

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E RECURSOS NATURAIS
DEPARTAMENTO DE BOTÂNICA
LABORATÓRIO DE FICOLOGIA

**Avaliação da Universalidade de *Primers*
de Marcadores Moleculares para Aplicação na Classe
Chlorophyceae
(Chlorophyta)**

Helena Henriques Vieira

São Carlos
2014

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA E RECURSOS NATURAIS
DEPARTAMENTO DE BOTÂNICA
LABORATÓRIO DE FICOLOGIA

**Avaliação da Universalidade de *Primers*
de Marcadores Moleculares para Aplicação na Classe
Chlorophyceae
(Chlorophyta)**

Helena Henriques Vieira

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação Em Ecologia e Recursos Naturais da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Ecologia e Recursos Naturais, área de concentração em Ecologia e Recursos Naturais.

São Carlos
2014

Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar

V658au Vieira, Helena Henriques.
Avaliação da universalidade de *primers* de marcadores moleculares para aplicação na classe Chlorophyceae (Chlorophyta) / Helena Henriques Vieira. -- São Carlos : UFSCar, 2014.
44 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2014.

1. Marcador molecular. 2. Chlorophyceae. 3. Filogenia. 4. *tufA*. 5. *rbcL*. I. Título.

CDD: 574.87328 (20^a)


Helena Henriques Vieira

Dissertação apresentada à Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Ecologia e Recursos Naturais.

Aprovada em 30 de maio de 2014

BANCA EXAMINADORA

Presidente




Prof. Dr. Armando Augusto Henriques Vieira
(Orientador)

1º Examinador



Prof. Dr. Pedro Manoel Galetti Júnior
PPGERN/UFSCar

2º Examinador



Profa. Dra. Célia Leite Sant'Anna
USP/São Paulo-SP

Agradecimentos

- À UFSCar e ao PPGERN pela infra-estrutura fornecida e pela formação acadêmica
- A Fapesp pelo financiamento do projeto temático e Capes pela bolsa concedida.
- Ao meu pai e orientador Dr. Armando Augusto Henriques Vieira, por ter me instruído profissionalmente nos momentos de inexperiência e pelos ensinamentos. Pela relação aluna/filha-orientador/pai, pois é um enorme privilégio poder trabalhar com quem possa te aconselhar sinceramente. E que segue me orientando há 29 anos.
- À Dra. Inessa Lacativa Bagatini, por ter me introduzido, com toda paciência, atenção e de maneira criteriosa e também por ter me ensinado todo o fundamental do mundo da biologia molecular. Por ter me acompanhado no passo-a-passo de cada novidade a ser aprendida. Por ter me criticado e elogiado quando preciso e por ter mantido toda atenção e disposição ao longo do meu mestrado.
- Ao professor Dr. Pedro Manuel Galetti Jr. Pela imensa contribuição no momento da avaliação de qualificação de mestrado, pelas sugestões e disponibilidade em ajudar quando fosse preciso.
- À Carla Guinart, pela importante participação para que este trabalho fosse concluído, sendo indispensável sua participação, e pela disposição em me receber por dias para trabalharmos juntas.
- À professora Dra. Mariana Cabral de Oliveira que me recebeu em seu laboratório para que pudesse acompanhar o trabalho de seus alunos e me inserir na rotina de um laboratório de biologia molecular e também pelas importantes sugestões.
- À professora Dra. Célia Leite Sant'Anna e à Thaís Garcia da Silva pela identificação e fotografias das cepas utilizadas neste estudo, cuja tarefa é extremamente especializada e importante para o desenvolvimento de estudos científicos com tais organismos.
- Ao Luís Aparecido Joaquim por nos levar as coletas, sem as quais não haveria cepas! E claro, pela agradabilíssima companhia e paciência em viajar conosco por uma semana variando momentos de bom humor com extrema falta de esportiva!
- Ao Guilherme Pavan de Moraes, amigo e namorado que vivenciou as mesmas ansiedades por estar fazendo mestrado ao mesmo tempo e no mesmo laboratório. Foi muito gostoso ter sua companhia, suas muitas críticas e sugestões que foram muito importantes, compartilhar momentos de preguiça, experimentos intensos, as cervejas no final do dia nas últimas semanas, troca de olhares de desespero, enfim, pela compreensão!

- Aos amigos de laboratório: Inessa, Letícia, Fabrício, Guilherme, Cilene, Thaís, Ingritt, Luíz, Naiara, Rodrigo (Chico), Rodrigo (Pepino) e Lucas, pelas ajudas, dicas, momentos de descontração, risadas, hora-do-lanche, hora-do-café, ou seja, por tornar esses anos no laboratório muito agradáveis pela companhia de vocês. Às saídas para uma cerveja e muitas comidas em grupo, sempre divertidíssimas!
- Aos amigos sempre presentes: Isabela Baratinha pela pausa pro cafezinho todos os dias, as maravilhosas massagens e amizade; Fá, Kiki, Vini, Fer, Aline, Lucão, Caipora, Pedro Mala, Camila, Cabeção, Mestre e claro, o Guilherme pelas reuniões, saídas, encontros, desde o primeiro ano de faculdade. Dizem que amizade que supera os 7 anos é amizade pra sempre. Tudo besteira, hahaha! Mas ainda estamos juntos. E à minha amiga de infância, Cecy, por estar até hoje presente!
- À minha família; Mãe, Pai, Dê, Rebeca, Essio e família incluí o cão Xigue e o gato Gomes. Família é pra quem a gente corre sempre. Sozinha não conseguiria nunca!
- Aos meus sogros e cunhados que estão sempre presentes, me consolando em momentos de desânimos ou celebrando pelo simples motivo de estarmos todos juntos, constituindo minha segunda família.
- Aos demais professores, familiares e amigos e que de alguma maneira estão ou estiveram presente e foram fundamentais para que pudesse concluir este trabalho.

Lista de Abreviaturas

BLAST	Basic Local alignment Search Tool
BSA	Bovine Serum Albumin
CBOL	Consortium for the Barcode of Life
COXI	Citocromo Oxidase I
CW	Clockwise
DNA	Ácido desoxirribonucleico
dNTP	Desoxinucleotídeos trifosfatados
DO	Diretamente opostos
EDTA	Ácido etilenodiamino tetra-acético
GTR-I-G	General Reversion Time – Invariable - Gamma
in/del	Insertions/deletions
ITS	Internal Transcribed Spacer
NGS	Next Generation Sequencing
NJ	Neighbor-Joining
MCMC	Monte Carlo Markov Chain
OCC	Oedogoniales Chaetopeltidales Chaetophorales
pb	Pares de base
PEG	Polietilenoglicol
PCR	Polymerase Chain reaction
<i>rbcL</i>	Large unit Ribulose Bifosfato Carboxilase (gene)
SC	Sphaeropleales Chlamydomonadales
TBE	TRIS-Borato-EDTA
TRIS	Base tris(hidroximetil)aminometano
<i>tufA</i>	Elongation Factor tu (gene)
UCP	Universal Chlorophyte Primers
WC	Wright's Chryptophyte
WDCM	World Data Microorganisms Center

Resumo

A abordagem polifásica de taxonomia se tornou indispensável na classificação especialmente de microorganismos, como os da classe Chlorophyceae. Esta classe é composta de organismos muito distintos, sendo a maioria microscópica, apresentando formação do talo unicelular, colonial e filamentosa e a classificação e identificação dos representantes desta classe são muitas vezes dificultadas pelo fato de não possuírem estruturas observáveis em microscópio usadas para discriminação. Alguns hábitos e tipos de reprodução, a ocorrência de espécies crípticas são alguns fatores que podem resultar em formas que também dificultam a taxonomia destes organismos. Associada a uma classificação mais robusta, a procura por métodos mais rápidos e práticos, como o DNA *Barcode*, se intensificou e muitos marcadores moleculares já se encontram disponíveis na literatura para serem testados. Para alguns grupos de animais, algas pardas, algas vermelhas e diatomáceas o marcador COXI é considerado oficial como DNA *Barcode*. Entretanto, outros grupos como Chlorophyceae este marcador não é aplicável devido a sua natureza conservada e, portanto, estão sob constante investigação, devido ao fato de que ainda não foram encontrados marcadores que possam ser aplicados satisfatoriamente como DNA *Barcode*. Marcadores plastidiais como *tufA*, *rbcl* e nucleares como o ITS são alguns dos mais aplicados em estudos com algas verdes, muitas vezes combinados entre si ou com outros marcadores. Entretanto para que figurem como marcadores moleculares devem ser amplificados por *primers* que possam ser aplicados de maneira ampla. Assim, o presente estudo teve como objetivo testar *primers* para os marcadores *tufA*, ITS, *rbcl*, disponíveis na literatura e já aplicados com outros grupos, buscando a universalidade destes *primers* também em Chlorophyceae. Foi testado também um par de *primers* proposto como universal para Chlorophyta (UCP4). Os *primers* para o gene *tufA* mostraram universalidade para a classe amplificando todas as 22 cepas testadas com falha no seqüenciamento apenas para a cepa de *Oedogonium sp.* Os *primers* para o gene *rbcl*, *Grbcl*, amplificou 12 e seqüenciou satisfatoriamente 7 cepas das 22 utilizadas, e o par *rbclFP*, o qual teve um de seus iniciadores desenhado neste estudo, amplificou e seqüenciou 16 cepas. *Primers* para o gene ITS amplificaram 5 cepas com bom seqüenciamento de apenas 3 e o par de *primers* UCP4 não amplificou nenhuma cepa. Além da universalidade observada do par de *primers* para o gene *tufA*, análises com as seqüências obtidas mostraram que este gene se mostrou um promissor candidato a marcado molecular para aplicação na classe Chlorophyceae.

Abstract

The polyphasic taxonomy approach became indispensable in classification of living beings, especially microorganisms, such as Chlorophyceae. This class is composed of very distinct organisms most microscopic, unicellular, colonial, filamentous, and the classification and identification is many times hampered due to the lack of structures observable that helps the discrimination. Some habits, reproduction, the occurrence of cryptic species are some factors that can lead to forms that make difficult the taxonomy of these organisms. Besides the traditional morphology, the molecular biology allows a robust concerning the classification of the studied organism. Associated to a better classification, the search for faster and more practical methods of identification, as DNA Barcode, became intensified and many molecular markers are available in literature to be tested. For some groups of animals, brown algae, red algae and diatoms the marker COXI is considered an official DNA Barcode. However other groups, as Chlorophyceae, are under constant investigation, due to the fact that molecular marker that can be used satisfactory as DNA Barcode have not been found yet. Plastid markers as *tufA*, *rbcL* and nuclear as ITS are some of the most proposed to be applied in green algae, often combined with each other or other markers. However, for act as a molecular marker, it is necessary to be amplified by primers broadly applied. The present study has tested primers for *tufA*, ITS and *rbcL* marker, available in literature and already in use in other groups, looking for universality in Chlorophyceae. We also tested proposed universal primers for Chlorophyta (UCP4). Moreover, preliminary analyses have been performed for application of the markers as barcode in the class. The primers used for *tufA* marker showed universality in Chlorophyceae amplifying for all 22 strains tested and sequencing have failed only for *Oedogonium* strain. From the *rbcL* marker, the primer pair *GrbcL* has amplified 12 and sequenced 7 of 22 strains and *rbcLFP* – which one the reverse primer was designed in this study – amplified and sequenced 16 lineages. Primers from the ITS marker amplified only 5 lineages and UCP4 primers have not amplified any strain. Beside the universality found for *tufA* primers, Analysis with the obtained sequences showed that the *tufA* gene is a promising molecular marker to be applied in Chlorophyceae.

Conteúdo

Lista de Abreviaturas	I
Resumo	II
Abstract	III
Introdução	1
Objetivo	9
Materiais e Métodos	9
<i>Organismos</i>	9
<i>Fotografias identificações</i>	11
<i>Cultivo</i>	11
<i>Extração de DNA</i>	12
<i>Amplificação do DNA</i>	13
<i>Primers</i>	15
<i>Análise das Sequências</i>	18
Resultados	19
<i>Universalidade dos primers testados e rendimento de seqüências</i>	19
UCP4:.....	19
ITS	20
GrbcL	22
rbcLFP.....	22
tufA.....	26
<i>Análises filogenéticas</i>	28
<i>Nephrocitium lunatum</i>	34
Conclusões	35
Referências	37

Introdução

As microalgas protagonizam inúmeros estudos mundiais em biotecnologia, muitos já em fase de aplicação comercial. São fontes de energias renováveis como biodiesel, bioetanol e biogás (metano), fonte de produção de proteínas para ração animal e muitos outros bioprodutos para uso em fármacos, cosméticos, alimentos, como a astaxantina, b-caroteno, etc. Exemplificando seu valor e eficiência, a produção de óleo de algumas espécies é de até 60% de seu peso seco em lipídios sendo muitas vezes mais produtivas que plantas terrestres oleaginosas (Sheehan *et al.* 1998).

Por serem eficientes assimiladores de carbono respondem pela maior produção primária dos corpos d'água, e são conseqüentemente, organismos adequados para utilização no mercado de crédito de carbono. (Hanagata *et al.* 1992).

