

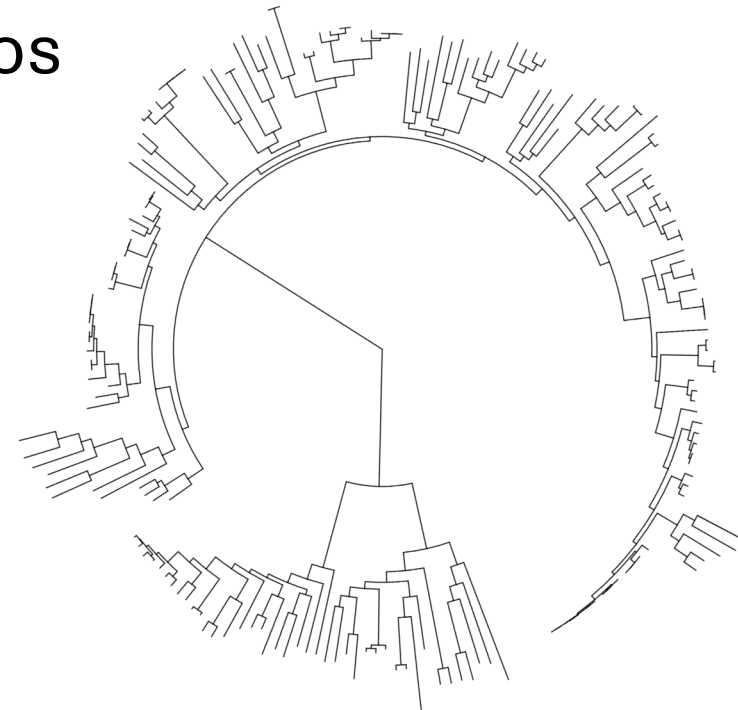


# Construcción y edición de árboles filogenéticos

## Día 3

# Contenido del curso

1. Conceptos clave
2. Preparación de datos y herramientas
3. Reconstrucción de árboles filogenéticos
4. Edición e interpretación de árboles filogenéticos



