J Mol Evol (1990) 31:432–442

Journal of Molecular Evolution

Phylogenetic Position of Some *Chlorella* Species within the Chlorococcales Based upon Complete Small-Subunit Ribosomal RNA Sequences

V.A.R. Huss[1] and M.L. Sogin[2]

[1] Institut für Botanik und Pharmazeutische Biologie der Universität, D-8520 Erlangen, Federal Republic of Germany
[2] Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

Summary. Complete small-subunit rRNA (16S-like rRNA) coding region sequences were determined for eight species of the Chlorococcales (Chlorophyceae). The genera investigated include *Prototheca, Ankistrodesmus, Scenedesmus,* and five *Chlorella* species. Distance matrix methods were used to infer a phylogenetic tree that describes evolutionary relationships between several plant and green algal groups. The tree exhibits a bifurcation within the Chlorococcales consistent with the division into Oocystaceae and Scenedesmaceae, but three of the five *Chlorella* species are more similar to other algae than to *Chlorella vulgaris.* All of the sequences contain primary and secondary structural features that are characteristic of 16S-like rRNAs of chlorophytes and higher plants. *Anikstrodesmus stipitatus,* however, contains a 394-bp group I intervening sequence in its 16S-like rRNA coding region.

Key words: Chlorococcales — *Ankistrodesmus* — *Chlorella* — *Prototheca* — *Scenedesmus* — Green algal phylogeny — Eukaryotic small-subunit rRNA — Secondary structure

Introduction

Unicellular green algae of the genus *Chlorella* Beijerinck were the first algae to have been grown extensively in axenic cultures. They serve as model organisms in physiological and biochemical plant research, e.g., photosynthesis and nitrate reduction. More recently, *Chlorella* has been mass cultured for production of human and animal feed, treatment of sewage, microbial energy conversion, etc. (Soeder 1980).

The cells of *Chlorella* are spherical or ellipsoidal, the size (2–12 μm) often varying with culture conditions. The chloroplast is parietal, with a pyrenoid in most species. Reproduction is asexual by release of autospores. The lack of significant morphological variation has forced taxonomists to rely upon similarities in physiological characters (Shihira and Krauss 1965) or a combination of morphological and structural features (Fott and Nováková 1969). Kessler (1982a, 1984, 1987) has presented a convincing classification of *Chlorella* strains that includes 16 taxa. They are differentiated by a combination of 12 physiological and biochemical species-specific properties. The observed diversity of such properties within a single genus is remarkable. Even more striking is the enormous variation of DNA base composition determined for 88 strains by Hellmann and Kessler (1974). The guanine + cytosine (G+C) content of *Chlorella* species ranges from 43 to 78 mol% G+C, which suggests that *Chlorella* represents an assemblage of morphologically similar species of polyphyletic origin rather than a natural genus (Kessler 1976). Indeed, several authors have nearly dismantled the taxonomic unit of *Chlorella,* and at least 6 out of the 16 taxa defined by Kessler have been renamed and classified in different genera or families (Fott et al. 1975; Hegewald 1982; Hindák 1982; Komárek 1987; Kalina and Punčochářová 1987). These investigations are based again primarily upon morphological criteria.

Quantitative DNA hybridization procedures have been employed to investigate intraspecific relation-

Offprint requests to: V.A.R. Huss

ships of *Chlorella* species (Huss et al. 1986, 1987a, b, 1988, 1989b). However, with few exceptions these studies were unable to reveal interspecies relationships (Huss et al. 1989a). Even though very useful for the delimitation of different taxa and for resolving the genealogical relationships within a species, the DNA hybridization data thus provided little conclusive information about the evolution of *Chlorella* species and their relation to other green algae. This has been discussed in terms of a relatively large phylogenetic divergency within the genus *Chlorella*.

Phylogenetic relationships between even the most divergent organisms can be examined by comparative sequence analysis of conserved macromolecules (Zuckerkandl and Pauling 1965). The larger rRNAs have proven to be particularly useful for this purpose. Similarities between small-subunit rRNA (16S-like rRNAs) demonstrated the existence of three primary kingdoms of organisms (Woese and Fox 1977). Because of advances in cloning and sequencing techniques, and more recently the widespread application of polymerase chain reaction (PCR) techniques for specifically amplifying rRNA coding regions, the database for eukaryotic and prokaryotic 16S-like rRNAs has grown dramatically. Although most studies within eukaryotes have focused upon phylogenetic relationships between divergent organisms (Sogin et al. 1989), it has been demonstrated that it is also possible to infer reliable relationships between species of a single genus (Sogin et al. 1986).

In this study we determined complete 16S-like rRNA sequences from five *Chlorella* species. In order to obtain a broader database for discussing *Chlorella* phylogeny [sequences of only two green algae, *Chlamydomonas reinhardtii* (Gunderson et al. 1987) and *Nanochlorum eucaryotum* (Sargent et al. 1988) have been reported], we also determined sequences for the chlorococcalean genera *Ankistrodesmus, Prototheca,* and *Scenedesmus. Prototheca,* a heterotrophic and colorless alga, and *Scenedesmus* were chosen to clarify the previously suggested relationship of *Prototheca* to *Chlorella protothecoides* (Kessler 1982a; Pore 1985) as well as of *Chlorella fusca* to *Scenedesmus* (Kessler 1982a; Huss et al. 1989a).

Materials and Methods

Organisms, Vectors, and Enzymes. Algal cultures were the same as used by Huss et al. (1989a): *Chlorella vulgaris* 211-11b, *Chlorella kessleri* 211-11g, *Chlorella minutissima* C-1.1.9, *C. protothecoides* 211-7a, *C. fusca* var. *vacuolata* 211-8b, *Scenedesmus obliquus* 276-3a, and by Huss et al. (1988): *Prototheca wickerhamii* 1283. *Ankistrodesmus stipitatus* 202-5 was from the Sammlung von Algenkulturen, Pflanzenphysiologisches Institut der Universität Göttingen, FRG (SAG). Algae were grown as described by Huss et al. (1986). *Escherichia coli* strain JM 109 was used as host for M13 vectors mp18 and mp19 and grown in 2× YT medium (Messing 1983). Restriction endonucleases, Klenow fragment of DNA polymerase I, and T_4 DNA ligase were from BRL, and *Thermus aquaticus* DNA polymerase was from New England BioLabs.

DNA Isolation, Purification, and Amplification of 16S-like rRNA Genes. Total DNA from algae was isolated according to Huss et al. (1986). DNA of some strains was further purified by CsCl density gradient centrifugation (Huss et al. 1988). The 16S-like rRNA coding region was amplified by a modification of the PCR method (Saiki et al. 1988) using heat-stable *Taq* DNA polymerase and eukaryote specific synthetic oligonucleotide amplification primers as previously described (Medlin et al. 1988). Thirty amplification cycles were carried out in a Perkin-Elmer Cetus DNA Thermal Cycler. Each cycle included a 2-min denaturation period at 94°C, 2-min primer annealing at 37°C, and 6-min primer extension at 72°C.

