J. Phycol. **45,** 1206–1212 (2009)

DOI: 10.1111/j.1529-8817.2009.00731.x

ANALYSIS OF A PLASTID MULTIGENE DATA SET AND THE PHYLOGENETIC POSITION OF THE MARINE MACROALGA *CAULERPA FILIFORMIS* (CHLOROPHYTA)[1]

G. C. Zuccarello[2]*, Natalie Price*

School of Biological Sciences, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand

Heroen Verbruggen and Frederik Leliaert

Phycology Research Group and Center for Molecular Phylogenetics and Evolution, Ghent University, Krijgslaan 281 (S8), B-9000 Gent, Belgium

Molecular phylogenetic relationships within the Chlorophyta have relied heavily on rRNA data. These data have revolutionized our insight in green algal evolution, yet some class relationships have never been well resolved. A commonly used class within the Chlorophyta is the Ulvophyceae, although there is not much support for its monophyly. The relationships among the Ulvophyceae, Trebouxiophyceae, and Chlorophyceae are also contentious. In recent years, chloroplast genome data have shown their utility in resolving relationships between the main green algal clades, but such studies have never included marine macroalgae. We provide partial chloroplast genome data (~30,000 bp, 23 genes) of the ulvophycean macroalga *Caulerpa filiformis* (Suhr) K. Herig. We show gene order conservation for some gene combinations and rearrangements in other regions compared to closely related taxa. Our data also revealed a pseudogene (*ycf62*) in *Caulerpa* species. Our phylogenetic results, based on analyses of a 23-gene alignment, suggest that neither Ulvophyceae nor Trebouxiophyceae are monophyletic, with *Caulerpa* being more closely related to the trebouxiophyte *Chlorella* than to *Oltmannsiellopsis* and *Pseudendoclonium*.

Key index words: ***Caulerpa*; Chlorophyta; chloroplast genome; phylogenetics; systematics; Ulvophyceae**

Abbreviations: **tufA, elongation factor Tu; UTC, Ulvophyceae+Trebouxiophyceae+Chlorophyceae**

The Viridiplantae (sensu Cavalier-Smith 1981) is a major eukaryotic lineage characterized by a primary plastid containing chl *a* and *b*. The lineage contains both the green algae and their descendents, the land plants. The taxonomic history of the green algae has gone through several major changes during the last few decades (reviewed in Lewis and McCourt 2004, Pröschold and Leliaert 2007).

With molecular and ultrastructural data certain evolutionary trends are clear. The Viridiplantae are divided into two distinct lineages, the Streptophyta and the Chlorophyta (Bremer 1985). The Streptophyta includes the land plants and their sister clades, a paraphyletic assemblage of green algae (known as charophyte green algae). The Chlorophyta includes the remaining green algae belonging to four classes. The Prasinophyceae are the earliest diverging Chlorophyta and form a paraphyletic assemblage at the base of the Chlorophyta (Steinkötter et al. 1994, Guillou et al. 2004, Nakayama et al. 2007). The Trebouxiophyceae was first proposed in 1995 based on SSU data (Friedl 1995), but support from molecular data has not always been satisfactory and depends strongly on the taxon sets analyzed (Krienitz et al. 2003, Lokhorst et al. 2004). The other two classes contain the more characteristic freshwater green algae (Chlorophyceae, taxa such as *Volvox*, *Chlamydomonas*) and the mainly marine Ulvophyceae, containing some of the best-known green seaweeds, such as *Acetabularia, Cladophora, Codium,* and *Ulva*. Phylogenies based on nuclear-encoded SSU rRNA gene data mostly support the monophyly of the Ulvophyceae+Trebouxiophyceae+Chlorophyceae (the so-called UTC clade), but the relationships among these classes, and the monophyly of Trebouxiophyceae and Ulvophyceae remain contentious (Watanabe and Nakayama 2007).

The Ulvophyceae was originally proposed based on the flagellar root system (offset counterclockwise and overlapping basal bodies) and cytokinesis (furrowing with a persistent interzonal spindle) (Mattox and Stewart 1984, O'Kelly and Floyd 1984, Sluiman 1989). Many members of the class have a haplodiplontic or a diplontic life cycle, distinct from the majority of Chlorophyta with a haplontic life cycle. The Ulvophyceae are morphologically diverse, ranging from unicellular algae to plants with filamentous, parenchymatous, siphonocladous, and siphonous thallus construction. Within the Ulvophyceae, six main groups are recognized and ranked as orders

[1]Received 24 August 2008. Accepted 25 February 2009.
[2]Author for correspondence: e-mail joe.zuccarello@vuw.ac.nz.

(Ulvales, Ulotrichales, Bryopsidales, Dasycladales, Cladophorales, and Trentepohliales), although some authors favored their recognition at the class level (van den Hoek et al. 1995). Early molecular studies based on partial rRNA data (Zechman et al. 1990) did not resolve a monophyletic Ulvophyceae. These data are especially problematic due to the highly divergent nature of the SSU in four ulvophycean orders (Bryopsidales, Cladophorales, Dasycladales, Trentepohliales), and subsequent studies only produced a moderate to poorly supported Ulvophyceae, depending on taxon sampling (see, e.g., Friedl and O'Kelly 2002, Lopez-Bautista and Chapman 2003, Watanabe and Nakayama 2007).