Devido à plasticidade fisiológica do grupo, são cultiváveis em diversos ambientes, desde baixas concentrações de nutrientes (Griffiths & Harrison 2009) até efluentes domésticos, reduzindo o conteúdo de nutrientes inorgânicos (Chevalier & Noüe 1985; Sawayama *et al.* 1995), e podem também ocupar áreas não produtivas para agricultura. Demandam pouca energia e água para sua produção e dispensam o uso de agrotóxicos que oferecem riscos de contaminação ao meio ambiente, muito freqüentes na agricultura.

Entretanto, o termo “microalgas” representa um universo enorme, com mais de 50.000 espécies (e aumentando). A maior biodiversidade de algas eucarióticas em águas continentais é representada pelas algas verdes da classe Chlorophyceae (Chlorophyta). Somente esta classe possui cerca de 2650 espécies, distribuídas em 355 gêneros (Hoek *et al.* 1995) e são majoritariamente microalgas de água doce.

Esta classe agrupa indivíduos com mitose fechada e citocinese por formação de

ficoplasto (Stewart, & Mattox, 1975) que apresentam reprodução sexuada com ciclo reprodutivo haplôntico e com meiose zigótica. As células destes representantes têm cloroplasto envolto em duas membranas com tilacóides dispostos em lamelas e os pigmentos fotossintetizantes são clorofila a e b, além de pigmentos acessórios (Hoek *et al.* 1995).

A morfologia desta classe é muito diversificada, encontrando-se organismos unicelulares, coloniais, cocóides, palmelóides e filamentosos, e até multicelulares (Lewis & McCourt 2004) com uma rudimentar divisão de papéis, como em *Volvox* e *Oedogonium*. Possuem diferentes formatos celulares, hábitos e tipos de reprodução.

Dados moleculares de filogenia mostraram a monofilia desta classe e sua subdivisão em dois grandes clados remetendo a evolução do aparato flagelar: clado OCC, com as ordens Oedogoniales, Chaetophorales e Chaetopeltidales e clado SC, com Sphaeropleales e Chlamydomonadales.

Indivíduos representantes da ordem Sphaeropleales, apresentam células vegetativas não móveis, cocóides, unicelulares ou colônias produtoras de zoósporos biflagelados com orientação basal diretamente oposta, do tipo 12/6 (DO) (Lewis & McCourt 2004).

A ordem Chlamydomonadales é formada por indivíduos unicelulares e coloniais com células vegetativas biflageladas com orientação *clockwise* 1/7 (CW) ou quadriflageladas podendo apresentar variadas conformações. (Nozaki *et al.* 2003).

Oedogoniales agrupa indivíduos exclusivamente filamentosos e apresentam algumas características conspícuas, como a citocinese, onde a parede celular se racha para o alongamento e a deposição do material de parede forma a nova parede celular, deixando estrias características da ordem. A conformação estefanoconte de seus

múltiplos flagelos é também uma característica única em toda a classe. (Hoek *et al.* 1995).

Chaetophorales tem divisão celular peculiar, onde ao final desta, as células filhas não estão completamente separadas e permanecem conectadas entre si, formando um organismo multicelular, taloso, filamentoso, ramificado ou não, e heterotrícos. Suas células móveis, zoósporos, são quadriflageladas com disposição das bases dos pares de flagelos DO/CW (Lokhorst *et al.* 1984; Watanabe *et al.* 1989)

Os organismos da ordem Chaetopeltidales quadriflagelados, entretanto apresentam somente orientação DO das bases dos dois pares de flagelos. Os zoósporos são recobertos por escamas, característica da ordem e única na classe.

A divisão em ordens e famílias desta classe tende a ser bem definida, a partir de características mais facilmente observáveis, como formação de colônias, formato celular, presença ou não de flagelos. Entretanto o delineamento de gêneros e espécies é uma tarefa extremamente trabalhosa, devido à semelhança morfológica e à plasticidade fenotípica destes organismos, agravada pelo fato de muitas não terem estruturas celulares óbvias, pela ocorrência de espécies crípticas e ocorrência de convergência morfológica facilitada pelo hábito unicelular, entre outros (Fučíková *et al.* 2011) (Bickford *et al.* 2007; Fawley *et al.* 2006).

Outros fatores dificultam a identificação desses organismos, como a observação dificultada de estruturas deterioradas, ciclos de vida heteromórficos e a variabilidade morfológica dentro das espécies, como heterotalos concomitantes no ambiente e nem sempre detectados, exigindo o olhar de um taxonomista experiente.

Inovações em técnicas de cultivo e observação por microscópio eletrônico podem revelar novos organismos ou até mesmo, novas formas de organismos já antes identificados que não puderam ser prontamente distinguíveis pela morfologia, como é o

caso das "clorococáceas". Entretanto, a microscopia eletrônica é uma tecnologia cara e muitas vezes com acesso restrito.

Ainda assim, há casos de gênero polifiléticos, onde espécies do mesmo gênero podem ser alocadas em diferentes classes quanto aos fatores de classificação acima citados. (e.g. *Neochloris*, com características que o enquadram em Chlorophyceae, Trebouxiophyceae e Ulvophyceae) (Lewis & McCourt 2004).

A precisa identificação destes organismos é indispensável para conhecimento de sua biodiversidade e levantamento da filogenia do grupo. Com o continuado aumento do emprego destes organismos na biotecnologia, é essencial, também, para que se tenha conhecimento de espécies mais eficientes na produção dos bioprodutos desejados, espécies que tenham bom desempenho quando cultivadas em ambientes adversos e que apresentem rápido crescimento visando maior rendimento da produção. É importante também para conhecimento de espécies nativas, evitando a introdução espécies exóticas no ambiente que podem causar desequilíbrio e problemas em estudos de biodiversidade, para evitar impactos ambientais e até para contenção da liberação de produtos tóxicos por microalgas em corpos d'água utilizados para abastecimento humano.

Desta maneira, é fundamental a busca por métodos que possam resolver os problemas de identificação tão comuns em Chlorophyceae de maneira rápida e precisa.

Uma técnica é a obtenção de marcadores moleculares que juntamente com métodos taxonômicos clássicos compõe uma abordagem polifásica de identificação.

Alguns marcadores moleculares já são aplicados eficientemente na distinção de espécies para alguns grupos de organismos. A extremidade 5' do Cox1, por exemplo, é um marcador oficial utilizado para insetos, aves, peixes, algas pardas (Saunders & McDevit 2012), algas vermelhas (Le Gall & Saunders 2010) e diatomáceas.

No entanto, em Viridiplantae este gene não é aplicável, devido a sua natureza conservada neste grupo (Fazekas *et al.* 2008) e *introns* já foram reportados em algas verdes (Turmel *et al.* 2002). Existem alguns marcadores moleculares propostos para aplicação em plantas terrestres, como os loci de genes codificantes *matK*, *rbcL*, *rpoB*, *rpoC1*, e regiões espaçadores de genes não codificantes, como *atpF-atpH*, *trnH-psbA* e *psbK-psbI*, todos plastidiais. Estes loci são muitas vezes combinados entre si, de modo a alcançar o melhor desempenho como identificador. Assim, dos marcadores propostos, os que têm se mostrado mais apropriados como marcadores moleculares para plantas terrestres são o *rbcL*, que apresenta universalidade e seqüências com boa qualidade, aplicado juntamente ao marcador *matK*, que apresenta boa resolução específica (CBOL Plant Working Group *et al.* 2009). Entretanto o marcador molecular *matK* está ausente em Chlorophyta (Pombert, 2005).

As algas verdes são uma linhagem antiga e muito diversificada (Yoon *et al.* 2004), ainda com pouca informação sobre seu genoma. No *GenBank* (National Center for Biotechnology Information) existem apenas oito genomas completos de Chlorophyceae, sendo todas da ordem Chlamydomonadales o que dificulta a busca por marcadores moleculares que possam ser aplicados como *DNA-barcode*.

Alguns genes vêm sendo freqüentemente usados em trabalhos de modo mais focado nos táxons internos de Chlorophyceae, investigando seu uso como *DNA barcode* ou em análises filogenéticas mostrando bons resultados.

Genes plastidiais são muito explorados, pois suas características favorecem seu uso como *barcode*, como a origem uniparental. O gene *rbcL*, já em uso com plantas terrestres é codificante para a maior subunidade da enzima rubisco e está presente no genoma plastidial de todos os eucariontes fotossintetizantes.

Este gene figura em muitos estudos com Chlorophyceae combinado geralmente com genes nucleares para esclarecer relações filogenéticas (Nozaki, 2000; Hoham *et al.* 2002; Assunção *et al.* 2011). Por ser codificante tem poucos *indels* (pequenas inserções ou deleções), porém a ocorrência de alguns *introns* é conhecida entre Chlorophyceae (Nozaki *et al.* 2002). A extremidade 3' deste gene é variável o suficiente para discriminação específica em Ulvophyceae (Saunders & Kucera 2010) o que o coloca como bom candidato para uso como *barcode*. Existem muitas seqüências disponíveis no *GenBank* de Chlorophyceae, sendo de diversas regiões componentes ou deste gene completo, que contém cerca de 1400pb.

O gene *tufA* também faz parte do genoma do cloroplasto em Chlorophyceae, mas em Charophyceae e plantas terrestres sofreu transferência para o núcleo, mantendo em alguns casos ambas as cópias nuclear e plastidial (Baldauf *et al.* 1990). Tal fato pode dificultar seu uso como marcador *barcode* nestes grupos, devido à possibilidade de haver pseudogenes e à homologia. Entretanto em Chlorophyceae está localizado apenas no genoma plastidial e *introns* são desconhecidos em algas verdes (Nozaki *et al.* 2002)

O *tufA* é codificante para uma molécula que faz a mediação da entrada de um amino-acil-tRNA no sítio aceptor ribossomal durante a síntese protéica, ditando a elongação da cadeia peptídica a ser formada, e devido à regulação, é um gene conservado (Delwiche *et al.* 1995). Este gene mostrou bom sinal filogenético possibilitando resolver questões de filogenia (Famà *et al.* 2002; O'Kelly *et al.* 2004), apresentando taxa intermediária de evolução (Saez *et al.* 2008), entretanto resolveu melhor questões supra-genéricas (Ludwig *et al.* 1990) (Delwiche *et al.* 1995).

Outro gene também constantemente explorado em estudos de filogenia com diversos organismos é o espaçador interno nuclear (ITS) (Buchheim *et al.* 2012; Moniz & Kaczmarska 2010). Seu uso para identificação em fungos é anterior ao conceito de

barcode (Hebert *et al.* 2003) sendo hoje reconhecido como tal para fungos (Seifert, 2009; Schoch *et al.* 2012).

A aplicação de marcadores do gene ITS em Chlorophyceae, portanto, acrescenta a cerca de sua universalidade. Pelo fato de ser um espaçador, sofre menor pressão evolutiva, sendo altamente variável e seu uso é adequado para resolver questões em nível específico em algas verdes (O’Kelly *et al.* 2010; Coleman, 2007; Verbruggen *et al.* 2006).

Na busca de marcadores moleculares para DNA *barcode* é conveniente trabalhar com genes que tenham bom sinal filogenético, ou seja, que remetam a filogenia dos grupos estudados, mas que também possibilitem a discriminação específica (apresente a variação genética necessária dentro das espécies). O sinal filogenético do marcador descarta a criação de uma seqüência para cada organismo catalogado, o que inviabilizaria seu uso em amostras ambientais, das quais freqüentemente não se tem conhecimento de quais táxons pertencem os organismos ali presentes e muitos não são facilmente cultiváveis para que se possa então trabalhar em catalogá-los.

Assim, trabalhar com marcadores moleculares com sinal filogenético possibilitaria o enquadramento do organismo desconhecido próximo de taxa semelhantes, facilitando a identificação.

Somado a praticidade de se trabalhar com marcadores moleculares, o seqüenciamento de nova geração (NGS do inglês Next-Generation Sequencing) é uma tecnologia nova que pode ser empregada globalmente, pois pode ser aplicada com qualquer produto de PCR, independente do marcador ou *primers* utilizados, provendo informações moleculares do *pool* de uma população (Fumagalli *et al.* 2014; Salipante *et al.* 2013; Reyes *et al.* 2012) e conseqüentemente sendo aplicável em amostras ambientais (Hajibabaei *et al.* 2011).

O seqüenciamento de diferentes fragmentos de uma só vez permite a identificação dos mais variados organismos, podendo ser pertencentes à Archea, Bacteria ou eukaryotos (Eiler *et al.* 2013), organismos como nano ou picoplanctônicos, muitas vezes não observáveis pela taxonomia clássica.

Portanto, os marcadores moleculares combinados ao poder massivo do NGS, tornam a identificação e catalogamento dos organismos ainda mais eficiente.

Todavia, os genes candidatos a marcadores para Chlorophyceae, como *rbcL*, ITS, *tufA*, entre outros, devem seguir as premissas básicas previstas por (Hebert *et al.* 2003) para que possam ser aplicados como identificador molecular.

Seguindo esses requisitos, a região marcadora deve apresentar variação suficiente para discriminação específica. Esta discriminação é verificada pela existência do *Barcode Gap*, a partir de medições de distâncias genéticas onde a maior distância genética intra-específica não deve sobrepor a menor distância interespecífica.

