal tBLASTn
656 searches (Camacho et al. 2009), and orthology was confirmed using gene trees built in IQ-
657 TREE (v2.2.2.3) (Minh et al. 2020). Hox annotations are consistent with previous work, where
658 available (Saadi et al. 2023).
659
660 **Phylogenetics**
661 Single-copy orthologs were identified using OrthoFinder (v2.5.4) (Emms and Kelly 2019).
662 Alignments for each set of orthologs were created with MAFFT (v7.520) (Katoh et al. 2002;
663 Katoh and Standley 2013), trimmed with ClipKIT (v1.4.1) (Steenwyk et al. 2020), and
664 concatenated with PhyKIT (v1.11.7) (Steenwyk et al. 2021). IQ-TREE (v2.2.2.3) (Minh et al.
665 2020) was used to construct maximum likelihood trees with automated substitution model
666 selection using ModelFinder (Kalyaanamoorthy et al. 2017) and 1,000 UFBoot2 ultra-fast
667 bootstrap replicates (Hoang et al. 2018). Trees were rendered with iTOL (Letunic and Bork
668 2021). A tree of Hox protein sequences was generated using the same IQ-TREE method.
669
670 **Macrosynteny analysis**
671 We developed SyntenyFinder, a pipeline comprising a collection of Python and R scripts
672 designed to facilitate macrosynteny analysis using chromosome-level genome assemblies.
673 SyntenyFinder is a straightforward pipeline with two main functions: first, it uses OrthoFinder
674 (v2.5.4) (Emms and Kelly 2019) to identify single-copy orthologs in a dataset of protein
675 sequence files. Second, it utilizes a custom R script to generate chromosome idiograms using
676 RIdeogram (v0.2.2) (Hao et al. 2020) and to create dot plots showing the positions of
677 orthologous genes between two genomes. SyntenyFinder is made available on Github
678 (https://github.com/symgenoevolab/SyntenyFinder) and can be run in two modes: either
679 automatically downloading and running on annotated genomes from NCBI by their accessions,
680 or on local files parsed directly to the program. It requires three types of file as an input: (1)
681 genome sequence FASTA file; (2) gene positions in the format of a GTF file; and (3) a protein
682 sequence FASTA file.
683 For the purposes of this paper, two runs were performed: first with only the five
684 bryozoan species, and second adding *B. floridae* (Chordata), *P. maximus* (Mollusca), *L.*
685 *longissimus* (Nemertea) and *L. anatina* (Brachiopoda). In both cases, only orthologs which
686 were single-copy in all study species were used for downstream analysis. No gene loss or
687 duplication was permitted. In the second case, to track the genomic position of bilaterian

688 ALGs, we used an input dataset of the genes used to define the 24 bilaterian ALGs (Simakov
689 et al. 2022) and used OrthoFinder to find orthologs of these genes in bryozoan genomes. The
690 restriction to genes with known ALG relationships reduces the noise within the dataset.
691 Macrosynteny plots for this manuscript were created in R (v4.3.0) (R Core Team 2023). The
692 same set of genes is used for each comparison in both idiograms and dot plots of
693 macrosynteny. In idiograms, links between chromosomes with fewer than five shared genes
694 were removed for visual clarity. These represent single-gene or few-gene translocations. All
695 genes, including translocated genes, are shown in dot plots. For plotting purposes and to
696 ensure alignment with orthologous chromosomes, the orientations of chr1, chr2, chr4, chr6,
697 chr7, chr10, and chr11 in *W. subatra*, and chr2, chr3, chr7, chr9, chr10, chr11, chr12 in *Cry.*
698 *pallasiana* were reversed from their orientations as presented in published datasets.
699
700 **Examination of large-scale duplications**
701 Two complementary methods were used to determine whether segmental duplication or whole
702 genome duplication (WGD) events have occurred in the bryozoan lineage. In each case, we
703 use the annelid *M. vulgaris* as a positive control for a species with a known WGD (Jin et al.
704 2020; Lewin et al. 2024a; Vargas-Chávez et al. 2024). First, synonymous substitutions per
705 site (K_s) distributions were calculated with wgd (v.2.0.26) (Chen et al. 2024). Briefly, paralogs
706 were identified with 'wgd dmd' using Diamond (v2.0.14.152) (Buchfink et al. 2015, 2021) and
707 then 'wgd ksd' was run to calculate the K_s distribution using PAML (v4.9) (Yang 2007). The K_s
708 between two paralogs is an estimate of their divergence time. The null expectation with no
709 WGD is that the number of paralogs decreases exponentially as K_s increases. When a WGD
710 occurs, an excess of duplicates are created simultaneously, creating a peak in the distribution
711 (Tiley et al. 2018). Paralogs situated in collinear blocks in the genome, so-called 'anchor
712 duplicates', are more likely to be produced by WGD than dispersed paralogs and are therefore
713 more reliable for WGD detection (Chen et al. 2024). Anchor duplicates were identified using
714 'wgd syn' to implement the collinearity search algorithm of i-ADHoRe (v3.0.01) (Proost et al.
715 2011). K_s distributions using only anchor duplicates were then plotted with 'wgd syn' (Chen et
716 al. 2024).
717 Second, MCScanX (Wang et al. 2012) was used to identify collinear blocks of
718 duplicated genes within bryozoan genomes and to classify all duplicates within each genome
719 as dispersed, proximal, tandem or WGD/segmental. The presence of large homologous
720 regions within a genome containing collinear blocks of duplicated genes is suggestive of WGD.
721 The MCScanX output was plotted as a Circos plot (Krzywinski et al. 2009) using shinyCircos
722 (v2.0) (Wang et al. 2023).
723
724 **Validation of chromosome fission**

