AlgaeCeuticals

Development of microalgae-based natural UV Sunscreens and Proteins as cosmeceuticals and nutraceuticals

PCR & DNA Barcoding

Dr. Evangelia Stavridou, postdoctoral researcher

This project has received funding from the [European Union's Horizon 2020 research and innovation programme] under the [Marie Skłodowska-Curie grant agreement No 778263]

CONTENT

- Polymerase Chain Reaction (PCR)
- Components of PCR
- The 3 stages of PCR
- DNA barcoding
- BAR-HRM

Polymerase Chain Reaction (PCR)

"*A technique used in molecular biology to* copy DNA, utilizing repeated cycles of three basic steps *to generate thousands to millions copies of that particular DNA sequence*"

→ Developed in 1983 by Kary Mullis

→ In 1993, Mullis was awarded the Nobel Prize in Chemistry for his work on PCR

COMPONENTS OF PCR

DNA template

DNA polymerase

DNA template = the DNA target sequence

The DNA molecule that contains the DNA region (segment) to be amplified

DNA polymerase = thermostable enzyme

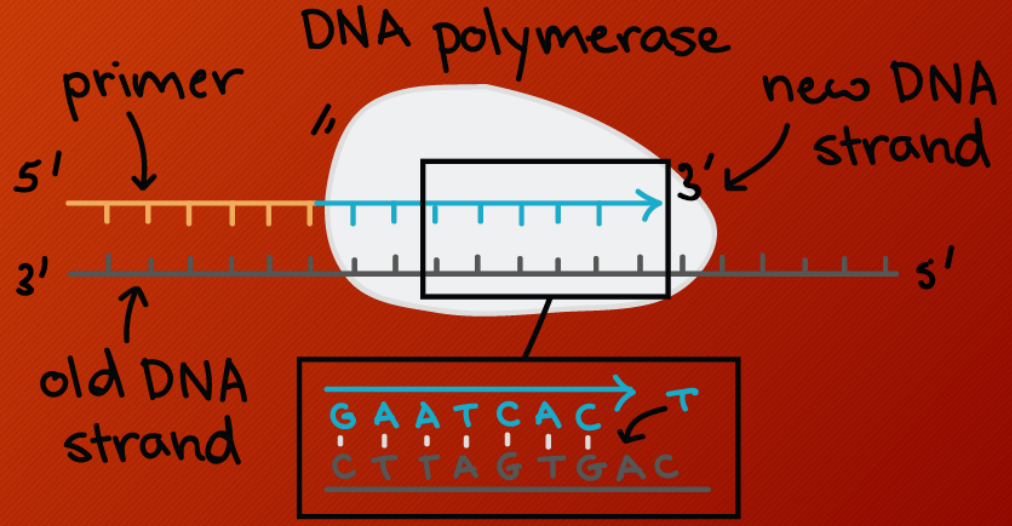
→ Always needs a template

→ Adds sequentially nucleotides to the 3' end of a DNA strand

→ **Requires a pre-existing chain or short stretch of nucleotides called a primer**

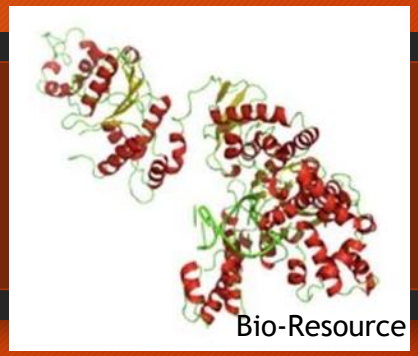
→ **Proofreads** their work

→ **Mismatch repairing,** removing the vast majority of "wrong" nucleotides that are accidentally added to the chain



By https://www.khanacademy.org/

COMPONENTS OF PCR



Taq DNA polymerase

Taq DNA polymerase (from *Thermus aquaticus*: a thermophilic bacterium) → **the most common enzyme**

High temperature stability
- facilitates high specificity of the **primers**
- reduces the production of nonspecific products (such as primer dimers)


Habitat of *Thermus aquaticus*

Bio-Resource: https://www.youtube.com/watch?v=tNR7Curs0N8

COMPONENTS OF PCR

Primers

- Synthetic DNA strands are of about 17 to 24 nucleotides long
- Complementary to 3’end of the template strand
- Two primers are required for PCR:
    - the forward primer complimentary to the 3’end of antisense strand (3’-5’) of a gene
    - the reverse primer complimentary to the 3’end of sense strand (5’-3’) of a gene
- Presence of Guanine (G) and Cytosine (C) bases at the 3′ end of the primer—the GC clamp—helps promote correct binding due to stronger bonding of G and C bases

  >3 C or G repeats in the first 5 bases from the 3’-end of primers may cause primer-dimer formation
- Thymine (T) or Adenine (A) residues should be avoided at the 3’end of primers as they pair with a weaker single H-bond

COMPONENTS OF PCR

dNTPs

Magnesium

Nucleotides (dNTPs or deoxynucleotide triphosphates)

All types of nucleotides are “building blocks” for the new DNA strands and essential for the PCR

