

Em busca do contaminante de *H. pluvialis*

Algologia Aplicada 18/05/2023

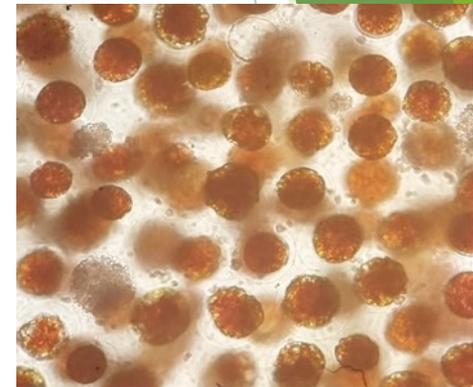
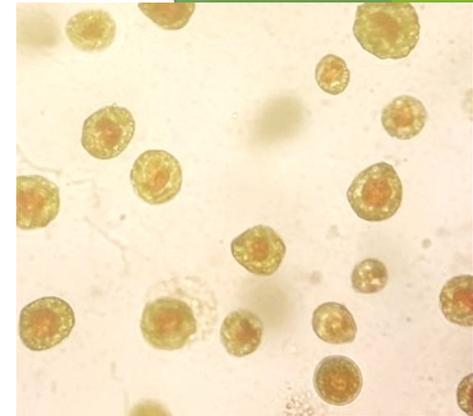
Helena David - hidavid@fc.ul.pt

Haematococcus pluvialis

(Chlorophyceae, Volvocales)

- ▶ Espécie de alto interesse para a biotecnologia
- ▶ Produz astaxantina - Pigmento Vermelho com potente propriedade antioxidante.
- ▶ Tem uma fase do ciclo de vida com células flageladas.
- ▶ Problema! Observámos a cultura e muitas células não nadavam e possuíam características diferentes das que estávamos à espera:

▶ **Cultura contaminada!**



Identificação taxonomica

- ▶ Foi-vos pedido que nos ajudassem a identificar qual seria o organismo com que iriam trabalhar.

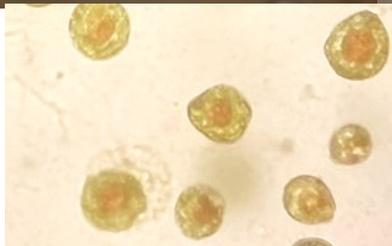
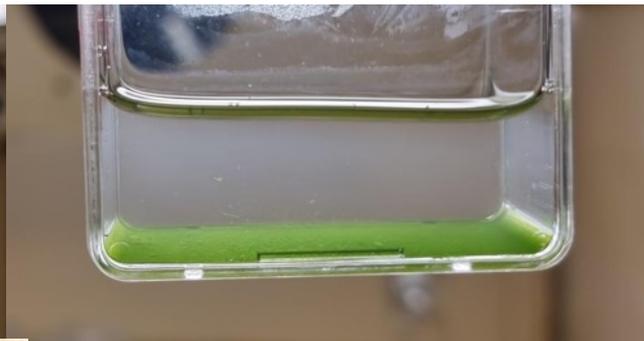


Morfologia

Biologia molecular

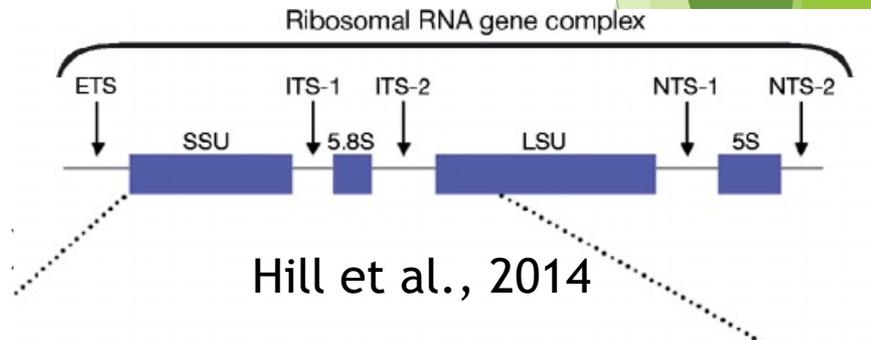
Em busca do contaminante de *H. pluvialis*

