# **BRUNA RONCON FAVARELLI**

# "ANÁLISE DO GENOMA DA MICROALGA PRODUTORA DE LIPÍDIOS *NEOCHLORIS OLEOABUNDANS*, VISANDO À PRODUÇÃO DE COMBUSTÍVEIS DE TERCEIRA GERAÇÃO."

CAMPINAS 2012



# UNIVERSIDADE ESTADUAL DE CAMPINAS

## INSTITUTO DE BIOLOGIA

## **BRUNA RONCON FAVARELLI**

# "ANÁLISE DO GENOMA DA MICROALGA PRODUTORA DE LIPÍDIOS *NEOCHLORIS OLEOABUNDANS*, VISANDO À PRODUÇÃO DE COMBUSTÍVEIS DE TERCEIRA GERAÇÃO".

Loi	e eve	inplet cours	sponue	e revação i	a rai
da	tere	dafendida	peic(a)	candidato	(a)
SR	UNA	RONCON	FAVA	RELLI	600.043

e oprovada pela Gamissão Julgadora.

Dissertação apresentada ao Instituto de Biologia para obtenção do Título de Mestre em Genética e Biologia Molecular, na área de Genética de Microorganismos.

Orientador: Prof. Dr. Gonçalo Amarante Guimarães Pereira

CAMPINAS, 2012

i

#### FICHA CATALOGRÁFICA ELABORADA POR ROBERTA CRISTINA DAL' EVEDOVE TARTAROTTI – CRB8/7430 BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

F277a	Favarelli, Bruna Roncon, 1986- Análise do genoma da microalga produtora de lipídios Neochloris oleoabundans, visando à produção de combustíveis de terceira geração / Bruna Roncon Favarelli. – Campinas, SP: [s.n.], 2012.	
	Orientador: Gonçalo Amarante Guimarães Pereira. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.	
	<ol> <li>Microalga. 2. Biocombustíveis. 3. Genoma. 4. Lipídios. 5. Plantas – Efeitos do nitrogênio. 6. Neochloris oleoabundans. I. Pereira, Gonçalo Amarante Guimarães, 1964 II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.</li> </ol>	

#### Informações para Biblioteca Digital

Título em Inglês: Genome analysis of the lipid-producing microalgae Neochloris oleoabundans, aiming the production of third generation fuel Palavras-chave em Inglês: Microalgae Biofuels Genome Lipids Plants - Effect of nitrogen on Neochloris oleoabundans Área de concentração: Genética de Microorganismos Titulação: Mestra em Genética e Biologia Molecular Banca examinadora: Gonçalo Amarante Guimarães Pereira [Orientador] Renato Vicentini dos Santos Paulo Luiz de Andrade Coutinho Data da defesa: 01-08-2012 Programa de Pós Graduação: Genética e Biologia Molecular

Campinas, 01 de agosto de 2012.

### BANCA EXAMINADORA

Prof. Dr. Gonçalo Amarante Guimarães Pereira (Orientador)

Assinatura

Assinatura

Assinatura

Prof. Dr. Renato Vicentini dos Santos

Dr. Paulo Luiz de Andrade Coutinho

Prof(a). Dr(a). Ana Maria Lima de Azeredo-Espin

Prof. Dr. Carlos Augusto Colombo

Assinatura

Assinatura

Dedico esse trabalho ao meu pai, João Carlos Favarelli, amor maior da minha vida e meu eterno incentivador, que infelizmente não pôde ver completada essa jornada.

## AGRADECIMENTOS

Ao professor Dr. Gonçalo Amarante Guimarães Pereira pela competência com que coordena o Laboratório de Genômica e Expressão, junto com as queridas Eliane Laranja Dias, Sílvia Kazue Missawa e Mariana Cerqueira. Obrigada pelo apoio.

Aos meus co-orientadores extra oficiais, os doutores Jorge Maurício Costa Mondego e Ana Carolina Deckmann. Muito obrigada pelos ensinamentos e discussões que muito enriqueceram o meu aprendizado.

Aos meus colegas do grupo das algas, em especial ao Gustavo Pereira de Almeida, meu amigo querido que compartilhou comigo a pesquisa, a bancada, a orientação, as alegrias e tristezas de todos os dias. Obrigada por me aguentar e por fazer meus dias mais alegres.

Ao CNPq e a Braskem-Ideom por viabilizarem financeiramente o projeto.

Aos meninos da Bioinformática, em especial ao Leandro Costa Nascimento e Marcelo Falsarella Carazzolle. Leleco, muito obrigada pela paciência e carinho comigo e pela eficiência de sempre. Tchelo, muito obrigada por acompanhar meu crescimento ao longo desse projeto e por mostrá-lo sempre para mim.

Aos amigos e colegas do LGE, que me ajudaram de diversas maneiras ao longo desses quase cinco anos no laboratório. Em especial, gostaria de agradecer à minha querida amiga Marcela Mendes Salazar pela imensa ajuda profissional e pessoal. Obrigada também à Daniela Toledo Thomazella por acompanhar meu trabalho e pelas ótimas sugestões.

Aos meus amigos e familiares pela torcida e apoio incondicional. Em especial agradeço aos meus queridos: Deise, Waldir, Isa, Heloísa, Sônia, Antonio Carlos e Patrícia.

Por fim, agradeço especialmente ao meu pai pelo amor incondicional e o exemplo de conduta que persistem mesmo após sua partida e são a razão de eu seguir em frente todos os dias.

À minha amada mãe Maria Lucia Roncon Favarelli, uma sobrevivente como eu, que com todo o amor e carinho me incentivou em todos os momentos e pacientemente aguentou todas as minhas crises.

Ao meu namorado, Leonardo Valentin Zeferino, agradeço por ter sido meu chão, quando esse me faltou. Jamais teria conseguido sem você.

# ÍNDICE

LISTA DE ABREVIAÇÕES	3
ÍNDICE DE FIGURAS E TABELAS	5
RESUMO	8
ABSTRACT	10
1. INTRODUÇÃO	12
1.1. Problemática das fontes fósseis de energia	12
1.2. Biocombustíveis - problemas e alternativas	14
1.2.1. Biocombustíveis de 1ª geração	14
1.2.2. Biocombustíveis de 2ª geração	15
1.2.3. Biocombustíveis de 3ª geração	16
1.3. Características gerais das microalgas	16
1.4. Neochloris oleoabundans	17
1.5. Estratégia de produção: fracionamento em três etapas	20
1.6. A fixação de carbono e o acúmulo lipídico	22
2. OBJETIVOS	25
2. OBJETIVOS	<b>25</b> 25
<ul><li>2. OBJETIVOS.</li><li>2.1. Geral.</li><li>2.2. Específicos.</li></ul>	<b>25</b> 25 25
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 25
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 26 26
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 26 26 26
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 26 26 27 27
<ul> <li>2. OBJETIVOS</li></ul>	25 25 26 26 26 27 27 27 27
<ul> <li>2. OBJETIVOS</li></ul>	25 25 26 26 26 27 27 27 29 29
<ul> <li>2. OBJETIVOS.</li> <li>2.1. Geral.</li> <li>2.2. Específicos.</li> <li>3. MATERIAL E MÉTODOS.</li> <li>3.1. Condições de crescimento.</li> <li>3.2. Extração de DNA.</li> <li>3.3. Sequenciamento de DNA e montagem do genoma.</li> <li>3.4. Predição gênica.</li> <li>3.5. Anotação e análise.</li> <li>4. RESULTADOS E DISCUSSÃO.</li> </ul>	25 25 26 26 26 27 27 27 29 29 29 29 29
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 26 26 27 27 27 27 29 29 29 29 23 23
<ul> <li>2. OBJETIVOS</li></ul>	25 25 25 26 26 26 27 27 27 29 29 29 29 29 31 31
<ol> <li>2. OBJETIVOS</li></ol>	25 25 25 26 26 26 27 27 27 29 29 29 29 31 31 31

4.5. Análise funcional	47
4.5.1. Privação de nitrogênio	51
4.5.2. Remodelamento de membranas celulares	52
4.5.3. Mecanismos de fixação de carbono	54
5. CONCLUSÕES E PERSPECTIVAS	56
6. REFERÊNCIAS BIBLIOGRÁFICAS	59
ANEXO I	66

# LISTA DE ABREVIAÇÕES

TAG	triacilglicerídeos
°C	grau Celsius
μL	microlitro
μM	micromolar
ACL	ATP citrato liase
ACP	proteína carreadora de acil
AIC	Critério de Informação de Akaike
AMP	ácido adenílico
ATP	adenosina trifosfato
B12	cobalamina ou vitamina B12
C3	Mecanismo que fixa carbono em uma molécula com três carbonos (ciclo de Calvin)
C4	Mecanismo que fixa carbono em uma molécula com quatro carbonos
CDD	Conserved Domain Database
CESA-like	celulose sintase like
Ciclo TCA	ciclo do ácido cítrico ou ciclo de Krebs
CO <sub>2</sub>	dióxido de carbono
CoA	coenzima A
DGAT	diacilglicerol acil-transferase
DNA	ácido desoxiribonucléico
FAT1	acil-ACP tioesterase
g	grama
Gt	Gigatonelada
GTP	guanosina trifosfato
h	horas
HCO <sup>3-</sup>	bicarbonato
нн	Hedgehog
HHIP	inibidor de Hedgehog
Kb	Kilobase
kD	kilodaltons
LGE	Laboratório de Genômica e Expressão
Mb	Megabases
metE	metionina sintase independente de cobalamina
mM	milimolar
NAD	dinucleotídio nicotinamida-adenina
NCBI	National Center for Biotechnology Information
NR	banco de dados de sequencias não redundantes
Pb	par de bases
Ptc	Patched

Rpm	rotações por minuto
rRNA	ácido ribonucléico ribossomal
Rubisco	ribulose-1,5-bisfosfato carboxilase/oxidade
Smo	Smoothened
V	volume

# ÍNDICE DE FIGURAS E TABELAS

## Figuras:

Figura 1: Consumo mundial de energia comercializada por tipo de combustível. A projeção do consumo foi feita de 2010 a 203012
Figura 2: Oscilação do preço do petróleo bruto em função de conflitos políticos ocorridos nos países com reservas mais expressivas13
Figura 3: Conteúdo lipídico de <i>Neochloris oleabundans</i> (valores médios) detectado por citometria de fluxo usando o corante vermelho do Nilo em diversas condições experimentais
Figura 4: Conteúdo e produtividade lipídica de <i>N. oleoabundans</i> cultivadas em meios de cultura contendo 5 mM de nitrato de sódio, ureia ou bicarbonato de amônio19
Figura 5: Perfil de ácidos graxos de Neochloris oleoabundans19
Figura 6: Curvas de crescimento de <i>Neochloris oleoabundans</i> em sistema fechado (a) e lagoa aberta (b) em diversas condições experimentais20
Figura 7: Representação do modelo de produção de cadeias carbônicas lipídicas em três etapas21
Figura 8: Esquema simplificado do ciclo de Calvin22
Figura 9: Via de biossíntese de ácidos graxos e triacilglicerídeos24
Figura 10: Esquema da separação dos contigs do genoma de alga, de bactéria e das organelas28
Figura 11: Classificação dos scaffolds por conteúdo GC e cobertura médios32
Figura 12: Correlação do genoma de diversas microalgas com sua respectiva densidade gênica (genes/Kb) Dados calculados a partir da divisão da quantidade de genes preditos no genoma de cada uma das espécies pelo respectivo tamanho de cada genoma. Os dados foram obtidos no NCBI (http://www.ncbi.nlm.nih.gov/genome)
Figura 13: Filogenia das espécies utilizadas para a detecção de família gênicas em <i>Neochloris oleoabundans</i> . Os valores próximos aos ramos internos são porcentagens de bootstraps. O ramo de <i>Nannochloropsis gaditana</i> (pontilhado) está fora de escala
Figura 14: As predições gênicas de <i>N. oleoabundans</i> foram comparadas com todos os genomas já sequenciados contidos no banco de dados de proteínas não-redundantes (NR) utilizando um Blastp. O número de vezes que um organismo foi o primeiro hit (e-value menor que e-5) são indicados. Os hits com <i>Chlorella</i>

- Figura 16: A proteína Hedgehog (HH) inteira é composta de dois domínios distintos: o domínio amino-terminal 'Hedge' (mostrado em verde escuro) e o domínio 'Hog' carboxi-terminal (mostrado em azul), ambos os quais são também encontrados em outras proteínas que não são membros da família HH......45

#### **Tabelas:**

Tabela 1: Composição do meio Bold's Basal Medium26
Tabela 2: Genes envolvidos no metabolismo de cobalamina presentes no genoma de Neochloris oleoabundans
Tabela 3: Estatísticas do genoma35
Tabela 4: Porcentagem do genoma de diversas algas e A. thaliana contidas em famíliasortólogas classificadas por tamanho
<ul> <li>Tabela 5: Anotação das famílias gênicas encontradas apenas em Nannochloropsis gaditana (N. g.) e Neochloris oleoabundans (N. o.) e o número de genes contido em cada família41</li> </ul>
Tabela 6: Anotação das famílias gênicas encontradas apenas em <i>Neochloris oleoabundans</i> e o número de genes contido em cada família. As famílias marcadas com um asterisco (*) foram obtidas a partir de análise do CDD e as demais foram obtidas a partir do OrthoMCL

#### **RESUMO**

A busca de fontes sustentáveis de biocombustíveis tem levado ao interesse renovado em microalgas como potencial matéria-prima. Esses organismos acumulam grandes quantidades de óleos sob a forma de triacilglicerídeos (TAG) quando em privação de nutrientes, entretanto, uma análise aprofundada dos mecanismos moleculares envolvidos nesse acúmulo ainda está no início. É improvável que uma espécie apresente naturalmente todas as características necessárias para a produção industrial de biocombustíveis, como altas taxas de crescimento e acúmulo de lipídios. Sendo assim, fazse necessária a ampliação do conhecimento da genética, bioquímica e das técnicas de manipulação desses organismos para que futuramente a cepa ideal seja obtida por engenharia genética. A microalga verde Neochloris oleoabundans foi escolhida como objeto de estudo por ser naturalmente capaz de acumular grandes quantidades de TAG quando em privação de nitrogênio. Seu genoma foi sequenciado, visando a obtenção de indícios da origem genética do fenótipo de interesse. Foram preditos 10.006 genes distribuídos ao longo de 1.544 scaffolds, totalizando aproximadamente 40 Mb de genoma nuclear. A análise comparativa com outras cinco espécies de microalgas mostrou que o conteúdo do genoma de Neochloris se assemelha ao de Chlorella variabilis. A presença de famílias gênicas ortólogas muito pequenas na espécie estudada motivou a expansão dessa análise para as demais espécies de alga já sequenciadas, indicando que essa ausência de expansões gênicas pode ser um padrão de microalgas. As famílias gênicas exclusivas de Neochloris mostraram que a parede celular dessa espécie pode ser formada por hemicelulose e quitina, já que foram encontrados genes da síntese desses polissacarídeos no genoma. As vias de metabolismo energético foram mapeadas, permitindo entender o caminho energético existente na célula estudada. Além disso, a comparação do número de cópias dos diversos genes envolvidos nos processos de interesse permitiu levantar vias e moléculas chaves para estudos mais detalhados. Uma enzima chave envolvida na sinalização de privação de nitrogênio, a AMP deaminase, foi encontrada em maior número de cópias e pode estar envolvida no direcionamento para acúmulo de TAG. Há indícios de que um possível resultado da ausência de nitrogênio seja a remodelação de membranas

celulares, principalmente do cloroplasto, e quebra de pigmentos fotossintéticos. Essas reações podem gerar ácidos graxos livres e substâncias tóxicas para a célula que são neutralizadas através de sua participação na síntese de TAG. O mecanismo de concentração de carbono empregado por um organismo fotossintetizante afeta o acúmulo de biomassa, pois plantas que utilizam o mecanismo C4-*like* são mais eficientes na fixação de carbono do que plantas C3. Não foi possível determinar qual o mecanismo preferencialmente empregado por *Neochloris*, mas grande parte dos genes envolvidos no mecanismo C4-*like* foram identificados no genoma. Tais genes consistem em alvos para manipulação genética, visando direcionar a fixação de carbono para C4-*like*. A maior eficiência fotossintética resultante pode ocasionar aumento da biomassa e do acúmulo de cadeias carbônicas na célula dessa alga, podendo gerar um maior acúmulo lipídico.

