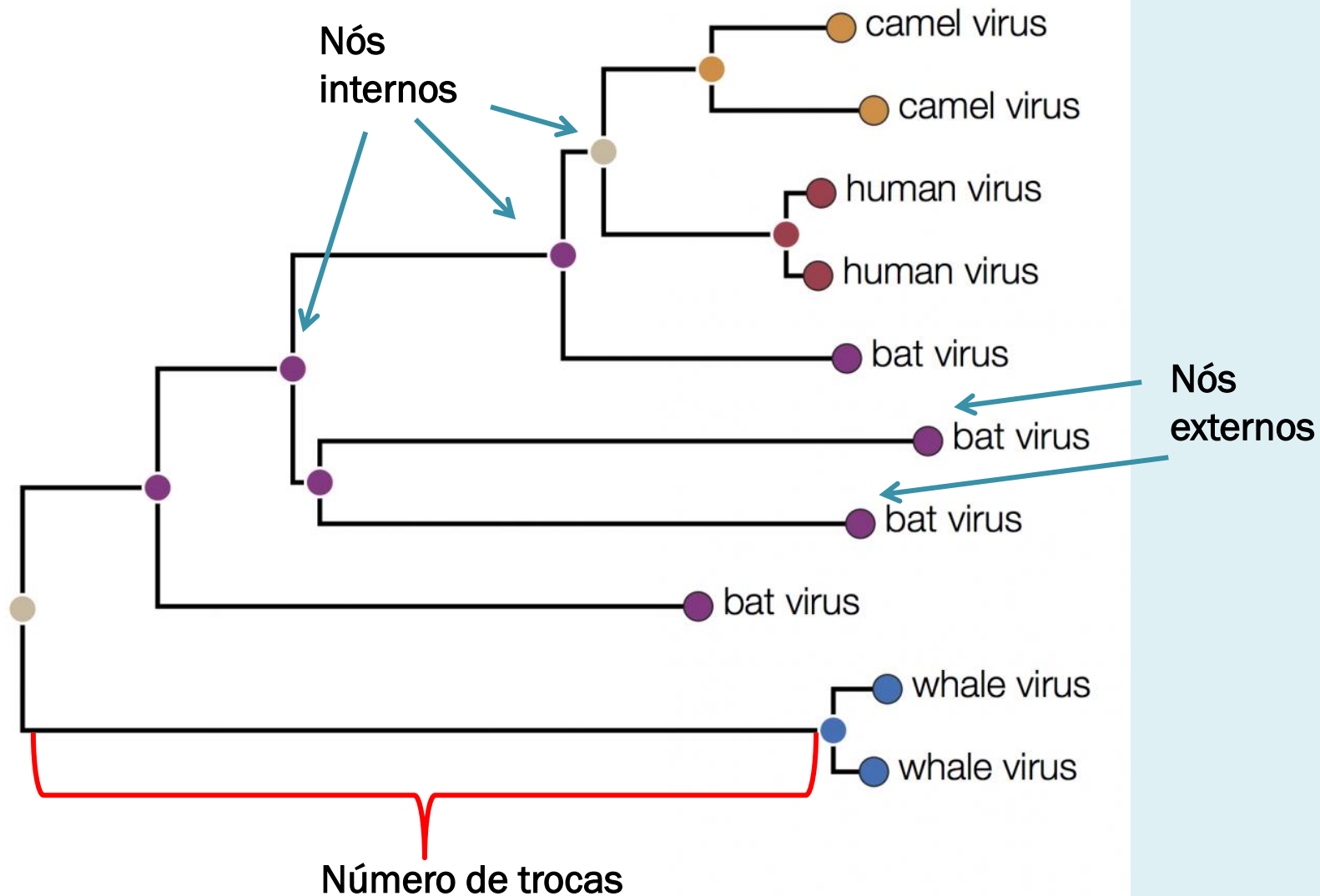


---

**GENE7029**  
**SEQUENCIAS DE DNA NA**  
**IDENTIFICAÇÃO DE ESPÉCIES E**  
**ANÁLISE FILOGENÉTICA**  
**(22/10/2020)**

---

# ÁRVORES FILOGENÉTICAS



# ARVORE FILOGENÉTICA - DEFINIÇÕES

- × **Monofilético:** Um grupo de taxa (indivíduos da análise) que compartilham o mesmo braço, também chamado de cluster.
- × **Parafilético:** Um grupo de taxa que não formam um cluster sem incluir linhagens adicionais

Cluster 1

Cluster 2

OTU / MOTU

A

B

C

D

E

H

I

J

F

G

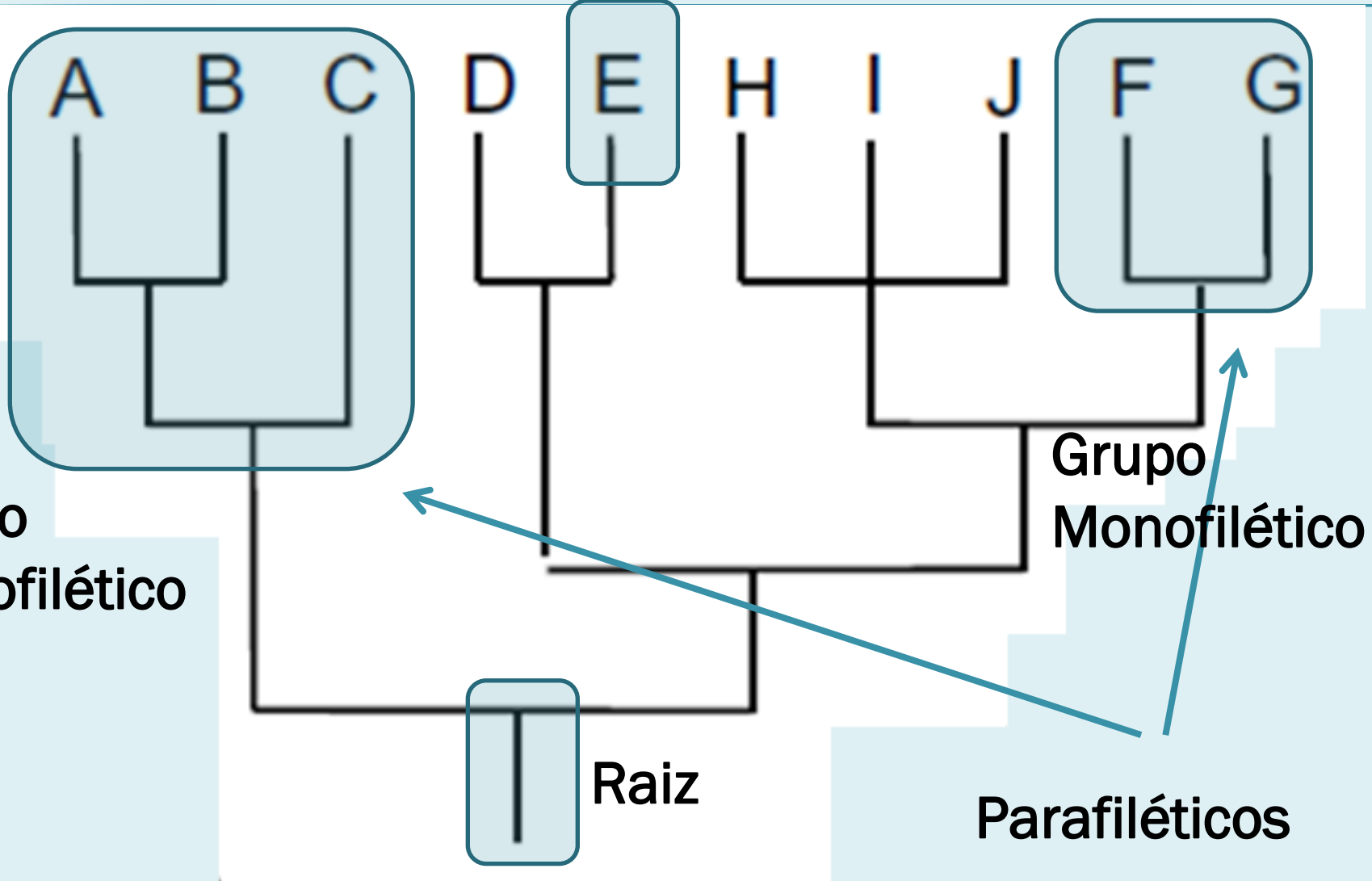
Grupo  
Monofilético

Grupo  
Monofilético

Raiz

Parafiléticos

Relações propostas: Topologia



# ÁRVORES FILOGENÉTICAS

## × Dois interesses principais:

- \* Obter a topologia da árvore - a forma como os nós internos se conectam uns com os outros e com nós da extremidade.
- \* Obter as distâncias entre todos os nós da árvore.

# ÁRVORES FILOGENÉTICAS

- × A raiz de uma árvore filogenética:
  - Na árvore com raiz (ou enraizada), a raiz representa o ancestral comum a todos os nós da árvore
  - Sem informações suficientes para determinar o ancestral comum a todos os nós - árvore sem raiz
    - Especifica somente as relações entre os taxa e não define a via evolutiva

