PALEONTOLOGY

Paleozoic origins of cheilostome bryozoans and their parental care inferred by a new genome-skimmed phylogeny

Russell J. S. Orr1*,†, Emanuela Di Martino1, Mali H. Ramsfjell1, Dennis P. Gordon7, Björn Berning3, Ismael Chowdhury4, Sean Craig4, Robyn L. Cumming5, Blanca Figuerola6, Wayne Florence7, Jean-Georges Harmelin8, Masato Hirose9, Danwei Huang10, Sudhanshi S. Jain10, Helen L. Jenkins11,12, Olga N. Kotenko13, Piotr Kuklinski14, Hannah E. Lee15, Teresa Madurell16, Linda McCann17, Hannah L. Mello16, Matthias Obst17, Andrew N. Ostrovsky13,18, Gustav Paulay19, Joanne S. Porter20, Andrea Waeschenbach12, Kamil Zágoršek22, Rachel C. M. Warnock23, Lee Hsiang Liow1,24*†

Phylogenetic relationships and the timing of evolutionary events are essential for understanding evolution on longer time scales. Cheilostome bryozoans are a group of ubiquitous, species-rich, marine colonial organisms with an excellent fossil record but lack phylogenetic relationships inferred from molecular data. We present genome-skimmed data for 395 cheilostomes and combine these with 315 published sequences to infer relationships and the timing of key events among c. 500 cheilostome species. We find that named cheilostome genera and species are phylogenetically coherent, rendering fossil or contemporary specimens readily delimited using only skeletal morphology. Our phylogeny shows that parental care in the form of brooding evolved several times independently but was never lost in cheilostomes. Our fossil calibration, robust to varied assumptions, indicates that the cheilostome lineage and parental care therein could have Paleozoic origins, much older than the first known fossil record of cheilostomes in the Late Jurassic.

INTRODUCTION

Quantifying macroevolutionary processes, for example, diversification rates, and testing macroevolutionary hypotheses, such as whether speciation rate shifts are driven by environmental changes and/or trait evolution, require robust reconstructions of the genealogical relationships of the members of the clade in question. It is increasingly clear that it is preferable to reconstruct evolutionary histories based on both extant and extinct organisms (1, 2) and to combine morphological and molecular data (3, 4). Yet, empirical datasets that combine molecular, morphological, and substantial amounts of fossil data are still rare. This is, in part, because molecular sequencing efforts have been disproportionately focused on organisms with relatively poor fossil records (e.g., birds, some insects, and plant groups), while those with good fossil records are somewhat neglected (e.g., foraminifers, ostracods, and bryozoans). In this contribution on cheilostome bryozoans, we present one of the largest species-level molecular phylogenies for any order of marine invertebrates to alleviate the lack of molecular phylogenetic hypotheses for fossil-rich groups and to answer long-standing evolutionary questions on timing and rates.

Members of the colonial phylum Bryozoa have had important roles as marine ecosystem constructors and ecological interactors since their origins (5–7). They are long known to have an evolutionary history visible in the fossil record since the Early Ordovician (8) that has very recently been extended to the Cambrian (9). The constituent clades of Bryozoa have waxed and waned over geological time, with three classes, Phylactolaemata, Stenolaemata, and Gymnolaemata, still extant today (Fig. 1). The latter two classes are largely marine and calcified and hence have rich fossil records. The order Cheilostomata within the Gymnolaemata have especially intricate skeletal morphologies that allow species-level delimitation, as shown using breeding experiments and allelic analyses (10). This suggests that cheilostome fossils are amenable to species-level identifications, an advantage for integrating data from molecular sequences with fossil remains to reconstruct macroevolutionary history and processes. Cheilostomes are also the most species-rich order within Bryozoa, with more than 6000 described extant species and likely about the same number yet to be formally described (11). They represent c. 80% of the phylum’s living species diversity (12). Likewise, there are c. 7900 described...
We use bryozoan (phylactolaemate, ctenostome, and cyclostome) and bilaterian out-groups and 18 fossil calibration points and present a time-calibrated bryozoan tree, asking how much of the early lineage leading to extant cheilostomes is now “invisible” (i.e., not detected) in the fossil record and when Bryozoa might have originated. Then, we ask when (and how often) parental care in the form of incubation (brooding) evolved in the history of cheilostome evolution. The transition from a nonbrooded embryo resulting in a long-lived, planktotrophic larva to a brooded embryo resulting in a short-lived, nonfeeding larva is hypothesized to have driven rapid speciation among cheilostomes displaying the brooding trait (16). It is thought that species with nonfeeding larvae disperse much shorter distances than those with feeding larvae and are hence associated with lower amounts of gene flow and consequently a higher speciation probability. Using our new time tree presented here, we ask whether there is evidence that species with nonfeeding larvae (note that all cheilostomes that brood have only nonfeeding larvae) are associated with higher speciation rates across the cheilostome clade.

