J. Phycol. *, ***–*** (2021)

DOI: 10.1111/jpy.13168-20-145

PHYLOGENOMIC ANALYSIS RESTRUCTURES THE ULVOPHYCEAE[1]

Øyvind Sætren Gulbrandsen

Section for Genetics and Evolutionary Biology (EVOGENE), Department of Biosciences, University of Oslo, Kristine Bonnevies Hus, Blindernveien 31, 0316 Oslo, Norway

Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Centre for Integrative Genetics, Norwegian University of Life Sciences, Ås, Norway

Ina Jungersen Andresen, *Anders Kristian Krabberød*

Centre for Integrative Microbial Evolution (CIME), Section for Genetics and Evolutionary Biology (EVOGENE), Department of Biosciences, University of Oslo, Kristine Bonnevies Hus, Blindernveien 31, 0316 Oslo, Norway

Jon Bråte[2]

Section for Genetics and Evolutionary Biology (EVOGENE), Department of Biosciences, University of Oslo, Kristine Bonnevies Hus, Blindernveien 31, 0316 Oslo, Norway

Department of Virology, Norwegian Institute of Public Health, Oslo, Norway

and Kamran Shalchian-Tabrizi[2]

Centre for Integrative Microbial Evolution (CIME), Centre for Epigenetics, Development and Evolution (CEDE), Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Kristine Bonnevies Hus, Blindernveien 31, 0316 Oslo, Norway

Here, we present new transcriptome sequencing data from seven species of Dasycladales (Ulvophyceae) and a phylogenomic analysis of the Chlorophyta with a particular focus on Ulvophyceae. We have focused on a broad selection of green algal groups and carefully selected genes suitable for reconstructing deep eukaryote evolutionary histories. Increasing the taxon sampling of Dasycladales restructures the Ulvophyceae by identifying Dasycladales as closely related to Scotinosphaerales and Oltmannsiellopsidales. Contrary to previous studies, we do not find support for a close relationship between Dasycladales and a group with Cladophorales and Trentepohliales. Instead, the latter group is sister to the remainder of the Ulvophyceae. Furthermore, our analyses show high and consistent statistical support for a sister relationship between Bryopsidales and Chlorophyceae in trees generated with both homogeneous and heterogeneous (heterotachy) evolutionary models. Our study provides a new framework for interpreting the evolutionary history of Ulvophyceae and the evolution of cellular morphologies.

Key index words: **Acetabularia; Bryopsidales; Chlorophyta; cytomorphology; Dasycladales; evolution; green algae; phylogeny; Scotinosphaerales; transcriptomics; Ulvophyceae**

***Abbreviations*: Burki250, 250 gene alignment from Burki et al. (2016); ML, maximum likelihood; MLBS, multilocus bootstrap support; MMETSP, Marine Microbial Eukaryote Transcriptome Sequencing Project; NPP, Abbreviation for the genera *Nephroselmis*, *Pycnococcus*, and *Picocystis*; UTC, Clade consisting of the classes Ulvophyceae, Trebouxiophyceae and Chlorophyceae.**

Dasycladales is a group of single-celled green algae with an astonishing adaptation to a macroscopic lifestyle. Species of Dasycladales consist of only a single cell, yet they can grow over ten centimeters in size and display highly elaborate morphological structures (Berger 2006). Dasycladales is part of the class Ulvophyceae which belongs to the Chlorophyta. The phylogeny of Chlorophyta has historically been subjected to significant changes and uncertainties (Fang et al. 2017). In the most recent revision of eukaryote systematics (Adl et al. 2019), Chlorophyta is divided into eleven classes. Among these, Ulvophyceae, Trebouxiophyceae, Chlorophyceae, Pedinophyceae, and Chlorodendrophyceae make up a monophyletic assemblage coined "Core Chlorophyta" (Leliaert et al. 2016, Fang et al. 2017, Del Cortona et al. 2020). Within the Core Chlorophyta, a large group composed of Ulvophyceae,

[1]Received 19 June 2020. Revised 26 January 2021. Accepted 3 February 2021.
[2]Author for correspondence: e-mails jon.brate@ibv.uio.no (J.B.); kamran@ibv.uio.no (K.S.).
Editorial Responsibility: H. Verbruggen (Associate Editor)

Trebouxiophyceae, and Chlorophyceae (the UTC clade) is especially well studied due to its large species diversity and ecological importance.

Ulvophyceae hosts a wide array of different cellular morphologies (cytomorphologies) and is therefore a unique model system for studying morphological and cytomorphological diversification (Verbruggen et al. 2009). Among the Ulvophyceae, cytomorphologies range from small unicellular species, such as Scotinosphaerales, to large multicellular species, such as the sea lettuce *Ulva*, to the gigantic single-celled species (up to meters in size) of Bryopsidales and Dasycladales. However, the evolutionary relationships between the different groups of Ulvophyceae are uncertain, and even the monophyly of the entire class has recently been contested (Leebens-Mack et al. 2019, Del Cortona et al. 2020).

Notably, the position of Dasycladales and their most closely related lineages within Ulvophyceae has never been fully resolved. Dasycladales have traditionally been regarded as closely related to Bryopsidales because of shared cytomorphological traits such as extremely large, single-celled siphonous species (i.e., macroscopic, single-celled organisms with cytoplasmic streaming and sometimes multiple nuclei; e.g., Verbruggen et al. 2009, Coneva and Chitwood 2015). A sister relationship between these two groups would imply that the single-celled macroscopic cytomorphology has only originated once within Ulvophyceae (Cocquyt et al. 2010b). However, several chloroplast and nuclear gene phylogenies over the past decade have shown discrepant placements of Bryopsidales, questioning their sister relationship to Dasycladales and their position within the UTC (Fučíková et al. 2014, Leliaert and Lopez-Bautista 2015, Leliaert et al. 2016, Turmel et al. 2017). A deeper position within UTC, as seen recently (Leebens-Mack et al. 2019, Del Cortona et al. 2020), would instead suggest independent origins of the macroscopic cytomorphologies of Dasycladales and Bryopsidales.

Furthermore, several ulvophycean lineages are unstable in molecular trees, such as the Ignatiales, Oltmannsiellopsidales, and the Cladophorales/Trentepohliales clade (Fučíková et al. 2014, Del Cortona et al. 2020). As a consequence, it is difficult to infer the ancestral cytomorphology of Ulvophyceae and the cytomorphological origins of Dasycladales. Resolving the uncertainties regarding the exact position of Dasycladales and its sister groups among Ulvophyceae, as well as the position of Bryopsidales in relation to Ulvophyceae, is a key to understand the evolution of large, siphonous cell forms and the different adaptations to a macroscopic lifestyle among green algae.

One reason for the instabilities in the molecular phylogenies of Ulvophyceae is likely the undersampling of Dasycladales, which are usually only represented by a single species if included at all. Additionally, extant Dasycladales are divided into two families, Dasycladaceae and Polyphysaceae, with species of very different cellular morphologies (Berger 2006), and it is therefore important to increase the sampling of taxa across Dasycladales and include representatives of both families. Another reason is the long and independent evolutionary histories of the different lineages within Ulvophyceae, exhibiting a pattern of rapid and ancient radiations (Del Cortona et al. 2020), possibly also followed by several extinction events (Berger et al. 2003, Fang et al. 2017). For these reasons, most ulvophycean groups have very long branches in molecular phylogenies, making them prone to long-branch attraction artifacts.