Define relações entre os taxa



Define via evolutiva



Kangaroo



Lemur



Human



Root



Rat



Mouse



Unrooted

Outgroup Rooted



Kangaroo

Lemur



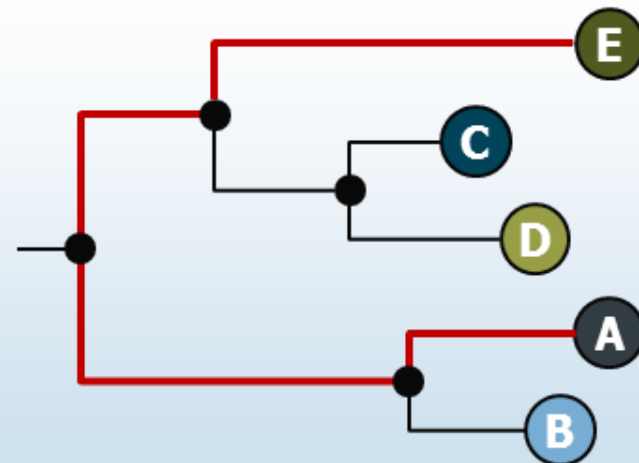
Human



Mouse



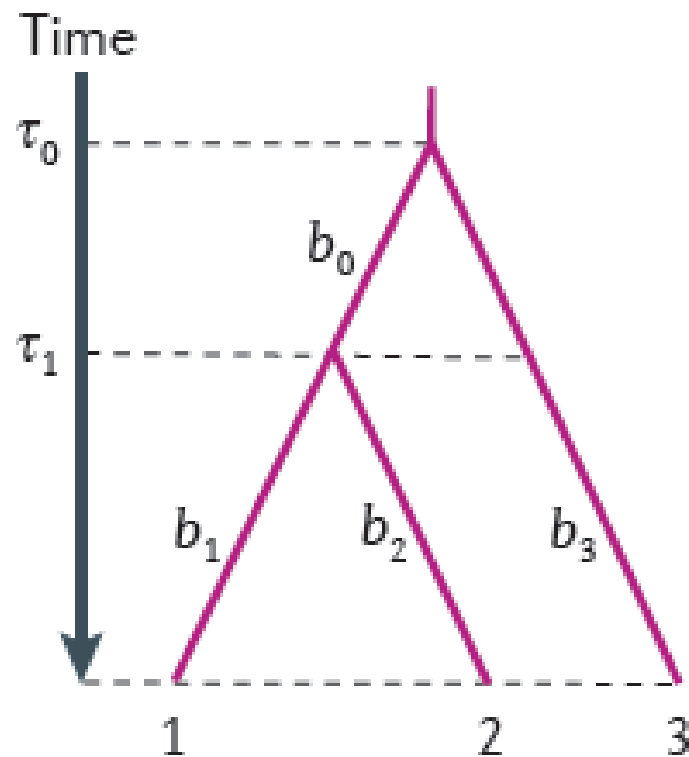
Rat



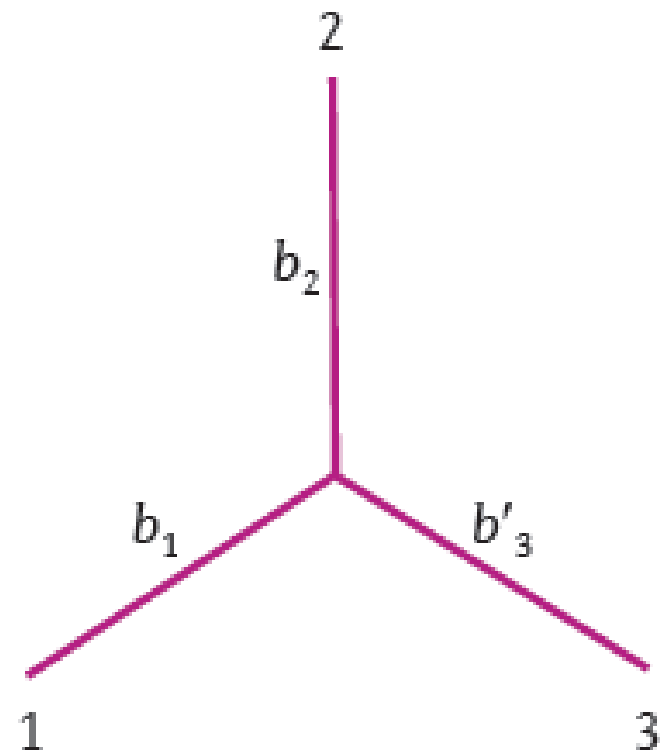
Mid-point  
Rooted

# ÁRVORES FILOGENÉTICAS

**a Rooted tree**

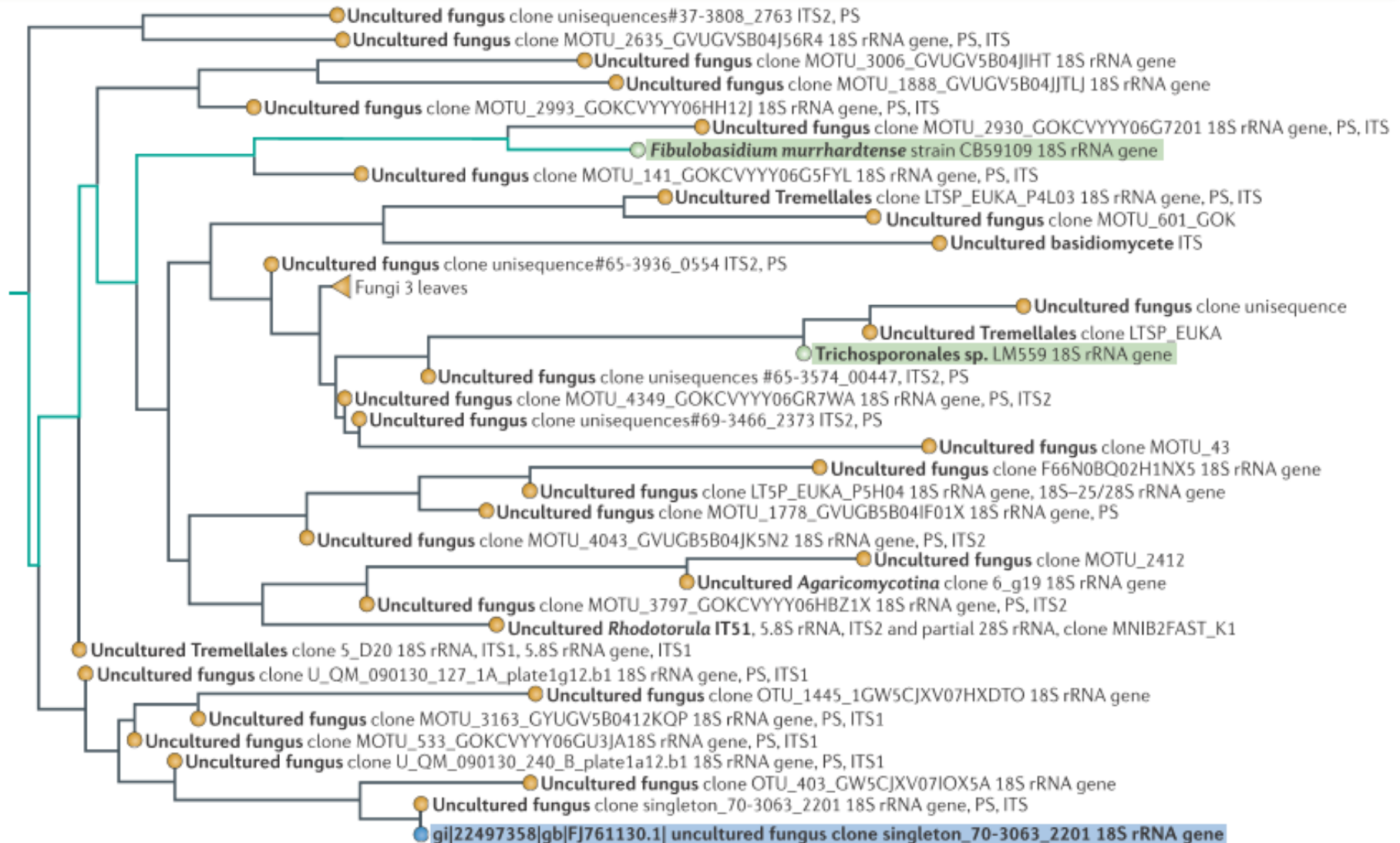


**b Unrooted tree**





# MOTU



# PLANEJANDO.....

---



- ✖ Qual o meu objetivo?
  - + O que eu quero responder????
  - + Identificação ou filogenia????
- ✖ Qual minha hipótese?
  - + O que eu sei sobre o meu fungo????
- ✖ Qual gene usar?
- ✖ Quais sequencias estão disponíveis?
- ✖ Quais linhagens eu devo ter?
- ✖ Onde achar?
- ✖ Como interpretar uma árvore filogenética?

# PLANEJANDO.....

---



× Qual o meu objetivo?

+ O que eu quero responder????

+ Identificação de gênero x espécie????

+ Quando eu preciso fazer análise filogenética????

# SEQUENCIAS DE DNA PARA FILOGENIA E TAXONOMIA MOLECULAR

- × O que usar para procariotos?
  - + Para gêneros: gene 16S do rRNA
  - + Para espécies: Sequenciamento parcial de genes house-keeping:
    - × Exemplos: genes *atpD* (ATP synthase F1, beta subunit), *gyrB* (DNA gyrase B subunit), *rpoB* (RNA polymerase beta subunit), *recA* (recombinase A) and *trpB* (tryptophan synthetase, beta subunit).

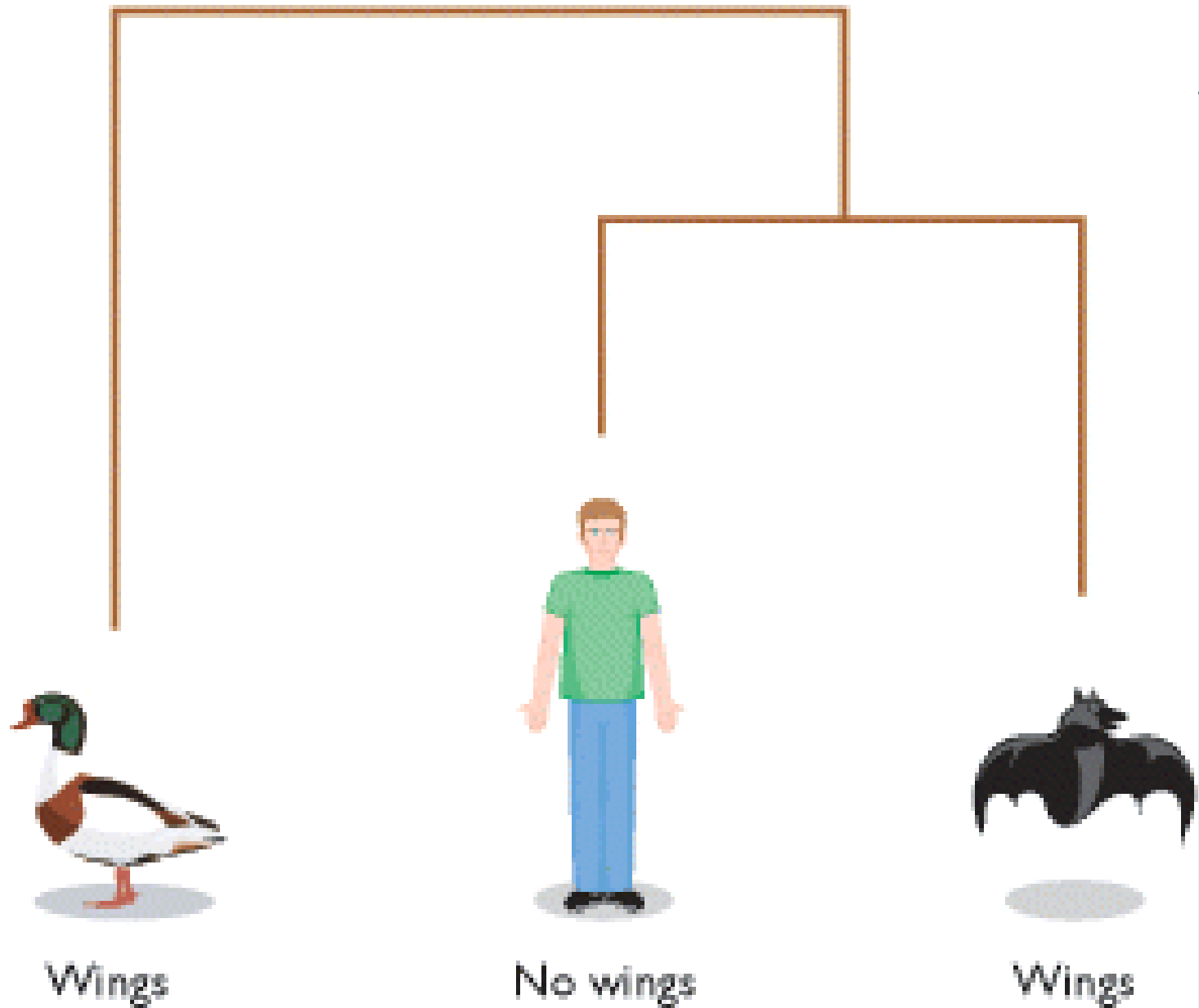
# SEQUENCIAS DE DNA PARA FILOGENIA E TAXONOMIA MOLECULAR

- × O que usar para eucariotos?
- × Para fungos:
  - + Para gêneros: LSU (28S do rRNA)
  - + ITS1-5.8S-ITS2 do rDNA, Segmento  $\alpha$ 1 do Fator de Elongação da Tradução; Sequencias parciais dos genes:  $\beta$ -tubulina, GPDH, Histona, Actina, rpb2 e Mating Type.

