Plant Gene 4 (2015) 20–28

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plant-gene

ITS2 and 18S rDNA sequence-structure phylogeny of *Chlorella* and allies (Chlorophyta, Trebouxiophyceae, Chlorellaceae)

Jaqueline S. Heeg, Matthias Wolf *

Department of Bioinformatics, Biocenter, University of Würzburg, 97074 Würzburg, Germany

ARTICLE INFO

Article history:
Received 19 June 2015
Received in revised form 11 August 2015
Accepted 26 August 2015
Available online 21 September 2015

Keywords:
Barcoding
Internal transcribed spacer
ITS2
Maximum likelihood
RRNA
Secondary structure
Sequence-structure alignment

ABSTRACT

In the last decade, the evolutionary diversity of *Chlorella* and allies has been discussed in a huge number of publications using internal transcribed spacer 2 (ITS2) and/or 18S ribosomal RNA gene sequences to infer the phylogenies. However, sister-group relations between different genera classified within the Chlorellaceae remained provisional, due to a lack of bootstrap support. In this study, using more than four hundred sequences, a comprehensive phylogenetic portrait of *Chlorella* and allies is presented and discussed; sixty key taxa are reconsidered by an analysis using primary sequences and their individual secondary structures simultaneously in inferring neighbor-joining, maximum parsimony and maximum likelihood trees, an approach most recently reviewed, with increasing robustness and accuracy of reconstructed phylogenies. While neighbor-joining and maximum parsimony analyses failed in inferring a robust phylogenetic tree, the maximum likelihood tree (in particular on a concatenated data set) provides a supported phylogeny preceding any taxonomic discussion.

1. Introduction

Chlorella is one of the best-studied phototrophic eukaryotes. The evolutionary diversity of *Chlorella* sensu stricto and related taxa, most recently reviewed by Krienitz et al. (2015), had been thoroughly discussed for the last decade (e.g. Bashan et al., 2015; Bock et al., 2010, 2011a, 2011b, 2011c; Hoshina, 2014; Hoshina et al., 2005, 2010; Hoshina and Fujiwara, 2013; Krienitz et al., 2004, 2010, 2012; Luo et al., 2006, 2010; Pröschold et al., 2010, 2011; Ustinova et al., 2001; Wolf et al., 2002). A growing number of coccoid green algae whose morphology is different from the spherical *Chlorella* have been classified as Chlorellaceae (Chlorophyta, Trebouxiophyceae). Within the family a genus and species concept, i.e., sister group relations between different genera, remained provisional due to a lack of bootstrap support in any molecular analysis (e.g. Hoshina and Fujiwara, 2013; Krienitz et al., 2012; Luo et al., 2010). Typically, the core Chlorellaceae consist of two clades, a well-supported *Parachlorella*- and a moderately-supported *Chlorella*-clade (cf. Krienitz et al., 2004). In most studies, due to high sequence variability, the alignment was guided by secondary structure information. Though, none of the studies used sequence-structure information simultaneously in inferring alignments and trees; an approach most recently reviewed, increasing robustness and accuracy of reconstructed phylogenies (Keller et al., 2010; Wolf et al., 2014; Wolf, 2015). Moreover, biased by taxon sampling, existing studies focus on specific subgroups classified within Chlorellaceae, and lack a complete picture presenting all currently available chlorellacean strains for which ITS2 and 18S sequences are available. In this study, a comprehensive phylogenetic portrait of *Chlorella* is presented and discussed; key taxa are reconsidered by a sequence-structure analysis using neighbor-joining (NJ), maximum parsimony (MP) and maximum likelihood (ML) approaches.

2. Material & methods

2.1. Taxon sampling and ITS2 sequence analysis

To get a feasible taxon sampling, all 421 currently available ITS2 rRNA gene sequences of Chlorellaceae were obtained from GenBank (Benson et al., 2013) and annotated using Hidden Markov Models (HMMs) (Keller et al., 2009) as implemented in the ITS2 databases I–IV (Ankenbrand et al., 2015; Koetschan et al., 2010, 2012; Merget et al., 2012; Schultz et al., 2006; Selig et al., 2008). Three hundred seventy five sequences were annotated. Strains unclassified as "sp." or strains not classified as Chlorellaceae sensu stricto were discarded. Two hundred seventeen sequences (cf. Table S1) were aligned by ClustalX (Larkin et al., 2007) and with ProfDistS (Wolf et al., 2008) a phylogenetic tree was reconstructed by neighbor-joining (Saitou and Nei, 1987) using a Jukes Cantor (JC) correction. Bootstrap support (Felsenstein, 1985) was estimated based on 100 pseudo-replicates. Outgroup taxa (*Chloroidium saccharophilum* (W. Krüger) Darienko, Gustavs, Mudimu, Menendez, Schumann, Karsten, Friedl *et* Pröschold 2010 and *Chloroidium ellipsoideum* (Gerneck) Darienko, Gustavs, Mudimu,

* Corresponding author.
E-mail address: matthias.wolf@biozentrum.uni-wuerzburg.de (M. Wolf).