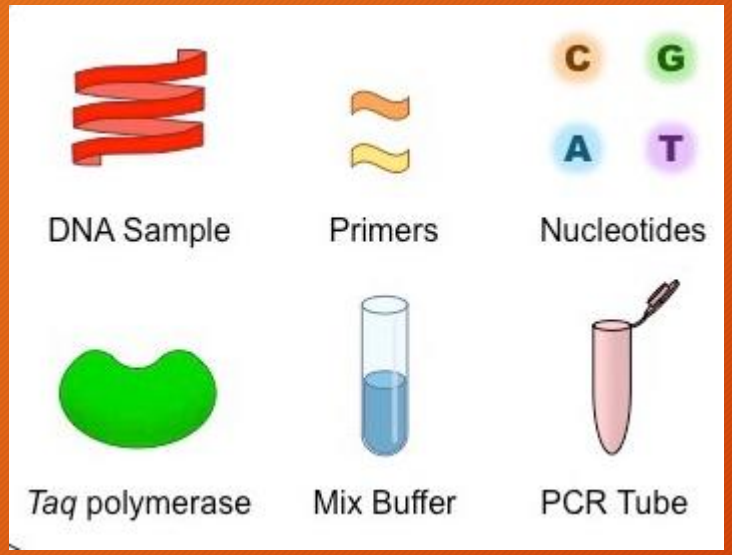
Adenine(A), Guanine(G), Cytosine(C) and Thymine(T)

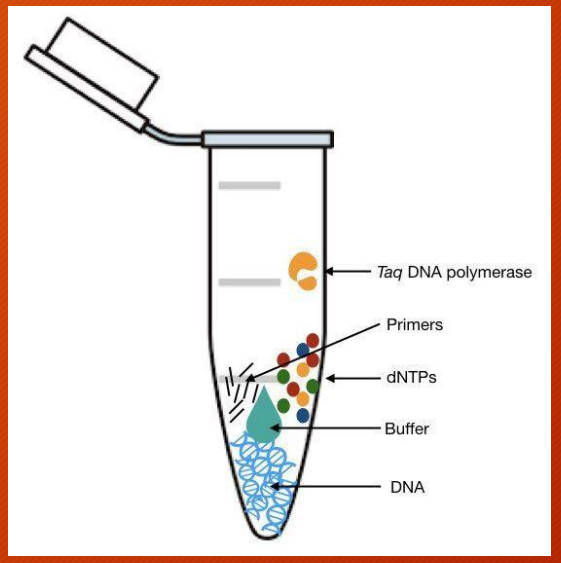
Magnesium (Mg^{2+})

- Affects primer annealing and template denaturation, as well as enzyme activity
- An excess of magnesium gives non-specific amplification products, while low magnesium yields lesser amount of desired product.
- Mg^{2+} ions in the buffer act as co-factor for DNA polymerase enzyme and hence are beneficial to the reaction

COMPONENTS OF PCR SUM-UP

PCR
SUM-UP

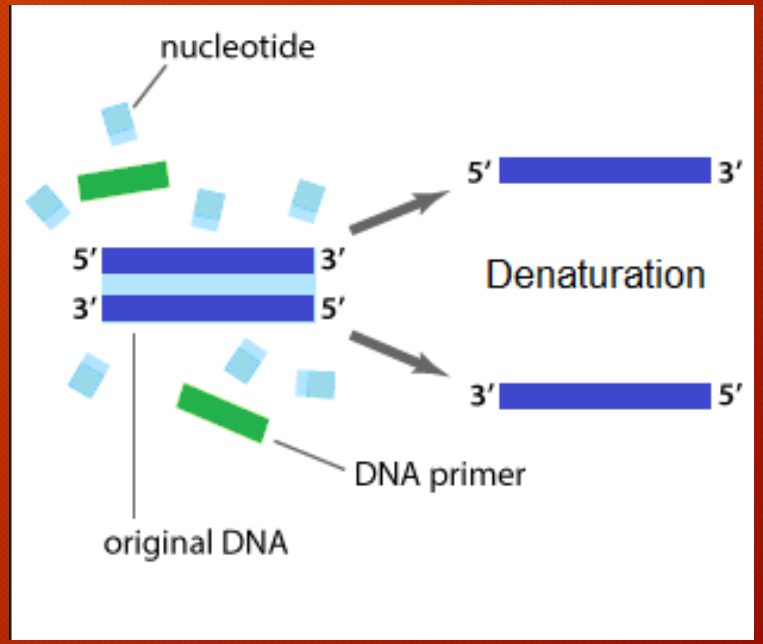




THE 3 STAGES OF PCR - Denaturation

Denaturation step: During the heating step (denaturation), the reaction mixture is heated to 94-95°C for 1 min, which causes separation of DNA double stranded

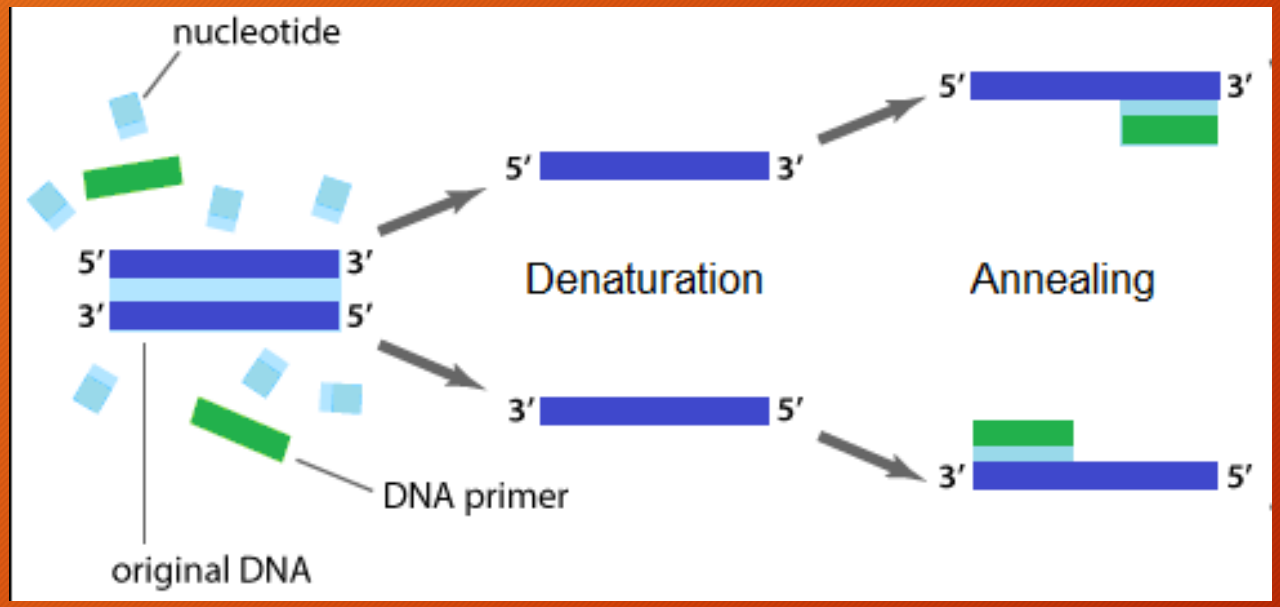
→ Now, each strand acts as template for synthesis of complimentary strand



