



BIOG1020: Bioinformática

Examen Final

Fecha: 29/08/2018

Alumno:

Observaciones: Preguntas A y B: Prof Julio Bonilla. Preguntas C y D: Profa Diana Cardenas. Pregunta E: Profa Fernanda Cordeiro

A. Los siguientes enumerados deben ser resueltos de manera individual. Cada respuesta debe ir respondida de forma escrita como gráfica (imágenes de la pantalla donde se identifique con claridad el resultado mostrado).

1. A continuación, se encuentran 4 secuencias de nucleótidos que codifican para el pseudogen GULO (L-gulonolactone oxidase).

>Macaca_fascicularis

```
AAAGAAGACCACAGGGGCCCTGCTGGAGATGAAGGCCATGCTGGAGGCCACCCTGAG
GTGGTGTCCCACTAACCGGTGGGGGTGCGCTTCACCCAAGGGATGACATCATACTGAGC
CCCTGCTTCCAGCAGGACAGCTGCTACCTGGACATCAACCTGTACAG
```

>Pan_troglodytes

```
AAAGAAGACCACGGAGGCCCTGCTGGAGCTGAAGGCCATGCTGGAGGCCACCCCGAG
GTGGTGTCCCACTACCTGGTGGGGGTACGCTTCACCTGGAG-
GATGACATCCTACTGAGCCCCTGCTTCCAGCGGGACAGCCGCTACCTGAACATCAACCT
GTACAG
```

>Pongo_pygmaeus

```
AAAGAAGACCACGGAGGCCCTGCTGGAGCTGAAGGCCATGCTGGAGGCCACCCTGAG
GTGGTGTCCCACTACCGGTGGGGGTGCGCTTCACCCAGAGGATGACGTCCTACTGAGC
CCCTGCTTCCAGCAGGACAGCCGCTATCTGAACATCAACCTGTACAG
```

>Homo_sapiens

```
AAAGAAGACCACGGAGGCCCTGCTGGAGCTGAAGGCCGTGCTGGAGGCCACCCTGAG
GTGGTGTCCCACTACCTGGTGGGGGTACGCTTCACCTGGAGGATGACATCCTACTGAGC
CCCTGCTTCCAGTGGGACAGCCGCTACCTGAACATCAACCTGTACAG
```

>Mus_musculus

```
GGAGAAGACCAAGGAGGCCCTGCTGGAGCTAAAGGCCATGCTGGAGGCCACCCCAAG
GTGGTAGCCCACTACCCCGTGGAGGTGCGCTTCACCCGAGGTGATGACATCCTGCTGAG
CCCGTCTTCCAGAGGGACAGCTGCTACATGAACATCATTATGTACAG
```

Examinar las 4 secuencias y alinearlas para identificar:

- a. La zona donde existe la más alta variabilidad entre las secuencias comparadas.
 - b. La zona donde existe la mayor cantidad de nucleótidos conservados
2. Identificar el mejor modelo que se debe utilizar para elaborar un árbol filogenético mediante un test de máxima similitud con los datos presentados.
 3. Elaborar el árbol filogenético usando el modelo sugerido con un “boostrapping” de 100 veces.
 4. Mostrar el árbol generado en:
 - a. Estilo radial
 - b. Estilo rectangular tradicional
 - c. ¿Qué organismo fue seleccionado como el “outgroup” para este análisis filogenético?
 - d. ¿Qué importancia tiene un “outgroup” en un estudio de filogenia?

BIOG1020: Bioinformática

5. ¿Qué organismos se encuentran más relacionados entre ellos?
6. ¿Cuál de los grupos formados posee el mejor nivel de confianza en su estimación?
7. Elaborar el árbol filogenético usando el modelo de Neighbor-Joining usando el modelo de sustitución de Juke-Cantor y con un “bootstrapping” de 100 veces.
 - a. Al comparar ambos árboles, ¿qué se puede decir de ellos?
 - b. ¿Cuál considera que es la mayor diferencia entre ambos arboles filogenéticos?

B) Los siguientes enumerados deben ser resueltos de manera individual. Cada respuesta debe ser respondida de forma escrita y de preferencia de manera gráfica. Imágenes de la pantalla donde se identifique con claridad el resultado mostrado.