#### ABSTRACT

The search for sustainable sources for the production of biofuels has resulted in an increased interest in microalgae as a potential feedstock. These microrganisms accumulate large amounts of lipids in the form of triglycerides (TAG) when subjected to some types of nutrient deprivation. However, the elucidation of the molecular mechanisms involved in these metabolic responses is far from necessary for the effective use of microalgae in industrial processes. It is unlikely that a species naturally present all the characteristics necessary for the industrial production of biofuels, as for example high rates of growth and lipid accumulation. Thus, it is necessary to expand the knowledge of genetics, biochemistry and genetic manipulation of microalgae in order to adapt the productivity of wild-strains to the demands of the bioenergy sector. In this context, the green algae *Neochloris oleoabundans* was chosen as the object of study because it is naturally capable of accumulating large amounts of TAG when subjected to nitrogen deprivation. Its genome was sequenced in order to obtain genetic data to increase the understanding and, in the future, to produce phenotypes of interest. A total of 10,006 predicted genes were identified over the 1,544 scaffolds generated, totalizing approximately 40 Mb of nuclear genome. The comparative analysis with five other microalgae species showed that the genetic content of *Neochloris* resembles that of other green algae, Chlorella variabilis. The presence of small gene families in the species studied led to the expansion of this analysis for some other sequenced species of algae, indicating that the absence of gene expansions can be a common pattern of microalgae. The unique Neochloris gene families showed that the cell wall of this species may be formed from hemicelluloses and chitin, since genes associated to the metabolism of these polysaccharides were found in the in the genome. Metabolic reconstruction of energetic pathways allowed an increase in the comprehension of energy routes in Neochloris. Moreover, comparing the number of copies of several genes involved in the processes of interest allowed us the identification of key candidates for further studies. A key enzyme involved in signaling nitrogen deprivation, AMP deaminase, was found as presenting

various gene copies in *Neochloris* genome and may be involved in the response of TAG accumulation in situations of nitrogen deprivation. Also, there are some results indicating that the lack of nitrogen may promote the remodeling of cell membranes, especially in the chloroplast, as well as the breakage of photosynthetic pigments. These responses would generate free fatty acids and toxic substances, which would be therefore neutralized by the synthesis of TAG. Other response studied was the carbon concentration after photosynthesis. The carbon concentration mechanism employed by a photosynthetic organism affect the accumulation of biomass; it is known that plants using C4-like mechanism are more efficient than C3 plants. Despite the *in silico* data is insufficient to determine the mechanism employed by Neochloris, it was observed that most of the genes involved in C4-like mechanism were identified in its genome. These genes are interesting targets for genetic manipulation of *Neochloris* in order to adjust its metabolism towards a C4-like metabolism. In that case, it is expected that the resulting higher photosynthetic efficiency can lead to an increased biomass and larger accumulation of carbon chains in the cell of this alga, resulting in an increase in lipid productivity rates.

## 1. INTRODUÇÃO

#### 1.1. Problemática das fontes fósseis de energia

O aumento contínuo no consumo de combustíveis derivados de petróleo (figura 1) já é considerado preocupação global por consumir sua fonte geradora e contribuir consideravelmente para a acumulação de gás carbônico na atmosfera, principal causador do efeito estufa. Assim, o grande acúmulo deste gás tornou-se uma questão importante em discussões sobre desenvolvimento do conceito de economia verde, já que o CO<sub>2</sub> está relacionado com 85% do efeito estufa e, anualmente, 9,4 Gton (9,4<sup>·</sup>10<sup>6</sup> toneladas) desse gás são liberadas na atmosfera, sendo 7,9 Gton de origem fóssil (Raupach, 2007). Considerando que nada seja feito, predições indicam que em 2030 as emissões de gás carbônico oriundas de combustível serão 50-110% maiores que as observadas em 2000, estimadas em 25,5 Gt de CO<sub>2</sub> (IPCC, 2007).



Figura 1: Consumo mundial de energia comercializada por tipo de combustível. A projeção do consumo foi feita de 2010 a 2030.

Além dos impactos ambientais cada vez maiores, as projeções indicam um aumento significativo no consumo de petróleo e outros compostos de origem fóssil até 2030 (figura 1). A demanda crescente acalenta ainda mais a preocupação com o esgotamento das reservas e com a dependência do crescimento econômico de uma só fonte de energia. Combustíveis fósseis fornecem mais de 90% da energia utilizada no mundo todo, mas as reservas são limitadas a algumas regiões (Boudghene Stambouli & Traversa, 2002) cujas questões políticas repercutem nos custos de produção e de bens de consumo em todo o mundo (Abbasi et al., 2011). O problema é agravado com a grande instabilidade que ocorre nos maiores produtores, como o Oriente Médio, ocasionando flutuações nos preços dos barris (figura 2).



Figura 2: Oscilação do preço do petróleo bruto em função de conflitos políticos ocorridos nos países com reservas mais expressivas. Fonte: BBC Brasil.com

Portanto, a busca de fontes sustentáveis e ambientalmente seguras de energia para a nossa economia e sociedade industrial de consumo tornou-se urgente nos últimos anos (Mabee *et al.*, 2005). Consequentemente, é renovado o interesse na produção e utilização de combustíveis a partir de outras fontes como plantas ou resíduos orgânicos (Naik *et al.*, 2010).

#### 1.2. Biocombustíveis – Problemas e alternativas

Atualmente, procura-se uma matéria-prima industrial alternativa e processos verdes para obter os produtos químicos de origem fóssil a partir de recursos renováveis, como a biomassa (Naik *et al.*, 2010). Biocombustíveis produzidos a partir de plantas ou resíduos orgânicos podem ajudar a reduzir a dependência mundial do petróleo e a produção de CO<sub>2</sub>, uma vez que as plantas consomem gás carbônico à medida que crescem (Osamu & Carl, 1989). Todavia, a melhoria na utilização de biomassa requer um esforço tremendo para desenvolver sistemas novos, nos quais conversão, produção e utilização de produtos de base biológica devem ser realizadas de forma eficiente e com baixo impacto ambiental (Osamu & Carl, 1989). Sendo assim, muitos pesquisadores e empresas de todo o mundo estão trabalhando no desenvolvimento de tecnologias integradas de conversão de biomassa.

#### 1.2.1. Biocombustíveis de 1ª geração

A primeira geração de biocombustíveis é caracterizada por sua capacidade de ser queimado na combustão interna dos motores e distribuído através da infraestrutura existente, além de alguns serem usados em veículos com tecnologia "Flex" e a gás natural (Naik *et al.*, 2010). Anualmente, são produzidos comercialmente quase 105 bilhões de litros (Worldwatch Institute).

Os três principais tipos utilizados comercialmente são biodiesel, etanol e biogás, embora o butanol também seja considerado um biocombustível de primeira geração. O biodiesel é um substituto do óleo diesel produzido através de transesterificação de óleos vegetais ou óleos e gorduras residuais. O bioetanol é um substituto da gasolina, derivado do açúcar ou do amido através da fermentação por microrganismos (Naik *et al.*, 2010). Nos Estados Unidos é produzido a partir do milho e, no Brasil, um dos maiores produtores mundiais, a matéria-prima é a cana de açúcar. Somos o único país a utilizar o bioetanol em larga escala, sendo que mais de 15% da frota nacional pode rodar com álcool (Favaro, 2007). Recentemente, incentivos à produção vem possibilitando a adição crescente desse biocombustível à gasolina.

Entretanto, existem preocupações sobre o fornecimento de matérias-primas, incluindo o impacto que ele pode ter sobre a biodiversidade, uso do solo e competição com culturas alimentares. A principal desvantagem dos biocombustíveis de primeira geração é o debate alimento *versus* combustível, sendo que uma das razões para a alta do preço dos alimentos é o aumento na produção destes combustíveis (Laursen, 2006). Além disso, é alegado que biodiesel não é uma tecnologia com custo eficaz para a redução das emissões de CO<sub>2</sub> (Naik *et al.*, 2010).

#### 1.2.2. Biocombustíveis de 2ª geração

Visando solucionar o problema da competição com alimentos e encontrar uma matéria-prima barata surgiram os biocombustíveis de segunda geração. São produzidos a partir de materiais lignocelulósicos da biomassa vegetal, que é composta em grande parte por paredes celulares de plantas das quais tipicamente 75% são polissacarídeos (Pauly, 2008). Tais moléculas representam uma valiosa matéria-prima para biocombustíveis avançados que podem ser obtidos por meio de hidrólise e fermentação (etanol, por exemplo) ou gaseificação. Típicos recursos para estes combustíveis são as culturas florestais de curta rotação (choupo, salgueiro e eucalipto), gramíneas perenes e resíduos provenientes da indústria da madeira, da silvicultura e da agricultura. No Brasil, pesquisadores tentam utilizar o bagaço da cana de açúcar como substrato para a fermentação que produz o etanol. Tratam-se dos mais baratos e abundantes materiais não alimentares disponíveis a partir de plantas (Gomez *et al.,* 2008; Zabaniotou *et al.,* 2008).

Atualmente, a produção de tais combustíveis ainda não é comercial e rentável porque há um número de barreiras técnicas que necessitam ser superadas antes que seu potencial possa ser concretizado, embora já existam instalações piloto e de demonstração sendo desenvolvidas. No entanto, a produção a partir de subprodutos agrícolas só poderia satisfazer uma proporção da crescente demanda por combustíveis líquidos. Isso tem gerado grande interesse em fazer uso de culturas de biomassa dedicadas como matéria-prima para produção de biocombustíveis (Gomez *et al.,* 2008).

Muitas pesquisas ainda estão em desenvolvimento para encontrar o microrganismo ou a enzima ideal, capaz de degradar os resíduos lignocelulósicos em moléculas mais simples que outros organismos utilizarão como substrato para a fermentação. Enquanto esse processo ainda é otimizado e viabilizado financeiramente, buscam-se alternativas mais simples e naturalmente mais eficientes.

#### 1.2.3. Biocombustíveis de 3ª geração

No contexto atual uma alternativa sustentável, e talvez a mais promissora, são as microalgas (Ratledge & Cohen, 2008). Trata-se de organismos fotossintetizantes capazes de utilizar energia solar para promover a fixação de CO<sub>2</sub> em cadeias carbônicas ricas em energia, as quais podem ser utilizadas como biocombustíveis de terceira geração.

Deve-se considerar também que a cultura de microalgas não compete por solos agricultáveis, já que cresce em substratos simples e de baixo custo, podendo se estender verticalmente, multiplicando seu rendimento (Tredici & Materassi, 1992). Além disso, diferentemente de combustíveis mais alternativos como o hidrogênio, o biodiesel de microalgas não implica na substituição da infra-estrutura existente, por exemplo, em motores de veículos e sistemas de abastecimento (Briggs, 2004).

As microalgas têm cada vez mais se mostrado promissoras, não apresentando riscos de impacto ambiental ou prejuízo às atividades humanas (Harwood & Guschina, 2009). Entretanto, só tornar-se-ão uma possibilidade real quando houver competitividade de custo de produção e visão estratégica de mercados (Benemann, 2008).

#### 1.3. Características gerais das microalgas

As microalgas são responsáveis pela liberação de cerca de 50% do oxigênio da atmosfera (Field *et al.,* 1998). Devido à sua grande diversidade, podem ser encontrados organismos no grupo capazes de produzir vários compostos de interesse comercial, como pigmentos, vitaminas, açúcares e óleos (Pruvost *et al.,* 2002; Borowitzka, 2006). São organismos unicelulares, cujo ciclo de vida é curto e de estruturas simples, sendo considerados de fácil manipulação e cultivo em laboratórios (Hallmann, 2007).

Ao contrário de plantas superiores oleaginosas, as microalgas apresentam altíssimas taxas de crescimento chegando a dobrar sua biomassa em 24h (Teixeira *et al.,* 2007). Em escala experimental, estima-se que elas possam produzir de 200 a 300 vezes mais óleo vegetal do que a maioria das oleaginosas em uma área 100 vezes menor (Teixeira *et al.,* 2007). São consideradas como potencial matéria-prima para a produção de biodiesel, pois os óleos que produzem possuem características físico-químicas e químicas similares às de óleos vegetais (Teixeira *et al.,* 2007). O conteúdo lipídico de muitas espécies pode chegar a 80% do peso seco (Metting Jr., 2005; Spolaore *et al.,* 2006), favorecendo seu emprego para fins energéticos (Chisti, 2007; Rocha *et al.,* 2003).

A busca por espécies e linhagens mais adequadas constitui o principal foco de diversas iniciativas para superação das limitações atuais do uso de microalgas (Sheehan *et al.*, 1998). É importante encontrar um organismo naturalmente capaz de acumular altas quantidades de triacilglicerídeos, o substrato para a produção de biodiesel. Sendo assim, diversas informações foram levantadas na literatura e a espécie *Neochloris oleoabundans* se destacou, tornando-se o objeto do presente trabalho.

#### 1.4. Neochloris oleoabundans

É uma microalga verde, unicelular, pertencente à classe *Chlorophyceae*, mesmo táxon da espécie modelo *Chlamydomonas reinhardtii*. Apesar de pouco estudada, é bastante citada devido a seu teor lipídico alto e foi escolhida com base em revisões bibliográficas que comprovam essa superioridade (Tornabene, 1983; Sheehan *et al.*, 1998; Li *et al.*, 2008; Gouveia *et al.*, 2009; da Silva *et al.*, 2009; Griffiths & Harrison, 2009; Pruvost *et al.*, 2009; Beal *et al.*, 2010; Huang *et al.*, 2010; Wahal & Viamajala, 2010; Murray *et al.*, 2011; Pruvost et al., 2011; Wang & Lan, 2011).

O primeiro estudo relatando a composição lipídica de *Neochloris oleoabundans* foi realizado por Tornabene e colaboradores em 1983. Eles observaram que em privação de nitrogênio, a alga apresenta um acúmulo significativo de triacilglicerídeos. Trabalhos mais recentes (Gouveia *et. al*, 2009) detalham a produção de biomassa, o crescimento e várias condições de cultivo que podem ser empregadas para induzir maior produção de óleo,

sendo que a maioria delas emprega meios sem carbono e com concentrações limitantes de compostos com nitrogênio e fósforo. A maior concentração de lipídios ocorre em meio sem nitrogênio após 6 dias de privação (figura 3), chegando a 56% g/g da biomassa produzida. Em meio suplementado com nitrato a produção cai para apenas 20% g/g (Gouveia *et. al*, 2009).



Figura 3: Conteúdo lipídico de Neochloris oleabundans (valores médios) detectado por citometria de fluxo usando o corante vermelho do Nilo em diversas condições experimentais (Gouveia *et. al*, 2009).

Culturas com diversas fontes de nitrogênio foram testadas e, dentre elas, nitrato mostrou a maior concentração de lipídios e produtividade global (figura 4, Li *et. al*, 2008).