Chloroplast phylogenomics has proved useful for elucidating relationships among early diverging lineages of green algae and land plants (e.g., Pombert et al. 2005, Turmel et al. 2008). The chloroplast genomes of two members of the Ulvophyceae have been sequenced so far, both of them being unicellular and hence not characteristic of this predominantly macroalgal class. *Pseudendoclonium akinetum* (for taxonomic authors, see Table S1 in the supplementary material) is a freshwater flagellate member of the Ulvales (Floyd and O'Kelly 1990), with unusual chloroplast genome architecture (e.g., large number of Group I introns) (Pombert et al. 2005). *Oltmannsiellopsis viridis* is a marine flagellate with a complex taxonomic history. First proposed to belong to the Chlorophyceae (Hargraves and Steele 1980, Chihara et al. 1986) and later believed to be an early diverging ulvophyte (Friedl and O'Kelly 2002, O'Kelly et al. 2004, Pombert et al. 2006). Phylogenetic analyses of chloroplast gene data from these two taxa support the monophyly of the Ulvophyceae.

The only Ulvophycean marine macroalgae in which the chloroplast genome has been investigated, based on restriction fragment analysis, mostly with reference to genome size, is *Codium fragile* having an estimated genome size of only 89 kb (Manhart et al. 1989), one of the smallest green algal chloroplasts known at the time, and *Caulerpa sertularioides* with an estimated size of 131.4 kb (Lehman and Manhart 1997).

Caulerpa is a conspicuous member of the marine ulvophycean order Bryopsidales. The large thallus is siphonous, consisting of a contiguous cytoplasm, with the millions of organelles circulated through the cytoplasm by cytoplasmic streaming. It is common in tropical and warm temperate seas and is composed of a morphologically diverse siphonous thallus. The phylogeny of many species has been performed using the plastid gene *tuf*A and/or nuclear ribosomal cistrons (Famà et al. 2002, Stam et al. 2006). The Bryopsidales also contains many other marine macroalgae common to coastal shores (e.g., *Codium*, *Bryopsis*, *Halimeda*).

We cloned and sequenced pieces of the chloroplast genome of *Caulerpa filiformis*, aiming to produce a multigene data set that can be used to improve our understanding of the phylogenetic relationships among UTC taxa and, more specifically, to have the Ulvophyceae represented by a more typical macroalga example. A secondary goal is to gain further insights into the evolution of the chloroplast genome in this marine alga (e.g., gene order, intergenic spacer size).

MATERIALS AND METHODS

Collection, chloroplast isolation, cpDNA extraction, and sequencing. Plants of *C. filiformis* were collected from Tamarama, New South Wales, Australia, on 4 November 1997. Assimilators of *C. filiformis* were picked that lacked visible epiphytes and looked "healthy" (i.e., dark green color). Chloroplasts were isolated following the procedure described by Palmer (1986) (see Appendix S1 in the supplementary material). cpDNA was subsequently extracted using a phenol-chloroform extraction protocol and further purified by CsCl ultracentrifugation (see Appendix S2 in the supplementary material). This resulted in a single high-molecular weight band, as visualized on an agarose gel. cpDNA was digested with two restriction enzymes, ligated into plasmids and subsequently cloned and sequenced (see Appendix S3 in the supplementary material). Sequencing of plasmids involved using a single vector primer (M13F) to determine if a plasmid contained a chloroplast insert. If a National Center for Biotechnology Information (NCBI) BLAST search (http://www.ncbi.nlm.nih.gov) produced matches to known chloroplast genes, the plasmid was set aside. A subset of these "chloroplast-positive" inserts was sequenced to completion, by primer-walking in both directions until the insert was completely sequenced. Sequences were compiled and edited using ABI software (Sequence Assembler; Applied Biosystems, Foster City, CA, USA).

The sequences were searched for open reading frames (ORFs) using ORF-Finder at the NCBI. The bacterial genetic code was selected for this search, and minimum ORF length was set at 100 nucleotides. Amino acid BLAST searches of the ORFs found were performed by BLASTp using the nonredundant database of the NCBI. From the BLASTp search result, it was usually obvious if the ORF was a true gene, as the sequence would have significant homology with genes in the database from chloroplast genomes. ORFs that were not true genes either returned a "no significant similarity found" result or had weak homology with a gene that was unrelated to the chloroplast. A threshold *E*-value of 10^{-6} was used, and ORFs with a score higher than this were disregarded.

tRNA genes were located using the program tRNAscan-SE v. 1.21 (Lowe and Eddy 1997; http://lowelab.ucsc.edu/tRNAscan-SE/) with default search mode and source selected as "mito/chloroplast."

ycf62 analysis. Our analysis of an ORF that corresponded to the putative protein ycf62 (RF62) indicated that a stop codon was present in the middle of the gene compared to all other green algae, with an alternate reading frame producing the 3′ of the protein. To show that this was not a cloning artifact, we analyzed this region further. Primers were produced to span this midgene stop codon and alternate reading frame, ycf62for (5′-TTAGGAGAACAGAGAGCTCGAAA-3′) and ycf62rev (5′-TCCAATCAGGTA GAATAGGCAAAT-3′). DNA was extracted using a cetyl trimethyl ammonium bromide (CTAB) procedure (Zuccarello and Lokhorst 2005). Standard PCR conditions with an annealing temperature of 45°C were performed on other samples of *C. filiformis* and other *Caulerpa* species (*C. cactoides*, *C. flexilis*, *C. geminata*, *C. scalpelliformis*) (GenBank accession numbers FJ565672–FJ565676).

