

Metaheurísticas aplicadas al alineamiento múltiple de secuencias de ADN y/o de aminoácidos

Proyecto: Big data optimization con algoritmos metaheurísticos utilizando frameworks de computación distribuida

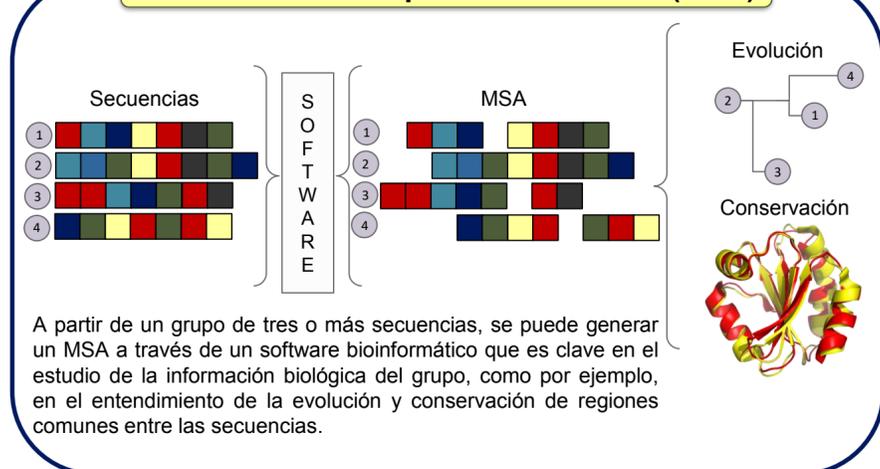
Directora: Carolina Salto

Investigadores: Adrián Díaz, Gabriela Minetti

Resumen

El desafío abordado en esta investigación, que forma parte de la tesis de maestría del primer autor en la Universidad Nacional de Quilmes, es desarrollar un algoritmo de alineamiento múltiple de secuencias (MSA), utilizando enfriamiento simulado (SA). Esta investigación contribuye a la formación académica y profesional del autor en el ámbito de la investigación científica.

Alineamiento múltiple de secuencias (MSA)



Desarrollo

Siguiendo el método científico se desarrolla un algoritmo meta-heurístico basado en SA para hibridar con heurísticas, que permita encontrar un MSA de manera eficiente y competitiva con respecto a otros enfoques vigentes en el estado del arte.

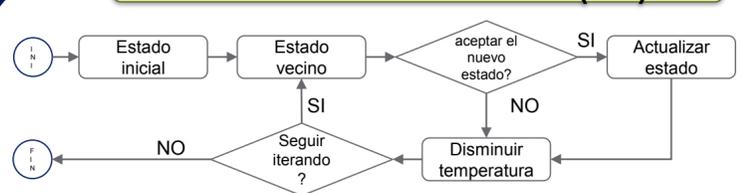
Este algoritmo es evaluado con secuencias de longitudes cortas, medianas, y largas, así como en conjuntos de secuencias pequeñas o numerosos, a fin de demostrar su aplicabilidad y efectividad en una amplia gama de escenarios de investigación en bioinformática.

Adicionalmente al algoritmo propuesto, esta investigación incluye un marco de trabajo utilizando *Nextflow* para comparar distintas herramientas MSA en un ambiente reproducible y fácilmente extensible.

Introducción

El MSA es un enfoque bioinformático esencial para comparar y analizar secuencias de nucleótidos o aminoácidos. Esta técnica permite identificar regiones conservadas y evolutivamente relacionadas entre distintas secuencias biológicas, lo que facilita la predicción de estructuras y funciones moleculares. El MSA es fundamental en investigaciones genómicas y proteómicas, siendo clave para el avance en la comprensión de procesos biológicos y además de la conservación y evolución de las secuencias.

Enfriamiento simulado (SA)



Para el problema de encontrar un MSA, un estado del sistema representa un alineamiento múltiple. Para evaluar el nuevo estado se utiliza un cálculo de la energía, que en este caso, es una función que determina heurísticamente la calidad del MSA generado. Cada nuevo estado vecino, es un MSA con una pequeña modificación.

Aplicación de SA en el problema MSA



Marco de trabajo propuesto

