Molecular Phylogenetics and Evolution 94 (2016) 730–736

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Multigene phylogeny of the red algal subclass Nemaliophycidae ☆

Daryl W. Lam [a], Heroen Verbruggen [b], Gary W. Saunders [c], Morgan L. Vis [a,*]

[a] Department of Environmental and Plant Biology, Ohio University, Athens, OH 45701, United States
[b] School of BioSciences, University of Melbourne, Victoria 3010, Australia
[c] Centre for Environmental & Molecular Algal Research, Department of Biology, University of New Brunswick, Fredericton, NB E3B 5A3, Canada

ARTICLE INFO

Article history:
Received 4 March 2015
Revised 11 October 2015
Accepted 13 October 2015
Available online 27 October 2015

Keywords:
Red algae
Tree of life
Freshwater macrophytes
Seaweeds
Simulation analysis
Hard polytomy

ABSTRACT

The red algae (Rhodophyta) are a lineage of primary endosymbionts whose ancestors represent some of the first photosynthetic eukaryotes on the planet. They primarily inhabit marine ecosystems, with only ~5% of species found in freshwater systems. The subclass Nemaliophycidae is very diverse in ecological and life history features and therefore a useful model to study these traits, but the phylogenetic relationships among the orders are, for the most part, poorly resolved. To elucidate the phylogeny of the Nemaliophycidae, we constructed a nine-gene dataset comprised of nuclear, plastid, and mitochondrial markers for 67 red algal specimens. The resulting maximum likelihood (ML) phylogeny confirmed the monophyly of all orders. The sister relationship of the Acrochaetiales and Palmariales received high support and the relationship of the Balliales with Balbianiales and Entwisleiales with Colaconematales was moderately supported. The Nemaliales, Entwisleiales, Colaconematales, Palmariales and Acrochaetiales formed a highly supported clade. Unfortunately, all other relationships among the orders had low bootstrap support. Although the ML analysis did not resolve many of the relationships, further analyses suggested that a resolution is possible. A Phycas analysis supported a dichotomously branching tree and Bayesian analysis showed a similar topology with all relationships highly supported. Simulations extrapolating the number of nucleotide characters beyond the current size of the dataset suggested that most nodes in the phylogeny would be resolved if more data become available. Phylogenomic approaches will be necessary to provide a well-supported phylogeny of this subclass with all relationships resolved such that the evolution of freshwater species from marine ancestors as well as reproductive traits can be explored.

1. Introduction

The red algae (Rhodophyta) are a diverse group of ca. 7000 species of photosynthetic eukaryotes with a fossil record dating back 1.2 billion years (Butterfield, 2000; Guiry and Guiry, 2015). Their chloroplasts derive from a primary endosymbiosis event (Delwiche, 1999) and possess light-harvesting complexes called phycobilisomes that contain the red and blue pigments phycoerythrin, phycocyanin and allophycocyanin (Gantt, 1990). These pigments allow marine red algae to tolerate a sizeable range in light levels and live in deeper waters as compared to other photoautotrophic organisms (Kain and Norton, 1990). Red algae lack typical microtubule structures such as flagella and centrioles (Woelkerling, 1990). Many red algae have pit connections between adjacent cells that may function to facilitate cell-to-cell communication (Pueschel, 1990). These connections feature a capped plug comprised primarily of protein, whose shape is an important diagnostic feature at the subclass and order level (Saunders and Bailey, 1997, 1999; Saunders and Hommersand, 2004).

Of the five subclasses of Florideophyceae, the Nemaliophycidae is the most biologically diverse providing an ideal test case to study red algal evolution. Traits of interest include, among others, habitat transitions between marine and freshwater ecosystems, evolution of complex thalli from filamentous forms and vice versa (Saunders and Kraft, 1997; Saunders and Hommersand, 2004), species with and without calcification (Huisman et al., 2004), and uncharacteristic variability in phycoerythrin type (e.g., Saunders et al., 1995). Although there are other freshwater red algae (e.g., *Bangia*, *Compsopogon*, *Hildenbrandia*), only the Nemaliophycidae contains orders comprised strictly of freshwater taxa (Balbianiales, Batrachospermales, Thoreales) along with strictly marine orders (Balliales, Colaconematales, Entwisleiales, Nemaliales, Palmariales,

☆ This paper was edited by the Associate Editor Elizabeth Zimmer.

* Corresponding author.
E-mail addresses: daryl.w.lam@gmail.com (D.W. Lam), heroen.verbruggen@gmail.com (H. Verbruggen), gws@unb.ca (G.W. Saunders), vis-chia@ohio.edu (M.L. Vis).

