Bayesian phylogenetics of bryozoa
Hayward, P; Porter, Joanne; Skibinski, D O F; Tsyganov-Bodounov, A

Published in:
Molecular Phylogenetics and Evolution

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Heriot-Watt University Research Portal

Citation for published version (APA):
Bayesian phylogenetics of Bryozoa

Anton Tsyganov-Bodounov,*, Peter J. Hayward, Joanne S. Porter, David O.F. Skibinski

Abstract

1. Introduction

The phylum Bryozoa is the largest phylum of the lophophorate invertebrates. The number of extant species has been estimated at around 5600 (Todd, 2000), but is probably greater (Hayward and Ryland, 1999). The phylum comprises three classes: Gymnolaemata, Stenolaemata and Phylactolaemata. The affinity of the Phylactolaemata with the rest of the Bryozoa is uncertain, and is disputed (Mundy et al., 1981). The Gymnolaemata consists of more than 3000 species classified in two orders, the Ctenostomata and the Cheilostomata. The Ctenostomata is further subdivided into two suborders, the Stolonifera and the Carnosa. The Cheilostomata first appears in Upper Jurassic horizons, and has remained the dominant bryozoan group (Taylor and Larwood, 1990). Four suborders are presently recognised within the Cheilostomata, the Inovicellata, Malacostegina, Neochelostomina and Scrupariina, and the Neochelostomina encompasses two infraorders, the Ascophora and the Flustrina (Gordon, unpublished). The class Stenolaemata is believed to have originated in the Lower Ordovician (approximately 480 mya) with the majority of taxa belonging to four extint orders. The order Cyclostomata remains the only extant order and includes five (Taylor, 2000) or six (Kluge, 1975) suborders: Articulata, Cancellata, Cerioporina, Isoporina, Rectangulata and Tubuliporina.

Phylogenetic studies of the Bryozoa are limited, and controversial in their findings. Despite the fact that numerous skeletal morphological characters are readily available, both for fossil and extant species, only a few studies have employed computerised cladistic methodologies in phylogenetic studies. However, these studies have been criticised by Todd (2000), who, on the basis of skeletal and morphological data for fossil and extant species, suggested that the Ctenostomata were paraphyletic, with Stenolaemata (Cyclostomata) and Cheilostomata nesting within the Ctenostomata. Polyphyly of the cheilostomes based on skeletal morphology has also been suggested by Gordon (2000) and Voigt (1991).

Until very recently only three molecular phylogenetics studies had been conducted (Dick et al., 2000; Hao et al., 2005, 2002) and results are inconclusive. Using 16S rRNA, Dick et al. (2000), found the Ctenostomata and Cheilostomata to be paraphyletic, with the Cyclostomata polyphyletic, and both cheilostomes and cyclostomes embedded within the Ctenostomata. These results bore little resemblance to the more commonly accepted tree topologies, in which Cyclostomata and Cheilostomata are monophyletic (Todd, 2000). However, the suitability of the 16S rRNA gene for phylogenetic studies is limited by its potential to resolve divergences only as far back as the mid-Cretaceous, which may...
be insufficient given the palaeontological record of the Bryozoa (Dick et al., 2000). Hao et al. (2002) employed 18S rRNA in their analysis, but used a limited number (12) of sequences, giving inconclusive results. Subsequently, it has been suggested, following phylogenetic reconstruction analysis, that some of the sequences used in that study (Membranipora sp., AF119081, and Lichenopora sp., AF119080) may have represented contaminants (Waeschenbach, 2003). Finally, Hao et al. (2005) re-evaluated cheilostome phylogenetic reconstruction analysis, that some of the sequences were inconclusive, but also presented results that conflicted with those of the morphological studies. A very recent study on the molecular phylogeny of Byzyozoa is now available (Fuchs et al., 2009) based on 18S rDNA, 28S rDNA and the mitochondrial CO1 gene and using 32 species. This suggests monophyly of bryozoan classes, but ambiguous results for the relationships amongst them.