🗀 Lipid Content (g/g) 🔳 Lipid Productivity (g/l·d)



Além da quantidade de TAG produzida, o perfil de ácidos graxos também faz dessa espécie uma promessa para a produção de biocombustíveis (figura 5). O ácido oléico (18:1ω9) é o mais abundante, seguido do palmítico (16:0) e do esteárico (18:0) (Gouveia *et. al*, 2009). As cadeias mais abundantes são médias e longas com poucas insaturações (Tornabene *et al.*, 1983) e a proporção de ácido linoleico (18:3ω3) está abaixo de 12%, atendendo aos requerimentos do padrão europeu EN 14214 para a produção de biodiesel (Gouveia *et. al*, 2009).



Figura 5: Perfil de ácidos graxos de Neochloris oleoabundans. Adaptado de Gouveia et. al, 2009.

O sistema de cultivo mais estudado é em pequena escala no laboratório, mas existem trabalhos relatando também crescimento em pequenos biorreatores (Gouveia *et. al*, 2009) ou em lagoas abertas (Murray *et. al*, 2010). O acúmulo de biomassa, como esperado, é mais lento em sistemas abertos (figura 6, b) levando de 15 a 20 dias para atingir a fase estacionária. Isso ocorre devido a grande variação das condições ambientais ao longo do crescimento (Murray *et. al*, 2010). Com maior controle sobre as variáveis que atuam no crescimento, a fase estacionária pode ser atingida em 5 dias, tanto em privação quanto em presença de nitrogênio (figura 6, a; Gouveia *et. al*, 2009).



Figura 6: Curvas de crescimento de Neochloris oleoabundans em sistema fechado (a; Gouveia *et. al*, 2009) e lagoa aberta (b; Murray *et. al*, 2010) em diversas condições experimentais.

#### 1.5. Estratégia de produção: fracionamento em três etapas

É possível observar (figura 6, a) que, quando sob privação de nutrientes, o crescimento e, consequentemente, o acúmulo de biomassa é comprometido. Sabe-se que em microalgas a duplicação celular e o acúmulo lipídico são metabolicamente divergentes. Enquanto no primeiro caso as cadeias carbônicas são desviadas para a síntese protéica, no segundo o objetivo é o armazenamento dessas cadeias (Rodolfi, *et al.*, 2008). Na tentativa de solucionar esse impasse dividiu-se o processo de cultivo em três etapas subsequentes (Figura 7): (1) a primeira etapa dedicada ao aumento da quantidade de células na cultura por divisão celular, (2) a segunda consistindo na acumulação efetiva de lipídios e (3) a terceira finalizaria o processo com a secreção espontânea dos lipídios.



Figura 7: Representação do modelo de produção de cadeias carbônicas lipídicas em três etapas.

A situação mais próxima da ideal seria aquela em que as células primeiramente se multiplicassem com alta velocidade, seguido de outro período no qual a quantidade de células não sofreria alteração. Nesta segunda etapa seria fornecida uma sinalização capaz de induzir o acúmulo lipídico e, consequentemente, aumento do volume celular. Ao final dessa etapa, ocorreria a secreção dos lipídios para o meio extracelular, facilitando a recuperação dos mesmos e evitando a perda das células. Um modelo semelhante foi anteriormente proposto por Walker e colaboradores (2005) para a produção de  $\omega$ -3, exceto pela etapa de excreção. O processo de Walker (2005) , assim como o nosso, consiste em uma etapa prévia de acumulação de biomassa e uma etapa subsequente onde ocorre o aumento de massa devido à acumulação do composto de interesse, mostrando que o processo é factível.

Mesmo considerando a importância de triagens da biodiversidade, admite-se que é bastante improvável a junção espontânea de todas as características industrialmente desejáveis em um único organismo selvagem (Alper & Stephanopoulos, 2009). Assim sendo, a principal plataforma tecnológica para a otimização da produção de lipídios e o estabelecimento da estratégia de cultivo de microalgas é, indiscutivelmente, a manipulação genética, destacando-se a engenharia metabólica (Walker *et al.*, 2005; Rosenberg *et al.*, 2008; Alper & Stephanopoulos, 2009) e o sequenciamento de genomas,

visando compreender os mecanismos moleculares que levam a acumulação de óleo e nortear a estratégia de manipulação genética.

#### 1.6. A fixação de carbono e o acúmulo lipídico

Para compreender o metabolismo energético de *Neochloris oleoabundans* as principais vias a serem estudadas são as de fixação de carbono e de síntese de TAG. As microalgas possuem processo fotossintético e de biossíntese de lipídios similares ao descrito para plantas terrestres (Guschina & Harwood, 2006), assim estes podem nos servir de ponto de partida. Em todos os organismos fotossintetizantes o ciclo dos três carbonos, ou de Calvin, (figura 8) é o ponto inicial para o metabolismo de carbono.



Figura 8: Esquema simplificado do ciclo de Calvin.

O composto inicial e final do ciclo é um açúcar de cinco carbonos, a ribulose-1,5bifosfato, que se liga covalentemente ao CO<sub>2</sub>. O componente resultante contendo seis carbonos sofre uma redução e quebra-se, formando uma molécula com três carbonos. A cada volta completa, uma molécula de dióxido de carbono entra no ciclo. Seis voltas são necessárias para produzir duas moléculas de três carbonos e, portanto, uma molécula de glicose (Raven, et al., 1992).

O ciclo de Calvin não é a única via fixadora de carbono. Em algumas plantas, o primeiro produto da fixação do  $CO_2$  é uma molécula com quatro carbonos, o oxaloacetato. As plantas que empregam essa via junto com o ciclo de Calvin são chamadas de C4, deferentes das C3, que utilizam apenas o ciclo descrito na figura 8. A anatomia das plantas C4 permite uma separação espacial entre a via C4 e o ciclo de Calvin (Raven, *et al.*, 1992). O mecanismo empregado por *N. oleabundans* é desconhecido e será discutido no presente trabalho.

O entendimento da via de síntese de lipídios na espécie estudada também é importante, para que se possa estabelecer uma possível relação entre a privação de nitrogênio e o alto acúmulo de TAG. Embora esse processo seja bastante conhecido, os fatores genéticos que controlam o acúmulo de lipídios em um organismo em particular ainda são desconhecidos (Zhang *et al.*, 2007).

Todos os átomos de carbono encontrados em um ácido graxo são derivados da acetil-coenzima A (acetil-CoA), que é um intermediário em muitas vias metabólicas celulares. Uma vez que as membranas são impermeáveis aos derivados CoA, pode-se inferir que a acetil-CoA é gerada em pelo menos quatro grupos metabólicos distintos representando os quatro compartimentos subcelulares onde o metabolismo de acetil-CoA ocorre: plastídios (local da síntese de ácidos graxos), mitocôndrias, peroxissomos e citosol (Buchanan *et al.*, 2002).

A via de síntese global de TAG (figura 9) nas células compreende três grandes etapas: (1) carboxilação do acetil-CoA para formar malonil-CoA, sendo esta uma etapa de comprometimento com a síntese de ácidos graxos; (2) Elongamento das cadeias acil; e (3) formação de TAG (Courchesne *et al.*, 2009).



Figura 9: Via de biossíntese de ácidos graxos e triacilglicerídeos (Courchesne et al., 2009).

A análise do genoma de *Neochloris oleoabundans* fornece uma base para a compreensão do funcionamento dessas e outras vias de interesse e a possível interação entre elas, levantando indícios da origem do fenótipo que desperta interesse industrial. Os dados obtidos podem ser empregados em trabalhos futuros, visando a domesticação, o aumento de produtividade da espécie estudada e sua adequação a processos de cultivo.

## 2. OBJETIVOS

#### 2.1. Geral

Avaliar o conteúdo do genoma da microalga *Neochloris oleoabundans* com enfoque em sua alta capacidade de acúmulo de lipídios ocasionada pela privação de nitrogênio, visando à produção de combustíveis de terceira geração.

### 2.2. Específicos

- a) Sequenciamento do genoma;
- b) Montagem do genoma;
- c) Predição gênica;
- d) Análise do genoma da espécie;
- e) Levantamento de alvos para manipulação genética.

## **3. MATERIAL E MÉTODOS**

#### 3.1. Condições de crescimento

*N. oleoabundans* (UTEX 1185) foi obtida da coleção de culturas da UTEX (The Culture Collection of Algae, University of Texas, TX - USA) e cultivada no meio definido Bold's Basal Medium (Nichols e Bold 1965; tabela 1). As culturas cresceram sob constante iluminação de 100µmol fótons.m<sup>-2</sup>s<sup>-1</sup>, temperatura de 27°C e agitação de 90 rpm, sem adição de CO<sub>2</sub>. Os cultivos foram inoculados com 10% v/v de pré-inóculo.

Componente	Concentração
NaNO <sub>3</sub>	2,9mM
MgSO <sub>4</sub> <sup>-</sup> 7H <sub>2</sub> O	0,3mM
NaCl	0,4mM
KH <sub>2</sub> PO <sub>4</sub>	1,3mM
K <sub>2</sub> HPO <sub>4</sub>	0,4mM
$CaCl_2 H_2O$	0,17mM
ZnSO <sub>4</sub> ·7H <sub>2</sub> O	30,7µM
MnCl <sub>2</sub> ·4H <sub>2</sub> O	7,3 μM
MoO <sub>3</sub>	4,9 μM
CuSO <sub>4</sub> ·5H <sub>2</sub> O	6,3 μM
CoNO <sub>2</sub> <sup>-6H</sup> 2O	1,7 μM
$H_3BO_3$	0,18mM
EDTA	0,17mM
КОН	0,18mM
FeSO <sub>4</sub> ·7H <sub>2</sub> O	17,9mM
Ácido cítrico	10mM

Tabela 1: Composição do meio Bold's Basal Medium (Nichols & Bold, 1965).

#### 3.2. Extração de DNA

As células foram precipitadas por centrifugação a 2.000g durante 30 minutos e congeladas em nitrogênio líquido. O rompimento da parede e membrana celular foi feito por maceração da biomassa congelada em cadinho e pistilo de porcelana. O protocolo de isolamento do DNA foi realizado conforme descrito por Magneschi (2008). Após a maceração das células, adicionou-se 750µL de tampão SDS-EB (1% SDS; 200mM NaCl, 20mM EDTA pH 8,0; e 50mM Tris-HCl pH8,0). Foram adicionados 750µl da mistura fenol:clorofórmio (1:1) e o resultante foi misturado em vortex por 30 segundos. Depois de centrifugação a 12.000g por cinco minutos a 4°C, o sobrenadante foi misturado à um volume de clorofórmio e submetido à nova centrifugação nas mesmas condições. Foi adicionado um volume de isopropanol ao novo sobrenadante que foi reservado por 1 hora à -20°C, depois de delicadamente misturado. Findo esse tempo, o DNA foi separado do meio líquido através de centrifugação a 13.000g por 20 minutos a 4°C. O precipitado foi lavado com 500µl de etanol 70% e ressuspendido em 50µl de água MilliQ depois de seco. O DNA extraído foi quantificado em NanoDrop 2000c e sua qualidade foi avaliada em gel de agarose 1%.

#### 3.3. Sequenciamento de DNA e montagem do genoma

A montagem, predição gênica e anotação automática foram feitas pela equipe de bioinformática do LGE – Laboratório de Genômica e Expressão – UNICAMP, sendo os principais envolvidos no trabalho os bioinformatas Lucas P. Parizzi e Leandro Costa Nascimento.

Aproximadamente 23,9 milhões de *reads paired-end* com 2x76 pb foram produzidos a partir de fragmentos de 400pb de DNA utilizando o sequenciador Genome Analyzer II<sub>x</sub> (Illumina). Adicionalmente, também foram produzidos aproximadamente 70,2 milhões de *reads mate-pairs* com 2x50 pb a partir de fragmentos de DNA com 3.000 pb em um sequenciador HiSeq 2000 (Illumina). Os *reads paired-end* foram montados inicialmente utilizando o software Velvet (Zerbino *et al.,* 2008) com um k-mer de 41. A

seguir, os *reads mate-pairs* foram utilizados para a produção de *scaffolds* com o software Sspace (Boetzer, *et al.*, 2010).

Foram encontradas nos *scaffolds* sequencias possivelmente oriundas de bactérias. Suspeitando que possa ter havido contaminação bacteriana no DNA enviado para sequenciamento, uma acurada busca por similaridade com genomas de bactérias e *Chlorophytas* foi conduzida (figura 10). Os *contigs* montados foram fragmentados em pequenas partes contendo 300 pb cada e foi feita uma busca por similaridade através de um Blastn (Altschul *et al.,* 1997) contra o nt com um *e-value* de corte e<sup>-30</sup>. O tamanho escolhido para os fragmentos além de possibilitar uma similaridade significativa confere alta resolução à separação. Os fragmentos foram posteriormente separados de acordo com a maior semelhança encontrada entre os fragmentos e o banco de dados.



Figura 10: Esquema da separação dos *contigs* do genoma de alga, de bactéria e das organelas.

#### 3.4. Predição gênica

O genoma de *Neochloris oleobundans* foi anotado utilizando uma combinação de métodos: predição *ab initio* e predição comparativa. Para predição comparativa as proteínas de *Chlorella variabilis* foram alinhadas no genoma de *Neochloris* utilizando o software Exonerate (Slater & Birney, 2005). No caso da predição ab *initio*, a equipe de bioinformática rodou primeiramente o preditor Genemark (Lomsadze et al., 2005) em todas as sequências do genoma para conseguir um catálogo inicial de genes. A partir do resultado do Genemark, foram selecionadas, de maneira aleatória, 1.000 proteínas que serviram de treinamento para o preditor Augustus (Stanke *et al.*, 2006). Para combinação dos resultados da predição comparativa e das predições *ab initio* foi utilizado o EVidenceModeler (Haas *et al.*, 2008) considerando de maior importância a predição gerada pelo Augustus.

#### 3.5. Anotação e análise

O programa Autofact (Koski *et al.*, 2005) foi escolhido para a realização de uma anotação automática dos genes preditos. A principal contribuição Autofact é a capacidade de retomar a anotação com base em pesquisas de similaridade de sequencia em várias bases de dados. Para isso, foi utilizado o procedimento BLASTX (Altschul *et al.*, 1997; *e-value* de corte de 1e<sup>-5</sup>) para alinhar os genes em bancos de dados de proteínas determinadas, incluindo: banco de dados não redundante (NR) do NCBI, uniref90 e uniref100 - bancos de dados contendo conjuntos agrupados de proteínas a partir do UniProt e o KEGG - um banco de dados de vias metabólicas (Kanehisa & Goto, 2000). O banco de dados de domínios conservados (CDD) juntamente com um RPSBLAST (e-value de corte: 1e<sup>-5</sup>) foi empregado para buscar homólogos dos genes de *N. oleoabundans* nas outras espécies de microalgas empregadas na análise de genômica comparativa.

Para avaliar a composição de famílias gênicas, as proteínas de *N. oleoabundans* foram agrupados por similaridade de sequências usando o OrthoMCL (Li *et al.*, 2003) versão 1.4. Para essa análise foram utilizados os datasets de proteínas dos seguintes organismos: *Neochloris oleoabundans, Chlamydomonas reinhardtii* (Merchant *et al.*,

2007), Volvox carteri (Prochnik et al., 2010), Chlorella variabilis (Blanc et al., 2010), Ostreococcus lucimarinus (Palenik et al., 2007) e Nannochloropsis gaditana (Radakovits et al., 2012).

Para predição de localização celular, o programa SignalP (Bendtesen *et al.*, 2004) foi utilizado para identificar as proteínas de *N. oleobundans* com alta probabilidade de conter o peptídeo sinal (>= 90%). Depois, o resultado do SignalP foi filtrado usando o WoLF PSORT (Horton *et al.*, 2007).

