

Biotecnologia e diagnósticos moleculares

Maria Cristina Rocha Cordeiro

Introdução

Em primeiro lugar, é necessário conceituar o que seja biotecnologia e diagnóstico molecular para, em seguida, explicar qual é a relação entre biotecnologia e diagnóstico molecular.

Biotecnologia é um termo relativamente novo, muito embora sua prática seja até antiga. A biotecnologia em si é a utilização de processos biológicos como geradores de tecnologias de forma que se agregue valor a um dado produto. A biotecnologia pode ser considerada um processo antigo quando nos referimos por exemplo à tecnologia de produção de vinhos, cerveja, queijo etc. O que são, em última instância, essas práticas? São processos biológicos que foram aproveitados para a geração de tecnologia tendo em vista a formação de um novo produto com valor agregado. No caso do vinho e da cerveja, o aproveitamento foi da fermentação anaeróbica do suco da uva ou da cevada e, no caso do queijo, do leite.

Diagnóstico molecular pode ser conceituado como sendo o uso de ferramentas moleculares que indique, ou demonstre, a solução de uma dada questão. Por exemplo, no campo da saúde, o termo diagnóstico é utilizado como um meio de esclarecer os sintomas de uma doença. Diagnóstico molecular é quando esse esclarecimento é realizado em nível molecular. No campo da biologia, o diagnóstico molecular pode ser utilizado, por exemplo, quando se quer identificar especificamente uma planta ou um animal. Na agropecuária, o diagnóstico molecular pode estar relacionado com a identificação de uma dada cultivar mais produtiva, mais resistente a uma dada doença, a identificação de um dado patógeno etc. O diagnóstico molecular pode ser considerado como um caminho que leva à biotecnologia e esta, como geradora de um processo, uma tecnologia nova com valor agregado. Como exemplo dessa relação estão os processos de identificação em

nível molecular de, por exemplo, plantas e animais transgênicos e (ou) kits de identificação de patógenos que podem resultar em processos patenteáveis (com valor agregado).

Para que um diagnóstico molecular seja efetivo, para qualquer situação, ele deve ser considerado dentro de quatro quesitos importantes como: a reprodutibilidade, a especificidade, a distinguibilidade e a sensibilidade. Reprodutibilidade significa que o diagnóstico deve ser sempre repetido exatamente da mesma maneira dentro da situação específica. Especificidade significa que deve ser específico para a questão estudada (ex.: planta, patógeno etc); distinguibilidade significa que o diagnóstico ou a identificação é distinta e não assume múltiplas formas (não há falsos positivos ou negativos), é única; e sensibilidade significa que o diagnóstico ou a identificação é capaz de ser realizado com quantidade muito reduzidas de matéria-prima (ex. sangue, tecido vegetal, animal, DNA, RNA, proteína etc). Esses quesitos estão diretamente relacionados com o “objeto” que conduz à identificação, que é chamado de biomarcador.

Biomarcador é a molécula que identifica, reconhece, diagnostica um determinado sistema como uma cultivar, um patógeno, uma enfermidade etc. Essa molécula deve ser selecionada e relacionada com o ser que se quer identificar, o agente causal da enfermidade ou um sistema biológico específico. Pode ser de origem proteica ou outra como o ácido desoxiribonucleico (DNA), ácido ribonucleico (RNA), compostos fenólicos etc.

Outra relação entre biotecnologia e diagnósticos moleculares diz respeito às ferramentas moleculares da biotecnologia moderna que são utilizadas para selecionar e relacionar uma molécula como biomarcadora de um determinado ser vivo, processo ou sistema. Essas ferramentas moleculares assim utilizadas podem ter origem na genômica, na transcriptômica, na proteômica ou no metaboloma.

