



UAX

UNIVERSIDAD ALFONSO X EL SABIO

Trabajo Opcional

GRADO EN BIOMEDICINA

Bioinformática

idelhgar@uax.es

2023

Trabajo Opcional

- Entrega como un cuaderno de Jupyter Notebook
- Mete comentarios, código y capturas de pantalla de la salida del código
- ENTREGA 29 Octubre 2023 23:59
- Entrega como NombreApellidos_TrabajoOpt.ipynb

Trabajo Opt

- Escoge un gen de tu interés que tenga estructura en el PDB!
- Guarda el archivo PDB
- Guarda la secuencia de nucleótidos del gen que has elegido
- Busca 6 ortólogos diferentes de tu gen y extrae sus secuencias de nucléotidos. No cambies los encabezados!!!!
- Crea un archivo multifasta
- Los headers serán algo parecido a esto

```
-----  
>lcl|CP018245.1_cds_API18669.1_5 [locus_tag=BFL23_00025] [protein=hypothetical protein]  
[protein_id=API18669.1] [location=5105..5254] [gbkey=CDS]
```

Trabajo Opcional

Empleando las herramientas que hemos dado como grep, sed y scripts en Bash obtén los siguientes resultados:

1. Extrae el número de secuencias que hay en el archivo
2. Extrae el identificador de las secuencias
3. Extrae el organismo de cada secuencia
4. Extrae las subsecuencias de tu archivo multifasta que empiecen por ATGC y que acaben por GTAA. ¿Hay alguna?
5. Las enzimas de restricción FokI, HphI y Tail unen las secuencias GGATG, GGTGA y ACGT. ¿Hay algún sitio de unión en tu archivo? Si hay, ¿cuántos son de cada una? ¿Puedes indicar en qué línea están cada uno?

Trabajo Opcional

En el archivo PDB que has obtenido haz los siguiente:

- 1. ¿Cuántas cadenas tiene tu proteína?**
- 2. Elimina todas las líneas que contenga información sobre puentes de hidrógeno. Guarda en un archivo nuevo**
- 3. Cambia las PHE por ALA. Guarda en un archivo nuevo**