

Tópicos Especiais em Genética Aplicada

Volume 3

Carla Resende Bastos
Janete Aparecida Desidério
Manoel Victor Franco Lemos
Maria Laura Viola Augusto



ISBN 978-85-7805-155-6



Tópicos Especiais em Genética Aplicada

Volume 3

Jaboticabal-SP
Funep - 2016

Copyright©: Fundação de Apoio à Pesquisa, Ensino e Extensão - Funep

Editoração e Impressão Gráfica:
Funep

T674 Tópicos especiais em genética aplicada / Coordenador Manoel Victor Franco Lemos... [et al.]. -- Jaboticabal : Funep, 2016
v. 3, x, 109 p. : il.

Inclui bibliografia

ISBN 978-85-7805-155-6

1. Genética. 2. Biorremediação. 3. Melhoramento genético.
4. RNA. 5. Microrganismos. I. Lemes, Manoel Victor Franco. II.
Título.

CDU 575

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação - Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.



2016

Proibida a reprodução total ou parcial.
Os infratores serão punidos na forma da lei.

Fundação de Apoio a Pesquisa, Ensino e Extensão - Funep
Via de Acesso Prof. Paulo Donato Castellane, s/nº
14884-900 - Jaboticabal - SP
Fone: (16) 3209-1300 / Fax: (16) 3209-1301
Site: <http://www.funep.org.br>

Prefácio

O Curso de Inverno de Genética é realizado pelos alunos de Pós-Graduação e docentes da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, FCAV/UNESP, desde 1994. A iniciativa da elaboração do livro “Tópicos Especiais em Genética” surgiu no ano de 2014, por ocasião da 10ª edição do Curso de Inverno de Genética.

Devido à aceitação pelos leitores, dos volumes I e II deste livro, decidimos continuar difundindo as informações científicas nas áreas apresentadas nos minicursos práticos pelos professores e pós-graduandos da Instituição. Em Tópicos Especiais em Genética Vol. III, os estudantes e profissionais com interesse em temas gerais sobre genética encontrarão assuntos como a molécula de RNA; biorremediação; melhoramento genético florestal, melhoramento clássico de alógamas e autógamias, genômica animal; genômica funcional; epidemiologia molecular de microrganismos; marcadores moleculares; PCR Quantitativa em Tempo Real e LAMP - LOOP MEDIATED ISOTHERMAL AMPLIFICATION (Amplificação Isotérmica Mediada por Loop).

Agradecemos aos docentes e discentes dos Programas de Pós-Graduação (Genética e Melhoramento de Plantas, Microbiologia Agropecuária, Genética e Melhoramento Animal, Zootecnia, Produção Vegetal, Entomologia e Medicina Veterinária) envolvidos na realização de todas as edições do Curso, à direção da FCAV/UNESP, à FUNEP, pelo apoio ao evento, e, em especial, à Comissão Organizadora da 12ª edição do Curso de Inverno de Genética pelo empenho e comprometimento no desenvolvimento desta obra.

*Carla Resende Bastos
Janete Aparecida Desidério
Manoel Victor Franco Lemos
Maria Laura Viola Augusto*

Coordenação

Profa. Dra. Janete Aparecida Desidério

UNESP Jaboticabal
janete@fcav.unesp.br

Prof. Dr. Manoel Victor Franco Lemos

UNESP Jaboticabal
mvictor@fcav.unesp.br

Carla Resende Bastos

Médica Veterinária
Doutoranda em Medicina Veterinária
(Medicina Veterinária Preventiva) - UNESP Jaboticabal
carlarbvet@gmail.com

Maria Laura Viola Augusto

Bióloga
Doutoranda em Agronomia
(Genética e Melhoramento de Plantas) - UNESP Jaboticabal
marialauraa@gmail.com

Comissão Organizadora

Claudio Damasceno Pavani

Biólogo
Doutorando em Microbiologia Agropecuária - UNESP Jaboticabal
claudio_dp913@hotmail.com

Carlos Henrique Caprio

Engenheiro Agrônomo
Mestrando em Agronomia
(Genética e Melhoramento de Plantas) - UNESP Jaboticabal
carloscaprio@hotmail.com

Daniele Araujo Pereira

Médica Veterinária
Doutoranda em Medicina Veterinária
(Clínica Médica Veterinária) - UNESP Jaboticabal
daniele.araujop@gmail.com

Jorge Augusto Petroli Marchesi

Biólogo

Mestrando em Genética e Melhoramento Animal - UNESP Jaboticabal
jorgea_petroli@hotmail.com

Paula Castanho Borges

Engenheira Agrônoma

Mestranda em Genética e Melhoramento de Plantas - UNESP Jaboticabal
paulacastanhob@gmail.com

Renan Furlan Gonsaga

Engenheiro Agrônomo

Mestrando em Agronomia
(Genética e Melhoramento de Plantas) - UNESP Jaboticabal
renanfurlan@gmail.com

Romeu Moreira dos Santos

Médico Veterinário

Doutorando em Medicina Veterinária
(Patologia Animal) - UNESP Jaboticabal
romeumdsantos@hotmail.com

Vânia Casari Mussato

Biomédica

Mestranda em Genética e Melhoramento Animal - UNESP Jaboticabal
vanielcasari@hotmail.com

Sumário

Capítulo 1

A Molécula de RNA: Manipulação, Extração e suas

Aplicações.....	1
2. Manipulação do RNA.....	3
3. Coleta de material e extração de RNA.....	3
4. Verificação da Concentração e Integridade de RNA.....	5
4.1 Espectrofotometria e Fluorometria.....	5
4.2 <i>Bioanalyzer</i>	6
5. Aplicações da Molécula de RNA.....	7
5.1 Northern blot.....	7
5.2 PCR Quantitativo em Tempo Real (qRT-PCR – Real Time Quantitative PCR).....	7
5.2.1 Síntese da Primeira Fita de cDNA.....	8
5.2.2 qRT-PCR.....	8
5.3 Sequenciamento de Nova Geração.....	9
6. Referências bibliográficas.....	11

Capítulo 2

Biorremediação de Solos Contaminados por Cádmio.....13

2. Uso da fitorremediação na limpeza de solos contaminados por cádmio	14
3. Bactérias como facilitadoras da fitoextração de metais.....	16
4. Referências bibliográficas.....	17