THE 3 STAGES OF PCR- Annealing

Annealing step: cooling of reaction mixture after denaturation step, which causes hybridization (annealing) of primers to the separated strands of DNA (template)

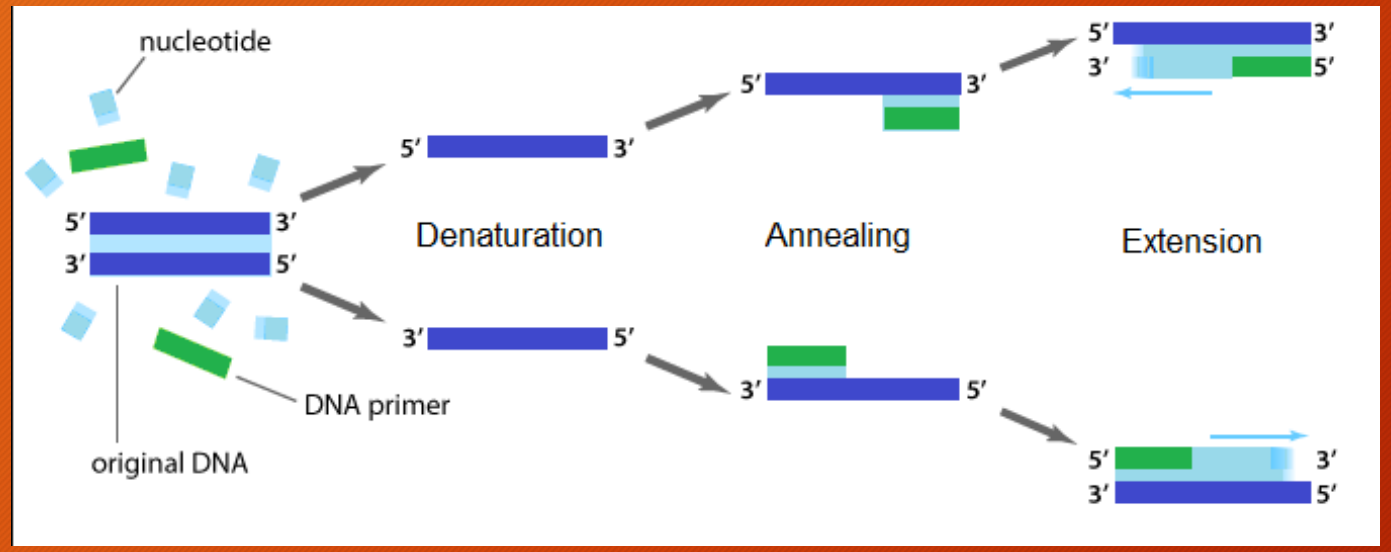
→ The length and GC-content of the primers should be sufficient for stable binding with template



THE 3 STAGES OF PCR- Extension

Extension step: The reaction mixture is heated to 72 °C which is the ideal working temperature for the Taq polymerase

→ The polymerase adds nucleotide (dNTP's) complimentary to template on 3'-OH of primers thereby extending the new strand



THE 3 STAGES OF PCR- Final hold

The first three steps are repeated 35-40 times

→ to produce millions of exact copies of the target DNA
→ exponential increase

Suppose there is only one copy of the desired gene before the PCR starts, after one cycle of PCR, there will be 2 copies, after 2 cycles of PCR, there will be 4 copies. After 3 cycles there will be 8 copies and so on…

Final hold: Once several cycles have been completed, during the hold step, a temperature of 4-10^{o} C is maintained for short-term storage of the amplified DNA sample

DNA BARCODING for species identification

In 2003, Paul Hebert (researcher at the University of Guelph in Ontario, Canada) proposed “DNA barcoding” as a way to identify species

DNA barcoding is a technique used to establish genetic relationships between organisms

This technology is widely used in eukaryotic organisms including algae for species identification

Plant DNA BARCODING

A good DNA barcoding locus should have:

- adequate internal variability to enable differentiation at the species level
- contain flanking regions that are conserved enough to study routine amplification across highly divergent taxa

  → Plastid regions (e.g. rbcL and matK, and the non-coding spacer trnH-psbA)

  → Internal transcribed spacer (ITS) region of nuclear ribosomal DNA

Algae species identification

The identification is based on the genetic diversity of specific genomic regions characterized by their universality across diverse taxa & their effectiveness in identifying inter-/ intra- species-specific differences

DNA barcoding of algae is commonly used for **species identification** and **phylogenetic studies**

→ Common primers used amplify the:

- *ITS1-ITS4* region
- the Cyanobacterial 16S rRNA region
- the ribulose biphosphate carboxylase (*rbcL*) chloroplast region restricted to the chloroplasts of the photosynthetic organism's

iBOL and BOLD

- The International Barcode of Life (iBOL) Consortium, an alliance of research organizations in more than 30 nations since 2010 the DNA-based biosurveillance system (https://ibol.org/)