http://dx.doi.org/10.1016/j.plgene.2015.08.001

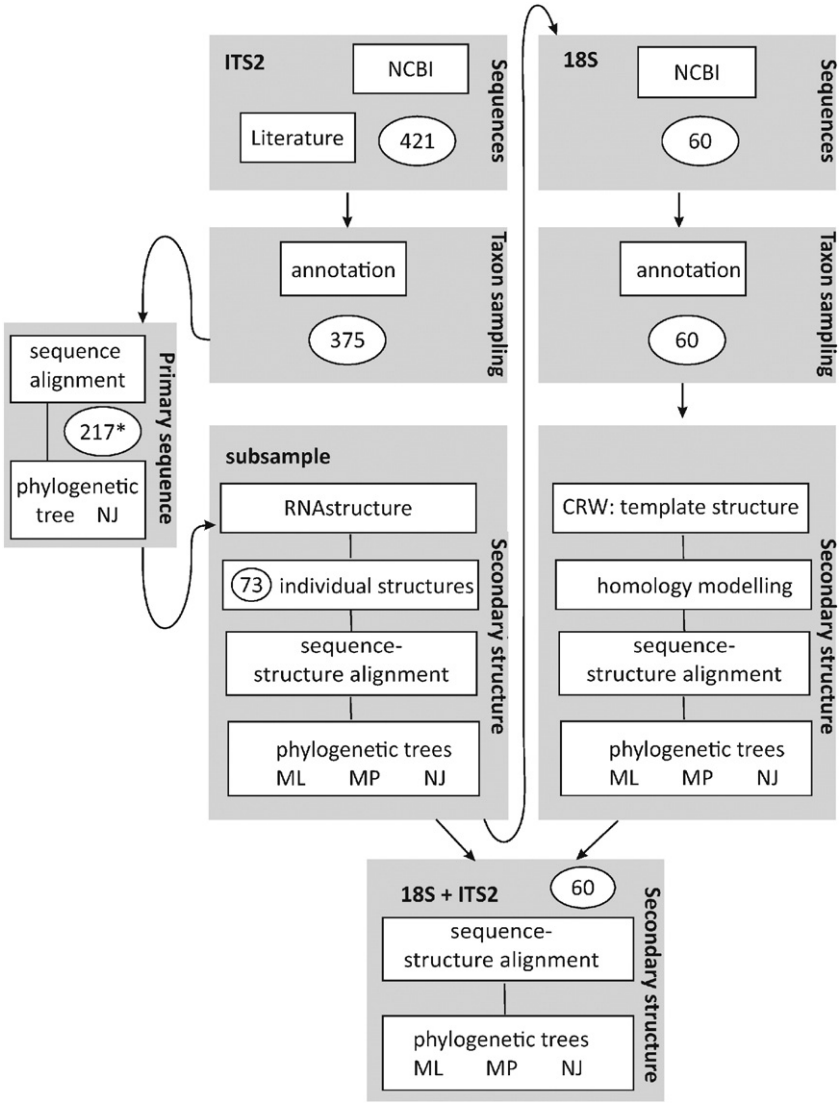

Fig. 1. Flowchart of the experimental setup. *After annotation of ITS2 and 18S rDNA sequences, strains unclassified as "sp." or strains not classified as Chlorellaceae sensu stricto were discarded. Clade specific key taxa were deduced from the overall picture (cf. Fig. 2) in order to realize a subsequent phylogenetic RNA sequence-structure analysis using a concatenated data set as well as each marker separately.

Menendez, Schumann, Karsten, Friedl *et* Pröschold 2010 were selected from among close allies in the Chlorellales (cf. Krienitz et al., 2004). Clade specific key taxa were deduced from the overall picture in order to reduce the number of taxa for a subsequent phylogenetic RNA sequence-structure analysis (Fig. 1).

2.2. ITS2 and 18S sequence-structure analysis

A total of 73 chlorellacean taxa were included in the phylogenetic analysis of ITS2 sequence-structure data (cf. Table S1). *C. saccharophilum* and *C. ellipsoideum* were used as outgroup. Secondary structures were predicted by homology modeling using a relevant template (cf. Wolf et al., 2005; Selig et al., 2008) or by RNAstructure using energy minimization and constraint folding (Mathews et al., 1999; Reuter and Mathews, 2010). In accordance with Keller et al. (2010); Wolf et al. (2014) and Wolf (2015), phylogenetic analysis of ITS2 followed the procedures outlined in Koetschan et al. (2012); Markert et al. (2012); Merget et al. (2012) and Schultz and Wolf (2009). Specifically, a global multiple sequence-structure alignment was automatically generated in 4SALE v1.7 (Seibel et al., 2006, 2008), whereby ITS2 sequences and their individual secondary structures were simultaneously aligned using an ITS2 sequence–structure specific scoring matrix (Seibel et al., 2006, reviewed in Wolf et al., 2014). 4SALE uses ClustalW (Larkin et al., 2007), but, with a specified scoring matrix, fitted to a 12-letter alphabet encoding the sequence–structure information and specifically trained on ITS2 sequence–structure data obtained from hundreds of thousands of sequence–structure pairs available at the ITS2 database. Hence, 4SALE does not use a 4×4 scoring matrix but rather a 12×12 matrix for each nucleotide, with its three structural states (paired left, paired right, or unpaired). Based on the simultaneous consideration of the primary sequence and the secondary structure information, phylogenetic relationships were reconstructed by neighbor-joining (NJ) through the use of an ITS2 sequence–structure specific, general time reversible (GTR) substitution model as implemented in ProfDistS v0.9.9 (Wolf et al., 2008). Using the ITS2 sequence and the ITS2 secondary structure information simultaneously (encoded by the 12-letter alphabet), a maximum parsimony tree (MP) (Camin and Sokal, 1965) was reconstructed by PAUP (Swofford, 2002) (with default settings) and a maximum likelihood tree (ML) (Felsenstein, 1981) was calculated using phangorn (Schliep, 2011) as implemented in the statistical framework R (R Core Team 2014). The R script is available from the 4SALE homepage at http://4sale.bioapps.biozentrum.uni-wuerzburg.de (cf. Wolf et al., 2014). Bootstrap support for the sequence-structure trees was estimated based on 1000 (NJ, MP) and 100 (ML) pseudo-replicates, respectively.

A total of 60 available 18S rRNA gene sequences >1690 nucleotides (without introns) accompanying the ITS2 data set were used for further processing (cf. Table S1). Secondary structures were predicted by homology modeling (at least 75% helix transfer, without pseudoknots) using the 18S rDNA secondary structure of *Heterochlorella luteoviridis* (Chodat) J. Neustupa, Y. Němcová, M. Eliáš *et* P. Škaloud 2009 as template (Selig et al., 2008; Wolf et al., 2005). The template structure (being the closest relative) was obtained from the Comparative RNA Web Site (CRW) (Cannone et al., 2002). Sequence-structure alignments and phylogenetic trees were reconstructed as described for the ITS2 data set.

The ITS2 and the 18S rDNA data sets (sequences and their individual secondary structures) were finally concatenated and processed as described for the single marker analyses.

2.3. Compensatory base change analysis

Concerning the ITS2 data set consisting of 73 taxa, species were further distinguished by a CBC analysis (cf. Coleman, 2000; Coleman, 2009; Mai and Coleman, 1997; Müller et al., 2007; Wolf et al., 2013). Compensatory base changes (CBCs) were identified using the CBCAnalyzer option implemented in 4SALE v1.7 (Seibel et al., 2006, 2008; Wolf et al., 2005).