To shed light on the evolution of macroscopic cytomorphologies in Ulvophyceae, the main aim of this study was to resolve the phylogeny of the whole group, with a particular focus on the position of Dasycladales and Bryopsidales. To address this, we have improved the representation of Dasycladales by sequencing the transcriptomes of seven species from five different genera covering both the families Polyphysaceae and Dasycladaceae. And to overcome the challenges posed by the long, independent evolutionary histories of the chlorophyte classes, we have elected for a conservative approach where we selected only slowly evolving single-copy eukaryote genes, removed fast-evolving sites, and minimized the amount of missing data.

METHODS

Dasycladales cultures. Seven species of Dasycladales (Table 1) were kept in culture at the Institute of Biosciences, University of Oslo, in Dasycladales seawater medium prepared after the recipe of UTEX Culture Collection of Algae. Cultures were kept in incubators with a 12:12 h light:dark cycle at a temperature of 20°C, and a light intensity of 2,500 lux.

Table 1. Species of Dasycladales investigated in this study.

Species	UTEXID	Family	Order
Acetabularia acetabulum	—	Polyphysaceae	Dasycladales
Acetabularia crenulata	—	Polyphysaceae	Dasycladales
Acetabularia peniculus	—	Polyphysaceae	Dasycladales
Parvocaulis polyphysoides[a]	LB 2707	Polyphysaceae	Dasycladales
Bornetella oligospora	LB 2689	Dasycladaceae	Dasycladales
Chloroclados australasicus	LB 2686	Dasycladaceae	Dasycladales
Neomeris dumetosa	LB 2691	Dasycladaceae	Dasycladales

Except for the *Acetabularia* species (identified and provided by Prof. W. Martin of the University of Düsseldorf), the species names and identification were taken from the UTEX culture collection.

[a]*Parvocaulis polyphysoides* was originally listed by UTEX as *Polyphysa polyphysoides*, but we have chosen to use *Parvocaulis polyphysoides* as this is listed as the currently accepted name in AlgaeBase (Guiry et al. 2014).

RNA isolation and sequence library preparation. Tissue from an entire culture bottle from each of the seven cultured species was extracted with a sterile transfer pipette and flash-frozen in liquid nitrogen. The culture of *Chloroclados australasicus* contained microscopic cells in addition to the typical macroscopic cells of the adult, vegetative stage. These possible contaminants were isolated by removing the macroscopic cells from the growth medium and sent, along with the other seven samples, to Vertis Biotechnologie AG (Freising, Germany) for RNA isolation and preparation of Illumina sequencing libraries.

Briefly, the library preparations were performed as follows: Total RNA was isolated using the peqGOLD TriFast kit (VWR, Monroeville, PA, USA), including a DNase treatment step, and the RNA integrity was confirmed using capillary gel electrophoresis. Polyadenylated RNA was extracted from the total RNA, followed by ultrasound fragmentation and ligation of a 3′ oligonucleotide adapter. Using the 3′ adapter as a primer, first-strand cDNA was synthesized using M-MLV reverse transcriptase and purified using the Agencourt AMPure XP kit (Beckman Coulter, Brea, CA, US). A 5′ Illumina TruSeq sequencing adapter was ligated to the 3′ end of the antisense cDNA and the resulting cDNA was PCR-amplified to about 9–14 $ng \cdot \mu L^{-1}$ per library. The eight sequencing libraries were sent to the Norwegian Sequencing Centre (www.sequencing.uio.no) for 150 bp insert paired-end sequencing on the Illumina HiSeq 4000 system (Illumina, Inc., San Diego, CA, USA).

Sequence processing and transcriptome assembly. In addition to the seven cultivated species and the possible contaminant from the *C. australasicus* culture, 35 paired-end sequence sets (Illumina HiSeq 2000, Illumina, Inc.) originating from 35 species of Chlorophyta (Table S1 in the Supporting Information) were downloaded from the EMBL-EBI nucleotide archive and processed alongside our newly generated sequence sets for a total of 42 paired-end sequence libraries. The data for *Acetabularia acetabulum* were merged with Illumina reads from an mRNA sequencing of adult cells performed in (I. J. Andresen, R. J. S. Orr, K. Shalchian-Tabrizi, & J. Bråte, unpub. data). Read qualities of the 42 paired-end Illumina data sets were evaluated using FastQC v.0.11.2 (Andrews 2010). Quality trimming was done with Trimmomatic v.0.36 (Bolger et al. 2014) with cutoffs set to $Q = 20$ for trailing and leading nucleotides, and a $Q = 20$ mean score over a 4-nucleotide interval. Trimmomatic was also used to remove sequencing adapters and sequences shorter than 36 bp.

De novo transcriptome assembly was performed using Trinity v.2.0.6 (Grabherr et al. 2011, Haas et al. 2013) using both paired and unpaired sequences and read-normalization, with otherwise default parameters. Only the longest isoforms of each predicted transcript were retained. Single and paired reads from the seven cultured species were also assembled using rnaSPAdes v.3.11.1 (Bankevich et al. 2012) with default parameters. Trinity and rnaSPades assemblies were analyzed together to maximize gene representation, following the procedure described in the section "Ortholog selection".

In addition to the cultured species, previously assembled transcriptomes from 21 different species were downloaded from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Keeling et al. 2014), as well as an assembled *Caulerpa taxifolia* transcriptome from Ranjan et al. (2015) and a *Scotinosphaera lemnae* transcriptome from Del Cortona et al. (2020; Table S2 in the Supporting Information). Species represented by multiple strains or accessions (see Table S2) were concatenated into a single transcriptome assembly. Altogether, a total of 65 transcriptomes were passed to the next step.

ORF prediction. Translation of the assembled transcripts and prediction of open reading frames (ORFs) was done using TransDecoder v.5.0.2 (Haas et al. 2013). The *Acetabularia* codon table (an option in TransDecoder) was used for species within the orders Dasycladales, Scotinosphaerales (see Appendix S1 in the Supporting Information for verification), Cladophorales, and Trentepohliales, while the standard eukaryote codon table was used for all other assemblies.

Ortholog selection. As a basis for selecting genes suitable for phylogenetic reconstruction, we used the 250 single-gene alignments (with ambiguously aligned sites removed) from Burki et al. (2016; hereafter referred to as Burki250), consisting of genes curated for large-scale analyses of deeply diverging eukaryote lineages. To identify which of these 250 genes that were present in our 65 transcriptomes, we used the BIR v1.0 pipeline (Kumar et al. 2015) on the University of Oslo LifePortal platform (https://lifeportal.uio.no). To ensure correct ortholog identification in our data and avoid the inclusion of hidden paralogs, we added paralogous genes from reference genomes covering all eukaryote supergroups (genes were identified by BLAST, e-value requirement < 10e−40), followed by single-gene phylogenetic reconstruction. Single-gene trees (in addition to the ones produced by BIR) were generated by aligning sequences using MAFFT v.7.309 (Katoh and Standley 2013) with a BLOSUM62 scoring matrix and the L-INS-I algorithm followed by FastTree v.2.15 (Price et al. 2010) using the Whelan and Goldman (WAG) substitution model and an optimized Gamma likelihood model with 20 rate categories. The single-gene trees were carefully examined to identify genes with a reliable phylogenetic signal and to remove assembly artifacts, gene paralogs, and potential contaminants. First, if a single-gene tree did not produce a monophyletic Viridiplantae with the Burki250 sequences, the gene was discarded altogether. Second, sequences from the assembled transcriptomes were removed from the alignments if they did not cluster within a monophyletic Chlorophyta (including the Burki250 sequences). Third, in cases where identical sequences from the same species were present (due to masking of ambiguous sites), the shortest sequence was removed. Fourth, obvious assembly artifacts such as very short sequences or sequences with extremely long branches arising due to long unaligned stretches (e.g., chimeric sequences) were removed. This process was repeated several times to identify additional paralogs or artifacts arising from removing gene sequences.

