Microbial Ecology

Molecular Diversity among Communities of Freshwater Microchlorophytes

M.W. Fawley[1], K.P. Fawley[1] and M.A. Buchheim[2]

(1) Department of Biological Sciences, North Dakota State University, Fargo, ND 58105-5517, USA
(2) Department of Biological Science, The University of Tulsa, Tulsa, OK 74104-3189, USA

Received: 30 September 2003 / Accepted: 29 January 2004 / Online publication: 9 November 2004

Abstract

Current hypotheses on the distribution of freshwater microchlorophytes lead to predictions of low diversity and wide environmental tolerances. Thus, the same few species should be found worldwide in many different habitats. However, these hypotheses are based on a morphospecies concept, which precludes the possibility of numerous cryptic species among these organisms. In this study, we examined the diversity of coccoid green microalgae and chlamydomonads (Chlorophyta) isolated from sites in Minnesota and North Dakota (USA) using techniques of 18S rDNA sequence analysis. Of 93 distinct 18S rDNA sequences identified from among 273 isolates examined by molecular techniques, all but four are new to science. The spatial distribution of organisms represented by these 18S rDNA sequences was not uniform, because some lakes and ponds yielded distinct 18S rDNA types not found at other sites. In addition, organisms generally considered to be cosmopolitan, such as *Chlamydomonas reinhardtii* and *Chlorella vulgaris*, were not found. These results challenge predictions of low species number and wide environmental tolerances among these eukaryotic microorganisms.

Introduction

The diversity of eukaryotic microorganisms is generally considered to be lower than that of many other organisms. For example, Finlay [14] contrasts the estimate of 10,000 to 20,000 species of free-living protozoa with ~5,000,000 estimated species of insects. Ubiquitous dispersal and wide environmental tolerances could explain the low diversity of eukaryotic microorganisms. With no geographic and few ecological barriers to gene flow, allopatric speciation is essentially impossible, and low species diversity and cosmopolitan distributions result [14]. However, not all researchers agree with Finlay's conclusions, and estimates of the number of species of microeukaryotes vary widely. For example, only about 12,000 species of diatoms have been described, but estimates range from 100,000 to 10,000,000 total species [1].

Estimating the diversity of microorganisms is difficult because of the lack of a definitive, operational species concept [23]. A morphological species concept is frequently used for eukaryotic microorganisms, but molecular data have also indicated patterns of local adaptation and endemism among some algae where morphological differences are difficult to detect [9]. Unfortunately, we know little about niche variation among sibling species of microorganisms, and as a result, the molecular data provide little information about local adaptation and speciation [15]. In order to clarify species concepts for microorganisms, we must gain a better understanding of their molecular diversity and ultimately determine if these differences are associated with unique niches.

Many microalgae possess few morphological characteristics that are useful for species characterization, leading to the possibility of numerous cryptic species. For example, some of the most commonly reported microalgae are coccoid organisms, often referred to as "little green balls" or "little round green things" [7]. These organisms are extremely difficult to identify because of their small size (often <5 μm) and simple morphologies (spherical or nearly so). As a result, many floristic studies, if they attempt to identify these organisms at all, often report the "little green balls" as *Chlorella vulgaris* Beij. or *Chlorella* spp. On the other hand, some microalgae, such as the green algal genus *Scenedesmus*, have previously been divided into many species based upon minor differences in morphology [19]. Some of these differences

Correspondence to: M.W. Fawley; E-mail: marvin.fawley@ndsu.nodak.edu

have subsequently been shown to represent environmentally induced plasticity and are unreliable for delimiting species [40].

Although the morphological species concept is problematic, it has been retained for microalgae because it is highly operational; that is, if this species concept is unequivocally applied, then an investigator can efficiently use microscopy to provide species names for microalgae. When this approach is applied in floristic studies, morphologically distinct microalgae are often characterized as minor variants of the most commonly described species. As a result, many species of freshwater and soil algae are thought to have world wide distributions and many genera of common algae have relatively few described species. Finlay's [14] hypotheses of cosmopolitan distribution, low diversity, and environmental tolerance are dependent on the "lumping" of these minor variants into a single species. However, minor morphological variants may, in fact, represent new species, and therefore many species may not be as widely distributed as the literature suggests.

Molecular techniques allow an assessment of the validity of the morphological species concept for common microalgae and ultimately niche partitioning and distribution of these organisms. For many types of microorganisms, the gene most commonly employed for diversity studies is the small-subunit ribosomal RNA gene (16S rDNA in prokaryotes and 18S rDNA in eukaryotes). However, studies of land plants have indicated that analyses of the more rapidly evolving sequences, such as ribosomal ITS and ETS regions, are usually necessary to resolve species-level phylogenies [2, 30]. Thus, differences of even a single nucleotide in the 18S rDNA sequence may differentiate distinct species. On the other hand, variations in ribosomal DNA sequences are likely to represent neutral mutations, and it will ultimately be necessary to correlate these variations to functional, reproductive, or morphological variations in order to employ sequence data as species markers. Until these studies can be performed, however, molecular variations can be used as markers that allow the study of both molecular diversity and distribution patterns of organisms that lack recognized stable morphological characters useful to differentiate species.

For the microalgae, some studies have characterized morphologically similar isolates as the same species even if these isolates exhibit 18S rDNA variability (e.g., [37]). In contrast, other studies have shown that 18S rDNA identity alone does not indicate that isolates are conspecific and that more rapidly evolving sequences (e.g., ITS, *rbc*L) are necessary for species characterization (e.g., [8, 22, 28, 38]).

We are studying the diversity of "little green balls" and other microalgae of freshwater lakes and ponds. These organisms are found in several protistan lineages, but our main focus here is the green algae (Chlorophyta). Happey-Wood [18] proposed hypotheses on the diversity and distribution of green microalgae based upon studies that employed a morphological species concept and identification using light microscopy. These hypotheses are based upon the assumption that very few species of these algae occur in a wide range of environmental conditions. Thus, the same few species of microalgae would be expected in eutrophic, mesotrophic, dystrophic, and oligotrophic waters, consistent with Finlay's [14] assessment of the diversity and distribution of eukaryotic microorganisms in general.

Our approach to test the hypotheses of Happey-Wood [18] and Finlay [14] is to isolate algae from ecologically distinct, freshwater sites in Itasca State Park, MN (USA) and Arrowwood National Wildlife Refuge, ND (USA) and characterize the isolates by 18S rDNA sequence analysis. According to the hypotheses of Finlay [14] and Happey-Wood [18], the overall diversity should be low and the same organisms should be widely distributed in a variety of freshwater habitats. In addition, these hypotheses lead to predictions that many of the organisms we isolate will share 18S rDNA sequences with algae that are already in culture from other locations, especially organisms that are considered cosmopolitan, such as *Chlamydomonas reinhardtii* Dang. and *Chlorella vulgaris*. We chose 18S rDNA sequence analysis for this initial study because it is the only gene with an extensive data set for green algae. The use of other, more rapidly evolving DNA, such as *rbc*L and the ribosomal RNA ITS regions, will ultimately be necessary for a more complete understanding of the diversity of these organisms.

