UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

PATRÍCIA VERDUGO PASCOAL

Caracterização genômica e biotecnologica das microalgas *Chlamydomonas biconvexa* e *Dunaliella viridis*

Brasília
2022

PATRÍCIA VERDUGO PASCOAL

Caracterização genômica e biotecnológica das microalgas *Chlamydomonas biconvexa* e *Dunaliella viridis*

Tese apresentada ao Programa de Pesquisa e Pós-graduação em Biotecnologia da Universidade Federal da Bahia, para obtenção do título de doutora em Biotecnologia.

Orientador: Dr. Bruno dos Santos A. F. Brasil
Coorientador: Dr. Eduardo Fernandes Formighieri

Brasília
2022

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA), com os dados fornecidos pelo(a) autor(a).

Pascoal, Patrícia Verdugo
Caracterização genômica e biotecnologica das microalgas Chlamydomonas biconvexa e Dunaliella viridis / Patrícia Verdugo Pascoal. -- Salvador, 2024.
97 f.

Orientador: Bruno dos Santos Alves Figueiredo Brasil.
Coorientador: Eduardo Fernandes Formighieri.
Tese (Doutorado - Programa de pós-graduação em Biotecnologia - UFBA) -- Universidade Federal da Bahia, Instituto de Ciências da Saúde, 2024.

1. Microalgas. 2. Genômica. 3. Extremófilas. 4. Carotenoides. I. Brasil, Bruno dos Santos Alves Figueiredo. II. Formighieri, Eduardo Fernandes. III. Título.

PATRÍCIA VERDUGO PASCOAL

"Caracterização genômica e biotecnológica das microalgas *Chlamydomonas biconvexa* e *Dunaliella viridis*"

Tese apresentada como requisito para obtenção do grau de Doutora em Biotecnologia pelo Instituto de Ciências da Saúde da Universidade Federal da Bahia.

Aprovada em 17 de maio de 2022.

BANCA EXAMINADORA:

Bruno dos Santos Alves Figueiredo Brasil – Orientador _____
Doutor em Ciências Biológicas pela Universidade Federal de Minas Gerais, UFMG, Brasil.
Empresa Brasileira de Pesquisa Agropecuária, Embrapa Agroenergia.

Luís Gustavo Carvalho Pacheco _____
Doutorado em Bioquímica e Imunologia pela Universidade de Federal de Minas Gerais.
UFMG, Brasil.
Universidade Federal da Bahia.

Betânia Ferraz Quirino _____
Doutora em Cellular And Molecular Biology pela University of Wisconsin - Madison, WISC, Estados Unidos.
Empresa Brasileira de Pesquisas Agropecuárias (Embrapa).

João Ricardo Moreira de Almeida _____
Doutor em Engenharia: Microbiologia Aplicada pela Lund University, LUND, Suécia.
Empresa Brasileira de Pesquisas Agropecuárias (Embrapa).

Priscila Grynberg _____
Doutora em Bioinformática pela Universidade Federal de Minas Gerais, UFMG, Brasil.
Empresa Brasileira de Pesquisa Agropecuária, Embrapa Agroenergia.

AGRADECIMENTOS

Ao meu querido e amado pai, Antonio Pascoal Filho, que esteve sempre ao meu lado, mesmo longe, dando todo tipo de suporte com palavras de estímulo e discussões acerca do meu projeto, definitivamente é meu maior incentivador! Obrigada por todo aparato na infância para me despertar a curiosidade e chegar ao ponto de querer me tornar parte do "fazer ciência", pois é isto que me move. Amo o senhor!

Aos meus irmãos: Nathália, Nicholas, Rhaquel, Raphael e Henderson, por me apoiarem e fazerem alguns momentos mais leves, vocês mal sabem que fizeram e fazem diferença nesta minha trajetória.

À minha grande amiga e irmã, Giselly Batista Alves, por todas as trocas de conhecimento, pelos estudos em conjunto, por todas as pitangas choradas e todo apoio durante o doutoramento. Você fez muita diferença!

Ao meu orientador, professor Bruno Brasil, por todo suporte, sugestões, apoio, compreensão nos momentos difíceis e por toda troca de conhecimento. Muito grata por confiar no meu trabalho e por ser este ser humano maravilhoso!

Ao meu coorientador, Edu =), por todo apoio técnico no LBB, por todas as conversas profissionais e pessoais, pela compreensão, pelo coração enorme e pelas muitas risadas no laboratório. Sentirei muita falta!

Às amigas da Embrapa Agroenergia: Tallytinha, Carol e Dani, vocês são demais! Depois da pandemia tomaremos um café juntas.

Agradeço também à doutora Letícia Jungmann Cançado, pelo apoio como supervisora, por todas as trocas e conversas, cresci muito, profissionalmente, com a nossa convivência.

À FAPESB pela bolsa concedida, ao Programa de Pesquisa e Pós-Graduação em Biotecnologia pela oportunidade e colaboração e à Embrapa Agroenergia, pela infraestrutura, recursos e acolhimento neste período.

“Sua saúde mental está nas suas mãos, na emoção de lidar. O trabalho não deve ser uma coisa servil, é algo que exprime a alma da pessoa”.

(Nise da Silveira)

PASCOAL, Patrícia Verdugo. Caracterização genômica e biotecnológica das microalgas *Chlamydomonas biconvexa* e *Dunaliella viridis.* 2022. Orientador: Bruno dos Santos Alves Figueiredo Brasil. 97 f. il. Tese (Doutorado em Biotecnologia) – Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador, 2022.

RESUMO

As microalgas têm potencial para produção de uma diversidade de ativos de interesse industrial, seja na indústria alimentícia, por seu alto teor de proteínas e carboidratos contidos na biomassa; seja no tratamento de efluentes agroindustriais, como o POME (do inglês *Palm oil mil effluent*). Também pode ser voltada para a indústria de ativos com alto valor agregado, como os pigmentos carotenoides, aplicados na área de cosméticos, nutrição animal alimentação humana. Dentre as espécies com potencial biotecnológico, destacam-se duas cepas isoladas de ambientes brasileiros, a dulcícola *Chlamydomonas biconvexa* Embrapa|LBA40 e a espécie halotolerante *Dunaliella viridis* Embrapa|LBAS001. No presente estudo, a cepa *C. biconvexa* Embrapa|LBA40, isolada e cultivada em efluente da indústria de óleo de palma (POME) foi capaz de alcançar a produtividade de biomassa de 190,60 mgDW • L^{-1} • d^{-1} em fotobiorretores *airlift* de placa plana de 15L. A espécie foi capaz de reduzir a amônia e o nitrito do resíduo de POME em 99%, assim como também reduziu o fosfato em 98% após 5 dias de cultivo. Além disso, o genoma mitocondrial foi obtido através de sequenciamento genético, revelando um mtDNA de 15,98 Kb, com 14 genes, dos quais 9 são genes codificadores de proteínas. Análises filogenéticas por meio do gene COX1 confirmaram a identificação taxonômica como *C. biconvexa*, abrindo oportunidade para futuros estudos genéticos de modificação e melhoramento da espécie. Já a cepa microalga *Dunaliella viridis* EMBRAPA| LBA#S001, isolada de lagoas salinas no Brasil, foi estudada quanto às suas características morfológicas e genômicas. A cepa foi identificada por marcadores moleculares, como os genes rbcL e ITS2, e demonstrou alta tolerância a salinidades elevadas, com crescimento consistente em concentrações de NaCl de até 4,8 M. A produção de β-caroteno foi avaliada, resultando em 18,7 mg/L, comparável aos níveis encontrados em outras espécies do gênero. O DNA mitocondrial revelou um genoma de 46,2 Kbp, enquanto o genoma cloroplastidial alcançou 197,1 Kb, ambos com genes essenciais para a fotossíntese. A análise do genoma nuclear, com 176,9 Mb, identificou 29.594 genes codificadores de proteínas, indicando um potencial significativo para a biotecnologia. O estudo analisou os genes de resistência na cepa, identificando mecanismos como a alteração de alvos antibióticos e o efluxo de antibióticos. A presença de genes como vanY e adeF sugere uma diversidade adaptativa em ambientes contaminados. A análise dos fatores de transcrição destaca a regulação de genes essenciais para a fotossíntese e respostas ao estresse. A manipulação desses fatores pode aumentar a produção de compostos valiosos, como carotenoides. O desenvolvimento de promotores sintéticos é crucial para otimizar microalgas como plataformas biotecnológicas, em comparação com organismos mais tradicionais.

Palavras-chave: Microalgas. Genômica. Extremófilas. Carotenoides.

PASCOAL, Patricia Verdugo. Genomic and biotechnological characterization of the microalgae *Chlamydomonas biconvexa* and *Dunaliella viridis*. 2022. Advisor: Bruno dos Santos Alves Figueiredo Brasil. 97 f. il. Thesis (Doctorate in Biotechnology) – Institute of Health Sciences, Federal University of Bahia, Salvador, 2022.

ABSTRACT

Microalgae have the potential to produce a variety of active ingredients of industrial interest, whether in the food industry, due to the high protein and carbohydrate content of their biomass; or in the treatment of agro-industrial effluents, such as POME (Palm oil mil effluent). It can also be used in the industry to produce assets with high added value, such as carotenoid pigments, which are used in cosmetics, animal nutrition, and human food. Among the species with biotechnological potential, two strains isolated from Brazilian environments stand out: the freshwater *Chlamydomonas biconvexa* Embrapa|LBA40 and the halotolerant species *Dunaliella viridis* Embrapa|LBAS001. This study showed that the strain *C. biconvexa* Embrapa|LBA40, isolated and cultivated in palm oil industry effluent (POME), was able to achieve biomass productivity of 190.60 mg DW • L^{-1} • d^{-1} in 15L flat plate airlift photobioreactors. The species was able to reduce ammonia and nitrite from the POME residue by 99%, as well as reducing phosphate by 98% after 5 days of cultivation. In addition, the mitochondrial genome was obtained through genetic sequencing, revealing an mtDNA of 15.98 Kb, with 14 genes, 9 of which are protein-coding genes. Phylogenetic analysis using the COX1 gene confirmed the taxonomic identification as *C. biconvexa*, opening opportunities for future genetic studies to modify and improve the species. The microalgae strain *Dunaliella viridis* EMBRAPA| LBA#S001, isolated from saline lagoons in Brazil, was studied for its morphological and genomic characteristics. The strain was identified by molecular markers, such as the rbcL and ITS2 genes, and demonstrated high tolerance to high salinities, with consistent growth in NaCl concentrations of up to 4.8 M. The production of β-carotene was evaluated, resulting in 18.7 mg/L, comparable to the levels found in other species of the genus. The mitochondrial DNA revealed a genome of 46.2 Kbp, while the chloroplast genome reached 197.1 Kb, both with genes essential for photosynthesis. The analysis of the nuclear genome, with 176.9 Mb, identified 29,594 protein-coding genes, indicating a significant potential for biotechnology. The study analyzed the resistance genes in the strain, identifying mechanisms such as antibiotic target switching and antibiotic efflux. The presence of genes such as vanY and adeF suggests adaptive diversity in contaminated environments. Analysis of transcription factors highlights the regulation of genes essential for photosynthesis and stress responses. Manipulating these factors could increase the production of valuable compounds such as carotenoids. The development of synthetic promoters is crucial to optimizing microalgae as biotechnological platforms, compared to more traditional organisms.

Keywords: Microalgae. Genomics. Extremophiles. Carotenoids.

SUMÁRIO

1. INTRODUÇÃO

As microalgas são organismos eucariontes, unicelulares fotossintetizantes, comumente isoladas de ambientes de água doce ou salinos, bem como solos úmidos; possuem eficiência na biofixação de carbono, são capazes de alcançar alta produtividade de biomassa de composição química variada. Possuem vantagens quanto à possibilidade de serem cultivadas em terras não agricultáveis, em períodos sem entressafra nos países tropicais, quando comparadas às culturas terrestres de interesse econômico (Dismukes et al., 2008; Sayre, 2010).

Podem ser usadas em diversas aplicações como produção de suplementos alimentares, ração animal, cosméticos, enzimas, materiais, biocombustíveis, dentre outros.

Os gêneros mais estudados são *Spirulina* (cianobactéria), *Chlorella*, *Chlamydomonas* e *Dunaliella*, podem ser aplicados industrialmente como fonte de ácidos graxos, pigmentos carotenoides e podem contribuir com a produção de biocombustíveis como bioetanol e biodiesel (Andrade & Colozzi Filho, 2014).

A microalga *Chlamydomonas biconvexa* é um organismo eucariótico pertence às algas verdes, classe das clorófitas e filo das clorofíceas. Dentro deste gênero a espécie amplamente estudada, *C. reinhardtii*, possui proximidade filogenética com as plantas terrestres e, portanto, por sua capacidade fotossintética foi escolhida como organismo modelo para pesquisas. A espécie *C. biconvexa* possui potencial aplicação na produção de biomassa com elevados teores de proteínas e carboidratos de interesse agrícola e indústria de biocombustíveis, uma vez que pode ser cultivada em efluentes agroindustriais, contribuindo com sua biorremediação (Nascimento, 2016).

Tanto o cloroplasto como a mitocôndria são organelas que possuem material genético próprio e a origem destas organelas é baseada na teoria endossimbiótica. As mitocôndrias são parcialmente autônomas e desenvolvem um importante papel na respiração celular. Muitas evidências indicam sua origem endossimbiótica, resultado do englobamento de um organismo procarioto, provavelmente uma α-proteobacteria, que foi fagocitada por um organismo unicelular de vida livre. Durante a evolução simultânea do endossimbionte e seu hospedeiro ocorreu a redução do conteúdo genômico do endossimbionte, onde parte dos genes foram perdidos e parte deles transferidos para o núcleo, gerando assim arquiteturas e tamanhos de mitogenomas diferenciados em diversas linhagens filogenéticas. Portanto, com o acúmulo de dados moleculares há uma maior possibilidade de entendimento sobre a origem, evolução e organização das mitocôndrias.

Dentre os compostos de interesse biotecnológico de maior valor agregado produzido por microalgas, destacam-se os carotenoides, pigmentos naturais comercializados como suplementos nutricionais e corantes para a indústria de alimentos, cosmética e de higiene pessoal. Na célula, os pigmentos carotenoides estão localizados no citoplasma como glóbulos hidrofóbicos e exercem a função de proteção aos centros de reação celular, onde a energia luminosa é convertida em energia química (Skjånes et al., 2013).

O mercado global de carotenoides naturais foi estimado em US$ 919 milhões em 2015, US$ 1,5 bilhões em 2021 e estimativas apontam um mercado de US$ 2 bilhões até o final de 2022 (Sofiana et al., 2021). O seu preço de venda pode ultrapassar US$ 12.500,00 por quilograma de produto, a depender de seu grau de pureza (Mustapa et al., 2011). A depender do tipo de carotenoide comercializado os preços podem variar de € 6,000 por quilograma, para todas as formas *trans*, por exemplo, até o valor de € 500,000 por grama para a forma 9-*cis* β-caroteno, em razão de sua dificuldade de síntese. Já o preço da luteína, outro carotenoide de interesse econômico, pode alcançar o valor de € 1470 por quilograma; estes bioativos de alto valor agregado podem ser obtidos de espécies do gênero *Dunaliella* (Harvey & Ben-Amotz, 2020). Industrialmente, estes pigmentos também podem ser obtidos a partir do fungo *Blakeslea trispora*, da palma de óleo (*Elaeis guineensis*), mas principalmente por espécies *D. salina* e *D. bardawi*l (Mustapa et al., 2011).

A espécie *D. salina*, modelo amplamente estudado, pode se manter ativa tanto em água com salinidade média da água do mar, cerca de 3,5% de NaCl, até águas com a concentração de saturação deste sal, ou seja, 31% de NaCl (Borowitzka & Silva, 2007). Tal fato confere a esta microalga poder competitivo em meio salino controlado, pela limitação do crescimento de outros microrganismos, inclusive predadores. Quando se trata do acúmulo destes pigmentos carotenoides há a interferência de eventos de estresse, seja por alta concentração de sal no meio, baixas temperaturas, alta intensidade luminosa ou mesmo a depleção de determinados nutrientes como nitrato e fontes de fósforo e enxofre (Lv et al., 2016).

Portanto, o gênero *Dunaliella* é alvo de estudos voltados à pesquisa da halotolerância e produção de carotenoides, uma vez que há variabilidade destes metabólitos e características em níveis intra e interespécie (Oren, 2005). Os mecanismos para este fenômeno ainda são pouco conhecidas e certamente envolvem regulação em níveis transcricionais e pós-transcricionais, exigindo maiores estudos da via de biossíntese destes carotenoides, especialmente quando da privação de nutrientes essenciais (Kim et al., 2013).

O Brasil importa betacaroteno em valores superiores a 1 milhão de dólares anuais, pois ainda conta com baixa produção local (Mesquita, 2013). Com a crescente demanda por

betacaroteno produzido de forma natural para aplicação em alimentos funcionais, surge como oportunidade o desenvolvimento de sistemas produtivos de *Dunaliella* para as condições brasileiras. Isto implica na seleção de cepas produtivas, bem como no desenvolvimento de tecnologias custo-eficientes de produção e colheita de biomassa algal.

Parte do mecanismo genético de controle da síntese de carotenoides é conhecido, bem como estudos iniciais de manipulação genética visando o aumento da sua produção foram reportados (Couso et al., 2012). Foi demonstrado, por exemplo, que a expressão do gene exógeno Betacaroteno Oxigenase (*bkt*), isolado de *Chlorella zofingiensis*, na microalga *C. reinhardtii* estimula a síntese de carotenoides que esta microalga não produzia inicialmente (León et al., 2004).

Além disso, embora estudos focados no genoma mitocondrial e plastidial de espécies do gênero *Dunaliella* estejam disponíveis, ainda não há genoma completo publicado para espécies deste gênero, além do *draft* nuclear de uma cepa de *Dunaliella salina* dos Estados Unidos (Polle et al., 2017).

A falta de um genoma de referência completo pode limitar, por exemplo, a estimativa do número de proteínas presentes no genoma e investigações acerca dos processos metabólicos envolvidos na carotenogênese e de sua adaptação à ambientes hipersalinos (Ramos et al., 2011).

1.1 Objetivos

Desta forma, os objetivos deste estudo orbitaram em realizar a montagem, anotação e análise filogenômica do mitogenoma da cepa *C. biconvexa*, para que futuramente o genoma possa ser utilizado para desenvolver ferramentas genômicas, como marcadores moleculares, sistemas de expressão gênica e técnicas de manipulação genética.

Posteriormente, o estudo foi voltado para o sequenciamento genômico da cepa halotolerante de *D. viridis* EMBRAPA | LBA#, utilizando-se as tecnologias Illumina e Nanopore, de forma a obter subsídios sobre a ocorrência dos genes envolvidos no controle genético das rotas de produção de carotenoides e resistência a condições de hipersalinidade, os quais poderão desdobrar em trabalhos futuros de melhoramento.

O presente estudo está dividido em dois capítulos, o primeiro conta com a publicação do genoma mitocondrial completo da cepa dulcícola *Chlamydomonas biconvexa* EMBRAPA | LBA40 no periódico PLOS ONE. O segundo capítulo compreende o manuscrito do genoma da mitocôndria e cloroplasto completos e o *draft* do genoma nuclear da cepa salina *Dunaliella viridis* EMBRAPA | LBA#S001 a ser submetido ao periódico Algal Research.

CAPÍTULO 1. Biochemical and phylogenetic characterization of the wastewater tolerant *Chlamydomonas biconvexa* Embrapa|LBA40 strain cultivated in palm oil mill effluent*

*Artigo publicado como Pascoal PV, Ribeiro DM, Cereijo CR, Santana H, Nascimento RC, Steindorf AS, Calsing LCG, Formighieri EF e Brasil BSAF (2021) Biochemical and phylogenetic characterization of the wastewater tolerant *Chlamydomonas biconvexa* Embrapa|LBA40 strain cultivated in palm oil mill effluent. PLoS ONE 16(4):e0249089. https://doi.org/10.1371/journal.

Biochemical and phylogenetic characterization of the wastewater tolerant *Chlamydomonas biconvexa* Embrapa|LBA40 strain cultivated in palm oil mill effluent

Patrícia Verdugo Pascoal[1,4], Dágon Manoel Ribeiro[1,2,5¶], Carolina R. Cereijo[1], Hugo Santana[1], Rodrigo C. Nascimento[1,3], Andrei Stecca Steindorf[1], Lorena C. G. Calsing[1], Eduardo Fernandes Formighieri[1], Bruno S. A. F. Brasil [1, 4]

[1]Embrapa Agroenergia, Brasília, Distrito Federal, Brazil
[2] Universidade de Brasília, Brasília, Distrito Federal, Brazil
[3] Universidade Federal do Tocantins, Gurupi, Tocantins,Brazil
[4] Universidade Federal da Bahia, Salvador, Bahia, Brazil
[5] Universidade Zambeze, Sofala, Mozambique

Abstract

The increasing demand for water, food and energy poses challenges for the world´s sustainability. Tropical palm oil is currently the major source of vegetable oil worldwide with a production that exceeds 55 million tons per year, while generating over 200 million tons of palm oil mill effluent (POME). It could potentially be used as a substrate for production of microalgal biomass though. In this study, the microalgal strain *Chlamydomonas biconvexa* Embrapa|LBA40, originally isolated from a sugarcane vinasse stabilization pond, was selected among 17 strains tested for growth in POME retrieved from anaerobic ponds of a palm oil industrial plant located within the Amazon rainforest region. During cultivation in POME, *C. biconvexa* Embrapa|LBA40 biomass productivity reached 190.60 mgDW • L^{-1} • d^{-1} using 15L airlift flat plate photobioreactors. Carbohydrates comprised the major fraction of algal biomass (31.96%), while the lipidic fraction reached up to 11.3% of dry mass. Reductions of 99% in ammonium and nitrite, as well as 98% reduction in phosphate present in POME were detected after 5 days of algal cultivation, it can contribute in saving time and treatment area once the aerobic pond stage in industrial plants could be substituted by photobioreactors. In addition, the complete mitochondrial genome of *C. biconvexa* Embrapa|LBA40 strain was sequenced, revealing a compact mitogenome, with 15.98 kb in size, a total of 14 genes, of which 9 are protein coding genes. Phylogenetic analysis confirmed the strain taxonomic status within the *Chlamydomonas* genus, opening up opportunities for future genetic modification and molecular breeding programs in these species.

Keywords: microalgae; biorefinery; Amazon rainforest; airliftflat plate photobioreactor; mitochondrial genome.

Introduction

Water, food and energy security, in a context of climatic changes, are major worldwide challenges for the 21st century. Therefore, the advancement of biofuel production and the integration of industrial processes in biorefineries will be pivotal in order to provide energy, food and fibers, such as cellulose for paper production, in a sustainable manner [1].

Brazil has one of the most advanced and well-established biofuel policies among BRICS nations and the developed economies of the world [2,3]. Since the adoption of the National Fuel Alcohol Program (proálcool), in 1975, and the National Program of Production and Use of Biodiesel (PNPB), in 2004, the country has leveraged its production of biofuels and has become the second largest bioethanol and biodiesel producer in the world. The country's light fleet is largely composed of flex-fuel vehicles and a National Biofuel Policy (RenovaBio) was established in 2017 to increase the production and use of biofuels through the commercialization of decarbonization credits [3,4].

Among the different feedstocks available for biodiesel and edible oil production, palm oil is one of the most efficient crops regarding land use and productivity in tropical regions. Worldwide palm oil production exceeds 55 million tons annually, representing a market of over $ 34 billion a year [5]. Currently, Brazil is the nineth world’s producer of palm oil concentrating the large majority of its plantations within the Amazon rainforest.

The Amazon region spans over 6.7 thousand square kilometers of tropical rainforest, comprising one of the most species-rich biodiversity hotspots in the world. In order to prevent the primary and secondary forest clearing, Brazil`s Sustainable Palm Oil Production Program, since 2010, regulates palm oil production by restricting new oil palm plantations in Amazonia, including legally protected parks, indigenous reserves and intact forest areas. As a result, it has been shown that direct conversion of intact forest to oil palm declined 4% from 2006 - 2010, to less than 1% from 2010-2014 [6]. However, despite these efforts to protect the forest areas, crude palm oil production can still pose as an important threat to the environment. For example, the production of 1 ton of crude palm oil requires between 5 and 7.5 tons of water, leading to the generation of large volumes of wastewater in the form of palm oil mill effluent (POME) [7]. Given the

continuous increase in palm oil production within the Amazon region, POME generation at large volumes might pose a challenge for the sustainable growth of this industry.

POME is composed of 95-96% water, 0.6-0.7% residual oil and 4-5% total solids. It has high chemical and biochemical oxygen demands and contains significant concentrations of nitrogen and phosphorus that can cause severe pollution to the environment [7]. Palm oil industries typically use stabilization pond-based systems to treat POME wastewater. This type of treatment is particularly common in tropical conditions due to the high temperatures, solar radiation, presence of oxygen and phytoplankton that help stabilize the effluent [8]. The main advantages are the high efficiency for organic and solid matter removal and the low cost of implementation and maintenance. However, pond systems require large areas, long stabilization periods (usually over 3 months) and do not completely eliminate the organic and inorganic load of the wastewater [9,10].

POME can potentially be used as a cultivation medium for the production of microalgae, opening up opportunities for integrated production of high value-added algal biomass within the palm oil industry [7,11]. The use of microalgae has been increasingly attracting interest worldwide as an alternative biomass source, due to its wide biotechnological applications in food, feed, energy and material production, its flexibility in terms of cultivation conditions, as well as its sustainability and contribution to the reduction of greenhouse gas emissions [4,12]. However, low cost culture systems, improvements in nutrient use efficiency and selection of productive strains, still pose as challenges for the economic viability of large scale production[13]. In this context, the use of industrial wastewater and residues stands as a potential low cost and abundant source for the formulation of culture media for algae [14].