http://dx.doi.org/10.1016/j.ympev.2015.10.015
1055-7903/

and Rhodachlyales), while Acrochaetiales is primarily marine with a few freshwater species in the genus *Audouinella* (Kumano, 2002). Reproductively, life histories have transitioned between alternations of heteromorphic and isomorphic generations (e.g., Huisman et al., 2004), and vary from the standard florideophycean triphasic pattern (a haploid gametophyte, a diploid carposporophyte, and a diploid tetrasporophyte; e.g., Huisman et al., 2004), to patterns lacking production of tetrasporangia (site of meiosis in most red algae), to biphasic patterns lacking a carposporophyte with (Van der Meer and Todd, 1980) and without (e.g., Stegenga, 1978) stark sexual dimorphism, and finally monophasic species, in which the gametophytic stage directly produces diploid tetrasporangia following fertilization (e.g., DeCew and West, 1982). In addition asexual reproduction in the midst of sexual lineages has been documented at the population (e.g., *Rhodophysema elegans*; Saunders et al., 1989), species (e.g., *Rhodophysema georgii*; Saunders and Bird, 1989), genus (e.g., *Camontagnea*; Womersley, 1994) and family (e.g., Meiodiscaceae; Clayden and Saunders, 2010) levels of taxonomy. Most nemaliophycidaean species are free living and grow attached to a variety of substrata, however, some taxa grow exclusively inside of other algae (e.g., *Rhododrewia porphyrae*; Clayden and Saunders, 2014) or in marine invertebrates (e.g., *Rubrointrusa membranacea* inside the hydroid *Dynamena pumila*; Clayden and Saunders, 2010) while *Rhodophysema kjellmanii* is unusual among red algal parasites by occurring in a lineage (Clayden and Saunders, 2014) that is essentially devoid of secondary pit connections (see Blouin and Lane, 2012).

To understand the evolutionary diversification of these traits, a solid phylogeny needs to be reconstructed. An association between the Acrochaetiales and Palmariales, and a complex of the previous two with the Colaconematales and Nemaliales was shown by the earliest SSU + LSU analyses of Harper and Saunders (2001, 2002), but remarkably few improvements have been made to the phylogeny since. The phylogenetic relationships among orders in the Nemaliophycidae (as currently defined) were not well-supported in the three-gene phylogeny by Le Gall and Saunders (2007) except for additional support for the relationships outlined previously, moderate evidence for an alliance between the Balbianiales and Balliales, and variable indications that the Batrachospermales were the deepest diverging lineage in this subclass. A data mining approach targeted at identifying research priorities for red algal phylogenetics based on 14 loci mined from GenBank highlighted the radiation among orders in the Nemaliophycidae as one of five unresolved regions (Verbruggen et al., 2010). Simulations suggested that the lack of support in this region was probably due to the lack of informative data (Verbruggen et al., 2010) and added nothing new to our understanding of relationships among the constituent orders. In adding the new order Entwisleiales to the Nemaliophycidae, Scott et al. (2013) provided the best phylogenetic resolution among orders in this subclass to date, but even their analyses only served to strengthen some of the previous results while providing an indication that this new order was sister to the Colaconematales. In short, previous studies have fallen short at resolving interordinal relationships in this diverse subclass of red algae.

The goal of this study was to bring more taxa and more gene regions to the conundrum of inferring relationships among orders of Nemaliophycidae. We present a nine-gene data set representing the plastid, nuclear and mitochondrial genomes of 67 specimens representing all of the constituent orders in an effort to resolve a molecular phylogeny. We use several data exploration methods including fast site removal and simulations studies to determine if a well-resolved phylogeny among the orders can be achieved.

2. Materials and methods

2.1. Taxon sampling

A total of 67 specimens representing each of the 10 Nemaliophycidae orders were selected for study (Table S1). The species richness of each order (based on the current taxonomy) was used to determine our final taxon sampling (Table 1). All genera were sampled for the Balbianiales and Entwisleiales. The freshwater order Thoreales has two species-rich genera, so multiple species were sequenced per genus to capture the diversity. When possible, all genes were sequenced from the same specimen. Voucher information is given in Table S1.

2.2. DNA extraction, amplification and sequencing

DNA was extracted from field-collected samples or culture strains (Table S1) using various standard methods. Extraction methods for Acrochaetiales, Balliales, Colaconematales, Entwisleiales, Nemaliales, Palmariales and Rhodachlyales were reviewed in Saunders and McDevit (2012a,b). For the Batrachospermales, Balbianiales and Thoreales the NucleoSpin® Plant II (Macherey-Nagel, Düren, Germany) kit was used according to the manufacturer's protocol. For *psa*A, *psb*A, *psa*B, EF2, *cox*1, *cob*, 18S rDNA, and 28S rDNA genes, the primers and amplification conditions outlined in Saunders and Moore (2013) were followed with a few exceptions. For the EF2 gene, the primer set (BatEF2F 5′ CTCGTATCATCGAGAC GGCGAATGT 3′ and BatEF2R 5′ GGAAGATCMGCYGGRTTYTTCGGCT 3′) was utilized for amplification of the batrachospermalean taxa and (ThorEF2F 5′ CAAGAATTATAGAATCTGCTAATGT 3′ and ThorEF2R 5′ GGAAGATCKGCRGGRTTYTTCGGCT 3′) for *Thorea hispida*. A few of the taxa in the Batrachospermales did not amplify using the standard *cob* primers and the following primer set was designed (COBdwlF 5′ AGCAYRTWATGMGVGAYGTDAAYTT 3′ and COBdwlR 5′ CWATWACHCCHCCYAATTTRTGWGG 3′). For marine taxa, the amplification of the *rbc*L gene followed Saunders and Moore (2013), for Balbianiales and Batrachospermales it followed (Vis et al., 1998), and for Thoreales it followed (Johnston, 2012). For freshwater specimens, PCR products were prepared for sequencing using UltraClean™ PCR Clean-up DNA purification kit (Mo Bio, Carlsbad, CA, USA) according to manufacturer's protocols. If multiple PCR bands were obtained, the product with the correct length was gel purified using GelElute extraction kit (5 Prime, Gaithersburg, MD). Sequence data for the freshwater Batrachospermales, Balbianiales and Thoreales were obtained using an ABI 3100 Genetic Analyzer (Applied Biosystems). We sequenced with PCR primers and internal primers to have adequate coverage for sense and antisense strands. Contigs were assembled and edited in Sequencher 4.10.1 (Gene Codes Corp., Ann Arbor, Michigan, USA). For the marine specimens, all PCR products were sent to Genome Quebec for cleaning and sequencing. All new sequence data generated were submitted to GenBank (Table S1).

