J. Eukaryot. Microbiol., 56(4), 2009 pp. 367–372

DOI: 10.1111/j.1550-7408.2009.00411.x

The Distribution of Elongation Factor-1 Alpha (EF-1α), Elongation Factor-Like (EFL), and a Non-Canonical Genetic Code in the Ulvophyceae: Discrete Genetic Characters Support a Consistent Phylogenetic Framework

GILLIAN H. GILE,[a] PHILIP M. NOVIS,[b] DAVID S. CRAGG,[a] GIUSEPPE C. ZUCCARELLO[c] and PATRICK J. KEELING[a]

[a]*Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4, and*
[b]*Allan Herbarium, Manaaki Whenua – Landcare Research, P.O. Box 40, Lincoln 7640, New Zealand, and*
[c]*School of Biological Sciences, Victoria University of Wellington, P.O. Box 600, Wellington 6140, New Zealand*

ABSTRACT. The systematics of the green algal class Ulvophyceae have been difficult to resolve with ultrastructural and molecular phylogenetic analyses. Therefore, we investigated relationships among ulvophycean orders by determining the distribution of two discrete genetic characters previously identified only in the order Dasycladales. First, *Acetabularia acetabulum* uses the core translation GTPase Elongation Factor 1α (EF-1α) while most Chlorophyta instead possess the related GTPase Elongation Factor-Like (EFL). Second, the nuclear genomes of dasycladaleans *A. acetabulum* and *Batophora oerstedii* use a rare non-canonical genetic code in which the canonical termination codons TAA and TAG instead encode glutamine. Representatives of Ulvales and Ulotrichales were found to encode EFL, while Caulerpales, Dasycladales, Siphonocladales, and *Ignatius tetrasporus* were found to encode EF-1α, in congruence with the two major lineages previously proposed for the Ulvophyceae. The EF-1α of *I. tetrasporus* supports its relationship with Caulerpales/Dasycladales/Siphonocladales, in agreement with ultrastructural evidence, but contrary to certain small subunit rRNA analyses that place it with Ulvales/Ulotrichales. The same non-canonical genetic code previously described in *A. acetabulum* was observed in EF-1α sequences from *Parvocaulis pusillus* (Dasycladales), *Chaetomorpha coliformis*, and *Cladophora cf. crinalis* (Siphonocladales), whereas Caulerpales use the universal code. This supports a sister relationship between Siphonocladales and Dasycladales and further refines our understanding of ulvophycean phylogeny.

Key Words. Caulerpales, Chlorophyta, Dasycladales, elongation factors, green algae, *Ignatius tetrasporus*, systematics, taxonomy, Ulotrichales, Ulvales.

Mattox and Stewart (1984) defined the green algal class Ulvophyceae mainly on the basis of the counterclockwise offset of the cruciate flagellar basal apparatus and the mode of cytokinesis, which involves neither a phycoplast (precluding placement in the Chlorophyceae) nor a phragmoplast (a Charophycean feature). Based on this definition, O'Kelly and Floyd (1984) included five orders in the Ulvophyceae: Ulvales, Ulotrichales, Siphonocladales, Dasycladales, and Caulerpales; the Trentepohliales were omitted on the basis of anomalous features reminiscent of the Charophyceae, such as a multilayered structure in the flagellar root system and plasmodesmata between vegetative cells (O'Kelly and Floyd 1984). While some molecular phylogenetic analyses of 18S rRNA weakly recover ulvophycean monophyly (López-Bautista and Chapman 2003; Watanabe and Nakayama 2007), those with all orders represented suggest two distinct, non-sister lineages within a clade that also includes members of the Chlorophyceae and Trebouxiophyceae (Watanabe, Kuroda, and Maiwa 2001; Zechman et al. 1990). These lineages have been referred to as the Ulvophyceae I, which includes the orders Siphonocladales, Dasycladales, Caulerpales, and Trentepohliales, and the Ulvophyceae II, which includes the orders Ulvales and Ulotrichales (Watanabe et al. 2001). Overall, it seems likely that the Ulvophyceae is not monophyletic, unless more narrowly described to include just the Ulvophyceae II clade (Ulvophyceae *sensu* van den Hoek, Mann, and Jahns 1995; Watanabe et al. 2001), but a new taxonomic revision awaits further evidence.

