

# Internal Transcribed Spacer 2 (nu ITS2 rRNA) Sequence-Structure Phylogenetics: Towards an Automated Reconstruction of the Green Algal Tree of Life

Mark A. Buchheim<sup>1\*</sup>, Alexander Keller<sup>2</sup>, Christian Koetschan<sup>2</sup>, Frank Förster<sup>2</sup>, Benjamin Merget<sup>2</sup>, Matthias Wolf<sup>2</sup>

**1** Department of Biological Science and the Mervin Bovaird Institute for Molecular Biology and Biotechnology, The University of Tulsa, Tulsa, Oklahoma, United States of America, **2** Department of Bioinformatics, Biocenter, University of Würzburg, Am Hubland, Würzburg, Germany

## Abstract

**Background:** Chloroplast-encoded genes (*matK* and *rbcl*) have been formally proposed for use in DNA barcoding efforts targeting embryophytes. Extending such a protocol to chlorophyten green algae, though, is fraught with problems including non homology (*matK*) and heterogeneity that prevents the creation of a universal PCR toolkit (*rbcl*). Some have advocated the use of the nuclear-encoded, internal transcribed spacer two (ITS2) as an alternative to the traditional chloroplast markers. However, the ITS2 is broadly perceived to be insufficiently conserved or to be confounded by introgression or biparental inheritance patterns, precluding its broad use in phylogenetic reconstruction or as a DNA barcode. A growing body of evidence has shown that simultaneous analysis of nucleotide data with secondary structure information can overcome at least some of the limitations of ITS2. The goal of this investigation was to assess the feasibility of an automated, sequence-structure approach for analysis of ITS2 data from a large sampling of phylum Chlorophyta.

**Methodology/Principal Findings:** Sequences and secondary structures from 591 chlorophycean, 741 trebouxiophycean and 938 ulvophycean algae, all obtained from the ITS2 Database, were aligned using a sequence structure-specific scoring matrix. Phylogenetic relationships were reconstructed by Profile Neighbor-Joining coupled with a sequence structure-specific, general time reversible substitution model. Results from analyses of the ITS2 data were robust at multiple nodes and showed considerable congruence with results from published phylogenetic analyses.

**Conclusions/Significance:** Our observations on the power of automated, sequence-structure analyses of ITS2 to reconstruct phylum-level phylogenies of the green algae validate this approach to assessing diversity for large sets of chlorophyten taxa. Moreover, our results indicate that objections to the use of ITS2 for DNA barcoding should be weighed against the utility of an automated, data analysis approach with demonstrated power to reconstruct evolutionary patterns for highly divergent lineages.

**Citation:** Buchheim MA, Keller A, Koetschan C, Förster F, Merget B, et al. (2011) Internal Transcribed Spacer 2 (nu ITS2 rRNA) Sequence-Structure Phylogenetics: Towards an Automated Reconstruction of the Green Algal Tree of Life. PLoS ONE 6(2): e16931. doi:10.1371/journal.pone.0016931

**Editor:** Jonathan Badger, J. Craig Venter Institute, United States of America

**Received:** October 3, 2010; **Accepted:** January 5, 2011; **Published:** February 10, 2011

**Copyright:** © 2011 Buchheim et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Financial support for this study was provided by the Deutsche Forschungsgemeinschaft (Mu-2831/1-1 to MW) and the National Science Foundation (MCB-0132083 and DEB-0129030 to MAB). AK was supported by the BIGGS Graduate School, Bavaria, Germany. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: mark-buchheim@utulsa.edu

## Introduction

Researchers for a host of organisms have turned to DNA barcoding as a powerful, new tool in the study of diversity. Although the literature is replete with cautionary statements regarding DNA barcoding [1,2,3,4,5,6], a large number of studies have suggested that the benefits of barcoding either outweigh the problems or that most problems can be addressed [7,8,9,10,11,12,13,14,15,16].

Much of our own research interests have focused less on the issue of species delimitation but rather more on the phylogenetics of chlorophyten green algae [17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32]. Nonetheless, our own work [17,18,33,34,35] and the work of many others [36,37,38,39,40,41,42] have revealed the utility of the nu ITS2 rRNA (ITS2) gene in studies of closely related

green algae. It has become abundantly clear that much of the data gathered in our purely phylogenetics efforts have tremendous potential for use in DNA barcoding for the Chlorophyta.

Barcoding efforts within the Viridiplantae (green plants) have, as one might expect, largely focused on vascular plants, in general, and flowering plants, in particular [8,14,43,44,45,46,47,48,49,50]. Genomic targets for potential land plant barcodes have included chloroplast (*rbcl*, *atpB*, *matK*, *psbA*, *rpoC1*, *rpoB*, *ndhJ*, *accD*), mitochondrial (COX [CO]1) and nuclear genes (various single copy genes, ITS1, ITS2, 5.8S) [51]. Chen et al. (2010) concluded that many of these potential markers are inappropriate for barcoding due to low variability (e.g., *rpoB*, *ndhJ*, *accD*, *atpB*, COX1, 5.8S rRNA) or suffer from difficulties in amplification (e.g., ITS1 rRNA and nuclear, single copy genes). The chloroplast encoded *matK* gene (with *rbcl*) has been formally selected as a

DNA barcoding candidate for the land plants [50]. However, the absence of *matK* from all green algae except the charophytes [52,53,54] renders moot, the question of its utility for the Chlorophyta.

It remains possible that one or more of the problematic genomic targets noted above could be useful for studies of chlorophyean barcoding. However, at present, only the 5.8S rRNA and ITS1 rRNA genes have been studied in more than fifty chlorophyean taxa (3025 GenBank citations). Moreover, if the goal is to identify and test a universal (at least for the Viridiplantae) barcoding candidate, it is important to target only those candidates that will be of use for the land plants. Of those potentially suitable genomic targets that remain, only the cp *rbcL* (2477 current GenBank citations) and nu ITS2 rRNA (3418 current GenBank citations) genes have been routinely targeted for assessing chlorophyean diversity. Investigations of the *rbcL* gene from Chlorophyta have failed to identify a set of universal primers that successfully yield amplicons for all Chlorophyta [17,55,56,57,58]. Moreover, attempts to obtain *rbcL* data from cladophoralean green algae (Ulvophyceae) have largely been unsuccessful (only 3 GenBank citations as of 10/10/2010). Because of the extreme heterogeneity in *rbcL* across the green algae, the *rbcL* is, effectively, a non-universal gene. In contrast, the nu ITS2 gene from virtually all Viridiplantae can be amplified with a single set of universal primers [59]. Despite a relatively short length (128–483 bases across the Chlorophyta), some have even suggested that the nu ITS2 rRNA may be useful for comparisons within much of the domain Eukarya [60,61,62,63,64]. On the basis of the efficiency of amplification, the nu ITS2 rRNA gene is preferable to the cp *rbcL*. In addition, as a nuclear gene, the nu ITS2 rRNA gene is likely to have broader taxonomic applicability (i.e., beyond Viridiplantae) should it be deemed a good DNA barcode.

Many of the limitations first associated with the nu ITS2 rRNA (e.g., too much variation, too few nucleotide sites) have been overcome by secondary structure analysis which has systematically identified regions of variability as well as areas of substantial conservation [61,62,64,65,66,67]. Furthermore, a simulation study recently confirmed the benefit of a sequence-structure approach [68]. Analyses of the simulated data resulted in the most robust trees, as assessed by the bootstrap, when secondary structure data were included in the phylogenetic reconstruction [68]. Moreover, the addition of sequence-structure permits the comparison of a much broader phylogenetic spectrum [68]. In reinforcing the conclusions from the simulation study, recent sequence-structure analyses of ITS2 data from lepidopterans permitted alignment of a broad taxonomic spectrum and yielded phylogenetic reconstructions that matched the resolution provided by analyses of COI and COII [69].