# Día 3: Reconstrucción de árboles filogenéticos

18:05 - 18:25

- Repaso
- Información preliminar

18:25 - 18:35

- Descarga de genes ortólogos
- Exploración de los datos

18:35 - 18:50 \*

- Alineamiento de secuencias
- Visualización del alineamiento

19:00 - 19:50

- Cálculo del método de sustitución
- Construcción del árbol filogenético

19:50 - 20:00

- Dudas y comentarios finales

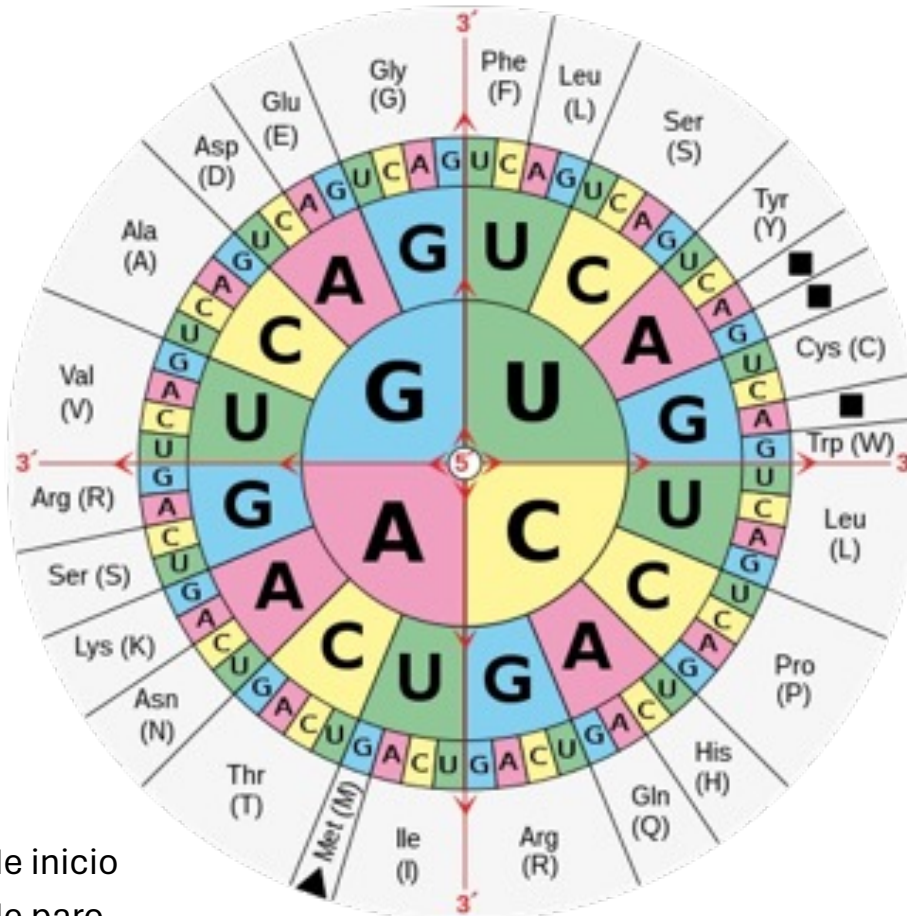


# Repaso



Información preliminar

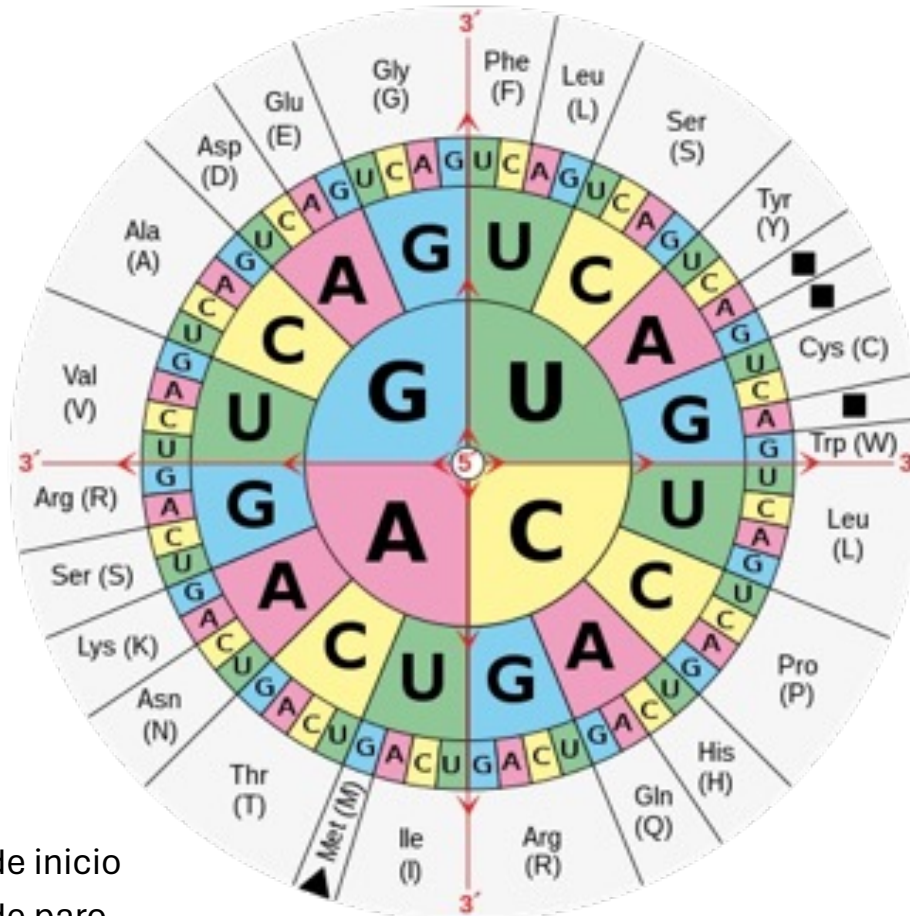
# Nucleótidos o proteínas para la reconstrucción de árboles, ¿cuál elijo?



- ▶ Codón de inicio
- Codón de paro

- **Objetivo del estudio**  
Pregunta biológica e historia evolutiva
- **Nivel de conservación**  
Las proteínas suelen estar más conservadas que los nucleótidos (código genético)
- **Resolución filogenética**  
Los **nucleótidos** pueden proporcionar una resolución filogenética más alta cuando se trabaja con especies estrechamente relacionadas  
Las **proteínas** son más útiles para resolver relaciones filogenéticas más profundas o divergencias evolutivas más antiguas

# Nucleótidos o proteínas para la reconstrucción de árboles, ¿cuál elijo?



- ▶ Codón de inicio
- Codón de paro

## Preferible nucleótidos:

- Casos de evolución reciente (especies estrechamente relacionadas)
- Análisis de mutaciones puntuales (SNPs, indels)

