**Navegação e busca de banco de dados – Comparação de sequências, 02/05/2022.**

**Busca de genes e suas proteínas:**

**(RESPONDA CADA PERGUNTA E ENTREGUE PELO EDISCIPLINAS- faça em duas pessoas. – mas entregue individualmente.)**

Hoje deveremos fazer comparação entre genes específicos, utilizando ferramentas de bioinformática do NCBI. A comparação será inicialmente evolutiva, considerando a sequência de DNA e de proteínas, mas faremos as comparações dois a dois, o que facilita a compreensão dos dados.

1. Vamos buscar inicialmente o gene da DNA polymerase gamma humana, de chimpanzé e de peixe.

Acesse a página do NCBI- (https://www.ncbi.nlm.nih.gov/).

a) Onde está marcado All database, escolha GENE- procure a Human Polymerase gamma

***Inicialmente vamos encontrar a sequência do mRNA***

Clique no POLG (*Homo sapiens*) e depois à direita (abaixo) RefSeq RNAs;

Vamos usar o transcrito variante 1 (clique nele)

Depois busque o FASTA e guarde essa sequencia no formato FASTA do mRNA.

***Agora busque a sequência da proteína (você também pode traduzir a sequência do RNA no ORF-Finder)***

Volte ao PolG (*Homo sapiens*), e depois à direita (abaixo) clique em RefSeq proteins.

Pode encontrar qualquer dessas proteínas.

Depois busque o FASTA e guarde essa sequencia no formato FASTA da proteína.

1b) Volte para 1a) e procure a sequência do mRNA e de proteína de chimpanzee DNA polymerase gamma (cuidado aqui você tem que encontrar o gene *POLG* de *Pan troglodytes*, a subunidade catalítica. Também busque o “Predicted transcript variant X1” e “isoform X1”)

1c) Volte para 1b) e procure a sequência do mRNA e de proteína de peixe paulistinha (Zebrafish)- *Danio rerio* (procure “Danio rerio DNA polymerase gamma”).

2) de posse das sequencias comece a comparar os RNAs. Para isso utilizaremos o BLAST two sequences: (https://blast.ncbi.nlm.nih.gov/Blast.cgi?BLAST\_SPEC=blast2seq&LINK\_LOC=align2seq&PAGE\_TYPE=BlastSearch)

Compare os RNAs do gene humano e de chimpanzé (em BLASTN) e veja os resultados:

Pergunta-se (deve ser respondido):

1. **Veja o alinhamento entre as sequencias: qual o índice de similaridade (%identities) entre os RNAs humano e de chimpanzé?**
2. **Veja o graphic summary e o dot plot. Faça um copy e paste dos resultados para colocar em seu relatório. Descreva o que entendeu de cada gráfico em no máximo 3 linhas para cada gráfico.**

Refaça a comparação para humano X peixe paulistinha.

1. **Veja agora o alinhamento das sequências de mRNA para humana e de peixe paulistinha, qual o índice de similaridade entre as duas sequências?**
2. **Compare as sequencias de proteínas (BLASTP) entre humano e chimpanzé . Qual o índice de similaridade (%positives) e identidade? Apresente de forma similar (pergunta 1) os gráficos do “Graphic Summary” e de “dot blot”.**
3. **Refaça a pergunta 4 para a proteína humana e de peixe paulistinha?**
4. **Seus resultados estão de acordo com a distância evolutiva entre essas espécies? Por que?**
5. **Por que a comparação das sequências humana e de peixe paulistinha apresentou resultados tão discrepantes entre mRNA e proteínas?**
6. **Busque a proteína DNA polymerase gamma de levedura (*Saccharomyces cereviseae*) (volte para 1a) e refaça a pergunta 4 comparando as sequencias de proteínas entre humano e levedura.**
7. **Faça uma conclusão geral de seus resultados.**

Para aqueles que querem “brincar” eu sugiro que tente achar e comparar outras sequencias de genes de seu interesse!

A partir da próxima segunda vamos avaliar a conservação de genes do genoma dos coronavírus, tendo como base o genoma do SARS-Cov2.

Faça a busca de genomas de SARS-Cov2 em Genome (e depois encontre a sequência do genoma da linhagem de Wuhan). Faça um ORF-Finder. Que genes são mais prováveis?

Veja no esquema apresentado pelo professor qual a estrutura do genoma do SARS-Cov2.