Much of the progress in establishing a nu ITS2 rRNA tool for diversity assessment, has been accomplished as a consequence of new bioinformatics applications, concepts and resources [35,64,65,67,70,71,72,73,74,75]. In particular, the ITS2 Database III has substantially advanced the effectiveness of phylogenetic analyses using ITS2 data. At present, the ITS2 Database III, mined from the NCBI database, comprises over 250,000 structures (both partial and complete) that covers the range of eukaryotic diversity [73]. One of the innovations that is coupled with the database is the use of Hidden Markov Models to more fully automate the annotation pipeline [73]. The final stage of the pipeline involves homology-modelling that provides the user with a sequence-structure assessment that is the product of a phylogenetically broad, comparative approach [73]. Given the bioinformatics support coupled with the relative ease of obtaining comparable data, the nu ITS2 rRNA appears to be a superior

candidate for use in phylogenetic reconstruction of large data arrays and as a DNA barcode for the Chlorophyta.

One goal of this study is to evaluate the use of an automated workflow that includes those analyses suggested by Schultz and Wolf [66] and that can be accomplished within a reasonable time frame on an ordinary desktop computer. The need for automated procedures without further manual corrections in phylogenetics and species delineation is obvious, as the number of available sequences on public databases grows daily.

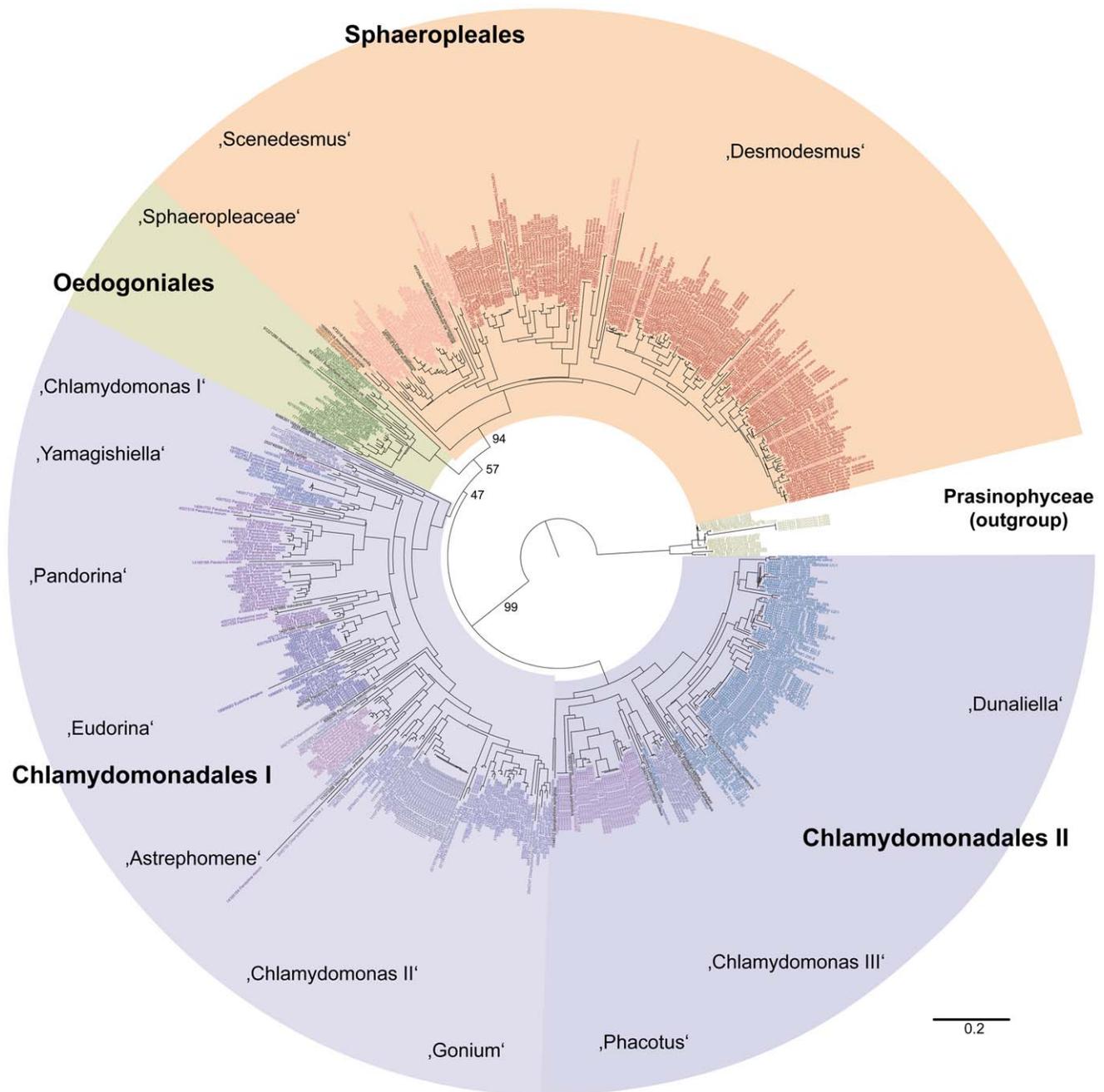
A secondary goal of this investigation is, however, a demonstration of the potential utility of the nu ITS2 rRNA as a DNA barcode for the Chlorophyta as tested against phylogenetic assessments based on other markers. The green algal class, Chlorophyceae, in particular, has been the target of numerous phylogenetic investigations in which the nu ITS2 rRNA gene was included as a genomic target [17,18,29,32,36,39,41,76,77,78,79,80,81,82,83]. These chlorophycean investigations, which represent only a portion of the total body of work in which the nu ITS2 rRNA gene has been used to study chlorophyean diversity (>80 published manuscripts), clearly show the utility of this marker in addressing species level questions. Our challenge is to determine if the use of automated analytical methods with both primary and secondary structural analysis yield robust trees that are largely congruent with other data (e.g., 18S rRNA, 26S rRNA, *rbcL*, *atpB*).

As part of the current investigation, we completed a pilot investigation of the potential for the ITS2 to serve as a DNA barcode for the class Chlorophyceae, which we, then, extended to include the whole of the phylum, Chlorophyta. Results from our tests of this approach clearly indicate that the nu ITS2 rRNA data possess considerable power to reconstruct reasonably robust hypotheses that are congruent with past work that employed markers that have been deemed “more conservative” than the nu ITS2 rRNA gene. Our results indicate that ITS2 has the potential to serve as a powerful tool for phylogenetics in an extraordinarily broad taxonomic context that may eventually encompass virtually the entirety of the domain Eukarya. Furthermore, the empirical results of our investigation suggest that the general antipathy to the implementation of ITS2 as a DNA barcode may not be wholly warranted.

## Results

The aligned nu ITS2 rRNA data for the class Chlorophyceae yielded a tree (Fig. 1) that resolved data representing the orders Oedogoniales (*Oedogonium*, *Bulbochaete* and *Oedocladium*), Sphaeropleales (*Desmodesmus*, *Scenedesmus*, *Atractomorpha* and *Sphaeroplea*), and Chlamydomonadales/Volvocales (*Chlamydomonas* [three non-monophyletic clades], *Yamagishiella*, *Pandorina*, *Eudorina*, *Astrephomene*, *Gonium*, *Phacotus* and *Dunaliella*). Two distinct chlamydomonad alliances were resolved (with only weak bootstrap support) by the ITS2 data (Fig. 1). The Sphaeropleales were resolved as monophyletic with high bootstrap support (94%). Furthermore, distinct lineages corresponding to putative chlorophycean species are preserved by the analytical protocol utilized in this experiment (Fig. 1).