To further refine our understanding of relationships among ulvophyceans, we have examined taxa from the five orders identified by O'Kelly and Floyd (1984) plus *Ignatius tetrasporus* for the presence of two discrete genetic characters: the presence of Elongation Factor 1α (EF-1α) versus Elongation Factor-Like (EFL) proteins, and the presence of a non-canonical genetic code where TAA and TAG encode glutamine. The eukaryotic EF-1α (also known as EF1A) plays an essential role in translation by bringing aminoacyl-tRNAs to the ribosome, and was thought to be ubiquitous. However, it was recently discovered that certain eukaryotic groups lack EF-1α altogether and instead possess a distinct paralog called EFL (Keeling and Inagaki 2004). Within the Chlorophyta, all investigated species possess EFL, with one intriguing exception: the ulvophycean *Acetabularia acetabulum* possesses EF-1α. The relationships between green algal EFL genes are not well resolved, but the EF-1α gene from *A. acetabulum* is clearly related to EF-1α from charophytes and land plants, suggesting that at least EF-1α was present in the ancestor of Viridiplantae (Noble, Rogers, and Keeling 2007). The Ulvophyceae are therefore at the center of the puzzle of EF-1α/EFL evolution in the green algae, but EFL and EF-1α have been characterized from only one ulvophycean order each: EFL in the Ulvales from two *Ulva* species (Noble et al. 2007), and EF-1α in the Dasycladales from *A. acetabulum* (Keeling and Inagaki 2004).

Another unusual molecular character in the Dasycladales is a non-canonical genetic code. This was first discovered in *A. acetabulum* (Schneider, Leible, and Yang 1989) and later in *Batophora oerstedii* (Schneider and de Groot 1991). In these genomes the canonical stop codons UAA and UAG specify glutamine. The same non-canonical code also occurs in the nuclear genomes of oxymonads (Keeling and Leander 2003; de Koning et al. 2008) and diplomonads (Keeling and Doolittle 1996, 1997), and it has likely arisen at least twice in ciliates (Baroin-Tourancheau et al. 1995; Lozupone, Knight, and Landweber 2001). In this study, we find that both characters are more broadly and informatively distributed than was previously recognized. Elongation factor 1α was found in Dasycladales, Siphonocladales, Caulerpales, and *I. tetrasporus*, whereas EFL was found in Ulvales and Ulotrichales. The non-canonical genetic code, in turn, was found in Dasycladales and Siphonocladales, but not in Caulerpales (or Ulvales or Ulotrichales). Together these characters support previous suggestions of a clade of Dasycladales, Siphonocladales, Caulerpales, and *I. tetrasporus* (Watanabe and Nakayama 2007; Watanabe et al. 2001) and a specific relationship between the Siphonocladales and Dasycladales within this clade, which is also consistent

Corresponding Author: P. Keeling, Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC Canada V6T 1Z4—Telephone number: 604 822 4906; FAX number: 604 822 6089; e-mail: pkeeling@interchange.ubc.ca

with ultrastructural characters (O'Kelly and Floyd 1984; Roberts, Stewart, and Mattox 1984; Sluiman 1989).

MATERIALS AND METHODS

Culture sources. Seven strains of ulvophycean green algae from five orders were used in this study. Four were ordered from either the Provasoli-Guillard National Center for Culture of Marine Phytoplankton (CCMP) or the Culture Collection of Algae at the University of Texas at Austin (UTEX): *Ochlochaete hystrix* Thwaites ex Harvey CCMP 2319, *Urospora* sp. CCMP 1082, *I. tetrasporus* Bold and MacEntee (UTEX B 2012), and *Parvocaulis pusillus* (Howe) Berger, Fettweiss, Gleissberg, Liddle, Richter, Sawitsky, and Zuccarello (UTEX LB 2710). *Caulerpa* cf. *racemosa* (Forsskål) Agardh was donated by the Vancouver Aquarium. *Chaetomorpha coliformis* (Montagne) Kuetzing was collected at Taylor's mistake, New Zealand, December 18, 2007, and *Cladophora cf. crinalis* Harvey was collected at Wainui, Akaroa Harbour, New Zealand, October 26, 2007. Voucher specimens for the latter two collections were deposited at the Allan Herbarium, Lincoln, New Zealand, numbers CHR585485 and CHR585488.

RNA extraction, polymerase chain reaction, and sequencing methods. Total RNA was extracted from algal cell pellets of *Ochlochaete*, *Urospora*, *Ignatius*, *Parvocaulis*, and *Caulerpa* using Trizol reagent (Invitrogen, Carlsbad, CA) or the RNeasy Mini kit (Qiagen, Mississauga, ON, Canada) for *Chaetomorpha* and *Cladophora*. Elongation Factor-Like and EF-1α cDNA sequences were obtained by 3′ RACE using the following nested degenerate forward primers designed to be universal for any eukaryotic EFL and EF-1α: 5′-GTCGARATGCAYCAY-3′ (outer) and 5′-CCG GGCGAYAAYGTNGG-3′ (inner) using the FirstChoice RLM-RACE kit (Ambion, Austin, TX). Gene-specific reverse primers were designed from EFL and EF-1α 3′ RACE sequences and used with the following nested degenerate forward primers using Superscript III One-Step RTPCR with Platinum Taq (Invitrogen): for EFL—5′- CTGTCGATCGTCATHTGYGGN-3′ and 5′-CATGT CGACTCGGGCAAGTCNACNACNACNGG-3′, and 5′-AACA TCGTCGTGATHGGNCAYGTNGA-3′; for EF-1α—5′- CATG TCGACTCGGGCAAGTCNACNACNACNGG-3′ and 5′-TTCG AGAAGGAGGCNGCNGARATGAA-3′. Polymerase chain reaction products of *Ochlochaete*, *Urospora*, *Ignatius*, *Parvocaulis*, and *Caulerpa* were cloned using the TOPO-TA cloning kit (Invitrogen) and for *Chaetomorpha* and *Cladophora* using the pGEM-T Easy vector system (Promega, Madison, WI). New sequences obtained in both directions using BigDye Terminator v. 3.1 (Applied Biosystems, Foster City, CA) were deposited in GenBank under Accession numbers FJ539138-FJ539144.

