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

Mark A. Buchheim1*, Alexander Keller2, Christian Koetschan2, Frank Förster2, Benjamin Merget2, Matthias Wolf2

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 chlorophyta 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 chlorophyta 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.

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 chlorophyta 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 [GO]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

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

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

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.
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 chlorophyta barcoding. However, at present, only the 5.8S rRNA and ITS1 rRNA genes have been studied in more than fifty chlorophyta 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 chlorophyta 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 Chlorophyta (Viridiplantae). The final stage of the analytical methods with both primary and secondary structural analyses 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 rDNA data for the class Chlorophyceae yielded a tree (Fig. 1) that resolved data representing the orders Oedogoniales (*Oedogonium, Balnochaete and Oedocladium*), Sphaero- pleales (*Desmodesmus, Stenodesmus, Atractomorpha and Sphaeroplea*), and Chlamydomonadales/Volvocales (*Chlamydomonas* [three non-monophyletic clades], *Famaguetiella, Pandorina, Eudorina, Astrophyton*, *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 rDNA 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 prasinophycean data as the outgroup. The PNJ analysis resolved three principal clades of trebouxiophycean taxa (Fig. 2) that correspond to two sets of microhminalean taxa (*the Trebouxia alliance*...
[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/Acrosiphonia clade was resolved with 79% bootstrap support. Neither of the two ulvalean alliances (Ulvales I: Bolbocelion, Blidingia, Monostroma, Umbraula and one group of Ulva taxa; Ulvales II: a second group of Ulva taxa) were robustly resolved. However, the Ulvales II clade formed a sister group with the Urospora/Acrosiphonia 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 Ulvales I and Chlorellales III (Coccomyxa), a subset of the

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
Sphaeropleales (Sphaeropleales II [Sphaeropleaceae]) are allied with Chlorellales I (Chlorella, Parachlorella, Micractinium, Didymogenses, Diaclathus, 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 + Urospor/Acrosporina, 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).
Discussion

The independent analyses for each chlorophytan 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], \textit{rbcL} [17,55,56,90,91,92,93,94] and \textit{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 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
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 chlorophytan 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 chlorophytan 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 chlorophytan 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 phylogenetics 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 chlorophyta 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 chlorophytan 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 embroyophytes 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. Introggression, 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 Rossello [101] persuasively argue in favor of a multilocus 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 chlorophyta 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 chlorophyta 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 Chorophyta 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 ProtDistS 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, 1983) 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 Chustal [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 trebouxioiphycean 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. Contributed reagents/materials/analysis tools: MW MAB. Wrote the paper: MAB MW.

References

1. Wheeler QD (2005) Losing the plot: DNA “barcodes” and taxonomy. Cladistics 21: 405–407.
2. Will KW, Musher BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. Systematic Biology 54: 193-195.
3. Ebach MC, Holdrege C (2005) More Taxonomy, Not DNA Barcoding. BioScience 55: 823-824.
4. DeSalle R, Egan MG, Siddall M (2005) The unhybrid trinity: taxonomy, species delimitation and DNA barcoding. Philosophical Transactions of the Royal Society of London B: Biological Sciences 360: 1905–1916.
5. Smith VS (2005) DNA barcoding: Perspectives from a “Partnerships for Enhancing Expertise in Taxonomy” (PEET) debate. Systematic Biology 54: 841–844.
6. Holdrege C, Ebach MC (2006) Response from Holdrege and Ebach: What about Taxa? Bioscience 56: 93.
7. Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society of London B: Biological Sciences 360: 1083–1087.
8. 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.
9. Jakopiczak PJ, Colwell RR (2009) Biological agent detection technologies. Molecular Ecology Resources 9: 51–57.
10. 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.
11. Blaxter ML (2004) The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society of London B: Biological Sciences 359: 669–679.
12. Engelmann JC, Rahman S, Wolf M, Schultz J, Fritzlaus 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.
13. Hebert PDM, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852–859.
14. Seborg O, Petersen G (2009) How many loci does it take to DNA barcode a Coccus? PLoS ONE 4: e4598.
15. Chase MW, Fay MF (2009) Barcoding of plants and fungi. Science 325: 602–603.
16. Wolf M, Schultz J (2009) ITS better than its reputation. Science (E-Letter, 10 December 2009).
17. Buchheim MA, Kirkwood A, Buchheim JA, Vergheze B, Henley WJ (2010) Hypersaline soil supports a diverse community of Dunaliella (Chlorophyceae). Journal of Phycology 46: 1038–1047.
18. 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.
19. Buchheim MA, Buchheim JA, Carlson T, Krogmann P (2002) Phylogeny of Laboschira (Chlorophyceae) and allies: A study of 18S and 26S rDNA data. Journal of Phycology 38: 376–383.
20. 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.
21. Buchheim MA, Buchheim JA, Chapman RL (1997) Phylogeny of Chloromonas (Chlorophyceae): A study of 18S ribosomal RNA gene sequences. Journal of Phycology 33: 289–293.
22. 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.
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, Lu C, Song J, et al. (2010) Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. PLoS ONE 5: e6163.

52. Sanders ER, Karel KG, McCourt RM (2003) Occurrence of mallot in a tree group II intron in Charophyte green algae and phylogeny of the Characeae. American Journal of Botany 90: 620–631.

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 Cambraia JC, Piumbert J-F, Lemieux C (2002) The complete chloroplast DNA sequence of Chlorella amphitricha: evidence that chlorellae green algae from an early-diverging lineage adapted to terrestrial life. Botany 2002. Madison, WI.

