

Primer Exámen Informática (Biología)

Pontificia Universidad Javeriana - Cali

Prof. Luis Garreta

12 de septiembre de 2018

1. Un **motivo biológico** es una secuencia corta que vamos a considerar debe repetirse en TODAS las secuencias de un genoma. Un motivo puede corresponder a una secuencia corta específica, por ejemplo "tata" o "cata".

Dado las siguientes secuencias de un genoma:

- s1="cgagtatccataaatataggttttaa"
- s2="ccccggttacctataggccata",
- s3="tgccatatataacgc",
- s4="tgccatatataacgc",
- s5="tgccatttatacgc"

- a) Realice un programa que imprima si la cadena "tata" es un motivo.
- b) Cambie su programa definiendo una función suya que ingresen dos parámetros: **una secuencia y una cadena**, y retorne verdadero o falso (True/False) si la cadena aparece en la secuencia.

```
## Retorna True/False si dentro de la secuencia existe la cadena
def existeCadena (secuencia , cadena):
    ...
    ...
```

- c) Pruebe su función con las secuencias anteriores para determinar primero si la cadena "tata" es un motivo, y segundo si cadena "cata" es otro motivo.

2. Vamos a considerar la **identidad entre dos secuencias** como el número de bases iguales que se repiten en ambas secuencias y en las mismas posiciones. Por ejemplo:

- "tata" y "tata" son 100 % idénticas,
- "tata" y "taca" son 75 % idénticas, y
- "tata" y "tgta" son 50 % idénticas

- a) Realice un programa que determina la identidad de las siguientes secuencias:
- "gacatatata" y "gacataaata"
 - "gacatatata" y "gacataaata"
 - "gacatatata" y "gacataaaaa"
- b) Transforme su programa en función de tal manera que le ingresen como parámetros las dos cadenas y retorne su porcentaje de identidad.

```
## Retorna el porcentaje de identidad de dos secuencias: sec1 y sec2
def porcentajeIdentidad (sec1 , sec2):
    ...
    ...
```

- c) Pruebe su función con los pares de secuencias anteriores
3. Realice un programa que dado una secuencia y una cadena, recorra la secuencia e imprima todas las posiciones donde aparece la cadena. Por ejemplo:
- Si la secuencia es `s1="cgagtatacatataagacatataggttttaa"` y la cadena `cad="tata"`, entonces imprimiría: 4, 10, 19
 - Si la secuencia es `s1="cgagtatacatataagacatataggttttaa"` y la cadena `cad="catata"`, entonces imprimiría: 8, 17
4. Modifique el programa anterior para que identifique cadenas incompletas, es decir con una **identidad** mayor o igual al 50 % (utilice la función de identidad de arriba). Por ejemplo:
- Si la secuencia es `s1="cgagtatacgtataagacagataggttttaa"` y la cadena `cad="catata"`, entonces imprimiría: 8, 17