Capítulo 3

Atualidades em Melhoramento Genético Florestal.....21

2. Clonagem de espécies florestais	22
3. Proteômica no melhoramento florestal	25
4. Transgênese florestal	26
5. Referências bibliográficas.....	28

Capítulo 4

Estudo da variabilidade genética aplicado

à genômica animal	31
2. Conceitos básicos de genética de população.....	32
3. Conceitos básicos em genética quantitativa	33

4. Conceitos básicos em genética molecular	36
5. Seleção genômica e estudos genômicos	37
6. Referências bibliográficas.....	39

Capítulo 5

Epidemiologia molecular e diversidade genética

de microrganismos	41
2. Identificação de espécies por PCR espécie-específico.....	42
3. Abordagem polifásica para classificação microbiana	42
4. Diversidade genética: principais marcadores moleculares e suas aplicações	44
4.1 AFLP.....	44
4.2 Microssatélites.....	45
4.3 REP-PCR.....	47
4.4 PFGE.....	47
5. Abordagem metagenômica	48
6. Referências bibliográficas.....	48

Capítulo 6

Genômica funcional: análise de transcriptoma

animal por RNA-Seq.....	53
2. Isolamento do RNA, preparação da biblioteca e estratégia de sequenciamento	54
3. Processamento e análise de dados para obter o perfil da expressão diferencial de genes.....	56
3.1 Anotação funcional	58
4. Metodologias de validação da expressão de genes (qRT-PCR).....	59
5. Referências bibliográficas.....	61

Capítulo 7

Principais marcadores moleculares utilizados

no melhoramento de plantas.....	65
2. Marcadores Baseados em Hibridização	66
2.1 RFLP (<i>Restriction Fragment Length Polymorphisms</i>).....	66
3. Marcadores Baseados em PCR.....	67
3.1 RAPD (<i>Random Amplified Polymorphic DNA</i>).....	67
3.2 AFLP (<i>Amplified Fragment Length Polymorphism</i>)	68
3.3 SSR (<i>Simple Sequence Repeats</i>)	69

4. Marcadores Baseados em Sequenciamento	70
4.1 SNP (<i>Single Nucleotide Polymorphism</i>)	70
5. Considerações Finais	71
6. Referências bibliográficas	72

Capítulo 8

Melhoramento Clássico de Alógamas	73
1.1 Equilíbrio de Hardy-Weinberg	74
1.2 Significado da seleção no melhoramento de espécies alógamas	75
2. Métodos de melhoramento populacional	76
2.1 Seleção Intrapopulacional	76
2.1.1 Seleção massal	76
<i>Seleção massal estratificada</i>	76
<i>Seleção massal estratificada geneticamente</i>	77
2.1.2 Seleção com teste de progênes	77
2.1.3 Seleção recorrente intrapopulacional	78
2.2 Seleção Interpopulacional	78
2.2.1 Seleção Recorrente Recíproca (SRR)	78
3. Métodos de obtenção de híbridos	79
3.1 Obtenção de linhagens endogâmicas	79
<i>Método genealógico ou padrão</i>	79
<i>Single seed descendent (SSD)</i>	80
<i>Cruzamento das linhagens</i>	80
4. Referências bibliográficas	81

Capítulo 9

Princípios e aplicações da PCR em Tempo Real Quantitativa e LAMP (Loop Mediated Isothermal Amplification)	83
1. Bases da PCR em Tempo Real Quantitativa	83
2. Sistemas de Detecção	84
2.1 Intercalantes DNA	84
2.2 Sondas de Hidrólise (Sistemas TaqMan®)	85
3. PCR em tempo real quantitativa relativa	87
3.1 Quantificação relativa <i>normalizada pela unidade de massa</i> (Exemplo 1)	88
3.2 Quantificação relativa <i>normalizada pelo gene de referência</i> (Exemplo 2)	89

<i>Método de Livak (Método $2^{-\Delta\Delta Cq}$)</i>	90
<i>Método do ΔCq</i>	91
<i>Método de Pfaffl</i>	91
4. PCR em tempo real quantitativa absoluta.....	92
5. LAMP (Loop Mediated Isothermal Amplification).....	94
6. Referências bibliográficas.....	96

Capítulo 10

Melhoramento Genético de Plantas Autógamas	99
2. Pré-Melhoramento	100
3. Hibridações Artificiais em Plantas Autógamas.....	101
4. Métodos de Condução de Populações Segregantes.....	102
5. Método Genealógico (<i>Pedigree</i>).....	103
6. Método Populacional (<i>Bulk</i>).....	104
7. Método SSD (<i>Single Seed Descent</i>).....	105
8. Método do Retrocruzamento	106
9. Ensaios de VCU	108
10.Referências bibliográficas.....	109

A Molécula de RNA: Manipulação, Extração e suas Aplicações

*Flávia Maria de Souza Carvalho
Juliana da Silva Vantini
Michelle Mendonça Pena
Tamiris Kempner
Maria Inês Tiraboschi Ferro*

1. Introdução

A partir da elucidação da estrutura em dupla hélice da molécula de DNA e sua função no armazenamento da informação genética, foi necessário descobrir como essa informação era transportada até os ribossomos (no citoplasma) para a síntese proteica. Foi em 1961 que este processo foi descrito, nomeando como RNA mensageiro a molécula carregadora da informação genética (BRENNER; JACOB; MESELSON, 1961).

O RNA é formado por uma longa cadeia não ramificada contendo quatro diferentes tipos de nucleotídeos unidos entre si por uma ligação fosfodiéster. A quantidade das bases púricas (adenina e guanina) e pirimídicas (citossina e uracila) não é complementar como ocorre no DNA, já que a molécula é de fita única. A instabilidade da molécula de RNA é devido à presença do grupamento hidroxil no carbono 2 da pentose. Por isso, é facilmente clivada em mononucleotídeos quando em contato com soluções alcalinas ou em condições de temperaturas acima de 65°C, além de ser degradada pelas ribonucleases ou RNases (NELSON; COX, 2014).

