Research Article

Algae 2016, 31(2): 155-165
http://dx.doi.org/10.4490/algae.2016.31.4.14
Open Access

*tuf*A gene as molecular marker for freshwater Chlorophyceae

Helena Henriques Vieira[1,*], Inessa Lacativa Bagatini[1], Carla Marques Guinart[2] and Armando Augusto Henriques Vieira[1]

[1]*Ecology and Natural Resources Post-graduation Program (PPGERN), Laboratory of Phycology, Botany Department, Federal University of São Carlos, São Carlos, SP 13565-905, Brazil*
[2]*Molecular Laboratory of Biodiversity and Conservation, Genetic and Evolution Department, Federal University of São Carlos, São Carlos, SP 13565-905, Brazil*

Green microalgae from the class Chlorophyceae represent a major biodiversity component of eukaryotic algae in continental water. Identification and classification of this group through morphology is a hard task, since it may present cryptic species and phenotypic plasticity. Despite the increasing use of molecular methods for identification of microorganisms, no single standard barcode marker is yet established for this important group of green microalgae. Some available studies present results with a limited number of chlorophycean genera or using markers that require many different primers for different groups within the class. Thus, we aimed to find a single marker easily amplified and with wide coverage within Chlorophyceae using only one pair of primers. Here, we tested the universality of primers for different genes (*tuf*A, ITS, *rbc*L, and UCP4) in 22 strains, comprising 18 different species from different orders of Chlorophyceae. The ITS primers sequenced only 3 strains and the UCP primer failed to amplify any strain. We tested two pairs of primers for *rbc*L and the best pair provided sequences for 10 strains whereas the second one provided sequences for only 7 strains. The pair of primers for the *tuf*A gene presented good results for Chlorophyceae, successfully sequencing 21 strains and recovering the expected phylogeny relationships within the class. Thus, the *tuf*A marker stands out as a good choice to be used as molecular marker for the class.

Key Words: ITS; molecular marker; phylogeny; *rbc*L; *tuf*A; UCP4

Abbreviations: BLAST, Basic Local alignment Search Tool; CBOL, Consortium for the Barcode of Life; CCMA, Freshwater Microalgae Culture Collection (in Portuguese acronymic); COXI, cytochrome oxidase I; GTR, general-time-reversible nucleotide substitution model; ISS, Index of Substitution Saturation; ISSc, Index of Substitution Saturations critic; ITS, internal transcribed spacer; MCMC, Monte Carlo Makov Chain; NCBI, National Center for Biotechnology Information; OCC, Oedogoniales Chaetopeltidales Chaetophorales; PCR, polymerase chain reaction; *rbc*L, large unit ribulose bisphosphate carboxylase (gene); SC, Sphaeropleales Chlamydomonadales; *tuf*A, elongation factor tu (gene); UCP, universal chlorophyte primers

Received December 20, 2015, **Accepted** April 14, 2016
***Corresponding Author**
E-mail: helenavieira84@gmail.com
Tel: +55-16-3351-8311, Fax: +55-16-3351-8308

INTRODUCTION

The class Chlorophyceae comprises approximately 3,496 described species, according to Algaebase and is one of the most relevant phytoplankton groups in continental waters. The classification of this group is often hampered by the predominance of microscopic cells, frequently lacking obvious structures used to discriminate species or genera. Moreover, life habits, morphologic convergence favored by the unicellular form, the occurrence of cryptic species and asexual reproduction, which keeps mutations that can lead to a large morphologic variability (Potter et al. 1997) are factors that make the classification task arduous (Krienitz et al. 2001, Fawley et al. 2006, Krienitz and Bock 2012, Leliaert et al. 2012).

The urgency of a faster and practical classification system drives many investigations for an efficient molecular marker attending the premises of barcode concept from Consortium for the Barcode of Life (CBOL). This concept comprises the idea that molecular identifications should be conducted using a single pair of primers applicable in the most diverse groups of organisms, recovering a short marker (~700 bp) with enough variation for specific discrimination (CBOL Plant Working Group et al. 2009).

There are many markers proposed for different groups, such as the widely used cytochrome oxidase I (COX I), an official marker for some groups of animals, like fishes (Ward et al. 2005), red (Sherwood et al. 2008, Le Gall and Saunders 2010), and brown algae (McDevit and Saunders 2010), as well as diatoms (Evans et al. 2007).

In green algae, COX I is too variable requiring specific primers to be recovered in different taxa (Fučíková et al. 2011). The amplification of this gene has failed for some chlorophycean taxa (Hall et al. 2010). Furthermore, it may present introns (Turmel et al. 2002), hindering the design of new primers (Saunders and Kucera 2010).

Other markers are frequently used for phylogeny and identification studies of some algal groups, such as *rbc*L (rubisco large subunit), ITS (internal transcribe spacer), *tuf*A (plastid elongation factor). Although widely used in phylogeny of green algae, 18S rDNA (Baldauf et al. 1990, Buchheim and Chapman 1991, An et al. 1999, Krienitz et al. 2001, 2002, Shoup and Lewis 2003, Hall et al. 2010, Buchheim et al. 2011) is a conserved gene (Luo et al. 2010, Fučíková et al. 2011) requiring other genes to solve closely phylogenetic relations in green algae. Moreover, many primers are necessary to recover it from different taxa (Garcia et al. in press), for example, used 12 primers to recover 18S rRNA gene from strains of one family within Chlorophyceae.

The most constructive results achieved so far have focused in phylogenetic questions for genera within the class (Van Hannen et al. 2000, Hall et al. 2010, Fučíková et al. 2011, McManus and Lewis 2011), therefore there is no known marker fulfilling the requirements of a universal barcode marker for Chlorophyceae.