Phylogenetic analyses. New and previously published EFL and EF-1α sequences were translated and aligned by Multiple Alignment using Fast Fourier Transform (MAFFT, Katoh et al. 2002) and edited in MacClade 4.08 (Maddison and Maddison 2003) to final matrix sizes of 32 taxa and 444 unambiguously aligned characters for EFL and 29 taxa and 426 characters for EF-1α. Phylogenetic trees were inferred using maximum likelihood (ML) and Bayesian methods. ProtTest 1.4 (Abascal, Zardoya, and Posada 2005) was used to determine the best amino acids replacement models and analysis parameters. Maximum likelihood trees were inferred using RAxML 7.0.3 (Stamatakis 2006) on the CIPRES portal (http://8ball.sdsc.edu:8889/cipres-web/Home.do) using the RtREV amino acids substitution matrix (Dimmic et al. 2002), four rate categories approximated by a Γ distribution with parameter α estimated from the data, amino acid frequencies calculated from the data, and in the case of EF-1α, a proportion of invariable sites also calculated from the data. The EFL alignment had a negligibly low proportion of invariable sites. Two hundred and fifty bootstrap replicates were performed for both datasets, as computed to be sufficient by RAxML (Stamatakis, Hoover, and Rougemont 2008). MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003) was used to perform Bayesian analyses using the RtREV amino acids substitution matrix and the same parameters as the likelihood analyses. Five independent analyses for each of the EFL and EF-1α datasets were carried out in order to test for convergence, and all five analyses for each protein produced identical topologies and near-identical posterior probabilities. One cold and three heated chains were run for all analyses, sampling one tree per thousand generations, and 50% majority rule consensus trees were computed after observing that log likelihood values stabilized at 5,000 generations and discarding the first five sampled trees. Consensus trees were also computed after discarding the first 100 sampled trees for the 5,000 generation runs and first 1,000 sampled trees for the five million generations runs, with no effect on the consensus topology or posterior probabilities (data not shown). All five EFL analyses were run for five million generations, and one consensus tree was arbitrarily chosen to indicate posterior probabilities on the ML topology. Four of the EF-1α analyses were run for one million generations and one for five million generations; the longer run was chosen to represent the topology and posterior probabilities. For the EF-1α tree, ML branch lengths were computed using TREE-PUZZLE 5.1 (Schmidt et al. 2002) and displayed on the Bayesian topology.

Approximately Unbiased (AU) tests (Shimodaira 2002) were carried out using CONSEL 1.19 (Shimodaira and Hasegawa 2001) to evaluate the likelihood of alternate EF-1α and EFL topologies in which the Ulvophyceae is constrained as monophyletic. Site likelihoods were calculated by TREE-PUZZLE 5.1 (Schmidt et al. 2002) using the –wsl option, the WAG amino acids substitution matrix (Whelan and Goldman 2001), and four Γ rate categories with parameter α, amino acid frequencies, and the proportion of invariable sites estimated from the data. Because the RtREV model is not available in this program, ML trees were also inferred with the WAG substitution model (Whelan and Goldman 2001), the second best model according to ProtTest. The resulting topologies were congruent with only minor, unsupported differences: *C. racemosa* and *I. tetrasporus* form an unsupported clade (24%) in the EF-1α tree, and the EFL tree was identical to the Bayesian topology (see ''Results'' for details).

RESULTS

Distribution of EFL and EF-1α and a non-canonical genetic code. In this study, seven species representing five ulvophycean orders were tested for the presence of EF-1α and EFL by RT-PCR. EFL sequences were determined from *O. hystrix* (Ulvales) and *Urospora* sp. CCMP1082 (Ulotrichales), both members of the Ulvophyceae II clade. EF-1α sequences were determined from all other species tested, all of which belong to the Ulvophyceae I: *Caulerpa cf. racemosa* (Caulerpales), *P. pusillus* (Dasycladales), *C. cf. crinalis* and *C. coliformis* (Siphonocladales). None of the species investigated was found to express both genes. EF-1α sequences for *C. crinalis* and *C. coliformis* in the Siphonocladales and *P. pusillus* from the Dasycladales were found to use a non-canonical genetic code. In all three genes, a UAA or UAG codon was found at one or more positions that are otherwise highly conserved for the amino acid glutamine (Fig. 1). The *C. racemosa* EF-1α sequence contained neither UAA nor UAG codons.