In the present study the 18S rRNA gene has been used in a study of Bryozoa phylogenetics. The use of this gene is hindered by the difficulty of aligning the sequences due to the presence of variable regions in the secondary structure. However, the importance of an accurate alignment of 18S rRNA for successful retention of homologous characters within the aligned sequences has been emphasised by many authors, see Xia et al. (2003) for review. In phylogenetic studies of Bryozoa, secondary structure has not been used during alignment (Dick et al., 2000; Hao et al., 2005; Fuchs et al., 2009), or if used, variable regions were excluded (Hao et al., 2002), an approach which has been criticised (Xia et al., 2003). In the present study the 18S rRNA was aligned using a secondary structure model.

2. Materials and methods

2.1. Sample collection

Samples were collected from sites in South Wales: Skomer Island, Dale Harbour, Watbyck Bay, Pembroke Ferry, Lydstep Bay, Bracelet Bay, Mumbles Pier and by diving in Skomer Island. To avoid possible contamination by foreign DNA when colony somatic tissue is used, DNA extraction from embryos was employed (Porter et al., 2001). Live larvae or embryos were dissected from colonies and washed three times in filtered sterile water to minimise contamination prior to DNA extraction. DNA was extracted by direct lysis: individual larvae were transferred into a 15 μl lysis solution (7.5 mM Tris–HCl, pH 8.3; 3.75 mM NH4Cl; 3.75 mM KCl; 1.5 mM MgCl2; 2 μg proteinase-K) then incubated for 1.5 h followed by deactivation of proteinase-K by heating at 99 °C for 10 min. Samples were stored at −20 °C.

2.2. PCR conditions

PCR was performed in a total volume of 15 μl using Taq DNA polymerase (ABgene), Buffer II (ABgene), 0.5–1.5 μl of 100 μM primers (see below), and MgCl2 (ABgene) adjusted per sample. Cycling conditions were as follows: initial denaturation at 95 °C for 1 min, 33 cycles of annealing for 1 min at 40–70 °C (depending on primers used), 90 s extension at 72 °C, terminated with a final extension at 72 °C for 10 min. Bryozoan-specific primers were designed and are given in Table 1. For each species, three sets of primers for overlapping segments were used because of the length of the gene (these are marked as Set 1, Set 2, Set 3, respectively, in Table 1).

PCR products were purified using the Promega DNA purification kit (Wizard® PCR Preps DNA Purification System). Direct sequencing (using the same set of primers as for the PCR) was done on an Applied Biosystems 3730 DNA Analyser automated DNA sequencer. For cloning, samples were purified using GFX™ PCR DNA and Gel Purification Kit (Amersham Biosciences). Cloning was done using pGEM®-T Easy Vector System (Promega). Transformation by electroporation was carried out using E. coli JM109 electro-competent cells in 1 mm electroporation cuvettes (HiMax EP-101 CellProjects) and MicroPulseTM electroporation apparatus (Bio-RAD). Plasmid extraction was done using the Promega DNA purification system (Wizard® Plus SV Minipreps DNA Purification System). Following cloning, sequencing was done using the same protocol as for direct sequencing, using T7 and SP6 primers. All sequences were obtained in both directions. The three overlapping segments were assembled manually into a consensus sequence.

2.3. Phylogenetic analysis

In total, 27 Bryozoan 18S rRNA sequences were obtained (17 Cheilostomata, 5 Ctenostomata and 5 Cyclostomata). Other sequences were taken from NCBI GenBank. Table 2 lists all sequences used here.

2.4. Alignment

Improved alignments using secondary structure have been achieved in many phylogenetic studies, see (Xia et al., 2003) for detailed discussion. Here, hypervariable regions of rRNA were included in the alignment as their omission may remove informative sites (Kjer, 1995) and lead to incorrectly inferred phylogenies (Xia et al., 2003). However, hypervariable regions are often impossible to align with available software.