Esta região marcadora deve também ser flanqueada por regiões conservadas para desenvolvimento dos *primers* para sua amplificação. Estes *primers* devem gerar seqüências bidirecionais de boa qualidade que possam ser aplicados universalmente recuperando as regiões marcadoras do maior número possível de taxa.

Desta maneira o primeiro passo é a obtenção de *primers* que possam amplificar as regiões marcadoras da maneira mais abrangente possível. Entretanto há pouca iniciativa de testar a universalidade de *primers* que tenham mostrado previamente bons resultados.

Objetivo

Tendo em vista as dificuldades de identificação dos organismos de Chlorophyceae acima expostas e a falta de pares de *primers* que possam recuperar toda a classe, este estudo teve como objetivo testar *primers* para marcadores moleculares disponíveis na literatura e já aplicados em outros grupos de organismos de modo encontrar candidatos universais para a classe Chlorophyceae.

Secundariamente foram avaliadas as seqüências geradas com os *primers* testados, buscando por candidatos a aplicação em reconstruções filogenéticas e futuras aplicações como DNA *barcode*.

Materiais e Métodos

Organismos

Os organismos utilizados neste trabalho são representantes da classe Chlorophyceae, todos de água doce. Foi utilizado apenas um representante da classe Trebouxiophyceae servindo como *outgroup* deste estudo. Todos os organismos se encontram em cultivo na Coleção de Culturas de Microalgas no laboratório de ficologia da Universidade Federal de São Carlos, (Centro de Dados Mundial de Microorganismos, WDCM, n° 835)

Segundo o banco de algas internacional Algaebase, as cepas utilizadas estão classificadas da seguinte maneira:

Classe Chlorophyceae

Ordem Sphaeropleales

Família Selenastraceae: *Ankistrodesmus densus* Korshikov 1953 (cepas 003, 128, 239), *Monoraphidium komarkovae* Nygaard 1979 (cepa 353), *Selenastrum bibraianum* Reinsch 1866 (cepa 350), *Selenastrum gracile* Reinsch 1866 (005 e 047),

Kirchneriella aperta Teiling 1912 (cepa 123), *Raphidoceles subcaptata* (Korshikov)
Nygaard, Komárek, J.Kristiansen & O.M.Skulberg 1987 (cepa 048).

Família Scenedesmaceae: *Scenedesmus* sp (cepa 439), *Scenedesmus ecornis*
(Ehrenberg) Chodat (cepa 088), *Scenedesmus spinosus* Chodat, 1913 (cepas 046, 062),
Coelastrum sphaericum Nägeli 1849 (cepa 060), *Desmodesmus communis*
(E.Hegewald) E.Hegewald 2000 (cepa 030).

Família Hydrodictyaceae: *Pediastrum duplex* Meyen 1829 (cepa 055).

Ordem Chlamydomonadales:

Família Volvocaceae: *Volvox* sp (cepa 193), *Pandorina morum* (O.F.Müller)
Bory de Saint-Vincent in Lamouroux, Bory de Saint-Vincent & Deslongschamps 1824
(cepa 095).

Família Chlamydomonadaceae: *Chlamydomonas chlorastera* Ettl 1968 (cepa
009).

Ordem Chaetophorales: *Chaetophora* sp. (cepa 552).

Ordem Oedogoniales: *Oedogonium* sp (cepa 570).

Classe Trebouxiophyceae

Ordem Chlorelalles

Família Oocystacea: *Nephrocytium lunatum* West 1892 (cepa 065)

Fotografias identificações

A documentação das diferenciações morfológicas foi feita por fotografias obtidas com microscopia óptica (Axioplan 2 Imaging com câmera acoplada AxioCam – Zeiss), com o auxílio de programa para análise de imagens e medidas de escala AxioVision 4.6 - Zeiss. As fotos foram cortadas (4X6 cm) utilizando o programa Picasa versão 3.9.137.

Cultivo

Os organismos foram isolados e axenizados, cultivados em *erlenmeyers* de 100 ml em meio WC (Guillard & Lorenzen 1972), em pH 7,0 e temperatura de $25\pm 1^{\circ}\text{C}$, sob intensidade luminosa de $300\ \mu\text{mol photons m}^{-2}\cdot\text{s}^{-1}$ e fotoperíodo claro/escuro de 12/12h.

As cepas das filamentosas Chaetophorales e Oedogoniales não puderam ser axenizadas por plaqueamento ou diluição e, portanto, foram lavadas antes da extração.

O método de lavagem consistiu em consecutivas centrifugações e adições de meio de cultivo limpo às células, sendo que no primeiro passo de lavagem a ressuspensão das células foi feita com solução de cloreto de lítio.

O cloreto de lítio tem ação caotrópica desestabilizando a mucilagem e em maiores concentrações desestabiliza a membrana celular, podendo levar a célula a morte (Nordi *et al.* 2006). A concentração utilizada foi de 1M, desejando eliminar o máximo possível da mucilagem e conseqüentemente possíveis fungos e bactérias aderidos a célula algal sem causar danos a esta.

As cepas de *N. lunatum* e *P. morum* são coloniais, caracterizadas por ter um espesso envoltório de polissacarídeo, o que após a extração do DNA pode deixar resquícios e dificultar a reação de amplificação. A fim de evitar isso, estas cepas também foram lavadas com cloreto de lítio.

Quando em fase de crescimento exponencial, que foi determinada por densidade ótica celular, os cultivos foram concentrados em *pellets* de 40-60mg em centrífuga *Eppendorf* 5415D a 3500g para extração do DNA.

Extração de DNA

O material concentrado foi homogeneizado com pérolas de vidro de 0,5mm de diâmetro inseridas nos tubos sobre os *pellets* e estes foram levados ao vórtex por cerca de 15 segundos. O DNA foi extraído com kit de extração *Invisorb Spin Plant Mini Kit*.

A extração de DNA da cepa de *C. chlorastera* com passo de extração mecânica para quebra da parede celular (pérolas de vidro) apresentou rastro no gel de extração, possivelmente indicando que houve degradação de DNA. As células de *Chlamydomonas* são células mais frágeis por não possuírem parede celular celulósica típica, mas sim uma cobertura glicoprotéica denominada *teca* e também devido ao formato de célula. Após essa medida, a extração desta cepa deu-se da maneira esperada e a amplificação pode, portanto, ser verificada.

A cepa de *Oedogonium sp* por outro lado apenas mostrou extração satisfatória de material genético quando no passo de homogeneização com as pérolas de vidro foi macerada com bastão de modo a romper mecanicamente sua espessa parede celular.

Após a extração, o material foi submetido à eletroforese em gel de agarose 1% em tampão TRIS-Borato-EDTA 1x (TBE), corado com *GelRed* (*GelRed Nucleic Acid Gel Stain-Biotium*, Hayward, CA, EUA) diluído em água na proporção 1:500 para verificação da qualidade da extração.

O material extraído foi separado em tubos de 1,5 ml para uso e armazenamento em freezers -20°C e -80°C respectivamente.

Amplificação do DNA

Para a reação de amplificação foi preparada solução dos reagentes necessários à PCR com as seguintes concentrações:

- 1 ul de cada *primer* (F e R) com concentração de 0,25 uM;
- 4 ul de solução de dNTP com concentração final de 0,2 mM cada,
- 0,75 ul de MgCl com concentração final de 1,5 mM,
- 2,5 u de tampão 10X de PCR sem Mg com concentração final de 1X,
- 1 ul de BSA (albumina bovina) (10 mg/ml),
- 0,2 ul de Taq (Taq DNA Polymerase, Recombinant, 5U/ul, Invitrogen)
- Quantidade variável de DNA (concentração final de 5-10 ng)

A solução de amplificação foi completada com água ultrapura até o volume final de 25ul.

Os programas usados no termociclador foram:

- Desnaturação inicial a 94°C por 4 minutos,
 - 45 segundos a 94°C,
 - 30 segundos T°C específica
 - 45 segundos de extensão a 72°C
- } 29 ciclos
- Extensão final de 10 minutos a 72°C.

As temperaturas de anelamento específicas para cada par de *primers* foram:
tufA: 55°C, ITS: 56°C, *rbcLFP*: 50-55°C, *GrbcL*: 50°C.

Cada cepa foi amplificada em duplicata e para cada reação de amplificação foi incluído um controle negativo (todos os reagentes menos o DNA) para garantir que não houvesse contaminação.

Foi verificado o sucesso de amplificação por eletroforese em gel de agarose 1% em tampão TBE 1x, a 90V por 35 minutos. Foi utilizado marcador de tamanho (1Kb DNA Ladder-Invitrogen, Carlsbad, CA, EUA) para verificar os tamanhos dos fragmentos amplificados, corados com corante *GelRed* na concentração 1:500.

Para todas as cepas que não mostraram amplificação alguns parâmetros foram alterados nas tentativas seguintes, sempre mantendo um controle positivo (uma cepa que tivesse sido prontamente amplificada com o par de *primers* em questão). Para tanto foram testados gradientes de temperatura de anelamento, aumento da concentração de magnésio, tempo de extensão, porém não foram eficientes para que houvesse amplificação de cepas que não haviam sido amplificadas com os parâmetros iniciais.

Entretanto a diminuição no número de ciclos do programa – de 35 para 29 - usado no termociclador otimizou a amplificação para todos os *primers*, principalmente em relação a intensidade da banda referente ao fragmento esperado.

Portanto, foi considerada ausência de amplificação quando não foram observadas bandas no gel de amplificação mesmo após os ajustes citados acima e ajustes na concentração de DNA extraído na reação de amplificação.

As extrações e amplificações foram quantificadas no fotodocumentador *BioRad Universal Hood*, pelo recurso Ferramentas de Quantificação, do software *ImageLab 4.0 build 17*.

O produto de PCR foi purificado manualmente usando solução de PEG (Polietilenoglicol) 20% e NaCl 1M (Lis and Schleif 1975) seguindo os seguinte protocolo (Bagatini, 2008):

Precipitação do DNA:

1. Acrescentar igual volume de solução de PEG e amostra: homogeneizar imediatamente
2. Levar a banho Maria por 37°-40°C por 15 minutos

3. Centrifugar a 16000 xg por 15 minutos

Lavagem do material

4. Retirar o sobrenadante com uma pipeta. Descartar o sobrenadante
5. Adicionar lentamente, escorrendo pela parede do tubo 125 ul de etanol 80% gelado.
6. Centrifugar a 16000 xg por 2 minutos.
7. Retirar lentamente o sobrenadante com uma pipeta. Descartar o sobrenadante.
8. Repetir passos 5 a 7.

Secagem

9. Secar em *estufa* a 37°C por aproximadamente 30 minutos a 1 hora. Olhar periodicamente para não deixar os tubos secarem em excesso. Ou Secar em banho Maria aberto a 50°C por 5 minutos
10. Adicionar água ou tampão 12 a 20 ul.

O seqüenciamento foi conduzido pela *Macrogen* (Seul-Korea).

Primers

Os *primers* testados neste estudo foram escolhidos na literatura de acordo com o grupo de organismos estudados pelos autores. Desta maneira, foram escolhidos aqueles já testados em Chlorophyceae ou outras classes de Chlorophyta. Foram testados *primers* para marcadores moleculares plastidiais como *tufA* e *rbcL* e nuclear, como o ITS (tabela 1) (**figura 1**).

Os pares escolhidos foram previamente alinhados com seqüências de Chlorophyceae que contivessem os genes alvo obtidas no *GenBank* de modo a verificar se seria possível o anelamento destes *primers* e estimar o tamanho do fragmento a ser gerado.

Tabela 1- Marcadores moleculares, nomes, trabalho de referência e seqüências dos *primers* testados (F=forward, R=reverse), tamanho esperado do fragmento em pares de bases (bp) e temperatura de anelamento (T°C anel) de cada par de *primer*.