Previous studies have shown the viability of cultivating microalgae in different concentrations of POME at a laboratory scale [15,16] as well as the use of acid or antibiotic pretreatment to control contaminants has also been proposed [17]. However, information regarding microalgae cultivation in POME retrieved from anaerobic stabilization ponds without nutrient supplementation using closed photobioreactors and under non-axenic conditions is still scarce. Furthermore, the chemical composition of algal biomass produced in POME has not yet been characterized in detail.

Recently, the strain *Chlamydomonas biconvexa* Embrapa|LBA40 has been shown to produce high amounts of protein-rich biomass when cultivated in sugarcane vinasse, an abundant wastewater from ethanol plants, uncover a promising biotechnological potential [18,19]. *C. biconvexa* is a freshwater species that presents characteristic oval shaped cells with two anterior flagella. Often forms colonial aggregates of immobile nonflagellated individuals, referred as palmella stage [20]. assigned to the *Chlamydomonas* genus solely based on morphologic and physiologic analysis, since genomic and phylogenetic information about this species is virtually absent from the literature.

The main purpose of this study is to evaluate whether *Chlamydomonas biconvexa* Embrapa|LBA40 is capable of producing large amounts of value-added biomass during non-axenic cultivation in POME without nutrient supplementation, while bioremediating its nitrogen and phosphorous load. Here, we show that *C. biconvexa* Embrapa|LBA40 strain presents the highest microalgae productivity reported to date when grown in airlift flat plate photobioreactors using non-supplemented POME retrieved from anaerobic ponds as cultivation medium. The strain was also efficient in wastewater bioremediation, potentially reducing the area used for POME stabilization and contributing to sustainable palm oil production within the Amazon region. Biochemical characterization of proteins, carbohydrates and lipids fractions within algae biomass was performed in order to evaluate algae bioproducts production potential in a biorefinery approach. Furthermore, the complete mitochondrial genome sequence and phylogenetic analysis of *C. biconvexa* Embrapa|LBA40 is presented.

Materials and methods

Microalgal strains

Seventeen (17) microalgae strains from the Collection of Microorganisms and Microalgae for Agroenergy and Biorefineries of the Brazilian Agricultural Research Corporation - Embrapa (Brasília-DF) [19,21] were used in this study (S1 table). Microalgae strains were maintained in erlenmeyer flasks containing 150 mL working volume of bold basal medium (BBM) at 26 ± 1 ° C, with aeration of 5 $L \cdot h^{-1}$of atmospheric air, light intensity of 50 μEm^{-2} s^{-1} (3750 lux) at and 12/12h light/dark cycle.

Palm oil mill effluent

The palm oil mill effluent used in this study was collected at the exit of the POME anaerobic stabilization pond at a palm oil industrial plant (Dendê do Pará S/A) located within the Brazilian Amazonian region at the Pará state (Fig 1). It is referred hereafter as "anaerobic pond POME", or simply, "AP-POME". The treatment system used in this industrial plant is composed of two sequential ponds, one anaerobic and the second aerobic (Fig 1).

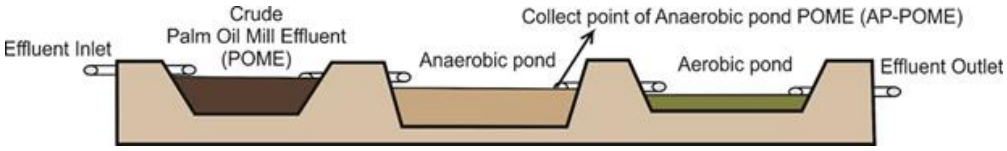


Fig 1. Representation of Palm Oil Mill Effluent (POME) wastewater treatment ponds (DENPASA - Dendê do Pará S/A).

Prior to experimentation, AP-POME was prepared by centrifugation at 4800 ×g for 10 minutes to remove suspended solids. This step resulted in a decrease in AP-POME turbidity from 396 NTU prior to centrifugation to 18 NTU after. No alteration in the color of AP-POME was detected. AP-POME was stored at 4ºC until use.

Microalgae screening for growth in anaerobic pond palm oil mill effluent

Screening was carried out in 500 ml erlenmeyer flasks containing 250 ml of AP-POME as culture medium. Batch culturing was independently conducted in triplicates for each of the 17 strains at a constant aeration of 5 $L \cdot h^{-1}$of atmospheric air, at 26 ± 1° C, light intensity of 100 $\mu Em^{-2}\ s^{-1}$ (7500 lux), and 12/12 h light/dark cycle. Microalgae growth was monitored through periodic measures of absorbance at 680 nm and microscopic inspection of culture samples during 10 days of cultivation. The absorbance of AP-POME prior to algal inoculation was used as blank and the basal growth rate (μ) was calculated according to the equation $\mu=[(OD_2 - OD_{AP\text{-}POME})/(OD_1 - OD_{AP\text{-}POME})]/t_2-t_1$.

***Chlamydomonas biconvexa* Embrapa|LBA40 strain cultivation**

Chlamydomonas biconvexa Embrapa|LBA40 strain was cultivated using 250 mL of AP-POME or BBM (control) under different conditions: i) axenic culture using BBM at 12/12 h light/dark cycles; ii) axenic culture using AP-POME at 12/12 h light/dark

cycles; iii) axenic culture using AP-POME in the dark; iv) non-axenic culture using AP-POME at 12/12 h light/dark cycles; v) non-axenic culturing using 50% AP-POME diluted in distillate water at 12/12 h light/dark cycles. For axenic cultures experiments, BBM or AP-POME were autoclaved at 121° C for 15 minutes and subsequently stored at 4 ° C until use. Batch cultures were conducted in triplicates under aeration of 5 $L \cdot h^{-1}$of atmospheric air, at 26 ºC ± 1 °C, at light intensity of 100 μEm^{-2} s^{-1}(7500 lux) (when applicable). Algae growth kinetics was evaluated through cell counting using a Neubauer chamber as described by Santana and collaborators [22].

Microalgae cultivation in airlift flat plate photobioreactors

Chlamydomonas biconvexa Embrapa|LBA40 was batch cultivated in 15 L acrylic airlift flat plate photobioreactors under non-axenic conditions. The working volume of 13 L of cultivation medium (BBM or AP-POME) was used, with constant aeration of $60L \cdot h^{-1}$ and CO_2 supplementation adjusted to 5% of the air flow. Experimentation was conducted at 12h/12h light/dark cycling regime, at a light intensity of 495 μEm^{-2} s^{-1} (35000 lux) and temperature of 25° C during dark periods and 35° C during the light periods. Biomass dry weight (DW) was gravimetrically determined [22] using 10 ml samples retrieved from cultures from each photobioreactor replicate at the initial time and at 5 days intervals during 15 days of culturing.

The chemical parameters of palm oil mill effluent (POME)

The chemical parameters of crude POME and AP-POME before and after 5, 10 and 15 days of microalgae cultivation using airlift flat plate photobioreactors were determined. Samples were collected from each photobioreactor replicate, centrifuged for 10 min at 4800 ×g and the supernatants were used for analysis. The physicochemical characteristic and methods used were: biochemical oxygen demand (BOD) - SM5210B; chemical oxygen demand (COD) - QAM.IT.FQ.16A; total organic carbon - SM 4500-O/D; nitrate (NO_3^-) - ABNT NBR 12620:1992; nitrite (NO_2^-) - SM 4500-NO2-B; ammoniacal nitrogen (NH_4^+) - SM 4500-NH3 F; phosphate (PO_4^{3-}) - SM 4500-P E and total potassium (K^+) - SM 3500-K B [22].

Composition of algal biomass

After cultivation, algal biomass was harvested by centrifugation for 10 min at 4800 × g and then freeze dried. Total protein content was measured by the quantification of nitrogen using micro-Kjeldahl with the microalgae-specific conversion factor of 4.78 [23]. Carbohydrate content, fractions, total solids and ashes were quantified following the protocol described by Van Wychen and Laurens [24].

The lipid content and fatty acid profiles were determined according to Van Wychen et al. [25]. Briefly, the lipid fraction was obtained by ether extraction (EE) with petroleum ether at 90 ° C for 90 minutes (Ankom XT15). Experimentation was carried out in triplicates using 100 mg of freeze-dried biomass. Samples were placed in hydrophobic filter bags (XT4 Filter Bags, ANKOM Technology), with porosity of 3 microns that only allows the extraction of nonpolar compost (lipid fraction) soluble in petroleum ether, the quantification is calculated by the gravimetric difference of the biomass after the extraction.

Fatty acid profiles were determined through gas chromatography after treatment of 10 mg of biomass for transesterification with 0.2 ml of chloroform: methanol (2:1, v/v) and 0.3 ml of 0.6 M HCl in methanol heated at 85 °C in dryi-block for 1 hour. After heating, 1 ml hexane was used to extract the Fatty acid methyl esters (FAMES). FAMEs were separated by gas chromatography using an Agilent 7890 A gas chromatograph (Agilent Technologies, California, USA), coupled with a flame ionization detector (FID) and a fused silica capillary column (100 m x 250 μm x 0.2 μm, Supelco SP). The operating parameters were set as follows: detector temperature, 260 ºC column temperature, 140 ºC for 5 minutes, programmed to increase 4 °C/min, up to 240 ºC, with a final running time of 48 minutes. The carrier gas was Helium at 1.2 mL•min-1, with injection of 1 μL sample. The retention times of fatty acids were compared to those of standard methylesters (Sigma-Aldrich, St. Louis, MO, USA). Retention times and percentages of the peak area were calculated automatically by the ChemStation Software. Fatty acids quantifications (FA) were performed using the methyl ester of nonadecanoate acid (Sigma-Aldrich, USA) as an internal standard.

Statistical analysis

Experiments were conducted using biological triplicates (n = 3). Algal biomass productivity ($mgDW \cdot L^{-1} \cdot day^{-1}$) was calculated using the equation described by Kishi and Toda [26]. POME and algal Biomass composition data were subjected to analysis of variance (ANOVA) at 5% probability followed by a Tukey test, using the software Action Stat version 3.5.

Genomic analysis

Sequencing

Sixty milligrams of *C biconvexa* Embrapa|LBA40 fresh biomass were used for genomic DNA extraction following the protocol previously described by Doyle [27]. The genomic DNA was further fragmented for the: construction of two libraries shotgun library (2 x 250 bp), sequenced on the Illumina MiSeq platform; and jump library 3 kb (2 x 125 bp), sequenced on the Illumina HiSeq 2500 platform according to the manufacturer's protocols. Subsequently, *in silico* analyses were performed at the Laboratory of Bioinformatics in Bioenergy (LBB) of Embrapa Agroenergy (Brasília, Brazil).

The FastQC tool [28] was used to evaluate the quality of the sequence data set. The adapters were trimmed using the Trimmomatic software [29] with the following adapted parameters: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 SLIDINGWINDOW:4:20 LEADING:15 TRAILING:10 MINLEN:30.

Mitochondrial genome assembly and annotation

De novo genome assembly was performed with whole genome shotgun assembler ALLPATHS-LG v.3 [29], using low paired-end read coverage (2X).The annotation process was accomplished using MITOS2 Web Server [30], Reference: 63 – Opisthokonta and Genetic code: 16 Chlorophycean. The visualization of mitogenome annotation was performed with OGDRAW [31], using the *Chlamydomonas reinhardtii* mitogenome as a model, available at GEO (NCBI's Gene Expression Omnibus) under accession number GSE101944 [66]. The *C. biconvexa* Embrapa|LBA40 mitogenome sequence data was deposited at NCBI GenBank under the accession number MG916975.1.

Phylogenetic Tree construction

Phylogenetic analysis was performed using the mitochondrial cox1 (cytochrome oxidase subunit 1) subunit gene as a barcode. The cox1 protein annotated sequence of *C. biconvexa* Embrapa|LBA40 mtDNA was concatenated to other twenty-nine green algae cox 1 protein sequences obtained from GenBank. Phylogenetic tree construction was conducted in MEGA X [32]. The sequences were aligned by ClustalW and Maximum Likelihood method and JTT matrix-based model was used to evolutionary inference. The bootstrap consensus tree inferred from 100 replicates. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (100 replicates) are shown next to the branches. This analysis involved 30 amino acid sequences. There were a total of 541 positions in the final dataset.

Results and Discussion

Screening of microalgae strains for growth in anaerobic pond palm oil mill effluent

The selection of robust strains adapted for wastewater growth and with high biomass productivity is a crucial step for the viability of the cultivation system [1,22]. Here, seventeen (17) strains of microalgae (S1 table) were screened for growth using AP-POME as substrate (Fig 2).

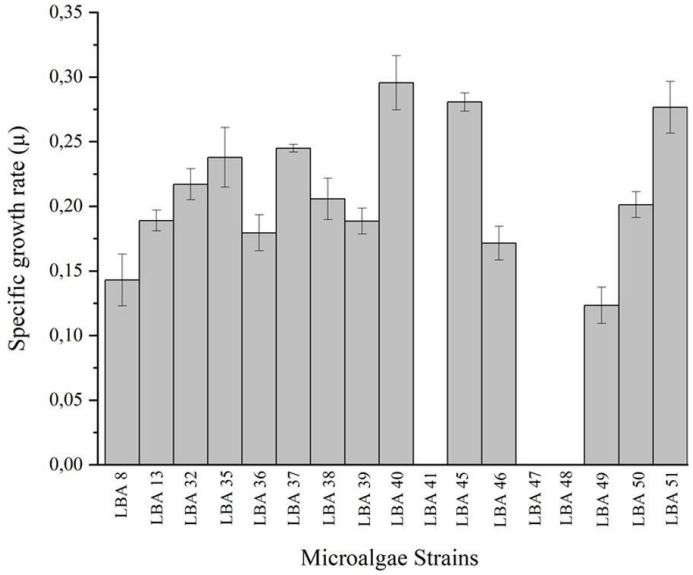


Fig 2. Screening of microalgae strains for growth in anaerobic pond palm oil mill effluent (AP-POME): Screening was carried out in 500 ml erlenmeyer flasks containing 250 ml of AP-POME as culture medium. Batch culturing was independently conducted in triplicates for each of the 17 strains at a constant aeration of 5 $L \cdot h^{-1}$of atmospheric air, at 26 ± 1° C, light intensity of 100 μEm^{-2} s^{-1} (7500 lux), and 12/12h light/dark cycle. Microalgal growth was monitored through periodic measures of absorbance at 680nm and microscopic inspection of culture samples during 10 days of cultivation. The absorbance of AP-POME prior to algal inoculation was used as blank. Initial and final absorbance were used to calculate basal growth rate (μ). *Results are presented as mean ± error bars of triplicate experiments (n = 3).

The strains used were originally collected from distinct neotropical biomes, including not only natural environments but also anthropizedones, such as rural wastewater and sugarcane vinasse stabilization ponds [19]. The strains *Chlamydomonas biconvexa* Embrapa|LBA40, *Chloromonas* sp. Embrapa|LBA45 and *Chlorococcum* sp. Embrapa|LBA51 presented the highest specific growth rates among the microalgae strains tested. On the other hand, the strains *Chlamydomonas* sp. Embrapa|LBA41, *Chlorococcum* sp. Embrapa|LBA47 and *Chlorococcum* sp. Embrapa|LBA48 did not present any growth (Fig 2). The growth rate differences observed among closely related strains (i.e.: *C. biconvexa* Embrapa|LBA40 and *Chlamydomonas* sp. Embrapa|LBA41; *Chlorococcum* sp. Embrapa|LBA51 and *Chlorococcum* sp. Embrapa|LBA48) indicates that tolerance to POME might be strain-specific.

C. biconvexa Embrapa|LBA40 strain has been recently shown to produce high amounts of protein-rich biomass when cultivated in sugarcane vinasse wastewater, either diluted 50% in water or chemically clarified [19,22]. The capability of achieving high

growth rates in two distinct industrial wastewaters (i.e.: sugarcane vinasse and POME) suggests that *C. biconvexa* Embrapa|LBA40 has an unusual versatility to tolerate adverse growth conditions and might have potential for large scale production in a biorefinery context. Therefore, this strain was selected for further biochemical and phylogenetic characterization.

Algal growth in anaerobic pond palm oil mill effluent

C. biconvexa Embrapa|LBA40 strain was grown in aerated erlenmeyer flasks containing AP-POME under different conditions in order to characterize the effect of light and microbial contamination upon algal growth (Fig 3).

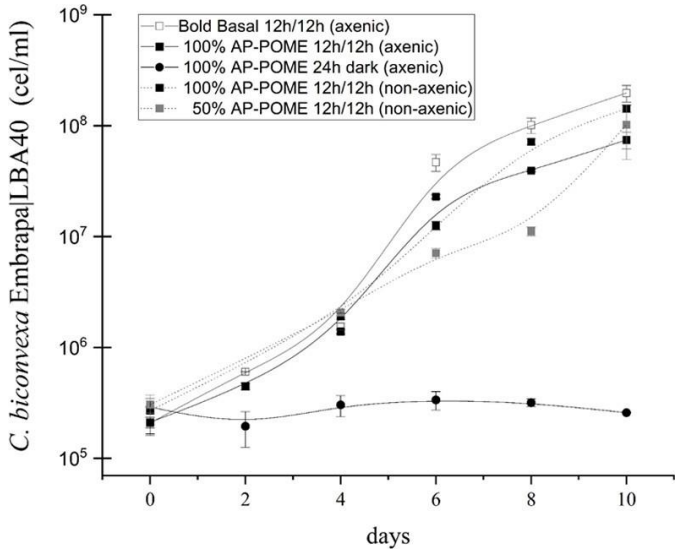


Fig 3. Growth dynamics of *Chlamydomonas biconvexa* Embrapa|LBA40 under different conditions. The strain was cultivated in erlenmeyer flasks using 250 mL of AP-POME or BBM (control) under different conditions: axenic culture using BBM and 12/12 h light/dark cycles (−□−), axenic culture using undiluted AP-POME in the light (−■−); axenic culture using undiluted AP-POME in the dark (−●−), non-axenic culture using undiluted AP-POME and 12/12 h light/dark cycles (··■··), non-axenic culturing using 50% diluted AP-POME in distillate water and 12/12 h light/dark cycles (··■··). Batch cultures were aerated with 5 $L \cdot h^{-1}$ of atmospheric air, at 26 ºC ± 1 °C, at light intensity of 100 μEm^{-2} s^{-1} (7500 lux) (when applicable). The growth kinetics of the algae was evaluated through cell counting using a Neubauer chamber. Results shown are the mean of biological triplicates of the experiment (n = 3).

No algal growth was detected in the absence of light. On the other hand, *C. biconvexa* Embrapa|LBA40 presented similar growth patterns either in 100% and 50% non-axenic cultures performed under light/dark cycles using both Bold´s Basal Medium (BBM) and AP-POME, indicating robust photo-dependent algae growth albeit the

presence of microbial competitors (Fig 3). Similarly, *C. biconvexa* Embrapa|LBA40 has been reported to perform light-dependent growth when cultivated in sugarcane vinasse [14].

Cultivation of *C. biconvexa* Embrapa|LBA40 in AP-POME was further scaled-up using 15L capacity airlift flat plate photobioreactors, under constant aeration with atmospheric air enriched with 5% CO_2 (Fig4). As shown on Figure 4, the exponential growth phase of the algae occurs until the 5th day of cultivation, followed by a phase of declining relative growth / stationary phase from the 5th to the 10th day of cultivation. Figure 4 also indicates that cultures are probably entering the death phase after the 10th day, which is characterized by a slight decrease in algal biomass until the 15th.

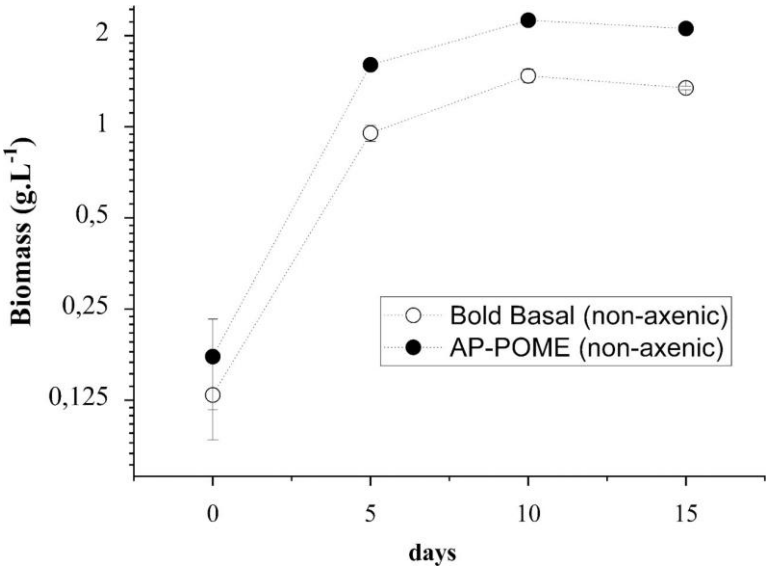


Fig 4. Growth dynamics of *Chlamydomonas biconvexa*Embrapa|LBA40 in airlift flat plate photobioreactors. The strain was batch cultivated using in airlift flat plate photobioreactors under non-axenic conditions. The working volume of 13 L of cultivation medium, bold basal medium (BBM) (··○··) or anaerobic pond palm oil mill effluent (AP-POME) (··●··), was used with constant aeration of 60 L·h-1 and CO2 supplementation adjusted to 5% of the air flow. Experimentation was conducted at 12h/12h light/dark cycling regimen, at a light intensity of 495 $\mu Em^{-2}\ s^{-1}$ (35000 Lux) and temperature of 25° C during dark periods and 35° C during the light periods. Biomass dry weight (DW) was gravimetrically determined using 10 ml samples retrieved from cultures from each photobioreactor replicate at the initial time and at 5 days intervals during 15 days of culturing. Results shown are the mean of biological triplicates of the experiment (n = 3).

Culture volume expansion testing is an essential step for large scale application, and the flat plate geometry represents one of the major types of photobioreactors available for production of microalgae [18]. After five days of cultivation, *Chlamydomonas*

biconvexa presented productivities of 190,6mgDW • L^{-1} • d^{-1} in POME (Fig 4 and Table 3). This is the highest algae biomass productivity reported to date using an aerobic pond derived POME without nutrient supplementation (Table 1).

Table 1. Comparison of biomass productivities and nutrient removal efficiencies obtained with microalgae species cultured in palm oil mill effluent (POME) under various conditions.

Strain	Culture conditions	Biomass Productivity ($mg.L^{-1}.d^{-1}$)	Nutrient removal	Reference
***Chlamydomonas biconvexa*. Embrapa|LBA40**	100% POME from anaerobic pond, supplemented with 5% of CO2 in 15L flat plate airlift photobioreactor	190.60	34.70% of Nitrate 99% of Nitrite 99% of Ammonium	This study
Chlorella* sp. *(UKM2)	100%POME from anaerobic pond supplemented with 10% of CO2 in 2 L flask.	64	80.90% of TN	[10]
***Scenedesmus* sp.**	84% diluted POME from anaerobic pond in 1L flask	20.40	74.40 % of COD 37 % of TN 52.42 % of Phosphate	[16]
***Chlorella sorokiniana* C 212**	75% diluted POME filtered supplemented with 60 mg/L of urea in 500ml	97	63% of COD	[17]
Chlorella sorokiniana	40% diluted POME from anaerobic pond, supplemented with 5% CO2 in 1L flask	90	64.30 % of Ammonium 62.30 % of Phosphate	[15]
Chlorella sorokiniana	40% diluted POME from anaerobic pond, supplemented with 5% CO2 in 1L flask	79.2	75.75% of TN 100% of TP 93.36% of Ammonium 94,50% of Phosphate	[11]
***Chlorella sorokiniana*CY-1**	30% diluted POME in BG11 medium supplemented with 2.5% of CO2 in 1L flask	150	47.09% of COD 62.07% of TN 30.77% of TP	[33]
Chlamydomonas sp UKM 6	16.6% diluted POME from anaerobic pond, supplemented with 2% CO2 in 2L flask	101	12.61% of COD 69.39% of TN 53.33% of TP 100% of Ammonium	[34]

Effect of algal cultivation on anaerobic pond palm oil mill effluent composition

Microalgae can be used to efficiently uptake inorganic compounds and heavy metals present in wastewater [35]. The presence of nitrogen and phosphorus is a determinant factor favoring eutrophication when the effluent is not properly disposed. On the other hand, these elements are essential for microalgae production [36]. Nitrogen sources such as nitrate, nitrite and ammonia are absorbed by different pathways that converge on the GS-GOGAT cycle for the production of amino acids and ultimately other nitrogenized metabolites [37,38]. Additionally, phosphorus is absorbed as soluble phosphate, in the form of hydrogen phosphate and dihydrogen phosphate, which is further directed to the synthesis of adenosine triphosphate (ATP), nucleic acids, phospholipids and coenzymes [37].

Here, the chemical composition of AP-POME prior and after *Chlamydomonas biconvexa* Embrapa|LBA40 cultivation in airlift flat plate photobioreactors was evaluated. Table 2 shows that there is high organic load (i.e.: Biochemical Oxygen Demand and Chemical Oxygen Demand) and ammonium in AP-POME used for *Chlamydomonas biconvexa* Embrapa|LBA40 cultivation.

Table 2. Physicochemical characterization of Anaerobic Pond Palm Oil Mill Effluent (AP-POME) before and after cultivation of *Chlamydomonas biconvexa* Embrapa|LBA40 in airlift flat plate photobioreactors for 5, 10 and 15 days.