M13 Cloning and Sequencing. Amplification products were ligated into the RF of M13 mp18 and M13 mp19, and single-stranded templates for primer extension sequencing were prepared from recombinant M13 phages (Medlin et al. 1988). Complete sequences for both the coding and noncoding DNA strand were determined by the dideoxynucleotide chain-termination sequencing protocols (Sanger and Coulson 1975) initiating DNA synthesis with oligonucleotide primers that are complementary to evolutionarily conserved regions of the 16S-like rRNA gene (Elwood et al. 1985).

Sequence Analysis and Construction of Phylogenetic Trees. For the inference of phylogenetic trees by distance matrix methods, similarity values must be computed for all possible pairwise comparisons of homologous nucleotide positions. Similarity is defined as

$$s = m/(m + u + g/2)$$

where m is the number of sequence positions with matching nucleotides, u is the number of positions with nonmatching nucleotides, and g is the number of sequence gaps (only the first five positions in a gap are considered in making the calculations; large insertions or deletions probably reflect single rare events). The similarity values are converted to distance values (the number of evolutionary changes per 100 positions) using the formula of Jukes and Cantor (1969), which compensates for the probability of multiple events at the same position. The distances are then converted to phylogenetic trees using a modification (Elwood et al. 1985) of the distance matrix methods (Fitch and Margoliash 1967). The evaluation of alternative phylogenetic trees is based upon the agreement of the distance data separating pairs of organisms and the sum of tree segment lengths joining the organisms in the tree. Because it is not practical to test all possible tree topologies, an algorithm was used in which the effects of a given set of rearrangements on a given phylogenetic tree were tested, and then the best of all tested alternatives (i.e., the most improved tree) was maintained and used as the starting point for another round of optimization. Two simple classes of tree rearrangements were tested by the optimization algorithm. Both regard the current tree as sets of subtrees connected by segments. A subtree can range from a single sequence to $N - 3$ sequences, where N is the number of organisms represented in the tree. A subtree can be moved to a new location by removing its nearest node from the tree and inserting this node into an alternative tree segment. In the first class of rearrangements, the effect of moving each possible subtree (one at a time) to every alternative location in the tree is systematically tested. The second class of rearrangements tested involves interchanging the locations of a pair of subtrees. The effect of all pairwise interchanges of subtrees (one pair at a time) is tested. The rearrangement that leads to the most improved tree is used as the starting point for a new round of optimization.

Table 1. Structural similarity and distance data between small subunit rRNA gene sequences

Organism	*Z.p.*	*O.s.*	*Z.m.*	*L.e.*	*G.m.*	*C.r.*	*P.w.*	*C.p.*	*C.m.*
Zamia pumila	—	0.932	0.924	0.942	0.931	0.879	0.884	0.871	0.888
Oryza sativa	0.071	—	0.980	0.960	0.952	0.882	0.882	0.871	0.886
Zea mays	0.080	0.020	—	0.951	0.948	0.877	0.878	0.868	0.885
Lycopersicon esculentum	0.060	0.041	0.050	—	0.968	0.882	0.883	0.874	0.892
Glycine max	0.072	0.049	0.054	0.033	—	0.880	0.877	0.867	0.889
Chlamydomonas reinhardtii	0.131	0.128	0.134	0.128	0.130	—	0.922	0.906	0.943
Prototheca wickerhamii	0.126	0.128	0.133	0.126	0.134	0.082	—	0.947	0.957
Chlorella prototheocoides	0.141	0.141	0.145	0.138	0.146	0.100	0.055	—	0.935
Chlorella minutissima	0.121	0.124	0.125	0.116	0.120	0.059	0.044	0.068	—
Nanochlorum eucaryotum	0.119	0.123	0.126	0.115	0.123	0.063	0.047	0.071	0.021
Chlorella vulgaris	0.112	0.116	0.121	0.111	0.115	0.060	0.040	0.063	0.020
Chlorella kessleri	0.118	0.121	0.126	0.115	0.118	0.061	0.046	0.068	0.024
Ankistrodesmus stipitatus	0.125	0.131	0.133	0.122	0.124	0.062	0.071	0.091	0.045
Chlorella fusca var. *vacuolata*	0.123	0.130	0.133	0.122	0.128	0.066	0.065	0.088	0.046
Scenedesmus obliquus	0.125	0.134	0.136	0.126	0.129	0.066	0.068	0.091	0.047

The upper right half of the table gives the structural similarity (fraction of sites that are identical), and the lower-left half of the table shows the distance data (average number of base changes per sequence position). Sequence data used for comparison are from *Zamia pumila* (Nairn and Ferl 1988), *Oryza sativa* (Takaiwa et al. 1984), *Zea mays* (Messing et al. 1984), *Lycopersicon esculentum* (Kiss et al. 1989), *Glycine max* (Eckenrode et al. 1985), *Chlamydomonas reinhardtii* (Gunderson et al. 1987), and *Nanochlorum eucaryotum* (Sargent et al. 1988)

Results

The 16S-like rRNA primary structures inferred from the coding region sequences are presented in Fig. 1. The sequence of *C. vulgaris* (Huss and Sogin 1989) is used as reference and aligned with sequences of the other chlorococcalean algae determined in this study. The extent of length variation ranges from 1792 nucleotides in the *A. stipitatus* mature small-subunit rRNA to 1831 in *C. prototheocoides*. The 16S-like rRNA coding region of *A. stipitatus* contains 394 extra bases inserted between nucleotides 1261 and 1262 (data not shown). The insert has all of the features characteristic for group I introns (Cech 1988). The primary and secondary structure of the intron found in the 16S-like rRNA gene of *A. stipitatus* and its splicing abilities will be published elsewhere.

The number of nucleotide differences between the 16S-like rRNA coding regions for the five *Chlorella* species ranges from 37 sites between *C. vulgaris* and *C. kessleri* to as many as 209 positions between *C. fusca* var. *vacuolata* and *C. prototheocoides*. The structural similarities and distance data for 16S-like rRNA sequences of all published higher plants and green algae are given in Table 1. Only nucleotides that could be aligned unambiguously according to the method described by Elwood et al. (1985) were included in the analysis. This corresponds to 96% of the *C. vulgaris* sequence. The structural distances of Table 1 were used to construct the phylogenetic tree shown in Fig. 2 (Elwood et al. 1985). The tree is rooted using the sequence of *Acanthamoeba castellanii* (Gunderson and Sogin 1986) as an outgroup.