Marcador molecular	<i>Primer</i>	Referência	Sequencia 5' → 3'	Tamanho frag/o (pb)	T°C anel
<i>rbcL</i>	<i>rbcL</i> -M379 F*	(McManus & Lewis, 2011)	GGTTTCAAAGCTYTWCGTGC	550	50 - 55
	<i>rbcLFP</i> R*	(Designed in this study)	GTAAATACCACGGCTACGRTCTT		
<i>rbcL</i>	<i>GrbcL</i> F	(Saunders & Kucera, 2010)	GCTGGWGTAAGATTAYCG	500 - 581	50
	<i>GrbcLR</i>		TCACGCCAACGCATRAASGG		
<i>Rpl5 – rpl14</i>	UCP4 F	(Provan et al.2004)	ACGATCTAAAAMGCATACAT	367 - 421	54
	UCP4 R		AATTGTWTCDDTDGCACCDGAAG		
<i>tufA</i>	<i>tufA</i> F	(Fama et al. 2002)	GGNGCNGCNCAAATGGAYGG	780	55
	<i>tufA</i> R		CCTTCNCGAATMGCRAAWCGC		
ITS	ITS5 F	(White et al. 1990)	GGAAGTAAAAGTCGTAACAAGG	574 - 601	56
	ITS5 R		TCCTCCGCTTATTGATATGC		

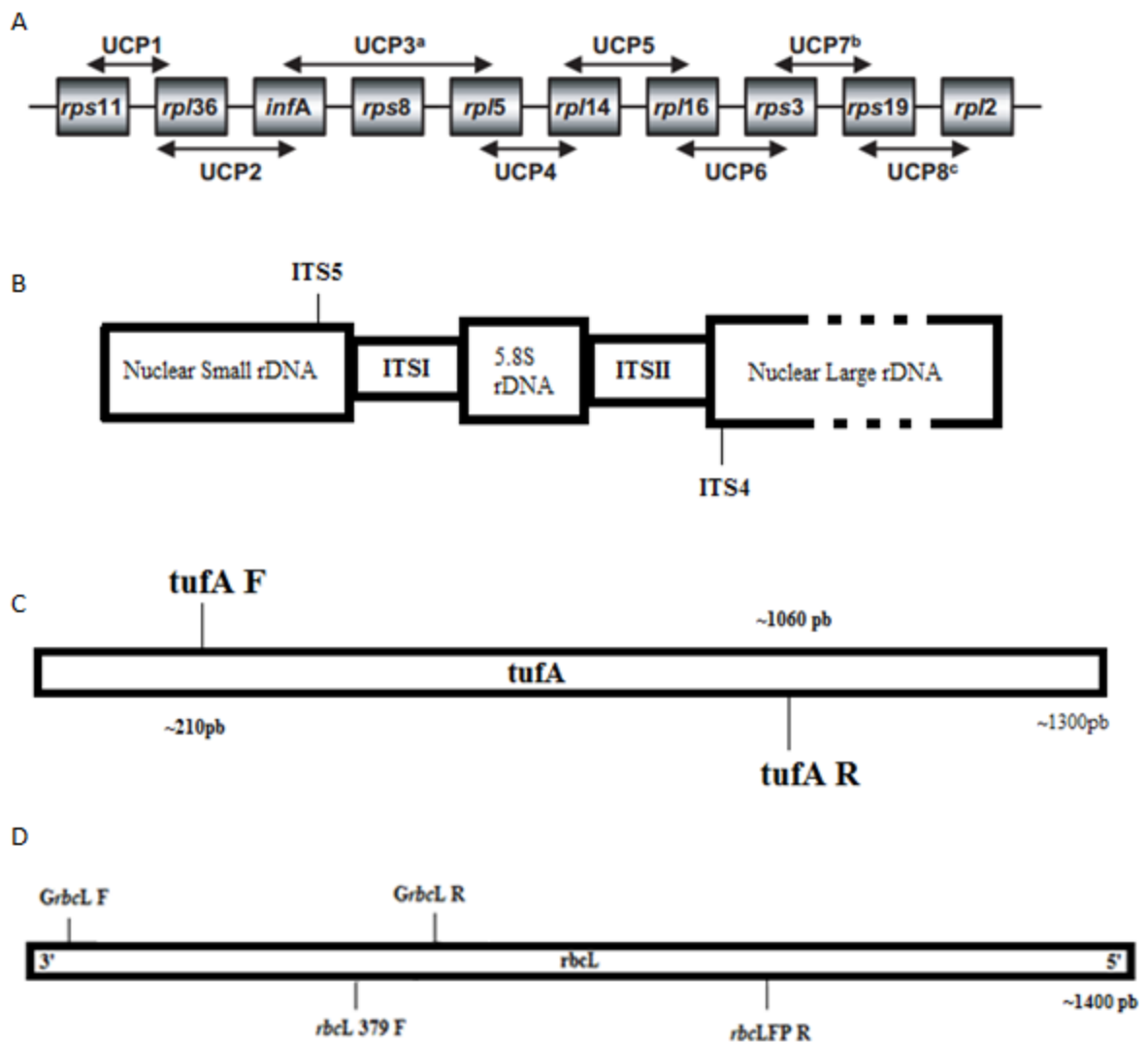


Figura 1 – Representações das regiões do gene amplificadas por cada par de *primer* testado. A – Cluster de gene plastídias onde se encontra a região amplificada pelo par de *primers* UCP4. Retirado de Provan, 2004. B – Região amplificada pelo par de *primers* ITS5- ITS4. C – Região do gene *tufA* amplificada pelo par de *primers* *tufA*. D – Regiões do gene *rbcL* amplificadas pelos dois pares testados: *GrbcL* e *rbcLFP*.

O sucesso de amplificação dos *primers* foi determinado com base no número de cepas amplificadas por cada par de *primers* testado e número de seqüências obtidas com boa qualidade.

Análise das Sequências

Os cromatogramas foram submetidos primeiramente à análise de *base-calling* no programa *Phred* (Ewing *et al.* 1998), implementado no programa *Geneious* version 6.1.7 (Biomatters), onde as seqüências que possuíam valores de qualidade (HQ%) inferiores a 40% foram descartadas, assim como seqüências com picos duplos e grande quantidade de ruídos. Quando possível foram trabalhadas as *consensus* das seqüências *forward* e *reverse* obtidas. O alinhamento das seqüências foi realizado com o *software Clustal W* (Thompson *et al.* 1994) também implementado no *software Geneious* v.6.1.7, assim como a edição das mesmas.

As seqüências geradas foram confrontadas com as depositadas no *GenBank* utilizando a ferramenta de busca de seqüências por alinhamento *Basic Local Alignment Search Tool* (BLAST) (Altschul *et al.* 1997) para verificar se correspondiam às linhagens esperadas ou ao menos táxons próximos e se não havia contaminação.

O *Frame* de leitura da proteína foi investigado no *ORF finder* ambos no *website National Center for Biotechnology and Information* (NCBI; Bethesda, MD, USA). A tradução e análises de *gaps*, *in/del* e *stop condóns* foram realizadas no *Geneious* v.6.1.7.

No *software Mega6* (Tamura *et al.* 2013) foram calculadas as distâncias genéticas (*p-distance*) par a par, entre famílias, entre ordens e entre classes, assim como a composição de nucleotídeos, composição de aminoácidos e “*codon usage bias*”. Dados de polimorfismo como sítios polimórficos, sítios parsimoniamente informativos, número de códons, substituições sinônimas e não sinônimas foram calculados no *DNAsp* 5.10 (Librado & Rozas 2009).

O modelo evolutivo foi obtido a partir do programa *jModelTest* v.0.1.1 (Darriba *et al.* 2012). Quando presentes, os *gaps* foram considerados deleção completa.

Para obtenção da árvore baseada no algoritmo Neighbour-Joining (*p-distance*) foram usados 1000 *bootstraps* no *Mega6* com completa deleção. A árvore de inferência bayesiana foi implementada no Mr. Bayes (Huelsenbeck & Ronquist, 2001) sob o modelo evolutivo GTR+I+G, utilizando MCMC com 3.000.000 gerações.

Resultados

Universalidade dos primers testados e rendimento de seqüências

UCP4:

Apesar de ser proposto como *primer* universal para Chlorophyta, o par de *primers* UCP4 não amplificou nenhuma das cepas utilizadas (tabela 2).

O par de *primers* UCP4 falhou para todas as tentativas de amplificação, entretanto não se sabia ao certo como estes *primers* iriam se comportar. (Provan et al. 2004) testa a universalidade de *primers* em Rhodophytas e Chlorophytas. Os *primers* por ele testados recuperam um cluster genético plastidial e o teste da universalidade destes pares em Chlorophyta foi feito com apenas 4 organismos, sendo dois deles pertencentes à classe Ulvophyceae, um da Chlorophyceae e um de Charophyceae.

O par UCP4 foi o escolhido por ter amplificado a Chlorophyceae (*Dunaliella salina*) e por ter gerado fragmentos mais constantes quanto ao número final de pares de base (367 pb a 421 pb) e também ao número de pares de bases não codificantes que se manteve proporcional nestes fragmentos entre indivíduos de Chlorophyta e Charophyta, sugerindo a ausência de *in/dels* e, portanto, maior facilidade para trabalhar como possível marcador.

Entretanto o sucesso de amplificação de apenas um organismo como representante de uma classe que contém mais de 2500 espécies não é suficiente para

afirmar que os *primers* sejam universais para a classe. O mesmo cabe ao teste de universalidade dos outros *primers* testados em seu estudo.

ITS

Os *primers* testados para recuperar a região do ITS amplificaram apenas 5 das 22 cepas testadas. As ampliações foram obtidas sem dificuldades gerando bandas intensas e únicas, porém renderam bom seqüenciamento de apenas 3 destas cepas: *Kirchneriella aperta*, *Desmodesmus communis* e *Pediastrum duplex* (Tabela 2).

O par de *primer* ITS testado neste estudo foi escolhido dentre os propostos por (White, *et al.* 1990) em seu trabalho avaliando genes para filogenia de fungos, mas já testados com organismos de Chlorophyceae (Van Hannen *et al.* 2000; Buchheim *et al.* 2012). Este par recupera as duas regiões espaçadoras ITS1 e ITS2 e também o gene 5.8S.

Por ser uma região espaçadora e, portanto, sofrer menor pressão seletiva, o ITS tende a ser mais variável, podendo apresentar *in/dels* e conseqüentemente tamanhos variáveis entre diferentes taxa. Em algas verdes este espaçador é comumente utilizado em filogenias de nível específico. (O'Kelly *et al.* 2010; Coleman 2007). Ainda assim a região do ITS2 tem melhor resolução específica do que o ITS inteiro (Chen *et al.* 2010). Desta maneira a natureza pouco conservada e o tamanho muito variável do gene ITS pode ter sido a causa do baixo sucesso de amplificação e baixa qualidade das seqüências obtidas neste estudo.

Este mesmo par de *primers* foi utilizado por Van Hannen *et al.* (2000) em um estudo com o gênero *Scenedesmus*. Entretanto no presente estudo não mostrou amplificação para nenhuma cepa do gênero *Scenedesmus* testada.

Tabela 2 – Amplificação e sequenciamento para cada cepa e par de *primer* testado Amp = amplificação, Seq = sequenciamento; (✓) sucesso (-) falha amplificação/sequenciamento; R/F = apenas seqüência reverse/forward com boa qualidade QH>40%).

Cepa	<i>tufA</i>		<i>GrbcL</i>		ITS		<i>rbcL</i> FP	
	Amp	Seq	Amp	Seq	Amp	Seq	Amp	Seq
<i>Ankistrodesmus densus</i> 003	✓	✓	-	-	-	-	-	-
<i>Ankistrodesmus densus</i> 128	✓	✓	-	-	-	-	✓	✓
<i>Ankistrodesmus densus</i> 239	✓	✓	-	-	-	-	-	-
<i>Selenastrum gracilis</i> 005	✓	✓	-	-	✓	-	✓	✓
<i>Selenastrum gracilis</i> 350	✓	✓	✓	-	-	-	✓	✓
<i>Selenastrum bibraianum</i> 047	✓	✓	✓	-	-	-	✓	✓
<i>Raphidoceles subcaptata</i> 048	✓	R	✓	-	-	-	✓	✓
<i>Nephrocytium lunatum</i> 065	✓	R	✓	R	-	-	✓	✓
<i>Kirchneriella aperta</i> 123	✓	✓	✓	✓	✓	✓	-	-
<i>Monoraphidium kormakokae</i> 353	✓	✓	✓	R	-	-	✓	✓
<i>Desmodesmus communis</i> 030	✓	✓	✓	✓	✓	✓	✓	✓
<i>Coelastrum sphaericum</i> 060	✓	✓	✓	-	-	-	-	-
<i>Scenedesmus ecornis</i> 088	✓	✓	-	-	-	-	✓	✓
<i>Scenedesmus</i> sp 439	✓	✓	✓	✓	-	-	✓	✓
<i>Desmodesmus spinosus</i> 046	✓	✓	✓	✓	-	-	✓	✓
<i>Desmodesmus spinosus</i> 062	✓	✓	✓	✓	-	-	✓	✓
<i>Pandorina morum</i> 095	✓	✓	-	-	-	-	✓	✓
<i>Volvox</i> sp. 193	✓	✓	-	-	✓	-	✓	✓
<i>Pediastrum duplex</i> 055	✓	✓	✓	-	✓	✓	✓	✓
<i>Chlamydomonas clorastera</i> 009	✓	✓	-	-	-	-	-	-
<i>Chaetophora</i> sp.	✓	✓	-	-	-	-	-	-
<i>Oedogonium</i> sp. 570	✓	-						

GrbcL

O par de *primers* GrbcL amplificou 12 das 22 linhagens testadas, resultando em bom seqüenciamento (HQ>40%) de 7 cepas: *D. communis*, *M. komarkovae*, *N. lunatum*, *K. aperta*, *Scenedesmus sp* (439), *S. spinosus* (046 e 062) (tabela 2).

As cepas que mostraram maior facilidade na amplificação foram *D. communis*, *S. spinosus* (046 e 062), *Scenedesmus sp* (439) e *K. aperta* que geraram uma única banda intensa de tamanho referente ao fragmento alvo na primeira tentativa de amplificação.

As cepas *C. sphaericum*, *M. komarkovae* e *N. lunatum* foram amplificadas duas vezes para gerar quantidade suficiente de material amplificado, pois na primeira amplificação geraram bandas únicas referentes ao fragmento alvo, porém fracas.