Physicochemical characterization	**Crude POME**	**AP-POME**	**Culture supernatant 5 days**	**Culture supernatant 10 days**	**Culture supernatant 15 days**
Biochemical Oxygen Demand (mg.O2/L^{-1})	45967.20 (± 3603.27)^{e}	20.80 (± 1.70)^{a}	38.70 (±0.57)^{b}	49.85 (±1.20)^{c}	133.95 (±1.63)^{d}
Chemical Oxygen Demand (mg.O2/L^{-1})	145890.00 (± 19643.43)^{d}	667.00 (± 7.07)^{a}	734.00 (±4.24)^{b}	768.00 (±11.31)^{b}	1272.00 (±28.28)^{c}
Total organic carbon (mg.L^{-1})	84073.60 (± 1751.36)	3680.70 (± 243.39)^{a}	4678.40 (±778.38)^{ab}	4816.00 (±194.60)^{ab}	5916.60 (±194.31)^{b}
Nitrate (mg.L^{-1})	24.80 (±0.05)^{c}	0.72 (±0.00)^{b}	0.71 (±0.01)^{b}	0.56 (±0.03)^{a}	0.47 (±0.04)^{a}
Nitrite (mg.L^{-1})	0.45 (±0.05)^{b}	0.22 (±0.01)^{a}	n.d**	n.d**	n.d**
Ammonium (mg.L^{-1})	95.01 (± 15.03)^{b}	44.85 (± 0.64)^{a}	n.d**	n.d**	n.d**
Phosphate (mg.L^{-1})	630.00 (± 14.14)^{c}	56.75 (±2.47)^{b}	1.18 (±0.01) ^{a}	1.55 (±0.06)^{a}	1.43 (±0.71)^{a}
Potassium (mg.L^{-1})	3668.50 (± 338.70)^{b}	1850.00 (±70.71)^{a}	1575.00 (±106.10)^{a}	1750.00 (±70.71)^{a}	1875.00 (±35.36)^{a}
pH	4.59 (± 0.01)^{a}	9.17 (± 0.01)^{e}	8.37 (±0.02)^{c}	8.52 (±0.07)^{d}	8.01 (±0.01)^{b}

*Results are presented as mean ± error bars of triplicate experiments (n = 3). Means followed by the same letter within a row are not significantly different by One-way ANOVA with Tukey test at the 5% probability level ($p \leq 0.05$). **n.d = not detected.

Additionally, even after the anaerobic stabilization stage, the derived AP-POME still has a residual high biochemical and chemical oxygen demand, as well as a high concentration of nitrogen, that are above the levels permitted by most environmental policies and regulations [10]. Table 2 shows that *Chlamydomonas biconvexa* Embrapa|LBA40 culturing in AP-POME leads to 34.7% reduction of the nitrate after 15 days of cultivation, as well as 99% reduction of the nitrite and ammoniacal nitrogen levels present in the effluent after 5 days of cultivation. Furthermore, phosphate and potassium levels are decreased by 97.9% and 14.8%, respectively, after 5 days of cultivation.

It is noteworthy that, although nitrogen and phosphorous are effectively removed from AP-POME, potassium and TOC, increased in concentration after 10 days of cultivation (Table 2). This observation is parallel with a slight reduction in algal biomass during the same period (Fig 4). Taken together, these findings suggest that the alga reach stationary phase between day 5 and 10, with some cell lysis occurring, potentially from essential nutrient starvation. Therefore, the increase of potassium and TOC detected in culture supernatants is, probably, a consequence of algal cell death and lysis. This hypothesis was further corroborated by the observation of a progressive increase in cell lysis could be observed through microscope inspection of cultures over time (data not shown).

The use of microalgae for wastewater treatment offers an opportunity for reducing environmental pollution at low costs [35]. Although it has been reported that the cultivation of microalgae in POME can lead to nitrogen and phosphorous compounds reduction levels, algal biomass productivity reported in these previous studies is lower than reported in here (Table 1). Additionally, Table 1 show that the reduction of POME inorganic load was achieved after only 5 days of cultivation. This suggests that the aerobic pond stage, usually used in palm oil industrial plants (Fig1), could be substituted by high rate photobioreactors, significantly reducing the time and area requirements for wastewater treatment. Such possibility is particularly desirable considering sustainable palm oil production within the Amazon region, since this system could contribute to reduce the pressure for the expansion (or change) of the cultivated area.

Biomass composition analysis

Medium nutrient recovery translates into the production of algal biomass that could be used to favor the economic viability of the process in a biorefinery strategy (Brasil et al., 2017a). Therefore, algal biomass harvested from *C. biconvexa* Embrapa|LBA40 cultures in airlift flat plate photobioreactors was chemically analyzed in order to characterize its potential use as feedstock for added value bioproducts (Table 3).

Table 3. Biomass content, yield and compounds of interest accumulated in *Chlamydomonas biconvexa* Embrapa|LBA40 grown in airlift flat plate photobioreactors using Bold Basal medium (BBM) and anaerobic pond palm oil mill effluent (AP-POME).

*Chlamydomonas biconvexa*Embrapa|LBA40	**BBM 5 days**	**BBM 10 days**	**BBM 15 days**	**AP-POME 5 days**	**AP-POME 10 days**	**AP-POME 15 days**
Biomass productivity ($mg.L^{-1}.d^{-1}$)	190.6 (± 15)^{a}	147.3 (± 7.5)^{b}	89.5 (± 5)^{c}	164.7 (± 15)^{ab}	156 (± 1.1)^{b}	104 (± 17.3)^{c}
Carbohydrate content (%)	42.53 (± 1.7)^{a}	37.17 (± 1.5)^{abc}	30.97 (± 2.4)^{c}	31.96 (± 3.4)^{c}	42.35 (± 4.0)^{ab}	34.31 (± 3.8)^{bc}
Carbohydrate productivity ($mg.L^{-1}.d^{-1}$)	81.1 (± 3.3)^{a}	54.7 (± 2.2)^{bc}	27.7 (± 2.2)^{d}	52.6 (± 5.6)^{c}	66.1 (± 6.3)^{b}	35.7 (± 4.0)^{d}
Protein content (%)	21.6(± 0.12)^{b}	18.9 (± 0.14)^{c}	17.9 (± 0.08)^{d}	26 (± 0.18)^{a}	19.3 (± 0.41)^{c}	19.0 (± 0.13)^{c}
Protein productivity ($mg.L^{-1}.d^{-1}$)	41.1 (± 0.2)^{b}	27.8 (± 0.21)^{d}	15.7 (± 0.07)^{f}	42.8 (± 0.31)^{a}	30 (± 0.65)^{c}	19.8 (± 0.13)^{e}
Lipids content (%)	12 (± 0.43)^{a}	12.3 (± 0.52)^{a}	6.42 (± 0.31)^{c}	11.3 (± 0.8)^{ab}	9.95 (± 0.45)^{b}	6.99 (± 0.60)^{c}
Lipids productivity ($mg.L^{-1}.d^{-1}$)	22.9 (± 0.82)^{a}	18 (± 0.77)^{b}	5.7 (± 2.8)^{c}	18.6 (± 1.3)^{b}	15.5 (± 0.71)^{b}	7.3 (± 0.62)^{c}
Ash content (%)	6.45 (± 0.56)^{a}	4.8 (± 0.04)^{b}	5.2 (± 0.12)^{b}	6.68 (± 0.2)^{a}	4.64 (± 0.08)^{bc}	4.06 (± 0.18)^{c}

*Results are presented as mean ± standard deviation of triplicate experiments (n = 3). The values with contrast in the table obtained higher content or productivity compared to the other treatments. Means followed by the same letter within a row are not significantly different by One-way ANOVA with Tukey test at the 5% probability level ($p \leq 0.05$). AP-POME: anaerobic pond palm oil mill effluent.

Total carbohydrates comprised the largest fraction detected in the analyzed biomass, reaching 42.53% in BBM and 31.9% in AP-POME cultures after 5 days of algal growth (Table 3). Additionally, *C. biconvexa* Embrapa|LBA40 biomass also achieved its highest protein content (26.0%) and lipid content (11.3%) after 5 days of culture in AP-POME (Table 3). Comparatively, *C. biconvexa* Embrapa|LBA40 cultivated in sugarcane vinasse achieved 13.5% of carbohydrates, 41.7% of proteins and 1.6% of lipid contents [22]. Indeed, factors like medium composition and cultivation conditions can play a major influence on the biochemical composition of algal biomass [4,15,39].

The contents of lipids and proteins are higher on day 5, both in BBM and AP-POME, decreasing on days 10 and 15 (Table 3). The carbohydrate content also decreases over time in algae biomass derived from BBM cultures. Although there is an increase in the content of carbohydrate in biomass derived from AP-POME on day 10, it falls on day 15. Algae cells have been shown to accumulate carbohydrates or lipids as sources of energy during the stationary phase [40]. The analysis of algae growth kinetics (Fig.4), together with the composition of algae biomass (Table 4), suggests that the stationary growth phase of *C. biconvexa* Embrapa|LBA40 in AP-POME occurs before the 10th day of cultivation, followed by a death phase caused by depletion of essential nutrients (ie: nitrogen and phosphorus – Table 2), and characterized by accelerated cell death and decrease in biomass content after day 10. It is worth mentioning that this finding is of practical relevance, since the rate of culture batches/harvest cycles must be optimized to achieve technical-economic viability on an industrial scale. Therefore, the implementation of a fed-batch culture process for *C. biconvexa* Embrapa|LBA40 should include harvesting the biomass before the 10th day of growth followed by the renewal of the medium (ie: AP-POME) to replace the depleted nutrients.

Biomolecule profiling is an important step in algal biomass characterization. Carbohydrate profiling revealed the presence of myo-inositol, galactose, mannose, ribose and glucose in *C. biconvexa* Embrapa|LBA40 biomass cultivated in AP-POME. Glucose represents the largest fraction corresponding to up to 70% of total carbohydrate (Table 4). Glucose represents the largest fraction corresponding to up to 70% of total carbohydrate (Table 4). It has been reported that algal biomass can be hydrolyzed leading to glucose release, which in turn, can be used for bioethanol and other chemicals production [41]. The second largest fraction observed in *C. biconvexa* Embrapa|LBA40

biomass is mannose (Table 4). This is expected since the cell wall from *Chlamydomonas* species is reported to be composed mainly of glycoproteins [42]. Indeed, mannose is the main monosaccharide involved in N-glycosylation and the formation of glycoproteins in *Chlamydomonas reinhardtii* [43,44].

Over the time course of algal culturing (Table 4), it can be observed a slight decrease in the glucose and ribose contents, followed by an increase in the content of mannose. These findings suggest that a reduction in the photosynthetic activity (Calvin – Benson cycle) due to the limitation of nitrogen and phosphorus is occurring after 10 days of algal growth (table 2) [45, 46].

Table 4. Profile of carbohydrate accumulated in *Chlamydomonas biconvexa*Embrapa|LBA40 grown in airlift flat plate photobioreactors using Bold Basal medium (BBM) and anaerobic pond palm oil mill effluent (AP-POME).

***Chlamydomonas biconvexa* Embrapa|LBA40**	**BBM 5 days**	**BBM 10 days**	**BBM 15 days**	**AP-POME 5 days**	**AP-POME 10 days**	**AP-POME 15 days**
Myo-inositol (%)	0.99 (±0,03)[a]	1.18 (±0,02)[a]	1.92 (±0,12)[a]	0.66 (0.05)[a]	1.29 (0.07)[a]	1.46 (0.26)[a]
Arabinose (%)	n.d.	0.52 (±0,03)	n.d.	n.d.	n.d.	n.d.
Galactose (%)	4.04 (±0,01)[a]	2.98 (±0,19)[b]	2.97 (±0,04)[b]	2.93 (0.03)[b]	2.72 (0.14)[b]	2.30 (0.08)[c]
Glucose (%)	77.73 (±0,27)[a]	76.59 (±1,67)[a]	71.48 (±0,45)[b]	77.73 (0.58)[a]	77.13 (0.57)[a]	70.92 (1.90)[b]
Mannose (%)	12.64 (±0,22)[b]	11.82 (±1,37)[b]	18.04 (±0,43)[a]	12.87 (0.04)[b]	11.94 (0.82)[b]	18.82 (1.97)[a]
Ribose (%)	2.22 (±0,13)[a]	2.06 (±0,11)[b]	1.62 (±0,02)[c]	2.16 (0.07)[a]	2.43 (0.07)[a]	2.04 (0.23)[b]
Others (%)	3.44 (±0,14)[c]	3.81 (±0,27)[bc]	4.65 (±0,07)[a]	3.61 (0.09)[b]	4.29 (0.23)[a]	4.04 (0.60)[a]

*Results are presented as mean ± standard deviation of triplicate experiments (n = 3). Means followed by the same letter letter within a row are not significantly different by One-way ANOVA with Tukey test at the 5% probability level ($p \leq 0.05$). AP-POME: anaerobic pond palm oil mill effluent.

As an important parameter for the quality of the oil, the fatty acid profile changes in algae biomass cultivated under different conditions were analyzed [47]. The most abundant fatty acid in the biomass of *C. biconvexa* Embrapa|LBA40 is palmitic acid (ranging from 37% to 42%), followed by oleic acid (ranging from 21% to 28%) (Table 5).

Table 5. Profile of Fatty acid methyl esters (FAME) accumulated in *Chlamydomonas biconvexa* Embrapa|LBA40 grown in airlift flat plate photobioreactors using Bold Basal medium (BBM) and anaerobic pond palm oil mill effluent (AP-POME).

***Chlamydomonas biconvexa*Embrapa|LBA40**	**BBM 5 days**	**BBM 10 days**	**BBM 15 days**	**AP-POME 5 days**	**AP-POME 10 days**	**AP-POME 15 days**
Caproic (%)	2.22 (0.09)^{a}	1.87 (0.06)^{cd}	1.81 (0.01)^{cd}	2.11 (0.12)^{ab}	1.73 (0.04)^{d}	1.99 (0.09)^{bc}
Palmitic (%)	37.80 (0.46)^{cd}	42.77 (0.65)^{a}	40.32 (0.47)^{b}	37.55 (0.24)^{d}	38.55 (0.31)^{cd}	38.90 (0.08)^{c}
Palmitoleic (%)	2.48 (0.06)^{d}	2.51 (0.03)^{d}	2.63 (0.04)^{c}	2.89 (0.02)^{b}	3.08 (0.02)^{a}	2.98 (0.03)^{b}
Stearic (%)	4.14 (0.04)^{d}	4.92 (0.09)^{b}	4.59 (0.05)^{c}	4.42 (0.07)^{c}	4.21 (0.05)^{d}	5.53 (0.02)^{a}
Oleic (%)	25.28 (0.48)^{c}	28.63 (0.24)^{a}	27.40 (0.26)^{b}	21.56 (0.20)^{d}	25.16 (0.41)^{c}	24.53 (0.17)^{c}
Linoleic (%)	12.99 (0.04)^{c}	8.92 (0.26)^{e}	12.50 (0.18)^{d}	13.91 (0.11)^{a}	13.31 (0.05)^{bc}	13.46 (0.05)^{b}
Gamma-Linolenic (%)	0.96 (0.02)^{d}	1.90 (0.04)^{c}	0.62 (0.01)^{e}	2.62 (0.02)^{a}	2.13 (0.07)^{b}	0.85 (0.01)^{d}
Alpha-Linolenic (%)	14.09 (0.06)^{b}	8.43 (0.28)^{e}	10.10 (0.21)^{d}	14.92 (0.12)^{a}	11.78 (0.08)^{c}	11.73 (0.09)^{c}

*Results are presented as mean ± standard deviation of triplicate experiments (n = 3). Means followed by the same letter within a row are not significantly different by One-way ANOVA with Tukey test at the 5% probability level ($p \leq 0.05$). AP-POME: anaerobic pond palm oil mill effluent.

Palmitic acid was reported as the largest fraction in *Chlamydomonas sp.* grown in BBM [48] and in other wastewaters, such as in vinasse (32%) and raw chicken manure (35%) [39]. Palmitic acid was also shown to be the main fatty acid produced under sodium acetate stress [49], at nutrient limitation [47, 50, 51], at high temperatures [52], and at high light intensity [53]. Indeed, a limitation in the content of nitrogen and phosphorous in the AP-POME culture supernatants is observed after 5 days of cultivation (Table 2). Furthermore, *C. biconvexa* Embrapa|LBA40 culturing in AP-POME was conducted at high temperatures, reaching 35º C during the light period, and high luminous intensity (i.e.: 35000 lux) (Fig 4). Thus, it seems reasonable to hypothesize that the cultivation conditions observed after 5 days of algal growth might promote an increase in the levels of palmitic acid and of oxygen reactive species [50]. This oxidative stress could also explain the increase in oleic acid (C18:1) levels, and the parallel decrease in the contents of unsaturated fatty acids such gamma-linolenic (C18: 1n9c), linolenic (18: 2n6c) and alpha-Linolenic (18: 2n3), observed over the time course of algal culturing (Table 5) [50]".

Taken together these results reveal that *C. biconvexa* Embrapa|LBA40 biomass could potentially be used as source for the production of biofuels, animal feed, fertilizers, nutraceutical and oleo chemical bio-products [4,12]. In particular, applications of algal biomass as animal feed for in land aquaculture, as soil fertilizers or growth promoters would be of interest in the context of a microalgae production integrated to palm oil industrial plants within the Amazonian region. Indeed, it has been shown that microalgae can be beneficial to aquaculture, as they provide a rich source of micronutrients, lipids and proteins essential for fish farming [54]. Furthermore, different microalgae species can improve soil quality, promote atmospheric nitrogen fixation and produce plant growth hormones [55]. Saadaoui and collaborators [56] have reported a positive effect of applications of algae-based biofertilizer on date palm (*Phoenix dactylifera* L.) cultivation. However, experimentation targeting the evaluation of *C. biconvexa* Embrapa|LBA40 biomass for such applications remain an issue to be pursued further.

Genomic analysis

The *Chlamydomonas* genus comprise hundreds of flagellated unicellular green algae species. Recently, however, phylogeny studies based on molecular markers have

unveiled polyphyletic origin for species originally assigned to Chlamydomonas genus based on morphological data [57,58].

Therefore, a significant reclassification of Chlamydomonas genus species is in course, with several species being reassigned to genera like Oogamochlamys, Chloromonas, Dangeardinia, Ixipapillifera, Rhysamphichloris and Lobochlamys [57,58]. Genomic information about *C. biconvexa* species are scarce in the literature and taxonomic assignment of this species has relied basically upon morphological and physiological data [20]. In order to confirm *C. biconvexa* Embrapa|LBA40 taxonomic status within Chamydomonas genus and to expand the genetic knowledge basis of this species, mitochondrial DNA (mtDNA) genomic sequencing, annotation and phylogenetic analysis was performed.

C. biconvexa Embrapa|LBA40 mitogenome was assembled through the generation of a consensus sequence with a total of 15980 nucleotides (maximum coverage of 30) and 44.51% of GC content. Fourteen (14) genes could be annotated within the mitogenome, including the protein-coding genes cob, cox1, nad1_0, nad1_1, nad2, nad4, nad5, nad6 and rtl (Fig 5; Table 6). The *C. biconvexa* Embrapa|LBA40 mitogenome sequence data was deposited at NCBI GenBank under the accession number MG916975.1.

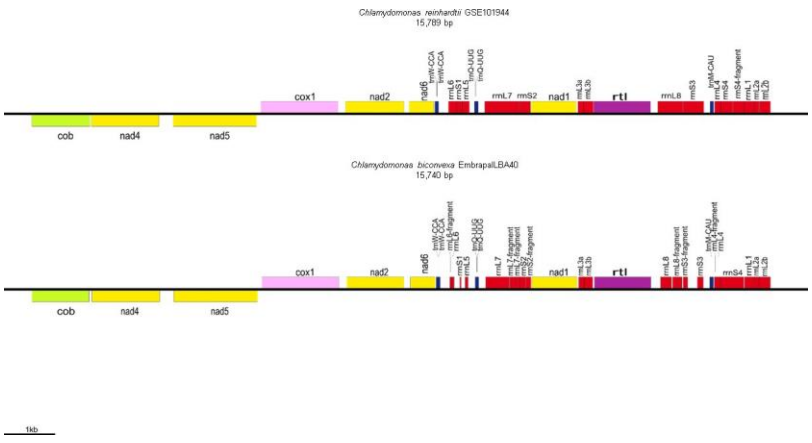


Fig 5. Representation and comparison of the mtDNA from *Chlamydomonas biconvexa*Embrapa|LBA40 strain and mtDNA from the green microalgae reference *Chlamydomonas reinhardtii* [66]. Genes: nad1 - Subunit 1 NADH dehydrogenase fragmented; rnL7, rrnL8 - Large subunit ribosomal RNA; trnQ-UUG - Transfer RNA, initiation codon amino acid tryptophan; rrnS1, rrnS3, rrnS4 - Small subunit ribosomal RNA; trnW-CCA - Transfer RNA, initiation codon amino acid glutamine; nad6 - Subunit 6 NADH dehydrogenase; nad2 - Subunit 2 NADH dehydrogenase; cox1 - Cytochrome c oxidase subunit I; cob - Apocytochrome b protein; nad5 - Subunit 5 NADH

dehydrogenase; nad4 - Subunit 4 NADH dehydrogenase; trnM - Transfer RNA, initiation codon amino acid methionine; nad1 - Subunit 1 NADH dehydrogenase.

Table 6. Main gene content and features of *Chlamydomonas biconvexa* Embrapa|LBA40 complete mitogenome.

Gene	Strand (+ or -)	Genome position		Length (bp)	Protein product or function
		Start	End		
nad1_0	-	0	199	199	Subunit 1 NADH dehydrogenase, fragment
rrnL	-	452	1108	657	Large subunit ribosomal RNA
trnQ (ttg)	-	1221	1294	73	Transfer RNA. Initiation codon amino acid tryptophan
rrnS	-	1572	1602	31	Small subunit ribosomal RNA
trnW (cca)	-	1882	1958	75	Transfer RNA. Initiation codon amino acid glutamine
nad6	-	2146	2527	382	Subunit 6 NADH dehydrogenase
nad2	-	2731	3574	844	Subunit 2 NADH dehydrogenase
cox1	-	3813	5316	1504	Cytochrome c oxidase subunit I
cob	-	5639	6521	883	Apocytochrome b protein
nad5	+	6997	9309	2313	Subunit 5 NADH dehydrogenase
nad4	+	9462	10605	1144	Subunit 4 NADH dehydrogenase,
trnM (cat)	+	11914	11987	74	Transfer RNA. Initiation codon amino acid methionine
nad1_1	+	12975	13731	757	Subunit 1 NADH dehydrogenase, fragment
rtl	+	14343	14928	586	Reverse Transcriptase-Like

The tRNAs coding genes encoding trnW, trnQ and trnM, as well rRNAs rrnS and rrnL, the small and large subunit ribosomal RNA, respectively, were also found. The protein-coding genes present in *C. biconvexa* Embrapa|LBA40 mitogenome are responsible for cellular respiration, specifically in the oxidative phosphorylation and ATP synthesis pathways, processed in mitochondrion.

Only three tRNAs are coded in mitochondrial genome of *C. biconvexa* Embrapa|LBA40, tRNAMet, tRNATrp and tRNAGln (Table 6). The same is observed in the mitochondrial genome of *C. reinhardtii* and *C. incerta* species, suggesting that there might be importation of other tRNAs to the mitochondrion *in vivo* [59].

On the other hand, the cox2 and cox3 subunit genes, which are characteristic of Reinhardtinia clade species, were not found in *C. biconvexa* Embrapa|LBA40 mtDNA. Gene transfer from the mitochondrion to the nucleus throughout the evolution process might explain this feature. According to Pérez-Martínez and collaborators [60], gene transfer process can be considered evolutionary advantageous, since nuclear genes exhibit lower mutation rates, mainly due to a better repair DNA system and genetic stability compared to the mitochondrial environment. Furthermore, the loss of genetic information can also be explained by the need to reduce the energy demand during protein synthesis, leaving only to cox1 subunit the role of cytochrome C oxidase coding inside the mitochondrion.

Chlamydomonas reinhardtii is the holotype species of the genus and has been studied for decades as a model for photosynthetic organisms [61]. There is considerable amount of information available about this species genome, mutant strains, culture media, preservation protocols, sexual life cycle and genetic modification methods [4]. Genes like rRNA rrnL and rrnS are fragmented, including three fragments for rrnL and five for rrnS in *C. reinhardtii*, a similar gene fragmentation has been reported for the species as *C. incerta*, though [59]. tRNAs secondary structures *in silico* predictions based on *C. biconvexa* Embrapa|LBA40 mtDNA revealed the presence of both amino acid and anticodon arms, consistent with the cloverleaf shape (S1 fig.). While both trnM and trnQ predicted structures from *C. biconvexa* Embrapa|LBA40 possess internal loops between the “T” arm, the *C. reinhardtii* trnM is the only gene with this feature (S1 fig.). This feature decreases thermodynamic stability of tRNA structure compared to perfect double-strand pairing [62].

The most conspicuous mitogenomic features reported for species of the Reinhardtinia clade are summarized in Table 7.

Table 7. Mitochondrial genome features of Reinhardtinia clade algae.

Genus and species	mtDNA Architecture	Size (kb)	Number of genes	GC Content (%)	GenBank Accession	Reference
Polytomella sp.	Linear	13.1	19	42	GU108480.1	(Smith et al. 2010)
Polytomauvella	Circular	17.4	19	55	NC_026572.1	(Smith et al. 2013)
Eudorina sp.	Circular	20.7	24	Not described	KY442294.1	(Hamaji et al. 2017)
Volvox carteri	Linear	29.9	26	44	EU760701.1	(Smith, Lee 2009)
Pleodorinastarrii	Circular	20.4	25	38	NC_021108.1	(Smith et al. 2013)
Chlamydomonas reinhardtii	Linear	15.8	13	45,2	NC_001638.1	(Vahrenholz et al. 1993)
C.biconvexa Embrapa|LBA40	Linear*	15.7	13	44,5	MG916975.1	This study

*Based on *in silico* analysis.