Given the success of the experiment with data from the Chlorophyceae, the test was extended to include a comprehensive sampling of nu ITS2 rRNA sequence data from the green algal classes, Trebouxiophyceae (741 sequences) and Ulvophyceae (938 sequences). These data were analyzed under the same analytical conditions as the Chlorophyceae, including the use of prasino-phycean data as the outgroup. The PNJ analysis resolved three principal clades of trebouxiophycean taxa (Fig. 2) that correspond to two sets of microthamnialean taxa (the *Trebouxia* alliance

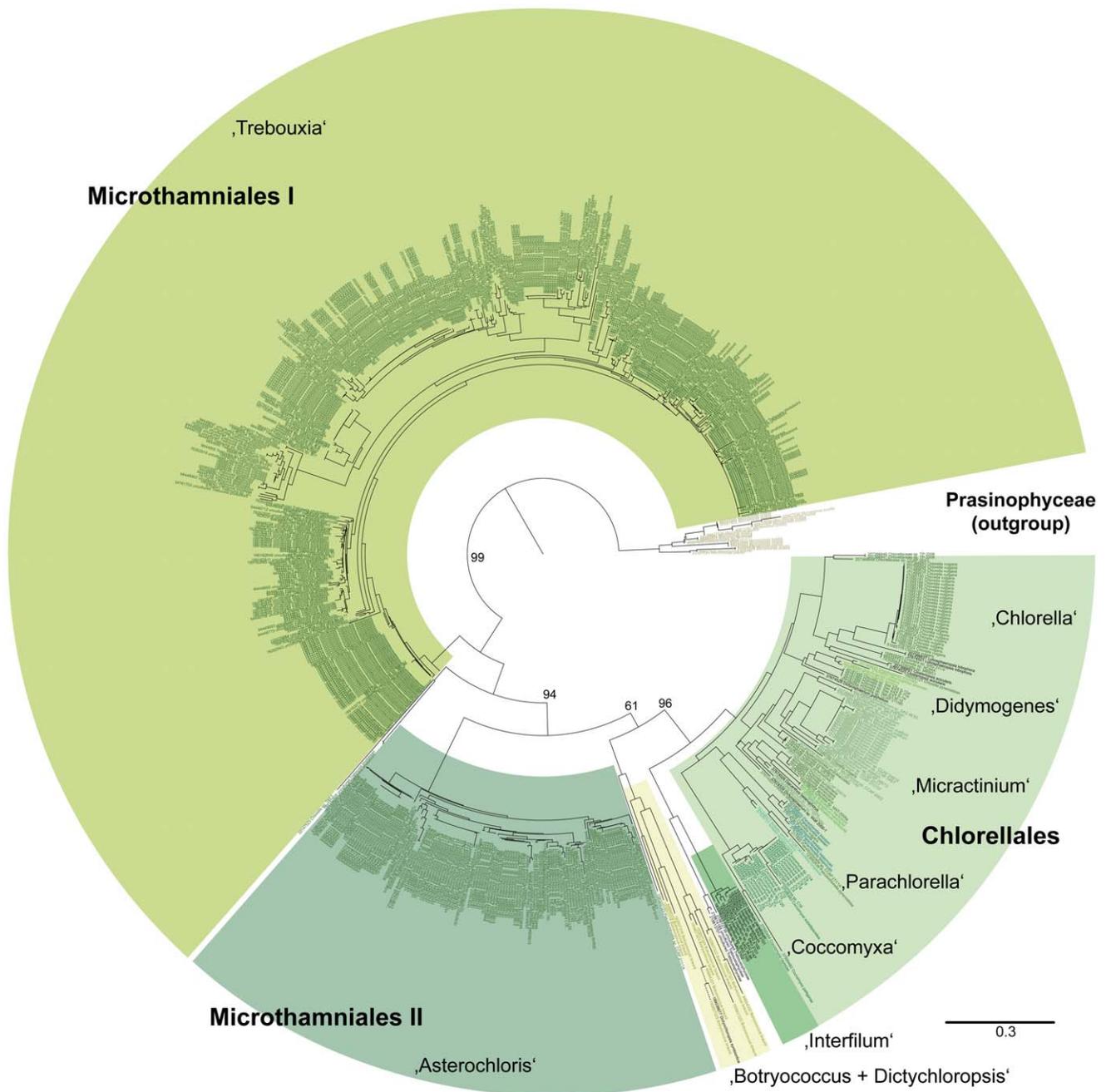


**Figure 1. PNJ tree (with bootstrap values from 100 replicates) for sequence-structure data from the nu ITS2 rRNA gene for a comprehensive sampling of the class Chlorophyceae.** Major taxonomic groups are labelled and highlighted using differential color coding. doi:10.1371/journal.pone.0016931.g001

[Microthamniales I] and the *Asterochloris* alliance [Microthamniales II] and the Chlorellales which includes *Chlorella*, *Parachlorella*, *Coccomyxa*, *Micractinium* and *Didymogenes*. Bootstrap values for these three clades are 99%, 94% and 96%, respectively. Results of a third PNJ analysis (Fig. 3) revealed high bootstrap support for a Bryopsidales clade (92% bootstrap support; *Halimeda* and *Caulerpa* alliances). A *Urospora/Acosiphonia* clade was resolved with 79% bootstrap support. Neither of the two ulvalean alliances (Ulvaes I: *Bolbocoelon*, *Blidingia*, *Monostroma*, *Umbraulva* and one group of *Ulva* taxa; Ulvaes II: a second group of *Ulva* taxa) were robustly resolved. However, the Ulvaes II clade formed a sister group with the *Urospora/Acosiphonia* alliance with 70% bootstrap support. As

with the chlorophycean data (Fig. 1), the trebouxiophycean (Fig. 2) and ulvophycean (Fig. 3) data revealed numerous distinct branches that correspond to putative species.

A composite, phylum-level analysis of ITS2 data (Fig. 4) derived from each of the class-level analyses reveals the same major clades for each class of green algae. However, the branching order of some of these clades differs between class-level and phylum-level analyses. The class level analyses, by default, present each class as monophyletic (Figs. 1–3). In contrast, the phylum level analysis challenges, albeit weakly, the monophyly of each of the classes (Fig. 4). For the Chlorophyceae, the Oedogoniales are allied with Ulvaes I and Chlorellales III (*Coccomyxa*), a subset of the



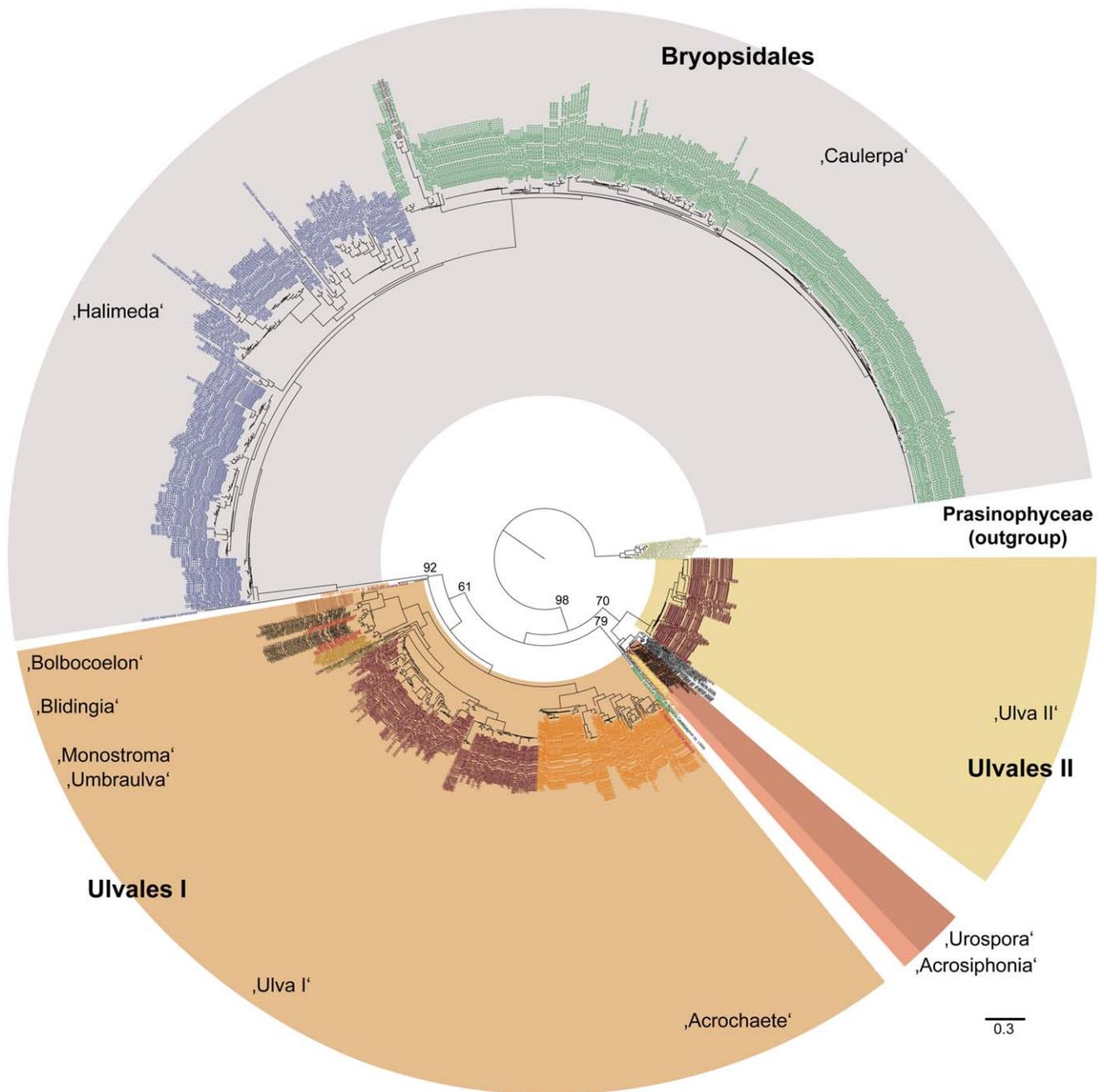
**Figure 2. PNJ tree (with bootstrap values from 100 replicates) for sequence-structure data from the nu ITS2 rRNA gene for a comprehensive sampling of the class Trebouxiophyceae.** Major taxonomic groups are labelled and highlighted using differential color coding.