Phylogenetic analysis. The data set consisted of the 23 genes that we sequenced in *Caulerpa* (*atp*A, *atp*B, *atp*H, *atp*I, *cem*A, *chl*B, *chl*I, *chl*L, *chl*N, *inf*A, *psa*B, *psb*C, *psb*D, *psb*Z, *rpl2*, *rpl5*, *rpl36*, *rps2*,

rps3, rps8, rps11, rps19, ycf4), aligned individually to available chloroplast genome data from the green plants (with only a small selection of land plants chosen) (Table S1) using Se-Al version2a11 (Rambaut 1996). Ambiguously aligned regions, mostly at the 5′ or 3′ end, were removed, and the sequences were then concatenated to produce a data set of 16,824 aligned nucleotide positions (Appendix S4 in the supplementary material), corresponding to 5,608 amino acid positions.

The concatenated data sets of protein and DNA sequences were analyzed with Bayesian inference (BI) and maximum likelihood (ML), using MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003) and PhyML v2.4.4 (Guindon and Gascuel 2003), respectively. Analyses at the amino-acid level assumed an evolutionary model consisting of the CpREV matrix of amino acid substitution (Adachi et al. 2000) with gamma distribution split into four categories to model the rate heterogeneity among sites (CpREV+G4). Analyses at the DNA level (only the first two codon positions used) assumed a general-time-reversible model with a proportion of invariable sites and among site rate heterogeneity following a gamma distribution split into eight discrete rate categories (GTR+I+G8). BI analyses consisted of two parallel runs of each four incrementally heated chains, and 3×10^6 generations with sampling every 1,000 generations. A burn-in sample of 1,000 trees (well beyond convergence of likelihood values) was removed before constructing the majority-rule consensus tree. For the ML trees, the reliability of internal branches was evaluated based on 500 nonparametric bootstrap replicates.

Shimodaira–Hasegawa tests (SH; Shimodaira and Hasegawa 1999) and approximately unbiased tests (AU; Shimodaira 2002) were performed with CONSEL 0.1i (Shimodaira and Hasegawa 2001) on the amino acid and nucleotide data sets to test: (1) the monophyly of Ulvophyceae (*Oltmannsiellopsis* + *Pseudendoclonium* + *Caulerpa* constrained as monophyletic) and (2) the monophyly of the Trebouxiophyceae (*Chlorella* + *Leptosira* constrained as monophyletic). To obtain the ML tree for a specified hypothesis, we used constrained trees and the "resolve multifurcations" option of TreeFinder version June 2008 (Jobb et al. 2004) with edge length estimation under the same evolutionary models as specified above. The site-specific likelihoods used as input for CONSEL were generated by ML model optimization in PAML 4.0 using the exact same model specifications mentioned above (Yang 2007).

RESULTS

Partial chloroplast genome data. Restriction enzyme digests of purified chloroplast DNA, visualized on 1% agarose gels and band sizes calculated, estimated the size of the *C. filiformis* chloroplast genome between 109.7 and 120.5 kb. Seven clones were analyzed corresponding to 30,348 bp of the chloroplast of *C. filiformis* (GenBank accessions FJ565677–FJ565683). The clones contained homology to the following chloroplast genes: *atp*A, *atp*B, *atp*H, *atp*I (ATP synthase); *chl*B, *chl*I, *chl*L (chl biosynthesis); *inf*A (translation factor); *psa*A, *psa*B (PSI); *psb*C, *psb*D, *psb*Z (PSII); *rpl2, rpl5, rpl36* (LSU ribosomal proteins); *rps2, rps3, rps8, rps11, rps19* (SSU ribosomal proteins); *ycf4, ycf62* (conserved proteins); *cem*A, two tRNA genes, *trn*R(ccg) and *trn*P(cca); plus two ORFs showing homology to group II intron reverse transcriptases. Three clones contained three or more genes. One clone (960 bp in length) contained the genes *rpl2, rps19, rps3*, a gene order that is also found in *Oltmannsiellopsis, Pseudendoclonium,* and *Chlorella* (although in *Chlorella* an ORF is found between *rpl2* and *rps19*) (Fig. 1a). Another clone (2,818 bp) contained seven genes in a very conserved order (Fig. 1b). This order is only different in that *Caulerpa* has the *psa*A gene linked to *rpo*A, while in the other taxa, it is found elsewhere in the genome. The intergenic spacer regions between the genes in this cluster are very short compared to the other analyzed gene clusters (Fig. 1b). The largest clone analyzed (15,430 bp) contained 11 genes in a highly rearranged order when compared to other closely related genomes, except for the groupings of *atp*H, *atp*I and *rps2*, and *psb*D and *psb*C. The *psb*D gene in *Caulerpa* has a single intron, while the *psb*C gene is followed with an ORF, showing homology to a putative group II maturase/reverse transcriptase, which falls at the end of the clone.