It can be observed that there is a wide genome architecture variation, including both linear and circular conformations and a wide range of mtDNA sizes, number of genes and GC content. This diversity is characteristic of Chlamydomonadalean algae [63]. It has been proposed that at least three conformational changes (from linear to circular and vice-versa) occurred during evolutionary divergence within this clade. Although, the biological significance of this architecture variation remains elusive, the GC content diversity observed within Chlamydomonadales an algae might be related to ecological adaptation for thermal and/or UV tolerance, as well as specific gene regulation mechanisms [64]. Therefore, Chlamydomonadales mtDNAs can provide a rich source of data to support the reconstruction of phylogenetic history of this group. It is important to highlight, though, that the putatively linear architecture of the mtDNA inferred here base on *in silico* analysis requires future experimental corroboration (e.g.: analysis of telomeres).

Previous species-level identification based on the analysis of the chloroplast marker, Ribulose Bisphosphate Carboxylase Large subunit gene (*rbcL*), and the nuclear markers, Internal Transcribed Spacers 1 and 2 of the nuclear rDNA (*nu*ITS1 and *nu*ITS2), as well as traditional morphological analysis, have assigned Embrapa|LBA40 to *C. biconvexa* species [19]. In order to assess *C. biconvexa* Embrapa|LBA40 identification through the use of mitochondrial markers, the cox1 subunit gene from Chlorophyta species was used to reconstruct a phylogenetic tree (Fig 6).

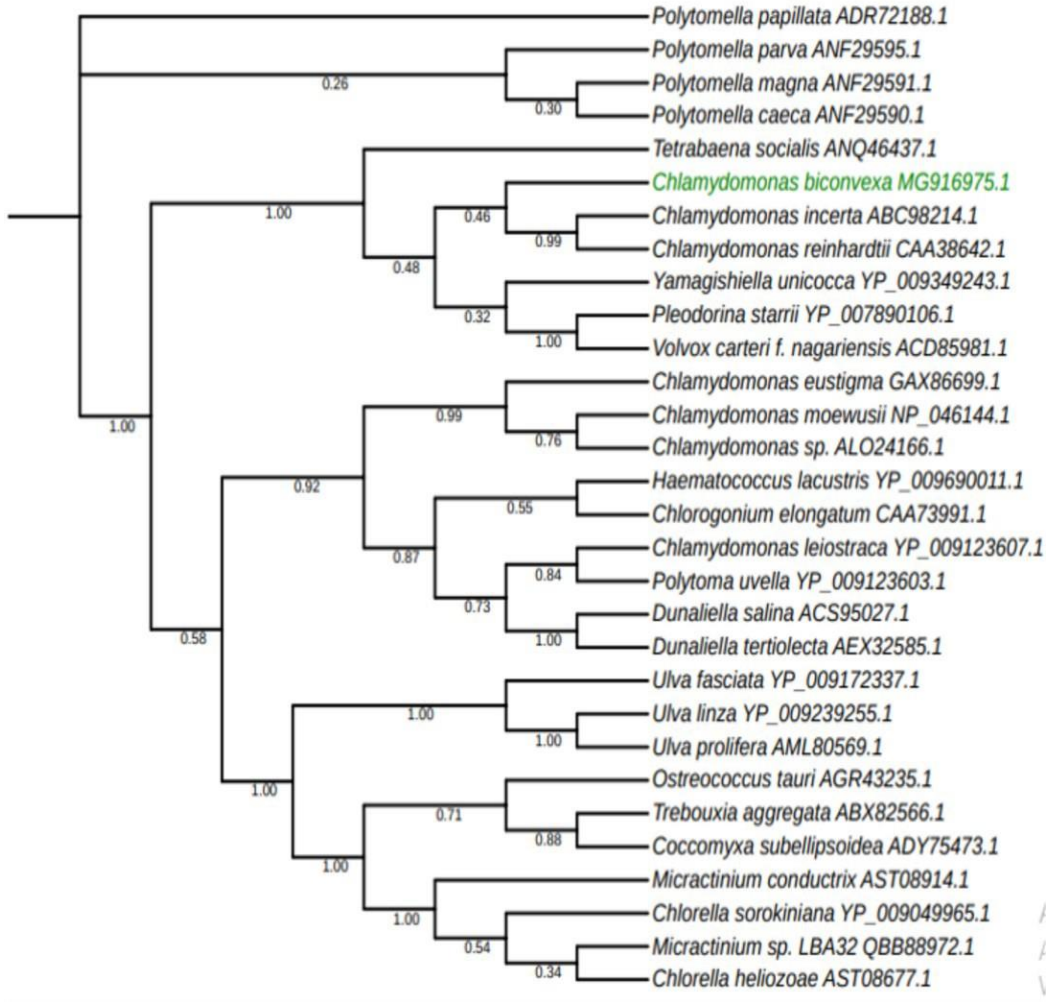


Fig 6. Phylogenetic tree based on mitochondrial cox1 gene sequence inferred by using Maximum Likelihood method and JTT matrix-based model. The bootstrap consensus tree inferred from 100 replicates. Analysis involved amino acid cox 1 gene sequences from 30 microalgae strains and were conducted in MEGA X.

Cytochrome oxidase subunit 1 is a single-copy coding gene informative for phylogenetic analysis in diverse taxa, since it evolves at moderate rates, conserving phylogenetic signatures during the evolutionary process [65]. The tree shows that *C. biconvexa* Embrapa|LBA40strain forms a monophyletic branch with *C. reinhardtii*, corroborating the taxonomic assignment of *C. biconvexa*Embrapa|LBA40 to the Chlamydomonas genus and reinhardtinia clade. This information also suggests that its evolutionary proximity to *C. reinhardtii* might allow the interchangeable use of molecular tools and protocols between both species.

Conclusions

The wastewater tolerant strain *Chlamydomonas biconvexa* Embrapa | LBA40 is capable of achieving high productivity when grown in POME, providing biomass that could potentially be used as a source for the production of biofuels, animal feed, fertilizers, nutraceutical or oleochemical bioproducts. The algae culture did not require any supplementation of nutrients other than CO_2 and was successfully carried out under non-axenic conditions, indicating that *C. biconvexa* Embrapa | LBA40 might be a robust strain for industrial scale production. Future studies should focus on scaling up the cultivation and harvesting processes. In addition, the inorganic load of POME is drastically reduced after five days of algae cultivation. This suggests that the aerobic pond stage normally used in industrial palm oil plants could be replaced by high-rate photobioreactors, significantly reducing the time and area required for wastewater treatment. Such a possibility is particularly desirable considering the sustainable production of palm oil in the Amazon region, since this system could contribute to reducing the pressure for the expansion (or change) of the cultivated area. Together, the results unveil a potential use of microalgae in biorefineries integrated with palm oil plants. It is worth mentioning that, once these integrated processes reach industrial scale, the biorefinery may benefit from the commercialization of decarbonization credits provided by the Brazilian National Biofuel Policy (RenovaBio). Therefore, key mechanisms related to technology, economics, environmental sustainability and policy support for algal integrated biorefineries might be available in Brazil in the future. Finally, mitogenomic analysis confirmed that *C. biconvexa* Embrapa|LBA40 forms a monophyletic branch with *C. reinhardti*, opening up opportunities for the use of molecular tools and protocols developed for this prototypic species in future breeding programs.

Acknowledgments

The authors are grateful to Dr. Thomas Christopher Rhys Williams (Botany department of Brasilia University) for the important help in biochemical orientations.

References

1. Brasil BSAF, Silva FCP, Siqueira FG. Microalgae biorefineries: The Brazilian scenario in perspective. N Biotechnol. 2017;39: 90–98. doi:10.1016/j.nbt.2016.04.007

2. Saravanan AP, Mathimani T, Deviram G, Rajendran K, Pugazhendhi A. Biofuel policy in India: A review of policy barriers in sustainable marketing of biofuel. J Clean Prod. 2018;193: 734–747. doi:10.1016/j.jclepro.2018.05.033

3. Saravanan AP, Pugazhendhi A, Mathimani T. A comprehensive assessment of biofuel policies in the BRICS nations: Implementation, blending target and gaps. Fuel. 2020;272: 117635. doi:10.1016/j.fuel.2020.117635

4. Brasil BSAF, Silva FCP, Siqueira FG. Microalgae biorefineries: The Brazilian scenario in perspective. N Biotechnol. 2017;39: 90–98. doi:10.1016/j.nbt.2016.04.007

5. Center For International Development At Harvard University. The Atlas of Economic Complexit. 2019.

6. Benami E, Curran LM, Cochrane M, Venturieri A, Franco R, Kneipp J, et al. Oil palm land conversion in Pará, Brazil, from 2006-2014: Evaluating the 2010 Brazilian Sustainable Palm Oil Production Program. Environ Res Lett. 2018;13. doi:10.1088/1748-9326/aaa270

7. Ahmad A, Buang A, Bhat AH. Renewable and sustainable bioenergy production from microalgal co-cultivation with palm oil mill effluent (POME): A review. Renew Sustain Energy Rev. 2016;65: 214–234. doi:10.1016/j.rser.2016.06.084

8. Glaz P, Bartosiewicz M, Laurion I, Reichwaldt ES, Maranger R, Ghadouani A. Greenhouse gas emissions from waste stabilisation ponds in Western Australia and Quebec (Canada). Water Res. 2016;101: 64–74. doi:10.1016/j.watres.2016.05.060

9. Von Sperling M. Waste Stabilisation Ponds. Water Intelligence Online. 2015. doi:10.2166/9781780402109

10. Hariz HB, Takriff MS, Ba-Abbad MM, Mohd Yasin NH, Mohd Hakim NIN. CO

2 fixation capability of Chlorella sp. and its use in treating agricultural wastewater. J Appl Phycol. 2018;30: 3017–3027. doi:10.1007/s10811-018-1488-0

11. Khalid AAH, Yaakob Z, Abdullah SRS, Takriff MS. Analysis of the elemental composition and uptake mechanism of Chlorella sorokiniana for nutrient removal in agricultural wastewater under optimized response surface methodology (RSM) conditions. J Clean Prod. 2019;210: 673–686. doi:10.1016/j.jclepro.2018.11.095

12. Khan MI, Shin JH, Kim JD. The promising future of microalgae: Current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. Microb Cell Fact. 2018;17: 1–21. doi:10.1186/s12934-018-0879-x

13. Bhujade R, Chidambaram M, Kumar A, Sapre A. Algae to economically viable low-carbon-footprint oil. Annu Rev Chem Biomol Eng. 2017;8: 335–357. doi:10.1146/annurev-chembioeng-060816-101630

14. Santana H, Cereijo CR, Teles VC, Nascimento RC, Fernandes MS, Brunale P, et al. Microalgae cultivation in sugarcane vinasse: Selection, growth and biochemical characterization. Bioresour Technol. 2017;228: 133–140. doi:10.1016/j.biortech.2016.12.075

15. Khalid AAH, Yaakob Z, Abdullah SRS, Takriff MS. Growth improvement and metabolic profiling of native and commercial Chlorella sorokiniana strains acclimatized in recycled agricultural wastewater. Bioresour Technol. 2018;247: 930–939. doi:10.1016/j.biortech.2017.09.195

16. Hariz HB, Takriff MS, Mohd Yasin NH, Ba-Abbad MM, Mohd Hakimi NIN. Potential of the microalgae-based integrated wastewater treatment and CO2 fixation system to treat Palm Oil Mill Effluent (POME) by indigenous microalgae; Scenedesmus sp. and Chlorella sp. J Water Process Eng. 2019;32. doi:10.1016/j.jwpe.2019.100907

17. Nwuche C. Use of Palm Oil Mill Effluent as Medium for Cultivation of Chlorella sorokiniana. Br Biotechnol J. 2014;4: 305–316. doi:10.9734/bbj/2014/6921

18. Santana H, Cereijo CR, Teles VC, Nascimento RC, Fernandes MS, Brunale P, et al. Microalgae cultivation in sugarcane vinasse: Selection, growth and biochemical characterization. Bioresour Technol. 2017;228: 133–140. doi:10.1016/j.biortech.2016.12.075

19. Hadi SIIA, Santana H, Brunale PPM, Gomes TG, Oliveira MD, Matthiensen A, et al. DNA barcoding green microalgae isolated from neotropical inland waters. PLoS One. 2016;11: 1–18. doi:10.1371/journal.pone.0149284

20. Pascher A. Die Siisswasserjlora Deutschlands, Osterreichs und der Schweiz. Nature. 1925.

21. Fernandes MS, Calsing LCG, Nascimento RC, Santana H, Morais PB, de Capdeville G, et al. Customized cryopreservation protocols for chlorophytes based on cell morphology. Algal Res. 2019;38: 101402. doi:10.1016/j.algal.2018.101402

22. Santana H, Cereijo CR, Teles VC, Nascimento RC, Fernandes MS, Brunale P, et al. Microalgae cultivation in sugarcane vinasse: Selection, growth and biochemical characterization. Bioresour Technol. 2017;228: 133–140. doi:10.1016/j.biortech.2016.12.075

23. Laurens LML. Summative Mass Analysis of Algal Biomass – Integration of Analytical Procedures Summative Mass Analysis of Algal Biomass – Integration of Analytical Procedures Laboratory Analytical Procedure (LAP). 2015.

24. Van Wychen S, Laurens LML. Determination of Total Solids and Ash in Algal Biomass: Laboratory Analytical Procedure (LAP). 2016. doi:10.2172/1118077

25. Van Wychen S, Ramirez K, Laurens LM. Determination of Total Lipids as Fatty Acid Methyl Esters (FAME) by in situ Transesterification. Contract. 2013;303: 275–3000. Available: www.nrel.gov/publications

26. Kishi M, Toda T. Carbon fixation properties of three alkalihalophilic microalgal strains under high alkalinity. J Appl Phycol. 2018;30: 401–410. doi:10.1007/s10811-017-1226-z

27. Doyle J. DNA Protocols for Plants. Mol Tech Taxon. 1991; 283–293.

doi:10.1007/978-3-642-83962-7_18

28. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010 p. http://www.bioinformatics.babraham.ac.uk/projects/.

29. Ebenezer TE, Zoltner M, Burrell A, Nenarokova A, Novák Vanclová AMG, Prasad B, et al. Transcriptome, proteome and draft genome of Euglena gracilis. BMC Biol. 2019;17: 1–23. doi:10.1186/s12915-019-0626-8

30. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: Improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69: 313–319. doi:10.1016/j.ympev.2012.08.023

31. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47: W59–W64. doi:10.1093/nar/gkz238

32. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35: 1547–1549. doi:10.1093/molbev/msy096

33. Cheah WY, Show PL, Juan JC, Chang JS, Ling TC. Microalgae cultivation in palm oil mill effluent (POME) for lipid production and pollutants removal. Energy Convers Manag. 2018;174: 430–438. doi:10.1016/j.enconman.2018.08.057

34. Ding GT, Yaakob Z, Takriff MS, Salihon J, Abd Rahaman MS. Biomass production and nutrients removal by a newly-isolated microalgal strain Chlamydomonas sp in palm oil mill effluent (POME). Int J Hydrogen Energy. 2016;41: 4888–4895. doi:10.1016/j.ijhydene.2015.12.010

35. Manzoor F, Karbassi A, Golzary A. Heavy Metal Contaminants Removal from wastewater by using Chlorella vulgaris: A Review. Curr Environ Eng. 2019;06: 1–13. doi:10.2174/2212717806666190716160536

36. Park JBK, Craggs RJ, Shilton a. N. Wastewater treatment high rate algal ponds for biofuel production. Bioresour Technol. 2011;102: 35–42. doi:10.1016/j.biortech.2010.06.158

37. Li G, Bai X, Li H, Lu Z, Zhou Y, Wang Y, et al. Nutrients removal and biomass production from anaerobic digested effluent by microalgae: A review. Int J Agric Biol Eng. 2019;12: 8–13. doi:10.25165/j.ijabe.20191205.3630

38. Ribeiro DM, Fernando L, Cunha G, Costa L, Jungmann L, Christopher T, et al. A low-cost approach for Chlorella sorokiniana production through combined use of urea , ammonia and nitrate based fertilizers. Bioresour Technol Reports. 2020;9: 100354. doi:10.1016/j.biteb.2019.100354

39. Calixto CD, da Silva Santana JK, de Lira EB, Sassi PGP, Rosenhaim R, da Costa Sassi CF, et al. Biochemical compositions and fatty acid profiles in four species of microalgae cultivated on household sewage and agro-industrial residues. Bioresour Technol. 2016;221: 438–446. doi:10.1016/j.biortech.2016.09.066

40. Qu W, Loke Show P, Hasunuma T, Ho SH. Optimizing real swine wastewater treatment efficiency and carbohydrate productivity of newly microalga Chlamydomonas sp. QWY37 used for cell-displayed bioethanol production. Bioresour Technol. 2020;305: 123072. doi:10.1016/j.biortech.2020.123072

41. de Farias Silva CE, Bertucco A. Bioethanol from microalgae and cyanobacteria: A review and technological outlook. Process Biochem. 2016;51: 1833–1842. doi:10.1016/j.procbio.2016.02.016

42. Miller DH, Mellman IS, Lamport DTA, Miller M. The chemical composition of the cell wall of chlamydomonas gymnogama and the concept of a plant cell wall protein. J Cell Biol. 1974;63: 420–429. doi:10.1083/jcb.63.2.420

43. Mathieu-Rivet E, Scholz M, Arias C, Dardelle F, Schulze S, Le Mauff F, et al. Exploring the N-glycosylation pathway in chlamydomonas reinhardtii unravels novel complex structures. Mol Cell Proteomics. 2013;12: 3160–3183. doi:10.1074/mcp.M113.028191

44. Vanier G, Lucas PL, Loutelier-Bourhis C, Vanier J, Plasson C, Walet-Balieu ML, et al. Heterologous expression of the N-acetylglucosaminyltransferase i dictates a reinvestigation of the N-glycosylation pathway in Chlamydomonas reinhardtii. Sci Rep. 2017;7: 1–12. doi:10.1038/s41598-017-10698-z

45. Kamalanathan M, Pierangelini M, Shearman LA, Gleadow R, Beardall J. Impacts of nitrogen and phosphorus starvation on the physiology of Chlamydomonas reinhardtii. J Appl Phycol. 2016;28: 1509–1520. doi:10.1007/s10811-015-0726-y

46. Talebi AF, Tohidfar M, Mousavi Derazmahalleh SM, Sulaiman A, Baharuddin AS, Tabatabaei M. Biochemical modulation of lipid pathway in microalgae Dunaliella sp. for biodiesel production. Biomed Res Int. 2015;2015. doi:10.1155/2015/597198

47. Lu N, Chen JH, Wei D, Chen F, Chen G. Global metabolic regulation of the snow alga chlamydomonas nivalis in response to nitrate or phosphate deprivation by a metabolome profile analysis. Int J Mol Sci. 2016;17. doi:10.3390/ijms17050694

48. Salama ES, Kim HC, Abou-Shanab RAI, Ji MK, Oh YK, Kim SH, et al. Biomass, lipid content, and fatty acid composition of freshwater Chlamydomonas mexicana and Scenedesmus obliquus grown under salt stress. Bioprocess Biosyst Eng. 2013;36: 827–833. doi:10.1007/s00449-013-0919-1

49. Yang L, Chen J, Qin S, Zeng M, Jiang Y, Hu L, et al. Growth and lipid accumulation by different nutrients in the microalga Chlamydomonas reinhardtii. Biotechnol Biofuels. 2018;11: 1–12. doi:10.1186/s13068-018-1041-z

50. Shi K, Gao Z, Shi TQ, Song P, Ren LJ, Huang H, et al. Reactive oxygen species-mediated cellular stress response and lipid accumulation in oleaginous microorganisms: The state of the art and future perspectives. Front Microbiol. 2017;8: 1–9. doi:10.3389/fmicb.2017.00793

51. Qari HA, Oves M. Fatty acid synthesis by Chlamydomonas reinhardtii in phosphorus limitation. J Bioenerg Biomembr. 2020;52: 27–38. doi:10.1007/s10863-019-09813-8

52. James GO, Hocart CH, Hillier W, Price GD, Djordjevic MA. Temperature modulation of fatty acid profiles for biofuel production in nitrogen deprived Chlamydomonas reinhardtii. Bioresour Technol. 2013;127: 441–447. doi:10.1016/j.biortech.2012.09.090

53. Chouhan N, Raju Devadasu E, Mohan Yadav R. Autophagy induced accumulation of lipids in pgrl1 and pgr5 of Chlamydomonas reinhardtii under high light 2 3. bioRxiv. 2020; 2020.09.14.296244. Available: https://doi.org/10.1101/2020.09.14.296244

54. Semary NA El. Algae and Chain Aquaculture: An Approach Towards Sustainable Agriculture. Handb Environ Chem. 2017;5: 1–12. doi:10.1007/698

55. Das P, Khan S, Chaudhary AK, AbdulQuadir M, Thaher MI, Al-Jabri H. Potential Applications of Algae-Based Bio-fertilizer. 2019; 41–65. doi:10.1007/978-3-030-18933-4_3

56. Saadaoui I, Sedky R, Rasheed R, Bounnit T, Almahmoud A, Elshekh A, et al. Assessment of the algae-based biofertilizer influence on date palm (Phoenix dactylifera L.) cultivation. J Appl Phycol. 2019;31: 457–463. doi:10.1007/s10811-018-1539-6

57. Pröschold T, Marin B, Schlösser UG, Melkonian M. Molecular phylogeny and taxonomic revision of Chlamydomonas (Chlorophyta). I. Emendation of Chlamydomonas ehrenberg and Chloromonas gobi, and description of Oogamochlamys gen. nov. and Lobochlamys gen. nov. Protist. 2001;152: 265–300. doi:10.1078/1434-4610-00068

58. Nakada T, Tomita M, Wu JT, Nozaki H. Taxonomic revision of Chlamydomonas subg. Amphichloris (Volvocales, Chlorophyceae), with resurrection of the genus Dangeardinia and descriptions of Ixipapillifera gen. nov. and Rhysamphichloris gen. nov. J Phycol. 2016;52: 283–304. doi:10.1111/jpy.12397

59. Popescu CE, Lee RW. Mitochondrial genome sequence evolution in chlamydomonas. Genetics. 2007;175: 819–826. doi:10.1534/genetics.106.063156

60. Pérez-Martínez X, Vázquez-Acevedo M, Tolkunova E, Funes S, Claros MG, Davidson E, et al. Unusual location of a mitochondrial gene. Subunit III of cytochrome c oxidase is encoded in the nucleus of Chlamydomonad algae. J Biol Chem. 2000;275: 30144–30152. doi:10.1074/jbc.M003940200

61. Kindle KL. High-frequency nuclear transformation of Chlamydomonas

reinhardtii. PNAS. 1990;87: 1228–1232. doi:10.1016/S0076-6879(98)97005-7

62. Lorenz C, Lünse CE, Mörl M. Trna modifications: Impact on structure and thermal adaptation. Biomolecules. 2017;7. doi:10.3390/biom7020035

63. Hamaji T, Kawai-Toyooka H, Toyoda A, Minakuchi Y, Suzuki M, Fujiyama A, et al. Multiple independent changes in mitochondrial genome conformation in chlamydomonadalean algae. Genome Biol Evol. 2017;9: 993–999. doi:10.1093/gbe/evx060

64. Šmarda P, Bureš P, Horová L, Leitch IJ, Mucina L, Pacini E, et al. Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc Natl Acad Sci U S A. 2014;111: E4096–E4102. doi:10.1073/pnas.1321152111

65. Li Z, De La Torre AR, Sterck L, Cánovas FM, Avila C, Merino I, et al. Single-copy genes as molecularmarkers for phylogenomic studies in seed plants. Genome Biol Evol. 2017;9: 1130–1147. doi:10.1093/gbe/evx070

66. Gallaher SD, Fitz-Gibbon ST, Strenkert D, Purvine SO, Pellegrini M, Merchant SS. High-throughput sequencing of the chloroplast and mitochondrion of Chlamydomonas reinhardtii to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. Plant J. 2018;93(3):545-565. doi: 10.1111/tpj.13788

CAPÍTULO 2. Whole genome sequencing of the halophilic green alga *Dunaliella viridis* unveils a unique set of genes for biotechnological exploitation*

*Manuscrito a ser submetido e publicado no periódico Algal Research como Pascoal PV, Formighieri EF, Cançado LJ, Miranda CHB, Calsing LCG, Brasil BSAF (2024).

Whole-genome sequencing of the halophilic green alga *Dunaliella viridis* unveils a set of genes and transcription factors for biotechnological exploitation

Patrícia Verdugo Pascoal [1,2], Eduardo Fernandes Formighieri [1], Cesar Heraclides Behling Miranda [1], Letícia Jungmann Cançado [1], Lorena Costa Garcia Calsing [1], Bruno S. A. F. Brasil [1, 2]

[1] Embrapa Agroenergia, Brasília, Distrito Federal, Brazil
[2] Universidade Federal da Bahia, Salvador, Bahia, Brazil

Abstract

The study investigated the microalgae *Dunaliella viridis* EMBRAPA| LBA#S001, isolated from salt marsh blooms in Brazil, highlighting its morphological and genomic characteristics. The strain was identified by molecular markers, such as the rbcL and ITS2 genes, and demonstrated high tolerance to high salinities, with consistent growth in NaCl concentrations of up to 4.8 M. The production of β-carotene was evaluated, resulting in 18.7 mg/L, comparable to the levels found in other species of the genus. The mitochondrial DNA revealed a genome of 46.2 Kbp, while the chloroplast genome reached 197.1 Kb, both with genes essential for photosynthesis. The analysis of the nuclear genome, with 176.9 Mb, identified 29,594 protein-coding genes, indicating significant potential for biotechnology. The study analyzed the resistance genes in the strain, identifying mechanisms such as antibiotic target switching and antibiotic efflux. The presence of genes such as vanY and adeF suggests adaptive diversity in contaminated environments. Analysis of transcription factors highlights the regulation of genes essential for photosynthesis and stress responses. Manipulating these factors can increase the production of valuable compounds such as carotenoids. The development of synthetic promoters is crucial to optimize microalgae as biotechnological platforms, compared to more traditional organisms.