3. Results

The neighbor-joining tree obtained from 217 annotated ITS2 sequences yielded a well-supported monophyletic *Parachlorella*- and a paraphyletic *Chlorella*-clade (Fig. 2). Most genera (e.g. *Actinastrum*,

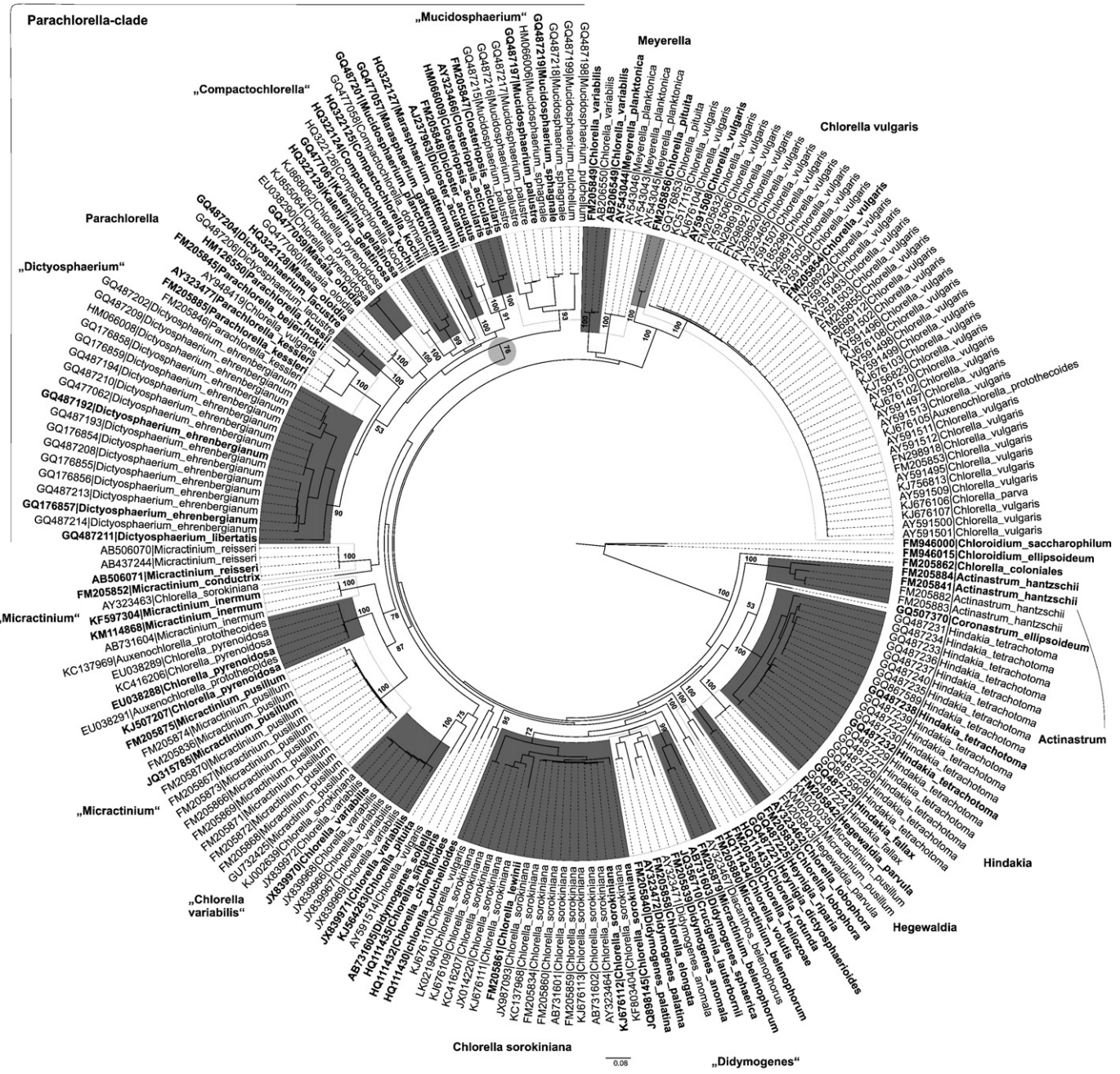

Fig. 2. Neighbor-joining tree based on the primary sequence information from all available and annotatable chlorellacean ITS2 sequences (strains unclassified as "sp." or strains not classified as Chlorellaceae sensu stricto were discarded). Bootstrap values >50 are added for monophyletic genera, if polyphyletic on species level. The tree is rooted with *Chloroidium saccharophilum* and *Chloroidium ellipsoideum*. GenBank accession numbers accompany each taxon name. Sequences used for a subsequent sequence-structure analysis are indicated in bold (cf. Fig. 1). The monophyletic *Parachlorella*-clade is highlighted. Key taxa are alternately marked in gray and white and additionally named alongside the tree. Non-monophyletic genera are indicated by quotation marks. The scale bar indicates evolutionary distances.

Hindakia, *Dicloster*, *Closteriopsis*, *Masaia*, *Kalenjinia*, *Maraspherium* or *Meyerella*) are highly supported being monophyletic. Others (e.g. *Dictyospherium*, *Parachlorella* or *Mucidosphaerium* are almost monophyletic (Fig. 2). How these genera are related to each other remained unresolved (Fig. 2).

ITS2 secondary structures (Figs. 3, 4) obtained for 73 key taxa (cf. Fig. 2) folded into the common core structure known for eukaryotes, consisting of four helices, the third being the longest (Schultz et al., 2005). Fig. 3 shows the ITS2 secondary structure for the type species *Chlorella vulgaris* Beyerinck (Beijerinck) 1890. Fig. 4 visualizes the complete sequence-structure alignment by a 51% consensus structure. Well-known sequence motifs like the U–U mismatch in helix II, an A-rich region between helices II and III, as well as the UGGU motif 5′ side to the apex of helix III are present in 100%, 100% and 95.77% respectively.

The maximum likelihood tree obtained from those 73 ITS2 sequence-structure pairs yielded a well-supported monophyletic *Parachlorella*- and a well-supported monophyletic *Chlorella*-clade (Fig. 5). Most genera are highly supported being monophyletic. Whereas maximum parsimony and neighbor-joining failed in supporting a robust tree topology, maximum likelihood yielded some further well supported monophyletic sub-groups within both clades, the *Parachlorella*- and the *Chlorella*-clade (Fig. 5).

18S secondary structures (not shown) obtained for 60 accompanying reference taxa folded into the common core structure known for eukaryotes (cf. Cannone et al., 2002).