While highlighting the continued need for an increased effort in systematics based on morphology and sequence data and broader taxon-sampling for phylogenetic inference, we underscore that this work is a considerable step toward establishing cheilostomes as a model macroevolutionary system.
RESULTS

The largest cheilostome molecular phylogeny to date

The four extant bryozoan groups (phylactolaemates, cyclostomes, ctenostomes, and cheilostomes) each form well-supported clades (Fig. 1; see Discussion on ctenostomes). Our full (fig. S1, based on all the cheilostome sequences included in this study) and trimmed molecular phylogenies (fig. S2, a subset of the full phylogeny trimmed with criteria listed in the methods) are illustrated in an abbreviated form in Fig. 1, highlighting the main inferred, and extant, clades of cheilostomes. Our cheilostome phylogeny has a well-supported backbone with seven highly supported [bootstrap support (BS) >90%; Fig. 1] ancestral nodes that gave rise to the depicted, and again highly supported, extant cheilostome clades (Fig. 1, A to G, and fig. S2), albeit with the exception of the ancestral node (BS of 64%) that resulted in the fully supported Conopeum genus (Fig. 1, clade B, and fig. S2). The overall mean BS support is also high, averaging at 88.94% per node (calculated based on fig. S2 with 721 taxa). The seven branches that led to the extant monophyletic A to G clades (Fig. 1) do not match the now available broad systematic framework of cheilostomes (17). Our inferred topology (see fig. S1A) substantially filled out regions of the cheilostome tree (i.e., Scruparia to Macropora) where key evolutionary transitions, including parental care (and hence nonfeeding larvae), are thought to have taken place (16, 18, 19). The metadata associated with our new sequences (N = 516, where 395 cheilostomes and five cyclostomes are associated with physical vouchers with museum accession numbers, and 116 cheilostomes and one ctenostome without; see section S5) are presented in table S1; genes used for phylogenetic inferences are tabulated in table S2 (mean = 14 of 17 genes for 854 taxa); and National Center for Biotechnology Information (NCBI) accession numbers for deposited sequences are presented in table S3.

Our inferred tree topology, based solely on molecular sequences, largely supports the phylogenetic coherence of morphological species and genus concepts used by bryozoan taxonomists (figs. S1 and S2). For instance, three specimens of Klugiastra vanhoffeni collected during two different expeditions in distinct locations and identified by three independent experts (table S1) are found to be monophyletic, with little to no genetic distance based on 13 to 15 genes (table S2 and fig. S1A). In a genus-level example, Stegnoporella, represented in the next most recent cheilostome molecular phylogeny by five species (20), is now represented by nine species: Stegnoporella was, and still is, monophyletic. The genus Microporella is here inferred by molecular sequences to include Diporula and Flastramorpha. These latter genera have been recently synonymized with Microporella based purely on morphological grounds (21, 22), likewise for the reassignment of Fenestrulina joannae to Microporella (22). In contrast, many cheilostome families are polyphyletic. For instance, genera of the family Smittiniidae (17) are scattered throughout clade G. Likewise, the genera of Bugulidae are scattered throughout clade C (Fig. 1 and figs. S1 and S2; see also section S4, Supplementary Text for Results and Discussion).

A fossil-calibrated bryozoan tree and deep origins of clades

Our phylogeny, fossil calibrated on 18 nodes with a relaxed independent rates molecular clock model (Fig. 2; see details of calibrations in table S4 and Supplementary Text and joint time priors in fig. S3) suggests that bryozoans originated (i.e., became distinct from other Lophotrochozoa) about 518 Ma ago (million years ago) (Fig. 2, node iii). This is the median value of the posterior distribution with 95% highest posterior density (HPD) between 495 and 547 Ma ago; i.e., bryozoans are inferred to have originated in the Cambrian or as early as the Ediacaran (fig. S4 for sensitivity analyses using different clock models and calibrations). Node iv in Fig. 2, where cyclostomes and ctenostomes plus cheilostomes diverged from their common ancestor shared with extant phylactolaemates, is estimated at 488 (HPD 471 to 517) Ma ago, i.e., Late Cambrian or Early Ordovician. Node v, the divergence of cheilostomes plus ctenostomes from other bryozoans, is estimated at 407 (HPD 353 to 457) Ma ago, i.e., Early Devonian. The cheilostome lineage is inferred to have diverged from ancestors shared with ctenostomes in the Carboniferous (345 Ma ago, HPD 292 to 398 Ma ago), c. 200 Ma earlier than the confirmed fossil record of cheilostomes in the Late Jurassic (Fig. 2, node vi; see fig. S4 for sensitivity analyses). Two of the seven deep splits within the cheilostome clade (Fig. 1, nodes A and B) are inferred to have happened in the Carboniferous, one in the Triassic (Fig. 1, node C), and four in the Jurassic (Fig. 1, nodes D to G), while many lineages leading to extant genera originated in the Cretaceous or Paleogene (see Fig. 2).

Evolution of parental care and speciation rates of brooders

Parental care in a form of embryonic incubation (brooding for short hereon) of nonfeeding larvae, internally (inside zooidal cavity or internal brood sacs) or with specially developed external structures (membranous brood sacs and skeletal brood chambers), has independently evolved c. five times according to an ancestral state reconstruction (23), given our cheilostome tree topology (Fig. 3 and fig. S5). The transitions to brooding (and hence nonfeeding larvae) are inferred to have occurred as early as the Permian (Fig. 3, transition 4). The brooding state is inferred never to have transitioned back to nonbrooding, while the nonbrooding state transitions to a brooding state at a rate of 0.1888 per 100 Ma (SE 0.0503).

