

TP5 – Alineamiento y BLAST

👉 **PARA PENSAR:** ¿Qué tipo de información se puede extraer de la comparación de secuencias?
¿Cómo esperas que se vea en una comparación? 😞

Podemos analizar semejanzas entre dos especies.

Esperaría ver una comparación letra por letra tratando de que coincidan la mayor cantidad posible.

👉 **PARA PENSAR:** ¿Por qué crees que es mejor evaluar las relaciones evolutivas lejanas comparando proteínas? 😞

Porque las proteínas se ven menos modificadas con el tiempo de lo que puede ser una secuencia de ácidos nucleicos, que pudo sufrir muchas alteraciones por distintos factores.

👉 **RETO I:** Intentemos, entonces alinear estas dos palabras, para comprender mejor el problema.
Alinea en la siguiente table de comparaciones las palabras "BANANA" y "MANZANA".

¡Tomá nota de tus observaciones y de las conclusiones que se desprendan de estas observaciones!

☒ **PREGUNTAS DISPARADORAS:** ¿Existe una única forma de alinearlas? ¿Es alguno de los posibles alineamientos mejor que otro? Si así fuera ¿Por qué?

B	A	N	A	N	A	-	-
M	A	N	Z	A	N	A	-
✗	✓	✓	✗	✗	✗	✓	✓


☒ **PREGUNTAS DISPARADORAS:** ¿Qué representan esos guiones?

El mejor alineamiento que se logro fue el siguiente:


B	A	N	-	A	N	A	-
M	A	N	Z	A	N	A	-
✗	✓	✓	✓	✓	✓	✓	✓

No existe una única forma para alinearlas. Sin embargo, hay algún alineamiento que resulta mejor que el otro. Por ejemplo, en el alineamiento que se realizó, coinciden mayor número de letras respecto del que se mostró como ejemplo.

Los guiones son espacios en blanco que pueden representar la falta de un aminoácido de la secuencia.









 **RETO II:** En la siguiente tabla probá distintos alineamientos para las palabras "ANA" y "ANANA". Verás que en el margen superior izquierdo aparece un valor de identidad calculado para cada alineamiento que intentes.


Tomá nota de los valores de identidad observados y de las conclusiones que se desprendan de estas observaciones.

 **PREGUNTAS DISPARADORAS:** ¿Son todos los valores iguales? ¿Qué consideraciones deberían tenerse en cuenta a la hora de realizar el cálculo? ¿Se te ocurre, distintas formas de calcularlo? ¿Serán todas ellas igualmente válidas en Biología?

Penalidad

Identidad


A	N	A	-	-	
A	N	A	N	A	
					



Después de probar distintos alineamientos, notamos que no son todos iguales. No es lo mismo cuando se comparan las mismas letras que cuando no se hace.

Se debería tener en cuenta si se comparan dos letras iguales, diferentes, o una letra y un espacio en blanco. Se podría calcular todos los alineamientos posibles para ambas secuencias y quedarse con el que mayor puntaje tenga. A la hora de realizar el cálculo se deben tener en cuenta varias cosas. Por ejemplo, la penalidad por comparar una letra con un gap o la penalidad de comparar dos letras distintas.

No todas serán igualmente validas, ya que cuanto menor sea el puntaje, seguramente sea mucho menos probable que sean la misma secuencia o tengan relación alguna.

 **RETO III:** En la siguiente tabla probá distintos alineamientos para las palabras "ANA" y "ANANA". Verás que en el margen superior izquierdo aparece un valor de identidad calculado para cada alineamiento que intentes y un botón para cambiar la penalidad que se le otorga a dicho para el cálculo de *identidad*.

Probá varias combinaciones, tomá nota de los valores de identidad observados y de las conclusiones que se desprendan de estas observaciones.

☒ **PREGUNTAS DISPARADORAS:** ¿Cómo se relacionan los valores de identidad obtenidos con las penalizaciones que se imponen al gap? ¿Qué implicancias crees que tiene una mayor penalización de gaps? ¿Se te ocurre alguna otra forma de penalización que no haya sido tenido en cuenta en este ejemplo?

Penalidad Identidad 0.2

A	N	A	-	-	<input checked="" type="checkbox"/>
A	N	A	N	A	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

☒

Se relacionan directamente.

A mayor penalización impuesta al gap, menor es el puntaje obtenido en la identidad.
Por ejemplo, el siguiente caso muestra lo que se comenta:


Penalidad Identidad -0.2

A	-	-	N	A	<input checked="" type="checkbox"/>
A	N	A	N	A	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>


☒

Al aumentar la penalidad, la identidad baja.

Se me ocurre que puede existir diferentes tipos de penalidades al comparar aminoácidos entre sí. No será igual compararlo con uno, que con otro.

 **PARA PENSAR:** Entonces, pensando en un alineamiento de ácidos nucleicos ¿Cuáles te parece que son las implicancias de abrir un gap en el alineamiento? ¿Qué implicaría la inserción o delección de una región de más de un residuo?

Cuanto más gaps se tengan que agregar en el alineamiento, mas nos alejamos de que sean secuencias iguales.

 **RETO IV:** En la siguiente tabla probá distintos alineamientos para las secuencias nucleotídicas. Podrás ver las traducciones para cada secuencia.

Probá varias combinaciones, tomá nota de las observaciones y de las conclusiones que se desprendan de estas.

Consigna: Alineá "TGCGAGG" y "TGCCGAAGG" y mirá las traducciones


Penalidad Identidad 0.3333333333333333

T	G	C	G	A	G	G	-	-	
C			E			-			
T	G	C	C	G	A	A	G	G	
C			R			R			

Consigna: *Intentá* alinear "AGGGGA" y "TGCAGAGGG" y mirá las traducciones

Penalidad Identidad 0.2222222222222222

A	G	G	G	G	A	-	-	-	
R			G			-			
T	G	A	G	A	G	G	G	-	
X			E			-			

 **PARA PENSAR:** ¿Dá lo mismo si el gap que introducís cae en la primera, segunda o tercer posición del codón? ¿Cómo ponderarías las observaciones de este ejercicio para evaluar el parecido entre dos secuencias?

Según las pruebas realizadas, da igual en que posición dentro del codón, este colocado el gap.

Este ejercicio nos hace dar cuenta de que alinear secuencias no es tan simple como parecía. Se deben tener en cuenta distintos factores de penalización o de acierto y en base a eso calcular la identidad del alineamiento. Creo que resulta muy útil este tipo de herramientas para detectar similitudes entre secuencias.

 **PARA PENSAR:** ¿En qué casos serán de utilidad uno u otro tipo de alineamientos? ¿Qué limitaciones tendrá cada uno?

Cada tipo de alineamiento será mejor o peor opción dependiendo del caso de estudio. Por ejemplo, un alineamiento de pares de secuencias puede servir para hacer una prueba de paternidad. En cambio, si se está averiguando a que especie pertenece una proteína, entiendo que será mucho mejor hacer un alineamiento de secuencias múltiples, para así tener muchas más posibilidades de comparación.

La principal limitación del alineamiento de a pares, creo que es la cantidad de secuencias que permite alinear. Este alineamiento en muchos estudios no resultaría muy útil.

En cambio, en el alineamiento múltiple, la gran limitante será la complejidad de calculo que debe tener el tener que comparar tantas secuencias unas con otras.