The maximum likelihood tree obtained from those 18S sequence-structure pairs yielded a well-supported monophyletic *Parachlorella*- and a well-supported *Chlorella*-clade (Fig. 6). Most genera are highly supported being monophyletic. Whereas maximum parsimony and neighbor-joining failed in supporting a robust tree topology (besides for the *Parachlorella*- and the *Chlorella*-clade), maximum likelihood yielded bootstrap support values >50 for all but one branches (just one terminal clade connecting *Micractinium inermum* and *Hindakia fallax* is supported with only 46) (Fig. 6). The 18S tree topology

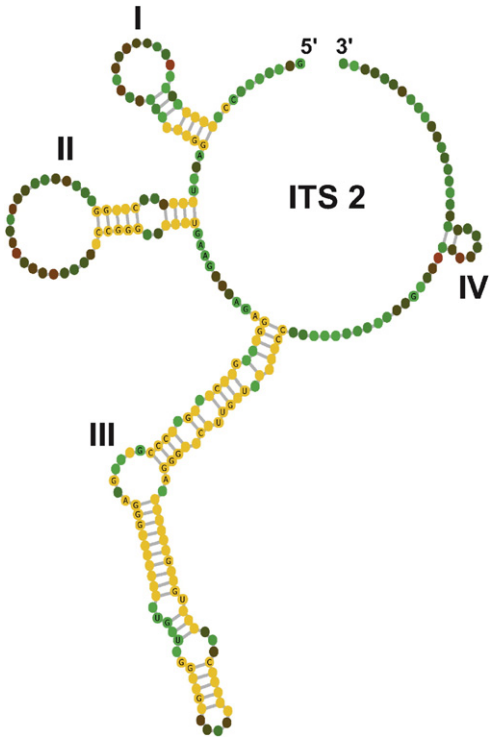

Fig. 4. Visualization of the complete sequence-structure alignment by a consensus structure (51%) for all ITS2 primary sequences obtained from the complete multiple sequence-structure alignment without gaps (71 chlorellacean sequence-structure pairs, outgroup taxa were excluded). Helices are numbered I–IV. Sequence conservation is indicated from red/brownish (not conserved) to green (conserved). Nucleotides which are 100% conserved in all sequences are written as A, U, G or C. Nucleotide bonds which are 90% conserved throughout the alignment are marked in yellow. The figure was generated with 4SALE.

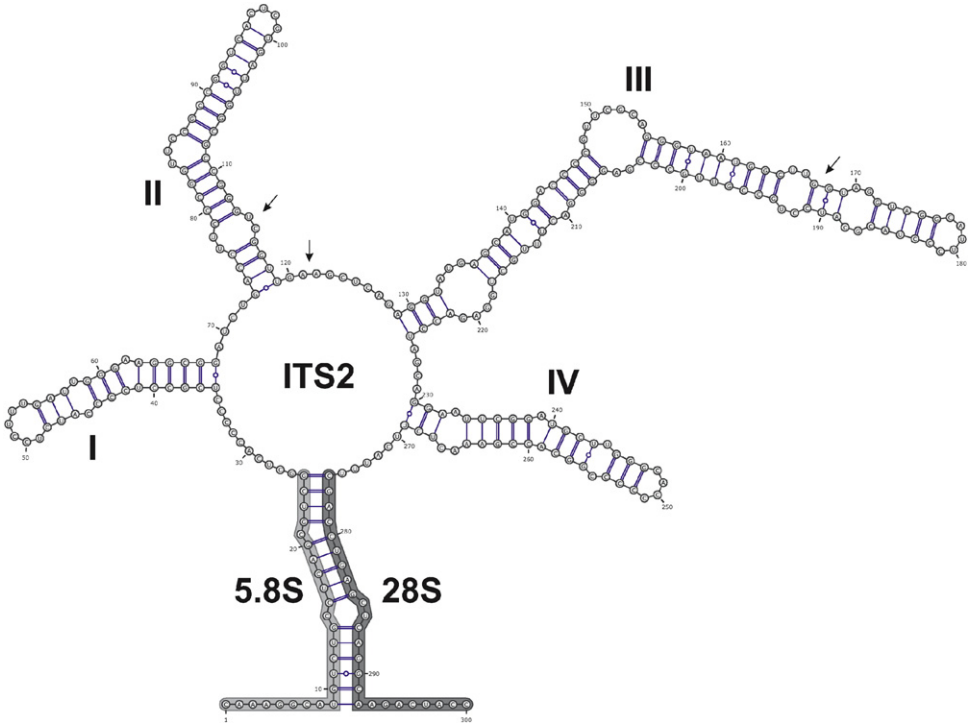

Fig. 3. 5.8S-28S rRNA gene hybridization (proximal stem region) and ITS2 secondary structure of *Chlorella vulgaris* (AY591505) visualized with VARNA (Darty et al., 2009). The 5.8S is indicated in light gray, the 28S in dark gray. ITS2 helices are numbered I–IV. Typical ITS2 motifs are highlighted: the pyrimidine–pyrimidine mismatch in helix II; the A-rich region between helices II and III, as well as the UGGU motif 5′ to the apex of helix III. The structure was generated by RNAstructure through energy minimization.