- ▶ Para o encontrarmos foi preciso:
 - ▶ Isolar o contaminante e isolar o *Haematococcus*
 - ▶ Esperar que as culturas se desenvolvam
 - ▶ Recolher biomassa num Eppendorf

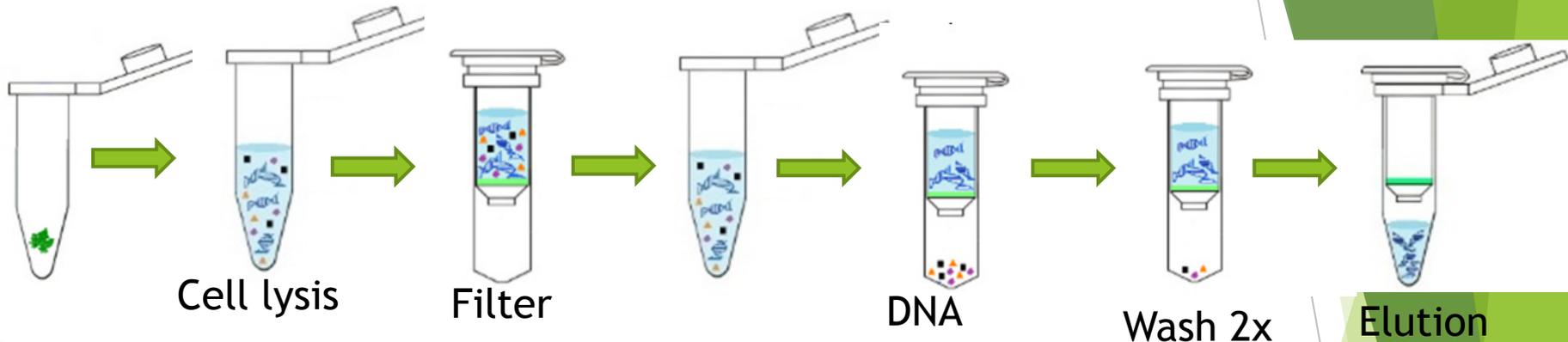


Em busca do contaminante de *H. pluvialis*

- ▶ Como as células são muito pequenas para conseguirmos identificar por morfologia, fomos procurar na biologia molecular.
 - ▶ Taxonomia moderna inclui a ajuda de análises moleculares.
- ▶ Nuclear ribosomal RNA genes:
 - ▶ Intergenic spacer (IGS)
 - ▶ Small subunit (SSU)
 - ▶ Internal transcribed spacer 1 (ITS1)
 - ▶ 5.8S rRNA subunit (5.8S)
 - ▶ Internal transcribed spacer 2 (ITS2)
 - ▶ Large subunit (LSU)