# SECUENCIAS DE DNA PARA FILOGENIA - Leveduras YeastIP

Marker	Description	Position in the gene*	Length (bp)
LSU	Complete sequence of the large subunit 26S ribosomal RNA gene	65–3364	3300
D1/D2 LSU	Partial sequence of the 26S ribosomal gene comprising the D1/D2 region	65–636	570
SSU	Complete sequence of the small subunit 18S ribosomal RNA gene	1–1800	1700
ITS	Ribosomal RNA region containing the intergenic region 1 (between 18S and 5.8S), the 5.8S ribosomal RNA gene and the intergenic region 2 (between 5.8S and 26S)	First base of ITS1 to last base of ITS2	400–600
mtSSU	Mitochondrial small subunit 15S ribosomal RNA gene	383–1006	400–600
<i>RPB1</i>	Partial sequence of the RNA polymerase II largest subunit coding gene	253–873	620
<i>RBP2</i>	Partial sequence of the RNA polymerase II second largest subunit coding gene	1645–2319	680–1000
<i>TEF1</i> -alpha	Partial sequence of the translation elongation factor 1-alpha coding gene	64–1190	930
<i>ACT1</i>	Partial sequence of the exon2 of the actin coding gene	405–1383	980 <sup>†</sup>
mtCOX II	Partial sequence of mitochondrial cytochrome C oxidase subunit 2 coding gene	121–707	590

(A)



# DNA BARCODING

A TOOL FOR SPECIMEN IDENTIFICATION AND SPECIES  
DISCOVERY



HOME

ABOUT ▼

PROGRAMS ▼

NEWS & MEDIA ▼

JOIN

## DNA BARCODING

A TOOL FOR SPECIMEN IDENTIFICATION AND SPECIES DISCOVERY

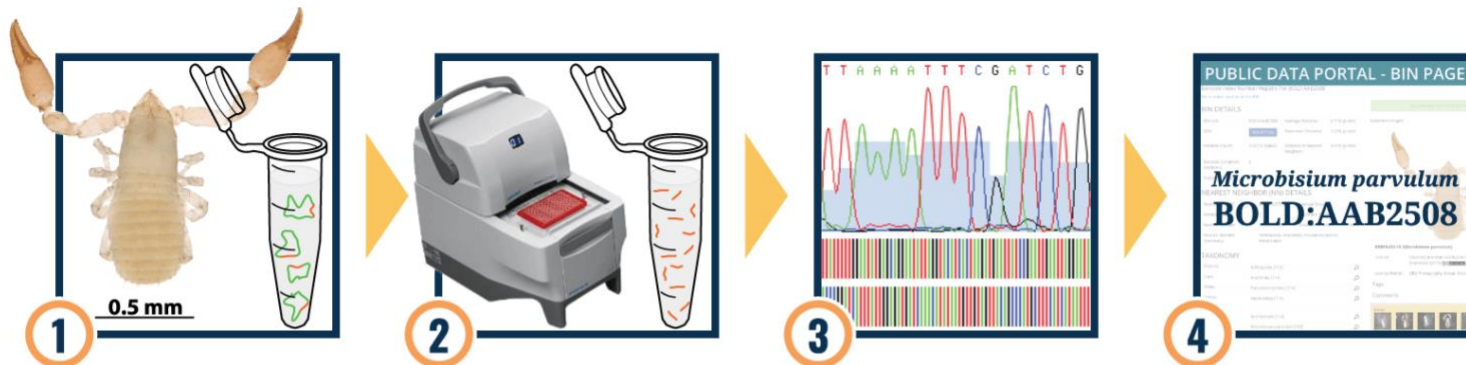


<https://ibol.org/about/dna-barcoding/>



# DNA BARCODING

Como funciona?



**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

<https://ibol.org/about/dna-barcoding/>

# THE DNA BARCODE

[HOME](#)[ABOUT](#) ▼[PROGRAMS](#) ▼[NEWS & MEDIA](#) ▼[JOIN](#)

Plant barcoding studies use one or a few plastid regions (e.g. *rbcL* and *matK*, and the non-coding spacer *trnH-psbA*) and the internal transcribed spacer (ITS) region of nuclear ribosomal DNA.



Animal barcoding studies use a region in the mitochondrial cytochrome *c* oxidase 1 gene (“CO1”).



Fungal barcoding studies use the internal transcribed spacer (ITS) region in the nuclear ribosomal cistron. This region shows reasonable discriminatory power at the species level in many groups.

[f](#)[t](#)[in](#)

# DNA BARCODING

[HOME](#)[ABOUT ▾](#)[PROGRAMS ▾](#)[NEWS & MEDIA ▾](#)[JOIN](#)

## THE LIBRARY

Barcode sequences are placed in the Barcode of Life Data Systems (BOLD) database – an online workbench that includes a reference library of DNA barcodes that can be used to assign identities to sequences of unknown origin.

BOLD is a searchable repository for barcode records, storing specimen data and images as well as sequences and trace files. It provides an identification engine based on the current barcode library and monitors the number of barcode sequence records and species coverage.



# BARCODE OF LIFE DATA SYSTEM v4

---

**Advancing biodiversity science through  
DNA-based species identification.**

<https://www.boldsystems.org/>

# QBOL – DNA Barcoding para pragas quarentenárias vegetais



Development of a new diagnostic tool using DNA barcoding to identify quarantine organisms in support of plant health

## Continue to

About QBOL

e-Newsletter

Training courses

Meeting 2012

DNA Barcoding

Activities

Q-lists

Databases

Collaboration

Partners

## QBOL project has come to an end

On 21 September 2012 the QBOL project ended officially. Many results were obtained in the different work packages.

> Read more

Centralizado na Europa

<https://www.qbol.org/en/qbol.htm>





DNA Barcoding

**Activities**

Q-lists

Databases

C

Barcoding of  
Arthropods

Barcoding of Bacteria

Barcoding of Fungi

Barcoding of  
Nematodes

Barcoding of  
Phytoplasm

Barcoding of Viruses

DNA Banks

Library/Database/Informa

Validation/Evaluation

Dissemination



## Q-BOL Activities

**QBOL Activities:**

# DNA Barcoding de fungos – CBS na Holanda



DNA Barcoding

**Activities**

Q-lists

Databases

Collaboration

Partners

Barcoding of  
Arthropods

Barcoding of Bacteria

**Barcoding of Fungi <**

Barcoding of  
Nematodes

Barcoding of  
Phytoplasmas

Barcoding of Viruses

DNA Banks

Library/Database/Informa

Validation/Evaluation

Dissemination



## Barcoding of Fungi

In this work package a short list of 19 Q-species were selected for barcoding.

### Contact

Ewald Groenewald

[View profile >>](#)

WP Coordinator: KNAW-CBS  
(Partner 9)

KNAW-CBS, PRI, FERA and CIP





DNA Barcoding

**Activities**

Q-lists

Databases

Collaboration

Partners

Barcoding of  
Arthropods

Barcoding of Bacteria

Barcoding of Fungi

**Barcoding of  
Nematodes** <

Barcoding of  
Phytoplasmas

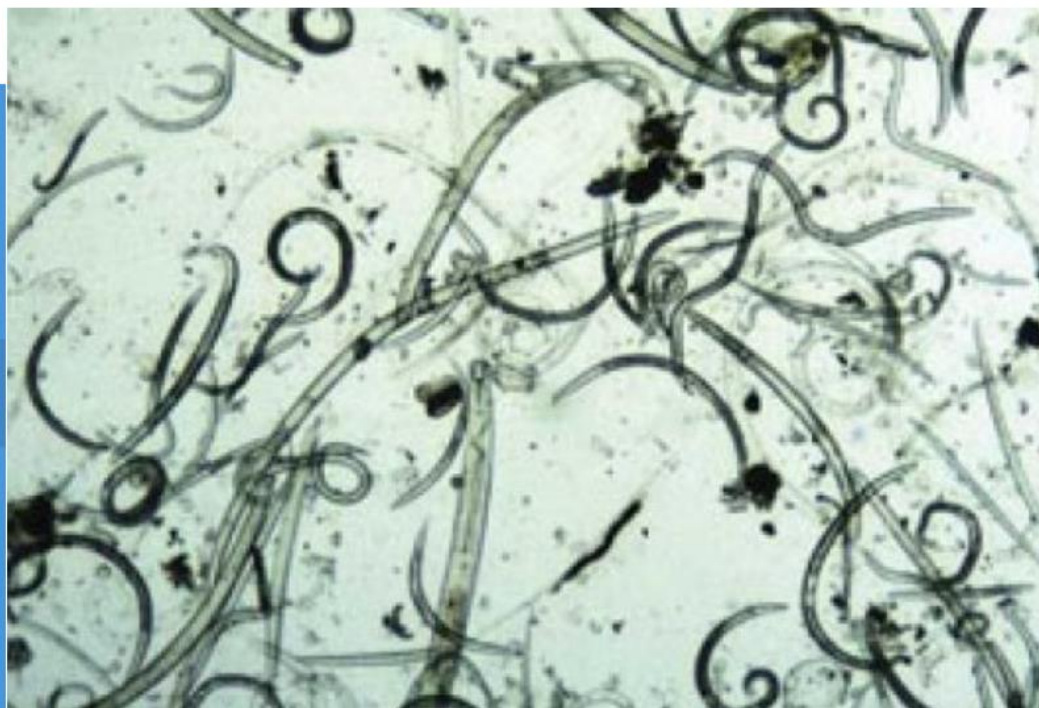
Barcoding of Viruses

DNA Banks

Library/Database/Informa

Validation/Evaluation

Dissemination



## Barcoding of Nematodes

In this work package a base list of 32 nematode species was created for which barcodes needed to be collected