Com exceção das seqüências das cepas de *N. lunatum* e *M. komarkovae* - que só foi possível analisar as seqüências *reverse* - as *consensus* das seqüências obtidas com este par de *primers* resultaram em fragmento de cerca de 500 - 581 bases de extensão.

rbcLFP

Os *primers* rbcLFP amplificaram satisfatoriamente 16 das 22 cepas testadas, com bandas intensas nas primeiras amplificações (tabela 2). Não foram verificadas bandas duplas. O seqüenciamento destas cepas foi satisfatório e o fragmento final das *consensus* das seqüências obtidas com este par de *primers* teve tamanho de 550 bases.

Este par mostrou amplificação com bandas únicas e intensas com temperaturas de anelamento de 50 a 55°C. Entretanto o teste de gradiente da temperatura de anelamento não resultou em sucesso de amplificação para nenhuma cepa que não havia sido até então amplificada.

Duas seqüências obtidas com este par de *primers* apresentaram diferenças em relação às restantes, sendo elas *A. densus* (cepa 128) e *D. communis*. Foram verificadas bandas de fragmentos maiores no gel de amplificação (figura 2) e de fato as seqüências destas cepas apresentaram 584 e 669 bases a mais em relação às outras amplificações com o mesmo par de *primers*.

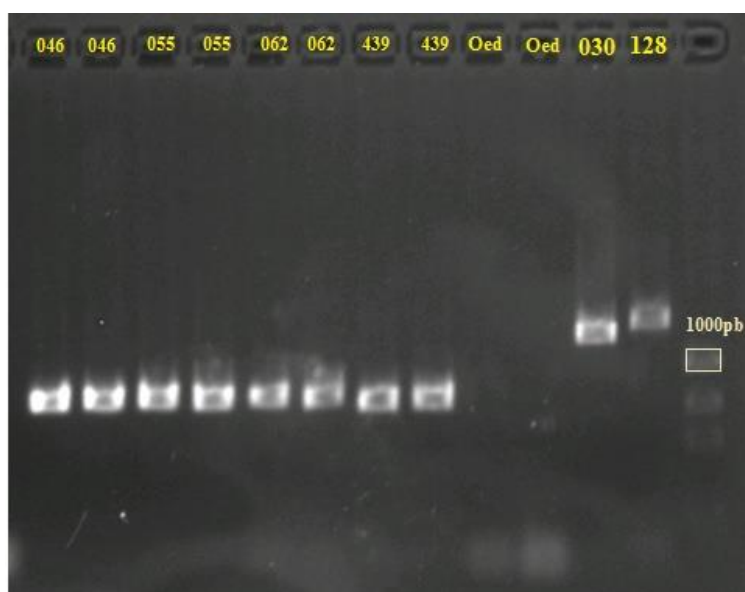


Figura 2 - Gel de amplificação com *primers rbcLFP*. Os oito primeiros poços são cepas que geraram bandas de tamanho esperado (700 a 800 pb). Os poços 9 e 10 são cepas que não foram amplificadas com estes *primers* e os poços 11 e 12 são as cepas *D. communis* e *A. densus*, respectivamente, que geraram bandas com cerca de 1200 pb. O ladder está no poço 13 e a primeira banda do ladder destacada corresponde a 1000 pb.

Os fragmentos finais de *A.densus* e *D.communis*, após o alinhamento, mostraram uma porção de cerca de 800 bases que não se alinharam com as seqüências de referência ou com as seqüências das outras cepas obtidas com o par *rbcLFP* neste estudo.

O fragmento do gene *rbcL* de *A. Densus* 128 obtido com o par de *primers rbcLFP* apresentou total de 1219 bases após a edição e encontra similaridade a partir de sua base 887 com a base 696 da seqüência de *Ankistrodesmus stipitatus* (EF113406)

usada como referência no alinhamento. Como esta cepa foi amplificada com o par de *primers GrbcL* foi possível verificar a posição desta seqüência de bases destoante no alinhamento com as outras cepas que não a apresentaram (figura 3).

O fragmento do gene *rbcL* de *D. communis* obtido com o par *rbcLFP*, resultou em 1134 bases e encontra similaridade com a seqüência de referência *Desmodesmus santosii* (GI281314250) apenas a partir de sua base 654 correspondente a base 801 da seqüência de referencia.

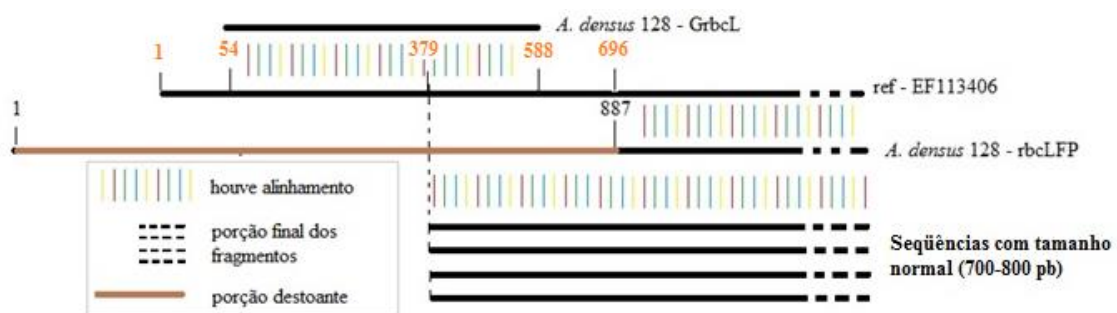


Figura 3 – Diagrama do alinhamento das seqüências obtidas com os pares de *primers rbcLFP* (Seqüências com tamanho normal 700 – 800pb), da seqüência de *A. densus 128* obtida com o par de *primer GrbcL* (*A. densus 128 – GrbcL*), da seqüência de referência de *Ankistrodesmus stipitatus* obtida no *GenBank* (ref – EF113406) e da seqüência de *A. densus 128* obtida com o par de *primer rbcLFP* (*A. densus 128 – rbcLFP*), mostrando em marrom a localização da porção de aproximadamente 800 pb que não se alinhou com a seqüência de referência. Os números em laranja representam as bases da seqüência de referência (EF113406) e os pretos representam as bases da seqüência de *A. densus 128 – rbcLFP*.

É possível que esta porção com ~800 bases corresponda a um *íntron*, cuja ocorrência em Chlorophyceae já foi mostrada neste gene por Nozaki et al. (1998; 2002) ou que se trate de um pseudogene.

Caso esta porção se trate de um *íntron*, isso dificultaria as análises de DNA *barcoding* uma vez que, sendo o *íntron* uma região não expressada, sua taxa de mutação pode ser mais alta, não refletindo a verdadeira distância entre as espécies. Da mesma

forma a amplificação de um pseudogene pode levar a resultados errôneos superestimando o número de espécies existente (Song *et al.* 2008).

Dessa forma independente da origem ou natureza desta porção de ~800pb no gene *rbcL*, o torna inapropriado para identificação de espécies através da abordagem do DNA *Barcoding*. A conclusão sobre o que de fato possa ser essa fração de seqüência de nucleotídeos não comum às outras cepas será possível apenas através de investigações direcionadas especificamente para a situação, o que não foi objetivo deste estudo.

O maior sucesso de amplificação do par *rbcLFP* em relação ao *GrbcL* deve-se ao fato de que o primeiro é composto de iniciadores desenhados a partir de seqüências de Chlorophyceae. O iniciador *forward* do par *rbcLFP* foi retirado de um estudo filogenético com *Pediastrum* (McManus and Lewis 2011) e seu *reverse* desenvolvido neste estudo a partir de seqüências de Chlorophyceae. Já o par *GrbcL* foi desenhado para aplicação com organismos de Ulvophyceae (Saunders & Kucera, 2010).

Saunders & Kucera (2010) testaram regiões das extremidades 3' e 5' do gene *rbcL* obtendo melhor discriminação específica com a região 3', e melhor universalidade com a região 5'. O par de *primers* testado neste estudo foi escolhido com foco na amplitude de aplicação e, portanto, foi testado o par que recuperasse a extremidade 5'. Assim, se este par mostrasse sucesso de amplificação, mas não de discriminação específica também em Chlorophyceae, poderia ser combinado com outros *primers* de outra região marcadora com melhor resolução no nível específico. Porém o baixo sucesso de amplificação em Chlorophyceae e suas seqüências de baixa qualidade levaram à sua exclusão como *primers* universais para este grupo.

Há grande disponibilidade de seqüências no Genbank do gene *rbcL* de organismos da classe Chlorophyceae, podendo ser seqüências com intron, seqüências que contém parcialmente a seqüência codificante, ou seja, diferentes região do gene,

além da possibilidade de serem depositadas seqüências de organismos com classificação equivocada, o que torna impreciso os estudos baseados nestas. Além disso, estas seqüências depositadas foram obtidas com diversos pares de *primers* (Hoham *et al.*2002), muitas vezes desenvolvido para ser trabalhado apenas com um gênero, o que não garante que possa ser aplicado a outros taxa.

tufA

O par de *primers* utilizado para o marcador *tufA* foi o único par que amplificou todas as 22 cepas testadas e com exceção da cepa de *Oedogonium sp* foi possível seqüenciar todas as cepas. De todas as seqüências obtidas com este par de *primers*, *Raphidoceles subcapitata* e *Nephrocytium lunatum* foram as únicas que não apresentaram bom seqüenciamento bidirecional, sendo possível trabalhar apenas suas seqüências *reverse* (tabela 2).

As cepas de *N. lunatum* e *P. morum* mostraram bandas fracas na primeira tentativa de amplificação do gene *tufA*. Entretanto, após a lavagem destas cepas com o cloreto de lítio para remoção do polissacarídeo das colônias seguida do protocolo de extração, a amplificação do gene *tufA* destas cepas resultou em bandas únicas e intensas. Assim, as reações de amplificação com os outros pares de *primers* foram realizadas com o DNA extraído destas células lavadas. Todas as outras cepas foram facilmente amplificadas.

Quando conferidas com seqüências disponíveis nos bancos internacionais através do BLAST, todas as seqüências obtidas neste trabalho tiveram maior similaridade com linhagens dentro de Chlorophyceae, muitas vezes com o mesmo gênero. É importante ressaltar que não há disponíveis seqüências de todos os marcadores para todas as espécies, sendo computada, portanto, a maior similaridade com as seqüências disponíveis.

As seqüências de *tufA* obtidas não apresentaram *gaps* após o alinhamento e o fragmento final teve após edição 743 pares de base, dos quais 305 sítios foram invariáveis. Os 438 sítios restantes mostraram-se polimórficos com total de 716 mutações e 364 sítios foram parsimoniamente informativos.

A região amplificada corresponde à seqüência codificadora integral, com 247 códon, apresentando apenas um *stop* códon, correspondente ao final do *frame* da leitura da proteína. O número de sítios com mutações sinônimas foi de 172,26 e de mutações não sinônimas foi de 568,74.

O modelo evolutivo de substituição nucleotídica selecionado foi o GTR+I+G, onde se assume o modelo Geral de Reversão ao Longo do Tempo (GTR) com uma proporção de sítios Invariáveis (I) e distribuição Gamma (G) para acomodar as diferentes taxas para sítios variáveis.

As seqüências foram depositadas no *GenBank* sob o número de acesso (ainda não depositadas).

O gene *tufa* já foi amplificado e seqüenciado para diversos grupos de organismos como plasmodium, cianobactérias, bactérias, outras classes de Chlorophyta e plantas terrestres, e destes, o grupo que apresenta maior número de seqüência disponíveis no *GenBank* são as bactérias.

Além do sucesso de amplificação e seqüenciamento mostrado por Saunders & Kucera (2010), este mesmo par de *primer* utilizado no presente trabalho é amplamente utilizado em Ulvophyceae, (Famà *et al.* 2002; O’Kelly *et al.* 2004; Wynne *et al.* 2009; Lawton *et al.* 2013). Combinado com *rbcL*, permitiu a distinção e sugestão de novas espécies de *Codium* (Verbruggen *et al.* 2012) e combinado com genes nucleares, além do *rbcL*, foi possível esclarecer o delineamento de alguns gêneros de Sphaeropleales

(Fučíková et al. 2014) e até sugeriu a ressurgência do gênero *Chromochloris* (Fučíková et al. 2012).

Foi aplicado também em diferentes grupos como macroalgas verdes (Du et al. 2014), criptofitas, (Garcia-Cuetos et al. 2010) e na identificação das microalgas fontes dos cloroplastos presentes no trato digestivo de gastrópodes (Christa et al. 2013). Entretanto estes mesmos *primers* falharam quando testado em Cladophoraceae (Saunders & Kucera, 2010).

Visto que apenas o par de *primers tufA* amplificou todas as cepas testadas as análises filogenéticas foram conduzidas apenas com as seqüências obtidas com ele.

Análises filogenéticas

Para análises de filogenia foram incluídas seqüências do *GenBank* de modo a melhorar a representação da ordem Chaetophorales (*Schizomeris leibleinii* 336169554), representar a ordem Chaetopeltidales, (*Floydiella terrestris* 304322936) com a qual não foi possível trabalhar com nenhum organismo pertencente por não dispormos de nenhuma cepa isolada em nosso banco, representar a ordem Oedogoniales, (*Oedogonium cardiacum* EF587375), por não ter sido possível gerar seqüência com o par de *primers* para o gene *tufA* e uma seqüência de *Ostreococcus tauri* (112806818), um representante da classe Mamiellophyceae para atuar como *outgroup*.