Pseudogene analysis of ycf62. The hypothetical protein ycf62 (RF62) has a reading frame with a stop codon at amino acid position 164 (492 bp) from the start methionine. Another ORF is found starting from position 482 bp from the start nucleotide. Both translated ORFs have homology to ycf62. To investigate this unusual pattern, we analyzed the region around this stop codon in other isolates of *C. filiformis* and different *Caulerpa* species. The sequences (350 bp) of the other *C. filiformis* samples were identical to the original chloroplast clone. The five other *Caulerpa* species all had a stop codon in a similar position. *C. scalpelliformis* had a stop codon at position 164, while *C. flexilis* had a stop codon at position 160, and *C. cactoides* and *C. geminata* both

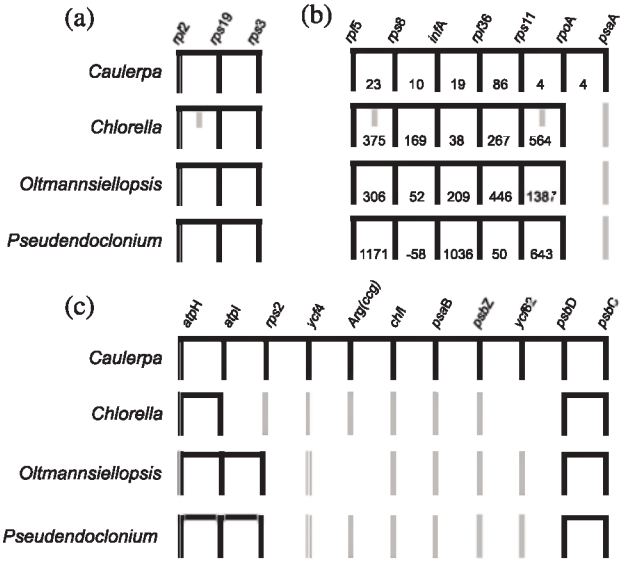

FIG. 1. Gene order of other UT chloroplast genomes compared to the order found in *Caulerpa filiformis*. Black vertical lines represent genes, names at top. If genes linked, connected by a horizontal line. Unlinked gene in vertical gray line. Short gray vertical lines are inserted open reading frames (ORFs). Gene missing from the genomes left empty. In cluster (b) length of spacer region, in bp, between the named genes from the respective genomes shown.

had stop codons at 162. Deletions followed the stop codon in these species, compared to *C. filiformis*, and there is no similarity match for the translated product after the stop codon.

Phylogenetic analysis. Analyses of amino acid and nucleotide data resulted in virtually identical tree topologies, though with very dissimilar branch support. The BI analysis of protein data resulted in an almost fully resolved phylogenetic tree (Fig. 2), while node support (notably among members of the UTC clade) was remarkably lower in the other three analyses. Relationships within the Streptophyta were well resolved in all analyses, with *Mesostigma* and *Chlorokybus* united in an early branching clade. *Chara* (Charophyceae) and *Chaetosphaeridium* (Coleochaetophyceae) branch off next with the two Zygnematophyceae (*Zygnema*, *Staurastrum*) forming a clade sister to the land plants. Within the Chlorophyta, the prasinophycean representatives (*Nephroselmis*, *Ostreococcus*) diverge first with the remaining taxa belonging to the UTC clade strongly supported in all analyses. While the Chlorophyceae samples form a strongly supported clade, our results indicate that neither the Trebouxiophyceae nor the Ulvophyceae are monophyletic. The two trebouxiophycean algae (*Leptosira* and *Chlorella*) do not group together. *Leptosira* is recovered as sister to either the Chlorophyceae (BI and ML protein tree and ML nucleotide analysis) or the *Caulerpa–Chlorella–Oltmannsiellopsis–Pseudendoclonium* clade (BI nucleotide analysis). While the ulvophycean *Pseudendoclonium* and *Oltmannsiellopsis* group together strongly, *Caulerpa* is recovered as the sister taxon of *Chlorella* in all analyses, although with only low support in the nucleotide analyses. Monophyly of Ulvophyceae and Trebouxiophyceae could not be ruled out by Shimodaira–Hasegawa topology tests based on both the amino acid and nucleotide data, but the approximately unbiased (AU) test did reject monophyly of the Ulvophyceae based on both data sets and monophyly of Trebouxiophyceae based on the amino acid data (Table 1).

DISCUSSION

This is the first, although only partial, chloroplast genome sequence of a marine macrophyte belonging to the Ulvophyceae. The estimated length of the chloroplast genome in *C. filiformis* (average 115.1 kb) is slightly smaller than estimates made of the chloroplast in *C. sertularioides* (131.4 kb) (Lehman and Manhart 1997) but larger than in *Codium* (89 kb) (Manhart et al. 1989). This length would indicate that we sequenced ~26% of the genome. Our data suggest several intriguing evolutionary insights that are revealed by the inclusion of *C. filiformis*, a morphologically quite different species from the other two known ulvophycean taxa from which chloroplast genomes have been determined (i.e., *Pseudendoclonium*, *Oltmannsiellopsis*) (Pombert et al. 2005, 2006).

Analysis of the three clones containing multiple genes shows a conserved gene order in two clones and scrambled gene order in another. The gene order (*rpl2*, *rps19*, *rps3*; Fig. 1a) is conserved in the three genera most closely related to *Caulerpa*, with only *Chlorella* having an ORF between genes *rpl2* and *rps19*. The gene order in the clone containing *rpl5* (Fig. 1b) is also conserved in other green algae and has been used for the construction of the universal green algal plastid primer combinations UCP1, UCP2, UCP3 (Provan et al. 2004). The gene