Methods

Water samples were collected from Arrowwood Lake, Mud Lake, and Jim Lake in Arrowwood National Wildlife Refuge, Pingree, North Dakota (USA) during the winters of 1994–1997 and throughout the year 1995. Samples were also taken from seven sites in Itasca State Park, Minnesota (USA), four times during 2000–2001. Table 1 provides basic information on sites. For a more complete description of the sites and their algal floras, see Meyer and Brook [32] and Phillips and Fawley [34–36]. Sites in Arrowwood National Wildlife Refuge and Itasca State Park are all located between approximately 47°10′ and 47°16′ N latitude, with Arrowwood 280 km west of Itasca.

Phytoplankton samples were collected from surface grabs from all sites during open water periods, and from just below the ice during periods of ice cover. Additional samples were collected from just above and just below the thermocline in Mary Lake when thermal stratification was present, and from mid-depth and 1 m above the bottom from Mary Lake when stratification did not occur

Table 1. Physical characteristics of sites included in this study

Site	*Lattitude/Longitude*	*Area (hectares)*	*pH*	*Conductivity (μS)*	*Nitrogen (mg/L)*	*SRP (mg/L)*
Arrowwood National Wildlife Refuge						
Arrowwood Lake (mesotrophic)	47 15.45′ N, 98 51.37′ W	648	7.6	542	0.20	0.35
Mud Lake (mesotrophic)	47 13.12′ N, 98 50.97′ W	190	7.8	542	0.34	0.36
Jim Lake (mesotrophic)	47 08.40′ N, 98 46.62′ W	350	8.0	542	0.11	0.29
Itasca State Park						
Lake Itasca (mesotrophic)	47 14.05′ N, 95 12.10′ W	437	8.2	326	0.07	0.05
Mary Lake (mesotrophic)	47 11.25′ N, 95 10.05′ W	22.6	8.2	278	0.07	0.06
"Picnic Pond" (eutrophic)	47 14.41′ N, 95 12.15′ W	<1	7.3	359	0.06	0.20
"Tower Pond" (dystrophic)	47 11.41′ N, 95 10.84′ W	<1	7.2	35	0.06	<0.01
North Deming Pond (dystrophic)	47 10.28′ N, 95 09.98′ W	<1	7.4	97	0.07	<0.01
West Twin Lake (dystrophic)	47 10.52′ N, 95 09.99′ W	2.5	7.8	256	0.08	0.04
Bog D	47 10.63′ N, 95 09.93′ W	ND	7.7	661	0.08	0.32

Nitrogen is nitrate + nitrite; SRP: soluble reactive phosphorus; ND: not determined. Conductivities for Arrowwood sites are summer means for all sites from Hanson [17].

and for all samples from Lake Itasca, where stratification was not detected. Tychoplankton samples were collected from Itasca sites by squeezing submerged vegetation. No tychoplankton samples were collected from Arrowwood sites or during the winter. Winter phytoplankton samples were not collected from Picnic Pond or BogD, which freeze through the water column. Algae were isolated from samples as previously described [34]. Isolations from Arrowwood focused only on coccoid organisms, whereas Itasca isolations were more general and included chlamydomonads. A total of 908 isolates from Itasca and 210 isolates from Arrowwood were acquired by these methods. This study focused on two groups of these isolates: (1) coccoid isolates from phytoplankton samples from both Arrowwood and ltasca, and (2) chlamydomonad isolates from both phytoplankton and tychoplankton from Itasca sites only. A few coccoid tychoplankton from Itasca have also been examined, but the bulk of the isolates from tychoplankton samples and some coccoid phytoplankton remain to be characterized. Thus far, 273 isolates have been characterized by 18S rDNA sequence analysis and/or PCR-RFLP of the 18S rDNA, including 61 chlamydomonad isolates and 212 coccoid isolates. All characterized isolates are available from the authors.

DNA was purified from each algal isolate and 18S rDNA was amplified as previously described [4, 5, 12]. Coccoid isolates were initially screened by PCR-RFLP of the 18S rDNA [34]. Sequences of the 18S rDNA were determined for coccoid isolates with distinctive RFLP patterns. If multiple coccoid isolates produced identical PCR-RFLP fragments, then sequences were generated from more than one of the isolates of each PCR-RFLP type. The 18S rDNA sequences were also determined for all 61 chlamydomonad isolates. A Beckman CEQ 2000 or ABI 373 Stretch automated DNA sequencer were used to generate sequences, according to manufacturers' instructions, and both strands were sequenced using standard primers [4, 5, 13]. All distinct 18S rDNA sequences generated are available from GenBank. Sequences used for analyses varied somewhat in length, depending on the template and primer sets used. Trebouxiophycean sequences (coding region only) ranged from 1740 to 1798 nucleotides, with a mean of 1793. Sphaeroplealean sequences ranged from 1756 to 1794 nucleotides, with a mean of 1782. The raw chlamydomonad sequences ranged from 1672 to 1758 nucleotides with a mean length of 1732 nucleotides. The chlamydomonad alignment used for this analysis was truncated to 1693 nucleotide sites because of missing data for a number of published sequences.

The 18S rDNA sequences were manually aligned with published sequences using MacClade 4.03 [31]. The alignments included all published sequences that were most similar to those of our coccoid and chlamydomonad isolates. All putative introns were excluded from the alignments. All sites were included for the analyses despite the fact that some regions could not be unequivocally aligned over the complete data set; however, these regions could be aligned at a more local level. If these

highly variable sites had not been included in the alignment, many of the differences between very similar sequences would have been excluded from subsequent analyses. Distance matrices were generated in PAUP* 4.0b10 [39] using absolute differences and with gaps ignored for pairwise comparisons. Neighbor-joining phenograms were constructed using PAUP* 4.0b10. Three separate analyses were performed for the Trebouxiophyceae, Sphaeropleales, and Chlamydomonadales. The phenograms (Figs. 1, 2, and 3) were rooted using data from members of sister clades, the Ulvophyceae, Chlamydomonadales, and Sphaeropleales, respectively. These analyses were designed to assess the sequence diversity among our isolates. Therefore, results are not to be interpreted as thorough phylogenetic analyses and bootstrap analyses or other tests of the robustness of lineages are not presented. Future studies that focus on individual genera or groups of genera will employ more critical phylogenetic analyses.