## Preferible aminoácidos:

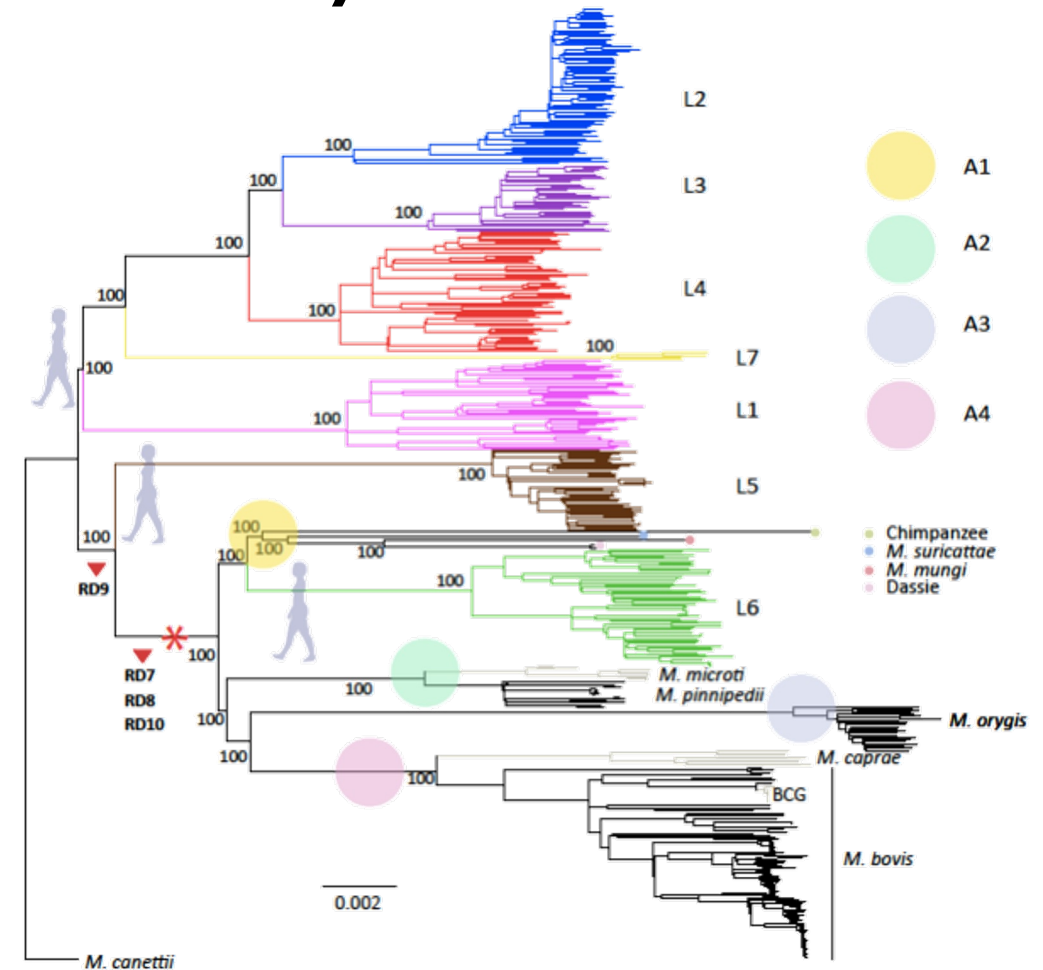
- Casos de divergencias evolutivas antiguas (especies lejanas)
- Análisis funcionales de proteínas



# ¿Qué es un **outgroup** (o grupo externo)?

- Es un taxón o grupo de organismos que se encuentra fuera del grupo principal de interés, pero que se utiliza como punto de referencia para determinar las relaciones evolutivas dentro del grupo principal
- Sirve como referencia externa que nos permite establecer la dirección y la naturaleza de los cambios evolutivos dentro de un grupo de interés