Phylogeny of EF-1α and EFL in the Ulvophyceae. Phylogenetic analyses of EFL and EF-1α were carried out in order to gain insight into the evolutionary history of these proteins, but they should not be interpreted as reflective of ulvophycean relationships as neither of these proteins is well-suited for inferring

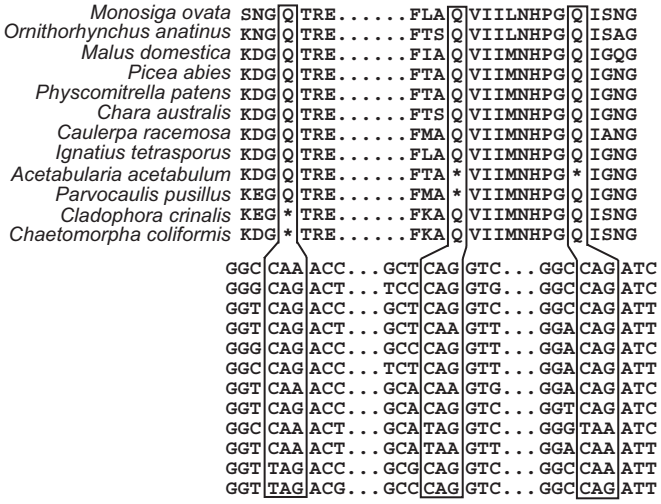

Fig. **1.** Selected regions of the elongation factor 1α (EF-1α) alignment indicating TAA and TAG codons at positions conserved for glutamine (Q) in the dasycladaleans *Parvocaulis pusillus* and *Acetabularia acetabulum* and the siphonocladaleans *Cladophora crinalis* and *Chaetomorpha coliformis*.

organismal phylogenies (Keeling and Inagaki 2004; Roger et al. 1999). In the EF-1α phylogeny, the branch uniting streptophytes is strongly supported (100% bootstrap support, 1.0 posterior probability), as is the branch uniting streptophyte and ulvophycean EF-1α sequences (97% bootstrap support, 1.0 posterior probability, Fig. 2). The ulvophycean sequences as a whole do not group together (only species from the same order form supported clades), but AU tests failed to reject the possibility of a monophyletic Ulvophyceae with ($p = 0.45$) or without ($p = 0.161$) Siphonocladales and Dasycladales constrained as sister groups (Fig. 3A, 3B). Therefore, while the best ML topology of the EF-1α tree does not recover accepted relationships among members of the Ulvophyceae and Streptophyta, it is not inconsistent with their monophyly (Fig. 2), suggesting that it was likely encoded in the common ancestor of these two groups.

Our Bayesian EFL analysis (Fig. 4) is congruent with previous analyses, but the new ulvophycean sequences are not monophyletic: *Urospora* and the *Ulva* species form a strongly supported clade, but *O. hystrix*, which is more closely related to *Ulva* than either are to *Urospora* (O'Kelly, Wysor, and Bellows 2004), is excluded. *Chlorococcum* sp. NEPCC 478 also groups robustly with the *Ulva* clade, but as the support is high and many other coccoid green algae have been transferred out of their traditional morphology-based genera, even into new classes (Lewis and McCourt 2004), we suspect that this *Chlorococcum* strain is more