doi:10.1371/journal.pone.0016931.g002

Sphaeropleales (Sphaeropleales II [Sphaeropleaceae]) are allied with Chlorellales I (*Chlorella*, *Parachlorella*, *Micractinium*, *Didymogenes*, *Diacanthos*, *Closteriopsis*, *Actinastrum*, *Dictyosphaerium*, *Auxenochlorella*, *Lobosphaeropsis*), II (*Pseudochlorella*, *Koliella*), and Microthamniales II (Fig. 4), and Sphaeropleales I (*Desmodesmus* and *Scenedesmus*) is sister to Ulvales I. The Chlamydomonadales are resolved as a monophyletic sister group to the latter alliance (Fig. 4). The Trebouxiophyceae form four distinct, non-monophyletic clades comprising the Microthamniales I, Microthamniales II, Chlorellales III, and Microthamniales II + Chlorellales I + Chlorellales II (Fig. 4). The Ulvophyceae also form four, non-monophyletic clades comprising

the Bryopsidales II (*Caulerpa*), Ulvales + *Urospora/Acosiphonia*, Bryopsidales I (*Halimeda*), and Ulvales I (Fig. 4).

Results from ML analyses of sequence data only (Supplemental files S1, S2, S3, and S4) indicate that the ML approach and the sequence-structure approach using NJ (Figs. 1–4) are topologically congruent in resolving the same major groups of taxa in each of the three class-level analyses and in the phylum-level analysis. However, the relative positions of a number of these groups differ in comparisons of the two distinct analyses for each of the four taxon sets (Chlorophyceae, Trebouxiophyceae, Ulvophyceae and Chlorophyta).



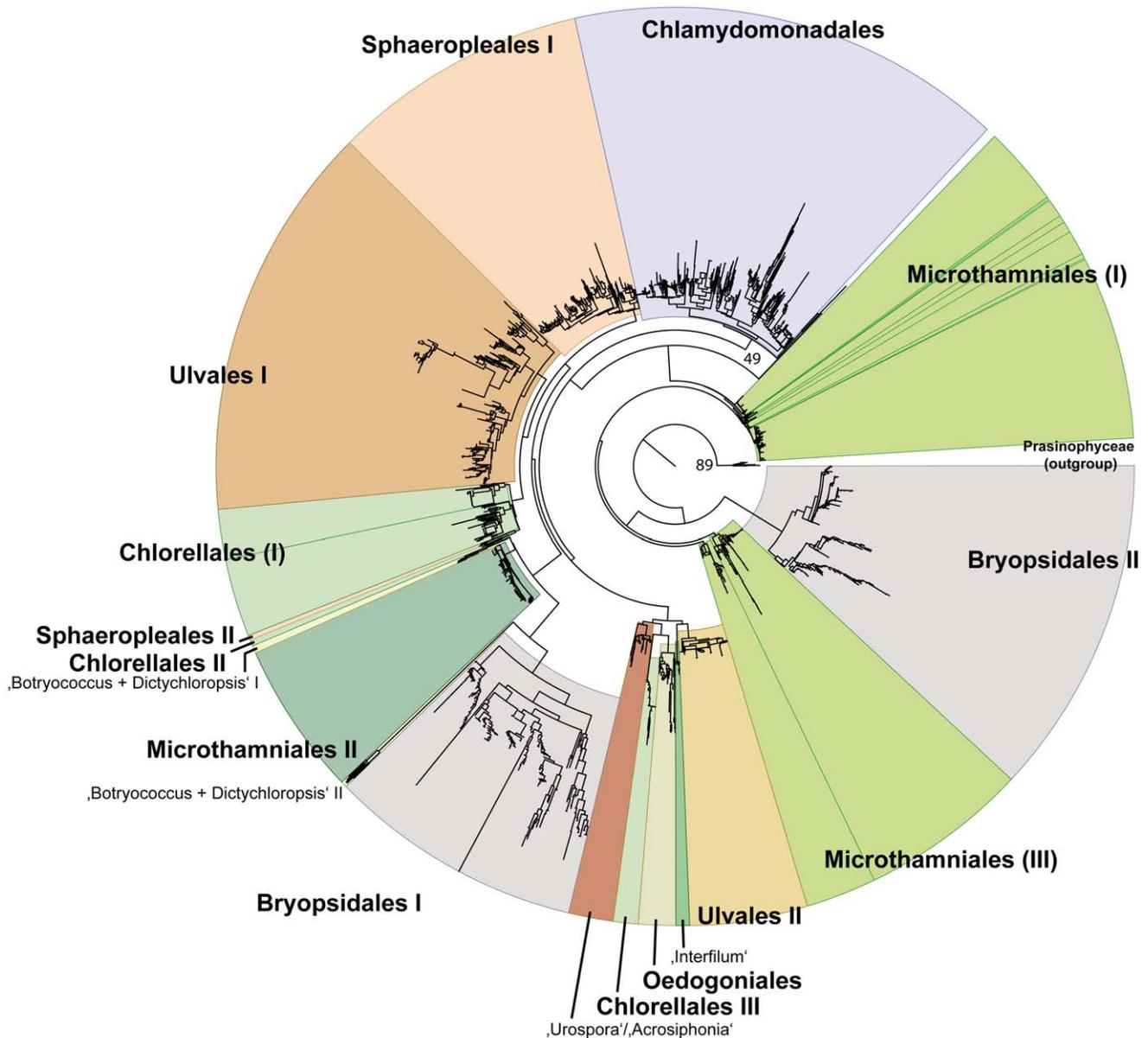
**Figure 3. PNJ tree (with bootstrap values from 100 replicates) for sequence-structure data from the nu ITS2 rRNA gene for a comprehensive sampling of the class Ulvophyceae.** Major taxonomic groups are labelled and highlighted using differential color coding. doi:10.1371/journal.pone.0016931.g003

## Discussion

The independent analyses for each chlorophyten class generally recover phylogenetic signal that is consistent with studies of 18S rRNA [17,18,19,25,26,27,28,29,30,31,32,41,84,85,86,87], 26S rRNA [18,19,25,88,89], *rbcL* [17,55,56,58,90,91,92,93,94] and *atpB* [17,55,56,93].

Topological differences do exist between results with ITS2 data and other data sets. For example, analyses of the ITS2 data for the Chlorophyceae place the Chlamydomonadales as a basal, paraphyletic assemblage in the class (Fig. 1), whereas, both 18S and 26S rRNA data place the Oedogoniales, Chaetophorales and/or Chaetopeltidales as

basal members of the class [19,25]. However, these differences can be attributed to (1) weak support in one or both sets of data, (2) substantial differences in taxon sampling (e.g., no ITS2 data for Chaetopeltidales or Chaetophorales are available), (3) substantial differences in outgroup rooting, or (4) some combination of these influences. In addition to differences between phylogenetic results from ITS2 and other data sets, differences between results from class-level and phylum-level analyses of ITS2 data were also observed. For example, the class level analysis challenges the monophyly of Chlamydomonadales (Fig. 1), but the phylum level analysis (Fig. 4) resolves the order as monophyletic. Again, these differences are not robust and, thus, can be attributed to weak support, taxon sampling error or both.