Therefore, secondary structure alignment was performed by hand, based on the procedure modified from Kjer (1995). To assist with the alignment several sequences of 18S rRNA from the European Ribosomal RNA Database (ERRD) were used: two bryozoan species, three entoprocts, and two brachiopod species. These sequences were converted from the ERM format using a set of tools X-stem and Y-stem (Telford et al., 2005). The species that had the sequence spanning the 1-stem of helix 1 to the 50-
of helix 50, and thus covering the entire 18S rRNA gene, viz. Bugula turbinate, was used as the model for bryozoan 18S rRNA. Helix numbering followed Van de Peer et al., 2000; numbering for helices E23_1 to E23_14 followed Wuys et al., 2000. The secondary structure model for Bryozoa was evaluated and built based on previous findings show the superiority of RNA16-based specific models over GTR models when used for stems (Telford et al., 2005) as a suitable model for the loop regions of the rRNA. A discrete gamma (Γ) distribution using four categories and a proportion of invariable sites (I) were estimated for both stem and loop models separately, thus the models are RNA16HKY + Γ + I and GTR + Γ + I.

2.5. Model selection for phylogenetic analysis

Separate models were used for stems and variable regions. Previous findings show the superior performance of RNA16-based specific models over GTR models when used for stems (Telford et al., 2005). MrBayes employs the RNA16B model (Savill et al., 2001) and its derivatives i.e. RNA16B-JC, RNA16B-HKY and RNA16B-GTR, corresponding, respectively, to JC-like, HKY-like and GTR-like models with respect to substitution rates. The RNA16B-GTR model appears to be only very slightly better for stems than RNA16B-HKY (Telford et al., 2005) and takes significantly more time to evaluate in MrBayes, hence the latter was chosen as a stem model. This model has 16 frequency parameters and two substitution rate parameters. For the variable regions a GTR model was chosen, based on the results of MrModeltest (Nylander, 2004) using AIC as criterion, and because this model has been suggested elsewhere (Telford et al., 2005) as a suitable model for the loop regions of the rRNA. A discrete gamma (Γ) distribution using four categories and a proportion of invariable sites (I) were estimated for both stem and loop models separately, thus the models are RNA16HKY + Γ + I and GTR + Γ + I.

2.6. Bayesian analysis

Phylogenetic trees were constructed using MrBayes 3.1 software (Huelsenbeck and Ronquist, 2001), compiled for MPI parallel use on Blue C, a 2.7 teraflops IBM Power 5 series cluster running AXI UNIX. Multi-node architecture allowed spreading of independent Markov Chain Monte Carlo (MCMC) MrBayes runs and individual chains throughout the cluster nodes and processor cores, respectively, thus speeding up calculations. Convergence diagnostics were performed by assessing the correlation between the posterior probabilities of the individual cladades found in separate chains and runs as well as use of log likelihoods. Separate plots of cumulative split frequencies for selected taxon bipartitions over an entire MCMC analysis run were
evaluated using AWTY (Wilgenbuch et al., 2004). The aligned data-set was divided into two partitions corresponding to stems and loops of the rRNA and assigned the models: RNA16HKY + Γ + I and GTR + Γ + I, respectively. The number of chains per individual run was limited to four (three hot chains and one cold chain per run). In total 16 chains, four chains for four parallel runs, were used. The heated chains “temperature” parameters, as well as prior settings, were left at the default values (chain temp = 0.2). The analysis was limited to 60,000,000 generations or approximately 1 month of real time and sampled each 1000th generation. The burn-in value was determined based on log likelihood values and AWTY analysis of variable bipartition stability.

2.7. Stratigraphic congruence

Stratigraphic congruence analysis was carried out to compare the general structure of the tree in terms of node positioning and branching to predictions obtained from the fossil record using SCI (Huelsenbeck, 1994), RCI (Benton, 1994) and GER (Wills, 1999) indices. Stratigraphic data were taken from Taylor (1993) for Bryozoa, Smirnova (1997) for Brachiopoda, and Todd and Taylor (1992) for Entoprocta, and assigned at the family level. When no fossil record was available the stratigraphic range was assigned as Recent to Recent. The significance levels for three indices were calculated using a permutation test (1000 permutations) (Huelsenbeck, 1994) with test values below 5% considered to indicate a significantly good fit between the molecular and stratigraphic data (Wills, 1999).

3. Results and discussion

3.1. Phylogenetic analysis

The 18S rRNA model of Bugula turbinata is presented in Fig. 1. This model was used here to partition the sequences into stem
and loop regions for the phylogenetic analysis. The model should be useful to assist alignments of 18S rDNA in future work on bryozoan phylogenetics. To this end the file created during the alignment can be used as a skeleton for folding of other bryozoan 18S rRNA sequences, and can be supplied upon request.