A filogenia foi construída por máxima verossimilhança usando as sequências de rRNA nuclear 18S e 23S disponíveis já alinhadas na SILVA rRNA database (Pruesse *et al.*, 2007). Foi usado o modelo de evolução GTR com correção gamma aproximada por quatro categorias e uma proporção de sítios invariantes estimada pelo programa. Este modelo foi escolhido por meio do Critério de Informação de Akaike (AIC) usando o script PERL MrAIC (Nylander 2019, distribuído pelo autor em <u>http://www.abc.se/~nylander/mraic/mraic.html</u>. O suporte dos ramos internos foi inferido com 1000 bootstraps. A filogenia foi computada usando o programa PhyML 3.0 (Guindon *et al.*, 2010).
# 4. RESULTADOS E DISCUSSÃO

#### 4.1. Implementação da linha de pesquisa

Em 2009 o LGE (Laboratório de Genômica e Expressão) passou a empregar a experiência adquirida ao longo dos anos em manipulação genética e biologia sintética no estudo das microalgas. Os primeiros passos na implementação dessa linha de pesquisa consistiram na escolha da espécie de microalga a ser estudada e na metodologia que seria empregada.

Com apoio da literatura científica, nosso grupo de pesquisa se propôs a trabalhar com a espécie com potencial biotecnológico *Neochloris oleoabundans*. Como se trata de uma espécie pouco estudada, a microalga *Chlamydomonas reinhardtii* foi escolhida para os experimentos envolvendo manipulação molecular fina, visto que é um organismo modelo para o estudo de vários processos biológicos e possui diversos protocolos de manipulação estabelecidos (Harris, 2009). Os protocolos de interesse para os estudos do grupo disponíveis para *C. reinhardtii* foram adaptados e padronizados para a espécie estudada. Os dados de crescimento e acúmulo de lipídios relatados na literatura foram validados experimentalmente em nosso laboratório.

## 4.2. Montagem do genoma

Após o sequenciamento do DNA, a montagem do genoma evidenciou a presença de sequencias bacterianas dentre as sequencias de *Neochloris*. Para a separação dos genomas os *contigs* obtidos na montagem tiveram fragmentos de 300pb identificados por similaridade como sendo de alga, bactéria, organelas (cloroplasto e mitocôndria) e *no hits*. Um mesmo fragmento pode ter similaridade com mais de uma destas categorias. *Contigs* pequenos com fragmentos mistos foram classificados como não identificados. Todos os *scaffolds* foram comparados quanto ao seu conteúdo GC e área de cobertura. Na figura 11 é possível perceber uma clara separação entre os *contigs* de alga, bactéria e das organelas. Além disso, todos os *no hits* encontrados são pertencentes à *Neochloris* o montagem final do genoma consiste em um *draft*, apresentando diversos *gaps* que podem limitar as análises realizadas. A principal limitação é a dificuldade em precisar se os genes não encontrados estão ausentes no genoma ou contidos em um desses *gaps*.



Figura 11: Classificação dos *scaffolds* por conteúdo GC e cobertura médios.

A análise de similaridade utilizando alinhamento de nucleotídeos (blastn) indicou alta probabilidade deste DNA ser proveniente de uma bactéria pertencente (ou próxima) ao gênero *Brevundimonas*. Apesar de análises aprofundadas buscando a identificação da espécie e seu posicionamento filogenético não terem sido realizadas, os resultados de similaridade indicam bastante semelhança com *Brevundimonas sp.* BAL3 (<u>http://www.ncbi.nlm.nih.gov/genome/13676</u>).

A relação de 1:1 entre a alga e a bactéria que pode ser observada na figura 11 despertou a curiosidade do grupo, pois contaminantes presentes nas culturas costumam crescer mais rápido que a alga, apresentando um maior número relativo de organismos. Uma vez detectada a presença da bactéria, foram realizadas diversas tentativas de descontaminação da cultura. A dificuldade encontrada ao separar esses dois organismos sugere que a presença da bactéria pode não ser um evento isolado resultante de contaminação.

As sequencias de origem bacteriana foram então utilizadas em uma nova montagem, resultando em 471 scaffolds e genoma de tamanho aproximado de 3,7 Mb. Este tamanho de genoma é condizente com o observado nas outras espécies Brevundimonas, В. sequenciadas do gênero diminuta (3,24 Mb; http://www.ncbi.nlm.nih.gov/genome/2821), В. subvibrioides (3,45 Mb; http://www.ncbi.nlm.nih.gov/genome/1931) e Brevundimonas sp. BAL3 (3,64 Mb; http://www.ncbi.nlm.nih.gov/genome/13676).

Park e colaboradores observaram, em 2008, a simbiose entre uma bactéria do gênero *Brevundimona*s e a alga verde *Chlorella ellipsoidea* e, apesar da natureza da relação entre as duas espécies não ter sido explorada neste estudo, os autores sugerem que a presença da bactéria pode ajudar a fornecer cobalamina (vitamina B12) para a alga. Eles também afirmam que as associações são espécie-específicas e uma bactéria que contribui para o crescimento de uma espécie de alga, como é o caso apresentado no estudo, pode comprometer grandemente o crescimento de outra espécie. Em 2005, Croft e colaboradores realizaram um amplo estudo em que constataram que, de mais de 300 espécies de algas analisadas, 171 espécies necessitam de uma fonte externa de vitamina B12. Assim, é possível que este mesmo tipo de relação simbiótica ocorra em outros grupos de algas.

Neste contexto, imaginou-se que poderia estar ocorrendo uma associação de natureza semelhante entre *Neochloris* e a bactéria identificada em nossa análise do genoma. O estudo de Croft e colaboradores indica que a principal função da cobalamina em algas é atuar como cofator da enzima metionina sintase, sendo que existe a versão dependente e independente dessa proteína. No estudo, verificou-se que as algas que dependem da vitamina perderam a versão independente da enzima, diferentemente do que ocorre nas espécies auxotróficas. Das 154 *Chlorophytas* estudadas, apenas 49

espécies necessitam de suplementação de B12, indicando que esse não é um fenômeno comum nas algas verdes.

Ao observar que cinco espécies do gênero *Neochloris* entraram no estudo e não apresentam a dependência do cofator, concluiu-se que esse tipo de associação não deve ocorrer com a espécie sequenciada pelo nosso grupo (Croft *et al.*, 2005). Essa hipótese foi corroborada ao identificarmos cópias do gene *metE* – metionina sintase independente de cobalamina no genoma de *Neochloris oleoabundans*. A versão dependente do cofator também está presente no genoma, como ocorre nas outras espécies estudadas por Croft e colaboradores. Além disso, há famílias de alguns genes relacionados à síntese de B12 em *N. oleoabundans* (tabela 2).

Gene	Anotação
Gene7387	cobalamin synthesis protein P47K
Gene8284	cobalamin synthesis protein/P47K
Gene9529	cobalamin synthesis protein P47K
Gene867	cobalamin synthesis protein P47K
Gene7968	Cobalamin biosynthesis CobW-like protein
Gene6748	Cobalamine-independent methonine synthase, MetE, EC:2.1.1.14
Gene9430	Cobalamine-independent methonine synthase, MetE, EC:2.1.1.14
Gene7184	METH1; cobalamin-dependent methionine synthase (EC:2.1.1.13); K00548
Gene1775	METH1; cobalamin-dependent methionine synthase (EC:2.1.1.13); K00548

Tabela 2: Genes envolvidos no metabolismo de cobalamina presentes no genoma de *Neochloris oleoabundans*.

Assim, apesar da presença da bactéria identificada como *Brevundimonas* ter sido observada nas culturas de *Neochloris*, a natureza desta relação parece constituir um tópico complexo. Sua investigação pode ser objeto de estudos futuros, de forma que este tema não será aprofundado no presente trabalho.

Para as etapas seguintes de anotação e análise do genoma, apenas as sequencias provenientes de *N. oleoabundans* foram utilizadas. Os genomas de cloroplasto e mitocôndria não foram estudados nesse trabalho.

### 4.3. Predição gênica

Após os procedimentos de montagem do genoma de *N. oleoabundans*, os *scaffolds* obtidos foram analisados por três programas de predição gênica, resultando na predição final. Os principais resultados desta análise estão apresentados na tabela 3.

Estatísticas do genoma nuclear de Neochloris oleoabundans						
Tamanho estimado do genoma	40.201,792 pb (40,2 Mb)					
Número de <i>scaffolds</i>	1.544					
N50	115.608 pb (99 <i>scaffolds</i> )					
Tamanho médio do <i>scaffold</i>	26.037,43 pb					
Maior <i>scaffold</i>	619.196 pb					
Conteúdo GC (genoma)	57,32%					
Conteúdo GC (codante)	62,56%					
Estatísticas de predição gênica de Neochloris oleoabundans						
Número de genes	10.006					
Tamanho médio do gene	1.269,95 pb					
Genes anotados (banco de dados NR)	8.526 (85,22%)					
Éxons por gene	7,71					
Tamanho médio do éxon	164,64 pb					
Íntrons por gene	6,71					

Tabela 3: Estatísticas do genoma.

Durante a análise do genoma de N. oleoabundans, houve certa dificuldade em obter dados relativos a outras algas, tanto devido à falta de estudos envolvendo transcriptomas e proteomas, como pela ocorrência de poucos projetos genomas disponíveis em bancos públicos. Na data 15/07/2012, consulta ao NCBI revelou a existência de 10 completos 10 andamento genomas e em (http://www.ncbi.nlm.nih.gov/genome?term=chlorophyta) e, destes, a maioria se encontra pouco anotada. Desta forma, buscou-se avaliar os resultados da presente montagem e predição gênica com base nestes dados disponíveis. A figura 12 apresenta uma avaliação da densidade gênica observada em outras espécies de alga e sua relação com o tamanho do genoma de cada uma destas espécies.



Figura 12: Correlação do genoma de diversas microalgas com sua respectiva densidade gênica (genes/Kb). Dados calculados a partir da divisão da quantidade de genes preditos no genoma de cada uma das espécies pelo respectivo tamanho de cada genoma. Os dados foram obtidos no NCBI (http://www.ncbi.nlm.nih.gov/genome).

É possível observar na figura 12 que existe uma tendência de que a densidade gênica sofra uma redução conforme os genomas sofram um aumento de tamanho. A densidade gênica de *N. oleoabundans* (0,249 genes/Kb) está perfeitamente inserida nesta curva, o que valida a estratégia utilizada em nosso trabalho para a montagem e predição gênica desta microalga.

As espécies utilizadas nesta análise constituem não apenas algas verdes, mas também vermelhas e pardas. Observou-se que o genoma de 40 Mb de *N.oleoabundans* é quase quatro vezes maior que o das algas verdes *Ostreococcus tauri* e *O. lucimarinus* (13 Mb) e duas vezes maior que *Micromonas* (21 Mb). Tal resultado é esperado, já que o gênero *Ostreococcus* possui um genoma e um tamanho celular bastante reduzido (Palenik *et al.*, 2007).

É interessante observar que as espécies com tamanhos de genoma mais próximos de *N. oleoabundans* são *Nannochloropsis gaditana* (34 Mb) e *Chlorella variabilis* (46 Mb).

*C. variabilis* é filogeneticamente bastante próxima de *N. oleoabundans,* portanto a semelhança das características genômicas de ambas já seria esperada. O mesmo não pode ser dito de *N. gaditana,* a qual pertence a um grupo filogenético mais distante. Entretanto, como esta espécie também apresenta comportamento oleaginoso, considerou-se bastante relevante a observação da similaridade da densidade gênica entre ambas as espécies. Isto poderia indicar uma convergência entre características genômicas destas algas oleaginosas.

#### 4.4. Genômica comparativa

Para identificar novas características e tentar encontrar genes que possam justificar a condição oleaginosa da microalga estudada, foram determinados quais genes possuem ortólogos nas algas verdes *Chlorella variabilis* NC64A38 (Blanc *et al.*, 2010), *Chlamydomonas reinhardtii* (Merchant *et al.*, 2007), *Ostreococcus lucimarinus* (Palenik *et al.*, 2007) e *Volvox carteri* (Prochnik *et al.*, 2010) e na Eustigmatophyta oleaginosa *Nannochloropsis gaditana* (Radakovits *et al.*, 2012).

Essas espécies de *Chlorophytas* foram escolhidas por serem filogeneticamente mais próximas da alga estudada, conforme apresentado pela figura 13. Enquanto *C. variabilis* e *C. reinhardtii* foram incluídas por possuírem genomas mais estudados e bem anotados, as demais algas verdes foram incluídas para auxiliarem na anotação das famílias e dos genes que não são comuns entre essas espécies.



Figura 13: Filogenia das espécies utilizadas para a detecção de família gênicas em *Neochloris oleoabundans*. Os valores próximos aos ramos internos são porcentagens de *bootstraps*. O ramo de *Nannochloropsis gaditana* (pontilhado) está fora de escala.

A análise filogenética mostra um maior parentesco de *Neochloris* com *Chlamydomonas* e *Volvox* seguido por *Chlorella* e *Ostreococus*. Este maior parentesco é visto também na similaridade entre os genes destas espécies, como indicado pelos resultados de Blastp contra o NR (figura 14) das proteínas preditas de *Neochloris*.



Figura 14: As predições gênicas de *N. oleoabundans* foram comparadas com todos os genomas já sequenciados contidos no banco de dados de proteínas não-redundantes (NR) utilizando um Blastp. O número de vezes que um organismo foi o primeiro hit (e-value menor que e-5) são indicados. Os hits com *Chlorella variabilis* não aparecem no gráfico por inviabilizara visualização do mesmo.

A maior similaridade para os genes de *N. oleoabundans* é encontrada em *C. variabilis* (dado não mostrado), com 7945 genes em primeiros *hits* com *N. oleoabundans*. Em seguida aparecem *Chlamydomonas* e *Volvox*, comprovando as relações mostradas na filogenia. Isso pode sugerir que *N. oleoabundans* manteve formas ancestrais dos genes, enquanto *Chlamydomonas* e *Volvox* passaram por um processo generalizado de evolução mais acelerada. Interessantemente, *Neochloris* apresenta ainda maior semelhança com plantas terrestres do que com *Ostreococcus*. Isso pode ser devido à grande redução que ocorreu no genoma desse gênero, ocasionando um maior distanciamento dessa espécie das demais que entraram na comparação, embora deva ser considerada a maior disponibilidade de genomas de plantas terrestres.

A comparação entre os genes destas cinco espécies de alga gerou um resultado bastante peculiar. As famílias gênicas obtidas são muito pequenas, sendo que a maior família de *N. oleoabundans* possui apenas oito genes.

As outras espécies de alga apresentaram dados semelhantes, levando em consideração o tamanho do genoma de cada espécie, fato esse que levou à análise das famílias de genes ortólogos em diversas espécies de microalgas. A planta *Arabidopsis thaliana* também foi analisada para fins de comparação, uma vez que apresenta grandes famílias gênicas (Swarbreck *et al.*, 2008). A tabela 4 apresenta estes resultados.

% de genes	Ν.	С.	С.			Micromonas	V. carteri
em famílias	oleoabundans	variabillis	reinhardtii	O. lucimarinus	O. tauri	sp.	
com							
2 genes	10,0	8,7	6,3	7,5	4,5	5,7	4,3
3 genes	1,6	1,9	1,8	0,9	0,5	1,1	1,1
mais de 3	1,1	3,5	3,5	0,3	0,1	1,2	2,8
% de genes		Т.	Р.	А.	Ε.		А.
em famílias	C. merolae	pseudonana	tricornutum	anophagefferens	siliculosus	N. gaditana	thaliana
com							
2 genes	5,9	5,2	5,5	6,4	4,7	5,1	6,0
3 genes	0,8	1,5	1,4	2,1	1,2	1,1	3,5
mais de 3	1,0	1,5	0,9	2,8	3,3	0,5	13,2

Tabela 4: Porcentagem do genoma de diversas algas e *A. thaliana* contidas em famílias ortólogas classificadas por tamanho.

