

# CURSO: CC471 - 2018 -1

## Examen Sustitutorio.

Tiempo: 120 minutos. : Genere un documento de texto <Su Nombre><Apellido>-cc471-2018-1.odt  
Incluya todas las respuestas en ese documento.

1. Encuentre la secuencia codificadora CDS de mRNA del gen que codifica la proteína “Cytochrome c Somatic” en el organismo Homo Sapiens.

**a) (2 Puntos)** Provea el Nombre del Gen, su Gene ID, los enlaces en la base de datos respectivas. Y la secuencia codificadora (CDS) de mRNA.

R: Nombre del Gene: CYCS; Gene ID: 54205; Enlace del Gene :

<https://www.ncbi.nlm.nih.gov/gene/54205>

Enlace de mRNA: [https://www.ncbi.nlm.nih.gov/nuccore/NM\\_018947.5](https://www.ncbi.nlm.nih.gov/nuccore/NM_018947.5)

Secuencia Codificadora CDS:

```
atgggtgatg
181 ttgagaaagg caagaagatt tttattatga agtggtccca gtgccacacc gttgaaaagg
241 gaggcaagca caagactggg ccaaattctc atggtctctt tgggcggaag acaggtcagg
301 cccctggata ctcttacaca gccgccaata agaacaaagg catcatctgg ggagaggata
361 cactgatgga gtatttggag aatcccaaga agtacatccc tggaacaaaa atgatctttg
421 tcggcattaa gaagaaggaa gaaagggcag acttaatagc ttatctcaaa aaagctacta
481 atgagtaa
```

**b) (2 Puntos)** Además provea y señale en su reporte cual es la secuencia codificadora en formato FASTA, así como la secuencia traducida. Que se puede decir de la secuencia traducida?

R: Secuencia codificadora FASTA:

```
>NM_018947.5:171-488 Homo sapiens cytochrome c, somatic (CYCS), mRNA
ATGGGTGATGTTGAGAAAGGCAAGAAGATTTTATTATGAAGTGTTCCAGTGCCACACCGTTGAAAAGG
GAGGCAAGCACAAGACTGGGCCAAATCTCCATGGTCTCTTTGGGCGGAAGACAGGTCAGGCCCTGGATA
CTCTTACACAGCCGCAATAAGAACAAAGGCATCATCTGGGGAGAGGATACACTGATGGAGTATTTGGAG
AATCCAAGAAGTACATCCCTGGAACAAAAATGATCTTTGTCTGGCATTAAAGAAGAAGGAAGAAAGGGCAG
ACTTAATAGCTTATCTCAAAAAAGCTACTAATGAGTAA
```

Secuencia Traducida:

```
translation="MGDVEKGKKIFIMKCSQCHTVEKGGKHKTGPNLHGLFGRKTGQA
PGYSYTAANKNKGIIWGEDTLMEYLENPKKYIPGTMIFVGIKKKEERADLIAYLKKA
TNE"
```

La secuencia traducida es la cadena de aminoácidos de la proteína “Cytochrome c Somatic”

**c) (3 Puntos)** Provea lo siguiente: Un archivo denominado “cytochromec-org.fasta” con las siguientes secuencias en formato FASTA.

- La secuencia de Cytochrome c Somatic para “Homo Sapiens”
- Las secuencias de Cytochrome c para “mus musculus”, “rattus norvegicus”, “pan troglodytes”, “equus caballus” y “gorilla”

Ademas, provea un archivo con el alineamiento multiple de estas secuencias. Pongale nombre “CytochromeC-alineamiento.clw”. ¿Que herramienta o método utilizo para hacer el alineamiento?

R: Se ha provisto el archivo cytochromec-org.fasta con las secuencias solicitadas

El alineamiento se ha provisto en el archivo CytochromeC-alineamiento.clw

La herramienta utilizada fue Muscle <https://www.ebi.ac.uk/Tools/msa/muscle/> tambien se puede utilizar el comando clustalw.