For the entire dataset of the 34 taxa, a total of 2046 characters was used in the data matrix with 643 unique characters for the loop partition and 259 unique characters for the stem partition. The phylogenetic tree rooted with *Neocrania anomala* is shown in Fig. 2. The general features of the tree are discussed in detail below but two seemingly anomalous features are considered first. *Bugula plumosa* (obtained here), appeared to be clustering abnormally with the Ctenostomata (specifically with Vesiculariidae), although other Bugulidae appeared to cluster with each other, and within the order Cheilostomata, as expected (Fig. 2). This is hard to explain. There is little support in the taxonomic literature for non-uniformity within this well-described genus, and on the basis of alignment and secondary structure there is no reason to doubt that it is a bona fide bryozoan 18S rRNA sequence. Misclassification or cross-contamination of the sample might have been the cause of this result. Because of this uncertainty the sequence is excluded from further discussion, and has not been submitted to GenBank.

The second anomaly is that *Scruparia chelata* is not positioned as expected with other cheilostomes. Only two species of *Scruparia* are known, and together within one other taxon, *Brettiopsis triple*, are classified within the suborder Scrupariina, based on the similarities of the distinctive bivalve brood chamber of the brooding zooids. There is very little information available on *Scruparia chelata*, its reproductive cycle and larval biology. Here, *Scruparia chelata* appeared as a sister taxon to the Cyclostomata with low posterior probability support. While the marked morphological differences between species of Scrupariina and other cheilostomes might be reflected in *S. chelata* being uniquely positioned relative to other cheilostomes, its position as a sister taxon to the Cyclostomata is not predicted. The *Bugula plumosa* and *Scruparia chelata* sequences were retained in the tree, however, because their bryozoan origins are assured.

If the position of *Scruparia chelata* is ignored then the phylogenetic tree (Fig. 2) shows the Gymnolaemata as a sister group to a clade containing Phylactolaemata, Entoprocta and Cyclostomata, thus making the Cyclostomata (Stenolaemata) a sister taxon to the Gymnolaemata. The above relationships contrast with the notion that ctenostomes are ancestral to stenolaemates (Larwood and Taylor, 1979). In particular, ctenostomes are believed to be
paraphyletic with both stenolaemates (cyclostomes) and cheilostomes nested within them (Todd, 2000). Cyclostomes (Stenolaemata) forming a monophyletic clade is in full agreement with other hypotheses (Taylor and Larwood, 1990; Todd, 2000).

The paraphyletic Ctenostomata containing monophyletic Cheilostomata is in agreement with Taylor and Larwood (1990) and Todd (2000), providing that the position of the stenolaemates (Cyclostomata) is ignored. The phylogenetic relationships are in agreement with the results of Fuchs et al. (2009) in that Cyclostomata are a well supported monophyletic group. In both studies Ctenostomata are not present as a monophyletic group. There is thus some general agreement even though the selection of species in the two studies differs. The sub-order structure is rather different between studies, though in general support for individual clades is higher in the present study. These differences could be due to methodological and technical differences between the studies in addition to species coverage. Thus the Fuchs et al. (2009) study used a shorter 18S sequence (1713 vs. 2046 nucleotides) whereas their total sequence length was longer (2724) comprising three gene regions. Other technical differences pertain to the methods of DNA extraction, the nature of the PCR primers and the use of secondary structure in alignment. It is of great value, however, to compare and contrast trees obtained by different groups using different approaches for cross-corroboration and identification of the phylogenetic uncertainties that still remain.

3.2. Cyclostomata

Within the monophyletic cyclostome clade there was a very strong posterior probability support for the nodes. Among the cyclostome sequences, the family Crisiidae was monophyletic although the genus Crisia was not fully resolved: Filicrisia geniculata showed polytomy with Crisia denticulata with the remaining two Crisia spp. fully resolved. Tubulipora liliacea, was at the root of the clade as a sister group to the rest of cyclostomes.