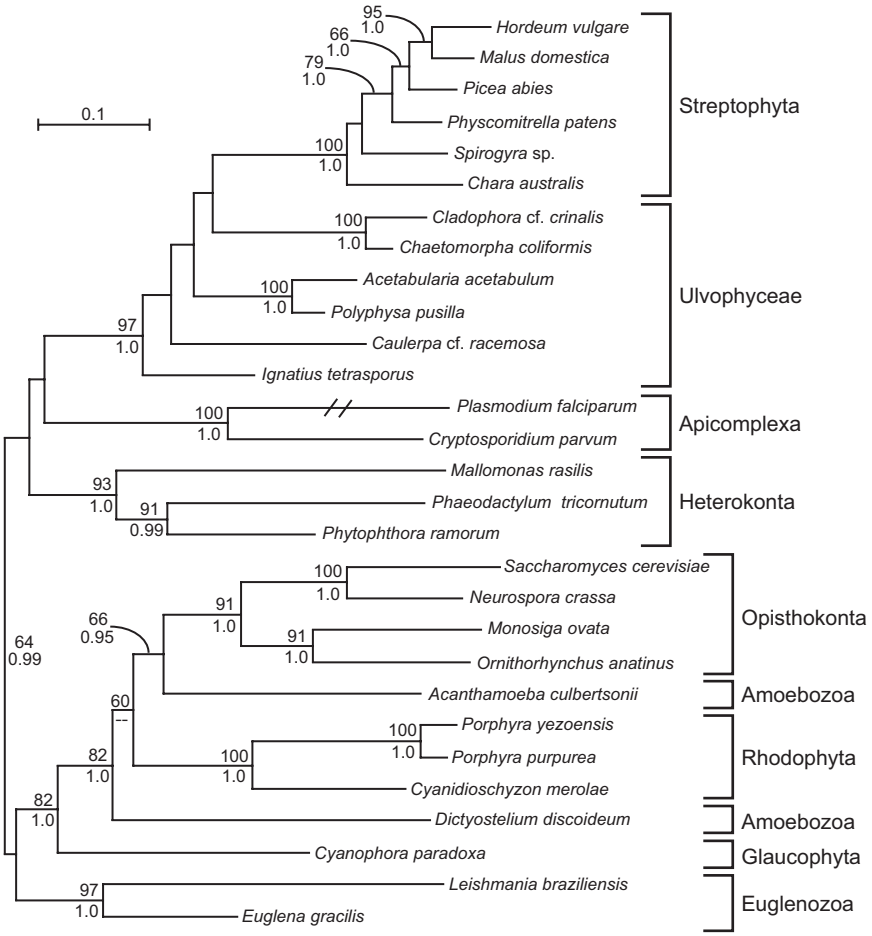

Fig. **2.** Maximum likelihood (ML) phylogeny of elongation factor 1α (EF-1α) with major lineages bracketed to the right. Hash marks indicate branches whose lengths have been reduced by precisely one half, while ML bootstrap values of 50% or greater (above) and Bayesian posterior probabilities of 0.9 or greater (below) are indicated at nodes.

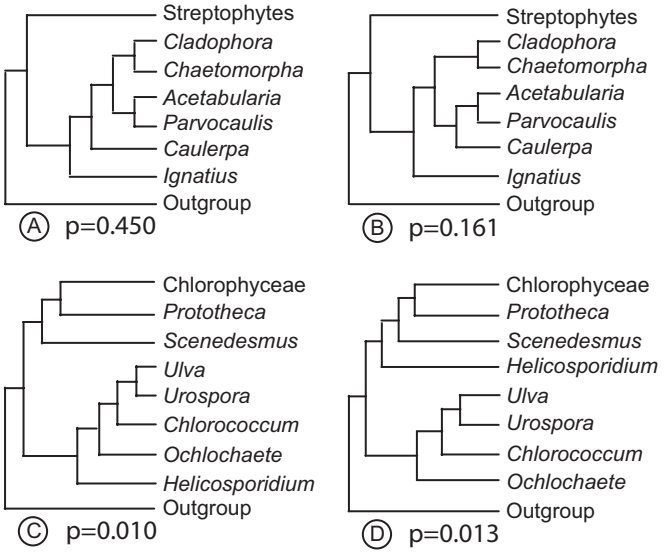

Fig. **3.** Alternate topologies evaluated by Approximately Unbiased (AU) tests and their associated *P*-values. **A, B** elongation factor 1α (EF-1α). **C, D** Elongation Factor-Like (EFL). Siphonocladales are represented by *Cladophora* and *Chaetomorpha* and Dasycladales by *Acetabularia* and *Parvocaulis*.

likely a misidentified member of the Ulvales or Ulotrichales than a true *Chlorococcum* species. As a result of the unexpected placement of *O. hystrix*, the phylogeny of EFL is both consistent with a single origin of EFL in the UTC clade and inconsistent with a single origin of EFL in the Ulvophyceae. Approximately Unbiased tests further supported the non-monophyly of ulvophycean EFL (Fig. 3C, D). The ML topology (not shown) was essentially identical to the Bayesian topology, having only unsupported differences. The position of *O. hystrix* is the same but lacks support (44%), *Helicosporidium* falls at the base of the *Ulva* clade with no support (6%), and *T. tetrathele* groups with *Raphidiophrys contractilis* with no support (8%).

DISCUSSION

Previously, the distribution of EFL and EF-1α in the Ulvophyceae was only known from *Ulva fenestrata*, *Ulva intestinalis*, and *A. acetabulum*. We have used seven species from five orders of the Ulvophyceae to determine the presence of EFL and EF-1α by 3′ RACE and RT-PCR, and their respective distributions were found to correspond to the two major groups within the class. Elongation Factor-Like was found in the Ulvales from *O. hystrix* and previous *Ulva* sequences and in the Ulotrichales from *Urospora* sp. CCMP 1082, the two orders that make up the Ulvophyceae II group. Elongation factor 1α, on the other hand, was found in the Caulerpales from *Caulerpa* cf. *racemosa*, Dasycladales from *P. pusillus*