- Barcode sequences are placed in the Barcode of Life Data Systems (BOLD) database

online workbench that includes a reference library of DNA barcode records for assigning identities to sequences of unknown origin (http://www.ibol.org/phase1/bold/)

An identification engine based on the current barcode library that monitors the number of barcode sequence records and species coverage

DNA BARCODING-procedure

Step 1: Isolate DNA from the sample

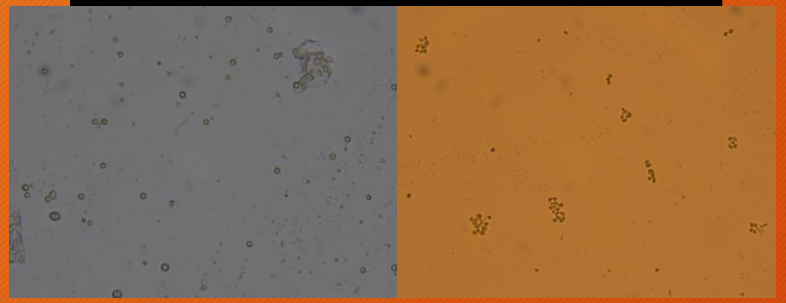
Step 2: Amplify the target DNA barcode region using PCR

Step 3: Sequence the PCR products

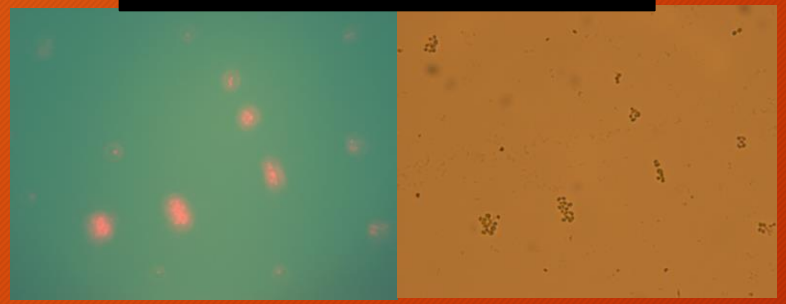
Step 4: Compare the resulting sequences against reference databases to find the matching species

DNA isolation from microalgae

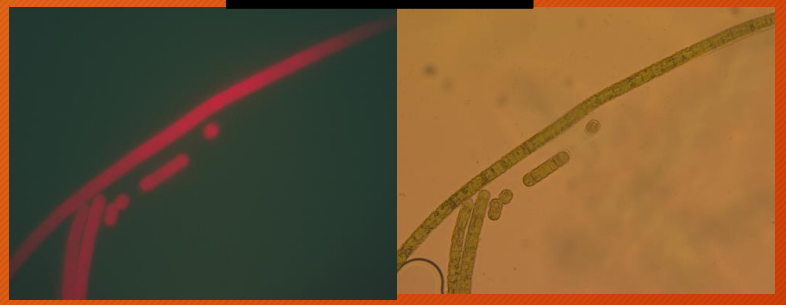
Nannochloropsis limnetica



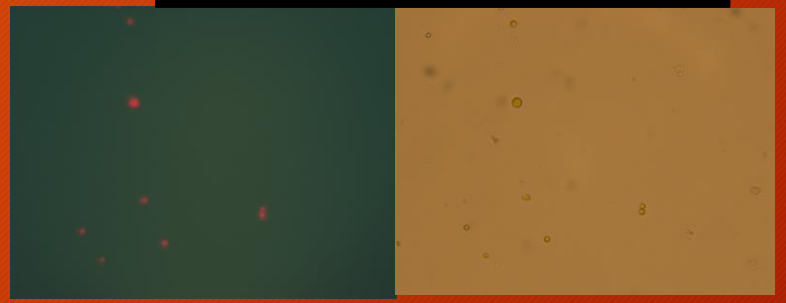
Chlorella minutissima



Arthrospira



Chlorella sorokiniana



CTAB

Commercial kit

DNA BARCODING

Step 1: Isolate DNA from the sample

for algae an ideal sequence is the 16s rRNA in the chloroplasts

Step 2: Amplify the target DNA barcode region using PCR

Step 3: Sequence the PCR products

Step 4: Compare the resulting sequences against reference databases to find the matching species

DNA BARCODING- PCR

PCR REACTION

PCR reaction reagents	Concentration
Buffer	1X
dNTPs	0.2 mM
Primer F	0.25 μM
Primer R	0.25 μM
Taq	1 U/μl
DNA	20-30ng
H_20	-

PCR CONDITIONS

Stage	Temperature (°C)	time (min:sec)	Cycles
Denaturation	94	03:00	1
Denaturation	94	00:30	40
Annealing	59	00:40	
Elongation	72	00:30	
Elongation	72	07:00	1

DNA BARCODING

Step 1: Isolate DNA from the sample

for algae an ideal sequence is the 16s rRNA in the chloroplasts

Step 2: Amplify the target DNA barcode region using PCR

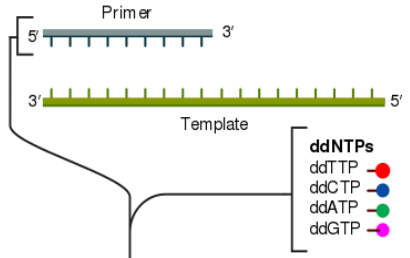
Step 3: Sequence the PCR products

Step 4: Compare the resulting sequences against reference databases to find the matching species