2.3. Phylogenetic analyses

The nine genes were aligned with MAFFT version 7.058beta (Katoh et al., 2002) each gene under the following options: – localpair – maxiterate 1000. These alignments were checked by eye for errors. Most genes aligned easily without indels, but 18S rDNA and 28S rDNA had indels. There were unalignable regions in the 28S rDNA primarily caused by taxa in the Thoreales. These regions were removed. There was also a 6 bp gap in EF2 due to the outgroup taxa having two extra codons; this gap was left in the alignment. The Nemaliales had unalignable introns in EF2 and

Table 1
Summary of Nemaliophycidae orders, number of specimens per order, gene regions utilized and number of sequences generated.

Order	Number of taxa sampled	*rbc*L	*psa*A	*psb*A	*psa*B	EF2	*cox*1	COB	18S rDNA	28S rDNA
Acrochaetiales	4	4/4	3/4	4/4	3/4	1/4	4/4[a]	3/4	2/4	4/4
Balbianiales	2	2/2	0/2	1/2	0/2	1/2	1/2	0/2	2/2	2/2
Balliales	2	2/2	2/2	1/2	1/2	1/2	2/2[a]	2/2	1/2	2/2
Batrachospermales	18	18/18	15/18	18/18	12/18	18/18	18/18[a]	16/18	4/18	16/18
Colaconematales	3	3/3	3/3	3/3	3/3	3/3	3/3	3/3	3/3	3/3
Entwisleiales	1	1/1	1/1	0	0	1/1	1/1	1/1	1/1	1/1
Nemaliales	18	18/18	13/18	18/18	11/18	13/18	18/18	17/18	5/18	16/18
Palmariales	8	8/8	6/8	8/8	6/8	7/8	7/8[a]	8/8	7/8	8/8
Rhodachylales	2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Thoreales	9	9/9	6/9	9/9	3/9	1/9	9/9[a]	2/9	0/9	7/9
Total	67	67/67 (100%)	51/67 (76%)	48/67 (72%)	41/67 (61%)	48/67 (72%)	65/67 (97%)	54/67 (81%)	27/67 (40%)	61/67 (91%)

[a] Full-length *cox*1 could not be generated for some taxa so for those taxa the COI-5P (~664 bp) used in analyses.

these introns were removed from the analysis. The aligned genes were combined into a single nexus file using Sequence Matrix (Vaidya et al., 2011).

All protein-coding genes were partitioned by codon position and the rDNA genes kept unpartitioned resulting in a total of 23 partitions. This scheme was used as input for the program PartitionFinder version 1.0 (Lanfear et al., 2012). The best scheme according to the greedy algorithm, analyzed with Bayesian Information Criterion (BIC) grouped the 23 partitions into 14 partitions (Table S2). The GTR substitution model + I + G was chosen for all 14 partitions.

The 14 partitions with the GTR + I + G were utilized for Maximum Likelihood (ML) analysis in RAxML HPC2 7.6.6 (Stamatakis et al., 2008) on the CIPRES server. Support for relationships was determined from 1000 rapid bootstrap replicates.

Bayesian phylogenetic inference was conducted with ExaBayes 1.4.2 (Aberer et al., 2014), using the same partitioning strategy and model as above. Two independent MCMC runs of 4 increasingly heated chains were run for 10 million generations, sampling every 500th generation. Runs were started from a parsimony tree and branch lengths were unlinked between partitions. All other settings were default. Stationarity of chains and convergence of the runs was assessed with Tracer v. 1.6 (Rambaut et al., 2013). We also determined a suitable burn-in of 6 million generation based on the stationarity of runs. A consensus topology was computed with consensus from the ExaBayes package, using the greedily refined MR consensus option and a burnin proportion of 0.6.

The program Phycas 2.2.0 (Lewis et al., 2015) was used to evaluate whether the polytomies among orders could be resolved. The analysis used the same parameters as the ML phylogenetic analysis (14 partitions, GTR + I + G). The MCMC analysis had 50,000 cycles and was sampled every 100 (Supplement S1).

2.4. Fast site removal

In order to study the impact of potentially noisy fast sites on early diverging relationships, we also analyzed alignments from which the fastest sites were removed. Starting from the complete concatenated alignment and the RAxML tree (cf. Section 2.3), we calculated sitewise rates of evolution with the HyPhy v.2.2 Site Rates standard analysis (JC69 model, fixed rates; Pond et al., 2005). Subsequently, we removed increasing numbers of fast sites from the original alignment using SiteStripper v.1.01 (Cocquyt et al., 2010; Verbruggen, 2012), retaining between 50% and 95% of the slowest sites in 5% increments.