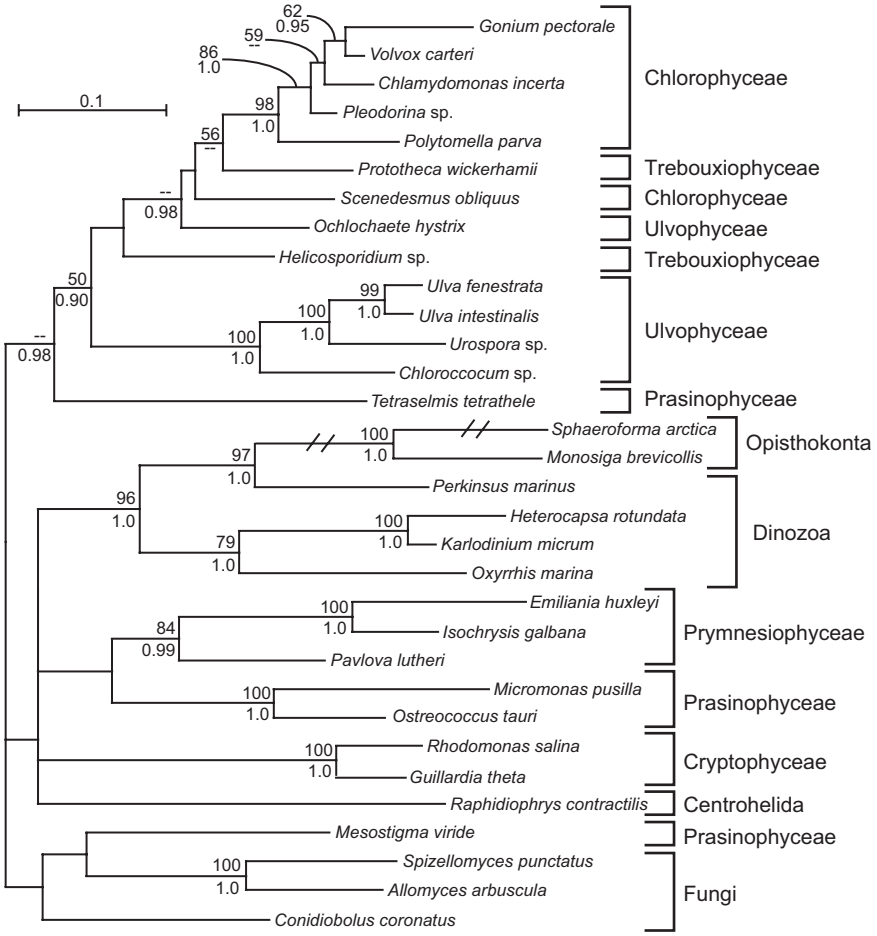

Fig. **4.** Bayesian phylogeny of Elongation Factor-Like (EFL) with maximum likelihood (ML) branch lengths. Major lineages are bracketed to the right. Hash marks indicate branches whose lengths have been reduced by precisely one half, while ML bootstrap values of 50% or greater (above) and Bayesian posterior probabilities of 0.9 or greater (below) are indicated at nodes.

and previously from *A. acetabulum*, and Siphonocladales from *C.* cf. *crinalis* and *C. coliformis*, all of which are members of the Ulvophyceae I group. *Ignatius tetrasporus* was also found to encode EF-1α, supporting its inclusion in Ulvophyceae I. None of the selected taxa was found to express both EFL and EF-1α. During the course of our research, other members of the Ulvophyceae I (i.e. *Ostreobium quekettii*, *Blastophysa rhizopus*, and *Codium*, *Derbesia*, and *Bryopsis* species) were independently found to encode EF-1α by examining gene fragments by PCR (Cocquyt, E. E., pers. commun.).

In the EF-1α sequences for *C. crinalis* and *C. coliformis* in the Siphonocladales and *P. pusillus* from the Dasycladales, the codons UAA or UAG were found at one or more positions that are otherwise highly conserved for the amino acid glutamine (Fig. 1). We interpret this as being due to the use of a non-canonical code because the same code has been identified in *A. acetabulum* and confirmed by protein sequencing (Schneider et al. 1989) and also identified in the closely related dasycladalean *B. oerstedii* (Schneider and de Groot 1991). It is therefore no surprise to find this code in *P. pusillus*, which is even more closely related to *A. acetabulum* than *B. oerstedi* is (Zechman 2003), but its use by two members of the Siphonocladales is informative. The EF-1α sequences from *C. racemosa* and *I. tetrasporus* use only canonical CAA and CAG codons for glutamine, and the coding sequences terminate with UGA codons, characteristics that are consistent with either genetic code. However, publicly available sequences from other Caulerpales (i.e. *Bryopsis hypnoides* (lectin, GenBank EU410470), *Bryopsis maxima* (RNase Bm2, AB164318), *Bryopsis plumosa* (bryohealin, EU769118), and *Flabellia petiolata* (P-type ATPase, AJ972675), all use canonical glutamine codons, and coding sequences terminate with UGA, UAA, or UAG codons, indicating that the Caulerpales as a whole use the universal code. Additional sequences from *Caulerpa* and other genera support this conclusion (unpubl. data, and Cocquyt, E. E., pers. commun.).