Keywords: Carotenogenesis; Halotolerance; ARGs; Transcription Factors.

Introduction

Microalgae comprise a diverse group of microscopic photosynthetic organisms that inhabit marine and freshwater environments. Due to their minimal nutritional requirements and high growth rates, several species can efficiently convert sunlight and inorganic carbon sources into carbohydrates, proteins and lipids-rich biomass useful for the industry (Brasil et al., 2017; Molino et al., 2018; Gharajeh et al., 2020). Indeed, microalgae are used in various applications, such as the production of human food, animal feed, cosmetics and wastewater treatment (Santos et al., 2021, Dismukes et al., 2008; Debowski et al., 2020).

Microalgae species from the *Dunaliella* genus have been used for decades to produce pigments such as β-carotene (Burak et al., 2019; Harvey & Ben-Amotz, 2020; Capa-Robles et al., 2021). These extremophile algae can tolerate stressful environments such as high light and salinity and, under these conditions, accumulate high value-added products such as carotenoids, lipids and enzymes industrial-interest (Silva et al., 2020; Roy et al., 2021). Additionally, it has been proposed that the enzymes and metabolites from *Dunaliella* species could be potentially applied as bio-inputs in agriculture, through the improvement of economic crops in saline soil, with halotolerant feature also due to the biostimulant and biofertilizer properties (Ronga et al., 2019, Colla and Rouphael, 2020). Currently, Nature Beta Technologies (NBT) in Israel produces around 25 tons per year of ash-free dry weight (AFDW) *D. salina* var. *bardawil* biomass from 10 hectares of raceways, making it one of the largest reliable sources of *D. salina* powder globally. However, this biomass is not cultivated for high 9-cis β-carotene content. Monzon Biotech S.R.L. (MB) in Spain produces approximately 400 kg per year (AFDW) of *D. salina* for established markets. But current *D. salina* biomass productivity ranges between 0.75 - 3.0 g/m2/day (AFDW), depending on seasonality and other factors. (Harvey & Ben-Amotz, 2020; Pourkarimi et al., 2020). The main cultivated species is *Dunaliella salina*.

Although recent microalgae genomics reports reveal a growing body of novel gene sequences that could be used for biotechnological applications, little is known about *Dunaliella* species genomics. Currently, there is only a draft nuclear genome sequence of the prototypic species of the genus, *D. salina*, available (Polle et al., 2017). Additionally, five mitochondrial and three chloroplast genomes have been reported for species within the *Dunaliella* genus (Smith et al., 2010; Del Vasto et al., 2015; Polle et al., 2017). A

recent comparative genomic, biochemical, and physiological analysis with *D. bioculata* and *D. quartiolecta* was performed. qPCR analysis unveils 382 and 85 novel genes related to salt stress response in *D. bioculata* and *D. quartolecta*, respectively (Gao et al., 2022).

Dunaliella organellar genomes are larger than other species from the Reinhardtinia clade, characterized by high noncoding content (> 40%) and the presence of introns. These are unique traits of Dunaliella species and have been rarely found in other species (Smith et al., 2010; Del Vasto et al., 2015). These genomic features indicate distinct evolutionary patterns for the genus; therefore, deeper knowledge of Dunaliella species genes might provide access to a rich reservoir of unexploited biodiversity.

This microalga was chosen for sequencing due to its prominence in growing and developing in saline media and presenting the production of carotenoid pigments. In addition, the absence of a cell wall makes it an interesting candidate for the accumulation of compounds of industrial interest, since it is possible to dispense with downstream steps in the context of production and recovery of these compounds.

Few recent studies deal with the growth of this species in media that can be reused after flocculation and biomass harvesting (Lai et al., 2023). It has already been reported that a species, isolated from salt pans, can tolerate severe environmental conditions and has an amino acid profile of interest to the human food industry (Bombo et al., 2023). In addition, it is also reported as an accumulator of β-carotene, a precursor of the antioxidant astaxanthin. Its metabolic pathway has been studied and the key enzymes CRTR-B and BKT from *Haematococcus pluvialis* were expressed in *D. viridis*, which was able to perform as an astaxanthin production biofactory (Lin et al., 2019).

So, this study aims to perform the whole genomic sequencing of a *Dunaliella viridis* strain isolated from naturally occurring algal blooms in solar salt pans used for sea salt extraction in Brazil. Genomic sequencing was performed using Illumina and Nanopore platforms coupled with the assembly and functional annotation of its genetic content and the phylogenetic analysis. For the first time, the nuclear genome sequence of *D. viridis* is revealed, as well as the complete mitochondrion (mtDNA) and chloroplast (ptDNA) genomes. Furthermore, putative genes related to the carotenogenesis pathway, as well as the production of other metabolites and enzymes of industrial interest, were identified.

Material and method

Water samples were collected from naturally occurring algal blooms in solar salt pans used for sea salt extraction at Mossoró/RN, Brazil. The collection was authorized and did not involve endangered or protected species. Isolation steps were carried out as described in Hadi et al. (2016), using modified J/L medium (NaCl: 34.3 gL^{-1}; $MgCl_2$: 1.51 gL^{-1}; $MgSO_4$: 0.52 gL^{-1}; KCl: 0.206 gL^{-1}; $CaCl_2$: 0.203 gL^{-1}; KNO_3: 1.0 gL^{-1}; $NaHCO_3$: 0.044 gL^{-1}; KH_2PO_4: 0.045 gL^{-1}) (Sathasivam & Juntawong, 2013). Axenic microalgae cultures of *Dunaliella viridis* EMBRAPA| LBA#S001 strain were deposited in the collection of Microorganisms and Microalgae of Agroenergy and Biorefineries interest (CMMAB) of EMBRAPA Agroenergy - (Brasília/DF, Brazil).

Strain identification

Microscopic morphologic identification at the genus level was performed according to (Bellinger & Sigee, 2015). Further identification to species levels was accomplished through molecular methods (Hadi et al., 2016). Molecular identification was performed based on the analysis of the chloroplast marker sequence, Ribulose Bisphosphate Carboxylase Large subunit gene (rbcL), and the nuclear markers sequences, Internal Transcribed Spacer 2 of the nuclear rDNA (nuITS2), as described in Hadi et al., 2016. The rbcL gene and ITS2 sequences from *Dunaliella viridis* EMBRAPA| LBA#S001 strain were included in the dataset together with reference sequences retrieved from NCBI GenBank. Sequences were aligned by ClustalW, by Neighbor-Joining method with 1000 bootstrap value. Tree reconstruction was performed using MEGA v. 11 tool (Tamura et al., 2021). Final trees were edited with FigTree tool (Rambaut, 2006).

Microalgae cultivation and carotenoid content determination

Dunaliella viridis EMBRAPA| LBA#S001 was cultured in 250 mL of modified J/L medium in 500 mL Erlenmeyer flasks with different NaCl concentrations (0.59 M, 1.2 M, 2.4 M, and 4.8 M). The flasks were shaken manually once a day, at 25°C±1°C, light intensity of 50 $\mu Em^{-2}s^{-1}$in a 12/12 h light/dark regimen for 20 days. The growth was monitored by OD analysis using a spectrophotometer at 400 and 750 range as well as Neubauer's chamber.

For biomass dry weight determination, 50 mL samples of the algal culture were collected, and centrifuged for 10 min at 10700RCF and the supernatant was discarded. The pellet was washed through three cycles of resuspension in distilled water followed by 10 min centrifugation at 10700 RCF. The washed pellet was dried overnight using a

dry oven at 105°C and weighed, as described by Santana et al, 2017. Carotenoids were determined by UV detection in 450 ηm, as described by Ribeiro et al., 2020. Statistical analysis was performed by ANOVA with 5% significance and Tukey test with ASSISTAT software v. 7.6

Genomic DNA preparation

The genomic DNA used for Illumina sequencing – short DNA fragments – was obtained using the CTAB 2% (Cethylthyrmethylammonium Bromide) method (Doyle & Doyle, 1987) with modifications. In brief, the method was modified by using two antioxidants β-Mercaptoethanol (140 mM) and Dithiothreitol (5 mM), three 3 mm metal beads were added to each microtube and mixed in the bead beater device for 30 seconds repeated three times. The Chloroform Isoamyl cleaning step was repeated three times. The samples were quantified and analyzed for purity in Nanodrop™ equipment and by the fluorometric method of quantification with Qubit™. The genomic DNA was fragmented with NEBNext® DNA Library Prep Kit construction.

For Nanopore sequencing – long DNA fragments – the DNA extraction was performed according to a modified PCI (Phenol:Chloroform:Isoamyl) method, shared protocol from PacBio, considering smooth pipetting, five times Chloroform:Isoamyl cleaning step, avoiding sudden movements and temperature variations, once the excess of fragmentation must be avoided. The samples were quantified and analyzed for purity in a Nanodrop™ device and by the fluorometric method of quantification with Qubit™ equipment. Each sample was transferred to tubes (GenTegra™), following the instructions for use. The genomic library was prepared with high molecular weight DNA with 16X of genome coverage. Both sequencing protocols were performed by GenOne Company (Rio de Janeiro, Brazil).

Genomic sequencing

Genomic sequencing was performed using both Illumina (short DNA fragments) and Nanopore (long DNA fragments) platforms to facilitate genomic sequence assembly. According to the manufacturer's protocols, the short reads genomic library was sequenced on the Illumina HiSeq 2000 platform with 100X of genome coverage. Oxford Nanopore sequencing was performed on PromethION. The raw reads quality control was conducted by NanoPlot software (De Coster et al., 2018) to remove adapter contamination and low-quality reads.

Genomic analysis

FastQC tool (Andrews, 2010) was used to evaluate the quality of the sequence data set. The adapters were trimmed using the Trimmomatic software (Bolger, 2010). Nanopore data were concatenated in fastq format and corrected with a hybrid approach, using Illumina data. This step was performed using the LoRDEC tool (Salmela and Rivals, 2014) with three k-mers iterations, k19, k31 and k41. The hybrid correction generated a 6.4 Gb output file, used in the next step. The LoRDEC output file was trimmed and assembled with Canu, considering the error rate to Nanopore data, 0.144. The genome size parameter was 360 Mb, considering the *Dunaliella salina* genome size as a reference to the assembly. Assemblies were performed using the Flye and SMARTdenovo tools, however, the hybrid approach with Lordec and Canu was more promising in the assembly and genome completeness metrics (Supplementary Table 1).

The genome annotation step was performed using a combination of tools, including Funannotate (Palmer & Stajich, 2020), InterProScan (Jones et al., 2014) and MAKER (Cantarel et al., 2008). The process involved running Funannotate to identify genes, coding regions and functional elements. For detailed functional annotation, InterProScan was used, which allows the identification of protein domains, and the prediction of functions based on protein sequences. In addition, the MAKER tool was used to perform additional and refined annotation.

The completeness of the genome was assessed by BUSCO (Manni et al., 2021), using the database chlorophyta_odb10.

The chloroplast genome was assembled *de novo* with GetOrnagelle tool (Jin et al., 2020), using a low coverage approach with 10% of the Illumina data pool previously interleaved and k-mers (21, 45, 65, 85, 105) parameter. Mitochondrial and chloroplastidial DNA are differentiated from the nuclear genome by their representativeness in the genetic material. They are considerably smaller and usually have a circular architecture in microalgae, while the nuclear genome is linear (Zhao et al., 2022). The annotation and map generation of the chloroplast was performed by GeSeq and OGDRAW tools from Chlorobox set of web tools (https://chlorobox.mpimp-golm.mpg.de/index.html), using *D. salina* chloroplast genome as a reference.

Mitochondrial genome was obtained with MitoBim software, with *D. viridis* mitochondrial as a reference and low coverage approach, with 10% of the Illumina data read pool. Annotation step and map generation were performed by GeSeq and OGDAW

tools from Chlorobox set of web tools (https://chlorobox.mpimp-golm.mpg.de/index.html), using *D. salina* and *D. viridis* mitochondrial genome as a reference.

Resistance Genes Analysis

The identification of the resistance genes was performed RGI (Resistance Gene Identifier) (Alcock et al., 2023) tools. RGI allows for the identification of resistance genes based on homologous sequences, also providing predictions about their functions. The genomic sequences were input and analyzed, allowing for the detection of resistance genes and the assessment of the reliability of the predictions.

Metabolic Pathway

The KEGG Mapper Reconstruction was obtained by BlastKOALA (Taxonomy group: Dunaliella and KEGG database: genus_eukaryotes) using the proteins fasta file generated by annotation step and then Mapper Reconstruction Result was performed. A total of 29,594 protein sequences were used as input and 5683 (19.2%) were mapped by KEGG (Kanehisa & Goto, 2000).

Promoters and Transcription Factors analysis

The search for promoters was conducted using the TBTools (Chen et al., 2023) software, utilizing the assembled genome sequence. The option to identify promoters was configured with a limit of 1000 bases upstream of industrially relevant genes. This setup allows the tool to analyze the region preceding the genes, identifying sequences that may act as promoters. With the list of promoters, the online tool PlantPAN 4.0 (Chow et al., 2024) was used to search for transcription factors associated with these promoters. The option used to search was Multiple promoter analysis with *Chlamydomonas reinhardtii* as database species. The tool allows for the analysis of the sequences in search of known binding motifs for transcription factors, providing information about which factors may be regulating the expression of the genes associated with the identified promoters.

Results and discussion

Dunaliella viridis isolation and identification

D. viridis EMBRAPA | LBA#S001 strain was isolated from naturally occurring algal blooms in solar salt pans used for sea salt extraction in Brazil (Figure 1 - Panel A). The collected environmental samples were submitted to an enrichment step in liquid modified J/L medium followed by two subsequent rounds of subculturing on agar plates supplemented with antibiotics. Individualized macroscopic colonies on agar plates were collected and inoculated into liquid modified J/L medium to derive axenic cultures (Figure 1 – panels B and C). Microscopic inspection of the isolated strain revealed typical *Dunaliella* species morphotypes, characterized by a single large bowl-shaped chloroplast that fills the base of the cell, presence of pyrenoids, two flagella of equal length and the lack of cell wall (Figure 1 – panels B and C).

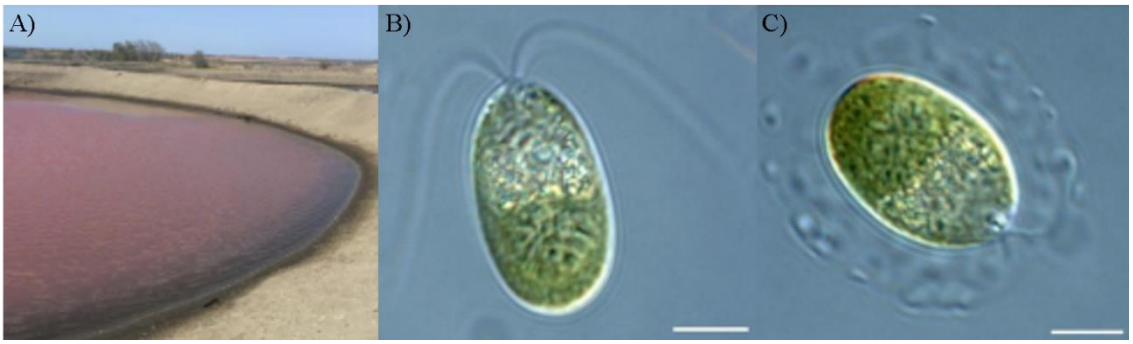


Figure 1. Representative images of water sample collection points at Mossoró/RN, Brazil and microalgae micrographs of *Dunaliella viridis* EMBRAPA| LBA#S001 isolated at the Algae Biotechnology Laboratory of Embrapa Agroenergy from the collected samples: Collection point – Algal bloom in solar salt pan (A). Vegetative cells of the microalgae *Dunaliella viridis* EMBRAPA| LBA#S001 (B and C). Micrographs were obtained by differential interference contrast (DIC) microscopy. Scale bar: 5 µm.

Dunaliella viridis was identified based on rbcL (95% of identity) and ITS2 (96.3% of identity) markers (SI.1 e SI.2). The rbcL gene codes for the large subunit of the enzyme RuBisCo, responsible for CO_2 fixation catalyzed (Patel et al., 2018). This marker is indicated by the Barcode of Life consortium as a universal barcode for chlorophytes (Puillandre et al., 2012). The ITS2 region of nuclear ribosomal DNA is a complementary marker also indicated to identify algal species (Tran et al., 2013; Hadi et al., 2016).

D. viridis EMBRAPA| LBA#S001 strain halotolerance was evaluated through cultivation in different NaCl concentrations, ranging from 2 to 8 times of the medium sea salinity (Figure 2). No statistical difference in algal growth was observed under these conditions at the end of cultivation. This finding is in line with previous reports on *D.*

viridis tolerance to high salt concentrations (Hadi et al., 2008; Tran et al., 2013), except for the fact that no decline in algal growth at 4.8 M of NaCl was observed (Figure 2). This extremophile characteristic is shared by other Dunaliella species, especially *D. salina*, that can even tolerate salinities close to NaCl saturation (~5 M NaCl) and *D. bardawil* (~5.5 M NaCl) (Borowitzka & Silva, 2007; Liang & Jiang, 2017).

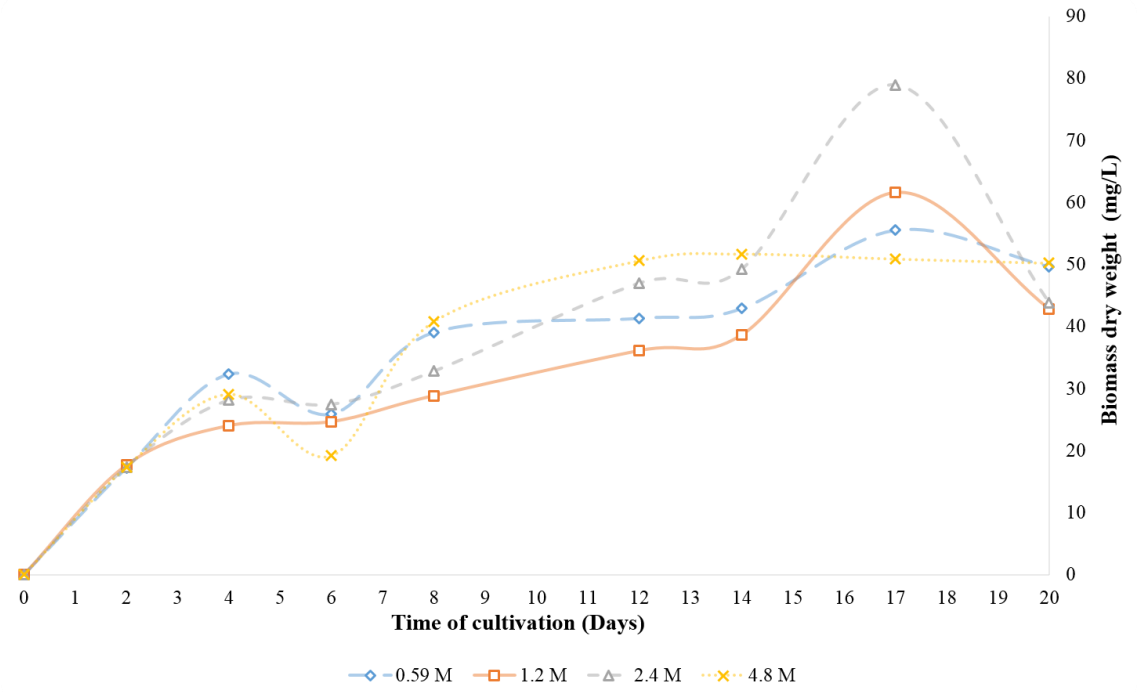


Figure 2. Biomass accumulation of *D. viridis* EMBRAPA| LBA#S001 in different NaCl concentrations for 20 days.

Under certain cultivation conditions, especially under salt stress or high light *Dunaliella* species accumulate carotenoid pigments (Hadi et al., 2008; Preetha et al., 2012). Thus, carotenoid concentration in *D. viridis* EMBRAPA| LBA#S001 strain was measured after 32 days of cultivation in a modified J/L medium at 0.59 M of NaCl. The results obtained revealed the presence of β-carotene at 18.7 mg/L. Comparable levels of β-carotene content have been reported for *D. salina* species that can achieve up to 25 mg/L under hypersalinity conditions in open photobioreactors (Wolf et al., 2020). Future experiments targeting carotenoid production optimization in *D. viridis* EMBRAPA| LBA#S001 will be required to fully reveal this species potential.

Organellar genomes

D. viridis EMBRAPA| LBA#S001 mitochondrial DNA sequencing revealed a 46.2 Kbp long genome with a putative circular architecture. Seven single-copy mitochondrial protein genes were found, such as cob, cox1, nad1, nad2, nad4, nad5, and nad6, as well as two rRNAs fragmented and 22 tRNAs (Figure 3).

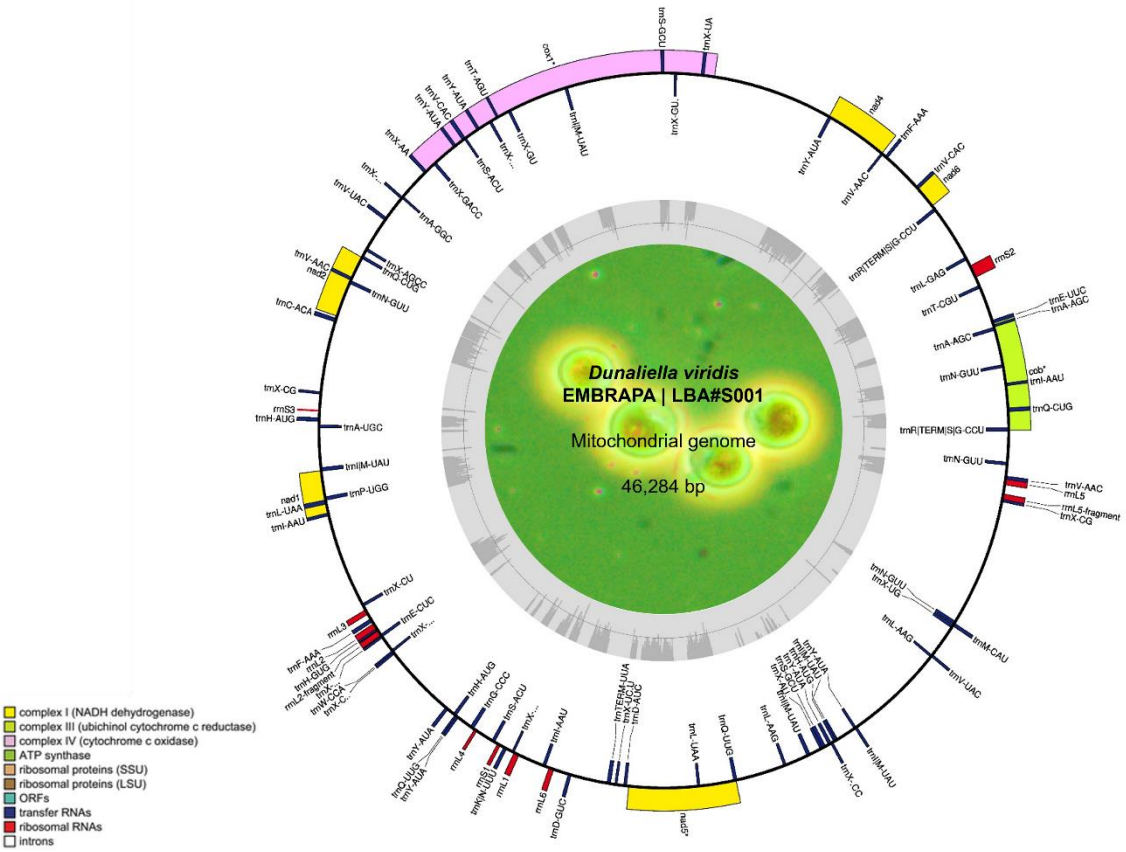


Figure 3. Representative mitochondrial complete genome of *Dunaliella viridis* EMBRAPA| LBA#S001. Genes oriented in the clockwise direction are positioned on the inner part of the genome map, while genes in the counterclockwise orientation are located on the outer part of the map. The annotated genes are colored according to the functional categories presented in the legend. The inner circle of the visualization depicts the GC content, with the midpoint line indicating the 50% GC content level.

Species from the Chlamydomonadales order, within which Dunaliellales clade takes place, display a variety of mitochondrial DNA architectures and sizes (Del Vasto et al, 2015). Indeed, the mtDNA size and content in Chlamydomonas species are generally reduced compared to Dunaliella species. These large organelle genomes may contain over 50% of noncoding DNA (Smith et al., 2010; Del Vasto et al., 2015). The *D. viridis* EMBRAPA| LBA#S001 mitochondrial genome presented here displays a total of eight

introns distributed across cob, cox1, nad1 and nad2 genes (Figure 3). This finding was also observed in *D. salina* (CCAP 19-18 and CONC-001) and *D. viridis* (CONC-002) mitochondrial genomes (Del Vasto et al., 2015).

The image shows the circular genome of The *D. viridis* EMBRAPA| LBA#S001 chloroplast genome, with a total length of 197.1 Kb (Figure 4). The chloroplast is represented as a circle, with the outer part showing genes and the inner part containing additional information about the regions of the genome, such as LSC (Large Single Copy) and SSC (Small Single Copy). Most of the genes related to photosynthesis are present, including components of the photosynthetic system I and II, such as the genes for psaA and psbA, crucial for light capture and energy conversion. Genes such as petA and petB are highlighted, which are essential for electron transport during photosynthesis. The atpB gene is present, responsible for the synthesis of ATP, essential for cellular energy. rbcL gene (large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase) is also included, essential for carbon fixation in photosynthesis. The rpoB gene is indicated, necessary for the transcription of DNA into RNA. Several tRNA (transfer RNA) genes are shown, such as trnA-UGC, trnI-GAU, and trnL-UAA, which are vital for the translation of messenger RNA during protein synthesis. Ribosome genes are also present, including rpl and rps, which code for ribosomal proteins. The table contains hypothetical genes (represented by open squares) that do not yet have a defined function but are important for future research. When comparing the chloroplast genome of *D. viridis* EMBRAPA| LBA#S001 chloroplast genome with *Dunaliella salina*, some important observations can be made.