Each of the resulting alignments was subjected to ML searches with bootstrapping (RAxML v.7.3.5, partitioned following aforementioned PartitionFinder results, GTR + G + I model, 1000 bootstrap replicates). The resulting trees were plotted and the bootstrap values along the backbone (i.e. branches situated below the class level but above the order level) were averaged and compared across site-stripped alignments.

2.5. Data requirement simulations

To estimate the amount of data that would be required to resolve the remaining uncertainties in the early branching patterns of the tree, we used two types of simulation experiments. The rationale is to generate alignments longer than the current alignments and study how bootstrap values along the backbone of the tree changes as more data are added. The two types of simulations are described in detail in Verbruggen et al. (2010), and we will only summarize the procedure used for our study.

The nonparametric simulations resample sites from the original alignment (with replacement) of various lengths, including lengths exceeding that of the alignment. We generated alignments ranging from 10^2 (100) to 10^7 (10 million) sites in 0.5 increments of the exponent (i.e. $10^{2.0}$, $10^{2.5}$, $10^{3.0}$, etc.). Each of the resulting alignments was analyzed with RAxML (GTR + G + I model, 100 bootstrap replicates, no partitioning). A majority rule consensus tree (all compatible branches) was generated from the 100 bootstrap trees and annotated with the bootstrap values using phyutility 2.2.1 (Smith and Dunn, 2008).

The parametric simulations use the estimated model parameters and the estimated tree to generate artificial datasets. First, we estimated a tree from the concatenated dataset (RAxML, cf. Section 2.3). Second, we optimized a GTR + G + I model (without partitioning) on that tree with RAxML. With that tree and the inferred model parameters, we simulated alignments of $10^{2.0}, 10^{2.5}, \ldots, 10^{7.0}$ nt in length with seq-gen (Rambaut and Grassly, 1997). The resulting alignments were analyzed as described for nonparametric simulation.

3. Results

3.1. Sequences and alignments

Following trimming and concatenation, the final alignment used for analyses comprised of 13,544 characters (*cob*: 941; *cox*1: 1,233; EF2: 1,709; LSU: 2,727; SSU: 1,782; *psa*A: 1,567; *psa*B: 1,261; *psb*A: 961; *rbc*L 1,363). PartitionFinder suggested a partitioning strategy consisting of 14 partitions that was used for phylogenetic inference (Table S2).

3.2. Phylogenetic analyses

The primary RAxML analysis resulted in a tree (Fig. 1) with excellent support for the order-level groups (100% bootstrap for