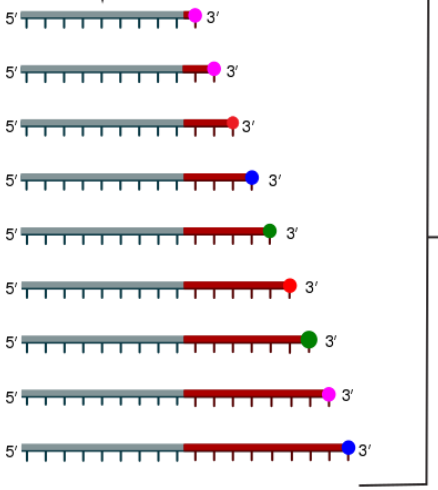
DNA BARCODING- sequencing

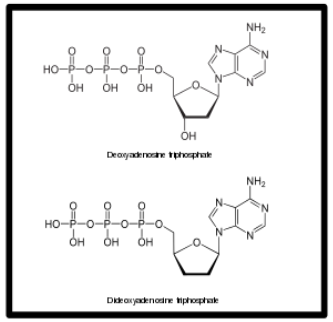
1. Reaction mixture

- Primer and DNA template
- DNA polymerase
- ddNTPs with flourochromes
- dNTPs (dATP, dCTP, dGTP, and dTTP)

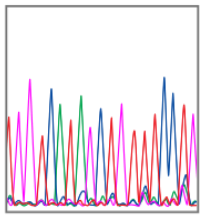


2. Primer elongation and chain termination



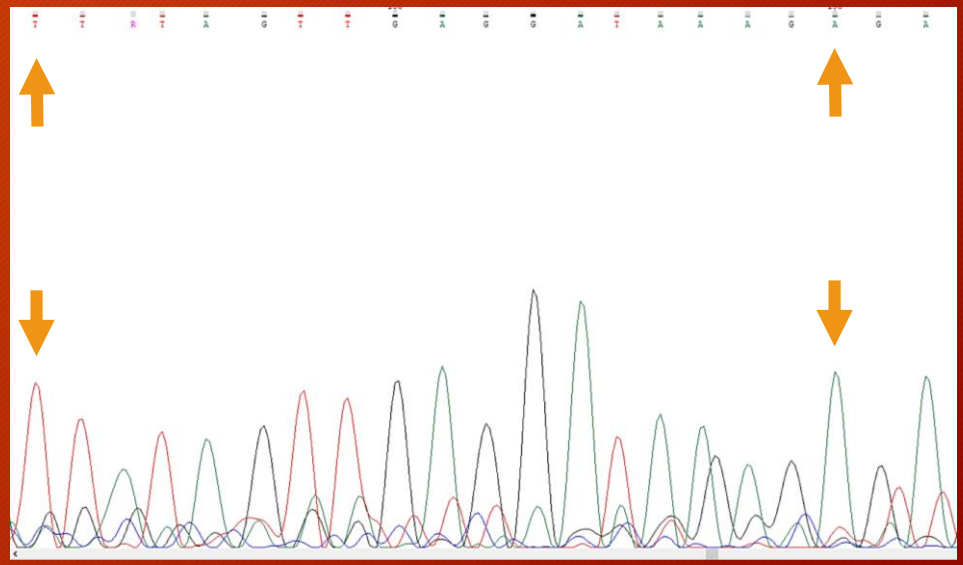


3. Capillary gel electrophoresis separation of DNA fragments



4. Laser detection of the fluorochromes and computational sequence analysis

The nucleotides of the sequence correspond to the highest fluorochrome peak of the chromatograph



By Estevezj - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=23264166

DNA BARCODING

Step 1: Isolate DNA from the sample

for algae an ideal sequence is the 16s rRNA in the chloroplasts

Step 2: Amplify the target DNA barcode region using PCR

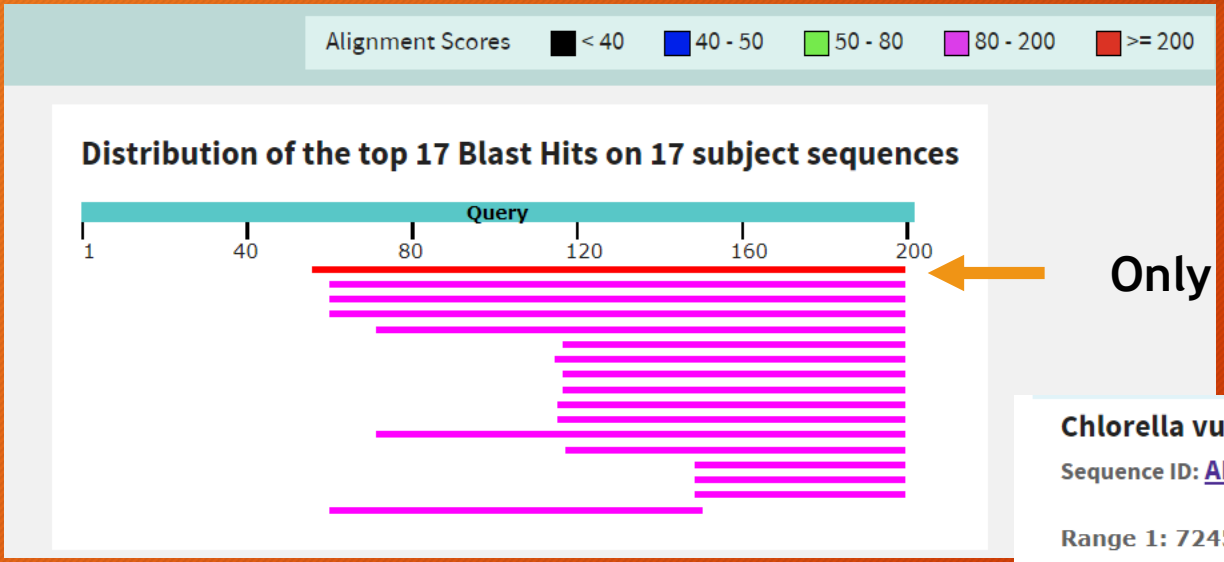
Step 3: Sequence the PCR products

Step 4: Compare the resulting sequences against reference databases to find the matching species

