J. Phycol. **46,** 743–750 (2010)

DOI: 10.1111/j.1529-8817.00851.x

DNA BARCODING OF CHLORARACHNIOPHYTES USING NUCLEOMORPH ITS SEQUENCES[1]

Gillian H. Gile,[2] *Rowena F. Stern,*[2] *Erick R. James, and Patrick J. Keeling*[3]

Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, British Columbia, Canada V6T 1Z4

Chlorarachniophytes are a small group of marine photosynthetic protists. They are best known as examples of an intermediate stage of secondary endosymbiosis: their plastids are derived from green algae and retain a highly reduced nucleus, called a nucleomorph, between the inner and outer pairs of membranes. Chlorarachniophytes can be challenging to identify to the species level, due to their small size, complex life cycles, and the fact that even genus-level diagnostic morphological characters are observable only by EM. Few species have been formally described, and many available culture collection strains remain unnamed. To alleviate this difficulty, we have developed a barcoding system for rapid and accurate identification of chlorarachniophyte species in culture, based on the internal transcribed spacer (ITS) region of the nucleomorph rRNA cistron. Although this is a multicopy locus, encoded in both subtelomeric regions of each chromosome, interlocus variability is low due to gene conversion by homologous recombination in this region. Here, we present barcode sequences for 39 cultured strains of chlorarachniophytes (>80% of currently available strains). Based on barcode data, other published molecular data, and information from culture records, we were able to recommend names for 21 out of the 24 unidentified, partially identified, or misidentified chlorarachniophyte strains in culture. Most strains could be assigned to previously described species, but at least two to as many as five new species may be present among cultured strains.

Key index words: ***Bigelowiella*; *Chlorarachnion*; culture collections; *Gymnochlora*; internal transcribed spacer; *Lotharella*; *Norrisiella*; *Partenskyella***

Abbreviations: **bp, base pairs; ITS, internal transcribed spacer**

Chlorarachniophytes are a small group of marine photosynthetic protists belonging to the phylum Cercozoa in the supergroup Rhizaria (Bhattacharya et al. 1995, Cavalier-Smith and Chao 1996, Keeling 2001, Nikolaev et al. 2004). They can take the form of amebae, coccoid cells, or flagellates, and many of them alternate among all of these forms and variations thereof during their complex life cycles (Fig. 1). A major distinguishing feature of the chlorarachniophytes is their retention of a nucleomorph. Chlorarachniophytes acquired their plastid by the secondary endosymbiotic uptake of a green alga, and the nucleomorph is a relict nucleus of that endosymbiont that persists between the inner and outer pairs of their four membrane-bound plastids (Ludwig and Gibbs 1989, Ishida et al. 1997, Rogers et al. 2007). Chlorarachniophyte nucleomorph genomes are invariably composed of three chromosomes ranging in size from roughly 100 to 200 kbp and bearing rRNA cistrons at each of the chromosome ends (Silver et al. 2007). The nucleomorph has attracted the attention of molecular evolutionary biologists as a window into processes of endosymbiotic organelle integration, such as endosymbiotic gene transfer, genome compaction, and development of protein-targeting systems (Gilson and McFadden 1996, 2002, Archibald et al. 2003, Rogers et al. 2004, Gilson et al. 2006, Gile and Keeling 2008). As a result, knowledge about chlorarachniophytes is heavily biased toward their molecular evolution: the entire chloroplast and nucleomorph genomes of *Bigelowiella natans*, the model chlorarachniophyte for molecular studies, have been sequenced (Gilson et al. 2006, Rogers et al. 2007).

The type species, *Chlorarachnion reptans* Geitler, was first described in 1930 by Lothar Geitler, but there were no further studies of chlorarachniophytes until it was rediscovered, described more fully with EM, and brought into culture by Hibberd and Norris in 1981 (Geitler 1930, Hibberd and Norris 1984). Today, seven genera and 12 species have been described, of which nine species from five genera are represented among the 48 strains in public culture collections worldwide. Chlorarachniophyte genera are defined mainly by the morphology of the bulbous pyrenoid that protrudes from the

[1]Received 15 May 2009. Accepted 28 April 2010.
[2]These authors contributed equally to this work.
[3]Author for correspondence: e-mail pkeeling@interchange.ubc.ca.