Given the absence of a high-quality Hi-C contact map for *Cry. pallasiana* (Bishop et al. 2023b), the fission event in this species remained uncertain. To validate the fission event that led to the formation of chromosomes CPA10 and CPA12, as shown in Fig. 2, we examined chromatin interactions by mapping Hi-C data to the chromosome-level assembly. Hi-C reads for *Cry. pallasiana* were downloaded from the NCBI SRA (accession ERR9866429) using the SRA Toolkit. The SRA file was fetched with prefetch and converted to FASTQ format using fasterq-dump. The Hi-C reads were then aligned to the genome assembly (accession GCA_945261195.1) using BWA (v0.7.17-r1198) (Li 2013), following the Arima Hi-C mapping pipeline (A160156 v03). A Hi-C contact map was generated using YaHS (v1.2a.1) (Zhou et al. 2023) and JuicerTools (v1.9.9), and visualized with JuiceBox (v1.11.08) (Durand et al. 2016).

Quantification of microsynteny mixing score

To assess the amount of intra-chromosomal gene shuffling, we first calculated Spearman's rank correlation coefficient (ρ). This coefficient quantifies the strength and direction of association between the ranked order of orthologous genes along orthologous chromosome pairs. The Spearman's coefficient is given by the following equation:

$$\rho = 1 - (6 * \Sigma d_i^2) / [n(n^2 - 1)] \quad (1)$$

where ρ is the coefficient, d_i represents the difference in ranks between corresponding values of the two variables, and n is the total number of observations. Subsequently, we defined the synteny mixing score, an indicator of chromosomal rearrangement, using the equation:

$$1 - |\rho| \quad (2)$$

where $|\rho|$ denotes the absolute value of the Spearman's coefficient. This metric allows the synteny mixing score to vary from 0, indicative of no intra-chromosomal rearrangement and thus a high degree of conserved microsynteny, to 1, representative of maximum rearrangement, signifying considerable disruption in the original gene order.

Estimation of species divergence time

A phylogenetic tree, constructed above from single-copy orthologs extracted from five bryozoan genomes and a mollusc outgroup (*P. maximus*), was used to estimate divergence times. We calibrated this tree with contemporary molecular and fossil data (Orr et al. 2022), focusing on the divergence between *P. maximus* and *Cri. mucedo* during the Cambrian Period (541 to 488 million years ago) (Erwin et al. 2011), employing the ape package in R (Paradis and Schliep 2018).

762

763 **Quantification of gene expression**

764 Gene expression was quantified in a published *Cri. mucedo* (PRJNA594616) RNA-seq
765 datasets. Raw reads were obtained from the NCBI SRA using the SRA toolkit (Leinonen et al.
766 2011) and GNU parallel (Tange 2021). Hox gene expression was quantified with Salmon
767 (v1.10.2) (Patro et al. 2017). *Hox4a*, *Hox4b*, *Lox5a*, and *Lox5b* gene expression levels were
768 compared using *t*-tests with a Benjamini-Hochberg correction for multiple testing.

769

770 **Data access**

771 Gene models for all five bryozoan species have been deposited in Dryad
772 (https://doi.org/10.5061/dryad.76hdr7t3f). Custom Python and R scripts used for
773 macrosynteny analysis, calculating mixing scores, and estimating divergence times are
774 available in our GitHub repository (https://github.com/symgenoevolab/SyntenyFinder) and as
775 Supplemental Code.

776

777 **Competing interest statement**

778 The authors declare no competing interests.

779

780 **Acknowledgments**

781 This work was funded by a Royal Society Newton International Fellowship (NIF\R1\201315)
782 and an Academia Sinica Career Development Award (AS-CDA-112-L06) to Y.-J.L. We
783 acknowledge the efforts in sample collection by the Marine Biological Association (MBA)–
784 Darwin Tree of Life collecting team: Patrick Adkins, Robert Mrowicki, Joanna Harley, and
785 Helen Jenkins, as well as Christine Wood of the MBA. We thank the members of the Symbiosis
786 Genomics & Evolution Lab for their assistance and support.