Genômica é a parte da biotecnologia que estuda os genomas em sentido macro, sua estrutura básica. São, por exemplo, os grandes trabalhos de sequenciamento de genomas que têm, por principal objetivo, revelar o tamanho preciso dos diversos genomas animais ou vegetais; inferir no número provável de genes que se encontra nesse genoma; bem como conhecer a sequência desses diversos genes. As

principais ferramentas utilizadas pela genômica são o sequenciamento automatizado de DNA e as ferramentas de bioinformática de anotação e análise das sequências gênicas. Porém, o sequenciamento completo dos genomas não informa sobre a funcionalidade de muitas das sequências gênicas conhecidas nos estudos da genômica e esta somente pode ser revelada por meio das ferramentas da transcriptômica e da proteômica.

A transcriptômica representa todos os estudos que visam à identificação e análise de genes expressos especificamente em um determinado sistema como, por exemplo, os genes expressos em tecidos vegetais infectados por diversos patógenos, em tecidos de plantas tolerantes a condições desfavoráveis como a seca ou a presença de metais pesados no solo e outras. Diversas metodologias moleculares são utilizadas nos estudos da transcriptômica como a síntese de bibliotecas de cDNA (DNA complementar), o PCR (Polymerase Chain Reaction), o RT-PCR (Reverse transcription PCR), o PCR quantitativo (RTqPCR) etc. Esses estudos inferem, em última análise, situações fisiológicas específicas e contribuem para respostas e identificação dos principais genes (biomarcadores potenciais), que são ativados em diversas dessas situações fisiológicas diferentes.

A proteômica representa os estudos de todas as proteínas que são expressas em células específicas (tecidos específicos, situações fisiológicas específicas). As principais ferramentas metodológicas utilizadas, hoje, nesses estudos são a eletroforese em duas dimensões, o sequenciamento de proteínas, a espectroscopia de massa e as análises de bioinformática.

E, por fim, a metabolômica é o conjunto de ferramentas utilizadas para estudar moléculas relacionadas à identificação de compostos não proteicos ou DNA/RNA como os alcaloides, flavonoides etc. e sua relação com seus ciclos de síntese. Algumas dessas metodologias são a cromatografia líquida de alta pressão (HPLC), a cromatografia em camada fina (TLC) e a ressonância nuclear magnética (NMR).

Outros estudos no campo da genômica são aqueles relacionados com a genética. Nesse tipo de abordagem, utilizam-se como ferramentas experimentais as tecnologias dos marcadores moleculares

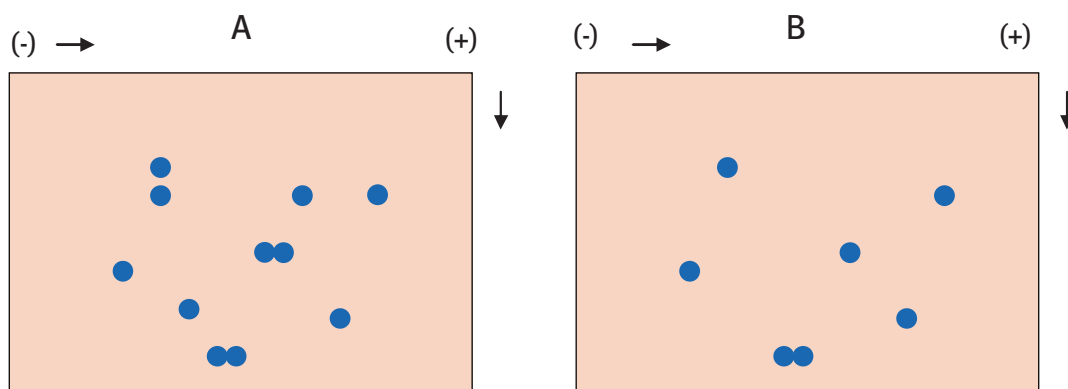


Figura 3. Esquema geral de uma análise de eletroforese em 2D demonstrando a expressão de proteínas diferentes nos tecidos A e B.

Recapitulando, as duas dimensões em que são separadas as proteínas em gel de eletroforese consistem na separação por pontos isoelétricos (pH em que a proteína tem carga nula e, portanto, estaciona no gel) em uma primeira etapa, e, em seguida, a separação em gel por peso molecular. Essa técnica parte da premissa de que as diversas proteínas que existem não têm, ao mesmo tempo, o mesmo ponto isoelétrico e o mesmo peso molecular.