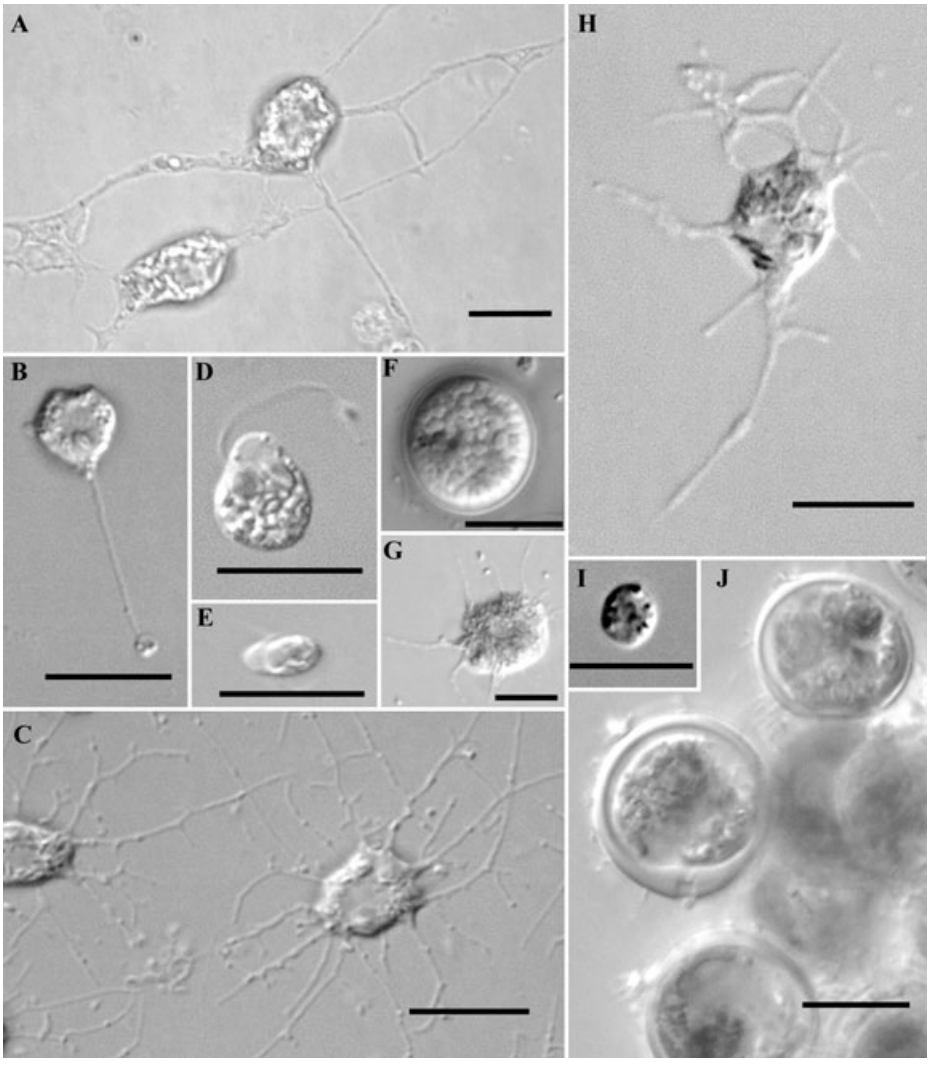

FIG. 1. Differential interference contrast light micrographs illustrating a diversity of chlorarachniophyte cell types. (A) *Chlorarachnion reptans* CCCM 449 ameboid cells, (B) *Bigelowiella longifila* characteristic ameboid cell with single extended filopodium, (C) *Gymnochlora stellata* CCMP 2057 ameboid cells, (D and E) *Lotharella oceanica* CCMP 622 flagellate cells, (F) *C. reptans* CCCM 449 coccoid cell, (G) *Lotharella amoeboformis* CCMP 2058 ameboid cell, (H) *C. reptans* CCMP 238 ameboid cell, (I) *Bigelowiella natans* CCMP 621 flagellate cell, (J) *Lotharella vacuolata* CCMP 240 thick-walled resting cysts (a sealed pore is visible in the wall of one cell). All scale bars represent 10 μm.

plastid toward the center of the cell and is capped on the cytoplasmic side by a carbohydrate storage vesicle. The pyrenoid can have a shallow cleft caused by an invagination of the inner pair of plastid membranes, as in *Norrisiella* S. Ota et Ishida (Ota et al. 2007a); a deep cleft, as in *Lotharella* Ishida et Y. Hara (Ishida et al. 1996, 2000, Ota et al. 2005, 2009a); or a deep cleft that encloses the nucleomorph, as in *Chlorarachnion* (Hibberd and Norris 1984). Otherwise, the pyrenoid can be invaded with small tubular invaginations of the innermost plastid membrane, as in *Gymnochlora* Ishida et Y. Hara (Ishida et al. 1996), or it may be lacking entirely, as in *Partenskyella* S. Ota, Vaulot et Ishida (Ota et al. 2009b). Two genera are exceptions to this scheme: there are no EM data concerning pyrenoid morphology in *Cryptochlora* E. Calderón et Schnetter (Calderon-Saenz and Schnetter 1987, 1989), and the genus *Bigelowiella* Moestrup (Moestrup and Sengco 2001) was defined by its dominant life-cycle stage being flagellate cells (pyrenoids of *Bigelowiella* are invaded by a shallow cleft, as in *Norrisiella*).

The principle behind DNA barcoding is to enable accurate species identification through a short signature sequence of DNA (Hebert et al. 2003a). Where feasible, barcoding has the advantage of allowing nonexperts to accurately and rapidly identify species even when diagnostic morphological features are missing (as for immature or damaged specimens), misleading (as for common convergent morphologies, such as coccoid green algae), or not easily observable (as for simple microscopic

organisms). Chlorarachniophytes exemplify the latter two points in that in their coccoid or flagellate forms they can easily be mistaken for other green-pigmented algae [CCMP 621 *B. natans* was initially misidentified as *Pedinomonas minutissima*, a chlorophyte (Moestrup and Sengco 2001)], and their diagnostic features can be observed only by EM. Moreover, a great many chlorarachniophyte cultures in available collections have not been identified to the species level and for many years were automatically assigned to the type genus *Chlorarachnion*. A barcoding system for chlorarachniophytes would therefore be a useful tool for clarifying the existing diversity of cultures as well as accurately distinguishing chlorarachniophyte species from each other and from other groups of algae.

The most commonly used DNA sequence marker for barcoding is the 5′ end of the cytochrome c oxidase I (*cox1*) gene. This locus can effectively discriminate species in many animal groups (Hebert et al. 2003b), as well as red algae (Saunders 2005), brown algae (Lane et al. 2007, Kucera and Saunders 2008), diatoms (Evans et al. 2007), and the ciliate genus *Tetrahymena* (Chantangsi et al. 2007). The *cox1* locus is not universally applicable, however, either because it does not display sufficient variation to distinguish species, as is the case with land plants (Chase et al. 2005, Fazekas et al. 2008), or because it is missing, as is the case with many anaerobic protists. A commonly used alternative is the ITS of the rRNA cistron. Although this locus can be problematic due to length variation among the multiple copies typically present per nuclear genome, the ITS locus has been used successfully for groups including sponges (Park et al. 2007), anthozoans (Sinniger et al. 2008), and fungi (Li et al. 2008, Nguyen and Seifert 2008, Ortega et al. 2008).