A árvore filogenética elaborada com as seqüências do gene mostra a monofila da classe Chlorophyceae e sua subdivisão nos clados SC e OCC remetendo à evolução do aparato flagelar (figura 4 e 5).

O clado SC com as ordens Sphaeropleales e Chlamydomonadales e o clado OCC foram mostrados com valor de suporte máximo na análise Bayesiana (figura 4 e 5).

As ordens Sphaeropleales (grupo DO) e Chlamydomonadales foram representadas com valor de suporte de 0,63 (análise Bayesiana). A monofilia de Sphaeropleales com altos valores de suporte também foi verificada com dados da estrutura secundária de ITS2 (Keller *et al.* 2008), mais conservada que a seqüência de nucleotídeos do mesmo gene (Coleman 2007).

A divisão das famílias Scenedesmaceae e Selenastraceae mostrou valor máximo Bayesiano de suporte (figuras 4 e 5) A forte sustentação da família Selenastraceae foi também observada por Krienitz *et al.* (2001). Entretanto houve sobreposição de alguns gêneros dentro dessas famílias, como é o caso de *Selenastrum*, em Selenastraceae.

As três cepas do gênero *Ankistrodesmus* foram agrupadas com valor máximo para ambas as análises. Este gênero forma um grupo coeso morfológicamente e facilmente distinguível dos outros gêneros amostrados (Figura 6).

Os gêneros de *Selenastrum*, *Raphidocelis* e *Monoraphidium* foram sobrepostos. Tais sobreposições de gêneros dentro das famílias de Sphaeropleales, principalmente de Selenastraceae já foram reportadas em estudos de filogenia com dados moleculares. (Fawley *et al.* 2006; Krienitz & Bock 2012).

Devido à dificuldade de delineamento das espécies pelo conceito morfológico é compreensível que os dados genéticos de ramos mais internos não se comportem coerentemente com a morfologia. A grande maioria destes organismos não tem reprodução sexuada, ou seja, a reprodução assexuada favorece mutações que não afetam o crescimento ou a capacidade do mutante competir, gerando uma grande variabilidade morfológica (Krienitz & Bock 2012). Deste modo, a discriminação é dificultada devido à continuidade dos critérios de diferenciação morfológica, não sendo precisos os limites entre espécies.

Ainda, a morfologia pode se mostrar polifilética, ou seja, organismos morfologicamente parecidos apresentam as diferenças genéticas observadas. (Krienitz *et al.* 2011) provou que a morfologia de foice, visível em Selenastraceae evoluiu polifileticamente em Chlorophyceae e Trebouxiophyceae.

Bock *et al.* (2010) e Darienko *et al.* (2010) mostraram que os genes 18S e 26S são genes muito conservados para separar espécies morfologicamente distintas, mas que são geneticamente próximas, frequente em linhagens de algas verdes.

Entretanto é necessário que se investigue a fundo, preferencialmente com dados de variados genes com poder de discriminação específica e já testados para grupos com classificação estabelecida, para que se possa primeiramente sincronizar as definições morfológicas com os dados genéticos das espécies componentes desta classe.

Estudos de filogenia com organismos de Chlorophyceae combinam mais de um gene, para que se tenha a melhor resolução possível das relações filogenéticas. Opta-se por genes mais conservados, como 18S, 26S, *rbcL*, que possam resolver níveis filogenéticos superiores, e genes menos conservados, como o ITS e alguns genes plastidiais, para relações mais internas entre os taxa estudados, como a monofilia a ordem Sphaeropleales.

Como já visto, a monofilia de Sphaeropleales foi bem sustentada, assim como as famílias representadas neste estudo.

Combinado ao gene 26S, o gene 18S resolveu com bom suporte os clados SC e OCC (Turmel *et al.* 2008), mas devido ao baixo sinal filogenético destes genes, não resolveram com clareza os ramos mais internos destes clados (Buchheim *et al.* 2012).

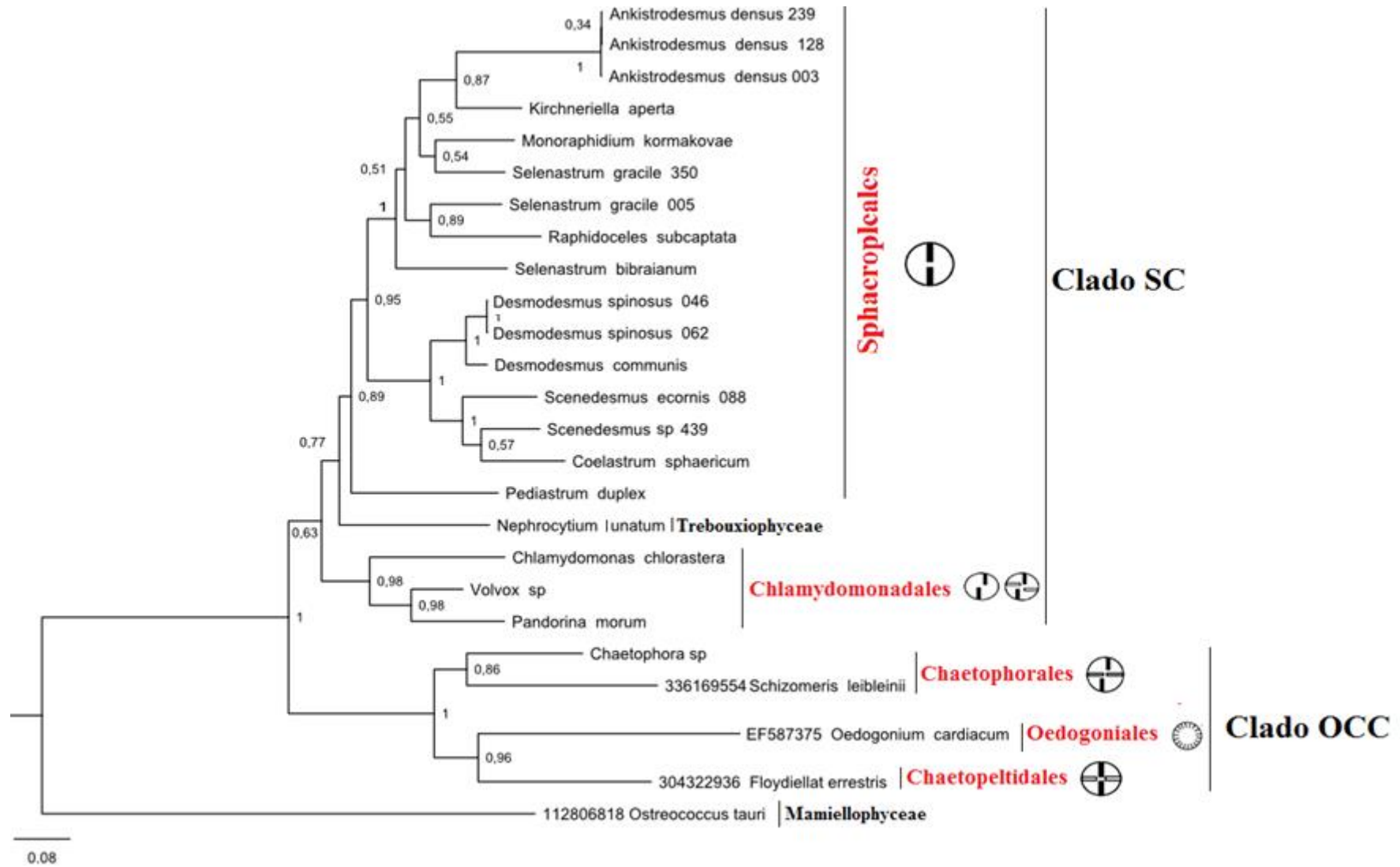


Figura 4 – Análise Bayesiana das seqüências obtidas do marcador *tufA* mostrando o agrupamentos dos clado OCC, SC, das ordens e famílias das cepas do banco WDCM nº835 da classe Chlorophyceae. Os valores bayesianos de probabilidade posterior estão indicados nos nós. As seqüências obtidas no *GenBank* estão indicadas pelo numero de acesso à esquerda do nome da espécie. Os desenhos representam a orientação do corpo basal do aparato flagelar como as células estivessem indicadas pelos círculos e fossem vistas de cima.

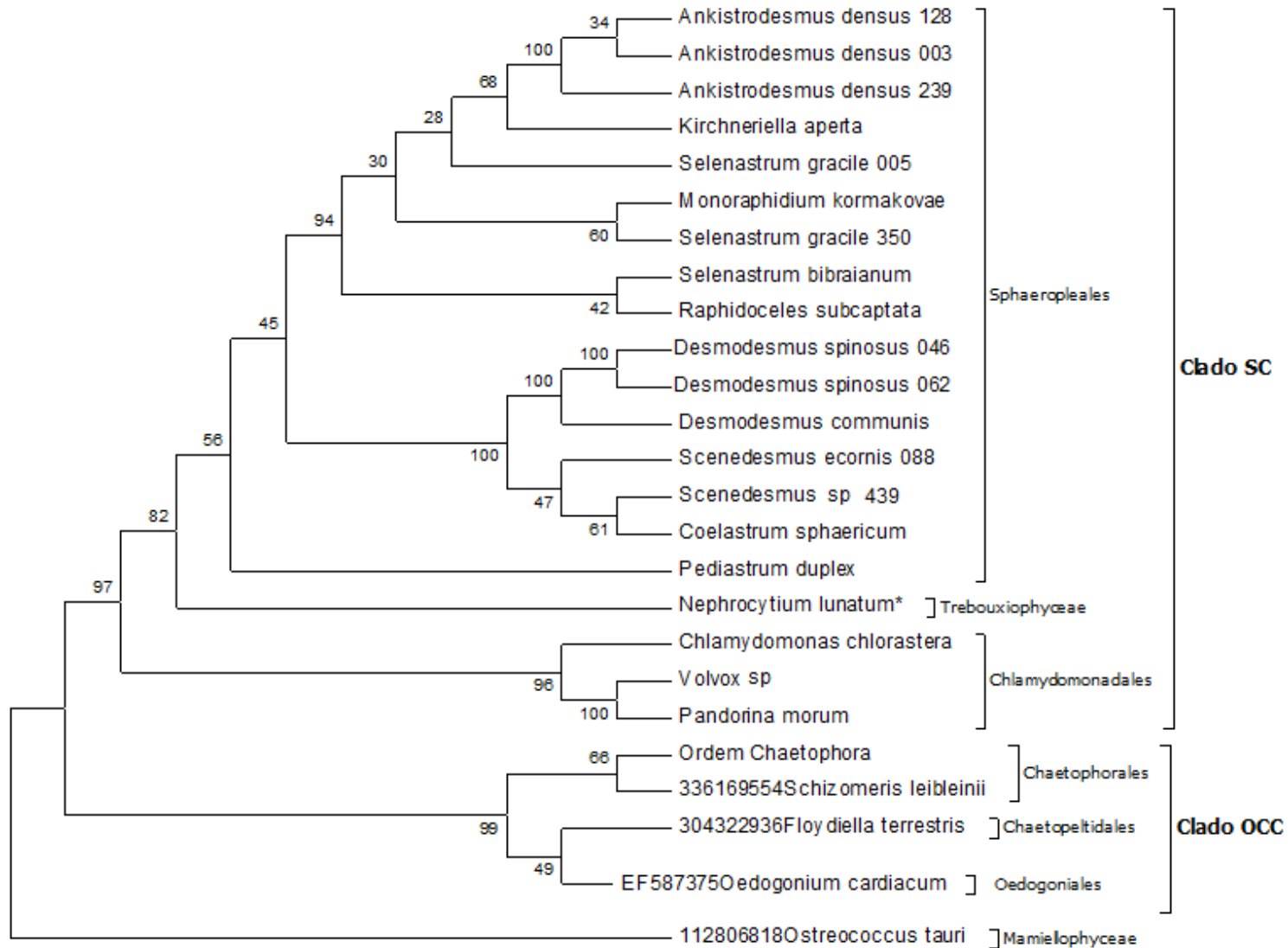


Figura 5 – Análise de Neighbor-Joining para as seqüências obtidas do marcador *tufA* mostrando o agrupamento dos clados OCC, SC, ordens e famílias das cepas do banco WDCM, n°835 da classe Chlorophyceae. Nos ramos estão plotados os valores de *Bootstrap*. As seqüências obtidas no *GenBank* estão representadas pelo número de acesso à esquerda do nome da espécie.