**Figure 4. PNJ tree for sequence-structure data from the nu ITS2 rRNA gene for a comprehensive sampling of the phylum Chlorophyta.** Major taxonomic groups are labelled and highlighted using differential color coding. doi:10.1371/journal.pone.0016931.g004

The similarities and differences between the results from a NJ analysis using sequence-structure data (Figs. 1–4) and a ML analysis using sequence data alone (Supplemental files S1, S2, S3, and S4) are difficult to interpret given that it is not possible to discriminate between the effects of the model, the method, and the influence of the secondary structure data. Nonetheless, one or more of these factors are influencing the outcome of phylogenetic reconstruction. These observations highlight the need to expand the sequence-structure approach to include character-based methods of tree-building (e.g., ML). A union of sequence-structure analysis with character-based tree-building methods will create new opportunities for hypothesis testing that have the potential to further enhance the use of a sequence-structure approach in standard phylogenetic analyses, as well as, for DNA barcoding.

Our results represent further evidence that the ITS2 data can be aligned for a taxonomically broad set of organisms and that the

alignment yields corroborated alliances of chlorophyten taxa. Most importantly, our results confirm that the analytic procedure does not lead to a loss of signal for the resolution of discrete, species level branches. The behavior of the ITS2 in conjunction with the automated, secondary-structure-based alignment compels us to conclude that the ITS2 data can be used to reconstruct chlorophyten phylogeny. As such, ITS2 has the potential to be a good choice for DNA barcoding in the Chlorophyta.

The remarkable results for the ITS2 gene from chlorophyten taxa raise the question: can these data and analytical approaches be applied to other organisms? Given that ITS2 data already exist for so many disparate groups of organisms, there is little doubt that this protocol could be easily extended to other members of the domain Eukarya. Recent work, which validates the use of ITS2 in barcoding embryophyte plants and animals, strongly supports this assertion [95]. As with most tools, there will be situations that may

negate the utility of the ITS2 for phylogenetic analysis or as a DNA barcode. For example, some parasitic taxa have been identified as possessing substantially shortened ITS2 genes [96]. The ability of the analytical method to recover data from shortened sequences has yet to be tested in a broad taxonomic context.

One of the more problematic issues for the use of ITS2 for phylogenetic reconstruction or as a DNA barcode is that of heterogeneity. As part of the rDNA array, multiple, homogeneous copies of the ITS2 are presumed to exist within all eukaryotic organisms (ironically, making it an excellent barcode candidate due to greater ease of amplification). An assumption of homogeneity, as a consequence of concerted evolution [97,98], may be unrealistic for a number of organisms [99], including at least some chlorophytes [42,100]. Since heterogeneity of the rDNA array is an issue for the use of ITS2 in an ordinary phylogenetic analysis [101], the problem is not merely a product of its use in DNA barcoding. Consequently, the same measures for identifying heterogeneity (cloning, mixing of multiple PCR reactions, see also below) can be applied for use in DNA barcoding. Nonetheless, addressing the problem of heterogeneity in the ITS2 clearly burdens the approach with additional time and expense. However, it is our contention that this extra burden is overshadowed by the significant savings in time and effort through the use of the automated analytical pipeline. No other phylogenetic marker or DNA barcoding candidate is similarly equipped for analytical high-throughput. Furthermore, no other potential barcode exhibits the same level of universality (i.e., in primers for PCR) than the ITS2. Thus, the ITS2 meets criterion one of the recommendations for a standard plant barcode [50]. Furthermore, our current assessment of primary and secondary sequence structure among an exhaustive survey of chlorophyten diversity indicates that ITS2 also meets Criteria Two (bi-directional sequencing with few or no ambiguities) and Three (enables the most species to be distinguished) of the CBOL recommendations [50].

With some notable exceptions [16,51,95,102], the ITS2 gene has largely been shunned by those investigators that are designing or promoting DNA barcodes for the land plants [15,50,103,104]. Concern about the confounding impact of pseudogenes and the potential presence of intraspecific or intra-individual variation (due to differing rates of homogenization of the rDNA tandem array or due to introgression) were cited as reasons for relegating ITS2 to, at best, a supporting role in DNA barcoding for the land plants [15,50,103]. The confounding influence of pseudogenes (from the aberrant secondary structures produced by ITS2 pseudogenes that have accumulated a substantive number of indels as a consequence of the loss of function of the ITS2 gene) can be minimized or eliminated by the use of DMSO during the PCR [104]. In addition, testing for the presence of conserved 5.8S rRNA motifs may be a relatively easy (i.e., amplifying the spacer region to include the 5.8S rRNA adds very little time and investment to an investigation of the ITS2) means of recognizing spacer pseudogenes [105]. At present, there have been no reports of ITS2 pseudogenes in the Chlorophyta, but this is likely to change as more chlorophyten taxa are scrutinized.

As was noted above, the issue of heterogeneity within a species or within an individual has the potential to be more problematic than the confounding issue of ITS2 pseudogenes. Regardless of the source, ITS2 heterogeneity has been deemed a liability for its use as a DNA barcode for the land plants [15,103]. However, life history differences between most Chlorophyta and the embryophytes may account, at least in part, for the antipathy towards the ITS2. Specifically, many Chlorophyta exhibit zygotic meiosis and,

thus, are vegetatively haploid. All embryophytes exhibit sporic meiosis and, thus, are vegetatively diploid. Therefore, the ITS2 in many Chlorophyta behaves more like an organellar gene that exhibits uniparental inheritance. Angiosperms will have two copies from each parent, thus doubling the opportunities for introducing heterogeneity. Introgression, which may play a role in the evolutionary history of a significant number of angiosperm taxa, is often cited as the culprit in producing multiple ITS alleles which, in turn, would likely confound a phylogenetic analysis [103,104]. Except for some marine macrophytes that may exhibit sporic meiosis [106,107,108,109], there seems to be little evidence of introgression [110] that could produce ITS2 heterogeneity in the Chlorophyta. Moreover, the positive results from the most recent and extensive investigations of ITS2 as a DNA barcode for plants [51,95] suggest that the concerns regarding ITS2 may be overstated.

Lastly, we confront the issue of pragmatism. Although their work did not specifically address a DNA barcoding approach, Feliner and Rosseló [101] persuasively argue in favor of a multi-locus line of attack if ITS2 is to be used for assessing organismal diversity. However, as we stated in the Introduction, virtually all of the other candidate genomic targets for DNA barcoding in the Chlorophyta exhibit one or more serious deficiencies. The *rbcL* gene may be able to play a role in DNA barcoding for select groups (e.g., the Chlamydomonadales), but a lack of universal primers coupled with difficult or intractable chlorophyten groups compromises a taxonomically broad use of *rbcL* for the near term. At present, the ITS2 gene is the only viable candidate for immediate use in DNA barcoding for the Chlorophyta. Despite objections to the routine use of ITS2 for land plants, our tests of the ITS2 data demonstrate that this marker resolves major green algal lineages (some with high bootstrap support). Most importantly, our results dramatically illustrate that ITS2 data from unknown chlorophyten organisms can be plugged into a high resolution tool for taxonomic assessment. If the ITS2 gene can serve as a powerful DNA barcode, then this approach has the potential to help address some of the most complex problems in microbial ecology and diversity including analyses of community structure, the paradox of plankton, issues of dispersal and the nature or existence of biogeographical patterns among algal microbes.

