

# Ancient gene linkages support ctenophores as sister to other animals

<https://doi.org/10.1038/s41586-023-05936-6>

Received: 30 April 2022

Accepted: 9 March 2023

Published online: 17 May 2023

Open access

 Check for updates

Darrin T. Schultz<sup>1,2,3</sup>✉, Steven H. D. Haddock<sup>2,4</sup>, Jessen V. Bredeson<sup>5</sup>, Richard E. Green<sup>3</sup>, Oleg Simakov<sup>1</sup>✉ & Daniel S. Rokhsar<sup>5,6,7</sup>✉

A central question in evolutionary biology is whether sponges or ctenophores (comb jellies) are the sister group to all other animals. These alternative phylogenetic hypotheses imply different scenarios for the evolution of complex neural systems and other animal-specific traits<sup>1–6</sup>. Conventional phylogenetic approaches based on morphological characters and increasingly extensive gene sequence collections have not been able to definitively answer this question<sup>7–11</sup>. Here we develop chromosome-scale gene linkage, also known as synteny, as a phylogenetic character for resolving this question<sup>12</sup>. We report new chromosome-scale genomes for a ctenophore and two marine sponges, and for three unicellular relatives of animals (a choanoflagellate, a filasterean amoeba and an ichthyosporean) that serve as outgroups for phylogenetic analysis. We find ancient syntenies that are conserved between animals and their close unicellular relatives. Ctenophores and unicellular eukaryotes share ancestral metazoan patterns, whereas sponges, bilaterians, and cnidarians share derived chromosomal rearrangements. Conserved syntenic characters unite sponges with bilaterians, cnidarians, and placozoans in a monophyletic clade to the exclusion of ctenophores, placing ctenophores as the sister group to all other animals. The patterns of synteny shared by sponges, bilaterians, and cnidarians are the result of rare and irreversible chromosome fusion-and-mixing events that provide robust and unambiguous phylogenetic support for the ctenophore-sister hypothesis. These findings provide a new framework for resolving deep, recalcitrant phylogenetic problems and have implications for our understanding of animal evolution.

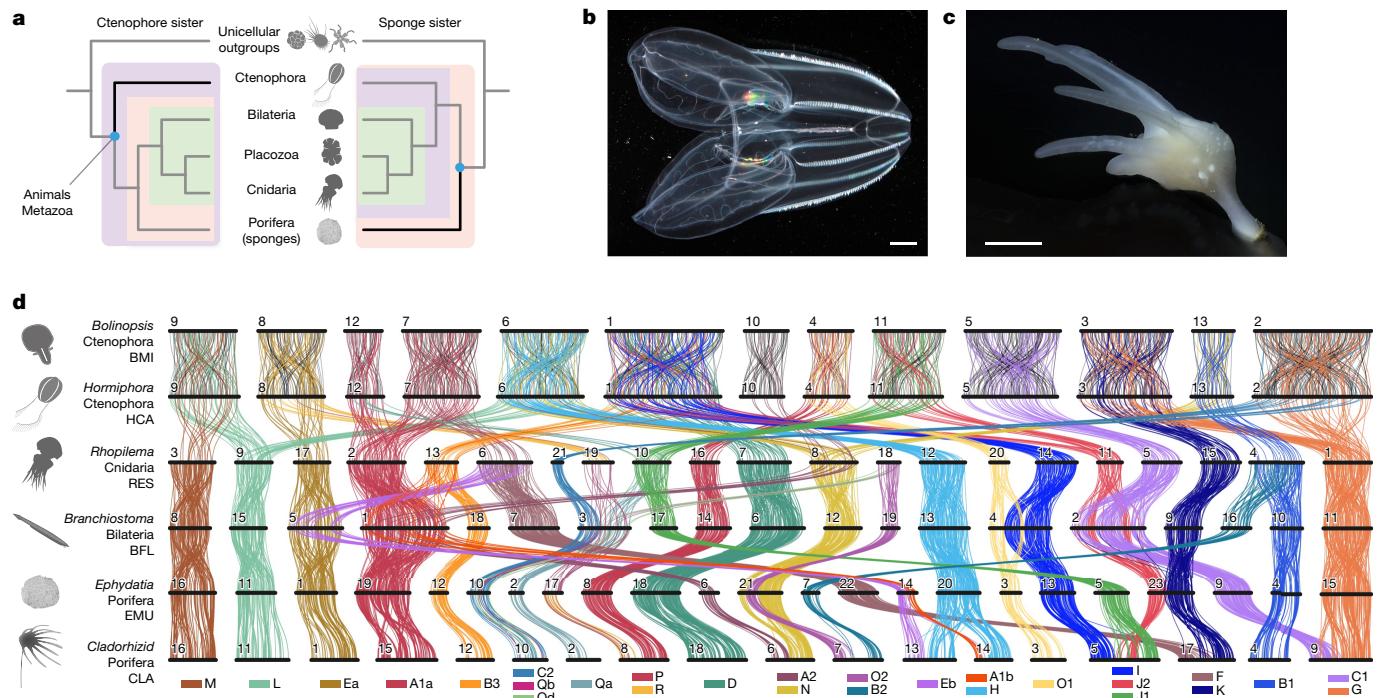
Five major lineages arose early in animal evolution and survive to the present day: sponges (poriferans), ctenophores (comb jellies), placozoans (microscopic flat animals), cnidarians (such as anemones, jellyfishes and hydra) and bilaterians (such as chordates, molluscs, arthropods and diverse worms)<sup>1,8,10,13,14</sup>. Although morphological and phylogenomic studies consistently unite bilaterians, cnidarians, and placozoans into a monophyletic clade (Parahoxozoa) that excludes sponges and ctenophores<sup>8,10,14</sup> the relationship between sponges, ctenophores and Parahoxozoa remains controversial. There are two competing scenarios—the sponge-sister hypothesis<sup>7,8</sup> and the ctenophore-sister hypothesis<sup>9,10</sup>—reflecting which lineage diverged first among animals (Fig. 1a).

As sponges and ctenophores are such disparate animals<sup>13</sup>, the nature of the first diverging animal lineage has implications for the evolution of fundamental animal characteristics. Adult sponges are generally sessile filter-feeding organisms with body plans organized into reticulated water-filtration channels, structures built out of silica or calcium carbonate, and specialized cell types and tissues used for feeding, reproduction and self-defence, but they lack neuronal and muscle cells<sup>15</sup>. By contrast, ctenophores are gelatinous marine predators that

move using eight longitudinal ‘comb rows’ of ciliary bundles<sup>16,17</sup>; they are superficially similar but unrelated to cnidarian medusae<sup>13,18</sup> and possess multiple nerve nets<sup>19</sup>. Thus, whereas the sponge-sister scenario suggests a single origin of neurons on the ctenophore–parahoxozoon stem, the ctenophore-sister scenario implies either that either ancestral metazoan neurons were lost in the sponge lineage, or that there was convergent evolution of neurons in the ctenophore and parahoxozoon lineages<sup>3,6</sup>. Similar considerations apply to other metazoan cell types<sup>18</sup>, gene regulatory networks, animal development<sup>13,18</sup> and other uniquely metazoan features.

Despite its importance for understanding animal evolution, the relative branching order of sponges, ctenophores and other animals has proven to be difficult to resolve<sup>2</sup>. The fossil record is largely silent on this issue as verified Precambrian sponge fossils are extremely rare<sup>20</sup> and putative fossils of the soft-bodied ctenophores are difficult to interpret<sup>21</sup>. Morphological characters of living groups (for example, choanocytes of sponges) are not sufficient to resolve the question because true homology is difficult to assign, and such characters are easily lost or can arise convergently<sup>13,22</sup>. The ctenophore-sister hypothesis is supported by a pair of gene duplications shared by sponges,

<sup>1</sup>Department of Neuroscience and Developmental Biology, University of Vienna, Vienna, Austria. <sup>2</sup>Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA. <sup>3</sup>Department of Biomolecular Engineering and Bioinformatics, University of California, Santa Cruz, CA, USA. <sup>4</sup>Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA. <sup>5</sup>Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. <sup>6</sup>Molecular Genetics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Japan. <sup>7</sup>Chan Zuckerberg Biohub, San Francisco, CA, USA. ✉e-mail: darrin.schultz@univie.ac.at; oleg.simakov@univie.ac.at; DS.Rokhsar@gmail.com



**Fig. 1 | Conserved synteny and the phylogenetic position of ctenophores and sponges.** **a**, Two alternative metazoan phylogenetic hypotheses, with either ctenophores (left) or sponges (right) as sister to all other animals. **b, c**, Specimens of species of which the genomes are reported here. Scale bars, 1 cm. **b**, The lobate ctenophore *B. microptera* from the Monterey Bay, California. **c**, Undescribed cladorhizid demosponge collected offshore of Big Sur, California at a depth of 3,975 m. **d**, Ribbon diagram showing conserved syntenies among animals ( $\alpha \leq 0.05$ , permutation test one-sided false-discovery rate), including (from top to bottom) two ctenophores (*B. microptera* (BMI) and *H. californensis*); the jellyfish *R. esculentum*; the bilaterian amphioxus *B. floridae*; and two demosponges

(*E. muelleri* and the cladorhizid demosponge). Each horizontal black bar represents a chromosome. The vertical lines between species represent orthologous genes, coloured according to the BCnS synteny groups<sup>12</sup>. Only groups of genes that have significantly conserved chromosome-scale linkage (synteny) between metazoan species are shown. There is extensive 1:1 conserved chromosomal synteny between the two ctenophores, consistent with the conserved ctenophore karyotype. Orthologous gene pairs in the two ctenophores that do not participate in conserved syntenies with BCnS are shown in grey. Photography credits: Shannon Johnson © 2019 MBARI (**b**), © 2021 MBARI (**c**).

bilaterians, placozoans and cnidarians but not ctenophores<sup>23</sup>. Although sophisticated methods for sequence-based phylogenomics have been developed and applied to increasingly large molecular datasets, there is still considerable debate about the relative position of sponges and ctenophores as results are sensitive to how sequence evolution is modelled<sup>11</sup>, which taxa or sites are included<sup>24,25</sup>, and the effects of long-branch artifacts and nucleotide compositional variation<sup>26</sup>. New approaches are needed.

We reasoned that patterns of synteny, classically defined as chromosomal gene linkage without regard to gene order<sup>27</sup>, could provide a powerful tool for resolving the ctenophore-sister versus sponge-sister debate. Chromosomal patterns of gene linkage evolve slowly in many lineages<sup>12,28–30</sup>, probably because it is improbable for interchromosomal translocations to be fixed in populations with large effective population sizes<sup>28,31,32</sup>. Notably, some changes in synteny are effectively irreversible. For example, when two distinct ancestral synteny groups are combined onto a single chromosome by translocation, and subsequent intra-chromosomal rearrangements mix these two groups of genes, it is very unlikely that the ancestral separated pattern will be restored by further rearrangement and fission, in the same sense that spontaneous reduction in entropy is improbable<sup>12</sup>. Such rare and irreversible changes are particularly useful for resolving challenging phylogenetic questions as they give rise to shared derived features that unambiguously unite all descendant lineages<sup>33–35</sup>. Deeply conserved syntenies observed between animals and their closest unicellular relatives<sup>12</sup> suggest that outgroup comparisons could be used to infer ancestral metazoan states and polarize changes within animals to address the sponge-sister versus ctenophore-sister debate. Yet, chromosome-scale genome sequences of the unicellular

or colonial eukaryotic outgroups closest to animals (choanoflagellates, filastereans and ichthyosporeans) have not been reported.

Here we show that conserved syntenies between animals and their closest unicellular relatives support ctenophores as the sister group to all other animals. Specifically, we find seven sets of genes for which (1) ctenophores share ancestral metazoan gene linkages with one or more unicellular eukaryotes; and (2) bilaterians, cnidarians, placozoans and sponges are united (to the exclusion of ctenophores) by shared derived patterns of synteny that arose by ancient interchromosomal translocations. In four of these cases, irreversible mixing after chromosome fusion evidently occurred on the bilaterian–cnidarian–sponge (BCnS) stem lineage, providing unambiguous support for the ctenophore-sister scenario. The alternative sponge-sister hypothesis is not supported by any synteny-based characters, and would require reversal of four sets of fusion-with-mixing events and/or extensive convergent fusion in both sponges and on the bilaterian/cnidarian stem to account for the observed patterns of synteny. To enable these analyses, we generated chromosome-scale genome sequences for three animal species (two sponges and a ctenophore), and three non-animal species (a filasterean, ichthyosporean and choanoflagellate) to serve as outgroups. Our analyses further reveal ancient syntenies conserved between animals and their closest unicellular relatives (animal plesiomorphies) as well as metazoan syntenies shared by all animals but not present in unicellular organisms (animal synapomorphies). These findings establish a phylogenetic framework for understanding the early evolution of metazoan genomes and characters.

To examine conserved syntenies across animals, we traced the chromosomal distribution of orthologous genes among diverse metazoan

lineages using previously and newly sequenced genomes (Fig. 1bc, Methods, Supplementary Information 1–3 and Supplementary Data 1). Figure 1d highlights conserved metazoan synteny groups, that is, groups of genes of which orthologues are linked on the same chromosome across multiple lineages, regardless of gene order. Syntenic groups shown in Fig. 1d are statistically significant (Methods). In Fig. 1d, lines connecting orthologous genes are coloured according to the previously identified BCnS ancestral linkage groups (ALGs)<sup>12,28–30</sup>. For example, the group on the far left represents the BCnS ALG\_M (comprising genes found on jellyfish *Rhopilema esculentum* chromosome 2 (RES2), amphioxus *Branchiostoma floridae* chromosome 8 (BFL8), and sponge chromosomes CLA16 (of a cladorhizid demosponge) and EMU16 (of *Ephydatia muelleri*)). Note that, by our definition, two different conserved synteny groups can coexist on the same chromosome in some species. For example, amphioxus chromosome BFL5 is seen to be a combination of BCnS ALGs Ea and Eb, which are found on distinct chromosomes in other species.

Our results extend previous findings<sup>12</sup> of BCnS ALGs by incorporating a new chromosome-scale genome sequence of a recently discovered bioluminescent deep-sea cladorhizid demosponge<sup>36</sup> (Fig. 1c, Supplementary Information 2 and Extended Data Fig. 1c–i) complementing the spongillid demosponge *E. muelleri*<sup>37</sup>. Although the cladorhizid and spongillid lineages diverged approximately 450 million years ago<sup>38</sup>, chromosomes of the two demosponges correspond simply with each other (Fig. 1d and Extended Data Fig. 2f–h) and with bilaterian and cnidarian chromosomes (Fig. 1d and Extended Data Fig. 2i–k), consistent with the previously described genome tectonic schema<sup>12</sup>. Further comparisons with other recently released chromosome-scale demosponge genome sequences<sup>39</sup> confirm the high degree of conserved synteny in this group, but show that one of the rearrangements that we found in the cladorhizid genome is the result of a fission in that lineage (ALG\_H; Extended Data Fig. 2f–k). We also sequenced the genome of a previously undescribed hexactinellid (glass) sponge (Extended Data Fig. 1j–n and Supplementary Information 2), but found it to be considerably rearranged. Despite many lineage-specific genomic changes in glass sponges, relicts of 10 out of 29 BCnS ALGs are detectable (Extended Data Fig. 3). Owing to the high degree of rearrangement, we do not consider hexactinellid genomes further.

However, in contrast to demosponges, genomic comparisons between the cydippid ctenophore *Hormiphora californensis*<sup>40</sup> and other metazoans reveal patterns of both conserved and altered synteny (Fig. 1d and Extended Data Fig. 2b–e). For example, whereas the BCnS group ALG\_Ea is localized to a single ctenophore chromosome (*H. californensis* chromosome 8 (HCA8)), the BCnS synteny group ALG\_A1a (comprising genes found on amphioxus chromosome BFL1, jellyfish chromosome RES2, and sponge chromosomes EMU1 and CLA15) is partitioned across two ctenophore chromosomes (HCA12 and HCA7). To test whether the observed patterns of ctenophore synteny are unique to the *H. californensis* lineage or common across ctenophores, we assembled and analysed the genome of the recently redescribed<sup>41</sup> lobate comb jelly *Bolinopsis microptera* (Fig. 1b; the assembly is reported in the Methods, Supplementary Information 1, Extended Data Fig. 1a,b and Supplementary Table 1.1–1.4). Despite the 160–260-million-year divergence between lobate and cydippid ctenophores<sup>10</sup> their  $n=13$  chromosomes show one-to-one correspondence (without gene order conservation) (Fig. 1d and Extended Data Fig. 2a). This finding implies that a common  $n=13$  karyotype is ancestral for the *Hormiphora–Bolinopsis* crown group, and that cross-metazoan patterns of synteny shown in Fig. 1d are general.