### Contact

Juerg Frey

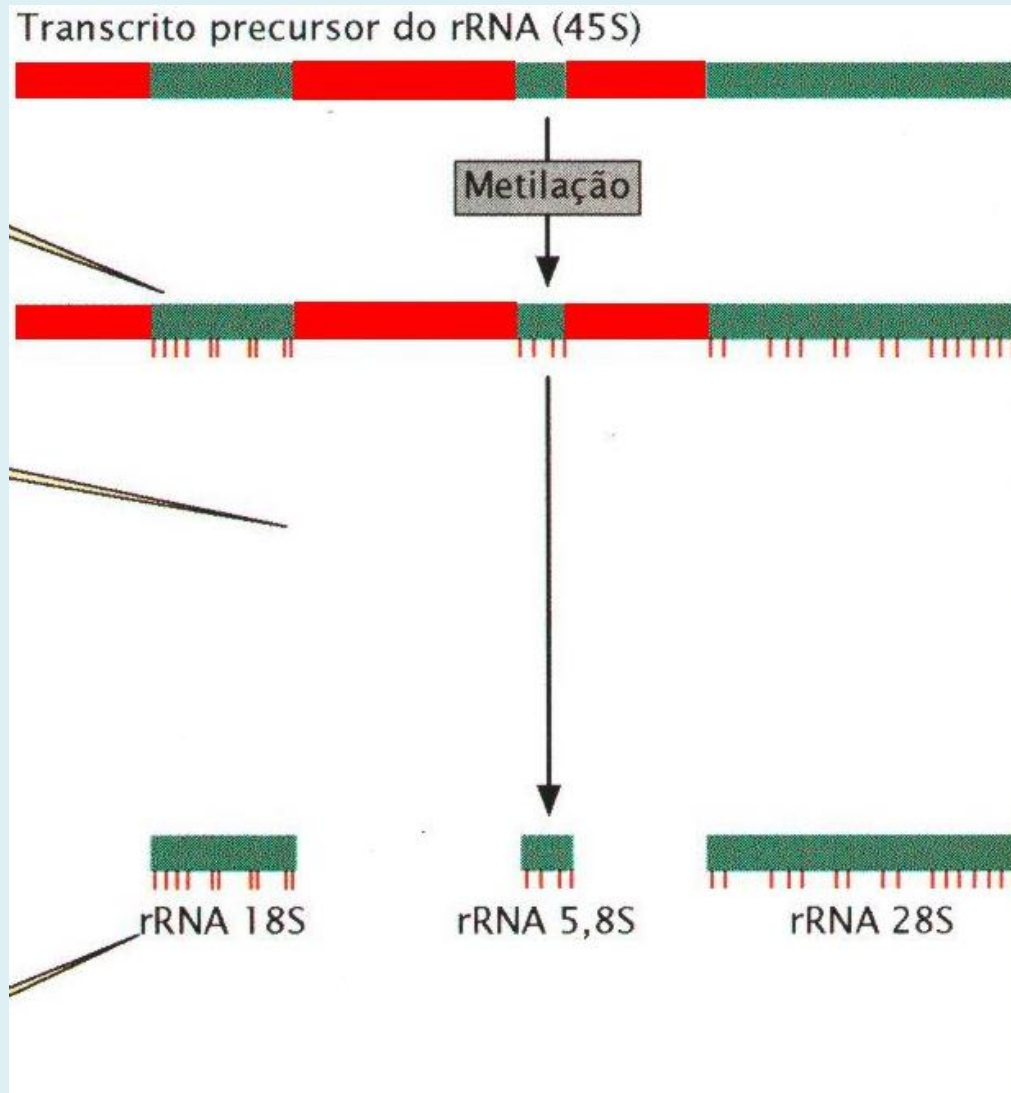
[View profile >>](#)

WP Coordinator: ACW (Partner 5)

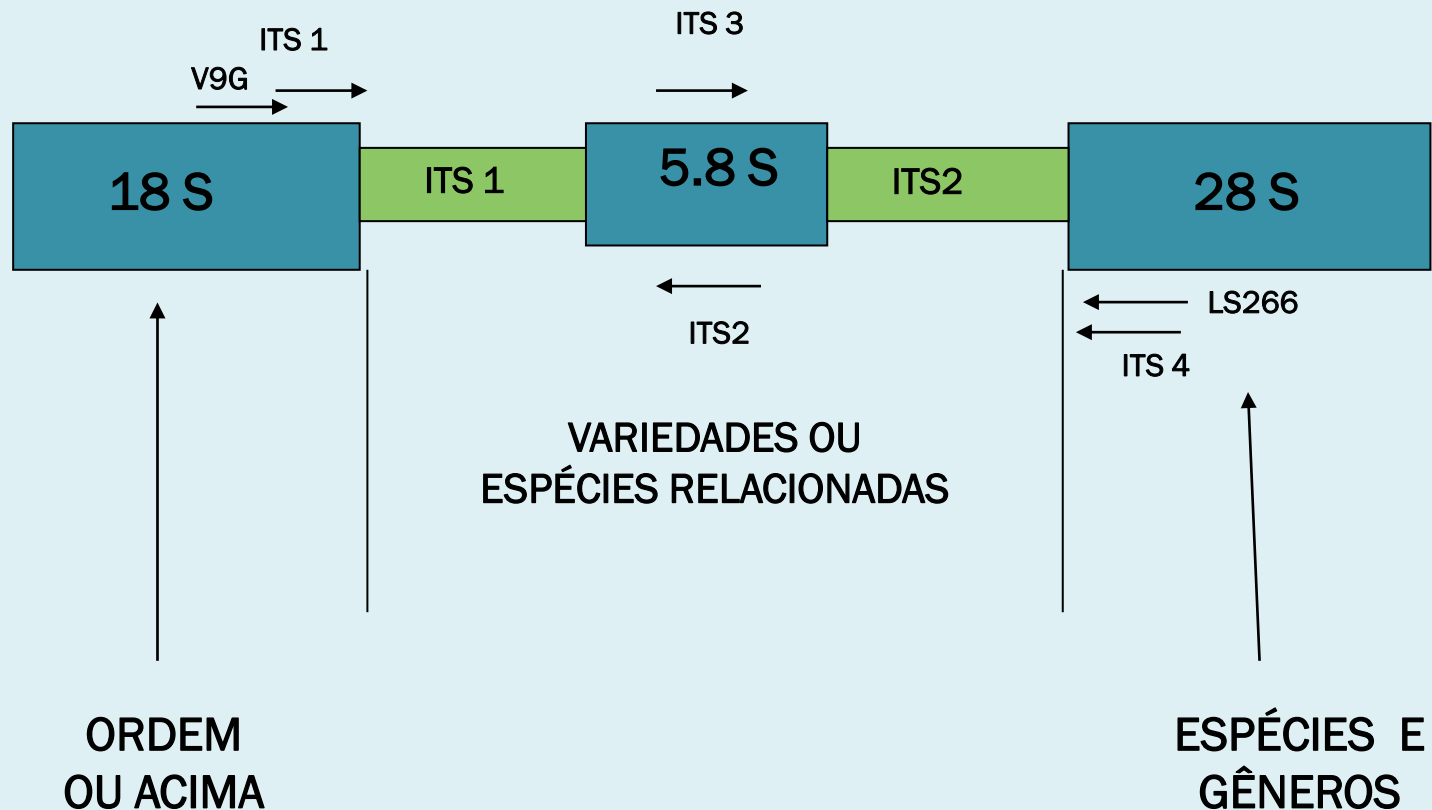
ACW, PRI, VLAGIEW (ILVO).  
INRA and CIP



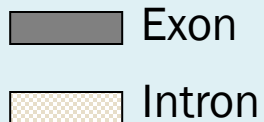
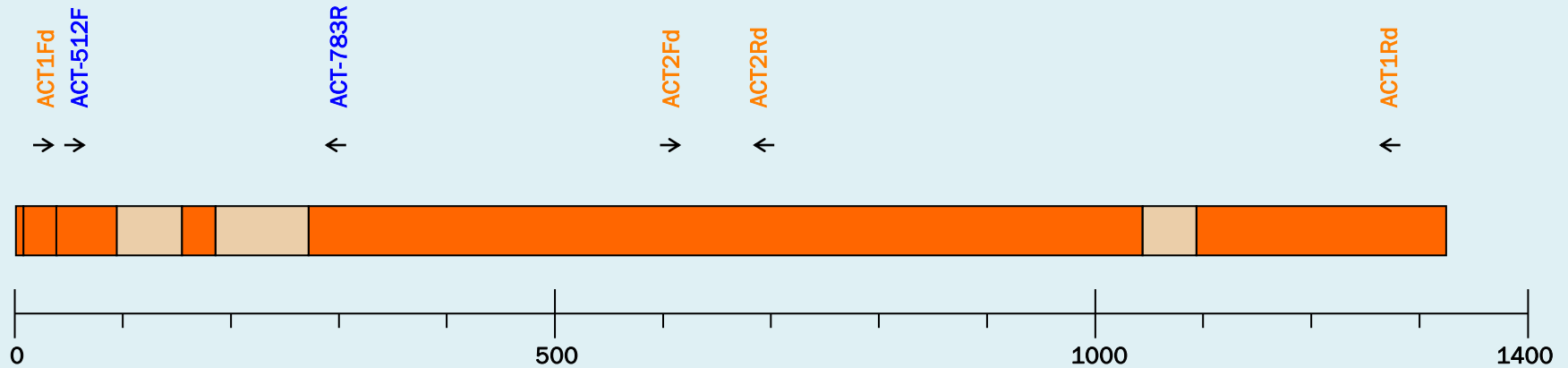
# RNA RIBOSSOMAL EUCARIÓTICO