Here, we present a method for barcoding chlorarachniophytes using nucleomorph ITS sequences, in which intragenomic variation is limited by gene conversion in nucleomorph genomes. We have determined barcode sequences for 39 out of 48 chlorarachniophyte strains in culture collections plus *Norrisiella sphaerica* S. Ota et Ishida, representing nine described species from seven genera (*Cryptochlora perforans* E. Calderón et Schnetter and *Lotharella polymorpha* C. Dietz, K. Ehlers, C. Wilh., Gil-Rodriguez et Schnetter are not available from culture collections). This marker clearly resolved described species, assigned most unidentified or provisionally named chlorarachniophyte strains to a previously described species, and revealed that one ''chlorarachniophyte'' strain, UTEX 2631, is a member of the Eustigmatophyceae. Up to 25 strains are redundant, being either directly derived from other strains or isolated at the same time and place and found to have identical barcode sequences, indicating a lower diversity of chlorarachniophytes in culture than the number of strains would suggest.

MATERIALS AND METHODS

DNA extraction, PCR, and sequencing. Chlorarachniophyte cultures (Table S1 in the supplementary material) were either purchased or donated from the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP, West Boothbay Harbor, ME, USA); the Roscoff Culture Collection (RCC, Roscoff, France); the Plymouth Culture Collection of Marine Algae (PCC, Plymouth, UK); the Culture Collection of Algae at the University of Texas (UTEX, Austin, TX, USA); the Culture Collection of Algae, Philipps-University Marburg (Marburg, Germany); the Canadian Centre for the Culture of Microorganisms (CCCM, Vancouver, Canada); and the Microbial Culture Collection at the National Institute for Environmental Studies (NIES, Tsukuba, Japan). Chlorarachniophyte cells from ~10 mL of culture were harvested by centrifugation at 1,150*g* in an Eppendorf 5415 D benchtop centrifuge (Eppendorf, Hamburg, Germany) and subjected to two rounds of snap-freeze treatment with liquid nitrogen. DNA was extracted using the DNeasy Plant mini kit (Qiagen, Mississauga, ON, Canada) according to the manufacturer's instructions. Previously published primers (Inagaki et al. 1998) failed to amplify the most commonly employed locus for barcoding, the 5′ region of cytochrome c oxidase subunit 1 (*cox1*) under various reaction conditions. Because *cox1* sequences are publicly available for only two chlorarachniophytes and only two other Cercozoa, primer optimization for this locus in the chlorarachniophytes is impracticable at this time, and we shifted our focus to the ITS between the LSU and SSU rRNAs, which includes the 5.8S rRNA. Previously published universal ITS primers, ITS1 (Gottschling et al. 2005) and ITS4 (White et al. 1990), amplified either the nuclear or nucleomorph ITS region, so we designed new primers based on published nuclear and nucleomorph SSU sequences (Silver et al. 2007) to selectively amplify either nuclear or nucleomorph ITS loci along with a short stretch of SSU sequence to aid in positive identification and orientation of this variable region. Nuclear ITS primers also include a short stretch of the LSU sequence. Nucleomorph ITS sequences were amplified using nmITSF, 5′-AACGAGGAATGCCTAGTAAGC-3′ and ITS4, 5′-TCCTCCGCTTATTGATATGC-3′ primers. Nuclear ITS sequences required a heminested PCR using nucITSF 5′-AACGAGGAATTTCTAGTAAAC-3′ as a forward primer in both reactions, and nucITSR-outer 5′-CAATCCCAAACAACACGACTCTTCG-3′ and nucITSR-inner 5′-TCTGTACGGGGTTCTCACCCT-3′ as reverse primers. Some of the sequences presented here were initially obtained using ITS1 and ITS4 primers; for each of these sequences, the nucleomorph-specific ITS primers were subsequently confirmed to amplify products of the expected size. PCR reactions were carried out using Econotaq DNA polymerase (Lucigen, Middleton, WI, USA) or PuReTaq Ready-to-Go PCR beads (GE Life Sciences, Piscataway, NJ, USA) with an initial denaturation at 94°C for 3 min followed by 35 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min for nucleomorph or 2 min for nuclear ITS amplification, and a final extension step at 72°C for 10 min. PCR products were purified using QIAquick PCR purification kit (Qiagen) and sequenced directly on both strands at Macrogen (Seoul, Korea). Barcode sequences determined in this study were deposited in GenBank under accession numbers FJ821375–FJ821382, FJ821384, FJ821385, FJ821387–FJ821428, FJ937328–FJ937333, and FJ937351.