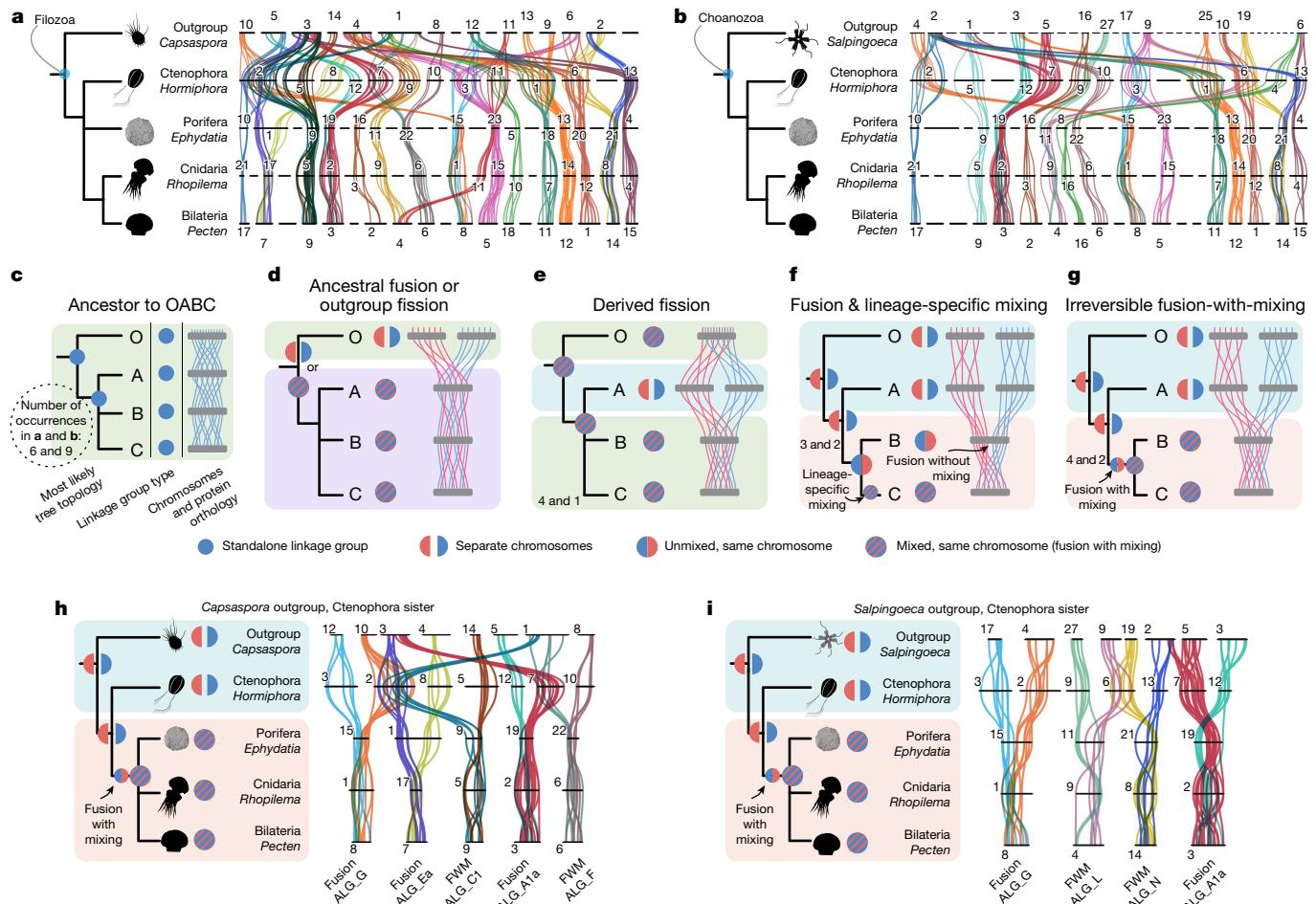
Interpreting the differences in synteny between ctenophores and other animals depends on the ancestral metazoan state (Fig. 2c–g, Extended Data Fig. 4 and Supplementary Information 4). If BCnS syntenies are ancestral to all metazoans, then the partitioned syntenies observed in ctenophores would have arisen by rearrangements that split the ancestral chromosomes in the ctenophore lineage (syntenic autapomorphies;

Fig. 2e) and would therefore be uninformative for discriminating between the ctenophore-sister and sponge-sister hypotheses. Alternatively, if the patterns of synteny found in ctenophores are ancestral to animals, the derived syntenies shared by BCnS to the exclusion of ctenophores could have arisen by fusion on the BCnS stem lineage, which would represent syntenic synapomorphies (Fig. 2f,g). In this case, ctenophores would be excluded from the BCnS clade and established as the sister clade of all other extant metazoans. Note that the extensive conservation of synteny between sponges, bilaterians, and cnidarians<sup>12</sup> confirmed here makes it improbable that ctenophores could share syntenies with cnidarians and bilaterians to the exclusion of sponges and, indeed, we did not find any such cases in analyses described below.

To provide outgroups for inferring ancestral metazoan syntenies, we assembled chromosome-scale sequences of representatives of three unicellular lineages closest to animals (collectively, outgroups): the choanoflagellate *Salpingoeca rosetta* (chromosome number,  $n=36$ ), the filasterean amoeba *Capsaspora owczarzaki* (chromosome number,  $n=16$ ) and the ichthyosporean *Creolimax fragrantissima* (chromosome number,  $n=26$ ). Chromosome-scale sequences and karyotypes were obtained by integrating previously reported subchromosomal draft sequences<sup>42–44</sup> with new chromatin conformation data (Methods, Supplementary Information 3 and Extended Data Fig. 5).

Chromosomal comparisons across animal and non-animal outgroup genomes revealed conserved ancestral metazoan synteny groups ranging in size from 5 to 29 genes, totalling 291 genes (out of 2,474 outgroup-metazoan orthology groups; Methods, Figs. 2 and 3, Extended Data Tables 1 and 2, Extended Data Figs. 5–8, Supplementary Information 5–9 and Supplementary Data 2). This finding extends previous observations based on subchromosomal assemblies of non-animal species<sup>12</sup>. Each such ancestral metazoan synteny group is a collection of genes of which the orthologues are consistently linked on single chromosomes in diverse metazoans and at least one outgroup (Fig. 2h–i and Extended Data Table 2). In contrast to the readily detected conserved syntenies among sponges, cnidarians and bilaterians, conserved syntenies involving ctenophores and non-animal outgroups are not visually evident in pairwise comparisons with other animals (Extended Data Figs. 2 and 6) but are statistically supported in multispecies comparisons (Methods and Supplementary Information 4 and 11). On the basis of permutation tests, the false-discovery rate of a conserved group of five linked genes in a four-species comparison is  $\alpha \leq 0.0003$ , and groups of eight or more linked genes never occurred in ten million permutations (Supplementary Information 8 and Extended Data Table 2). To maximize coverage of lineages relevant for the branching order of sponges and ctenophores, we considered orthologous genes across quartets of the form {outgroup, sponge, ctenophore, cnidarian/bilaterian}, which does not presuppose either the ctenophore-sister or sponge-sister hypothesis (Fig. 3 and Supplementary Information 4, 8 and 9). The extensive conservation of synteny across BCnS and within ctenophores makes our analysis insensitive to which genomes are used to represent these major metazoan clades (Supplementary Information 8). Here we used the scallop *Pecten maximus*<sup>45</sup>, the fire jellyfish *R. esculentum*<sup>46</sup>, the freshwater sponge *E. muelleri*<sup>37</sup> and the ctenophore *H. californensis* to represent the bilaterian, cnidarian, sponge and ctenophore genomes (Methods), although our findings do not depend on these choices (Figs. 2 and 3). We used two different methods for identifying orthologues—a simple mutual-best-hits method and an alternative orthologue-clustering approach (OrthoFinder<sup>47</sup>; Supplementary Information 10 and 11), and obtained comparable results using both approaches (Extended Data Fig. 9).

Although choanoflagellates are considered to be the closest living relatives of animals<sup>48,49</sup>, we found that the more distantly related filasterean *Capsaspora* shares 29 conserved synteny groups with metazoans, compared to 20 between the choanoflagellate *Salpingoeca* and metazoans (Extended Data Figs. 6 and 7), perhaps indicating more rapid interchromosomal rearrangement in the *Salpingoeca* lineage. The even



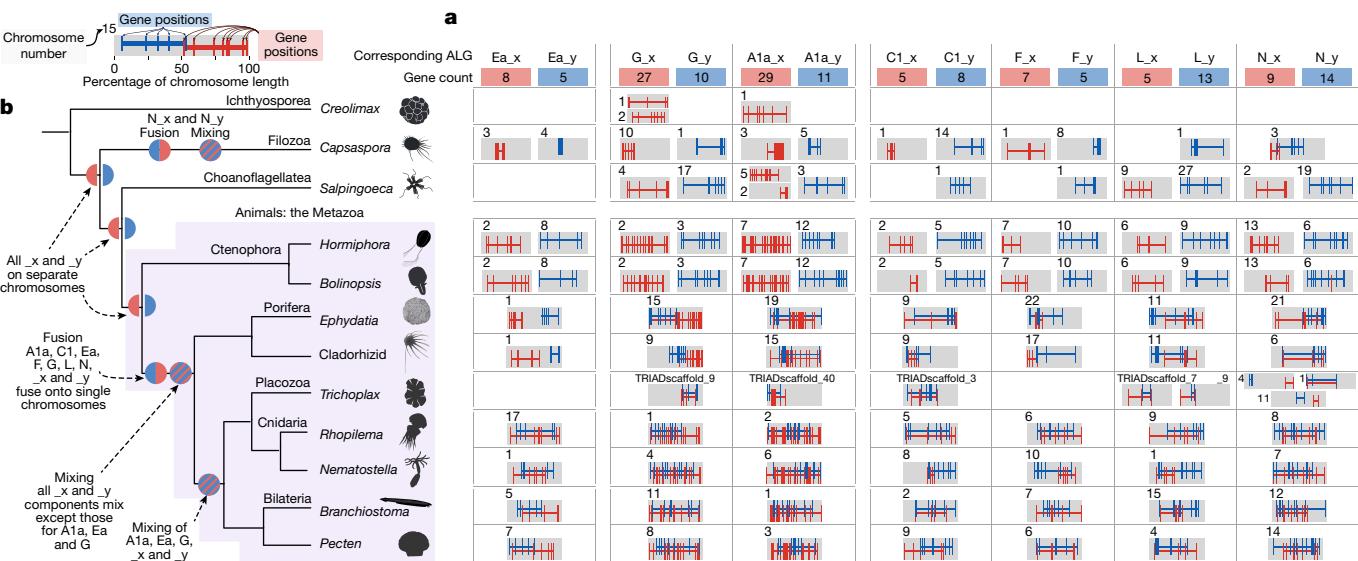
**Fig. 2 | Patterns of conserved synteny between animals and outgroups, and their implications. a,b**, Conserved linkages between the chromosomes of animals and two non-animal outgroups ( $\alpha \leq 0.05$ , permutation test one-sided false-discovery rate). **a**, The filasterean amoeba *C. owczarzaki*. **b**, The choanoflagellate *S. rosetta*. Each synteny group (conserved between metazoans and an outgroup) was assigned a distinct colour (different from Fig. 1). **c–g**, Schematics showing phylogenetic information (ancestor to OABC (c); ancestral fusion or outgroup fission (d); derived fission (e); fusion and lineage-specific mixing (f); and irreversible fusion with mixing (g)) conveyed by patterns of conserved synteny based on a quartet analysis. Node O designates the outgroup, and nodes A, B and C are ingroups of which the phylogenetic branching is to be determined. The thin red and blue lines in **d–g** represent genes of distinct synteny groups on different chromosomes in at least one species. Changes in synteny characters are indicated schematically on

parsimonious phylogenies on the left of each diagram. **d–g**, Single-species differences are phylogenetically uninformative. **f,g**, Shared chromosomal distributions between outgroup O and one of the ingroups (labelled taxon A) imply that the other two ingroups (taxa B and C) are related by fusion of ancestral synteny groups. **f** and **g** differ in whether all fusions have subsequently mixed. Fusion with mixing (**g**) is the strongest phylogenetic character because it represents an irreversible change, as discussed in the main text. **h,i**, Subsets of the synteny groups shown in **a** and **b** (*Capsaspora* (**h**) and *Salpingoeca* (**i**)) that match the phylogenetically informative patterns indicated in **f** and **g**. In all such cases, ctenophore syntenies match the outgroup and sponges share fusions with bilaterians and cnidarians. Note that groups A1a and G are found in both outgroups. We did not observe any cases in which sponge syntenies match the outgroup to the exclusion of ctenophores, bilaterians and cnidarians.

more distantly related ichthyosporean *Creolimax* still retains eight conserved synteny groups with metazoans. Although we considered each outgroup-plus-metazoan comparison separately, we found widespread overlap between the ancient synteny groups defined independently by comparison with *Capsaspora* and *Salpingoeca* (Extended Data Table 1). The extensive conservation of synteny within BCnS implies that the ancestral metazoan synteny groups correspond to subsets of the BCnS groups, and we name them using the BCnS notation with the suffixes \_x and \_y. If we relax the condition that an outgroup gene must be present, more metazoan genes can be added to these ancestral metazoan syntenic units (Extended Data Fig. 8).

Conservation of synteny between animals and their unicellular relatives may at first seem surprising, as these lineages diverged more than 800 million years ago<sup>50</sup>. Within animals, it has been estimated

that ongoing small-scale translocations between chromosomes typically transfer 1% of genes to a different chromosome every ~40 million years<sup>12</sup>. The limited residual conservation of synteny between animals and close unicellular relatives suggests that small-scale translocations have occurred at similarly low rates along both choanoflagellate and filasterean lineages. The more extensive conservation observed between animals and *Capsaspora* versus *Salpingoeca* may be due to variations in this rate or differences in other chromosomal rearrangements over deep time. The *Capsaspora* karyotype is predominantly metacentric and, notably, we find that 11 of the 29 ancient synteny groups found in *Capsaspora* are concentrated on single chromosome arms, rather than dispersed across whole *Capsaspora* chromosomes, based on estimates of centromere position using chromatin conformation contacts. This raises the possibility that *Capsaspora* chromosome arms preserve ancient filozoon units and suggests further attention to the chromosome biology of non-metazoan relatives. We found no significant functional



**Fig. 3 | Phylogenetically informative syntenies support ctenophores as the sister clade to other animals.** **a**, The rows represent the species considered in our analyses. The top three rows are non-metazoan outgroups. Columns show pairs of phylogenetically informative and significantly large metazoan syntenies ( $\alpha \leq 0.05$ , permutation test one-sided false-discovery rate), labelled according to their BCNs names (A1a, C1 and so on) with the suffix \_x or \_y denoting subgroups shared across metazoans. The number of genes participating in each metazoan synteny group is indicated in red and blue rectangles at the top of each column. Only genes with defined orthologues in outgroups are shown here. Extended Data Fig. 8 shows a larger set of genes requiring only metazoan

orthologues. Inset: the convention for representing gene distributions on chromosomes (top left). The grey rectangles represent chromosomes (or large scaffolds) in the placozoan *Trichoplax*. The chromosome number or scaffold name is located above or to the left of the grey rectangle. Red and blue vertical hashes represent the relative position of genes participating in phylogenetically informative pairs of metazoan synteny groups. **b**, The most parsimonious phylogeny according to the logic of Fig. 2c–g (Extended Data Fig. 4 and Supplementary Information 4), the results of ref. 12 and the accepted monophyly of demosponges.

associations of anciently linked groups of genes (Supplementary Information 12 and Supplementary Data 3), consistent with a general slow rate of synteny loss due to the infertility of translocation heterozygotes<sup>28,31</sup>, which allows only small-scale interchromosomal translocation<sup>32</sup>.

With conserved ancestral metazoan syntenies in hand, we tested the ctenophore-sister versus sponge-sister hypotheses by identifying shared, derived synteny characters using standard phylogenetic methods. As noted above, two or more metazoan synteny groups can co-occur on the same chromosome in one or more genomes, corresponding to ancient fusions (that is, translocations<sup>51</sup>) (Fig. 2h,i). As only shared derived characters are phylogenetically informative, changes that are unique to a single lineage can be disregarded (Fig. 2e and Extended Data Fig. 4b,d,e,g). There are two different types of chromosomal fusions between two ancestrally linked groups of genes: without mixing (Fig. 2f) or with subsequent intermixing (Fig. 2g and Supplementary Data 4 and 5). Fusion-without-mixing is potentially reversible, as observed in Robertsonian fusions and fissions involving whole chromosome arms<sup>51</sup>. However, in the fusion-with-mixing case, reversion is extremely unlikely, comparable to the spontaneous reduction of entropy after mixing of two fluids<sup>12</sup>.

We encoded the state of each potential fusion into a phylogenetic character matrix as 0 (no fusion, that is, ALGs found on separate chromosomes), 1 (fused but unmixed) or 2 (fused and mixed). The mixed/unmixed status of a fusion was determined on the basis of the likelihood of the observed gene arrangement under a model of random rearrangement (that is, entropy of mixing of the two fused groups) (Methods, Supplementary Information 13 and Supplementary Data 4 and 5). The same fusion character states were obtained using orthology defined by mutual-best-hits or OrthoFinder. We then applied the machinery of Bayesian phylogenetics<sup>52</sup> to this character matrix, using asymmetric transition probabilities to reflect the highly improbable unmixing transition (Methods and Supplementary Data 6).

Bayesian phylogenetic analyses of the fusion character matrix strongly support the ctenophore-sister topology (Fig. 4 and Supplementary

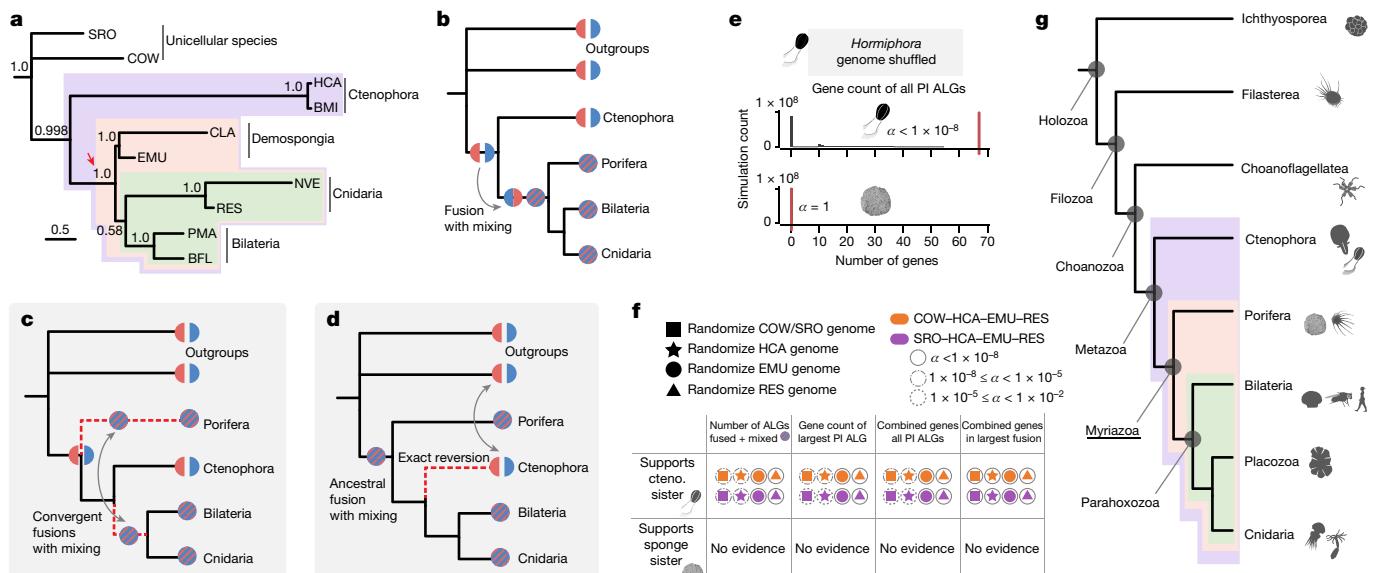
Information 14). The same conclusion is clear from direct examination of the fusions identified in our data. Specifically, there are seven derived fusions shared by bilaterians, cnidarians and sponges to the exclusion of ctenophores (Extended Data Table 1). Of these seven derived fusions, four are accompanied by mixing of genes from two different ancestral chromosomes—a process that is essentially irreversible (Figs. 2e and 4b–d); the other three are mixed only in bilaterians and cnidarians (Fig. 3).