1. Identificar y descargar la estructura 3D de la proteína 2DFT en la base de datos RCSB PDB.
2. ¿Cuántos y cuáles son los ligandos que se encuentran en la proteína identificada?
3. ¿Cuántas subunidades posee la proteína? Generar una imagen donde cada subunidad se encuentre de un color diferente.
4. ¿Cuántas láminas betas posee toda la proteína? ¿Estas se encuentran de forma paralela o antiparalela?
5. Identifique 3 en una de las subunidades, 3 aminoácidos que se encuentren a 7 Å del ligando e interactuando mediante puentes de Hidrógeno.

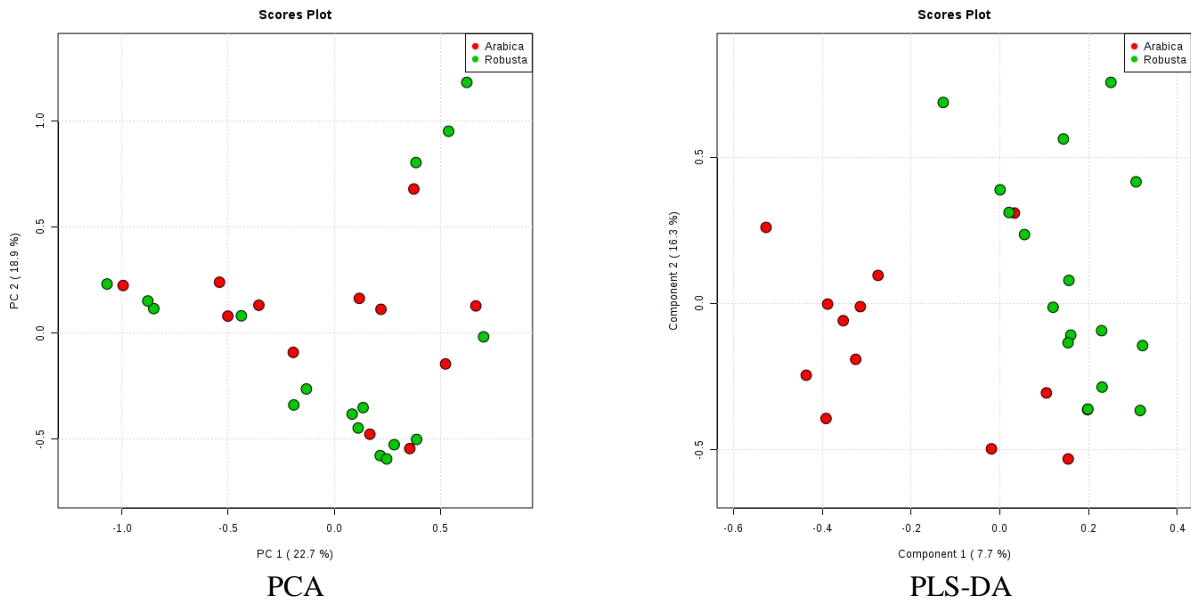
C. De acuerdo a la alineación en Geneious, llenar el siguiente cuadro:

	Secuencia (5' a 3')	Tamaño de primer (pb)	G/C %	> 4 nt = (si/no)	> 4nt GC en 3' (si/no)	GC Clamp (2) (si/no)	Tamaño de amplicon	Puntaje de degeneraci
Forward								
Reverse								

D. En base a las características evaluadas, considera que estos primers son eficientes?

E. En un estudio sobre tipos de café, algunos estudiantes de la ESPOL hicieron un análisis metabolómico con la técnica de MALDI-TOF, buscando descubrir si los ejemplares de café Robusta y Arábica tenían perfiles de metabolitos diferentes. Después de coleccionar los datos del espectrómetro de masas, los estudiantes analizaron los datos a través de PCA y PLS-DA, utilizando el software Metaboanalyst.

Como resultado, obtuvieron las siguientes graficas:



Responda las preguntas:

1. ¿Hay diferencia entre los perfiles de metabolitos de cada tipo de café?
2. ¿Cuál análisis (PCA o PLS-DA) muestra un mejor resultado? ¿Hay diferencias? Explique.
3. ¿Es posible detectar *outliers*? En caso de que si, indique donde en la(s) grafica(s).
4. ¿Usted cree que las muestras analizadas pueden pertenecer a solo uno tipo de café? ¿Por qué?