On the basis of an information criterion-based comparison of binary state speciation and extinction (BiSSE) (24) and hidden state speciation and extinction (HiSSE) models (25) and their null versions, we rejected a model where brooding is associated with differential rates of speciation. Here, a null BiSSE model (character independent model with two states, “cid2”; see the “Ancestral state reconstruction and HiSSE” section) has the highest Akaike information criterion (AIC) model weight (0.581) of the five models that we compared (Table 1; see table S6 for all parameter estimates). We also compared the same models with two alternative topologies; one where Lunularia is removed and one where Conopeum is alternatively placed as sister to all other cheilostomes (see Methods and the Supplementary Materials for reasoning). In both latter cases, we rejected a model where brooding is directly associated with differential rates of speciation but found strong support for a model where unmeasured states associated with the brooding state drove higher speciation rates (see table S7 for model weights).

DISCUSSION

Cheilostome bryozoans have exceptionally useful traits for tackling some long-standing questions in evolutionary biology. Such traits include a calcified skeleton that renders these marine organisms very fossilizable (7), external, calcified brooding structures that allow fecundity (a fitness component) to be quantified in the fossil record (15), a colonial and modular nature that allows the estimation of sources of phenotypic variation among and within individual genotypes and environments (26, 27), polymorphic structures that...
represent ergonomically partitioned divisions of labor (28, 29), and ecological interactions “frozen in time” (6, 14). Analyzing molecular sequence data, independent of morphological traits used to identify species, to infer evolutionary relationships, we lend strong support to important assumptions often invoked in the cheilostome literature with limited empirical support. The first is that skeletal traits can be used to identify cheilostome species (10), as separate specimens identified as the same species (based only on morphology) have little genetic distance in our inferred tree. The second is that cheilostome genera are natural groupings (monophyletic or paraphyletic clades).
Fig. 3. Lower section of cheilostome tree with parental care states. The topology shows the lower part of the cheilostome tree where brooders with nonfeeding larvae are marked in dark blue and nonbrooders with planktotrophic (feeding) larvae are marked in light blue. For the probability of transition of every node, including those not shown here, see fig. S5. Numbers show the transitions to a brooding state that are inferred, where transition 1 (as early as the Carboniferous; Fig. 2 and fig. S5) led to Scruparia (with a skeletal ovicell-like brood chamber), transition 3 (as early as the Jurassic) led to Eucratea (with external membranous brooding sacs), transition 4 (as early as the Triassic) led to the clade including Steginoporella (some with internal brooding sacs and others with skeletal brood chambers), and transition 5 (as early as the Triassic) led to “neocheilostomes” (cheilostomes with brooding structures called ovcicells or brooding sacs). See the Supplementary Materials for a discussion of transition 2.
and can arguably be used as unit of evolutionary analysis (30, 31). In addition, by generating a large volume of molecular sequences for cheilostome species, we are primed for an integration of such data with their morphological characters, moving one step closer to total evidence analyses (3). Such a phylogeny will allow us to answer other long-standing general evolutionary questions, including whether higher rates of morphological evolution happen close to speciation events (32).

Age and rates are two major features of evolution, and we contribute information with regards to both. The phylum Bryozoa has been considered enigmatic, not least because it is the only potentially fossilizable metazoan phylum with no body fossil representation in the Cambrian record (33), until very recently (9). Our main analysis and our sensitivity analyses with alternative calibration and clock assumptions (fig. S4) have inferred Bryozoa to have originated in the Cambrian, an idea first proposed by Hyman (34), or even as early as the late Ediacaran, despite the lack of fossil remains (see the Supplementary Materials for a discussion of Pywackia, a controversial Cambrian fossil). The lineage ancestral to living cheilostomes and ctenostomes has two peaks in its posterior age distribution (Fig. 2, node v), where the older peak overlaps the calibration and the younger peak does not. This may be an artifact of using boring ctenostomes as a calibration point, while our molecular data are represented by perhaps very distantly related, nonboring ctenostomes, i.e., there is a conflict between the fossil calibration and the molecular data for this node. This can be resolved by including boring ctenostomes in the phylogeny, although extracting sequences, given their life habit, is now challenging.

The cheilostome fossil record is long and rich (35), and we might have expected cheilostome origins to be on the order of only a few tens of million years earlier than the oldest cheilostome fossil from the Late Jurassic (36). However, given the tree topology, gene sampling and multiple fossil calibrations and sensitivity analyses (fig. S4), we have estimated the evolutionary origin of cheilostomes to be Paleozoic, somewhat earlier than the only other study based on sequence data to estimate cheilostome origins (37). Hao et al. (37) estimated cheilostomes to have originated in the Permian to the Early Triassic, based on only one nuclear gene (16S), 40 taxa, and one calibration point, which we did not include, as we did not have sequences pertaining to that node. While one might postulate that extinct bryozoan groups that are contemporary with this invisible stem lineage could be possible cheilostome progenitors, we now have no clear candidates that we can reasonably suggest from the fossil record (7). Both our tree topology and our understanding of their morphology points to the gymnolaemate order Ctenostomata as the most likely ancestor for crown group Cheilostomata. However, ctenostomes are known only from borings in the Paleozoic, and there are only a few fossils of ctenostomes, even in the more recent fossil record (38). This hints at largely uncalcified stem and ancestral crown cheilostome lineages (e.g., calcified skeletons in crown group cheilostomes may have multiple origins) and/or perhaps encrusting, calcified taxa favoring substrates that do not easily preserve. It is plausible that increasing taxon sampling and/or applying more complex models that allow for total evidence analyses could help us refine this and other age estimates in our bryozoan tree (39), but data for such analyses are not yet available. Four of the seven deep branches emerging from the backbone (Fig. 1) emerged throughout the Jurassic, at the end of which the fossil record of cheilostomes began with a trickle. Similarly, many lineages leading to extant genera are inferred to have originated in the Cretaceous (Fig. 2), when the fossil record of cheilostomes exploded in its morphological disparity and observed abundance. This suggests that, when cheilostomes lineages are observed in the fossil record, they are likely to have been extant for a substantial amount of time, perhaps at lower abundances or in cryptic habitats that enter the fossil record at a much lower rate. The continued exploration of the fossil record of bryozoans may yet reveal surprises, as even origination of groups as well studied as land plants continue to astonish (40), and the relationships among metazoan groups remain elusive (41).