Existem quatro classes principais de moléculas de RNA: RNA mensageiro (mRNA), RNA ribossômico (rRNA), RNA transportador (tRNA) e microRNA (miRNA). Cada uma dessas moléculas desempenha uma função específica: o mRNA é responsável por transmitir a informação genética do núcleo até o citoplasma, direcionando a

síntese proteica; os rRNAs são transportados para o citoplasma após se acumularem no nucléolo e associam-se com proteínas ribossômicas para formar o ribossomo; os tRNAs são moléculas adaptadoras da síntese proteica que se ligam covalentemente a um aminoácido em uma extremidade, pareando com um mRNA de forma a unir os aminoácidos a um polipeptídeo crescente na sequência correta; os miRNAs, por sua vez, são RNAs não codificantes com cerca de 22 nucleotídeos, complementares a determinadas regiões de mRNAs que estão envolvidos na regulação gênica pela clivagem do mRNA ou suprimindo sua tradução (NELSON; COX, 2014).

Na maioria dos organismos, os RNAs presentes nas células são sintetizados ou transcritos a partir de uma fita molde de DNA, e o produto será um RNA de fita simples. Diferentemente do processo de replicação, a transcrição não é feita ao longo de toda a molécula de DNA, mas, sim, em genes individuais ou grupos de genes selecionados. Para iniciar a transcrição, a enzima RNA polimerase tem de se ligar ao sítio promotor da molécula de DNA a ser transcrita (NELSON; COX, 2014).

A molécula de RNA transcrita em procariotos a partir de um gene constitui-se no RNA propriamente dito, sendo utilizada, posteriormente, na síntese de proteínas e, em seguida, degradadas. Entretanto, em eucariotos, a situação é bem mais complexa, já que o RNA sofre uma série de modificações antes de ser utilizado como mediador na síntese de proteínas (NELSON; COX, 2014).

Estudos realizados com a utilização da molécula de RNA têm-se tornado cada vez mais frequentes e promissores no meio científico. Diversas são suas aplicações: na área agrícola, pode ser aplicada na identificação de genes candidatos para suscetibilidade e resistência à doença, predição da função de genes e regiões regulatórias (MOCHIDA; SHINOZAKI, 2010) entre outras; na área médica, em tratamento de câncer, identificação de doenças, manipulação de fármacos, etc. Recentemente, tanto na medicina como na agronomia, o uso de RNA de interferência - RNAi (moléculas de RNA complementares a mRNAs, podendo levar ao silenciamento gênico) tem sido uma das mais novas e grandes promessas da biologia molecular, provocando uma verdadeira revolução na forma como é estudado o funcionamento dos genes nas células e, consequentemente, levando a novas perspectivas de intervenção direta na ação gênica *in vivo* (MENCK, 2010).

2. Manipulação do RNA

Uma das maiores preocupações na extração de RNA é a ação das ribonucleases ou RNases, que são responsáveis pela degradação do RNA. Tais enzimas estão presentes em bactérias, fungos, protozoários, plantas superiores e vertebrados (ZIMMER et al., 2009). Ao contrário do DNA, as RNases são altamente estáveis, permanecendo ativas mesmo após fervuras prolongadas ou autoclavagem; além disso, não necessitam de íons metálicos como cofatores como as DNases (SCHAEFER, 2006). Diante disso, para evitar a degradação do RNA pelas RNases, são necessários alguns cuidados durante a manipulação das amostras.

O laboratório deve ser de uso exclusivo para a manipulação do RNA e deve estar sempre limpo e climatizado (16-20°C). Os materiais utilizados na coleta e na extração do RNA devem ser tratados previamente, para ficarem livres de RNases. As bancadas, pipetadores, entre outros materiais utilizados, devem ser de uso exclusivo para RNA e limpo com solução de NaOH ou RNaseZap (Ambion), antes do início da extração. As ponteiros devem ser livres de RNases, e de preferência, com filtro. As vidrarias devem ser lavadas com detergente especial (Alconox 1%), enxaguadas várias vezes em água destilada e queimadas em estufa a 180 °C, por 8 horas, enquanto os plásticos devem ser tratados com NaOH 0,5M por, no mínimo, 10 minutos ou com H₂O₂ (10%) por, no mínimo, 30 minutos, enxaguando no final com água tratada com DEPC (dietilpirocarbonato). É recomendado o uso de luvas, de máscaras, bem como evitar falar durante o processo de extração (POÇAS-FONSECA, 2003; SCHAEFER, 2006).

3. Coleta de material e extração de RNA

São necessários alguns cuidados na coleta e no processamento do material para a extração do RNA. Antecedendo a coleta, todos os materiais e utensílios a serem utilizados devem estar previamente tratados contra RNases, considerando todos os cuidados citados acima.

Após a coleta, a amostra deve ser imediatamente congelada em nitrogênio líquido antes de ser congelada a -70°C ou inferior, e

mantida nessa temperatura, independentemente do tempo de duração do armazenamento. A amostra também pode ser mantida em solução estabilizadora de RNA, como, por exemplo, o RNAlater (ThermoFisher) e o RNA holder (BioAgency Biotecnologia). As amostras também podem ser processadas para extração de RNA logo após a coleta (máximo uma hora após a coleta) desde que mantidas refrigeradas (MELO et al., 2010).

A extração do RNA pode ser realizada, utilizando protocolos seguindo orientações de seus fabricantes, como o método do Trizol® (Invitrogen) e do Brazol® (LGC Biotecnologia). Além disso, a extração pode ser realizada por meio de kits comerciais, como, por exemplo, PureLink® RNA Mini Kit (Ambion), SV Wizard Isolation RNA System (Promega®), Invisorb® Spin Plant RNA Mini Kit (Invitex®), entre outros kits disponíveis comercialmente.

Os métodos de extração de RNA são baseados na lise celular (liberação dos ácidos nucleicos totais) e na proteção dos RNAs (inibidores de RNases). Desse modo, a primeira etapa, na maioria dos métodos de extração de RNA após a pulverização dos tecidos em baixa temperatura, é a imediata exposição deste material a tampões de extração, que irão promover a lise celular. Esses tampões apresentam substâncias como o cloreto de lítio, que auxilia na precipitação do RNA, e o isotiocianato de guanidina, que permite a manutenção do RNA intacto nas etapas seguintes por meio da degradação das ribonucleases endógenas (BITENCOURT et al., 2011).