Besides the universality, if the recovered marker has a good phylogenetic signal, it will allow a correct identification of a completely unknown organism, based on its phylogeny among others organisms already described. Thus, although unknown or undescribed, organisms can be classified in lower taxonomic levels if species discrimination is not possible, helping in culturing independent community studies, such as studies using massive sequencing platforms (Reyes et al. 2012, Salipante et al. 2013, Fumagalli et al. 2014).

According to the CBOL criteria of barcode applicability, the first step is to find primers that can recover those candidate molecular markers from the largest possible number of *taxa*. Thus, we aimed to test the universality of primers from published studies, already tested in other groups, for molecular markers in different orders of freshwater Chlorophyceae. Furthermore, we have built a phylogenetic tree with successfully sequenced marker, in order to investigate the possibilities of its application in the class.

MATERIALS AND METHODS

Strain cultures

All organisms are maintained in pure cultures in the Microalgae Collection at the Phycology Laboratory of the Federal University of São Carlos–Freshwater Microalgae Culture Collection (CCMA) (Portuguese acronym). Most strains were cultured in axenic conditions. The strains used in this study were classified and identified according to Algaebase *sensu* Komárek and Fott (Komárek and Fott 1983) (Table 1). *Chaetophora* sp. (CCMA-UFSCar 548) and *Oedogonium* sp. (CCMA-UFSCar 570) strains could not be classified further than genera. The only order from the Chlorophyceae that could not be tested was the Chaetopeltidales, due to the lack of isolates from this order in the culture collection.

Microalgae strains were cultivated in 100 mL Erlenmeyer flasks, with Wright's Chryptophyte medium (Guillard and Lorenzen 1972), pH 7.0, 25 ± 1°C, light intensity of 300 µmol photons m^{-2} s^{-1} and a 12 : 12 light : dark cycle. Cultures in exponential growth phase, determined by op-

Table 1. Microalgae strains from CCMA–UFSCar collection and the amplification and sequencing results for each marker

	Strain	*tuf*A	G*rbc*L	*rbc*LFP	ITS
Chlorophyceae					
Ankistrodesmus densus Korshikov 1953	CCMA-UFSCar 003	KT003399	n/a	n/a	n/a
Ankistrodesmus densus Korshikov 1953	CCMA-UFSCar 128	KT003380	n/a	KT003371	n/a
Ankistrodesmus densus Korshikov 1953	CCMA-UFSCar 239	KT003398	n/a	n/a	n/a
Selenastrum bibraianum Reinsch 1866	CCMA-UFSCar 047	KT003389	KT003379[a]	KT003379[a]	n/a
Selenastrum gracile Reinsch 1866	CCMA-UFSCar 005	KT003390	n/a	n/a	n/a
Selenastrum gracile Reinsch 1866	CCMA-UFSCar 350	KT003391	n/a	n/a	n/a
Kirchneriella aperta Teiling 1912	CCMA-UFSCar 123	KT003385	KT003372	n/a	KT003363
Monoraphidium komarkovae Nygaard 1979	CCMA-UFSCar 353	KT003386	KT003377[a]	KT003377[a]	n/a
Raphidocelis subcaptata (Korshikov) Nygaard, Komárek, J. Kristiansen & O. M. Skulberg 1987	CCMA-UFSCar 048	KT003400	n/a	n/a	n/a
Scenedesmus ecornis (Ehrenberg) Chodat 1926	CCMA-UFSCar 439	KT003393	KT003373[a]	KT003373[a]	n/a
Scenedesmus sp. E. Hegewald 1978	CCMA-UFSCar 088	KT003392	n/a	n/a	n/a
Coealastrum sphaericum Nägeli 1849	CCMA-UFSCar 060	KT003382	n/a	n/a	n/a
Desmodesmus communis (Hegewald) E. Hegewald 2000	CCMA-UFSCar 030	KT003384	KT003374[a]	KT003374[a]	KT003361
Desmodesmus spinosus (Chodat) E. Hegewald 2000	CCMA-UFSCar 046	KT003394	KT003375[a]	KT003375[a]	n/a
Desmodesmus spinosus (Chodat) E. Hegewald 2000	CCMA-UFSCar 062	KT003395	KT003376[a]	KT003376[a]	n/a
Chlamydomonas chlorastera Ettl 1968	CCMA-UFSCar 009	KT003381	n/a	n/a	n/a
Pandorina morum (O. F. Müller) Bory de Saint-Vincent 1824	CCMA-UFSCar 095	KT003388	n/a	n/a	n/a
Pediastrum duplex Meyen 1829	CCMA-UFSCar 055	KT003387	n/a	KT003378	KT003362
Volvox sp. Linnaeus 1758	CCMA-UFSCar 193	KT003396	n/a	KT003369	n/a
Chaetophora sp. F. Shrank 1973	CCMA-UFSCar 548	KT003383	n/a	n/a	n/a
Oedogonium sp. Link ex Hirn 1900	CCMA-UFSCar 570	ns	n/a	n/a	n/a
Trebouxiophyceae					
Nephrocytium lunatum West 1892	CCMA-UFSCar 065	KT003397	n/a	KT003366	n/a

Successfully sequenced marker is indicated by its accession number at GenBank.
n/a, not acquired; ns, not sequenced.
[a]Accession number for overlapping fragments obtained with both *rbc*L pair of primer.

tical density, were harvested in a centrifuge (Eppendorf 5415D; Eppendorf, Hamburg, Germany) under 3,500 ×g resulting in pellets of 40-60 mg of cells for DNA extraction.

DNA extraction and marker gene amplification

The concentrated material was homogenized by mixing in vortex for 15 seconds with glass beads (0.5 mm diameter) (Ningbo Utech International, Formosa, Taiwan) for mechanical cell disruption. The DNA was further extracted with Invisorb Spin Plant Mini Kit (Invitek, Hayward, CA, USA).