A planktotrophic larva, associated with a nonbrooding state, is found to be the ancestral condition in cheilostomes, based on the topology of our inferred tree, as long hypothesized in the literature (19, 34). This is despite the living members of ctenostomes (putative ancestors of cheilostomes) displaying varied levels of parental care, ranging from planktotrophic larvae to complex forms of embryonic incubation (42, 43).

Parental care is thought to not only confer fitness advantages (44) but also, in the case of cheilostomes, hypothesized to be associated with increased speciation rates (16). The given reason is the distribution of brooding with nonfeeding larvae that are unable to survive in the water column for extended periods of time and that thus settle close to their parental colonies. The evolutionary reversal of nonfeeding back to feeding larvae is thought to be uncommon among marine invertebrates (45), and we have shown here for cheilostomes that this is true: Feeding larvae, once lost, never reevolved. Although brooding and nonfeeding larvae are “irreversibly” evolved very early in their history, the apparent higher speciation rates of cheilostomes that brood are unlikely due (directly) to the brooding/nonfeeding larvae. This is in contrast to other empirical studies based on different taxa, which suggest that larva dispersal modes or geographic range sizes associated with them directly influence diversification rates (46).

Rather, it could be “external” factors, such as the macroevolutionary influence from a competing clade, the cyclostomes, which drove their diversification (13), as suggested by a character-independent model of speciation and extinction (Table 1). Alternatively, an unmeasured trait that is associated with brooding/nonfeeding larvae could be responsible for differential rates of diversification in cheilostomes (Table S7). One such trait could be increased polymorphism. For example, spines, considered as modified polymorphic zooids, can (evolutionarily) develop into brood chambers or frontal shields, i.e., morphological structures with functions different from

### Table 1. Comparison of trait-(in)dependent models of diversification

Models of speciation and extinction rates of nonbrooding and brooding cheilostomes are compared using the Akaike criteria. The italicized model (cid2, a character-independent model that is the null version of a BISSE model) has the AIC highest model weight in this set of models, followed closely by a more complex character-independent model (cid4). See table S7 for results based on other topologies.

| Model          | Log likelihood | AIC model weight |
|----------------|----------------|------------------|
| Null           | −254.163       | 5.20 × 10^−6     |
| BISSE          | −248.735       | 4.22 × 10^−4     |
| cid2 (BISSE null) | −240.472     | 0.581            |
| cid4 (HiSSE null) | −238.748     | 0.402            |
| HiSSE          | −239.791       | 0.017            |
the original ones (29). A diversity of traits, derived from polymorphs in a modular construction, which permit varied or even novel ecological function, could allow the occupation of new niches and thus promote macroevolutionary diversification (28).

Phylogenetic topologies and inferences made from them are limited by both taxon and gene sampling (47, 48). Although our phylogeny is the largest in terms of both taxon and gene sampling for cheilostome bryozoans, the inferences presented here are far from final. However, our data are the seed for new data accumulation and our inferred tree is a starting point for many more sophisticated macroevolutionary analyses. Our estimated node ages are subject to the well-known and well-studied limitations of the molecular clock (49), our knowledge of the evolution of the group and its fossil record. The ancestral state reconstructions that we performed did not incorporate trait information from fossil taxa, whose inclusion would most definitely improve such analyses (50). While HiSSE rate models (25) overcome some statistical issues inherent in earlier related models (51), extinction rates estimated from phylogenies based only on extant taxa are still nonide al and also limit the interpretation of our analyses (24, 25), although we focused on trait and speciation rate estimation. Despite the limitations listed, this largest cheilostome tree to date has provided first glimpses of the timing and tempo of evolution of main clades and a key trait for an ecologically and evolutionarily important order that has been overlooked for too long.

This work emphasizes that continued collaborative research between molecular phylogeneticists, systematists, paleontologists, and macroevolutionary biologists can confirm and elucidate relationships, identify important gaps, understand timing and rates of evolution, and open a window into evolution itself, even before integrating substantial data from the fossil record.