The single-gene alignments from the Burki250 dataset used to identify genes in our de novo assembled transcriptomes had ambiguously aligned sites removed. Because we included a narrower selection of taxa than Burki et al. (2016), we realigned the original full-length sequences prior to making the final alignments so that we could include more sites. Orthologs that could not be clearly identified using our above-mentioned criteria for ortholog selection were re-evaluated using these untrimmed alignments from Burki250.

Removing potential contaminants. Sequences from the potential contaminant library of the *C. australasicus* culture often clustered with species from Ulvales/Ulotrichales. This sequencing library was therefore discarded from the analyses. Sequences from the *C. australasicus* library were deleted if they clustered with sequences originating from the potential contaminant, or did not cluster with Dasycladales. The sequences from *Trentepohlia jolithus* (downloaded from EBI) also appeared to contain contaminants. However, these were easily identifiable as they clustered separately from all other Chlorophyta or with Embryophyta and were therefore removed.

Phylogenetic reconstruction. The untrimmed sequences were aligned with MAFFT as before, erroneously inserted end-gaps

were removed, and the sequences realigned. Ambiguously aligned residues, inserts containing unknown residues, and end-gaps were manually masked. Finally, the gene alignments were concatenated to a single multigene alignment using Geneious v.11.0.3 (http://www.geneious.com; Kearse et al. 2012). Maximum likelihood (ML) analysis was conducted using IQ-TREE v.1.6.8 (Nguyen et al. 2015), with the ModelFinder (Kalyaanamoorthy et al. 2017) automatic model selection and up to 1,000 Ultrafast Bootstrap Approximate searches (Hoang et al. 2017). The various alignments with fast-evolving sites and missing data removed were analyzed with the same evolutionary model as the original dataset, LG+F+R8, and otherwise the same parameters. IQ-TREE was also run using the recently implemented GHOST model (Crotty et al. 2020) which can model heterotachously evolved sequences and allows separate sets of model parameters and edge lengths on the same tree topology. GHOST was run with the LG+F*H7 model chosen as the most fitting according to the Akaike Information Criterion.

Bayesian phylogenetic inference was conducted using PhyloBayes MPI v.1.8 (Lartillot et al. 2013), with two chains using the CATGTR model with four gamma categories (CATGTR+Γ4). The chains were run until meandiff and maxdiff were less than 0.3 (with a burn-in of 1,000) as estimated by the bpcomp program in PhyloBayes.

A coalescence analysis was performed following a similar procedure as Del Cortona et al. (2020). ASTRAL v5.7.3 (Zhang et al. 2018) was run directly on the 95 single-gene trees produced by IQ-TREE described above without any additional options, in addition to performing multi-locus bootstrapping (MLBS). For the MLBS, 100 replicates were run and each of the 1,000 bootstrap replicate single-gene trees from IQ-TREE were included.

Testing the impact of fast-evolving sites and missing data. Evolutionary rates of sites in the concatenated multigene alignment were estimated by flagging the -wsr option in IQ-TREE v.1.6.8. Sites were removed from the alignment in 5% increments, starting with the fastest evolving sites, using Sitestripper v.1.0.3 (Verbruggen 2018).

Genes were removed from the same alignment according to the percentage of total taxa in which each gene was present (i.e., 100% = gene present in all taxa). Starting with the least representation of taxa, seven alignments were generated by removing genes in 10% increments ranging from 30% to 90% taxa present for each gene (no genes had less than 20% of taxa represented).

Two additional alignments were generated, one with *Ignatius tetrasporus* removed and one with Oltmannsiellopsidales removed, to assess their effects on the branching pattern and support values within Ulvophyceae.

RESULTS

Relationship between classes outside Core Chlorophyta. The final alignment consisted of 78 taxa and 95 genes (35,307 aa positions), covering all major groups of Chlorophyta (Fig. S1 in the Supporting Information). Both Bayesian, Maximum Likelihood (ML), and the coalescence inferences of the main dataset resulted in a robust and overall well-supported topology (Fig. 1). Branching pattern of the major groups were identical.

Prasinococcales (Palmophyllophyceae), Pyramimonadales, and Mamiellophyceae were all retrieved as maximally supported clades (100% ML and MLBS support and 1.00 Bayesian support). Prasinococcales were placed distinct from all other Chlorophyta, followed by Pyramimonadales and Mamiellophyceae which were clustered as sister clades (all with maximum support). The three prasinophyte lineages *Nephroselmis*, *Pycnococcus,* and *Picocystis* (abbreviated NPP) branched as the closest relatives to the core chlorophytes but were not monophyletic (*Nephroselmis* and *Pycnococcus* clustered as sisters, except in the coalescence analysis where the three lineages were divided into separate branches.). Although fully supported in the Bayesian analysis, their interrelationships were only moderately supported in ML. When removing genes present in less than 70% of taxa, NPP was weakly supported (55-68%) as monophyletic, with *Pycnococcus* and *Picocystis* placing as sisters.

Phylogenetic relationships between classes of Core Chlorophyta. Our phylogenetic analyses consistently recovered the core Chlorophyta as a major monophyletic assemblage (maximum support) consisting of Pedinophyceae, Chlorodendrophyceae, Trebouxiophyceae, Chlorophyceae, and Ulvophyceae (Fig. 1). *Pedinomonas tuberculata* (Pedinophyceae) was separated from the rest of the core Chlorophyta in the ML and Bayesian analyses (98% and 99% ML support and 1.00 Bayesian support), with Chlorodendrophyceae as a sister group to the UTC classes (maximum support). However, in the coalescence analysis, *P. tuberculata* was sister to the UTC (results not shown).

Within the UTC group, Trebouxiophyceae were recovered as a fully supported clade, distinct from the other UTC classes. Trebouxiophyceae were divided into two distinct sub-clades, one consisting of Chlorellales and the other containing members of Trebouxiales, Prasiolales, and taxonomically uncertain trebouxiophyceans. *Ettlia oleoabundans*, which is not annotated as Trebouxiophyceae in AlgaeBase (Guiry et al. 2014) but as Chlorophyceae, clustered with full support as sister to *Chlorella vulgaris* within the Chlorellales.

Bryopsidales, an order traditionally classified within the Ulvophyceae, clustered as sister to Chlorophyceae with high support in the ML and Bayesian analyses (94% and 89% ML support and 1.00 Bayesian support), but with only 15% MLBS support in the coalescence analysis (the normalized quartet score was 0.79, showing an overall very high concordance between the single-gene trees and the topology shown in Fig. 1). Removal of fast-evolving sites increased the ML support of this relationship to 100% (Fig. 2A), and the removal of poorly represented genes increased the ML support to 96% (Fig. 2B).