DNA Kit extraction

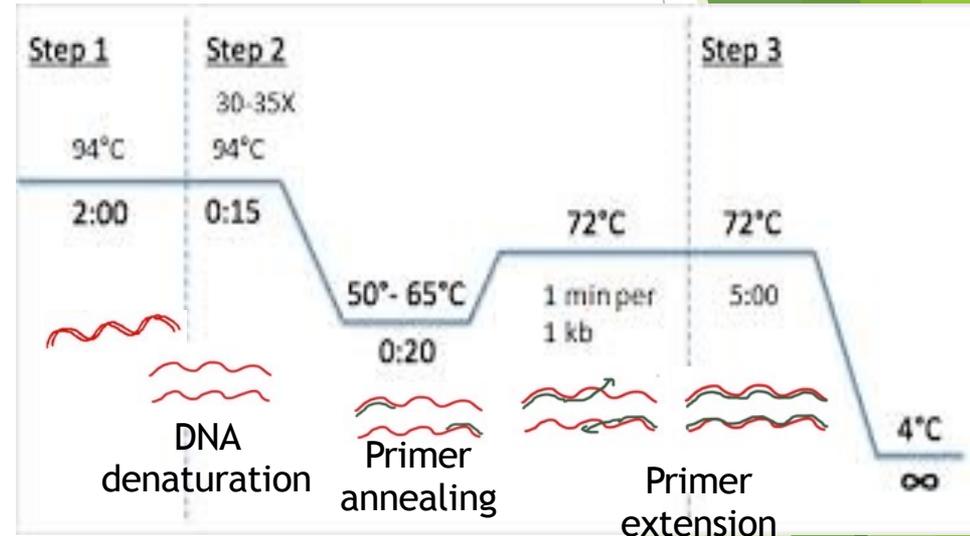
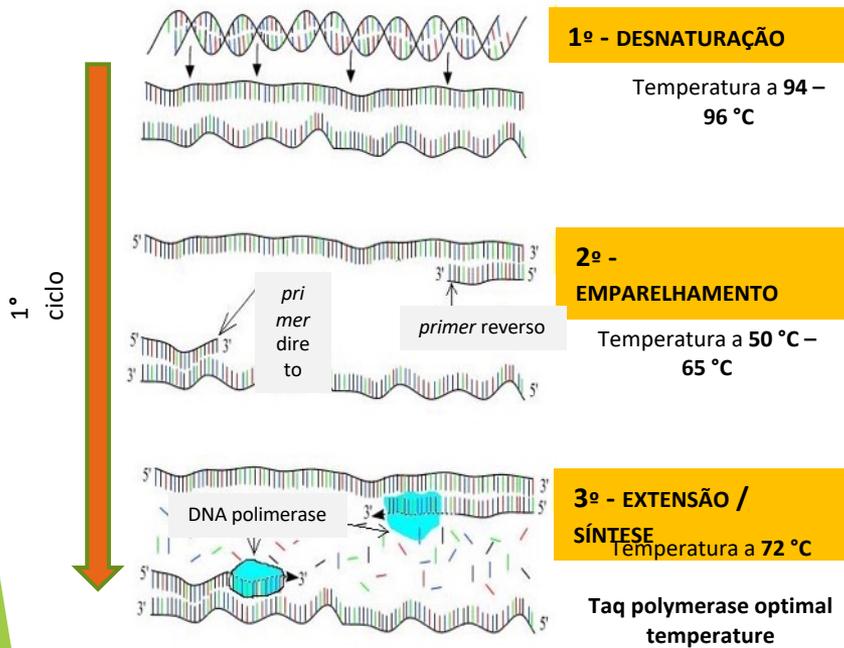


- Breaking cells open
- Isolation and purification of nucleic acid
- Recovering nucleic acids
- Quantifying the amount of nucleic acid

- Existem vários Kits e formas de extração

PCR

- ▶ Polymerase chain reaction/Reação de Polimerização em Cadeia
- ▶ Um processo cíclico e rápido que termina com várias cópias de uma parte específica do ADN.



PCR

Mistura de reação:

- DNA molde
- BioMix (nucleótidos, Tampão de reação, DNA polimerase termoestável, Magnésio)
- *Primers* (direto e reverso)
- Água ultrapura



Termociclador

O TERMOCICLADOR faz variar **ciclicamente** a temperatura e os tempos de reação de acordo com os três passos da PCR

Que sets de primers usar?

- ▶ Extraímos o ADN e para amplificar...

