

Programação dinâmica aplicada ao alinhamento de sequências

ALINHAMENTO DE SEQUÊNCIAS

O que é Alinhamento de Sequências?

Comparativo entre sequências para **inferir homologia (Ancestrais Comuns)**.

- Entender qual a funcionalidade de determinada proteína;

Existem as **Mutações** que são as diferenças entre as sequências, podendo ser por:

- **Inserções;**
- **Deleções;**
- **Substituições.**

O estudo do estudo dessas mutações está relacionado as **doenças ligadas a mutações**, devido a grande variedade de doenças que existem, e a **Indústria** (por exemplo, melhorias de enzimas utilizadas nos processos industriais).

Como realizar um alinhamento?

Para fazer um alinhamento, precisamos de pelo menos duas sequências. Ainda precisa-se lembrar dos três conceitos básicos:

- **Match:** quando os mesmos nucleotídeos se alinham (+);
- **Mismatch:** quando nucleotídeos diferentes se alinham (-);
- **Gap:** quando o nucleotídeo não se alinha a ninguém (- -);

(OBS: Cada um desses conceitos é pontuado para definir qual é o melhor alinhamento).

Ainda existe a **Matrizes de pontuação** que, pontuando os alinhamentos, caso aconteça alguma mutação, tem-se uma ideia do tamanho do impacto da mutação na pontuação.

Tipos de Alinhamento

Global

- Exemplo: Algoritmo de Needleman-Wunsch;
- Considera-se o tamanho completo das sequências.

Local

- Exemplo: BLAST;
- Posso alinhar as partes idênticas das sequências.

Múltiplos

- Quando temos várias sequências e quero detectar o que tem de similar entre elas.

PROGRAMAÇÃO DINÂMICA

Como realizar a programação dinâmica?

I. Definir uma pontuação

- a) Tomando alguns valores como exemplo
 - o Match = 5 | Mismatch = -2 | Gap = -6

II. Construir a Matriz de Pontuação

- a) Sempre iniciar com 0
- b) A primeira linha e colunas devem ficar vazias
- c) Por estarem vazias, o preenchimento sempre será somando -6 em cada célula.

III. Preenchendo a Matriz de Pontuação

- a) Com um quadrante, tendo 3 deles preenchidos, verifica-se qual foi o maior possível das três possíveis soluções:
 - a. Match ou Mismatch: cria uma seta para na diagonal superior esquerda e soma-se o valor a origem (onde a seta aponta);
 - b. Gap: se for da sequência superior, cria-se a seta apontada para cima e soma-se ao valor atribuído ao Gap.
- b) Após isso, segue-se preenchendo os valores na vertical e, após preencher a linha, repete-se os mesmos passos na linha de baixo

IV. Realizando o Rastreio

- a) Encontre o maior valor e caminhe pelas setas até chegar ao 0 novamente.

Construindo o Alinhamento

A partir do rastreio, verifica-se nas duas sequências o que ocorreu ao final e verifique a pontuação final (sempre dando prioridade a maior pontuação).

T G C T C G T A

T - - T C A T A

→ No final, o que temos é $5-6-6+5+5-2+5+5$, tendo como resultado o valor 11.