Ulvophyceae, with the exclusion of Bryopsidales, was highly supported as a separate clade (94%–98% ML-, 1.00 Bayesian- and 100% MLBS-support), with *Cladophora glomerata* (Cladophorales) and Trentepohliales clustering together as a maximally supported clade, sister to all other ulvophyceans. The

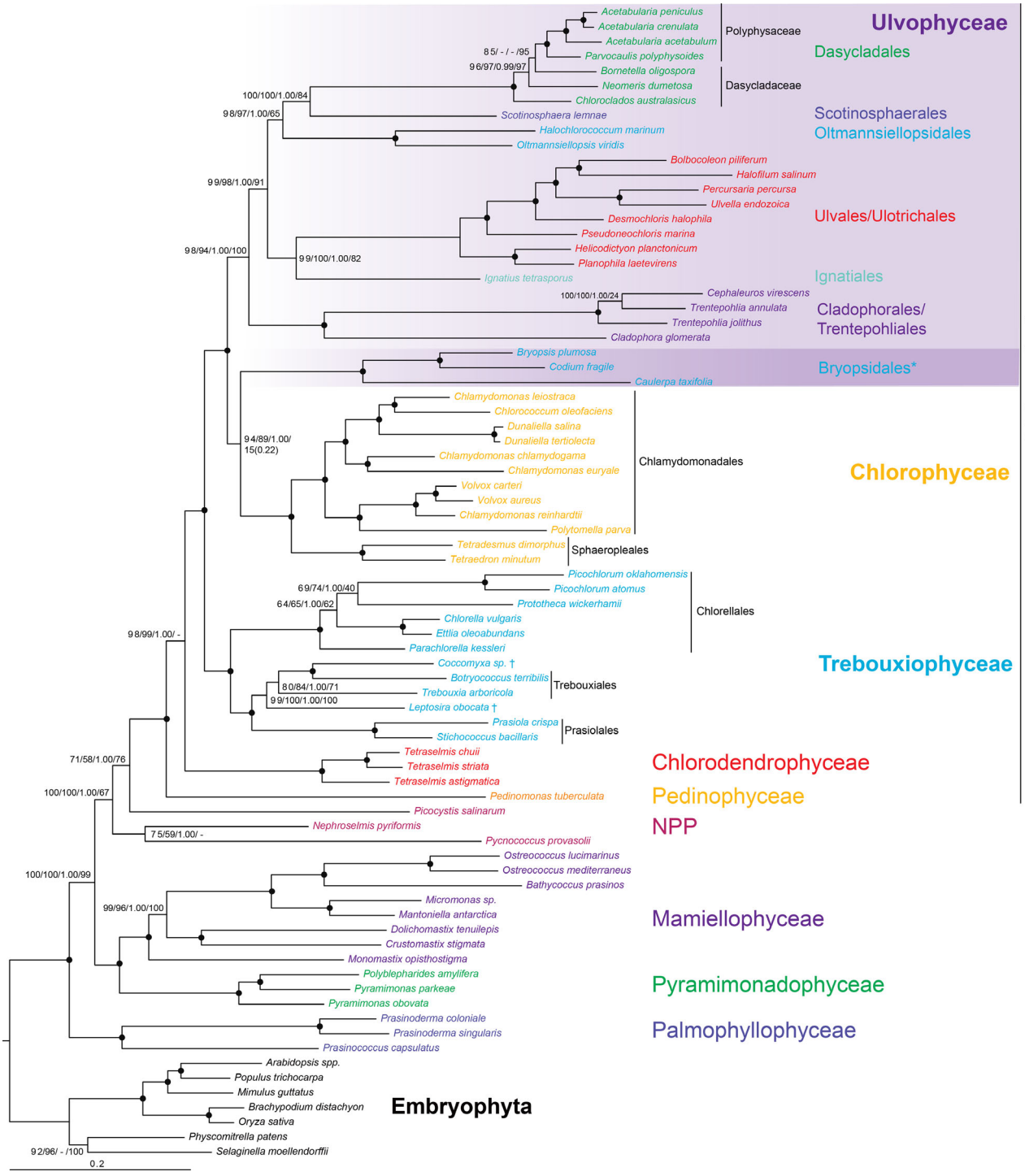

FIG. 1. Molecular phylogeny of Chlorophyta. Phylogenetic tree of Chlorophyta, with Embryophyta as outgroup, made from Maximum Likelihood analysis of a gene matrix consisting of 78 taxa and 95 genes. Support values on the branches are from the Maximum Likelihood (ML) analysis (both with and without modeling heterotachy), Bayesian analysis, and the coalescence analysis written in the following order on the branches: ML/ML-heterotachy/Bayes/coalescence. A black dot on a branching point indicates maximum support in all analyses (100/100/1.00/100). A "-" symbol indicates that the node was not recovered in the respective analysis. The number in parenthesis on the bifurcation between Bryopsidales and Chlorophyceae is the p-value from the polytomy performed with ASTRAL. See the Methods section for details. NPP is an abbreviation of the three species *Picocystis salinarum*, *Nephroselmis pyriformis*, and *Pycnococcus provasolii*. *Bryopsidales is currently classified among Ulvophyceae. †*Coccomyxa* sp. and *Leptosira obocata* are classified as Trebouxiophyceae *incertae sedis*.

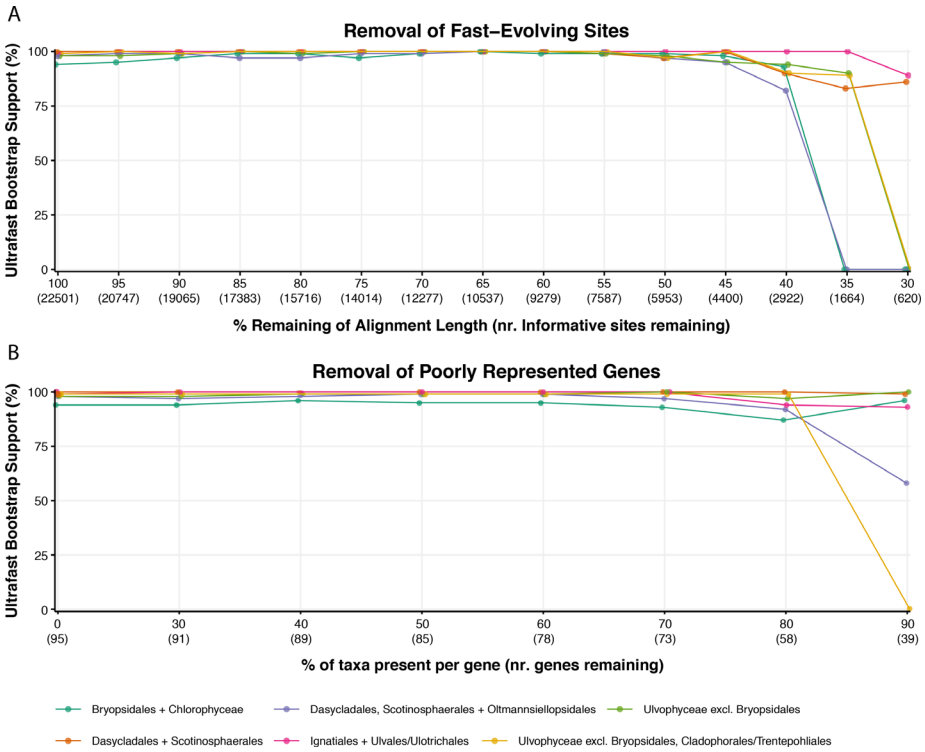

FIG. 2. Removing fast-evolving sites and genes poorly represented among taxa. (A) Sites in the alignment used to generate Figure 1 were ranked according to evolutionary rate (see Methods) and removed in 5% increments, starting with the fastest evolving and down to the 30% fastest. For each increment, a phylogenetic tree was generated with IQ-TREE. The IQ-TREE ultrafast bootstrap support for selected nodes is shown on the *y*-axis and the number of sites removed is shown on the *x*-axis (shown as sites remaining, i.e., 100% remaining = 0% removed). Informative sites are parsimony-informative, that is, sites that contain at least two different amino acids and each of them occur in at least two taxa. (B) Genes in the same alignment as in A were ranked according to the percent of taxa in which the gene was present (i.e., 100% = gene present in all taxa). Starting with the most poorly represented genes (i.e., genes present in 30% of the taxa), genes were removed in 10% increments up to 90%. For each increment, the same phylogenetic analysis as in A was performed. The support values for selected nodes are shown on the *y*-axis, and the proportion of genes removed is shown on the *x*-axis, with the number of alignment sites remaining shown in parentheses (0% on the *x*-axis indicates that no genes were removed).

remainder of Ulvophyceae (98%–99% ML-, 1.00 Bayesian- and 91% MLBS-support) were separated into two groups: One group consisted of *Ignatius tetrasporus* (Ignatiales) and the orders Ulvales, Ulotrichales, and Chlorocystidales (99%–100% ML-, 1.00 Bayesian- and 82% MLBS-support). Although Chlorocystidales (represented by *Desmochloris halophila*) is classified as a separate order in AlgaeBase (Guiry et al. 2014), it clustered within Ulvales as previously noted by Watanabe et al. (2001), and we hereafter refer to these three orders together as Ulvales/Ulotrichales. The other group was made up of Oltmannsiellopsidales, Scotinosphaerales, and Dasycladales (97%–98% ML-, 1.00 Bayesian- and 65% MLBS-support), placing Scotinosphaerales and Dasycladales as sisters (maximum support, except for 84% MLBS support).