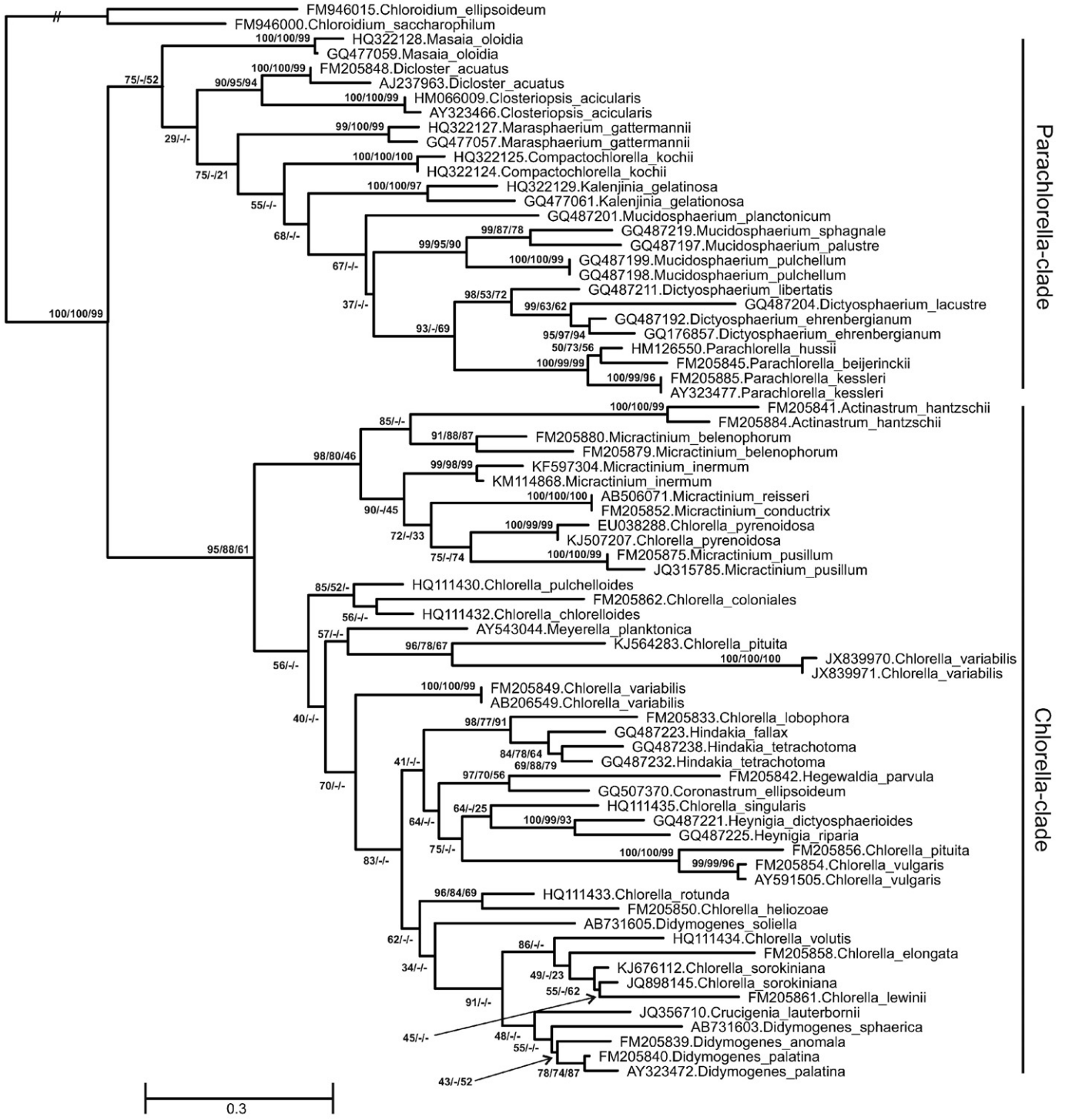

Fig. 5. ITS2 sequence-structure maximum likelihood tree reconstructed with phangorn using 73 chlorellacean ITS2 sequence-structure pairs. Bootstrap values are from ML, MP and NJ analyses. The tree is rooted with *Chloroidium saccharophilum* and *Chloroidium ellipsoideum*. GenBank accession numbers accompany each taxon name. The *Parachlorella*- and the *Chlorella*-clade are highlighted. The scale bar indicates evolutionary distances.

considerably differs from the ITS2 tree topology (Fig. 6), i.e., how monophyletic genera are related to each other differs throughout the trees.

The maximum likelihood tree obtained from the concatenated data set yielded a tree topology in agreement with the ITS2 tree topology. Whereas maximum parsimony and neighbor-joining failed in supporting a robust tree topology (besides for the *Parachlorella*- and the *Chlorella*-clade), maximum likelihood yielded bootstrap support values >50 for all but five branches (there is no support for two large subgroups within the *Parachlorella* clade and sister group relations for *Heynigia* and *Crucigenia* remained unclear) (Fig. 7). The concatenated sequence-structure alignment is available as supplementary material (File S1).

For Chlorellaceae, the probability of two ITS2 sequences belonging to different species, given at least one CBC, is 0.74. The probability for two