Sequence alignment is substantially facilitated by considering the evolutionary conservation of secondary structures. We therefore constructed secondary structure models for all of the algae studied. As an example, the model for *C. vulgaris* 16S-like rRNA is shown in Fig. 3a. It is a modification of the *Zea mays* structure proposed by Gutell et al. (1985). However, it takes into account a substantially larger database.

For eukaryotic 16S-like rRNAs it is difficult to model the region that corresponds to helix 21 in prokaryotes [helix numbering is according to Dams et al. (1988)]. This helix is replaced by a much longer and highly variable sequence. Various secondary structure models were proposed (Atmadja et al. 1984; Choi 1985; Gonzalez and Schmickel 1986; Herzog and Maroteaux 1986; Hendriks et al. 1988) and proven either by testing for compensatory base variations (Woese et al. 1983) and/or by the base-paired fragment approach (Atmadja et al. 1984). The results, however, are different even for similar sequences of, e.g., *Homo sapiens* and *Xenopus laevis*. Proving a helix by the comparative method presumes perfect alignment of the sequences to be compared. Because of its highly variable sequence and substantial length variation, it has not been possible to develop a consensus structure for the eukaryotic analogue of helix 21 (Raué et al. 1988).

Our model for the analogue of helix 21 of green algae and higher plants is shown in Fig. 3a. It consists of five eukaryote-specific helices E21-1 to E21-6 [E21-3 (Dams et al. 1988) is missing in our model]. The existence of the hypervariable helices E21-1 and E21-2 is well proven by numerous compensating

Table 1. Extended

N.e.	*C.v.*	*C.k.*	*A.s.*	*C.f.*	*S.o.*
0.889	0.895	0.890	0.885	0.886	0.885
0.886	0.892	0.888	0.879	0.880	0.877
0.884	0.888	0.884	0.878	0.878	0.876
0.893	0.896	0.893	0.887	0.887	0.884
0.886	0.893	0.890	0.885	0.882	0.881
0.939	0.941	0.941	0.940	0.937	0.936
0.954	0.960	0.955	0.932	0.937	0.935
0.932	0.940	0.934	0.914	0.917	0.914
0.978	0.980	0.975	0.956	0.954	0.954
—	0.975	0.967	0.946	0.947	0.947
0.024	—	0.984	0.952	0.956	0.956
0.033	0.016	—	0.949	0.952	0.951
0.055	0.049	0.052	—	0.960	0.962
0.055	0.045	0.049	0.041	—	0.992
0.054	0.044	0.050	0.039	0.008	—

base exchanges in our large database of relatively closely related green algae (Fig. 4). In addition, helix E21-2 has been proven by Atmadja et al. (1984) for *X. laevis* by secondary structure mapping techniques of ribonuclease T_1 and/or RNase A-resistant base-paired RNA fragments. The 16S-like rRNA of *C. protothecoides* contains a unique insertion in helix E21-1 resulting in a large loop (Fig. 3b). Because of the lack of homologous sequences no base pairing can be proposed for this region by the method of comparative base exchanges. Helix E21-4 is almost invariable within the organisms studied. Base changes occur only in the loop and at the base of the stem where a U·G pair is replaced by A·G in *A. stipitatus*. Helices E21-5 and E21-6 are relatively conserved within the green algae. In this case, the sequences of the higher plants, *Saccharomyces cerevisiae* (Rubtsov et al. 1980), and *H. sapiens* (Gonzalez and Schmickel 1986) provide evidence for the existence of these helices.

In the helix E21 region, our model is most similar with that proposed for human 18S rRNA by Gonzalez and Schmickel (1986). Helices E21-2, E21-4, and E21-5 are almost identical. There is, however, no evidence within the plants for the extension of helices E21-1 and E21-6 as proposed by these authors. The human sequence, on the other hand, can be easily fitted in our model.

Discussion

The 16S-like rRNA distance data in Table 1 and the inferred phylogenetic tree shown in Fig. 2 demonstrate the heterogeneity of the genus *Chlorella*, and the relationship of some *Chlorella* species to other green algae. The tree topology in a parsimony analysis using the PAUP program (D. Swofford; Illinois Natural History Survey) is identical to that shown in Fig. 2.

Within the genus *Chlorella*, *C. vulgaris* and *C. kessleri* are closely related, but other *Chlorella* species are more closely related to different green algal taxa. A common ancestry of *C. minutissima* and *Nanochlorum eucaryotum* is significant although they are separated only by a short branch from the true chlorellae. The chlorophyte *N. eucaryotum* has been described as a new, very small alga with minimal eukaryotic features including a small cell size, a low DNA content, a possible total lack of histones along with deviations in the protein properties and morphology of the mitotic apparatus (Wilhelm et al. 1982; Zahn 1984). It has been proposed that *Nanochlorum* should be considered as a new genus in the family Oocystaceae that might represent an early divergence in the eukaryotic line of descent (Zahn 1984). The 16S-like rRNA sequence of *N. eucaryotum*, however, is similar to other green plants favoring evolutionary reduction as the source of the minimal features (Sargent et al. 1988).

Our data show that *Nanochlorum* is closely related to *C. minutissima*, which contradicts the original description where a relationship to *C. minutissima* was discounted (Wilhelm et al. 1982). Based upon comparative ultrastructural studies, *N. eucaryotum* was recently included in the genus *Nannochloris*, which contains algae similar to *Chlorella* but with smaller cell size (<5 μm) (Menzel and Wild 1989). It will be important to include *Nannochloris* in the rRNA analyses in order to determine if these microalgae form a coherent sister group of *Chlorella* and whether *Nanochlorum* is more closely related to *C. minutissima* or *Nannochloris*.

Prototheca is generally considered to be the achlorophyllous equivalent of *Chlorella*. These heterotrophic algae are found in sewage, feces, and soil (Pore et al. 1983) and are of medical and economic interest. Some are pathogenic for man and animals (Sudman 1974), others can degrade hydrocarbons including oil (Walker et al. 1975; Walker and Pore 1978). They formerly were assumed to be fungi but the presence of leucoplasts and propagation by means of autospores showed them to belong to the algae (Nadakavukaren and McCracken 1973). Shared features between *Prototheca* and *C. protothecoides* include a thiamine dependency and the incapability to assimilate nitrate. These features are not found in other *Chlorella* species. It has been proposed that *C. protothecoides* is the closest present-day representative of the algal progenitor that gave rise to the genus *Prototheca* (Pore 1972). *Prototheca wickerhamii* was included in this study because both taxa are similar with respect to shape, size, nutritional characteristics, and DNA base composition (Kessler 1982b; Pore 1985; Huss et al. 1988).