Após a lise, faz-se a separação dos componentes celulares através de centrifugação. Nesta fase, agentes como trizol, fenol, clorofórmio e álcool isoamílico, causam a precipitação dos resíduos de proteínas pré-quebradas no primeiro passo, além de outros resíduos celulares, deixando o RNA livre no sobrenadante. Nesta fase, é importante tomar cuidado com o pH da solução de purificação, que deve estar entre 4,0 - 5,0. O sobrenadante contendo o RNA é então separado dos contaminantes. Em seguida, o RNA deve ser precipitado com isopropanol ou etanol. O RNA não é solúvel em álcool, portanto ele precipita na presença de alta concentração deste. Quando submetido à centrifugação, o ácido nucleico forma um pequeno sedimento no tubo (*pellet*), permitindo separá-lo do álcool utilizado. O *pellet* é então lavado com etanol (70-80%) para remoção de sais e restos de reagentes. O RNA purificado é então ressuspensionado em água tratada com DEPC (*Diethyl Pyrocarbonate*),

água ultrapura ou algum outro tampão eluente.

Quando a extração é realizada por kits, após a separação em fases com fenol, o sobrenadante é colocado em uma coluna, composta por sílica e com carga positiva. Como os ácidos nucleicos possuem carga negativa, eles ligam-se à superfície da coluna, enquanto proteínas, reagentes, lipídios e outros contaminantes passam através de seus poros. A coluna, então, passa por um tratamento com DNA-se e lavagens com etanol (70-80%), eliminando o DNA, enquanto o RNA se mantém preso à coluna. Por fim, adiciona-se uma solução aquosa que desfaz a afinidade da sílica pelo RNA, eluindo-o.

4. Verificação da Concentração e Integridade de RNA

4.1 Espectrofotometria e Fluorometria

Após a extração de RNA total, a concentração e a pureza das amostras devem ser analisadas. É possível fazer essas análises através de equipamentos do tipo espectrofotômetros e/ou fluorômetros.

A determinação da concentração e da contaminação do RNA pode ser realizada através da leitura da absorbância em equipamento do tipo espectrofotômetro, verificando os valores nos comprimentos de onda de 260 nm, 280 nm e 230 nm. De modo que, no comprimento de 260 nm, lê-se: ácidos nucleicos; 280 nm lê-se; proteínas, e em 230 nm o equipamento lê: contaminações por sais, polissacarídeos e compostos orgânicos (por exemplo, fenol e tiocianato de guanidina). Os valores de referência para as razões 260/280 é de 1,8 a 2,0, e para 260/230 é ≥ 2 (SAMBROOK; FRITSCH; MANIATS, 1989).

Outra maneira de quantificar a amostra de RNA é através do equipamento Qubit® 2.0 Fluorometer (Life Technologies) e o kit específico para quantificar RNAs (Qubit RNA Assay Kit). Este kit possui corantes fluorescentes que se ligam especificamente na molécula de RNA, e o equipamento, por sua vez, é capaz de detectar a emissão de fluorescência. Este equipamento, além de altamente preciso, permite identificar eventuais contaminações com DNA genômico e proteínas quando utilizado com os respectivos kits de quantificações (Qubit dsDNA Assay Kit ou Qubit Protein Assay Kit). Desta forma, a quantificação por fluorometria possui a vantagem, em relação ao

espectrofotômetro, de poder quantificar os ácidos nucleicos (DNA e RNA) separadamente. Por outro lado, o método por fluoróforos não consegue estimar contaminações por compostos orgânicos na amostra, o que é possível através da espectrofotometria.

4.2 Bioanalyzer

Para algumas técnicas moleculares, é de grande importância a utilização de RNAs íntegros. Nesses casos, a integridade do RNA (total ou mensageiro) pode ser analisada por eletroforese microfluidica através do equipamento *Agilent 2100 Bioanalyzer* (Agilent Technologies). Este equipamento também é capaz de analisar DNAs e proteínas. A integridade do RNA pode ser visualizada através de um eletroferograma gerado pelo equipamento. A avaliação tem como base a razão entre as subunidades ribossomais 28S e 18S para amostras provenientes de eucariotos e 23S e 16S para procariotos (geram-se 2 picos bem definidos), que devem ser de 2:1 em caso de RNAs com alta integridade. Quando ocorre degradação dessas subunidades de RNA, os respectivos picos diminuem e aparecem outros picos correspondentes aos fragmentos de RNA. Em função desta razão 2:1 (28S:18S, ou 23S:16S) e dos picos referentes à degradação da amostra, em comparação a um marcador molecular, o equipamento determina o número de integridade do RNA, ou RIN (*RNA integrity number*), que varia de 1 a 10, sendo 10 o maior valor de integridade. Abaixo, ilustração de amostras de RNA apresentando integridade e degradação (Figuras 1A e 1B).

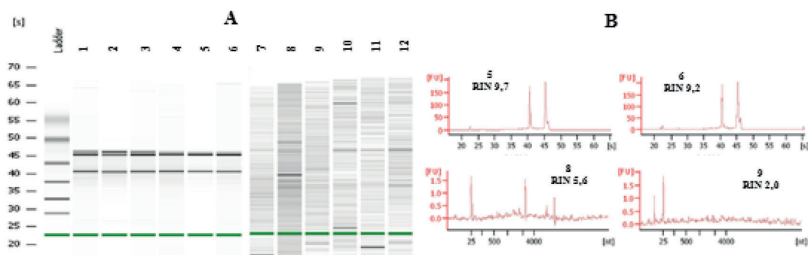


Figura 1. A) Resumo de uma corrida de eletroforese microfluidica no equipamento *Agilent 2100 Bioanalyzer*. Ladder corresponde ao marcador molecular. Colunas de 1 ao 6 correspondem a amostras de RNA íntegras, e colunas de 7 ao 12 correspondem a amostras de RNA degradadas. **B)** Exemplos de eletroferogramas gerados pelo *Agilent 2100 Bioanalyzer*.

5. Aplicações da Molécula de RNA

5.1 *Northern blot*

O *Northern blot* é uma técnica simples, de baixo custo e muito eficiente para estudar e quantificar a expressão gênica de um determinado organismo, tanto em situações normais como em anormais de doença ou de estresses bióticos e abióticos. Essa técnica possui este nome devido à similaridade de seu procedimento com o *Southern blot*, o qual foi descrita pelo biólogo britânico Edwin Southern, sendo que a diferença-chave entre elas é o tipo de ácido nucleico estudado. No *Southern blot*, utiliza-se o DNA, e no *Northern blot*, o RNA total.