Results

We have thus far characterized 273 isolates of green algae with 93 distinct 18S rDNA sequences. Analyses of these sequences indicated that the isolates are distributed among the green algal classes Chlorophyceae and Trebouxiophyceae (Figs. 1, 2, and 3). Of the 93 distinct sequences, 29 were Trebouxiophyceae (Fig. 1), 15 were allied with the Sphaeropleales (Chlorophyceae, Fig. 2), and 49 were Chlamydomonadales (Chlorophyceae, Fig. 3). Only four of the 93 distinct sequences matched any sequence published in GenBank.

Most of the distinct trebouxiophycean sequences were similar to published sequences of *Chlorella, Nannochloris*, and *Choricystis*, but some sequences were resolved in the *Oocystis* lineage (Fig. 1). One sequence, that of Tow 9/21 P-16w, was profoundly different from any published sequence, with 93 substitutions and three indels different from the most similar sequence. Only two sequences from trebouxiophycean isolates exactly matched published sequences, those of JL11-11, which was identical to the published sequence of *Nannochloris coccoides* Naumann, and AS5-1, which matched the published sequence of *Choricystis* sp. Two isolates, MDL5-9 and Itas 9/21 14-5w, had identical 18S rRNA coding sequences, but MDL5-9 possessed an intron that Itas 9/21 14-5w lacked. The shared coding sequence from these two isolates differed from that of the published sequence of *Choricystis minor* (Skuja) Fott by only one indel.

For the Sphaeropleales and allied lineages (Chorophyceae, Fig. 2), the sequences of several coccoid isolates were most similar to *Desmodesmus* and *Scenedesmus* species. Four sequences were also similar to *Mychonastes* and *Pseudodictyosphaerium* species. The 18S rDNA sequences from Mary 9/21 T-3w and Itas 9/21 14-1d differed from published sequences by >40 substitutions, and the sequence from Tow 2/24 P-12d had >100 differences from any published sequence, None of the sequences from sphaeroplealean taxa were identical to any published sequences; however, Tow 9/21 P-1w and AS 7-9 differed from published sequences by only a single indel.

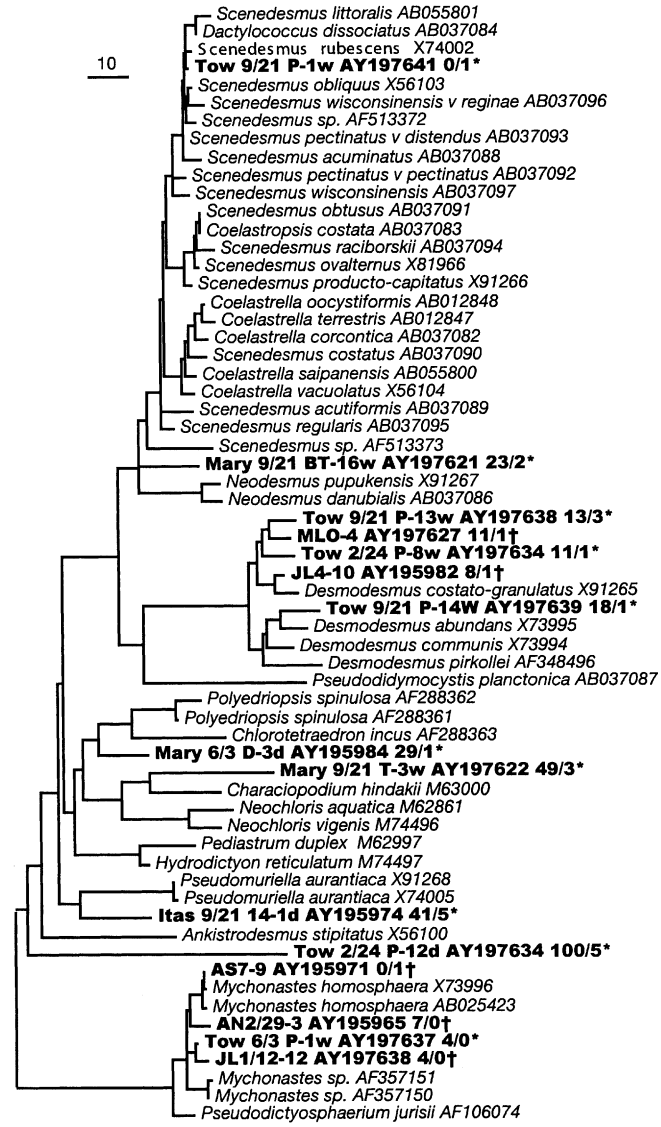

Figure 1. Phenogram from neighbor-joining analyses of 69 trebouxiophycean 18S rDNA sequences. The phenogram was rooted using data from members of a sister clade, the Ulvophyceae. Branch lengths are proportional to absolute nucleotide differences (see scale). Each terminal branch is labeled with the isolate name, the accession number for the 18S sequence of that isolate, and the number of nucleotide substitutions and indels associated with the most similar isolate as identified from pairwise analysis of absolute distances. New isolates are indicated by bold font, with Itasca isolates indicated by an asterisk and isolates from Arrowwood indicated by a cross. All other isolates derive from published 18S rDNA data and are labeled with GenBank accession numbers.