**METHODS**

**Sampling and taxon identification**

The procedure summarized herein follows (20) closely with minimum modifications. Colonies whose sequences are presented here were collected and preserved in 70 to 96% ethanol (table S1). Each colony, preliminarily identified to the lowest possible taxonomic level using a stereoscope, was subsampled for DNA isolation and scanning electron microscopy (SEM). While we aimed at sequencing a colony for each distinct species, uncertainty in initial taxon identification using a stereo scope combined with a realization that within-taxon replicates are important for sequence verification compelled us to include such replicates. The scanning electron micrographs (see SEM cards deposited in Zenodo, https://zenodo.org/record/5721078#, YZz39VMo_Fy), taken with a Hitachi TM4000PLUS after bleaching to remove tissue where appropriate, are required for species-level identification and serve as digital vouchers, in addition to physical vouchers deposited at the Natural History Museum in Oslo (table S1). Taxonomic identifications were made independently of but are subsequently verified using the phylogenetic inference and metadata.

Because of the microscopic nature of cheilostomes and their benthic and often encrusting lifestyle, each visible colony (see the previous paragraph) is often a mixed tissue sample that could consist of other organisms including nontarget cheilostome species. Rather than treating nontarget species as contaminants to be discarded, we leverage these to lend clarity to the cheilostome phylogeny (fig. S1). There are three classes of nontarget specimens. The first is where we have found enough macroscopic remains of the nontarget cheilostome, after sequencing, for imaging. In the second class, we did not find any remaining macroscopic material, but given our taxon sampling and observed sequences, we are certain that the contaminant belongs to a given taxon (these are labeled with taxon names and “SEQ” in fig. S1). In yet other cases, given the tree topology and observed sequences, we do not assign the unvouchedered sequences to any known taxon name (these are labeled “UNKNOWN” in fig. S1; see the Supplementary Materials for criteria and examples of all three types of nontarget sequences and table S1 for their metadata).

**DNA isolation, sequencing, and assembly**

The subsamples of colonies (henceforth “samples”) were dried before genomic DNA isolation using the DNeasy Blood and Tissue kit (QIAGEN, Germantown, MD, USA). Samples were homogenized with a pestle in lysis buffer in the presence of proteinase K. Genomic DNA were sequenced at the Norwegian Sequencing Centre (Oslo, Norway) using Illumina HiSeq 4000 150–base pair (bp) paired-end sequencing with a 350-bp insert size. Approximately 20 samples were genome sk immed (multiplexed) on a single lane. Illumina HiSeq reads were quality-checked using FastQC v0.11.8 (52) and then quality- and adapter-trimmed using Trim Galore v0.4.4 with a Phred score cutoff of 30 (53). Trimmed reads were de novo assembled with SPAdes 3.13 (54) using k-mers of 21, 33, 55, 77, 99, and 127. The mitogenome and rRNA operon of each sample were identified separately with BLASTN (55) using BLAST+ against a database constructed from cheilostome sequences available in NCBI (20). An E value of 1.00 × 10−153 and maximum target sequence of 1 were used to filter any blast hits of non-cheilostome origin.

**Annotation**

Mitogenomes for each of the samples were annotated with MitoS2 using a metazoa reference (RefSeq 89) and the invertebrate genetic code (56) to identify two rRNA [mitochondrial large-subunit (16S) ribosomal RNA (rrnL) and mitochondrial small-subunit (12S) ribosomal RNA (rrnS)] and 13 protein coding genes (atp6, atp8, cox1, cox2, cox3, cob, nad1, nad2, nad3, nad4, nad4l, nad5, and nad6). Two nuclear rRNA operon genes (18S and 28S) were also identified and annotated using RNAmer (57). A total of 315 published cheilostome sequences (20, 58–60) and the mitogenomes and rRNA operons of 31 non-cheilostome out-group taxa, both bryozoan and nonbryozoan, were aligned with our sequences to compile a broader out-group taxon sample (table S3).

**Sequence alignment**

MAFFT (61) was used for alignment with default parameters: For the four rRNA genes (nucleotide), the Q-INS-i model, considering secondary RNA structure, was used; for the 13 protein-coding genes, in amino acid format, the G-INS-I model was used. The 17 separate alignments were edited manually using Mesquite v3.61 to remove any uncertain characters (62). Ambiguously aligned characters were removed from each alignment using gblocks (63) with least stringent parameters. The single-gene alignments were concatenated to a supermatrix using the castastaphym tool perl script (64). The alignments (both masked and unmasked) are available at Dryad (https://doi.org/10.5061/dryad.2v6wwpzp9).

**Datasets for phylogenetic reconstruction**

As mentioned in the “Sampling and taxon identification” section, cheilostomes are small and attached to substrata, so even the most macroscopically pristine sample may have sequences of nontarget
species included. Where the contaminants are non-cheilostome, they are removed bioinformatically (see the “DNA isolation, sequencing, and assembly” section). Here, we use the nontarget cheilostome sequences (see also the “Sampling and taxon identification” section). Hence, two concatenated datasets are presented: (i) “Full alignment” is the alignment with our largest taxon sample. It includes both UNKNOWN [cheilostomes lacking both a voucher (physical and/or SEM) and an inferred taxonomic identity] and SEQ [cheilostomes lacking a voucher (SEM) but with an inferred taxonomic identity], constructed to show the hidden diversity within the phylum (fig. S1). Full alignment also includes cheilostomes previously sequenced and available from NCBI and non-cheilostome and non-bryozoan out-groups (table S2). (ii) “Trimmed alignment” is the alignment where “UNKNOWN”, SEQ, and those taxa with less than three genes and “rogue taxa” are pruned using RogueNaRok (65). We picked a three-gene cutoff after preliminary analyses showed that this is the best compromise between the number of taxa included and BS support for our tree inference. Rogue taxa are those with unstable phylogenetic affinities based on evaluation of the extended majority rule consensus threshold, optimized for support and with a maximum dropset size of 1. Those with a sum >0.2 were pruned from the trimmed alignment dataset. Note that the ML tree topologies are termed as “full tree” and “trimmed tree” from the full alignment and trimmed alignment datasets, respectively. For each of the two datasets, ambiguously aligned characters were removed from each single-gene alignment using gblocks (63) with least stringent parameters before concatenation.