Sequenciamento de DNA

A metodologia do sequenciamento de DNA foi descrita por Frederick Sanger na década de 1970. O procedimento original é realizado por meio da amplificação em PCR da sequência alvo com oligonucleotídeos (*primers*) específicos localizados no vetor (DNA plasmidial) onde está clonado a sequência de interesse alvo (gene) e a utilização de dideoxynucleotídeos (ddNTPs). ddNTPs são nucleotídeos que contêm duas oxidrilas ao invés de apenas uma, normalmente encontrada nos dNTPs. Esses dideoxinucleotídeos são adicionados um a um no processo de amplificação em PCR (*Polymerase Chain Reaction*). Na verdade, fazem-se 4 reações diferentes em que 3 dos dNTPs são normais e apenas 1 é ddNTP. A reação de polimerização das novas cadeias de DNA onde está o ddNTP é parada com a adição do mesmo gerando como produto da PCR uma gama de fragmentos de DNA de tamanhos diferentes de acordo com a adição do ddNTP. Assim, cada uma das reações possuem fragmentos cujas pontas contêm ddATP, outra ddTTP, outra ddCTP e outra ddGTP. Além da adi-

ção diferencial dos ddNTPs nas reações do PCR, é adicionado a cada reação um dos dNTPs (geralmente dATP ou dCTP) marcados com radioatividade. Portanto, os fragmentos gerados em cada reação terão um diferencial na ponta (por causa do ddNTP) e são radioativos.

Essas quatro reações são adicionadas em um gel de poliacrilamida com uréia que permite a resolução dos diferentes fragmentos de DNA sintetizados em cada reação. Após a corrida eletroforética, o gel é seco e exposto a um filme radiográfico. Esse procedimento é chamado de autorradiografia, que representa uma imagem negativa dos fragmentos resolvidos no gel que imprimem sua imagem no filme radiográfico por estarem radioativos. A leitura dessa autorradiografia faz-se de baixo para cima, pois os fragmentos mais baixos são menores do que os fragmentos que estão acima. Como se aplica as quatro reações no mesmo gel, a leitura é feita da banda mais baixa, seja na linha do ddATP, ou ddCTP, ou ddTTP, ou ddGTP, e vai-se alternando conforme as diferentes posições (Figura 4).

