

UNIVERSIDADE FEDERAL DO PAMPA
CAMPUS SÃO GABRIEL
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

EVELISE LEIS CARVALHO

**GENOMAS ACESSÓRIOS DA ALGA ANTÁRTICA *PRASIOLA CRISPA*: INFERÊNCIAS
ESTRUTURAIS E FILOGENÉTICAS**

SÃO GABRIEL

2015

EVELISE LEIS CARVALHO

**GENOMAS ACESSÓRIOS DA ALGA ANTÁRTICA *PRASIOLA CRISPA*: INFERÊNCIAS
ESTRUTURAIS E FILOGENÉTICAS**

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Orientador: Prof. Dr. Paulo Marcos Pinto

Coorientador: Dr. Gabriel da Luz Wallau

São Gabriel

2015

EVELISE LEIS CARVALHO

**GENOMAS ACESSÓRIOS DA ALGA ANTÁRTICA *PRASIOLA CRISPA*: INFERÊNCIAS
ESTRUTURAIS E FILOGENÉTICAS**

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Dissertação defendida e aprovada em 19 de Maio de 2015.

Banca examinadora:

Prof. Dr. Paulo Marcos Pinto
Orientador - UNIPAMPA

Prof^ª. Dr^ª. Alessandra Loureiro Morassutti
Membro Titular - PUCRS

Prof. Dr. Valdir Marcos Stefenon
Membro Titular - UNIPAMPA

À meu pai, Evaldo Carvalho, por me ensinar a jamais desistir dos meus objetivos, por mais árduo que seja o caminho.

AGRADECIMENTOS

Agradeço primeiramente aos meus pais Evaldo e Indiara, pelo incentivo constante, pela confiança depositada, por todo o afeto e amor. Sou e serei eternamente grata pelo exemplo que são em minha vida. Palavras apenas não são seriam suficientes para expressar a admiração, o amor e a gratidão que sinto. Amo vocês incondicionalmente!

Aos meus avós José Carlos e Elisa, pelo apoio constante, de todas as formas. É impossível descrever o tamanho do amor e da veneração que sinto. Obrigada por tudo!

Ao meu padrasto, Antônio, por me acolher como mais uma das tuas gurias. Obrigada pelo carinho e apoio.

Ao meu orientador Prof. Paulo Marcos Pinto pelos ensinamentos, paciência, pela amizade e por acreditar no meu potencial, muitas das vezes mais que eu mesma. Durante estes cinco anos de convívio, desde a minha iniciação científica, pude ver o quão brilhante profissional és. Sempre serei grata pela tamanha contribuição que destes a minha carreira profissional. Obrigada por me dar “asas para voar na carreira científica”.

Ao meu co-orientador Gabriel Wallau pela enorme participação no desenvolvimento deste trabalho. Obrigada pelo convívio, paciência, incentivo, ensinamentos, amizade, pelos churrascos com coração de ovelha, por me proporcionar o convívio com a tua linda família. Torço muito para que consigas voltar ao nosso LPA, pois a tua ausência será muito sentida por todos nós.

Aos colegas do Laboratório de Proteômica Aplicada #TrupedoPaulo #TityusdoAsfalto #BonedoLPA: Matheus, Alexandre, Laís, Luiz Fernando, Tainah, Filipe, Pablo, Bruno, Darlene, Jéssica pelos momentos de convívio e almoços no RU.

As gurias da formação “Megazord”, Darlene e Laís, pela ajuda com as intermináveis, e por vezes desesperadoras, tabelas.

Ao Bruno, nosso salvador quando o assunto é informática. Muito obrigada pela disponibilidade e vontade de ajudar sempre.

Ao Alexandre, meu primeiro IC e companheiro na tentativa frustrada no cultivo de algas.

Ao Luiz Fernando, vulgo Luiz “Filipe”, o guri da primeira capital farroupilha, “especialista” em organização de aniversários surpresa e em coletar ovos de *Drosophila*. Obrigada pela disposição em ajudar sempre, seja no que for.

Aos professores do Programa de Pós-Graduação em Ciências Biológicas.

À comissão examinadora desta dissertação pela disponibilidade.

Aos órgãos financiadores CNPq, CAPES, FAPERGS, INCT-APA.

À todos que contribuíram para a realização deste trabalho.

E finalmente agradeço ao homem que tornou a minha vida mais feliz, mais completa. Meu marido Paulo. Obrigada por ser tão dedicado, amoroso, companheiro, por tolerar os meus muitos defeitos, por seres o meu ponto de equilíbrio, por tudo que me proporciona ao viver ao teu lado. Obrigada por abrir meus olhos quando tudo parece tão difícil e nebuloso e por acreditar, mais que ninguém, no meu potencial. Tu és a página mais linda que o destino escreveu na minha vida. Eu te amo!

"Se as coisas são inatingíveis... ora!
Não é motivo para não querê-las...
Que tristes os caminhos, se não fora
A presença distante das estrelas!"

(Mário Quintana)

RESUMO

Algas verdes da classe Trebouxiophyceae estão entre os organismos presentes no continente Antártico, onde a espécie mais relatada é a macroalga verde *Prasiola crispa* (Lightfoot) Kützinger. Considerada um organismo extremófilo, pois se desenvolve com muito sucesso no habitat extremo da Antártica, ainda são raros na literatura dados moleculares sobre esta espécie, o que impede uma avaliação sobre sua taxonomia e posição filogenética. Com o advento das tecnologias de sequenciamento de nova geração, os genomas de organelas tornaram-se uma grande ferramenta para estudos de filogenia, pois fornecem inúmeros dados filogenéticos, sequências de proteínas e nucleotídeos e também informações sobre conteúdo gênico e arquitetura. Neste trabalho, foi determinada a sequência dos genomas do cloroplasto (cpDNA) e mitocondrial (mtDNA) de *P. crispa*, com o intuito de inferir as relações evolutivas deste organismos com outras espécies de plantas verdes, bem como uma análise estrutural. Os genomas plastidial e mitocondrial foram sequenciados por MacroGen Service (Solexa-Illumina Hi-Seq 2500). A montagem, anotação, alinhamento, construção da filogenia e análise sintênica foram realizados *in silico* com softwares específicos. O cpDNA e mtDNA *P. crispa* apresentam 196.502 pb e 89.819 pb, respectivamente. Estes genomas acessórios apresentam 21 genes putativos relacionados com a fotossíntese e 18 genes relacionados com o metabolismo oxidativo. A análise filogenômica baseada no cpDNA demonstrou que *P. crispa* agrupou com alga trebouxiophyceae *Prasiolopsis sp.* formando o clado *Prasiola* juntamente com *Stichococcus bacillaris*. Nossos resultados para filogenômica embasada no mtDNA revelam que *P. crispa* agrupa com as outras espécies da classe Trebouxiophyceae. A análise de sintenia do cpDNA e mtDNA de *P. crispa* com a espécies de plantas verdes relacionadas evolutivamente demonstram que estes organismos apresentam poucos blocos gênicos sintênicos. Este trabalho pioneiro com a alga *P. crispa*, demonstra que os genomas acessórios suprem uma gama de dados moleculares que podem ser utilizados para estudos filogenômicos. Além disto, as informações geradas a partir do sequenciamento do cpDNA e mtDNA de *P. crispa* fornecem um aporte para estudos futuros mais aprofundados.

Palavras-chave: *Prasiola crispa*, genoma mitocondrial, genoma do cloroplasto, análise filogenômica, sequenciamento de nova geração.

ABSTRACT

Green algae from Trebouxiophyceae class are among the organisms in the Antarctic continent, where the most reported species is the green macroalga *Prasiola crispa* (Lightfoot) Kützing. This algae is considered an extremophile organism because develops successfully in the harsh Antarctic habitat, however studies reporting molecular data of this species are still lacking in the literature, which prevents an assessment of their correct taxonomy and phylogenetic position. With the advent of next generation sequencing technologies, it became easier to obtain molecular information as for example from organelle genomes making them a great tool for taxonomic studies because they provide a great number of, phylogenetic data, nucleotides, protein sequences, gene content and architecture information. In this study, we determined the sequence of the chloroplast (cpDNA) and mitochondrial (mtDNA) genome of *P. crispa* in order to infer the evolutionary relationships of the organisms with other species of green plants, as well as a structural analysis. Plastid and mitochondrial genome was sequenced by Macrogen Service (Illumina Solexa Hi-Seq 2500). The genome assembly, annotation, sequences alignment, phylogeny construction, and structural analyses were performed *in silico* with specific softwares. Plastid and Mitochondrial genomes have a total length of 196,502 bp and 89,819 bp, respectively. These genomes presented 21 putative photosynthesis related genes and 18 oxidative metabolism related genes, respectively. Phylogenetic analysis based on the cpDNA demonstrated that *P. crispa* grouped with Trebouxiophyceae algae *Prasiolopsis* sp. forming the *Prasiola* clade along with *Stichococcus bacillaris*. Our results for phylogenetic analysis grounded in mtDNA show that *P. crispa* groups with other species of Trebouxiophyceae alga. Synteny analysis of *P. crispa* cpDNA and mtDNA with evolutionarily related species of green plants shows that these organisms have few syntenic gene blocks. This pioneering work with *P. crispa* provided the accessory genomes which supplied a range of molecular data that can be employed to taxonomic studies. In addition, the information generated from the sequencing of cpDNA and mtDNA of *P. crispa* provide a contribution for further studies.

Keywords: *Prasiola crispa*, mitochondrial genome, chloroplast genome, phylogenomic analysis, next generation sequencing.

LISTA DE FIGURAS

INTRODUÇÃO

Figura 1 - Alga Antártica <i>Prasiola crispa</i>	14
Figura 2 - Estrutura quadripartida do cromossomo plastidial.....	17
Figura 3 - Mapas representativos dos genomas plastidiais das algas verdes <i>Floydiella terrestris</i> (Chlorophyceae) e <i>Helicosporidium sp.</i> (Trebouxiophyceae).....	20
Figura 4 - Mapa representativo do genoma mitocondrial de <i>Chlorella sp.</i>	24
Figura 5 - Posição filogenética <i>Chlorokybus atmophyticus</i> dentro do reino Viridiplantae.....	26

ARTIGO 2

Figure 1 - Phylogenomic analysis of 70 chlorophytes chloroplast genomes using 14469 amino acid positions from 79 cpDNA-encoded proteins.....	56
Figure 2 - Phylogenomic analysis of 27 mitochondrial genomes encompassing all 18 chlorophyte mtDNA genomes available.....	57
Figure 3 - Chloroplast synteny blocks analysis performed using blastn with default parameters and Artemis ACT software.....	58
Figure 4 - Mitochondrial synteny blocks analysis performed using blastn with default parameters and Artemis ACT software.....	59

LISTA DE TABELAS

INTRODUÇÃO

Tabela 1 - Exemplos da classe Trebouxiophyceae com genoma plastidial completo sequenciado.....	19
Tabela 2 - Espécies da classe Trebouxiophyceae com genoma plastidial sequenciado não depositado na plataforma <i>Organelle Genome Resources</i> (NCBI).....	20
Tabela 3 - Exemplos da classe Trebouxiophyceae com genoma mitocondrial sequenciado.....	23

ARTIGO 2

Table 1 - Genes identified in the chloroplast and mitochondrial genome of <i>P. crista</i>	43
--	----

Additional files

Additional file 1 - Numbers and the description of cpDNA genomes found using different search approach.....	60
Additional file 2 - Numbers and the description of mtDNA genomes found using different search approach genome.....	62

SUMÁRIO

1. INTRODUÇÃO	13
1.1. A alga <i>Prasiola crispera</i>	13
1.2. Análise de genomas plastidiais como ferramenta filogenética.....	16
1.3. Análise de genomas mitocondriais como ferramenta filogenética.....	22
2. OBJETIVOS	27
2.1. Objetivo geral	27
2.2. Objetivos específicos.....	27
3. ARTIGO 1	28
4. ARTIGO 2	32
5. CONSIDERAÇÕES FINAIS	63
6. PERSPECTIVAS FUTURAS	63
7. REFERÊNCIAS BIBLIOGRÁFICAS	64

1. INTRODUÇÃO

1.1. A alga *Prasiola crispera*

Viridiplantae, ou plantas verdes, é um clado com aproximadamente 500 mil espécies que exibem uma diversidade espantosa de formas de vida (Ruhfel et al., 2014). Esta linhagem fracionou-se em duas divisões principais: Chlorophyta, que contém a maioria das algas verdes descritas e Streptophyta, contendo as algas verdes do filo Charophyta e plantas terrestres (Embriófitas) (Lewis; McCourt, 2004).

Algas verdes representam uma linhagem antiga de eucariotos fotossintetizantes, estimando-se sua origem entre 700 e 1.500 milhões de anos atrás (Leliaert et al., 2012). Enquanto as algas verdes são onipresentes em oceanos do mundo e ecossistemas de água doce, as plantas terrestres são os principais componentes estruturais dos ecossistemas terrestres (Lewis; McCourt, 2004).

Chlorophyta constitui um filo morfologicamente e ecologicamente diversificado de algas verdes. As espécies constituintes deste táxon foram classificadas com base na ultraestrutura do aparelho flagelar e citocinese durante a mitose (Mattox; Stewart, 1984). Filocostas reconhecem quatro classes de algas verdes pertencentes ao filo Chlorophyta: Chlorophyceae, Trebouxiophyceae, Ulvophyceae e Prasinophyceae, este último o único grupo parafilético (Turmel et al., 1999).

Trebouxiophyceae abrange as algas das águas marinhas e doces (Servín-Garcidueñas; Martínez-Romero, 2012). Dentre as diversas ordens que compõe a classe Trebouxiophyceae, Prasiolales compreende um dos mais generalizados e versáteis grupos de algas verdes. Esta ordem inclui espécies distribuídas em ambientes marinhos, de água doce, habitats terrestres de regiões polares e frio temperado (Rodríguez et al., 2007). A morfologia destas algas varia desde filamentos unisseriados para talos em forma de fita, lâminas expandidas ou de pacotes como colônias, que são caracterizadas por uma grande plasticidade fenotípica relacionada a fatores ambientais (Rindi et al., 2007).

Espécies da família Prasiolaceae são caracterizadas por um talo frondoso de folhas com forma variável, com ou sem uma nítida *holdfast*, células dispostas em linhas verticais ou horizontais e um cloroplasto axial estrelado contendo um pirenóide central (John, 2002).

O gênero *Prasiola* está entre as algas presentes nas áreas de degelo do continente Antártico, onde a espécie mais relatada é *Prasiola crispera* (Lightfoot) Kützing (figura 1), um

importante produtor primário (Wiencke; Clayton, 2002; Convey, 2007; Kováčik; Pereira, 2001). Esta macroalga verde talóide está amplamente distribuída na Antártica Marítima e na Antártica Continental, podendo ser encontrada em regiões supralitorais e terrestres, principalmente próximo a colônias de pinguins, onde o solo é fertilizado com o guano dessas aves (Graham et al., 2009).

Figura 1. Alga Antártica *Prasiola crispa*.



Fonte: Graciele Alves.

P. crispa é um organismo extremófilo que, não apenas sobrevive, mas se desenvolve com muito sucesso em um dos mais extremos ambientes da terra onde se acreditava não haver vida (Javaux, 2006). Por tolerar repetidos ciclos de descongelamento e congelamento durante a primavera e outono, temperaturas negativas durante o inverno e altos níveis de radiação ultravioleta (UV) durante o verão (Lud et al., 2001; Kosugi et al., 2010), supõe-se que esta espécie apresente expressão diferencial de genes responsivos ao frio. A expressão destes genes durante a exposição de *P. crispa* a baixas temperaturas ainda permanece pouco estudada, bem como se estes genes também estão envolvidos na proteção da mesma a alta radiação UV do continente Antártico.

Devido às características extremas e impacto humano limitado, a Antártica tem atraído considerável interesse de pesquisadores (Moniz et al., 2012). Em recente estudo, Zemolin et al. (2014) demonstraram a ação bioinseticida de *P. crispa* sobre a mosca da fruta *Drosophila melanogaster* e a barata *Neuphoeta cinerea*. Contudo, informações a nível molecular e taxonômico desta espécie ainda são escassos (Moniz et al., 2012).

Os dados moleculares disponíveis atualmente para *P. crispa* foram produzidos nos últimos 13 anos e consistem principalmente de sequências dos genes *rbcL* (Rindi et al., 2007; Moniz et al., 2012), 18S rRNA (Rindi et al., 2004; Friedl; O’Kelly, 2002), *tufA* (Moniz; Guiry; Rindi, 2014), *psaB* (Moniz et al., 2012; Novis et al., 2013), *rpoB* (Novis et al., 2013) e *atpB* (Moniz et al., 2012).