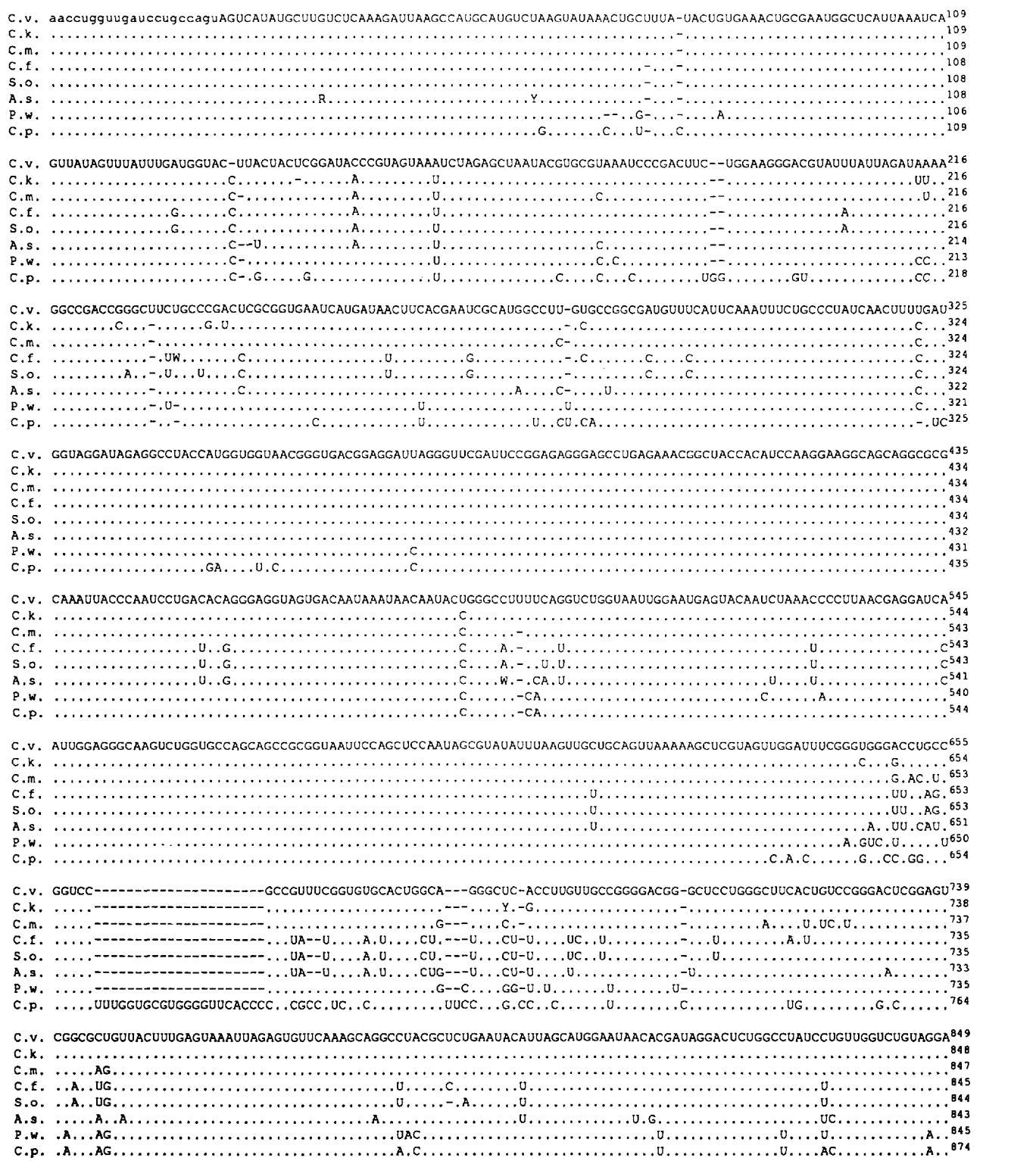

Fig. 1. Small subunit rRNA sequences of chlorococcalean algae aligned with *Chlorella vulgaris* 16S-like rRNA (Huss and Sogin 1989). Only the nucleotides that differ from those of *C. vulgaris* are shown (dots indicate nucleotide identity and dashes indicate alignment gaps). Ambiguity of nucleotides may be caused by PCR errors or by microheterogeneity within the 16S-like rRNA genes, as up to 12 different clones have been combined for preparing the sequencing template (R = A/G; W = A/U; Y = C/U; K = G/U). The absolute nucleotide positions are indicated for each sequence. For abbreviations of organisms see Table 1.