Phylogenetic analysis. Nuclear and nucleomorph ITS sequences were aligned separately using ClustalX (Thompson et al. 1997) with the gap opening penalty reduced to five and gap extension penalty reduced to two. Sequences from strains belonging to the same species were nearly identical, but each species was quite divergent relative to the others. To improve the resolution of among-species relationships, ambiguously aligned regions were removed by GBlocks 0.91b (Castresana

2000) with all gap positions allowed and all other parameters set to default, which reduced the alignment sizes from 903 to 610 and 1,459 to 679 characters for nucleomorph and nuclear data sets, respectively. Modeltest v3.7 (Posada and Crandall 1998, Posada and Buckley 2004) was used to determine the best-fit model of nucleotide substitution using the Akaike information criterion (AIC) from log likelihoods computed in PAUP* 4.0b10 (Swofford 2002). Phylogenetic trees were inferred using distance, maximum-likelihood (ML), and Bayesian methods using the HKY+I model for the nuclear data set and the TVMef+I+G model for the nucleomorph data set and specifying transition/transversion ratios, equilibrium base frequencies, substitution category rates, proportion of invariable sites, and gamma shape parameter alpha as specified by Modeltest. Distance trees were computed using the neighbor-joining (NJ) method from 1,000 bootstrap replicate data sets, and a consensus tree was generated in PAUP*. ML trees with 1,000 bootstrap replicates were computed in PhyML v3.0 (Guindon and Gascuel 2003) using a BIONJ starting tree and NNI search procedure. Bayesian analyses were performed with MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001) using six substitution categories and a proportion of invariable sites for the nucleomorph data set and two substitution categories and a proportion of invariable sites for the nuclear data set. For each data set, two independent chains, sampled each 100 generations, were run until they converged (the average standard deviation of partition frequency values between the chains dropped below 0.01) with 25% of the trees discarded as burn-in. The nuclear data set runs continued for 230,000 generations, and the nucleomorph data set runs continued for 650,000 generations. Consensus trees were computed from the final 75% of saved trees of both runs, for a total of 3,450 nuclear trees and 9,750 nucleomorph ITS trees. Uncorrected "*p*" pair-wise distances were computed in PAUP* 4.0b10 (Swofford 2002).

RESULTS AND DISCUSSION

Assessment of nucleomorph ITS for barcoding. Here, we present barcodes for chlorarachniophyte strains in culture worldwide using the nucleomorph ITS locus. Nucleomorph ITS was chosen over nuclear ITS and the more commonly used *cox1* locus due to its ability to clearly differentiate described chlorarachniophyte species and the efficiency with which it can be amplified and directly sequenced. Nucleomorph ITS PCR products are ~700 base pairs long, an ideal length for directly sequencing both strands, and are expected to display little or no interlocus variability on the basis of the *B. natans* nucleomorph genome. In chlorarachniophyte nucleomorphs, the rRNA cistrons are typically located in both subtelomeric regions of all three chromosomes, and their sequence homogeneity is maintained by gene conversion via homologous recombination (Gilson et al. 2006). In the nucleomorph genome of *B. natans*, five out of six ITS regions are identical in sequence, and the sixth, located on one end of chromosome 1 (1a), differs only by four nucleotides (Fig. 2). In contrast, the nuclear ITS-amplified product length is roughly 1,300 base pairs, and for many chlorarachniophytes, it could not be sequenced directly on both strands due to a variable region near the 3′ end of the amplified fragment. We searched the incomplete nuclear genome of *B. natans* by BLASTn and found four distinct loci with an indel in the ITS2 region, which accounts for difficulties sequencing the antisense strand and the 3′ end of the sense strand. More rRNA cistrons may exist in the genome, as the genome is incomplete and the nuclear ITS sequence in GenBank (accession AF289035) is not identical to any of the four loci we identified. Nuclear ITS sequences were also more easily sequenced from *B. natans* strains than others, so interlocus length polymorphism may be more pronounced in other taxa.

The relationships between both nuclear and nucleomorph ITS regions (Figs. 2 and 3) were roughly as expected based on phylogenies of SSU rRNA (Silver et al. 2007), and both molecules effectively distinguish the species represented in culture. Because the majority of chlorarachniophyte species in culture are represented by only one strain (see below), we do not yet know whether either locus will effectively discriminate between chlorarachniophyte species in nature. There is cause to be optimistic that the nucleomorph ITS will remain a useful barcode marker, however, because the minimum interspecies distance is more than five times greater than the maximum intraspecies distance (Table 1). This gap is more pronounced in the nucleomorph than it is in the nucleus, mainly due to the lower intraspecies maximum distance in the nucleomorph.

Assignment of cultured strains to described species. In this study, we have determined barcode sequences for 39 out of 48 chlorarachniophyte strains in culture collections, plus *N. sphaerica*, representing >80% of available strains. From this sample, we were able to assign all but two strains to previously described species (Table S1). Our analyses suggest that CCMP 2014 could be a new species. Based on the barcode, it is most likely a member of the genus *Gymnochlora*, but this assignment requires EM to confirm that its pyrenoid is invaded by tubular invaginations of the innermost plastid membrane. Furthermore, confirmation that CCMP 2014 is morphologically distinct from *Gymnochlora stellata* would help to support our inference that the nucleomorph ITS locus can assign species to genera in addition to distinguishing species from one another. The other unassignable strain, UTEX 2631, yielded an anomalous ITS sequence with 90% sequence identity to *Nannochloropsis limnetica*, a freshwater eustigmatophyte. Because this strain was collected from the sporangiophores of the lichen-like fungus *Multiclavula corynoides*, and the plastids visible in light micrographs on the UTEX Web site (http://www.sbs.utexas.edu/utex/algaeDetail.aspx?algaeID=5115) do not show the bulbous pyrenoid characteristic of chlorarachniophytes, we conclude that this culture is not a chlorarachniophyte and should be reassigned to the Eustigmatophyceae.