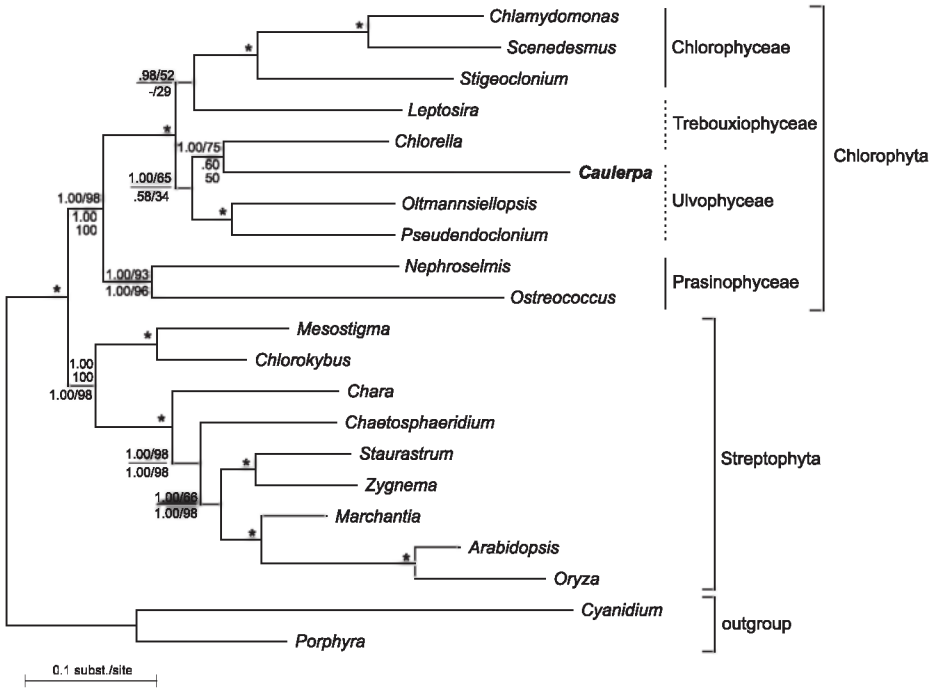

FIG. 2. Phylogenetic tree resulting from Bayesian inference (BI) of 23 chloroplast genes at the amino acid level, showing the position of *Caulerpa filiformis* among the Ulvophyceae+Trebouxiophyceae+Chlorophyceae (UTC) taxa. Values above branches indicate BI posterior probabilities and maximum-likelihood (ML) bootstrap support from the analyses of protein data under a CpREV+G8 model; values below branches indicate BI posterior probabilities and ML bootstrap support from the analyses of nucleotide data (1st and 2nd codon position) under a GTR+G8 model. The nodes that received full support in BI and ML analyses of both protein and nucleotide data are denoted by asterisks. Solid brackets on the right indicate monophyletic taxa; dashed brackets indicate nonmonophyletic taxa. GTR, general time reversible.

TABLE 1. Topology test results. AU, approximately unbiased test, SH, Shimodaira–Hasegawa test (*P*-values).

	Amino acid data		Nucleotide data	
Phylogenetic hypothesis	AU	SH	AU	SH
Ulvophyceae monophyly	0.035	0.086	0.002	0.111
Trebouxiophyceae monophyly	0.025	0.050	0.185	0.225

order is slightly rearranged in *Caulerpa*, as this genus has *psa*A linked to *rpo*A, while in the other taxa, it is elsewhere in the genome. In *Chlorella* and *Oltmannsiellopsis*, for example, *psa*A is linked to *psa*B. In *Caulerpa*, *psa*B is found in another gene cluster associated with *chl*I and *psb*Z (Fig. 1c).

A comparison of the lengths of the spacer between the four genome segments shows that the *Caulerpa* genes have markedly shorter spacers between them. While the chloroplast genome size as a whole is only approximate (115.1 kb), it is significantly smaller than the genomes of *Chlorella*, *Oltmannsiellopsis*, and *Pseudendoclonium* (150 kb, 151 kb, and 156 kb, respectively) (Wakasugi et al. 1997, Pombert et al. 2005, 2006). Variation in noncoding organelle DNA spacers length has been shown to be nonrandom in land plants with various evolutionary forces leading to short spacers (Duminil et al. 2008). It is tempting to speculate that in a siphonous thallus with millions of circulating plastids, selection for more rapidly copied (i.e., shorter) plastid genome, and hence plastids, may be favored and lead to short intergenic spacers. A comparison of chloroplasts in other siphonous or coenocytic thalli, compensating for phylogenetic relationship, in the orders Bryopsidales, Dasycladales, and Cladophorales are needed.

In the third clone, the gene order is quite scrambled compared to the other genome investigated, with only the proximity of *psb*D and *psb*C to each other conserved in all taxa and the grouping of *atp*H and *apt*I, a gene order also conserved in most green and red algae. Interestingly, *atp*H, *apt*I, and *rps2* is a gene order conserved in all three Ulvophyceae studied, and not present in *Chlorella*.

The *ycf62* hypothetical gene has been identified in many green algal genomes (e.g., *Leptosira*, *Nephroselmis*, *Pseudendoclonium*) (Turmel et al. 1999, Pombert et al. 2005, de Cambiaire et al. 2007). Its product is the tRNA (Ile)-lysidine synthase; it ligates, in an ATP-dependent manner, lysine onto the cytidine present at position 34 of the AUA codon-specific tRNA (Ile) that contains the anticodon CAU. In *Caulerpa* species, this gene appears to be a pseudogene as it contains a stop codon, and in some species, indels of different lengths are found. Pseudogenes are not common in chloroplast genomes but have been reported (e.g., Steane 2005, Koch et al. 2007, Tsuji et al. 2007). It would be useful to look at the evolution of this gene within the Bryopsidales and related orders to determine when the loss of function occurred.