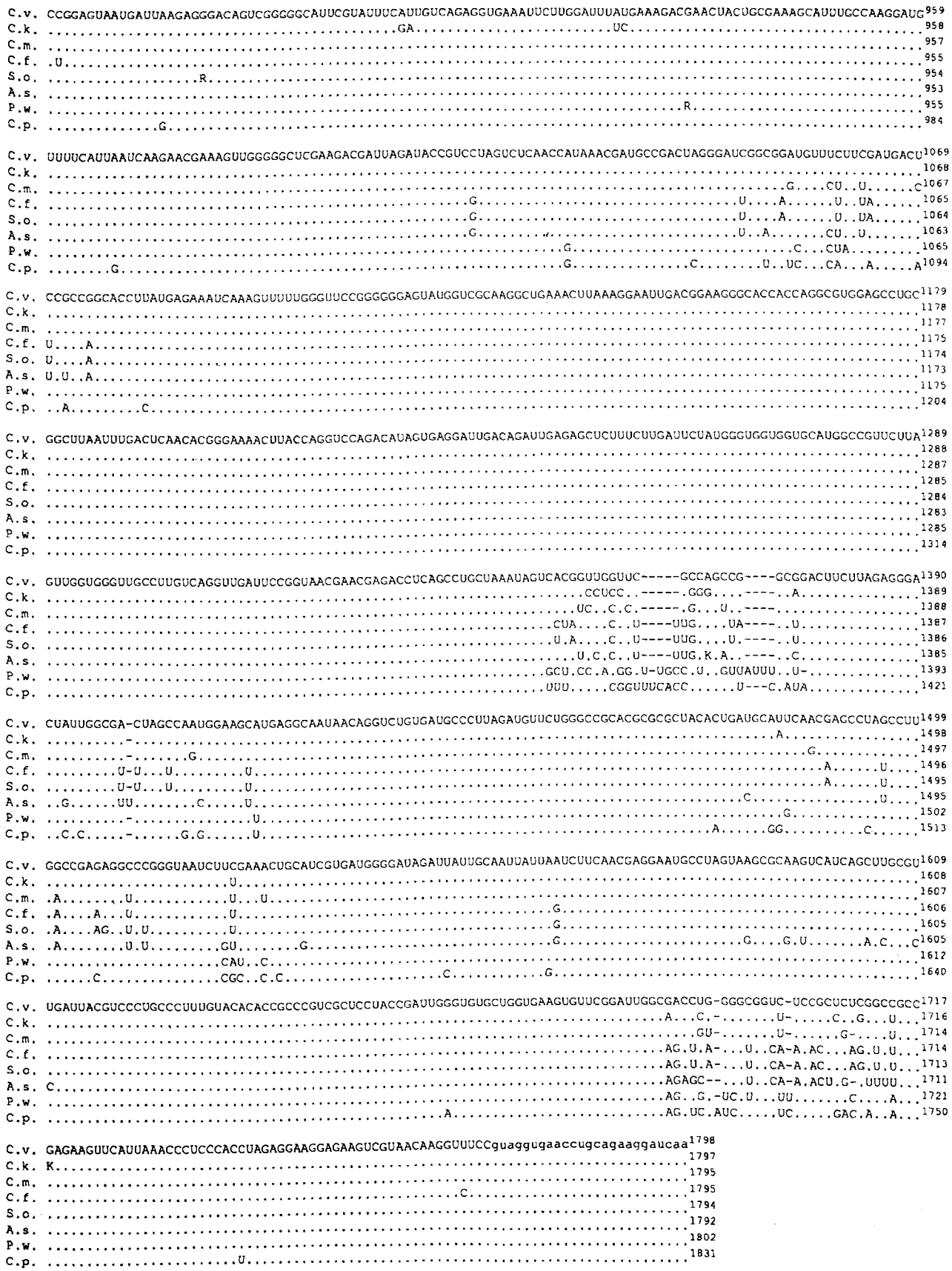

Fig. 1. Continued

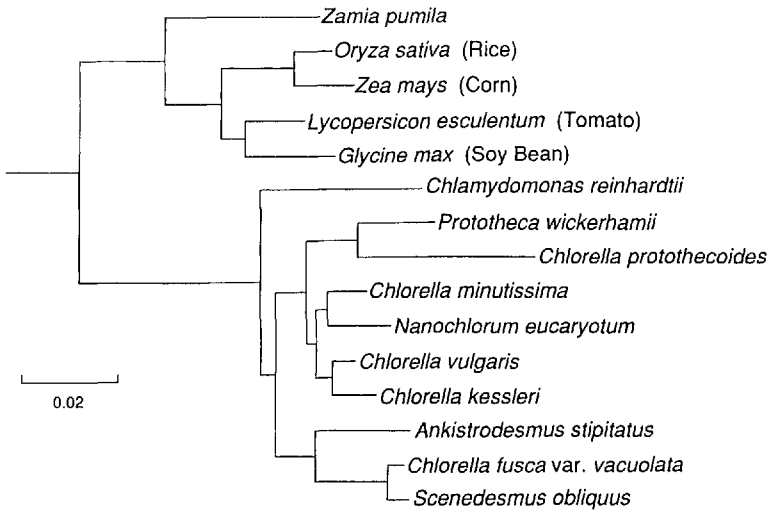

Fig. 2. Phylogenetic relationships within higher plants and green algae. The tree is inferred from the distance data in Table 1 and rooted using the sequence of *Acanthamoeba castellanii* (Gunderson and Sogin 1986). The horizontal component of separation represents the evolutionary distance between organisms. The scale indicates fixed mutations per sequence position.

Our phylogenetic tree indeed shows *C. protothecoides* and *P. wickerhamii* grouped together on a separate branch diverging prior to the separation of the microalgae and *Chlorella* as represented by *C. vulgaris* and *C. kessleri*. The question of the ancestry of *Prototheca*, however, cannot be regarded as being completely solved because the analysis of the relationship between both algae is complicated by two facts: First, the genus *Prototheca* is almost as heterogeneous as is *Chlorella*. DNA base compositions range from 60 to 76 mol% G+C, and no significant DNA similarities were found between most species (Kerfin and Kessler 1978; Huss et al. 1988). Second, the long branch in our phylogenetic tree leading to *C. protothecoides* shows it to be a fast clock organism, meaning that its rRNA has evolved more rapidly than in other algae. Fast clock sequences are known to cause problems with respect to branching orders in a phylogenetic tree (Felsenstein 1988). Additional sequences for other *Prototheca* species will be required to determine if *Prototheca* should be regarded as a monophyletic group of algae that lost the ability to synthesize chlorophyll after they diverged from an ancestor represented by *C. protothecoides*.

Electron microscopical investigations of cell wall structures have shown that *C. fusca* var. *fusca* belongs to the genus *Scenedesmus* (Fott et al. 1975). In that study, it is explicitly stated that the varieties *vacuolata* and *rubescens* are not related to *Scenedesmus*. The similarities of sterols, ribosomal proteins, and cytochromes c-553 (Patterson 1974; Götz and Arnold 1980a,b; Kümmel and Kessler 1980), however, indicate a relationship of all *C. fusca* varieties with *Scenedesmus* as suggested by Kessler (1984). This is also supported by DNA reassociation studies (Huss et al. 1989a). The 16S-like rRNA analysis confirms such a relationship for *C. fusca* var. *vacuolata*. It is surprising, however, that the RNA sequences of these morphologically different algae are almost identical with variation at only 15 positions. This shows that similarities and dissimilarities in morphology can be misleading in the inference of phylogenetic relationships.

The phylogenetic tree in Fig. 2 is consistent with the classification of *Chlorella* and *Prototheca* into the family Oocystaceae, and *Scenedesmus* into the family Scenedesmaceae (Bold and Wynne 1985). The genus *Ankistrodesmus* is commonly also placed into the Oocystaceae but *A. stipitatus* 202-5 is more closely related to *Scenedesmus* (Fig. 2). However, the DNA base composition of this strain (77.7 mol% G+C) is different from that characteristic for *Ankistrodesmus* (63–70 mol% G+C), and its taxonomic position has been questioned (Kessler 1980). Therefore, it may not be representative for the genus *Ankistrodesmus*. It is interesting to note that the 16S-like rRNA of *A. stipitatus* 202-5 includes a type I intron that is capable of a self-splicing process (Cech, personal communication).

Our data confirm the view of Kessler (1976, 1982a) that *Chlorella* in the traditional sense cannot be regarded as a natural genus. The evolutionary distance between *C. fusca* var. *vacuolata* and *C. protothecoides* is comparable with that between gymnosperms (represented by *Zamia pumila* in Table 1 and Fig. 2) and angiosperms. Excluding the *Scenedesmus* species "*C. fusca* var. *vacuolata*," there still remains a distance like that characteristic for the separation of mono- and dicotyledons. This, of course, also shows that the evolution of morphological diversity took quite a different path in the higher plants and in unicellular green algae like *Chlorella*. On the other hand, the only recently published 16S-like rRNA gene sequence from the colony-forming, coenobic green alga *Volvox carteri* dif-