## Materials and Methods

All phylogenetic analyses followed the procedure outlined in Schultz and Wolf (2009). Data were obtained (2009/09/30) from the ITS2 Database [65,73,111]. A global, multiple sequence-structure alignment of all available (591) chlorophycean ITS2 sequences with available secondary structures was generated in 4SALE v1.5 [70,112]. Sequences and secondary structures were synchronously aligned, making use of an ITS2 sequence-structure specific scoring matrix [70,112]. Accordingly, alignments were calculated for the Ulvophyceae (938 sequences) and Trebouxiophyceae (741 sequences). Further, a global Chlorophyta tree was calculated that includes all the sequences described above for the individual class-specific trees. For each of the alignments, a set of all *Micromonas* (Prasinophyceae) sequences available in the ITS2 database was used as the outgroup. Based on primary and secondary structure information, phylogenetic relationships were reconstructed by Profile Neighbor-Joining (PNJ) [72], through the use of an ITS2 sequence-structure-specific, General Time Reversible (GTR) substitution model, in ProDistS v0.9.8 [71,74,75]. In addition to the usual Windows/Mac/Linux GUIs, all of the methods described above may be used from a UNIX

command line shell and thus be incorporated in any type of automated scripts. The complete procedure of data acquisition, alignment calculation and tree reconstruction took less than one hour of computational time for the three class-specific trees and 3.5 h for the complete Chlorophyta tree on a conventional 2.0 GHz single core computer.

In a second manual step we obtained bootstrap support values (Felsenstein, 1985) for the major taxonomic clades within the trees. For this step, manual profiles were set in ProfDistS with the Cartoon2Profile tool (<http://profdist.bioapps.biozentrum.uni-wuerzburg.de/cgi-bin/index.php?section=cart2prof>), after rooting and visualizing the distance trees with FigTree v1.2.3 [113]. Cartoon2Profile is a Perl script that converts cartoons as set in FigTree into a ProfDistS compatible profile file. Cartoon2Profile has been explicitly developed for this study, but may be used for any investigation that uses FigTree and ProfDistS. Calculation of bootstrap values with these profiles required less than 10 minutes of computational time using a desktop computer. We visualized a concatenated topology of the three class-specific trees in a hyperbolic tree based on the HyperGeny tree browser (<http://bioinformatics.psb.ugent.be/hypergeny>). The hyperbolic tree is publicly available as a supplement to this study at the ITS2-Database Supplements Page and at <http://hypertree.bioapps.biozentrum.uni-wuerzburg.de>.

At the present time, we are aware of no sequence-structure approach using individual secondary structures that can accommodate treeing methods other than the algorithmic approach of NJ. However, in order to provide an alternative context in which to evaluate the sequence-structure trees, a second set of analyses of nucleotide data only for each of the three classes of green algae and a composite analysis for the Chlorophyta was completed. These analyses employed an approximately maximum likelihood approach (ML) using FastTree 2 [114] with default settings. The sequence alignment was determined using Clustal [115].

## Supporting Information

**File S1** Phylogenetic tree (in Newick format) from ML analysis (using FastTree 2) of sequence data only from the same set of

chlorophycean taxa used in the sequence-structure analysis. This file is best viewed using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

(TREE)

**File S2** Phylogenetic tree (in Newick format) from ML analysis (using FastTree 2) of sequence data only from the same set of trebouxiophycean taxa used in the sequence-structure analysis. This file is best viewed using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

(TREE)

**File S3** Phylogenetic tree (in Newick format) from ML analysis (using FastTree 2) of sequence data only from the same set of ulvophycean taxa used in the sequence-structure analysis. This file is best viewed using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

(TREE)

**File S4** Phylogenetic tree (in Newick format) from ML analysis (using FastTree 2) of sequence data only from the same set of chlorophycean taxa used in the sequence-structure analysis. This file is best viewed using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

(TREE)

## Acknowledgments

We thank Danica Sutherland for comments on an early draft of this manuscript. We also gratefully acknowledge comments from two anonymous reviewers that allowed us to substantively improve the manuscript.

## Author Contributions

Conceived and designed the experiments: MW. Performed the experiments: MW AK CK FF BM. Analyzed the data: AK CK FF BM MW MAB. Contributed reagents/materials/analysis tools: MW MAB. Wrote the paper: MAB MW.

## References

- Wheeler QD (2005) Losing the plot: DNA “barcodes” and taxonomy. *Cladistics* 21: 405–407.
- Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* 54.
- Ebach MC, Holdrege C (2005) More Taxonomy, Not DNA Barcoding. *BioScience* 55: 823–824.
- DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 360: 1905–1916.
- Smith VS (2005) DNA barcoding: Perspectives from a “Partnerships for Enhancing Expertise in Taxonomy” (PEET) debate. *Systematic Biology* 54: 841–844.
- Holdrege C, Ebach MC (2006) Response from Holdrege and Ebach: What about Taxa? *Bioscience* 56: 93.
- Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 360: 1805–1811.
- Hollingsworth ML, Clark AA, Forrest LL, Richardson J, Pennington RT, et al. (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Molecular Ecology Resources* 9: 439–457.
- Jakupciak JP, Colwell RR (2009) Biological agent detection technologies. *Molecular Ecology Resources* 9: 51–57.
- Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, et al. (2005) Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 360: 1935–1943.
- Blaxter ML (2004) The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 359: 669–679.
- Engelmann JC, Rahmann S, Wolf M, Schultz J, Fritzilas E, et al. (2009) Modeling cross-hybridization on phylogenetic DNA microarrays increases the detection power of closely related species. *Molecular Ecology Resources* 9: 83–93.
- Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. *Systematic Biology* 54: 852–859.
- Seberg O, Petersen G (2009) How many loci does it take to DNA barcode a *Crocus*? *PLoS ONE* 4: e4598.
- Chase MW, Fay MF (2009) Barcoding of plants and fungi. *Science* 325: 682–683.
- Wolf M, Schultz J (2009) ITS better than its reputation. *Science (E-Letter)*, 10 December 2009.
- Buchheim MA, Kirkwood A, Buchheim JA, Verghese B, Henley WJ (2010) Hypersaline soil supports a diverse community of *Dunaliella* (Chlorophyceae). *Journal of Phycology* 46: 1038–1047.
- Buchheim MA, Buchheim JA, Carlson T, Braband A, Hepperle D, et al. (2005) Phylogeny of the Hydrodictyaceae (Chlorophyceae): Inferences from rDNA data. *Journal of Phycology* 41: 1039–1054.
- Buchheim MA, Buchheim JA, Carlson T, Kugrens P (2002) Phylogeny of *Lobocharacium* (Chlorophyceae) and allies: A study of 18S and 26S rDNA data. *Journal of Phycology* 38: 376–383.
- Buchheim MA, Buchheim JA, Chapman RL (1997) Phylogeny of the VLE-14 *Chlamydomonas* (Chlorophyceae) group. A study of 18S rRNA gene sequences. *Journal of Phycology* 33: 1024–1030.
- Buchheim MA, Buchheim JA, Chapman RL (1997) Phylogeny of *Chloromonas* (Chlorophyceae): A study of 18S ribosomal RNA gene sequences. *Journal of Phycology* 33: 286–293.
- Buchheim MA, Chapman RL (1991) Phylogeny of the colonial green flagellates: A study of 18S and 26S ribosomal RNA sequence data. *BioSystems* 25: 85–100.