The relationships observed in our phylogeny based on 23 chloroplast genes for the Viridiplantae in general and streptophytes in particular are congruent with recently published green plant phylogenies based on whole chloroplast genome sequences (Lemieux et al. 2007). The inclusion of *C. filiformis* produced a phylogeny in which the monophyly of the Ulvophyceae is not supported, although our data cannot significantly reject monophyly, except in the AU test. The two ulvophytes *Pseudendoclonium* and *Oltmannsiellopsis* group together as they do in other chloroplast gene phylogenies (Robbens et al. 2007, Turmel et al. 2008), while *Caulerpa* groups with the trebouxiophycean *Chlorella*, although with only moderate support. Interestingly, *Leptosira*, the other trebouxiophyte in the analysis, does not group with *Chlorella*, and this lack of monophyly is again only supported in the AU test.

Molecular data supporting the monophyly of both Trebouxiophyceae and the Ulvophyceae are limited. Most analyses using SSU nrDNA data show only moderate support for the classes (e.g., Lopez-Bautista and Chapman 2003, Watanabe and Nakayama 2007). O'Kelly and Floyd (1984) hypothesized that the ancestral orientation of the flagellar apparatus in the UTC clade was counterclockwise (CCW) and further evolved to a direct-opposite (DO) and clockwise (CW) orientation in the Chlorophyceae. This prediction was recently supported and refined by Turmel et al. (2008). The unexpected phylogenetic relationships between *Caulerpa*, *Chlorella*, and *Leptosira* (all with a CCW orientation) do not put a different complexion on this evolutionary hypothesis.

The relationships among the UTC classes have never been fully resolved. The evolutionary scenario in which Trebouxiophyceae branch first, leaving Chlorophyceae and Ulvophyceae as closest relatives, is the preferred relationship based on several organellar genomic characters (Pombert et al. 2004, 2005). Our data suggest that the idea of a monophyletic class Ulvophyceae should be questioned. With the inclusion of *C. filiformis* chloroplast data, it becomes clear that the quest for the branching order among UTC taxa is more complex than previously assumed and that it is critically dependent on taxon sampling (Hillis et al. 2003, Heath et al. 2008). In order to come to a reasonable level of understanding of the evolution of the UTC clade, chloroplast genomic data of other members of the class, including groups such as the Cladophorales, Dasycladales, and Trentepohliales, will have to be gathered and analyzed.

The apparent contradiction between the results of the approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests requires attention because it would lead to considerably different interpretations concerning the monophyly of the Ulvophyceae and Trebouxiophyceae. Whereas the SH tests uniformly suggested that monophyly of the Ulvophyceae and Trebouxiophyceae was not significantly

worse than the ML result (nonmonophyly), all but one of the AU tests rejected the monophyly of these classes. As applied in our study, SH and AU both set out to test whether the likelihood of the ML phylogeny (nonmonophyly of the classes) is significantly better than that of the alternative phylogeny (monophyly of the classes), but they are based on different philosophies, which gives them different properties. The SH test is known to be biased toward the null hypothesis, meaning that it can be misleading in indicating that the likelihood of the ML tree and the alternative topology are not statistically different (Shimodaira 2002). The AU test was designed to counter this problem and is generally recommended for tree selection problems (Shimodaira 2002). We follow this recommendation and conclude that monophyly of the Ulvophyceae is rejected by our data set. As for the Trebouxiophyceae, the AU tests suggest that its monophyly is doubtful: the analysis of the amino-acid alignment rejects its monophyly, but the analysis of the nucleotide alignment does not. We cannot presently explain this conflict between the amino acid and nucleotide data of the same genes, and this conflict certainly warrants future attention.

In conclusion, our study casts doubt on the monophyly of the chlorophytan classes Ulvophyceae and Trebouxiophyceae. This insight will need to be explored further using much greater sampling within the marine seaweeds. Further study using genomic approaches, including full plastid sequencing of macroalgae, will prove useful in not only clarifying evolutionary relationships within green algae but also in understanding the patterns and mechanisms of plastid rearrangements, and the evolution of chloroplast genomes.

We thank the many people that have helped with this project through the years: Patrizia Famà, Torsten Meissner, Dierdre Sharkey, and Jeff Wright. We thank two anonymous reviewers for helpful comments. Phylogenetic analyses were run on the KERMIT computing cluster (Ghent University) and the Computational Biology Service Unit (Cornell University and Microsoft Corporation, http://cbsuapps.tc.cornell.edu/mrbayes.aspx). H. V. and F. L. are indebted to the Research Foundation – Flanders (FWO) for postdoctoral fellowship grants. G. C. Z. received funding from BOF (Ghent University) for a sabbatical at the Phycology Research Group in Ghent.

Adachi, J., Waddell, P. J., Martin, W. & Hasegawa, M. 2000. Plastid genome phylogeny and a model of amino acid substitution for proteins encoded by chloroplast DNA. *J. Mol. Evol.* 50:348–58.

Bremer, K. 1985. Summary of green plant phylogeny and classification. *Cladistics* 1:369–85.

de Cambiaire, J. C., Otis, C., Turmel, M. & Lemieux, C. 2007. The chloroplast genome sequence of the green alga *Leptosira terrestris*: multiple losses of the inverted repeat and extensive genome rearrangements within the Trebouxiophyceae. *BMC Genomics* 8:213.