BLAST and phylogenetic tree

- ‘Clean’ the sequences by editing the chromatograph for inconsistences (using software such as Chromas 2.2.6)
- Blast and select the most appropriate (▲%) sequence hit from the database
- Perform multiple sequences alignment
- Construct the phylogenetic tree (using software such as MEGAX)

BLAST



Only one hit showed 95% identity to the query sequence

Chlorella vulgaris C-27 chloroplast DNA, complete sequence

Sequence ID: AB001684.1 **Length:** 150613 **Number of Matches:** 1

Range 1: 72457 to 72599 GenBank Graphics ▼ Next Match ▲ Previous Match

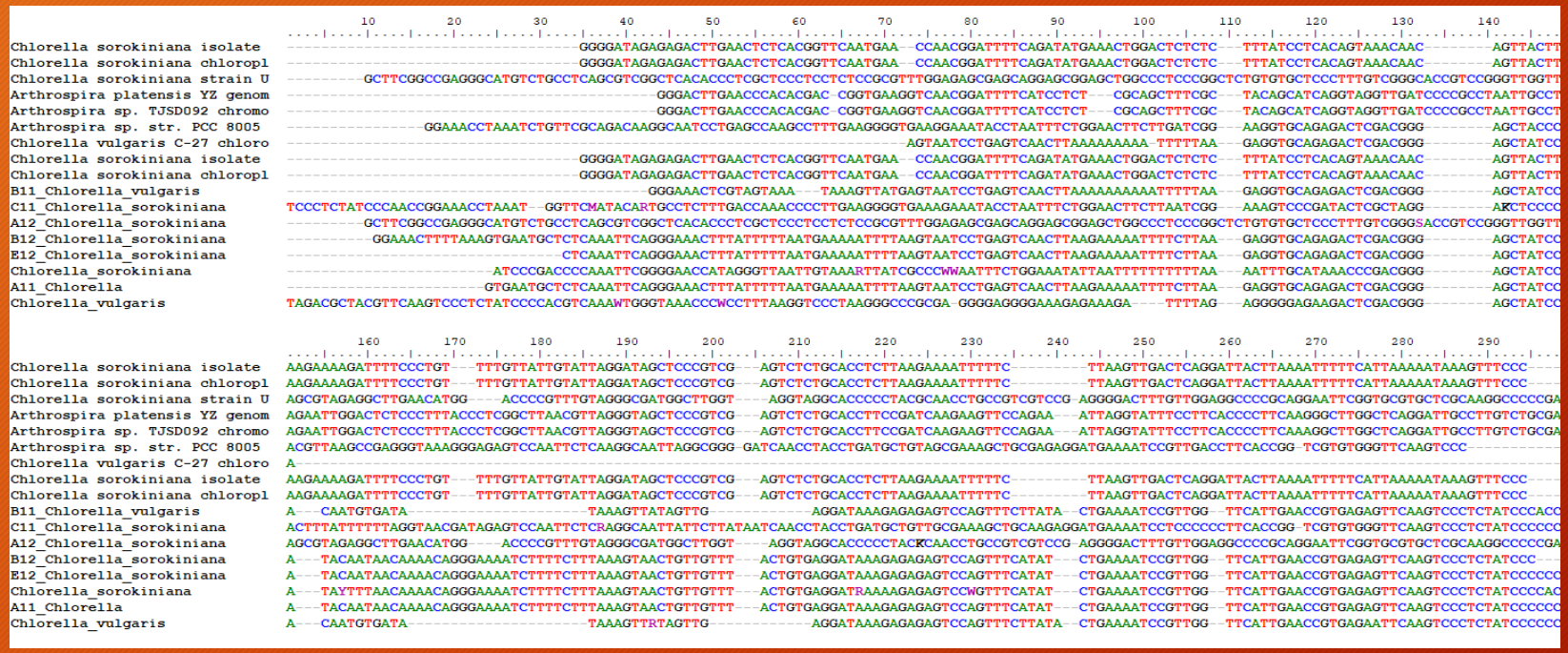
Score	Expect	Identities	Gaps	Strand
228 bits(123)	1e-55	137/144(95%)	2/144(1%)	Plus/Plus

```
Query  57     TTTT-AGAGGGGGAGAAGACTCGACGGGAGCTATCCTAACAATGTGATATAAAGTTRTAG  115
              |||| ||||| | || ||||||||||||||||||||||||||||||||||||||| |||
Sbjct  72457  TTTTAAGAGGTGCAG-AGACTCGACGGGAGCTATCCTAACAATGTGATATAAAGTTATAG  72515

Query  116    TTGAGGATAAAGAGAGAGTCCAGTTTCTTATACTGAAAATCCGTTGGTTCATTGAACCGT  175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  72516  TTGAGGATAAAGAGAGAGTCCAGTTTCTTATACTGAAAATCCGTTGGTTCATTGAACCGT  72575

Query  176    GAGAATTCAAGTCCCTCTATCCCC  199
              |||| |||||||| ||||||||||
Sbjct  72576  GAGAGTTCAAGTCTCTCTATCCCC  72599
```