Strains of *Nephrocytium lunatum* and *Pandorina morum* form colonies with a thick polysaccharide envelope, which may avoid DNA extraction and hamper the polymerase chain reaction (PCR) reaction. For that reason, these strains were previously washed with lithium chloride to remove this envelope (Nordi et al. 2006).

Primers and PCR reaction

The primers tested for *tuf*A, *rbc*L, and ITS (covering ITS1, 5.8S gene, and ITS2) markers, were chosen from published studies with organisms from class Chlorophyceae (Table 2). We tested two primers for *rbc*L gene, and their resulting fragments are overlapping each other. When both fragments were amplified from the same strain, they were submitted as a unique sequence with one access number.

One of the pairs of primers tested for *rbc*L gene, *rbc*LFP, had the reverse primer designed in this study from sequences of Chlorophyceae available on the National Center for Biotechnology Information (NCBI). We also tested a pair of Universal Plastid Primers for Chlorophyta (UCP4) which recovers a portion of a plastidial gene, proposed by Provan et al. (2004).

The PCR mix was made as recommended by the Taq polymerase manufacturer (DNA polymerase, recombinant, 5 U μL^{-1}; Invitrogen, Carlsbad, CA, USA) with 0.5 µM of each primer. The DNA was quantified by agarose gel electrophoresis using the *ImageLab* 4.0 (BioRad, Hercules, CA, USA) software and ranged from 5 to 10 ng.

PCR profiles were the same for all markers: initial denaturation for 4 min at 94°C; 29 cycles of 45 s at 94°C, annealing temperature specific for each pair of primers (Table 2) and 45 s of extension at 90°C followed by a final extension at 72°C for 7 min. Amplification was verified through electrophoresis in 1% agarose gel. In the case of amplification failure, changes in concentration of PCR reagents, DNA quantity and gradient of annealing temperature were tested, but none of these tests resulted in success of amplification (data not shown). PCR products were purified with polyethylene glycol 20% (polyethylene glycol) solution and NaCl 1 M (Lis and Schleif 1975) and the DNA sequencing was performed by Macrogen (Seoul, Korea).

Sequence analysis

Sequences were aligned with the CLUSTAL W software (Thompson et al. 1994) and the edition and protein frame reading translation, analysis of gaps, in/del and stop codons were performed at GENEIOUS version 6.1.7. Sequences were checked for contamination using the Ba-

Table 2. Molecular markers, names, and sequence of the tested primers

Molecular marker	Primer	Reference	Sequence 5′ → 3′	Fragment size (bp)	Annealing temperature (°C)
*rbc*L	*rbc*L-M379 F	McManus and Lewis (2011)	GGTTTCAAAGCTYTWCGTGC	653-679	50-55
	*rbc*LFP R	Designed (in this study)	GTAAATACCACGGCTACGRTCTT		
*rbc*L	G*rbc*L F	Saunders and Kucera (2010)	GCTGGWGTAAAAGATTAYCG	417-591	50
	G*rbc*L R		TCACGCCAACGCATRAASGG		
*Rpl*5-*rpl*14	UCP4 F	Provan et al. (2004)	ACGATCTAAAAAMGCATACAT	367-421[a]	54
	UCP4 R		AATTGTWTCDTTDGCACCDGAAG		
*tuf*A	*tuf*A F	Fama et al. (2002)	GGNGCNGCNCAAATGGAYGG	758-901	55
	*tuf*A R		CCTTCNCGAATMGCRAAWCGC		
ITS1, 5.8S, ITS2	ITS5 F	White et al. (1990)	GGAAGTAAAAGTCGTAACAAGG	657-737	56
	ITS5 R		TCCTCCGCTTATTGATATGC		

Expected fragment size in base pairs (bp) and annealing temperature (°C) used for each pair of primers.
F, forward; R, reverse; *rbc*L, large unit ribulose bisphosphate carboxylase; UCP, universal chlorophyte primers; ITS, internal transcribed spacer.
[a]Fragment size obtained in the original work.

sic Local Alignment Search Tool (BLAST) (Altschul et al. 1990). Polymorphisms data, polymorphic sites, number of codons, synonym and non-synonym mutations, and parsimony informative sites were calculated with DNAsp 5.10 (Librado and Rozas 2009). Index of Substitution Saturation (ISS) and the Index of Substitution Saturations critic (ISSc) were calculated with the DAMBE5 v5.3.27 software (Xia et al. 2003) to evaluate if there was loss of phylogenetic signal by saturation of substitutions. Sequences were deposited in GenBank under the accession numbers found in Table 1.

Phylogenetic analysis

Phylogeny reconstruction was performed at Mr. Bayes (Huelsenbeck and Ronquist 2001) using a Monte Carlo Makov Chain (MCMC) with 3,000,000 generations, under the general-time-reversible nucleotide substitution model (GTR) (Rodríguez et al. 1990) including parameters for invariable sites (I) and gamma distributed rate variation (G), which was found using *jModelTest* v.0.1.1 (Darriba et al. 2012). Bootstraps values were obtained through neighbor-joining analysis, using 1,000 bootstrap replicates and genetic distances (p-distance) were calculated with MEGA 6 (Tamura et al. 2013).

For phylogenetic analysis with fragments of the *tuf*A gene, sequences from GenBank were included to improve the representation of the order Chaetophorales (*Schizomeris leibleinii* UTEX LB 1228, accession number NC015645) and to represent the orders Chaetopeltidales (*Floydiella terrestris* UTEX 1709, accession number NC014346) which is lacking in our microalgae collection, and Oedogoniales (*Oedogonium cardiacum* UTEX 40, accession number EF587375), due to failure in sequencing the *tuf*A gene of our representative strain. Furthermore, a sequence of *Ostreococcus tauri* (OTTH0595, accession number CR954199), class Mamiellophyceae, was included as outgroup.

RESULTS AND DISCUSSION

DNA amplification and sequencing

The *tuf*A gene was easily amplified in all 22 strains. Only the strain *Oedogonium* sp. did not yield good sequences probably due to contamination, since this strain was not axenic (Table 1).