A técnica baseia-se na separação das moléculas de RNA por peso molecular através de uma electroforese (com adição de formaldeído no gel de agarose, que funciona como desnaturante para a molécula de RNA). O RNA do gel será transferido para uma membrana de náilon (Hybond H+) para que o mesmo fique fixo permanentemente, e a detecção da molécula de RNA mensageiro de interesse é realizada pela hibridação por complementariedade total ou parcial de uma sonda específica (DNA ou RNA). A sonda geralmente é marcada com átomos radioativos; neste caso, utiliza-se uma autoradiografia em filme de raios-x, a mesma é mantida em freezer a -80°C por aproximadamente uma semana. No caso de não ocorrer hibridação por não existir a sequência de RNA mensageiro de interesse na amostra em estudo, apenas o marcador molecular emitirá o sinal. Para maiores detalhes, Vantini et al. (2008) e Sambrook et al. (1989).

5.2 PCR Quantitativo em Tempo Real (qRT-PCR – Real Time Quantitative PCR)

Para a realização da técnica de PCR quantitativo em tempo real (qRT-PCR), os RNAs totais extraídos do organismo (e/ou tecido e/ou célula) de interesse devem ser transcritos para DNA complementar, através da síntese da primeira fita de cDNA. Cabe aqui dizer que algumas outras técnicas moleculares que avaliam a expressão gênica também utilizam cDNA como *template*, como, por exemplo, as técnicas de cDNA-AFLP, macroarranjo, microarranjo, biblioteca subtrativa, SAGE, RNASeq, etc.

5.2.1 Síntese da Primeira Fita de cDNA

A síntese da primeira fita do DNA complementar (cDNA) é dependente da enzima *transcriptase reversa* (uma DNA polimerase dependente de RNA). Este processo de polimerização de moléculas de DNA a partir de moléculas de RNA foi verificado primeiramente em vírus (*in vivo*). O isolamento desta enzima possibilitou a síntese *in vitro* de DNA (complementar) usando-se o RNA como molde.

Há vários kits comerciais que proporcionam a síntese da primeira fita de cDNA. Os kits possuem todos os componentes necessários para que a síntese ocorra corretamente, tais como: a enzima *transcriptase reversa*, dNTPs, tampão, inibidor de ribonuclease, primers (oligo d(T) e *random hexamers*) e o corante fluorescente que se liga ao DNA. Normalmente, a *transcriptase reversa* atua em temperaturas que variam de 42°C a 60°C, e é capaz de sintetizar cDNA a partir de RNA total ou RNA mensageiro [neste caso, capturam-se os RNAm que possuem cauda poli A, através de oligos d(T)]. O cDNA fita simples deve ser estocado a -20°C até o momento de uso na reação da qRT-PCR.

5.2.2 qRT-PCR

A técnica de PCR Quantitativo em Tempo Real derivou da tradicional técnica de PCR (*Polymerase Chain Reaction* ou Reação em Cadeia da Polimerase). Entretanto, na qRT-PCR, os produtos amplificados (*amplicons*) da sequência de interesse podem ser monitorados em tempo real. A quantificação dos produtos é determinada durante cada ciclo (*cycle threshold* = CT) na fase exponencial da reação, através da fluorescência emitida, sendo que este sinal de fluorescência está diretamente relacionado com a quantidade de *amplicons* gerados (WONG; MEDRANO, 2005). A emissão da fluorescência dá-se através do fluoróforo intercalante da dupla fita de DNA ou de uma sonda que possui em uma extremidade um fluoróforo e, na outra extremidade, outra molécula que aceita a energia do fluoróforo. Esta metodologia utiliza um equipamento com sistema de monitoramento da emissão da fluorescência.

A qRT-PCR pode quantificar de duas formas diferentes: quan-

tificação absoluta e quantificação relativa. No ensaio de quantificação absoluta, é possível obter-se o número de cópias do transcrito de interesse através da comparação de uma curva-padrão que é construída a partir de quantidades conhecidas (número de cópias) e diluídas em série. Este ensaio permite saber o número de cópias do RNA de interesse em uma determinada amostra biológica, enquanto na quantificação relativa, a expressão de um gene de interesse em uma determinada amostra (chamada de experimental) é verificada comparando-se com a expressão do mesmo gene, numa outra amostra (chamada de controle ou calibradora). Como resultado, é possível verificar o aumento ou a diminuição da expressão de genes da situação experimental em relação ao controle. E por fim, faz-se necessário comparar esses valores obtidos com os resultados de expressão de genes de referência nas mesmas amostras. Os genes de referência, também conhecidos como endógenos ou genes constitutivos, são aqueles genes que não devem variar ou variar muito pouco sua expressão, independentemente da situação em que se encontra (se controle ou experimental). Esta comparação entre a expressão de genes de referência com os genes de interesse é um procedimento importante, pois controla as possíveis variáveis experimentais (LIFE TECHNOLOGIES, 2012).

Para que um ensaio de qRT-PCR seja bem-sucedido, primeiramente o experimento precisa ser muito bem planejado e delineado. Em seguida, é preciso realizar as várias etapas (purificação dos ácidos nucleicos, transcrição reversa, desenho de *primers*, titulação dos *primers*, controle da reação, normalização, curva-padrão e eficiência da amplificação, sensibilidade da técnica e reprodutibilidade) que fazem parte de um ensaio de qRT-PCR, pois conduzem à otimização e precisão da técnica, bem como a normalização dos dados. Para maiores detalhes dessas etapas, utilize o “Real Time Handbook” (Life Technologies).

5.3 Sequenciamento de Nova Geração

As tecnologias de sequenciamento de nova geração, implantadas desde 2005, estão permitindo o sequenciamento de DNA e/ou cDNA (RNA) em plataformas capazes de gerar informação so-

bre milhões de pares de bases em uma única corrida. São plataformas que geram um poder de informação muitas vezes maior que o sequenciamento de Sanger, com diminuição nos custos e no tempo de trabalho laboratorial devido ao uso da clonagem *in vitro* e de sistemas de suporte sólido para as unidades de sequenciamento. Além do mais, estas técnicas não estão limitadas à detecção de transcritos que correspondam a sequências genômicas existentes. Dentre estas vantagens, o RNA-seq requer menos amostra de RNA por não necessitar de passos de clonagem (WANG et al., 2008).