We were unable to barcode 10 chlorarachniophyte strains due to inefficient DNA extraction from

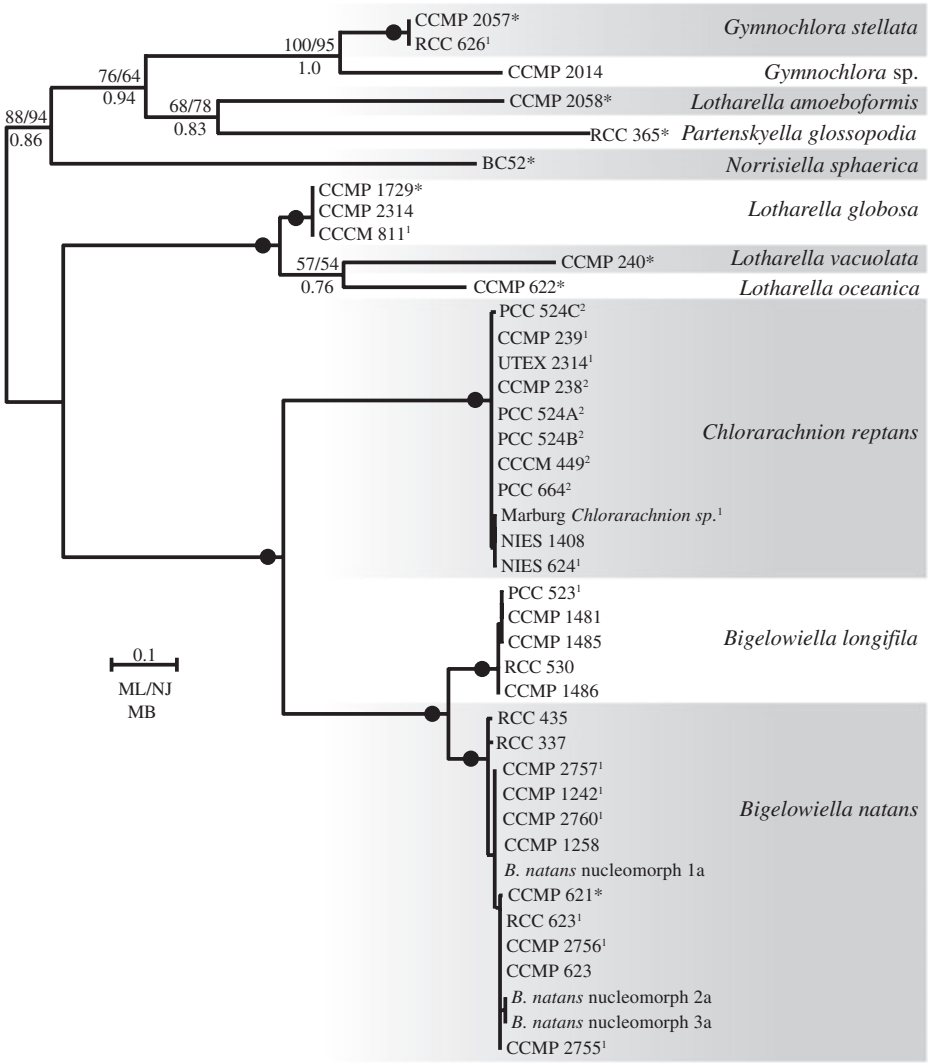

FIG. 2. Maximum-likelihood (ML) phylogenetic tree of nucleomorph-encoded internal transcribed spacer (ITS) sequences. Filled circles at nodes indicate 100% bootstrap support for both ML and distance analyses and a Bayesian posterior probability of 1.0. An asterisk indicates the type species; ([1]) indicates strain synonyms of the type species; ([2]) indicates strains from the same isolation date and location as the type species. MB, MrBayes; NJ, neighbor joining.

sparse culture samples. For six of these strains (CCMP 242, CCMP 3166, Marburg CCMP 621, CCMP 1259, CCMP 1408, and CCMP 1487), we have, nonetheless, recommended updated names. One (CCMP 242) has been formally named as *B. longifila* (Ota et al. 2007b), and two (CCMP 3166 and Marburg CCMP 621) are synonyms of the type strain for their species (*P. glossopodia* and *B. natans*)—that is, their collection records indicate that they are each directly derived from another strain. We have also recommended updated names for CCMP 1259, CCMP 1408, and CCMP 1487 mainly on the basis of previously published molecular data (Silver et al. 2007, Ota et al. 2009b), but supported in the case of CCMP 1259 and CCMP 1487 by collection information indicating they are not independent isolates from other strains we have barcoded here. For each of these six strains, therefore, we expect the nucleomorph ITS barcodes to be identical with those of their related strains. However, one or more new species may be represented among the remaining four strains that we were unable to barcode (RCC 375, RCC 376, CCMP 2285, and RCC 531). Previous analyses have indicated that RCC 375 is likely a new species of *Lotharella*, and a formal description is underway (Ota et al. 2009c). A potentially related strain, RCC 376, from the same collection date and location, has no published molecular data. Similarly, there is no molecular data for CCMP 2285, but it was collected at the same time and place as CCMP 2314, which we have here assigned to *Lotharella globosa*. Without molecular data, we cannot determine whether RCC 376 and CCMP 2285 are identical to the barcoded strains bearing the same isolation information, but in all other cases so far, strains sharing the same