A compreensão da diversidade e biogeografia de algas verdes vem se tornando um importante desafio. Embora muitos aspectos da biologia de *P. crispa* tenham sido completamente investigados, a escassez de dados moleculares impede uma avaliação precisa da sua taxonomia e posição filogenética (Moniz et al., 2012).

Genomas de organelas fornecem inúmeros dados filogenéticos, incluindo as sequências de proteínas e nucleotídeos, bem como informações sobre o conteúdo genético e arquitetura (Letsch; Lewis, 2012).

A compreensão da evolução do genoma mitocondrial (mtDNA) e do genoma do cloroplasto (cpDNA) dentro do Streptophyta é de grande interesse, uma vez que o uso de sequências de aminoácidos em substituição de sequências de nucleotídeos favorecem os estudos de filogenômica devido a riqueza de dados que portam (Lemieux; Otis; Turmel, 2014b).

As informações geradas nos últimos anos têm reformulado consideravelmente a visão sobre estes organismos, revelando uma diversidade genética maior do que sugere sua morfologia simples (Leliaert et al., 2012).

1.2. Análise de genomas plastidiais como ferramenta filogenética

Plastídeos são uma das principais características distintivas da célula vegetal (Wicke et al., 2007). Estas organelas fotossintetizantes fornecem a energia essencial para algas, plantas terrestres e alguns protozoários. Além da fotossíntese, outras vias metabólicas estão presentes nos plastídeos, incluindo a biossíntese de ácidos graxos, aminoácidos, pigmentos e vitaminas (Wang et al., 2013). Provavelmente a origem dos plastídeos está relacionada com um evento de endossimbiose, pelo qual um organismo protozoário unicelular, através do processo de fagocitose, englobou e manteve uma cianobactéria fotossintetizante (Reyes-Prieto et al., 2007), permitindo assim a transição de heterotrofia para autotrofia, adquirindo a capacidade de utilizar fotoenergia (Wicke et al., 2007). Em nível genômico, esta integração envolveu a perda de genes e a transferência de muitos destes genes para o genoma nuclear do hospedeiro fenômeno semelhante ao mtDNA (deKonin;, Keeling, 2006).

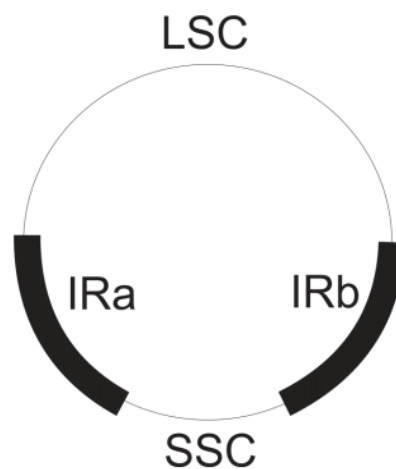
Comumente o cloroplasto desenvolve-se a partir de proplastídeos subdesenvolvidos (progenitor de todos os plastídeos), que contém apenas vesículas, não existindo estruturas diferenciadas. Durante a diferenciação, tilacóides são formados e organizam-se uns sobre os outros, formando estruturas denominadas *grana*. Os tilacóides são membranas internas lipídicas entrelaçadas com complexos de proteínas, que fornecem a plataforma para as reações de fotossíntese (Pribil; Labs; Leister, 2014). O espaço interno do cloroplasto é preenchido por um fluido viscoso denominado estroma onde estão presentes o material genético (DNA), enzimas e ribossomos. Possuem membrana dupla, fruto da origem dos plastídeos (Mota, 2012). Assim como as mitocôndrias, os plastídeos são, em geral, de origem materna (Harrison; Kidner, 2011).

Os plastídeos possuem um genoma próprio, circular e independente do DNA nuclear. O cromossomo plastidial é dividido em quatro segmentos principais, sendo duas regiões invertidas e repetidas (IR) que separam a região grande de cópia única (LSC), onde estão presentes a maioria dos genes plastidiais, e a região curta de cópia única (SSC) (figura 2) (Kolodner; Tewari, 1979).

A organização e a expressão do genoma de plastídeos estão sendo exaustivamente estudados. A estrutura e a sequência do genoma do cloroplasto são notavelmente conservadas em Streptophyta, divisão que compreende as algas verdes da classe Charophyceae e as plantas terrestres (Turmel; Otis; Lemieux, 2006). No entanto, estudos com genomas de cloroplastos de Chlorophyta, revelaram que este padrão não se aplica para este grupo de algas (Turmel;

Otis; Lemieux, 2009; deCambiaire et al., 2007). Por exemplo, as algas verdes Prasinophytes *Nephroselmis olivacea*, *Pyramimonas parkeae*, *Prasinoderma coloniale* e *Prasinococcus sp.*, abrigam em seu genoma plastidial genes conservados (*ndhJ*, *rbcR*, *rpl21*, *rps15*, *rps16*, *ycf66*, *rpl22* e *ycf65*) que nunca antes foram descritos para outras Chlorophyta. O genoma do cloroplasto de *N. olivacea* e *P. parkeae* exibe ainda um gene que codifica uma DNA primase, que se supõe ter sido adquirida de um vírus (Lemieux; Otis; Turmel, 2014b).

Figura 2 - Estrutura quadripartida do cromossomo plastidial.



Fonte: Motta, 2012.

A Trebouxiophyceae *Leptosira terrestris*, apresenta genoma plastidial circular de 195.081 pb e não possui uma região IR, fato este que se encontra também no genoma do cloroplasto de *Chlorella vulgaris* (150.613 pb) (Wakasugi et al., 1997), *Chlorella sorokiniana* (109.811 pb) (Orsini et al.; 2014) e *Helicosporidium sp.* (37.454 pb) (deKoning; Keeling, 2006). Por mais que estas espécies apresentem conteúdo gênico semelhante, elas diferem significativamente em densidade gênica, ordem dos genes e quantidade de íntrons. As características semelhantes apresentada por *L. terrestris* e seus homólogos em Chlorophyceae sugerem que a mesma força evolutiva gerou a ausência da região IR nessas duas linhagens de algas (deCambiaire et al., 2007).

Existem evidências que as plantas terrestres evoluíram de algas verdes e que durante esta evolução, vários rearranjos ocorreram no genoma dos cloroplastos (Graham, 1996). Portanto, para compreender o processo de evolução do genoma do cloroplasto, informações sobre sequências repetidas, regiões intergênicas e pseudogenes são extremamente úteis (Wakasugi et al., 1997).

O tamanho e a sintenia dos genomas dos plastídeos variam muito entre os grupos de algas (Kim et al., 2014). O genoma do cloroplasto da alga verde *Floydiella terrestris* (Chlorophyceae) com 521.168 pb é o maior genoma de plastídeo já sequenciado, apresentando 97 genes, 26 íntrons e teor A+T de 65,5% (Brouard et al., 2010) (figura 3A).

N. olivacea exibe um genoma plastidial de 200.799 pb e apresenta o maior repertório gênico relatado até o momento para uma Chlorophyta, 128 genes, enquanto que a maioria dos genomas dos cloroplastos de Chlorophyta apresentam um conjunto reduzido de 86 a 88 genes (Lemieux; Otis; Turmel, 2014a).

A disponibilidade de dados sobre genomas de organelas em Trebouxiophyceae é ainda limitada (Jeong et al., 2014). Contudo, informações retiradas do *Organelle Genome Resources* do NCBI (<http://www.ncbi.nlm.nih.gov/genomes/GenomesHome.cgi?taxid=2759>) mostram que atualmente 29 genomas plastidiais de algas verdes Trebouxiophyceae estão disponíveis (tabela 1). Adicionalmente, há pelo menos outros três genomas não inclusos na plataforma do NCBI, os de *Chlorella sp.* ArM0029B, (Jeong et al., 2014), *Prasiolopsis sp.* (Lemieux; Otis; Turmel, 2014b), *Prototheca wickerhamii* (Yan et al., 2015), (tabela 2), totalizando-se 32 genomas plastidiais completos sequenciados desta classe de algas verdes.

O conteúdo gênico e número de íntrons de espécies desta classe é bastante diversificado (deCambiaire et al., 2007). Além disto, estes genomas exibem uma variabilidade de tamanhos de 306.152 pb de *Prasiolopsis sp.* (Lemieux; Otis; Turmel, 2014b) a 37.454 pb da alga verde parasita não fotossintética *Helicosporidium sp.*, que codifica 26 proteínas, 3 rRNAs e 25 tRNA (figura 3B), carecendo de todos os genes envolvidos com a fotossíntese. Esta última possui apenas um íntron no grupo I do gene tRNA^{Leu} (UAA) e espaços intergênicos minúsculos. Ainda assim, a densidade gênica é alta, com apenas 5,1% de DNA não codificante. *Helicosporidium sp.* possui o menor genoma plastidial sequenciado de qualquer Viridiplantae conhecido (deKoning; Keeling, 2006).

A ampla quantidade de informações contida no genoma dos plastídeos tem demonstrado que o material genético desta organela pode ser empregado como uma ferramenta adequada e de valor inestimável para a filogenia molecular (Wang et al., 2013). Análises filogenéticas inferidas a partir do gene nuclear 18S rRNA muitas vezes são incapazes de desvendar as inter-relações das linhagens de algas verdes. Contudo, filogenias com base na análise de genomas plastidiais são bem sucedidas na resolução de questões distintas referentes às relações de algas (Brouard et al., 2010). Além de oferecer um grande

conjunto de genes para análises filogenômicas, os genomas de cloroplastos revelam as características estruturais do genoma para validação das filogenias (Turmel et al., 2008).

Tabela 1 – Exemplos da classe Trebouxiophyceae com genoma plastidial completo sequenciado.

Espécie	Acesso NCBI	Tamanho do genoma (pb)
<i>Auxenochlorella protothecoides</i>	NC_023775	84.576
<i>Botryococcus braunii</i>	NC_025545	172.826
<i>Chlorella sorokiniana</i>	NC_023835	109.811
' <i>Chlorella</i> ' <i>mirabilis</i>	NC_025528	167.972
<i>Chlorella variabilis</i>	NC_015359	124.579
<i>Chlorella vulgaris</i>	NC_001865	150.613
<i>Choricystis parasitica</i>	NC_025539	94.206
<i>Coccomyxa</i> sp.	NC_015084	175.731
<i>Diclosteraquatus</i>	NC_025546	169.201
<i>Dictyochloropsis reticulata</i>	NC_025524	289.394
<i>Elliptochloris bilobata</i>	NC_025548	134.677
<i>Fusochloris perforata</i>	NC_025543	148.459
<i>Helicosporidium</i> sp.	NC_008100	37.454
<i>Koliella longiseta</i>	NC_025531	197.094
<i>Leptospira terrestres</i>	NC_009681	195.081
<i>Lobosphaera incisa</i>	NC_025533	156.031
<i>Marvania geminata</i>	NC_025549	108.470
<i>Microthamnion kuetzingianum</i>	NC_025537	158.609
<i>Myrmecia israelensis</i>	NC_025525	146.596
<i>Neocystis brevis</i>	NC_025535	211.747
<i>Pabia signiensis</i>	NC_025529	236.463
<i>Parachlorella kessleri</i>	NC_012978	123.994
<i>Paradoxia multiseta</i>	NC_025540	183.394
<i>Planctonema lauterbornii</i>	NC_025541	114.128
<i>Pseudochloris wilhelmii</i>	NC_025547	109.775
<i>Stichococcus bacillaris</i>	NC_025527	116.952
<i>Trebouxiophyceae</i> sp	NC_018569	149.707
<i>Watanabea reniformis</i>	NC_025526	201.425
<i>Xylochloris irregularis</i>	NC_025534	181.542

Fonte: <http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt=plastid&taxid=3041> modificado pelo autor.

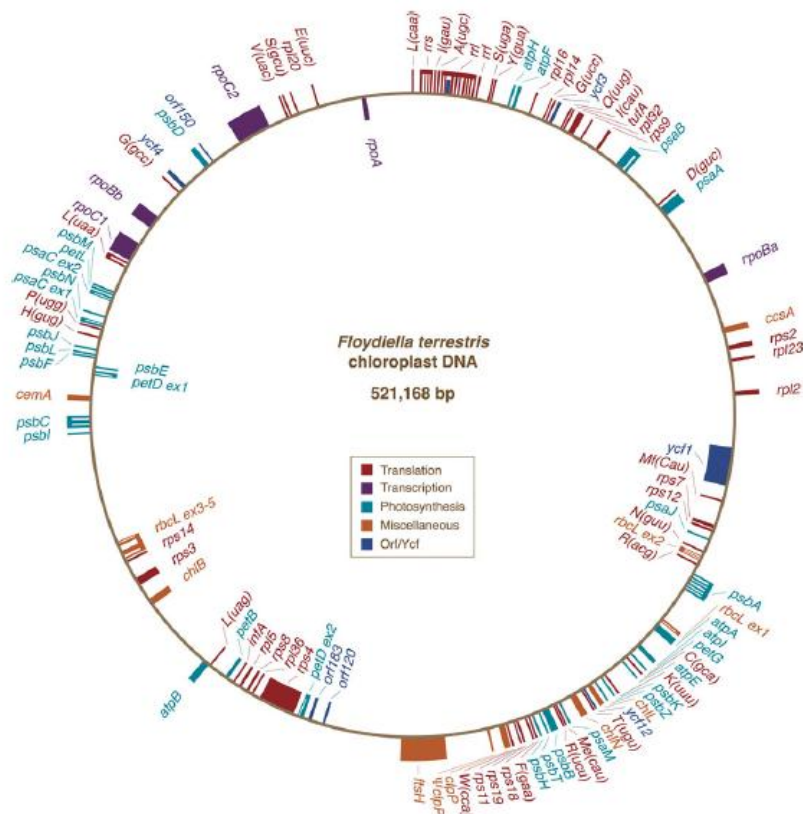
Tabela 2 – Espécies da classe Trebouxiophyceae com genoma plastidial sequenciado não depositado na plataforma *Organelle Genome Resources* (NCBI).

Espécie	Acesso NCBI	Tamanho do genoma (pb)
<i>Chlorella sp. ArM0029B</i>	KF554427.1	119.989
<i>Prasiolopsis sp.</i>	KM462862.1	306.152
<i>Prototheca wickerhamii</i>	KJ001761.1	55.636

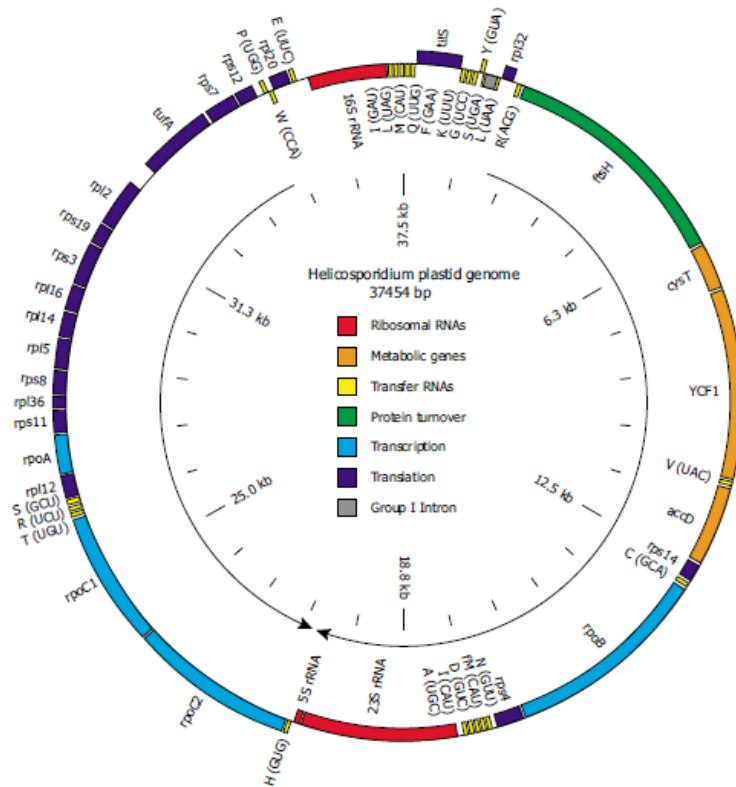
Fonte: do autor.