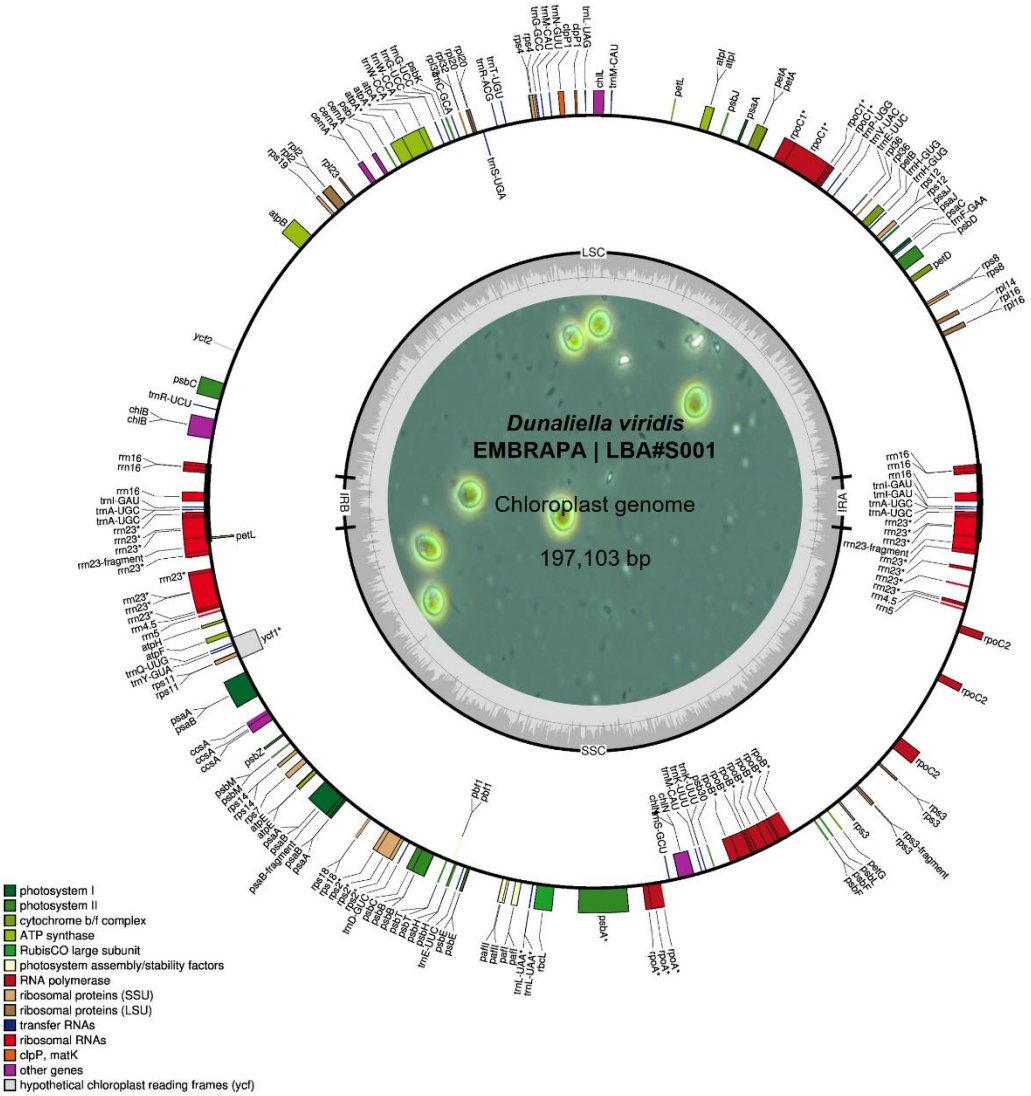


Figure 4. Representative chloroplast complete genome of *Dunaliella viridis* EMBRAPA| LBA#S001. Genes oriented in the clockwise direction are positioned on the inner part of the genome map, while genes in the counterclockwise orientation are located on the outer part of the map. The annotated genes are colored according to the functional categories presented in the legend. The inner circle of the visualization depicts the GC content, with the midpoint line indicating the 50% GC content level.

There are two chloroplast genome sequences available for Dunaliella species and both present similar characteristics, such as circular architecture and large genome size. Indeed, it has been reported that the ptDNA can be equally or more inflated with intronic sequences than mtDNA in Dunaliella species (Smith et al., 2010; Del Vasto et al., 2015).

The chloroplast genome of *Dunaliella salina* is 269 kb, is the largest complete plastid DNA (ptDNA) sequence currently deposited in GenBank (Smith et al., 2010). Both species share many genes critical for photosynthesis and metabolism. However, *Dunaliella salina* may contain additional genes that can confer resistance to salt stress, an essential factor for its survival in hypersaline environments.

The secondary metabolites of industrial interest, such as carotenoids, are produced inside the chloroplast of Dunaliella genus strains, but the related enzymes are encoded by the nuclear genetic content and transported to the plastid in polypeptides form (Henriquez et al., 2016).

In addition, unicellular photosynthetic microalgae are gaining recognition as significant hosts for sustainable industrial biotechnology because of their capability to effectively convert CO2 into biomass and valuable bioproducts. The microalga *Chlamydomonas reinhardtii* has been utilized as a framework to produce Rhamnolipid (RL) biosurfactants through the engineering of its chloroplast genome, facilitating the stable expression of the gene that encodes the acyltransferase RhlA from *P. aeruginosa*, for instance. Therefore, engineering the chloroplasts of microalgae strains is expected to aid in creating a clean, safe, and cost-effective alternative platform for the sustainable production of RLs (Miró-Vinyals et al., 2023).

There are limitations in the existing methodologies that lead to relatively low expression yields in chloroplasts for the desired products in microalgae. Therefore, it is essential to enhance the understanding of the transcriptional and post-transcriptional control of gene expression in the chloroplast, as well as the role of RNases in algal plastids. To date, *Chlamydomonas* remains the only alga routinely used in transplastomic technology for metabolic pathway engineering and molecular agriculture (Esquível et al., 2024).

Nuclear genome

The organellar genomic features found in *D. viridis* are indicative of distinct evolutionary patterns. This hypothesis becomes particularly sound considering the phenotypic characteristics of this species, such as halotolerance and carotenoid production, and the current unavailability of *D. viridis* species nuclear genomic sequences.

The nuclear genomic sequence of *D. viridis* EMBRAPA| LBA#S001 was assembled using a hybrid approach, coupling Illumina and Oxford Nanopore platforms. The Illumina sequencer generated 74.0 Gb of raw data, 246,683,582 raw reads with an effectiveness of 99.74%, and Nanopore sequencing results in 11.7 Gb of raw data, 559,267 reads and effectiveness of 94.3% (>Q7). The use of a hybrid approach can increase the quality of the assembly, annotation, and completeness of genomic sequences (Hanschen & Starkenburg, 2020). Assessment of the assembled nuclDNA sequence with BUSCO showed a genome with 176.9 Mb a total of 29,594 protein-coding genes with 93.6% completeness (SRA Submission: SUB11183448) indicating a set of 1,519 genes inside of Chlorophyta database, with 1,421 genes considered complete and single-copy and 79 missing genes. The BUSCO analysis assesses genome completeness based on the presence of conserved single-copy genes. A high percentage of complete BUSCO single-copy genes indicates the assembly captures the core genome, suggesting high quality. The single-copy BUSCO genes also provide insights into the organism's evolution and genome characteristics (Simão et al., 2015).

The search for enzymes in this *D. viridis* strain can be innovative due to the metabolic diversity of the genus, which can reveal enzymes of industrial interest with different applications and properties compared to other model organisms. In addition, the extreme conditions in which this species can develop provide them with unique characteristics. Adaptations to the environment also involve enzymatic adaptation, specialized metabolism, and, consequently, secretion of compounds of interest (Barera & Forlani, 2023; Guermazi et al., 2023). These possibilities could generate the potential to produce pigments, biofuels, and other bioproducts. Studying their enzymes could reveal new biocatalysts with useful properties for these applications. For example, in microalgae, acetyl-CoA carboxylase (ACCase) is the rate-limiting enzyme in lipid biosynthesis, converting acetyl-CoA to malonyl-CoA. The β-carboxyl transferase (β-CT) subunit, encoded by the plastidic accD gene, is crucial for the heteromeric ACCase in

chloroplasts. Lower accD expression has been shown to decrease ACCase activity in tobacco plants. Therefore, understanding the accD expression pattern in the target microalgae under different nutrient conditions could aid genetic engineering strategies to maximize lipid production. (Riyazat Khadim et al., 2023).

The Dunaliella genus comprises several genetically distinct species and strains, which increases the likelihood of finding enzymes with diverse characteristics (Liu et al., 2024; Stravidou et al., 2024). Despite the biotechnological interest, knowledge about enzymatic repertoire in Dunaliella is still limited, with more studies covering the *D. salina* species, opening opportunities for discoveries about other potential species (Souza Celente et al., 2023).

The annotation step of the 176.9 Mb *D. viridis* nuclear genome resulted in 29,594 protein-coding genes. On the other hand, the 343.7 Mbp draft nuclear genome of *D. salina* strain CCAP19/18 presented 18,800 genes (Polle et al., 2017), while *C. reinhardtii* 120 Mbp nuclear genome presents 3,000 protein-coding genes (Grossman et al., 2003).

Functional annotation was performed using KEGG and returned 5,683 entries (19.2% from annotation) involved in several metabolic pathways (Figure 5). It was observed that processes related to genetic information and carbohydrate metabolism are predominant.

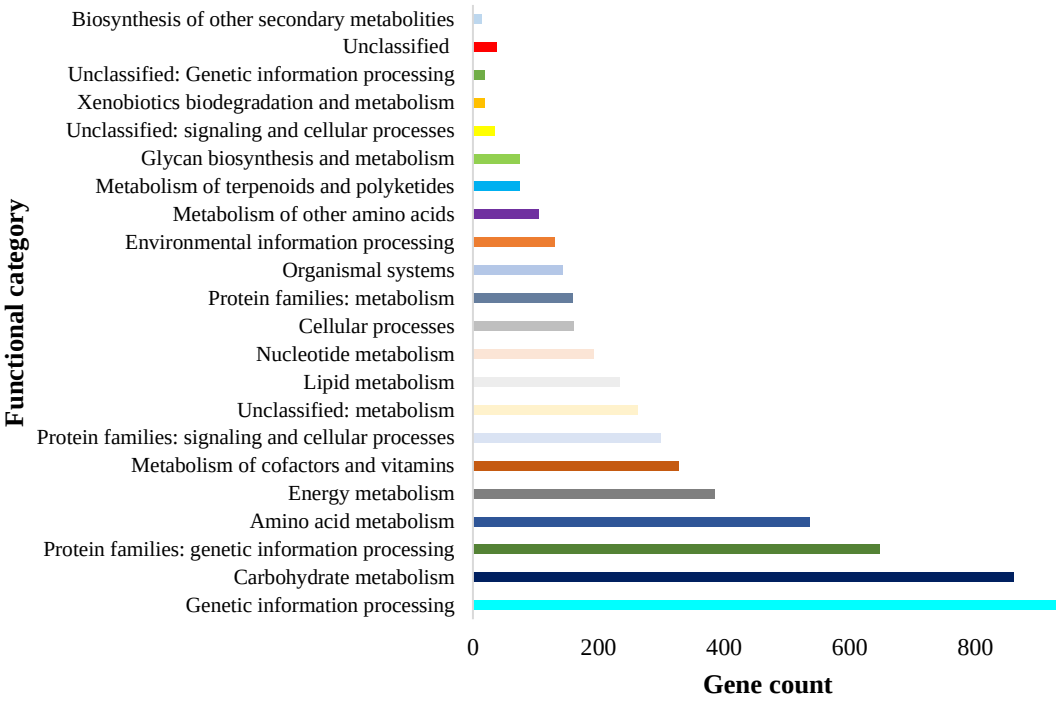


Figure 5. Overview of gene representation in KEGG functional categories of *Dunaliella viridis* EMBRAPA| LBA#S001.

The figure 6 shows the distribution of the number of genes associated with different groups of CAZy enzymes (Carbohydrate Active Enzymes). These data indicate a significant predominance of glycosyl transferases among the groups studied, suggesting an important role of these enzymes in the biological function of *Dunaliella viridis* EMBRAPA| LBA#S001 (Polle et al., 2020; Dehghani et al., 2024).

In addition, enhancing biomass production is crucial for applications in nutrition, pharmaceuticals, and biofuels, and knowledge of CAZy enzymes can inform strategies to achieve this. Understanding the role of these enzymes also provides insights into how Dunaliella adapts to various environmental conditions, such as salinity and light availability, which contributes to its survival in extreme habitats (van der Loos et al., 2024). Furthermore, identifying specific CAZy enzymes opens up possibilities for genetic engineering, allowing for the optimization of metabolic pathways and improving the efficiency of desired product synthesis.

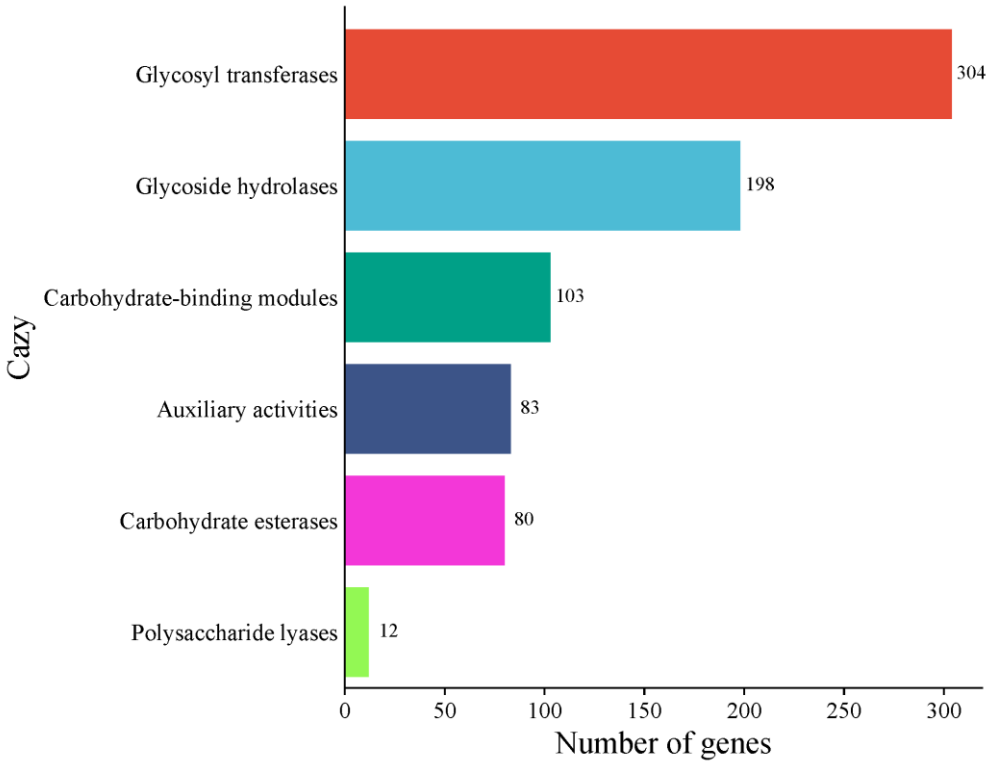


Figure 6. Distribution of CAZy genes across enzyme classes in *Dunaliella viridis* EMBRAPA| LBA#S001.

Furthermore, the nuclear genome of *D. viridis* codes for enzymes and proteins related to carotenogenesis, salt-tolerance, and even potential industrial enzymes. Genes coding for IPI (Isopentenyl diphosphate isomerase), violaxanthin de-epoxidase, FPPS

(Farnesyl), and CYB5 (Cytochrome), which are related to the carotenogenesis pathway in cyanobacteria (Pagels et al., 2021) are present in *D. viridis* nuclear genome. Genes such as SKIP, CYP, and Spermidine synthase, which are linked to halotolerance in Arabidopsis (Feng et al., 2015), were also found in the *D. viridis* genome. Furthermore, given the halophilic characteristic of this species. Indeed, the production of a variety of enzymes of industrial interest (e.g., amylase, galactosidase, phytase, protease, and peroxidase) has been reported in halophilic bacteria strains (Shirazian et al., 2016; Huo et al., 2017; Mokashe et al., 2018; Ruginescu et al., 2020). These enzymes can be applied in bioenergy production, food processing, and biodegradation of organic pollutants (Ruginescu et al., 2020). Furthermore, genes coding for proteins reported to have immunological proprieties, such as Glutamine amidotransferase, Glutaminyl-peptide cyclotransferase, and Asparaginase (Shirazian et al., 2016), have also been identified.

A detailed mapping of the carotenoid biosynthesis pathways in *Dunaliella viridis* EMBRAPA| LBA#S001 and *Dunaliella salina* (CCAP 19/18) can be observed in Figure 7.

The metabolic pathways of biosynthesis vary significantly between the two species *Dunaliella viridis* EMBRAPA| LBA#S001 and *Dunaliella salina* (CCAP 19/18), emphasizes the biosynthesis of certain carotenoids, such as β-carotene, featuring several highlighted enzymes and interactions. The presence of specific pathways not observed in *Dunaliella viridis* suggests evolutionary adaptations to different environments or metabolic needs. In contrast, *Dunaliella salina* shows a stronger focus on the biosynthesis of zeaxanthin and neurosporaxanthin, indicating a potential adaptation to stress conditions or protection against UV radiation.

D. viridis EMBRAPA| LBA#S001 possesses a greater number of enzymes involved in the conversion of farnesyl pyrophosphate into carotenoids, reflecting a robust capacity to produce a variety of compounds (Huang et al., 2021; Kang et al., 2022). Meanwhile, *D. salina* displays a broader diversity in the pathways for the degradation and recycling of carotenoids, which may represent an adaptation to specific environments where nutrient availability fluctuates (Souza Celente et al., 2023; Huang et al., 2024). Additionally, in *D. viridis* EMBRAPA| LBA#S001, there are biosynthetic pathways for other secondary metabolites, suggesting a diverse secondary production that could have implications for stress adaptation. Secondary metabolites are valuable products with significant potential for global biotrends. Enhancing their production can be achieved by utilizing nutritional, environmental, and physiological stress as key stimulating strategies.

Over the past decade, there has been a growing interest in algal stress biology and omics technologies, which have expanded our understanding of this field (Kolackova et al., 2023). On the other hand, *D. salina* appears to have a greater integration of lipid metabolism pathways, which may be related to its ability to survive in extreme saline environments (He et al., 2020; Zhang et al., 2023).

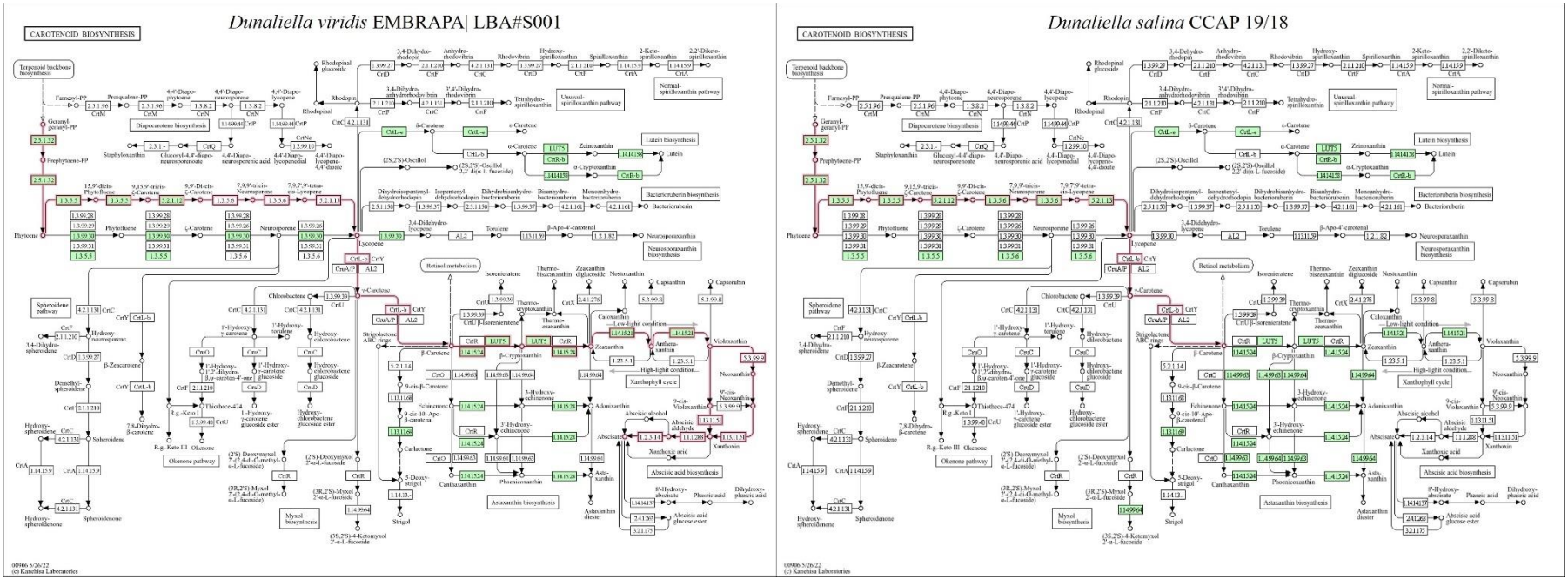


Figure 7. Comparative analysis of carotenoid biosynthesis between *Dunaliella viridis* EMBRAPA| LBA#S001 and *Dunaliella salina* (CCAP 19/18). Highlighted enzymes and products (Green) were annotated by KEGG.

Current studies primarily concentrate on the initial and final stages of β-carotene synthesis, involving enzymes like phytoene synthase, phytoene desaturase, and lycopene cyclase. These investigations utilize both gene expression and engineering methods. Additionally, the discussion includes various strategies, such as genetic and metabolic engineering, aimed at enhancing β-carotene production (Félix-Castro et al., 2023).

The gene *crtZ* is responsible for generating both lutein, astaxanthin and zeaxanthin in carotenoids biosynthesis pathway. This enzyme can be blocked or silenced to increase the content of β-carotene, for example. CRISPR/Cas9 system was successfully applied to test this hypothesis in *Dunaliella salina* and the zeaxanthin decreased while carotene yield increased (Hu et al., 2021). Carotenoids metabolism products as phytoene, phytofluene, lutein, zeaxanthin, lycopene, α-carotene and β-carotene have great industrial importance as high added-value products in medical, cosmetic, food and animal feed areas. Furthermore, projects as D-factory (Harvey & Ben-Amotz, 2020) have been developed in Europe to cultivate *D. salina* for industrial-scale production of carotenoids and derivate from this species in a biorefinery model. Similar approaches could be developed for *D. viridis* species in the future.

In addition to carotenoids, photosynthesis is a process of great importance for biotechnology and industrial applications. The main components of the photosynthesis include the antenna complexes and the photosystems, which play a crucial role in capturing light. The antenna proteins capture light and transfer this energy to photosystems I and II, which assist in the photolysis of water and the generation of energy in the form of ATP and NADPH. The Calvin cycle, in turn, is responsible for fixing CO2, using the stored energy to produce sugars. The supplementary figure (SI3) can detail this information of Dunaliella viridis EMBRAPA| LBA#S001

The interconnection between these processes is essential, as the energy generated in the photosystems is vital for the Calvin cycle. There is a direct relationship between light absorption, the production of ATP and NADPH, and the subsequent carbon fixation (Yao et al., 2024). Additionally, the image also illustrates different metabolic pathways, such as the C4 cycle and other carbon fixation pathways, highlighting the adaptability of plants in various environments, such as those using C3 and C4 cycles, and how this influences the efficiency of photosynthesis.

Photosynthesis is vital in areas like agriculture and biofuel production. The intricate and remarkable nature of the photosynthetic process highlights the need to comprehend how these

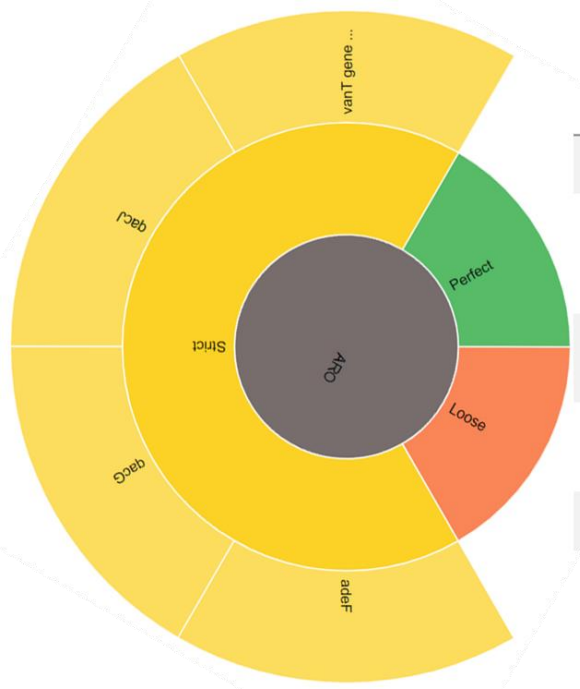
mechanisms sustain life and can be utilized in biotechnology and environmental conservation (Zeng et al., 2021).

