



Universidad Católica
San Pablo

CIENCIA DE LA COMPUTACIÓN

Biología Molecular Computacional

Laboratorio - 02a: Estructura Secundaria en ARN

Rodrigo Alonso Torres Sotomayor

CCOMP 9-1

"El alumno declara haber realizado el presente trabajo de acuerdo a las normas de la Universidad Católica San Pablo"

Laboratorio 02a: Estructura Secundaria en ARN

Algoritmo de Predicción Estructura Secundaria: Pruebas y ajustes

0.1 Ejercicio 01

Probar su implementación en el ejemplo visto en clase GGAAAUCC y usando las siguientes funciones de energía.

$$\alpha(r_i, r_j) = \begin{cases} -5 & \text{si } r_i r_j = \text{CG ó GC} \\ -4 & \text{si } r_i r_j = \text{AU ó UA} \\ -1 & \text{si } r_i r_j = \text{GU ó UG} \\ 0 & \text{en los demás casos} \end{cases} \quad \alpha(r_i, r_j) = \begin{cases} -1 & \text{si } r_i r_j = \text{CG ó GC} \\ -1 & \text{si } r_i r_j = \text{AU ó UA} \\ 0 & \text{en los demás casos} \end{cases}$$

```
Matriz de Scores (E):
0 0 0 0 0 -4 -9 -14
0 0 0 0 0 -4 -9 -9
0 0 0 0 0 -4 -4 -4
0 0 0 0 0 -4 -4 -4
0 0 0 0 0 -4 -4 -4
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0

Matriz de Caminos (P):
(0,-1) (1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (1,6)
(0,-1) (0,-1) (2,2) (2,3) (2,4) (2,5) (2,5) (1,6)
(0,-1) (0,-1) (0,-1) (3,3) (3,4) (3,5) (3,6) (3,7)
(0,-1) (0,-1) (0,-1) (0,-1) (4,4) (4,5) (4,6) (4,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (5,4) (4,5) (4,6)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (6,6) (6,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (7,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1)

Estructura:
G - C
G - C
A - U
A - U
```

Figure 1: Emparejamiento de cadena GGAAAUCC con la 1ra función score.

```
Matriz de Scores (E):
0 0 0 0 0 -1 -2 -3
0 0 0 0 0 -1 -2 -2
0 0 0 0 0 -1 -1 -1
0 0 0 0 0 -1 -1 -1
0 0 0 0 0 -1 -1 -1
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0

Matriz de Caminos (P):
(0,-1) (1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (1,6)
(0,-1) (0,-1) (2,2) (2,3) (2,4) (2,5) (2,5) (1,6)
(0,-1) (0,-1) (0,-1) (3,3) (3,4) (3,5) (3,6) (3,7)
(0,-1) (0,-1) (0,-1) (0,-1) (4,4) (4,5) (4,6) (4,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (5,4) (4,5) (4,6)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (6,6) (6,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (7,7)
(0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1) (0,-1)

Estructura:
G - C
G - C
A - U
A - U
```

Figure 2: Emparejamiento de cadena GGAAAUCC con la 2da función score.

0.2 Ejercicio 02

Probar con otras secuencias de ARN como ACUCGAUCCGAG.

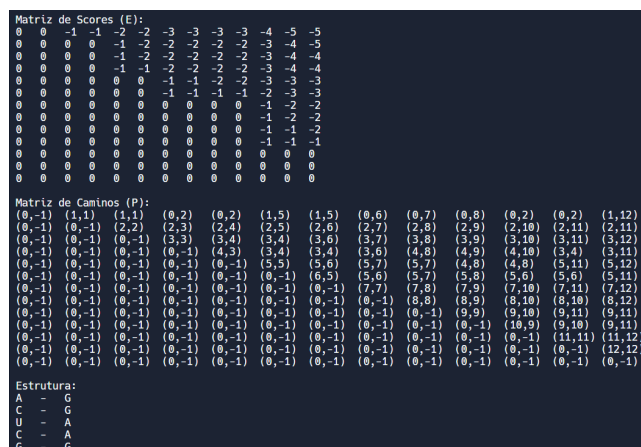


Figure 3: Emparejamiento de cadena ACUCGAUUCCGAG con la 1ra función score.