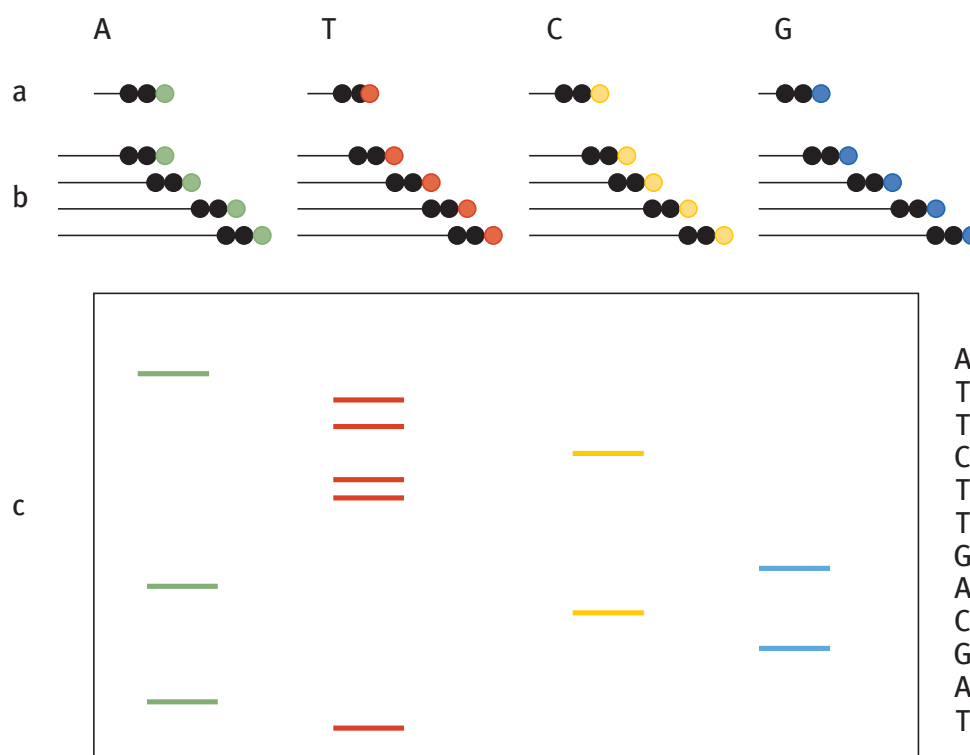


Figura 4. Esquema do procedimento de sequenciamento automático. (a) dideoxinucleotídeos marcados com fluorescência diferencial; (b) reação de polimerização de fragmentos; e (c) separação eletroforética dos fragmentos gerados na reação do PCR.

A sequência lida estará no sentido direto ou reverso conforme a orientação do *primer* utilizado no plasmídeo para a reação de síntese dos fragmentos homólogos gerados do gene alvo. Com a sequência lida, pode-se buscar uma sequência homóloga a ela em banco de dados gênicos ou proteicos utilizando ferramentas de bioinformática como o Blast em suas múltiplas formas como nBlast, tBlast, Blastx etc. Ou várias sequências podem ser analisadas em comparação utilizando o programa Clustal W ou ainda montadas de forma que se encontre a continuidade das sequências gerando uma sequência maior.

Em resumo, como fundamentos teóricos da metodologia do sequenciamento, podemos apresentá-los assim: (1) para a cadeia do DNA crescer em uma reação de polimerização em PCR, é necessária a adição de um nucleotídeo do tipo dNTP, caso contrário ela é parada (adição de ddNTP); (2) se um dos dNTPs na reação de polimerização for marcado de alguma maneira (com radioatividade ou fluorescência), toda a cadeia será marcada da mesma forma; (3) fragmentos de DNA de tamanhos diferentes migram em um gel submetido a um campo elétrico em geral para o pólo positivo e os menores com mais velocidade do que os maiores (princípio da eletroforese); (4) partículas radioativas, porque emitem radioatividade na forma de elétrons livres da camada mais externa dos átomos, podem imprimir filmes radiográficos permitindo uma autorradiografia.

A metodologia do sequenciamento atual mais utilizada é a do sequenciamento automático. Ela está baseada na metodologia original de Sanger, contudo é realizada em um mesmo tubo onde são adicionados os quatro ddNTPs marcados com quatro diferentes fluoróforos que emitem intensidade luminosa diferente quando submetidos à luz. Os picos de diferentes fluorescências são relacionados à presença de um diferente ddNTP. Essa metodologia é mais vantajosa porque tem o poder maior de leitura em relação ao tamanho da sequência, que pode ser lida em uma única reação, e também por não utilizar radioatividade.

O sequenciamento de DNA auxilia um diagnóstico molecular, uma vez que permite o conhecimento de sequências gênicas espécie-específicas que podem ser comparadas por programas de bioinformática e, a partir dessa comparação, *primers* específicos podem ser sintetizados de forma a permitirem a amplificação específica de fragmentos espécie-específicos que funcionem para identificar especificamente

patógenos, plantas, animais, plantas transgênicas (geram biomarcadores) (Figuras 5 e 6). Esses diagnósticos têm sido alvo de muitos trabalhos (FERRI, 2009).