# REGIÃO ITS1-5.8S-ITS2 DO rDNA

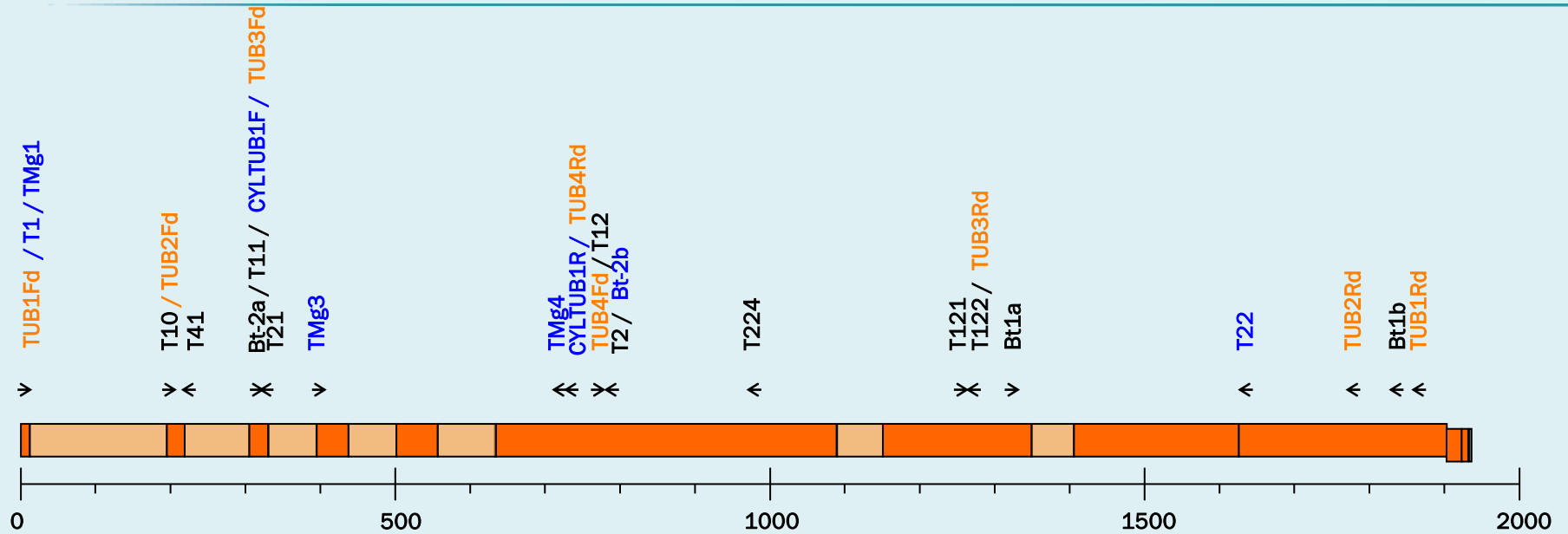


# ACTIN (1325 bp total; 1128 coding for 376 amino acids)



Primer available  
New primer  
Published primer

# BETA-TUBULIN (1933 bp total; 1344 coding for 448 amino acids)

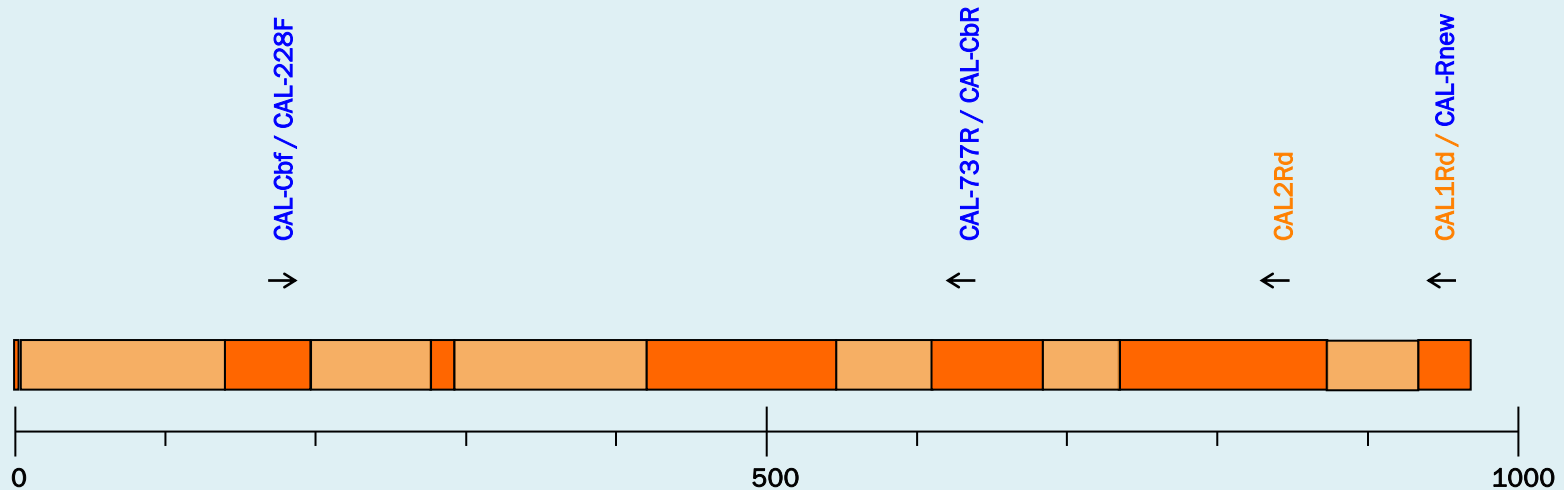


Exon

Intron

Primer available  
New primer  
Published primer

# CALMODULIN (970 bp total; 450 coding for 150 amino acids)

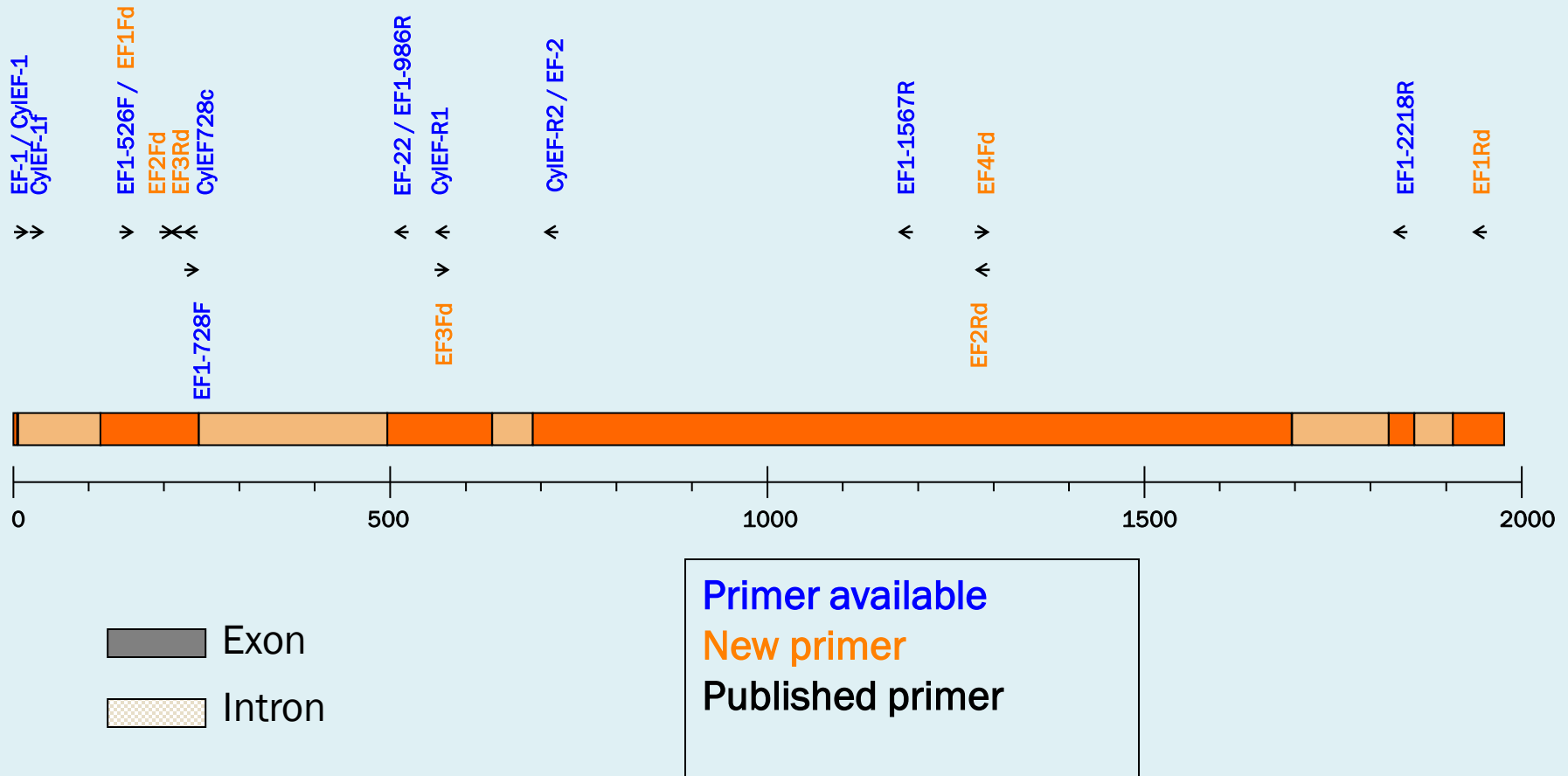


Exon  
Intron

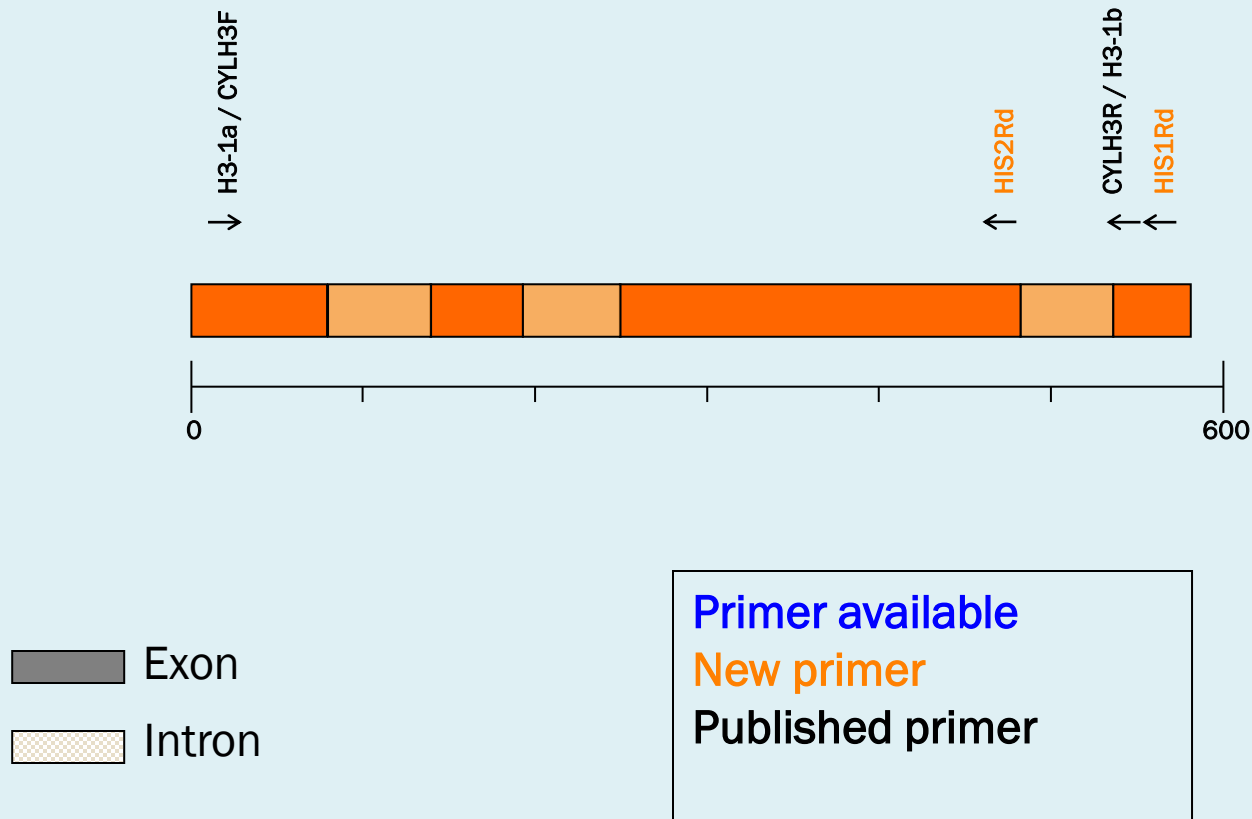
Primer available  
New primer  
Published primer