The discovery of a non-canonical genetic code in the Siphonocladales expands the known distribution of this character within the Ulvophyceae I, and combined with ultrastructural evidence, supports a sister relationship between the Siphonocladales and Dasycladales. Genetic code changes in nuclear genomes are quite rare, having occurred in only a handful of eukaryotic lineages. Although the conversion of UAA and UAG codons from specifying stop to specifying glutamine has happened twice within the ciliates (Baroin-Tourancheau et al. 1995; Lozupone et al. 2001), this is unlikely to be the case in the Ulvophyceae for two main, interrelated reasons. First, Siphonocladales and Dasycladales undoubtedly share a recent common ancestor; the only question is whether they are sisters or whether one of them is closer to the Caulerpales (Roberts et al. 1984; Sluiman 1989; Watanabe and Nakayama 2007; Watanabe et al. 2001; Zechman et al. 1990). The two lineages of ciliates known to share this code are far more distantly related. Second, sisterhood of Siphonocladales and Dasycladales has been proposed previously on the basis of shared ultrastructural characters. These include a somewhat flattened cruciate arrangement of basal bodies and roots, a striated distal fiber connecting the two distal basal bodies, and a transverse septum in the flagellar transition zone (Roberts et al. 1984; Sluiman 1989). Molecular analyses neither support nor refute this hypothesis: they have either failed to resolve the relationships among these three orders (Zechman et al. 1990) or failed to include sequences from the Caulerpales (López-Bautista and Chapman 2003; Watanabe et al. 2001). Because the evidence from molecular phylogeny, morphology, and now the shared retention of EF-1α all support a monophyletic Ulvophyceae I, and Dasycladales and Siphonocladales share ultrastructural features, the most straightforward interpretation of the distribution of genetic codes is that Dasycladales and Siphonocladales share a common ancestor to the exclusion of Caulerpales.

The distribution of discrete genetic characters can be useful in inferring phylogenetic relationships, if they are interpreted correctly and ideally are consistent with other forms of data, especially when evidence from ultrastructural features and molecular phylogenies is in conflict or inconclusive. Previous molecular phylogenetic analyses provided conflicting placements of *I. tetrasporus*: a Bayesian analysis of 18S rRNA placed *I. tetrasporus* and its sister taxon *Pseudocharacium americanum* as early-diverging members of the Ulvophyceae II clade, while a distance analysis placed them at the base of the Ulvophyceae I, though neither placement was strongly supported (Watanabe and Nakayama 2007). The authors hypothesized that *I. tetrasporus* belongs with the Ulvophyceae I clade on the basis of their ultrastructural analysis, which is consistent with our findings. *Ignatius tetrasporus* is only one of several putative ulvophyceans with uncertain affinities, however, Trentepohliales are hypothesized to be sisters to the Dasycladales and/or Siphonocladales, but their exact placement is uncertain (López-Bautista and Chapman 2003). If this hypothesis is correct, we would predict that this group also possesses EF-1α, and their genetic code may be especially informative. The affinities of *Oltmannsiellopsis viridis* may also be clarified by determining which elongation factor it encodes. It has been shown to branch at the base of the Ulvophyceae with strong support (Friedl and O'Kelly 2002), but because members of the Ulvophyceae I clade were omitted, its precise position remains unclear. Finally, certain trebouxiophytes show a weak affinity to the Ulvophyceae I clade in small subunit rRNA trees (Watanabe et al. 2001). It would be of interest to determine whether these taxa also possess EF-1α and, by extension, whether they might be better placed in the Ulvophyceae I.

ACKNOWLEDGMENTS

This work was supported by a Discovery Grant (227301) from the Natural Sciences and Engineering Research Council of Canada (NSERC) awarded to P.J.K. G.H.G. is supported by an NSERC postgraduate doctoral scholarship and D.S.C. was supported by an NSERC undergraduate scholarship. P.J.K. is a Fellow of the Canadian Institute for Advanced Research and a Senior Investigator of the Michael Smith Foundation for Health Research. P.M.N. was supported by the Foundation for Research, Science and Technology, New Zealand (contract LCRX0701). N. Bolstridge (Landcare Research) provided technical support in New Zealand. We thank Wairewa and Ōnuku Rūnanga and Te Hapū o Ngāti Wheke (Râpaki) for access to collecting sites in Aotearoa/New Zealand and permission to undertake molecular work.

LITERATURE CITED

Abascal, F., Zardoya, R. & Posada, D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, **21**:2104–2105.

Baroin-Tourancheau, A., Tsao, N., Klobutcher, L. A., Pearlman, R. E. & Adoutte, A. 1995. Genetic code deviations in the ciliates: evidence for multiple and independent events. *EMBO J.*, **14**:3262–3267.