All of the relationships were robust to the removal of fast-evolving sites and poorly represented genes. Only at the strictest removal of poorly represented genes (39 genes and 16,696 amino acids remaining) did the Cladophorales/Trentepohliales clade cluster as sister to the clade of Ignatiales and Ulvales/Ulotrichales, but this topology was weakly supported (50%). Removing *Ignatius tetrasporus* from the alignment changed the position of Oltmannsiellopsidales as sister to Ulvales/Ulotrichales (89%; Fig. S2 in the Supporting Information). No topology changes were seen when removing Oltmannsiellopsidales from the analyses, although the support for a Chlorophyceae/Bryopsidales sister relationship was reduced to 58% (Fig. S3 in the Supporting Information).

Dasycladales. The seven newly sequenced species of Dasycladales were fully supported as a monophyletic group. Within Dasycladales, the family Dasycladaceae (represented by *C. australasicus*, *Neomeris dumetosa*, and *Bornetella oligospora*) was paraphyletic with all inference methods. *C. australasicus* was always separate from the rest of the dasycladaleans, but whether *N. dumetosa* and *B. oligospora* formed a clade or not differed between inference methods (they formed a clade in the Bayesian tree topology and when modeling heterotachy in the ML analysis, results not shown). The family Polyphysaceae (represented by *Parvocaulis polyphysoides* and the genus *Acetabularia*) clustered as the most derived clade with full support.

DISCUSSION

Relationship between the different lineages of Chlorophyta. Separated from the rest of the chlorophytes, our phylogeny supports a sister relationship between *Prasinococcus* and *Prasinoderma*, confirming analyses based on mitochondrial (Turmel et al. 2020), chloroplast (Lemieux et al. 2014b), and nuclear data (Li et al. 2020). Additionally, and in line with both chloroplast- and nuclear-based phylogenies (Leliaert et al. 2016, Del Cortona et al. 2020), Pyramimonadales and Mamiellophyceae are sister clades and separated from NPP (*Nephroselmis pyriformis*, *Pycnococcus provasolii*, and *Picocystis salinarum*) and the core Chlorophyta. The relationships between the NPP lineages are not fully resolved, although *N. pyriformis* and *P. provasolii* are probably sister species, and *P. salinarum* is most likely the most closely related to the core Chlorophyta. Nonetheless, *P. provasolii* is strongly excluded from the Pyramimonadophyceae in all our analyses, which is in line with recently suggested taxonomical revisions excluding Pseudocourfieldiales and *Pycnococcus* from Pyramimonadophyceae (Adl et al. 2019, Daugbjerg et al. 2020). Considering the long branches of the three NPP taxa, increased sampling of these lineages is probably required to confidently establish their precise relationships.

We recovered core Chlorophyta as monophyletic with Pedinophyceae and Chlorodendrophyceae separated from the other groups, although. Some chloroplast phylogenies have suggested the placement of Pedinophyceae as sister to Chlorellales (Lemieux et al. 2014a, Leliaert and Lopez-Bautista 2015, Lemieux et al. 2015). In the UTC group (Ulvophyceae, Trebouxiophyceae, Chlorophyceae), Trebouxiophyceae is excluded from Ulvophyceae and Chlorophyceae, in line with several other phylogenies (Cocquyt et al. 2010b, Lemieux et al. 2014a, Leliaert et al. 2016, Turmel et al. 2017, Del Cortona et al. 2020). The position among Trebouxiophyceae for both *Coccomyxa* and *Leptosira obocata* is uncertain (Guiry et al. 2014). However, our analyses suggest that *Coccomyxa* belongs to Trebouxiales, as it is sister to *Botryococcus terribilis*. Otherwise, Trebouxiales is paraphyletic. It seems that our inclusion of Chlorodendrophyceae and the trebouxiophyte genera *Picochlorum* and *Prototheca* was important for recovering a monophyletic Trebouxiophyceae. Furthermore, Chlorellales is in our phylogeny strongly supported as sister to the other Trebouxiophyceae, instead of making up a separate clade outside the UTC group as seen in some chloroplast-based studies (Lemieux et al. 2014a,b, Leliaert and Lopez-Bautista 2015).

Interestingly, *E. oleoabundans* is fully supported as sister to *C. vulgaris*. This suggests that it does not belong to the Chlamydomonadales (Chlorophyceae) as it is currently classified (Guiry et al. 2014), but should instead be considered part of Chlorellales (Trebouxiophyceae) in line with earlier phylogenetic and comparative proteomics studies (Pegg et al. 2015, Garibay-Hernández et al. 2016).

Phylogeny of the Ulvophyceae. Resolving the phylogenetic relationships between the different groups of Ulvophyceae has been notoriously difficult using molecular phylogenetics (Fang et al. 2017). As already mentioned, the position of Bryopsidales has been difficult to infer, and several clades, such as Dasycladales, Oltmannsiellopsidales, and Ignatiales, often have very long branches and unstable positions in molecular phylogenies.

The phylogenetic discrepancies of Ulvophyceae are probably largely due to differences in taxon sampling between studies and low sampling of taxa within groups. For instance, with the exception of Del Cortona et al. (2017), Cladophorales have been absent from multigene phylogenies based on chloroplast data, and available data are sparse for both Trentepohliales and Dasycladales. Molecular phylogenetic analyses of Dasycladales have usually only included a single species, *A. acetabulum*, which has a very long branch. In addition, both Ignatiales and Oltmannsiellopsidales have often only been represented by a single or very few species, if present at all. All these lineages have very different cytomorphologies and therefore probably represent very different cellular evolutionary histories of the Ulvophyceae.

Although Del Cortona et al. (2020) shortened the branches leading to Dasycladales with their inclusion of Scotinosphaerales, and to Cladophorales/Trentepohliales (C/T) with their inclusion of *Blastophysa rhizopus*, they still received conflicting support for the relationship between these lineages. Our increased taxon sampling of Dasycladales, combined with a conservative selection of genes, seems to have alleviated these instabilities. As a result, our phylogeny of Ulvophyceae is overall highly supported and is not sensitive to the presence or absence of fast-evolving sites or poorly represented genes. We firmly recover Dasycladales among Ulvophyceae with the Scotinosphaerales as the sister group, in agreement with Del Cortona et al. (2020). In contrast, however, we find that Oltmannsiellopsidales is most closely related to the clade of Scotinosphaerales and Dasycladales. Hence, we do not find support for a close relationship between Dasycladales and Cladophorales and/or Trentepohliales. Instead, our results support that C/T evolved from an ancestor distinct from that of all other ulvophyceans. Interestingly, an early radiation of these groups is further supported by a recently described one-billion-years-old chlorophyte fossil hypothesized to resemble modern Cladophorales (Tang et al. 2020).