Figura 3 - Mapas representativos dos genomas plastidiais das algas verdes *Floydiella terrestris* (Chlorophyceae) e *Helicosporidium sp.* (Trebouxiophyceae). (A) *F. terrestris* apresenta o maior genoma cloroplastial já sequenciado até o momento, com 521.168 pb (B) contrariamente ao que é observado em *Helicosporidium sp.*, que com 37.454 pb possui o menor genoma de plastídeos já sequenciado do Viridiplantae.

A



B



Fonte: Brouard et al.; 2010; deKoning; Keeling, 2006.

Em recente estudo, Lemieux, Otis e Turmel (2014b) valeram-se das sequências de aminoácidos de 79 genes codificantes de proteínas de 63 espécies de algas verdes para análises filogenômicas entre as principais linhagens de Trebouxiophyceae.

O sequenciamento completo de genomas plastidiais com tamanho relativamente pequeno (aproximadamente 150 kb) foram tecnicamente possíveis desde meados da década de 1980. Com uso das tecnologias de sequenciamento de nova geração (NGS), o custo e a dificuldade para o sequenciamento de genomas plastidiais foram drasticamente reduzidos o que consequentemente aumentou o número de genomas plastidiais disponíveis (Ruhfel et al., 2014).

Os dados da sequência do genoma dos plastídeos transformaram a sistemática vegetal e contribuíram para visão atual das relações entre plantas (Ruhfel et al., 2014), oferecendo uma riqueza de dados filogeneticamente informativos que são relativamente fáceis de obter e utilizar (Wicke et al., 2007), se tornando uma abordagem valiosa para inferir relações entre os eucariontes fotossintetizantes (Lemieux; Otis; Turmel, 2014b).

1.3. Análise de genomas mitocondriais como ferramenta filogenética

Mitocôndrias são onipresentes em células eucarióticas e executam um gama de funções celulares essenciais (Douce, 1985). Consideradas a “casa de força” das células, fornecem a energia necessária para as atividades celulares (Hammani; Giege, 2014), mas também estão envolvidas nos processos de sinalização celular, regulação da proliferação celular, diferenciação e sobrevivência (Merrill; Strack, 2014). Nas plantas, além de prover a energia celular e a respiração, estas organelas estão envolvidas em outras vias metabólicas incluindo a assimilação do nitrogênio, fotorrespiração, metabolismo do dióxido de carbono, fotossíntese em plantas C4, metabolismo ácido das crassuláceas, armazenamento de carbono e nitrogênio durante a germinação das sementes (Douce, 1985) e também desempenham um papel na biossíntese de aminoácidos e ácidos graxos (Picault et al., 2004).

Supõe-se que as mitocôndrias originaram-se a partir de um procarioto de vida livre consumidor de oxigênio que por endossimbiose foi capturado por outra célula hospedeira procariótica. Durante a simbiose na célula hospedeira, o genoma mitocondrial de plantas sofreu uma perda massiva de conteúdo gênico, onde parte foi transferido para o núcleo da célula hospedeira, codificando apenas um conjunto parcial dos componentes das membranas de transdução de energia e componentes da maquinaria de expressão gênica (Hammani; Giege, 2014). Um bom exemplo desta perda é o genoma mitocondrial da Streptophyta *Marchantia polymorpha*, o qual carece de alguns tRNAs, proteínas ribossomais e subunidades do complexo succinato-ubiquinona redutase (complexo respiratório II), o que condiz com a hipótese de transferência dos genes mitocondriais para o núcleo ao longo da evolução (Schuster; Brennicke, 1994).

Dois padrões distintos de evolução do mtDNA foram atribuídos as espécies do filo Chlorophyta (Pombert et al., 2006). Os genomas com um padrão ancestral de evolução mantiveram vestígios claros da sua ancestralidade eubacteriana (Gray; Burger; Lang, 1999) com a estrutura do genoma compacta, repertório gênico substancial e sequências gênicas conservadas. Em contraste, o padrão derivado tem sido atribuído aos mtDNA de Chlorophyta que se afastam radicalmente do padrão ancestral, com pouca ou nenhuma evidência dos traços primitivos, apresentando uma perda extensa de genes, gerando um genoma diminuto em tamanho e conteúdo gênico, divergência acentuada no DNA ribossomal e estrutura do rRNA (manifestada no truncamento da sequência e fragmentação dos genes do rRNA) e acelerada

taxa de divergência da sequência, tanto em codificação de proteínas e genes do rRNA (Gray; Burger; Lang, 1999; Pombert et al., 2006).

O mtDNA da alga verde Prasinophyceae *Nephroselmis olivacea* (Turmel et al., 1999) e da Trebouxiophyceae *Prototheca wickerhamii* (Wolff et al., 1994) apresentam o padrão ancestral de evolução, enquanto que o mtDNA da Chlorophyceae *Chlamydomonas reinhardtii* (Michaelis; Vahrenholz; Pratje, 1990) e da Pedinophyceae *Pedinomonas minor* (Turmel et al., 1999) foram classificados como sendo do padrão derivado (Pombert et al., 2006). Assim, os dados disponíveis apontam para a existência de padrões distintos da evolução do genoma mitocondrial, não só entre Streptophyta e Chlorophyta, mas também dentro do filo Chlorophyta (Turmel et al., 1999).

As mitocôndrias de Streptophyta e Chlorophyta adquiriram características únicas que as diferem do seu ancestral procariótico assim como das mitocôndrias de outros reinos (Hammani; Giege, 2014). Em Chlorophyta o mtDNA sofreu mudanças radicais na estrutura, conteúdo gênico, organização dos genes, quantidade de íntrons e são altamente variáveis em tamanho (Lewis; McCourt, 2004; Pombert et al., 2004).

De acordo com dados do *National Center for Biotechnology Information* (NCBI), até o momento, quarenta genomas de mitocôndrias de Chlorophyta foram sequenciados, sendo sete exemplares da classe Trebouxiophyceae (Organelle Genome Resources) (tabela 3). Além destes, o genoma mitocondrial de *Chlorella sp.* ArM0029B também encontra-se sequenciado, apresentando 65.049 pb de tamanho e depositado na plataforma GenBank do NCBI sobre o número de acesso KF554428.1 (Jeong et al., 2014).

Tabela 3 – Exemplares da classe Trebouxiophyceae com genoma mitocondrial sequenciado.

Espécie	Acesso NCBI	Tamanho do genoma (pb)
<i>Auxenochlorella protothecoides</i>	NC_026009	57.274
<i>Chlorella sorokiniana</i> isolate 1230	NC_024626	52.528
<i>Chlorella variabilis</i> isolate NC64A	NC_025413	78.500
<i>Coccomyxa sp.</i> C-169	NC_015316	65.497
<i>Helicosporidium sp.</i> ex <i>Simulium jonesi</i>	NC_017841	49.343
<i>Prototheca wickerhamii</i>	NC_001613	55.328
<i>Trebouxiophyceae sp.</i> MX-AZ01	NC_018568	74.423

Fonte: <http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt=organelle&taxid=3041> modificado pelo autor.

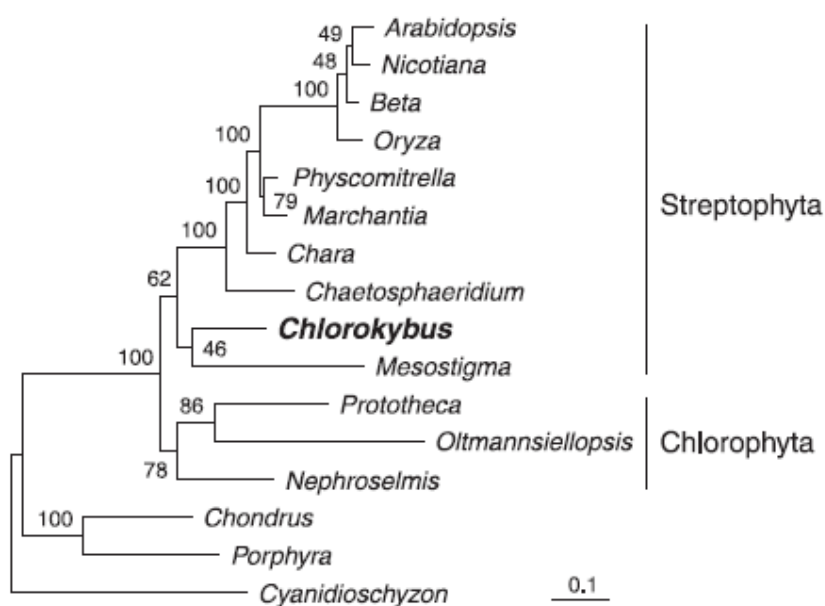
Genomas mitocondriais contém um conjunto limitado de genes que codificam proteínas e RNAs (Gray; Burger; Lang, 1999). As mitocôndrias que possuem os componentes clássicos da fosforilação oxidativa, ou seja, os complexos respiratórios I (NADH-ubiquinona redutase), II (succinato-ubiquinona redutase), III (ubiquinol-citocromo *c* redutase) e IV (citocromo *c* oxidase) e ATP sintase F1-F0 (complexo V) que normalmente contém os genes *atp6*, *atp8*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* e *nad6* no mtDNA. No entanto, o mtDNA de vários exemplares da linhagem Chlorophyceae e afins normalmente, não exibem os genes *atp6*, *atp8*, *cox2*, *cox3*, *nad3* e *nad4L* (Fan; Lee, 2002; Denovan-Wright et al., 1998; Vahrenholz, 1993). O sistema de tradução das organelas, pelo qual os mRNA mitocondriais são codificados, também é composto em parte por componentes especificados pelo mtDNA. Em plantas, protistas e na maioria dos fungos, algumas proteínas ribossômicas são codificadas no mtDNA (Gray; Burger; Lang, 1999).

O mtDNA é considerado um marcador molecular útil para análises filogenéticas e identificação de espécies. Recentemente, o mtDNA também tem sido utilizado como DNA *barcoding* para identificação a nível de espécies e populações (Ko et al., 2013).

Nas últimas três décadas, genes individuais do mtDNA, particularmente o gene que codifica uma citocromo *c* oxidase (*cox1*), foram utilizados para desvendar as relações filogenéticas (Hajibabaei et al., 2007). Análises tendo como referência um único gene geraram filogenias atualmente aceitas, no entanto, o poder de resolução da análise de um único gene é limitada pelo pequeno conteúdo de informações obtidas, o que no caso particular das mitocôndrias, é ainda mais afetado pelas extremas diferenças na composição gênica e na taxa de divergência da sequência do mtDNA de diferentes linhagens eucarióticas (Gray; Burger; Lang, 1999).

As informações obtidas a partir do sequenciamento do mtDNA completos estão sendo utilizadas por pesquisadores para o desenvolvimento de filogenias. A fim de determinar a posição filogenética de *Chlorokybus atmophyticus*, uma Streptophyta, foi analisado um conjunto de aminoácidos obtidos de dezoito genes que codificam proteínas (*atp4*, *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *mttB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rps3* e *rps12*), comuns à dezesseis espécies de algas verdes e plantas terrestres (figura 5) (Turmel; Otis; Lemieux, 2007). Sequências de aminoácidos traduzidas dos genes codificadores de proteínas *cob*, *cox1*, *nad1*, *nad2*, *nad4*, *nad5* e *nad6*, foram aplicados para estudos filogenéticos da alga verde trebouxiophyceae *Chlorella* sp. ArM0029B (Jeong et al., 2014).

Figura 5 - Posição filogenética *Chlorokybus atmophyticus* dentro do reino Viridiplantae. Análise filogenética embasada nos genes mitocondriais codificadores de proteínas *atp4*, *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *mttB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rps3* e *rps12*.



Fonte: Turmel; Otis; Lemieux, 2007.

As sequências codificantes de proteínas e informações do mtDNA podem auxiliar a desvendar relações filogenéticas que sequências de genes nucleares seriam incapazes de resolver. Genomas mitocondriais compreendem um repositório de genes codificadores de proteínas, cuja origem está bem estabelecida e cuja evolução parece acompanhar a do hospedeiro eucarioto. A determinação de uma maior variedade de sequências de genomas mitocôndrias de algas verdes poderá permitir uma reconstrução mais rigorosa de filogenias embasadas em sequências de aminoácidos de genes codificantes de proteínas (Gary; Burger; Lang, 1999).

2. OBJETIVOS

2.1. Objetivo Geral

Este trabalho tem como objetivo determinar a sequência do genoma do cloroplasto e o genoma mitocondrial da alga Antártica *Prasiola crista* visando inferir as relações evolutivas deste organismo com outras espécies de plantas verdes, além de analisar a estruturas destes genomas acessórios.

2.2. Objetivos Específicos

- Sequenciar, montar e anotar o genoma do cloroplasto de *Prasiola crista*;
- Sequenciar, montar e anotar o genoma mitocondrial de *Prasiola crista*;
- Com base nas sequencias genômicas realizar uma análise evolutiva (filogenômica) de *Prasiola crista*;
- Analisar a estrutura dos genomas do clado *Prasiola*.

3. ARTIGO 1

O artigo intitulado "Draft plastid and mitochondrial genomes from Antarctic algae *Prasiola crisper*" está *no prelo* para publicação no periódico Genome Announcements (ISSN: 2169-8287).



Evelise Leis Carvalho <eveliseleis@gmail.com>

Final Decision made for genomeA00435-15

1 mensagem

asm@msubmit.net <asm@msubmit.net>

17 de abril de 2015

14:11

Responder a: aesplin@asmusa.org, genomeAjournal@gmail.com

Para: eveliseleis@gmail.com

Dear Mrs. Carvalho:

Here is a copy of the decision letter for manuscript "Draft plastid and mitochondrial genomes from Antarctic algae *Prasiola crisper*" by Evelise Carvalho, Gabriel Wallau, Darlene Rangel, Laís Machado, Alexandre Silva, Antonio Pereira, Filipe Victoria, Juliano Boldo, Cháriston Dal Belo, and Paulo Pinto (genomeA00435-15), for which you were a contributing author.

Sincerely,
Genome Announcements

Subject: genomeA00435-15 Decision Letter

Dr. Paulo Marcos Pinto
Universidade Federal do Pampa
Av. Antonio Trilha, 1847
II Piso - Sala 7
São Gabriel, RS 97300000
Brazil

Re: genomeA00435-15 (Draft plastid and mitochondrial genomes from Antarctic

algae *Prasiola crisper*)

Dear Dr. Pinto:

Your manuscript has been accepted, and I am forwarding it to the ASM Journals Department for publication in Genome Announcements (genomeA). For your reference, ASM Journals' address is given below. Before it can be scheduled for publication, your manuscript must be checked by ASM production staff to make sure that all elements meet the technical requirements for publication. Amber Esplin, assistant production editor for genomeA, will contact you if anything needs to be revised before copyediting and production can begin. Otherwise, you will be notified when your proofs are ready to be viewed.

ASM charges a flat fee for Genome Announcements. The 2014 fee is \$330 for a corresponding author who is an active ASM member at the Contributing or Premium level or \$500 for a nonmember or Supporting member corresponding author. A corresponding author who is not a Contributing or Premium member may [upgrade or join ASM](#) to obtain the member rate. You will receive payment instructions by email along with the article proofs. Please note that waivers of publication charges are not considered for Genome Announcements.

Thank you for submitting your paper to genomeA.

Sincerely,
Philip Matsumura
Editor, Genome Announcements

Journals Department
American Society for Microbiology
1752 N St., NW
Washington, DC 20036
E-mail: aesplin@asmusa.org
Phone: [1-202-942-9336](tel:1-202-942-9336)
Fax: [1-202-942-9355](tel:1-202-942-9355)

Draft Plastid and Mitochondrial Genome Sequences from the Antarctic Alga *Prasiola crispa*

Evelise Leis Carvalho,^a Gabriel da Luz Wallau,^a Darlene Lopes Rangel,^a Laís Ceschini Machado,^a Alexandre Freitas da Silva,^a Antonio Batista Pereira,^b Filipe de Carvalho Victoria,^b Juliano Tomazzoni Boldo,^a Cháriston André Dal Belo,^a Paulo Marcos Pinto^a

Applied Proteomics Laboratory, University of Pampa, São Gabriel, Rio Grande do Sul, Brazil^a; Antarctic Studies Plant Core, National Institute of Antarctic Science and Technology for Environmental Research, University of Pampa, São Gabriel, Rio Grande do Sul, Brazil^b

The organelle genomes of the Antarctic alga *Prasiola crispa* (Lightfoot) Kützinger have been sequenced. The plastid and mitochondrial genomes have total lengths of 196,502 bp and 89,819 bp, respectively. These genomes have 21 putative photosynthesis-related genes and 18 oxidative metabolism-related genes, respectively.