All the sequences obtained with *tuf*A are new entries in GenBank, although there are *tuf*A sequences deposited for the species *K. aperta*, *P. duplex*, and *P. morum*. The remaining 18 sequences which correspond to 15 species, since there are species with more than one strain, are new entries in the database for this marker.

After alignment of *tuf*A sequences, gaps were not found and the final trimmed fragment had 743 bp, of which 305 were invariable sites, 438 were polymorphic sites displaying 716 mutations and 364 were parsimony informative sites. Amplified region was 247 codons, and the number of sites with synonym mutations was 172.26 and non-synonym mutation was 568.74. Sequences set ISS value (0.32) was significantly lower ($p = 0.001$) than ISSc values (0.75 and 0.50) for symmetric and asymmetric trees, respectively, thus the phylogenetic signal was not hampered by the substitution saturation (Xia et al. 2003) also seen by (Fama et al. 2002, Fučíková et al. 2011).

Considering a lower taxonomic level, for example the family Selenastraceae which has more representatives (9 strains), the highest variation between two strains was 170 bases in a fragment of 826 bp (~20%), and the lowest variation was found between the three strains of the same species, *Ankistrodesmus densus*, 0-10 bases. Thus, the *tuf*A marker was more variable than 18S rRNA gene for this family, since (Garcia et al. in press), for example, using 44 sequences of 18S rDNA (1,511 bp) of different genera of Selenastraceae, found the highest divergence of 76 bp. This higher variability, already shown in other studies of green algae (Hall et al. 2010), could make this gene more useful than the 18S rDNA for delimitation of lower taxonomic levels within the class.

The *tuf*A gene codes for a molecule that mediates the entry of an amino-acyl-tRNA in the ribosome acceptor site during protein synthesis, dictating the peptide chain elongation to be formed. Due to its regulation function, it is a conserved gene (Delwiche et al. 1995), with intermediate evolution rate (Sáez et al. 2008).

The obtained fragment of the *tuf*A gene is a partial coding sequence, being less vulnerable to major mutations that could have caused insertions, deletions or introns, which are unknown in green algae in this gene (Nozaki et al. 2002). Indeed, we have found no indications of introns, making this marker suitable to be tested as DNA barcoding for green algae, and appropriate for phylogenetic reconstruction.

The wide covering and sequencing success of the *tuf*A gene with the primers tested here improves the results for the application of this marker in different groups, since it is already used for plasmodium, cyanobacteria and other bacteria, and terrestrial plants, with sequences available at the NCBI. This pair of primers has also been used in

groups of macro (Du et al. 2014) and microalgae, such as cryptophytes (Garcia-Cuetos et al. 2010) and in the identification of microalgae present in the digestive tract of gastropods (Christa et al. 2013).

Furthermore, it has been widely applied in Ulvophyceae in different studies (Fama et al. 2002, O'Kelly et al. 2004, Wynne et al. 2009, Lawton et al. 2013) presenting great performance as DNA barcode for this class, except for the family Cladophoraceae (Saunders and Kucera 2010). In previous studies, species discrimination power of the *tuf*A marker was observed for Ulvophycean (Fama et al. 2002, Saunders and Kucera 2010) and chlorophyceaan algae albeit they have used few genera from the class Chlorophyceae.

Although we have found that it is possible to recover *tuf*A fragments from diverse chlorophycean *taxa* using a single pair of primers, the same could not be verified for the other markers tested (Supplementary Table S1).

For *rbc*L, it was not possible to perform the amplifications for all the strains using just one pair of primers. The G*rbc*L primers yielded good sequences for only 7 strains (Table 1), whereas the *rbc*LFP primers yielded 15 successful bidirectional sequences (Table 1). The *rbc*LFP primers had good performance from 50 to 55°C of annealing temperature (Table 2), although variations in the annealing temperature did not result in DNA amplification of the strains that failed to amplify in the first test.

A. densus (128) and *Desmodesmus communis* (030) yielded larger fragments (1,188 and 1,114 bp, respectively) than other strains when amplified with the *rbc*LFP primers. Comparing to a reference fragment from the NCBI, these larger sequences had an intermediate region (~800 bp) that could not be aligned with other sequences obtained.

This nucleotide sequence could correspond to an intron, what has already been reported for the *rbc*L in Chlorophyceae (Nozaki et al. 2002, McManus et al. 2012) (Supplementary Table S1). The presence of introns is not wanted in a candidate as a molecular marker since it hampers the design of primers and yields variable length fragments, complicating the sequence alignment. However, the nature of the intermediate portion can only be asserted through specific investigations, which were not the objective of this study.

The greater success of *rbc*LFP primers over G*rbc*L primers may be due to the fact that the first pair was specifically developed to be applied in class Chlorophyceae, using a forward primer chosen from a phylogenetic study with *Pediastrum duplex* (McManus and Lewis 2011) and a reverse primer designed in this study, from Chlorophyceae sequences available at GenBank.

The G*rbc*L primers were designed for application in Ulvophycean organisms (Saunders and Kucera 2010), in which authors tested different regions of the *rbc*L gene, finding better specific discrimination with the 3′ region, but more success of amplification with the 5′ region. Thus, the chosen pair of primers, aiming for universality, was the one that recovered the 5′ region.

However, the low amplification success and low quality sequences led to the exclusion of both *rbc*L primers as universal for class Chlorophyceae.

It must be noticed that although there is a large number of *rbc*L sequences available in GenBank for class Chlophyceae and other groups, they were often obtained using different primers and may be different regions of the gene, which makes their use as genetic markers for phylogeny or barcode less practical (Supplementary Table S1).

For the ITS region only 3 strains showed good sequencing (Table 1). The pair of primers ITS4-ITS5 for ITS region was chosen among proposed primers in a study with fungi phylogeny (White et al. 1990) and has already been tested with organisms from Chlorophyceae (Van Hannen et al. 2000, Buchheim et al. 2012).