787

788 **Author contributions**

789 T.D.L. and Y.-J.L. designed research. T.D.L., I.J.-Y.L., M.-E.C., and Y.-J.L. performed
790 research. T.D.L., I.J.-Y.L., J.D.D.B., and Y.-J.L. contributed new reagents/analytic tools. T.D.L.
791 and Y.-J.L. analyzed data. T.D.L., P.W.H.H., and Y.-J.L. wrote the paper. All authors read and
792 approved the final manuscript.

793

794 **References**

795 Ahola V, Lehtonen R, Somervuo P, Salmela L, Koskinen P, Rastas P, Välimäki N, Paulin L,
796 Kvist J, Wahlberg N, et al. 2014. The Glanville fritillary genome retains an ancient
797 karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat Commun* **5**:
798 4737.

799 Akam M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell*
800 **57**: 347–349.

801 Andrews S, Others. 2010. FastQC: a quality control tool for high throughput sequence data.

802 Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. BamTools: a C++
803 API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**: 1691–1692.

804 Berdan EL, Aubier TG, Cozzolino S, Faria R, Feder JL, Giménez MD, Joron M, Searle JB,
805 Mérot C. 2023. Structural variants and speciation: Multiple processes at play. *Cold*
806 *Spring Harb Perspect Biol*. http://dx.doi.org/10.1101/cshperspect.a041446.

807 Bishop J, Adkins P, Wood C, Jenkins H, Marine Biological Association Genome Acquisition
808 Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life
809 programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,
810 Tree of Life Core Informatics collective, Darwin Tree of Life Consortium. 2023a. The
811 genome sequence of the sea mat, *Membranipora membranacea* (Linnaeus, 1767).
812 *Wellcome Open Res* **8**: 38.

813 Bishop J, Wood C, Adkins P, Jenkins H, Marine Biological Association Genome Acquisition
814 Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life
815 programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective,
816 Tree of Life Core Informatics collective, Darwin Tree of Life Consortium. 2023b. The
817 genome sequence of an encrusting bryozoan, *Cryptosula pallasiana* (Moll, 1803).
818 *Wellcome Open Res* **8**: 128.

819 Bishop J, Wood CA, Marine Biological Association Genome Acquisition Lab, Darwin Tree of
820 Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples
821 and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing
822 Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life
823 Core Informatics collective, Darwin Tree of Life Consortium. 2024. The genome
824 sequence of the red ripple bryozoan, *Watersipora subatra* (Ortmann, 1890). *Wellcome*
825 *Open Res* **9**: 458.

826 Bock PE, Gordon DP. 2013. Phylum Bryozoa Ehrenberg, 1831. *Zootaxa* **3703**: 67–74.

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

829 Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale
830 using DIAMOND. *Nat Methods* **18**: 366–368.

831 Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND.
832 *Nat Methods* **12**: 59–60.

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

835 Cheng C, White BJ, Kamdem C, Mockaitis K, Costantini C, Hahn MW, Besansky NJ. 2012.
836 Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-
837 resequencing approach. *Genetics* **190**: 1417–1432.

838 Chen H, Zwaenepoel A, Van de Peer Y. 2024. wgd v2: a suite of tools to uncover and date
839 ancient polyploidy and whole-genome duplication. *Bioinformatics* **40**: btae272.

840 Coyle S, Kroll E. 2007. Starvation induces genomic rearrangements and starvation-resilient
841 phenotypes in yeast. *Mol Biol Evol* **25**: 310–318.

842 Dean A. 2011. In the loop: long range chromatin interactions and gene regulation. *Brief*
843 *Funct Genomics* **10**: 3–10.

844 de Vos JM, Augustijnen H, Bätscher L, Lucek K. 2020. Speciation through chromosomal
845 fusion and fission in Lepidoptera. *Philos Trans R Soc Lond B Biol Sci* **375**: 20190539.

846 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
847 Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–
848 21.

849 Duboule D, Dollé P. 1989. The structural and functional organization of the murine HOX
850 gene family resembles that of *Drosophila* homeotic genes. *EMBO J* **8**: 1497–1505.

851 Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. 2002.
852 Characteristic genome rearrangements in experimental evolution of Saccharomyces
853 cerevisiae. *Proc Natl Acad Sci USA* **99**: 16144–16149.

854 Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016.

855 Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom.
856 *Cell Syst* **3**: 99–101.

857 Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative
858 genomics. *Genome Biol* **20**: 238.

859 Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. 2011. The
860 Cambrian conundrum: early divergence and later ecological success in the early history
861 of animals. *Science* **334**: 1091–1097.

862 Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020.
863 RepeatModeler2 for automated genomic discovery of transposable element families.
864 *Proc Natl Acad Sci USA* **117**: 9451–9457.

865 Fuchs J, Martindale MQ, Hejnol A. 2011. Gene expression in bryozoan larvae suggest a
866 fundamental importance of pre-patterned blastemic cells in the bryozoan life-cycle.
867 *Evodevo* **2**: 13.