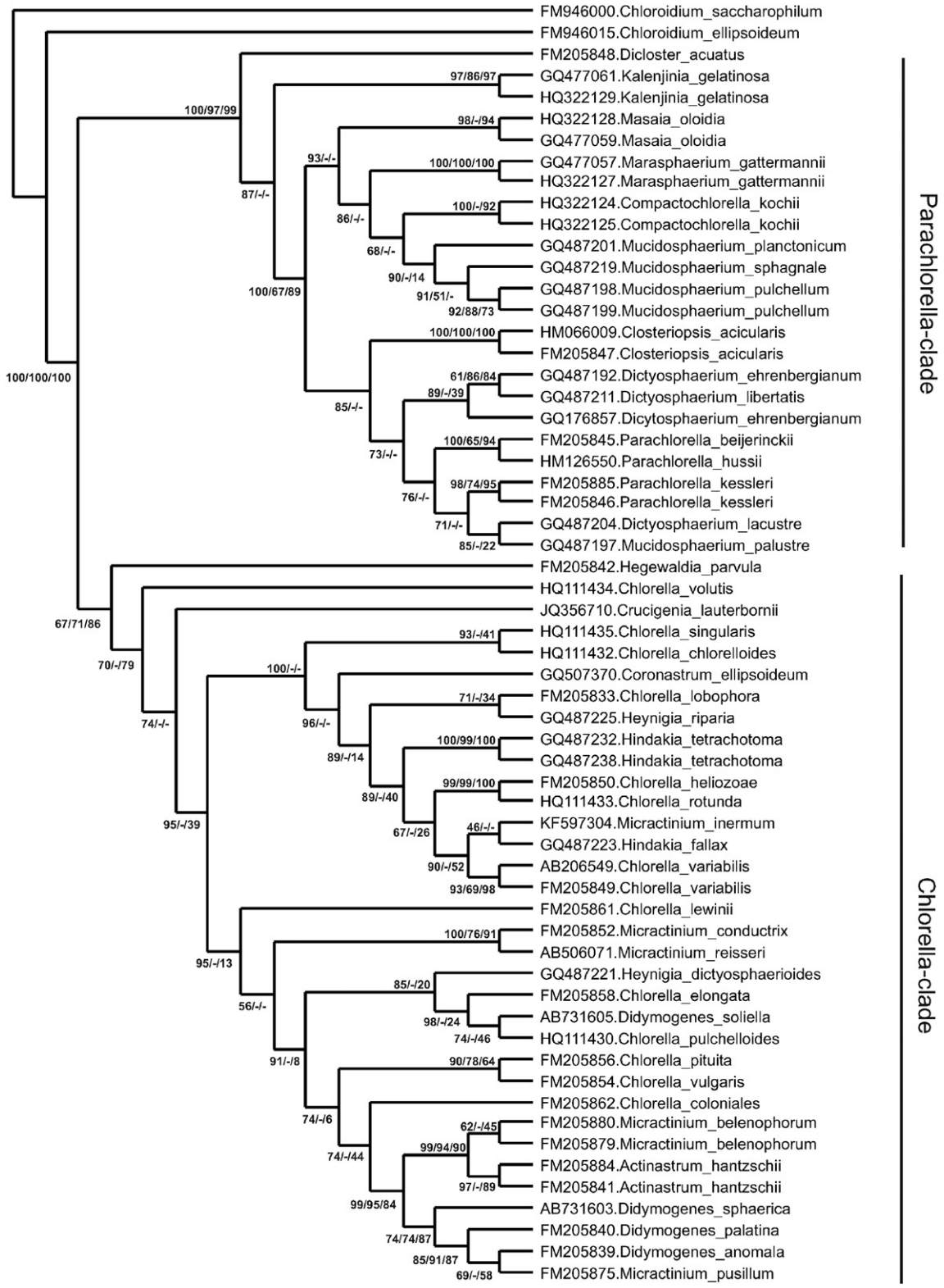

Fig. 6. 18S sequence-structure maximum likelihood tree reconstructed with phangorn using 60 chlorellacean 18S sequence-structure pairs. Bootstrap values are from ML, MP and NJ analyses. The tree is rooted with *Chloroidium saccharophilum* and *Chloroidium ellipsoideum*. GenBank accession numbers accompany each taxon name. The *Parachlorella*- and the *Chlorella*-clade are highlighted. For evolutionary distances see supplementary material Fig. S1.

ITS2 sequences belonging to the same species, given no CBC, is 0.93. The probabilities slightly differ from the average probabilities known from eukaryotes (cf. Müller et al., 2007). However, for Chlorellaceae, using only the more conserved helices II and III, the probabilities are 0.91 and 0.95, respectively.

4. Discussion

Remarkable newly-supported sub-clades might be (i) for ITS2, e.g. *Dicloster* + *Closteriopsis*, *Dictyosphaerium* + *Parachlorella*, *Hegewaldia* + *Coronastrum*, *Micractinium* (polyphyletic) + *Actinastrum*,

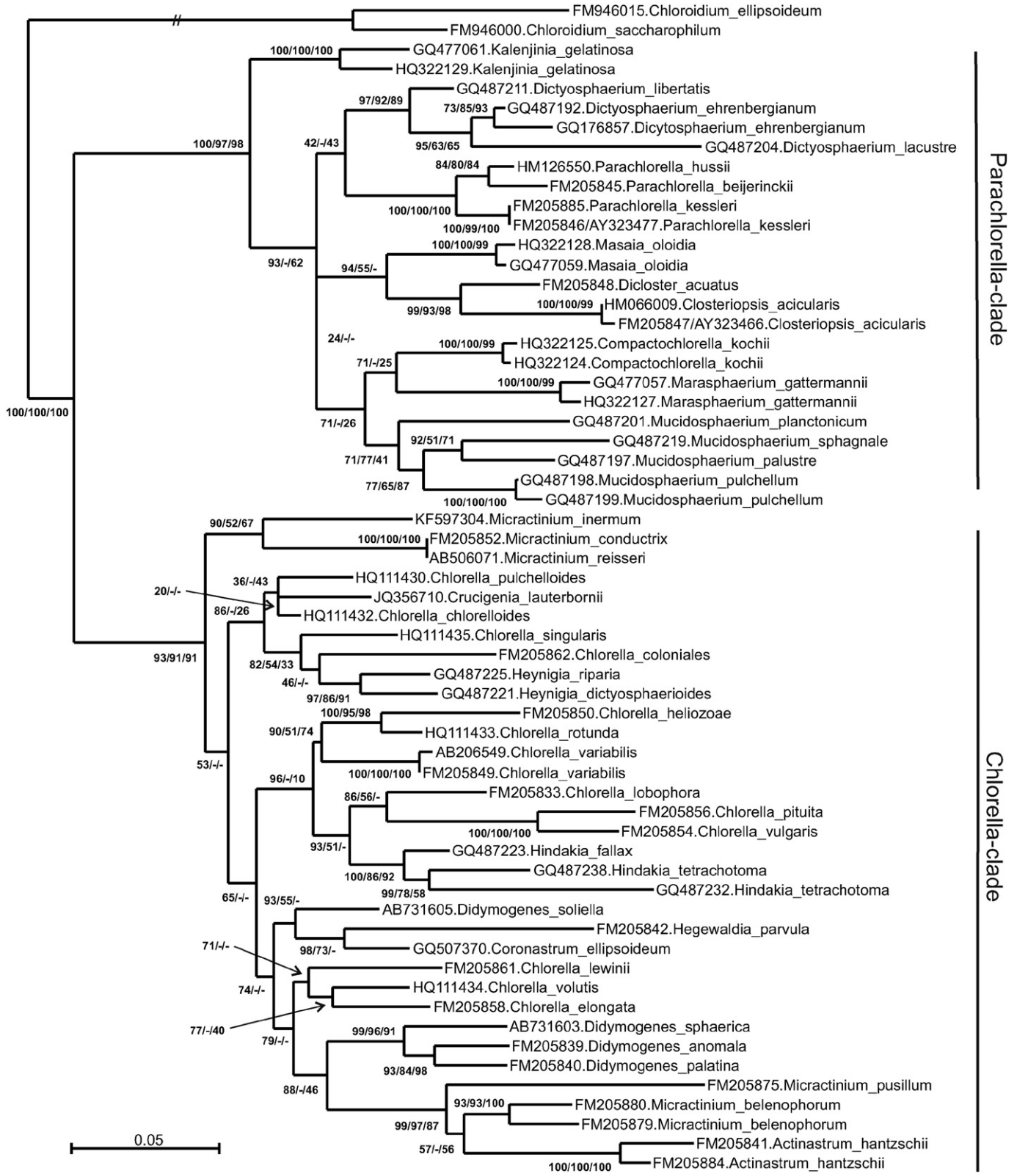

Fig. 7. ITS2 + 18S sequence-structure maximum likelihood tree reconstructed with phangorn using a concatenated data set of 60 chlorellacean ITS2 and 18S sequence-structure pairs. Bootstrap values are from ML, MP and NJ analyses. The tree is rooted with *Chloroidium saccharophilum* and *Chloroidium ellipsoideum*. GenBank accession numbers accompany each taxon name. The *Parachlorella*- and the *Chlorella*-clade are highlighted. The scale bar indicates evolutionary distances.