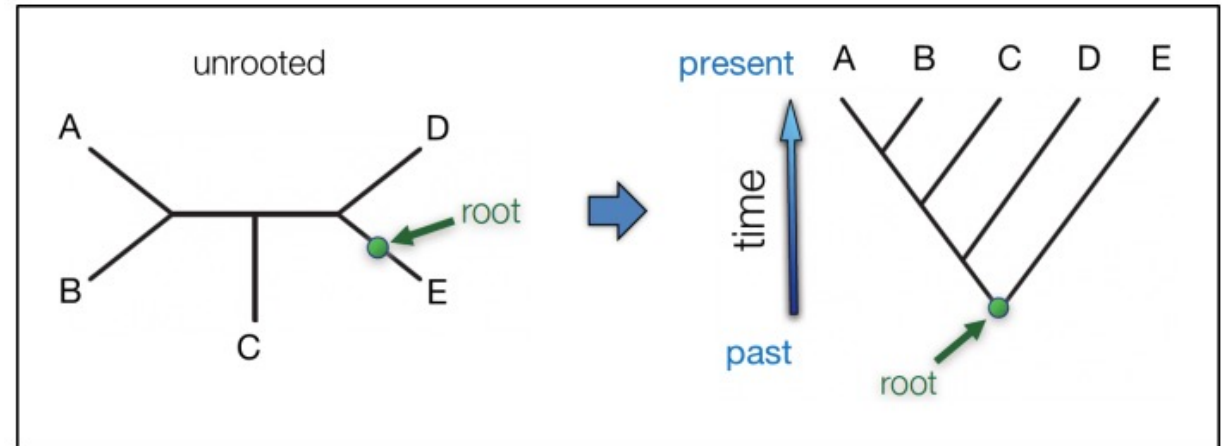
ML topology of 322 human-adapted and 529 animal-adapted MTBC members. Branch lengths are proportional to nucleotide substitutions and the topology is **rooted** with *Mycobacterium canettii*.





# ¿Qué son los *árboles enraizados* (o rooted tree)?

- Son árboles que contienen una “raíz” o outgroup
- La raíz representa al ancestro común más reciente de todos los taxas del árbol filogenético
- La raíz es la parte más antigua del árbol, por lo tanto, nos indica la dirección de la evolución desde la raíz hacia las ramas

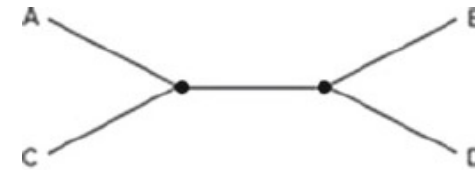


# Estrategias para obtener un árbol enraizado

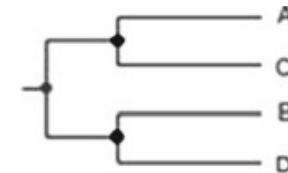
## Midpoint rooting

- Se asume que todas las secuencias del árbol evolucionan con la misma tasa. Esto no aplica a muchos casos biológicos
- En este caso, la raíz se coloca en el punto medio de las dos ramas más largas
- Limita la interpretación evolutiva en comparación con otras alternativas

Unrooted tree



Midpoint rooted

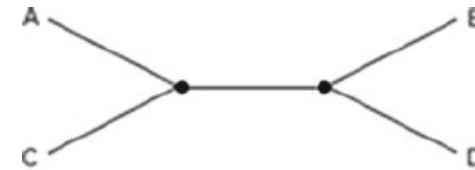


# Estrategias para obtener un árbol enraizado

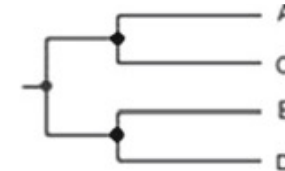
## Outgroup rooting

- Es el método más recomendado
- Incluir una o más secuencias “outgroup” al análisis que sabemos son más distantes a las del grupo de interés
- La raíz estimada será entonces el punto en el cual se conecta el outgroup con el resto de las secuencias de interés

