

No encontro dessa semana vamos falar de banco de dados de proteínas e vamos aproveitar para abordar um pouco alguns dos dilemas na anotação de função e domínios de proteínas. Além de passar um pouco sobre as funcionalidades de alguns bancos de dados e ferramentas embutidas neles.

Além dos materiais de apoio de cada encontro, vamos disponibilizar também um Desafio da Semana. Caso você já se sinta confortável com algum dos tópicos da seção “Materiais de Apoio”, pode ignorá-lo sem problemas, a única parte obrigatória é a instalação dos programas da parte prática antes do encontro em si.

Terceiro encontro (08/10/2020)

TEMA:

Banco de dados de Proteínas

DESAFIO DA SEMANA:

Para o desafio dessa semana decidimos incentivar diferentes formas de responder a mesma pergunta. Usando a sequência de aminoácidos abaixo gostaríamos que vocês tentassem encontrar o nome da proteína formada por ela e sua função biológica.

```
>Desafio_LIB  
MVHLTPVEKSAVTALWGKVNVDVGGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGN  
PKVKAHGKKVLGAFSDGLAHLNLTGKTFATLSSEHCDKLVDPENFRLLGNVLVCVLAHH  
FGKEFTPPVQAAYQKVVAGVASALAHKYH
```

PROGRAMAS DA PARTE PRÁTICA:

Durante esse encontro usaremos algumas plataformas WEB além de algumas ferramentas do Biopython, que pode ser baixado em Biopython.org ou instalado, tanto no Linux quanto no Windows, através do comando:

pip install biopython

Entre alguns dos pacotes do python ou biopython que usaremos para a parte prática está a Bio.SeqIO, urllib e brevemente pela Bio.AlignIO.

MATERIAIS DE APOIO

1. Biologia de Proteínas

Separamos um capítulo do livro [Bioquímica Básica](#) para orientá los sobre o conhecimento biológico de proteínas. Destacamos que não é necessário a leitura de todo o capítulo mas indicamos um bom conhecimento dos seguintes tópicos:

2.3 - Polímero de aminoácidos: Peptídeos e Proteínas

2.4 - Estruturas das proteínas

2.8 - Alterações estruturais de proteínas.

Há também um canal no youtube muito bom com uma playlist de vídeo aulas sobre o assunto que talvez possa ajudar a entender melhor o conteúdo também. Eles aborda o conteúdo acima em dois vídeos: [Aula 5 - Aminoácidos - estrutura](#) e [Aula 6 - Proteínas - estrutura](#)

2. Alinhamento de sequências e BLAST

Se você não sabe nada de alinhamento talvez um google e wikipedia já te dê uma situada no que se trata.

Indicamos depois uma nesse material da embrapa [Guia prático de BLAST](#) ou no capítulo 6 do livro [Bioinformatics for Begginers](#) em que apresenta alguns conceitos básicos de alinhamento e Blast. Ao qual alguns podem ser útil para o encontro de hoje:

6.3 - Sequence Identity and Sequence Similarity

6.4 - Global Versus Local Alignment

6.6 - Alignment Algorithms, Gaps, and Gap Penalties

6.8 - Database Searching with the Heuristic Versions of the SmithWaterman Algorithm—BLAST and FASTA

3. Uniprot

Durante a prática dessa semana usaremos como bancos de dados principais: [Uniprot](#), que possui um [Quick Guide](#) para se conhecer um pouco mais.

A [EMBL-EBI Training](#) tem vários treinamentos em bioinfo entre eles um chamado [“UniProt : Quick Tour”](#) falando o básico sobre a plataforma. Dá pra traduzir o conteúdo com a extensão do Chrome de Google Tradutor.

4. Artigo para discussão

O seguinte artigo [Functional annotation of putative hypothetical proteins from Candida dubliniensis](#) será utilizado como base para a discussão do nosso encontro.