868 Gabriel L, Brůna T, Hoff KJ, Ebel M, Lomsadze A, Borodovsky M, Stanke M. 2024.
869 BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence
870 with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res* **34**: 769–777.

871 Gordon D, Costello M. 2016. Bryozoa—not a minor phylum. *NZ Sci Rev* **73**: 63–66.

872 Graham A, Papalopulu N, Krumlauf R. 1989. The murine and Drosophila homeobox gene
873 complexes have common features of organization and expression. *Cell* **57**: 367–378.

874 Gregory TR, Mable BK. 2005. CHAPTER 8 - Polyploidy in Animals. In *The Evolution of the*
875 *Genome* (ed. T.R. Gregory), pp. 427–517, Academic Press, Burlington.

876 Haag CR, Roze D. 2007. Genetic load in sexual and asexual diploids: segregation,
877 dominance and genetic drift. *Genetics* **176**: 1663–1678.

878 Hao Z, Lv D, Ge Y, Shi J, Weijers D, Yu G, Chen J. 2020. RIdeogram: drawing SVG
879 graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput Sci*
880 **6**: e251.

881 Harewood L, Fraser P. 2014. The impact of chromosomal rearrangements on regulation of
882 gene expression. *Hum Mol Genet* **23**: R76–82.

883 Hejnol A, Obst M, Stamatakis A, Ott M, Rouse GW, Edgecombe GD, Martinez P, Baguñà J,
884 Bailly X, Jondelius U, et al. 2009. Assessing the root of bilaterian animals with scalable

885 phylogenomic methods. *Proc Biol Sci* **276**: 4261–4270.

886 Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the
887 ultrafast bootstrap approximation. *Mol Biol Evol* **35**: 518–522.

888 Hoencamp C, Dudchenko O, Elbatsh AMO, Brahmachari S, Raaijmakers JA, van Schaik T,
889 Sedeño Cacciatore Á, Contessoto VG, van Heesbeen RGHP, van den Broek B, et al.
890 2021. 3D genomics across the tree of life reveals condensin II as a determinant of
891 architecture type. *Science* **372**: 984–989.

892 Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. 2016. BRAKER1: Unsupervised
893 RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS.
894 *Bioinformatics* **32**: 767–769.

895 Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-genome annotation with
896 BRAKER. *Methods Mol Biol* **1962**: 65–95.

897 Höök L, Näsvall K, Vila R, Wiklund C, Backström N. 2023. High-density linkage maps and
898 chromosome level genome assemblies unveil direction and frequency of extensive
899 structural rearrangements in wood white butterflies (*Leptidea* spp.). *Chromosome Res*
900 **31**: 2.

901 Hubert KA, Wellik DM. 2023. Hox genes in development and beyond. *Development* **150**.
902 http://dx.doi.org/10.1242/dev.192476.

903 Ivankovic M, Brand JN, Pandolfini L, Brown T, Pippel M, Rozanski A, Schubert T, Grohme
904 MA, Winkler S, Robledillo L, et al. 2023. A comparative analysis of planarian genomes
905 reveals regulatory conservation in the face of rapid structural divergence. *bioRxiv*
906 2023.12.22.572568. https://www.biorxiv.org/content/10.1101/2023.12.22.572568v1
907 (Accessed January 17, 2024).

908 Jenner RA, Littlewood DTJ. 2008. Problematica old and new. *Philos Trans R Soc Lond B*
909 *Biol Sci* **363**: 1503–1512.

910 Jenner RA, Timothy D, Littlewood J. 2009. Invertebrate Problematica: kinds, causes, and
911 solutions. http://dx.doi.org/10.1093/acprof:oso/9780199549429.003.0011 (Accessed
912 January 16, 2024).

913 Jin F, Zhou Z, Guo Q, Liang Z, Yang R, Jiang J, He Y, Zhao Q, Zhao Q. 2020. High-quality
914 genome assembly of *Metaphire vulgaris*. *PeerJ* **8**: e10313.

915 Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder:
916 fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587–589.

917 Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple
918 sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059–3066.

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

921 Kenny NJ, McCarthy SA, Dudchenko O, James K, Betteridge E, Corton C, Dolucan J, Mead
922 D, Oliver K, Omer AD, et al. 2020. The gene-rich genome of the scallop *Pecten*
923 *maximus*. *Gigascience* **9**: giaa037.

924 Khalturin K, Shunatova N, Shchenkov S, Sasakura Y, Kawamitsu M, Satoh N. 2022. Polyzoa
925 is back: The effect of complete gene sets on the placement of Ectoprocta and
926 Entoprocta. *Sci Adv* **8**: eabo4400.

927 King M. 1995. *Species Evolution: The Role of Chromosome Change*. Cambridge University
928 Press.

929 Kocot KM, Struck TH, Merkel J, Waits DS, Todt C, Brannock PM, Weese DA, Cannon JT,
930 Moroz LL, Lieb B, et al. 2017. Phylogenomics of Lophotrochozoa with consideration of
931 systematic error. *Syst Biol* **66**: 256–282.