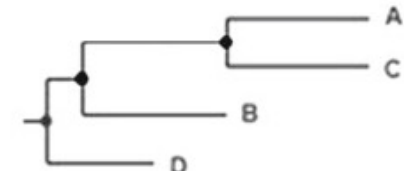
Unrooted tree



Midpoint rooted

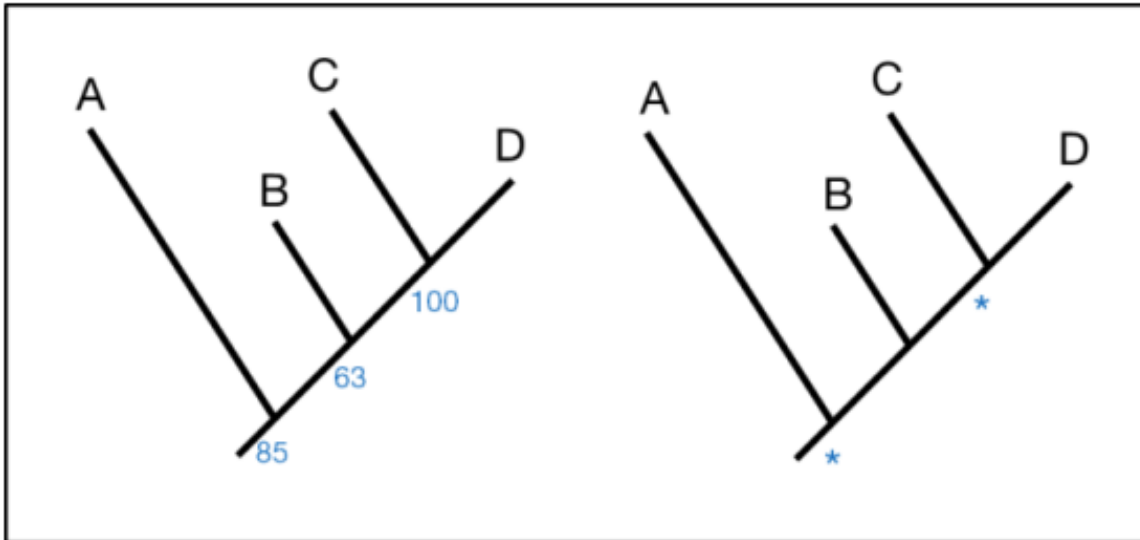


Outgroup rooted



# ¿Qué es el **bootstrap**?

Es una estrategia para estimar la confianza de un árbol filogenético inferido

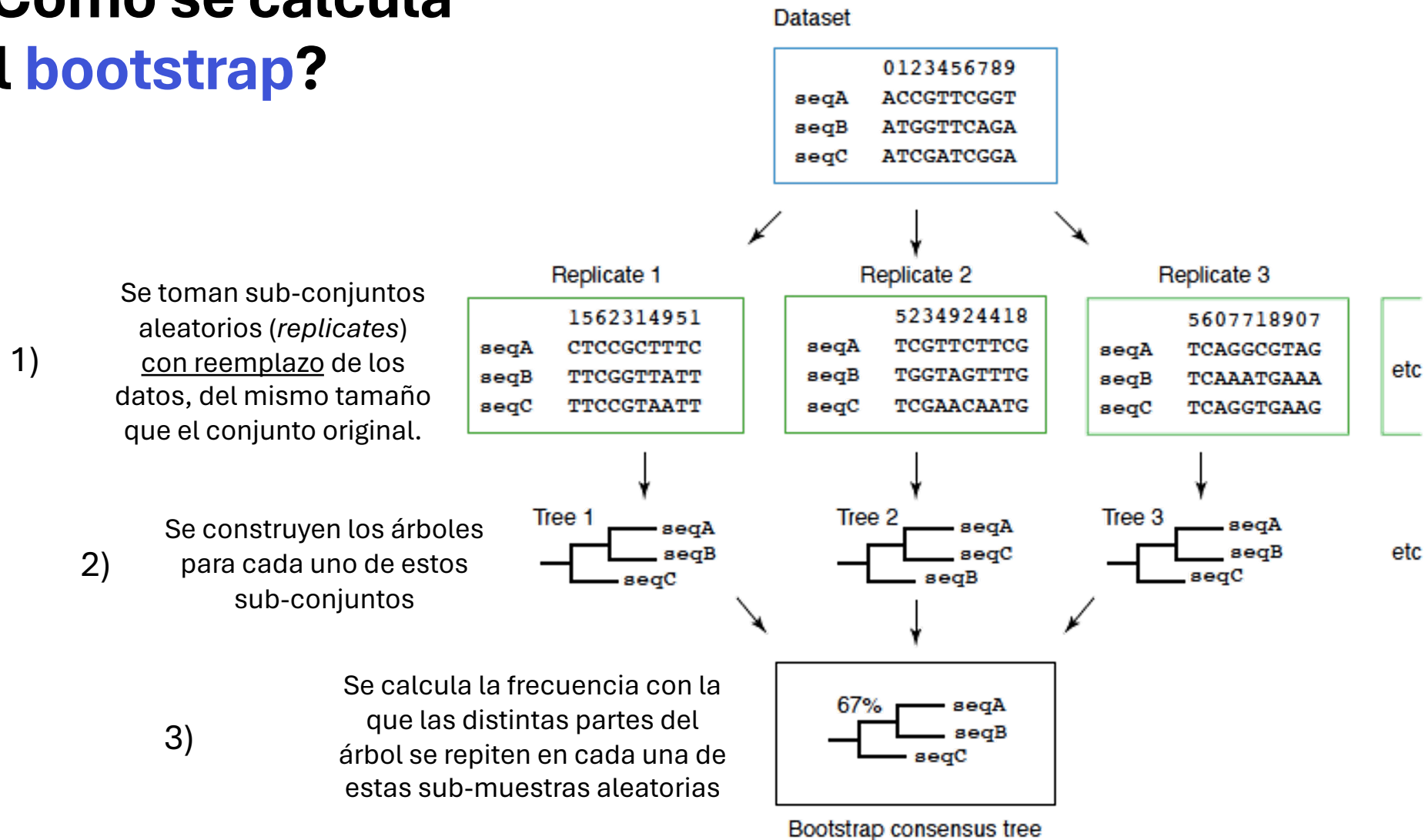


**Valores  $\geq 70\%$  de bootstrap probablemente indican agrupaciones confiables**

*"Existe un soporte constante (100% bootstrap) de que los taxa C y D están más estrechamente relacionados entre sí que con B"*