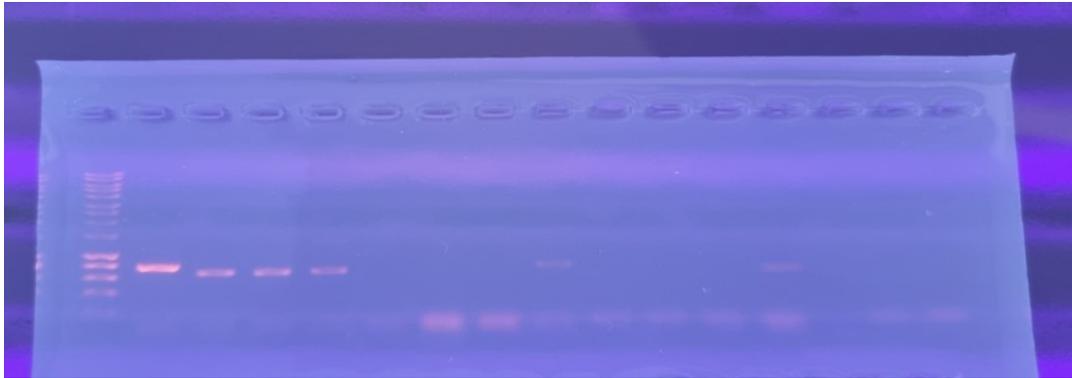
- ▶ Que pares de primers usar?



- ▶ Verificar quais as regiões mais apropriadas
 - ▶ Procurar em artigos científicos
 - ▶ Se existem sequências nas bases de dados

Região amplificada	Primers
ITS1-5.8S-ITS2	ITS1F/ITS1R
SSU	SSU1F/SSU1R

Gel Electroforese



M A. B. C. T.

A B C T

M = Marcador molecular 1Kb

A. = Haematococcus isolado

B. = Contaminante isolado

C. = Controlo da vossa experiência

T. = Tratamento da vossa experiência

Região amplificada	Primers
ITS1-5.8S-ITS2	ITS1F/ITS1R
SSU	SSU1F/SSU1R

Temos resultados, e agora o que fazer?

- ▶ Purificar as amostras de PCR

- ▶ Enviar para sequenciar:



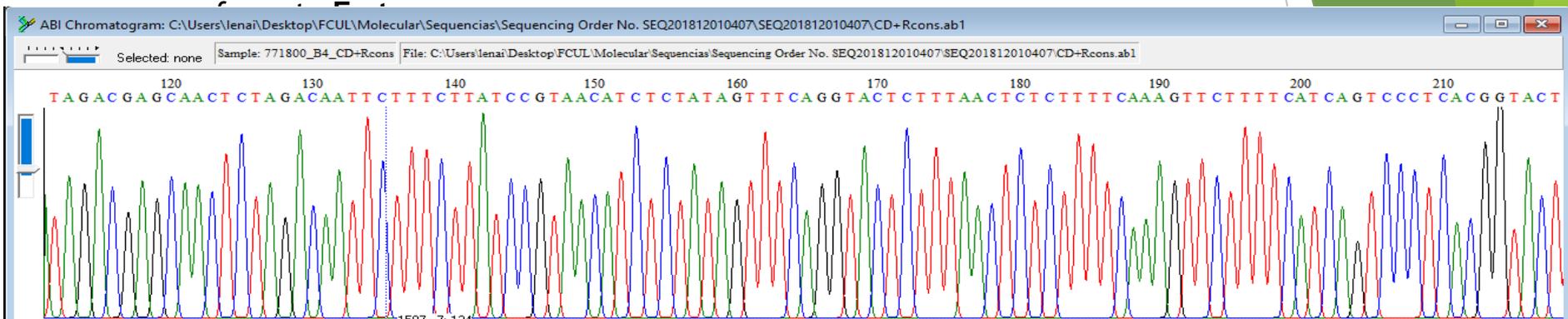
OUR SERVICES

<p>SANGER SEQUENCING</p>	<p>NEXT-GEN SEQUENCING</p>	<p>FRAGMENT ANALYSIS</p>	<p>OLIGOS & POLYMERASES</p>	<p>CLINICAL SEQUENCING</p>
<p>PATERNITY TESTING</p>	<p>ELISAs</p>	<p>ANTIBODIES</p>	<p>GENOMIC IDENTIFICATION</p>	<p>BIRD SEXING</p>