Cavalier-Smith, T. 1981. Eukaryote kingdoms: seven or nine? *Biosystems* 14:461–81.

Chihara, M., Inouye, I. & Takahata, N. 1986. *Oltmannsiellopsis*, a new genus of marine flagellate (Dunaliellaceae, Chlorophyceae). *Arch. Protistenkd.* 132:313–24.

Duminil, J., Grivet, D., Ollier, S., Jeandroz, S. & Petit, R. J. 2008. Multilevel control of organelle DNA sequence length in plants. *J. Mol. Evol.* 66:405–15.

Famà, P., Wysor, B., Kooistra, W. H. C. F. & Zuccarello, G. C. 2002. Molecular phylogeny of the genus *Caulerpa* (Caulerpales, Chlorophyta) inferred from chloroplast *tuf*A gene. *J. Phycol.* 38:1040–50.

Floyd, G. L. & O'Kelly, C. J. 1990. Phylum Chlorophyta: class Ulvophyceae. *In* Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J. [Eds.] *Handbook of Protoctista.* Jones and Bartlett Publishers, Boston, pp. 17–635.

Friedl, T. 1995. Inferring taxonomic positions and testing genus level assignments in coccoid green lichen algae: a phylogenetic analysis of 18S ribosomal RNA sequences from *Dictyochloropsis reticulata* and from members of the genus *Myrmecia* (Chlorophyta, Trebouxiophyceae cl. nov.). *J. Phycol.* 31:632–9.

Friedl, T. & O'Kelly, C. J. 2002. Phylogenetic relationships of green algae assigned to the genus *Planophila* (Chlorophyta): evidence from 18S rDNA sequence data and ultrastructure. *Eur. J. Phycol.* 37:373–84.

Guillou, L., Eikrem, W., Chretiennot-Dinet, M. J., Le Gall, F., Massana, R., Romari, K., Pedros-Alio, C. & Vaulot, D. 2004. Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* 155:193–214.

Guindon, S. & Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696–704.

Hargraves, P. E. & Steele, R. L. 1980. Morphology and ecology of *Oltmannsiella virida*, sp. nov. (Chlorophyceae: Volvocales). *Phycologia* 19:96–102.

Heath, T. A., Hedtke, S. M. & Hillis, D. M. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *J. Syst. Ecol.* 46:239–57.

Hillis, D. M., Pollock, D. D., McGuire, J. A. & Zwickl, D. J. 2003. Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* 52:124–6.

van den Hoek, C., Mann, D. G. & Jahns, H. M. 1995. *Algae. An Introduction to Phycology.* Cambridge University Press, Cambridge, UK, 623 pp.

Jobb, G., von Haeseler, A. & Strimmer, K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* 4:18.

Koch, M. A., Dobes, C., Kiefer, C., Schmickl, R., Klimes, L. & Lysak, M. A. 2007. Supernetwork identifies multiple events of plastid trnF(GAA) pseudogene evolution in the Brassicaceae. *Mol. Biol. Evol.* 24:63–73.

Krienitz, L., Hegewald, E., Hepperle, D. & Wolf, M. 2003. The systematics of coccoid green algae: 18S rRNA gene sequence data versus morphology. *Biologia* 58:437–46.

Lehman, R. L. & Manhart, J. R. 1997. A preliminary comparison of restriction fragment patterns in the genus *Caulerpa* (Chlorophyta) and the unique structure of the chloroplast genome in *Caulerpa sertularioides*. *J. Phycol.* 33:1055–62.

Lemieux, C., Otis, C. & Turmel, M. 2007. A clade uniting the green algae *Mesostigma viride* and *Chlorokybus atmophyticus* represents the deepest branch of the Streptophyta in chloroplast genome-based phylogenies. *BMC Biol.* 5:2.

Lewis, L. A. & McCourt, R. M. 2004. Green algae and the origin of land plants. *Am. J. Bot.* 91:1535–56.

Lokhorst, G. M., Star, W. & Zuccarello, G. C. 2004. New genus *Koliellopsis* (Trebouxiophyceae, Chlorophyta): its phylogenetic position inferred from ultrastructure and nuclear ribosomal DNA sequences. *Phycol. Res.* 52:235–43.

Lopez-Bautista, J. M. & Chapman, R. L. 2003. Phylogenetic affinities of the Trentepohliales inferred from small-subunit rDNA. *Int. J. Syst. Evol. Microbiol.* 53:2099–106.

Lowe, T. M. & Eddy, S. R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–64.

Manhart, J. R., Kelly, K., Dudock, B. S. & Palmer, J. D. 1989. Unusual characteristics of *Codium fragile* chloroplast revealed by physical and gene mapping. *Mol. Gen. Genet.* 216:417–21.

Mattox, K. R. & Stewart, K. D. 1984. Classification of the green algae: a concept based on comparative cytology. *In* Irvine, D. E. G. & John, D. M. [Eds.] *The Systematics of the Green Algae.* Academic Press, London, pp. 29–72.

Nakayama, T., Suda, S., Kawachi, M. & Inouye, I. 2007. Phylogeny and ultrastructure of *Nephroselmis* and *Pseudoscourfieldia* (Chlorophyta), including the description of *Nephroselmis anterostigmatica sp. nov.* and a proposal for the *Nephroselmidales ord. nov. Phycologia* 46:680–97.

O'Kelly, C. J. & Floyd, G. L. 1984. Flagellar apparatus absolute orientations and the phylogeny of the green algae. *Biosystems* 16:227–51.