(ii) for 18S, e.g. *Masaia* + (*Marasphaerium* + (*Compactochlorella* + most strains of *Mucidosphaerium*)), and (iii) for a concatenated data set, e.g. *Masaia* + (*Dicloster* + *Closteriopsis*). Furthermore, *Meyerella*, *Diacanthos*, *Kalenjinia*, *Hindakia* as well as *Heynigia*, each genus appears as monophyletic and/or is represented by just a single species. However, first of all *Chlorella* itself, and additionally *Micractinium*, *Mucidosphaerium* and *Didymogenes* appear as polyphyletic. ITS2 and 18S are not congruent concerning a robust tree topology; using a concatenated data set the ITS2

obviously prevail the 18S rRNA gene because the tree obtained from the concatenated data set is much more in agreement with the ITS2 tree than with the 18S tree. Beside the *Parachlorella*- and the *Chlorella*-clade and the detection of monophyletic genera, the available literature (always using concatenated data sets) comes up with bootstrap support values <50 and a huge amount of chlorellacean sequences available in GenBank is not classified (cf. Fig. 1). Inferring the phylogeny of *Chlorella* and allies is still a nightmare. Better bootstrap support not necessarily means better trees. Bootstrap is about robustness, not about accuracy. The outgroup selection, the taxon sampling, and the phylogenetic reconstruction method do influence inferred tree topologies, not to speak about designation errors in morphological species concepts.

In this study using RNA sequence and secondary structure information simultaneously we come up with moderately supported but contradicting trees. Assuming the tree based on the concatenated data set being the most accurate (Fig. 7), we have maximum likelihood support (>50) for all but five branches. Sub-clades of morphological distinct species supported within the *Parachlorella*- and the *Chlorella*-clade strengthen the need for a profound genus concept, flanked by a CBC species concept, and obtained by future analysis using additional marker genes and/or genomic approaches. Technical updates concerning the methods (e.g. currently phangorn uses nearest neighbor interchange (NNI) instead of tree bisection reconnection (TBR)) are needed as much as morphologists trained in species designation. Moreover, we need evolutionary developmental biologists arguing for apomorphies deduced from the evolution of Chlorellaceae. Nevertheless, with all trees presented in this little research note, especially concerning the overall picture, here we present a phylogenetic framework and a basis at least for future barcoding approaches distinguishing *Chlorella* and allies.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.plgene.2015.08.001.

Acknowledgments

We thank Volker AR Huss (University of Erlangen-Nuremberg) for pre-review and helpful comments on the manuscript as well as Tobias Müller (University of Würzburg) for his help with R.

References

Ankenbrand, M.J., Keller, A., Wolf, M., Schultz, J., Förster, F., 2015. ITS2 database V: twice as much. Mol. Biol. Evol. http://dx.doi.org/10.1093/molbev/msv174.

Bashan, Y., Lopez, B.R., Huss, V.A.R., Amavizca, E., de- Bashan, L.E., 2015. *Chlorella sorokiniana* (formerly *C. vulgaris*) UTEX 2714, a non-thermotolerant microalga useful for biotechnological applications and as a reference strain. J. Appl. Phycol. http://dx.doi.org/10.1007/s10811-015-0571-z.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013. GenBank. Nucleic Acids Res. 41, D36–D42.

Bock, C., Krienitz, L., Pröschold, T., 2011a. Taxonomic reassessment of the genus *Chlorella* (Trebouxiophyceae) using molecular signatures (barcodes), including description of seven new species. Fottea 11, 293–312.

Bock, C., Pažoutová, M., Krienitz, L., 2011c. Phylogenetic position of *Coronastrum ellipsoideum* and description of *Parachlorella hussii* sp. nov. Biologia 66, 585–594.

Bock, C., Pröschold, T., Krienitz, L., 2010. Two new *Dictyosphaerium*-morphotype lineages of the Chlorellaceae (Trebouxiophyceae): *Heynigia* gen. nov. and *Hindakia* gen. nov. Eur. J. Phycol. 45, 267–277.

Bock, C., Pröschold, T., Krienitz, L., 2011b. Updating the genus *Dictyosphaerium* and description of *Mucidosphaerium* gen. nov. (trebouxiophyceae) based on morphological and molecular data. J. Phycol. 47, 638–652.

Camin, J.H., Sokal, R.R., 1965. A method for deducing branching sequences in phylogeny. Evolution 19, 311–326.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Müller, K.M., Pande, N., Shang, Z., Yu, N., Gutell, R.R., 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinf. 3, 2.

Coleman, A.W., 2000. The significance of a coincidence between evolutionary landmarks found in mating affinity and a DNA sequence. Protist 151, 1–9.

Coleman, A.W., 2009. Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide. Mol. Phylogenet. Evol. 50, 197–203.

Darty, K., Denise, A., Ponty, Y., 2009. VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25, 1974–1975.

Felsenstein, J., 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution 35, 1229–1242.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Hoshina, R., 2014. DNA analyses of a private collection of microbial green algae contribute to a better understanding of microbial diversity. BMC Res. Notes 7, 592.

Hoshina, R., Fujiwara, Y., 2013. Molecular characterization of *Chlorella* cultures of the National Institute for Environmental Studies culture collection with description of *Micractinium inermum* sp. nov., *Didymogenes sphaerica* sp. nov., and *Didymogenes soliella* sp. nov. (Chlorellaceae, Trebouxiophyceae). Phycol. Res. 61, 124–132.

Hoshina, R., Iwataki, M., Imamura, N., 2010. *Chlorella variabilis* and *Micractinium reisseri* sp. nov. (Chlorellaceae, Trebouxiophyceae): redescription of the endosymbiotic green algae of *Paramecium bursaria* (Peniculia, Oligohymenophorea) in the 120th year. Phycol. Res. 58, 188–201.

Hoshina, R., Kato, Y., Kamako, S., Imuamura, N., 2005. Genetic evidence of "American" and "European" type symbiotic algae of *Paramecium bursaria* Ehrenberg. Plant Biol. 7, 526–532.