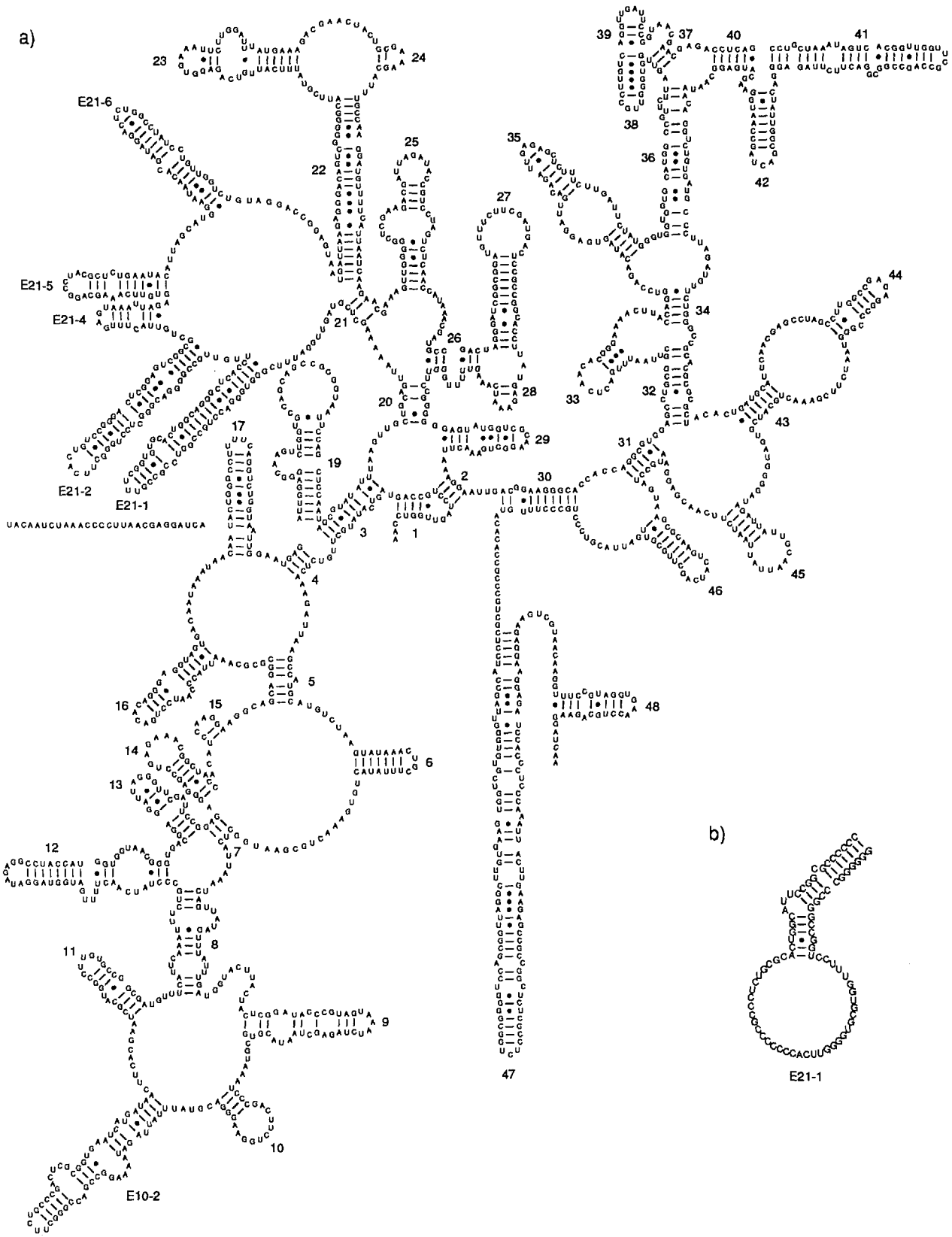

Fig. 3. **a** Secondary structure model for small subunit rRNA of *Chlorella vulgaris*. The model is based on that for *Zea mays* by Gutell et al. (1985). A new model for the eukaryote-specific region E21 is proposed for green algae and higher plants. Helix numbering is according to Dams et al. (1988). **b** Secondary structure model for helix E21-1 of *C. prototheocoides* containing a unique insertion of 26 nucleotides.