932 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA.
933 2009. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**:
934 1639–1645.

935 Kwiatkowski D, Blaxter M, Darwin Tree of Life Barcoding collective, Wellcome Sanger
936 Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA
937 Pipelines collective, Tree of Life Core Informatics collective, Darwin Tree of Life
938 Consortium. 2021. The genome sequence of the bootlace worm, *Lineus longissimus*
939 (Gunnerus, 1770). *Wellcome Open Res* **6**: 272.

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

943 Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database
944 Collaboration. 2011. The sequence read archive. *Nucleic Acids Res* **39**: D19–21.

945 Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic
946 tree display and annotation. *Nucleic Acids Res* **49**: W293–W296.

947 Lewin TD, Liao IJ-Y, Luo Y-J. 2024a. Annelid comparative genomics and the evolution of
948 massive lineage-specific genome rearrangement in bilaterians. *Mol Biol Evol* **41**:
949 msae172.

950 Lewin TD, Shimizu K, Liao IJ-Y, Chen M-E, Endo K, Satoh N, Holland PWH, Wong YH, Luo
951 Y-J. 2024b. Brachiopod genome unveils the evolution of the BMP–Chordin network in
952 bilaterian body patterning. *bioRxiv* 2024.05.28.596352.

953 Liao IJ-Y, Lu T-M, Chen M-E, Luo Y-J. 2023. Spiralian genomics and the evolution of animal
954 genome architecture. *Brief Funct Genomics* **22**: 498–508.

955 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
956 MEM. *arXiv [q-bioGN]*. http://arxiv.org/abs/1303.3997.

957 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin
958 R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–
959 2079.

960 Lin C-Y, Marlétaz F, Pérez-Posada A, Martínez-García PM, Schloissnig S, Peluso P,
961 Conception GT, Bump P, Chen Y-C, Chou C, et al. 2024. Chromosome-level genome
962 assemblies of 2 hemichordates provide new insights into deuterostome origin and
963 chromosome evolution. *PLoS Biol* **22**: e3002661.

964 Lombardi C, Taylor PD, Cocito S. 2020. Bryozoans: The “forgotten” bioconstructors. in
965 *Perspectives on the marine animal forests of the world*, pp. 193–217, Springer
966 International Publishing, Cham.

967 Lomsadze A, Burns PD, Borodovsky M. 2014. Integration of mapped RNA-Seq reads into
968 automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res* **42**: e119.

969 Loxton J, Wood CA, Bishop JDD, Porter JS, Spencer Jones M, Nall CR. 2017. Distribution of
970 the invasive bryozoan Schizoporella japonica in Great Britain and Ireland and a review
971 of its European distribution. *Biol Invasions* **19**: 2225–2235.

972 Luo Y-J, Kanda M, Koyanagi R, Hisata K, Akiyama T, Sakamoto H, Sakamoto T, Satoh N.
973 2017. Nemertean and phoronid genomes reveal lophotrochozoan evolution and the
974 origin of bilaterian heads. *Nat Ecol Evol* **2**: 141–151.

975 Mackintosh A, de la Rosa PMG, Martin SH, Lohse K, Laetsch DR. 2023a. Inferring inter-
976 chromosomal rearrangements and ancestral linkage groups from synteny. *bioRxiv*
977 2023.09.17.558111. https://www.biorxiv.org/content/10.1101/2023.09.17.558111v1
978 (Accessed January 8, 2024).

979 Mackintosh A, Vila R, Laetsch DR, Hayward A, Martin SH, Lohse K. 2023b. Chromosome
980 fissions and fusions act as barriers to gene flow between *Brenthis* fritillary butterflies.
981 *Mol Biol Evol* **40**. http://dx.doi.org/10.1093/molbev/msad043.

982 Mackintosh A, Vila R, Martin SH, Setter D, Lohse K. 2023c. Do chromosome
983 rearrangements fix by genetic drift or natural selection? Insights from *Brenthis*
984 butterflies. *Mol Ecol*. http://dx.doi.org/10.1111/mec.17146.

985 Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: Novel
986 and streamlined workflows along with broader and deeper phylogenetic coverage for
987 scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* **38**: 4647–4654.

988 Marlétaz F, Couloux A, Poulain J, Labadie K, Da Silva C, Mangenot S, Noel B, Poustka AJ,
989 Dru P, Pegueroles C, et al. 2023. Analysis of the *P. lividus* sea urchin genome highlights
990 contrasting trends of genomic and regulatory evolution in deuterostomes. *Cell Genom* **3**:
991 100295.

992 Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. 2019. A new spiralian
993 phylogeny places the enigmatic arrow worms among gnathiferans. *Curr Biol* **29**: 312–
994 318.e3.