We reject the alternative sponge-sister hypothesis as it would require either (1) multiple convergent fusions (that is, involving the same groups of genes) in both the sponge and bilaterian-cnidarian lineages (Fig. 4c and Supplementary Information 15) or (2) the precise reversal of multiple fusions-with-mixings in the ctenophore lineage to match the original patterns found in the ancestral metazoan lineage (Fig. 4d). The extreme unlikelihood of recovering the observed synteny patterns by chance is shown by simulations in which we permuted the configuration of the genes in each of the *C. owozcarzaki*, *S. rosetta*, *H. californensis*, *E. muelleri* and *R. esculentum* genomes (Supplementary Information 15 and Extended Data Fig. 10). Across one hundred million randomized *Hormiphora* genomes, we never found synteny signals comparable to those observed with the actual genome, indicating that synteny support for ctenophore-sister is unlikely to have arisen by chance (Fig. 4e). We also note the complete absence of synteny synapomorphies of a hypothetical ctenophore–bilaterian–cnidarian clade that excludes sponges, both in the actual data and in genome-shuffling simulations (Fig. 4f, Extended Data Fig. 10 and Extended Data Tables 1 and 2). This lack of homoplasy allows for a simple interpretation of the results (Fig. 4).

## Conclusions

### Support for the ctenophore-sister hypothesis

Our findings provide strong support for the ctenophore-sister scenario and reject the sponge-sister hypothesis. Although we encoded synteny states as a character matrix and analysed it using a Bayesian



**Fig. 4 | Phylogenetic analysis of patterns of conserved synteny and alternative interpretations.** **a**, Bayesian phylogenetic analysis of conserved synteny supports monophyly of the group comprising demosponges, Cnidaria and Bilateria, to the exclusion of Ctenophora, with high posterior probability (red arrow: 1.0, 100,000 generations with 25% burn-in). Bayesian analysis was run on both constrained (per-phylum constrained tree shown) and unconstrained tree topologies (Supplementary Information 14 and Supplementary Data 6). This panel corresponds to Supplementary Fig. 14.1c. **b**, Character transitions involving ALGs C1, F, L and N in Fig. 3 are most parsimoniously interpreted as fusion with mixing on the myriozoan stem after divergence from ctenophores, which retain the ancestral metazoan state as inferred from outgroup comparisons. **c,d**, To interpret the observed patterns under the alternative sponge-sister hypothesis would require unlikely convergent chromosomal changes (either convergent fusions (**c**) or exact unmixing and fissions to the ancestral state (**d**)) that were not seen in our genomes. **e**, The number of genes in the genome-shuffling simulations

( $n = 1 \times 10^8$ ) that support the ctenophore-sister (upper) or sponge-sister (lower) hypothesis. For the ctenophore *Hormiphora*, the number of fusion-with-mixing events is significantly higher in the observed genomes (vertical red bars) than in the *Hormiphora* genome-shuffling simulations (vertical grey histogram bars). Significance is shown as the one-sided false-discovery rate,  $\alpha$ , of a genome-shuffling permutation test. There were no groups of genes that supported the sponge-sister hypothesis in the real genomes, and none occurred in the genome-shuffling simulations. **f**, Additional statistical measures also support only the ctenophore (cteno.)-sister hypothesis in genome-shuffling simulations of *Hormiphora*, *Capsaspora* (COW), *Salpingoeca* (SRO), *Ephydatia* (EMU) and *Rhopilema*. PI ALG, phylogenetically informative linkage groups. The shape indicates the treatment; the colour indicates the outgroup. The full figure is shown in Extended Data Fig. 10. **g**, Summary of phylogenetic relationships among animals and close outgroups including syntetic characters. Myriozoa (underlined) is the name proposed for the clade containing extant animals, except Ctenophora. Outgroup topology follows ref. 49.

phylogenetic framework (Fig. 4a and Supplementary Fig. 14.1), the cladistic logic supporting our conclusions is easily appreciated, as emphasized above (Fig. 2, Extended Data Fig. 4 and Supplementary Information 4). Previous phylogenetic analyses of sequence-based characters have not resolved the sponge-sister versus ctenophore-sister hypotheses because the phylogenetic signal is weak and distributed across thousands of individual amino acid positions that are often saturated or subject to confounding evolutionary forces<sup>11</sup>. By contrast, the synteny-based characters that support ctenophores as sister to other animals in our analysis are clear: sponges, bilaterians, and cnidarians share multiple irreversible changes in synteny to the exclusion of ctenophores (BCnS syntetic synapomorphies) (Figs. 2 and 3). Support for the ctenophore-sister hypothesis is directly testable by future genome sequencing, as it is a strong prediction of our model that all bilaterian, placozoan, cnidarian or sponge genomes should share the four fusion-with-mixing syntetic synapomorphies shown in Figs. 2 and 3 and, to a lesser extent, the three fusion-without-mixing events (pending considerations of sponge monophyly; Supplementary Information 7.2.6). The placement of ctenophores as sister to other animals also rejects the old notion of a Coelenterata clade that would unite ctenophores with cnidarians<sup>53</sup>.

## Myriozoans

The clade containing all sponges, bilaterians, cnidarians and placozoans is diverse, accounting for all living animals other than ctenophores. In recognition of this morphological diversity, we propose that this

clade be called Myriozoa, from the Greek myria (extremely great in number) and zoa (animals) (Fig. 4g). While Myriozoa is supported by shared derived chromosomal fusions, there are currently no obvious morphological characters that unite them. The name Benthozoa was proposed for this clade<sup>23</sup> on the basis of the inference of a pelagic ancestral metazoan and a derived benthic adult ancestor of the clade sister to ctenophores, but a benthic life history stage may not be a shared derived feature of this clade. In particular, it would be just as parsimonious for the ancestor of Metazoa to have had a benthic stage, and for most ctenophores to have lost it. We therefore prefer to avoid any assumption of the ancestral life history strategy in referring to the clade.

## Parahoxozoans, sponges and placozoans

A clade grouping bilaterians, placozoans and cnidarians to the exclusion of sponges and ctenophores<sup>54</sup> has been recovered in multiple phylogenetic studies<sup>8,10</sup> and is now called Parahoxozoa on the basis of the shared presence of Hox/ParaHox-class genes<sup>14</sup>. Parahoxozoa is supported in our analysis by the disposition of the ancestral myriozoan linkage groups Ea and G, which are each partitioned across two chromosomes in non-metazoan outgroups and ctenophores. The pre-myriozoan partitions of Ea and G are fused in demosponges and parahoxozoans, but are mixed only in parahoxozoans, providing a candidate parahoxozoan synapomorphy. The most parsimonious interpretation is that fusions forming Ea and G occurred without mixing on the myriozoan stem, a state that is preserved in demosponges, but that mixing occurred on the parahoxozoan stem lineage so that the

mixed state is shared by all bilaterians, cnidarians and placozoans. However, a detailed understanding of the history of Ea and G linkages in sponges will require chromosome-scale genome sequences from other sponge classes beyond demosponges and lyssacinosid glass sponges (Supplementary Information 7.2.2, 7.2.5 and 7.2.6). If sponges are monophyletic (as supported by recent phylogenomic studies<sup>8,10,55</sup>), then the four fusions-with-mixing that are found in demosponges and parahoxozoans must be shared by all sponges. However, if one or more sponge classes branched before the split between the demosponge and parahoxozoan lineages, it is possible that the descendants of the early-branching sponges might not possess one or more of these myriazoan fusions-with-mixing.

Although the subchromosomal assemblies currently available for *Trichoplax* preclude its full integration into the present analysis, Fig. 3 shows that placozoans share the diagnostic myriazoan fusion-with-mixing characters related to ALG\_C1 and the two bilaterian–cnidarian fusions-with-mixing related to ALGs A1a and G. The placozoa-sister-to-other-animals hypothesis<sup>56</sup> is rejected by the placement of placozoans within Myriazoa using synteny. It is therefore a strong prediction of our overall approach that chromosome-scale assemblies of placozoans will show that they share the fusions and mixing events that define Myriazoa. Furthermore, we previously showed that cnidarians and placozoans are united as sister lineages to the exclusion of bilaterians and sponges based on the mixing of genes from ALG\_Ea and ALG\_F found on cnidarian chromosomes and placozoan scaffolds<sup>57</sup>, consistent with recent gene trees<sup>55</sup>. These characters do not appear in the present analysis owing to the stringent requirement that syntenies considered here are also preserved in outgroups to Metazoa. If placozoans are nested within Parahoxozoa, homologies between the mouth, gut and nervous systems of cnidarians and bilaterians imply that placozoans are secondarily flattened and have lost an ancestral nervous system, rather than representing the ancestral parahoxozoan state.

### Implications for early animal evolution

Finally, we consider implications of the ctenophore-sister hypothesis for early animal evolution<sup>1,2</sup>. Comparisons among diverse genomes have identified numerous genes that are present in myriazoans but are absent in ctenophores<sup>1,5,58</sup>. Under the ctenophore-sister scenario, these are most parsimoniously interpreted as arising on the myriazoan stem after the divergence of ctenophores<sup>1,5,59</sup>, and include genes associated, in bilaterians and cnidarians, with neuronal function<sup>1,5,59</sup>, development<sup>58</sup> and cell adhesion<sup>60</sup>. However, as gene loss is common throughout animal evolution<sup>61</sup>, it is also possible that some of these genes were present in the ancestral metazoan but lost in ctenophores. Similarly, some genes are present in ctenophores and parahoxozoans but absent in sponges<sup>58</sup>, and these must be interpreted as gene losses on the sponge lineage.

Perhaps the most intriguing suite of metazoan characters pertain to neuromuscular systems, which are present in varying complexity in ctenophores, bilaterians, and cnidarians but are absent in sponges<sup>3,6,59</sup>. In sponge-sister scenarios, these characters are interpreted as being primitively absent, arising after the divergence of sponges on the stem lineage leading to other animals. However, in the ctenophore-sister scenario supported here by deeply conserved syntenies, there are two possible alternatives explaining the evolution of neurons: either complex neural systems arose more than once<sup>3,59,62</sup> but were elaborated differently in ctenophores, cnidarians and bilaterians<sup>3,4,59</sup>, or neuronal cell types were present in the metazoan ancestor but were lost in the sponge lineage<sup>4,9,63</sup>.

Sponge-sister and ctenophore-sister hypotheses are sometimes erroneously interpreted as suggesting that the most recent common ancestor of animals was sponge-like or ctenophore-like. We must be mindful, however, that the living representatives of sponges, ctenophores, bilaterians and placozoans may be poor surrogates for the earliest members of each stem-lineage, as the crown group of each clade arose hundreds of millions of years after their divergence from

each other, let alone from the common metazoan ancestor<sup>2</sup>. Although living sponges are often defined by the cellular, morphological and developmental characters that they lack relative to other animals, they are complex animals in their own right, successfully adapted to a unique benthic filter-feeding lifestyle<sup>13</sup>. Consistent with a neuron-bearing metazoan ancestor, sponges possess secretory cell types<sup>15</sup> and extensive molecular components associated with presynaptic function that could be derived from a primitive neurosecretory cell. Conversely, the elaborate and divergent nervous systems of living ctenophores, bilaterians, and cnidarians do not represent the stem ancestors of these groups, which would have had very different lifestyles in the Ediacaran. The nervous systems of living ctenophores, cnidarians, and bilaterians each have unique properties<sup>6,19,59</sup>, and could represent divergent evolution from a simpler neuron-bearing common ancestor. With the ctenophore-sister topology in hand, reconstructing the characters of this metazoan ancestor will require an improved understanding of molecular, cellular and system homologies and specializations across the full range of animal diversity.

### Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-05936-6>.

1. Ryan, J. F. et al. The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science* **342**, 1242592 (2013).
2. Halanych, K. M. The ctenophore lineage is older than sponges? That cannot be right! Or can it? *J. Exp. Biol.* **218**, 592–597 (2015).
3. Jékely, G., Paps, J. & Nielsen, C. The phylogenetic position of ctenophores and the origin(s) of nervous systems. *Evodevo* **6**, 1 (2015).
4. Ryan, J. F. & Chiodin, M. Where is my mind? How sponges and placozoans may have lost neural cell types. *Philos. Trans. R. Soc. Lond. B* **370**, 20150059 (2015).
5. Moroz, L. L. et al. The ctenophore genome and the evolutionary origins of neural systems. *Nature* **510**, 109–114 (2014).
6. Burkhardt, P. Ctenophores and the evolutionary origin(s) of neurons. *Trends Neurosci.* **45**, 878–880 (2022).
7. Philippe, H. et al. Phylogenomics revives traditional views on deep animal relationships. *Curr. Biol.* **19**, 706–712 (2009).
8. Simion, P. et al. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr. Biol.* **27**, 958–967 (2017).
9. Dunn, C. W. et al. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749 (2008).
10. Whelan, N. V. et al. Ctenophore relationships and their placement as the sister group to all other animals. *Nat. Ecol. Evol.* **1**, 1737–1746 (2017).
11. Li, Y., Shen, X.-X., Evans, B., Dunn, C. W. & Rokas, A. Rooting the animal tree of life. *Mol. Biol. Evol.* **38**, 4322–4333 (2021).
12. Simakov, O. et al. Deeply conserved synteny and the evolution of metazoan chromosomes. *Sci. Adv.* **8**, eabi5884 (2022).
13. Dunn, C. W., Leys, S. P. & Haddock, S. H. D. The hidden biology of sponges and ctenophores. *Trends Ecol. Evol.* **30**, 282–291 (2015).
14. Ryan, J. F. et al. The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxozoa. *Evodevo* **1**, 9 (2010).
15. Musser, J. M. et al. Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *Science* **374**, 717–723 (2021).
16. Harbison, G. R. in *The Origins and Relationships of Lower Invertebrates* Systematics Association Special Vol. 28 (eds Morris, S. C. et al.) 78–100 (Clarendon Press, 1985).
17. Tamm, S. L. Formation of the statolith in the ctenophore *Mnemiopsis leidyi*. *Biol. Bull.* **227**, 7–18 (2014).
18. Burton, P. M. Insights from diploblasts; the evolution of mesoderm and muscle. *J. Exp. Zool. B* **310**, 5–14 (2008).
19. Sachkova, M. Y. et al. Neuropeptide repertoire and 3D anatomy of the ctenophore nervous system. *Curr. Biol.* **31**, 5274–5285 (2021).
20. Antcliffe, J. B., Callow, R. H. T. & Brasier, M. D. Giving the early fossil record of sponges a squeeze. *Biol. Rev. Camb. Philos. Soc.* **89**, 972–1004 (2014).
21. O'Brien, L. J. & Caron, J.-B. A new stalked filter-feeder from the middle Cambrian Burgess Shale, British Columbia, Canada. *PLoS ONE* **7**, e29233 (2012).
22. Mah, J. L. & Leys, S. P. Think like a sponge: the genetic signal of sensory cells in sponges. *Dev. Biol.* **431**, 93–100 (2017).
23. Erives, A. & Fritzs, B. A screen for gene paralogies delineating evolutionary branching order of early metazoa. *G3* **10**, 811–826 (2020).
24. Shen, X.-X., Hittinger, C. T., Rokas, A., Minh, B. Q. & Braun, E. L. Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat. Ecol. Evol.* **1**, 0126 (2017).

25. Francis, W. R. & Canfield, D. E. Very few sites can reshape the inferred phylogenetic tree. *PeerJ* **8**, e8865 (2020).

26. Kapli, P., Yang, Z. & Telford, M. J. Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* **21**, 428–444 (2020).

27. Renwick, J. H. The mapping of human chromosomes. *Annu. Rev. Genet.* **5**, 81–120 (1971).

28. Muller, H. J. in *The New Systematics* (ed. Huxley, J.) 185–268 (Oxford Univ. Press, 1940).

29. Drosophila 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**, 203–218 (2007).

30. Hane, J. K. et al. A novel mode of chromosomal evolution peculiar to filamentous Ascomycete fungi. *Genome Biol.* **12**, R45 (2011).

31. Wright, S. On the probability of fixation of reciprocal translocations. *Am. Nat.* **75**, 513–522 (1941).

32. Lv, J., Havlak, P. & Putnam, N. H. Constraints on genes shape long-term conservation of macro-synteny in metazoan genomes. *BMC Bioinform.* **12**, S11 (2011).

33. Hillis, D. M. SINEs of the perfect character. *Proc. Natl. Acad. Sci. USA* **96**, 9979–9981 (1999).

34. Rokas, A. & Holland, P. W. Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* **15**, 454–459 (2000).

35. Okada, N. SINEs: Short interspersed repeated elements of the eukaryotic genome. *Trends Ecol. Evol.* **6**, 358–361 (1991).

36. Martini, S., Schultz, D. T., Lundsten, L. & Haddock, S. H. D. Bioluminescence in an undescribed species of carnivorous sponge (Cladorhizidae) from the deep sea. *Front. Mar. Sci.* **7**, 576476 (2020).

37. Kenny, N. J. et al. Tracing animal genomic evolution with the chromosomal-level assembly of the freshwater sponge *Ephydatia muelleri*. *Nat. Commun.* **11**, 3676 (2020).

38. Schuster, A. et al. Divergence times in demosponges (Porifera): first insights from new mitogenomes and the inclusion of fossils in a birth-death clock model. *BMC Evol. Biol.* **18**, 114 (2018).

39. McKenna, V. et al. The Aquatic Symbiosis Genomics Project: probing the evolution of symbiosis across the tree of life. *Wellcome Open Res.* **6**, 254 (2021).

40. Schultz, D. T. et al. A chromosome-scale genome assembly and karyotype of the ctenophore *Hormiphora californiensis*. *G3* **11**, jkab302 (2021).