Resultados da sequenciação

Sequências

- Cromatogramas:



BioEdit



Análise Bioinformática

Comparação de
sequências com banco
de dados

(Blast - Basic Local
Alignment Search Tool)

The screenshot shows the NCBI BLAST website. At the top, there are logos for NIH (U.S. National Library of Medicine) and NCBI (National Center for Biotechnology Information). The main heading is "BLAST®". Below this, it says "Basic Local Alignment Search Tool". A brief description states: "BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance." There is a "Learn more" link. To the right, a "NEWS" box announces "IgBLAST version 1.12 is now available" and provides details about its use for immunoglobulin and T cell receptor variable domain sequences, dated Thu, 06 Dec 2018 09:00:00 EST. Under the heading "Web BLAST", there are three main options: "Nucleotide BLAST" (nucleotide to nucleotide), "blastx" (translated nucleotide to protein), and "tblastn" (protein to translated nucleotide). To the right of these is "Protein BLAST" (protein to protein). At the bottom, there is a section for "BLAST Genomes". The browser's address bar shows the URL: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. The Windows taskbar at the bottom shows the time as 4:35 PM on 13-Dec-18.

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Blast

Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments Download Select columns Show 100 ?

select all 11 sequences selected GenBank Graphics Distance tree of results MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Halochlorella rubescens strain SAG 5.95 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Halochlorella rube...	1086	1086	100%	0.0	100.00%	2198	MK975491.1
<input checked="" type="checkbox"/>	Scotiellopsis reticulata strain ATFG small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scotiellopsis retic...	1086	1086	100%	0.0	100.00%	2620	MT151679.1
<input checked="" type="checkbox"/>	Scotiellopsis reticulata genomic DNA sequence contains 5.8S rRNA gene, ITS2, 28S rRNA gene, strain Scotiellopsi...	Scotiellopsis retic...	1086	1086	100%	0.0	100.00%	714	LR215791.1
<input checked="" type="checkbox"/>	Scenedesmus sp. SM8_2 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scenedesmus sp...	1086	1086	100%	0.0	100.00%	1240	KT778097.1
<input checked="" type="checkbox"/>	Scenedesmus sp. SM8_3 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scenedesmus sp...	1086	1086	100%	0.0	100.00%	1225	KT778096.1
<input checked="" type="checkbox"/>	Scenedesmus sp. SM8_1 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scenedesmus sp...	1086	1086	100%	0.0	100.00%	1244	KT778095.1
<input checked="" type="checkbox"/>	Scenedesmus rubescens genomic DNA containing ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, strain SAG 5.95	Halochlorella rube...	1086	1086	100%	0.0	100.00%	659	HG514422.1
<input checked="" type="checkbox"/>	Halochlorella rubescens strain SBB-20 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Halochlorella rube...	1086	1086	100%	0.0	100.00%	706	OP810416.1
<input checked="" type="checkbox"/>	Scotiellopsis reticulata strain CCALA 474 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scotiellopsis retic...	1086	1086	100%	0.0	100.00%	2412	JX513885.1
<input checked="" type="checkbox"/>	Scotiellopsis reticulata isolate MJRU_PIRS4_UNSA small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scotiellopsis retic...	1086	1086	100%	0.0	100.00%	733	MZ836487.1
<input checked="" type="checkbox"/>	Scotiellopsis sp. strain DB-2-1 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scotiellopsis sp...	1086	1086	100%	0.0	100.00%	663	MW412766.1
<input type="checkbox"/>	Scenedesmus sp. SSKV3 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Scenedesmus sp...	1081	1081	100%	0.0	99.83%	680	KF731761.1
<input type="checkbox"/>	Tetrademus sp. NAMSU 221a small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene and internal transcribed spacer 2, partial sequence	Tetrademus sp...	1079	1079	100%	0.0	99.83%	684	MZ126807.1

Então o que é?

Halochlorella rubescens?
Scenedesmus rubescens?
Scotiellopsis reticulata?