Como pode ser visto na tabela 4, em todas as espécies de alga analisadas a maior fração do genoma não apresenta expansões. Dentre os genes que apresentam expansão, a maioria constitui famílias com apenas dois genes. Esta condição se inverte em *A. thaliana*, em que a maioria das famílias expandidas apresenta mais de três genes. Este resultado sugere que a ocorrência de famílias pequenas pode ser uma peculiaridade de algas.

As espécies utilizadas na comparação possuem diversas famílias ortólogas comuns a todas elas. Assim, foi realizada uma comparação mais detalhada entre *Neochloris*,

*Chlamydomonas, Chlorella* e *Nannochloropsis,* cujos resultados estão apresentados na figura 15.



Figura 15: Diagrama de Venn mostrando as famílias de genes ortólogos encontradas na espécie estudada *N. oleoabundans*, na espécie filogeneticamente mais próxima *C. variabilis*, no organismo modelo *C. reinhardtii* e na alga oleaginosa *N. gaditana*.

Foram encontradas 2023 famílias presentes nas quatro espécies. Ao analisar os genes que compõem estas famílias, observou-se que as principais funções desses genes são a fotossíntese e síntese de pigmentos. Este resultado constitui uma observação esperada, uma vez que todas as espécies avaliadas são fotossintetizantes.

Após a divergência de *Chlorella, Chlamydomonas* e *Neochloris*, 2.225 famílias se mantiveram comuns apenas entre essas espécies, mostrando que embora ainda exista uma grande semelhança genética entre elas, ela não está relacionada com a capacidade de acúmulo de lipídios.

A espécie que possui o maior número de famílias exclusivas é *Chlamydomonas reinhardtii*, com 4060 famílias, seguida por *Nannochloropsis* (355 famílias), *Chlorella* (190) e *Neochloris* (108). Esta observação é esperada, uma vez que *Chlamydomonas* possui o maior genoma.

A semelhança existente entre *N. oleoabundans* e *C. variabilis* aparece mais uma vez ao observarmos 1.566 famílias em comum entre as duas espécies. Tais famílias

apresentam as mais diversas funções, mas quase 1.000 delas são formadas apenas por proteínas hipotéticas.

A figura 15 também mostra 19 famílias compartilhadas apenas por *Nannochloropsis* e *Neochloris*. A tabela 5 apresenta a descrição destas famílias.

Tabela 5: Anotação das famílias gênicas encontradas apenas em *Nannochloropsis gaditana* (N. g.) e *Neochloris oleoabundans* (N. o.) e o número de genes contido em cada família.

N.g.	N. o.	Anotação	Função
5	1	Fator de transcrição nuclear Y subunidade alpha	Fator de transcrição
3	1	Proteína hipotética	Não identificada
2	1	Proteína da família peptidase C53 - tripeptidil-peptidase I (EC:3.4.14.9) K01279	Hidrolase lisossomal
1	1	serina/treonina quinase 19 (EC:2.7.11.1) K08880	Quinase
1	1	Proteína de manutenção estrutural do cromossomo 4	Manutenção estrutural de cromossomos
1	1	Cassete de ligação ao ATP, sub-família E, membro 1	Biogênese de ribossomos
1	1	Fenilalanil-tRNA sintetase (EC:6.1.1.20) K01889	Síntese de RNA transportador
1	1	Dineína flagelar cadeia gama	Proteína de ligação ao GTP
1	1	RNA helicase DHX8/PRP ATP-dependente (EC:3.6.4.13) K12818	Splicing de RNA
1	1	Proteína de ligação ao GTP	Proteína de ligação ao GTP
1	1	Esterase / lipase (EC:3.1.1) K01066	Metabolismo de lipídios
1	1	N4-(beta-N-acetilglicosaminil)-L-asparaginase (EC:3.5.1.26) K01444	Hidrolase lisosomal
1	1	p300/CBP acetil-transferase; (EC:2.3.1.48) K04498	Ciclo celular
1	1	glutamato sintase (NADPH/NADH) small chain K00266 (EC:1.4.1.13/1.4.1.14)	Metabolismo de nitrogênio
1	1	Histona H4 K11254	Histona
1	1	Batenina; proteína cln3-like K12389	Hidrolase lisossomal
1	1	2-succinil-6-hidroxi2,4-ciclohexadieno-1-carboxilato sintase (EC 4.2.99.20)	Síntese de vitamina K2
1	1	Oxidase de álcool graxo de cadeia longa	Metabolismo de lipídios

Ao analisar as categorias funcionais representadas por estas famílias, observou-se que os genes compartilhados entre as duas algas oleaginosas apresentam funções biológicas diversas, como organização de cromossomos, regulação de processos celulares e regulação da transcrição do DNA. Em relação ao metabolismo de lipídios, o qual se esperava estar bem representado entre as famílias gênicas conservadas nas duas espécies oleaginosas, foram encontradas apenas duas famílias: uma esterase/lípase e uma oxidase de álcool graxos.

Foi encontrada também uma glutamato sintase (EC: 1.4.1.13, 1.4.1.14), enzima que utiliza o glutamato para sintetizar glutamina e  $\alpha$ -ketoglutarato, sendo o último um composto intermediário no ciclo de Krebs. A amônia é utilizada como cofator da reação que é uma das vias responsáveis pelo metabolismo de nitrogênio na célula. Essa semelhança pode estar relacionada com a resposta bastante significativa das duas espécies à privação de nitrogênio, resultando em maior acúmulo lipídico.

A presença de três famílias com função de hidrolases lisossomais é curiosa, em especial a oxidase de álcool graxo que atua na β-oxidação de ácidos graxos e a família de esterase/lípase que também está envolvida na degradação desses compostos.

A proteína hipotética encontrada deveria ser alvo de maiores estudos, uma vez que não pôde ser identificada, pois sua função poderia estar relacionada com o fenótipo de interesse.

A ausência de famílias que indiquem a natureza do comportamento comum entre as duas espécies no acúmulo de óleo pode ter ocorrido pela grande distância evolutiva existente entre elas. Houve uma convergência de fenótipo, mas não no genoma, sendo que as duas mantiveram, diferentemente das outras espécies, apenas genes relacionados com a degradação dos compostos de interesse e diversas famílias relacionadas com funções basais da célula.

A busca por indícios da causa do fenótipo de interesse passou então para as famílias gênicas exclusivas de *Neochloris oleoabundans*, ou seja, cujos ortólogos não estão presentes nas demais algas estudadas. Também foram analisadas famílias gênicas exclusivas obtidas através do banco de dados de domínios conservados (CDD), de modo que diversas outras famílias encontradas somente no genoma estudado foram identificadas.

De acordo com os resultados desta análise, foram encontradas 68 famílias exclusivas a *N. oleoabundans*, das quais apenas 12 puderam ser identificadas (tabela 6). Das famílias restantes, 20 são formadas por proteínas hipotéticas e 36 por *no hits*.

Tabela 6: Anotação das famílias gênicas encontradas apenas em *Neochloris oleoabundans* e o número de genes contido em cada família. As famílias marcadas com um asterisco (\*) foram obtidas a partir de análise do CDD e as demais foram obtidas a partir do OrthoMCL.

Número de genes	Anotação
6	serina/treonina proteína quinase
4	ATPases da classe AAA+
3	glicosiltransferase putativa
2	Quitinase
2	Hedgehog
2	RNase T2
2	Preniltransferase
2	PAS/PAC sensor protein
2	Quitosanase
2	cisteína peptidase
2	fosfolipase A2
2	6-fosfogliconolactonase (EC: 3.1.1.31)
3	CESA like *
1	Glutamato sintase (EC: 1.4.1.13)*

Dentre os resultados obtidos, foi considerada mais interessante a identificação de uma família semelhante à celulose sintase (CESA-like). O homólogo da proteína de *Arabidopsis* CESA-*like* está presente em duas cópias no genoma de *N. oleoabundans*. Essa família é membro da superfamília celulose sintase e ainda não tem sua função completamente determinada, mas em algumas plantas sabe-se que está envolvida na síntese de hemicelulose (Yin *et al.,* 2009). Este resultado sugere que a parede celular de *N. oleoabundans* poderia ser formada por esse polissacarídeo. Isto constituiria uma característica particular desta espécie, pois sabe-se que tanto *Chlamydomonas* quanto *Chlorella* não apresentam hemicelulose em suas paredes celulares (Harris, 2009; Blanc *et al.,* 2010).

Outra peculiaridade de *Neochloris* é a presença de quitinases e quitosanases, uma vez que quitina e compostos *quitina-like* são raros em paredes celulares de alga (Herth *et al.*, 1986), mas já foram observados em *Chlorella*. Nessa última, a parede celular é

formada por quitina e quitosana, sendo que ela possui tanto os genes necessários para sua síntese como para seu remodelamento (Blanc *et al.*, 2010). Curiosamente, apesar de terem sido identificados os genes de remodelamento em *Neochloris* (quitosanase e quitinase), só foi possível identificar genes que codificam para a síntese de quitina. A presença de duas cópias de quitina sintase no genoma sugere que esse polissacarídeo, assim como a hemicelulose, pode ser um constituinte da parede celular da alga estudada.

Uma vez que não foi possível identificar os genes responsáveis pela síntese de quitosana, imaginou-se qual seria a função exercida por essa enzima. Uma possibilidade é sua atuação, semelhante ao que ocorre em planta (Wang, *et al.*, 2007), como mecanismo de defesa contra outros microrganismos contidos no plâncton, como fungos e pequenos crustáceos, os quais possuem quitosana na parede celular. Deve-se ressaltar que a não observação de genes associados à síntese de quitosana não significa necessariamente a sua ausência, uma vez que o genoma possui diversos *gaps* na sua montagem. Além disso, é necessário um estudo mais detalhado dessa quitosanase para precisar sua função nas células de *Neochloris*.

Outro resultado considerado bastante interessante foi a identificação de duas cópias do gene que codifica a proteína sinalizatória Hedgehog (HH). A via de sinalização HH é fundamental para a diferenciação e proliferação celular no desenvolvimento animal. Há relatos de que mutações nessa via causam defeitos congênitos e vários tipos de câncer (Ingham *et al.*, 2011). A molécula chave dessa via é HH, um ligante secretado que atua como morfógeno. A mais difundida e bem caracterizada resposta das células à sinalização HH é a regulação positiva da transcrição de genes-alvo cuja identidade varia entre as espécies (Ingham *et al.*, 2011).

Estruturalmente, HH é uma proteína de 45kD composta por dois domínios (figura 16): o de sinalização amino-terminal e o de autoprocessamento carboxi-terminal (Bürglin, 2008). Sua atividade sinalizatória depende da porção N-terminal, denominada Hedge ou HH-N, a qual possui a pouco usual propriedade de se ligar covalentemente a uma molécula de colesterol. A porção C-terminal é denominada Hog ou HH-C, e sua única função conhecida é promover uma reação de autoprocessamento. Este domínio possui

uma região de similaridade com a sequência de inteínas de *self-splicing* denominado Hint (Ingham *et al.*, 2011), o qual está presente nos dois genes de *Neochloris*.



Figura 16: A proteína Hedgehog (HH) inteira é composta de dois domínios distintos: o domínio aminoterminal 'Hedge' (mostrado em verde escuro) e o domínio 'Hog' carboxi-terminal (mostrado em azul), ambos os quais são também encontrados em outras proteínas que não são membros da família HH (Ingham *et al.*, 2011).

A capacidade do domínio Hog de processar proteínas secretadas parece ser evolutivamente ancestral. Já foram identificados domínios Hog em diversos organismos como algas vermelhas, dinoflagelados, musgos, Alveolata e entre os metazoa, mas não em plantas superiores (Koonin, 1995). Somente eumetazoários fazem a associação dos domínios Hedge e Hog na forma em que as proteínas HH apareceram pela primeira vez (Bürglin, 2008).

Apesar de haver relatos da presença de proteínas da via HH em eucariotos ancestrais, nunca foi relatada a presença desses genes em algas verdes (Ingham *et al.*, 2011). Entretanto, a anotação automática dos genes de *N. oleoabundans* indicou a presença de vários componentes desta via de sinalização (figura 17). Para confirmar tais resultados, toda a via HH foi analisada manualmente.



Figura 17: Via de sinalização Hedgehog como foi descrita em organismos superiores. As proteínas marcadas em verde estão presentes no genoma de *N. oleoabundans* e as marcadas em branco estão ausentes. A barra dupla contínua representa a membrana plasmática e a barra cinza clara pontilhada representa a parede nuclear.

A figura 17 mostra que, embora muitos genes da via estejam presentes em *N. oleoabundans*, não foram identificados componentes-chave da sinalização, como os receptores Ptc (*Patched*) ou Smo (*Smoothened*), tampouco o inibidor HHIP (*HH-inhibitor protein*). Também não foram identificados os fatores de transcrição *zinc-finger*, como Ci ou Zic2. Apesar de não ser possível afirmar que estes genes não estão presentes no genoma, devido à presença de *gaps* na montagem, é provável que esta via não esteja funcional em *N. oleoabundans* da maneira como é descrita para organismos superiores.

Ao investigar o genoma de *Chlorella variabilis*, a espécie evolutivamente mais próxima de *N. oleoabundans*, foram encontrados os mesmos genes presentes na espécie estudada, excetuando o inibidor negativo de HH *cos2*, que está ausente em *C. variabilis*.

## 4.5. Análise funcional

A última análise conduzida utilizando os dados do genoma de *N. oleoabundans,* e uma das principais do ponto de vista de seu uso como plataforma biotecnológica na produção de biocombustíveis, foi sua reconstrução metabólica.

Em resumo, com base nas informações relativas à presença ou não de genes no genoma de *N. oleoabundans* foi possível identificar quais as vias metabólicas presentes nesta espécie. A figura 18 apresenta as principais vias de interesse deste trabalho.



Figura 18: Vias do metabolismo energético de *Neochloris oleoabundans*. As reações em vermelho representam a possível reposta à privação de nitrogênio e os números indicam as enzimas que realizam as reações detalhadas no Anexo I.

Foi possível mapear as principais vias da fixação e metabolismo de carbono, nitrogênio e da síntese e degradação de triacilglicerídeos, além de vias como fotossíntese, ciclo de Krebs, ciclo celular, mTOR, fosforilação oxidativa e via das pentose fosfato. Além disso, buscou-se identificar a ocorrência de famílias gênicas entre os genes que codificam as enzimas presentes nestas vias. A tabela apresentada no Anexo I detalha estes resultados.

Tal mapeamento permitiu comparar o número de cópias dos diversos genes envolvidos nos processos de interesse permitiu levantar vias e moléculas-chave para estudos mais detalhados.

Na figura 18 foram correlacionadas as vias da fotossíntese, fixação de carbono, glicólise, ciclo TCA, fixação de nitrogênio, síntese de ácidos graxos e síntese de TAG. Através da observação da mesma é possível compreender o caminho mais provável do carbono dentro de uma célula de *N. oleoabundans*, desde sua entrada como dióxido de carbono (CO<sub>2</sub>) ou íons bicarbonato (HCO<sub>3</sub><sup>-</sup>) dissolvidos na água, até sua incorporação em uma molécula de triacilglicerol (TAG).