# ELONGATION FACTOR 1-ALPHA

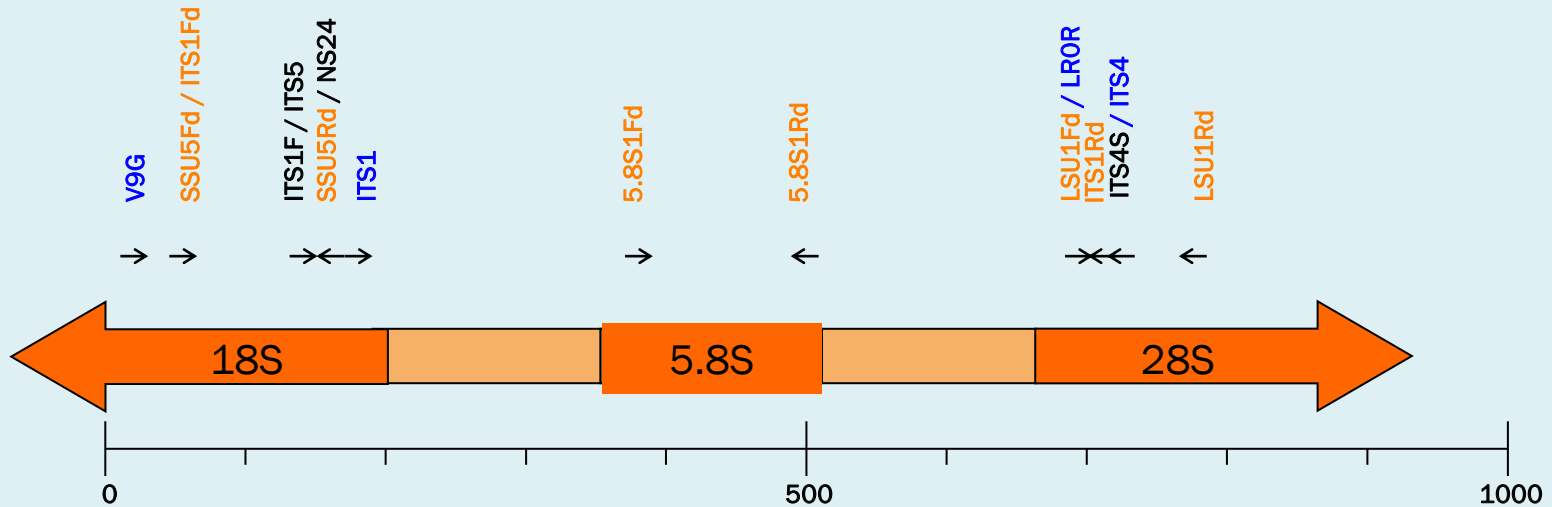
(1975 bp total; 1383 coding for 461 amino acids)



# HISTONE H3 (582 bp total; 411 coding for 137 amino acids)



# ITS1, ITS2 & 5.8S rRNA GENE (475 bp, 158 coding)

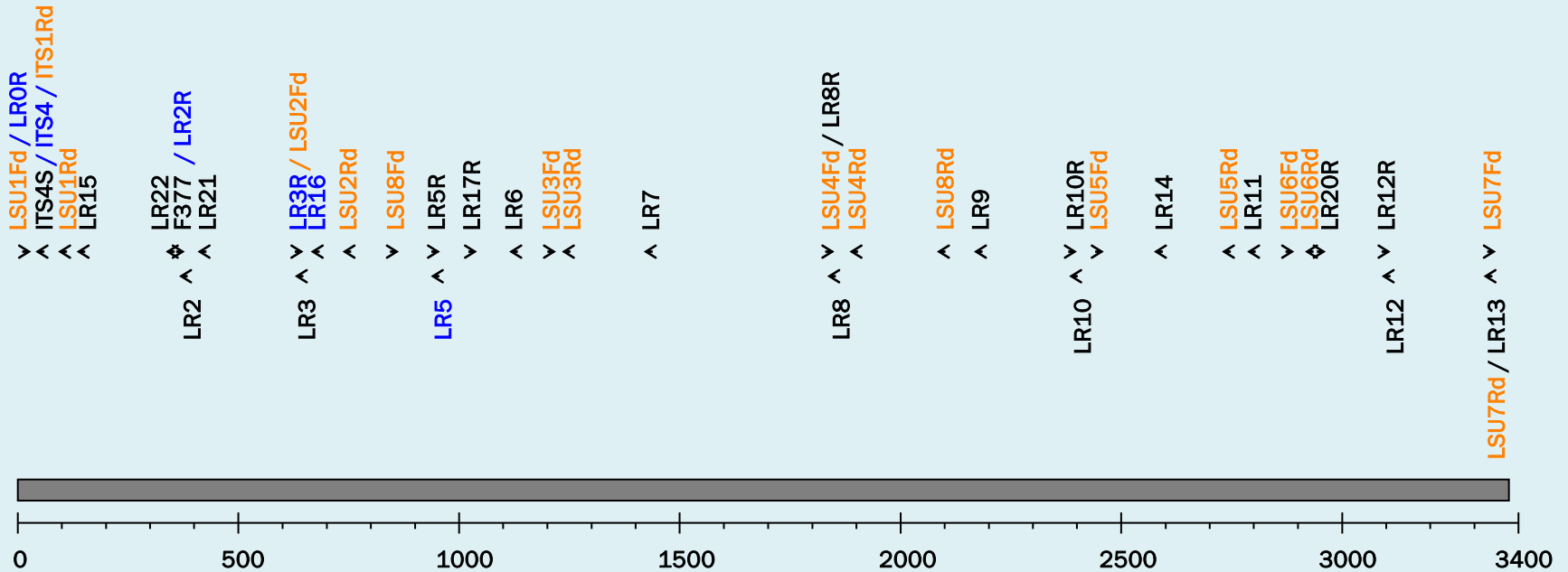


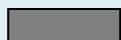
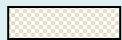
■ Coding  
■ Insertion

Primer available  
New primer  
Published primer



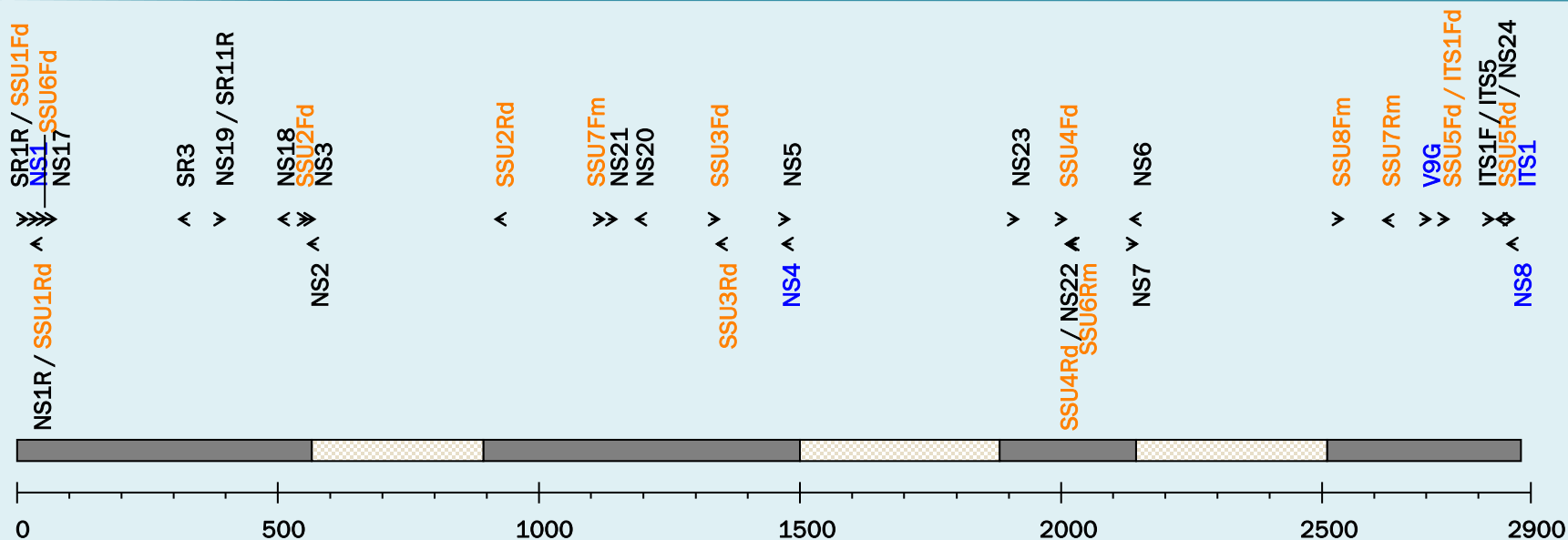
# LARGE SUBUNIT (28S) rRNA GENE (3378 bp)