**Phylogenetic reconstruction and congruence test**

Maximum likelihood phylogenetic analyses were carried out for each single-gene alignment using the “AUTO” parameter in RAxML v8.0.26 (66) to establish the evolutionary model with the best fit. The general time reversible (GTR + G) was the preferred model for the four rRNA genes (18S, 28S, rrnS, and rrnL), and MtZoa + G for all 13 protein coding genes. The two concatenated datasets (see the “Datasets for phylogenetic reconstruction” section) were divided into four separate rRNA and 13 protein gene partitions each (17 partitions in total), with its own distinct gamma distribution to accommodate for different substitution patterns among sites, and were analyzed using RAxML. For comparison, a partitioning scheme based on Akaike Information Criteria with small sample correction (AICc) and a greedy search scheme suggested by PartitionFinder2 (67) was also analyzed using RAxML. The topology with the highest likelihood score of 100 heuristic searches was chosen, and BS values were calculated from 500 pseudo replicates. As the first partition scheme (17 partitions) gave a higher likelihood score, we present the topology based on that, rather than the one suggested by PartitionFinder2. BS values were calculated from 500 pseudo replicates.

The topology of the phylogenetic tree in this contribution was compared to that from (20) to gauge whether a substantial increase in sampled taxa had any detectable bearing on the inferred topology. To this end, we trimmed samples not represented in (20) from the full tree (fig. S1) and using Dendroscope (68) compared the topology of their remaining 263 shared taxa using the $I_{long}$ index (69).

**Fossil calibration and Bayesian divergence time estimation**

We use MCMCTree v4.9 (70) for divergence time estimation, as it allows us to analyze amino acid and nucleotide partitions simultaneously and takes relatively less computational power than other comparable software. As input to MCMCTree, we use the trimmed tree (fig. S2); but to reduce computational burden further, we removed the following taxa: (i) those lacking species and/or genus designations and those assigned “cf. and aff.” and (ii) species duplicates with the largest number of alignment gaps. If a genus is represented by multiple species, then a maximum of three different named species were retained, choosing those with the least number of alignment gaps. Note that this dataset was created before minor changes detailed in the Supplementary Materials (section S4 in Supplementary Text for Methods). Excluded taxa were deleted from the amino acid and nucleotide alignments, while corresponding leaves for the same taxa were removed using Dendroscope (68), thus maintaining the topological branching pattern of the original rooted input tree (fig. S2). The resulting dataset consisted of 363 taxa where 335 are cheilostomes.

We applied a hard upper limit of 636 Ma ago to the root, representing the bilaterian maximum (71, 72). We used 18 internal fossil calibrated nodes (see table S4 and fig. S3 for details). To explore the impact of calibration prior choice, we ran different sets of analyses: (i) “L,” with minimum constraints only (table S4), using the truncated Cauchy distribution with a soft minimum and a diffuse tail; (ii) “B,” with uniform constraints with soft minimum and maximum bounds corresponding to the fossil ages; and (iii) “ST,” with a skew $T$ distribution, such that the 1 and 99% probability tails correspond to the minimum and maximum constraints (table S4). In all cases, we always used “soft bound,” where there is a 1% chance that a node could be younger or older than the specified constraints. Upon examining initial results for B and ST, we find that we had to impose a hard minimum constrain within our out-group on the Pancrustacea node; otherwise, we recovered an unreasonably young posterior distributions (although we note that this does not affect the ages recovered for the in-group nodes; see fig. S4). We hence also present the B and ST analyses with a hard constraint on the Pancrustacea node only (BL and STL, respectively). For all five sets of analyses (B, BL, L, ST, and STL), we ran both independent and autocorrelated molecular clock models. The main results we present use the independent clock model and STL, as branches close to the root of the tree represent huge evolutionary distances, and because it seems logical to put prior weight around the ages of the fossil calibrations, because the fossil record of cheilostomes is considered excellent. Mixing was checked by inspecting the trace plots and ensuring that the effective sample sizes were greater than 200 for all node ages and model parameters. In addition, we ensured that independent chains converged on the same values. For details of the substitution model, Markov chain Monte Carlo (MCMC) settings, see section S3 in Supplementary Text for Methods.