Because it is a spacer region and is under a relaxed selection, mutations may not be strictly selected, which means it is very variable and may present in/dels and inconsistent sizes among the *taxa*, being commonly used for phylogeny within genus and species in green algae (Verbruggen et al. 2006, O'Kelly et al. 2010) (Supplementary Table S1). Thus, the highly variable nature of the ITS region may have contributed to its failure as a universal primer for Chlorophyceae, probably requiring particularly designed primers for each case.

Although the UCP4 primers have been proposed as universal for application in Chlorophyceae (Provan et al. 2004), no strain could be amplified following the protocol used in the original study, even when different annealing temperatures were tested. Provan et al. (2004) have tested the universality of primers for plastidial DNA using four organisms representing the Division Chlorophyta, with only one organism of the class Chlorophyceae, the specie *Dunaliella salina*.

The pair UCP4 was chosen in their study because the targeted region had the best combinations of characteristics for DNA Barcoding among the proposed regions, like constancy of non-coding sites number and the fragment size in the amplified lineages. Although the pair of UCP4 primers had worked for *D. salina*, it did not work for any of our strains.

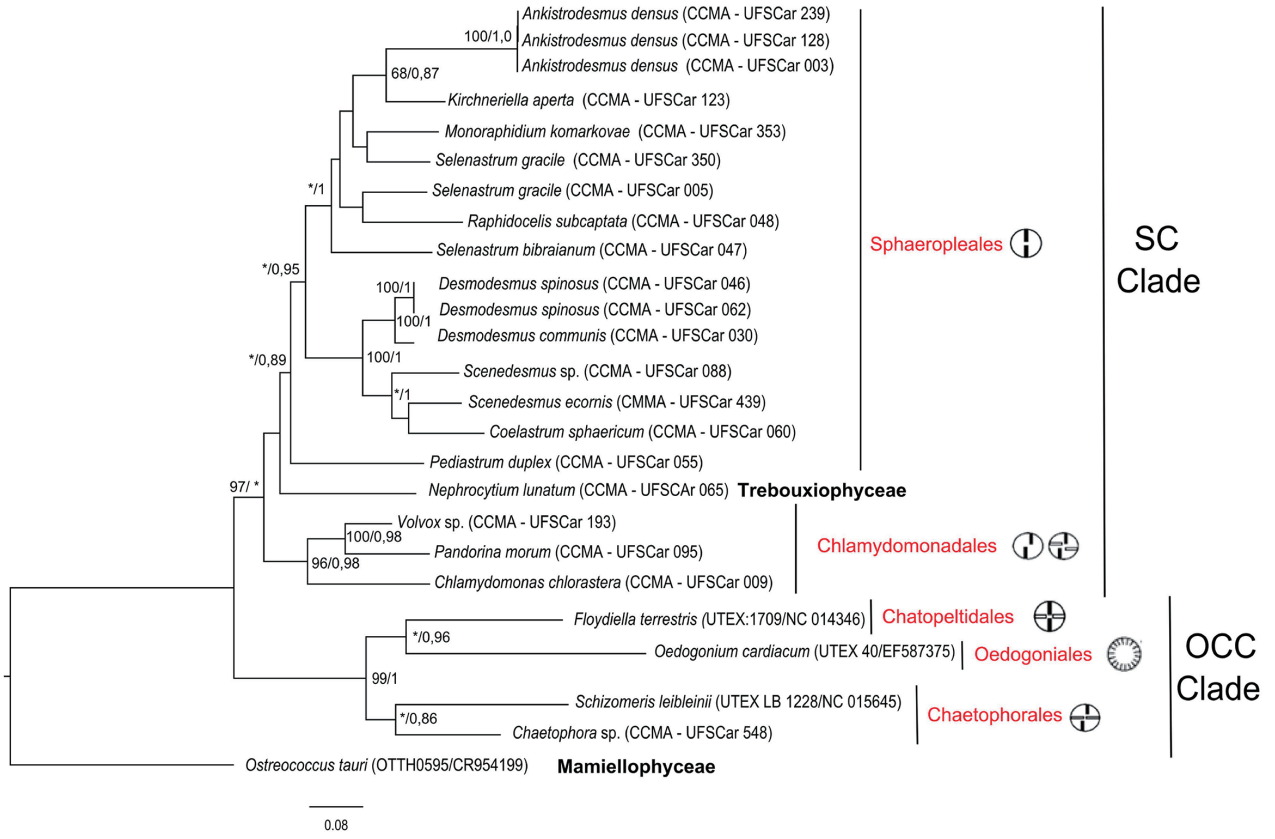

Fig. 1. Bayesian analysis for *tuf*A sequences of Chlorophyceae. Sequences obtained in this study are indicated with "CCMA-UFSCar" and the four sequences obtained from GenBank are indicated with the name strain and the accession number. *Ostreococcus tauri* (Mamiellophyceae) was used as outgroup. Support values at the nodes are the bootstrap values (%) for neighbor-joining, followed by Bayesian posterior probability. Values lower than 95% for bootstrap value and 0.75 for Bayesian probability are represented by asterisk, or not presented when both were lower. The draws represent the basal body orientation of flagella apparatus, as the cells are represented as circles and seen from above. OCC, Oedogoniales Chaetopeltidales Chaetophorales; SC, Sphaeropleales Chlamydomonadales.

Phylogeny

Concerning the use of the genes studied for phylogeny, the Bayesian tree topology with sequences of the *tuf*A gene showed the monophyly of class Chlorophyceae and the five represented orders: Sphaeropleales, Chlamydomonadales, Oedogoniales, Chaetopeltidales, and Chaetophorales (Fig. 1).