995 Martín-Zamora FM, Liang Y, Guynes K, Carrillo-Baltodano AM, Davies BE, Donnellan RD,
996 Tan Y, Moggioli G, Seudre O, Tran M, et al. 2023. Annelid functional genomics reveal
997 the origins of bilaterian life cycles. *Nature* **615**: 105–110.

998 Mayr E. 1961. Cause and effect in biology. *Science* **134**: 1501–1506.

999 Miele A, Dekker J. 2008. Long-range chromosomal interactions and gene regulation. *Mol*
1000 *Biosyst* **4**: 1046–1057.

1001 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear
1002 R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in
1003 the genomic era. *Mol Biol Evol* **37**: 1530–1534.

1004 Moriyama Y, Koshiba-Takeuchi K. 2018. Significance of whole-genome duplications on the
1005 emergence of evolutionary novelties. *Brief Funct Genomics* **17**: 329–338.

1006 Muffato M, Louis A, Nguyen NTT, Lucas J, Berthelot C, Roest Crollius H. 2023.
1007 Reconstruction of hundreds of reference ancestral genomes across the eukaryotic
1008 kingdom. *Nat Ecol Evol* **7**: 355–366.

1009 Nesnidal MP, Helmkampf M, Meyer A, Witek A, Bruchhaus I, Ebersberger I, Hankeln T, Lieb
1010 B, Struck TH, Hausdorf B. 2013. New phylogenomic data support the monophyly of
1011 Lophophorata and an Ectoproct-Phoronid clade and indicate that Polyzoa and
1012 Kryptrochozoa are caused by systematic bias. *BMC Evol Biol* **13**: 253.

1013 Nishizawa A, Sarashina I, Tsujimoto Y, Iijima M, Endo K. 2010. Artificial fertilization, early
1014 development and chromosome numbers in the brachiopod *Lingula anatina*. *Special*
1015 *Papers in Palaeontology* **84**: 309–316.

1016 Orr HA. 1990. “Why polyploidy is rarer in animals than in plants” revisited. *Am Nat* **136**: 759–
1017 770.

1018 Orr RJS, Di Martino E, Ramsfjell MH, Gordon DP, Berning B, Chowdhury I, Craig S,
1019 Cumming RL, Figuerola B, Florence W, et al. 2022. Paleozoic origins of cheilostome
1020 bryozoans and their parental care inferred by a new genome-skimmed phylogeny. *Sci*
1021 *Adv* **8**: eabm7452.

1022 Orr RJS, Sannum MM, Boessenkool S, Di Martino E, Gordon DP, Mello HL, Obst M,
1023 Ramsfjell MH, Smith AM, Liow LH. 2021. A molecular phylogeny of historical and
1024 contemporary specimens of an under-studied micro-invertebrate group. *Ecol Evol* **11**:
1025 309–320.

1026 Paradis E, Schliep K. 2018. ape 5.0: an environment for modern phylogenetics and
1027 evolutionary analyses in R. *Bioinformatics* **35**: 526–528.

1028 Parey E, Louis A, Montfort J, Bouchez O, Roques C, Iampietro C, Lluch J, Castinel A,
1029 Donnadieu C, Desvignes T, et al. 2023. Genome structures resolve the early
1030 diversification of teleost fishes. *Science* **379**: 572–575.

1031 Passarge E, Horsthemke B, Farber RA. 1999. Incorrect use of the term synteny. *Nat Genet*
1032 **23**: 387.

1033 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-
1034 aware quantification of transcript expression. *Nat Methods* **14**: 417–419.

1035 Plessy C, Mansfield MJ, Bliznina A, Masunaga A, West C, Tan Y, Liu AW, Grašič J, Del Río
1036 Pisula MS, Sánchez-Serna G, et al. 2024. Extreme genome scrambling in marine

1037 planktonic *Oikopleura dioica* cryptic species. *Genome Res* **34**: 426–440.

1038 Pratt C, Denley D, Metaxas A. 2022. Ocean warming and multiple source populations
1039 increase the threat of an invasive bryozoan to kelp beds in the northwest Atlantic
1040 Ocean. *Mar Ecol Prog Ser* **695**: 65–81.

1041 Proost S, Fostier J, De Witte D, Dhoedt B, Demeester P, Van de Peer Y, Vandepoele K.
1042 2011. i-ADHoRe 3.0—fast and sensitive detection of genomic homology in extremely
1043 large data sets. *Nucleic Acids Res* **40**: e11–e11.

1044 Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi
1045 M, Shoguchi E, Terry A, Yu J-K, et al. 2008. The amphioxus genome and the evolution
1046 of the chordate karyotype. *Nature* **453**: 1064–1071.

1047 Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H,
1048 Lindquist E, Kapitonov VV, et al. 2007. Sea anemone genome reveals ancestral
1049 eumetazoan gene repertoire and genomic organization. *Science* **317**: 86–94.

1050 R Core Team. 2023. R: A language and environment for statistical computing. Vienna: R
1051 foundation for statistical computing. *Vienna: R foundation for statistical computing*.
1052 https://cir.nii.ac.jp/crid/1370576118723163397 (Accessed January 9, 2024).

