

Guião 2 – Sequence Search

Duração: 2 semanas

Objectivos programáticos:

- Alinhamento de sequências
- Algoritmo de *Smith–Waterman*

Descrição da tarefa:

Implemente do algoritmo de *Smith–Waterman* de forma a encontrar padrões no alinhamento de sequências.

Parte 1:

1. Utilize a base de dados GeneBank para obter sequências de genes (DNA) e do seu respetivo transcrito (mRNA). Exemplo:
 - a. <http://www.ncbi.nlm.nih.gov/nuccore/6453568?report=fasta>
 - b. <http://www.ncbi.nlm.nih.gov/nuccore/6448578?report=fasta>
2. Implemente o algoritmo *Smith–Waterman*
3. Utilizando as sequências do ponto 1, anote o gene indicando os exões e intrões.

Parte 2:

1. Altere o código anterior de forma a suportar aminoácidos, utilizando a matriz BLOSUM62 (disponível online).
2. Utilize a base de dados Uniprot para obter a sequência de aminoácidos de duas proteínas homologas à sua escolha. Exemplo:
 - a. <http://www.uniprot.org/uniprot/P51587>
 - b. [http://www.uniprot.org/uniprot/?sort=&desc=&compress=no&query=cluster:\(UniRef50_P51587\)&fil=&force=no&preview=true&format=fasta](http://www.uniprot.org/uniprot/?sort=&desc=&compress=no&query=cluster:(UniRef50_P51587)&fil=&force=no&preview=true&format=fasta)