According to the flagella evolution (orientation of the basal body and number of flagella), it is possible to observe the Oedogoniales Chaetophorales Chaetopeltidales (OCC) clade, containing Oedogoniales, Chaetophorales, and Chaetopeltidales, and Sphaeropleales Chlamydomonadales (SC) clade, with Sphaeropleales and Chlamydomonadales. It is also in agreement with other studies that used the 18S rRNA gene (Alberghina et al. 2006, Němcová et al. 2011), 18S and 28S rRNA genes (Shoup and Lewis 2003) and nuclear and plastidial genes combined (Turmel et al. 2008, Tippery et al. 2012).

Despite not strongly supported (Bootstrap / Bayesian probability = 45/0.95), the monophyly of Sphaeropleales was shown with clear delineation of the families Selenastraceae and Scenedesmaceae (94 / 1.0 and 100 / 1.0, respectively) (Fig. 1). However, it is important to remember that some Sphaeropleales families were not represented here, and future works with the *tuf*A gene must include them.

As many authors have already found using other genes (Fawley et al. 2006, Krienitz et al. 2011, Krienitz and Bock 2012), some internal branches were not clearly solved with superimposed genera, reflecting that genetic data may not behave consistently with morphology and leading to ambiguity in species delimitation. For example, the sickle morphology visible in Selenastraceae and used as identification also occurs in Trebouxiophyceae, indicating morphological convergence.

In summary, the *tuf*A marker, standing alone, rebuilt the class Chlorophyceae phylogeny, which is often obtained with different genes combined, also at the internal branches, commonly addressed in specific investigations. Besides the overlap of some genera within Sphaeropleales, another issue that must be addressed is the position of *N. lunatum*. This species is currently classified as a Trebouxiophyceae member, but according to our phylogenetic reconstruction with the *tuf*A marker, *N. lunatum* was positioned among Chlorophyceae, within the SC clade, close to Sphaeropleales and Chlamydomonadales (Fig. 1).

The *Nephrocytium* genus has already been classified in class Chlorophyceae, order Chlorococcales previously (West 1892, Pascher 1915), but families from this order were reorganized and redistributed. However, the transfer of the family to Trebouxiophyceae was based on the analysis of other genera (Friedl 1995) and the genus *Nephrocytium* was, apparently, passively transferred together with the other Chlorellaceae. Such taxonomic transferences have been already investigated, suggesting the resurrection of a Chlorophyceae genus to accommodate linages that were transferred to Trebouxiophyceae (Fučíková and Lewis 2012).

Nevertheless, the *Nephrocytium* genus is often missing in studies of phylogeny of Chlorophyceae and Trebouxiophyceae (Friedl 1995, Krienitz et al. 2002), and is underrepresented in this study, making a focused study with combined genes an essential procedure to elucidate its classification.

CONCLUSION

One of the critical characteristics for molecular markers is its applicability in as many organisms as possible. Among the 5 molecular markers tested here, *tuf*A seems to comply with this objective for chlorophycean microalgae.

The easy amplification, sequencing and alignment of sequences, the crescent amount of available sequences on data bases summed with the good phylogenetic signal allowing a realistic phylogenetic reconstruction, despite the higher variability than 18S rRNA gene, indicate the *tuf*A gene as a promising molecular marker for the class. However, its utilization as a DNA barcode in Chlorophyceae, alone or combined with others markers, need to be tested in further studies, comprising problematic taxa, such as family Selenastraceae.

Despite the *rbc*L primers not amplifying all the strains they could be applied for green algae genera / species in focused studies. The primers tested for ITS and UCP4 regions were not appropriate for universal application in Chlorophyceae due to their low amplification / sequencing success rate.

SUPPLEMENTARY MATERIAL

Supplementary Table S1. Advantages and disadvantages of each molecular marker tested in this study and other principal markers used in studies with green algae (http://e-algae.org).

ACKNOWLEDGEMENTS

We would like to thank Thaís Garcia da Silva for the morphological identification of the microalgae strains. We also wish to thank Dr. Pedro Manoel Galetti Junior for the suggestions made for the development of this work.

REFERENCES

Alberghina, J. S., Vigna, M. S. & Confalonieri, V. A. 2006. Phylogenetic position of the Oedogoniales within the green algae (Chlorophyta) and the evolution of the absolute orientation of the flagellar apparatus. Plant Syst. Evol. 261:151-163.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

An, S. S., Friedl, T. & Hegewald, E. 1999. Phylogenetic relationships of *Scenedesmus* and *Scenedesmus*-like coccoid green algae as inferred from ITS-2 rDNA sequence comparisons. Plant Biol. 1:418-428.

Baldauf, S. L., Manhart, J. R. & Palmer, J. D. 1990. Different fates of the chloroplast *tuf*A gene following its transfer to the nucleus in green algae. Proc. Natl. Acad. Sci. U. S. A. 87:5317-5321.

Buchheim, M. A. & Chapman, R. L. 1991. Phylogeny of the colonial green flagellates: a study of 18S and 26S rRNA sequence data. Biosystems 25:85-100.

Buchheim, M. A., Keller, A., Koetschan, C., Förster, F., Merget, B. & Wolf, M. 2011. Internal transcribed spacer 2 (nu ITS2 rRNA) sequence-structure phylogenetics: towards an automated reconstruction of the green algal tree of life. PLoS One 6:e16931.

Buchheim, M. A., Sutherland, D. M., Schleicher, T., Förster,

F. & Wolf, M. 2012. Phylogeny of Oedogoniales, Chaetophorales and Chaetopeltidales (Chlorophyceae): inferences from sequence-structure analysis of ITS2. Ann. Bot. 109:109-116.