1053 Renwick JH. 1971. The mapping of human chromosomes. *Annu Rev Genet* **5**: 81–120.

1054 Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O,
1055 Mayrose I. 2015. The Chromosome Counts Database (CCDB) - a community resource
1056 of plant chromosome numbers. *New Phytol* **206**: 19–26.

1057 Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**:
1058 351–358.

1059 Román-Palacios C, Medina CA, Zhan SH, Barker MS. 2021. Animal chromosome counts
1060 reveal a similar range of chromosome numbers but with less polyploidy in animals
1061 compared to flowering plants. *J Evol Biol* **34**: 1333–1339.

1062 Saadi AJ, de Oliveira AL, Kocot KM, Schwaha T. 2023. Genomic and transcriptomic survey
1063 of bryozoan Hox and ParaHox genes with emphasis on phylactolaemate bryozoans.
1064 *BMC Genomics* **24**: 711.

1065 Sacerdot C, Louis A, Bon C, Berthelot C, Roest Crollius H. 2018. Chromosome evolution at
1066 the origin of the ancestral vertebrate genome. *Genome Biol* **19**: 166.

1067 Schultz DT, Haddock SHD, Bredeson JV, Green RE, Simakov O, Rokhsar DS. 2023.
1068 Ancient gene linkages support ctenophores as sister to other animals. *Nature* **618**: 110–
1069 117.

1070 Schultz DT, Heath-Heckman EAC, Winchell CJ, Kuo D-H, Yu Y-S, Oberauer F, Kocot KM,
1071 Cho S-J, Simakov O, Weisblat DA. 2024. Acceleration of genome rearrangement in
1072 clitellate annelids. *bioRxiv* 2024.05.12.593736.
1073 https://www.biorxiv.org/content/10.1101/2024.05.12.593736v1 (Accessed July 10,
1074 2024).

1075 Schwaha T. 2020. *Phylum Bryozoa*. ed. T. Schwaha. De Gruyter, Berlin, Germany

1076 Schwaha TF, Ostrovsky AN, Wanninger A. 2020. Key novelties in the evolution of the
1077 aquatic colonial phylum Bryozoa: evidence from soft body morphology. *Biol Rev Camb*
1078 *Philos Soc* **95**: 696–729.

1079 Simakov O, Bredeson J, Berkoff K, Marletaz F, Mitros T, Schultz DT, O'Connell BL, Dear P,
1080 Martinez DE, Steele RE, et al. 2022. Deeply conserved synteny and the evolution of
1081 metazoan chromosomes. *Sci Adv* **8**: eabi5884.

1082 Simakov O, Marletaz F, Cho S-J, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo D-H,
1083 Larsson T, Lv J, Arendt D, et al. 2013. Insights into bilaterian evolution from three
1084 spiralian genomes. *Nature* **493**: 526–531.

1085 Simakov O, Marlétaz F, Yue J-X, O'Connell B, Jenkins J, Brandt A, Calef R, Tung C-H,
1086 Huang T-K, Schmutz J, et al. 2020. Deeply conserved synteny resolves early events in
1087 vertebrate evolution. *Nat Ecol Evol* **4**: 820–830.

1088 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO:
1089 assessing genome assembly and annotation completeness with single-copy orthologs.
1090 *Bioinformatics* **31**: 3210–3212.

1091 Smit AFA, Hubley R, Green P. 2015. RepeatMasker Open-4.0.

1092 Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically
1093 mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**: 637–
1094 644.

1095 Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes
1096 with a generalized hidden Markov model that uses hints from external sources. *BMC*
1097 *Bioinformatics* **7**: 62.

1098 Steenwyk JL, Buida TJ 3rd, Li Y, Shen X-X, Rokas A. 2020. ClipKIT: A multiple sequence
1099 alignment trimming software for accurate phylogenomic inference. *PLoS Biol* **18**:
1100 e3001007.

1101 Steenwyk JL, Buida TJ, Labella AL, Li Y, Shen X-X, Rokas A. 2021. PhyKIT: a broadly
1102 applicable UNIX shell toolkit for processing and analyzing phylogenomic data.
1103 *Bioinformatics* **37**: 2325–2331.

1104 Steenwyk JL, King N. 2024. The promise and pitfalls of synteny in phylogenomics. *PLoS Biol*
1105 **22**: e3002632.

1106 Sun Y, Sun J, Yang Y, Lan Y, Ip JC-H, Wong WC, Kwan YH, Zhang Y, Han Z, Qiu J-W, et
1107 al. 2021. Genomic signatures supporting the symbiosis and formation of chitinous tube
1108 in the deep-sea tubeworm *Paraescarpia echinospica*. *Mol Biol Evol* **38**: 4116–4134.

1109 Tange O. 2021. GNU Parallel 2210822 (’Kabul'). *Zenodo*.
1110 https://play.google.com/store/books/details?id=sKdSDwAAQBAJ.

