

Introducción a la Bioinformática

TP3

Enrique Alonso

1) **DESAFIO I:** Enumerá las diferencias que existen entre una célula procariota y eucariota.

Las células procariotas no tiene núcleo y son mas pequeñas. La transcripción y traducción ocurren de manera simultánea, mientras que en las eucariotas sucede por separado. Las procariotas tienen pilis en su exterior. Las eucariotas a diferencia de las procariotas tienen organelas como lisosomas, mitocondrias, etc.

2) **DESAFIO II:** Dado el código genético como se muestra en la tabla, Crea un script en Python que tome como argumento una secuencia proteica e imprima una cadena de ARN codificante. Podés usar de ejemplo el siguiente péptido (cadena corta de aminoácidos):

Sec1: 'ATVEKGGKHKHTGPNEKGKKIFVQKCSQCHTVLHGLFGRKTGQA'

```
1  import re
2
3  sec1= "ATVEKGGKHKHTGPNEKGKKIFVQKCSQCHTVLHGLFGRKTGQA"
4
5  def get3CodonsFromGeneticCode(aminoacid):
6
7      aminoDictionary = {
8          "A": ["GCG", "GCA", "GCC", "GCT"],
9          "E": ["GAG", "GAA"],
10         "D": ["GAC", "GAT"],
11         "L": ["ATG", "ATA", "ATC", "ATT", "TTG", "TTA"],
12         "R": ["CGA", "CGT", "CGC", "CGG", "AGG", "AGA"],
13         "P": ["CCC", "CCG", "CCT", "CCA"],
14         "Q": ["CAG", "CAT"],
15         "H": ["CAC", "CAT"],
16         "M": ["ATG"],
17         "I": ["ATA", "ATC", "ATT"],
18         "S": ["AGT", "AGC", "TCC", "TCG", "TCT", "TCA"],
19         "T": ["ACC", "ACG", "ACT", "ACA"],
20         "K": ["AAA", "AAG"],
21         "N": ["AAC", "AAT"],
22         "F": ["TTC", "TTT"],
23         "W": ["TGG"],
24         "*": ["TGA"],
25         "C": ["TGC", "TGT"],
26         "#": ["TAG", "TAA"],
27         "Y": ["TAC", "TAT"],
28         "V": ["GTC", "GTG", "GTT", "GTA"],
29         "G": ["GGC", "GGG", "GGT", "GGA"],
30
31     }
32
33     return aminoDictionary[aminoacid][0] #TODO
34
35
36
37
38 out=""
39
40 for aminoacid in sec1:
41     out= out + get3CodonsFromGeneticCode(aminoacid)
42
43 print (out)
```

- 3) **DESAFIO III:** En muchos de los genes codificados en el ADN existe un motivo recurrente ubicado antes de la secuencia codificante del gen que direcciona la unión de la ARN Polimerasa II, la proteína encargada de copiar el ADN a un ARN mensajero. Ésta secuencia denominada caja TATA (consistente en una secuencia de nucleótidos 'TATAAA') se encuentra presente en lo que se denomina región promotora de diversos genes, en organismos de todos los reinos (Smale and Kadonaga 2003; Lifton et al. 1978). Créa un script en Python que, tomando como input un archivo con una secuencia de ADN, permita identificar las regiones promotoras de un gen, considerando que tal región comienza y termina con la caja TATA.

```
1 import re
2
3 clave="TATAAA"
4 p = re.compile(r'%s'%clave)
5
6 archivo = open("secuencia.txt","r")
7
8 nroLinea=0
9 for linea in archivo:
10     nroLinea=nroLinea+1
11     for m in p.finditer(linea):
12         print("Se ha hallado una region promotora en la linea ",nroLinea, " desde: ",m.start(), " hasta ", m.start()+len(clave))
13
14
15 archivo.close()
```

- 4) **DESAFIO IV:** Vamos a divertirnos un poco mientras aprendemos, ¡y no hay mejor modo de hacer esto que jugando! Diseña un juego rpg interactivo sobre la expresión génica que se muestre en la consola (que se ejecute [mediante CLI](#) de manera similar a lo visto en el Bashaton). Tené en cuenta que lo vas a tener que compartir con la clase.

El fuente del juego se encuentra en el repositorio: <https://github.com/alonsoem/bioinformatica/tree/main/TP3/cellShip>