CBOL Plant Working Group, Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M. W., Cowan, R. S., Erickson, D. L., Fazekas, A. J., Graham, S. W., James, K. E., Kim, K. J., Kress, W. J., Schneider, H., van AlphenStahl, J., Barrett, S. C., van den Berg, C., Bogarin, D., Burgess, K. S., Cameron, K. M., Carine, M., Chacón, J., Clark, A., Clarkson, J. J., Conrad, F., Devey, D. S., Ford, C. S., Hedderson, T. A., Hollingsworth, M. L., Husband, B. C., Kelly, L. J., Kesanakurti, P. R., Kim, J. S., Kim, Y. D., Lahaye, R., Lee, H. L., Long, D. G., Madriñán, S., Maurin, O., Meusnier, I., Newmaster, S. G., Park, C. W., Percy, D. M., Petersen, G., Richardson, J. E., Salazar, G. A., Savolainen, V., Seberg, O., Wilkinson, M. J., Yi, D. K. & Little, D. P. 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. U. S. A. 106:12794-12797.

Christa, G., Wescott, L., Schäberle, T. F., König, G. M. & Wägele, H. 2013. What remains after 2 months of starvation? Analysis of sequestered algae in a photosynthetic slug, *Plakobranchus ocellatus* (Sacoglossa, Opisthobranchia), by barcoding. Planta 237:559-572.

Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9:772.

Delwiche, C. F., Kuhsel, M. & Palmer, J. D. 1995. Phylogenetic analysis of *tuf*A sequences indicates a cyanobacterial origin of all plastids. Mol. Phylogenet. Evol. 4:110-128.

Du, G., Wu, F., Mao, Y., Guo, S., Xue, H. & Bi, G. 2014. DNA barcoding assessment of green macroalgae in coastal zone around Qingdao, China. J. Ocean Univ. China 13:97-103.

Evans, K. M., Wortley, A. H. & Mann, D. G. 2007. An assessment of potential diatom "barcode" genes (cox1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158:349-364.

Fama, 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. 38:1040-1050.

Fawley, M. W., Dean, M. L., Dimmer, S. K. & Fawley, K. P. 2006. Evaluating the morphospecies concept in the Selenastraceae (Chlorophyceae, Chlorophyta). J. Phycol. 42:142-154.

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. 31:632-639.

Fučiková, K. & Lewis, L. A. 2012. Intersection of *Chlorella*, *Muriella* and *Bracteacoccus*: resurrecting the genus *Chromochloris* Kol et Chodat (Chlorophyceae, Chlorophyta). Fottea 12:83-93.

Fučíková, K., Rada, J. C., Lukešová, A. & Lewis, L. A. 2011. Cryptic diversity within the genus *Pseudomuriella* Hanagata (Chlorophyta, Chlorophyceae, Sphaeropleales) assessed using four Barcode markers. Nova Hedwigia 93:29-46.

Fumagalli, M., Vieira, F. G., Linderoth, T. & Nielsen, R. 2014. ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics 30:1486-1487.

Garcia-Cuetos, L., Moestrup, Ø., Hansen, P. J. & Daugbjerg, N. 2010. The toxic dinoflagellate *Dinophysis acuminata* harbors permanent chloroplasts of cryptomonad origin, not kleptochloroplasts. Harmful Algae 9:25-38.

Garcia, T. S., Bagatini, I. L., Bock, C., Sant'Anna & Vieira, A. A. H. Selenastraceae (Sphaeropleales, Chlorophyceae): *rbc*L, 18S rDNA and ITS-2 secondary structure enlightens traditional taxonomy, with description of two new genera, *Messastrum* gen. nov. and *Curvastrum* gen. nov. Fottea (in press).

Guillard, R. R. L. & Lorenzen, C. J. 1972. Yellow-green algae with chlorophyllide C2. J. Phycol. 8:10-14.

Hall, J. D., Fučíková, K., Lo, C., Lewis, L. A. & Karol, K. G. 2010. An assessment of proposed DNA barcodes in freshwater green algae. Cryptogam. Algol. 4:529-555.

Huelsenbeck, J. P. & Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.

Komárek, J. & Fott, B. 1983. *Das Phytoplankton des Süsswassers: Systematik und Biologie. Teil 7*. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, 1044 pp.

Krienitz, L. & Bock, C. 2012. Present state of the systematics of planktonic coccoid green algae of inland waters. Hydrobiologia 698:295-326.

Krienitz, L., Bock, C., Nozaki, H. & Wolf, M. 2011. SSU rRNA gene phylogeny of morphospecies affiliated to the bioassay alga "*Selenastrum capricornutum*" recovered the polyphyletic origin of crescent-shaped Chlorophyta. J. Phycol. 47:880-893.

Krienitz, L., Ustinova, I., Friedl, T. & Huss, V. A. R. 2001. Traditional generic concepts versus 18S rRNA gene phylogeny in the green algal family Selenastraceae (Chlorophyceae, Chlorophyta). J. Phycol. 37:852-865.

Krienitz, L., Wolf, M., Hegewald, E. & Hepperle, D. 2002. Sys-

tematics of coccoid green algae: morphology vs. 18S rRNA gene phylogeny. J. Phycol. 38(Suppl 1):19-20.

Lawton, R. J., Mata, L., de Nys, R. & Paul, N. A. 2013. Algal bioremediation of waste waters from land-based aquaculture using ulva: selecting target species and strains. PLoS One 8:e77344.

Le Gall, L. & Saunders, G. W. 2010. DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora. J. Phycol. 46:374-389.

Leliaert, F., Smith, D. R., Moreau, H., Herron, M. D., Verbruggen, H., Delwiche, C. F. & De Clerck, O. 2012. Phylogeny and molecular evolution of the green algae. Crit. Rev. Plant Sci. 31:1-46.

Librado, P. & Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Lis, J. T. & Schleif, R. 1975. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. Nucleic Acids Res. 2:383-389.

Luo, W., Pröschold, T., Bock, C. & Krienitz, L. 2010. Generic concept in *Chlorella*-related coccoid green algae (Chlorophyta, Trebouxiophyceae). Plant Biol. (Stuttg.) 12:545-553.