1111 Taylor PD, Lombardi C, Cocito S. 2015. Biomineralization in bryozoans: present, past and
1112 future: Bryozoan biomineralization. *Biol Rev Camb Philos Soc* **90**: 1118–1150.

1113 Taylor PD, Waeschenbach A. 2015. Phylogeny and diversification of bryozoans.
1114 *Palaeontology* **58**: 585–599.

1115 Telford MJ, Budd GE, Philippe H. 2015. Phylogenomic insights into animal evolution. *Curr*
1116 *Biol* **25**: R876–87.

1117 The Darwin Tree of Life Project Consortium, Blaxter M, Mieszkowska N, Palma FD, Holland
1118 P, Durbin R, Richards T, Berriman M, Kersey P, Hollingsworth P, et al. 2022. Sequence
1119 locally, think globally: The Darwin Tree of Life Project. *Proc Natl Acad Sci USA* **119**:
1120 e2115642118.

1121 Tiley GP, Barker MS, Burleigh JG. 2018. Assessing the performance of *Ks* plots for
1122 detecting ancient whole genome duplications. *Genome Biol Evol* **10**: 2882–2898.

1123 Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. 2002. Looping and interaction
1124 between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**: 1453–1465.

1125 Vargas-Chávez C, Benítez-Álvarez L, Martínez-Redondo GI, Álvarez-González L, Salces-
1126 Ortiz J, Eleftheriadi K, Escudero N, Guiglielmoni N, Flot J-F, Novo M, et al. 2024. A
1127 punctuated burst of massive genomic rearrangements by chromosome shattering and

1128 the origin of non-marine annelids. *bioRxiv* 2024.05.16.594344.
1129 https://www.biorxiv.org/content/10.1101/2024.05.16.594344v2.full (Accessed August 10,
1130 2024).

1131 Wang S, Zhang J, Jiao W, Li J, Xun X, Sun Y, Guo X, Huan P, Dong B, Zhang L, et al. 2017.
1132 Scallop genome provides insights into evolution of bilaterian karyotype and
1133 development. *Nat Ecol Evol* **1**: 120.

1134 Wang Y, Jia L, Tian G, Dong Y, Zhang X, Zhou Z, Luo X, Li Y, Yao W. 2023. shinyCircos-
1135 V2.0: Leveraging the creation of Circos plot with enhanced usability and advanced
1136 features. *Imeta* **2**: e109.

1137 Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T-H, Jin H, Marler B, Guo H, et al.
1138 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and
1139 collinearity. *Nucleic Acids Res* **40**: e49.

1140 Weaver H, Cook P, Bock P, Gordon D. 2018. *Australian Bryozoa Volume 1: Biology,*
1141 *Ecology and Natural History*. Csiro Publishing.

1142 White MJD. 1954. *Animal cytology & evolution*. Cambridge university press.

1143 Wood ACL, Probert PK, Rowden AA, Smith AM. 2012. Complex habitat generated by marine
1144 bryozoans: a review of its distribution, structure, diversity, threats and conservation.
1145 *Aquat Conserv* **22**: 547–563.

1146 Wood C, Bishop J, Adkins P, Jenkins H, Marine Biological Association Genome Acquisition
1147 Lab, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute
1148 Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective,
1149 Darwin Tree of Life Consortium. 2023. The genome sequence of an erect bryozoan,
1150 *Bugulina stolonifera* (Ryland, 1960). *Wellcome Open Res* **8**: 26.

1151 Wright CJ, Stevens L, Mackintosh A, Lawniczak M, Blaxter M. 2024. Comparative genomics
1152 reveals the dynamics of chromosome evolution in Lepidoptera. *Nat Ecol Evol* **8**: 777–
1153 790.

1154 Wright S. 1941. On the probability of fixation of reciprocal translocations. *Am Nat* **75**: 513–
1155 522.

1156 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**:
1157 1586–1591.

1158 Yoshida K, Rödelsperger C, Röseler W, Riebesell M, Sun S, Kikuchi T, Sommer RJ. 2023.
1159 Chromosome fusions repatterned recombination rate and facilitated reproductive
1160 isolation during *Pristionchus* nematode speciation. *Nat Ecol Evol* **7**: 424–439.

1161 Zdobnov EM, von Mering C, Letunic I, Bork P. 2005. Consistency of genome-based methods
1162 in measuring Metazoan evolution. *FEBS Lett* **579**: 3355–3361.

1163 Zhang Z, Zhang Z, Ma J, Taylor PD, Strotz LC, Jacquet SM, Skovsted CB, Chen F, Han J,
1164 Brock GA. 2021. Fossil evidence unveils an early Cambrian origin for Bryozoa. *Nature*
1165 **599**: 251–255.

1166 Zhou C, McCarthy SA, Durbin R. 2023. YaHS: yet another Hi-C scaffolding tool.
1167 *Bioinformatics* **39**: btac808.