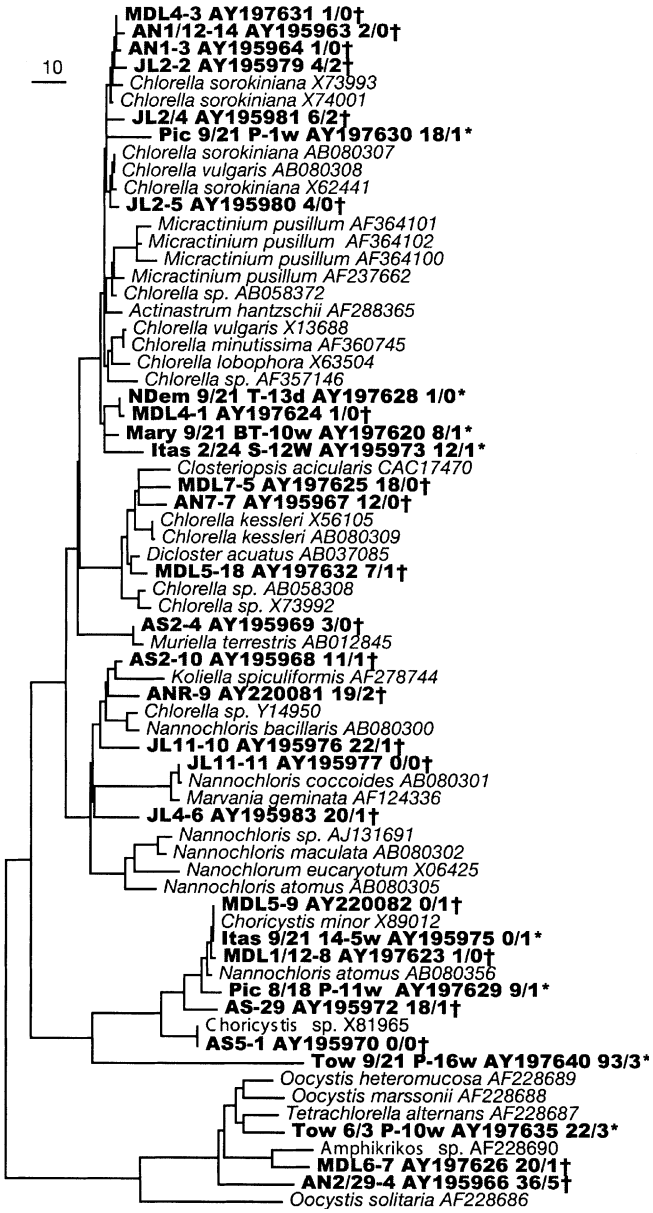

Figure 2. Phenogram from neighbor-joining analysis of 62 sphaeroplealean 18S rDNA sequences. The phenogram was rooted using data from members of a sister clade, the Chlamydomonadales. Branch lengths are proportional to absolute nucleotide differences (see scale). Labeling as in Fig. 1.

More sequence diversity was detected among chlamydomonad isolates (Fig. 3) than for the Trebouxiophyceae and Sphaeropleales combined. Several of the isolates had sequences similar to *Lobochlamys* species, including two isolates that had sequences identical to the published sequence of *Lobochlamys segnis* (Ettl) Pröschold, Marin, Schlösser et Melkonian. Two isolates with identical sequences were similar to *Chloromonas* species. One isolate, BogD T-1d, had a sequence identical to the published sequence of *Heterochlamydomonas inaequalis* Cox et Deason. The sequences of almost all of the other isolates differed from any published sequences by at least 20 nucleotides. The 18S rDNA sequence of NDem 9/21 T-11d differed from the most similar published sequence, that of *Chlamydomonas pseudogloeogama* Gerloff, by 95 substitutions and four indels.

All of the sequences generated from the 61 chlamydomonad isolates are represented in Fig. 3. Of these 61 sequences, 49 are distinct. The Chlamydomonadales were not specifically isolated from Arrowwood samples and therefore only a single Arrowwood chlamydomonad isolate was characterized. As a result, the chlamydomonad community of the Arrowwood lakes cannot be compared to the community from the Itasca sites. For the chlamydomonad isolates from Itasca, identical sequences were obtained from multiple isolates from the same site six times, but identical sequences from isolates from different sites were only found twice. However, these data are insufficient for critical analysis of the distribution of these organisms among the Itasca sites.

Culturing methods for the coccoid algae were similar for both Arrowwood and Itasca sites. For the 44 distinct 18S rDNA sequences that were generated from coccoid algae, only nine sequences were identical for three or more isolates (Table 2). Of these nine sequences, only two were from both Arrowwood and Itasca isolates. Some sequences types that were very common among Itasca isolates, such as Itas 2/24 S-12 W, Itas 9/21 14-5w, and Tow 6/3 P-1w, were either rare or unknown among Arrowwood isolates. Conversely, isolates with the sequences represented by ANR-9, AS-29, and JL1/12-12 were frequently cultured from Arrowwood sites, but never from Itasca sites.

It is also interesting that entire lineages that were well represented in the Arrowwood sites were totally absent from the Itasca sites. Sequences similar to the sequences of *Chlorella kessleri* Fott et Nováková (Trebouxiophyceae) and related organisms, represented by MDL7-5, AN7-7, and MDL5-18 (Fig. 1), were found for six isolates from Arrowwood, but no isolates from this lineage were cultured from Itasca sites. Also, 10 isolates, with six distinct sequences, from the *Nannochloris* lineage (Trebouxiophyceae) were cultured from Arrowwood sites, but no organisms from this lineage were isolated from Itasca sites.

The distribution of coccoid isolates with particular sequences also varied among the Itasca sites. In particular, some sequence types, which were frequently isolated from Lake Itasca, were rarely or never isolated from other sites. Twenty-three isolates with an 18S rDNA sequence identical to that of Tow 6/3 P-1w (Chlorophyceae, Fig. 2) were cultured from Lake Itasca, but only three were cultured from other Itasca sites. All 20 Itasca isolates of the trebouxiophycean represented by Itas 2/24 S-12w were from Lake Itasca.