Received 27 March 2015 Accepted 17 April 2015 Published XXX

Citation Leis Carvalho E, da Luz Wallau G, Lopes Rangel D, Ceschini Machado L, Freitas da Silva A, Batista Pereira A, de Carvalho Victoria F, Tomazzoni Boldo J, André Dal Belo C, Pinto PM. 2015. Draft plastid and mitochondrial genome sequences from the Antarctic alga *Prasiola crispa*. *Genome Announc* 3(3):e00435-15. doi:10.1128/genomeA.00435-15.

Copyright © 2015 Leis Carvalho et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Paulo Marcos Pinto, paulopinto@unipampa.edu.br.

Antarctica has attracted considerable interest from biologists interested in understanding the evolutionary adaptation of extremophile organisms (1). *Prasiola* spp. are the best-known Antarctic algae found at many terrestrial and supralittoral sites, representing the most important primary producers (1–3). The species most commonly reported is *Prasiola crispa* (Lightfoot) Kützinger. *P. crispa* has several interesting biological features, typically growing on moist soils fertilized by penguin guano (4), and it tolerates repeated thaw cycles in the spring and fall, freezing over winter, and high levels of UV radiation during the summer (5).

The organelle DNA was sequenced by the Macrogen Service Center in a Solexa-Illumina HiSeq 2500 next-generation sequencing device, according to the manufacturer's instructions. A paired-end approach with a read size of ~100 bp was employed. Sequence assembly was performed with the SOAPdenovo2 software version 2.01 (6). All open reading frames (ORFs) were annotated using CpGAVAS (7) and Mitofy (8) for the chloroplast and mitochondrial genomes, respectively.

The *P. crispa* plastid genome (cpDNA) information resides on a single molecule with a total length of 196,502 bp and a G+C content of 29.32%. Compared with other species from the *Prasiola* clade, the genomes of *P. crispa* are among the largest (*Prasiolopsis* sp., 306.1 kb; *Pabia signiensis* T. Friedl & O'Kelly, 236.5 kb; *Koliella longiseta* (Vischer) Hindák, 197.1 kb; *Stichococcus bacillaris* Nägeli, 116.9 kb) (9) or even compared to those of other species close to *Prasiolales*, such as *Chlorella mirabilis* V. M. Andreyeva (168.0 kb) (9). The cpDNA comprises 32 putative coding genes, 27 tRNAs, and 2 rRNAs. Among those, we were able to annotate at least 21 putative protein-coding genes related to photosynthesis, such as photosystems I and II (*ycf3*, *psaB*, *psaC*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbH*, *psbL*, *psbM*, *psbN*, and *psbT*).

The *P. crispa* mitochondrial genome (mtDNA) is the first mitochondrial sequenced genome from the *Prasiolales* order. It has a total of 89,819 bp and a G+C content of 29.29%; compared with other species from the *Trebouxiophyceae* class, *P. crispa* has the largest mtDNA (*Trebouxiophyceae* sp., 74.4 kb; *Chlorella* sp.

ArM0029B, 65.0 kb; *Oltmannsiellopsis viridis* [P. E. Hargraves & R. L. Steele] M. Chihara & I. Inouye, 56.8 kb; *Prototheca wickerhamii* Tubaki & Soneda, 55.3 kb; and *Chlorella sorokiniana* Shihira & R. W. Krauss, 52.5 kb) (10–14). The *P. crispa* mtDNA genome has the smallest gene content, with only 43 genes, comprising 18 putative protein-coding genes, 22 tRNAs, and 3 rRNAs. All of the annotated protein-coding genes were related to mitochondrial oxidative metabolism, such as mitochondrial respiratory chain complexes I, III, and IV (*nad1*, *nad2*, *nad3*, *nad4*, *nad5*, *nad6*, *nad7*, *cobA*, and *cox1*).

These organelle genomes are the first draft genome sequences obtained from Antarctic *Trebouxiophyceae* algae. Mitochondria and chloroplast organelles play a pivotal role in energy metabolism and are great tools for taxonomic analysis in higher plants and algae. The plastid and mitochondrial genome data would be useful for further genetic studies, phylogenetic analysis, and resource protection of *P. crispa* and phylogenetic analysis of *Trebouxiophyceae* green algae.

Nucleotide sequence accession numbers. The cpDNA and mtDNA contig sequences were deposited in GenBank under accession numbers KR017748, KR017749, and KR017750, and KR017746 and KR017747, respectively.

ACKNOWLEDGMENTS

This work was supported by the National Council for Scientific and Technological Development (CNPq-Brazil), the Coordination for the Improvement of Higher Education Personnel (CAPES-Brazil), the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS-Brazil), and the National Institute of Science and Technology - Antarctic Environmental Research (INCT-APA). E. Leis Carvalho received research fellowships from Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS-Brazil).

REFERENCES

1. Convey P. 2007. Biogeography, p 154–160. In Riffenburgh, B (ed), *Encyclopedia of the Antarctic*, vol 1, Routledge, New York, NY.
2. Kováčik L, Pereira AB. 2001. Green alga *Prasiola crispa* and its lichenized

- form *Mastodia tessellata* in Antarctic environment: general aspects. *Nova Hedwigia Beiheft* 123:465–478.
3. Wiencke C, Clayton MN 2002. In Antarctic seaweeds. Synopses of the Antarctic benthos, vol 9. A. R. G. Gantner Verlag K. G., Ruggell, Liechtenstein.
 4. Graham LE, Graham JM, Wilcox LW. 2009. *Algae*. Pearson Education, San Francisco, CA.
 5. Kosugi M, Katashima Y, Aikawa S, Tanabe Y, Kudoh S, Kashino Y, Koike H, Satoh K. 2010. Comparative study on the photosynthetic properties of *Prasiola* (*Chlorophyceae*) and *Nostoc* (*Cyanophyceae*) from Antarctic and non-Antarctic sites. *J Phycol* 46:466–476. <http://dx.doi.org/10.1111/j.1529-8817.2010.00831.x>.
 6. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
 7. Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated Web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* 13:715. <http://dx.doi.org/10.1186/1471-2164-13-715>.
 8. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (*Cucurbitaceae*). *Mol Biol Evol* 27:1436–1448. <http://dx.doi.org/10.1093/molbev/msq029>.
 9. Lemieux C, Otis C, Turmel M. 2014. Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class *Trebouxiophyceae*. *BMC Evol Biol* 14:211. <http://dx.doi.org/10.1186/s12862-014-0211-2>.
 10. Wolff G, Plante I, Lang BF, Kück U, Burger G. 1994. Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. Gene content and genome organization. *J Mol Biol* 237:75–86. <http://dx.doi.org/10.1006/jmbi.1994.1210>.
 11. Pombert JF, Beauchamp P, Otis C, Lemieux C, Turmel M. 2006. The complete mitochondrial DNA sequence of the green alga *Olmamsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the *Ulvo-phyceae*. *Curr Genet* 50:137–147. <http://dx.doi.org/10.1007/s00294-006-0076-z>.
 12. Servín-Garcidueñas LE, Martínez-Romero E. 2012. Complete mitochondrial and plastid genomes of the green microalga *Trebouxiophyceae* sp. strain MX-AZ01 isolated from a highly acidic geothermal lake. *Eukaryot Cell* 11:1417–1418. <http://dx.doi.org/10.1128/EC.00244-12>.
 13. Jeong H, Lim JM, Park J, Sim YM, Choi HG, Lee J, Jeong WJ. 2014. Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B. *BMC Genomics* 15:286. <http://dx.doi.org/10.1186/1471-2164-15-286>.
 14. Orsini M, Costelli C, Malavasi V, Cusano R, Concas A, Angius A, Cao G. 2014. Complete genome sequence of mitochondrial DNA (mtDNA) of *Chlorella sorokiniana*. *Mitochondrial DNA* 4:1–3. <http://dx.doi.org/10.3109/19401736.2014.953128>.

4. ARTIGO 2

O artigo intitulado como “Evolutionary and structural analyses of accessories genomes from an Antarctic algae *Prasiola crisper*” é um manuscrito que está disposto na forma na qual deverá ser submetido para o periódico BMC Plant Biology (ISSN:1471-2229).

Evolutionary and structural analyses of accessories genomes from an Antarctic algae *Prasiola crisper*

Evelise Leis Carvalho^{1,2}

Email: eveliseleis@gmail.com

Gabriel da Luz Wallau^{1,2}

Email: gabriel.wallau@gmail.com

Darlene Lopes Rangel¹

Email: darlenelrangel@gmail.com

Luiz Fernando Duarte¹

Email: luizfernandoduarte92@hotmail.com

Laís Ceschini Machado¹

Email: laisceschini@gmail.com

Antonio Batista Pereira^{2,3}

Email:antoniopereira@unipampa.edu.br

Filipe de Carvalho Victoria^{2,3}

Email: filipevictoria@gmail.com

Cháriston André Dal-Belo^{2,4}

Email:charistonbelo@unipampa.edu.br

Paulo Marcos Pinto^{1,2,*}

* Corresponding author

paulopinto@unipampa.edu.br

¹Applied Proteomics Laboratory, University of Pampa, São Gabriel, RS, Brazil;

² Post Graduate in Biological Sciences, University of Pampa, São Gabriel, RS, Brazil;

³Studies Plant Core, National Institute of Antarctic Science and Technology for Environmental Research, University of Pampa, São Gabriel, RS, Brazil

⁴LANETOX, University of Pampa, São Gabriel, RS, Brazil

*Correspondence to Av. Antonio Trilha, 1847. ZIP Code: 97300-000. São Gabriel, RS, Brazil.

Abstract

Background

Antarctica is one of the most difficult habitats for sustaining life on earth, organisms that live there can develop different strategies for survival. Within this extremophiles is the green algae *Prasiola crispera*, belonging to class trebouxiofiace. The literature on *P. crispera* taxonomy is rare and leaves many gaps in the evolutionary relationship with closest relatives. A strategy to

solve taxonomic problems faithfully is the use of chloroplasts and mitochondrial genomes information. The aim of this study was to analyze and establish the evolutionary and structural patterns of *P. crispera* using plastid and mitochondrial genomes.

Results

Seventy chloroplast genomes and twenty-seven mitochondrial genomes were analyzed for a phylogenomics study. Based on the plastid genome analysis of *P. crispera* grouped with *Prasiolopsis sp.* and *Stichoccus bacillaris* forming the *Prasiola* clade. In the mitochondrial genome analysis *P. crispera* grouped with other trebouxiofrican analyzed. The structure of *P. crispera* chloroplast genome has low synteny with *Prasiolopsis sp.* with some conserved genes blocks. The same can be observed in the mitochondrial genome compared to *Coccomixia sp.* In addition, this structure with many rearrangements and low synteny was observed in all analyzed genomes even when used species of the same genus as *Chorella sp.* ArM0029B and *Chlorella vulgaris*.

Conclusions

We can establish the phylogenetic position of *P. crispera* with other species of algae Trebouxiofrican using accessories genomes. In addition, we describe the plasticity of these genomes by a structural analysis.

The Plastid and Mitochondrial genomes data would be useful for further genetics studies, phylogenetic analysis, and resource protection of *P. crispera* and phylogenetic analysis of Trebouxiofrican green algae.

Keywords

Prasiola crispera, mitochondrial genome, chloroplast genome, phylogeny, antarctic algae.

Background

With only 0.32% of its surface ice-free, Antarctica is the coldest, windiest, and driest continent. The terrestrial ecosystems and climatic characteristics of the three zones, Sub-Antarctic, Maritime Antarctic and Continental Antarctic are distinctly different. The Continental Antarctic zone is subject to harsher conditions and more extreme temperatures, and is characterized by lower species diversity than the other two zones [1]. Due to its extreme characteristics and limited human impact, Antarctica has attracted considerable interest from biologists. However, logistical difficulties related to sampling and other fieldwork activities has hindered several studies about Antarctica biological diversity [2].

In the Antarctic continent, microorganisms including prokaryotes and eukaryotes (fungi, algae, protists) thrive in ice, cold water, even within rocks where they protect themselves from desiccation and UV radiation [3]. Trebouxiophyceae green algae of the genus *Prasiola* are among the best known Antarctic algae and at many terrestrial and supralittoral sites represent the most important primary producers [1,4,5].

The species most commonly reported is *Prasiola crispa* (Lightfoot) Kützing. This nitrophilic macroalga typically grows on moist soils that are fertilized by bird guano [6]. *P. crispa* is widely distributed in Continental Antarctica (Upper Garwood Valley and Marshall Valley) and Maritime Antarctica (Torgensen I., Saunders I. and King George I.) [2]. It tolerates repeated freeze/thaw cycles in spring and fall and high levels of UV radiation during summer [7, 8].

The study of accessory organelles genomes can provide a major impact in the fields of evolution, biology and biotechnology [9] and hence can help us to understand which evolutionary processes generated adaptation to extreme environments at the molecular level.

In addition, such genetic information also provides many molecular characters (gene order, nucleotide and amino acid sequences) that can be used for estimating phylogenetic trees and discover the positioning of some species in a broader phylogenetic context [10,11]. The phylogenetic positioning of extremophiles species, as *P. crista*, is of extreme importance to understand if its adaptations are unique or shared among other close related species. What can indicate when and in which ancestral organisms such adaptations appeared in the evolutionary history of green algae [2].

The chloroplast phylogenomic studies of green algae reported so far have provided valuable insights into the phylogeny of Prasiophytes [12, 13], Streptophytes [14, 15] and Chlorophyceae [16, 17], but only limited information is currently available regarding the relationships within the Trebouxiophyceae [18].

Complete mitochondrial genomes (mtDNA) have been increasingly used to address phylogenetic questions where multi-gene analyses have been either unresolved or poorly supported [19]. Since next-generation sequencing technologies have become more accessible, many algal genomics researchers have used these techniques to generate sequence data for organellar genomes [9, 20-23]. However, relative to the total number of algae, few mitochondrial and plastid genomes have been sequenced, with many large taxonomic groups unrepresented [9].

Investigations on the taxonomy of Antarctic representative of *Prasiola* genus are nonexistent at the molecular level in recent literature [2]. In addition, studies on molecular characteristics that are involved in extreme environments adaptations are still lacking. Instigated by those open questions we have studied the evolutionary relationship of *P. crista* (with organellar genomes recently sequenced by our group) and other Trebouxiophycian alga using a chloroplast genome (cpDNA) and mtDNA phylogenomic approach. In addition, we

performed a structural analysis and evaluated which groups of genes are missing in different green algae clades in order to understand its adaptation to different environments.

Results and Discussion

Species of Trebouxiophyceae and amino acid datasets

Searches in the Organelle Genomes Resource and NCBI database resulted in an addition of 6 new species with complete chloroplast genomes to the amount of species of Trebouxiophyceae analysed by Limeux et al. 2014. All new analyzed genomes can be seen in Figure 1 (*Auxenochlorella protothecoides* - NC_023775, *Chlorella sorokiniana* - NC_023835, *Chlorella sp.* - KF554427.1, *Coccomyxa sp.* - NC_015084, *Helicosporidium sp.* - NC_008100 and *Prototheca wickerhamii* - KJ001761.1) for chloroplast genomes.

Searches in the Organelle Genomes Resource database and BLASTn analyses resulted in an addition of 7 species with complete mitochondrial genomes to the amount of species of Trebouxiophyceae analysed by Turmel et al. 2007. All new analyzed genomes can be seen in Figure 2 (*Auxenochlorella protothecoides* - NC_026009, *Chlorella sorokiniana* - NC_024626, *Chlorella sp.* - KF554427.1, *Chlorella variabilis* - NC_025413, *Coccomyxa sp. C-169* - NC_015316, *Helicosporidium sp.* - NC_017841 and *Trebouxiophyceae sp.* - NC_018568) for mitochondrial genomes.

Seventy-nine coding genes were used for the chloroplast phylogenomic analysis that after Gblocks analysis resulted in a final alignment with 14460 amino acids. For the mitochondrial dataset eighteen genes were used resulting in a final alignment 2516 amino acids.

Phylogenetic positioning of *P. crispera* to other trebouxiophycean algae

Among the new species of Trebouxiophyceae analyzed in this work all 7 species, including *P. crispera* (sequenced by our group) and 6 species obtained from NCBI website, has an unknown phylogenetic positioning based on molecular data.