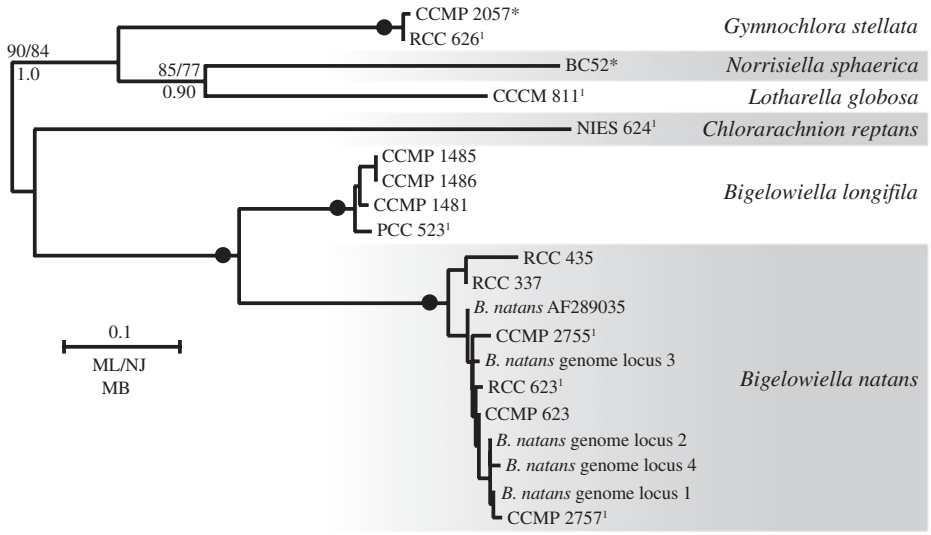

FIG. 3. Maximum-likelihood (ML) phylogenetic tree of nucleus-encoded internal transcribed spacer (ITS) sequences. Filled circles at nodes indicate 100% bootstrap support for both ML and distance analyses and a Bayesian posterior probability of 1.0. An asterisk indicates the type species; ([1]) indicates strain synonyms of the type species. MB, MrBayes; NJ, neighbor joining.

TABLE 1. Uncorrected *p* distances.

Genome	Minimum interspecies distance	Maximum intraspecies distance	Ratio
Nucleomorph	0.09886046	0.01622718	5.692957558
Nucleus	0.13570732	0.06469521	2.097640923

isolation information have yielded identical barcodes. Finally, RCC 531 represents an independent isolation for which there are no molecular or micrographical data yet available.

On the basis of barcode sequences, other published molecular data, and collection records, we here recommend names for all but three of the 48 strains in culture (Table S1), the vast majority of which belong to previously described species. Half of these (24 strains) have been previously identified to species and are correctly labeled, but 16 are currently unidentified, five are identified only to genus, and three are incorrectly labeled. We consider the 13 strains currently labeled either ''*Chlorarachnion* sp.'' or ''*Chlorarachnion* cf. sp.'' to belong to the unidentified category, rather than identified to genus, as these names were equivalent to ''unidentified chlorarachniophyte'' until 2001 when a second genus was described, and each of these cultures was deposited before 2000. Only three strains were determined to be incorrectly named according to our data. The barcode for CCMP 2314, currently labeled ''*Chlorarachnion reptans,*'' is identical to that of *L. globosa*, and nuclear and nucleomorph SSU sequences support this designation (Silver et al. 2007), so we recommend that this strain be renamed. Another misidentification, UTEX 2631, is actually a eustigmatophyte. Finally, the Culture Collection of Algae at Philipps University in Marburg holds a strain derived from CCMP 621 labeled *C. reptans* that should be updated to *B. natans*. Altogether, the number of potentially new chlorarachniophyte species present among the unidentified strains in culture collections is between two (CCMP 2014 and RCC 375, descriptions underway) and five (if all of the strains lacking molecular data turn out to be new species), bringing the total number of cultured species to between 14 and 17, a surprisingly low range given the total of 48 strains.

Aside from the number of species, how much diversity is represented among the chlorarachniophytes in culture? Many strains are synonymous, meaning that they were directly derived from another strain, either to be shared with another culture collection or to establish a clonal strain. In addition, many strains share the same collection date and location, so although they are not strictly synonyms, they were likely derived from the same population of chlorarachniophytes. Taken together, only 23 out of the 48 strains can be considered truly independent isolates. This is reflected in the unexpectedly high sequence similarity among nucleomorph ITS barcodes from each species (Fig. 2). For the most extreme example, *C. reptans*, there are 11 strains in culture, of which nine were collected at Puerto Peñasco, Mexico, by Richard Norris, on June 20, 1966. Four of these are labeled as strain synonyms of the type culture, CCAP 815/1, now deceased, although it is possible that all 10 derive from the same collection vessel. One, NIES 1408, was collected independently at Amami Island, Japan (Fig. 2, Table S1), and another was inferred to have been collected in Tunisia, as it was deposited

by K. Grell, who published observations of *C. reptans* collected at the Gulf of Hammamet (Grell 1990). Although only three independent isolations cannot provide insight into within-species genetic diversity, the nucleomorph ITS sequence similarity among the Japan, Tunisia, and Mexico isolations shows that this species is widespread. Likewise, there are 13 strains of *B. natans* in culture, derived from only four independent isolations. Six of these strains are synonymous with the type strain CCMP 621 according to the collection records. Two separate isolations from the Mediterranean Sea, RCC 337 and RCC 435, show some genetic differentiation relative to the remaining strains, all of which were isolated from the Sargasso Sea (Fig. 2). The eight strains of *B. longifila* are derived from five independent isolations, although four of these came from the Sargasso Sea, and the lone Mediterranean Sea collection, RCC 530, shows no sequence divergence. *L. globosa* has been isolated from the coasts of both France and Guam. The remaining seven chlorarachniophyte species in culture (seven out of 11, including the two strains whose descriptions are underway) are represented by one isolation each, thereby providing no information as to their diversity or distribution.

CONCLUSIONS

Nucleomorph ITS sequences are sufficiently homogeneous within a culture to be efficiently amplified and sequenced, and they contain sufficient variability to distinguish chlorarachniophyte species. This will be a useful tool for future species identifications, as it removes the need for EM. Based on these data, other published molecular data, and inferences from strain information, we have recommended updated names for 21 strains that are currently unidentified, identified only to genus, or, in only three cases, misidentified. There are fewer distinct chlorarachniophytes in culture than the number of strains would suggest, due mainly to the same strain being held at multiple culture collections, but also due to strains derived from the same collection. At least some chlorarachniophyte species are widely distributed, but an understanding of chlorarachniophyte diversity awaits future environmental sequencing studies.