Common genomic and ultrastructural characters are not suitable as phylogenetic markers for Ulvophyceae. A noncanonical codon usage, where TAA and TAG encode glutamine instead of a stop signal, is found in Dasycladales, Trentepohliales, Cladophorales, and *Blastophysa* (Schneider et al. 1989, Gile et al.

2009, Cocquyt et al. 2010a), and we have here verified the presence of this alternate genetic code in Scotinosphaerales as well (Appendix S1). A single clade of these five lineages (e.g. Del Cortona et al. 2020) is therefore the most parsimonious evolutionary scenario considering the distribution of the alternative codon. And the presence of the 1-alpha paralog of the elongation factor gene (EF1A) in both Cladophorales and Dasycladales (Gile et al. 2009), as well as ultrastructural similarities in the flagellar apparatus of these two groups (O'kelly and Floyd 1984), have traditionally also supported a close relationship between Cladophorales and Dasycladales. However, the position of Cladophorales and Trentepohliales as separate from the rest of the Ulvophyceae, which is highly supported in our phylogeny, rather suggests that these genomic and ultrastructural characters have either evolved early in the radiation of Ulvophyceae, or even long before, and was lost on several occasions, or have originated independently. In either case, they are unsuitable as phylogenetic markers. As noted by Gile et al. (2009), EF1A is quite widespread among Ulvophyceae and is even present in the more distant Bryopsidales. Additionally, the same noncanonical codons have also evolved in other eukaryotes such as the ciliates and the metamonads (in the oxymonads and the diplomonads; Gile et al. 2009) and could have originated multiple times semi-independently in Ulvophyceae (see the stepwise acquisition hypothesis proposed by Cocquyt et al. 2010a).

Evolution of Dasycladales. Our broad taxonomic sampling of Dasycladales has provided a better opportunity to understand the evolution of macroscopic cell morphologies within this order. Our phylogeny fully supports the monophyly of Dasycladales and confirms the two families Dasycladaceae and Polyphysaceae. However, similar to previous studies, we find that Polyphysaceae has emerged from the paraphyletic Dasycladaceae (Olsen et al. 1994, Berger et al. 2003, Zechman 2003, Verbruggen et al. 2009). All species of Dasycladales have a similar cytomorphology consisting of a single large siphonous cell. However, in the family Dasycladaceae the entire stalk is covered by lateral branching hairs, or whorls, with the gametangia spread across the entire cell length, giving the algae an elongated and "bushy" or sausage-looking morphology (Berger 2006). An example of this is *C. australasicus*. This species is sister to all other dasycladaleans in our analysis and possesses a large number of gametophores resembling small spheres, or grapes, covering the entire length of the cell. Species of Polyphysaceae on the other hand have a markedly different morphology. The whorls along the stalk are lost during development, leaving the entire stalk naked except for the apical tip where highly characteristic and species-specific caps are formed, such as in the genus *Acetabularia*. The fact that Polyphysaceae evolved from Dasycladaceae suggests that the ancestral morphology of Dasycladales was a "bushy" cell with gametangia along the entire cell body and that the distinct caps and bare stalks of the polyphysacean species evolved from a reduction of the multiple radial whorls of the stalk present in Dasycladaceae.

But what was the cytomorphology of the ancestral dasycladalean? Although Dasycladales is commonly described as a group of single-celled algae with a single nucleus (Berger 2006, Verbruggen et al. 2009, Del Cortona et al. 2020), there are uncertainties regarding the number of transcriptionally active, or primary, nuclei in several species. The presence of a single, primary nucleus is well established in species of *Acetabularia* and *Batophora* (Burr and West 1971, Spring et al. 1974). Likewise, *Cymopolia* and the closely related *Neomeris annulata* are cited in Chapman and Chapman (1973) as having a single primary nucleus (although Liddle et al. 1982 could not fully confirm this when investigating *Cymopolia barbata*). *Bornetella*, however, is often described as being multinucleate (Cocquyt et al. 2010b), without specifying whether these are multiple primary nuclei or not. *Bornetella*, and certain other members of Dasycladales such as *C. barbata*, have holocarpic reproduction (Liddle et al. 1982), continuous production of secondary haploid nuclei and release of gametes throughout growth, instead of production of secondary nuclei only during certain periods of the life cycle (as is the case in *Acetabularia*). Cells with holocarpic reproduction will therefore have multiple nuclei present most of the time, and the descriptions of multinucleate Dasycladales may be misinterpretations of these continuously produced secondary nuclei. However, it is currently not possible to rule out the presence of true multinucleate dasycladaleans as several genera do not appear to have been investigated. Little is known about the cytological organization of *C. australasicus*, but its close relationship to *Batophora* (Berger et al. 2003, Verbruggen et al. 2009) suggests that it is also uninucleate. Considering the uninucleate nature of the majority of Dasycladales, including the most derived genera such as *Acetabularia*, it is plausible that the ancestral dasycladalean was uninucleate and evolved from a microscopic single-celled ancestor, perhaps resembling present-day Scotinosphaerales and Oltmannsiellopsidales. On the other hand, given that decoupling of nuclear division (karyokinesis) and cellular division (cytokinesis) is common across Chlorophyta (Niklas et al. 2013), it is not entirely unlikely that Dasycladales were ancestrally multinucleate.

A new framework for interpreting the evolution of ulvophycean cytomorphologies. We recover Bryopsidales as a lineage separated from other Ulvophyceae, rendering the Ulvophyceae paraphyletic. The exact placement of Bryopsidales has largely been uncertain until now, either because of weak statistical

support or discrepant topologies between analysis of concatenated gene matrices and coalescence analysis. A recent phylogenetic analysis based on more than 500 nuclear genes from 55 green algal and plant species, with a dense taxon sampling of Chlorophyta, obtained high support for a sister relationship between Bryopsidales and Chlorophyceae in analyses of a concatenated gene matrix (Del Cortona et al. 2020). However, based on a coalescence analysis they could not reject the possibility of a hard polytomy between Chlorophyceae, Bryopsidales, and Ulvophyceae. Similarly, coalescence analysis on our data could not reject such a hard polytomy or support a sister relationship between Bryopsidales and Chlorophyceae. However, it should be noted that coalescence did not reject this topology nor support alternative topologies. And in contrast to these unresolved nodes in the coalescence analysis, all our concatenated analyses were consistent in supporting the sister relationship between Bryopsidales and Chlorophyceae. This is also congruent with several recent phylogenies based on both chloroplast and nuclear genes (Lü et al. 2011, Smith et al. 2011, Leliaert and Lopez-Bautista 2015, Leliaert et al. 2016, Turmel et al. 2017, Leebens-Mack et al. 2019). A sister relationship between Bryopsidales and Chlorophyceae was also robust to the removal of fast-evolving sites in our alignment and to the presence of missing data.

Our study and the studies of Leebens-Mack et al. (2019) and Del Cortona et al. (2020) show discrepancies between concatenated and coalescence analyses on the position of Bryopsidales among core chlorophytes. But the low resolution in the coalescence analysis questions the usefulness of this method with the currently available sequence data for these algal groups. Resolving this question probably requires better sampling of taxa among deeply diverging Ulvophyceae, Bryopsidales, and Chlorophyceae. Increasing the number of genes could be useful as well, but since similar results have been obtained with highly different gene numbers in our work and in Del Cortona et al. (2020), it is likely less important than adding more taxa.

