

## Bioinformática

### Exercícios TP5 : Pesquisa de sequências

1. Recolha as sequências de tripsina dos seguintes organismos: homem, chimpanzé, boi, ratazana, salmão, mosquito da malária (*Anopheles gambia*), fungo *Phusarium solanum*, bactéria *Streptomyces griseus* (Nota: sempre que possível, usa a sequência da tripsina 1).
  - a) Copie as sequências das várias proteínas no formato FASTA para um ficheiro de texto.
  - b) Abra a página da aplicação [CLUSTAL OMEGA](#) no EBI (European Bioinformatics Institute). Carregue as sequências para a caixa de input, e execute o programa com os parâmetros por defeito.
  - c) Observe o alinhamento obtido. As colunas completamente conservadas estão marcadas com um “\*”, enquanto níveis de conservação mais baixos são marcados com os símbolos “:”, “.” ou “ ”.
  - d) Para observar a árvore de guia (“guide tree”) usada para construir o alinhamento, carregue no botão “Guide Tree”. Observe a árvore. Parece-lhe que o agrupamento está de acordo com distância entre os grupos taxonómicos aos quais pertencem os diversos organismos?
  - e) Para ver as percentagens de identidade entre as diferentes sequências no alinhamento, carregue no botão “Result Summary” e escolha o link na caixa “Percent Identity Matrix”. Observe o intervalo de valores de % de identidade e compare-os com os agrupamentos observados na árvore de guia.
  - f) Carregue no botão “Alignments” para visualizar de novo o alinhamento. Identifique as regiões conservadas (caracterizadas por sequências de “\*\*”) e identifique as inserções e deleções entre os vários subgrupos de sequências.
  - g) As regiões muito conservadas através de um alinhamento múltiplo de uma família de proteínas estão frequentemente associadas à função ou conservação da estrutura da proteína. No caso presente, e tratando-se de uma família de enzimas, é expectável que as zonas conservadas estejam relacionadas com o centro ativo e mecanismo catalítico do enzima.
  - h) Para verifica a hipótese de g), vamos tentar encontrar os resíduos de aminoácido do centro ativo da tripsina no alinhamento múltiplo. Estes resíduos são uma Histidina (H), um Ácido Aspártico (D) e uma Serina (S). Para encontrar os números destes três resíduos dentro da sequência da Tripsina humana (P07477), abra a página Uniprot correspondente e loca a secção “Features”.
  - i) No alinhamento gerado por ClustalO, localize a sequência humana (TRY1\_HUMAN) e encontre os resíduos através dos números obtidos na alínea anterior (use os números no final de cada linha para iniciar as contagens. O que se observa quanto à conservação destes resíduos?
  - j) Vamos comparar o alinhamento obtido com ClustalO com aquele que se obtém usando um outro software de alinhamento múltiplo, [T-Coffee](#).
  - k) Após abrir a página de T-Coffee no EBI, faça pasta das sequências anteriormente usadas na janela de input do programa e execute com os parâmetros por defeito.
  - l) O output de T-Coffe está formatado de maneira um pouco diferente de ClustalO, o que dificulta um pouco a comparação. Vamos facilitar o processo usando uma ferramenta de visualização de alinhamentos do EBI, chamada MVIEW. Para a usar, use o botão “Result Viewers” em ambas os alinhamentos (CustalO e T-Coffee) e selecione a opção “View in MView” em ambos os casos. Agora os alinhamentos estão formatados e coloridos com os mesmos parâmetros, sendo mais fácil a sua comparação.
  - m) Observe as diferenças claras que existem na formação de algumas regiões dentro do alinhamento.

- n) Vamos comparar os alinhamentos obtidos com um alinhamento múltiplo obtido por sobreposição das estruturas de algumas das tripsinas usadas no alinhamento (aquelas para as quais foi resolvida a estrutura cristalográfica), usando o software [DALI](#). Para tal, necessitamos dos códigos PDB das estruturas correspondentes, a saber: 4WWY.A (tripsina humana), 1BTY.A (tripsina bovina), 1UTK.A (salmão), 1FN8.A (*Fusarium oxysporum*), 1SGT.A (*Streptomyces griseus*). Use a opção “Pairwise” e coloque na primeira caixa do código da estrutura tripsina humana, apendendo o código da cadeia (“4WWYA”). Usa o botão “+” para criar múltiplos caixas de entrada para colocar todos os códigos da estruturas da tripsina. Carregue em “Submit”. Na página de resultados escolha a opção “Interactive (html)” e marque as “checkboxes” de todas as estruturas. Carregue em “Structural Alignment” para ver o alinhamento das sequências guiado pela estrutura.