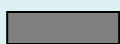



 Coding  
 Insertion

Primer available  
New primer  
Published primer

# SMALL SUBUNIT (18S) rRNA GENE (2880 bp total, 1803 bp coding)



 Coding  
 Insertion

Primer available  
 New primer  
 Published primer

---

# Exemplos

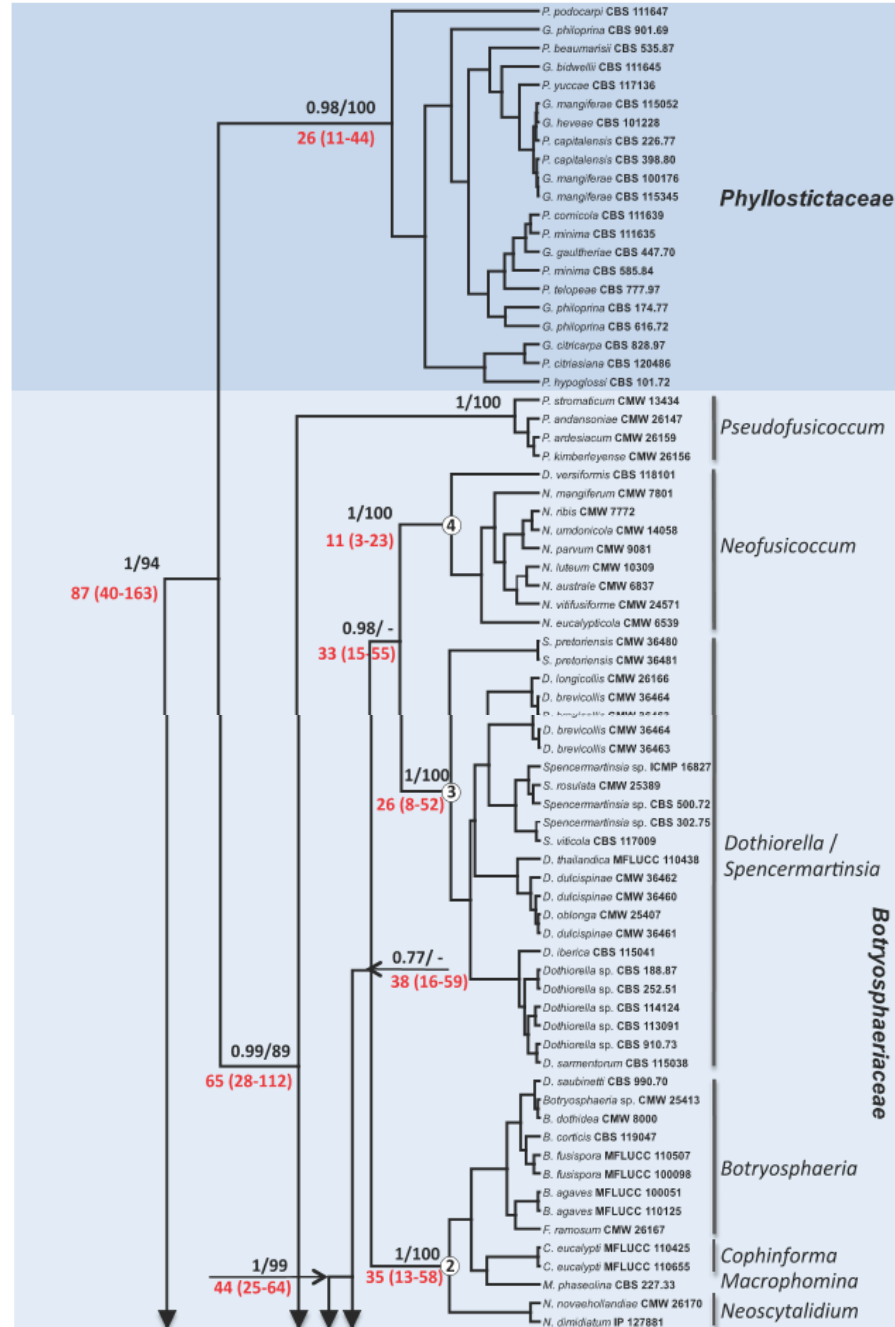
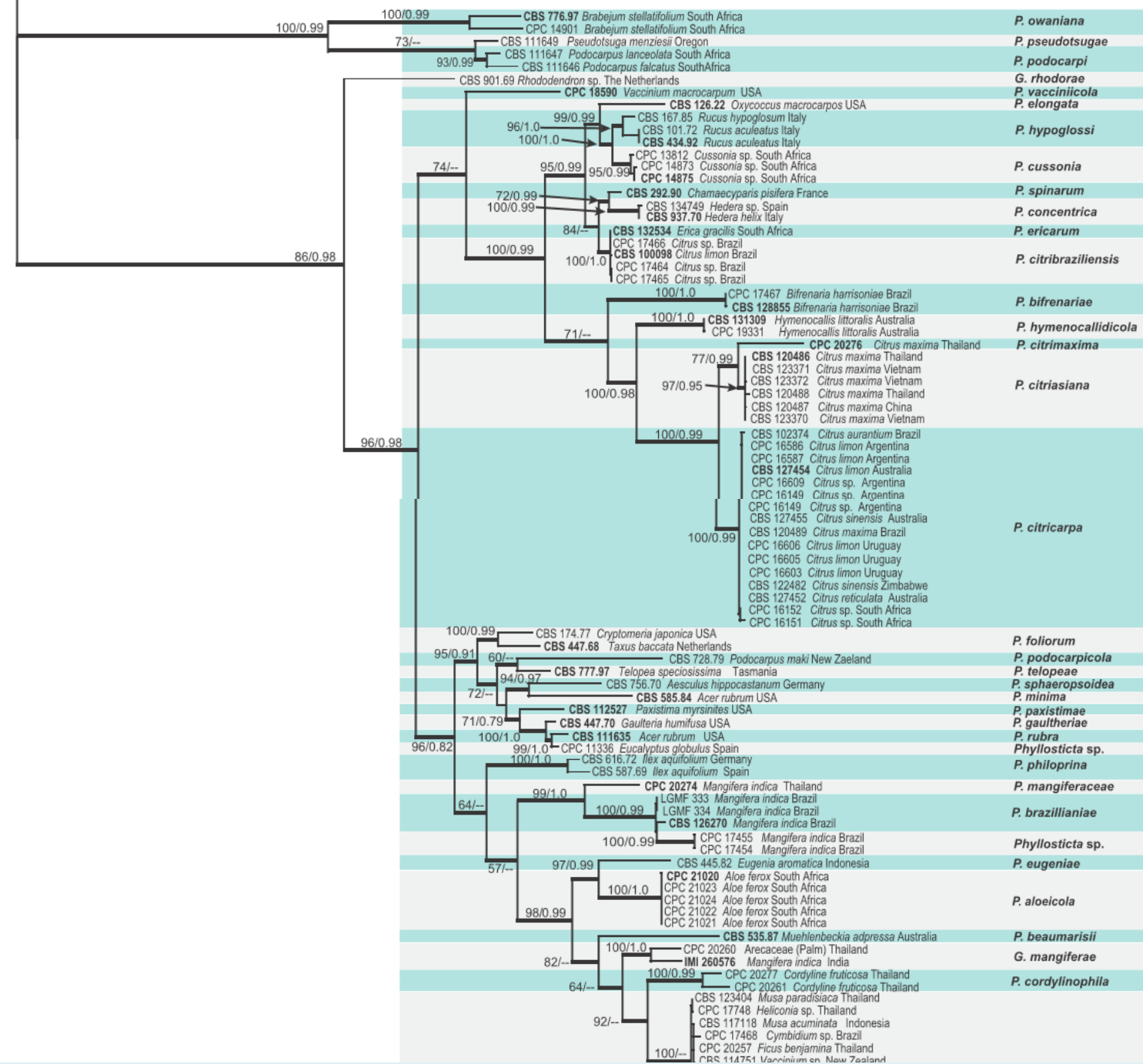


Fig. 1. Phylogenetic relationships of the Botryosphaerales using Bayesian reconstruction and six gene portions (LSU, SSU, ITS, EF1, BT and mtSSU). Numbers above branches indicate bootstrap values/posterior probabilities. Numbers highlighted in red below branches indicate estimated dates in million years with the 95 % Highest Posterior Density interval given in brackets. Clades 1–4 in the Botryosphaeriaceae are indicated by a circled number on the corresponding node.



# SEQUENCIANDO....



# COMO FUNCIONA O SEQUENCIAMENTO?

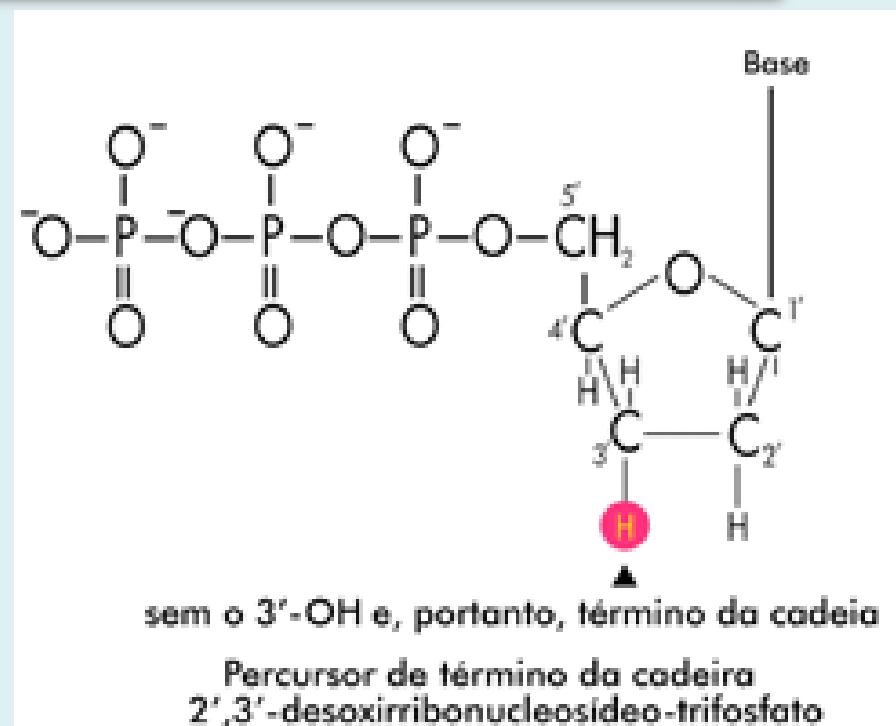
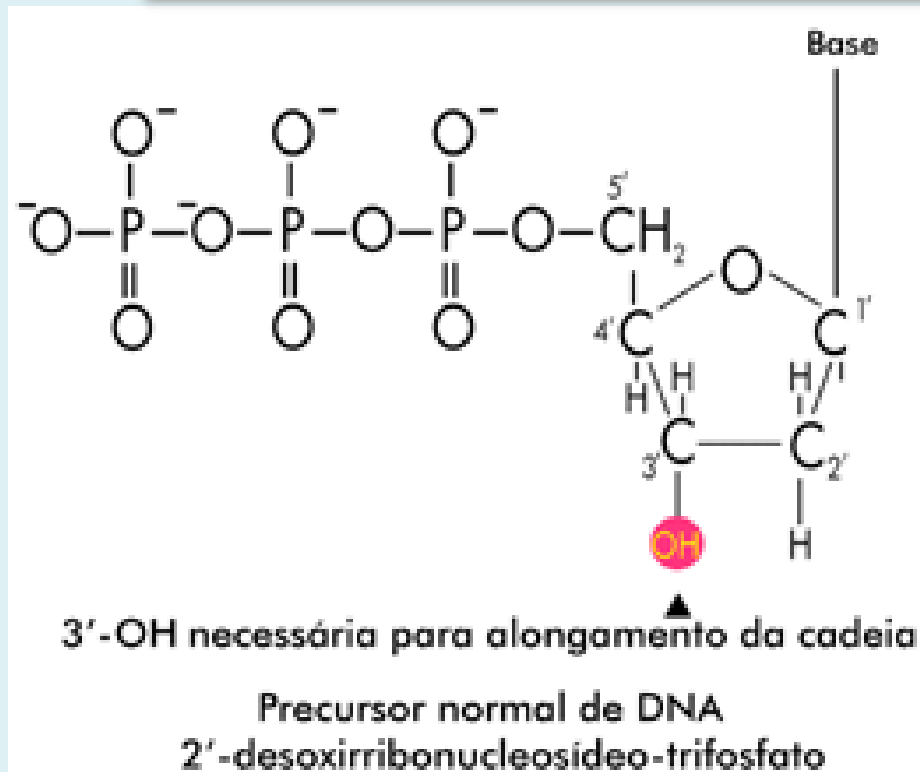
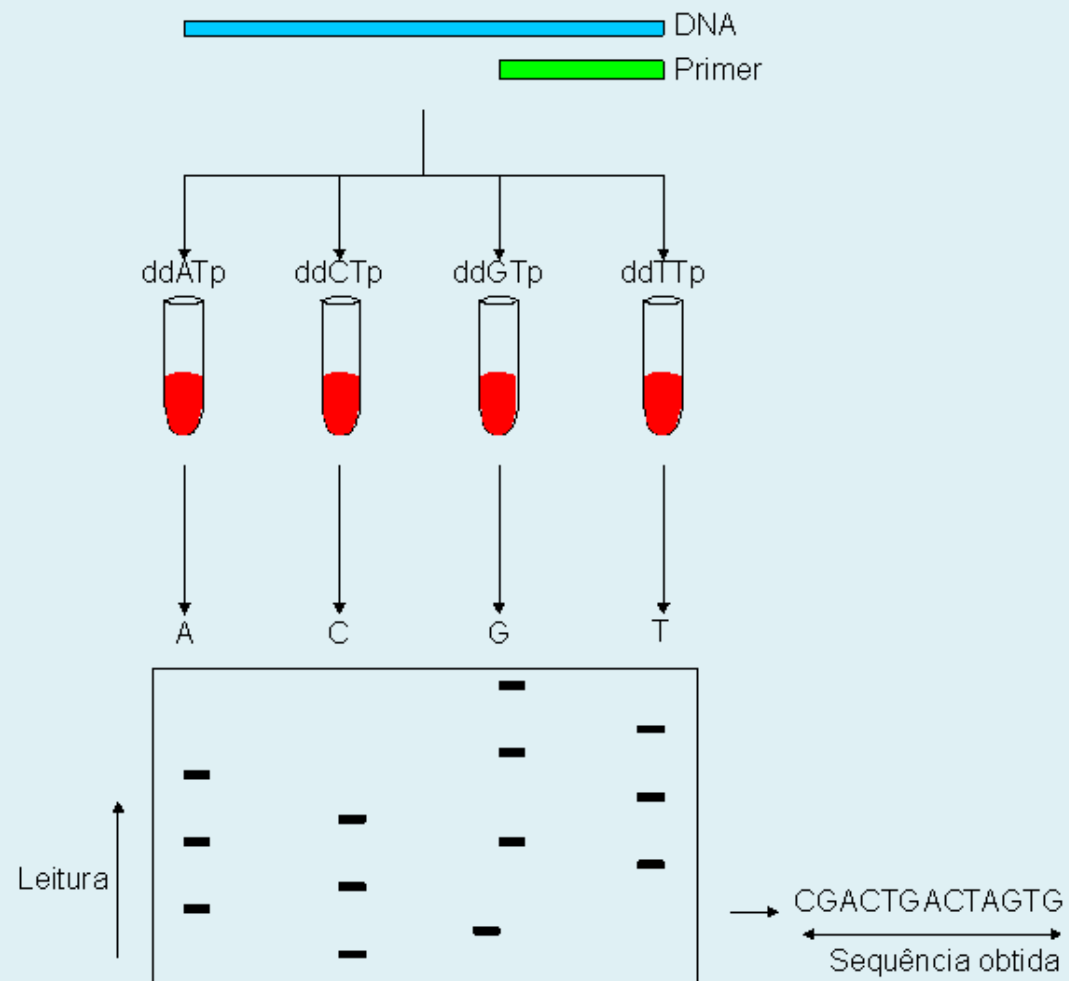