41. Johnson, S. B. et al. Speciation of pelagic zooplankton: invisible boundaries can drive isolation of oceanic ctenophores. *Front. Genet.* **13**, 970314 (2022).

42. Fairclough, S. R. et al. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate *Salpingoeca rosetta*. *Genome Biol.* **14**, R15 (2013).

43. Suga, H. et al. The *Capsaspora* genome reveals a complex unicellular prehistory of animals. *Nat. Commun.* **4**, 2325 (2013).

44. de Mendoza, A., Suga, H., Permanyer, J., Irimia, M. & Ruiz-Trillo, I. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *eLife* **4**, e08904 (2015).

45. Kenny, N. J. et al. The gene-rich genome of the scallop *Pecten maximus*. *Gigascience* **9**, giao037 (2020).

46. Li, Y. et al. Chromosome-level reference genome of the jellyfish *Rhopilema esculentum*. *Gigascience* **9**, giao036 (2020).

47. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).

48. Lang, B. F., O'Kelly, C., Nerad, T., Gray, M. W. & Burger, G. The closest unicellular relatives of animals. *Curr. Biol.* **12**, 1773–1778 (2002).

49. Torruella, G. et al. Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Mol. Biol. Evol.* **29**, 531–544 (2012).

50. Strasser, J. F. H., Irisarri, I., Williams, T. A. & Burki, F. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat. Commun.* **12**, 1879 (2021).

51. Schubert, I. & Lysak, M. A. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet.* **27**, 207–216 (2011).

52. Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).

53. Hyman, L. H. The phylogeny of the lower metazoa. *Q. Rev. Biol.* **51**, 344–347 (1976).

54. Collins, A. G. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. *Proc. Natl. Acad. Sci. USA* **95**, 15458–15463 (1998).

55. Laumer, C. E. et al. Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. Biol. Sci.* **286**, 20190831 (2019).

56. Schierwater, B. My favorite animal, *Trichoplax adhaerens*. *Bioessays* **27**, 1294–1302 (2005).

57. Simakov, O. et al. Deeply conserved synteny resolves early events in vertebrate evolution. *Nat. Ecol. Evol.* **4**, 820–830 (2020).

58. Fernández, R. & Gabaldón, T. Gene gain and loss across the metazoan tree of life. *Nat. Ecol. Evol.* **4**, 524–533 (2020).

59. Moroz, L. L. Multiple origins of neurons from secretory cells. *Front. Cell Dev. Biol.* **9**, 669087 (2021).

60. Belahbib, H. et al. New genomic data and analyses challenge the traditional vision of animal epithelium evolution. *BMC Genom.* **19**, 393 (2018).

61. Albalat, R. & Cañestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **17**, 379–391 (2016).

62. Parker, G. H. *The Elementary Nervous System* (J.B. Lippincott Company, 1919).

63. Leys, S. P. Elements of a 'nervous system' in sponges. *J. Exp. Biol.* **218**, 581–591 (2015).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

## Methods

A full description of the methods is provided in the Supplementary Information.

### Unicellular outgroup species genome scaffolding

Chromatin conformation capture (Hi-C) libraries were generated<sup>64</sup> from frozen cell cultures obtained directly from the American Type Culture Collection. The cultures used were of the species *C. owczarzaki* (ATCC, 30864), *C. fragrantissima* (ATCC, PRA-284) and *S. rosetta* (ATCC, PRA-366). The strains used were the same as those sequenced in the original genome assembly projects for each species<sup>42–44</sup>. The Hi-C libraries were sequenced at a depth of over 500× for each species on the Illumina NovaSeq 6000 system at MedGenome.

Previously published draft genome assemblies<sup>42,44</sup> were scaffolded to chromosome-scale using a combination of HiRIse (v.Aug2019)<sup>65</sup> and SALSA2 (v.2.3)<sup>66</sup>. The genomes were manually curated using PretextView v.0.2.4 (<https://github.com/wtsi-hpag/PretextView>), HiGlass v.1.10.0104<sup>67</sup>, Juicebox Assembly Tools (GitHub Commit 46c7ed1105)<sup>68</sup>, the Juicebox visualization system (v.1.11.08106)<sup>69</sup> and artisanal (<https://bitbucket.org/bredeson/artisanal/src>). For *C. owczarzaki*, we used the most recent ‘v4’ assembly as input for scaffolding<sup>70</sup>. The Hi-C data were used as evidence to remove several megabases of the original *C. fragrantissima* assembly that, after further analysis, appeared to be fungal contaminants. We identified the general location of the centromeres in *C. fragrantissima* and *C. owczarzaki* using the Hi-C data as described in Supplementary Information 3.

### Sponge and ctenophore genome assembly

Samples of *B. microptera*<sup>41</sup> were collected in Monterey Bay, California (36.63° N, 121.90° W) from surface waters and were reared to an F<sub>3</sub> population at the Monterey Bay Aquarium, from which one adult was sequenced. One individual cladorhizid sponge<sup>36</sup> was collected off the coast of Big Sur, California (35.49° N, 124° W) from the seafloor at 3,975 m. One hexactinellid ‘tulip’ sponge (HEX) was collected near Southern California (34.57° N, 122.56° W) from the seafloor at 3,852 m. This species of ctenophore, and presumably these species of sponges, are hermaphroditic. Sponge and ctenophore samples were collected under the State of California Department of Fish and Wildlife collecting permits SC-2026 (*Bolinopsis*) and SC-4029 (sponges).

DNA and RNA were isolated from these species to generate Pacific Biosciences (PacBio) CLR WGS, HiFi WGS libraries or PacBio Iso-Seq libraries at the Brigham Young University DNA Sequencing Center. These libraries were sequenced on the PacBio Sequel II system. Illumina WGS libraries, Chicago libraries and Hi-C libraries were generated at UC Santa Cruz and sequenced at MedGenome on the Illumina HiSeq X system. PacBio WGS library coverage was over 70× for all three species, and Hi-C coverage was over 190× for all three species. Genome sizes were estimated using jellyfish (v.2.2.10)<sup>71</sup>, then using the resulting spectrum in GenomeScope (v.2)<sup>72</sup>.

The genome of *B. microptera* was assembled using wtdbg (v.2.4)<sup>73</sup>, and the sponge genomes were assembled using hifiasm (v.0.16.1-r375)<sup>74</sup>. Hi-C reads were mapped using bwa mem (v.0.7.17)<sup>75</sup>, processed using pairtools (v.0.3.0)<sup>76</sup>, pairix (v.0.3.7; <https://github.com/4dn-dcic/pairix>) and Cooler (v.0.8.10)<sup>77</sup>, and scaffolding was performed using HiRIse (v.Aug2019)<sup>65</sup>. In *B. microptera*, gaps were closed using TGS-Gapcloser (v.1.1.1)<sup>78</sup>, haplotigs were removed using Purge Haplontigs (v.1.0.4)<sup>79</sup> and the assembly was polished using Illumina WGS reads and pilon (v.1.23)<sup>80</sup>. In both the sponge and *B. microptera* genomes, bacterial scaffolds were removed using Diamond (v.0.9.24)<sup>81</sup> and Blobtools (v.1.0)<sup>82</sup>. The genomes were manually curated with Hi-C data as described above. The haplotypes of the hifiasm-based assemblies were compared to one another using D-Genies (v.1.4.0)<sup>83</sup>.

### Genome annotations

The unicellular outgroup genome assemblies were annotated by mapping their transcripts from the original assemblies to the Hi-C scaffolded

assemblies using minimap2 (v.2.23)<sup>84</sup>. To clarify demosponge macrosyntenic relationships, we produced putative *Ephydatia* protein coordinates in the cladorhizid sponge using tblastn (v.2.10.0+)<sup>85</sup>. To annotate the hexactinellid sponge genome, we mapped the proteins of closely-related hexactinellid species<sup>86,87</sup> using miniprot (v.0.2)<sup>88</sup> (Supplementary Information 2.1.5). The *Bolinopsis* genome was annotated using BRAKER (v.2.14)<sup>89</sup> supplied with evidence from RNA-seq reads mapped with STAR (v.2.7.1a)<sup>90</sup> and minimap2 (v.2.23)<sup>84</sup>. Iso-Seq reads processed with lima (v.2.2.0; <https://github.com/PacificBiosciences/barcoding>) and isoseq3 (v.3.4.0; <https://github.com/PacificBiosciences/IsoSeq>) then mapped with minimap2 (v.2.23)<sup>84</sup>, and protein orthology identified using ProtHint (v.2.6.0)<sup>91</sup> from ctenophore transcriptomes<sup>92–94</sup> assembled with Trinity (v.2.5.1)<sup>95</sup> and translated using TransDecoder (v.5.5; <https://github.com/TransDecoder/TransDecoder>). We assessed genome sequence and protein datasets using BUSCO (v.5)<sup>96</sup>.

### Orthologue Inference

Orthologues were inferred between species by finding reciprocal-best BLASTp<sup>97</sup> hits between the proteins in the genomes, or with OrthoFinder (v.2.3.7)<sup>98</sup>. The reciprocal-best BLASTp hits were used to identify macrosyntenic chromosomes between species by performing Bonferroni-corrected one-sided Fisher’s exact tests<sup>57</sup>. To determine the provenance of the ALG\_H in sponges, the genomes of *Chondrosia* and *Petrosia*<sup>39,99</sup>, *Opsacas*<sup>86</sup>, CLA and HEX were compared using the odp software suite.

Orthologues shared between three, four or more species were selected by finding groups of proteins that were *n*-way reciprocal best BLASTp hits. In this conservative method, each orthogroup has a single protein from each of the *n* species. We performed this analysis for three-way and four-way comparisons of combinations of the species CFR, COW, SRO, HCA, EMU, CLA, RES, BFL, NVE and *P. maximus*.

### Gene linkage group identification

Orthologues from three-way or four-way reciprocal-best BLASTp searches were grouped by the chromosomes on which the genes occurred in the *n* species. To identify which sets of orthologues were larger than expected by random chance, we shuffled the genome coordinates of the *n* species and measured the frequency of finding sets of orthologues of size *k* on the same chromosomes in the *n* species. By performing this for 10 million iterations, we calculated the false-discovery rate (*α*) of finding an orthologue set of size *k* given the *n* input genomes.

### Combined unicellular outgroup analysis

Sets of orthologues with a false-discovery rate of less than 0.05 were retained from the four-way reciprocal best hit searches of COW–HCA–EMU–RES, CFR–HCA–EMU–RES and SRO–HCA–EMU–RES. The remaining orthogroups were joined based on gene identity in HCA–EMU–RES, such that each orthologue contained a protein from at least one of the unicellular outgroup species. This yielded 291 sets of orthologues.

### Identification of orthologues in other species

For each of the 291 orthologues, we aligned the proteins using MAFFT (v.7.310)<sup>100</sup>, built a hidden Markov Model using hmmbuild in hmmer (v.3.3.2)<sup>101</sup>, then found the best match using hmmsearch in the proteins of the genomes of other species, including the ctenophore *B. microptera*, the cladorhizid sponge, *T. adhaerens*<sup>102</sup>, *H. vulgaris*<sup>12</sup>, *N. vectensis*<sup>103</sup>, *B. floridae*<sup>57</sup>, *P. maximus*<sup>45</sup> and *E. muelleri*<sup>37</sup>. To test for Gene Ontology enrichment of the sets of orthogroups using PANTHER (v.17)<sup>104</sup>, we also searched for the orthologues in *Homo sapiens*<sup>105</sup>.

### Mixing analysis

To test whether the \_x and \_y gene sets present on single chromosomes were well-mixed, we used a metric that counts the number of transitions between a gene in \_x to a gene in \_y and vice versa. To provide an

objective measure of mixing, we computed the  $\alpha$  value (false-discovery rate) that the two sets of genes are unmixed by building a distribution of mixing scores from randomly sorted groups of the same size of the  $_x$  and  $_y$  groups in question. We consider  $\alpha < 0.05$  to be unmixed.

### Simulations testing the ctenophore-sister and sponge-sister hypotheses

We applied this methodology to test whether the findings supporting the ctenophore-sister hypothesis were due to the arrangement of any of the observed genomes, implemented as part of the odp software suite. For both the SRO-HCA-EMU-RES and COW-HCA-EMU-RES four-way reciprocal best hit results, we performed four analyses. One analysis shuffles the genome chromosome labels of one species 100 million times. Each time the genome chromosome labels are shuffled, we perform the gene linkage group identification analysis described above, and measure the quantity and size of gene linkage groups that support either the ctenophore-sister or sponge-sister hypothesis. The distribution of these results compared with the observed data of the real genomes is used to estimate the false-discovery rate of finding support for the ctenophore-sister hypothesis or sponge-sister hypothesis. We modelled fusion-with-mixing events in the animal genomes as state transitions, and used RevBayes (v.1.1.1)<sup>106</sup> and MrBayes (v.3.2.7a)<sup>52</sup> to estimate the likelihood of the ctenophore-sister hypothesis, and we used FigTree (v.1.4.4; <https://github.com/rambaut/figtree>) to visualize the trees.

### Software

We implemented a suite of tools for identifying orthologues, plotting synteny relationships and performing synteny-based phylogenetic analyses using a tool called odp, implemented in snakemake (v.7)<sup>107</sup> for scalability. To confirm the validity of these methods, we applied them to several genome quartets and showed that odp recovers previously identified synapomorphic chromosomal fusion-with-mixing events<sup>12</sup> in bilaterians and cnidarians (Supplementary Information 6).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data presented in this Article are available in public repositories. The sequencing reads are available in the NCBI database under BioProject accession numbers PRJNA818620, PRJNA818630, PRJNA903214 and PRJNA818537. The genomes for each species are available through the above BioProject accession codes, with the exception of the genomes of *C. fragrantissima*, *C. owczarzaki* and *S. rosetta*, which are available at Dryad (<https://doi.org/10.5061/dryad.dncjsxm47>). The results shown in the Supplementary Information, when not contained in figures, are also available in the Dryad repository (<https://doi.org/10.5061/dryad.dncjsxm47>). Publicly available sequencing data and genomes were downloaded from NCBI from BioProject accession numbers PRJNA168, PRJDB8655, PRJNA12874, PRJNA20249, PRJNA20341, PRJEB28334, PRJNA30931, PRJNA31257, PRJNA37927, PRJEB56075, PRJEB56892, PRJNA64405, PRJNA193541, PRJNA193613, PRJNA213480, PRJNA278284, PRJNA281977, PRJNA283290, PRJNA377365, PRJNA396415, PRJNA512552, PRJNA544471, PRJNA576068, PRJNA579531, PRJNA625562, PRJNA667495, PRJNA761294 and PRJNA814716. The *E. muelleri* genome was downloaded from <https://spaces.facsci.ualberta.ca/ephybase/>.

### Code availability

The scripts and software developed for this manuscript are available at Dryad (<https://doi.org/10.5061/dryad.dncjsxm47>) and Zenodo (<https://doi.org/10.5281/zenodo.7857390>). Long-term development of odp is available at GitHub (<https://github.com/conchoecia/odp>).

64. Adams, M. et al. One fly-one genome: chromosome-scale genome assembly of a single outbred *Drosophila melanogaster*. *Nucleic Acids Res.* **48**, e75 (2020).
65. Putnam, N. H. et al. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* **26**, 342–350 (2016).
66. Ghurye, J., Pop, M., Koren, S., Bickhart, D. & Chin, C.-S. Scaffolding of long read assemblies using long range contact information. *BMC Genom.* **18**, 527 (2017).
67. Kerpedjiev, P. et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125 (2018).
68. Dudchenko, O. et al. The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under \$1000. Preprint at *bioRxiv* <https://doi.org/10.1101/254797> (2018).
69. Durand, N. C. et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).
70. Denbo, S. et al. Revision of the *Capsaspora* genome using read mating information adjusts the view on premetazoan genome. *Dev. Growth Differ.* **61**, 34–42 (2019).
71. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of  $k$ -mers. *Bioinformatics* **27**, 764–770 (2011).
72. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
73. Ruan, J. & Li, H. Fast and accurate long-read assembly with wtdbg2. *Nat. Methods* **17**, 155–158 (2020).
74. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
75. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://arxiv.org/abs/1303.3997> (2013).
76. Open2C et al. PairTools: from sequencing data to chromosome contacts. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.02.13.528389> (2023).
77. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genetically labeled arrays. *Bioinformatics* **36**, 311–316 (2020).
78. Xu, M. et al. TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience* **9**, giaa094 (2020).
79. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotts: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinform.* **19**, 460 (2018).
80. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).
81. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
82. Laetsch, D. R. & Blaxter, M. L. BlobTools: interrogation of genome assemblies. *F1000Res.* **6**, 1287 (2017).
83. Cabanettes, F. & Klopp, C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **6**, e4958 (2018).
84. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
85. Gertz, E. M., Yu, Y.-K., Agarwala, R., Schäffer, A. A. & Altschul, S. F. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol.* **4**, 41 (2006).
86. Santini, S. et al. The compact genome of the sponge *Oopscacas minuta* (Hexactinellida) is lacking key metazoan core genes. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.07.26.501511> (2022).
87. Whelan, N. V., Kocot, K. M., Moroz, L. L. & Halanych, K. M. Error, signal, and the placement of Ctenophora sister to all other animals. *Proc. Natl Acad. Sci. USA* **112**, 5773–5778 (2015).
88. Li, H. Protein-to-genome alignment with miniprot. *Bioinformatics* **39**, btad014 (2023).
89. Hoff, K. J., Lomsadze, A., Borodovsky, M. & Stanke, M. Whole-genome annotation with BRAKER. *Methods Mol. Biol.* **1962**, 65–95 (2019).
90. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
91. Brúna, T., Lomsadze, A. & Borodovsky, M. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genom. Bioinform.* **2**, lqaa026 (2020).
92. Francis, W. R., Shaner, N. C., Christianson, L. M., Powers, M. L. & Haddock, S. H. D. Occurrence of isopenicillin-N-synthase homologs in bioluminescent ctenophores and implications for coelenterazine biosynthesis. *PLoS ONE* **10**, e0128742 (2015).
93. Townsend, J. P. et al. A mesopelagic ctenophore representing a new family, with notes on family-level taxonomy in Ctenophora: *Vampyroctenidae*, fam. nov. *Mar. Biodivers.* **50**, 34 (2020).
94. Babonis, L. S. et al. Integrating embryonic development and evolutionary history to characterize tentacle-specific cell types in a ctenophore. *Mol. Biol. Evol.* **35**, 2940–2956 (2018).
95. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
96. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
97. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
98. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
99. The Darwin Tree of Life Project. Sequence locally, think globally: The Darwin Tree of Life Project. *Proc. Natl Acad. Sci. USA* **119**, e2115642118 (2022).
100. Katoh, K., Misawa, K., Kuma, K.-I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

# Article

101. Potter, S. C. et al. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
102. Srivastava, M. et al. The *Trichoplax* genome and the nature of placozoans. *Nature* **454**, 955–960 (2008).
103. Zimmermann, B. et al. Sea anemone genomes reveal ancestral metazoan chromosomal macrosynteny. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.10.30.359448> (2020).
104. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566 (2013).
105. Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
106. Höhna, S. et al. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* **65**, 726–736 (2016).
107. Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).