Verificar as sequências que o Blast nos deu!

244

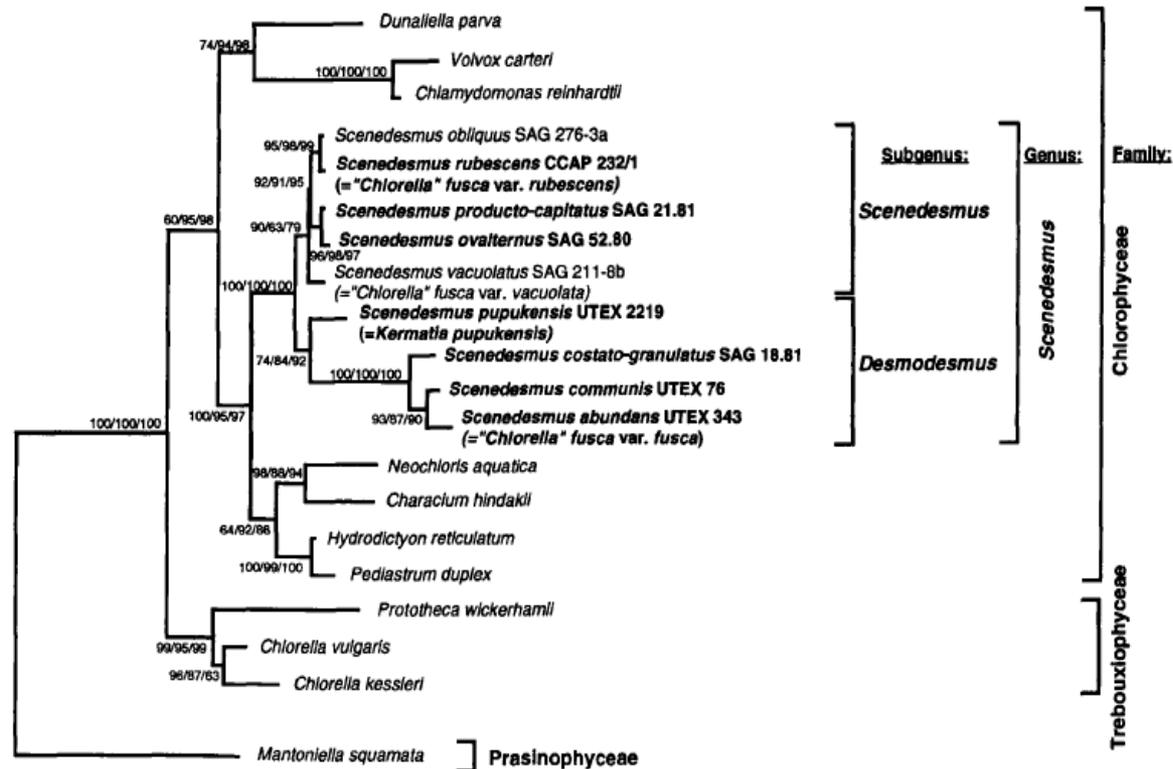
Original Paper

**Physiological, Biochemical, and Molecular Characters
for the Taxonomy of the Subgenera of *Scenedesmus*
(Chlorococcales, Chlorophyta)**

E. Kessler, Michaela Schäfer, Carola Hümmer, Annette Kloboucek, and V. A. R. Huss
Institut für Botanik und Pharmazeutische Biologie der Universität, Erlangen, Germany

Received: July 10, 1996; Accepted: November 6, 1996

Artigo de Kessler et al., 1996



Scenedesmus rubescens comb. nov., subgenus *Scenedesmus*;
Basionym: *Halochlorella rubescens* (Dangeard, 1965; cf. Kalina
 and Punčochářová, 1987); *Synonyms*: *Chlorella fusca* var.
rubescens (Kessler et al., 1968); *C. emersonii* var. *rubescens*
 (Fott et al., 1975).