Complementando esta análise com a investigação do número de cópias de cada gene (tabela 7), foi possível perceber que algumas vias, como a glicólise, possuem praticamente todos os seus genes expandidos em *Neochloris* em comparação com as outras algas analisadas. Isso pode ser um indicativo de que essa espécie possui uma grande eficiência na quebra de glicose, além de alguns dos genes contidos na via também participarem da fixação de carbono, que é outra via metabólica em que quase todos os genes envolvidos aparecem com mais cópias na espécie estudada.

			nerepeie gaa		(				
Reação	Função	КО	EC number	N.o	V. c	C.r	C.v	0.1	N.g.
2	Hexokinase	K00844	2.7.1.1	1	0	0	0	0	0
4	6-phosphofructokinase	K00850	2.7.1.11	3	2	2	2	2	0
8	glyceraldehyde 3-phosphate dehydrogenase	K00134	1.2.1.12	4	2	2	3	1	2
10	glyceraldehyde-3-phosphate dehydrogenase (NADP)	K00131	1.2.1.9	2	1	1	1	1	1
11	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	K15633	5.4.2.1	4	3	3	2	1	0
21	pyruvate carboxylase	K01958	6.4.1.1	2	1	1	1	1	1
22	NAD-dependent malate dehydrogenase	K00025	1.1.1.37	2	1	0	0	1	0
31	aconitate hydratase	K01681	4.2.1.3	1	0	0	0	0	0
37	succinate dehydrogenase flavoprotein subunit	K00234	1.3.5.1	2	1	1	1	1	1
14	phosphoenolpyruvate carboxykinase (ATP)	K01610	4.1.1.49	1	0	0	0	0	1
41	glutamate synthase (ferredoxin)	K00284	1.4.7.1	2	0	0	1	0	0
47	nitrate reductase (NADH)	K00360	1.7.1.1	4	1	2	1	3	1
50	acetyl-CoA carboxylase	K11262	6.4.1.2	6	5	3	5	1	4
51	malonyl-CoA:ACP-trans-acylase	K00645	2.3.1.39	3	2	2	2	2	1
62	glycerol-3-phosphate dehydrogenase (NAD(P)+)	K00057	1.1.1.94	3	2	2	2	2	2
64	AMP deaminase	K01490	3.5.4.6	2	1	1	1	1	0

Tabela 7: Genes das vias mostradas na figura 18 que aparecem em maior número de cópias no genoma de *Neochloris oleoabundans* (N.o) quando comparada com *Volvox carteri* (V.c.), *Chlamydomonas reinhardtii* (C.r.), *Chlorella variabilis* (C.v.), *Ostreococcus lucimarinus* (O.l.) e *Nannochloropsis gaditana* (N.g.).

Caso o maior número de cópias esteja refletindo um aumento da intensidade de expressão gênica dos genes que codificam cada uma dessas enzimas, fato esse que é impossível determinar apenas com a análise do genoma, então seria possível supor que *N. oleoabundans* apresenta uma alta capacidade de fixação do CO<sub>2</sub> atmosférico em cadeias carbônicas, além de uma alta eficiência na quebra de tais moléculas para obtenção de energia.

Além disso, essas vias fornecem os substratos para as mais diversas reações metabólicas celulares, entre elas a síntese de TAG. As duas primeiras enzimas da via de síntese de lipídios (figura 18; 50,51) são as responsáveis pelo comprometimento do acetil-CoA com a síntese de ácidos graxos, já que essa substância também participa de diversas outras reações metabólicas importantes. O fato de essas enzimas apresentarem um maior número de cópias na espécie estudada pode indicar que a alta capacidade de produção de acetil-CoA pode ser canalizada para a maior síntese de lipídios e, consequentemente, TAG. É importante ressaltar a grande distância evolutiva entre a espécie *Nannochloropsis gaditana* e a alga verde estudada. Consequentemente, muitos dos genes de *N. gaditana*, embora apresentem as mesmas funções, possuem sequencias tão distintas das presentes na espécie estudada que não permitem seu agrupamento como sequencias ortólogas. Tal fato é a razão de *N. gaditana* apresentar tão poucas cópias dos genes de síntese e ainda assim ser produtiva. Essa espécie apresenta também genes que realizam essas mesmas funções, mas são muito distintos de *Neochloris* para entrar na comparação.

#### 4.5.1. Privação de nitrogênio

Foi possível identificar a possível relação entre a privação de nitrogênio e a síntese de TAG (figura 18, vias marcadas em vermelho) a partir de um modelo descrito para leveduras oleaginosas (Rossi *et al.*, 2011). De acordo com este modelo, a limitação de nitrogênio ativa a enzima AMP deaminase que fornece amônia para a célula em privação a partir de AMP. Como consequência, a concentração de AMP mitocondrial diminui, causando queda na atividade da enzima isocitrato desidrogenase. O ciclo de TCA é então bloqueado no nível de isocitrato, que se acumula e é convertido a citrato através da aconitase. O excesso de citrato no ciclo TCA é exportado para fora da mitocôndria através do antiporte malato / citrato. A enzima citosólica ATP-citrato liase (ACL) cliva o citrato em oxaloacetato e acetil-CoA, que é posteriormente convertido em piruvato. Esse é transportado para o cloroplasto onde é direcionado para a síntese de lipídios (Ratledge & Wynn, 2002).

O fato de que os genes codantes da AMP deaminase e da aconitase estarem presentes em maior número de cópias em *Neochloris oleoabundans*, mostram que esse deve ser um caminho de sinalização importante para essa alga. Considerando que essa alga possui famílias gênicas pequenas e justamente esses genes aparecem em maior número de cópias que as demais espécies com baixo acúmulo lipídico, esse parece ser um forte indício da origem de sua alta capacidade para acúmulo de óleo.

#### 4.5.2. Remodelamento de membranas celulares

Foi demonstrado que a privação de nitrogênio, além de levar ao acúmulo de TAG, ocasiona mudanças estruturais e quebra de membranas intracelulares como as dos tilacóides e do retículo endoplasmático (Martin *et al.*, 1976; Moellering & Benning, 2010). Essas alterações foram comprovadas pelo transcriptoma de *Chlamydomonas reinhardtii* realizado em privação de nitrogênio (Miller *et al.*, 2010) e há muitos indícios de que o mesmo ocorra em *Neochloris oleoabundans*. Após algum tempo de privação, a cultura que normalmente é verde-clara passa a ter uma coloração levemente amarelada, indicando que pode estar havendo degradação de membranas e pigmentos do cloroplasto.

Há relatos de que em plantas, durante a senescência ou stress por privação de nitrogênio, membranas de tilacóides no cloroplasto são desintegradas e a clorofila e os galactolipídios são quebrados, resultando no acúmulo de intermediários tóxicos, como tetrapirroles, fitol e ácidos graxos livres (Lippold *et al.*, 2012). Cerca de 80% do nitrogênio contido em uma folha está localizado no cloroplasto e essa quebra serve para liberar esse e outros compostos para os novos órgãos da planta (Liu & Bassham, 2012). Nessas condições, uma grande proporção do fitol e dos ácidos graxos são convertidos em ácidos graxos fitilo ésteres e TAG para evitar os efeitos tóxicos de tais substâncias.

Consistente com a teoria de que o remodelamento de membranas é uma das mudanças ocasionadas pela privação de nitrogênio, no transcriptoma de *Chlamydomonas* os transcritos que tiveram maior aumento de expressão codificam genes para lipases putativas (Miller *et al.*, 2010). A presença de 66 lipases putativas encontradas no genoma de *Neochloris* sugere que esse fenômeno poderia ocorrer de maneira similar ao descrito em *Chlamydomonas* e em folhas senescentes. Infelizmente, entretanto, não é possível determinar somente com base na análise do genoma quais são as lipases envolvidas nesse processo, uma vez que as mesmas não são bem caracterizadas em microalgas e nem o transcriptoma de *Chlamydomonas* conseguiu determinar quais, dentre as muitas lipases encontradas nessa alga, atuariam no processo de quebra de membranas celulares.

Outras enzimas podem estar envolvidas com a relação entre privação de nitrogênio e o remodelamento de paredes celulares, como a acil-ACP tioesterase (FAT1) e a diacilglicerol acil-transferase (DGAT).

O gene FAT1 foi encontrado no genoma de *N. oleoabundans* e foi reportado um aumento de seus níveis de expressão em *Chlamydomonas* após a privação de nitrogênio (quase 4 vezes; Miller *et al.*, 2010). A reação catalisada por FAT1 termina a biossíntese de ácidos graxos pela separação da cadeia acil do ACP. Essa reação compete com a direta transacilação do ACP pela glicerol-3-fosfato aciltransferase para a formação de fosfatidato. Um aumento na atividade de FAT1, desse modo, poderia indicar um aumento da exportação de ácidos graxos do cloroplasto para o retículo endoplasmático, onde a montagem de TAG ocorre, já que acil-ACPs precisam ser hidrolisados antes da exportação (Pollard & Ohlrogge, 1999).

Apesar da impossibilidade de inferir os níveis de atividade dessa enzima nesse trabalho, a presença no genoma de *N. oleabundans* de um gene FAT1 completo e, portanto, potencialmente funcional, sugere que o aumento da síntese de lipídios poderia ocorrer também a partir da quebra de ácidos graxos de membrana.

O outro gene chave identificado é o que codifica DGAT, enzima responsável pela etapa de comprometimento dos ácidos graxos livres com a síntese de TAG. Este gene parece estar expandido em todas as espécies de algas estudadas, estando presente em três cópias em *Neochloris*. Dentre os cinco genes putativos identificados na versão 4.0 do genoma de *Chlamydomonas reinhardtii*, foram identificados somente quatro destes genes transcritos nos transcriptomas estudados por Miller e colegas, 2010. Interessantemente, a expressão de um desses genes foi quase completamente suprimida em condições de suplementação de N, mas mostrou um grande aumento na abundância de transcritos após a privação de nitrogênio (Miller *et al.*, 2010). Assim, apesar da expressão desta família de genes parecer ser regulada pela presença ou não de nitrogênio, a natureza desta regulação pode não depender da quantidade de cópias presentes no genoma.

Isso indica a possibilidade de que apenas um gene putativo, dos três encontrados no genoma de *N. oleoabundans*, atuaria na condição de stress fazendo com que com os

ácidos graxos livres, tanto da síntese *de novo* quanto oriundos da quebra de membranas sejam direcionados para a síntese de TAG.

#### 4.5.3. Mecanismos de fixação de carbono

Foram investigados também os mecanismos pelos quais a alga *N. oleoabundans* assimila o CO<sub>2</sub> em compostos de interesse. Os possíveis mecanismos empregados para a fixação de carbono variam entre as microalgas. Nos organismos mais estudados, como *Chlamydomonas* e *Chlorella*, o processo conta com apenas uma etapa, em que a ribulose-1,5-bifosfato é o aceptor primário do CO<sub>2</sub>, sendo convertido em duas moléculas de três carbonos (3-fosfoglicerato) pela enzima Rubisco (Kutschera & Niklas, 2007).

Este mecanismo, denominado C3, foi descrito pela primeira vez em uma espécie do gênero *Chlorella* no final dos anos 50, por Calvin e colaboradores, fazendo que o ciclo por ele descrito levasse seu nome (Calvin, 1958). Recentemente, com o aumento na quantidade de genomas sequenciados de microalgas, foi possível observar que o mecanismo C3 não é o único existente entre esses organismos. Em algumas espécies foram encontrados genes que participam de um mecanismo mais complexo de fixação de carbono, em que o CO<sub>2</sub> é convertido em uma molécula de quatro carbonos em duas etapas.

Esse mecanismo, conhecido como C4, foi descrito por Karpilov e colaboradores no começo dos anos 60 e está presente em diversas plantas superiores, como milho, cana-deaçúcar e sorgo (Karpilov, 1960). No mecanismo C4 o  $CO_2$  é inicialmente fixado em células do mesófilo em oxaloacetato, o qual é posteriormente reduzido a malato pela NADPmalato desidrogenase (30; figura 18). O malato é então transportado para as células da bainha vascular e, no cloroplasto dessas células, é decarboxilado pela enzima málica NADdependente, formando piruvato,  $CO_2$  e NADPH (Pessarakli, 2005).

Primeiramente, pensava-se haver a necessidade de separação espacial das reações que caracterizam o mecanismo C4, mas algumas espécies de microalgas aparentemente possuem um mecanismo semelhante ao descrito acima, utilizando bombas de transporte ativo de bicarbonato para concentrar o  $CO_2$  no cloroplasto, região de atuação da enzima

Rubisco. Neste caso, portanto a separação espacial ocorre entre compartimentos intracelulares, e não entre células específicas. Esse mecanismo é chamado C4-*like* e foi descrito no genoma de *N. gaditana* (Radakovits *et al.,* 2012).

Assim, buscou-se investigar se o mecanismo C4-*like* poderia ser o mecanismo de fixação de carbono em *Neochloris oleoabundans*, apesar de ser esperado encontrar apenas o mecanismo C3, já que as espécies filogeneticamente mais próximas de *N. oleoabundans* utilizam esse mecanismo. Entretanto, foram encontradas todas as enzimas necessárias para a fixação de carbono por C4-*like* descritas para *Nannochloropsis gaditana*, sugerindo ser este o mecanismo preferencialmente adotado por *N. oleoabundans*. Foram identificados tanto os genes que viabilizam a fixação de carbono por C4-*like*, como as anidrases carbônicas, enzimas responsáveis pela interconversão do CO<sub>2</sub> em bicarbonato, além de um transportador de bicarbonato no cloroplasto, entretanto não foi localizada a anidrase carbônica localizada no cloroplasto, a qual parece ser necessária para converter o bicarbonato em CO<sub>2</sub> neste compartimento e, assim, concentrando esse gás para a Rubisco.

Sendo assim, não é possível afirmar somente com a análise do genoma qual o mecanismo mais empregado por *Neochloris*. Esta espécie de alga poderia empregar o mecanismo C3 e utilizar os genes de C4 como anapleróticos (reposição) dos intermediários do ciclo TCA, conforme descrito em *Chlamydomonas* (Reinfelder *et al.*, 2004), ou ainda poderia utilizar ambos os mecanismos, C3 e C4, sendo C3 o principal e C4 em baixas concentrações de CO<sub>2</sub>, como no caso descrito em *Ostreococcus* (Xu *et al.*, 2012). Poderia ainda apresentar exclusivamente o mecanismo C4, como descrito em *N. gaditana*, ou exclusivamente C3 como descrito em *Chlorella*.

É importante investigar esse mecanismo a fundo, pois plantas C4 são uma sofisticação evolutiva de C3, garantindo altas taxas de fotossíntese mesmo quando a concentração de CO<sub>2</sub> está baixa (Pessarakli, 2005). Outra vantagem seria o baixo índice de fotorrespiração, que ocorre quando a Rubisco, considerada uma enzima "promíscua" (pois pode atuar tanto como carboxilase como oxidase), fixa oxigênio em detrimento de gás carbônico. Como no mecanismo C4 há concentração de CO<sub>2</sub> no local de atuação da

Rubisco, ocorre uma redução natural da competição com O<sub>2</sub>, diminuindo assim a possibilidade de ocorrer a fotorrespiração (Xu *et al.,* 2012; Pessarakli, 2005). Nesse contexto, o uso de engenharia genética é de extrema importância para direcionar uma maquinaria celular já existente no genoma de *Neochloris* para seu uso de forma mais eficiente na forma do mecanismo C4-*like*.