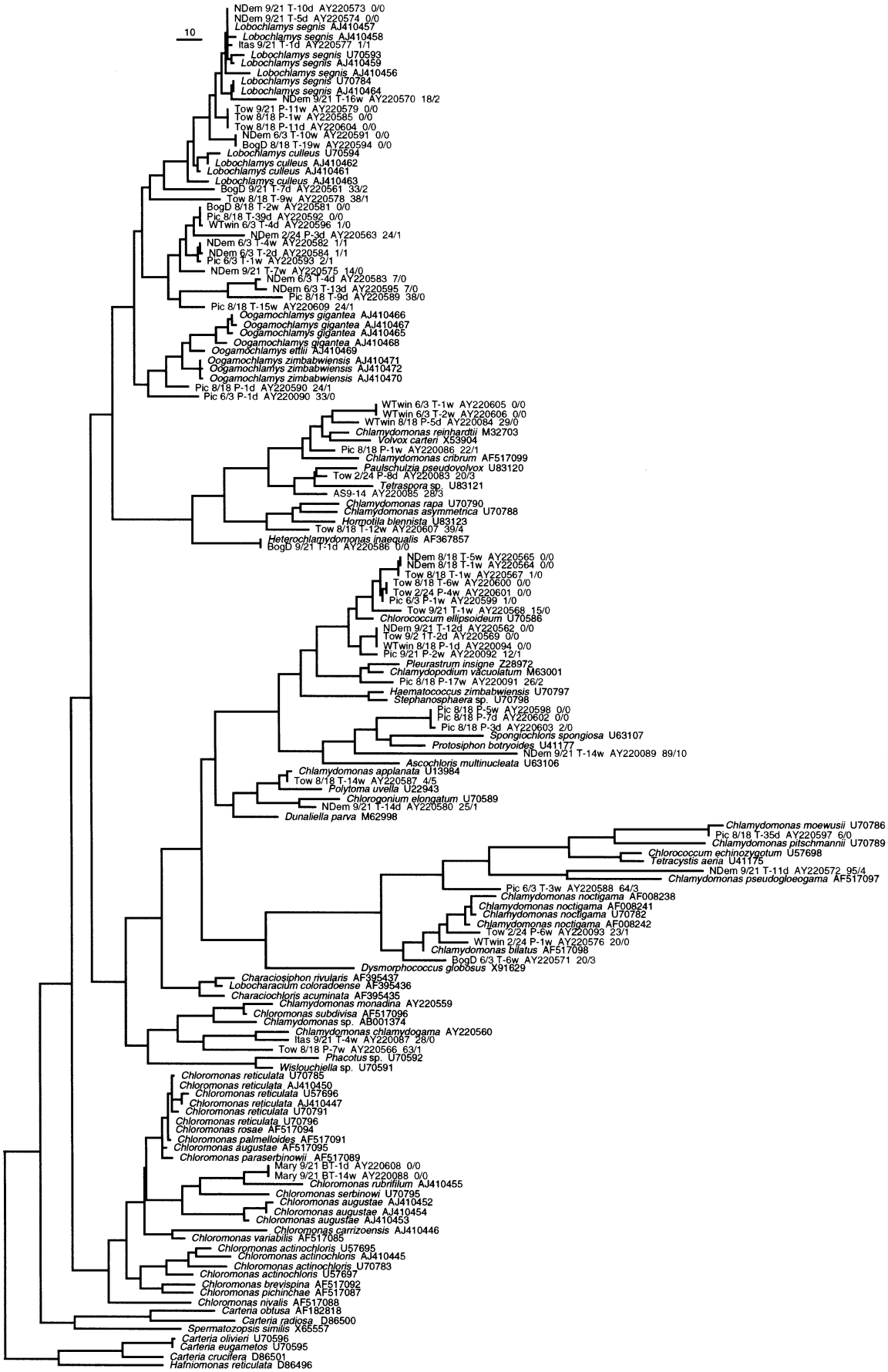

Figure 3. Phenogram from neighbor-joining analysis of 152 chlorophycean 18S rDNA sequences. The phenogram was rooted using data from a member of a sister clade (*Ankistrodesmus*, Sphaeropleales). Branch lengths are proportional to absolute nucleotide differences (see scale). Labeling as in Fig. 1.

Table 2. Number of coccoid isolates from the Itasca State Park and Arrowwood National Wildlife Refuge sites that possessed a common 18S rDNA sequence[a]

Tentative identification and representative isolate	*Number of isolates with identical sequences*	
	Itasca sites	*Arrowwood sites*
Trebouxiophyceae		
Chlorella sp., NDem 9/21 T-13d	3	0
Chlorella sp., Mary 9/21 BT-10w	1	4
Chlorella sp., Itas 2/24 S-12w	20	4
Chlorella sp., AN7-7	0	3
Nannochloris sp., ANR-9	0	5
Choricystis sp., Itas 9/21 14-5w	25	0
Choricystis sp., AS-29	0	7
Chlorophyceae		
Mychonastes sp., Tow 6/3 P-1w	26	0
Mychonastes sp., JL1/12-12	0	16

[a]Only sequences found in three or more isolates are shown.

Even with the large number of isolates examined, we still do not have data sufficient to estimate the actual diversity of these organisms at either site. For example, almost all the chlamydomonad isolates from each new sampling were previously undetected organisms. In fact, of the 48 distinct chlamydomonads isolated and characterized from Itasca sites, only five were detected in multiple samples (Fig. 3). We found more new taxa of chlamydomonads from our last sample (17, 18 August 2001) than we did from our first sampling (15, 21 Sept 2000). Moreover, of our 93 distinct sequence types, including those from the Sphaeropleales, Chlamydomonadales, and Trebouxiophyceae, 28 are represented by only a single isolate (data not shown).

Discussion

The results of this study indicate a large amount of diversity among chlamydomonads at the Itasca sites and coccoid green algae at both the Arrowwood and Itasca sites. Only four of the 93 distinct 18S rDNA sequences from these isolates match any sequences already present in GenBank. These results suggest that many potentially new taxa of green algae are present among our isolates. Our results are similar to those of Lewis and Flechtner [29], who determined the 18S rDNA sequences of 11 green algal isolates from desert soils. None of their sequences matched any published sequence and the authors also concluded that their isolates are likely to be previously undescribed species.

However, the conclusion that our isolates may be new species remains somewhat equivocal because numerous species of green algae have not been characterized by any molecular method. Notwithstanding, there are some green algal genera for which 18S rDNA sequences are available for many, if not all, described species. For example, only a single freshwater species has been described for *Mychonastes* (Chlorophyceae), a unicellular coccoid organism that is only known to reproduce asexually. The 18S rDNA genes of *M. homosphaera* (Skuja) Kalina et Pun·hochá·ová isolates from Lake Erkin (Sweden) and Lake Kinneret (Israel) differ by only a single nucleotide [16]. *Mychonastes homosphaera* and its close relative, the colonial *Pseudodictyosphaerium jurisii* Hindák, are thought to be widely distributed [20, 21, 27]. We have identified four different 18S rDNA sequences among our isolates with *Mychonastes* morphology (morphological data not shown). Although one of these sequences differs from the published sequence for the Swedish isolate of *M. homosphaera* by a single indel, a difference that could be due to sequencing error, our other isolates differ from published sequences by at least seven substitutions (Fig. 2). We have many additional isolates from Itasca State Park that form loose colonies in mucilage and can be identified as either *Pseudodictyosphaerium* spp. or *Korshpalmella* spp. These isolates have 18S rDNA sequences identical to that of the *Mychonastes* isolate from Itasca State Park, Tow 6/3 P-1w, but differ from the published sequence for *Pseudodictyosphaerium jurisii* by eight substitutions and five indels. None of our isolates from the *Mychonastes/Pseudodictyosphaerium* lineage have an 18S sequence that is identical to any published sequence.

Hepperle and Schlegel [21] recently examined diversity of eukaryotic picoalgae from three Swiss lakes. Four of the 10 isolates that they characterized were identified as *Mychonastes* spp. by both morphological and 18S rDNA analyses. Among these four isolates there were three distinct 18S rDNA sequences, none of which were identical to any published sequence. Hepperle and Schlegel also noted that one of their *Mychonastes* isolates occurred as a mixture of unicells and colonies, similar to our results with Itasca *Mychonastes/Pseudodictyosphaerium* isolates. Two of the three distinct sequences from Hepperle and Schlegel are included in Fig. 2. None of their sequences match any of the sequences from our isolates.

Additional coccoid isolates from the Chlorophyceae are associated primarily with the *Scenedesmus* and *Desmodesmus* lineages (Sphaeropleales). Small-subunit sequence data are available for several of the described species from these lineages and it is evident (Fig. 2) that small, or even no, variations in the 18S sequences separate these morphospecies. For example, *Coelastropsis costata* and *Scenedesmus obtusus* sequences do not differ

at all, and sequences from several other members of the *Scenedesmus* lineage differ by three or fewer substitutions. None of our 18S rDNA sequences from these lineages, or other isolates from the Sphaeropleales, match any published sequences. However, the sequence of one isolate, Tow 9/21 P-1w, differs from the published sequence of *Scenedesmus rubescens* by only a single nucleotide. This isolate is morphologically very similar to *S. rubescens*, although cells are narrower than the specimens illustrated by Kalina and Pun·hochá·ová [25]. Critical comparison of the morphology, physiology, and additional sequences with the type isolate will be necessary in order to determine if our isolate is actually a new species.