**Acknowledgements** We acknowledge the support of the David and Lucile Packard Foundation, the Monterey Bay Aquarium Research Institute and the United States National Science Foundation (NSF). D.T.S. was supported by the NSF GRFP DGE 1339067; S.H.D.H. by NSF DEB-1542679; D.T.S. and O.S. by the European Research Council's Horizon 2020: European Union Research and Innovation Programme, grant no. 945026; D.S.R. by internal funds of the OIST Molecular Genetics Unit, the Chan Zuckerberg Biohub and the Marthella Foskett Brown Chair in Biology. We thank W. Patry and the staff at Monterey Bay Aquarium for providing the *Bolinopsis* specimen; and C. Dunn, J. Eizenga, R. Revilla-i-Domingo and G. Genikhovich for

discussions. The *Petrosia* and *Chondrosia* genomes from the Aquatic Symbiosis Genomics project of the Tree of Life Programme, Wellcome Sanger Institute, were funded by the Gordon and Betty Moore Foundation and the Wellcome Trust.

**Author contributions** D.T.S., O.S. and D.S.R. designed the scientific objectives of the study. D.T.S. planned and carried out experiments and analyses, wrote the code for the analyses and the odp package, and wrote the first draft of the manuscript. O.S. and D.T.S. performed the Bayesian analyses. D.T.S. and S.H.D.H. collected samples and created figures and tables. J.V.B. assisted in genome assemblies and identified putative centromeres. D.T.S. and D.S.R. wrote the first drafts of the manuscript with input from O.S. and S.H.D.H. All of the authors contributed to the interpretation, presentation and writing of the Article and the Supplementary Information.

**Competing interests** R.E.G. and D.S.R. are paid consultants and equity holders of Dovetail Genomics. D.T.S. is a shareholder of Pacific Biosciences of California. The other authors declare no competing interests.

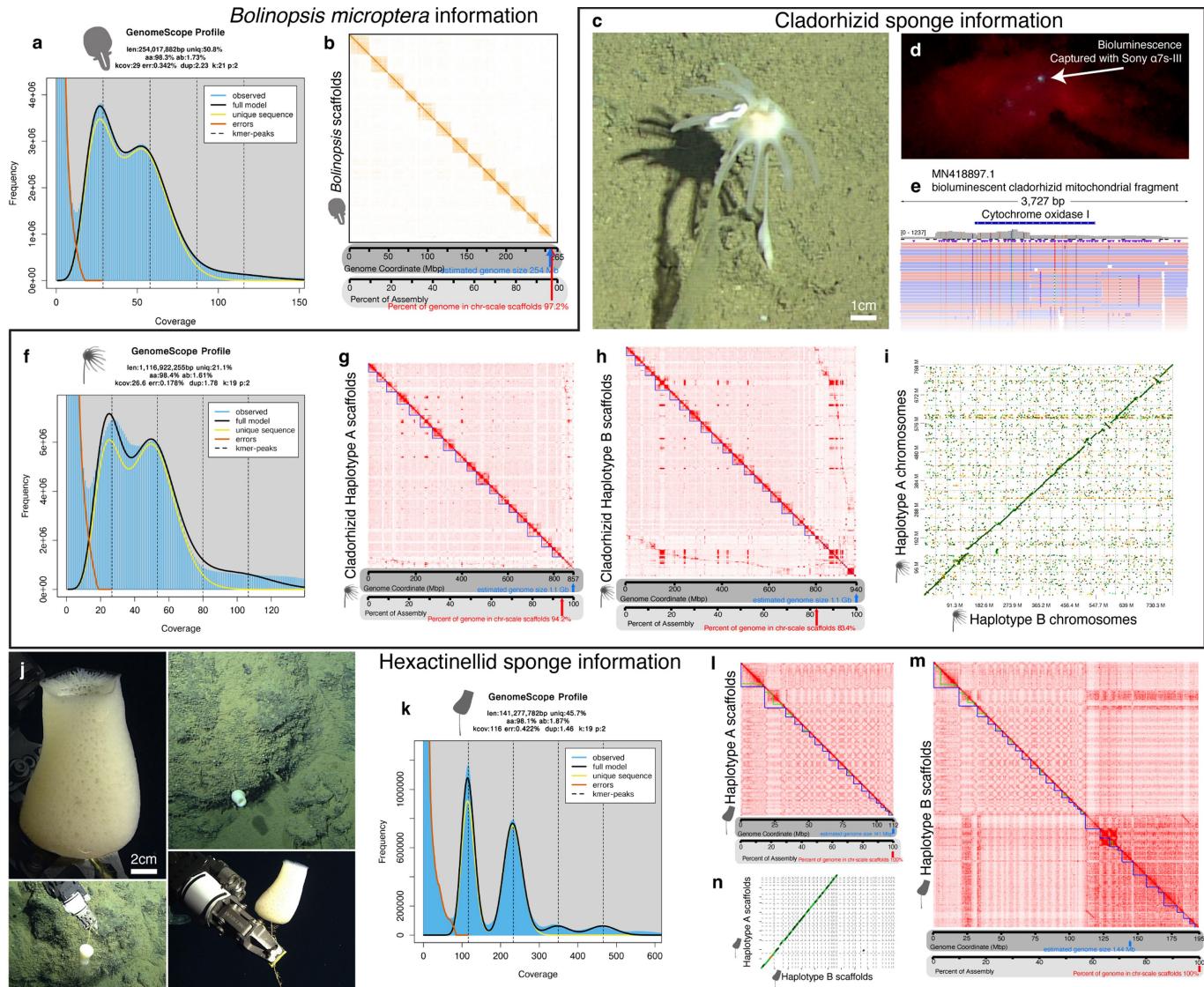
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-05936-6>.

**Correspondence and requests for materials** should be addressed to Darrin T. Schultz, Oleg Simakov or Daniel S. Rokhsar.

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

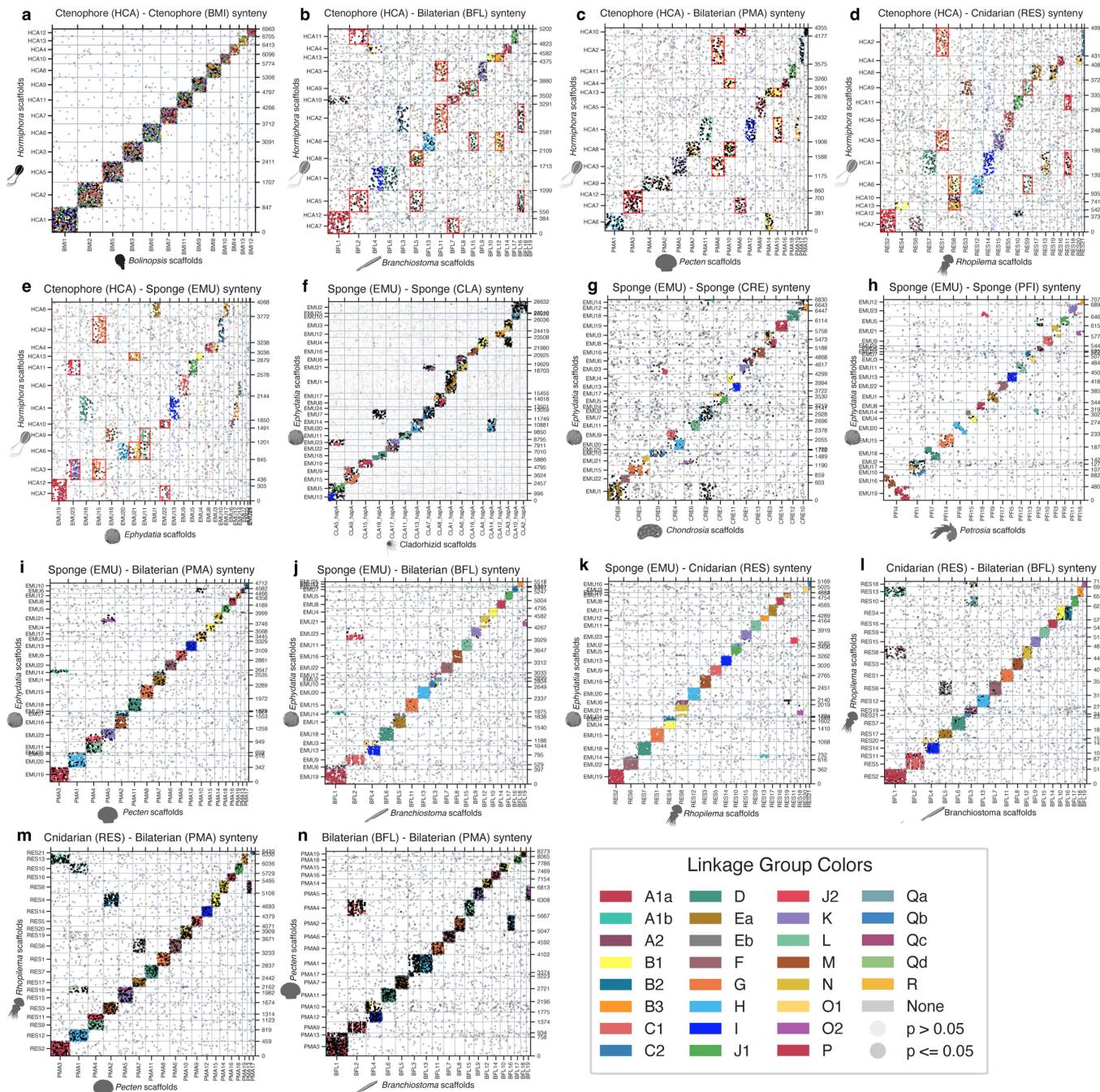
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Genomes of one ctenophore and two sponges.** **a.** The k-mer spectrum of the *Bolinopsis* data suggests that the animal is diploid, and the  $\text{In}$  genome size is approximately 254 Mbp. **b.** The *Bolinopsis microptera* genome assembly contains 13 chromosome-scale scaffolds, which account for 97.23% of the total bases in the assembly. Panel shows the Hi-C contact map. **c.** The cladorhizid sponge individual used in the genome sequencing at its collection site. **d.** This sponge was bioluminescent when mechanically disturbed. **e.** Its mitochondrial sequence is 99.2% identical to the previously identified bioluminescent cladorhizid sponge<sup>36</sup>. **f.** The estimated genome size of this sponge is 1.11 Gb, and the spectrum is consistent with diploid organisms. **g.,h.** Each haplotype assembly has 18 chromosome-scale scaffolds based on chromatin confirmation data as shown. In haplotype A 94.2% of bases

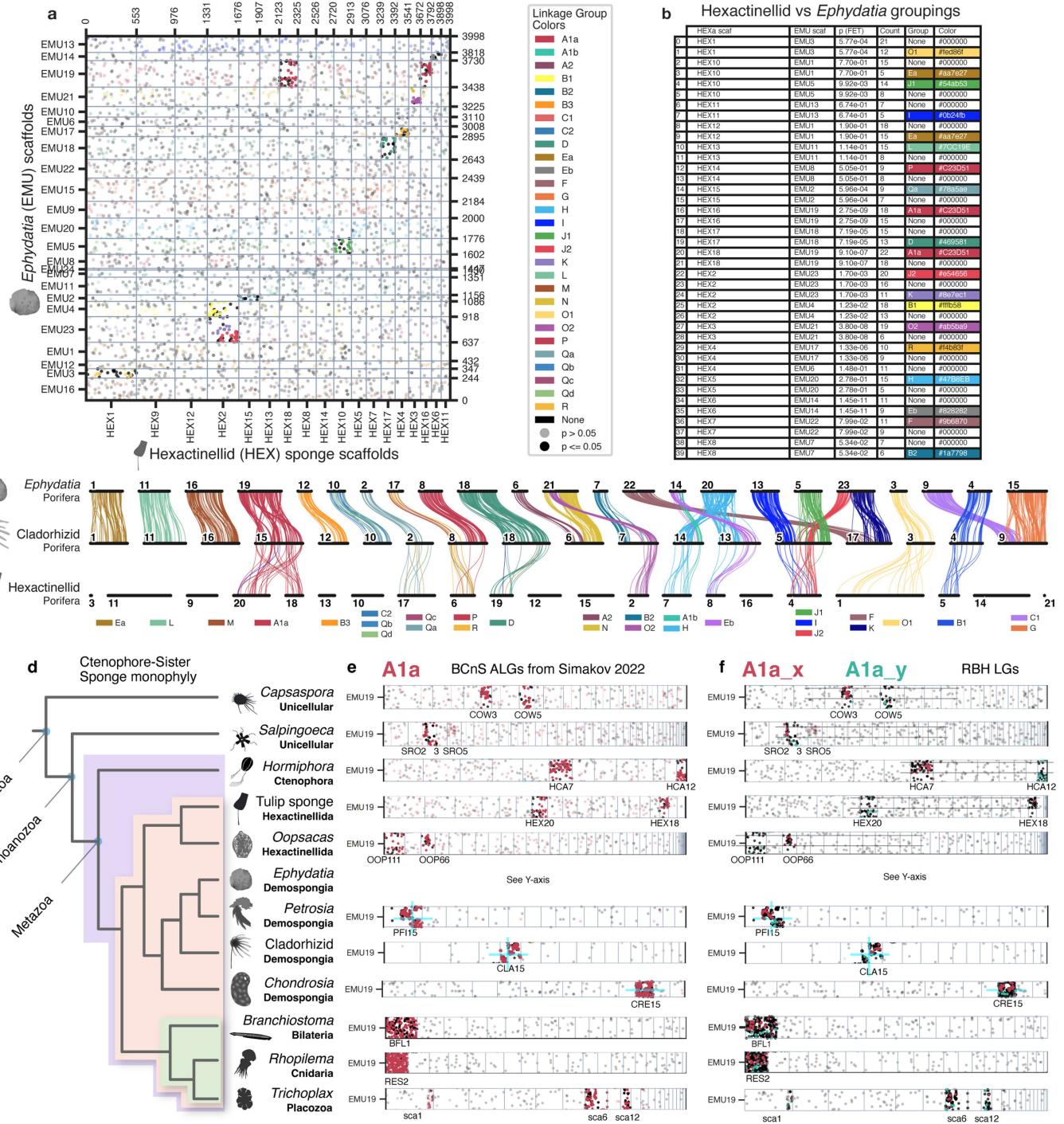
are in the chromosome-scale scaffolds. **i.** A whole-genome alignment of haplotypes A and B showed a high degree of concordance. **j.** The hexactinellid sponge collected and sequenced for this study. **k.** The estimated genome size is  $\text{In} = 141$  Mb. **l.** Haplotype A contains only one haplotype of chromosome-scale scaffolds orthologous with the scaffolds of the closely-related sponge *Opsacas minuta*. Panel shows chromatin conformation capture contact map of haplotype A. **m.** In addition to the alternate haplotype of chromosome-scale scaffolds from haplotype A, the haplotype B assembly contains the large, gene-poor, unplaced scaffolds that lack detectable homology to other sponges. The Hi-C contact map for haplotype B shown. **n.** Whole-genome alignments of the two haplotypes show colinearity. Photograph credits: (d.) Darrin Schultz, (e.,j.) © 2021 MBARI.