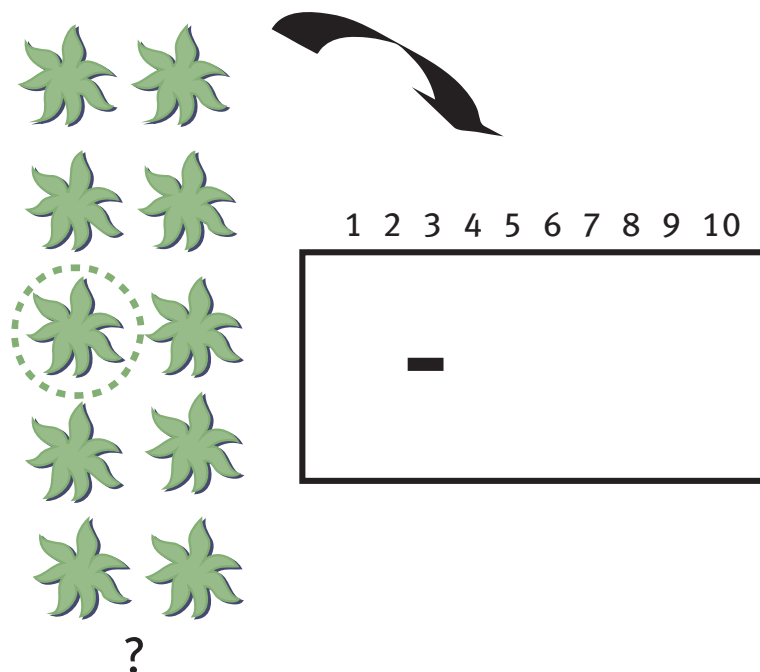


Figura 5. Esquema geral do diagnóstico molecular de uma planta.

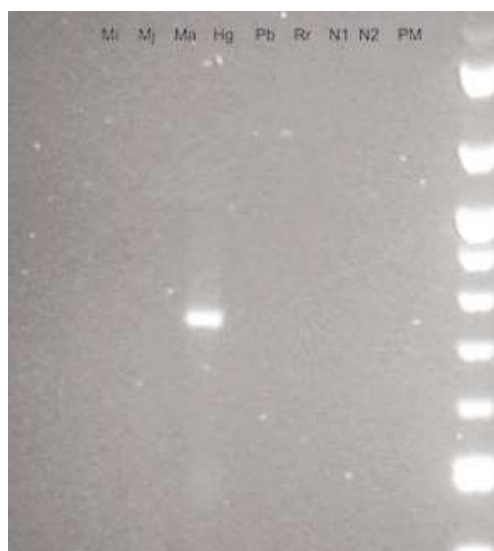


Figura 6. Gel de agarose 2% corado com brometo de etídio que demonstra o diagnóstico molecular específico para o nematoide *Heterodera glycines*. Mi-Meloidogynes incognita; Mj-M. javanica; Ma-M. arenaria; Hg-Heterodera glycines; Pb-Pratylenchus brachyurus e Rr-Rotylenchulus reniformis.

Tecnologias principais utilizadas em diagnóstico molecular

PCR e variações

A reação do Polymerase Chain Reaction (PCR) é uma metodologia que mimetiza a duplicação do DNA em um tubo de ensaio e esse é o seu fundamento teórico. Ela foi descrita pela primeira vez por Mullins em 1983 (SAIKI et al., 1988), provavelmente após a identificação de uma enzima DNA polimerase (enzima que polimeriza cadeias novas de DNA) chamada Taq polimerase, que tem esse nome por ter sido purificada pela primeira vez de uma bactéria chamada *Thermophilus aquaticus*. A Taq polimerase é capaz de resistir a ciclos de alta temperatura periódicos, por ser termoresistente. E essa alta temperatura (92 °C a 94 °C) é necessária na reação de PCR na etapa de desnaturação do DNA.

O PCR é constituído de três etapas fundamentais (Figura 7): desnaturação do DNA; o anelamento de oligonucleotídeos (*primers*); e a polimerização de novas cadeias. A etapa de desnaturação é necessária para que a dupla cadeia do DNA seja aberta, permitindo assim o anelamento dos *primers*; a etapa de polimerização é aquela que polimeriza novas cadeias de DNA a partir de uma cadeia que serve como molde (a duplicação do DNA é semiconservativa). Esse ciclo se repete sucessivas vezes de forma a aumentar a quantidade do fragmento amplificado de maneira que obedeça uma progressão geométrica. Assim, a partir de quantidades ínfimas de DNA, podem-se detectar fragmentos específicos e analisá-los. A grande vantagem do PCR é que é uma metodologia muito sensível, rápida e de relativo pouco custo, dessa forma um diagnóstico molecular que utiliza essa reação tem um grande impacto.