Anthropogenic carbon dioxide (CO_2) emissions are one of the most serious global risks facing humanity in the 21st century. The biodegradation of CO_2 is gaining increasing attention as an environmentally friendly technology. Microalgae are receiving global focus due to their rapid growth rates, strong adaptability to extreme environments, and low production costs (Goswami et al., 2022). Additionally, microalgae are vital CO_2 fixers in the ecological environment, converting carbon into biomass through photosynthesis. The efficiency of CO_2 capture using microalgae can vary based on species state, physiology, and environmental conditions such as CO_2 concentrations, pH, light intensity, and dissolved oxygen. Recent developments in trends and strategies to further enhance CO_2 fixation by microalgae include mutagenesis, genetic engineering, and interventions using nanomaterials. Among these, nanomaterials could significantly increase the relative electron transport rates in photosystem II and boost levels of reactive oxygen species in microalgae, thereby improving overall photosynthesis and carotenoid production (Li et al., 2022).

Resistance Genes Analysis

Table 1 identifies the main findings of the analysis performed by RGI, highlighting different mechanisms by *D. viridis* EMBRAPA| LBA#S001 can develop resistance. Among these mechanisms, antibiotic target alteration is observed in the vanY and vanT genes, which are associated with resistance to glycopeptides. Additionally, antibiotic efflux is identified in the adeF, qacG, and qacJ genes, which are responsible for the efflux of fluoroquinolones, tetracyclines, and disinfectants. The percentages of identity among the genes vary, with the adeF gene showing the highest identity at 52.38% and a coverage of 99.53%, indicating a strong match with the reference sequence. In contrast, the qacG gene presents an unusual length percentage of 636.45%, suggesting a possible overlap or multiple copies of the gene in the analyzed sequence. The presence of various families of resistance genes indicates a significant diversity of resistance mechanisms in *D. viridis* EMBRAPA| LBA#S001.

Table 1. Genetic resistance profiles in *D. viridis* EMBRAPA| LBA#S001 predicted by RGI (Resistance Gene Identifier) using protein homolog model and RGI strict as detection criteria.



ARO Term	AMR Gene Family	Drug Class	Resistance Mechanism	% Identity of Matching Region	% Length of Reference Sequence	Antibiotic
vanY gene in vanB cluster	vanY, glycopeptide resistance gene cluster	glycopeptide antibiotic	antibiotic target alteration	34.71	80.22	vancomycin
adeF	resistance-nodulation-cell division (RND) antibiotic efflux pump	fluoroquinolone antibiotic, tetracycline antibiotic	antibiotic efflux	52.38	99.53	tetracycline
qacG	small multidrug resistance (SMR) antibiotic efflux pump	disinfecting agents and antiseptics	antibiotic efflux	44.76	636.45	benzalkonium chloride
qacJ	small multidrug resistance (SMR) antibiotic efflux pump	disinfecting agents and antiseptics	antibiotic efflux	36.63	103.74	benzalkonium chloride
vanT gene in vanG cluster	glycopeptide resistance gene cluster, vanT	glycopeptide antibiotic	antibiotic target alteration	33.33	67.56	vancomycin

Each gene may be associated with a specific resistance mechanism, such as modification of the antibiotic, alteration of the cell membrane, or the activity of efflux pumps that remove the antibiotic from the cell.

The presence of resistance genes in microalgae offers several advantages, including a better understanding of environmental resilience These genes can indicate how microalgae adapt and survive in adverse environments, especially those contaminated by antimicrobial agents (Jareonsin & Pumas, 2021). Additionally, microalgae with resistance genes can be explored for the development of biofuels, biopharmaceuticals, or other industrial products that require resilience in hostile environments. The study of resistance in microalgae can also provide insights into the evolution of antibiotic resistance in eukaryotic organisms, especially in comparison to bacteria, which are often the focus of such research, once they share a common evolutionary lineage. Identifying resistance mechanisms can help develop biological control methods in aquatic environments (Fayaz et al., 2023).

The analysis revealed a diversity of resistance genes, tetracycline and vancomycin suggesting that microalgae may have developed multiple mechanisms to cope with environmental stresses, including the presence of antibiotics, favoring antibiotic removal and degradation in wastewater (Leng et al., 2020). The data obtained can serve as a foundation for future investigations into antibiotic resistance, including studies on how resistance genes spread among different organisms and ecosystems.

Also, this analysis can facilitate the selection of transformants, allowing for the evaluation of gene expression, and contributing to the creation of modified organisms that have applications in medicine and biotechnology (Li et al., 2021; Shi et al., 2022).

In general, residual antibiotics from sources such as municipal and industrial wastewater, sewage discharges, and agricultural runoff are continuously released into freshwater environments, creating reservoirs that facilitate the development and spread of antibiotic resistance. Understanding the effects of these antibiotic residues on aquatic organisms, particularly microalgae and cyanobacteria, is crucial because they play vital roles as primary producers in the ecosystem (Le et al., 2022).

Promoters and Transcription factors

The figure 8 provides a detailed overview of the transcription factors (TFs) associated with various genes in the microalga *D. viridis* EMBRAPA| LBA#S001. Different colors

indicate specific families of transcription factors. The AP2 family is known for its involvement in stress responses, developmental processes, and metabolic regulation. It often binds to specific DNA sequences to activate or repress gene expression, playing a vital role in plant and algal adaptation to environmental stresses (Shang et al., 2022). Myb transcription factors are essential for regulating various biological processes, including cell differentiation, metabolism, and stress responses. In microalgae, they can influence pathways related to photosynthesis and secondary metabolite production (Xu et al., 2024). The GATA factors are primarily involved in the regulation of nitrogen and carbon metabolism. They play a significant role in the response to nutrient availability and can influence the expression of genes related to photosynthesis and respiration in microalgae (Virolainen & Chekunova, 2024).

The WRKY family is known for its role in plant defense mechanisms and stress responses. In microalgae, WRKY factors can regulate responses to biotic and abiotic stresses, helping the organism adapt to changing environmental conditions (Wani et al., 2021). TATA-binding protein (TBP) is a general transcription factor that is essential for the initiation of transcription by RNA polymerase II. It plays a fundamental role in the basic transcription machinery and is crucial for the expression of many genes in microalgae (Akinyemi et al., 2021; Dorantes-Torres et al., 2024).

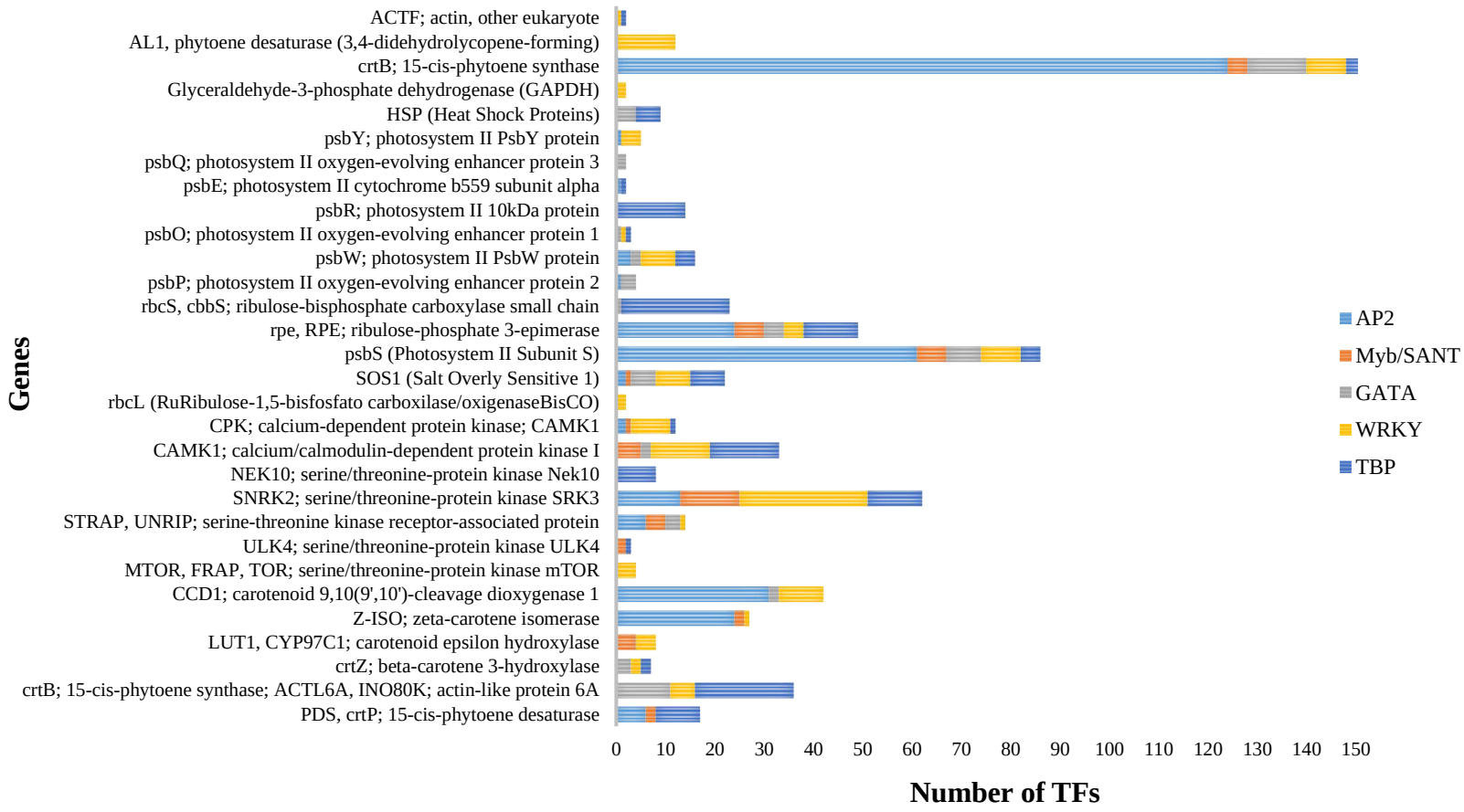


Figure 8. Regulatory Networks: Analyzing Transcription Factor associations in key genes of the microalga *D. viridis* EMBRAPA| LBA#S001

The chart includes a variety of genes, such as those involved in photosynthesis, like *psbY, psbO,* and *psbP*, as well as those related to stress responses, such as *HSP* and *SOS1*, and metabolic functions, including *GAPDH, rbcL* and *rbcS*. Certain genes, such as *rbcS* and *psb* proteins, show a high number of associated transcription factors, indicating a complex regulatory network. Additionally, genes like *ACTL6A* and *crtB* exhibit notable numbers of transcription factors from different families, suggesting their importance in metabolic and regulatory processes (Mitsis et al., 2020). The presence of diverse transcription factors linked to specific genes implies that these genes may be subject to intricate regulatory mechanisms, which could be essential for the adaptation and functionality of Dunaliella under varying environmental conditions (Liang et al., 2020).

A MYB-type transcription factor, Pi Starvation Response1 (PSR1), regulates phosphate starvation signaling by up-regulating phosphatases and Pi transporters. In *C. reinhardtii*, a PSR1 mutation inhibits lipid biosynthesis and abolishes starch production during phosphate starvation. The overexpressing PSR1 increased starch biosynthesis by enhancing the expression of specific genes like starch synthase (SSS1) and phosphorylases (SP1). Additionally, PSR1 is identified as a regulator of TAG biosynthesis in response to nitrogen starvation, where its overexpression boosts TAG accumulation without hindering growth. However, the study did not investigate the expression levels of downstream target genes or the mechanisms of PSR1 under phosphate starvation. Furthermore, it was reported that MYB-type transcription factors also influence CO2-responsive genes in *C. reinhardtii* and salt tolerance in *Dunaliella bardawil*, although this remains unconfirmed by direct genetic experiments (Sun et al., 2018).

Additionally, manipulating transcription factors can stimulate the production of valuable compounds, such as carotenoids and omega-3 fatty acids, with applications in various industries (Li et al., 2019).

In this scenario, bioengineering strains of microalgae will be essential for optimizing and modifying their metabolic outputs to make them competitive with established industrial biotechnology hosts, such as bacteria or yeast. To achieve this, precise and adjustable control over transgene expression will be crucial, requiring the rational development and design of synthetic promoters as a key strategy. Among green microalgae, *Chlamydomonas reinhardtii* serves as the reference species for bioengineering and synthetic biology; however, the repertoire of functional synthetic promoters for this species, and for microalgae in general, is limited compared to other

commercial chassis, highlighting the need to expand the current toolkit for gene expression in microalgae (Milito et al., 2023).

Complete information on the transcription factors with their binding sequences, position and strand can be found in the additional information sheet.

Conclusion

The study of *Dunaliella viridis* EMBRAPA| LBA#S001 provides valuable insights into the biotechnological potential of this microalga, which stands out for its extremophilic capabilities and production of industrially relevant compounds, such as carotenoids. The identification of genes related to salt tolerance and β-carotene production, along with the complete sequencing of mitochondrial, chloroplast, and nuclear genomes, establishes a solid foundation for future applications in biotechnology.

The present study on *Dunaliella viridis* EMBRAPA| LBA#S001 reveals significant insights into the mechanisms of antibiotic resistance and the genetic regulation of this microalga. The identification of resistance genes such as vanY and adeF demonstrates the adaptive capacity of this species in adverse environments, especially in contexts contaminated by antimicrobial agents. This not only reinforces the importance of microalgae in environmental biotechnology but also opens opportunities for the development of biopharmaceuticals.

Furthermore, the analysis of transcription factors associated with critical genes highlights the complexity of gene regulation in *D. viridis*. Factors from the MYB, Myb/SANT, GATA WRKY and AP2 families play essential roles in responding to environmental stresses, regulating metabolic processes and photosynthesis. The ability to manipulate these regulatory networks is a powerful tool for increasing the production of industrially relevant compounds, such as carotenoids and omega-3 fatty acids.

The need to develop synthetic promoters for microalgae is evident, as current tools are limited compared to model organisms like *Chlamydomonas reinhardtii*. This advancement will allow for more precise regulation of gene expression, enabling microalgae to be competitively utilized in biotechnology. The combination of genetic engineering with an understanding of resistance and regulatory mechanisms could transform *D. viridis* into a viable platform for the sustainable production of bioproducts, contributing to innovative solutions in sectors such as health, nutrition, and energy. Thus,

the research not only deepens knowledge about microalgae but also proposes new directions for their industrial use, promoting more sustainable practices adapted to contemporary demands.

References

Akinyemi, T. S., Shao, N., Lyu, Z., Drake, I. J., Liu, Y., & Whitman, W. B. (2021). Tuning gene expression by phosphate in the methanogenic archaeon Methanococcus maripaludis. ACS Synthetic Biology, 10(11), 3028-3039.

Alcock *et al*. 2023. CARD 2023: Expanded Curation, Support for Machine Learning, and Resistome Prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, 51, D690-D699.

Alet AI, Sánchez DH, Cuevas JC, Marina M, Carrasco P, Altabella T, Tiburcio AF, Ruiz OA. New insights into the role of spermine in Arabidopsis thaliana under long-term salt stress. Plant Sci. 2012 Jan; 182:94-100. doi: 10.1016/j.plantsci.2011.03.013. Epub 2011 Mar 30. PMID: 22118620.

Babraham. A. S. FastQC: a quality control tool for high throughput sequence data, 2010. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Barera, S., & Forlani, G. (2023). The role of proline in the adaptation of eukaryotic microalgae to environmental stress: An underestimated tool for the optimization of algal growth. Journal of Applied Phycology, 35(4), 1635-1648.

Bolger, A. M.; Lohse, M.; Usade, B. Advance Access publication Trimmomatic: a flexible trimmer for Illumina sequence data. Genome analysis, v. 30, p. 2114–2120, 2014.

Bombo, G., Cristofoli, N. L., Santos, T. F., Schüler, L., Maia, I. B., Pereira, H., ... & Varela, J. (2023). Dunaliella viridis TAV01: A Halotolerant, Protein-Rich Microalga from the Algarve Coast. Applied Sciences, 13(4), 2146.

Brasil, B. D. S. A. F., Siqueira, F. G., Salum, T. F. C., Zanette, C. M., & Spier, M. R. (2017). Microalgae and cyanobacteria as enzyme biofactories. Algal research, 25, 76-89.

Burak, H., Dunbar, A., & Gilmour, D. J. (2019). Enhancement of Dunaliella salina growth by using wavelength shifting dyes. Journal of Applied Phycology, 31(5), 2791-2796.

Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 2008 Jan;18(1):188-96. doi: 10.1101/gr.6743907.

Capa-Robles, W., García-Mendoza, E., & Paniagua-Michel, J. D. J. (2021). Enhanced β-carotene and Biomass Production by Induced Mixotrophy in Dunaliella salina across a Combined Strategy of Glycerol, Salinity, and Light. Metabolites, 11(12), 866.

Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., ... & Xia, R. (2023). TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. Molecular plant, 16(11), 1733-1742.

Chow, C. N., Yang, C. W., Wu, N. Y., Wang, H. T., Tseng, K. C., Chiu, Y. H., ... & Chang, W. C. (2024). PlantPAN 4.0: Updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. Nucleic acids research, 52(D1), D1569-D1578.

Dębowski, M., Zieliński, M., Kazimierowicz, J., Kujawska, N., & Talbierz, S. (2020). Microalgae cultivation technologies as an opportunity for bioenergetic system development—Advantages and limitations. Sustainability, 12(23), 9980.

Dehghani, J., Balieu, J., Perruchon, O., Mathieu-Rivet, E., Mati-Baouche, N., Lerouge, P., & Bardor, M. (2024). Exploring protein N-glycosylation in the green microalga Dunaliella salina. Algal Research, 103711.

Dismukes, G.C.; Carrieri, D.; Bennette, N.; Ananyev, G.M.; Posewitz, M.C. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. Curr Opin Biotechnol, v. 19, n. 3, p. 235-40, 2008.

Dorantes-Torres, C., Carrera-Reyna, M., Santos, W., Sánchez-López, R., & Merino, E. (2024). PhyloString: A web server designed to identify, visualize, and evaluate functional relationships between orthologous protein groups across different phylogenetic lineages. Plos one, 19(1), e0297010.

Doyle, J. J.; J. L. Doyle. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, v. 19, p.11-15, 1987.

Esquível, M. D. G., Matos, R. G., & Arraiano, C. M. (2024). Plastome Engineering in Microalgae: The Future of the Green Biotechnology. In Microalgal Bioengineering (pp. 167-180). Cham: Springer International Publishing.

Fayaz, T., Renuka, N., & Ratha, S. K. (2023). Antibiotic occurrence, environmental risks, and their removal from aquatic environments using microalgae: Advances and future perspectives. Chemosphere, 140822.

Félix-Castro, M. B., Arredondo-Vega, B. O., Rojas, M., & Gómez-Anduro, G. A. (2023). Synthesis of β-carotene in Dunaliella: From genome analysis to metabolic engineering. Algal Research, 72, 103135.

Feng, J., Li, J., Gao, Z., Lu, Y., Yu, J., Zheng, Q., ... & Zhu, Z. (2015). SKIP confers osmotic tolerance during salt stress by controlling alternative gene splicing in Arabidopsis. Molecular plant, 8(7), 1038-1052.

Gao, F., Nan, F., Feng, J., Lv, J., Liu, Q., Liu, X., & Xie, S. (2021). Comparative morphological, physiological, biochemical and genomic studies reveal novel genes of

Dunaliella bioculata and D. quartolecta in response to salt stress. Molecular Biology Reports, 1-13.

Gharajeh, N. H., Valizadeh, M., Dorani, E., & Hejazi, M. A. (2020). Biochemical profiling of three indigenous Dunaliella isolates with main focus on fatty acid composition towards potential biotechnological application. Biotechnology Reports, v. 26, e00479.

Giudicelli, G. C. Avaliação do marcador nuclear ITS para estudos evolutivos de espécies do gênero Passiflora L. (Passifloraceae). 81 f. Dissertation (Master Degree), Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2015.

Goswami, R. K., Agrawal, K., & Verma, P. (2022). Microalgae Dunaliella as biofuel feedstock and β-carotene production: An influential step towards environmental sustainability. Energy Conversion and Management: X, 13, 100154.

Guermazi, W., Masmoudi, S., Trabelsi, N. A., Gammoudi, S., Ayadi, H., Morant-Manceau, A., & Hotos, G. N. (2023). Physiological and biochemical responses in microalgae Dunaliella salina, Cylindrotheca closterium and Phormidium versicolor NCC466 exposed to high salinity and irradiation. Life, 13(2), 313.

Hadi, M.R., Shariati, M. & Afsharzadeh, S. Microalgal biotechnology: Carotenoid and glycerol production by the green algae Dunaliella isolated from the Gave-Khooni salt marsh, Iran. Biotechnol Bioproc E 13, 540 (2008). https://doi.org/10.1007/s12257-007-0185-7

Hadi, S. I.; Santana, H.; Brunale, P. P.; Gomes, T. G.; Oliveira, M. D.; Matthiensen, A.; Brasil, B. S. DNA barcoding green microalgae isolated from neotropical inland waters. PloS one, v. 11, n. 2, p. e0149284, 2016.

Hamaji T, et al. 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, vol. 8 pg. e57177

Harvey, P. J., & Ben-Amotz, A. (2020). Towards a sustainable Dunaliella salina microalgal biorefinery for 9-cis β-carotene production. Algal Research, 50, 102002.

He, Q., Lin, Y., Tan, H., Zhou, Y., Wen, Y., Gan, J., ... & Zhang, Q. (2020). Transcriptomic profiles of Dunaliella salina in response to hypersaline stress. BMC genomics, 21, 1-17.

Henríquez, V., Escobar, C., Galarza, J., & Gimpel, J. (2016). Carotenoids in microalgae. Carotenoids in Nature, 219-237.

Huang, P. W., Wang, L. R., Geng, S. S., Ye, C., Sun, X. M., & Huang, H. (2021). Strategies for enhancing terpenoids accumulation in microalgae. Applied Microbiology and Biotechnology, 105(12), 4919-4930.

Huang, J. J., Xu, W., Lin, S., & Cheung, P. C. K. (2024). The bioactivities and biotechnological production approaches of carotenoids derived from microalgae and cyanobacteria. Critical Reviews in Biotechnology, 1-29.

Huo, Y. Y., Rong, Z., Jian, S. L., Xu, C. D., Li, J., & Xu, X. W. (2017). A novel halotolerant thermoalkaliphilic esterase from marine bacterium Erythrobacter seohaensis SW-135. Frontiers in microbiology, 8, 2315.

Hu, L., Feng, S., Liang, G., Du, J., Li, A., & Niu, C. (2021). CRISPR/Cas9-induced β-carotene hydroxylase mutation in Dunaliella salina CCAP19/18. AMB Express, 11(1), 1-8.

Jareonsin, S., & Pumas, C. (2021). Advantages of heterotrophic microalgae as a host for phytochemicals production. Frontiers in Bioengineering and Biotechnology, 9, 628597.

Jones P, Binns D, Chang HY, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics (Oxford, England). 2014 May;30(9):1236-1240. DOI: 10.1093/bioinformatics/btu031.

Kanehisa M, Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 28(1), pp. 27-30.

Kang, N. K., Baek, K., Koh, H. G., Atkinson, C. A., Ort, D. R., & Jin, Y. S. (2022). Microalgal metabolic engineering strategies for the production of fuels and chemicals. Bioresource Technology, 345, 126529.

Kolackova, M., Janova, A., Dobesova, M., Zvalova, M., Chaloupsky, P., Krystofova, O., ... & Huska, D. (2023). Role of secondary metabolites in distressed microalgae. Environmental Research, 224, 115392.

Lai, Y. C., & Ducoste, J. J. (2023). Growth of Dunaliella viridis in multiple cycles of reclaimed media after repeated high pH-induced flocculation and harvesting. Science of The Total Environment, 891, 164087.

Le, V. V., Tran, Q. G., Ko, S. R., Lee, S. A., Oh, H. M., Kim, H. S., & Ahn, C. Y. (2023). How do freshwater microalgae and cyanobacteria respond to antibiotics? Critical Reviews in Biotechnology, 43(2), 191-211.

Lee RW, Dumas C, Lemieux C, Turmel M. Cloning and characterization of the Chlamydomonas moewusii mitochondrial genome. Mol Gen Genet. 1991 Dec;231(1):53-8. doi: 10.1007/BF00293821. PMID: 1753945.

Leng, L., Wei, L., Xiong, Q., Xu, S., Li, W., Lv, S., ... & Zhou, W. (2020). Use of microalgae based technology for the removal of antibiotics from wastewater: A review. Chemosphere, 238, 124680.

Li, D. W., Balamurugan, S., Yang, Y. F., Zheng, J. W., Huang, D., Zou, L. G., ... & Li, H. Y. (2019). Transcriptional regulation of microalgae for concurrent lipid overproduction and secretion. Science advances, 5(1), eaau3795.

Li, Y., Xu, Z., Han, W., Cao, H., Umarov, R., Yan, A., ... & Gao, X. (2021). HMD-ARG: hierarchical multi-task deep learning for annotating antibiotic resistance genes. Microbiome, 9, 1-12.

Li, S., Li, X., & Ho, S. H. (2022). How to enhance carbon capture by evolution of microalgal photosynthesis?. Separation and Purification Technology, 291, 120951.

Liang, M. H., & Jiang, J. G. (2017). Analysis of carotenogenic genes promoters and WRKY transcription factors in response to salt stress in Dunaliella bardawil. Scientific reports, 7(1), 37025.

Liang, M. H., Jiang, J. G., Wang, L., & Zhu, J. (2020). Transcriptomic insights into the heat stress response of Dunaliella bardawil. Enzyme and Microbial Technology, 132, 109436.

Lin, B., Cui, Y., Yan, M., Wang, Y., Gao, Z., Meng, C., & Qin, S. (2019). Construction of astaxanthin metabolic pathway in the green microalga Dunaliella viridis. Algal research, 44, 101697.