Keller, A., Förster, F., Müller, T., Dandekar, T., Schultz, J., Wolf, M., 2010. Including RNA secondary structures improves accuracy and robustness in reconstruction of phylogenetic trees. Biol. Direct 5, 4.

Keller, A., Schleicher, T., Schultz, J., Müller, T., Dandekar, T., Wolf, M., 2009. 5.8S–28S rRNA interaction and HMM-based ITS2 annotation. Gene 430, 50–57.

Koetschan, C., Förster, F., Keller, A., Schleicher, T., Ruderisch, B., Schwarz, R., Müller, T., Wolf, M., Schultz, J., 2010. The ITS2 database III — sequences and structures for phylogeny. Nucleic Acids Res. 38, D275–D279.

Koetschan, C., Hackl, T., Müller, T., Wolf, M., Förster, F., Schultz, J., 2012. ITS2 database IV: interactive taxon sampling for internal transcribed spacer 2 based phylogenies. Mol. Phylogenet. Evol. 63, 585–588.

Krienitz, L., Bock, C., Kotut, K., Pröschold, T., 2012. Genotypic diversity of Dictyosphaerium morphospecies (Chlorellaceae, Trebouxiophyceae) in African inland waters, including the description of four new genera. Fottea 12, 231–253.

Krienitz, L., Bock, C., Luo, W., Pröschold, T., 2010. Polyphyletic origin of the *Dictyosphaerium* morphotype within Chlorellaceae (Trebouxiophyceae). J. Phycol. 46, 559–563.

Krienitz, L., Hegewald, E., Hepperle, D., Huss, V.A.R., Rohr, T., Wolf, M., 2004. Phylogenetic relationship of *Chlorella* and *Parachlorella* gen. nov. (Chlorophyta, Trebouxiophyceae). Phycologia 43, 529–542.

Krienitz, L., Huss, V.A.R., Bock, C., 2015. *Chlorella*: 125 years of the green survivalist. Trends Plant Sci. 20, 67–69.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

Luo, W., Pflugmacher, S., Pröschold, T., Walz, N., Krienitz, L., 2006. Genotype versus phenotype variability in *Chlorella* and *Micractinium* (Chlorophyta, Trebouxiophyceae). Protist 157, 315–333.

Luo, W., Pröschold, T., Bock, C., Krienitz, L., 2010. Generic concept in *Chlorella*-related coccoid green algae (Chlorophyta, Trebouxiophyceae). Plant Biol. 12, 545–553.

Mai, J.C., Coleman, A.W., 1997. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J. Mol. Evol. 44, 258–271.

Markert, S.M., Müller, T., Koetschan, C., Friedl, T., Wolf, M., 2012. 'Y' *Scenedesmus* (Chlorophyta, Chlorophyceae): the internal transcribed spacer 2 rRNA secondary structure re-revisited. Plant Biol. 14, 987–996.

Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288, 911–940.

Merget, B., Koetschan, C., Hackl, T., Förster, F., Dandekar, T., Müller, T., Schultz, J., Wolf, M., 2012. The ITS2 database. J. Vis. Exp. 61, e3806.

Müller, T., Philippi, N., Dandekar, T., Schultz, J., Wolf, M., 2007. Distinguishing species. RNA 13, 1469–1472.

Pröschold, T., Bock, C., Luo, W., Krienitz, L., 2010. Polyphyletic distribution of bristle formation in Chlorellaceae: *Micractinium*, *Diacanthos*, *Didymogenes* and *Hegewaldia* gen. nov. (Trebouxiophyceae, Chlorophyta). Phycol. Res. 58, 1–8.

Pröschold, T., Darienko, T., Silva, P.C., Reisser, W., Krienitz, L., 2011. The systematics of Zoochlorella revisited employing an integrative approach. Environ. Microbiol. 13, 350–364.

Reuter, J.S., Mathews, D.H., 2010. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinf. 11, 129.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Schliep, K.P., 2011. Phangorn: phylogenetic analysis in R. Bioinformatics 27, 592–593.

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.

Schultz, J., Müller, T., Achtziger, M., Seibel, P.N., Dandekar, T., Wolf, M., 2006. The internal transcribed spacer 2 database — a web server for (not only) low level phylogenetic analyses. Nucleic Acids Res. 34, W704–W707.

Schultz, J., Wolf, M., 2009. ITS2 sequence-structure analysis in phylogenetics: a how-to manual for molecular systematics. Mol. Phylogenet. Evol. 52, 520–523.

Seibel, P.N., Müller, T., Dandekar, T., Schultz, J., Wolf, M., 2006. 4SALE — a tool for synchronous RNA sequence and secondary structure alignment and editing. BMC Bioinf. 7, 498.

Seibel, P.N., Müller, T., Dandekar, T., Wolf, M., 2008. Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4 SALE. BMC Res. Notes 1, 91.

Selig, C., Wolf, M., Müller, T., Dandekar, T., Schultz, J., 2008. The ITS2 database II: homology modelling RNA structure for molecular systematics. Nucleic Acids Res. 36, D377–D380.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
Ustinova, I., Krienitz, L., Huss, V.A.R., 2001. *Closteriopsis acicularis* (G. M. Smith) Belcher et Swale is a fusiform alga closely related to *Chlorella kessleri* Fott et Nováková (Chlorophyta, Trebouxiophyceae). Eur. J. Phycol. 36, 341–351.
Wolf, M., 2015. ITS so much more. Trends Genet. 31, 175–176.
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.
Wolf, M., Chen, S., Song, J., Ankenbrand, M., Müller, T., 2013. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences — a proof of concept. PLoS ONE 8, e66726.
Wolf, M., Koetschan, C., Müller, T., 2014. ITS2, 18S, 16S or any other RNA—simply aligning sequences and their individual secondary structures simultaneously by an automatic approach. Gene 546, 145–149.
Wolf, M., Krienitz, L., Hepperle, D., 2002. Phylogenetic position of *Actinastrum hantzschii* LAGERHEIM 1882 (Chlorophyta, Trebouxiophyceae). Algol. Stud. 104, 59–67.
Wolf, M., Ruderisch, B., Dandekar, T., Schultz, J., Müller, T., 2008. ProfDistS: (profile-) distance based phylogeny on sequence-structure alignments. Bioinformatics 24, 2401–2402.