# Article



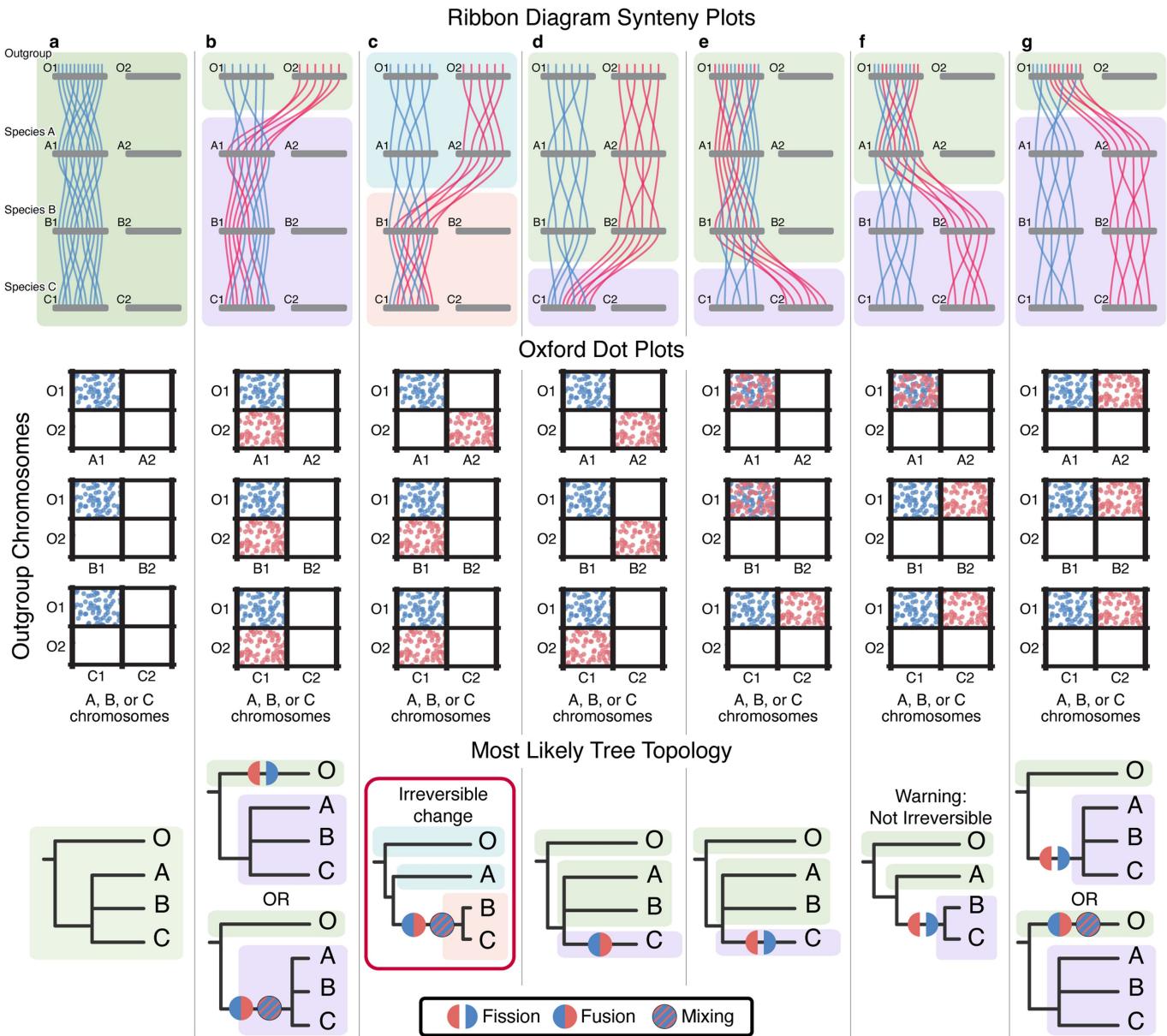
**Extended Data Fig. 2 | Chromosomes are largely conserved among metazoans.** The chromosome position of orthologous proteins plotted in panels a-n are coloured by orthologs in the previously identified ancestral bilaterian, cnidarian, and sponge linkage groups (BCnS-ALGs)<sup>12</sup>. Significant inter-species chromosome pairs ( $p \leq 0.05$ , Bonferroni-corrected one-sided Fisher's exact test<sup>37</sup>) are opaque. **a.** The karyotype of the Pleurobrachiid and lobate ctenophores is conserved ( $In = 13$ ). **b-e.** Ctenophore chromosomes share macrosynteny with BCnS-ALGs, but many BCnS-ALGs are split onto

several ctenophore chromosomes (red dotted boxes). There are many ctenophore-specific chromosome fusions. **f-h.** Macrosynteny is highly conserved between distantly-related demosponges. The sponge lineages shown diverged an estimated 358 Mya - 500 Mya<sup>38</sup>. **f-k.** Macrosynteny is also conserved between sponge, bilaterian, and cnidarian genomes. Many chromosomes in a species of one clade have a one-to-one homologous chromosome in the other clade. The genomes of species in these clades can be described by 29 constituent BCnS-ALGs<sup>12</sup>.



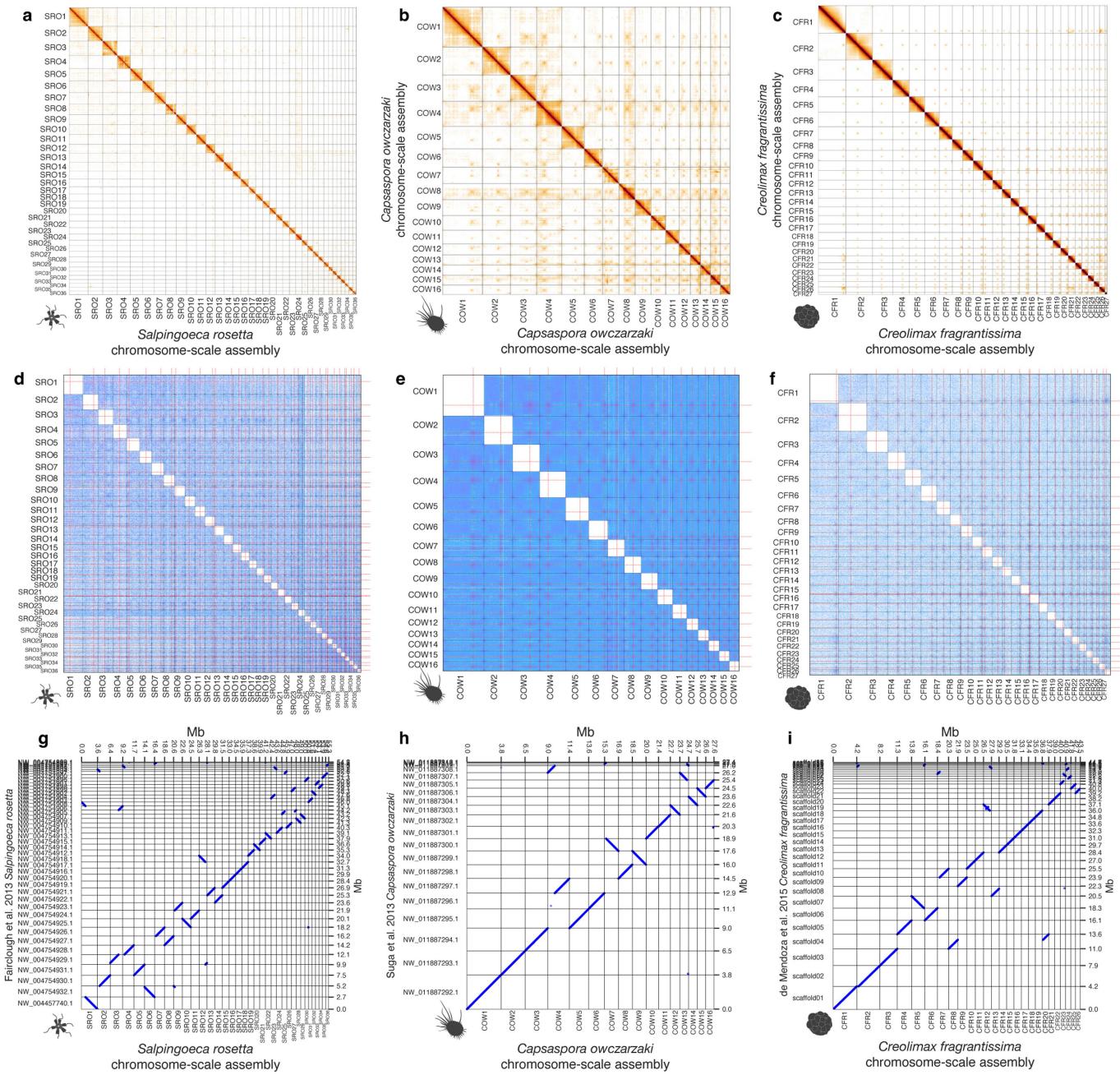
**Extended Data Fig. 3 | Sponge macrosynteny.** **a-c.** There have been many genome rearrangements since the divergence of the demosponge *Ephydatia* and the tulip hexactinellid genome, and they share macrosynteny of only some BCnS linkage groups ( $p \leq 0.05$ , Bonferroni-corrected one-sided Fisher's exact test<sup>57</sup>, opaque dots in **a**, rows in **b**, interspecies lines in **c**). **d.** The sponge cladogram is based on Schuster et al. 2018<sup>38</sup>. **e.** The orthologs in A1a\_x and A1a\_y are predominantly present on separate chromosomes in both the tulip hexactinellid and in *Oopsacas minuta*. **f.** A1a\_x and A1a\_y are on partly

overlapping regions of single demosponge chromosomes, but are mixed on a single *Chondrosia* chromosome. However, the linkage groups A1a\_x and A1a\_y are on separate chromosomes in the ctenophores and the unicellular outgroup species. This evidence suggests that hexactinellid sponges retain the ancestral state of A1a\_x and A1a\_y being present on separate chromosomes. The possible evolutionary scenarios explaining this karyotype will require further chromosome-scale sequencing of sponge genomes.



**Extended Data Fig. 4 | Seven basic ALG configurations in species quartets.**  
**a-g.** The seven configurations of ALGs found in four species highlight the evolutionary history of chromosomes. The cartoon ribbon plot in each panel shows chromosomes (horizontal bars) and the positions of genes in two ALGs

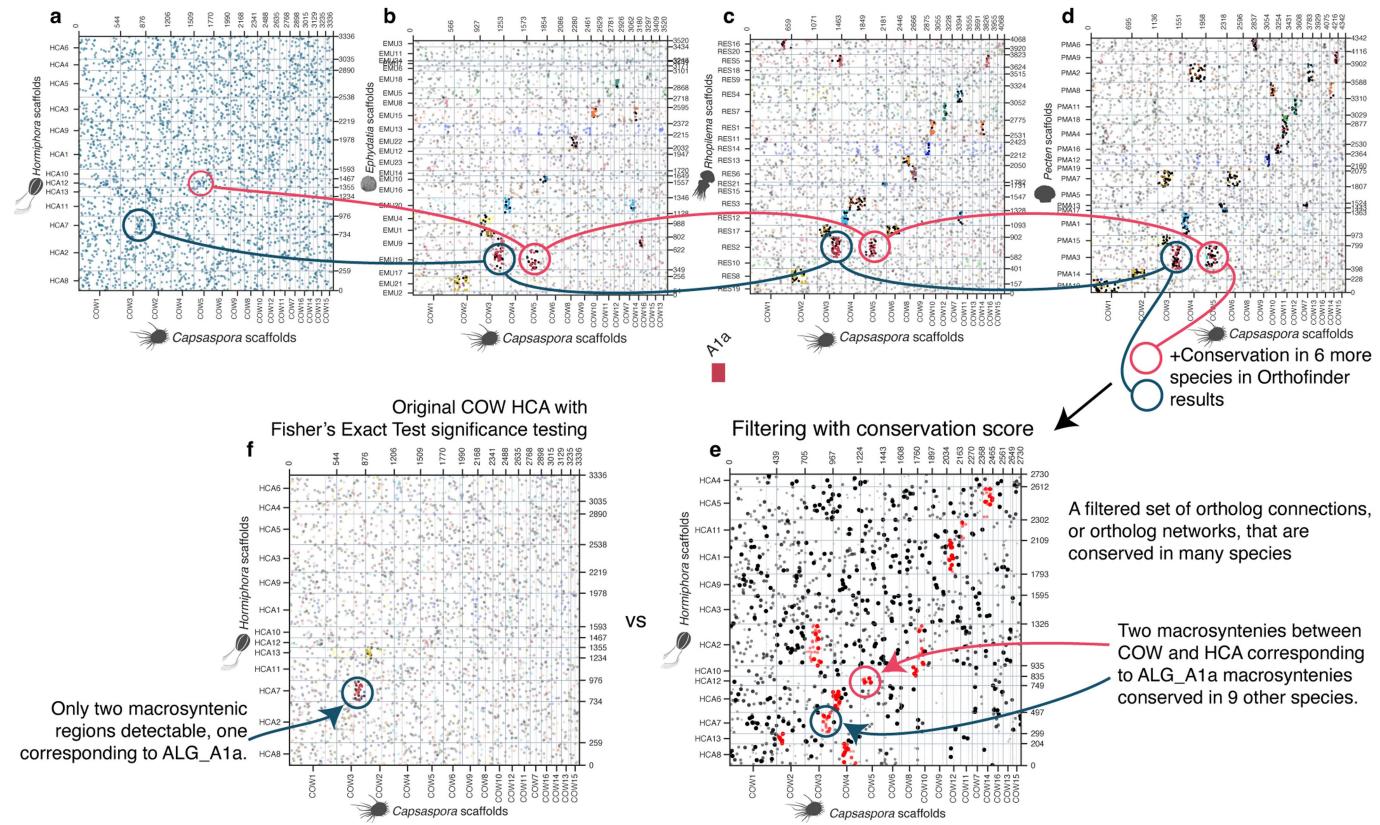
along those chromosomes (vertical blue or red lines, respectively). The cartoon Oxford dot plot in each panel shows the same information as the ribbon plot, but only in the context of the outgroup genome. The most parsimonious tree topology based on the ALG evolutionary history is also pictured.



**Extended Data Fig. 5 | Unicellular species chromosome-scale genome assemblies.** Hi-C heatmaps of **a**, *Salpingoeca rosetta*, **b**, *Capsaspora owczarzaki*, and **c**, *Creolimax fragrantissima* show that the assemblies are consistent with chromosome-scale assemblies of other unicellular species. **d-e.** Genome-wide ICE-normalized  $10^8$  observed count contact maps for **(d)**, *Salpingoeca*, **(e)**, *Capsaspora*, and **(f)**, *Creolimax* are shown at MapQ0 and 10 kb resolution. Chromosome boundaries are drawn as solid black lines. The

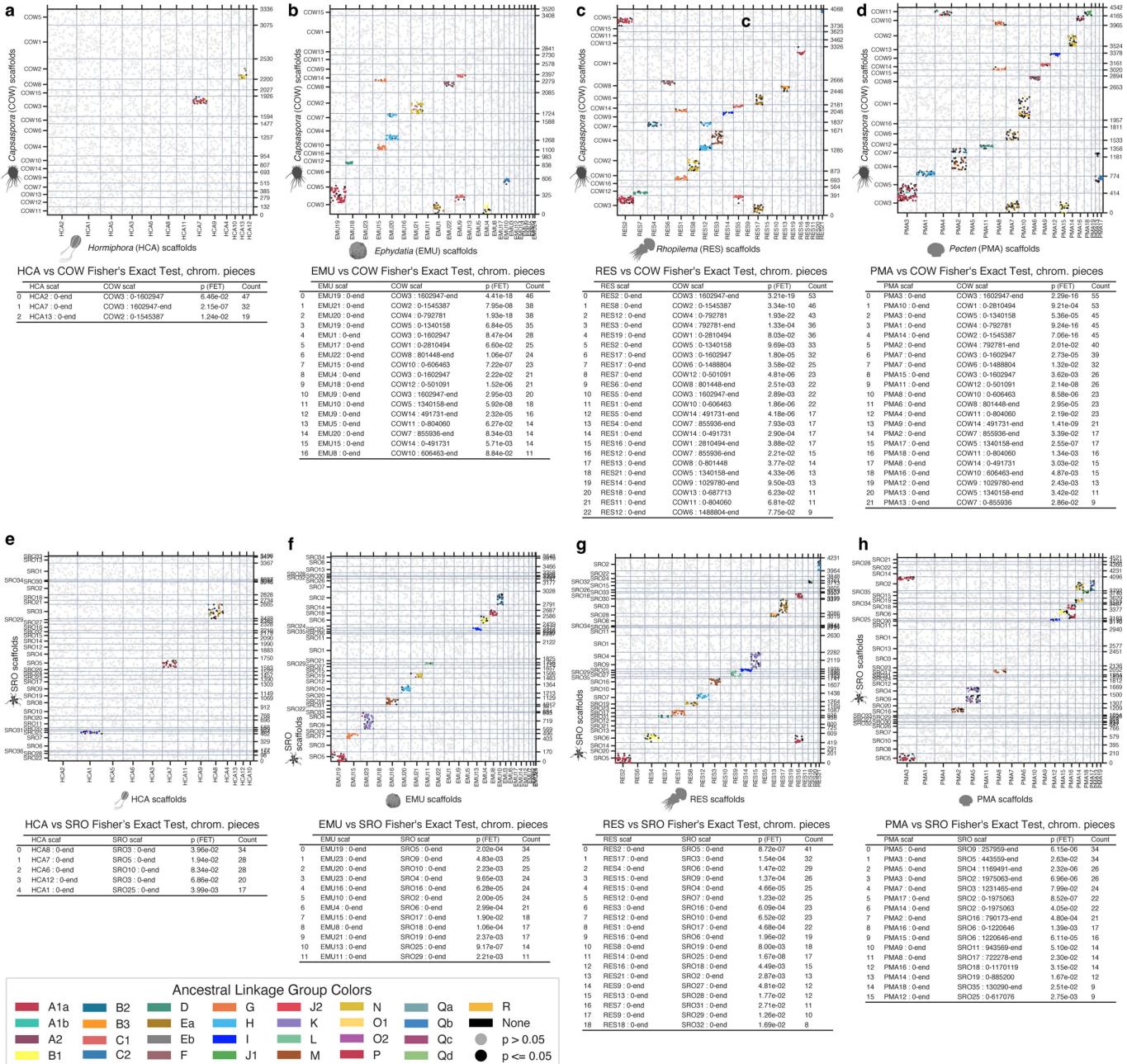
intersections of horizontal and vertical red lines mark the Centurion-estimated centromere positions. The Hi-C heatmaps of *Capsaspora* and *Creolimax* both contain inter-chromosomal hotspots that are consistent with centromeres in other species. **g-i.** Protein orthology plots (Oxford dot plots) of the chromosome-scale genome assemblies compared to the previously published assemblies. Despite the lack of Hi-C data, the original scaffold assemblies for all three species were nearly chromosome-scale.

# Article



**Extended Data Fig. 6 | Visual representation of multi-species gene linkage conservation score.** **a-d.** The dot plots of the *C. owczarzaki* genome show that there is conservation of the ALG\_A1a linkage group in ctenophore, sponge, cnidarian, and bilaterian genomes. The conservation score can be calculated from shared gene linkages across many species. **f.** Due to the highly rearranged state of both the *Hormiphora* and *Capsaspora* genomes, a Bonferroni-corrected

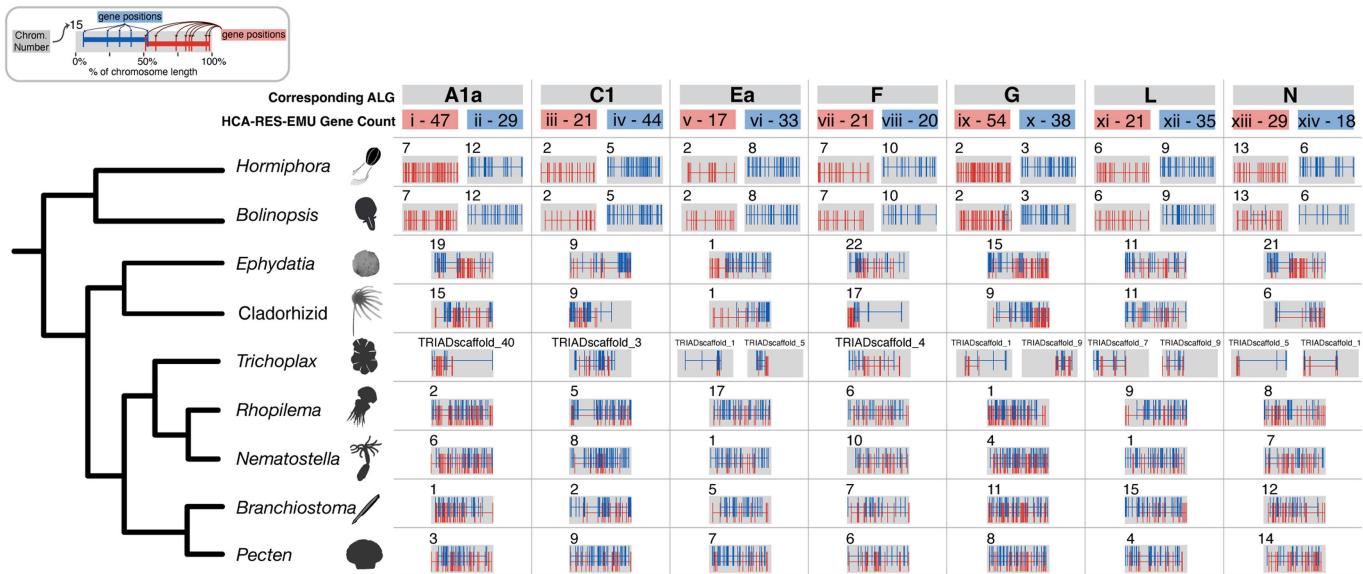
one-sided Fisher's exact test<sup>57</sup> only distinguishes three chromosome relationships as significant ( $p \leq 0.05$ ). **e.** Calculating the orthology conservation score for the relationships in these two genomes reveals more gene linkages that have been conserved across Filozoa. Red dots here are orthologs that are in significantly-conserved ortholog networks ( $\alpha \leq 0.05$ , permutation test). See complete results in Supplementary Information 11.