# **5. CONCLUSÕES E PERSPECTIVAS**

As microalgas são consideradas uma fonte promissora de matéria-prima para a produção de biocombustíveis. Recentemente, a diminuição das reservas petrolíferas e o aumento dos níveis atmosféricos de CO<sub>2</sub> despertaram o interesse e aumentaram as pesquisas envolvendo esses organismos. Um dos principais desafios é encontrar uma espécie capaz de concentrar todas as características de interesse, como grande acúmulo de biomassa e alta produção de lipídios. Para superá-lo será necessário empregar técnicas de engenharia genética para manipular a bioquímica desses organismos em favor do aumento da produtividade e maior adequação às plataformas de cultivo.

Por ser naturalmente capaz de grande acúmulo lipídico em resposta à privação de nitrogênio e possuir um perfil de ácidos graxos ideal para a produção de biodiesel, *Neochloris oleoabundans* é uma espécie de grande interesse biotecnológico para a produção de biocombustíveis de terceira geração. A caracterização de seu genoma permitiu levantar insumos para a compreensão das bases moleculares do fenótipo de interesse. Foram identificados os genes envolvidos no metabolismo energético, no acúmulo de lipídios e na resposta à privação de nitrogênio, permitindo levantar genes-chave que podem ser objeto de estudos mais detalhados ou até mesmo de manipulação genética.

A montagem do genoma gerou 10.006 genes preditos distribuídos ao longo de 1.544 *scaffolds*. Devido à presença de vários *contigs* muito fragmentados, a montagem atual apresenta *gaps*, que devem ser fechados para que o genoma deixe de ser um draft e

seja possível precisar que os genes não encontrados estão efetivamente ausentes no genoma de *Neochloris*.

O estudo do genoma também mostrou que, embora *Neochloris* seja filogeneticamente mais próxima de C. *reinhardtii* e *V. carteri*, seu genoma se assemelha mais com o de *C. variabilis*. As espécies mais próximas devem ter sofrido uma evolução global mais rápida, enquanto *Neochloris* parece ter mantidos genes ancestrais semelhantes aos de *Chlorella*.

Foram encontradas 18 famílias gênicas ortólogas presentes apenas nas espécies oleaginosas *Nannochloropsis gaditana* e *Neochloris oleoabundans*. Entre elas, apenas duas estão relacionadas com o metabolismo de lipídios, uma esterase/lipase e uma oxidase de álcool graxos, ambas participando de sua degradação de ácidos graxos. As poucas semelhanças entre essas espécies se deve a grande distância filogenética, muito embora apresentem a uma resposta similar à privação de nitrogênio.

A análise das famílias gênicas exclusivas de *Neochloris* deram indícios de que sua parede celular pode ser composta por hemicelulose e quitina. Este último é um componente pouco usual de paredes celulares de alga, estando presente apenas em *Chlorella*.

As vias metabólicas envolvidas no fenótipo de interesse foram integralmente mapeadas e foi possível identificar genes chave em maior número de cópias, indicando a natureza das reações que ocasionam o maior acúmulo lipídico e a ampla resposta à privação de nitrogênio. Para que as hipóteses sugeridas nesse trabalho sejam confirmadas faz-se necessária a realização de um transcriptoma, para verificar se o maior número de cópias desses genes acarreta uma maior expressão gênica, indicando se as vias propostas no presente trabalho são as responsáveis pela resposta de *N. oleoabundans* à privação de nitrogênio.

Através dos dados levantados na análise do genoma de *Neochloris* pode-se concluir que o aumento significativo no acúmulo de TAG decorrente da privação de nitrogênio possivelmente ocorre pela queda de AMP celular, pela ação da AMP deaminase, direcionando um intermediário do ciclo TCA para a síntese lipídica. Além disso, pode

ocorrer também um remodelamento de membranas do cloroplasto e quebra de pigmentos fotossintetizantes, como a clorofila. Como essa quebra pode liberar substâncias tóxicas para as células, as mesmas são neutralizadas tornando-se TAG sem a necessidade do uso de poder redutor da célula para que esse aumento aconteça.

Diante da existência de tantos mecanismos distintos para a fixação de carbono, bem como sua importância na definição da eficiência energética dos organismos fotossintetizantes, é bastante relevante a compreensão deste metabolismo em microalgas a serem utilizadas como plataformas biotecnológicas. Assim, este tópico é indicado como um assunto a ser objeto de estudos aprofundados em futuros projetos. Devido às vantagens apontadas acima, caso *N. oleoabundans* atue como C3, tanto exclusivamente como de forma complementar ao mecanismo C4, este seria um alvo importante à sua manipulação genética. Neste caso, as ferramentas de transformação genética poderiam atuar no sentido de direcionar esta alga para um metabolismo majoritariamente C4 já que a maior eficiência fotossintética pode ocasionar maior biomassa, maior produção de cadeias carbônicas e, consequentemente, aumento na sua capacidade potencial de acumular lipídios. Há casos na literatura (Sheriff, *et al.*, 1998) de tentativas de manipular geneticamente plantas superiores C3 para que passem a atuar como C4, diminuindo assim as perdas devido à promiscuidade da Rubisco e aumentando a eficiência fotossintética.

# **6. REFERÊNCIAS BIBLIOGRÁFICAS**

- Abbasi, T., Premalatha, M., & Abbasi, S. A. (2011). The return to renewables: Will it help in global warming control? Renewable and Sustainable Energy Reviews, 15(1), 891-894.
- Alper, H. & Stephanopoulos, G. (2009). Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential?. *Nature Reviews Microbiology* 7, 715-723; doi:10.1038/nrmicro2186.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.
- Beal, C. M., M. E. Webber, *et al.* (2010). Lipid analysis of *Neochloris oleoabundans* by liquid state NMR. Biotechnology and Bioengineering 106(4): 573-583.
- Bendtsen, J. D., Nielsen, H., Heijne, G., Brunak, S. (2004). Improved Prediction of Signal Peptides: SignalP 3.0. Journal of Molecular Biology. 340 (4): 783–795.
- Benemann, J. R. (2008). Opportunities and Challenges in Algae Biofuel Production. Algae World 2008 (p. 15). Singapore.
- Blanc, G., et al. (2010). The Chlorella variabilis NC64A Genome Reveals Adaptation to Photosymbiosis, Coevolution with Viruses, and Cryptic Sex. The Plant Cell, Vol. 22: 2943–2955.
- Boetzer, M., Henkel, C. V., Jansen, H. J., Butlerand, D., Pirovano, W. (2010) Scaffolding preassembled contigs using SSPACE. Bioinformatics; btq683.
- Borowitzka, M .A. (2006). Microalgae as sources of pharmaceuticals and other biologically active compounds. Journal of Applied Phycology. 7(1): 3-15.
- Boudghene Stambouli, A., & Traversa, E. (2002). Fuel cells, an alternative to standard sources of energy. Renewable and Sustainable Energy Reviews, 6(3), 295-304.
- Briggs, M. (2004). Widescale Biodiesel Production from Algae. Disponível em: <a href="http://www.unh.edu/p2/biodiesel/article\_alge.html">http://www.unh.edu/p2/biodiesel/article\_alge.html</a>.
- Buchanan, B., Gruissem, W., Jones, R. (2002). Biochemistry & Molecular Biology of Plants. Plant Physiol, 130: 740-756.
- Bürglin, T. R. (2008). Evolution of hedgehog and hedgehog-related genes, their origin from Hog proteins in ancestral eukaryotes and discovery of a novel Hint motif. BMC Genomics, 9:127.
- Calvin, M. (1958). From Microstructure and Function in the Photochemical Apparatus. Lawrence Berkeley National Laboratory.

- Chen, W., C. Zhang, et al. (2009). "A high throughput Nile red method for quantitative measurement of neutral lipids in microalgae." Journal of Microbiological Methods 77(1): 41-47.
- Chisti, Y. (2007). Biodiesel from microalgae. Biotechnol.Adv. 25(3): 294-306.

Contribution of working group 3 to the Fourth Assessment Report. Paris.

- Courchesne, N. M. D., Parisien, A., Wang, B., Lan, C.Q. (2009). Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches. Journal of Biotechnology 141 (31–41).
- Croft, M. T., Lawrence, A. D. *et al.* (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438(7064): 90-93.
- da Silva, T., Reis, A., *et al.* (2009). Oil Production Towards Biofuel from Autotrophic Microalgae Semicontinuous Cultivations Monitorized by Flow Cytometry. Applied Biochemistry and Biotechnology 159(2): 568-578.
- European Standard EN 14214 (2004) Automotive fuels—fatty acid methyl esters (FAME) for diesel engines—requirements and test methods. AFNOR, Saint-Denis
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281, 237-240.
- Gomez, L. D., Clare, G. S., McQueen-Mason, J. (2008). Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol;178:473–85.
- Gouveia, L., Marques, A. E., da Silva, T. L. (2009) *Neochloris oleabundans* UTEX #1185: a suitable renewable lipid source for biofuel production. J Ind Microbiol Biotechnol 36:821–826.
- Griffiths, M. J. & Harrison, S. T. L. (2009). Lipid productivity as a key characteristic for choosing algal species for biodiesel production, J Appl Phycol 21:493–507.
- Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology, 59(3):307-21.
- Guschina, I. A. and Harwood, J. L. (2006). Lipids and lipid metabolism in eukaryotic algae. Progress in Lipid Research 45(2): 160-186.
- Haas *et al.* (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology* 2008, **9**:R7.
- Hallmann, A. (2007). Algal Transgenics and Biotechnology. Transgenic Plant Journal. 1(1): 81-98.
- Harris, E. H. (2009). The Chlamydomonas Sourcebook (2nd ed., p. 2032). Elsevier Inc.
- Harwood, J. L., Guschina, I. A. (2009). The versatility of algae and their lipid metabolism. Biochimie.91(6):679-84.

- Herth, W., Mulisch, M., and Zugenmaier, P. (1986). Comparison of chitin fibril structure and assembly in three unicellular organisms. *In* Chitin in Nature and Technology (R. Muzzarelli, C. Jeuniaux, and G. W. Gooday, Eds.), pp. 107±120. Plenum, New York.
- Horton, P., Park, K., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J. and Nakai, K. (2007).
  WoLF PSORT: protein localization predictor. Nucl. Acids Res. 35 (suppl 2): W585-W587.doi: 10.1093/nar/gkm259.
- Huang, Y. Y., C. M. Beal, *et al.* (2010). Micro-Raman spectroscopy of algae: Composition analysis and fluorescence background behavior. Biotechnology and Bioengineering 105(5): 889-898.
- Ingham, P. W., Nakano, Y. and Seger, C. (2011). Mechanisms and functions of Hedgehog signalling across the metazoan. *Nature reviews Genetics* 12: 393.
- IPCC, Intergovernmental Panel on Climate Change (2007). Climate change 2007: mitigation.
- Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res28:27-30.
- Karpilov, Y. S. (1960). The distribution of radioactive carbon 14 amongst the products of photosynthesis of maize. Trudy Kazansk Sel'shokoz Institute 41: 15–24.
- Koonin, E. V. (1995). A protein splice-junction motif in hedgehog family proteins. *Trends Biochem. Sci.* 20, 141–142.
- Koski, L. B., Gray, L. W., Lang, B. F. and Burger, G. (2005). AutoFACT: An automatic functional annotation and classification tool. BMC Bioinformatics 6:e151.
- Kutschera, U., Niklas, K. J. (2007) Photosynthesis research on yellowtops: Macroevolution in progress. Theory in Biosciences 125 81–92.
- Kutschera, U., Niklas, K. J. (2007). Photosynthesis research on yellowtops: Macroevolution in progress. Theory in Biosciences 125 81–92.
- Laursen W. (2006). Students take a green initiative. Chem Eng;32–4.
- Li, Y., Horsman, M., *et al.* (2008). Effects of nitrogen sources on cell growth and lipid accumulation of green alga *Neochloris oleoabundans*. Applied Microbiology and Biotechnology 81(4): 629-636.
- Li, Y., Horsman, M., Wang, B., Wu, N., Lan, C. Q. (2008) Effects of nitrogen sources on cell growth and lipid accumulation of green alga *Neochloris oleoabundans*. Appl Microbiol Biotechnol 81:629–636
- Li,L., Stoeckert, Jr.,C. J. and Roos, D. S. (2003) OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Res. 13:2178-2189.
- Lippold, F., *et al.* (2012). Fatty Acid Phytyl Ester Synthesis in Chloroplasts of *Arabidopsis*. The Plant Cell. 24:2001-2014.

- Liu, Y. and Bassham, D. C. (2012). Autophagy: Pathways for Self-Eating in Plant Cells. Annu. Rev. Plant Biol. 63:215–37.
- Lomsadze *et al.* (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. Nucleic Acids Research, 2005, Vol. 33, No. 20.
- Mabee, W. E., Gregg, D. J., Saddler, J. N. (2005) Assessing the emerging biorefinery sector in Canada. Appl Biochem Biotechnol;121–124:765–78.
- Magneschi, L. (2008). *Chlamydomonas* RNA and DNA extraction. Acessado em 11 June 2011, 2011, from <u>http://www.plantlab.sssup.it/chlamydomonas-protocols</u>.
- Magneschi, L. (2008). Chlamydomonas RNA and DNA extraction. Retrieved June 11, 2011, from
- Martin, N.C., Chiang, K. S., Goodenough, U. W. (1976). Turnover of chloroplast and cytoplasmic ribosomes during gametogenesis in *Chlamydomonas reinhardtii*. Dev Biol 51: 190–201.
- Merchant, S S., *et* al. (2007). The Chlamydomonas Genome reveals the Evolution of Key Animal and Plant Functions. Science 318, 245; DOI: 10.1126/science.1143609.
- Metting Jr., F.B. (2005). Biodiversity and application of microalgae. Journal of Industrial Microbiology and Biotechnology. 17(5-6): 477-489.
- Miller, R., *et al.* (2010). Changes in Transcript Abundance in *Chlamydomonas reinhardtii* following Nitrogen Deprivation Predict Diversion of Metabolism. Plant Physiol. Vol. 154.
- Moellering. E.R., Benning, C. (2010). RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. Eukaryot Cell 9: 97–106.
- Murray, K. E., Healy, F. G., McCord, R. S. and Shields, J. A. (2011) Biomass production and nutrient uptake by *Neochloris oleoabundans* in an open trough system. Appl Microbiol Biotechnol 90:89–95.
- Murray, K., Healy, F., *et al.* (2011). Biomass production and nutrient uptake *by Neochloris oleoabundans* in an open trough system. Applied Microbiology and Biotechnology 90(1): 89-95.
- Naik, S. N., Goud, V.V., Rout, P. K. and Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review. Renewable Sustainable Energy Rev., 14: 578-597.
- Nichols, H. W. and H. C. Bold (1965). "Trichosarcina polymorpha Gen. et Sp. Nov." Journal of Phycology 1(1): 34-38.
- Osamu, K., Carl, H. W. (2007). Biomass Handbook. Gordon Breach Science Publisher;1989.
- Palenik, B., et AL (2007). The tiny eukaryote Ostreococcus provides genomic insights into the paradox of plankton speciation. PNAS 104:18.
- Park, Y., Je, K., Lee, K., Jung, S., Choi, T. (2008). Growth promotion of *Chlorella ellipsoidea* by coinoculation with *Brevundimonas sp.* isolated from the microalga. Hydrobiologia 598(1): 219-228.