FIGURA 1. Deoxiribonucleotídeos (dNTP) e dideoxiribonucleotídeos trifosfatados (ddNTP).

# SEQUENCIAMENTO

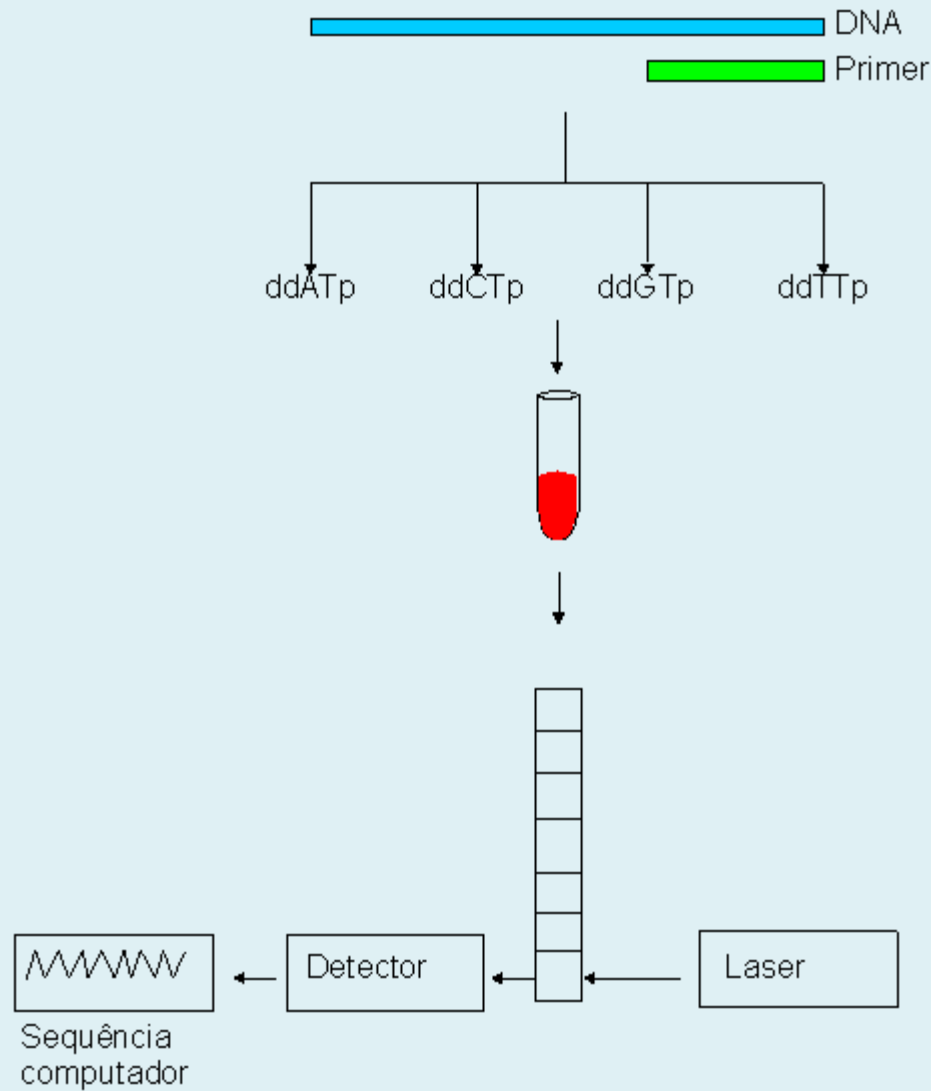




# SEQUENCIAMENTO

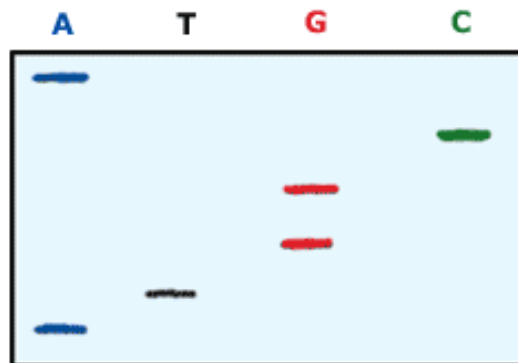


# SEQUENCIAMENTO



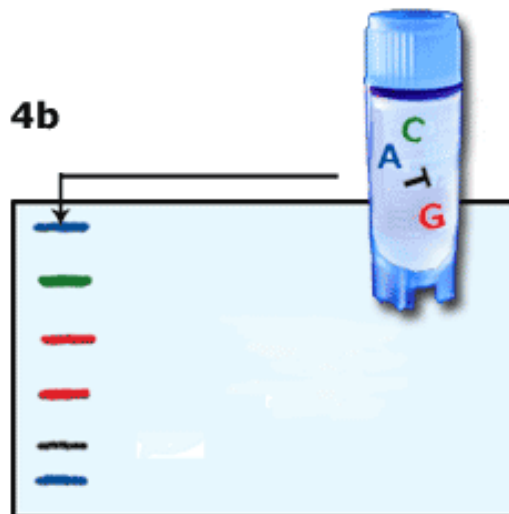
# SEQUENCIAMENTO

4a



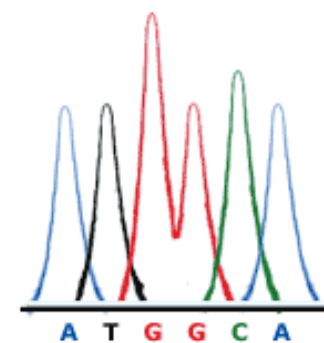
Análise dos produtos fluorescentes de sequenciamento em canais separados do gel de poliacrilamida.

4b



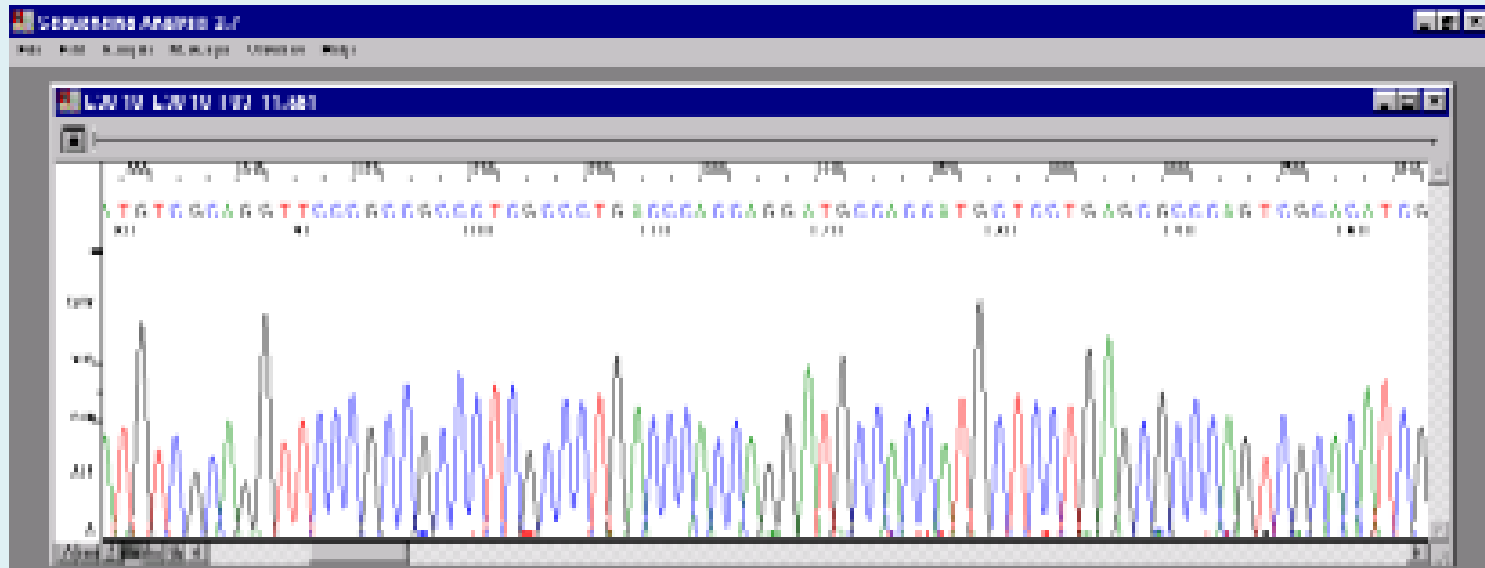
Análise dos produtos fluorescentes de sequenciamento em um único canal do gel de poliacrilamida.

4c



Eletroforegrama do gel

# SEQUENCIAMENTO



# SEQUENCIAMENTO

```
1  gtgcccagtg  agcgaggact  gca
61  ggggacccga  acaacatgga  tag
121  acgttccatg  gggccaagga  ggt
181  atgggcctca  tatacaacag  gat
241  tgtgccactt  gtgaacagat  tgc
301  accaccaatc  cactaatcag  gca
361  gctatggaac  agatggctgg  atc
421  cagactaggc  agatggtaca  tgc
481  ggtctgaaag  atgaccttct  tga
541  atgcagcgat  tcaagtgatc  ctc
```

(A) A simple sequence alignment

```
AGCAATGGCCAGACAATAATG
AGCTATGGACAGACATTAATG
***  ****  *****  ****
```

(B) A more difficult sequence alignment

```
GACGACCATAGACCAGCATAG
GACTACCATAGA - CTGCAAAG
***  ****  *  ***  **
```

```
GACGACCATAGACCAGCATAG
GACTACCATAGACT - GCAAAG
***  ****  ***  **
```

Two possible positions  
for the indel