**d) (2 Puntos)** A apartir del alineamiento anterior, calcule la matriz de distancias utilizando Bio Python.

Provea dicha matriz en su reporte, indicando claramente las columnas y filas con los organismos solicitados, es decir: (“Homo sapiens”, “mus musculus”, “rattus norvegicus”, “pan troglodytes”, “equus caballus” y “gorilla” )

R: La matriz de distancias es la siguiente:

[Homo	0					
[Gorilla	0	0				
[Pan	0	0	0			
[Equus	0.104761905	0.104761905	0.104761905	0		
[Mus	0.085714286	0.085714286	0.085714286	0.047619048	0	
[Rattus	0.085714286	0.085714286	0.085714286	0.047619048	0	0
	[Homo	[Gorilla	[Pan	[Equus	[Mus	[Rattus

**e) (3 Puntos)** Provea el código Python utilizado con comentarios en cada linea Ponga como nombre al programa “alineamiento.py”. Incluya una cabecera con los siguientes datos:

Examen sustitutorio CC471-2018-1

Problema N.º < 1e >

Autor: <Nombre> <Apellidos>

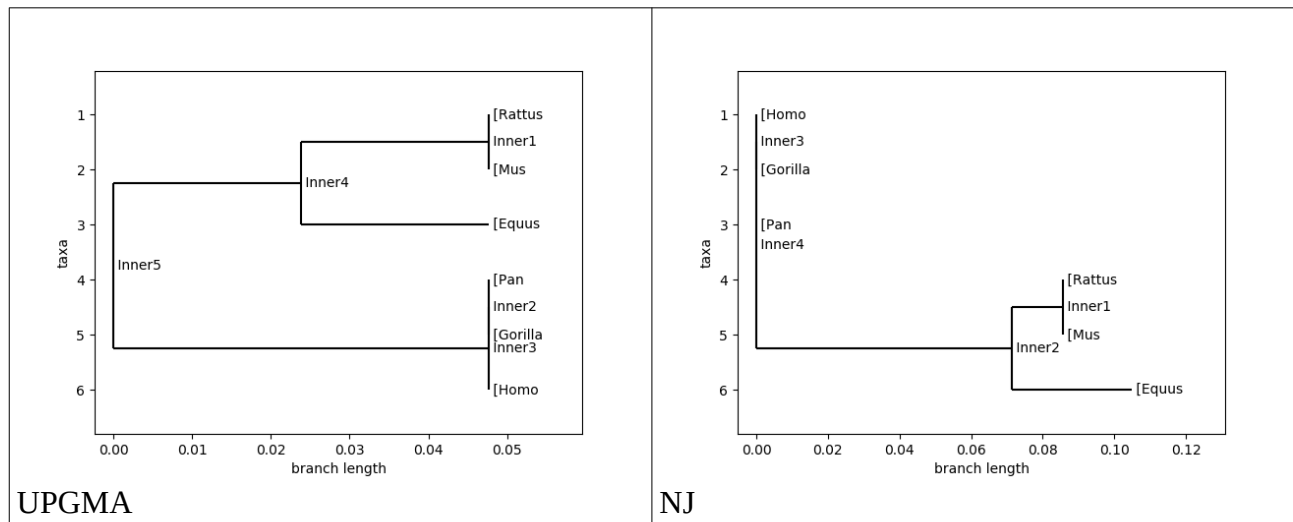
Nombre del programa: \_\_\_\_\_.

R: se provee el archivo “alineamiento.py” con el código solicitado.

**f)(4 Puntos)** Construya un arbol filogenetico utilizando los métodos UPGMA y Neighbour Joining (NJ) con la matriz de distancias obtenidas anteriormente. Puede tomar como base el código de “alineamiento.py” y completarlo para generar los árboles solicitados, incluya los comentarios correspondientes.

Provea el código en python con el nombre “arbol-filogenetico.py” (Actualice la cabecera del programa y comente el código) incluya en su reporte del examen los 2 gráficos en formato de imagen (p.ej. UPGMA.png y NJ.png)

R: se provee el archivo “arbol filogenetico.py” con el código solicitado



2. (4 Puntos) Describa y provea las características estructurales de Human Cytochrome C Y48h incluya los enlaces web correspondientes. Provea el archivo con la estructura correspondiente en formato xml. ¿ Cual es el elemento que se encuentra dentro de los tags:

<OrgName\_lineage> ... </OrgName\_lineage>

R:

Enlaces:

<https://www.ncbi.nlm.nih.gov/Structure/pdb/5EXQ>

Archivo provisto: 5EXQ.cgi

Elemento dentro del tag indicado:

<OrgName\_lineage>Eukaryota; Metazoa;  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires;  
Primates; Haplorrhini; Catarrhini; Hominidae;  
Homo</OrgName\_lineage>

PDB ID 5EXQ: Human Cytochrome C Y48h