We thank Ken Ishida for providing a culture of *N. sphaerica* from his private collection; Tia Silver and John Archibald for sharing the nucleomorph ribosomal cistron sequence from *Lotharella oceanica* CCMP 622; the Joint Genome Institute's Community Sequencing Program (http://www.jgi.doe.gov/sequencing/why/50026.html) for their ongoing efforts to sequence the nuclear genomes of *Guillardia theta* and *B. natans*; K. Barry and E. Lindquist of the JGI for project management and data availability; and M. W. Gray, J. Archibald, G. I. McFadden, and C. E. Lane for their contributions to the project. We also thank the Provasoli-Guillard National Center for Culture of Marine Phytoplankton, USA; the Roscoff Culture Collection, France; the Plymouth Culture Collection of Marine Algae, UK; the Culture Collection of Algae at the University of Texas, USA; the Culture Collection of Algae, Philipps-University Marburg, Germany; the Canadian Centre for the Culture of Microorganisms, Canada; and the Microbial Culture Collection at the National Institute for Environmental Studies, Japan for their kind donation of cultures for this project. This work was supported by a grant from Genome Canada to P. J. K. as part of the Canadian Barcode of Life Network. G. H. G. is supported by a postgraduate doctoral fellowship from the Natural Sciences and Engineering Research Council of Canada. P. J. K. is a Fellow of the Canadian Institute for Advanced Research and a Senior Investigator of the Michael Smith Foundation for Health Research.

Archibald, J. M., Rogers, M. B., Toop, M., Ishida, K. & Keeling, P. J. 2003. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc. Natl. Acad. Sci. U. S. A.* 100:7678–83.

Bhattacharya, D., Helmchen, T. & Melkonian, M. 1995. Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphina and the Chlorarachniophyta. *J. Eukaryot. Microbiol.* 42:65–9.

Calderon-Saenz, E. & Schnetter, R. 1987. *Cryptochlora perforans*, a new genus and species of algae (Chlorarachniophyta), capable of penetrating dead algal filaments. *Plant Syst. Evol.* 158:69–71.

Calderon-Saenz, E. & Schnetter, R. 1989. Morphology, biology, and systematics of *Cryptochlora perforans* (Chlorarachniophyta), a phagotrophic marine alga. *Plant Syst. Evol.* 163:165–76.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540–52.

Cavalier-Smith, T. & Chao, E. E. 1996. Sarcomonad ribosomal RNA sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. *Arch. Protistenkd.* 147:227–36.

Chantangsi, C., Lynn, D. H., Brandl, M. T., Cole, J. C., Hetrick, N. & Ikonomi, P. 2007. Barcoding ciliates: a comprehensive study of 75 isolates of the genus *Tetrahymena*. *Int. J. Syst. Evol. Microbiol.* 57:2412–25.

Chase, M. W., Salamin, N., Wilkinson, M., Dunwell, J. M., Kesanakurthi, R. P., Haidar, N. & Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360:1889–95.

Evans, K. M., Wortley, A. H. & Mann, D. G. 2007. An assessment of potential diatom ''barcode'' genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* 158:349–64.

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M. & Barrett, S. C. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 3:e2802.

Geitler, L. 1930. Ein grünes Filarplasmodium und andere neue Protisten. *Arch. Protisenkd.* 69:221–30.

Gile, G. H. & Keeling, P. J. 2008. Nucleus-encoded periplastid-targeted EFL in chlorarachniophytes. *Mol. Biol. Evol.* 25:1967–77.

Gilson, P. R. & McFadden, G. I. 1996. The miniaturized nuclear genome of eukaryotic endosymbiont contains genes that overlap, genes that are cotranscribed, and the smallest known spliceosomal introns. *Proc. Natl. Acad. Sci. U. S. A.* 93:7737–42.

Gilson, P. R. & McFadden, G. I. 2002. Jam packed genomes – a preliminary, comparative analysis of nucleomorphs. *Genetica* 115:13–28.

Gilson, P. R., Su, V., Slamovits, C. H., Reith, M. E., Keeling, P. J. & McFadden, G. I. 2006. Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 103:9566–71.

Gottschling, M., Keupp, H., Ptoener, J., Knop, R., Willems, H. & Kirsch, M. 2005. Phylogeny of calcareous dinoflagellates as inferred from ITS and ribosomal sequence data. *Mol. Phylogenet. Evol.* 36:444–55.