*"A partir de estos datos no queda claro que B, C y D sean los parientes más cercanos entre sí".*

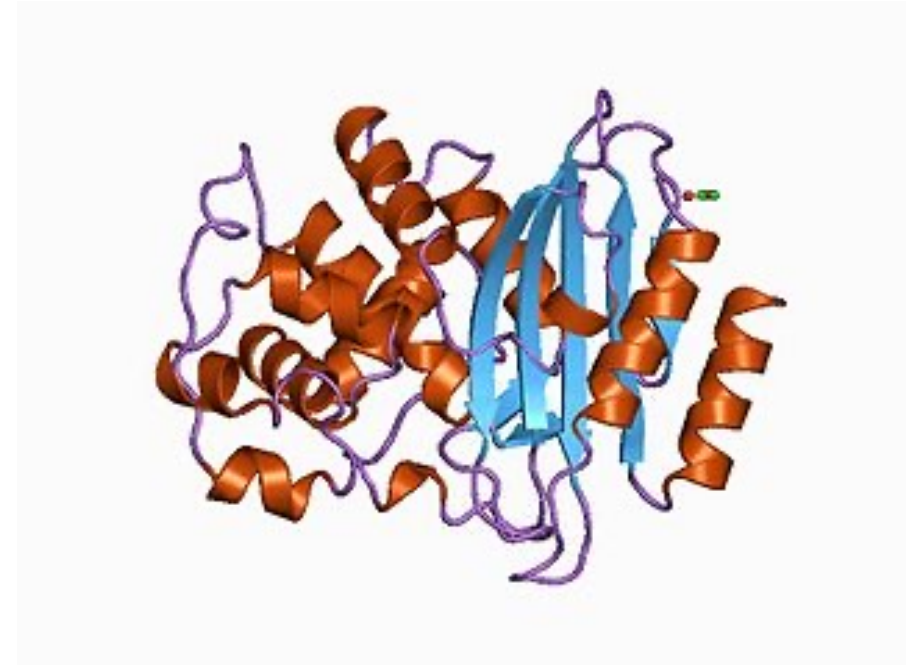
# ¿Cómo se calcula el bootstrap?



# Beta-lactamasas

# Árbol de genes: **Beta-lactamasas**

- Es una proteína presente en algunas bacterias que se encarga de degradar los anillos beta-lactámicos, que forman parte de varios antibióticos incluyendo las penicilinas, cefalosporinas, monobactámicos y carbapenémicos





