



Biología Molecular Computacional CB309 Laboratorio 01a: Neddleman-Wunsch

Yván Jesús Túpac Valdivia
Universidad Católica San Pablo, Arequipa – Perú

29 de agosto de 2024

1 Objetivos

- Implementar el algoritmo de alineamiento de dos secuencias de [Needleman and Wunsch, 1970] en un lenguaje de programación.
- Verificar su funcionamiento y comportamiento computacional ante diversos tamaños de cadenas
- Analizar la calidad de los resultados obtenidos.
- Valorizar la utilidad de este algoritmo.

2 Desarrollo

2.1 Algoritmo de alineamiento global: Implementación

- Conforme se vio en clase, se tiene el algoritmo de alineamiento global [Needleman and Wunsch, 1970] que permite tratar dos secuencias s y t .
- Implementar el algoritmo en C++ preferentemente. De acuerdo a lo visto en clase, analice qué recursos del lenguaje deberían aprovecharse y utilícelos. [SO1, SO2]

2.2 Algoritmo de alineamiento global: Pruebas y ajustes

- Probar su implementación en el ejemplo visto en clase: alinear las cadenas AAAC con AGC, verificando que se obtengan todas las soluciones posibles e indicando el *score* obtenido. [SO6]
- Probar su implementación en las cadenas del archivo [Secuencias.txt], tenga en cuenta que puede necesitar adecuar la lectura de archivo para poder capturar correctamente las cadenas. Verifique el correcto funcionamiento del algoritmo ante este tamaño de cadenas. [SO6]
- Caso se presenten dificultades, encuentre una forma de obtener el resultado de alineamiento probando para las combinaciones $\binom{2}{3}$ de las cadenas Bacteria, Sars-cov e Influenza. [SO2]
- Analice e implemente una forma de detectar la solución de alineamiento que ofrece menos rupturas en las moléculas. [SO1, SO2]
- Evalúe e implemente una forma práctica de visualizar las cadenas alineadas principalmente si se trata de cadenas de gran tamaño, se sugiere usar la idea de la Matriz de Puntos (revisar diapositivas que fueron actualizadas) [SO2]

2.3 Algoritmo de alineamiento global: Valorización

- Haga una revisión en la literatura buscando qué logros ha permitido alcanzar este algoritmo en el área de Biología Molecular Computacional y revise también las diversas mejoras que se han investigado y logrado al mismo algoritmo. Indique cuál de las mejoras le parece la más relevante. [S07]

3 Entrega

- Deberá entregar dentro del plazo que aparece en el Aula Virtual un informe (PDF) con lo realizado respondiendo las preguntas y mostrando los funcionamientos.
- Las implementaciones deberán estar en repositorio, cuyo URL debe indicarse en el informe.

References

[Needleman and Wunsch, 1970] Needleman, S. B. and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3):443–453.