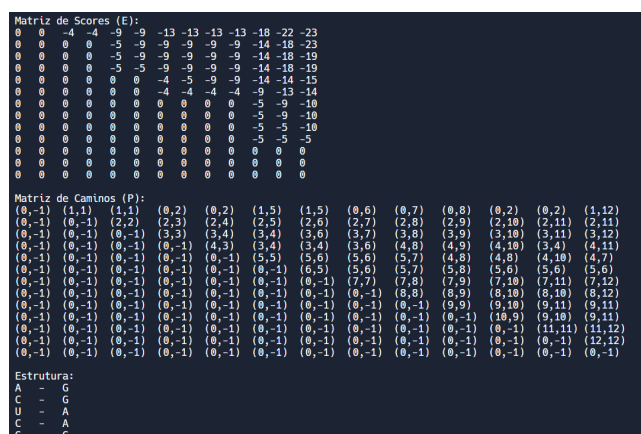


Figure 4: Emparejamiento de cadena ACUCGAUUCCGAG con la 2da función score.

0.3 Emparejamiento de pares ARN: Valorización

0.3.1 Encuentre una relación entre los modelos de predicción de estructuras ARN y la búsqueda de palíndromos.

Los modelos de predicción de estructuras de ARN, como el algoritmo de Nussinov, y la búsqueda de palíndromos tienen en común la identificación de secuencias que pueden formar estructuras secundarias. Un palíndromo en el contexto del ARN es una secuencia de bases que se lee igual en ambas direcciones y puede formar una horquilla o "hairpin". Estas estructuras son fundamentales en la conformación de la estructura secundaria del ARN.

El algoritmo de Nussinov utiliza una matriz de puntuación para evaluar las posibles estructuras secundarias del ARN, incluyendo las horquillas y otros motivos estructurales que pueden ser derivados de secuencias palindrómicas. La búsqueda de palíndromos puede ser un paso preliminar en la predicción de estructuras secundarias, ayudando a identificar regiones candidatas que podrían formar estructuras estables.

0.3.2 ¿Cómo la energía libre afecta a los resultados encontrados?

La energía libre es un factor crucial en la predicción de la estructura secundaria del ARN. Las estructuras con una menor energía libre son más estables y, por lo tanto, más probables de encontrarse en la naturaleza. El algoritmo de Nussinov, en su versión básica, no considera directamente la energía libre, sino que maximiza el número de pares de bases apareadas. Sin embargo, las versiones mejoradas del algoritmo incorporan modelos de energía termodinámica para evaluar la estabilidad de las estructuras predichas. Al incluir la energía libre en el cálculo, los algoritmos pueden predecir estructuras más biológicamente relevantes y precisas.

0.3.3 Logros del algoritmo de Nussinov en Biología Molecular Computacional

El algoritmo de Nussinov ha sido una herramienta fundamental en la biología molecular computacional para:

- Predicción de la estructura secundaria del ARN: Ha permitido a los científicos comprender mejor cómo las secuencias de ARN se pliegan en estructuras funcionales.
- Diseño de ARN sintético: Facilita el diseño de secuencias de ARN con estructuras deseadas para aplicaciones biotecnológicas y terapéuticas.
- Análisis evolutivo: Permite el estudio de la conservación estructural de ARN a través de diferentes especies.

0.3.4 Mejoras del algoritmo de Nussinov

A lo largo del tiempo, el algoritmo de Nussinov ha sido mejorado en diversas maneras:

- Incorporación de modelos de energía termodinámica: Para predecir estructuras más estables basándose en la energía libre.
- Consideración de pseudonudos: Aunque el algoritmo original no puede manejar pseudonudos, se han desarrollado extensiones para incluir estas complejas estructuras.
- Optimización computacional: Mejora en la eficiencia y velocidad del algoritmo para manejar secuencias más largas y complejas.

0.3.5 Mejora más relevante

La incorporación de modelos de energía termodinámica parece ser la mejora más relevante, ya que permite una predicción más precisa y biológicamente relevante de las estructuras secundarias del ARN. Esto ha ampliado considerablemente la utilidad del algoritmo de Nussinov, permitiendo su aplicación en investigaciones más avanzadas y prácticas en biología molecular.

Link del repositorio

<https://github.com/RodATS/Molecular.git>