

CURSO: CC471 - 2019 -1

Practica Calificada N°1.

1. (4 puntos)

- Presentacion de la asignacion de la Lec02: si [], no [].

2.(2 puntos)

Cual es la secuencia de ADN que dio origen al siguiente diagrama de SANGER?

Diagrama de Sanger.

R: La secuencia solicitada es TGCACGTCATGCT

3. (4 puntos)

Encuentre el alineamiento optimo Local entre de las secuencias ATACTGGG y TGAAGG usando el sistema de scoring: Match = 1, mismatch = -1, gap = -2. Que algoritmo utilizó?. Cual es el máximo score?. Muestre el camino del alineamiento en la matriz correspondiente.

R: Se utilizo Smith Waterman

		T	G	A	C	T	G	A	G
A	0	-1	-1	1	-1	-1	-1	1	-1
T	0	1	-1	-1	-1	1	-1	-1	-1
A	0	-1	-1	1	-1	-1	-1	1	-1
C	0	-1	-1	-1	1	-1	-1	-1	-1
T	0	1	-1	-1	-1	1	-1	-1	-1
G	0	-1	1	-1	-1	-1	1	-1	1
G	0	-1	1	-1	-1	-1	1	-1	1
G	0	-1	1	-1	-1	-1	1	-1	1

SCORE: 5x1 + (-1) = 4

4.(4puntos) Tenemos una secuencia de proteina en formato fasta:

```
>>sp|A1AU17|NDK_PELPD Nucleoside diphosphate kinase OS=Pelobacter propionicus (strain DSM 2379 / NBRC 103807 / OttBd1) OX=338966 GN=ndk PE=3 SV=1
```

```
MERTFAIIKPDAVERRLAGTVIDRIEANGFTIVGMKKIKLSKEQAGGFYCVHRERPFFGE
LCDFMSRSPVIVLCLEKENAIADWRKLMGATNPANAEPGTIRRDFALSLENSAHGSDAP
ETAAFEIAYFFNALELV
```

a)Cual seria la secuencia de nucleotidos necesaria para codificar los 4 primeros aminoacidos:

methionine - Met – M; Glutamic Acid – glu – E; Arginine – Arg – R; Threonine – Thr - T

b) Muestre la secuencia de la cadena de ADN de dos hebras que dio origen a la secuencia del punto (a) indique el direccionamiento correspondiente

c) Cuantas variantes de la secuencia solicitada en el punto (a) pueden existir?

R:Hay varias posibilidades: Una de ellas sería

a) AUG GAA CGA ACG

	U	C	A	G	
U	phenyl- alanine	serine	tyrosine	cysteine	U
	leucine		stop	stop	C
			stop	tryptophan	A
					G
C	leucine	proline	histidine	arginine	U
			glutamine		C
					A
					G
A	isoleucine	threonine	asparagine	serine	U
	* methionine		lysine	arginine	C
					A
					G
G	valine	alanine	aspartic acid	glycine	U
			glutamic acid		C
					A
					G

Utilizando la tabla de codificacion proporcionada, * and start sabiendo que 3 nucleotidos (codones) codifican un aminoacido.

b) La secuencia de ADN correspondiente a la del punto (a) es

5' - ATG GAA CGA ACG - 3'

3' - TAC CTT GCT TGC -5'

c) para codificar MERT pueden existir $1 \times 2 \times 6 \times 4 = 48$ variantes por cuanto existen:

1 variante para **M**, 2 variantes para **E**, 6 variantes para **R**, 4 variantes para **L**.

5. 1. (6 puntos) Defina Formalmente el algoritmo Needleman Wunsch – incluya el pseudo codigo.

R:

Entrada:

Dos secuencias P1 y P2

Matriz de scoring correspondiente a los valores de match, mismatch, y gap.

Se inicializan las secuencias de la matriz de similitud

$s(0,j) = j \cdot g$ para $j = 0$ hasta m ; g es el valor de gap; m Numero de elementos de la secuencia P1

$s(i,0) = i \cdot g$ para $i = 0$ hasta n ; g es el valor de gap; n Numero de elementos de la secuencia P2

1. Inicializar la matriz

Para todo $i = 1$ hasta m

Para todo $j = 1$ hasta n

if $p1(j) = p2(i)$ entonces $s(i,j)=\text{match}$

if $p1(j) \neq p2(i)$ entonces $s(i,j)=\text{mismatch}$

2. Calcular los valores de $F(i,j)$

Para todo $i = 1$ hasta m

Para todo $j = 1$ hasta n

$F(i,j) = \max(F(i-1,j) -g; F(i,j-1)-g; F(i-1,j-1)+s(i,j))$

3. Efectuar el alineamiento desde $s(m,n)$ hasta $s(0,0)$, buscando la ruta inversa con el mayor score.

Cada paso en la ruta corresponde a un par de elementos de las secuencias alineadas. Cuando la ruta no pasa por la diagonal, entonces se agregan espacios en blanco a las secuencias alineadas correspondientes.