Chloroplast phylogenomic analysis

Overall the phylogenomic analysis presented high branch support all over the tree, with some few exception as: a basal node grouping Prasinophyceae, Pedinophyceae, Trebouxiophyceae, Ulvophyceae and Chlorophyceae (aLRT-like = 0.604); the basal positioning of *Parietochloris pseudoalveolaris* in relation to the *Prasiola* clade and *Neocystis brevis* (aLRT-like = 0.716); and the grouping of *Pabia signensis*/*Kaliella longiseta* and '*Chlorella*' *mirabilis* (aLRT-like = 0.355) (Figure 1).

Based on the chloroplast phylogenomic analysis, *P. crispera* grouped consistently (aLRT-like branch support of 1) as a sister species of *Prasiolopsis sp.*, a species of the *Prasiola* clade, which includes more four species: *Stichoccus bacillaris*, *Pabia signensis*, *Kaliella longiseta* and '*Chlorella*' *mirabilis* (Figure 1).

Coccomyxa sp. grouped as a sister species of *Coccomyxa subellipsoidea* also with a high branch support (aLRT-like = 1) inside of the *Choricystis/Botryococcus* clade (Figure 1).

The two species from the *Chlorella* genus (*Chlorella sp.* and *Chlorella sorokiniana*) grouped closely to other two *Chlorella* species: *Chlorella vulgaris* and *Chlorella variabilis* (with a high branch support of 1 and 0.995) which belong to the order Chlorellales. However, both *Chlorella sp.* and *C. sorokiniana* branched before the sister species *C. vulgaris*/*C. variabilis* being the last one that branched first (Figure 1).

The last three species (*Auxenochlorella protothecoides*, *Protetheca wickerhamii* and *Heliscosporium sp.*) formed a sister monophyletic group of the clade consisting of *Marvania germinata*/*Pseudochloris wilhelmii*, the species from the *Chlorella* genus and *Diclostera acuatus*/*Parachlorella kessleri*, with high branch supports (aLRT-like = 0.918 and 1) (Figure 1). *Heliscosporidium sp.* showed a long branch length suggesting that this species have been evolving under a high evolutionary rate. It is in agreement with the parasitic life style of this species, once it is known that parasitic species usually have a genome reduction and high evolutionary rates [24]. However, even with this high variability of amino acid sequences, this species still could be confidently positioned in the Trebouxiophyceae tree, since no long branch attraction could be observed in the chloroplast phylogenomic analysis.

Mitochondrial phylogenomic analysis

In addition to the chloroplast phylogenomic analysis we also performed a mitochondrial phylogenomic analysis including 9 Trebouxiophyceae species. In this analysis we also observed a high branch support for most nodes with the exception of the node joining Streptophyta and Prasinophyceae/Ulvophyceae/Trebouxiophyceae clades (aLRT-like = 0.745); the node joining Prasinophyceae, Ulvophyceae and Trebouxiophyceae (aLRT-like = 0.257); and the node supporting *Marchantia polymorpha*/*Phycomitrella patens*/*Chara vulgaris* and *Arabidopsis thaliana* (aLRT-like = 0.166).

Even though, a much smaller number of mitochondrial genomes are available for Trebouxiophyceae species compared with plastid genomes, we still could reconstruct the phylogenetic positioning of some branches. Overall, our mitochondrial phylogenomic tree (Figure 2) matches well with tree reconstructed plastid tree (Figure 1). All three species from the *Chlorella* genus (*Chlorella sp.*, *Chlorella variabilis* and *Chlorella sorokiniana*) formed a

monophyletic group having as a sister clade *Auxenochlorella protothecoides*, *Protetheca wickerhamii* and *Heliscosporium sp.* Which is the reconstruction of the Chlorella clade observed in the plastid tree (Figure 1). The two most basal lineages comprise *P. crista* and a clade leading to *Trebouxiophyceae sp.* and *Coccomyxa sp.* All those nodes inside of the Trebouxiophyceae family presented aLTR-like supports higher than 0.900 (Figure 2).

We also established the phylogenetic placing of *Micromonas sp.* This species formed a monophyletic group with *Nephroselmis olivaceae* another species from the Prasinophyceae class [25].

In addition, we analyzed the phylogenetic positioning of *Neochloris aquatica* regarding the analyzed mitochondrial genomes. This species is considered as belonging to the Chlorophyta phylum and Chlorophyceae class [26]. However, in our analysis this species was placed as an early diverged lineage that branched before the split between Chlorophyta and Streptophyta phylum.

Incongruences between the plastid and mitochondrial analysis

The first and most evident incongruence between the mitochondrial and plastid phylogenomic analysis in the Trebouxiophyceans species. In the plastid tree species from the class Trebouxiophyceae form a paraphyletic clade with the “core Trebouxiophyceans” named by Limeux 2014 forming a monophyletic group and the Trebouxiophyceans from the Chlorellales order forming a basal monophyletic group in respect to the monophyletic group of species from Ulvophyceae/Chlorophyceae/core Trebouxiophyceae classes (Figure 1) [18]. Whereas, in the mitochondrial tree all core Trebouxiophyceans (*Coccomyxa sp.* and *Trebouxiophyceae sp.*) and the Trebouxiophyceae from the Chlorellales order forming a monophyletic group (Figure 2) having *Oltmannsiellopsis viridis*, and Ulvophyceae species as

a basal outgroup. Based on the plastid dataset this species was grouped with the *Pseudendoclonium akinetum* species and several species from the Chlorophyceae class forming a sister clade of the core Trebouxiophyceae class. But in the mitochondrial dataset this species was positioned as a basal lineage of all core Trebouxiophyceae and the Trebouxiophyceae from the Chlorellales order. Such positioning may be due to the low number of mitochondrial genomes from species of the Ulvophyceae class. However, the high branch supports observed in Figure 2 (aLTR-like support = 0.993) suggests that it is a real placing based on the mitochondrial data.

Two species of the Prasinophyceae class also were included in the mitochondrial phylogenomic analysis and were positioned as a sister group of Ulvophyceae/Trebouxiophyceae class which is incongruent with what was observed for the plastid dataset (a basal paraphyletic group). However, such positioning presented a weak node support of only 0.257, hindering any other conclusion about this clade.

Incongruences between plastid and mitochondrial phylogenetic analysis is a related phenomenon among species from the Streptophyta phylum [27, 28], however, among Chlorophyta species this is the first phylogenomic work to show such incongruences. It has been hypothesized that processes that can generate such patterns is interspecific hybridization is a well-established source of phylogenetic incongruence [29], and incongruence is expected when plastid and mitochondrial genomes exhibit contrasting paternal and maternal modes of inheritance. Furthermore, directional differences in the diploid-only organellar topology indicate that homoplasy is not exclusively caused by reticulation/hybridization associated with allopolyploid events, but also by stochastic processes including incomplete lineage sorting via biparental inheritance of organelles and/or parallel identical mutations [28].

Genome structure and organization

The draft plastid genome of *P. crispa* encodes 32 putative protein coding genes, 3 oxidative related genes, 2 rRNA and 27 tRNA genes (Table 1). No genes from the oxidative phosphorylation - complex II have been found. However, the draft aspect of the plastid genomes hinder any conclusion about losses of genes or genes complexes. The draft mitochondrial genome presented 18 protein coding genes, 3 rRNA and 22 tRNA genes [30].

In order to understand the gene rearrangement that took place during the evolutionary history of plastid and mitochondrial genomes of *P. crispa* and related species we performed syntheny analysis and placed it in an evolutionary context (Figure 3 and 4).

Regarding to the chloroplast syntheny analysis we can observe a low syntheny between *Prasiolopsis sp.* and *P. crispa*. However, some conserved blocks still can be detected in the 5' (*psbE*, *mind* and *fstH* genes), in the middle (*tufA* gene) and 3' (*infA*, *rps8*, *rpl5*, *rpl14*, *rpl16*, *rps3*, *rps19*, *rpl2*, *rpl23* genes) portion of the *P. crispa* plastid genome. In addition, we also can observe three major inversions in blue involving the block of genes *chlN*, *chlB*, *psbK*, *ycf12* and *psaM*; *atpB* and *atpE*; and *rpoB*, *rpoC1*, *rpoC2*.

In a broader phylogenetic context, we could observe a very conserved synthenic pattern between '*Chlorella*' *miriabilis*, *Pabia signensis* and *Koliella longiseta*. These synthenic blocks match perfectly with the phylogenetic relationship of those species (Figure 3). In addition, we can observe that two major inversions (in the middle and 3' region of the genome) can be detected between '*Chlorella*' *miriabilis* and *Koliella longiseta*/*Pabia signensis* clade and only one inversion between *Koliella longiseta* and *Pabia signensis*.

cpDNAs in most land plants share a highly conserved gene content and similar gene orders [31]. Most cpDNAs include two identical regions in opposite orientations called the inverted repeat (IR), flanked by large single copy (LSC) and small single copy (SSC) regions. The IRs generally contain the bacterial-like rRNA gene clusters, and the genes involved in photosynthesis (photosystem I/II, cytochrome b₆f, and ATP synthase) are arranged similarly in chloroplast and cyanobacterial genomes [32-34].

Table 1 Genes identified in the chloroplast and mitochondrial genome of *Prasiola crispa*

Photosynthesis-related	
Photosystem I	<i>psaB</i> , <i>psaC</i> , <i>ycf3</i>
Photosystem II	<i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbH</i> , <i>psbL</i> , <i>psbN</i> , <i>psbT</i> , <i>psbB</i> , <i>psbA</i> , <i>psbM</i>
Cytochrome b ₆ f complex	<i>petA</i> , <i>petB</i> , <i>petD</i>
ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpH</i> , <i>atpI</i>
Calvin cycle	<i>rbcL</i>
<i>Ribosomal proteins</i>	
Large subunit	<i>rpl14</i> , <i>rpl16</i> , <i>rpl36</i>
Small subunit	<i>rps12</i> , <i>rps19</i>
<i>RNA polymerases</i>	<i>rpoB</i> , <i>rpoC1</i> , <i>rpoC2</i>
<i>Others</i>	<i>accD</i> , <i>clpP</i> , <i>ccsA</i>
<i>Hypothetical proteins</i>	
<i>rRNAs</i>	<i>rrn16S</i> , <i>rrn23S</i>
<i>tRNAs</i>	<i>trnC-GCA</i> , <i>trnfM-CAT</i> , <i>trnG-TCC</i> , <i>trnH-GTG</i> , <i>trnI-CAT</i> , <i>trnI-GAT</i> , <i>trnL-CAA</i> , <i>trnM-CAT</i> , <i>trnN-GTT</i> , <i>trnP-TGG</i> , <i>trnQ-TTG</i> , <i>trnR-ACG</i> , <i>trnS-GCT</i> , <i>trnT-TGT</i> , <i>trnV-CAC</i> , <i>trnW-CCA</i> , <i>trnV-TAC</i> , <i>trnA-TGC</i> , <i>trnD-GTC</i> , <i>trnF-GAA</i> , <i>trnK-TTT</i> , <i>trnL-TAA</i> , <i>trnG-GCC</i> , <i>trnS-TGA</i> , <i>trnL-TAG</i> , <i>trnY-GTA</i> , <i>trnR-TCT</i>
Oxidative metabolism-related	
NADH: ubiquinone oxidoreductase (complex I)	<i>nad1</i> , <i>nad2</i> , <i>nad3</i> , <i>nad4</i> , <i>nad4L</i> , <i>nad5</i> , <i>nad6</i> , <i>nad7</i>
Cytochrome <i>bc</i> ₁ complex (complex III)	<i>cob_a</i> , <i>cob_b</i>
Cytochrome <i>c</i> oxidase (complex IV)	<i>cox1_0_a</i> , <i>cox1_0_b</i> , <i>cox1_1_a</i> , <i>cox1_1_b</i>
ATP synthase	<i>atp1</i> , <i>atp6</i> , <i>atp8</i> , <i>atp9</i>
<i>rRNAs</i>	<i>rrnL</i> , <i>rrnS</i> , <i>rrnS</i>
<i>tRNAs</i>	<i>trnR-ACG</i> , <i>trnN-GTT</i> , <i>trnQ-TTG</i> , <i>trnF-GAA</i> , <i>trnD-GTC</i> , <i>trnE-TTC</i> , <i>trnG-TCC</i> , <i>trnH-GTG</i> , <i>trnL-TAA</i> , <i>trnK-TTT</i> , <i>trnM-CAT</i> , <i>trnF-AAA</i> , <i>trnP-TGG</i> , <i>trnS-GCT</i> , <i>trnT-TGT</i> , <i>trnV-TAC</i> , <i>trnA-TGC</i> , <i>trnC-GCA</i> , <i>trnL-TAG</i> , <i>trnM-CAT</i> , <i>trnS-TGA</i> , <i>trnW-CCA</i>

Despite these well-characterized patterns, it is unknown to what extent the conserved gene order reflects a slow intrinsic rate of neutral chromosomal rearrangements, rather than selection against alternative gene orders [35].

Whereas gene order is generally conserved among land plant cpDNAs, very little synteny is observed between this group and cpDNAs of the chlorophytic green algae *Chlamydomonas reinhardtii* [36, 37] and *Chlorella vulgaris* [38]. Jeong et al., (2014) describes in cp genomes of different *Chlorella* species many conserved gene clusters in comparison with one Pedinomonadales and two Chlorellales species. However, the gene order of “*trnC-rpoB-rpoC1-rpoC2-rbcL-rps14*” is well conserved among the four analysed *Chlorella* species but not in related species *Coccomyxa* sp. and *Trebouxiophyceae* sp., suggesting that the gene order is well conserved and may be specific to *Chlorella* species. Other phenomenon described in *Chlorella* species is two genes overlapping, that occurs frequently in the genomes of viruses, prokaryotes, mitochondria, and eukaryotes, including humans [9].

Our results highlight the very dynamic plastid genome of *P. crispera* and its related species (*Prasiolopsis* sp. and *Stichococcus bacillaris*) when compared with *Pabia signensis/Koliella longiseta/Chlorella mirabilis* clade. It suggests that the ancestral of those species presented a stable cpDNA genome in terms of gene order and in the ancestral lineage leading to *P. crispera/S. bacillaris* and *Prasiolopsis* sp. several genome arrangements took place and are still happening after the split of those species.

For the mitochondrial genomes we have found only minor regions of syntenic blocks comparing *P. crispera* and *Coccomyxa* sp. species (Figure 4). Those region comprises the genes *rrnL* and *coxI* and an inversion of the gene *rrnS*. Moreover, well conserved syntenic blocks can be seen between *Coccomyxa* sp. and *Trebouxiophyceae* sp. and among all species from the Chlorellales order (Figure 4). The only comparison that resulted in only one small

synthetic block was between *Helicosporidium sp.* and *Trebouxiophyceae sp.* what is in agreement with the phylogenetic positioning of such species (Figure 1 and 2) and with the mtDNA genome reduction in the first species.

Based on the low levels of gene order conservation in other green algal mitochondrial genomes, we might expect to see few blocks of conservation between mtDNA analyzed. This close resemblances probably is best explained by a recent split between these two species. In the Chlorophyceae, the gene-poor mtDNAs of *Chlorogonium* and *C. eugametos* also display an appreciable level of synteny, with 8 of their 12 genes (66%) being located in shared clusters, although this percentage is still lower than that observed between *Helicosporidium sp.* and *Prototheca* mtDNAs (75%) [39]. Our data is incongruent with Pombert et al., (2010) describing surprisingly high level of similarity between *Helicosporidium* and *Prototheca* mtDNA [39]. The first mitochondrial sequenced genome from the Prasiolales order, *P. crispera* mtDNA [30], as expected to green algae, has few conserved blocks.

Conclusion

In the present work we established the phylogenetic positioning of *P. crispera* as a sister species of *Pransiolopsis sp.* In addition, we also established the phylogenetic positioning of other Trebouxiophycean algae based on a phylogenomic cpDNA analysis and compared it with the mtDNA tree. Several incongruences emerged, but future studies will be necessary on the biology of these algae species in order to understand the source of such incongruences. Moreover, we observed that *P. crispera* accessory genomes are very plastic with several inversions and rearrangements what is in agreement with other findings of Trebouxiophyceans accessory genomes.

Methods

Chloroplast and mitochondrial genomes

Most of mitochondrial and chloroplast genomes sequences were obtained from Turmel et al. 2007 and Lemieux et al. 2014, respectively [18, 40].