Chlorophyta is usually classified into four types of cellular organizations, or cytomorphologies; unicellular: small single-celled species with a single nucleus, siphonous: giant single-celled species with either one or several nuclei (i.e,. uninucleate- or multinucleate siphonous species), multicellular: species composed of multiple uninucleate cells, and siphonocladous: multicellular species with multinucleate cells (Cocquyt et al. 2010b). Both Bryopsidales and Dasycladales are currently placed in the same category of siphonous cytomorphology. This placement made sense when they were seen as sister clades sharing a common siphonous ancestor, and the siphonous cytomorphology was regarded as a synapomorphy of the group. However, if Bryopsidales is sister to Chlorophyceae, and therefore distantly related to other Ulvophyceae, the giant single-celled siphonous cytomorphology has either evolved twice and hence cannot be used to infer the evolutionary history of the ulvophycean cytomorphologies, or it represents the ancestral cytomorphological state of Ulvophyceae.

We thank Prof. W. Martin of the University of Düsseldorf for kindly providing cultures of Dasycladales, UNINETT Sigma2 AS for providing HPC resources, and the Norwegian Sequencing Centre for sequencing services. This work has been funded by the CIME Research Network at UiO through A grant to Shalchian-Tabrizi, and by the Norwegian Research Council through a grant to JB (grant no. 240284 and 306571).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

Øyvind Sætren Gulbrandsen: Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Validation (lead); Visualization (equal); Writing-original draft (lead); Writing-review & editing (equal). **Ina Jungersen Andresen:** Data curation (supporting); Formal analysis (supporting); Methodology (supporting); Supervision (supporting); Validation (supporting); Visualization (supporting); Writing-original draft (supporting); Writing-review & editing (equal). **Anders Kristian Krabberød:** Data curation (supporting); Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Validation (supporting); Visualization (supporting); Writing-review & editing (supporting). **Jon Bråte:** Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Methodology (equal); Project administration (lead); Supervision (lead); Validation (lead); Visualization (lead); Writing-original draft (supporting); Writing-review & editing (lead). **Kamran Shalchian-Tabrizi:** Conceptualization (lead); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (supporting); Project administration (equal); Supervision (equal); Writing-original draft (lead).

DATA AVAILABILITY

Transcriptome data generated in this study are available at the European Nucleotide Archive under the accession nr: PRJEB40441. Sequence alignments used in this study are available at Mendeley Data (data.mendeley.com) with the https://doi.org/10.17632/89xh5zp6zr.1.

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., Berney, C., Brown, M. W. & Burki, F. 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* 66:4–119.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S. & Prjibelski, A. D. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19:455–77.

Berger, S. 2006. Photo-Atlas of living Dasycladales. *Carnets de Géologie*. Brest, 348 pp.

Berger, S., Fettweiss, U., Gleissberg, S., Liddle, L. B., Richter, U., Sawitzky, H. & Zuccarello, G. C. 2003. 18S rDNA phylogeny and evolution of cap development in Polyphysaceae (formerly Acetabulariaceae; Dasycladales, Chlorophyta). *Phycologia* 42:506–61.

Bolger, A. M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–20.

Burki, F., Kaplan, M., Tikhonenkov, D. V., Zlatogursky, V., Minh, B. Q., Radaykina, L. V., Smirnov, A., Mylnikov, A. P. & Keeling, P. J. 2016. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc. R. Soc. B.* 283:20152802.

Burr, F. A. & West, J. A. 1971. Comparative ultrastructure of the primary nucleus in *Bryopsis* and *Acetabularia*. *J. Phycol.* 7:108–13.

Chapman, V. & Chapman, D. 1973. *The Algae*. Macmillan Press, London, UK, 497 pp.

Cocquyt, E., Gile, G. H., Leliaert, F., Verbruggen, H., Keeling, P. J. & De Clerck, O. 2010a. Complex phylogenetic distribution of a non-canonical genetic code in green algae. *BMC Evol. Biol.* 10:327.

Cocquyt, E., Verbruggen, H., Leliaert, F. & De Clerck, O. 2010b. Evolution and cytological diversification of the green seaweeds (Ulvophyceae). *Mol. Biol. Evol.* 27:2052–61.

Coneva, V. & Chitwood, D. H. 2015. Plant architecture without multicellularity: quandaries over patterning and the soma-germline divide in siphonous algae. *Front. Plant Sci.* 6:287.

Crotty, S. M., Minh, B. Q., Bean, N. G., Holland, B. R., Tuke, J., Jermiin, L. S. & Haeseler, A. V. 2020. GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Syst. Biol.* 69:249–64.

Daugbjerg, N., Fassel, N. M. & Moestrup, Ø. 2020. Microscopy and phylogeny of *Pyramimonas tatianae* sp. nov. (Pyramimonadales, Chlorophyta), a scaly quadriflagellate from Golden Horn Bay (eastern Russia) and formal description of Pyramimonadophyceae classis nova. *Eur. J. Phycol.* 55:49–63.

Del Cortona, A., Jackson, C. J., Bucchini, F., Van Bel, M., D'hondt, S., Škaloud, P., Delwiche, C. F. et al. 2020. Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *Proc. Natl. Acad. Sci. USA* 117:2551–9.

Del Cortona, A., Leliaert, F., Bogaert, K. A., Turmel, M., Boedeker, C., Janouškovec, J., Lopez-Bautista, J. M., Verbruggen, H., Vandepoele, K. & De Clerck, O. 2017. The plastid genome in Cladophorales green algae is encoded by hairpin chromosomes. *Curr. Biol.* 27:3771–82.e6.

Fang, L., Leliaert, F., Zhang, Z. H., Penny, D. & Zhong, B. J. 2017. Evolution of the Chlorophyta: insights from chloroplast phylogenomic analyses. *J. Sys. Evol.* 55:322–32.

Fučíková, K., Leliaert, F., Cooper, E. D., Škaloud, P., D'Hondt, S., De Clerck, O., Gurgel, C. F. D. et al. 2014. New phylogenetic hypotheses for the core Chlorophyta based on chloroplast sequence data. *Front. Ecol. Evol.* 2:63.

Garibay-Hernández, A., Barkla, B. J., Vera-Estrella, R., Martinez, A. & Pantoja, O. 2016. Membrane proteomic insights into the physiology and taxonomy of a green microalga. *Plant Physiol.* 173:390–416.

Gile, G. H., Novis, P. M., Cragg, D. S., Zuccarello, G. C. & Keeling, P. J. 2009. The distribution of elongation factor-1 alpha (EF-1α), elongation factor-like (EFL), and a non-canonical genetic code in the Ulvophyceae: discrete genetic characters support a consistent phylogenetic framework. *J. Eukaryot. Microbiol.* 56:367–72.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R. & Zeng, Q. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29:644–52.

Guiry, M. D., Guiry, G. M., Morrison, L., Rindi, F., Miranda, S. V., Mathieson, A. C., Parker, B. C., Langangen, A., John, D. M. & Bárbara, I. 2014. AlgaeBase: an on-line resource for algae. *Cryptogamie Algol.* 35:105–15.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B. & Lieber, M. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Prot* 8:1494–512.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. 2017. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35:518–22.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A. & Jermiin, L. S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14:587–9.

Katoh, K. & Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–80.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S. & Duran, C. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–9.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K. & Bell, C. J. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* 12: e1001889.

Kumar, S., Krabberød, A. K., Neumann, R. S., Michalickova, K., Zhao, S., Zhang, X. & Shalchian-Tabrizi, K. 2015. BIR pipeline for preparation of phylogenomic data. *Evol. Bioinform* 11:79–83.

Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* 62:611–5.

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W. & Grosse, I. et al. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574:679–85.

