

Atividade 08 - blast

O arquivo *lista_blast.fasta* possui 5 sequências de DNA desconhecidas. A presente atividade tem como objetivo encontrar quais são os organismos que possuem essas informações genéticas desconhecidas. Para isto vamos utilizar o algoritmo BLAST.

Desenvolva essa atividade em um único arquivo py. Coloque comentários para indicar as respostas dos itens

- a) Escreva uma rotina que faça a leitura do arquivo *lista_blast.fasta*. (utilizar o **biopython**) e faça a impressão da seqB
- b) Qual forma de blast (*Blastp*, *Blastn*, *Blastx*) irá utilizar para resolver este problema. Justifique (**coloque a resposta no arquivo py em forma de comentário**)
- c) Desenvolva uma rotina em biopython para fazer o alinhamento BLAST de todas as sequências contidas no arquivo (*lista_blast.fasta*). Salvar os resultados em um arquivo XML. Segue abaixo os parâmetros que precisam ser utilizados para fazer o alinhamento:

Database: nucleotide collection(NR/NT)

Optimize for: Megablast.

Número max de alinhamentos: 60

D) escreva uma rotina que faça a leitura do arquivo XML e imprima as seguintes informações, para alinhamentos com score maiores que 40: sequência, tamanho, score e gaps

Análise dos resultados

- d) Qual é o organismo que pertence as sequências? Justifique sua resposta. (**coloque a resposta no arquivo py em forma de comentário**)
- e) Cite a identificação do genbank que melhor alinha a sequência. (**coloque a resposta no arquivo py em forma de comentário**)
- f) Cite o trabalho na qual está envolvido a sequência do item e. (**coloque a resposta no arquivo py em forma de comentário**)

Entregar os arquivos XML resultados dos alinhamentos e o arquivo PY .