Alinhar várias sequências

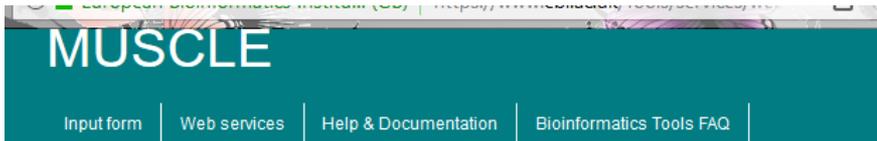
MUSCLE
(Multiple Sequence Comparison
by Log- Expectation)

1. Inserir sequências.
Importante ter o símbolo “>”
antes do nome para o
algoritmo perceber que é
uma nova sequência.
2. Mudar o parâmetro do
output de ClustalW para
Pearson/Fasta
3. Click Submit



The screenshot shows the MUSCLE web interface. The browser address bar displays "https://www.ebi.ac.uk/Tools/msa/muscle/". The page title is "MUSCLE" and the navigation menu includes "Input form", "Web services", "Help & Documentation", and "Bioinformatics Tools FAQ". The main heading is "Multiple Sequence Alignment" with a sub-heading "MUSCLE stands for Multiple Sequence Comparison by Log- Expectation. MUSCLE is claimed to achieve both better average accuracy and better speed than ClustalW2 or T-Coffee, depending on the chosen options." Below this, there is an "Important note" stating "This tool can align up to 500 sequences or a maximum file size of 1 MB." The interface is divided into three steps: "STEP 1 - Enter your input sequences" with a large text area for pasting sequences and a "Browse..." button; "STEP 2 - Set your Parameters" with a dropdown menu for "OUTPUT FORMAT" set to "ClustalW" and a "More options..." link; and "STEP 3 - Submit your job" with a checkbox for "Be notified by email" and a "Submit" button.

Alinhar várias sequências



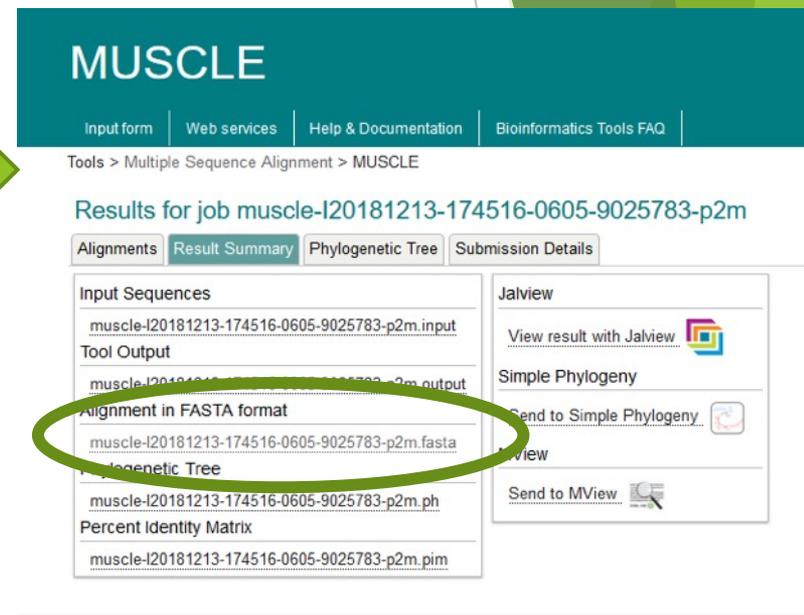
Tools > Multiple Sequence Alignment > MUSCLE