3.3. Ctenostomata

Ctenostomes showed paraphyly supporting previous hypotheses based on palaeontological findings (Todd, 2000), containing monophyletic Cheilostomata. Within the ctenostomes themselves there was a slightly lower support for families. Thus the four stoloniferous ctenostomes (three Bowerbankia spp. and one Walkeria) did not form a clade nor was the genus Bowerbankia monophyletic. Finally, Flustrellidra hispida together with Alcyonidium gelatinosum were at the root of the clade as a sister group to the rest of ctenostomes.

3.4. Cheilostomata

The cheilostomes showed monophyly, if Scruparia chelata and Bugula plumosa were excluded, which is in line with the common assumption of monophyletic Cheilostomata (Gordon, 2000; Taylor and Larwood, 1990; Todd, 2000). However, the current classification of higher taxa within the Cheilostomata is based largely on the morphology and structure of the frontal wall (Gordon, 2000), and thus the possibility of homoplasy has to be considered, using further molecular data, as it becomes available. The grouping of several species with very high posterior probability support still did not recover all expected taxonomic grades within the Ascophora. For instance, the Lepraliomorpha were still paraphyletic.

The polyphyly of the Ascophora, and the assumption of a common ctenostome ancestor, appears to be consistent with other findings based on different evolutionary models of frontal shields (Gordon, 2000) and their ontogeny and structure (Voigt, 1991). Further, the nesting of Celleporidae (Cellepora) and Smittiniidae (Schizomavella) within the Umbonulomorpha (Escharella, Escharo-

ides and Umbonula) is in agreement with the morphological model for an umbonuloid origin of lepralioid frontal shields (Gordon, 2000). However, the ancestor of the clade including all ascophorans could be an ascophoran, in which case ascophorans would be paraphyletic. However, a complete recovery of all other cheilostomes within the paraphyletic Ascophora contradicts the paradigm of flustrine cheilostomes giving rise to the more complicated and advanced ascophorans (Gordon, 2000). Nesting of the monophyletic Flustrina within the Ascopora is hard to explain, and requires more anascan sequences from 18S rRNA, and possibly other genes, in order to test this result further.

One interesting finding is that within cheilostomes the infraorder Flustrina, was monophyletic, and had a very strong support and the two monophyletic families of Flustrina, Bugulidae and Calloporidae, were recovered with equally high support.

3.5. Stratigraphic congruence

The stratigraphic congruence analysis was carried out to determine whether the structure of the tree is consistent with information from the fossil record currently available in the literature. Because of the lack of genus level information, the stratigraphic ranges were assigned at the family level.

The results of the congruence tests are given together with significance levels (Huelsnack, 1994; Wills, 1998), based on 1000 permutations. The results for the SCI index (0.56; significance 0.3%) are in line with previously published data for stratigraphic congruence for molecular trees, i.e. 0.4–0.6 (Benton, 1998). The RCI values (64.90; significance 0.1%) are lower than expected (ca. 80 for published data), but still considered to indicate good fit between the two types of data (M.A. Wills, personal communication). The value of the GER index is high (0.95; significance 0.1%). Despite the high value of the GER index, the values of the SCI and RCI indices were lower than expected for a very good agreement between the two data types. This might reflect either inaccuracies in the structure of the tree, which might be resolved by adding more taxa and clarifying the position of Scruparia chelata, or might reflect incompleteness of the stratigraphic record.

The present study will enable further 18S rRNA bryozoan sequences to be obtained with relative ease through use of the primers developed here. The 18S rRNA secondary structure presented here can be used for better alignment of newly obtained sequences. The question remains open on the position of Scruparia chelata, and specific attention has to be given to acquiring further Alcyonidiidae sequences in order to clarify relationships within the Ctenostomata–Cheilostomata assemblage. The addition of ctenostome sequences, in particular for the Alcyonidiidae, may change the topology of the tree, in particular with reference to relationships within the Gymnoalaemata.

Acknowledgments

We thank NERC (grantNER/S/J/2002/12005) and Swansea University for financial support. The computational aspects of this work were made possible by the IBM Blue C Supercomputer at the School of Medicine, Swansea University. We thank Mike Gravenor, Sergei Kosakovskiy-Pond, and IBM staff for advice regarding the implementation of the programs used on Blue C. We thank Dr. Paul Taylor for valuable suggestions and comments on the manuscript.

References