In order to include as many as Trebouxiophyceae genomes available we performed three types of searches at the NCBI website. A text based search using the string “Chloroplast genome” AND “Trebouxiophyceae” in the nucleotide databank of NCBI. A “Browse by Organism” search in the Genome resources (<http://www.ncbi.nlm.nih.gov/genome/>) of the NCBI. And a BLASTn [41] search with default parameter using the *P. crista* contigs against the non-redundant database.

For the accession numbers and the description from which search approach each genome was obtained see Additional files 1 and 2.

Phylogenomic analyses

The phylogenomic analysis was performed with 79 chloroplast coding genes and 18 mitochondrial coding genes. Those genes were aligned individually using Mafft v. 7.220 [42] with default parameters. Following, each alignment were computationally edited using Gblocks v. 0.91 [43] with the “Allowed gap positions” set up with “With Half” option. All genes were concatenated with Phyutility 2.2.6 [44].

The phylogenetic reconstruction was performed by maximum likelihood using PhyML 3.0 [45] webserver (<http://www.atgc-montpellier.fr/phyml/>). The amino acid substitution model used were the CpREV+G and LG+I+G for the mitochondrial and plastid dataset as suggested

by AIC and BIC parameter of Protest 2.4 [46]. The branch support was evaluated by the aLRT SH-like method, a fast likelihood-based method described by Anisimova 2006 [47].

Comparative analysis of organelle genomes

Syntheny analysis of the *P. crista* accessories genomes and close related ones based on the reconstructed trees where performed with BLASTn with default parameters. Then the Artemis Comparison Tool (ACT) was used to plot the synthenic blocks with more than 100 bp in size and its corresponding annotation.

Abbreviations

cpDNA, chloroplast genome; mtDNA, mitochondrial genome; NCBI, National Center for Biotechnology Information; ACT, Artemis Comparison Tool.

Competing interests

The authors declare they have no competing interests.

Acknowledgments

This work was supported by the National Council for Scientific and Technological Development (CNPq-Brazil), the Coordination for the Improvement of Higher Education Personnel (CAPES-Brazil), the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS-Brazil) and National Institute of Science and Technology - Antarctic Environmental Research (INCT-APA). EL Carvalho received research fellow ships from Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS-Brazil).

References

1. Convey P. Biogeography. In Riffenburgh, B. [Ed.] Encyclopedia of the Antarctic, 1, AK. Routledge, New York; 2007. p. 154-60.
2. Moniz MJB, Rindi F, Novis PM, Broady PA, Guiry MD. Molecular phylogeny of Antarctic *Prasiola* (Prasiolales, Trebouxiophyceae) reveals extensive cryptic diversity. Journal of Phycology. 2012;48(4):940-55.
3. Feller G, Gerday C, Psychrophilic enzymes: hot topics in cold adaptation. Nature Reviews Microbiology. 2003;1:200-08.
4. Kováčik L, Pereira AB. Green alga *Prasiola crispa* and its lichenized form *Mastodia tessellata* in Antarctic environment: general aspects. Nova Hedwigia, Beiheft. 2002. p. 465-78.
5. Wiencke C, Clayton MN. Antarctic seaweeds. Synopses of the Antarctic benthos, A.R.G. Gantner Verlag KG, Ruggell, Lichtenstein. 2002. 9, p. 239.
6. Graham LE, Graham JM, Wilcox LW. Algae. Pearson Education, Inc., San Francisco, p. 616, 2009.
7. Lud D, Buma AGJ, Van De Poll W, Moerdijk TCW, Huiskes AHL. DNA damage and photosynthetic performance in the Antarctic terrestrial alga *Prasiola crispa* ssp. *Antarctica* (Chlorophyta) under manipulated UV-B radiation. J. Phycol. 2001;37:459-67.
8. Kosugi M, Katashima Y, Aikawa S, Tanabe Y, Kudoh S, Kashino Y, Koike H, Satoh K. Comparative study on the photosynthetic properties of *Prasiola* (Chlorophyceae) and *Nostoc* (Cyanophyceae) from Antarctic and non-antarctic sites. J. Phycol. 2010;46:466-76.

9. Jeong H, Lim JM, Park J, Sim YM, Choi HG, Lee J, Jeong WJ. Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B. BMC Genomics, 2014;15(286):2-14.
10. Gray MW, Burger G, Lang BF. Mitochondrial evolution. Science. 1999;283: 1476-81.
11. Wang L, Mao Y, Kong F, Li G, Ma F, Zhang B, Sun P, Bi G, Zhang F, Xue H, Cao M. Complete sequence and analysis of plastid genomes of two economically important red algae: *Pyropia haitanensis* and *Pyropia yezoensis*. PLoS ONE. 2013;8:1-10.
12. Lemieux C, Otis C, Turmel M. Six newly sequenced chloroplast genomes from prasinophyte Green algae provide insights into the relationships among prasinophyte lineages and the diversity of streamlined genome architecture in picoplanktonic species. BMC Genomics. 2014;15(857):1-20.
13. Turmel, M.; Otis, C.; Lemieux, C. The chloroplast genomes of Green algae *Pedinomonas minor*, *Parachlorella kessleir* and *Oocystis solitaria* reveal a Shared ancestry between the Pedinomomadales and Chlorellales. Mol Biol Evol. 2009; 26(10):2317-31.
14. Civan P, Foster PG, Embley MT, Seneca A, Cox CJ. Analyses of charophyte chloroplast genomes help characterize the ancestral chloroplast genome of land plants. Gen Biol Evol. 2014;6(4):897-11.
15. Zhong B, Xi Z, Goremykin VV, Fong R, McLaenachan PA, Novis PM, Davis CC, Penny D. Streptophyte algae and the origin of land plants revisited using heterogeneous models with three new algal chloroplast genomes. Mol Biol Evol. 2014, 31(1):177-83.

16. Brouard JS, Otis C, Lemieux C, Turmel M. The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the chlorophyceae. *Gen Biol Evol.* 2010; 2:240-56.
17. Turmel M, Brouard JS, Gangnon C, Otis C, Lemieux C. Deep division in the Chlorophyceae (Chlorophyta) revealed by chloroplast phylogenomic analyses. *J Phycol.* 2008;44:739-50.
18. Lemieux C, Otis C, Turmel M. Chloroplast phylogenomic analysis resolves deep-level relationships within the Green algal class Trebouxiophyceae. *BMC Evol Biol.* 2014,14(211):1-15.
19. Yang X, Cameron SL, Lees DC, Xue D, Han H. A mitochondrial genome phylogeny of owl moths (Lepidoptera: Noctuoidea), and examination of the utility of mitochondrial genomes for lepidopteran phylogenetics. *Mol Phylog Evol.* 2015;85:230-37.
20. Hamaji T, Smith DR, Noguchi H, Toyoda A, Suzuki M, Kawai-Toyooka H, Fujiyama A, Nishii I, Marriage T, Olson BJ, Nozaki H. Mitochondrial and plastid genomes of the colonial green alga *Gonium pectorale* give insights into the origins of organelle DNA architecture within the volvocales. *PLoS One.* 2013;8(2):571-77.
21. Campbell MA, Presting G, Bennett MS, Sherwood AR. Highly conserved organellar genomes in the Gracilariales as inferred using new data from the Hawaiian invasive alga *Gracilaria salicornia* (Rhodophyta). *Phycol.* 2014;53:109-16.
22. DePriest MS, Bhattacharya D, López-Bautista JM. The mitochondrial genome of *Grateloupia taiwanensis* (Halymeniaceae, Rhodophyta) and comparative mitochondrial genomics of red algae. *Biol Bulletin.* 2014;227(2):191-00.

23. Kim SY, Yang EC, Boo SM, Yoon HS. Complete mitochondrial genome of the marine red alga *Grateloupia angusta* (Halymeniales). *Mitochondrial DNA*. 2014;25: 269-70.
24. deKoning AP, Keeling PJ. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. Is highly reduced and structured. *BMC Biol*. 2006;4(12):1-10.
25. Turmel M, Lemieux C, Burger G, Lang BF, Otis, C. Plante I. Gray MW. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within green algae. *Plant Cell*, 1999;11:1717-29.
26. Fucíková K, Lewis PO, González-Halphen D, Lewis L. Gene arrangement convergence, diverse intron content, and genetic code modifications in mitochondrial genomes Sphaeropleales (Chlorophyta). *Gen Biol Evolu*. 2014;6: 2170-80.
27. Tsutsui K, Suwa A, Sawada K, Kato T, Ohsawa TA, Watano Y. Incongruence among mitochondrial, chloroplast and nuclear gene trees in *Pinus* subgenus *Strobus* (Pinaceae). *J Plant Research*. 2009;122(5):509-21.
28. Govindarajulu R, Parks M, Tennessen JA, Liston A, Ashman TL. Comparison of nuclear, plastid, and mitochondrial phylogenies and the origin of wild octoploid strawberry species. *Amer J Botany*. 2015;102(4):544-45.
29. Rieseberg LH, Whitton J, Linder CR. Molecular marker incongruence in plant hybrid zone and phylogenetic trees. *Acta Botanica Neerlandica*. 1996;24:243-62.
30. Carvalho EL, Wallau, GL, Rangel DL, Machado LC, Silva AF, Pereira AB, Victoria FC, Boldo JT, Dal Belo CA, Pinto PM. Draft plastid and mitochondrial genomes from Antarctic algae *Prasiola crispa*. *Genome Announcements*. 2015;*No prelo*.

31. Palmer JD. Evolution of chloroplast and mitochondrial DNA in plants and algae. In Molecular Evol Genet. Edited by: Mac Intyre RJ. New York, Plenum Press. 1985.p. 131-40.
32. Kallas T, Spiller S, Malkin R. Primary structure of co transcribed genes encoding the Rieske Fe-S and cytochrome f proteins of the cyanobacterium Nostoc PCC 7906. PNAS. 1988;85:5794-98.
33. Pancic PG, Strotmann H, Kowallik KV. Chloroplast ATPase genes in the diatom *Odontella sinensis* reflect cyanobacterial characters in structure and arrangement. J Mol Biol. 1992;224:529-36.
34. Stoebe B, Kowallik KV. Gene-cluster analysis in chloroplast genomics. T Genet. 1999;15:344-47.
35. Cui L, Leebens-Mack J, Wang LS, Tang J, Rymarquis L, Stern DB, dePamphilis CW. Adaptive evolutions of chloroplast genome structure inferred using parametric bootstrap approach. BMC Evol Biol. 2006;9:6-13.
36. Boudreau E, Turmel M. Extensive gene rearrangements in the chloroplast DNA of *Chlamydomonas* species featuring multiple dispersed repeats. Mol Biol Evol. 1996;13:233-43.
37. Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, Harris EH, Stern DB. Chlamydomonas chloroplast chromosome: islands of genes in a sea of repeats. Plant Cell. 2002;14:2659-79.
38. Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsudzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K, Sugiura M. Complete

nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. PNAS. 1997;94:5967-72.

39. Pombert JF, Keeling PJ. The mitochondrial genome of the entomoparasitic green alga helicosporidium. PLoS ONE. 2010;5:1.

40. Turmel, M.; Otis, C.; Lemieux, C. An unexpectedly large and loosely packed mitochondrial genome in the charophycean Green alga *Chlorokybus atmophyticus*. BMC Genomics. 2007;8(137):1-12.

41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. J Mol Biol. 1990;215:403-10.

42. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013;30(4):772-80.

43. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540-52.

44. Smith SA, Dunn CW. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics. 2008;24(5):715-16.

45 Guindon S, Gascuel O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst Biol. 2003;52(5):696-04.

46. Abascal F, Zardoya R, Posada D. Prottest: selection of best-fit models of protein evolution. Bioinformatics. 2005;21(9):2104-05.

47. Anisimova M, Gascuel O. Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. *Syst Biol.* 2006;55(4):539-522.
48. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 2000;17(8):754-55.

Figures

Figure 1 Phylogenomic analysis of 70 chlorophytes chloroplast genomes using 14469 amino acid positions from 79 cpDNA-encoded proteins. The tree presented was reconstructed using a ML approach under the LG+I+G amino acid substitution model. Branch support values were estimated (over each branch) using the approximate likelihood ratio test (aLRT). Bars at the right side of the figure represents the taxonomic levels quoted in the text. ¹ Species obtained in the Organelle genome database (NCBI).

Figure 2 Phylogenomic analysis of 27 mitochondrial genomes encompassing all 8 chlorophyte mtDNA genomes available. The tree presented was reconstructed using a ML approach under the LG+G amino acid substitution model using 2277 amino acid positions from 18 mtDNA-encoded proteins. Branch support values were estimated (over each branch) using the approximate likelihood ratio test (aLRT). Bars at the right side of the figure represents the taxonomic levels quoted in the text. ¹ Species obtained in the Organelle genome database (NCBI); ² Species obtained by blast using the *P. crista* contigs as query.

Figure 3 Chloroplast synteny blocks analysis performed using blastn with default parameters and Artemis ACT software. Only regions with ≥ 100 bp were showed. Red and blue traces connecting cpDNA genomes represents synteny blocks in one or other strand.

Figure 4 Mitochondrial synteny blocks analysis performed using blastn with default parameters and Artemis ACT software. Only regions with ≥ 100 bp were showed. Red and blue traces connecting mtDNA genomes represents synteny blocks in one or other strand.