Results for job muscle-I20181213-174516-0605-9025783-p2m

Alignment | **Result Summary** | Phylogenetic Tree | Submission Details

Download Alignment File | View result with Jalview | Send to Simple Phylogeny | Send to MView

```
>IO96-08
-----TTTGT
ACCATGCTGA-GTTTGTGTATCTTGCACATGCAATGTAACACATAT-----GCTT
CACTATAA--GTTTGCATATCTGATCAGTGTGCTGTCTTGCATTCATTTTAAAGTGCAGC
AATGAATGCATCAATTCAAACAAATGAAGAGTGCAGCCAAATGCATTAATTATGCAA
ATTGCAGAATCCGTGAATCATTCTTCTTGAATGAGGCATACCTCTCTGGA-----
-----
>FM244634
-----TTTGTG
-ACAATGCTCATGTCATG-----ATGCTTGGTGGCATGCACTTGTAGTTG
TGCATGACAGCTTG-ATA-CITATCAAAGCCTTTCATCAACTGCTCTCTG-----ACAGC
AATGAATGCATCAATTCAAACAAATGAAGAGTGCAGCCAAATGCATTAATTATGCAA
ATTGCAGAATCCGTGAATCATTCTTCTTGAATGAGGCATGTTCTTGGCCACCCCTAA
GGACATCATTACTCAGTGCATCTTGTAAATGAGTGTGCTCCACATCATGCAATTATGTT
GTGTGATGGTATACCTCTTC---TTTGCAGTTATGCAATACACATTCG---TTTGT-----
-----
>JX065551
-----TGAACCTGCAGAAGGATCATTGTG
AACATGCTCATGTCATG-----ATGCTTGGTGGCATGCACTTGTAGTTG
TGCATGACAGCTTG-ATA-CITATCAAAGCCTTTCATCAACTGCTCTCTG-----ACAGC
AATGAATGCATCAATTCAAACAAATGAAGAGTGCAGCCAAATGCATTAATTATGCAA
ATTGCAGAATCCGTGAATCATTCTTCTTGAATGAGGCATGTTCTTGGCCACCCCTAA
GGACATCATTACTCAGTGCATCTTGTAAATGAGTGTGCTCCACATCATGCAATTATGTT
GTGTGATGGTATACCTCTTC---TTTGCAGTTATGCAATACACATTCGTTTGTGCCAAC
ATCTGAATGAATGCTTAA-----
>JX065557
```



- Ir para “Results Summary” e em “Alignment in FASTA format”, clicar no botão direito e guardar o alinhamento para o PC.

Analises Filogenéticas

- ▶ Filogenia → Estudo da relação evolutiva entre grupos de organismos
 - ▶ O resultado de estudos filogenéticos é a história evolutiva dos grupos taxonómicos.
- ▶ Árvore Filogenética → representação gráfica que apresenta as relações evolutivas entre varias espécies ou outras entidades que possam ter um ancestral comum.
 - ▶ Cada nó na árvore representa o mais recente antepassado comum
 - ▶ O comprimento dos ramos representa estimativas do tempo evolutivo
- ▶ Analises estatísticas robustas:
 - ▶ Neighbor-Joining
 - ▶ Maximum Likelihood
 - ▶ Maximum Parsimony



MEGA

MEGA 7.0.26(7170509-x86_64)

File Analysis Help

Align Data Models Distance Div

Close Data

MEGA release #7170509-x86_64

TTGG
AAT
TTT
GGTGTGCTCCACAATGCAATATATGTGTGTGTGATGTGGTATACTCTCTTTGCAGTTATGCAATA

M7: Analysis Preferences

Options Summary

Option	Selection
Analysis	Phylogeny Reconstruction
Scope	All Selected Taxa
Statistical Method	Neighbor-joining
Phylogeny Test	
Test of Phylogeny	Bootstrap method
<i>No. of Bootstrap Replications</i>	1000
Substitution Model	
Substitutions Type	Nucleotide
Model/Method	Tamura 3-parameter model
Substitutions to Include	d: Transitions + Transversions
Rates and Patterns	
Rates among Sites	Uniform rates
<i>Gamma Parameter</i>	Not Applicable
Pattern among Lineages	Same (Homogeneous)
Data Subset to Use	
Gaps/Missing Data Treatment	Complete deletion
<i>Site Coverage Cutoff (%)</i>	Not Applicable

? Help Compute Cancel

E que é o contaminante?

Scenedesmus rubescens