Another interesting coccoid genus that has been examined by 18S rDNA analysis is the minute, unicellular *Choricystis* (Trebouxiophyceae). Isolates of *Choricystis* previously thought to be different species based on morphological criteria have been shown to possess identical 18S rDNA sequences and one species, *C. minor*, is thought to be cosmopolitan [20, 21, 26]. We have isolates of *Choricystis* from all of our study sites. To date, we have characterized five unique 18S rDNA sequence types, often with numerous isolates of a particular type. One of these 18S rDNA sequences matches the sequence for *C. minor*, and another is identical to the sequence of an undescribed *Choricystis* sp. (Fig. 1). The remaining 18S sequences are new to science. We have also detected variation in the gene encoding the large subunit of the RUBISCO enzyme (*rbc*L) among those *Choricystis* isolates with identical 18S rDNA sequences. For example, MDL5-9 (from Arrowwood) and Itas 9/21 14-5w (from Itasca) have identical 18S rDNA coding sequences, but MDL5-9 possesses an intron that is not present in Itas 9/21 14-5w, and the partial *rbc*L sequences (1063 nucleotides, AY234380 and AY234381) differ by 10 substitutions (not shown). Although our analyses of *Choricystis* morphology are not complete, we have not detected any notable morphological variation among the different *Choricystis* genotypes (data not shown).

The most commonly reported genera of freshwater coccoid green algae are *Chlorella* and *Nannochloris* (both Trebouxiophyceae). Simple coccoid organisms have often been described as *Chlorella* species, and the genus as originally described is now known to be polyphyletic. We have characterized 14 different 18S sequences of coccoid algae that are similar to the sequence from *Chlorella vulgaris*, the type species of the genus, and other closely related *Chlorella* species. These genotypes differ from each other and from published sequences by as little as one substitution (Fig. 1). However, not a single 18S rDNA sequence from our *Chlorella*-like isolates is identical to any published sequence. Our isolates do have some morphological variation, and also differ slightly from named species (data not shown). Five additional sequences from our isolates are found in the lineage that includes *Nannochloris* species. One of these isolates, which is morphologically similar to *Marvania geminata* Hindák, has an 18S rDNA sequence identical to that of *Nannochloris coccoides* (Fig. 1). The other sequences from this lineage are distinct from published sequences.

Additional trebouxiophycean isolates are basal to the *Choricystis* lineage (Tow 19/21 P-16w) or allied with the *Oocystis* lineage (Tow 6/3 P-10w, MDL6-7 and AN2/29-4) (Fig. 1b). More than 90 species of *Oocystis* have been described [24] and most have not been examined by sequence analysis. Thus we can not determine if the sequences of our isolates are the same as described species.

For all the coccoid algae examined, initial molecular characterization was performed by PCR-RFLP analysis of the 18S rDNA. This technique, although rapid enough to permit screening the large number of isolates, is conservative in that organisms may present the same PCR-RFLP patterns, but actually have some sequence differences. Whenever more than one isolate produced the same PCR-RFLP fragments, we sequenced the 18S rDNA from multiple isolates. However, it is possible that additional 18S rDNA diversity remains to be discovered among those organisms with similar PCR-RFLP patterns. We have also examined more rapidly evolving genes from some of our organisms and have detected variation among ribosomal ITS regions and *rbc*L genes for isolates with identical 18S rDNA sequences (data not shown). We will focus on these variations in future studies on specific lineages of algae.

The chlorophycean flagellates (Chlamydomonadales), which include the model organism, *Chlamydomonas reinhardtii*, have also been rather intensively studied by both morphological and molecular methods. Meyer and Brook [32] noted 50 chlorophycean flagellate species in their morphology-based survey of Itasca algae. In the present study, sampling for chlorophycean flagellates focussed only on Itasca State Park. As a result, 60 unknown chlorophycean flagellates from the Itasca sites and only one chlamydomonad isolate from Arrowwood were characterized by 18S rDNA sequence analysis. Results from analyses of these sequence data along with 85 published chlamydomonad sequences (Fig. 3) show that the unidentified isolates are scattered among various chlamydomonad lineages. Each of the 61 18S rDNA sequences from our isolates falls into one of 49 distinct types. Of the 49 distinct sequences, all but two are new to science. Moreover, of the 47 new 18S rDNA types, 38 are represented by only a single Itasca isolate. Although Pröschold et al. [37] permit some 18S rDNA diversity in their determinations of chlamydomonad species, the evidence from *Chlamydomonas noctigama* Korshikov [4, 6], a microalga for which a biological species concept can be applied, suggests that a unique 18S rDNA sequence is correlated with effective reproductive isolation for that alga. We are currently exploring 18S rDNA sequence

diversity among heterothallic strains from national culture collections (e.g., various UTEX strains of *Chlamydomonas moewusii*) and have determined that, in every case, the "plus" and "minus" mating types have identical 18S rDNA sequences (unpublished observations). If this correlation between 18S rDNA sequence and species delineation can be broadly applied to chlamydomonad algae, then our 48 new 18S sequences likely represent 48 distinct species.

Unlike most microalgae, the genus *Chlamydomonas* could be cited as an example of a speciose unicellular green alga (>400 species [10]). However, recent efforts at *Chlamydomonas* synonymy has potentially reduced the number of species to as few as 15 [11]. This reduction in species would appear to be more consistent with the hypotheses of Finlay and Happey-Wood, which lead to predictions of low species diversity. On the other hand, molecular data for chlamydomonad isolates ([3, 22, 37], present investigation) make it apparent that the original, higher species tally may be closer to the mark than the more recent revisionist efforts. Moreover, a number of *Chlamydomonas* species have been placed in new genera as a consequence of molecular phylogenetic analysis [37]—further evidence that a morphospecies approach underestimates diversity in these organisms.

