

# Bioinformática

## Exercícios TP4: Algoritmos de alinhamento e pesquisa

1. Considere as duas sequências:

HEAGAWGHEE

PAWHEAE

- Construa a matriz de alinhamento para estas sequências e calcule o alinhamento global ótimo pelo algoritmo de Needleman-Wunsch, considerando uma gap penalty de -8 e uma matriz de score BLOSUM50.
- Construa um alinhamento semelhante, mas usando o método de alinhamento local de Smith-Waterman.

2. Obtenha as sequências da tripsina humana e de ratazana. Produza o seu alinhamento global com o programa “needle” ([www.ebi.ac.uk/emboss/Tools/psa](http://www.ebi.ac.uk/emboss/Tools/psa)). O software usa por defeito uma matriz BLOSUM62 para fazer este alinhamento, mas qual lhe parece que deveria ser a matriz apropriada para esta situação. Volte a correr “needle” com esta matriz e compare com o resultado anterior. Obtenha a sequência da tripsina de *Litopenaeus vannamei* (white leg shrimp) e faça o alinhamento desta com a humana. Qual seria a matriz de alinhamento mais apropriada nesta situação?... Repita o alinhamento com esta última e compare.

3. Obtenha uma lista de sequências da “Matrix Gla Protein” (MGP) a partir do portal Uniprot ([www.uniprot.org](http://www.uniprot.org)), (**Search in:** UniprotKB ; **Query:** matrix gla protein)

- Examine os resultados. Parece-lhe que todas as sequências obtidas pertencem à família das MGP's. Porquê ?
- Repita a pesquisa anterior mas, sem preencher a caixa **Query**, use a opção **Advanced Search** usando como **Field** “Protein Name” e preenchendo a caixa correspondente com “matrix gla protein”. Compare com os resultados obtidos anteriormente.
- Selecione a MGP humana, faça uma pesquisa BLAST com esta sequência contra toda a base de dados Uniprot usando o portal SRS (<https://www.ebi.ac.uk/Tools/sss/ncbiblast/>) (no Step3 – Parameters, mude o valor “Scores” de 50 para 100. Uma das sequências obtidas na alínea a não deveria aparecer na lista agora obtida, a sequência *MGP\_PRIGL* do “blue shark”. Qual será a razão por que esta sequência não é agora listada?
- Use a sequência *MGP\_PRIGL* para pesquisar com blast a base de dados UniprotKB. Analise os resultados e copie uma das sequências obtidas, a *A8YQS4\_PRIGL*
- Use a sequência gravada na alínea anterior para uma nova pesquisa blast2 contra UniprotKB.
- De acordo com os resultados obtidos em c), não é claro se a sequência *D2BNG8\_LACLK* pertence de facto à família das MGPs. Para responder a esta questão, faça uma nova pesquisa blast desta sequência contra a base de dados UniprotKB e analise o resultado obtido.