A superfície de clonagem (*flow cells*) é dividida em oito linhas que podem ser utilizadas para o sequenciamento. Em cada linha, adaptadores são fixados à superfície pela extremidade 5', deixando a extremidade 3' livre para servir na iniciação da reação de sequenciamento dos fragmentos imobilizados no suporte por hibridização. Os fragmentos de cDNA da amostra também são ligados aos adaptadores em ambas as extremidades, o que permite sua fixação ao suporte de sequenciamento por hibridização a um dos adaptadores fixados. No primeiro ciclo de amplificação, nucleotídeos não marcados são fornecidos para que haja a síntese da segunda fita do fragmento imobilizado no suporte. A alta densidade de adaptadores no suporte facilita a hibridização do adaptador livre dos fragmentos imobilizados a sua sequência complementar fixa perto do clone inicial, durante o ciclo de anelamento. Após o ciclo de anelamento, o fragmento forma uma estrutura em "ponte" na superfície de sequenciamento, e a extensão ocorre, formando a fita complementar, também em "ponte".

No ciclo de desnaturação, as fitas são separadas e linearizadas. Esses ciclos são repetidos 35 vezes e, assim, as cerca de mil cópias geradas de cada fragmento nessa PCR de fase sólida permanecem próximas umas das outras, formando um cluster de sequenciamento. Etapas de desnaturação são necessárias para a separação dos dúplex formados e, nos próximos ciclos de amplificação, nucleotídeos terminadores marcados são fornecidos para as reações de sequenciamento que ocorrem dentro de cada cluster (BELESINI, 2015). A alta densidade dos clusters de sequenciamento possibilita que o sinal de fluorescência gerado com a incorporação de cada um dos nucleotídeos terminadores tenha uma intensidade suficiente para garantir sua detecção exata. Até

50 milhões de clusters podem ser produzidos por linha, correspondendo a uma representação satisfatória da biblioteca. Após a incorporação de cada nucleotídeo no fragmento em síntese, a leitura do sinal de fluorescência é realizada. Em seguida, ocorre uma etapa de lavagem para a remoção dos reagentes excedentes e a remoção do terminal 3' bloqueado e do fluoróforo do nucleotídeo incorporado no ciclo anterior para que a reação de sequenciamento prossiga. A leitura das bases é feita pela análise sequencial das imagens capturadas em cada ciclo de sequenciamento (BELESINI, 2015; SHENDURE; JI, 2008).

6. Referências bibliográficas

BELESINI, A.A. Análise do transcriptoma de folhas de cana-de-açúcar submetidas à prolongada limitação hídrica usando RNA-Seq. 2015. 124p. **Dissertação** (Mestrado em Agronomia/Genética e Melhoramento de Plantas) Universidade Estadual Paulista - Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal, Jaboticabal.

BITENCOURT, G.A.; CHIARI, L.; VALLE, C.B.; LAURA, V.A.; MORO, J.R. Avaliação de diferentes métodos para extração de RNA total de folhas e raízes de braquiária. **Boletim de Pesquisa e Desenvolvimento** – EM-BRAPA Gado de Corte, 2011, 22p.

BRENNER, S.; JACOB, F.; MESELSON, M. An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. **Nature**, 190, p. 576-581, 1961.

LIFE TECHNOLOGIES. **Real Time Handbook**, Foster, 2012. Disponível em: https://lifetech.postclickmarketing.com/Global/FileLib/qPCR/RealTimePCR_Handbook_Update_FLR.pdf. Acessado em: 13 fev. 2016.

MELO, M.R.; MARTINS, A.R.; BARBOSA, I.V.; ROMANO, P.; SHCOLNIK, W. Coleta, transporte e armazenamento de amostras para diagnóstico molecular. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 46, n. 5, p. 375-381, 2010.

MENCK, C. F. M. A nova grande promessa da inovação em fármacos: RNA interferência saindo do laboratório para a clínica. **Estud. av.**, São Paulo, v. 24, n. 70, p. 99-108, 2010. Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142010000300007-&lng=en&nrm-iso>. Access on 12 Feb. 2016. <http://dx.doi.org/10.1590/S0103-40142010000300007>.

MOCHIDA, K.; SHINOZAKI, K. Genomics and Bioinformatics Resources for Crop Improvement. *Plant Cellular Physiology*, v.51, p.497-523, 2010.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. Porto Alegre-RS: Editora Artmed, 6. ed., 2014.

POÇAS-FONSECA, M. J. In: AZEVEDO, M. de O.; FELIPE, M. S. S.; BRÍGIDO, M. de M. (Orgs.). **Técnicas Básicas em Biologia Molecular**. Brasília-DF: Universidade de Brasília, 2003. p. 212.

SAMBROOK, J.; FRITCH, E. F.; MANIATS, T. **Molecular cloning: a laboratory manual**. Cold Spring Harbor: Press. 1989.

SCHAEFER, R. **Técnicas em biologia molecular**. Documentos / Embrapa Suínos e Aves, ISSN 0101-6245; 116, 24 p., Concórdia-SC, Dez. 2006.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. *Nat biotechnology*, v.26, n.10, p.1135-1145, 2008.

VANTINI, J. S.; FERRO, M. I. T.; FERRO, J. A.; OLIVEIRA, C. F. Expressão gênica diferencial da b-1,3-Glucanase (PR-2) nas interações compatível e incompatível entre *Xanthomonas axonopodis* e *Citrus sinensis*. **Científica**, Jaboticabal, v.36, n.2, p.139 - 147, 2008.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**. Advance Online Publication, 2008.

WONG, M.L.; MEDRANO, J. Real Time PCR for mRNA quantification. **Bio-techniques**, v. 39, p. 75-85, 2005.

ZIMMER, K.R.; BORRÉ, G.L.; TRENTIN, D.S.; WOICICKOSKI JÚNIOR, C.; FRASSON, A.P.; GRAEFF, A.A.; GOMES, P.; MACEDO, A.J. Enzimas microbianas de uso terapêutico e diagnóstico clínico. **Revista Liberato**, Novo Hamburgo, v. 10, n. 14, p. 123-137, 2009.