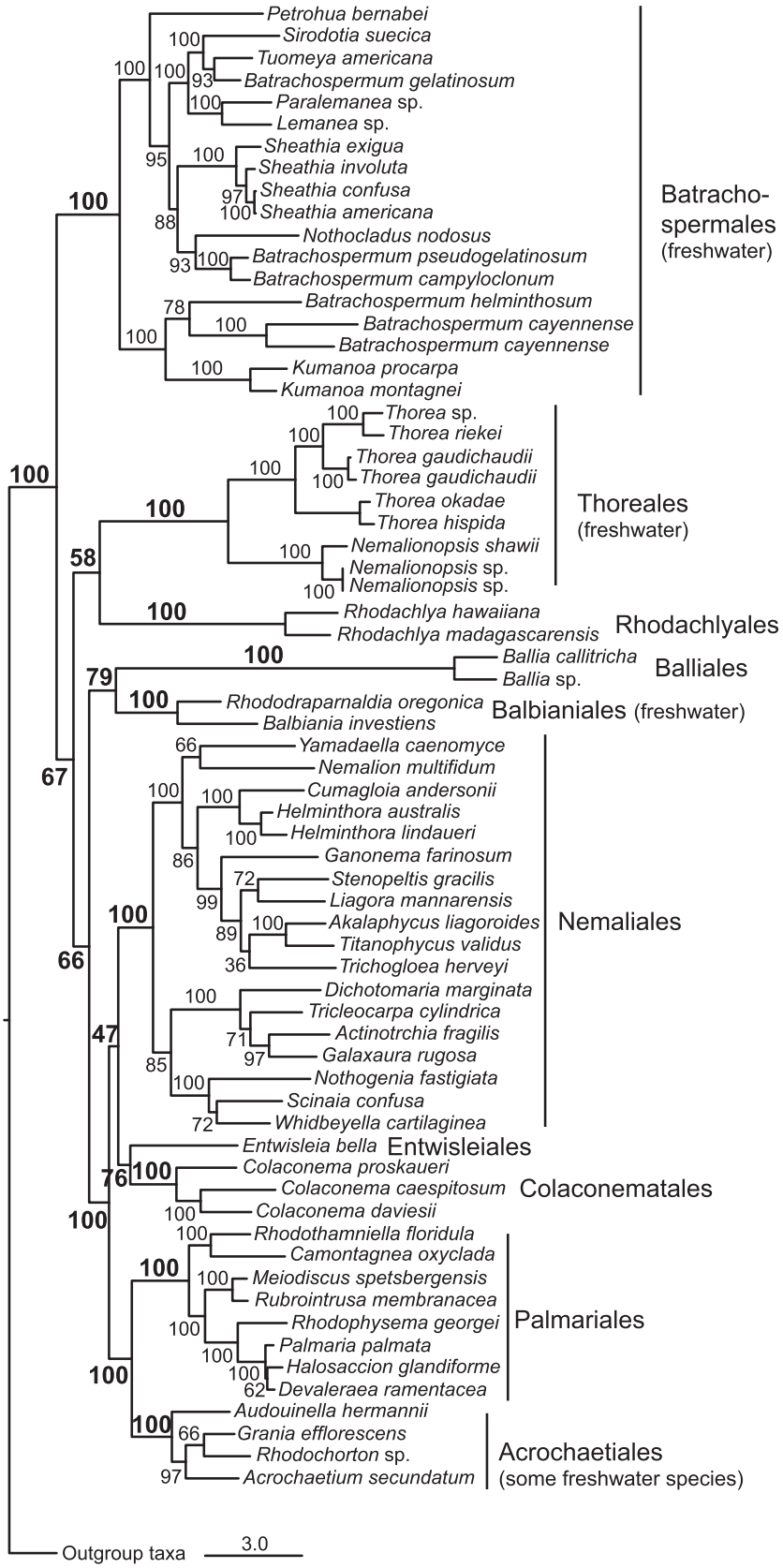

Fig. 1. ML tree (ln −209655.052203) with outgroups pruned for presentation. Numbers on branches are bootstrap support values. Values in bold are those for order monophyly and relationships among orders. Orders that are strictly freshwater or have a few freshwater species are noted; all other orders are marine.

all orders) and mixed support for relationships between orders. Acrochaetiales and Palmariales were resolved as sister lineages with 100% bootstrap support, and the group consisting of Acrochaetiales, Colaconematales, Entwisleiales, Nemaliales, and Palmariales and also received 100%, but the remaining higher-level relationships had support values below 80%. Balbianiales + Balliales were the only other sister relationship that, with a bootstrap support value of 79%, could be considered reasonably supported, dropping to 76% for an association of Colaconematales + Entwisleiales. Relationships within orders were generally well resolved, with the great majority of nodes receiving >90% bootstrap support.

Bayesian posterior probabilities for the relationships between orders were relatively high, with only two branches of the consensus tree receiving less than 0.95 BPP (Fig. S1). The relationship between Colaconematales and Entwisleiales received 0.98 BPP in this analysis. A noteworthy difference between the Bayesian consensus topology and the ML topology is that the Colaconematales + Entwisleiales were a strongly supported (0.99 BPP) sister group of the Acrochaetiales + Palmariales. In the ML topology, the Colaconematales + Entwisleiales clustered with Nemaliales albeit with poor support (47% BS) (Figs. 1 and S1).

The Phycas analysis returned a tree with most nodes resolved (Fig. S2). This result suggested that a tree with dichotomies is more likely than a tree with polytomies among the orders. Therefore, this analysis provides more evidence that the relationships among the orders is not a 'hard' polytomy, but may be resolved.

3.3. Site stripping

To evaluate whether removing noisy sites would improve the resolution of deeper relationships in the tree, i.e. those among orders, we gradually removed the fastest-evolving sites from the alignment and repeated the RAxML analysis. While some nodes did gain stronger support when noisy sites were removed, other nodes remained unchanged or decreased slightly (Table 2, Supplement S2).

The sister relationship between Colaconematales + Entwisleiales exceeded 80% bootstrap support when only the 70–80% slowest sites were analyzed and reached a peak of 90% bootstrap support when the phylogeny was inferred from the 75% slowest sites. A similar situation was observed for the first dichotomy in the tree, separating the Batrachospermales from the remaining nine orders. This relationship exceeded 80% bootstrap support when the 70–75% slowest sites were analyzed. The association between Acrochaetiales + Palmariales remained strong irrespective of site removal, and the same is true of the association among the five orders Acrochaetiales, Colaconematales, Entwisleiales, Nemaliales and Palmariales. The sister relationship between Balbianiales + Balliales remained doubtful with support values between 70% and 80% for most conditions. The sister relationship between Rhodachlyales and Thoreales was poorly supported (<50% bootstrap) except in the condition with the 70% slowest sites retained, where it reached 88% bootstrap support.

3.4. Simulations

Simulations were carried out to estimate how much data would be required to resolve the early radiation of the Nemaliophycidae to a pre-determined level of confidence. These results are illustrated in Fig. 2. In the left panel, the *y*-axis represents the average bootstrap value across all early nodes (i.e. relationships among orders). If we set 80% as a target for the average bootstrap support across these nodes, parametric simulations (blue[1] line) indicate that this target is reached even with very short alignments. For non-parametric simulations (orange line), it is estimated that an alignment of 145,000 nucleotides will be needed to reach this target. In the right panel, the *y*-axis represents the proportion of early nodes exceeding 80% in bootstrap support. If we use a target of 75% (i.e. we wish to resolve three quarters of the early nodes with bootstrap support >80%), non-parametric simulations indicate that alignments exceeding 3.1 million nucleotides will be needed.