**Ancestral state reconstruction and HISSE**

Data for nonbrooding species, with planktotrophic larvae (state = 0) and brooding species with nonfeeding larvae (state = 1) of all the cheilostomes species included in the calibration tree (N = 335), are provided in table S5. To estimate brooding states of the internal nodes, we use a standard Markov model of binary character evolution (23) implemented in ape (73), where a maximum likelihood joint estimation procedure was performed. Note that, although the tree for the analyses described here is pruned, the nonbrooding/brooding states that we are concerned are conserved at the genus level. To detect possible differences in diversification rates associated with the nonbrooding or brooding state, we applied
trait-dependent speciation and extinction models implemented in the R package HiSSE (25) to the fossil calibrated tree (N = 355), where we used the STL calibration with an independent clock model (see the “Fossil calibration and Bayesian divergence time estimation” section). We estimate that we have sampled 0.7 and 9.5% of species of nonbrooders and brooders (17), respectively, in our calibration tree and use this as information to account for biases due to incomplete sampling. We ran three different null models (“null,” cid2, and “cid4”), a BiSSE model, and a HiSSE model to investigate whether brooding might be associated with higher speciation rates. The null constrains speciation and extinction to be equal regardless of brooding state. A BiSSE model allows speciation and extinction to be different for the nonbrooding versus brooding state. The first character-independent model (cid2) allows two different sets of speciation and extinction to be estimated but does not link these to the observed traits, such that it has the same level of complexity as the null version of the BiSSE model. A HiSSE model assumes that there are unmeasured states that display distinct rates of speciation and extinction but that these states are associated with the coded state. In other words, HiSSE allows for the scenario in which a state coassociated with brooding drives the differences between the observed differences in speciation among species with and without brooding. The second character-independent model (cid4) allows four different sets of speciation and extinction, such that the model has the same level of complexity as a HiSSE model and serves as its null model. The five models are compared using AIC model weights, and their parameter estimates are also presented.

Because the topological placement of Lunularia is starkly incongruent with morphology as we understand it (see the Supplementary Materials), we also compared the same five models with a time tree where Lunularia is removed. In addition, as the BS support for Conopeum is weaker than for other major nodes in our tree and because it is a key taxon, we again compared the same models but using an alternative topology where Conopeum is placed as sister to all other cheilostomes.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abm7452

View/request a protocol for this paper from Bio-protocol.

REFERENCES AND NOTES

1. R. M. D. Beck, C. Baillie, Improvements in the fossil record may largely resolve current conflicts between morphological and molecular estimates of mammal phylogeny. Proc. R. Soc. B Biol. Sci. 285, 20181632 (2018).
2. N. M. Koch, L. A. Parry, Death is on our side: Paleontological data drastically modify conflicts between morphological and molecular estimates of mammal phylogeny. Proc. R. Soc. B Biol. Sci. 285, 20181632 (2018).
3. F. Ronquist, L. Vilhelmsen, S. Schulmeister, D. L. Murray, A. P. Rasnitsyn, A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenopera. Syst. Biol. 69, 1052–1067 (2020).
4. D. W. Bapst, H. A. Schreiber, S. J. Carlson, Combined analysis of extant Rhynchonellida (Brachiopoda) using morphological and molecular data. Syst. Biol. 67, 32–48 (2018).
5. A. B. Wood, C. D. Orton, S. M. Smith, Complex habitat generated by brooding. The second character-independent model (cid2) allows two different sets of speciation and extinction, such that the model has the same level of complexity as a HiSSE model and serves as its null model. The five models are compared using AIC model weights, and their parameter estimates are also presented.

Before the topological placement of Lunularia is starkly incongruent with morphology as we understand it (see the Supplementary Materials), we also compared the same five models with a time tree where Lunularia is removed. In addition, as the BS support for Conopeum is weaker than for other major nodes in our tree and because it is a key taxon, we again compared the same models but using an alternative topology where Conopeum is placed as sister to all other cheilostomes.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abm7452

View/request a protocol for this paper from Bio-protocol.

REFERENCES AND NOTES

1. R. M. D. Beck, C. Baillie, Improvements in the fossil record may largely resolve current conflicts between morphological and molecular estimates of mammal phylogeny. Proc. R. Soc. B Biol. Sci. 285, 20181632 (2018).
2. N. M. Koch, L. A. Parry, Death is on our side: Paleontological data drastically modify conflicts between morphological and molecular estimates of mammal phylogeny. Proc. R. Soc. B Biol. Sci. 285, 20181632 (2018).
3. F. Ronquist, L. Vilhelmsen, S. Schulmeister, D. L. Murray, A. P. Rasnitsyn, A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenopera. Syst. Biol. 69, 1052–1067 (2020).
4. D. W. Bapst, H. A. Schreiber, S. J. Carlson, Combined analysis of extant Rhynchonellida (Brachiopoda) using morphological and molecular data. Syst. Biol. 67, 32–48 (2018).
5. A. B. Wood, C. D. Orton, S. M. Smith, Complex habitat generated by brooding. The second character-independent model (cid2) allows two different sets of speciation and extinction, such that the model has the same level of complexity as a HiSSE model and serves as its null model. The five models are compared using AIC model weights, and their parameter estimates are also presented.

Before the topological placement of Lunularia is starkly incongruent with morphology as we understand it (see the Supplementary Materials), we also compared the same five models with a time tree where Lunularia is removed. In addition, as the BS support for Conopeum is weaker than for other major nodes in our tree and because it is a key taxon, we again compared the same models but using an alternative topology where Conopeum is placed as sister to all other cheilostomes.
99. P. D. Taylor, B. Berning, M. A. Wilson, Reinterpretation of the Cambrian “bryozoan” Pywackia as an octocoral. *J. Paleol.* **87**, 964–990 (2013).