Multiple Sequence Alignment

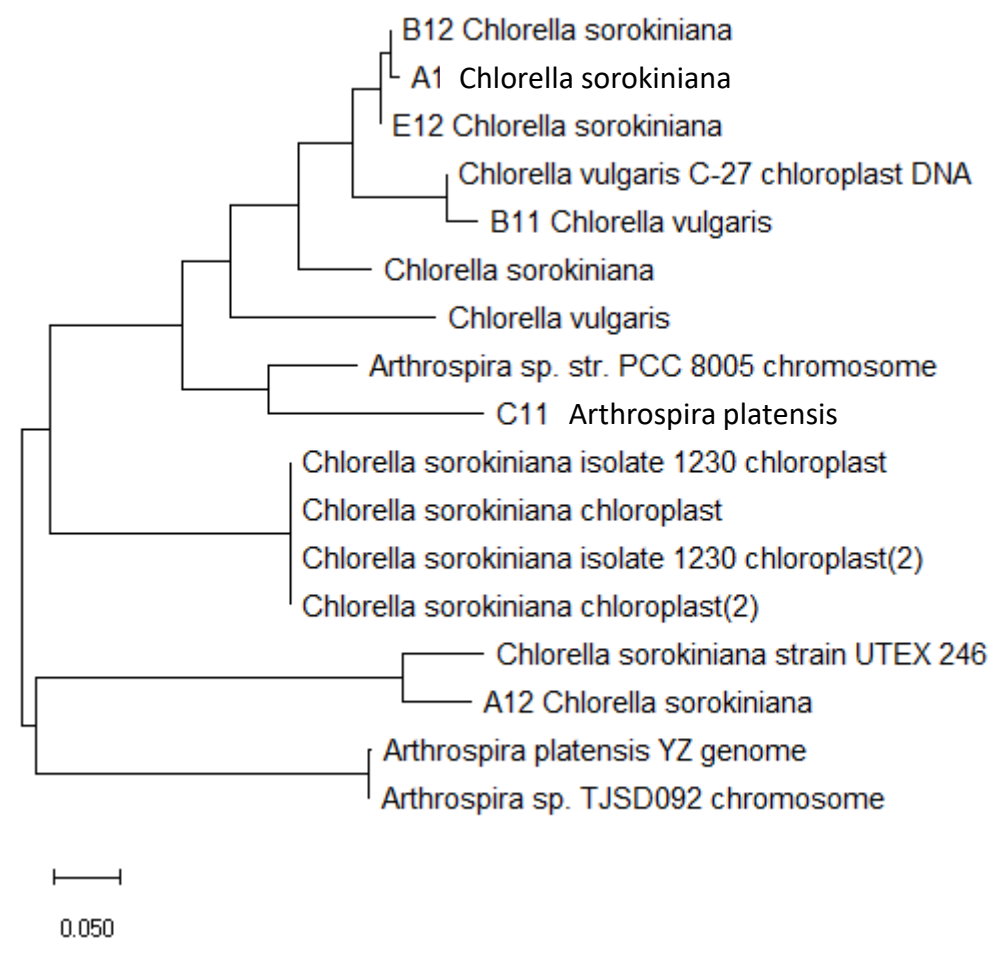
An arrangement in which two or more sequences are positioned in parallel to each other

→ to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences



Alignments are also used to aid in establishing evolutionary relationships by constructing phylogenetic trees

Phylogenetic tree



A branching diagram showing the evolutionary relationships among various biological species based upon similarities and differences in their genetic characteristics

Construct a **neighbor joining** tree:

→ a bottom-up (agglomerative) clustering method for the creation of phylogenetic trees

High Resolution Melting analysis (HRM)

High Resolution Melting analysis (HRM) →

- measures the rate of double stranded DNA dissociation to single stranded DNA with increasing temperature
- Requires a fluorescent dye (homogenously intercalated into DNA)

to follow the dissociation of the double stranded DNA

The amplicon is analysed by observing the change in fluorescence that is caused by the release of intercalating dye from a DNA as it is being denatured by the increasing temperature

Bar-HRM analysis for molecular identification of algae species

- HRM offers a rapid high-throughput method for various analyses as only DNA isolation followed by the PCR steps are needed

- HRM may be coupled with the DNA barcoding (Bar-HRM) using universal regions for the rapid detection, authenticity analysis, taxonomical identification, quantification and adulteration studies!