23. Buchheim MA, Chapman RL (1992) Phylogeny of *Carteria* (Chlorophyceae) inferred from molecular and organismal data. *Journal of Phycology* 28: 362–374.
24. Buchheim MA, Lemieux C, Otis C, Gutell RR, Chapman RL, et al. (1996) Phylogeny of the Chlamydomonadales (Chlorophyceae): A comparison of ribosomal RNA gene sequences from the nucleus and the chloroplast. *Molecular Phylogenetics and Evolution* 5: 391–402.
25. Buchheim MA, Michalopoulos EA, Buchheim JA (2001) Phylogeny of the Chlorophyceae with special reference to the Sphaeropleales: A study of 18S and 26S rDNA data. *Journal of Phycology* 37: 819–835.
26. Wolf M, Buchheim M, Hegewald E, Krienitz L, Hepperle D (2002) Phylogenetic position of the Sphaeropleaceae (Chlorophyta). *Plant Systematics and Evolution* 230: 161–171.
27. Krienitz L, Hegewald E, Hepperle D, Wolf M (2003) The systematics of coccoid green algae: 18S rRNA gene sequence data versus morphology. *Biologia* 58: 437–446.
28. Krienitz L, Hegewald EH, Hepperle D, Huss VAR, Rohr T, et al. (2004) Phylogenetic relationship of *Chlorella* and *Parachlorella gen. nov.* (Chlorophyta, Trebouxiophyceae). *Phycologia* 43: 529–542.
29. Wolf M, Hepperle D, Krienitz L (2003) On the phylogeny of *Radiococcus*, *Planktosphaeria* and *Schizochlamydelia* (Radiococcaceae, Chlorophyta). *Biologia* 58: 759–765.
30. Wolf M, Hegewald E, Hepperle D, Krienitz L (2003) Phylogenetic position of the Golenkiniaceae (Chlorophyta) as inferred from 18S rDNA sequence data. *Biologia* 58: 433–436.
31. Wolf M, Krienitz L, Hepperle D (2002) Phylogenetic position of *Actinastrum hantzschii* Lagerheim 1882 (Chlorophyta, Trebouxiophyceae). *Algalogical Studies* 104: 59–67.
32. Hegewald E, Wolf M (2003) Phylogenetic relationships of *Scenedesmus* and *Acutodesmus* (Chlorophyta, Chlorophyceae) as inferred from 18S rDNA and ITS-2 sequence comparisons. *Plant Systematics and Evolution* 241: 185–191.
33. Hegewald E, Hanagata N (2000) Phylogenetic studies on Scenedesmeae (Chlorophyta). *Archiv für Hydrobiologie Supplement* 136: 29–49.
34. Hegewald E, Wolf M, Keller A, Friedl T, Krienitz L (2010) ITS2 sequence-structure phylogeny in the Scenedesmeae with special reference to *Coelastrum* (Chlorophyta, Chlorophyceae), including the new genera *Comasiella* and *Pectinodesmus*. *Phycologia* 49: 325–335.
35. Müller T, Philippi N, Dandekar T, Schultz J, Wolf M (2007) Distinguishing species. *RNA* 13: 1469–1472.
36. An SS, Friedl T, Hegewald E (1999) Phylogenetic relationships of *Scenedesmus* and *Scenedesmus*-like coccoid green algae as inferred from ITS-2 rDNA sequence comparisons. *Plant Biology* 1: 418–428.
37. Bakker FT, Olsen JL, Stam WT (1995) Evolution of nuclear rDNA ITS sequences in the *Cladophora albidalsericea* clade (Chlorophyta). *Journal of Molecular Evolution* 40: 640–651.
38. Coat G, Dion P, Noailles MC, De Reviers B, Fontaine JM, et al. (1998) *Ulva americana* (Ulvales, Chlorophyta) from the coasts of Brittany (France): II. Nuclear rDNA ITS sequence analysis. *European Journal of Phycology* 33: 81–86.
39. Coleman AW (2001) Biogeography and speciation in the *Pandorina/Volvulina* (Chlorophyta) superclade. *Journal of Phycology* 37: 836–851.
40. Fabry S, Koehler A, Coleman AW (1999) Intraspecific analysis: Comparison of ITS sequence data and gene intron sequence data with breeding data for a worldwide collection of *Gonium pectorale*. *Journal of Molecular Evolution* 48: 94–101.
41. Lewis LA, Flechtner VR (2004) Cryptic species of *Scenedesmus* (Chlorophyta) from desert soil communities of Western North America. *Journal of Phycology* 40: 1127–1137.
42. Pillmann A, Woolcott GW, Olsen JL, Stam WT, King RJ (1997) Inter- and intraspecific genetic variation in *Caulerpa* (Chlorophyta) based on nuclear rDNA ITS sequences. *European Journal of Phycology* 32: 379–386.
43. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, et al. (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon* 56: 295–299.
44. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, et al. (2005) Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 360: 1889–1895.
45. Fazekas AJ, Burgess KS, Kesanakurthi PR, Graham SW, Newmaster SG, et al. (2008) Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. *PLoS ONE* 3: e2802.
46. Kress WJ, Erickson DL (2007) A Two-Locus Global DNA Barcode for Land Plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. *PLoS One* 2: Article No.: e508.
47. Newmaster SG, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. *Canadian Journal of Botany* 84: 335–341.
48. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* 102: 8369–8374.
49. Ledford H (2008) Botanical identities: DNA barcoding for plants comes a step closer. *Nature* 451: 616.
50. CBOL Plant Working Group (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 12794–12797.
51. Chen S, Yao H, Han J, Liu C, Song J, et al. (2010) Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. *PLoS ONE* 5: e8613.
52. Sanders ER, Karol KG, McCourt RM (2003) Occurrence of *matK* in a *trnK* group II intron in Charophyte green algae and phylogeny of the Characeae. *American Journal of Botany* 90: 628–633.
53. Lemieux C, Otis C, Turmel M (2000) Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature* 403: 649–652.
54. Turmel M, Otis C, De Cambiaire J-C, Pombert J-F, Lemieux C (2002) The complete chloroplast DNA sequence of *Chlorokybus atmophyticus*: evidence that charophycean green algae from an early-diverging lineage adapted to terrestrial life. *Botany* 2002. Madison, WI.
55. Nozaki H (2001) Chloroplast multigene phylogeny and systematics of the advanced genera of the Volvocaceae (Chlorophyceae). *Phycologia* 40: 37.
56. Nozaki H, Misawa K, Kajita T, Kato M, Nohara S, et al. (2000) Origin and evolution of the colonial Volvocales (Chlorophyceae) as inferred from multiple, chloroplast gene sequences. *Molecular Phylogenetics and Evolution* 17: 256–268.
57. Nozaki H, Motomi I, Ryosuke S, Hidenobu U, Makoto MW, et al. (1995) Phylogenetic relationships within the colonial Volvocales (Chlorophyta) inferred from *rbcL* gene sequence data. *Journal of Phycology* 31: 970–979.
58. Nozaki H, Ohta N, Takano H, Watanabe MM (1999) Reexamination of phylogenetic relationships within the colonial Volvocales (Chlorophyta): An analysis of *atpB* and *rbcL* gene sequences. *Journal of Phycology* 35: 104–112.
59. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR Protocols*. San Diego: Academic Press. pp 315–322.
60. Mai JC, Coleman AW (1997) The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. *Journal of Molecular Evolution* 44: 258–271.
61. Coleman AW (2003) ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends in Genetics* 19: 370–375.
62. Coleman AW (2007) Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Research* 35: 3322–3329.
63. Hershkovitz MA, Lewis LA (1996) Deep-level diagnostic value of the rDNA-ITS region. *Molecular Biology and Evolution* 13: 1276–1295.
64. Schultz J, Maisel S, Gerlach D, Müller T, Wolf M (2005) A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA* 11: 361–364.
65. Schultz J, Müller T, Achtziger M, Seibel PN, Dandekar T, et al. (2006) The internal transcribed spacer 2 database - a web server for (not only) low level phylogenetic analyses. *Nucleic Acids Research* 34: W704–W707.
66. Schultz J, Wolf M (2009) ITS2 sequence-structure analysis in phylogenetics: A how-to manual for molecular systematics. *Molecular Phylogenetics and Evolution* 52: 520–523.
67. Wolf M, Achtziger M, Schultz J, Dandekar T, Müller T (2005) Homology modeling revealed more than 20,000 rRNA internal transcribed spacer 2 (ITS2) secondary structures. *RNA* 11: 1616–1623.
68. Keller A, Förster F, Müller T, Dandekar T, Schultz J, et al. (2010) Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. *Biology Direct* 5: 4.
69. Wiemers M, Keller A, Wolf M (2009) ITS2 secondary structure improves phylogeny estimation in a radiation of blue butterflies of the subgenus *Agrodiaetus* (Lepidoptera: Lycaenidae: Polymmatinae). *BMC Evolutionary Biology* 9: 300.
70. Seibel PN, Müller T, Dandekar T, Schultz J, Wolf M (2006) 4SALE - A tool for synchronous RNA sequence and secondary structure alignment and editing. *BMC Bioinformatics* 7: 498.
71. Wolf M, Ruderisch B, Dandekar T, Schultz J, Müller T (2008) ProfDistS: (profile-) distance based phylogeny on sequence-structure alignments. *Bioinformatics* 24: 2401–2402.
72. Müller T, Rahmann S, Dandekar T, Wolf M (2004) Accurate and robust phylogeny estimation based on profile distances: a study of the Chlorophyceae (Chlorophyta). *BMC Evolutionary Biology* 4: 20.
73. Koetschan C, Förster F, Keller A, Schleicher T, Ruderisch B, et al. (2010) The ITS2 Database III - sequences and structures for phylogeny. *Nucleic Acids Research* 38: D275–279.
74. Friedrich J, Dandekar T, Wolf M, Müller T (2005) ProfDist: a tool for the construction of large phylogenetic trees based on profile distances. *Bioinformatics* 21: 2108–2109.
75. Rahmann S, Müller T, Dandekar T, Wolf M (2006) Efficient and robust analysis of large phylogenetic datasets. In: Hsu H-H, ed. *Advanced Data Mining Technologies in Bioinformatics*. Hershey, PA: Idea Group, Inc. pp 104–117.
76. Angeler DG, Schagerl M, Coleman AW (1998) Does the infraspecific disjunct occurrence of the carotenoid loroxanthin reflect phylogenetic relationships within syngens of *Pandorina morum* (Volvocales, Chlorophyta)? *Biologia* 53: 567–575.