We have also found that the communities of planktonic, coccoid green algae vary among our sites. Thirty-five different 18S rDNA sequence types of trebouxiophycean and sphaeroplealean coccoid algae were characterized in this study, but only two sequence types were found at both Arrowwood and Itasca sites (Table 2). In addition, the numerically dominant organisms among our isolates are differentially distributed between the two locations. For the Arrowwood sites, *Mychonastes* species were most frequently found, with 16 isolates of one genotype (JL1/12-12 in Fig. 1). However, only one *Mychonastes* isolate has been found from an Itasca site, and this isolate possessed an 18S rDNA sequence different from the Arrowwood *Mychonastes* isolates. Many isolates of *Choricystis* spp. were cultured from all sites, but among the five distinct 18S rDNA coding sequences, only one was found in both Itasca and Arrowwood sites. Additional studies have shown that these isolates vary in 18S introns and *rbc*L sequences (see above). Five different 18S sequence types related to *Nannochloris*, *Koliella*, and *Marvania* were isolated from Arrowwood lakes, but no members of this lineage have been detected from among our Itasca isolates. For the *Chlorella* lineage, we have thus far characterized 14 different 18S rDNA sequences; however, only two types have been found at both locations.

The numerically dominant coccoid algae among our isolates from Lake Itasca are also quite distinct from those of other lakes within Itasca State Park. We have 20 isolates of one *Chlorella* relative (represented by Itas 2/24 S-12w in Fig. 1) from Lake Itasca that has not been isolated from any other Itasca site. Isolates identified by morphological criteria as *Pseudodictyosphaerium* spp. and *Korshpalmella* spp., with 18S rDNA sequence identical to *Mychonastes* sp. (Tow 6/3 P-1d in Fig. 2) were frequently cultured from Lake Itasca (22 isolates). In contrast, only three *Korshpalmella* isolates were collected from one other Itasca site (Mary Lake) and no other site produced a *Pseudodictyosphaerium* isolate. There are also five different 18S rDNA sequence types that have been isolated only from Mary Lake. On the other hand, *Choricystis* spp. were widely distributed among all Itasca sites. All other distinct 18S rDNA types from Itasca were represented by three or fewer isolates and more are needed to clarify the distributions of these organisms. Although this study was not designed to be quantitative or exhaustive, the numerous examples of exclusive distributions (e.g., *Pseudodictyosphaerium*) cannot be easily dismissed as inadequate sampling.

If the hypothesis of cosmopolitan distribution is correct, then algae regarded as cosmopolitan, widely adapted and abundant, such as *Chlamydomonas reinhardtii* and *Chlorella vulgaris* [10, 24], should be frequently isolated from almost any freshwater system. *Chlamydomonas reinhardtii* is considered "one of the most frequently encountered species found in a wide range of water types" ([24], p 311) and *Chlorella vulgaris* is thought to be "very widely distributed in the plankton, associated with surfaces in a wide range of aquatic habitats" ([24], p 336). Indeed, both of these organisms are reported from Itasca based on morphological analysis [32]. However, we have not encountered *Chlamydomonas reinhardtii* among our 61 chlamydomonad isolates, or *Chlorella vulgaris* from among the coccoid isolates. Although we may find both organisms with additional sampling, these microalgae do not appear to be as overwhelmingly common in all freshwater systems as the literature suggests. *Mychonastes homosphaera*, *Pseudodictyosphaerium jurisii*, and *Choricystis minor* are also considered to be widely distributed [20, 21, 26, 27], and some of our isolates do have 18S sequences identical, or nearly identical, to *M. homosphaera* and *C. minor.* However, we have many isolates with distinct 18S rDNA sequences that are morphologically indistinguishable from these coccoid species. Moreover, neither *Mychonastes* nor *Pseudodictyosphaerium* have been found at several Itasca sites.

A key observation for this discussion is the tremendous diversity at the 18S rDNA level found among the Itasca and Arrowwood isolates. For example, the 18S rDNA sequence of Tow 2/24 P-12d differs from the most similar chlorophycean sequence by 100 substitutions and five indels out of 1793 nucleotides, a difference of nearly 6%. This difference is greater than the differences in 18S sequences generally observed among all angiosperms,

which vary at about 5% of the nucleotides [33]. Overall, the sequences from six of our isolates differed from other sequences by 60 or more substitutions. In addition, sequence analysis revealed what may be an entirely new chlamydomonadalean lineage comprised of 12 Itasca isolates (Fig. 3, the lineage containing Pic 8/18/T-15w). The profound differences between the 18S rDNA sequences of some of our isolates and published sequences suggest that much fundamental diversity among the Chlorophyceae and Trebouxiophyceae remains to be discovered.

In summary, our results do not support the hypothesis of low species diversity. For each genus we have examined, additional 18S rDNA diversity has been revealed that could not be detected by morphology alone. In addition, the coccoid communities of different lake types are clearly different. In fact, our results indicate that the coccoid communities of similar lakes can be quite different. As a result, we fully expect to find additional new organisms as we examine more freshwater systems. Preliminary results from another mesotrophic lake near Itasca State Park and from lakes and ponds in Tulsa (Oklahoma, USA) have already revealed additional 18S rDNA diversity (unpublished). The diversity that we have detected suggests that green algae may be responding to their habitat in a fine-grained manner. Even though ecological tolerances of microalgae may be broad in culture, individual niches may be narrow under natural conditions. Ecological studies of microalgae have been constrained by the morphological species concept and an inability to differentiate among potential taxa. Culturing and molecular studies such as ours allow a more rigorous approach to these problems than has been possible in the past, and will likely result in reassessment of niche partitioning.

Although these results challenge hypotheses that the diversity of eukaryotic microalgae is low and that the same organisms are found in many different environments, they are not sufficiently comprehensive to allow evaluation of the major thesis that eukaryotic microorganisms are cosmopolitan in distribution. We do find some green algae with identical sequences from diverse locations, such as *Lobochlamys segnis* and *Choricystis* spp., which are found in both Europe and North America. Whether all of the organisms we have found are similarly distributed, or whether some may be endemic, remains to be determined. Additional culture and molecular studies from other sites are planned and will provide insights into the distribution and evolution of microalgae and microorganisms in general.

Acknowledgments

We thank Mohamed Fakhr, Julie Buchheim, Mariah Veit, Lesley Bryden, Jamica Offord, and Matt Hoffman for technical assistance. We thank Phillip McClean and Rian Lee for the use of the Beckman sequencer and Mark Vanniman at Arrowwood National Wildlife Refuge for his interest and assistance. We thank Dr. Robert Andersen and anonymous reviewers for helpful comments. This material is based upon work supported by the National Science Foundation under Grant Nos. DBI-0070387, EPS-0132289, DEB-0129030, DEB-0128952, MCB 0132083, DBI 0139653, and MCB-0084188. Additional funding was provided by the North Dakota Water Resources Research Institute.