Bar-HRM protocol

PCR REACTION

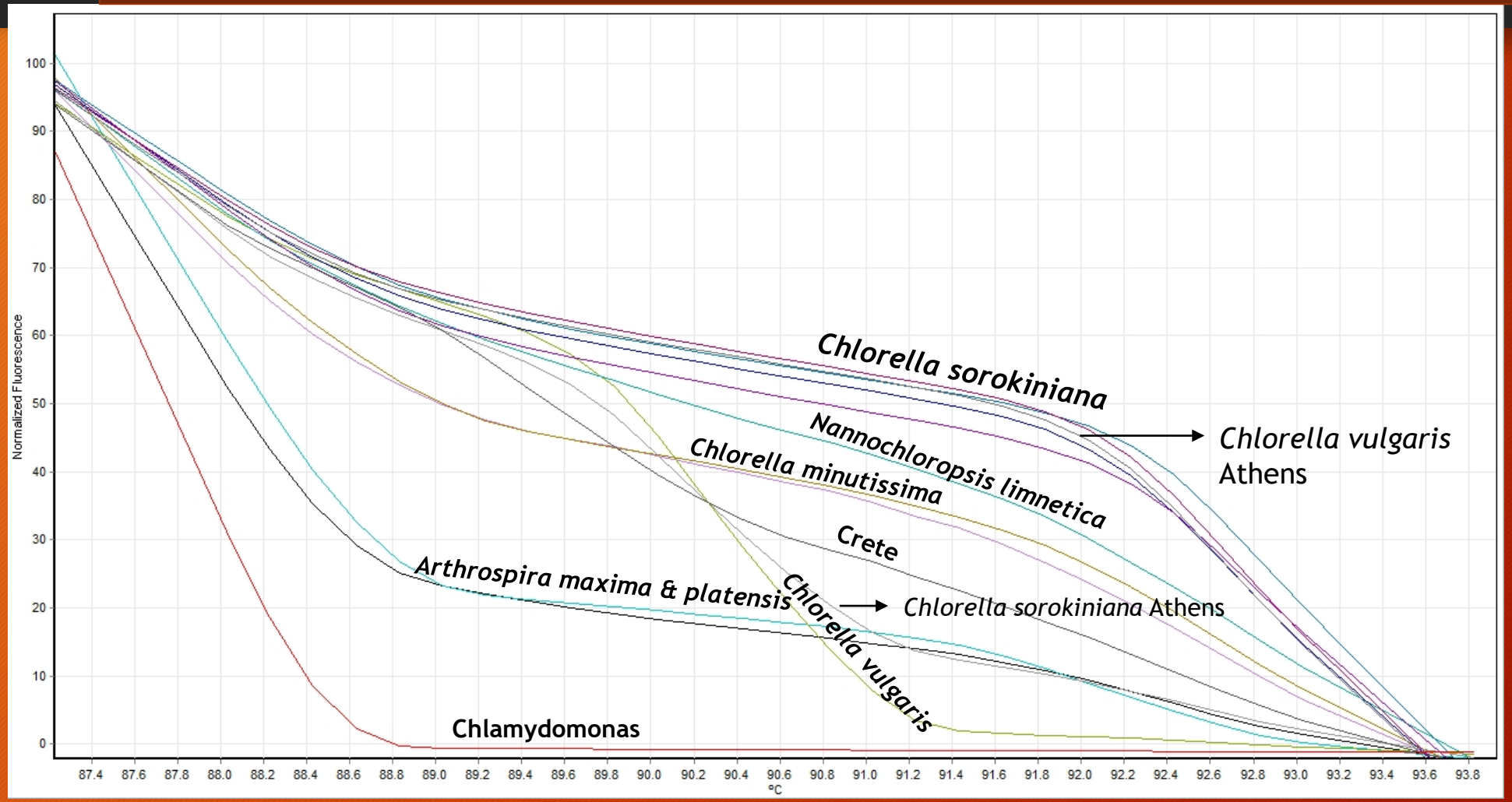
PCR reaction reagents	Concentration
Buffer	1X
dNTPs	0.2 mM
Primer F	0.25 μM
Primer R	0.25 μM
Taq	1 U/μl
Syto	0.6 μl
DNA	20-30ng
H_20	-

PCR CONDITIONS

Stage	Temperature (°C)	time (min:sec)	Cycles
Denaturation	94	03:00	1
Denaturation	94	00:30	40
Annealing	59	00:40	
Elongation	72	00:30	

HRM: melt at 75-95 °C in increments of 0.2 °C/step every 2 sec

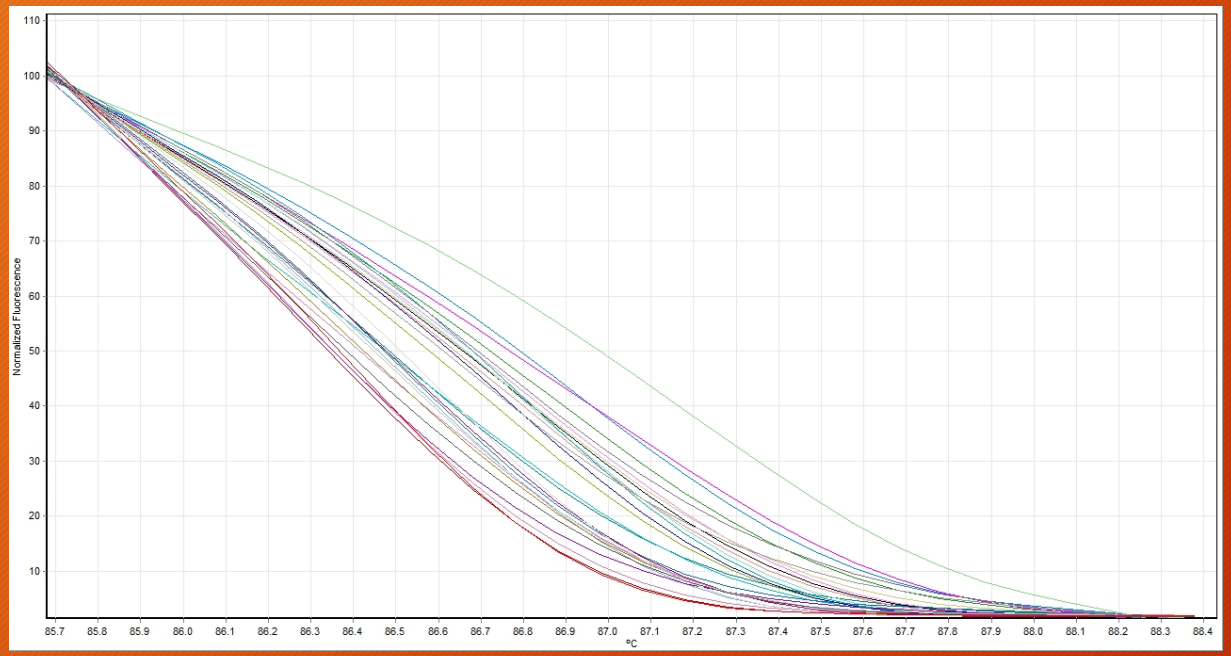
Bar-HRM: *ITS2*



Species are differentiated based on the *ITS2* region

Here we observe that the samples *Chlorella sorokiniana* Athens and *Chlorella vulgaris* Athens were probably mislabelled as they group with the opposite species

Bar-HRM: *trnL*



Species are differentiated based on the chloroplast *trn*L region

Here we observe that the species were not differentiated based on this locus