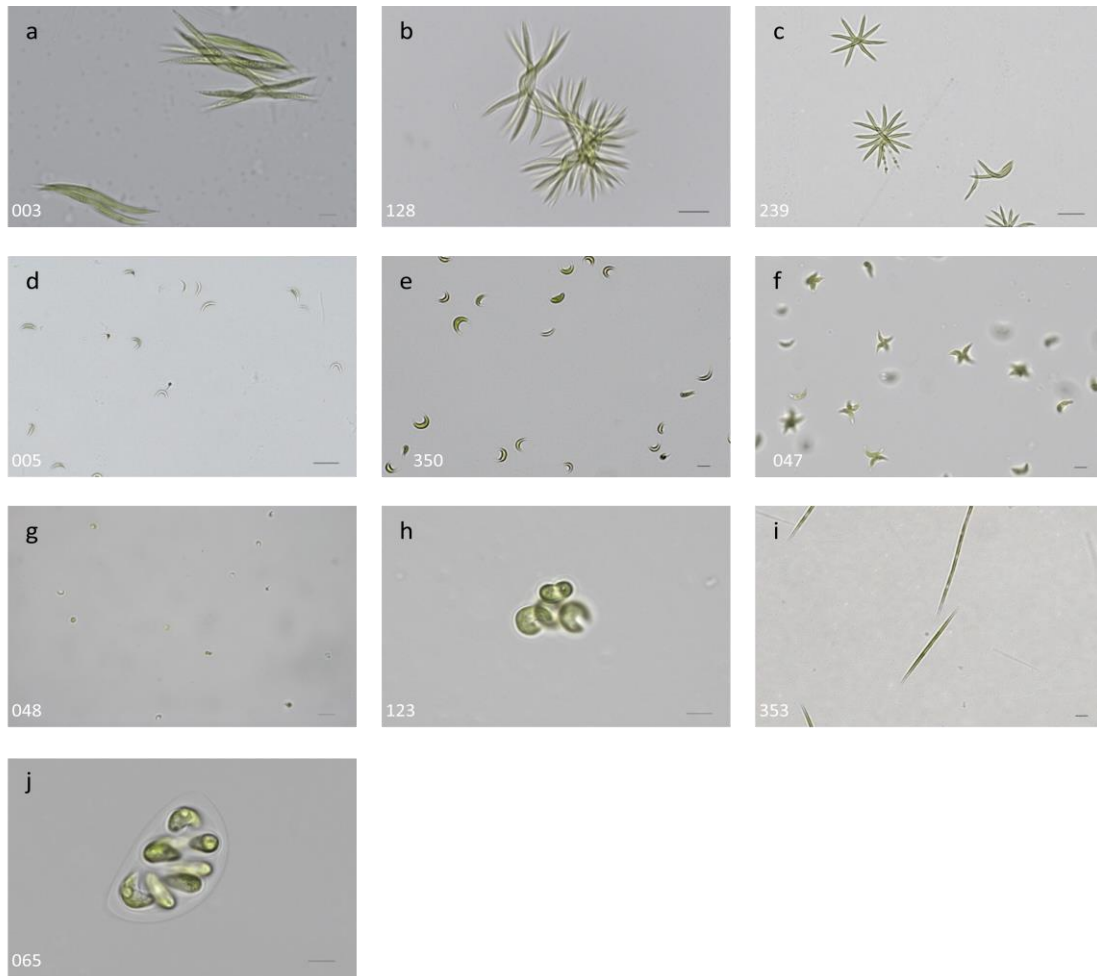


Figura 6 - Fotografias registradas no momento das identificações das cepas representantes da família de Sphaeropleales. Os números das cepas estão indicados no canto inferior esquerdo de cada fotografia: (a, b, c) *A. densus*. (d, e) *S. gracile*. (f) *S. bibraianum*. (g) *R. Subcapitata*. (h) *K. aperta*. (i) *M. komarkovae*. (j) *N. lunatum*. A barra de escala (10 μ m) está indicada no canto inferior direito.

Sozinho, o gene nuclear 18S mostrou bom suporte para os clados SC e OCC (Němcová *et al.* 2011), mas com baixo valor de suporte para a monofilia da ordem Sphaeropleales (Buchheim *et al.* 2001), o mesmo ocorrendo quando combinado com 28S (Shoup and Lewis 2003). Quando combinado com *rbcL* e outros genes plastidiais, foi possível verificar bom valor de suporte para a monofilia de Sphaeropleales (Tippery *et al.* 2012).

Os dados obtidos neste trabalho com o gene *tufA* isoladamente, ou seja, sem estar combinado a outros genes, possibilitaram bom suporte das topologias já encontradas em outros trabalhos que utilizaram dados de outros genes combinados entre si. Em outras palavras, o gene *tufA* foi capaz de reproduzir sozinho a filogenia esperada para o grupo.

Nephrocitium lunatum

Nas análises com os marcador *tufA* *Nephrocitium lunatum* encontrou-se dentro da classe Chlorophyceae. Considerando que este gênero seja atualmente representante de outra classe (Trebouxiophyceae), e que o gene *tufA* separou bem famílias e ordens, era esperado que *Nephrocitium lunatum* fosse alocada num ramo externo à classe Chlorophyceae, como ocorrido com *Ostreococcus tauri*.

Analisando as distâncias genéticas é possível que *N. lunatum* pertença à outra ordem componente do clado SC, visto que as ordens Chlamydomonadales e Sphaeropleales divergem 19% entre si e a cepa de *N. lunatum* tem 19% e 20% de diferença entre Sphaeropleales e Chlamydomonadales respectivamente.

Ordens componentes do clado OCC têm no mínimo 24% de diferença genética com o clado SC (Chaetophorales x Chlamydomonadales). Esta diferença é maior que a encontrada entre as classes Chlorophyceae e Trebouxiophyceae (considerando que *N. lunatum* seja da classe Trebouxiophyceae) (21%). Porém é necessário ressaltar que o gênero *Nephrocitium* não está suficientemente representado com apenas uma cepa, sendo, portanto crucial a inclusão de mais espécies do gênero e a análises com outros genes para elucidar sua posição em relação à classe Chlorophyceae.

O gênero *Nephrocytium* já esteve classificado em Chlorophyceae (West, 1892), quando então a classe teve algumas ordens redistribuídas. A ordem Chlorococcales foi desfeita e redistribuída, com alguns gêneros mantidos em Chlorophyceae. *Nephrocytium lunatum* junto com outros gêneros restantes de Chlorococcales passaram a integrar outra ordem, Chlorellales, agora da classe Trebouxiophyceae (Friedl, 1995). Entretanto não há dados suficientemente esclarecedores da transferência do gênero *Nephrocytium* para a classe Trebouxiophyceae.

Por não possuírem nenhum aparato flagelar, não é possível a classificação do gênero *Nephrocytium* a partir desta característica (Huss and Sogin 1990). Situação semelhante foi encontrada com o gênero *Chlorella*, a qual pôde ser classificada tanto em Chlorophyceae como em Trebouxiophyceae a partir de dados obtidos com 18S (Huss *et al.* 1999).

É importante ressaltar que foram usados poucos representantes dos táxons utilizados, e a melhor maneira de se elaborar uma filogenia robusta e avaliar a aplicação dos marcadores como *barcodes*, é ampliar o número representativo de espécimes dos táxons a serem testados ou ainda realizar a análise com múltiplos *loci* (Hajibabaei *et al.* 2007), como já utilizado para plantas terrestres (Newmaster *et al.* 2006) (Kress & Erickson, 2007).

Conclusões

O presente estudo mostrou que dos marcadores testados, o gene *tufA* pode ser amplificado e seqüenciado de maneira satisfatória com apenas um par de *primers* em todos os organismos testados da classe Chlorophyceae.

Os dados obtidos apenas com o gene *tufA* renderam reconstruções filogenéticas satisfatórias, com bom suporte dos principais clados observados em outros trabalhos que utilizaram dados de diversos genes combinados. Ainda assim, muitos destes estudos foram feitos com pouca amostragem da classe ou quando dispunham de amostragem mais ampla, tratava-se de seqüências obtidas no *GenBank*.

Desta maneira, o presente trabalho contribuiu também para a disponibilidade de mais seqüências de espécies e diferentes linhagens de Chlorophyceae no *GenBank*,

Além da possibilidade de reconstrução dos modelos de filogenia esperada para a classe com as seqüências obtidas com os *primers* para o gene *tufA*, as análises das características destas seqüências mostraram que se trata de um gene apropriado para estudos de filogenia, sendo usado sozinho, e que se mostra um promissor marcador molecular principal para a classe.

Para que se possam resolver as problemáticas encontradas nos ramos internos das famílias de Sphaeropleales, sugere-se o uso de outros genes menos conservados para serem aplicados nos gêneros que apresentaram dificuldade de separação de modo a auxiliar a conciliar os conceitos morfológicos com os dados moleculares.

Entretanto o gene *tufA* deve ser testado também nos ramos mais internos de grupos de Chlorophyceae onde a classificação esteja bem estabelecida, a fim de responder se há a possibilidade de sua aplicação também como *barcode* na classe.

Referências

- Altschul SF, Madden TL, Schäffer AA, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- Assunção P, Jaén-Molina R, Caujapé-Castells J, et al. (2011) Phylogenetic position of *Dunaliella acidophila* (Chlorophyceae) based on ITS and *rbcL* sequences. *J Appl Phycol* 24:635–639. doi: 10.1007/s10811-011-9676-1
- Bagatini, I. L. (2008) Associação de Bactérias à Capsula de *Anabaena spiroides* (Cyanobateria) em Cultura. Universidade Federal de São Carlos
- Baldauf SL, Manhart JR, Palmer JD (1990) Different fates of the chloroplast *tufA* gene following its transfer to the nucleus in green algae. *Proc Natl Acad Sci U S A* 87:5317–5321.
- Bickford D, Lohman DJ, Sodhi NS, et al. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155. doi: 10.1016/j.tree.2006.11.004
- Bock C, Pröschold T, Krienitz L (2010) Two new *Dictyosphaerium* -morphotype lineages of the Chlorellaceae (Trebouxiophyceae): *Heynigia* gen. nov. and *Hindakia* gen. nov. *Eur J Phycol* 45:267–277. doi: 10.1080/09670262.2010.487920
- Brouard JS, Otis C, Lemieux C, Turmel M (2010) The Exceptionally Large Chloroplast Genome of the Green Alga *Floydiella terrestris* Illuminates the Evolutionary History of the Chlorophyceae. *Genome Biol Evol* 2:240–256. doi: 10.1093/gbe/evq014
- Buchheim MA, Michalopoulos EA, Buchheim JA (2001) Phylogeny of the Chlorophyceae with special reference to the Sphaeropleales: a study of 18S and 26S rDNA data. *J Phycol* 37:819–835.
- Buchheim MA, Sutherland DM, Schleicher T, et al. (2012) Phylogeny of Oedogoniales, Chaetophorales and Chaetopeltidales (Chlorophyceae): inferences from sequence-structure analysis of ITS2. *Ann Bot* 109:109–116. doi: 10.1093/aob/mcr275
- CBOL Plant Working Group, Hollingsworth PM, Forrest LL, et al. (2009) A DNA barcode for land plants. *Proc Natl Acad Sci* 106:12794–12797. doi: 10.1073/pnas.0905845106
- Chen S, Yao H, Han J, et al. (2010) Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. *PLoS ONE* 5:e8613. doi: 10.1371/journal.pone.0008613
- Chevalier P, Noüe J (1985) Efficiency of immobilized hyper concentrated algae for ammonium and orthophosphate removal from wastewaters. *Biotechnol Lett* 7:395–400. doi: 10.1007/BF01166210

- Christa G, Wescott L, Schäberle TF, et al. (2013) What remains after 2 months of starvation? Analysis of sequestered algae in a photosynthetic slug, *Plakobranchnus ocellatus* (Sacoglossa, Opisthobranchia), by barcoding. *Planta* 237:559–572. doi: 10.1007/s00425-012-1788-6
- Coleman AW (2007) Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res* 35:3322–3329. doi: 10.1093/nar/gkm233
- Darienko T, Gustavs L, Mudimu O, et al. (2010) *Chloroidium*, a common terrestrial coccoid green alga previously assigned to *Chlorella* (Trebouxiophyceae, Chlorophyta). *Eur J Phycol* 45:79–95. doi: 10.1080/09670260903362820
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772–772. doi: 10.1038/nmeth.2109
- Delwiche CF, Kuhsel M, Palmer JD (1995) Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. *Mol Phylogenet Evol* 4:110–128. doi: 10.1006/mpev.1995.1012
- Du G, Wu F, Mao Y, et al. (2014) DNA barcoding assessment of green macroalgae in coastal zone around Qingdao, China. *J Ocean Univ China* 13:97–103. doi: 10.1007/s11802-014-2197-1
- Eiler A, Drakare S, Bertilsson S, et al. (2013) Unveiling Distribution Patterns of Freshwater Phytoplankton by a Next Generation Sequencing Based Approach. *PLoS ONE* 8:e53516. doi: 10.1371/journal.pone.0053516
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–185.
- Famà P, Wysor B, Kooistra WH, Zuccarello GC (2002) Molecular Phylogeny Of The Genus *Caulerpa* (Caulerpales, Chlorophyta) Inferred From Chloroplast *tufA* GENE1. *J Phycol* 38:1040–1050.
- Fawley MW, Dean ML, Dimmer SK, Fawley KP (2006) Evaluating The Morphospecies Concept In The Selenastraceae (Chlorophyceae, Chlorophyta). *J Phycol* 42:142–154. doi: 10.1111/j.1529-8817.2006.00169.x
- Fazekas AJ, Burgess KS, Kesanakurti PR, et al. (2008) Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. *PLoS ONE* 3:e2802. doi: 10.1371/journal.pone.0002802
- 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. doi: 10.1111/j.1529-8817.1995.tb02559.x
- Fucikovà, et al. (2012) Intersection of *Chlorella*, *Muriella* and *Bracteacoccus*: Resurrecting the genus *Chromochloris* KOL et CHODAT (Chlorophyceae, Chlorophyta). *Fottea* 12:83–93.