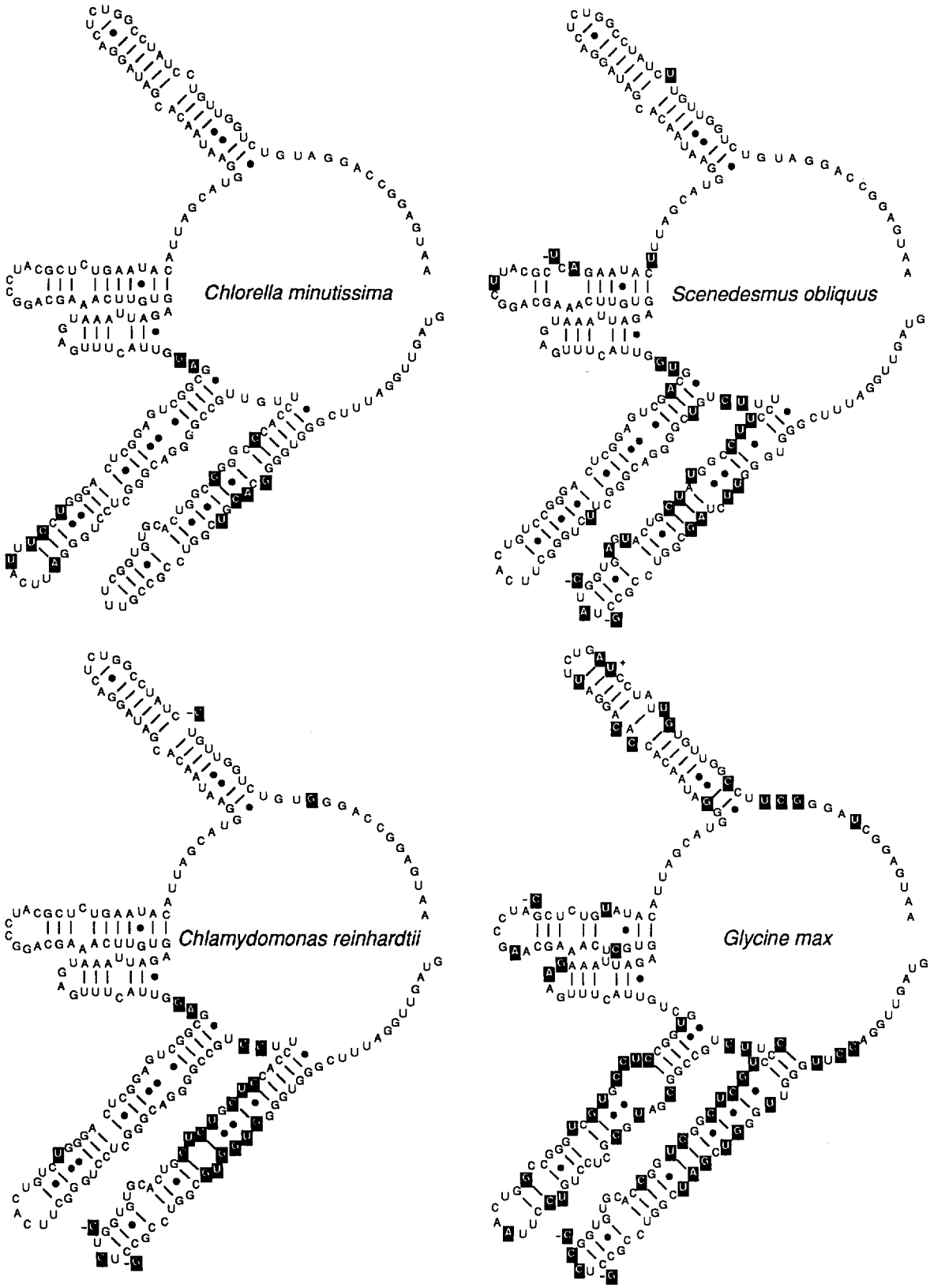

Fig. 4. Comparative secondary structure models for helix E21 of *Chlorella minutissima, Scenedesmus obliquus, Chlamydomonas reinhardtii,* and *Glycine max.* Nucleotide exchanges compared to *Chlorella vulgaris* are indicated by reversed letters; insertions and deletions by + and −, respectively. Compensatory exchanges that maintain the proposed helical structure are indicative of its phylogenetic conservation.

fers by only 27 positions compared with *C. reinhardtii* (Rausch et al. 1989), which is placed in a different family of the order Volvocales (Bold and Wynne 1985). This is less than found for the most closely related *Chlorella* species detected in our analysis.

The problems of applying classical taxonomical criteria to morphologically simple and asexually reproducing green algae are apparent not only from this study but also from the confusing continuous reclassification of strains throughout the past literature. The sequence analysis of RNA not only allows an insight into the evolution of green algae but may also solve many taxonomic problems even on the genus level.

Acknowledgments. We are grateful to E. Kessler for critically reading the manuscript. This work was supported by a National Institutes of Health Grant (GM32964) to M.L.S. and a stipend from the Deutscher Akademischer Austauschdienst to V.A.R.H.

References

Atmadja J, Brimacombe R, Maden BEH (1984) *Xenopus laevis* 18S ribosomal RNA: experimental determination of secondary structural elements, and locations of methyl groups in the secondary structure model. Nucleic Acids Res 12:2649–2667

Bold HC, Wynne MJ (1985) Introduction to the algae. Prentice-Hall, Englewood Cliffs NJ

Cech TR (1988) Conserved sequences and structures of group I introns: building an active site for RNA catalysis—a review. Gene 73:259–271

Choi YC (1985) Structural organization of ribosomal RNAs from Novikoff hepatoma. I. Characterization of fragmentation products from 40S subunit. J Biol Chem 260:12769–12772

Dams E, Hendriks L, Van de Peer Y, Neefs JM, Smits G, Vandenbempt I, De Wachter R (1988) Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res 16(suppl):r87–r173

Eckenrode VK, Arnold J, Meagher RB (1985) Comparison of the nucleotide sequence of soybean 18S rRNA with the sequences of other small-subunit rRNAs. J Mol Evol 21:259–269

Elwood HJ, Olsen GJ, Sogin ML (1985) The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. Mol Biol Evol 2: 399–410

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. Annu Rev Genet 22:521–565

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279–284

Fott B, Lochhead R, Clémençon H (1975) Taxonomie der Arten *Chlorella ultrasquamata* Clém. et Fott und *Chlorella fusca* Shih. et Krauss. Arch Protistenkd 117:288–296

Fott B, Nováková M (1969) A monograph of the genus *Chlorella*. The fresh water species. In: Fott B (ed) Studies in phycology. Academia, Prague, pp 10–74

Gonzalez IL, Schmickel RD (1986) The human 18S ribosomal RNA gene: evolution and stability. Am J Hum Genet 38: 419–427

Götz H, Arnold CG (1980a) Comparative electrophoretic study on ribosomal proteins from algae. Planta 149:19–26

Götz H, Arnold CG (1980b) Analysis of ribosomal proteins from various species of algae. Comparative electrophoretic study on proteins from chloroplast ribosomes. Biochem Physiol Pflanz 175:1–8

Gunderson JH, Sogin ML (1986) Length variation in eukaryotic rRNAs: small-subunit rRNAs from the protists *Acanthamoeba castellanii* and *Euglena gracilis*. Gene 44:63–70

Gunderson JH, Elwood HJ, Ingold A, Kindle K, Sogin ML (1987) Phylogenetic relationships between chlorophytes, chrysophytes and oomycetes. Proc Natl Acad Sci USA 84:5823–5827

Gutell RR, Weiser B, Woese CR, Noller HF (1985) Comparative anatomy of 16-S-like ribosomal RNA. Prog Nucleic Acid Res Mol Biol 32:155–216

Hegewald E (1982) Taxonomisch-morphologische Untersuchung von *Scenedesmus*-Isolaten aus Stammsammlungen. Algol Stud 29:375–406

Hellmann V, Kessler E (1974) Physiologische und biochemische Beiträge zur Taxonomie der Gattung *Chlorella*. VIII. Die Basenzusammensetzung der DNS. Arch Microbiol 95: 311–318

Hendriks L, De Baere R, Van Broeckhoven C, De Wachter R (1988) Primary and secondary structure of the 18S ribosomal RNA of the insect species *Tenebrio molitor*. FEBS Lett 232: 115–120

Herzog M, Maroteaux L (1986) Dinoflagellate 17S rRNA sequence inferred from the gene sequence: evolutionary implications. Proc Natl Acad Sci USA 83:8644–8648

Hindák F (1982) Taxonomic position of the chlorococcal alga *Chlorella zofingiensis* Dönz 1934 (Chlorophyceae). Algol Stud 30:13–23

Huss VAR, Sogin ML (1989) Primary structure of the *Chlorella vulgaris* small subunit RNA coding region. Nucleic Acids Res 17:1255

Huss VAR, Dörr R, Grossmann U, Kessler E (1986) Deoxyribonucleic acid reassociation in the taxonomy of the genus *Chlorella*. I. *Chlorella sorokiniana*. Arch Microbiol 145:329–333

Huss VAR, Schwarzwälder E, Kessler E (1987a) Deoxyribonucleic acid reassociation in the taxonomy of the genus *Chlorella*. II. *Chlorella saccharophila*. Arch Microbiol 147:221–224

Huss VAR, Hehenberger A, Kessler E (1987b) Deoxyribonucleic acid reassociation in the taxonomy of the genus *Chlorella*. III. *Chlorella fusca* and *Chlorella kessleri*. Arch Microbiol 149:1–3