[1] For interpretation of color in Fig. 2, the reader is referred to the web version of this article.

4. Discussion

Verbruggen et al. (2010) identified the early radiation of nemaliophycidaean orders as one of the five big remaining questions in the phylogenetic history of red algae. Based on simulations similar to those we have carried out here, their study estimated how much data would be needed to resolve the radiation in question. While parametric simulations indicated that alignments of 1950 nucleotides would be sufficient, non-parametric simulations never reached the target of 80% average bootstrap support, not even with alignments of up to a million nucleotides (Verbruggen et al., 2010). It was argued that while parametric simulations were too optimistic, non-parametric simulations were probably too pessimistic and the true amount of data needed would probably lie in between the estimates of these two types of simulations.

With an alignment of 13,544 nucleotides containing all orders and families brought to bear on the question, our study is certainly the most comprehensive phylogenetic analysis of the Nemaliophycidae. While some relationships between orders are strongly supported, the radiation as a whole is not resolved beyond question. Fast site removal suggests some additional relationships Colaconematales + Entwisleiales and placement of Batrachospermales as sister to the remaining orders as also variously noted in previous studies (Le Gall and Saunders, 2007; Scott et al., 2013), but most remained poorly supported.

Our simulation experiments shed more light on how much data may be needed to resolve the radiation more satisfactorily. As in Verbruggen et al. (2010), parametric simulations return complete resolution of the radiation for datasets in the order of magnitude of one to a few thousand bases. This is obviously not realistic as our dataset of >13,000 sites did not resolve the radiation. The reason parametric simulations underestimate the amount of data needed is that they fail to capture the complex processes of sequence evolution in real populations.

The non-parametric simulations, which resample sites from our "real" dataset beyond its original size, may give a more realistic picture of the amount of data needed to achieve resolution. It suggests that alignments of approximately 145,000 nucleotides are needed to resolve the phylogeny with an average of 80% bootstrap support across early-branching lineages (i.e. among order relationships). This would be achievable with high-throughput sequencing of organellar genomes or transcriptomes. Red algal chloroplast genomes are approximately 180k and mitochondrial genomes ca. 25k bases, while preliminary transcriptome data have reached ~2000k bases (Saunders and Jackson, unpublished data).

While this is a promising outlook, our non-parametric simulations also indicate that some nodes may never be resolved. The proportion of nodes exceeding 80% bootstrap support never reaches 100% even though our simulated datasets reached 10 million nucleotides in length. In this context it is important to note that the non-parametric simulation resamples the present dataset and no new data are introduced, so its predictions can be expected to be on the pessimistic side.

The Phycas and ExaBayes analyses results also lend support to the resolution of relationships among the orders of the Nemaliophycidae. The Phycas analysis showed that a dichotomous tree including the nodes among the orders was favored over a tree with

Table 2
Bootstrap support for selected clades as a function of the percentage of slow-evolving retained in the analysis. Support values equal to or exceeding 80% are indicated in boldface.

Sites retained (%)	Balliales + Balbianiales	Rhodachlyales + Thoreales	Colaconematales + Entwisleiales	Acrochaetiales + Palmariales	All except Batrachospermales	5 orders
65	**85**	54	49	**96**	56	**87**
70	53	**88**	**80**	**99**	**83**	**90**
75	72	47	**90**	**100**	**81**	**95**
80	77	29	**82**	**100**	28	**98**
85	77	23	42	**100**	64	**100**
90	69	22	48	**100**	46	**99**
95	77	39	53	**100**	58	**99**
100	**82**	34	56	**100**	53	**99**

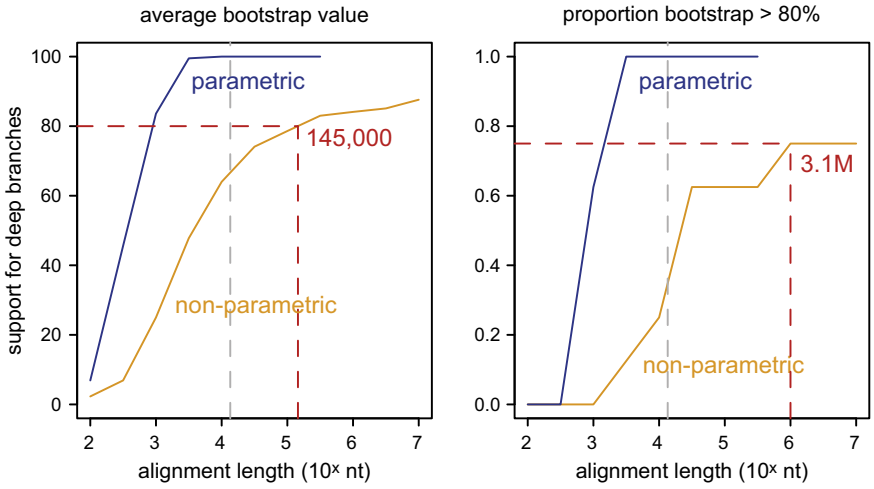

Fig. 2. Simulations of how much data may be required to resolve the radiation of nemaliophycidaean orders. The left panel shows the evolution of average bootstrap support across all nodes above the order level as the simulated alignment length is increased. The right panel shows the evolution of the proportion of nodes exceeding 80% bootstrap as the simulated alignment length is increased. Note the logarithmic scale on the *x*-axis. The gray dashed line indicates the alignment length used in this paper. The red dashed lines shows how much data are estimated to be needed for resolution of the radiation according to the non-parametric simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

polytomies. Likewise, the posterior support in the Bayesian analysis was high. However, this result should be tempered, as posterior probabilities tend to overestimate support. In addition, this analysis retrieved an alternative relationship with the Colaconematales + Entwisleiales being more closely related to the Acrochaetiales + Palmariales rather than allied with the Nemaliales as in our ML topology and resolved with varying support in Scott et al. (2013). These alternative relationships will need more data to be fully resolved.