**Extended Data Fig. 7 | Filozoa and choanoflagellate genomes share macrosynteny with metazoans.** Two-way reciprocal best hits blast searches between the filasterean amoeba *Capsaspora* and animals (a-d.), or between the choanoflagellate *Salpingoeca* and animals (d-g.) show that the chromosomes of these unicellular species are rearranged relative to animal chromosomes, that some regions of synteny remain, and that some ALGs are split across

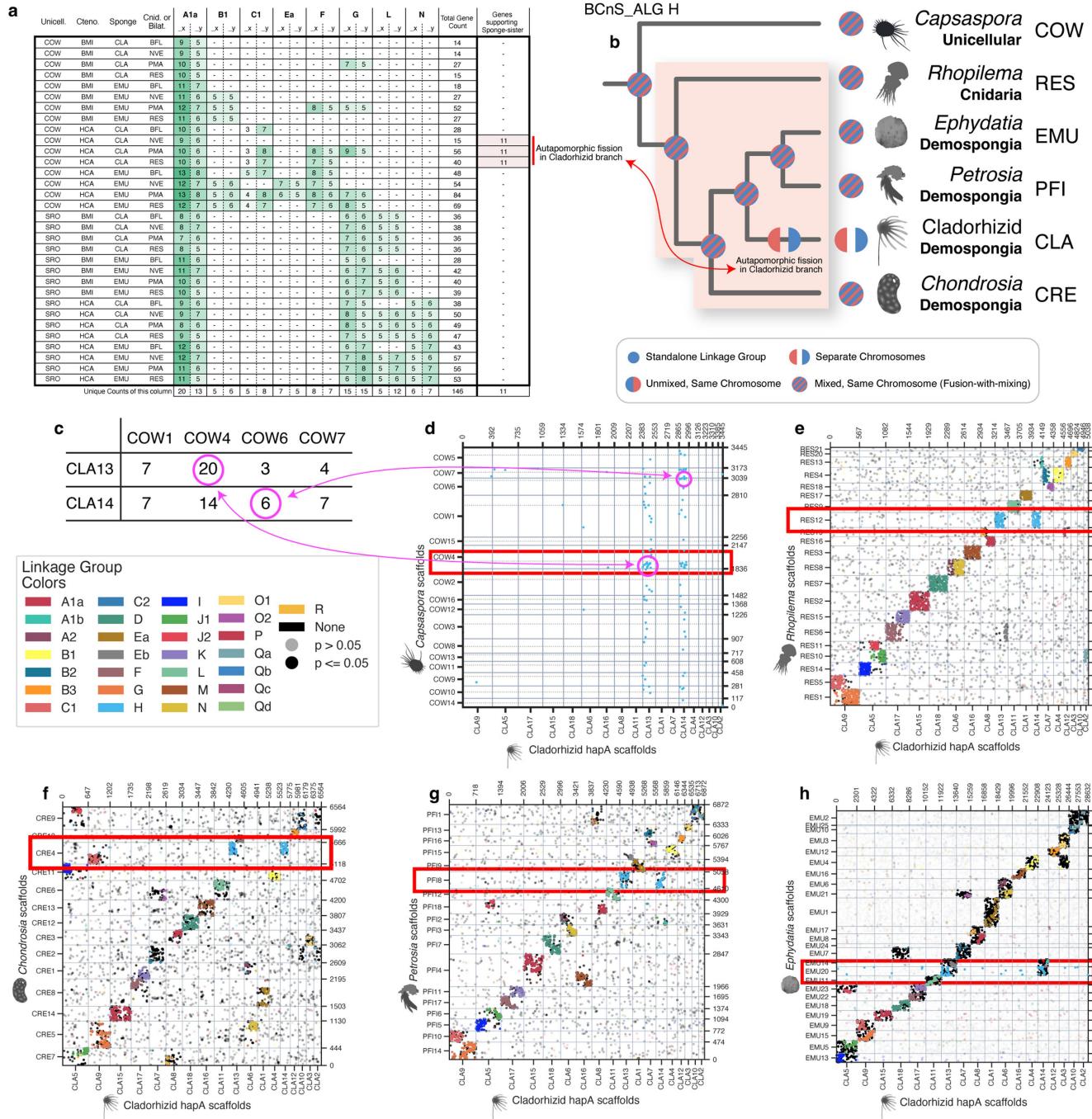
multiple chromosomes of the unicellular species. Orthologs are coloured based on BCnS-ALGs from Simakov et al. 2022<sup>12</sup>, and chromosome pairs with significantly-conserved macrosynteny ( $p \leq 0.05$ , Bonferroni-corrected one-sided Fisher's exact test<sup>57</sup>) have opaque dots. Axis labels show cumulative number of orthologs. Putative centromeres are marked by dotted lines.

# Article



**Extended Data Fig. 8 | Mixing plots of HCA-EMU-RES reciprocal best blastp results.** This figure parallels Fig. 3 of the main text, but includes more genes by requiring orthology between metazoans without requiring orthologs in corresponding outgroups. Limiting the macrosynteny search to animals shows many genes participating in the extension of metazoan ALGs to the ctenophores.

The *\_x* and *\_y* components of ALG\_Ea and ALG\_G are mixed and widely distributed across single sponge chromosomes, while the (COW/SRO)-HCA-EMU-RES results show no *\_x* and *\_y* overlap for ALG\_Ea, and little overlap for ALG\_G. We placed placozoans as the sister clade to cnidarians based on the findings of Simakov et al. 2022<sup>12</sup>. See also Supplementary Information 13.2.2.

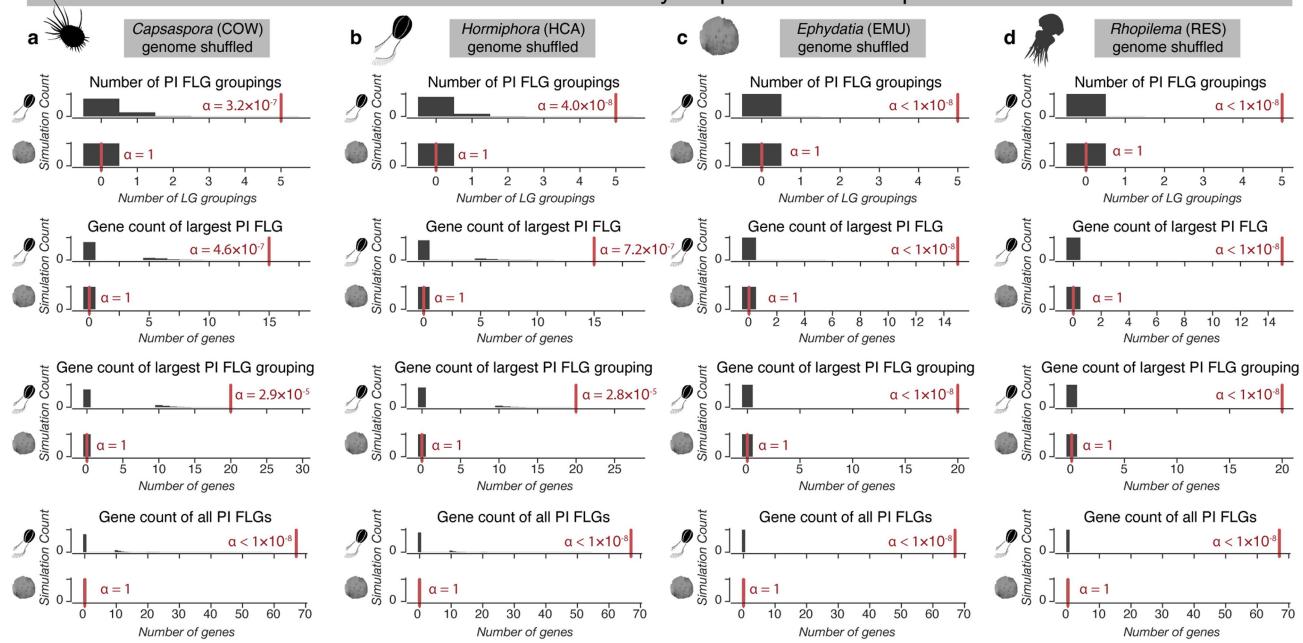


**Extended Data Fig. 9 | OrthoFinder results are consistent with the ctenophore-sister hypothesis.** **a.** Each green cell shows how many orthogroups support the ctenophore-sister hypothesis from each ALG in each species quartet. The Total Gene Count column is the total number of orthogroups supporting the ctenophore-sister hypothesis for that species quartet. The bottom row shows the number of unique orthogroups in each column. There are 146 orthogroups that support ctenophore-sister. **b.** The 11 orthologs that support CLA-sister in three analyses are due to a lineage-specific fission of ALG\_H that is only found in the cladorhizid genome, but not in the genome of

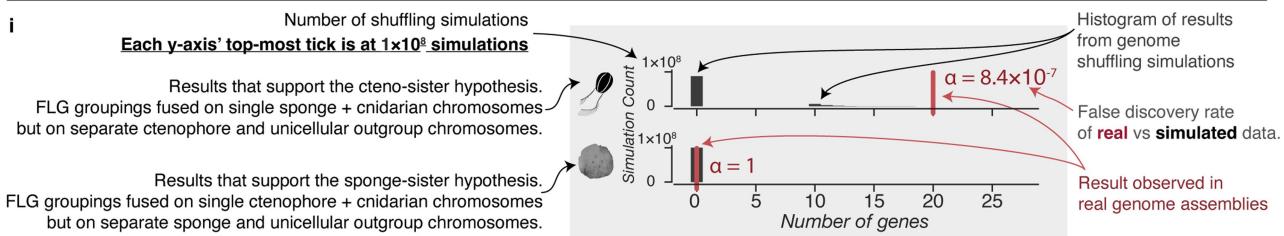
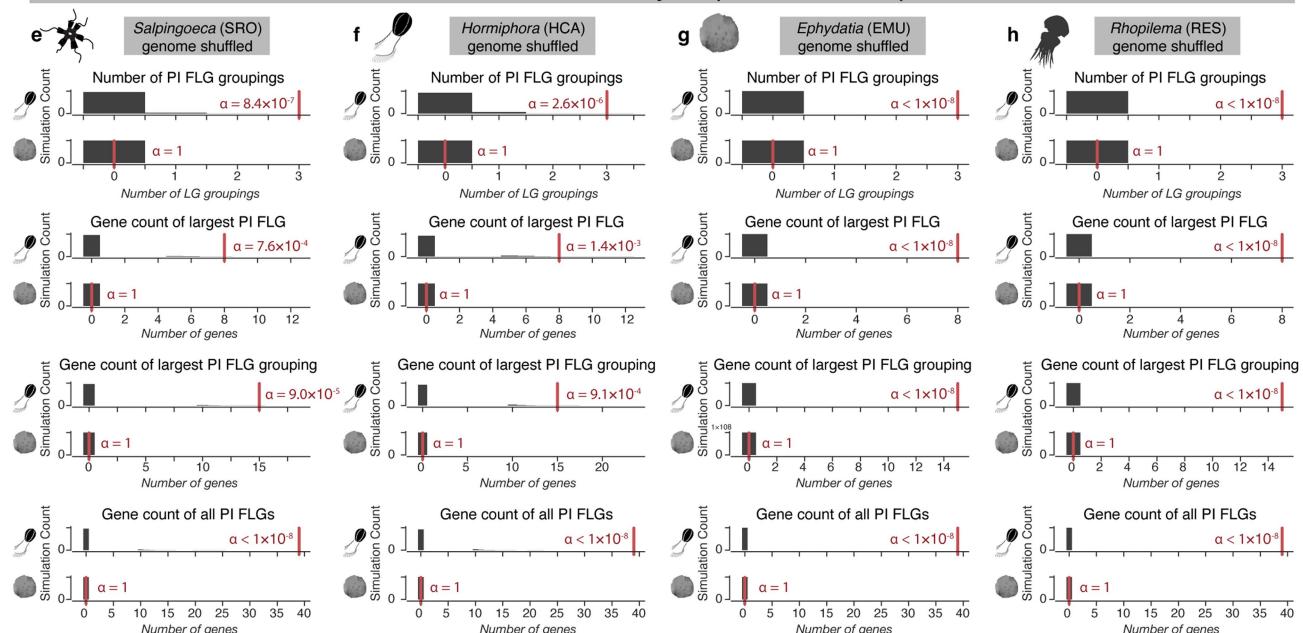
other sponges. Tree topology based on previous studies<sup>38,109</sup>. **c.** The *Capsaspora*-cladorhizid chromosome pairs with the most genes from ALG\_H (COW4-CLA13, COW4-CLA14) are not the chromosome pairs supporting sponge-sister (magenta circles, COW4-CLA13, COW4-CLA14). **d-h.** The fission of ALG\_H is specific to the cladorhizid sponge genome and is not found in the unicellular organism *Capsaspora* (COW), in other demosponges (EMU, CRE, PFI), in cnidarians (RES), or in bilaterians (not shown). Chromosome pairs that have significantly-conserved macrosynteny ( $p \leq 0.05$ , Bonferroni-corrected one-sided Fisher's exact test<sup>57</sup>) have opaque dots.

# Article

## COW-HCA-EMU-RES 4-way reciprocal best blastp hits



## SRO-HCA-EMU-RES 4-way reciprocal best blastp hits



### Extended Data Fig. 10 | Results of genome shuffling simulations.

**a-d**. Shuffling one of the genomes before the COW-HCA-EMU-RES comparison shows that the rearranged state of the ctenophore genome, let alone the other species in the analysis, cannot explain the signal supporting the ctenophore-

sister hypothesis (vertical red lines). **e-h**. Shuffling simulations using SRO as the outgroup independently support the ctenophore-sister hypothesis.