References

1. Andersen, RA (1992) Diversity of eukaryotic algae. Biodiversity Conserv 1: 267–292
2. Baldwin, BG, Markos, S (1998) Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). Mol Phyl Evol 10: 449–463
3. Buchheim, MA, Buchheim, JA, Carlson, T, Kugrens, P (2002) Phylogeny of *Lobocharacium* (Chlorophyceae) and allies: a study of 18S and 26S rDNA data. J Phycol 38: 376–383
4. Buchheim, MA, Buchheim, JA, Chapman, RL (1997) Phylogeny of the VLE-14 *Chlamydomonas* (Chlorophyceae) group: a study of 18S rRNA gene sequences. J Phycol 33: 1024–1030
5. Buchheim, MA, Chapman, RL (1992) Phylogeny of *Carteria* (Chlorophyceae) inferred from molecular and organismal data. J Phycol 28: 362–374
6. Burrascano, CG, VanWinkle-Swift, KP (1984) Interspecific matings of five *Chlamydomonas* species. Genetics 107S: 15
7. Callieri, C, Stockner, JG (2002) Freshwater autotrophic picoplankton: a review. J Limnol 61: 1–14
8. Coleman, AW (2001) Biogeography and speciation in the *Pandorina/Volvulina* (Chlorophyta) superclade. J Phycol 37: 836–851
9. Coleman, AW (2002) Microbial eukaryote species. Science 297: 337
10. Ettl, H (1983) Süßwasserflora von Mitteleuropa, vol. 9: Chlorophyta I. Phytomonadina. Gustav Fischer Verlag, New York
11. Ettl, H, Schlösser, UG (1992) Towards a revision of the systematics of the genus *Chlamydomonas* (Chlorophyta). Bot Acta 105: 323–330
12. Fawley, MW, Fawley, KP (2004) A simple and rapid technique for the isolation of DNA from microalgae. J Phycol 40: 223–225
13. Fawley, MW, Yun, Y, Qin, M (2000) Phylogenetic analyses of 18S rDNA sequences reveal a new coccoid lineage of the Prasinophyceae (Chlorophyta). J Phycol 36: 387–393
14. Finlay, BJ (2002) Global dispersal of free-living microbial eukaryote species. Science 296: 1061–1063
15. Finlay, BJ, Fenchel, T (2002) Microbial eukaryote species—response. Science 297: 337
16. Hanagata, N, Malinsky-Rushansky, N, Dubinsky, Z (1999) Eukaryotic picoplankton, *Mychonastes homosphaera* (Chlorophyceae, Chlorophyta), in Lake Canard, Israel. Phycol Res 47: 263–270
17. Hanson, DA (1978) *Chemical and Physical Characteristics of the James River from Arrowwood Refuge to the Outflow of Jamestown Dam, North Dakota.* M.S. Thesis, North Dakota State University, Fargo
18. Happey-Wood, CM (1991) Ecology of freshwater planktonic green algae. In: Sandgren, CD (Ed.) Growth and Reproductive Strategies of Freshwater Phytoplankton, Cambridge University Press, New York, pp 53–68
19. Hegewald, E, Silva, P (1988) An annotated catalogue of *Scenedesmus* and nomenclaturally related genera including original descriptions and figures. Bibl Phycol 80: 1–587

20. Hepperle, D, Krienitz, L (2001) Systematics and ecology of chlorophyte picoplankton in German inland waters along a nutrient gradient. Int Rev Hydrobiol 86: 269–284
21. Hepperle, D, Schlegel, I (2002) Molecular diversity of eucaryotic picoalgae from three lakes in Switzerland. Int Rev Hydrobiol 87: 1–10
22. Hoham, RW, Bonome, TA, Martin, CW, Leegens-Mack, JH (2002) A combined 18S rDNA and rbcL phylogenetic analysis of *Chloromonas* and *Chlamydomonas* (Chlorophyceae, Volvocales) emphasizing snow and other cold-temperature habitats. J Phycol 38: 1051–1064
23. John, DM, Maggs, CA (1997) Species problems in eukaryotic algae: a modern perspective. In: Claridge, MF, Dawah, HA, Wilson, MR (Eds.) Species: The Units of Biodiversity, Chapman and Hall, New York, pp 83–107
24. John, DM, Whitton, BA, Brook, AJ (2002) The Freshwater Algal Flora of the British Isles. University Press, New York
25. Kalina, T, Pun·hochá·ová, M (1987) Taxonomy of the subfamily Scotiellocystoideae Fott 1976 (Chlorellaceae, Chlorophyceae). Arch Hydrobiol Suppl 73: 473–521
26. Krienitz, L, Huss, VAR, Hümmer, C (1996) Picoplanktonic *Choricystis* species (Chlorococcales, Chlorophyta) and problems surrounding the morphologically similar '*Nannochloris*-like algae'. Phycologia 35: 332–341
27. Krienitz, L, Takeda, H, Hepperle, D (1999) Ultrastructure, cell wall composition, and phylogenetic position of *Pseudodictyosphaerium jurisii* (Chlorococcales, Chlorophyta) including a comparison with other picoplanktonic green algae. Phycologia 38: 100–107
28. LaJeunesse, TC (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: in search of a ''species'' level marker. J Phycol 37: 866–880
29. Lewis, LA, Flechtner, VR (2002) Green algae (Chlorophyta) of desert microbiotic crusts: diversity of North American taxa. Taxon 51: 443–451
30. Linder, CR, Goertzen, LR, Heuvel, BV, Francisco-Ortega, J, Jansen, RK (2000) The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. Mol Phyl Evol 14: 285–303
31. Maddison, DR, Maddison, WP (2000) MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA
32. Meyer, RL, Brook, AJ (1968) Freshwater algae from the Itasca State Park, Minnesota. I. Introduction and Chlorophyta. Nova Hedwigia 16: 251–266
33. Nickrent, DL, Soltis, DE (1995) A comparison of angiosperm phylogenies from nuclear 18S rDNA and *rbc*L sequences. Ann MO Bot Gard 82: 208–234
34. Phillips, KA, Fawley, MW (2000) Diversity of coccoid algae in shallow lakes during winter. Phycologia 39: 498–506
35. Phillips, KA, Fawley, MW (2002) Winter phytoplankton community structure in three shallow temperate lakes during ice cover. Hydrobiologia 470: 97–113
36. Phillips, KA, Fawley, MW (2002) Winter phytoplankton blooms under ice associated with elevated oxygen levels. J Phycol 38: 1068–1073
37. Pröschold, T, Marin, B, Schlösser, UG, Melkonian, M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (Chlorophyta). I. Emendation of *Chlamydomonas* Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and *Lobochlamys* gen. nov. Protist 152: 265–300
38. Suda, S, Atsumi, M, Miyashita, H (2002) Taxonomic characterization of a marine *Nannochloropsis* species, *N. oceanica* sp. Nov. (Eustigmatophyceae). Phycologia 41: 273
39. Swofford, DL (2002) PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA
40. Trainor, FR, Egan, PF (1991) Discovering the various ecomorphs of *Scenedesmus*: the end of a taxonomic era. Arch Protistenkd 139: 125–132