Liu, C., Wen, X., Pan, H., Luo, Y., Zhou, J., Wu, Y., ... & Li, H. (2024). Bioremoval of Co (II) by a novel halotolerant microalgae Dunaliella sp. FACHB-558 from saltwater. Frontiers in Microbiology, 15, 1256814.

Magdaleno, D., Lopez, H., & Stephano Hornedo, J. L. (2017). The complete mitochondrial genome of the green microalgae Dunaliella salina strain SQ. Mitochondrial DNA Part B, 2(1), 311-312.

Milito, A., Aschern, M., McQuillan, J. L., & Yang, J. S. (2023). Challenges and advances towards the rational design of microalgal synthetic promoters in Chlamydomonas reinhardtii. Journal of Experimental Botany, 74(13), 3833-3850.

Miró-Vinyals, B., Artigues, M., Wostrikoff, K., Monte, E., Broto-Puig, F., Leivar, P., & Planas, A. (2023). Chloroplast engineering of the green microalgae Chlamydomonas reinhardtii for the production of HAA, the lipid moiety of rhamnolipid biosurfactants. New Biotechnology, 76, 1-12.

Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G. P., & Eliopoulos, E. (2020). Transcription factors and evolution: an integral part of gene expression. World Academy of Sciences Journal, 2(1), 3-8.

Mokashe, N., Chaudhari, B., & Patil, U. (2018). Operative utility of salt-stable proteases of halophilic and halotolerant bacteria in the biotechnology sector. International journal of biological macromolecules, 117, 493-522.

Molino, A., Iovine, A., Casella, P., Mehariya, S., Chianese, S., Cerbone, A., ... & Musmarra, D. (2018). Microalgae characterization for consolidated and new application in human food, animal feed and nutraceuticals. International journal of environmental research and public health, 15(11), 2436.

Nawkarkar, P., Chugh, S., Sharma, S., Jain, M., Kajla, S., & Kumar, S. (2020). Characterization of the chloroplast genome facilitated the transformation of parachlorella kessleri-I, a potential marine alga for biofuel production. Current genomics, 21(8), 610-623.

Pagels F, Vasconcelos V, Guedes AC. Carotenoids from Cyanobacteria: Biotechnological Potential and Optimization Strategies. Biomolecules. 2021;11(5):735. Published 2021 May 15. doi:10.3390/biom11050735.

Palmer, J. M., and J. Stajich, 2020 Funannotate v1.8.1: Eukaryotic genome annotation. 10.5281/zenodo.4054262.

Patel, A.; Chaudhary, S.; Syed, B. A.; Gami, B.; Patel, P.; Patel, B. RbcL Marker Based Approach for Molecular Identification of Arthrospira and DunaliellaIsolates from Non-Axenic Cultures. Journal of Genetics and Genetic Engineering, v. 2, n. 2, p. 24-34, 2018.

Polle, J. E. W.; Polle, J. E.; Barry, K.; Cushman, J.; Schmutz, J.; Tran, D.; Hathwaik, L. T.; Lindquist, E. Draft nuclear genome sequence of the halophilic and beta-carotene-accumulating green alga *Dunaliella salina* strain CCAP19/18. Genome Announc., v. 5, n. 43, p. e01105-17, 2017.

Pourkarimi, S., Hallajisani, A., Alizadehdakhel, A., Nouralishahi, A., & Golzary, A. (2020). Factors affecting production of beta-carotene from Dunaliella salina microalgae. Biocatalysis and Agricultural Biotechnology, 29, 101771.

Preetha, K.; John, L.; Subin, C. S.; Vijayan, K. K. Phenotypic and genetic characterization of Dunaliella (Chlorophyta) from Indian salinas and their diversity. Aquatic biosystems, v. 8, n. 1, p. 27, 2012.

Puillandre, N.; Bouchet, P.; Boisselier-Dubayle, M. C.; Brisset, J.; Buge, B.; Castelin, M.; Lozouet, P. New taxonomy and old collections: integrating DNA barcoding into the collection curation process. Molecular Ecology Resources, v. 12, n. 3, p. 396-402, 2012.

Rambaut, A. 2006. FigTree, <http://tree.bio.ed.ac.uk/software/figtree/>.

Riyazat Khadim, S., Mohanta, A., Singh, P., Maurya, P., Kumar Singh, A., Kumar Singh, A., & Asthana, R. K. (2023). A study on Dunaliella salina under selected nutrient manipulation with reference to the biomass, lipid content along with expression of ACCase and RuBisCO genes. BioEnergy Research, 16(1), 622-637.

Roy, U. K., Nielsen, B. V., & Milledge, J. J. (2021). Antioxidant production in Dunaliella. Applied Sciences, 11(9), 3959.Hamaji T, Kawai-Toyooka H, Toyoda A, Minakuchi Y, Suzuki M, Fujiyama A, et al. Multiple independent changes in mitochondrial genome conformation in chlamydomonadalean algae. Genome Biol Evol.2017; 9: 993–999. https://doi.org/10.1093/gbe/evx060 PMID: 31972029.

Ruginescu R, Gomoiu I, Popescu O, Cojoc R, Neagu S, Lucaci I, Batrinescu-Moteau C, Enache M. Bioprospecting for Novel Halophilic and Halotolerant Sources of Hydrolytic Enzymes in Brackish, Saline and Hypersaline Lakes of Romania. Microorganisms. 2020 Nov 30;8(12):1903. doi: 10.3390/microorganisms8121903. PMID: 33266166; PMCID: PMC7760675.

Santos, M. G. B., Duarte, R. L., Maciel, A. M., Abreu, M., Reis, A., & de Mendonça, H. V. (2021). Microalgae biomass production for biofuels in brazilian scenario: a critical review. BioEnergy Research, 14(1), 23-42.

Sathasivam, R.; Juntawong, N. Modified medium for enhanced growth of Dunaliella Strains. Int J Curr Sci, v. 5, P. 67-73, 2013.

Shang, C., Pang, B., Zhang, J., Yu, L., Gan, S., Li, Y., & Wu, H. (2022). Identification of interacting proteins of transcription factor DpAP2 related to carotenoid biosynthesis from marine microalga Dunaliella parva. Frontiers in Marine Science, 9, 907065.

Shi, X., Xia, Y., Wei, W., & Ni, B. J. (2022). Accelerated spread of antibiotic resistance genes (ARGs) induced by non-antibiotic conditions: roles and mechanisms. Water research, 224, 119060.

Shirazian P, Asad S, Amoozegar MA. The potential of halophilic and halotolerant bacteria for the production of antineoplastic enzymes: L-asparaginase and L-glutaminase. EXCLI J. 2016; 15:268-279. Published 2016 Apr 18. doi:10.17179/excli2016-146.

Silva, S. C., Ferreira, I. C., Dias, M. M., & Barreiro, M. F. (2020). Microalgae-derived pigments: A 10-year bibliometric review and industry and market trend analysis. Molecules, 25(15), 3406.7.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), 3210-3212.

Smith, D. R.; Lee, R. W.; Cushman, J. C.; Magnuson, J. K.; Tran, D.; Polle, J. E. The Dunaliella salina organelle genomes: large sequences, inflated with intronic and intergenic DNA. BMC plant biology, v. 10, n. 1, p. 83, 2010.

Soto, J. O. Dunaliella identification using DNA fingerprinting intron-sizing method and species-specific oligonucleotides: new insights on Dunaliella molecular identification. In: Handbook of Marine Microalgae. Academic Press, p. 559-568, 2015.

Souza Celente, G., de Souza Schneider, R. D. C., Rizzetti, T. M., Lobo, E. A., & Sui, Y. (2023). Using wastewater as a cultivation alternative for microalga Dunaliella salina: Potentials and challenges. Science of the Total Environment, 168812.

Sun, X. M., Ren, L. J., Zhao, Q. Y., Ji, X. J., & Huang, H. (2018). Microalgae for the production of lipid and carotenoids: a review with focus on stress regulation and adaptation. Biotechnology for biofuels, 11, 1-16.

Stavridou, E., Karapetsi, L., Nteve, G. M., Tsintzou, G., Chatzikonstantinou, M., Tsaousi, M., ... & Madesis, P. (2024). Landscape of microalgae omics and metabolic engineering research for strain improvement: An overview. Aquaculture, 740803.

Tamura K, Stecher G, and Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. Molecular Biology and Evolution 38:3022-3027.

Tran, D., Louime, C., Võ, T., Giordano, M., Portilla, S., Doan, N., ... & Bui, L. (2013). Identification of Dunaliella viridis using its markers. International Journal of Applied, 3(4).

van der Loos, L. M., Steinhagen, S., Stock, W., Weinberger, F., D'hondt, S., Willems, A., & De Clerck, O. (2024). Functional stability despite high taxonomic turnover characterizes the Ulva microbiome across a 2,000 km salinity gradient. bioRxiv, 2024-06.

Virolainen, P. A., & Chekunova, E. M. (2024). GATA family transcription factors in alga Chlamydomonas reinhardtii. Current Genetics, 70(1), 1.

Xu, Y., Yao, H., Lan, Y., Cao, Y., Xu, Q., Xu, H., ... & Cao, Y. (2024). Genome-Wide Identification and Characterization of CCT Gene Family from Microalgae to Legumes. Genes, 15(7), 941.

Yao, D., Wu, L., Tan, D., Yu, Y., Jiang, Q., Wu, Y., ... & Liu, Y. (2024). Enhancing CO2 fixation by microalgae in a Photobioreactor: Molecular mechanisms with exogenous carbonic anhydrase. Bioresource Technology, 408, 131176.

Wani, S. H., Anand, S., Singh, B., Bohra, A., & Joshi, R. (2021). WRKY transcription factors and plant defense responses: latest discoveries and future prospects. Plant Cell Reports, 40, 1071-1085.

Wu, M., Zhu, R., Lu, J., Lei, A., Zhu, H., Hu, Z., & Wang, J. (2020). Effects of different abiotic stresses on carotenoid and fatty acid metabolism in the green microalga Dunaliella salina Y6. Annals of Microbiology, 70(1), 1-9.

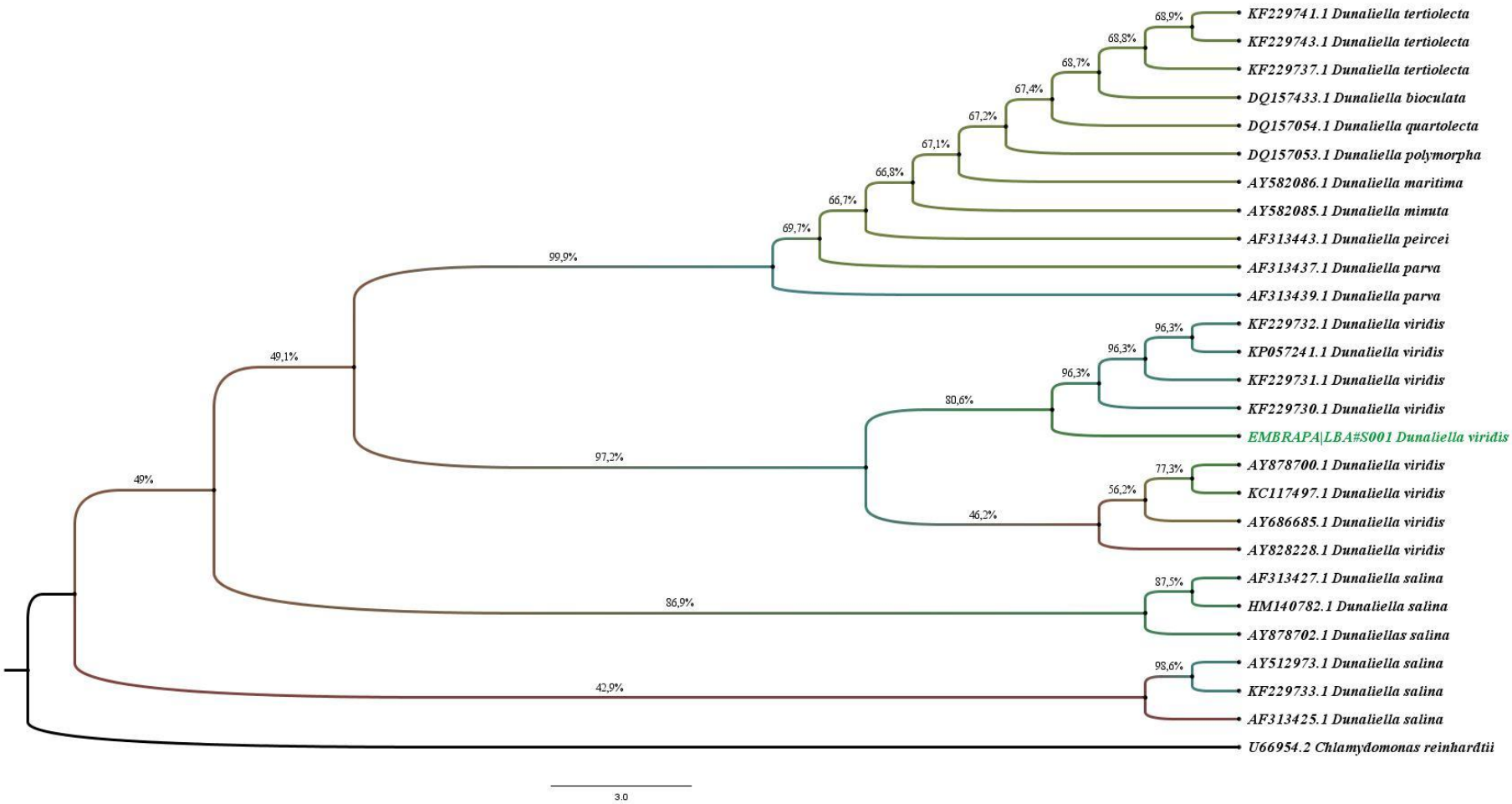
Zeng, J., Wang, Z., & Chen, G. (2021). Biological characteristics of energy conversion in carbon fixation by microalgae. Renewable and Sustainable Energy Reviews, 152, 111661.

Zhao, X., Liu, C., He, L., Zeng, Z., Zhang, A., Li, H., ... & Lou, S. (2022). Structure and Phylogeny of Chloroplast and Mitochondrial Genomes of a Chlorophycean Algae Pectinodesmus pectinatus (Scenedesmaceae, Sphaeropleales). Life, 12(11), 1912.

Supporting Information

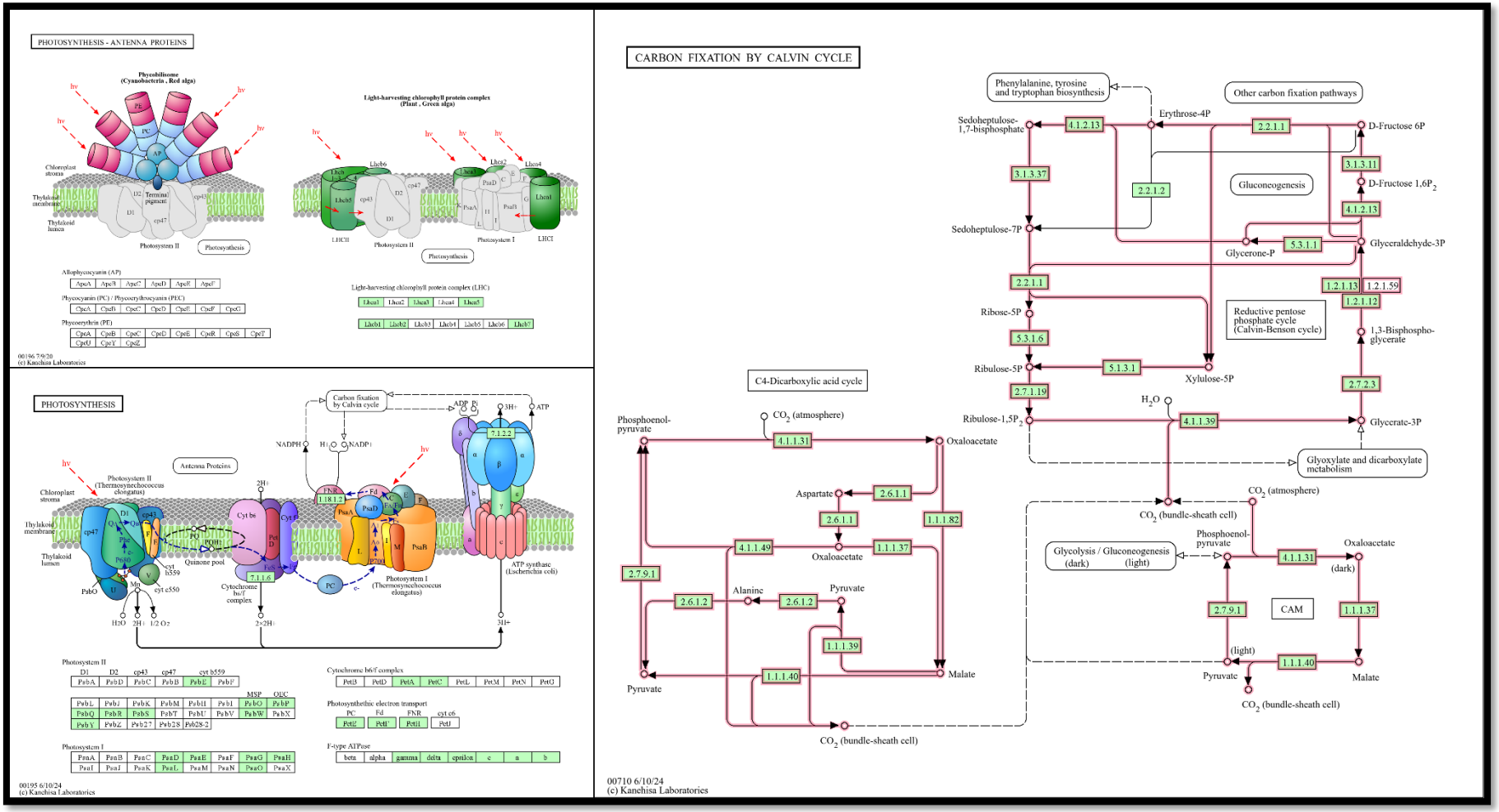
SI.1. Performance of different genome assembly tools and metrics.

Assembly metrics	Canu + LorDEC	SMARTdenovo	Flye
Number of sequences	968	600	276
Total length	176.948.172	161.880.255	126.840.971
Max scaffold length	4.629.997	4.630.206	4.599.249
L50	49	52	41
N50	1.016.379	673.140	755.992
N90	63.824	99.810	222.754
GC content	53,70%	53,96%	53,79%



SI.1. Phylogenetic tree of *D. viridis* EMBRAPA| LBA#S001 reconstructed with MEGA 11 tool. ITS2 sequences alignment with ClustalW, Neighbor-Joining Method and1000 bootstrap value.

SI.2. Phylogenetic tree of *D. viridis* EMBRAPA| LBA#S001 reconstructed with MEGA 11 tool. rbcL gene sequences alignment with ClustalW Codons, Maximun-Likelihood Method and 1000 bootstrap value.



SI.3. Photosynthesis and Carbon Fixation Pathways in *D. viridis* EMBRAPA| LBA#S001.

CONSIDERAÇÕES FINAIS

As espécies de microalgas, *Chlamydomonas biconvexa* e *Dunaliella viridis*, possuem potencial aplicação industrial, especialmente por serem passíveis de serem cultivadas em ambientes como resíduos agroindustriais e meios com hipersalinidade. Tal característica confere vantagens para o cultivo em larga escala, como a redução nos custos de produção e nos contaminantes externos.

Os estudos com *C. biconvexa* Embrapa|LBA40 revelaram que esta cepa possui capacidade de reduzir compostos inorgânicos, como nitrito, amônia e fosfato, presentes no efluente POME em até 99% após 5 dias de cultivo em fotobiorreatores. Assim, as lagoas de tratamento aeróbico normalmente utilizadas em plantas da indústria de palma de óleo poderiam ser substituídas por fotobiorreatores para o cultivo algal, diminuindo o tempo de tratamento e a área necessária para o tratamento convencional. A análise mitogenômica, por sua vez, confirmou a identidade da cepa dentro do gênero *Chlamydomonas*, agrupada em um clado monofilético com *C. reinhardtii*, espécie referência nos estudos de melhoramento de algas. Dessa forma, abrem-se possibilidades para o uso das ferramentas moleculares e protocolos utilizados nesta espécie para o desenvolvimento futuro de melhoramento genético da cepa robusta brasileira.

Já o estudo sobre *Dunaliella viridis* EMBRAPA| LBA#S001 revelou seu potencial biotecnológico, destacando suas capacidades extremófilas e a produção de compostos relevantes, como carotenoides. A identificação de genes relacionados à tolerância ao sal e à produção de β-caroteno, juntamente com o sequenciamento completo dos genomas mitocondrial, cloroplastidial e nuclear, estabelece uma base sólida para futuras aplicações.

Além disso, a análise de genes de resistência, como vanY e adeF, demonstra a adaptabilidade da microalga em ambientes adversos, reforçando sua importância na biotecnologia ambiental e na possibilidade de desenvolvimento de bioindutores e biofármacos. Os fatores de transcrição, incluindo MYB, GATA e WRKY, são fundamentais para a regulação genética em resposta a estresses ambientais, oferecendo oportunidades para aumentar a produção de compostos industriais.

A necessidade de promotores sintéticos para microalgas é evidente, permitindo uma regulação mais precisa da expressão gênica. Assim, a pesquisa propõe novas direções para o uso industrial de *D. viridis*, promovendo práticas mais sustentáveis e inovadoras em setores como saúde, nutrição e energia.

REFERÊNCIAS

ANDRADE, D. V.; COLOZZI FILHO, A. Produção de biomassa e coprodutos. **Microalgas de águas continentais**, v. 2, p. 78, 2014.

BOROWITZKA, M. A. Biology of Microalgae. In: Microalgae in Health and Disease Prevention. **Academic Press**, p. 23-72, 2018.

COUSO, I.; VILA, M.; RODRIGUEZ, H.; VARGAS, M. A.; LEON, R. 2011. EMBRAPA| LBA#S001Overexpression of an exogenous phytoene synthase gene in the unicellular alga *Chlamydomonas reinhardtii* leads to an increase in the content of carotenoids. **Biotechnology Progress**, v. 27, p. 54–60, 2011.

DISMUKES, G.C.; CARRIERI, D.; BENNETTE, N.; ANANYEV, G.M.; POSEWITZ, M.C. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. **Curr Opin Biotechnol,** v. 19, n. 3, p. 235-40, 2008.

HARVEY, Patricia J.; BEN-AMOTZ, Ami. Towards a sustainable Dunaliella salina microalgal biorefinery for 9-cis β-carotene production. **Algal Research**, v. 50, p. 102002, 2020.

KIM, SO-HYUN; LIU, K. H.; LEE, S. Y.; HONG, S. J.; CHO, B. K.; LEE, H.; CHOI, H. K. Effects of light intensity and nitrogen starvation on glycerolipid, glycerophospholipid, and carotenoid composition in *Dunaliella tertiolecta* culture. **PLoS One**, v. 8, n. 9, p. e72415, 2013.

LEON, R.; GONZALES, D.; GALVAN, A.; FERNANDEZ, E. Transgenic microalgae as green cell factories. **Trends in Biotechnology**, v. 22, p. 45–52, 2004.

LV, H., CUI, X., WAHID, F., XIA, F., ZHONG, C., JIA, S. Analysis of the physiological and molecular responses of Dunaliella salina to macronutrient deprivation. **PLoS One**, v. 11, n. 3, p. 1-19, 2016.

MESQUITA, S. Avaliação de estratégias para a indução da produção de βcaroteno por *Dunaliella bardawil*. Dissertação (mestrado) – Universidade Federal do Rio de Janeiro, Escola de Química, Programa de Pós-Graduação em Tecnologia de Processos Químicos e Bioquímicos, Rio de Janeiro, 73 f., 2013.

MUSTAPA, A.N.; MANAN, Z.A.; AZIZI, C.Y.; SETIANTO, W.B.; OMAR, A.K. 2011. Extraction of b-carotenes from palm oil mesocarp using sub-critical R134a. **Food Chem**., v. 125, p. 262-267, 2011.

NASCIMENTO, R. C. Cultivation of microalgae in flat plate photobioreactors for production of biomass and bioremediation of palm oil effluent from agroindustry. Dissertation, 55 p., 2016.

OREN, A. A hundred years of Dunaliella research: 1905–2005. **Saline systems,** v. 1, n. 1, p. 2, 2005.

POLLE, J. E. W.; POLLE, J. E.; BARRY, K.; CUSHMAN, J.; SCHMUTZ, J.; TRAN, D.; HATHWAIK, L. T.; LINDQUIST, E. Draft nuclear genome sequence of the halophilic and beta-carotene-accumulating green alga Dunaliella salina strain CCAP19/18. **Genome Announc.**, v. 5, n. 43, p. e01105-17, 2017.

RAMOS, A. A.; POLLE, J.; TRAN, D.; CUSHMAN, J. C.; JIN, E. S.; VARELA, J. C. The unicellular green alga Dunaliella salina Teod. as a model for abiotic stress tolerance: genetic advances and future perspectives. **Algae**, v. 26, n. 1, p. 3-20, 2011.

SAYRE, R. Microalgae: the potential for carbono capture. **BioScience**, v. 60, n. 9, p. 722-727, 2010.

SKJÅNES, K.; REBOURS, C.; LINDBLAD, P. Potential for green microalgae to produce hydrogen, pharmaceuticals and other high value products in a combined process. **Critical reviews in biotechnology**, v. 33, n. 2, p. 172-215, 2013.

SOFIANA, M. S. J., & HELENA, S. Optimization of β-carotene production in *Dunaliella salina* using led and different culture media. **Jurnal Ilmu Kelautan SPERMONDE**, v. 7, n. 1, p. 36-41, 2021.