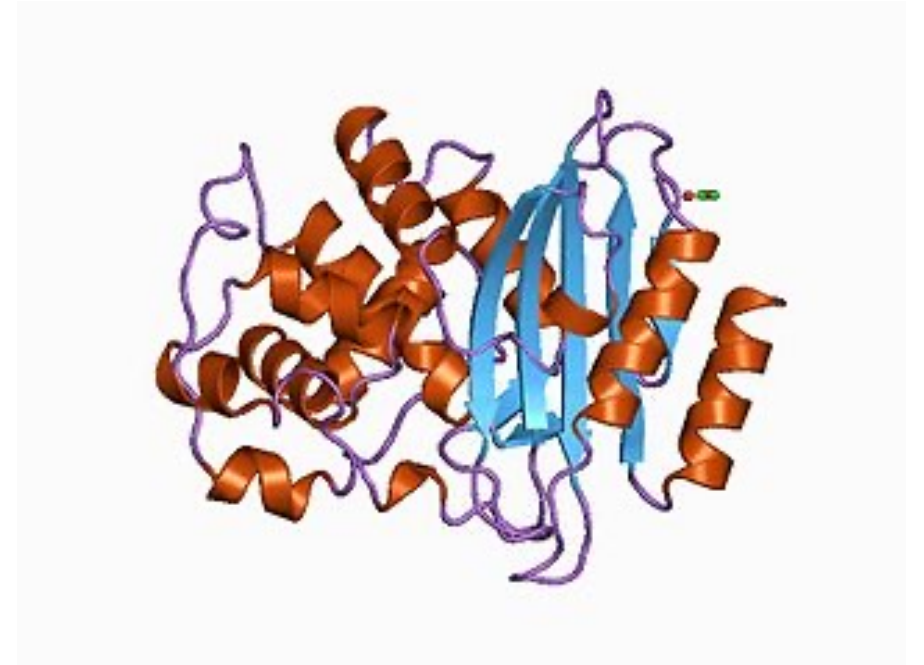
# Árbol de genes: **Beta-lactamasas**

## Pregunta:

¿Cómo han evolucionado las beta-lactamasas en bacterias?

## Objetivo:

Crear un árbol filogenético del gen de las beta-lactamasas considerando únicamente genes ortólogos en distintas bacterias



Dirígete el siguiente link, donde encontrarás las secuencias que usaremos para generar el árbol filogenético de las beta-lactamasas:

<https://github.com/paumayell/Phylogenetic-tree-reconstruction>

# Exploración de los datos

# Visualizar las secuencias en MEGA

- Abrir el programa MEGA
- Abrir o arrastrar el archivo FASTA de proteínas de las beta-lactamasas a la ventana principal de MEGA

# Alineamiento de las secuencias

# Alinear las secuencias usando Muscle

- Selecciona el algoritmo de alineamiento (Muscle)
- Selecciona el tipo de secuencias (proteínas)
- Parámetros del alineamiento (default)

# Visualización del alineamiento



# Exploración de la calidad del alineamiento

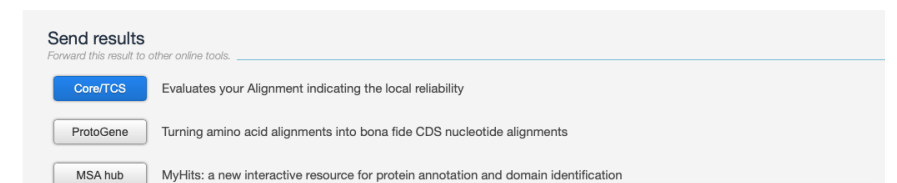
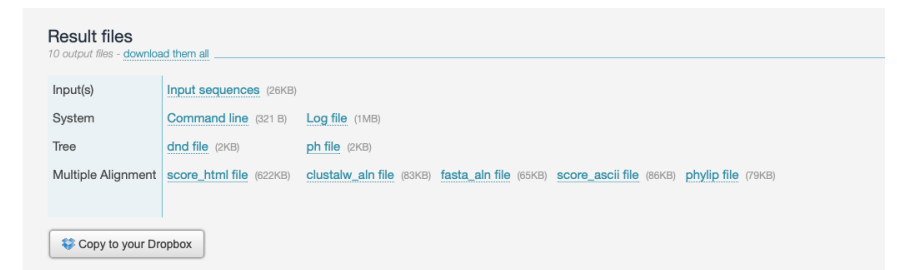
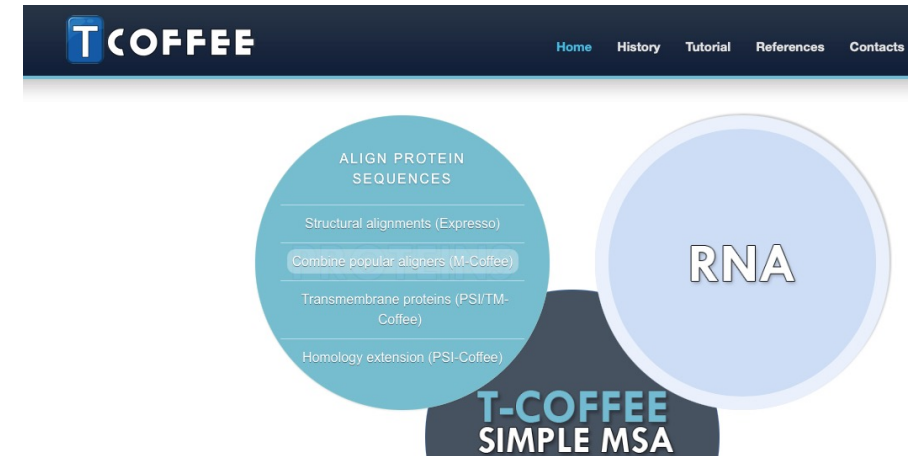
- Explora todos los taxos y revisa todo el alineamiento para ubicar la presencia de gaps
- Elimina las regiones no conservadas (exceso de gaps) que puedan meter ruido al análisis filogenético