100. J. Ma, P. D. Taylor, F. Xia, R. Zhan, The oldest known bryozoan: Prophyllodictya (Cryptostomata) from the lower Tremadocian (Lower Ordovician) of Liujiachang, south-western Hubei, central China. *Palaeontology* **58**, 925–934 (2015).

101. E. Landing, J. B. Antcliffe, M. D. Brasier, A. B. English, Distinguishing Earth’s oldest known bryozoan (Pywackia, late Cambrian) from pennatulacean octocorals (Mesozoic—Recent). *J. Paleol.* **89**, 292–317 (2015).

102. S. J. Hageman, A. Ernst, The last phylum: Occupation of Bryozoa morpho-ecospace (colony growth habits) during the early phase of the Great Ordovician Biodiversification Event. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **534**, 109270 (2019).

103. G. Hillmer, A 300-million-year gap in the bryozoan fossil record. *Naturwissenschaften* **78**, 123–125 (1991).

104. A. Ernst, in *Handbook of Zoology: Phylum Bryozoa*, T. Schwaha, Ed. (De Gruyter, 2021), p. 11–55.

**Acknowledgments:** We thank C. Seid (Scripps Institution of Oceanography), A. M. Bernis (University of Florida), S. Mills (NIWA), K. Moore (Tasmanian Museum and Art Gallery), U.S. Coast Guard, California Department of Fish and Wildlife, and all collectors named in the attached SEM cards, but especially M. Capa, M. Dick, M. Zabala, C. Avila, and R. Tan. The staff at Heron Island Research Station (University of Queensland) and at the Research and Educational Station “Belomorskaya” (Saint Petersburg State University) supported us for fieldwork. P. D. Taylor and J. E. Winston discussed earlier versions of this phylogeny with us. We thank A.-H. Rønning, B. L. G. Thorbek, and the Norwegian Sequencing Centre and the High-Performance Computing Cluster at the University of Oslo for assistance. We thank J. Beaulieu for help on the HISSE package. See the Supplementary Materials for permits. **Funding:** This project is funded by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 724324 to L.H.L.). Extra funding for specific sampling expeditions were provided by the National Science Centre of Poland (grant PANIC/2016/23/B/ST10/01936 to P.K.), the ASSEMBLE Plus (Horizon 2020), the Leverhulme Trust (Research Project Award RPG-2016-429 to A.W.), the Russian Science Foundation (grant 18-14-00086 to A.N.O.), Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil [PQ-CNpq 308768/2018-3], Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil [FAPESP 19/17721-9] and the DISTANTCOM project [ICTM2013-42667/ANT to C. Avila]. B.F. was supported by Beatriu de Pinós (2019-BP-00183), funded by the AGAUR (Government of Catalonia) and by the Horizon 2020 programme under the Marie Skłodowska-Curie grant (no. 801370). The CAML-CEAMARC cruise of RV Aurora Australis (IPY project no. 53) were supported by the Australian Antarctic Division, the Japanese Science Foundation, the French Polar Institute IPEV, and the Muséum national d’Histoire naturelle, led by voyage leader, M. Riddle. **Author contributions:** R.J.S.O. performed the preparation for sequencing, the bioinformatics, and phylogenetic analyses. E.D.M. and M.H.R. did the vouchering work. E.D.M. provided the calibration points. E.D.M. and D.P.G. drafted the interpretations. R.C.M.W. contributed to the fossil calibration. E.D.M., D.P.G., L.M.V., B.B., B.F., P.K., O.N.K., N.N.S., A.N.O., J.-G.H., J.S.-D., R.L.C., K.L.V., and L.M. identified the specimens. All authors contributed to the collecting of samples. L.H.L. obtained the funding, coordinated the project, performed analyses downstream of phylogenetic analyses with R.J.S.O. and R.C.M.W., and cowrote the first draft of the manuscript with R.J.S.O., E.D.M., and D.P.G., which all coauthors revised. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, Dryad, and Zenodo.

Submitted 15 October 2021
Accepted 2 February 2022
Published 30 March 2022
10.1126/sciadv.abm7452
Paleozoic origins of cheilostome bryozoans and their parental care inferred by a new genome-skimmed phylogeny

Russell J. S. Orr, Emanuela Di Martino, Mali H. Ramsfjell, Dennis P. Gordon, Björn Berning, Ismael Chowdhury, Sean Craig, Robyn L. Cumming, Blanca Figuerola, Wayne Florence, Jean-Georges Harmelin, Masato Hirose, Danwei Huang, Sudhanshi S. Jain, Helen L. Jenkins, Olga N. Kotenko, Piotr Kuklinski, Hannah E. Lee, Teresa Madurell, Linda McCann, Hannah L. Mello, Matthias Obst, Andrew N. Ostrovsky, Gustav Paulay, Joanne S. Porter, Natalia N. Shunatova, Abigail M. Smith, Javier Souto-Derungs, Leandro M. Vieira, Kjetil L. Voje, Andrea Waeschenbach, Kamil Zágoršek, Rachel C. M. Warnock, Lee Hsiang Liow

Sci. Adv., 8 (13), eabm7452. • DOI: 10.1126/sciadv.abm7452

View the article online
https://www.science.org/doi/10.1126/sciadv.abm7452
Permissions
https://www.science.org/help/reprints-and-permissions