5. Conclusions

The Nemaliophycidae, as a whole, is an exciting clade to study many aspects of red algal biology including the transition from marine to freshwater, unique life histories and unusual associations with other organisms. Systematists have realized the value of this subclass and have studied the phylogeny using all tools available. The present study generated a large amount of new sequence data from all orders and from the nuclear, plastid and mitochondrial genomes in order to produce a well-supported phylogeny. Even with this substantial sequence set, there was little resolution of the relationships among orders for this important subclass. The data exploration methods employed using the new data set from this study, suggest that a resolved phylogeny is tractable, but phylogenomics methodologies will need to be applied.

Acknowledgments

We wish to acknowledge all collectors of specimens listed in Table S1. Emily Johnston and Eric Salomaki are thanked for aid in sequencing some of the freshwater red taxa. Tanya Moore and other Saunders' Lab partners are especially thanked for generating the molecular data for the marine species in this study. For the Phycas analysis, help in optimizing from Paul Lewis is gratefully acknowledged. HV is a Future Fellow of the Australian Research Council (FT110100585). Funding for this study was provided by NSF DEB 0936855, 0936884 (RedToL). In addition to RedToL funding, Saunders was supported by Discovery and Accelerator Grants from the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chair Program, the Canada Foundation for Innovation and the New Brunswick Innovation Foundation.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.10.015.

References

Aberer, A.J., Kassian, K., Stamatakis, A., 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31, 2553–2556.
Blouin, N.A., Lane, C.E., 2012. Red algal parasites: models for a life history evolution that leaves photosynthesis behind again and again. BioEssays 34, 226–235.
Butterfield, N.J., 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology 26, 386–404.
Clayden, S.L., Saunders, G.W., 2010. Recognition of *Rubrointrusa membranacea* gen. et comb. nov., *Rhodonematella subimmersa* gen. et comb. nov (with a reinterpretation of the life history) and the Meiodiscaceae fam. nov within the Palmariales (Rhodophyta). Phycologia 49, 283–300.