O'Kelly, C. J., Wysor, B. & Bellows, W. K. 2004. Gene sequence diversity and the phylogenetic position of algae assigned to the genera *Phaeophila* and *Ochlochaete* (Ulvophyceae, Chlorophyta). *J. Phycol.* 40:789–99.

Palmer, J. D. 1986. Isolation and structural analysis of chloroplast DNA. *Methods Enzymol.* 118:167–86.

Pombert, J. F., Lemieux, C. & Turmel, M. 2006. The complete chloroplast DNA sequence of the green alga *Oltmannsiellopsis viridis* reveals a distinctive quadripartite architecture in the chloroplast genome of early diverging ulvophytes. *BMC Biol.* 4:3.

Pombert, J. F., Otis, C., Lemieux, C. & Turmel, M. 2004. The complete mitochondrial DNA sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) highlights distinctive evolutionary trends in the Chlorophyta and suggests a sister-group relationship between the Ulvophyceae and Chlorophyceae. *Mol. Biol. Evol.* 21:922–35.

Pombert, J. F., Otis, C., Lemieux, C. & Turmel, M. 2005. Chloroplast genome sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) reveals unusual structural features and new insights into the branching order of chlorophyte lineages. *Mol. Biol. Evol.* 22:1903–18.

Pröschold, T. & Leliaert, F. 2007. Systematics of the green algae: conflict of classic and modern approaches. *In* Brodie, J. & Lewis, J. [Eds.] *Unravelling the Algae: The Past, Present and Future of the Algae Systematics.* CRC Press, Taylor & Francis, Boca Raton, Florida, pp. 123–53.

Provan, J., Murphy, S. & Maggs, C. A. 2004. Universal plastid primers for Chlorophyta and Rhodophyta. *Eur. J. Phycol.* 39:43–50.

Rambaut, A. 1996. *Se-Al: Sequence Alignment Editor.* Available at http://evolve.zoo.ox.ac.uk/ (accessed 9 August 2002).

Robbens, S., Derelle, E., Ferraz, C., Wuyts, J., Moreau, H. & Van de Peer, Y. 2007. The complete chloroplast and mitochondrial DNA sequence of *Ostreococcus tauri*: organelle genomes of the smallest eukaryote are examples of compaction. *Mol. Biol. Evol.* 24:956–68.

Ronquist, F. & Huelsenbeck, J. P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–4.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51:492–508.

Shimodaira, H. & Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114–6.

Shimodaira, H. & Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–7.

Sluiman, H. J. 1989. The green algal class Ulvophyceae: an ultrastructural survey and classification. *Cryptogam. Bot.* 1:83–94.

Stam, W. T., Olsen, J. L., Zaleski, S. F., Murray, S. N., Brown, K. R. & Walters, L. J. 2006. A forensic and phylogenetic survey of *Caulerpa species* (Caulerpales, Chlorophyta) from the Florida coast, local aquarium shops, and e-commerce: establishing a proactive baseline for early detection. *J. Phycol.* 42:1113–24.

Steane, D. A. 2005. Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus* (Myrtaceae). *DNA Res.* 12:215–20.

Steinkötter, J., Bhattacharya, D., Semmelroth, I., Bibeau, C. & Melkonian, M. 1994. Prasionophytes form independent lineages within the Chlorophyta – evidence from ribosomal RNA sequence comparison. *J. Phycol.* 30:340–5.

Tsuji, S., Ueda, K., Nishiyama, T., Hasebe, M., Yoshikawa, S., Konagaya, A., Nishiuchi, T. & Yamaguchi, K. 2007. The chloroplast genome from a lycophyte (microphyllophyte), *Selaginella uncinata*, has a unique inversion, transpositions and many gene losses. *J. Plant. Res.* 120:281–90.

Turmel, M., Brouard, J. S., Gagnon, C., Otis, C. & Lemieux, C. 2008. Deep division in the Chlorophyceae (Chlorophyta) revealed by chloroplast phylogenomic analyses. *J. Phycol.* 44:739–50.

Turmel, M., Otis, C. & Lemieux, C. 1999. The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. *Proc. Natl. Acad. Sci. U. S. A.* 96:10248–53.

Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., et al. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proc. Natl. Acad. Sci. U. S. A.* 94:5967–72.

Watanabe, S. & Nakayama, T. 2007. Ultrastructure and phylogenetic relationships of the unicellular green algae *Ignatius tetrasporus* and *Pseudocharacium americanum* (Chlorophyta). *Phycol. Res.* 55:1–16.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* 24:1586–91.

Zechman, F. W., Theriot, E. C., Zimmer, E. A. & Chapman, R. L. 1990. Phylogeny of the Ulvophyceae (Chlorophyta): cladistic analysis of nuclear-encoded rRNA sequence data. *J. Phycol.* 26:700–10.

Zuccarello, G. C. & Lokhorst, G. M. 2005. Molecular phylogeny of the genus *Tribonema* (Xanthophyceae) using *rbc*L gene sequence data: monophyly of morphologically simple algal species. *Phycologia* 44:384–92.

Supplementary Material

The following supplementary material is available for this article:

Table S1. Species used for chloroplast genome analysis and GenBank accession numbers.

Appendix S1. Chloroplast isolation method.

Appendix S2. DNA extraction method.

Appendix S3. DNA cloning method.

Appendix S4. Concatenated DNA alignment of 23 chloroplast genes and 16,824 aligned nucleotide positions.

This material is available as part of the online article.