Leliaert, F. & Lopez-Bautista, J. M. 2015. The chloroplast genomes of *Bryopsis plumosa* and *Tydemania expeditiones* (Bryopsidales, Chlorophyta): compact genomes and genes of bacterial origin. *BMC Genom.* 16:204.

Leliaert, F., Tronholm, A., Lemieux, C., Turmel, M., DePriest, M. S., Bhattacharya, D., Karol, K. G., Fredericq, S., Zechman, F. W. & Lopez-Bautista, J. M. 2016. Chloroplast phylogenomic analyses reveal the deepest-branching lineage of the Chlorophyta. Palmophyllophyceae class. nov. *Sci. Rep.* 6:25367.

Lemieux, C., Otis, C. & Turmel, M. 2014a. Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. *BMC Evol. Biol.* 14:211.

Lemieux, C., Otis, C. & Turmel, M. 2014b. Six newly sequenced chloroplast genomes from prasinophyte green algae provide insights into the relationships among prasinophyte lineages and the diversity of streamlined genome architecture in picoplanktonic species. *BMC Genom.* 15:857.

Lemieux, C., Vincent, A. T., Labarre, A., Otis, C. & Turmel, M. 2015. Chloroplast phylogenomic analysis of chlorophyte green algae identifies a novel lineage sister to the Sphaeropleales (Chlorophyceae). *BMC Evol. Biol.* 15:264.

Li, L., Wang, S., Wang, H., Sahu, S. K., Marin, B., Li, H., Xu, Y. et al. 2020. The genome of Prasinoderma coloniale unveils the existence of a third phylum within green plants. *Nat. Ecol. Evol.* 4:1220–31.

Liddle, L. B., Thomas, J. P. & Scott, J. 1982. Morphology and distribution of nuclei during development in *Cymopolia barbata* (Chlorophycophyta, Dasycladales). *J. Phycol.* 18:257–64.
Lü, F., Xü, W., Tian, C., Wang, G., Niu, J., Pan, G. & Hu, S. 2011. The Bryopsis hypnoides plastid genome: multimeric forms and complete nucleotide sequence. *PLoS One* 6:e14663.
Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32:268–74.
Niklas, K. J., Cobb, E. D. & Crawford, D. R. 2013. The evo-devo of multinucleate cells, tissues, and organisms, and an alternative route to multicellularity. *Evol. Dev.* 15:466–74.
O'kelly, C. & Floyd, G. 1984. Correlations among patterns of sporangial structure and development, life histories, and ultrastructural features in the Ulvophyceae. *In* Irvine, D. & John, D. [Eds.] *Systematics of the Green Algae.* Academic Press, London, UK, pp. 121–56.
Olsen, J. L., Stam, W. T., Berger, S. & Menzel, D. 1994. 18S rDNA and evolution in the Dasycladales (Chlorophyta): modern living fossils. *J. Phycol.* 30:729–44.
Pegg, C., Wolf, M., Alanagreh, L. A., Portman, R. & Buchheim, M. A. 2015. Morphological diversity masks phylogenetic similarity of Ettlia and Haematococcus (Chlorophyceae). *Phycologia* 54:385–97.
Price, M. N., Dehal, P. S. & Arkin, A. P. 2010. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490.
Ranjan, A., Townsley, B. T., Ichihashi, Y., Sinha, N. R. & Chitwood, D. H. 2015. An intracellular transcriptomic atlas of the giant coenocyte *Caulerpa taxifolia*. *PLoS Genet.* 11: e1004900.
Schneider, S. U., Leible, M. B. & Yang, X. P. 1989. Strong homology between the small subunit of ribulose-1, 5-bisphosphate carboxylase/oxygenase of two species of *Acetabularia* and the occurrence of unusual codon usage. *Mol. Gen. Genet.* 218:445–52.
Smith, D. R., Burki, F., Yamada, T., Grimwood, J., Grigoriev, I. V., Van Etten, J. L. & Keeling, P. J. 2011. The GC-rich mitochondrial and plastid genomes of the green alga Coccomyxa give insight into the evolution of organelle DNA nucleotide landscape. *PLoS One* 6:e23624.
Spring, H., Trendelenburg, M. F., Scheer, U., Franke, W. W. & Herth, W. 1974. Structural and biochemical studies of the primary nucleus of two green algal species, *Acetabularia mediterranea* and *Acetabularia major*. *Cytobiologie* 10:1–65.
Tang, Q., Pang, K., Yuan, X. & Xiao, S. 2020. A one-billion-year-old multicellular chlorophyte. *Nat. Ecol. Evol.* 4:543–9.
Turmel, M., Otis, C. & Lemieux, C. 2017. Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophycean green algae. *Sci. Rep.* 7:994.
Turmel, M., Otis, C. & Lemieux, C. 2020. Complete mitogenomes of the marine picoplanktonic green algae *Prasinoderma* sp. MBIC 10622 and *Prasinococcus capsulatus* CCMP 1194 (Palmophyllophyceae). *Mitochondrial DNA B* 5:166–8.
Verbruggen, H. 2018. SiteStripper. v1.03. https://github.com/hverbruggen/SiteStripper
Verbruggen, H., Ashworth, M., LoDuca, S. T., Vlaeminck, C., Cocquyt, E., Sauvage, T., Zechman, F. W., Littler, D. S., Littler, M. M. & Leliaert, F. 2009. A multi-locus time-calibrated phylogeny of the siphonous green algae. *Mol. Phylogenet. Evol.* 50:642–53.
Watanabe, S., Kuroda, N. & Maiwa, F. 2001. Phylogenetic status of *Helicodictyon planctonicum* and *Desmochloris halophila* gen. et comb. nov. and the definition of the class Ulvophyceae (Chlorophyta). *Phycologia* 40:421–34.
Zechman, F. W. 2003. Phylogeny of the Dasycladales (Chlorophyta, Ulvophyceae) based on analyses of RuBisCO large subunit (*rbc*L) gene sequences. *J. Phycol.* 39:819–27.
Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19:153.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. Presence and absence of genes in the multigene alignment. The multigene alignment used to generate Figure 1 consisted of 95 genes and 78 taxa. A white box indicates that the gene is absent in the alignment from that taxon. The color intensity of blue indicates the length of the gene sequence compared to the longest representative sequence of the gene. The taxa (rows) are ordered according to the clustering pattern in Figure 1. The phylogeny from Figure 1 is represented by the dendrogram.

Figure S2. Testing the impact of removing Ignatiales from the phylogenetic reconstruction. *Ignatius tetrasporus* was removed from the alignment used to generate Figure 1, and the alignment was analyzed using IQ-TREE in the same way as for Figure 1. Labels and values are the same as for Figure 1. Only the portion of the tree containing Ulvophyceae and Chlorophyceae is shown here for simplicity, but the branching pattern of the remaining taxa is the same as in Figure 1.

Figure S3. Testing the impact of removing Oltmannsiellopsidales from the phylogenetic reconstruction. See Figure S2 for explanation. The only difference is that *Halochlorococcum marinum* and *Oltmannsiellopsis viridis* were removed from the alignment instead of *Ignatius tetrasporus*.

Table S1. Downloaded paired-end reads from EBI and their taxonomic origin according to the NCBI taxonomy database. * *Picocystis salina* is a synonym for *Halofilum salinum*.

Table S2. Downloaded transcriptome assemblies from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) and their taxonomic origins according to AlgaeBase. *The transcriptome of *Caulerpa taxifolia* originates from Ranjan et al. (2015). **The transcriptome of *Scotinosphaera lemnae* originates from Del Cortona et al. (2020).

Appendix S1. Supplementary note regarding the non-canonical codon usage in *Scotinosphaera lemnae*.