Huss VAR, Wein KH, Kessler E (1988) Deoxyribonucleic acid reassociation in the taxonomy of the genus *Chlorella*. IV. *Chlorella protothecoides* and its relationship to the genus *Prototheca*. Arch Microbiol 150:509–511

Huss VAR, Huss G, Kessler E (1989a) Deoxyribonucleic acid reassociation and interspecies relationships of the genus *Chlorella* (Chlorophyceae). Plant Syst Evol 168:71–82

Huss VAR, Scharpf TK, Kessler E (1989b) Deoxyribonucleic acid reassociation in the taxonomy of the genus *Chlorella*. V. *Chlorella vulgaris*, *C. luteoviridis*, *C. minutissima*, and *C. zofingiensis*. Arch Microbiol 152:512–514

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–132

Kalina T, Punčochářová M (1987) Taxonomy of the subfamily Scotiellocystoideae Fott 1976 (Chlorellaceae, Chlorophyceae). Algol Stud 45:473–521

Kerfin W, Kessler E (1978) Physiological and biochemical contributions to the taxonomy of the genus *Prototheca*. II. Starch hydrolysis and base composition of DNA. Arch Microbiol 116:105–107

Kessler E (1976) Comparative physiology, biochemistry, and

the taxonomy of *Chlorella* (Chlorophyceae). Plant Syst Evol 125:129–138

Kessler E (1980) Physiological and biochemical contributions to the taxonomy of the genera *Ankistrodesmus* and *Scenedesmus*. V. Starch hydrolysis and new assignment of strains. Arch Microbiol 126:11–14

Kessler E (1982a) Chemotaxonomy in the Chlorococcales. In: Round FE, Chapman DJ (eds) Progress in phycological research, vol. 1. Elsevier Biomedical Press BV, Amsterdam, pp 111–135

Kessler E (1982b) Physiological and biochemical contributions to the taxonomy of the genus *Prototheca*. III. Utilization of organic carbon and nitrogen compounds. Arch Microbiol 132: 103–106

Kessler E (1984) A general review on the contribution of chemotaxonomy to the systematics of green algae. In: Irvine DEG, John DM (eds) Systematics of the green algae. Academic Press, London, pp 391–407

Kessler E (1987) Separation of *Chlorella ellipsoidea* from *C. saccharophila* (Chlorophyceae): no growth on mannitol and cadmium sensitivity. Plant Syst Evol 157:247–251

Kiss T, Szkukálek A, Solymosy F (1989) Nucleotide sequence of a 17S (18S) rRNA gene from tomato. Nucleic Acids Res 17:2127

Komárek J (1987) Species concept in coccal green algae. Algol Stud 45:437–471

Kümmel H, Kessler E (1980) Physiological and biochemical contributions to the taxonomy of the genus *Chlorella*. XIII. Serological studies. Arch Microbiol 126:15–19

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71:491–499

Menzel K, Wild A (1989) A comparative ultrastructural investigation of some *Nannochloris* species (Chlorococcales) with particular reference to the systematic position of *Nanochlorum eucaryotum*. Botanica Acta 102:152–158

Messing J (1983) New M13 vectors for cloning. Methods Enzymol 101:20–78

Messing J, Carlson J, Hagen G, Rubenstein I, Oleson A (1984) Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. DNA 3:31–40

Nadakuvakaren MJ, McCracken DA (1973) *Prototheca*: an alga or a fungus? J Phycol 9:113–116

Nairn CJ, Ferl RJ (1988) The complete nucleotide sequence of the small-subunit ribosomal RNA coding region for the cycad *Zamia pumila*: phylogenetic implications. J Mol Evol 27: 133–141

Patterson GW (1974) Sterols of some green algae. Comp Biochem Physiol 47B:453–457

Pore RS (1972) Nutritional basis for relating *Prototheca* and *Chlorella*. Can J Microbiol 18:1175–1177

Pore RS (1985) *Prototheca* taxonomy. Mycopathologia 90:129–139

Pore RS, Barnett EA, Barnes WC Jr, Walker JD (1983) *Prototheca* ecology. Mycopathologia 81:49–62

Raué HA, Klootwijk J, Musters W (1988) Evolutionary conservation of structure and function of high molecular weight ribosomal RNA. Prog Biophys Mol Biol 51:77–129

Rausch H, Larsen N, Schmitt R (1989) Phylogenetic relationships of the green alga *Volvox carteri* deduced from small-subunit ribosomal RNA comparisons. J Mol Evol 29:255–265

Rubtsov PM, Musakhanov MM, Zakharyev VM, Krayev AS, Skryabin KG, Bayev AA (1980) The structure of the yeast ribosomal RNA genes. I. The complete nucleotide sequence of the 18S ribosomal RNA gene from *Saccharomyces cerevisiae*. Nucleic Acids Res 8:5779–5794

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94:441–448

Sargent M, Zahn R, Walters B, Gupta R, Kaine B (1988) Nucleotide sequence of the 18S rDNA from the microalga *Nanochlorum eucaryotum*. Nucleic Acids Res 16:4156

Shihira I, Krauss RW (1965) *Chlorella*. Physiology and taxonomy of forty-one isolates. University of Maryland, College Park

Soeder CJ (1980) Massive cultivation of microalgae: results and prospects. Hydrobiologia 72:197–209

Sogin ML, Ingold A, Karlok M, Nielsen H, Engberg J (1986) Phylogenetic evidence for the acquisition of ribosomal RNA introns subsequent to the divergence of some of the major *Tetrahymena* groups. EMBO J 5:3625–3630

Sogin ML, Gunderson JH, Elwood HJ, Rogelio AA, Peattie DA (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science 243: 75–77

Sudman MS (1974) Protothecosis. Am J Clin Pathol 61:10–19

Takaiwa F, Oono K, Sugiura M (1984) The complete nucleotide sequence of a rice 17S rRNA gene. Nucleic Acids Res 12: 5441–5448

Walker JD, Colwell RR, Vaituzis Z, Meyer SA (1975) Petroleum-degrading achlorophyllous alga *Prototheca zopfii*. Nature 254:423–424

Walker JD, Pore RS (1978) Growth of *Prototheca* isolated on n-hexadecane and mixed hydrocarbon substrate. Appl Environ Microbiol 35:694–697

Wilhelm C, Eisenbeis G, Wild A, Zahn R (1982) *Nanochlorum eucaryotum*: a very reduced coccoid species of marine Chlorophyceae. Z Naturforsch 37c:107–114

Woese CR, Gutell R, Gupta R, Noller HF (1983) Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol Rev 47:621–669

Woese CR, Fox E (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Zahn RK (1984) A green alga with minimal eukaryotic features: *Nanochlorum eucaryotum*. Origins Life 13:289–303

Zuckerkandl E, Pauling L (1965) Molecules as documents of evolutionary history. J Theor Biol 8:357–366

Received February 6, 1990/Revised May 23, 1990