Biorremediação de Solos Contaminados por Cádmio

Gabriela Cabral Fernandes

Claudio Damasceno Pavani

Wellington Pine Omori

Jackson Antônio Marcondes de Souza

1. Introdução

Os metais pesados são um grupo composto de elementos metálicos com propriedades que incluem os que são essenciais para a saúde humana, como o ferro (Fe), o cobalto (Co), o cobre (Cu), o manganês (Mn), o molibdênio (Mo) e o zinco (Zn); os que são elementos-traço e outros que não são essenciais e potencialmente tóxicos, especialmente se eles se acumulam, tais como Mercúrio (Hg), Arsênio (As), Chumbo (Pb), Plutônio (Pu), Vanádio (V), Tungstênio (W) e Cádmio (Cd) (JOHRI et al., 2010).

O Cd é um metal com uma pequena demanda biológica e pouco degradável por processos abióticos e bióticos. Esse pode ser encontrado na natureza associado a sulfetos de minérios de Zn, Cu e Pb. Ele é utilizado principalmente como anticorrosivo em aço galvanizado; já o sulfeto de cádmio e o selenito, como pigmentos em plásticos, e os compostos de cádmio, na manufatura de pilhas e baterias recarregáveis de níquel-cádmio, em componentes eletrônicos e reatores nucleares (CETESB, 2012).

A contaminação por Cd pode resultar de fontes naturais, como a atividade vulcânica, a erosão de rochas sedimentares e fosfáticas e os incêndios florestais, bem como das atividades antropogênicas, tais como mineração, indústria e agricultura. Tipicamente, o Cd do solo é facilmente incorporado em plantas comestíveis e posteriormente entra na cadeia alimentar, provocando efeitos adversos na saúde de seres humanos através do consumo dessas plantas. Esse

metal em solo agrícola é um problema ambiental grave e merece atenção imediata (PRAPAGDEE et al., 2013), pois afeta fortemente as plantas em diferentes processos, tais como: o crescimento, a fotossíntese, a absorção e a distribuição de nutrientes, podendo levar a planta à morte celular (GALLEGO et al., 2012).

Ao manter contato com plantas, humanos e animais, o Cd causa bioacumulação, que se caracteriza pelo acúmulo de metais pesados em níveis milhares de vezes maiores que os presentes no ambiente, podendo acumular-se no corpo por um período superior a 10 anos. Assim, quando as concentrações deste metal atingem níveis acima de dez vezes do detectado naturalmente, efeitos deletérios sobre o DNA e estruturas celulares podem ser evidenciados, com danos irreversíveis (PIZZAIA, 2013). Outro meio de o Cd entrar na cadeia alimentar é através da água potável que pode ser contaminada devido à presença do metal, como impureza no zinco de tubulações galvanizadas, soldas e acessórios metálicos; pela adsorção à argila ou material orgânico ou mesmo pela bioacumulação em plantas aquáticas, invertebrados, peixes e mamíferos.

Diante dos fatores estressores que o Cd causa às plantas, essas desenvolveram mecanismos de tolerância e adaptação, os quais envolvem processos e reações de natureza nutricional, fisiológica e bioquímica que possibilitam seu uso na recuperação de áreas com excesso de metais pesados (SOUZA et al., 2011).

“A fitorremediação refere-se basicamente ao uso de plantas e microrganismos associados ao solo, para reduzir as concentrações ou efeitos tóxicos de contaminantes nos ambientes” (GREIPSSON, 2011). Essa é uma tecnologia relativamente recente, com pesquisas realizadas de 1990 em diante. O conceito de fitorremediação (como fitoextração) foi sugerido por Chaney (1983) (ALI et al., 2013).

2. Uso da fitorremediação na limpeza de solos contaminados por cádmio

Os danos causados por metais sobre o ecossistema agrícola podem, possivelmente, ser resolvidos através da utilização de plantas tolerantes que “limpam” as áreas contaminadas com metais (SINGH; PRASAD, 2011; AZEVEDO et al., 2012), radionuclí-

deos, bem como os poluentes orgânicos (tais como hidrocarbonetos aromáticos polinucleares, bifenilos policlorados e pesticidas) (ALI et al., 2013). Porém, para que o uso da fitorremediação seja eficiente, é necessário investigar e entender como as plantas funcionam, bem como quais rotas metabólicas e genes estão envolvidos em tal processo (VERBRUGGEN et al., 2009). Os mecanismos de desintoxicação, acumulação e tolerância a Cd, tornaram-se a base para o uso de plantas na remediação de solos contaminados por este metal (MENDOZA-CÓZALT et al., 2011).

As plantas manuseiam os contaminantes sem afetar o solo, conservando assim sua utilidade e fertilidade (MENCH et al., 2009). Até agora, diferentes abordagens físicas, químicas e biológicas foram empregadas para este propósito. Os métodos convencionais de remediação incluem a vitrificação *in situ*, a incineração do solo, a escavação e o aterro sanitário, a lavagem do solo, o nivelamento do solo, a solidificação e a estabilização de sistemas eletrocinéticos (SHEORAN et al., 2011; WUANA; OKIEIMEN, 2011). Esses métodos sofrem limitações como: alto custo, mão de obra intensiva, mudanças irreversíveis em propriedades do solo e perturbação da microflora nativa do solo, bem como podem criar problemas de poluição secundária (ALKORTA et al., 2004; ALI et al., 2013).

Tabela 1: Diferentes técnicas de fitorremediação (ALI et al., 2013)

Técnica	Descrição
Fitoextração	Acúmulo de poluentes na biomassa disponível, por exemplo, nos brotos
Fitofiltração	Sequestro de poluentes de águas contaminadas por plantas
Fitoestabilização	Limitação da mobilidade e da biodisponibilidade de poluentes no solo por raízes de plantas
Fitovolatilização	Conversão de poluentes para a forma volátil e sua posterior liberação para a atmosfera
Fitodegradação	Degradação de xenobióticos orgânicos por enzimas contidas nos tecidos das plantas
Rizodegradação	Degradação de xenobióticos orgânicos na rizosfera por microrganismos rizosféricos
Fitodessalinização	Remoção do excesso de sal de solos salinos por halófitos

3. Bactérias como facilitadoras da fitoextração de metais

Nos últimos anos, os pesquisadores têm-se esforçado para utilizar plantas para facilitar a remoção de contaminantes orgânicos e inorgânicos do meio ambiente, especialmente a partir do solo. Essas abordagens de fitorremediação já percorreram longo caminho em curto espaço de tempo. No entanto, a maior parte deste trabalho foi realizada em condições laboratoriais controladas e não no campo. Como um complemento às várias estratégias de fitorremediação e como parte de um esforço para tornar esta tecnologia mais eficaz, um número de cientistas começou a explorar a possibilidade de utilizar várias bactérias do solo, juntamente com as plantas. Estas bactérias incluem bactérias biodegradativas, bactérias promotoras do crescimento de plantas e bactérias que facilitam a fitorremediação por outros meios (GLICK, 2010).