McDevit, D. C. & Saunders, G. W. 2010. A DNA barcode examination of the Laminariaceae (Phaeophyceae) in Canada reveals novel biogeographical and evolutionary insights. Phycologia 49:235-248.

McManus, H. A. & Lewis, L. A. 2011. Molecular phylogenetic relationships in the freshwater family Hydrodictyaceae (Sphaeropleales, Chlorophyceae), with an emphasis on *Pediastrum duplex*. J. Phycol. 47:152-163.

McManus, H. A., Lewis, L. A., Fučíková, K. & Haugen, P. 2012. Invasion of protein coding genes by green algal ribosomal group I introns. Mol. Phylogenet. Evol. 62:109-116.

Němcová, Y., Eliáš, M., Škaloud, P., Hodač, L. & Neustupa, J. 2011. *Jenufa* gen. nov.: a new genus of coccoid green algae (Chlorophyceae, incertae sedis) previously recorded by environmental sequencing. J. Phycol. 47:928-938.

Nordi, C. S. F., Cavagliere, T. G. W. F., Vieira, A. A. H. & Nascimento, O. R. 2006. Chaotropic effect of lithium íon on permeability of polysaccharide capsule of the microalga *Ankistrodesmus gracilis* (Reinsch) Korsikov (Chlorophyceae). Acta Bot. Bras. 20:449-454.

Nozaki, H., Takahara, M., Nakazawa, A., Kita, Y., Yamada, T., Takano, H., Kawano, S. & Kato, M. 2002. Evolution of *rbc*L group IA introns and intron open reading frames within the colonial Volvocales (Chlorophyceae). Mol. Phylogenet. Evol. 23:326-338.

O'Kelly, C. J., Kurihara, A., Shipley, T. C. & Sherwood, A. R. 2010. Molecular assessment of *Ulva* spp. (Ulvophyceae, Chlorophyta) in the Hawaiian islands. J. Phycol. 46:728-735.

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.

Pascher, A. 1915. Chlorophyceae II Tetrasporales, Protococcales, einzellige gattungen unsicherer stellung. *In* Lemmermann, E., Brunnthaler, J. & Pascher, A. (Eds.) *Die Süswasser-Flora Deutschlands, Österreichs und der Schweiz*. Gustav Fischer, Jena, p. 170.

Potter, D., Saunders, G. W. & Anderson, R. A. 1997. Convergente evolution masks extensive biodiversity among marine coccoid picoplankton. Biodivers. Conserv. 6:99-107.

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

Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F. & Gordon, J. I. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. Nat. Rev. Microbiol. 10:607-617.

Rodríguez, F., Oliver, J. L., Marín, A. & Medina, J. R. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142:485-501.

Sáez, A. G., Zaldivar-Riverón, A. & Medlin, L. K. 2008. Molecular systematics of the Pleurochrysidaceae, a family of coastal coccolithophores (Haptophyta). J. Plankton Res. 30:559-566.

Salipante, S. J., Sengupta, D. J., Rosenthal, C., Costa, G., Spangler, J., Sims, E. H., Jacobs, M. A., Miller, S. I., Hoogestraat, D. R., Cookson, B. T., McCoy, C., Matsen, F. A., Shendure, J., Lee, C. C., Harkins, T. T. & Hoffman, N. G. 2013. Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. PLoS One 8:e65226.

Saunders, G. W. & Kucera, H. 2010. An evaluation of *rbc*L, *tuf*A, UPA, LSU and ITS as DNA barcode markers for the marine macroalgae. Cryptogam. Algol. 31:487-528.

Sherwood, A. R., Vis, M. L., Entwisle, T. J., Necchi, O. Jr. & Presting, G. G. 2008. Contrasting intra versus interspecies DNA sequence variation for representatives of the Batrachospermales (Rhodophyta): insights from a DNA barcoding approach. Phycol. Res. 56:269-279.

Shoup, S. & Lewis, L. A. 2003. Polyphyletic origin of parallel basal bodies in swimming cells of chlorophycean green algae (Chlorophyta). J. Phycol. 39:789-796.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S.

2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30:2725-2729.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Tippery, N. P., Fučíková, K., Lewis, P. O. & Lewis, L. A. 2012. Probing the monophyly of the Sphaeropleales (Chlorophyceae) using data from five genes. J. Phycol. 48:1482-1493.

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. 44:739-750.

Turmel, M., Otis, C. & Lemieux, C. 2002. The complete mitochondrial DNA sequence of *Mesostigma viride* identifies this green alga as the earliest green plant divergence and predicts a highly compact mitochondrial genome in the ancestor of all green plants. Mol. Biol. Evol. 19:24-38.

Van Hannen, E. J., Lürling, M. & van Donk, E. 2000. Sequence analysis of the ITS-2 region: a tool to identify strains of *Scenedesmus* (Chlorophyceae). J. Phycol. 36:605-607.

Verbruggen, H., De Clerck, O., N'yeurt, A. D. R., Spalding, H. & Vroom, P. S. 2006. Phylogeny and taxonomy of *Halimeda incrassata*, including descriptions of *H. kanaloana* and *H. heteromorpha* spp. nov. (Bryopsidales, Chlorophyta). Eur. J. Phycol. 41:337-362.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360:1847-1857.

West, W. 1892. Algae of the English Lake District. J. R. Microsc. Soc. Lond. 14:1-7.

White, T. J., Bruns, T., Lee, S. & Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. (Eds.) *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York, pp. 315-322.

Wynne, M. J., Verbruggen, H. & Angel, D. L. 2009. The recognition of *Caulerpa integerrima* (Zanardini) comb. et stat. nov. (Bryopsidales, Chlorophyta) from the Red Sea. Phycologia 48:291-301.

Xia, X., Xie, Z., Salemi, M., Chen, L. & Wang, Y. 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26:1-7.