## Alternativa:

- Alinear con el software online

[T-coffee](#) → Proteins → M-coffee

- Una vez que se obtiene el alineamiento, correr el análisis de **Core/TCS** análisis para obtener solo las regiones conservadas del alineamiento
- Guardar el resultado y cargarlo en MEGA



# Cálculo del modelo de sustitución

# Cálculo del modelo de sustitución

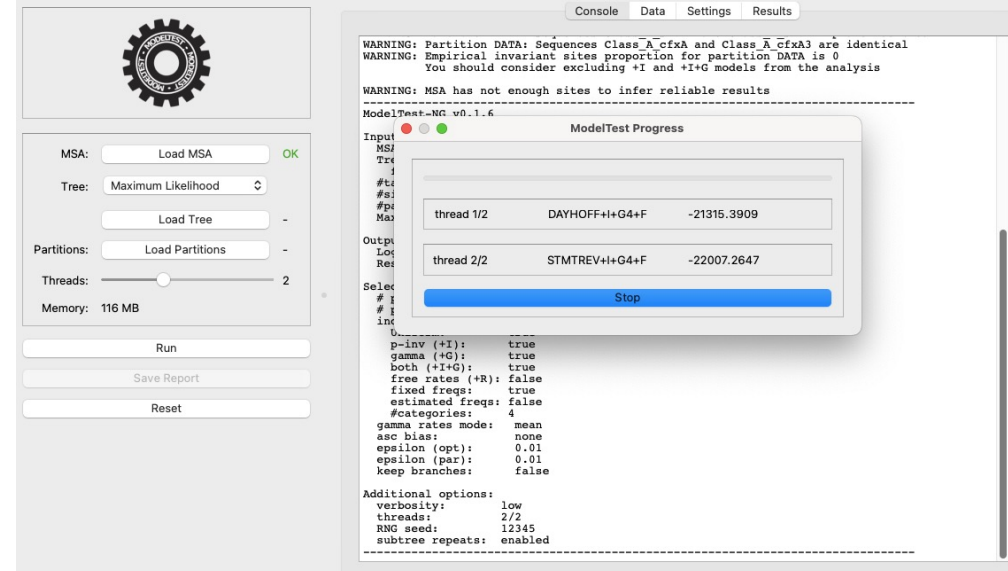
## MEGA

- Cargar el alineamiento de regiones conservadas y ejecutar el análisis de alineamiento
- Guardar el alineamiento en formato MEGA
- Ejecutar el análisis del modelo de sustitución en la pestaña “Modelo”

**Nota:** Este proceso puede tardar varios minutos

## ModelTest-NG

- <https://github.com/ddarriba/modeltest>



# Reconstrucción del árbol filogenético

# Reconstrucción del árbol filogenético

## MEGA

- Una vez que hemos obtenido el modelo de sustitución para nuestros datos, elegimos el método para la reconstrucción del árbol filogenético (Maximum Likelihood) y ajustamos los parámetros indicando el método de sustitución calculado
- Exportar el resultado del árbol en formato newick (.nwk)

**Nota:** Este proceso puede tardar varios minutos

## RAXML-NG

- <https://github.com/amkozlov/raxml-ng>

Dudas y comentarios finales

# Contacto

Dra. Paulina M. Mejía-Ponce



[paumayell@gmail.com](mailto:paumayell@gmail.com)