55. Nozaki H (2001) Chloroplast mtDNA phylogeny and systematics of the superfamilial genera of the Volvocales (Chlorophyceae). Phycologia 40: 315–322.

56. Nozaki H, Motomi I, Hidemori U, Makoto MW, et al. (1999) Phylogenetic relationships within the colonial Volvocales (Chlorophyta) inferred from ITS rDNA gene sequence data. Journal of Phycology 35: 970–979.

57. Nozaki H, Ohita N, Takahashi H, Watanabe MM (1999) Rerexamination of phylogenetic relationships within the colonial Volvocales (Chlorophyta): An analysis of ap2 and rL gene sequences. Journal of Phycology 35: 104–112.

58. 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.

59. Mai JC, Coleman AW (1997) The internal transcribed spacer 2 accommodates a common secondary structure in green algae and flowering plants. Journal of Molecular Evolution 44: 250–271.

60. Coleman AW (2003) ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends in Genetics 19: 370–375.

61. Coleman AW (2007) Pain-er-oyte ITS2 homologues revealed by RNA secondary structure. Nucleic Acids Research 35: 3322-3329.

62. Hershkovitz MA, Lewis LA (1996) Deep-level diagnostic value of the rDNA ITS region. Molecular Biology and Evolution 13: 1276–1295.

63. Schultz J, Mansel S, Gerlach D, Müller T, Wolf M (2005) A common core of secondary structures in the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 11: 361–364.

64. 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.

65. Schulz 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, 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.

67. Wolf M, Achtziger M, Schultz J, Dandekar T, Müller T (2005) Homology of Phylogenetic position of the Sphaeropleaceae (Chlorophyta) superclade. Journal of Phycology 37: 836–851.

68. Wolf M, Achtziger M, Schultz J, Dandekar T, Müller T (2005) Homology of Phylogenetic position of the Sphaeropleaceae (Chlorophyta) superclade. Journal of Phycology 37: 836–851.

69. Wolf M, Achtziger M, Schultz J, Dandekar T, Müller T (2005) Homology of Phylogenetic position of the Sphaeropleaceae (Chlorophyta) superclade. Journal of Phycology 37: 836–851.

70. Wolf M, Achtziger M, Schultz J, Dandekar T, Müller T (2005) Homology of Phylogenetic position of the Sphaeropleaceae (Chlorophyta) superclade. Journal of Phycology 37: 836–851.
77. Angerler DG, Schaefer M, Coleman AW (1999) Phylogenetic relationships among isolates of *Eutrochium* 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, Ruderisch B, Dandelkar T, et al. (2008) ITS2 data corroborate a monophyletic chlorophycean DOG-group (Sphaeropleales). *BMC Evolutionary Biology* 8: 218.
81. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyta). *Biologia* 58: 425–431.
82. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogenetics, morphological variation and colony-form evolution in the family Hydrodictyaceae (Sphaeropleales, Chlorophyta). *Phylogenetica* 44: 502–505.
83. Proschold T, Lachance MA, Proschold T, Priscu JC, Kim SS, et al. (2004) Identification of a psychrophilic green alga from Lake Bonney Antarctica: *Chlamydomonas nausenic* Ehrl. (UWO 241). *Chlorophyceae*. *Journal of Phycology* 40: 1138–1148.
84. 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.
85. Hepperle D, Hegewald E, Krienitz L (2000) Phylogenetic position of the Oogamochlamys gen nov. (Papillulophyceae, Chlorophyta) in Ireland and southern Britain. *Phycologia* 47: 416–429.
86. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyta). *Biologia* 58: 425–431.
87. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogenetics, morphological variation and colony-form evolution in the family Hydrodictyaceae (Sphaeropleales, Chlorophyta). *Phylogenetica* 44: 502–505.
88. Proschold T, Lachance MA, Proschold T, Priscu JC, Kim SS, et al. (2004) Identification of a psychrophilic green alga from Lake Bonney Antarctica: *Chlamydomonas nausenic* Ehrl. (UWO 241). *Chlorophyceae*. *Journal of Phycology* 40: 1138–1148.
89. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Oogamochlamys* Ehrl. *Phycologia* 58: 425–429.
90. 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.
91. 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.
92. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
93. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
94. Hepperle D, Hegewald E, Krienitz L (2000) Phylogenetic position of the Oogamochlamys gen nov. (Papillulophyceae, Chlorophyta) in Ireland and southern Britain. *Phycologia* 47: 416–429.
95. 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.
96. Edlind TD, Sharetsky C, Cha ME (1990) Ribosomal RNA of the primitive eukaryote *Gaudia lanulosa* large subunit I and potential processing signals. *Genes* 96: 299–307.
97. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
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 30–61.
99. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on 28S rDNA sequences. *Acta Hydrobiologica Sinica* 31: 492–498.
100. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
101. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on 28S rDNA sequences. *Acta Hydrobiologica Sinica* 31: 492–498.
102. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
103. Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on 28S rDNA sequences. *Acta Hydrobiologica Sinica* 31: 492–498.
104. Chase MW, Killian A, Mauzi A, Jansen RK (2005) Evidence for the gene tree and species tree in the clade Eudorina (Chlorophyceae) based on analyses of RUBISCO large subunit (*rbcL*) gene sequences. *European Journal of Phycology* 38: 233–246.
105. 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.
106. Proschold T, Marin B, Schloesser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (chlorophyta). *Phylogenetica* 44: 502–505.
107. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
108. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
109. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
110. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
111. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
112. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
113. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
114. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.
115. Kapraun DF (1994) Cytophotometric estimation of nuclear DNA contents in thirteen species of the Caulerpales (Chlorophyta). *Cryptogram Botany* 4: 410–418.