**i**. contains a legend to interpret panels a-h.

**Extended Data Table 1 | Linkage groups conserved in animals and unicellular outgroups**

		Group Number	ALG	No. of Shared Orthologs	Unicellular Outgroups			Ctenophora		Porifera		Placozoa		Cnidaria		Bilateria		
					Creolimax	Capsaspora	Salpingoeca	Hormiphora	Bolinopsis	Ephydalia	cladorthizid	Trichoplax	Nematostella	Rhopilema	Branchiostoma	Pecten		
From EMU-RES-BFL or HCA-RES-EMU searches	BCnS only	-	A1b	52	N/A	N/A	N/A	N/A	N/A	EMU14R	CLA14	complex	NVE2	RES13	BFL1	PMA3		
		-	Qabcd	-	N/A	N/A	N/A	N/A	N/A	complex	complex	complex	complex	complex	complex	complex		
	Metazoan Only	-	A2	10	N/A	N/A	N/A	HCA10	BMI10	EMU6	CLA6R	TRIADsc scaffold_5	NVE7	RES8	BFL1	PMA13		
		-	O1	21	N/A	N/A	N/A	HCA4	BMI4	EMU3	CLA3	complex	NVE5	RES20	BFL4L	PMA10		
		-	R	20	N/A	N/A	N/A	HCA8	BMI8	EMU17	CLA8R	TRIADsc scaffold_1	NVE3	RES19	-	PMA10		
		-	Eb	12	N/A	N/A	N/A	HCA5	BMI5	EMU14L	CLA13	TRIADsc scaffold_4	NVE10	RES6	BFL5	PMA7		
		-	O2	19	N/A	N/A	N/A			EMU21mid	CLA7R	complex	NVE15	RES18	BFL19	PMA5		
		-	B2	14	N/A	N/A	N/A	HCA1	BMI1	EMU7	CLA7L	TRIADsc scaffold_15	NVE2	RES4	BFL16	PMA2		
		-	B3	14	N/A	N/A	N/A			EMU12	CLA12	TRIADsc scaffold_3		RES13	BFL18	PMA19		
From (CFR, COW, SRO)-HCA-EMU-RES searches	Single-Chromosome in Filozoa	1	P	6	N/A	N/A	SRO6	HCA4	BMI4	EMU8	CLA8	TRIADsc scaffold_6	NVE3	RES16	BFL14	PMA16		
		2	C2	8	N/A	COW5		HCA2	BMI2	EMU10	CLA10	TRIADsc scaffold_5		RES21	BFL3	PMA17		
		3	D	15	N/A	COW12	SRO2	HCA1	BMI1	EMU18	CLA18	TRIADsc scaffold_5	NVE11	RES7	BFL6	PMA11		
		4	J1	5	N/A	COW1		HCA11	BMI11	EMU5	CLA5	TRIADsc scaffold_7	NVE12	RES10	BFL17	PMA18		
		5	J2	7	N/A	COW11				EMU23	CLA5	TRIADsc scaffold_20		RES11	BFL2	PMA4		
		6	M	10	N/A	COW4	SRO16	HCA9	BMI9	EMU16	CLA16	TRIADsc scaffold_2	NVE9	RES3	BFL8	PMA2		
	Holozoan, Translocation-participating, not phylogenetically informative	7	I_z1	11	CFR1 CFR3 COW4 COW9 COW13	N/A	HCA1	BMI1	EMU13	CLA5	TRIADsc scaffold_8	TRIADsc scaffold_18	TRIADsc scaffold_23	TRIADsc scaffold_27	NVE5	RES14	BFL4	PMA12
		8	I_z2	7														
		9	I_z3	6														
		10	K_z1	7	N/A CFR4 COW4 COW6 SRO9	COW4 COW9 COW13	HCA3	BMI3	EMU23	CLA17	TRIADsc scaffold_4	TRIADsc scaffold_17	NVE14	RES15	BFL9	PMA5		
		11	K_z2	14														
		12	B1_z1	6	N/A COW1 SRO6	COW1 SRO6	HCA13	BMI13	EMU4	CLA4	TRIADsc scaffold_3	N/A	N/A	NVE2	RES4	BFL10	PMA15	
		13	B1_z2	7														
		14	C1_z1	6	N/A COW3 (SRO1)	N/A COW3	HCA2	BMI2	EMU9	CLA9	TRIADsc scaffold_2	N/A	N/A	NVE8	RES5	BFL2	PMA9	
		15	C1_z2	8														
		16	H_z1	14	CFR1 CFR3 COW4 COW6	COW4 SRO10	HCA6	BMI6	EMU20	CLA13 CLA14	TRIADsc scaffold_2	N/A	N/A	NVE13	RES12	BFL13	PMA1	
		17	H_z2	10														
Ctenophore and unicellular outgroups with fusion on the BCnS stem	with fusion on the BCnS stem	18	A1a_x	29	CFR1 COW3	SRO2 SRO5	HCA7	BMI7	EMU19	CLA15	TRIADsc scaffold_40	N/A	N/A	NVE6	RES2	BFL1	PMA3	
		19	A1a_y	11			HCA12	BMI12										
		20	Ea_x	8	N/A COW3	N/A	HCA2	BMI2	EMU1	CLA1	N/A	N/A	NVE1	RES17	BFL5	PMA7		
		21	Ea_y	5														
	share ancestral state with fusion and mixing on the BCnS stem	22	G_x	27	CFR1 CFR2	COW10	SRO4	HCA2	BMI2	EMU15	CLA9	TRIADsc scaffold_9	N/A	N/A	NVE4	RES1	BFL11	PMA8
		23	G_y	10														
		24	C1_x	5	N/A COW14	N/A SRO1	HCA5	BMI5	EMU9	CLA9	TRIADsc scaffold_3	N/A	N/A	NVE8	RES5	BFL2	PMA9	
		25	C1_y	8														
		26	F_x	5	N/A COW8	N/A SRO1	HCA7	BMI7	EMU22	CLA17	N/A	N/A	NVE10	RES6	BFL7	PMA6		
		27	F_y	7														
		28	L_x	5	N/A COW1	N/A SRO9	HCA6	BMI6	EMU11	CLA11	TRIADsc scaffold_7	N/A	N/A	NVE1	RES9	BFL15	PMA4	
		29	L_y	13														
		30	N_x	9	N/A COW2	N/A SRO2	HCA13	BMI13	EMU21	CLA6	TRIADsc scaffold_4	N/A	N/A	NVE7	RES8	BFL12	PMA14	
		31	N_y	14														

The gene linkage groups only found in BCnS or Metazoans (group number "-"), and the merged OG-metazoan four-way reciprocal best blastp results (Group Number 1 through 31). Fusions-with-mixing events in the ancestor of the Choanozoa, or the ancestor of the Metazoa, or the ancestor of the BCnS clade are represented by rows of different colours, joined by striped cells. There is evidence for four fusion-with-mixing events uniting sponges, cnidarians, placozoans, and bilaterians, to the exclusion of ctenophores and unicellular OGs.

# Article

**Extended Data Table 2 | Ancestral linkage groups found by four-way outgroup-animal comparisons**

a		<i>H. californensis</i> Ctenophora		<i>R. esculentum</i> Cnidaria				b		<i>H. californensis</i> Ctenophora		<i>R. esculentum</i> Cnidaria					
ALG	No. of Shared Orthologs	<i>C. owczarzaki</i> Unicell	HCA	<i>E. muelleri</i> Porifera	EMU	RES	False Discovery Rate	ALG	No. of Shared Orthologs	<i>S. rosetta</i> Unicell	HCA	<i>E. muelleri</i> Porifera	EMU	RES	False Discovery Rate		
A1a_x	15	COW3	HCA7	EMU19	RES2	$\alpha < 1.0e-07$		A1a_x	6	SRO2	HCA7	EMU19	RES2	$\alpha = 2.0e-07$			
A1a_y	5	COW5	HCA12	EMU19	RES2	$\alpha = 3.2e-04$		A1a_x	11	SRO5	HCA7	EMU19	RES2	$\alpha < 2.0e-07$			
C1_x	5	COW1	HCA2	EMU9	RES5	$\alpha = 3.2e-04$		A1a_y	7	SRO3	HCA12	EMU19	RES2	$\alpha < 2.0e-07$			
C1_y	5	COW14	HCA5	EMU9	RES5	$\alpha = 3.2e-04$		G_x	8	SRO4	HCA2	EMU15	RES1	$\alpha < 2.0e-07$			
Ea_x	8	COW3	HCA2	EMU1	RES17	$\alpha < 1.0e-07$		G_y	7	SRO17	HCA3	EMU15	RES1	$\alpha < 2.0e-07$			
Ea_y	5	COW4	HCA8	EMU1	RES17	$\alpha = 3.2e-04$		L_x	5	SRO9	HCA6	EMU11	RES9	$\alpha = 2.3e-05$			
F_x	5	COW1	HCA7	EMU22	RES6	$\alpha = 3.2e-04$		L_y	6	SRO27	HCA9	EMU11	RES9	$\alpha = 2.0e-07$			
F_y	5	COW8	HCA10	EMU22	RES6	$\alpha = 3.2e-04$		N_x	7	SRO19	HCA6	EMU21	RES8	$\alpha < 2.0e-07$			
G_x	8	COW10	HCA2	EMU15	RES1	$\alpha < 1.0e-07$		N_y	6	SRO2	HCA13	EMU21	RES8	$\alpha = 2.0e-07$			
G_y	6	COW12	HCA3	EMU15	RES1	$\alpha = 7.1e-06$		I_z2	6	SRO3	HCA1	EMU13	RES14	$\alpha = 2.0e-07$			
N_x	7	COW2	HCA6	EMU21	RES8	$\alpha = 1.0e-07$		I_z3	7	SRO25	HCA1	EMU13	RES14	$\alpha < 2.0e-07$			
N_y	7	COW2	HCA13	EMU21	RES8	$\alpha = 1.0e-07$		B1_z1,2	5	SRO6	HCA13	EMU4	RES4	$\alpha = 2.3e-05$			
C1_z1	6	COW3	HCA2	EMU9	RES5	$\alpha = 7.1e-06$		C1	6	SRO1	HCA5	EMU9	RES5	$\alpha = 2.0e-07$			
C1_z2	8	COW3	HCA5	EMU9	RES5	$\alpha < 1.0e-07$		C2	7	SRO2	HCA2	EMU10	RES21	$\alpha < 2.0e-07$			
B1_z1	7	COW3	HCA13	EMU4	RES4	$\alpha = 1.0e-07$		D	10	SRO2	HCA1	EMU18	RES7	$\alpha < 2.0e-07$			
B1_z2	6	COW1	HCA13	EMU4	RES4	$\alpha = 7.1e-06$		F	5	SRO1	HCA10	EMU22	RES6	$\alpha = 2.3e-05$			
H_z1	10	COW4	HCA6	EMU20	RES12	$\alpha < 1.0e-07$		H	9	SRO10	HCA6	EMU20	RES12	$\alpha < 2.0e-07$			
H_z2	5	COW6	HCA6	EMU20	RES12	$\alpha = 3.2e-04$		K	7	SRO9	HCA3	EMU23	RES15	$\alpha < 2.0e-07$			
I_z1,2,3	5	COW13	HCA1	EMU13	RES14	$\alpha = 3.2e-04$		M	6	SRO16	HCA9	EMU16	RES3	$\alpha = 2.0e-07$			
I_z1,2,3	6	COW4	HCA1	EMU13	RES14	$\alpha = 7.1e-06$		P	6	SRO6	HCA4	EMU8	RES16	$\alpha = 2.0e-07$			
I_z1,2,3	6	COW9	HCA1	EMU13	RES14	$\alpha = 7.1e-06$											
K_z1	7	COW4	HCA3	EMU23	RES15	$\alpha = 1.0e-07$											
K_z2	8	COW6	HCA3	EMU23	RES15	$\alpha < 1.0e-07$											
C2	5	COW5	HCA2	EMU10	RES21	$\alpha = 3.2e-04$											
D	11	COW12	HCA1	EMU18	RES7	$\alpha < 1.0e-07$											
J1	5	COW1	HCA11	EMU5	RES10	$\alpha = 3.2e-04$											
J2	7	COW11	HCA11	EMU23	RES11	$\alpha = 1.0e-07$											
L_y	8	COW1	HCA9	EMU11	RES9	$\alpha < 1.0e-07$											
M	6	COW4	HCA9	EMU16	RES3	$\alpha = 7.1e-06$											
c		<i>C. fragran.</i> Unicell		<i>H. californensis</i> Ctenophora		<i>R. esculentum</i> Cnidaria											
ALG	No. of Shared Orthologs	<i>C. fragran.</i> CFR	Unicell	HCA	EMU	<i>E. muelleri</i> Porifera	RES	False Discovery Rate	ALG	No. of Shared Orthologs	<i>C. fragran.</i> CFR	HCA	<i>E. muelleri</i> Porifera	EMU	RES	False Discovery Rate	
G_x	5	CFR1		HCA2	EMU15	RES1	$\alpha = 8.0e-05$										
G_x	8	CFR2		HCA2	EMU15	RES1	$\alpha < 4.0e-08$										
H_z1,2	5	CFR1		HCA6	EMU20	RES12	$\alpha = 8.0e-05$										
H_z1	5	CFR3		HCA6	EMU20	RES12	$\alpha = 8.0e-05$										
I_z1,2,3	5	CFR1		HCA1	EMU13	RES14	$\alpha = 8.0e-05$										
I_z1,2,3	7	CFR3		HCA1	EMU13	RES14	$\alpha = 4.0e-08$										
K_z2	5	CFR4		HCA3	EMU23	RES15	$\alpha = 8.0e-05$										
A1a_x	7	CFR1		HCA7	EMU19	RES2	$\alpha = 4.0e-08$										

Each row represents a conserved syntenic group, with false discovery rate (a, b, n=1×10<sup>7</sup>, c, n=1×10<sup>8</sup>) estimated by comparison with genome shuffling simulations (Suppl. Information 8). We use three-letter code for species in each column. **a.** The COW-HCA-EMU-RES search yielded 10 gene groups where *Hormiphora* shares an ancestral partitioned state with *Capsaspora*, but are fused onto 5 chromosomes in the sponge and cnidarian. Part of ALG C1 appears to have a derived split in ctenophores. Linkage groups corresponding to ALGs B1, H, I, and K appear to have each become established on single chromosomes by time of the common ancestor of metazoans. In this comparison, there is no evidence that ALGs C2, D, J1, J2, L, and M do not participate in any clade-specific fusions that are informative to cteno- vs sponge-sister. **b.** The SRO-HCA-EMU-RES search yielded nine gene groups where *Hormiphora* shares an ancestral partitioned state with *Salpingoeca*, but are fused onto four chromosomes in the sponge and cnidarian. There is no evidence for ctenophore-derived splits in this search. Linkage groups corresponding to ALG I appear to have merged onto single chromosomes by the common ancestor of metazoans. This table suggests that ALGs B1, C1, C2, D, F, H, K, M and P do not participate in any clade-specific fusions that are informative to cteno- vs sponge-sister. **c.** A comparison of *Creolimax fragrantissima* to HCA-EMU-RES shows limited conservation of gene linkages between animals and *Creolimax*, an ichthyosporean.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection The software developed for this manuscript is available at <https://github.com/conchoecia/odp> and at <https://doi.org/10.5061/dryad.dncjsxm47>.

Data analysis The software packages used to analyze data in this manuscript were: odp v0.2.0 and v0.3.0, HiRise vAug2019, SALSA2 v2.3, wtdbg v2.4, hifiasm v0.16.1-r375, TGS-Gapcloser v1.1.1, Purge Haplotype v1.0.4, pilon v1.23, Diamond v0.9.24, Blobtools v1.0, minimap2 v2.17 and v2.23, tblastn v2.10.0+, miniprot v0.2, BRAKER v2.14, STAR v2.7.1a, ProtHint v2.6.0, Trinity v2.5.1, OrthoFinder v2.3.7, hmmer v3.3.2, PANTHER v17, snakemake v7, BUSCO v5, lima v2.2.0, isoseq3 v3.4.0, jellyfish v2.2.10, GenomeScope 2, bwa mem v0.7.17, PretextView v0.2.4, HiGlass v1.10.0104, Juicebox Assembly Tools github commit 46c7ed1105, Juicebox visualization system v1.11.08106, TransDecoder v5.5, D-genies v1.4.0, pairtools v0.3.0 pairix v0.3.7, Cooler v0.8.10, MAFFT v7.310, RevBayes version 3.2.7a, and FigTree v1.4.4.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The authors confirm that all data presented in this manuscript are available in public repositories. The sequencing reads are available in the NCBI database under BioProject accession numbers PRJNA818620, PRJNA818630, PRJNA903214, and PRJNA818537. The genomes for each species are available through the above BioProjects, with the exception that the genomes of *C. fragrantissima*, *C. owczarzaki*, and *S. rosetta*, which are available on Dryad: <https://doi.org/10.5061/dryad.dncjsxm47>. The scripts and results of the supplementary information, when not contained in figures, are also available in the aforementioned Dryad repository. Publicly available sequencing data and genomes were downloaded from NCBI from BioProject accession numbers PRJNA168, PRJDB8655, PRJNA12874, PRJNA20249, PRJNA20341, PRJEB28334, PRJNA30931, PRJNA31257, PRJNA37927, PRJEB56075, PRJEB56892, PRJNA64405, PRJNA193541, PRJNA193613, PRJNA213480, PRJNA278284, PRJNA281977, PRJNA377365, PRJNA396415, PRJNA512552, PRJNA544471, PRJNA576068, PRJNA579531, PRJNA625562, PRJNA667495, PRJNA761294, and PRJNA814716. The *Ephydatia muelleri* genome was downloaded from <https://spaces.facsci.ualberta.ca/ephybase/>. All of the above information is also included in the main text of the manuscript.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

N/A

Population characteristics

N/A

Recruitment

N/A

Ethics oversight

N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

These categories are not applicable to this study, which involves analysis of genome sequences and chromosome architecture. Sample sizes of permutation tests and the resulting statistics are included in the manuscript.

Data exclusions

DNA sequences were filtered (for quality) during the assembly process as described in the methods.

Replication

Our findings were replicated by (1) performing many analyses using combinations of different species, and (2) performing the same range of analyses using multiple orthology inference techniques. Otherwise, because this was a comparative genomics study using replicates of the same was not relevant.

Randomization

Randomization is not relevant to this study, as the genomes used in comparisons are selected at the point of experimental design. It is not possible to randomize these selections as we are dependent on the phylogenetic relationships between the species in the comparisons to draw conclusions.

Blinding

As above, the conclusions of this comparative genomics study were dependent on selecting groups of species with particular phylogenetic relationships. Therefore, blinding was not relevant to this study.

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

## Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

### Laboratory animals

No laboratory animals were used in this study.

### Wild animals

Samples of adult *Bolinopsis microptera* were collected on May 24th, 2015 in the Monterey Bay, California (36.63°N, 121.90°W) with jars from the surface waters, with permission under the State of California Department of Fish and Wildlife collecting permit SC-2026 to the Monterey Bay Aquarium. These samples were transported to the Monterey Bay Aquarium for rearing.

One individual of an undescribed hexactinellid sponge was collected on June 1st, 2021, in the Monterey Bay, California (34.57°N, 122.56°W) from the seafloor at 3,852 meters depth using the MBARI ROV Doc Ricketts aboard the R/V Western Flyer. On the following day one individual of an undescribed bioluminescent cladorhizid sponge was collected from a nearby site (35.49°N, 124°W) from the seafloor at 3,975 meters depth. The collection temperature of both samples was 1.5°C. The cladorhizid sample was consistent in morphology and locale with previously reported bioluminescent, carnivorous, cladorhizid sponges (1). Upon retrieval from the ROV, the samples were washed gently with 1°C filtered seawater to remove debris and maintained in the dark at 1°C for no more than 30 minutes after being collected from the ROV. Then, both samples were flash-frozen in liquid nitrogen. The samples were collected with the State of California Department of Fish and Wildlife collecting permit SC-4029 granted to the Haddock Laboratory at the Monterey Bay Aquarium Research Institute.

(1) Martini, S., Schultz, D. T., Lundsten, L. & Haddock, S. H. D. Bioluminescence in an Undescribed Species of Carnivorous Sponge (Cladorhizidae) From the Deep Sea. *Front. Mar. Sci.* 7, 576476 (2020).

### Reporting on sex

N/A. The organisms sequenced in this study are hermaphrodites, or the biology of the species' sex is unknown.

### Field-collected samples

A community culture was founded with 20 *Bolinopsis microptera* individuals in pseudokreisel tanks and diffusion tubes in 12°C seawater at the Monterey Bay Aquarium in Monterey, California. The culture was reared according to the published protocol (1) for three generations, and an F3 adult, called Bmic1, was selected and flash-frozen in liquid nitrogen on November 18th, 2019 for DNA sequencing for genome assembly and annotation. Four other F3 adults were placed into a spawning tank and spawned according to the published protocol(1). Fertilized eggs were collected 18 hours post-spawning for RNA sequencing.

(1) Patry, W. L., Bubel, M., Hansen, C. & Knowles, T. Diffusion tubes: a method for the mass culture of ctenophores and other pelagic marine invertebrates. *PeerJ* 8, e8938 (2020).

### Ethics oversight

No vertebrates or cephalopods were used in the study, and the organisms included in this study are unrestricted other than the collecting permits listed above.

Note that full information on the approval of the study protocol must also be provided in the manuscript.