- Pauly, M., Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J; 54:559–68.
- Pessarakli, M. (2005). Handbook of Photosynthesis, second edition.
- Pollard, M., Ohlrogge, J. (1999. Testing models of fatty acid transfer and lipid synthesis in spinach leaf using in vivo oxygen-18 labeling. Plant Physiol. 121: 1217–1226.
- Pruesse, E., C. Quast, K. Knittel, B. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Opens external link in new window. Nuc. Acids Res. 2007; Vol. 35, No. 21, p.7188-7196.
- Pruvost, J., Legrand, J., Legentilhomme, P., Muller-Feuga, A. (2002). Simulation of Microalgae Growth in Limiting Light Conditions: Flow Effect. AIChE Journal. 48(5): 1109-1120.
- Pruvost, J., Van Vooren, G., *et al.* (2009). Investigation of biomass and lipids production with *Neochloris oleoabundans* in photobioreactor. Bioresource Technology 100(23): 5988-5995.
- Pruvost, J., Van Vooren, G., *et al.* (2011). Systematic investigation of biomass and lipid productivity by microalgae in photobioreactors for biodiesel application. Bioresource Technology 102(1): 150-158.
- Radakovits, R., et al. (2012). Draft genome sequence and genetic transformation of the oleaginous alga Nannochloropis gaditana. Nature Communications 3:686.
- Ratledge, C. & Wynn, J. P. (2002). The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Adv Appl Microbiol* 51, 1–51.
- Ratledge, C., Cohen, Z. (2008). Microbial and algal oils: Do they have a future for biodiesel or as commodity oils? Lipid Technology. 20(7): 155-160.
- Raupach, M. et al. (2007). Global and regional drivers of accelerating CO2 emissions. PNAS. 104(24): 10288–10293.
- Raven, P., Evert, R., Eichhorn, S. (1992). Biology of plants. Worth Publishers, Inc., New York, NY, USA
- Reinfelder, J. R., Milligan, A. J., Morel, F. M. M. The Role of the C4 Pathway in Carbon Accumulation and Fixation in a Marine Diatom. Plant Physiology, August 2004, Vol. 135, pp. 2106–2111.
- Reinfelder. J. R. (2011). Carbon Concentrating Mechanisms in Eukaryotic Marine Phytoplankton. Annual Review of Marine Science Vol. 3: 291-315.
- Rocha, J.M., Garcia, J.E., Henriques, M.H. (2003). Growth aspects of the marine microalga *Nannochloropsis gaditana*. Biomol.Eng. 20(4-6): 237-242.
- Rodolfi, L., Chini, Z.G., Bassi, N., Padovani, G., Biondi, N., Bonini, G., Tredici, M.R. (2008). Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. Biotechnol.Bioeng.

- Rosenberg, J. N., Oyler, G. A., Wilkinson, L., & Betenbaugh, M. J. (2008). A green light for engineered algae: redirecting metabolism to fuel a biotechnology revolution. Current Opinion in Biotechnology, 19, 430-436.
- Rossi, M., Amaretti, A., Raimondi, S. and Leonardi, A. (2011). Getting Lipids for Biodiesel Production from Oleaginous Fungi. *Biodiesel – Feddstocks and Processing Technologies, ISBN: 978-953-307-713-0.*
- Sheehan, J., Dunahay, T., Benemann, J., Roessler, P. (1998). A Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae Relatório nº 1. pp. 1-328.
- Sheriff, A., Meyer, H., Riedel, E., Schmitt, J. M., Lapke, C. (1998) The influence of plant pyruvate, orthophosphate dikinase on a C 3 plant with respect to the intracellular location of the enzyme. Plant Sci.; 136:43–47.
- Simon E. Prochnik. S. E., *et al.* (2010). Genomic Analysis of Organismal Complexity in the Multicellular Green Alga Volvox carteri. Science 329 (5988): 223-226.
- Slater and Birney (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 2005, 6:31.
- Spolaore, P., Joannis-Cassan, C., Duran, E., Isambert, A. (2006). Commercial applications of microalgae. J.Biosci.Bioeng. 101(2): 87-96.
- Stanke *et al.* (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* 2006, **7**:62.
- Stanke, M., M. Diekhans, et al. (2008). "Using native and syntenically mapped cDNA alignments to improve de novo gene finding." Bioinformatics 24(5): 637-644.
- Swarbreck, D., *et al.* (2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. Nucl. Acids Res. 36(suppl 1): D1009-D1014.
- Teixeira, C.M., Teixeira, P.C., Rocha, H., Almeida, A.G., BRITO, G.F. (2007). Um Novo Sistema de Cultivo de Microalgas para a Produção de Biodiesel. II Congresso da Rede Brasileira de Tecnologia de Biodiesel.
- Tornabene, T. G., Holzer, G., Lien, S., Burris, N. (1983). Lipid composition of the nitrogen starved green alga Neochloris oleoabundans. Enzyme Microb Technol 5: 5.
- Tornabene, T.G., Holzer, G., Lien, S., Burris, N. (1983) Lipid composition of the nitrogen starved green alga *Neochloris oleabundans*. Enzyme Microb Technol 5:435–440.
- Tredici, M. R., Materassi, R. (1992). From open ponds to vertical alveolar panels: the Italian experience in the development of reactors for the mass cultivation of phototrophic microorganisms. Journal of Applied Phycology (3): 221-231.
- Wahal, S. and Viamajala S. (2010). Maximizing Algal Growth in Batch Reactors Using Sequential Change in Light Intensity. Applied Biochemistry and Biotechnology 161(1): 511-522.
- Walker, T., Purton, S., et al. (2005). Microalgae as bioreactors. Plant Cell Reports 24(11): 629-641.

- Wang, B. and Lan, C. Q. (2011). Biomass production and nitrogen and phosphorus removal by the green alga *Neochloris oleoabundans* in simulated wastewater and secondary municipal wastewater effluent. Bioresource Technology 102(10): 5639-5644.
- Wang, Y., *et al.*, (2007). Antimicrobial effect of Chitooligosaccharides Produced by Chitosanase from *Pseudomonas* CUY8. Asia Pac J Clin Nutr 2007;16 (Suppl 1):174-177.
- Worldwatch Institute. Disponível em: http://www.greencarcongress.com/2011/08/wwi-20110831.html#more.
- Xu J, Fan X, Zhang X, Xu D, Mou S, et al. (2012) Evidence of Coexistence of C 3 and C 4 Photosynthetic Pathways in a Green-Tide-Forming Alga, *Ulva prolifera*. PLoS ONE 7(5): e37438. doi:10.1371/journal.pone.0037438.
- Xu, J., Fan, X., Zhang, X., Xu, D., Mou, S., et al. (2012) Evidence of Coexistence of C<sub>3</sub> and C₄Photosynthetic Pathways in a Green-Tide-Forming Alga, Ulva prolifera. PLoS ONE 7(5): e37438. doi:10.1371/journal.pone.0037438.
- Yin,Y., Huang, J. and Xu, Y. (2009). The cellulose synthase superfamily in fully sequenced plants and algae. *BMC Plant Biology*, **9**:99.
- Zabaniotou, A., Ioannidou, O., Skoulou, V. (2008). Rapeseed residues utilization for energy and 2nd generation biofuels. Fuel;87:1492–502.
- Zerbino, D. R. & Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.
- Zerbino, D. R., McEwen, G. K., Margulies, E. H. & Birney, E. (2009) Pebble and Rock Band: heuristic resolution of repeats and scaffolding in the Velvet short-read de novo assembler. PLoS ONE 4, 8407.
- Zhang, Y. C., Rossow, W.B., Stackhouse Jr., P.W. (2007). Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the surface. J. Geophys. Res., 112.

# **ANEXO I**

Anexo I: Enzimas que catalisam as reações mostradas na figura X, sendo que o número da reação representa o número mostrado na figura, seguido do nome da enzima, sua identificação no KEGG, seu EC number e o número de cópias presentes no genoma de *Neochloris oleoabundans* (N.o) quando comparada com *Volvox carteri* (V.c.), *Chlamydomonas reinhardtii* (C.r.), *Chlorella variabilis* (C.v.), *Ostreococcus lucimarinus* (O.l.) e *Nannochloropsis gaditana* (N.g.).

Reação	Função	КО	EC number	N.o	V. c	C.r	C.v	0.1	N.g.
1	phosphoglucomutase	K01835	5.4.2.2	3	1	3	2	1	1
2	hexokinase	K00844	2.7.1.1	1	0	0	0	0	0
3	glucose-6-phosphate isomerase	K01810	5.3.1.9	1	1	1	1	1	1
4	6-phosphofructokinase	K00850	2.7.1.11	3	2	2	2	2	0
5	fructose-1,6-bisphosphatase I	K03841	3.1.3.11	1	0	0	1	0	0
6	fructose-bisphosphate aldolase, class I	K01623	4.1.2.13	6	4	6	3	3	4
7	triosephosphate isomerase (TIM)	K01803	5.3.1.1	3	1	1	2	3	1
8	glyceraldehyde 3-phosphate dehydrogenase	КОО134	1.2.1.12	4	2	2	3	1	2
10	glyceraldehyde-3-phosphate dehydrogenase (NADP)	K00131	1.2.1.9	2	1	1	1	1	1
9	phosphoglycerate kinase	К00927	2.7.2.3	2	2	2	3	1	3
11	2,3-bisphosphoglycerate-independent	K15633	5.4.2.1	4	3	3	2	1	0
12	enolase	K01689	4.2.1.11	1	1	1	1	1	2
13	pyruvate kinase	K00873	2.7.1.40	6	6	6	7	2	1
14	phosphoenolpyruvate carboxykinase (ATP)	K01610	4.1.1.49	3	6	1	3	0	1
15	ribulose-phosphate 3-epimerase	K01783	5.1.3.1	1	1	2	2	3	0
16	ribose 5-phosphate isomerase A	K01807	5.3.1.6	1	0	0	0	1	0
17	phosphoribulokinase	K00855	2.7.1.19	1	1	1	1	1	1
18	ribulose-bisphosphate carboxylase large chain	K01601	4.1.1.39	1	-	-	-	-	-
19	carbonic anhydrase	K01673	EC:4.2.1.1	6	7	8	6	0	2
20	phosphoenolpyruvate carboxylase	К01595	4.1.1.31	3	2	4	2	1	0
21	pyruvate carboxylase	К01958	6.4.1.1	2	1	1	1	1	1
22	NAD-dependent malate dehydrogenase	КООО25	1.1.1.37	2	1	0	0	1	0
22	malate dehydrogenase	К00026	1.1.1.37	3	3	3	3	3	0
23	malate dehydrogenase (NAD $^{\scriptscriptstyle +}$ )	КООО28	1.1.1.39	1	-	-	-	-	-
24	pyruvate dehydrogenase E1 component	K00162	1.2.4.1	5	5	4	5	3	2
25	pyruvate dehydrogenase E2 component	K00627	2.3.1.12	3	2	4	1	3	3
26	dihydrolipoamide dehydrogenase	K00382	1.8.1.4	1	1	1	1	1	1
----	---	--------	---------------------	---	---	---	---	---	---
27	citrate synthase	K01647	2.3.3.1	2	3	2	1	1	2
28	ATP citrate (pro-S)-lyase	K01648	2.3.3.8	1	1	1	1	0	1
30	malate dehydrogenase (NADP+)	K00051	1.1.1.82	1	1	1	1	1	0
29	citrate lyase, citryl-ACP lyase (beta) subunit	K01644	4.1.3.6 4.1.3.34	1	0	0	1	0	0
31	aconitate hydratase	K01681	4.2.1.3	1	0	0	0	0	0
32	isocitrate dehydrogenase (NAD+)	K00030	1.1.1.41	1	1	1	1	0	0
33	2-oxoglutarate dehydrogenase E1 component	K00164	1.2.4.2	1	0	0	1	3	0
34	2-oxoglutarate dehydrogenase E2 component	K00658	2.3.1.61	1	0	1	0	1	3
35	dihydrolipoamide dehydrogenase	K00382	1.8.1.4	1	1	1	1	1	1
36	succinyl-CoA synthetase alpha subunit	K01899	6.2.1.4/	1	1	1	1	1	1
			6.2.1.5						
36	succinyl-CoA synthetase beta subunit	K01900	6.2.1.4/ 6.2.1.5	1	1	1	1	1	1
37	succinate dehydrogenase flavoprotein subunit	K00234	1.3.5.1	2	1	1	1	1	1
38	fumarate hydratase, class II	K01679	4.2.1.2	1	1	1	1	1	1
39	malate dehydrogenase (NADP+)	K00029	1.1.1.40	5	2	7	2	5	2
40	glutamate synthase (NADPH/NADH)	K00264	1.4.1.13/	2	2	2	2	1	2
			1.4.1.14						
41	glutamate synthase (ferredoxin)	K00284	1.4.7.1	2	0	0	1	0	0
42	asparagine synthase (glutamine- hydrolysing)	K01953	6.3.5.4	1	1	1	1	2	1
43	glutamate dehydrogenase (NAD(P)+)	K00261	1.4.1.3	2	2	2	2	0	0
44	glutamate dehydrogenase (NADP+)	K00262	1.4.1.4	2	0	0	2	0	1
45	glutamine synthetase	K01915	6.3.1.2	2	4	4	2	0	1
46	ferredoxin-nitrite reductase	K00366	1.7.7.1	1	1	1	1	1	0
47	nitrate reductase (NADH)	K00360	1.7.1.1	4	1	2	1	3	1
48	3-oxoacyl-ACP synthase III	K00648	2.3.1.180	1	1	1	1	1	0
49	3-oxoacyl-ACP synthase II	K09458	2.3.1.179	2	2	2	2	3	0
50	acetyl-CoA carboxylase	K11262	6.4.1.2	6	5	3	5	1	4
51	malonyl-CoA:ACP-trans-acylase	K00645	2.3.1.39	3	2	2	2	2	1
52	3-oxoacyl-ACP reductase	K00059	1.1.1.100	2	2	2	1	1	1
53	3-hydroxy acyl-CoA dehydratase	K02372	4.2.1	1	1	1	1	1	1
54	enoyl-ACP reductase I	K00208	1.3.1.9	1	1	1	1	1	0

55	acyl-ACP thioesterase	K10782	3.1.2.14	1	1	1	1	1	0
56	glycerol-3-phosphate acyltransferase	K00629	2.3.1.15	1	0	0	1	0	0
56	glycerol-3-phosphate O-acyltransferase	K00630	2.3.1.15	1	2	1	0	1	0
57	1-acylglycerol-3-phosphate O- acyltransferase-	K00655	2.3.1.51	2	2	2	2	2	0
58	phosphatidate phosphatase	K01080	3.1.3.4	3	3	2	3	1	0
59	diacylglycerol kinase	K00901	2.7.1.107	2	1	2	2	2	2
60	diacylglycerol O-acyltransferase	K11155	2.3.1.20	3	5	5	5	3	9
61	glycerol-3-phosphate dehydrogenase	K00111	1.1.5.3	1	1	1	1	1	1
62	glycerol-3-phosphate dehydrogenase (NAD(P)+)	K00057	1.1.1.94	3	2	2	2	2	2
63	glycerol-3-phosphate dehydrogenase	K00006	1.1.1.8	1	1	3	1	0	0
64	AMP deaminase	K01490	3.5.4.6	2	1	1	1	1	0

## DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado intitulada "Análise do genoma da microalga produtora de lipídeos *Neochloris oleabundans*, visando à produção de combustíveis de terceira geração"

( ) não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

(X) CIBio – Comissão Interna de Biossegurança, projeto No.2011/03, Instituição: Instituto de Biologia (IB/UNICAMP).

) CEUA – Comissão de Ética no Uso de Animais , projeto No. \_\_\_\_\_, Instituição:

) CEP - Comissão de Ética em Pesquisa, protocolo No. \_\_\_\_\_, Instituição:

\* Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.

Aluno: (Bruna Roncon Favarelli)

Orientador: (Gonçalo Amarante Quimahaes Pereira)

Para uso da Comissão ou Comitê pertinente:

Carimbo e assinatura

Prof. Dr. MARCELO LANCELLOTTI Presidente da Comissão Interna de Biossegurança Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente: () Deferido () Indeferido

Carimbo e assinatura