Dimmic, M. W., Rest, J. S., Mindell, D. P. & Goldstein, R. A. 2002. RtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. *J. Mol. Evol.*, **55**:65–73.

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**:373–384.

van den Hoek, C., Mann, D. & Jahns, H. 1995. Algae: An Introduction to Phycology. Cambridge University Press, Cambridge, UK. 623 p.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, **30**:3059–3066.

Keeling, P. J. & Doolittle, W. F. 1996. A non-canonical genetic code in an early diverging eukaryotic lineage. *EMBO J.*, **15**:2285–2290.

Keeling, P. J. & Doolittle, W. F. 1997. Widespread and ancient distribution of a non-canonical genetic code in diplomonads. *Mol. Biol. Evol.*, **14**:895–901.

Keeling, P. J. & Inagaki, Y. 2004. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1-alpha. *Proc. Natl. Acad. Sci. USA*, **101**:15380–15385.

Keeling, P. J. & Leander, B. S. 2003. Characterization of a non-canonical genetic code in the oxymonad *Streblomastix strix*. *J. Mol. Biol.*, **326**:1337–1349.

de Koning, A. P., Noble, G. P., Heiss, A. A., Wong, J. & Keeling, P. J. 2008. Environmental PCR survey to determine the distribution of a non-canonical genetic code in uncultivable oxymonads. *Environ. Microbiol.*, **10**:65–74.

Lewis, L. A. & McCourt, R. M. 2004. Green algae and the origin of land plants. *Am. J. Bot.*, **91**:1535–1556.

López-Bautista, J. M. & Chapman, R. L. 2003. Phylogenetic affinities of the Trentepohliales inferred from small-subunit rDNA. *Int. J. Syst. Evol. Microbiol.*, **53**:2099–2106.

Lozupone, C. A., Knight, R. D. & Landweber, L. F. 2001. The molecular basis of nuclear genetic code change in ciliates. *Curr. Biol.*, **11**:65–74.

Maddison, D. R. & Maddison, W. P. 2003. MacClade 4: Analysis of Phylogeny and Character Evolution [Computer Program]. Version 4.08. Sinauer Associates, Sunderland, MA.

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. (ed.), Systematics of the Green Algae. Academic Press, London. p. 29–72.

Noble, G. P., Rogers, M. B. & Keeling, P. J. 2007. Complex distribution of EFL and EF-1alpha proteins in the green algal lineage. *BMC Evol. Biol.*, **7**.

O'Kelly, C. J. & Floyd, G. L. 1984. Correlations among patterns of sporangial structure and development, life histories, and ultrastructural features in the Ulvophyceae. *In*: Irvine, D. E. G. & John, D. M. (ed.), Systematics of the Green Algae. Academic Press, London. p. 121–156.

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**:789–799.

Roberts, K. R., Stewart, K. D. & Mattox, K. R. 1984. Structure and absolute configuration of the flagellar apparatus in the isogametes of *Batophora* (Dasycladales, Chlorophyta). *J. Phycol.*, **20**:183–191.

Ronquist, F. & Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**:1572–1574.

Roger, A. J., Sandblom, O., Doolittle, W. F. & Philippe, H. 1999. An evaluation of elongation factor 1alpha as a phylogenetic marker for eukaryotes. *Mol. Biol. Evol.*, **16**:218–233.

Schmidt, H. A., Strimmer, K., Vingron, M. & von Haeseler, A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**:502–504.

Schneider, S. U. & de Groot, E. J. 1991. Sequences of two *rbcS* cDNA clones of *Batophora oerstedii*: structural and evolutionary considerations. *Curr. Genet.*, **20**:173–175.

Schneider, S. U., Leible, M. B. & Yang, X. P. 1989. Strong homology between the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase of two species of *Acetabularia* and the occurrence of unusual codon usage. *Mol. Gen. Genet.*, **218**:445–452.

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

Shimodaira, H. & Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, **17**:1246–1247.

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

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**:2688–2690.

Stamatakis, A., Hoover, P. & Rougemont, J. 2008. A rapid bootstrap algorithm for the RAxML web-servers. *Syst. Biol.*, **57**:758–771.

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.

Watanabe, S., Kuroda, N. & Maiwa, F. 2001. Phylogenetic status of *Helicodictyon planctonicum* and *Desmochloris halophila* gen. et comb. nov. and the definition of the class Ulvophyceae (Chlorophyta). *Int. J. Syst. Evol. Microbiol.*, **40**:421–434.

Whelan, S. & Goldman, N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, **18**:691–699.

Zechman, F. W. 2003. Phylogeny of the Dasycladales (Chlorophyta, Ulvophyceae) based on analyses of RUBISCO large subunit (*rbcL*) gene sequences. *J. Phycol.*, **39**:819–827.

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**:700–710.

Received: 08/15/08, 12/15/08, 02/24/09; accepted: 02/25/09