77. Angeler DG, Schagerl M, Coleman AW (1999) Phylogenetic relationships among isolates of *Eudorina* species (Volvocales, Chlorophyta) inferred from molecular and biochemical data. *Journal of Phycology* 35: 815–823.
78. Coleman AW, Mai JC (1997) Ribosomal DNA ITS-1 and ITS-2 sequence comparisons as a tool for predicting genetic relatedness. *Journal of Molecular Evolution* 45: 168–177.
79. Coleman AW (2009) Is there a molecular key to the level of “biological species” in eukaryotes? A DNA guide. *Molecular Phylogenetics and Evolution* 50: 197–203.
80. Keller A, Schleicher T, Förster F, Ruderich B, Dandekar T, et al. (2008) ITS2 data corroborate a monophyletic chlorophycean DO-group (Sphaeropleales). *BMC Evolutionary Biology* 8: 218.
81. McManus HA, Lewis LA (2005) Molecular phylogenetics, morphological variation and colony-form evolution in the family Hydrodictyaceae (Sphaeropleales, Chlorophyta). *Phycologia* 44: 582–595.
82. Pockock T, Lachance M-A, Pröschold T, Priscu JC, Kim SS, et al. (2004) Identification of a psychrophilic green alga from Lake Bonney Antarctica: *Chlamydomonas raudensis* Ettl. (UWO 241) Chlorophyceae. *Journal of Phycology* 40: 1138–1148.
83. van Hanne E, Fink P, Lurling M (2002) A revised secondary structure model for the internal transcribed spacer 2 of the green algae *Scenedesmus* and *Desmodesmus* and its implication for the phylogeny of these algae. *European Journal of Phycology* 37: 203–208.
84. Friedl T (1996) Evolution of the polyphyletic genus *Pleurastrum* (Chlorophyta): Inferences from nuclear-encoded ribosomal DNA sequences and motile cell ultrastructure. *Phycologia* 35: 456–469.
85. Hepperle D, Hegewald E, Krienitz L (2000) Phylogenetic position of the Oocystaceae (Chlorophyta). *Journal of Phycology* 36: 590–595.
86. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on exhaustive 18S rRNA phylogenetic analyses. *Molecular Phylogenetics and Evolution* 48: 281–291.
87. Pröschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). I. Emendation of *Chlamydomonas* Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and *Lobochlamys* gen. nov. *Protist* 152: 265–300.
88. Mei H, Liu G-X, Hu Z-Y (2007) Phylogenetic studies of Oedogoniales (Chlorophyceae, Chlorophyta) based on 28S rDNA sequences. *Acta Hydrobiologica Sinica* 31: 492–498.
89. Leliaert F, Rousseau F, de Reviere B, Coppejans E (2003) Phylogeny of the Cladophorophyceae (Chlorophyta) inferred from partial LSU rRNA gene sequences: Is the recognition of a separate order Siphonocladales justified? *European Journal of Phycology* 38: 233–246.
90. Nakazawa A, Yamada T, Nozaki H (2004) Taxonomic study of *Asterococcus* (Chlorophyceae) based on comparative morphology and *rbcl* gene sequences. *Phycologia* 43: 711–721.
91. Loughnane CJ, McIvor LM, Rindi F, Stengel DB, Guiry MD (2008) Morphology, *rbcl* phylogeny and distribution of distromatic *Ulva* (Ulvophyceae, Chlorophyta) in Ireland and southern Britain. *Phycologia* 47: 416–429.
92. Nozaki H (2003) Origin and evolution of the genera *Pleodorina* and *Volvox* (Volvocales). *Biologia* 58: 425–431.
93. Nozaki H, Misumi O, Kuroiwa T (2003) Phylogeny of the quadriflagellate Volvocales (Chlorophyceae) based on chloroplast multigene sequences. *Molecular Phylogenetics and Evolution* 29: 58–66.
94. Zechman FW (2003) Phylogeny of the Dasycladales (Chlorophyta, Ulvophyceae) based on analyses of RUBISCO large subunit (*rbcl*) gene sequences. *Journal of Phycology* 39: 819–827.
95. Yao H, Song J, Liu C, Luo K, Han J, et al. (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS One* 5: e13102.
96. Edlind TD, Sharetzky C, Cha ME (1990) Ribosomal RNA of the primitive eukaryote *Giardia lamblia*: large subunit domain I and potential processing signals. *Gene* 96: 289–293.
97. Zimmer EA, Martin SL, Beverly SM, Kan YW, Wilson AC (1980) Rapid duplication and loss of genes coding for the  $\alpha$  chains of hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America* 77: 2518–2522.
98. Arnheim N (1983) Concerted evolution of multigene families. In: Nei M, Koehn M, eds. *Evolution of Genes and Proteins*. Sunderland, MA: Sinauer Associates. pp 38–61.
99. Harpke D, Peterson A (2006) Non-concerted ITS evolution in *Mammillaria* (Cactaceae). *Molecular Phylogenetics and Evolution* 41: 579–593.
100. Famà P, Olsen JL, Stam WT, Procaccini G (2000) High levels of intra- and inter-individual polymorphism in the rDNA ITS1 of *Caulerpa racemosa* (Chlorophyta). *European Journal of Phycology* 35: 349–356.
101. Feliner GN, Roselló JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* 44: 911–919.
102. Gile GH, Stern RF, James ER, Keeling PJ (2010) DNA Barcoding of the Chlorarachniophytes using nucleomorph ITS sequences. *Journal of Phycology* 46: 743–750.
103. Chase MW, Fay MF (2009) Response to J. Schultz and M. Wolf's E-Letter. *Science Online*.
104. Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, et al. (2003) Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). *Annals of Botany (London)* 92: 107–127.
105. Harpke D, Peterson A (2008) 5.8S motifs for the identification of pseudogenic ITS regions. *Botany* 86: 300–305.
106. Kapraun DF (1993) Karyology of marine green algae. *Phycologia* 32: 1–21.
107. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogamic Botany* 4: 410–418.
108. Kapraun DF, Buratti JR (1998) Evolution of genome size in the Dasycladales (Chlorophyta) as determined by DAPI cytophotometry. *Phycologia* 37: 176–183.
109. Durand C, Manuel M, Boudouresque CF, Meinesz A, Verlaque M, et al. (2002) Molecular data suggest a hybrid origin for the invasive *Caulerpa racemosa* (Caulerpales, Chlorophyta) in the Mediterranean Sea. *Journal of Evolutionary Biology* 15: 122–133.
110. Verbruggen H, De Clerck O, Schils T, Kooistra WHCF, Coppejans E (2005) Evolution and phylogeography of *Halimeda* section *Halimeda* (Bryopsidales, Chlorophyta). *Molecular Phylogenetics and Evolution* 37: 789–803.
111. Selig C, Wolf M, Müller T, Dandekar T, Schultz J (2008) The ITS2 Database II: homology modeling RNA structure for molecular systematics. *Nucleic Acids Research* 36: D377–380.
112. Seibel P, Müller T, Dandekar T, Wolf M (2008) Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. *BMC Research Notes* 1: 91.
113. Rambaut A (2009) FigTree v1.2.3. Program distributed by the author.
114. Price MN, Dehal PS, Arkin AP (2010) FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490.
115. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.