Figure 2

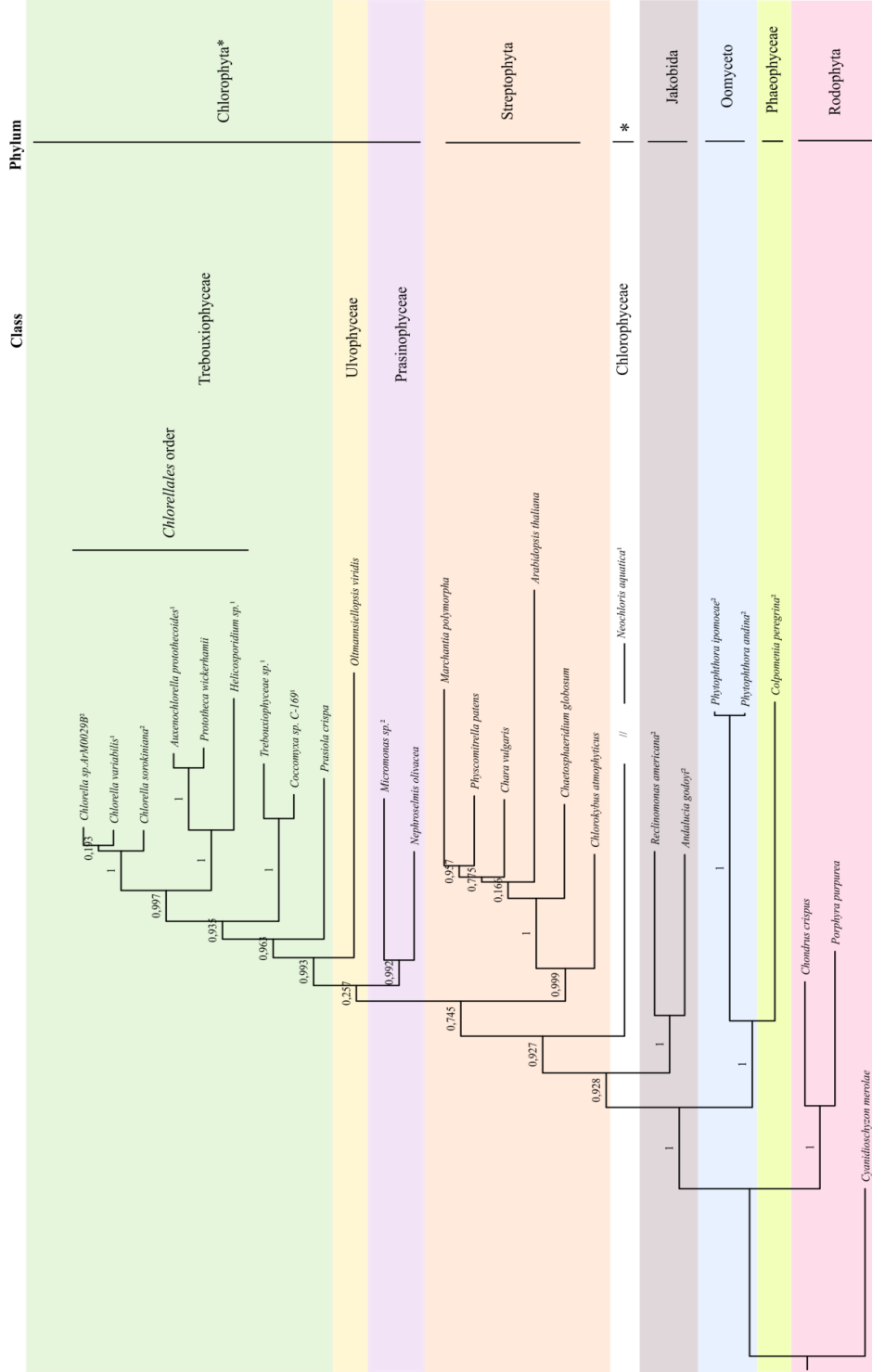


Figure 3

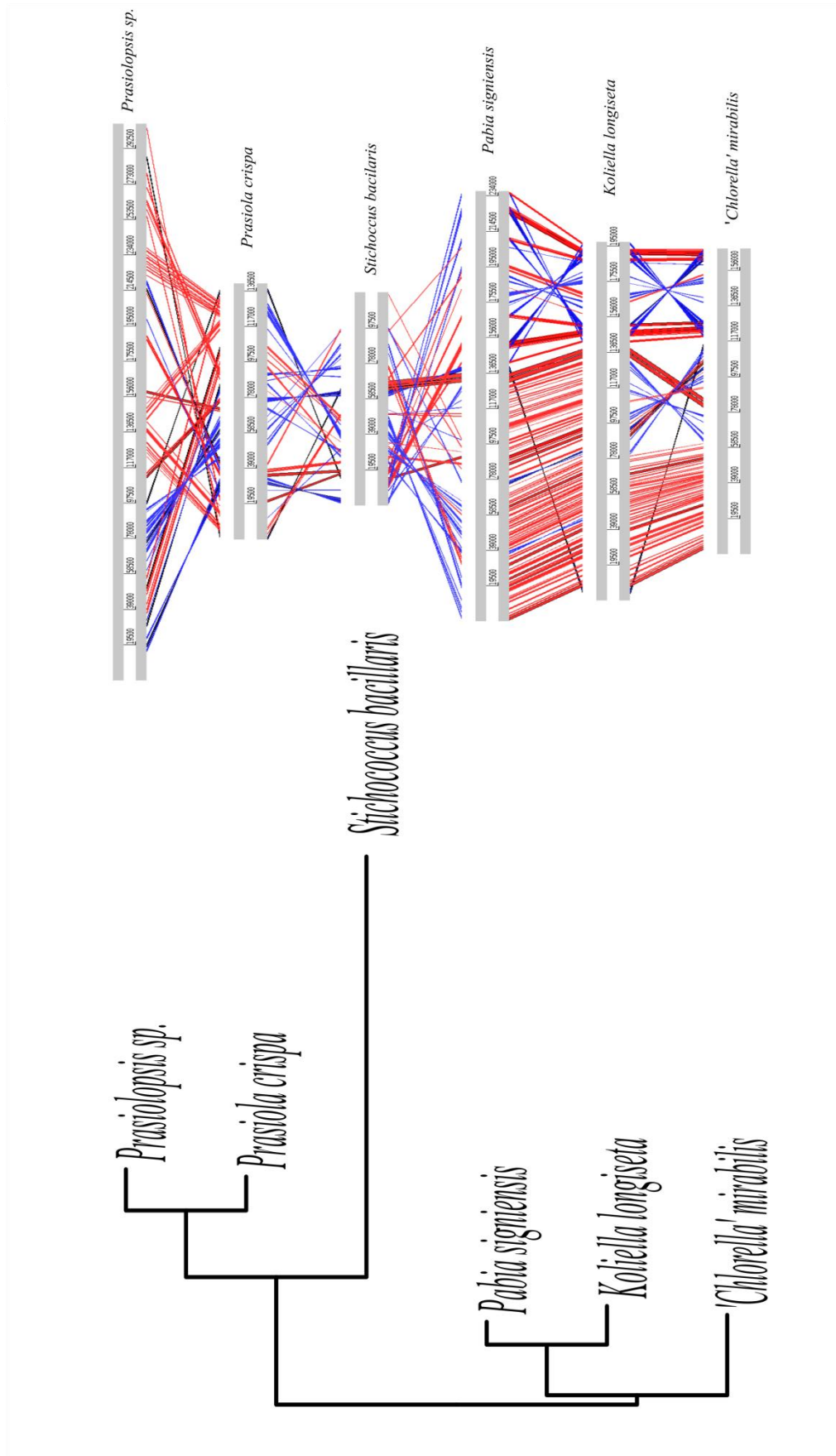
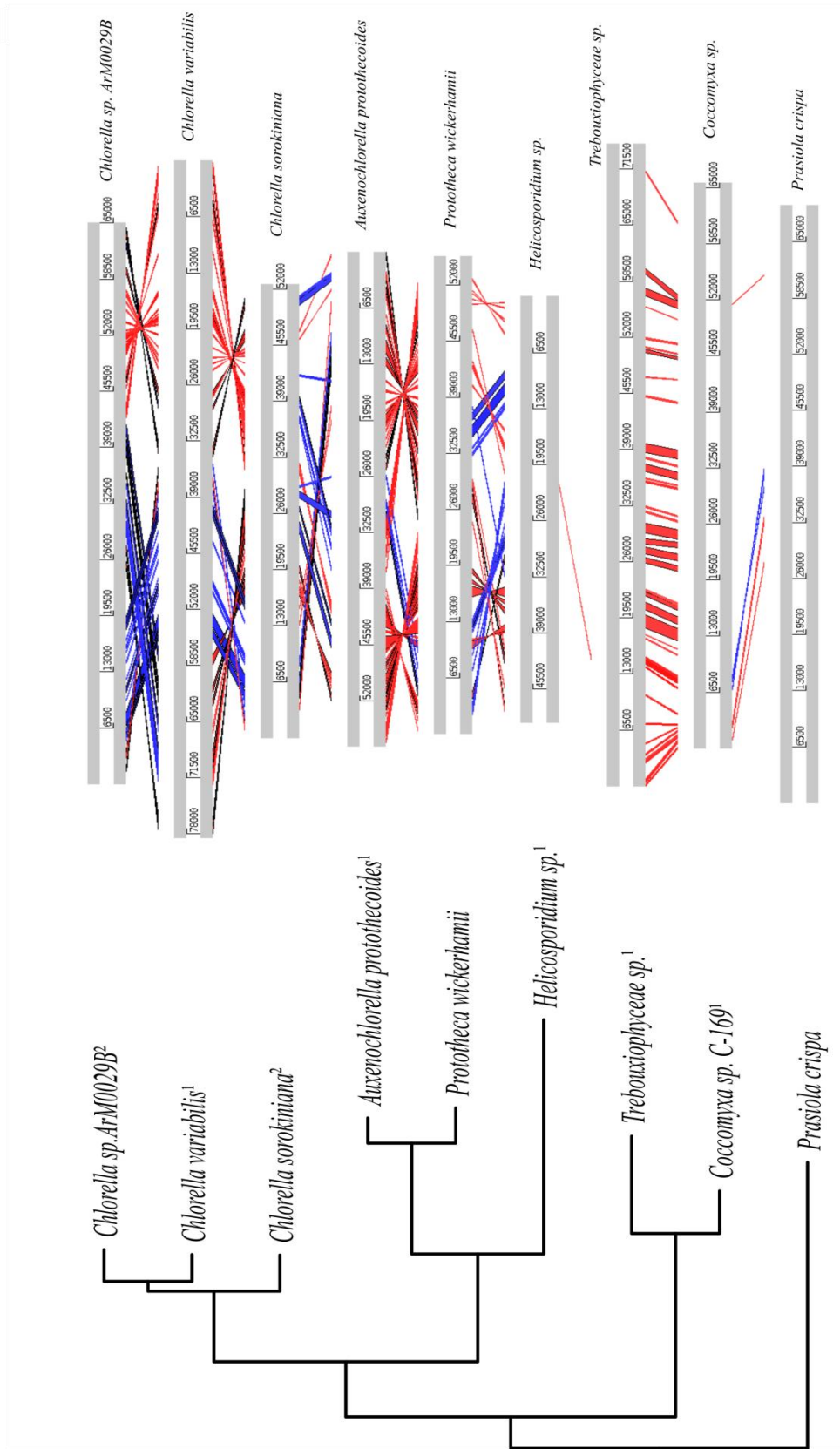


Figure 4



Additional files

Additional files 1. Numbers and the description of cpDNA genomes found using different search approach.

Species Chloroplast Genome	Access number NCBI
<i>Auxenochlorella protothecoides</i> ¹	NC_023775
<i>Botryococcus braunii</i> SAG 807-1 ²	KM462884
<i>Chlamydomonas moewusii</i> ²	EF587443 - EF587503
<i>Chlamydomonas reinhardtii</i> ²	NC_005353
' <i>Chlorella</i> ' <i>mirabilis</i> SAG 38.88 ²	KM462865
<i>Chlorella sorokiniana</i> ¹	NC_023835
<i>Chlorella</i> sp. ArM0029B ¹	KF554427.1
<i>Chlorella variabilis</i> NC64A ²	NC_015359
<i>Chlorella vulgaris</i> C-27 ²	NC_001865
<i>Chlorokybus atmophyticus</i> ²	NC_008822
<i>Choricystis minor</i> SAG 17.98 ²	KM462878
<i>Coccomyxa</i> sp. ¹	NC_015084
<i>Coccomyxa subellipsoidea</i> NIES 216 ²	NC_015084
<i>Dicloster acuatus</i> SAG 41.98 ²	KM462885
<i>Dictyochloropsis reticulata</i> SAG 2150 ²	KM462860
<i>Dunaliella salina</i> ²	NC_016732
<i>Elliptochloris bilobata</i> CAUP H7103 ²	KM462887
<i>Floydiella terrestris</i> ²	NC_014346
<i>Fusochloris perforata</i> SAG 28.85 ²	KM462882
<i>Helicosporidium</i> sp. ¹	NC_008100
<i>Geminella minor</i> SAG 22.88 ²	KM462883
<i>Geminella terricola</i> SAG 20.91 ²	KM462881
<i>Gloeotilopsis sterilis</i> UTEX 1704 ²	KM462877
<i>Koliella longiseta</i> UTEX 339 ²	KM462868
' <i>Koliella</i> ' <i>corcontica</i> SAG 24.84 ²	KM462874
<i>Leptosira terrestris</i> UTEX 333 ²	NC_009681
<i>Lobosphaera incisa</i> SAG 2007 ²	KM462871
<i>Marsupiomonas</i> sp. NIES 1824 ²	KM462870
<i>Marvania geminata</i> SAG 12.88 ²	KM462888
<i>Mesostigma viride</i> ²	NC_002186
<i>Micromonas</i> sp. RCC 299 ²	NC_012575
<i>Microthamnion kuetzingianum</i> UTEX 318 ²	KM462876
<i>Monomastix</i> sp. OKE-1 ²	NC_012101
<i>Myrmecia israelensis</i> UTEX 1181 ²	KM462861
<i>Neocystis brevis</i> CAUP D802 ²	KM462873
<i>Nephroselmis astigmatica</i> ²	KJ746600
<i>Nephroselmis olivacea</i> ²	NC_000927
<i>Oedogonium cardiacum</i> ²	NC_011031
<i>Oltmannsiellopsis viridis</i> ²	NC_008099
<i>Oocystis solitaria</i> SAG 83.80 ²	FJ968739
<i>Ostreococcus tauri</i> ²	NC_008289
<i>Pabia signiensis</i> SAG 7.90 ²	KM462866
<i>Parachlorella kessleri</i> SAG 211-11 ²	NC_012978

<i>Paradoxia multiseta</i> SAG 18.84 ²	NC_025540
<i>Parietochloris pseudoalveolaris</i> UTEX 975 ²	KM462869
<i>Pedinomonas minor</i> UTEX LB 1350 ²	NC_016733
<i>Picocystis salinarum</i> ²	KJ746599
<i>Planctonema lauterbornii</i> SAG 68.94 ²	KM462880
<i>Pleuraestrosarcina brevispinosa</i> UTEX 1176 ²	KM462875
<i>Prasinococcus</i> sp. CCMP 1194 ²	KJ746597
<i>Prasinoderma coloniale</i> CCMP 1220 ²	KJ746598
<i>Prasinophyceae</i> sp. CCMP1205 ²	KJ746601
<i>Prasinophyceae</i> sp. MBIC 106222 ²	KJ746602
<i>Prasiola crispa</i>	KR017748, KR017749, KR011750
<i>Prasiolopsis</i> sp. SAG 84.81 ²	KM462862
<i>Prasiolopsis</i> sp. SAG 84.81 ²	KM462862
<i>Prototheca wickerhamii</i> ¹	KJ001761.1
<i>Pseudendoclonium akinetum</i> ²	NC_008114
<i>Pseudochloris wilhelmii</i> SAG 1.80 ²	KM462886
<i>Pycnococcus provasolii</i> ²	NC_012097
<i>Pyramimonas parkeae</i> ²	NC_012099
<i>Scenedesmus obliquus</i> (<i>Acutodesmus obliquus</i>) ²	NC_008101
<i>Schizomeris leibleinii</i> ²	NC_015645
<i>Stichococcus bacillaris</i> UTEX 176 ²	KM462864
<i>Stigeoclonium helveticum</i> ²	NC_008372
<i>Trebouxia aggregata</i> SAG 219-1D ²	EU123962, EU124002
<i>Trebouxiophyceae</i> sp. MX-AZ01 ²	NC_018569
<i>Volvox carteri</i> f. <i>nagariensis</i> ²	GU084820
<i>Watanabea reniformis</i> SAG 211-9b ²	KM462863
<i>Xylochloris irregularis</i> CAUP H7801 ²	KM462872

¹Species obtained in the Organelle genome database (NCBI)

²Species obtained Lemieux et al. 2014.

Additional files 2. Numbers and the description of mtDNA genomes found using different search approach genome.

Species Mitochondrial Genome	Acess number NCBI
<i>Auxenochlorella protothecoides</i> ¹	NC_026009
<i>Chlorella variabilis</i> ¹	NC_025413
<i>Coccomyxa</i> sp. C-169 ¹	NC_015316
<i>Helicosporidium</i> sp ¹	NC_017841
<i>Neochloris aquatica</i> ¹	NC_024761
<i>Trebouxiophyceae</i> sp ¹	NC_018568
<i>Andalucia godoyi</i> ²	NC_021124.1
<i>Chlorella sorokiniana</i> ²	KM241869.1
<i>Chlorella</i> sp. ArM0029B ²	KF554428.1
<i>Colpomenia peregrina</i> ²	NC_025302.1
<i>Micromonas</i> sp.RCC 299 ²	FJ859351.1
<i>Phytophthora andina</i> ²	NC_015619.1
<i>Phytophthora ipomoeae</i> ²	NC_015622.1
<i>Reclinomonas americana</i> ²	NC_001823.1
<i>Arabidopsis thaliana</i> ³	NC_001284
<i>Chaetosphaeridium globosum</i> ³	NC_004118
<i>Chara vulgaris</i> ³	NC_005255
<i>Chlorokybus atmophyticus</i> ³	EF463011
<i>Chondrus crispus</i> ³	NC_001677
<i>Cyanidioschyzon merolae</i> ³	NC_000887
<i>Marchantia polymorpha</i> ³	NC_001660
<i>Nephroselmis olivacea</i> ³	NC_008239
<i>Oltmannsiellopsis viridis</i> ³	NC_008256
<i>Physcomitrella patens</i> ³	NC_007945
<i>Porphyra purpurea</i> ³	NC_002007
<i>Prototheca wickerhamii</i> ³	NC_001613
<i>Prasiola crispa</i>	KR017746, KR017747

¹Species obtained in the Organelle genome database (NCBI).

²Species obtained by blast using the *P. crispa* contigs as query.

³Species obtained Turmel et al. 2007.

5. CONSIDERAÇÕES FINAIS

Considerando os dados obtidos no presente estudo pode-se concluir que os genomas de organelas fornecem uma riqueza de dados filogeneticamente informativos, tornando-os uma abordagem valiosa para embasar estudos filogenômicos. Contudo, a determinação de uma maior variedade de sequências de genomas mitocondriais de algas verdes poderá permitir uma reconstrução mais rigorosa das relações taxonômicas destes organismos. Os genomas plastidiais da maioria das plantas terrestres compartilham genes conservados, ordem e conteúdo gênico semelhantes, fato este que não foi observado em nosso estudo, visto que a baixa sintonia observada entre os genomas plastidiais do clado *Prasiola* não condiz com o encontrado em plantas terrestres. Os dados gerados a partir do sequenciamento dos genomas acessórios da alga Antártica *P. crisper* fornecem um aporte para estudos futuros mais aprofundados.