- Fučíková K, Lewis PO, Lewis LA (2014) Putting incertae sedis taxa in their place: a proposal for ten new families and three new genera in Sphaeropleales (Chlorophyceae, Chlorophyta). *J Phycol* 50:14–25. doi: 10.1111/jpy.12118
- Fumagalli M, Vieira FG, Linderoth T, Nielsen R (2014) ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinforma Oxf Engl*. doi: 10.1093/bioinformatics/btu041
- Garcia-Cuetos L, Moestrup Ø, Hansen PJ, Daugbjerg N (2010) The toxic dinoflagellate *Dinophysis acuminata* harbors permanent chloroplasts of cryptomonad origin, not kleptochloroplasts. *Harmful Algae* 9:25–38. doi: 10.1016/j.hal.2009.07.002
- Griffiths MJ, Harrison STL (2009) Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *J Appl Phycol* 21:493–507. doi: 10.1007/s10811-008-9392-7
- Guillard RRL, Lorenzen CJ (1972) Yellow-Green Algae With Chlorophyllide C². *J Phycol* 8:10–14. doi: 10.1111/j.1529-8817.1972.tb03995.x
- Hajibabaei M, Shokralla S, Zhou X, et al. (2011) Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. *PLoS ONE* 6:e17497. doi: 10.1371/journal.pone.0017497
- Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet* 23:167–172. doi: 10.1016/j.tig.2007.02.001
- Hanagata N, Takeuchi T, Fukuju Y, et al. (1992) Tolerance of microalgae to high CO₂ and high temperature. *Phytochemistry* 31:3345–3348. doi: 10.1016/0031-9422(92)83682-O
- Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270:313–321.
- Hoek C van den (1995) *Algae: an introduction to phycology*. Cambridge University Press, Cambridge ; New York
- Hoham RW, Bonome TA, Martin CW, Leebens-Mack JH (2002) A Combined 18s Rdna And Rbcl Phylogenetic Analysis Of Chloromonas And Chlamydomonas (Chlorophyceae, Volvocales) Emphasizing Snow And Other Cold-Temperature Habitats1. *J Phycol* 38:1051–1064. doi: 10.1046/j.1529-8817.2002.t01-1-01227.x
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinforma Oxf Engl* 17:754–755.
- Huss VA, Sogin ML (1990) Phylogenetic position of some *Chlorella* species within the chlorococcales based upon complete small-subunit ribosomal RNA sequences. *J Mol Evol* 31:432–442.

- Huss VAR, Frank C, Hartmann EC, et al. (1999) Biochemical Taxonomy And Molecular Phylogeny Of The Genus *Chlorella* Sensu Lato (Chlorophyta). *J Phycol* 35:587–598. doi: 10.1046/j.1529-8817.1999.3530587.x
- Keller A, Schleicher T, Förster F, et al. (2008) ITS2 data corroborate a monophyletic chlorophycean DO-group (Sphaeropleales). *BMC Evol Biol* 8:218. doi: 10.1186/1471-2148-8-218
- Kress WJ, Erickson DL (2007) A Two-Locus Global DNA Barcode for Land Plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. *PLoS ONE* 2:e508. doi: 10.1371/journal.pone.0000508
- Krienitz L, Bock C (2012) Present state of the systematics of planktonic coccoid green algae of inland waters. *Hydrobiologia* 698:295–326. doi: 10.1007/s10750-012-1079-z
- 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 Chlorophyta1: Crescent-Shaped Chlorophyta. *J Phycol* 47:880–893. doi: 10.1111/j.1529-8817.2011.01010.x
- Krienitz L, Ustinova I, Friedl T, Huss VAR (2001) Traditional Generic Concepts Versus 18s Rrna Gene Phylogeny In The Green Algal Family Selenastraceae (Chlorophyceae, Chlorophyta). *J Phycol* 37:852–865. doi: 10.1046/j.1529-8817.2001.01004.x
- Lawton RJ, Mata L, de Nys R, Paul NA (2013) Algal Bioremediation of Waste Waters from Land-Based Aquaculture Using *Ulva*: Selecting Target Species and Strains. *PLoS ONE* 8:e77344. doi: 10.1371/journal.pone.0077344
- Le Gall L, Saunders GW (2010) DNA Barcoding Is A Powerful Tool To Uncover Algal Diversity: A Case Study Of The Phylloporaceae (Gigartinales, Rhodophyta) In The Canadian Flora. *J Phycol* 46:374–389. doi: 10.1111/j.1529-8817.2010.00807.x
- Lewis LA, McCourt RM (2004) Green algae and the origin of land plants. *Am J Bot* 91:1535–1556.
- Librado & Rozas (2009) DnaSP v5: A software for comprehensive analysis of DNA poly-morphism data. *Bioinforma*. 101093bioinformaticsbt187
- Lis JT, Schleif R (1975) Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res* 2:383–389.
- Lokhorst GM, Bakker ME, Star W (1984) Ultrastructure of *Draparnaldia glomerata* (Chaetophorales, Chlorophyceae) II. Mitosis and cytokinesis. *Nord J Bot* 4:553–562. doi: 10.1111/j.1756-1051.1984.tb02063.x
- Ludwig W, Weizenegger M, Betzl D, et al. (1990) Complete nucleotide sequences of seven eubacterial genes coding for the elongation factor Tu: functional,

- structural and phylogenetic evaluations. Arch Microbiol 153:241–247. doi: 10.1007/BF00249075
- McManus HA, Lewis LA (2011) Molecular Phylogenetic Relationships In The Freshwater Family Hydrodictyaceae (Sphaeropleales, Chlorophyceae), With An Emphasis On *Pediastrum Duplex*: Hydrodictyaceae Phylogeny. J Phycol 47:152–163. doi: 10.1111/j.1529-8817.2010.00940.x
- Moniz MBJ, Kaczmarek I (2010) Barcoding of Diatoms: Nuclear Encoded ITS Revisited. Protist 161:7–34. doi: 10.1016/j.protis.2009.07.001
- Němcová Y, Eliáš M, Škaloud P, et al. (2011) Jenufa Gen. Nov.: A New Genus Of Coccoid Green Algae (Chlorophyceae, Incertae Sedis) Previously Recorded By Environmental Sequencing1: Jenufa Gen. Nov.: A New Genus Of Coccoid Green Algae. J Phycol 47:928–938. doi: 10.1111/j.1529-8817.2011.01009.x
- Newmaster SG, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Can J Bot 84:335–341. doi: 10.1139/b06-047
- Nordi CSF, Cavagliere TGWF, Vieira AAH, Nascimento OR (2006) Efeito caotrópico do íon lítio na permeabilidade da cápsula polissacarídica da microalga *Ankistrodesmus gracilis* (Reinsch) Korsikov (Chlorophyceae). Acta Bot Bras 20:449–454. doi: 10.1590/S0102-33062006000200020
- Nozaki H (2000) Origin and Evolution of the Colonial Volvocales (Chlorophyceae) as Inferred from Multiple, Chloroplast Gene Sequences. Mol Phylogenet Evol 17:256–268. doi: 10.1006/mpev.2000.0831
- Nozaki H, Misumi O, Kuroiwa T (2003) Phylogeny of the quadriflagellate Volvocales (Chlorophyceae) based on chloroplast multigene sequences. Mol Phylogenet Evol 29:58–66. doi: 10.1016/S1055-7903(03)00089-7
- Nozaki H, Ohta N, Yamada T, Takano H (1998) Characterization of *rbcL* group IA introns from two colonial volvoclean species (Chlorophyceae). Plant Mol Biol 37:77–85.
- Nozaki H, Takahara M, Nakazawa A, et al. (2002) Evolution of *rbcL* group IA introns and intron open reading frames within the colonial Volvocales (Chlorophyceae). Mol Phylogenet Evol 23:326–338.
- O’Kelly CJ, Kurihara A, Shipley TC, Sherwood AR (2010) Molecular Assessment Of *Ulva Spp.* (Ulvophyceae, Chlorophyta) In The Hawaiian Islands1: Diversity Of Hawaiian *Ulva*. J Phycol 46:728–735. doi: 10.1111/j.1529-8817.2010.00860.x
- O’Kelly CJ, Wysor B, Bellows WK (2004) Gene Sequence Diversity And The Phylogenetic Position Of Algae Assigned To The Genera *Phaeophila* And *Ochlochaete* (Ulvophyceae, Chlorophyta): Phylogeny Of *Phaeophila*. J Phycol 40:789–799. doi: 10.1111/j.1529-8817.2004.03204.x
- Pombert J-F (2005) The 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* 22:1903–1918. doi: 10.1093/molbev/msi182
- Provan J, Murphy S, Maggs CA (2004) Universal plastid primers for Chlorophyta and Rhodophyta. *Eur J Phycol* 39:43–50. doi: 10.1080/09670260310001636668
- Reyes A, Semenkovich NP, Whiteson K, et al. (2012) Going viral: next-generation sequencing applied to phage populations in the human gut. *Nat Rev Microbiol* 10:607–617. doi: 10.1038/nrmicro2853
- Saez AG, Zaldivar-Riveron A, Medlin LK (2008) Molecular systematics of the Pleurochrysidaceae, a family of coastal coccolithophores (Haptophyta). *J Plankton Res* 30:559–566. doi: 10.1093/plankt/fbn023
- Salipante SJ, Sengupta DJ, Rosenthal C, et al. (2013) Rapid 16S rRNA Next-Generation Sequencing of Polymicrobial Clinical Samples for Diagnosis of Complex Bacterial Infections. *PLoS ONE* 8:e65226. doi: 10.1371/journal.pone.0065226
- Saunders & Kucera (2010) An evaluation of *rbcL*, *tufA*, *UPA*, *LSU* and *ITS* as DNA barcode markers for the marine green macroalgae. *Cryptogam Algal* 31(4), 487–528.
- Saunders GW, McDevit DC (2012) Methods for DNA Barcoding Photosynthetic Protists Emphasizing the Macroalgae and Diatoms. In: Kress WJ, Erickson DL (eds) *DNA Barcodes*. Humana Press, Totowa, NJ, pp 207–222
- Sawayama S, Inoue S, Dote Y, Yokoyama S-Y (1995) CO₂ fixation and oil production through microalga. *Energy Convers Manag* 36:729–731. doi: 10.1016/0196-8904(95)00108-P
- Schoch CL, Seifert KA, Huhndorf S, et al. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci* 109:6241–6246.
- Seifert KA (2009) Progress towards DNA barcoding of fungi. *Mol Ecol Resour* 9:83–89. doi: 10.1111/j.1755-0998.2009.02635.x
- Sheehan et al (1998) A look back at the US Department of Energy's Aquatic Species Program: Biodiesel from algae. Golden, CO: National Renewable Energy Laboratory. Report NREL/TP-580–24190.
- Shoup S, Lewis LA (2003) Polyphyletic Origin Of Parallel Basal Bodies In Swimming Cells Of Chlorophycean Green Algae (Chlorophyta)1. *J Phycol* 39:789–796. doi: 10.1046/j.1529-8817.2003.03009.x
- Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci* 105:13486–13491. doi: 10.1073/pnas.0803076105

- Stewart, K. D., & Mattox, K. R. (1975) Comparative cytology, evolution and classification of the green algae with some consideration of the origin of other organisms with chlorophylls a and b. *Bot Rev* 41:104–135.
- Tamura K, Stecher G, Peterson D, et al. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 30:2725–2729. doi: 10.1093/molbev/mst197
- Thompson JD, Higgins DG, Gibson TJ (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.
- Tippary NP, Fučíková K, Lewis PO, Lewis LA (2012) Probing the Monophyly of the Sphaeropleales (Chlorophyceae) Using Data From Five Genes. *J Phycol* 48:1482–1493. doi: 10.1111/jpy.12003
- Turmel M, Brouard J-S, Gagnon C, et al. (2008) Deep Division In The Chlorophyceae (Chlorophyta) Revealed By Chloroplast Phylogenomic Analyses. *J Phycol* 44:739–750. doi: 10.1111/j.1529-8817.2008.00510.x
- 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 EJ, Lüring 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 ADR, et al. (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. doi: 10.1080/09670260600709315
- Verbruggen H, Pauly K, De Clerck O (2012) The new species *Codium recurvatum* from Tanzania. *Eur J Phycol* 47:216–222. doi: 10.1080/09670262.2012.690106
- Watanabe S, Tsuchimoto K, Floyd GL (1989) Light and electron microscopy of *Brachiomonas submarina* Bohlin (Chlamydomonadales, Chlorophyceae). *Phycologia* 28:188–196. doi: 10.2216/i0031-8884-28-2-188.1
- West (1892) Algae of the English Lake District. *J R Microsc Soc Lond* 14:1–7.
- White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. W. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*. *PCR Protoc Guide Methods Appl* 315–322.
- Wynne MJ, Verbruggen H, Angel DL (2009) The Recognition of *Caulerpa Integerrima* (Zanardini) comb. et stat. nov. (Bryopsidales, Chlorophyta) from the Red Sea. *Phycologia* 48:291–301. doi: 10.2216/08-78.1

Yoon HS, Hackett JD, Ciniglia C, et al. (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* 21:809–818. doi: 10.1093/molbev/msh075