Grell, K. G. 1990. Some light microscope observations on *Chlorarachnion reptans* Geitler. *Arch. Protistenkd.* 138:271–90.
Guindon, S. & Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696–704.
Hebert, P. D., Cywinska, A., Ball, S. L. & deWaard, J. R. 2003a. Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 270:313–21.
Hebert, P. D., Ratnasingham, S. & deWaard, J. R. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B* 270:S96–9.
Hibberd, D. J. & Norris, R. E. 1984. Cytology and ultrastructure of *Chlorarachnion reptans* (Chlorarachniophyta divisio nova, Chlorarachniophyceae classis nova). *J. Phycol.* 20:310–30.
Huelsenbeck, J. P. & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–5.
Inagaki, Y., Ehara, M., Watanabe, K. I., Hayashi-Ishimaru, Y. & Ohama, T. 1998. Directionally evolving genetic code: the UGA codon from stop to tryptophan in mitochondria. *J. Mol. Evol.* 47:378–84.
Ishida, K., Cao, Y., Hasegawa, M., Okada, N. & Hara, Y. 1997. The origin of chlorarachniophyte plastids, as inferred from phylogenetic comparisons of amino acid sequences of EF-Tu. *J. Mol. Evol.* 45:682–7.
Ishida, K., Ishida, N. & Hara, Y. 2000. *Lotharella amoeboformis* sp. nov.: a new species of chlorarachniophytes from Japan. *Phycol. Res.* 48:221–9.
Ishida, K., Nakayama, T. & Hara, Y. 1996. Taxonomic studies on the Chlorarachniophyta. II. Generic delimitation of the chlorarachniophytes and description of *Gymnochlora stellata* gen. et sp. nov. and *Lotharella* gen. nov. *Phycol. Res.* 44:37–45.
Keeling, P. J. 2001. Foraminifera and Cercozoa are related in actin phylogeny: two orphans find a home? *Mol. Biol. Evol.* 18: 1551–7.
Kucera, H. & Saunders, G. W. 2008. Assigning morphological variants of *Fucus* (Fucales, Phaeophyceae) in Canadian waters to recognized species using DNA barcoding. *Botany-Botanique* 86:1065–79.
Lane, C. E., Lindstrom, S. C. & Saunders, G. W. 2007. A molecular assessment of northeast Pacific *Alaria* species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding. *Mol. Phylogenet. Evol.* 44:634–48.
Li, H. C., Bouchara, J. P., Hsu, M. M., Barton, R., Su, S. & Chang, T. C. 2008. Identification of dermatophytes by sequence analysis of the rRNA gene internal transcribed spacer regions. *J. Med. Microbiol.* 57:592–600.
Ludwig, M. & Gibbs, S. P. 1989. Evidence that the nucleomorphs of *Chlorarachnion reptans* (Chlorarachniophyceae) are vestigial nuclei: morphology, division, and DNA-DAPI fluorescence. *J. Phycol.* 25:385–94.
Moestrup, Ø. & Sengco, M. 2001. Ultrastructural studies on *Bigelowiella natans*, gen. et sp. nov., a chlorarachniophyte flagellate. *J. Phycol.* 37:624–46.
Nguyen, H. D. T. & Seifert, K. A. 2008. Description and DNA barcoding of three new species of *Leohumicola* from South Africa and the United States. *Persoonia* 21:57–69.
Nikolaev, S. I., Berney, C., Fahrni, J. F., Bolivar, I., Polet, S., Mylnikov, A. P., Aleshin, V. V., Petrov, N. B. & Pawlowski, J. 2004. The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 101:8066–71.
Ortega, A., Suarez-Santiago, V. N. & Reyes, J. D. 2008. Morphological and ITS identification of *Cortinarius* species (section Calochroi) collected in Mediterranean *Quercus* woodlands. *Fungal Divers.* 29:73–88.
Ota, S., Silver, T. D., Archibald, J. M. & Ishida, K. 2009a. *Lotharella oceanica* sp. nov., a new planktonic chlorarachniophyte studied by light and electron microscopy. *Phycologia* 48:315–23.
Ota, S., Ueda, K. & Ishida, K. 2005. *Lotharella vacuolata* sp. nov., a new species of chlorarachniophyte algae, and time-lapse video observations on its unique post-cell division behavior. *Phycol. Res.* 53:275–86.
Ota, S., Ueda, K. & Ishida, K. 2007a. *Norrisiella sphaerica* gen. et sp. nov., a new coccoid chlorarachniophyte from Baja California, Mexico. *J. Plant. Res.* 120:661–70.
Ota, S., Ueda, K. & Ishida, K. 2007b. Taxonomic study of *Bigelowiella longifila* sp. nov. (Chlorarachniophyta) and a time-lapse video observation of the unique migration of amoeboid cells. *J. Phycol.* 43:333–43.
Ota, S., Vaulot, D., Gall, F. L., Yabuki, A. & Ishida, K. 2009b. *Partenskyella glossopodia* gen. et sp. nov., the first report of a chlorarachniophyte that lacks a pyrenoid. *Protist* 160:137–50.
Ota, S., Viprey, M. & Vaulot, D. 2009c. Diversity of chlorarachniophytes in the Mediterranean Sea. *Phycologia* 48:100 (meeting abstract).
Park, M. H., Sim, C. J., Baek, J. & Min, G. S. 2007. Identification of genes suitable for DNA barcoding of morphologically indistinguishable Korean Halichondriidae sponges. *Mol. Cells* 23: 220–7.
Posada, D. & Buckley, T. R. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53:793–808.
Posada, D. & Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–8.
Rogers, M. B., Archibald, J. M., Field, M. A., Li, C., Striepen, B. & Keeling, P. J. 2004. Plastid-targeting peptides from the chlorarachniophyte *Bigelowiella natans*. *J. Eukaryot. Microbiol.* 51: 529–35.
Rogers, M. B., Gilson, P. R., Su, V., McFadden, G. I. & Keeling, P. J. 2007. The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol. Biol. Evol.* 24:54–62.
Saunders, G. W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360:1879–88.
Silver, T. D., Koike, S., Yabuki, A., Kofuji, R., Archibald, J. M. & Ishida, K. 2007. Phylogeny and nucleomorph karyotype diversity of chlorarachniophyte algae. *J. Eukaryot. Microbiol.* 54:403–10.
Sinniger, F., Reimer, J. D. & Pawlowski, J. 2008. Potential of DNA sequences to identify zoanthids (Cnidaria: Zoantharia). *Zool. Sci.* 25:1253–60.
Swofford, D. L. 2002. *PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods). 4.0.* Sinauer, Sunderland, Massachusetts.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876–82.
White, T. J., Bruns, T., Lee, S. & Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. [Eds.] *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Diego, California, pp. 315–22.

Supplementary Material

The following supplementary material is available for this article:

Table S1. Chlorarachniophyte strain information and recommended names.

This material is available as part of the online article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.