Clayden, S.L., Saunders, G.W., 2014. A study of two *Acrochaetium* complexes in Canada with distinction of *Rhododrewia* gen. nov. (Acrochaetiales, Rhodophyta). Phycologia 53, 221–232.
Cocquyt, E., Verbruggen, H., Leliaert, F., De Clerck, O., 2010. Evolution and cytological diversification of the green seaweeds (Ulvophyceae). Mol. Biol. Evol. 27, 2052–2061.
DeCew, T.C., West, J.A., 1982. A sexual life history in *Rhodophysema* (Rhodophyceae): a re-interpretation. Phycologia 21, 67–74.
Delwiche, C.F., 1999. Tracing the thread of plastid diversity through the tapestry of life. Am. Nat. 154 (Suppl.), S164–S177.
Gantt, E., 1990. Pigmentation and photoacclimation. In: Cole, K.M., Sheath, R.G. (Eds.), Biology of the Red Algae. Cambridge University Press, Cambridge, pp. 203–219.
Guiry, M.D., Guiry, G.M., 2015. AlgaeBase. World-wide Electronic Publication, National University of Ireland, Galway. <http://www.algaebase.org>.
Harper, J.T., Saunders, G.W., 2001. Molecular systematics of the Florideophyceae (Rhodophyta) using nuclear large and small subunit rDNA sequence data. J. Phycol. 37, 1073–1082.
Harper, J.T., Saunders, G.W., 2002. A re-classification of the Acrochaetiales based on molecular and morphological data, and establishment of the Colaconematales ord. nov. (Florideophyceae, Rhodophyta). Eur. J. Phycol. 37, 463–476.
Huisman, J.M., Harper, J.T., Saunders, G.W., 2004. Phylogenetic study of the Nemaliales (Rhodophyta) based on large-subunit ribosomal DNA sequences supports segregation of the Scinaiaceae fam. nov. and resurrection of *Dichotomaria* Lamarck. Phycol. Res. 52, 224–234.
Johnston, E.J., 2012. The Systematic Revision of the Freshwater Red Algal Order Thoreales (Nemaliophycidae, Rhodophyta). Ohio University, p. 84.
Kain, J.M., Norton, T.A., 1990. Marine ecology. In: Cole, K.M., Sheath, R.G. (Eds.), Biology of the Red Algae. Cambridge University Press, New York, pp. 377–423.
Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.
Kumano, S., 2002. Freshwater Red Algae of the World. Biopress Ltd, Bristol.
Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Le Gall, L., Saunders, G.W., 2007. A nuclear phylogeny of the Florideophyceae (Rhodophyta) inferred from combined EF2, small subunit and large subunit ribosomal DNA: establishing the new red algal subclass Corallinophycidae. Mol. Phylogenet. Evol. 43, 1118–1130.
Lewis, P.O., Holder, M.T., Swofford, D.L., 2015. Phycas: software for bayesian phylogenetic analysis. Syst. Biol. http://dx.doi.org/10.1093/sysbio/syu132.
Pond, S.L.K., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679.
Pueschel, C.M., 1990. Cell structure. In: Cole, K.M., Sheath, R.G. (Eds.), Biology of the Red Algae. Cambridge University Press, Cambridge, pp. 7–42.
Rambaut, A., Grassly, N.C., 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic frees. Comput. Appl. Biosci. 13, 235–238.
Rambaut, A., Suchard, M., Drummond, A., 2013. Tracer v1.6 <http://tree.bio.ed.ac.uk/software/tracer/>.
Saunders, G.W., Bailey, J.C., 1997. Phylogenesis of pit-plug associated features in the Rhodophyta: inferences from molecular systematic data. Can. J. Bot. 75, 1436–1447.
Saunders, G.W., Bailey, J.C., 1999. Molecular systematic analyses indicate that the enigmatic *Apophlaea* is a member of the Hildenbrandiales (Rhodophyta, Florideophycidae). J. Phycol. 35, 171–175.
Saunders, G.W., Bird, C.J., 1989. Considerations of life history, morphology and taxonomy in *Rhodophysema georgii* Batters (Rhodophyta, Palmariales). Brit. Phycol. J. 24, 63–71.
Saunders, G.W., Hommersand, M.H., 2004. Assessing red algal supraordinal diversity and taxonomy in the context of contemporary systematic data. Am. J. Bot. 91, 1494–1507.
Saunders, G.W., Kraft, G.T., 1997. A molecular perspective on red algal evolution: focus on the Florideophycidae. Plant Syst. Evol. 11 (Suppl.), 115–138.
Saunders, G.W., Maggs, C.A., McLachlan, J.L., 1989. Life-history variation in Rhodophysema elegans (Palmariales, Rhodophyta) from the North Atlantic and crustose Rhodophysema spp. from the North Pacific. Can. J. Bot. 67, 2857–2872.
Saunders, G.W., McDevit, D.C., 2012a. Methods for DNA barcoding photosynthetic protists emphasizing the macroalgae and diatoms. Methods Mol. Biol. 858, 207–222.
Saunders, G.W., McDevit, D.C., 2012b. Acquiring DNA sequence data from dried archival red algae (Florideophyceae) for the purpose of applying available names to contemporary genetic species: a critical assessment. Botany-Botanique 90, 191–203.
Saunders, G.W., Moore, T.E., 2013. Refinements for the amplification and sequencing of red algal DNA barcode and Red TOL phylogenetic markers: a summary of current primers, profiles and strategies. Algae 28, 31–43.
Saunders, G.W., Bird, C.J., Ragan, M.A., Rice, E.L., 1995. Phylogenetic relationships of species of uncertain taxonomic position within the Acrochaetiales/Palmariales complex (Rhodophyta): inferences from phenotypic and 18S rDNA sequence data. J. Phycol. 31, 601–611.
Scott, F.J., Saunders, G.W., Kraft, G.T., 2013. *Entwisleia bella*, gen. et sp. nov., a novel marine 'batrachospermaceous' red alga from southeastern Tasmania representing a new family and order in the Nemaliophycidae. Eur. J. Phycol. 48, 398–410.
Smith, S.A., Dunn, C.W., 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics 24, 715–716.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771.
Stegenga, H., 1978. Life histories of *Rhodochorton purpureum* and *Rhodochorton floridulum* (Rhodophyta, Nemaliales). Brit. Phycol. J. 13, 279–289.
Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180.
Van der Meer, J.P., Todd, E.R., 1980. The life history of *Palmaria palmata* in culture. A new type for the Rhodophyta. Can. J. Bot. 58, 1250–1256.
Verbruggen, H., 2012. SiteStripper. <http://phycoweb.net/software/SiteStripper>.
Verbruggen, H., Maggs, C.A., Saunders, G.W., Le Gall, L., Yoon, H.S., De Clerck, O., 2010. Data mining approach identifies research priorities and data requirements for resolving the red algal tree of life. BMC Evol. Biol. 10, 16.
Vis, M.L., Saunders, G.W., Sheath, R.G., Dunse, K., Entwisle, T.J., 1998. Phylogeny of the Batrachospermales (Rhodophyta) inferred from *rbc*L and 18S ribosomal DNA gene sequences. J. Phycol. 34, 341–350.
Woelkerling, W., 1990. An introduction. In: Cole, K.M., Sheath, R.G. (Eds.), Biology of the Red Algae. Cambridge University Press, Cambridge, pp. 1–6.
Womersley, H.B.S., 1994. The Marine Benthic Flora of Southern Australia. Rhodophyta—Part IIIA. Flora of Australia Supplementary Series Number 1. Australia Biological Resources Study, Canberra, Australia.