6. PERSPECTIVAS FUTURAS

- Gerar mais sequências dos genomas das organelas na tentativa de fechar os *gaps* entre os *contigs* de cpDNA e mtDNA.
- Após o fechamento do genoma, reanotaremos os genomas para confirmar a atual anotação e buscar possíveis novos genes.
- Realizar uma análise e categorização funcional dos genes pertencentes ao cpDNA e mtDNA de *P. crisper*.
- Sequenciar o genoma nuclear de *P. crisper*, em busca de genes relacionados com o metabolismo oxidativo e fotossistemas que possam ter sido transferidos durante a endossimbiose.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- Brouard, J.S.; Otis, C.; Lemieux, C.; Turmel, M. **The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the chlorophyceae.** *Genome Biology and Evolution*, 2, p. 240-256, 2010.
- Convey, P. **Biogeography.** In: *Encyclopedia of the Antarctic*, 1, AK. Routledge, New York, p. 154-160, 2007.
- deCambiaire, J.C.; Otis, C.; Turmel, M.; Lemieux, C. **The chloroplast genome sequence of the green alga *Leptosira terrestris*: multiple losses of the inverted repeat and extensive genome.** *BMC Genomics*, 8:213, p. 1-13, 2007.
- deKoning, A.P.; Keeling, P.J. **The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured.** *BMC Biology*, 4:12, p.1-10, 2006.
- Denovan-Wright, E.M.; Nedelcu, A.M.; Lee, R.W. **Complete sequence of the mitochondrial DNA of *Chlamydomonas eugametos*.** *Plant Molecular Biology*, 36, 285-295, 1998.
- Douce, R. **Mitochondria in Higher Plants. Structure, Function and Biogenesis.** New York, Academic Press, 1985.
- Fan, J.; Lee, R.W. **Mitochondrial genome of the colorless green alga *Polytomella parva*: two linear DNA molecules with homologous inverted repeat Termini.** *Molecular Biology and Evolution*, 19, 999-1007, 2002.
- Friedl, T.; O'Kelly, C.J. **Phylogenetic relationships of green algae assigned to the genus *Planophila* (Chlorophyta) evidence from 18S rDNA sequence data and ultrastructure.** *European Journal of Phycology*, 37, p. 373–384, 2002.
- Gissi, C.; Pesole, G.; Mastrototaro, F.; Iannelli, F.; Guida, V.; Griggio, F. **Hypervariability of ascidian mitochondrial gene order: exposing the myth of Deuterostome organelle genome stability.** *Molecular Biology and Evolution*, 27, p. 211-215, 2010.
- Graham, L.E. **Green algae to land plants: an evolutionary.** *Journal of Plant Research*, 109, p. 241-251, 1996.
- Graham, L.E.; Graham, J.M.; Wilcox, L.W. **Algae.** Pearson Education, Inc., San Francisco, p. 616, 2009.
- Gray, M.W.; Burger, G.; Lang, B.F. **Mitochondrial evolution.** *Science*, 283, p. 1476-1481, 1999.
- Hajibabaei, M.; Singer, G.A.C, Herbert, P.D.N.; Hickey, D.A. **DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics.** *Trends IN Genetics*, 23, p.167-172, 2007.

- Hammani, K.; Giege, P. **RNA metabolism in plant mitochondria.** Trends in Plant Science, 19(6), p. 380-389, 2014.
- Harrison, N.; Kidner, C.A. **Next-generation sequencing and systematics: what can a billion base of DNA sequence data do for you?** Taxon, 60, p. 1552-1566, 2011.
- Javaux, E.J. **Extreme life on Earth - past, present and possibly beyond.** Research Microbiology, 157, p. 37-48, 2006.
- Jeong, H.; Lim, J.M.; Park, J.; Sim, Y.M.; Choi, H.G.; Lee, J.; Jeong, W.J. **Plastid and mitochondrion genomic sequences from Arctic *Chlorella sp.* ArM0029B.** BMC Genomics, 15:286, p. 2-14, 2014.
- John, D.M. Order Prasiolales. **The freshwater algal flora of the British Isles.** Cambridge University Press, Cambridge, p. 433-468, 2002.
- Kim, K.M.; Park, J.H.; Bhattacharya, D.; Yoon, H.S. **Applications of next-generation sequencing to unraveling the evolutionary history of algae.** International Journal of Systematic and Evolutionary Microbiology, 64, p. 333-345, 2014.
- Ko, H.L., Wang, Y.T., Chiu, T.S.; Lee, M.A.; Leu, M.Y.; Chang, K.Z.; Chen, W.Y.; Shao, K.T. **Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding.** PLOS ONE, 8, p. 1-7, 2013.
- Kolodner, R.; Tewari, K.K. **Inverted repeats in chloroplast DNA from higher plants.** PNAS, 76, p.41-45, 1979.
- Kosugi, M.; Katashima, Y.; Aikawa, S.; Tanabe, Y.; Kudoh, S.; Kashino, Y.; Koike, H.; Satoh, K. **Comparative study on the photosynthetic properties of *Prasiola* (Chlorophyceae) and *Nostoc* (Cyanophyceae) from Antarctic and non-antarctic sites.** Journal of Phycology, 46, p. 466-476, 2010.
- Kováčik, L.; Pereira, A.B. **Green alga *Prasiola crispa* and its lichenized form *Mastodia tessellata* in Antarctic environment: general aspects.** Nova Hedwigia, Beiheft, 123, p. 465-478, 2001.
- Leliaert, F.; Smith, D.R.; Moreau, H.; Herron, M.D.; Verbruggen, H.; Delwiche, C.F.; deClerck, O. **Phylogeny and molecular evolution of the green algae.** Critical Reviews in Plant Sciences, 31, p. 1-46, 2012.
- Lemieux, C.; Otis, C.; Turmel, M. **Six newly sequenced chloroplast genomes from prasinophyte Green algae provide insights into the relationships among prasinophyte lineages and the diversity of streamlined genome architecture in picoplanktonic species.** BMC Genomics, 15:857, p. 1-20, 2014a.
- Lemieux, C.; Otis, C.; Turmel, M. **Chloroplast phylogenomic analysis resolves deep-level relationships within the Green algal class Trebouxiophyceae.** BMC Evolutionary Biology, 14(211), p. 1-15, 2014b.

- Letsch, M.; Lewis, L.A. **Chloroplast gene arrangement variation within a closely related group of green algae (Trebouxiophyceae, Chlorophyta)**. *Molecular Phylogenetics and Evolution*, 64:3, p. 524-532, 2012.
- Lewis, L.A.; McCourt, R.M. **Green algae and the origin of land plants**. *American Journal of Botany*, 91:10, p. 1535-1556, 2004.
- Lud, D.; Buma, A.G.J.; Van De Poll, W.; Moerdijk, T.C.W.; Huiskes, A.H.L. **DNA damage and photosynthetic performance in the Antarctic terrestrial alga *Prasiola crispa* ssp. *antarctica* (Chlorophyta) under manipulated UV-B radiation**. *Journal of Phycology*, 37, p. 459-467, 2001.
- Mattox, K.R.; Stewart, K.D. **Classification of the green algae: a concept based on comparative cytology**. In: Irvine DEG, John DM, editors. *The systematic of the green algae*. London: Academic Press, p. 29-72, 1984.
- Merrill, R.A.; Strack, S. **Mitochondria: a kinase anchoring protein 1, a signaling platform for mitochondrial form and function**. *International Journal of Biochemistry & Cell Biology*, 48, p. 92-96, 2014.
- Michaelis, G.; Vahrenholz, C.; Pratje, E. **Mitochondrial DNA of *Chlamydomonas reinhardtii*: the gene for apocytochrome *b* and the complete functional map of the 15.8 kb DNA**. *Molecular and General Genetics*, 223, p. 211-216, 1990.
- Moniz, M.B.J., Rindi, F.; Novis, P.M.; Broady, P.A.; Guiry, M.D. **Molecular phylogeny of Antarctic *Prasiola* (Prasiolales, Trebouxiophyceae) reveals extensive cryptic diversity**. *Journal of Phycology*, 48:4, p. 940-955, 2012.
- Moniz, M.B.J.; Guiry, M.D.; Rindi, F. ***tufA* phylogeny and species boundaries in the green algal order Prasiolales (Trebouxiophyceae, Chlorophyta)**. *Phycologia*, 53 :4, p. 369-406, 2014.
- Mota, A.P.Z. Sequência completa do genoma cloroplasmático do feijão-caupi [*Vigna unguiculata* (L.) Walp] e diversidade genética de variedades tradicionais brasileiras e africanas. <http://repositorio.unb.br/handle/10482/12969?mode=full>, 2012.
- Novis, P.M.; Smissen, R.; Buckley, T.R.; Gopalakrishnan, K.; Visnovsky, G. **Inclusion of chloroplast gene that have undergone expansion misleads phylogenetic reconstruction in the Chlorophyta**. *American Journal of Botany*, 100:11, p. 2194-2209, 2013.

Organelle	Genome	Resources
	http://www.ncbi.nlm.nih.gov/genomes/GenomesHome.cgi?taxid=2759	
	Orsini, M.; Costelli, C.; Malavasi, V.; Cusano, R.; Concas, A.; Angius, A.; Cao, G. Complete genome sequence of mitochondrial DNA (mtDNA) of <i>Chlorella sorokiniana</i> . <i>Mitochondrial DNA</i> . No prelo. DOI: 10.3109/19401736.2014.953128, 2014.	
	Picault, N.; Hodges, M.; Palmieri, L.; Palmieri, F. The growing family of mitochondrial carriers in <i>Arabidopsis</i> . <i>Trends Plant Science</i> , 9, p. 138-146, 2004.	

- Pombert, J.F.; Beauchamp, P.; Otis, C.; Lemieux, C.; Turmel, M. **The complete mitochondrial DNA sequence of the green alga *Oltmannsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the Ulvophyceae.** *Current Genetics*, 50, p. 137-147, 2006.
- Pombert, J.F.; Otis, C.; Lemieux, C.; Turmel, M. **The complete mitochondrial DNA sequence of the Green alga *Pseudendoclonium akinetum* (Ulvophyceae) Highlights distinctive evolutionary trends in the Chlorophyta and suggests a sister-group relationships between the Ulvophyceae and Chlorophyceae.** *Molecular Biology and Evolution*, 21, p. 922-935, 2004.
- Pribil, M.; Labs, M.; Leister, D. **Structure and dynamics of thylakoids in land plants.** *Journal of Experimental Botany*, 12 p. 1-18, 2014.
- Reyes-Prieto, A.; Webwe, A.P.M.; Bhattacharya, D. **The origin and establishment of the plastid in algae and plants.** *Annual Review of Genetics*, 41, p.147-168, 2007.
- Rindi, F.; McIvor, L.; Guiry, M.D. **The Prasiolales (Chlorophyta) of the Atlantic Europe: an assessment based on morphological, molecular, and ecological data, including the characterization of *Rosenvingiella radicans* (Kützing) comb. nov.** *Journal of Phycology*, 40:5, p. 977-997, 2004.
- Rindi, F.; McIvor, L.; Sherwood, A.R.; Friedl, T.; Guiry, M.D.; Sheath, R.G. **Molecular phylogeny of the green algal order Prasiolales (Trebouxiophyceae, Chlorophyta).** *Journal of Phycology*, 43, p. 811-822, 2007.
- Rodríguez, R.R., Jiménez, J.C.; Delgado, C.M. **Microhabitat and morphometric variation in two new species of *Prasiola* (Prasiolales, Chlorophyta) from streams in central Mexico.** *Aquatic Ecology*, 41, p. 161 – 168, 2007.
- Ruhfel, B.R.; Gitzendanner, M.A.; Soltis, P.S.; Soltis, D.E.; Burleigh, G. **From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes.** *BMC Evolutionary Biology*, 14:23, p. 1-27, 2014.
- Schuster, W.; Brennicke, A. **The plant mitochondrial genome: physical structure, information content, RNA editing, and gene migration to the nucleus.** *Annual Review of Plant Physiology and Plant Molecular Biology*, 45, p. 61-78, 1994.
- Servín-Garcideueñas, L.E.; Martínez-Romero, E. **Complete mitochondrial and plastid genomes of the Green Microalga *Trebouxiophyceae* sp. strain MX-AZ01 isolated from a Highly Acidic geothermal lake.** *Eukaryotic Cell*, 11, p. 1417-1418, 2012.
- Shamblin, B.M.; Bjorndal, K.A.; Bolten, A.B.; Hillis-Starr, Z.M.; Lundgren, I.; Naro-Maciel, E.; Nairn, C.J. **Mitogenomic sequences better resolve stock structure of southern Greater Caribbean green turtle rookeries.,** *Molecular Ecology*, 21, p. 2330-2340, 2012.
- Turmel, M.; Brouard, J.S.; Gangnon, C.; Otis, C.; Lemieux, C. **Deep division in the Chlorophyceae (Chlorophyta) revealed by chloroplast phylogenomic analyses.** *Journal of Phycology*, 44, p. 739-750, 2008.

- Turmel, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Otis, C.; Plante, I.; Gray, M.W. **The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within green algae.** *The Plant Cell*, 11, p. 1717-1729, 1999.
- Turmel, M.; Otis, C.; Lemieux, C. **An unexpectedly large and loosely packed mitochondrial genome in the charophycean Green alga *Chlorokybus atmophyticus*.** *BMC Genomics*, 8:137, p.1-12, 2007.
- Turmel, M.; Otis, C.; Lemieux, C. **The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closet Green algal relatives of land plants.** *Molecular Biology and Evolution*, 23:6, p. 1324-1338, 2006.
- Turmel, M.; Otis, C.; Lemieux, C. **The chloroplast genomes of Green algae *Pedinomonas minor*, *Parachlorella kessleir* and *Oocystis solitaria* reveal a Shared ancestry between the Pedinomonadales and Chlorellales.** *Molecular Biology and Evolution*, 26:10, p. 2317-2331, 2009.
- Vahrenholz, C.; Riemen, G.; Pratje, E.; Dujon, B.; Michaelis, G. **Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication.** *Current Genetics*, 24, 241-247, 1993.
- Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Suzuki, Y.; Hamada, A.; Ohta, T.; Inamura, A.; Yoshinaga, K.; Sugiura, M. **Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division.** *Proceedings of the National Academy of Sciences*, 94, p. 5967-5972, 1997.
- Wang, L.; Mao, Y.; Kong, F.; Li, G.; Ma, F.; Zhang, B.; Sun, P.; Bi, G.; Zhang, F.; Xue, H.; Cao, M. **Complete sequence and analysis of plastid genomes of two economically important red algae: *Pyropia haitanensis* and *Pyropia yezoensis*.** *PLOS ONE*, 8, p. 1-10, 2013.
- Wicke, S.; Schenneweiss, G.; Pamphilis, C.; Müller, K.; Quandt, D. **The evolution of the plastids chromosome in land plants: gene content, gene order, gene function.** *Plant Molecular Biology*, 76:3, p. 237-297, 2007.
- Wiencke, C.; Clayton, M.N. **Antarctic seaweeds.** *Synopses of the Antarctic benthos*, 9, A.R.G. Gantner Verlag KG, Ruggell, Lichtenstein, p. 239, 2002.
- Wolff, G.; Plante, I.; Lang, B.F.; Kück, U.; Burger, G. **Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca Wickerhamii*.** *Journal of Molecular Biology*, 237, p. 75-86, 1994.
- Yan, D.; Wang, Y.; Murakami, T.; Shen, Y.; Gong, J.; Gao, C.; Jiang, H.; Smith, D.R.; Pombert, J.F.; Wu, Q.; Dai, J. **The plastid genomes of microalgae *Chlorella protothecoides* and *Prototheca wickerhamii* shed light on the evolutionary transition from autotroph to heterotrophy.** Unpublished. <http://www.ncbi.nlm.nih.gov/nuccore/KJ001761.1>, 2015

Zemolin, A.P.; Cruz, L.C.; Paula, M.T.; Pereira, B.K.; Albuquerque, M.P.; Victoria, F.C.; Pereira, A.B.; Posser, T.; Franco, J.L. **Toxicity induced by *Prasiola crispa* to fruit fly *Drosophila melanogaster* and cockroach *Nauphoeta cinerea*: evidende for bioinsecticide action.** Journal of Toxicology and Environmental Health, 77:1-3, p. 115-124, 2014.