## O que é Blast?

Blast é uma ferramenta web anexada a um banco de dados contendo sequencias biológicas e suas regiões. Através dele, é possível realizar uma pesquisa de query onde o input é uma determinada sequência escolhida pelo usuário e o output são as sequencias registradas ordenadas por similaridade e região.

## Qual o objetivo do Blast?

É retornar a similaridade entre as sequencias genéticas pesquisadas pelo usuário e as registradas na base de dados, ordenadas por similaridade e região. É utilizado para ajudar pesquisadores e biologia molecular ao extrair dados genéticos de sua pesquisa.

## Qual o princípio de funcionamento de seu algoritmo?

Comparação entre sequencias genéticas e exposição ordenada das mais similares encontradas.

## 2 - Responda: Dentro da pesquisa que fez no Blast o que significa a informação dá o Escore, Escore Máximo, E-value e Identidade? (Explicar o que é cada um deles e a importância para a análise do alinhamento de sequência.) Quanto estes parâmetros foram úteis para análise da sequência identificadas.

## **Escore (Score)**:

## O escore é uma medida da qualidade do alinhamento entre duas sequências. Quanto maior o escore, mais similaridade há entre as sequências. É calculado com base na quantidade de correspondências, mismatches e gaps encontrados durante o alinhamento. Em geral, um escore mais alto indica uma relação mais significativa entre as sequências.

## **Escore Máximo (Max Score)**:

## O escore máximo é o maior escore encontrado em uma busca de similaridade de sequências. Ele representa o melhor alinhamento possível entre as sequências analisadas. É útil para determinar a qualidade geral da correspondência entre as sequências e pode ser usado para comparar diferentes resultados de busca.

## **E-value (Valor-p esperado)**:

## O E-value é uma medida estatística que indica a probabilidade de encontrar um alinhamento com um escore igual ou maior por acaso, apenas devido à aleatoriedade. Quanto menor o valor de E, mais significativo é o alinhamento. Um E-value baixo sugere que a correspondência observada entre as sequências não é resultado do acaso, mas sim de uma relação biológica real.

## **Identidade**:

## A identidade representa a porcentagem de nucleotídeos ou aminoácidos idênticos entre as sequências alinhadas. Em outras palavras, é a proporção de posições na sequência que têm o mesmo nucleotídeo ou aminoácido. A identidade é uma medida direta da similaridade entre as sequências e é útil para avaliar o grau de conservação evolutiva entre elas.