

### 1) Descreva o método de sequenciamento de genoma de Sanger?

O método Sanger envolve a produção de muitas cópias de uma região alvo de DNA. Os ingredientes são similares àqueles usados para replicação de DNA em um organismo, ou para a reação em cadeia da polimerase (PCR), os quais fazem cópias de DNA *in vitro*. Eles incluem:

- Uma enzima DNA polimerase
- Um primer, que é um pequeno fragmento de DNA de fita simples que se liga ao DNA molde e atua como um "iniciador" para a polimerase
- Os quatro nucleotídeos do DNA (dATP, dTTP, dCTP, dGTP)
- O DNA molde a ser sequenciado

Entretanto, uma reação de sequenciamento de Sanger contém um ingrediente único:

- Versões Dideoxi, ou terminadores de cadeia, para os quatro nucleotídeos (ddATP, ddTTP, ddCTP, ddGTP), cada uma marcada com um corante de uma cor diferente

Uma vez que um dideoxynucleotídeo é adicionado à cadeia, não há hidroxila disponível e nenhum outro nucleotídeo pode ser adicionado. A cadeia termina com o dideoxynucleotídeo, o qual é marcado com uma cor particular de corante dependendo da base (A, T, C ou G) que carrega.

A amostra de DNA a ser sequenciada é combinada em um tubo com primer, DNA polimerase e nucleotídeos de DNA. Os quatro nucleotídeos terminadores de cadeia marcados com corantes são adicionados também, mas em concentração muito menor que a dos nucleotídeos comuns.

A mistura é primeiro aquecida para a denaturação do DNA molde, então resfriada para que os primers possam se ligar ao molde de fita simples. Uma vez que o primer se ligou, a temperatura é aumentada novamente, permitindo que a DNA polimerase sintetize um novo DNA a partir do primer. A DNA polimerase continuará a adicionar nucleotídeos à cadeia até que aconteça a adição de um dideoxynucleotídeo ao invés de um normal. Nesse ponto, os nucleotídeos não podem mais ser adicionados, e portanto a fita terminará em um dideoxynucleotídeo.

Este processo é repetido em um número de ciclos. Quando o ciclo se completa, é praticamente garantido que um dideóxinucleotídeo terá se incorporado em todas as posições do DNA alvo em pelo menos uma reação. Ou seja, o tubo irá conter fragmentos de diferentes comprimentos, terminando em cada uma das posições de nucleotídeos no DNA original. As extremidades dos fragmentos serão rotuladas com corantes que indicam o nucleotídeo final.

Depois que a reação é executada, os fragmentos são levados através de um longo e fino tubo, contanto uma matriz de gel em um processo chamado de electroforese capilar em gel. Pequenos fragmentos movem-se rapidamente através dos poros do gel, enquanto longos fragmentos movem-se mais devagar. Assim que cada fragmento cruza a "linha de chegada" no final do tubo, ele é iluminado por um laser, permitindo que o corante anexado seja detectado.

O menor fragmento (que termina apenas um nucleotídeo após o primer) atravessa a linha de chegada primeiro, seguido do próximo menor fragmento (terminando dois nucleotídeos após o primer), e assim por diante. Portanto, a partir das cores dos corantes registradas uma após a outra no detector, a sequência do pedaço original de DNA pode ser construída com um nucleotídeo por vez. Os dados registrados pelo detector consistem em uma série de picos em intensidade de fluorescência. A sequência de DNA é lida a partir dos picos no cromatograma.

## **2) Cite quais possíveis utilidades do alinhamento de sequências no campo da bioinformática?**

Existem centenas de aplicações do alinhamento de seqüências, tanto na identificação de genes e proteínas desconhecidas, quanto na comparação da ordem de genes em genomas de organismos proximamente relacionados (sintenia), no mapeamento de seqüências expressas dentro de um genoma para identificação de genes, etc...

**3) Qual a diferença entre alinhamento global e local? Cite o nome de ferramentas que fazem um ou outro alinhamento.**

Um alinhamento global é uma forma de otimização global que "força" o alinhamento a cobrir todo o comprimento de todas as sequências interrogadas, AMAP (*Annealing* de sequências) é um exemplo de ferramenta de alinhamento global.

Os alinhamentos locais identificam regiões de similaridade dentro de sequências longas que são geralmente bastante divergentes em um todo. Os alinhamentos locais são frequentemente preferíveis, mas podem ser difíceis de calcular por causa do problema adicional de identificar regiões internas de similaridade, ALE (alinhamento manual) é um exemplo de ferramenta de alinhamento local.

**4) É possível realizar alinhamentos utilizando uma sequência de DNA e outra de proteína? Como você acha que isso poderia ser feito?**

Sim, basta converter o DNA em proteína ou o contrário

**5) Qual a diferença entre alinhamentos simples e múltiplos? Quais são as ferramentas de alinhamento (ótimo ou heurístico) mais indicadas para trabalhar com cada um desses tipos de alinhamento? Por quê?**

Um alinhamento múltiplo de sequências é um alinhamento de três ou mais sequências biológicas, geralmente proteínas, ADN ou ARN, já quando apenas duas sequências (DNA ou nucleotídeos) são comparadas entre si, diz-se que o alinhamento é simples.

Alinhamento heurístico traz o resultado mais próximo encontrado mas não o valor específico, ideal para trabalhar com o alinhamento múltiplo uma vez que o mesmo analisa uma gama muito maior de sequências.

Alinhamento ótimo traz como resultado o melhor alinhamento computacional possível, ideal para trabalhar com o alinhamento simples já que o mesmo possui uma apenas duas sequências para serem comparadas, permitindo maior especificidade.