A fitoextração de metais que precisam ser removidos do ambiente é uma técnica complexa, e suas principais limitações são (i) a biodisponibilidade do metal-alvo e (ii) a capacidade das plantas para acumular metais na parte aérea (RASKIN; ENSLEY, 2000; GLICK, 2010).

Embora a utilização de bactérias do solo (geralmente de bactérias promotoras do crescimento de plantas) como auxiliares na fitorremediação de metais possa facilitar significativamente o crescimento de plantas na presença de níveis elevados (inibitórios) de metais, as bactérias normalmente pouco ou nada fazem para aumentar a biodisponibilidade do metal (GLICK, 2010).

Após a demonstração inicial de que uma bactéria promotora do crescimento em plantas e resistente ao níquel diminuiu a toxicidade do metal às plantas de canola (BURD et al., 1998), houve aumento no número de pesquisas demonstrando que as bactérias promotoras do crescimento em plantas auxiliam na fitoextração de metal. Nos estudos relatados, a adição das bactérias facilitou o crescimento das plantas e, na presença de níveis baixos dos metais, aumentou a absorção de metal pela planta (GLICK, 2010).

As bactérias utilizadas para estes estudos são primeiramente selecionadas quanto à resistência ao metal tóxico alvo e, em seguida, elas são muitas vezes selecionadas ou testadas para a presença da enzima *ACC-deaminase* e/ou para a capacidade de sintetizar

AIA e sideróforos. Estes estudos são realizados com a expectativa de que as bactérias resistentes ao metal selecionadas serão capazes de crescer e promover o crescimento de plantas na presença de níveis elevados de metais tóxicos. Embora poucos estudos tenham identificado o(s) mecanismo(s) exato(s) da promoção do crescimento de plantas na presença do metal, a totalidade dos dados está relacionada com a presença/produção de AIA, sideróforos e *ACC-deaminase*. Assim, todas ou algumas dessas atividades já são suficientes para elaborar o mecanismo (GLICK, 2010).

4. Referências bibliográficas

ALL, H.; KHAN, E.; SAJAD, M. A. Phytoremediation of heavy metals-concepts and applications. **Chemosphere**, v. 91, n. 7, p. 869–81, 2013. Disponível em: <<http://doi:10.1016/j.chemosphere.2013.01.075>>.

ALKORTA, I.; HERNÁNDEZ-ALLICA, J.; BECERRIL, J.; AMEZAGA, I.; ALBIZU, I., GARBISU, C. Recent findings on the phytoremediation of soils contaminated with environmentally toxic heavy metals and metalloids such as zinc, cadmium, lead, and arsenic. **Review in Environmental Science and Biotechnology**, v.3, p. 71–90, 2004.

AZEVEDO, R. A.; GRATÃO, P. L.; MONTEIRO, C. C.; CARVALHO, R. F. What is new in the research on cadmium-induced stress in plants? **Food and Energy Security**, v.1, n. 2, p. 133–140, 2012. Disponível em <doi:10.1002/fes3.10>

BURD, G.I.; DIXON, D.G.; GLICK, B.R. A plant growth promoting bacterium that decreases nickel toxicity in plant seedlings. **Applied and Environmental Microbiology**, v. 64, p. 3.663-3.668, 1998.

CETESB, 2012 - Divisão de Toxicologia, Genotoxicidade e Microbiologia Ambiental. Cádmio e seus compostos. Ficha de informação toxicológica – FIT. Disponível em: <<http://laboratorios.cetesb.sp.gov.br/wp-content/uploads/sites/47/2013/11/cadmio.pdf>>. Acessado em: 17 fev. 2016.

CHANEY, R.L. Plant uptake of inorganic waste constituents. In: Parr, J.F.E.A. (Ed.), Land Treatment of Hazardous Wastes. **Noyes Data Corporation**, Park Ridge, NJ, p. 50–76, 1983.

GALLEGO, S.M.; PENA, L.B.; BARCIA, R.A.; AZPILICUETA, C.E. et al., 2012. Unravelling cadmium toxicity and tolerance in plants: Insight into regulatory mechanisms. **Environmental and Experimental Botany**, v. 83, p.33–46, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.envexpbot.2012.04.006>>.

GLICK, B.R. Using soil bacteria to facilitate phytoremediation. **Biotechnology advances**, v. 28(3), p. 367–374, 2010. Disponível em: <doi:10.1016/j.biotechadv.2010.02.001.>.

GREIPSSON, S. Phytoremediation. **National Educational Knowledge**, v. 2, p. 7, 2011.

JOHRI, N.; JACQUILLET, G.; UNWIN, R. Heavy metal poisoning: the effects of cadmium on the kidney. **Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine**, v. 23, n. 5, p. 783–92, 2010. Disponível em: <http://doi10.1007/s10534-010-9328-y>.

MENCH, M.; SCHWITZGUEBEL, J.P.; SCHROEDER, P.; BERT, V.; GAWRONSKI, S.; GUPTA, S. Assessment of successful experiments and limitations of phytotechnologies: contaminant uptake, detoxification and sequestration, and consequences for food safety. **Environmental Science and Pollution Research**, v. 16, p. 876–900, 2009.

MENDOZA-CÓZATL, D.; LOZA-TAVERA, H.; HERNÁNDEZ-NAVARRO, A.; MORENO-SÁNCHEZ, R. Sulfur assimilation and glutathione metabolism under cadmium stress in yeast, protists and plants. **FEMS Microbiology Reviews**, v. 29, p. 653–671, 2005.

PIZZAIA, D. Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz” **Genotoxicidade do cádmio em tomateiro (*Solanum lycopersicum* L.)**, Tese, 2013.

PRAPAGDEE, B.; CHANPRASERT, M.; MONGKOLSUK, S. 2013. Bioaugmentation with Cadmium-Resistant Plant Growth-Promoting Rhizobacteria to Assist Cadmium Phytoextraction by *Helianthus Annuus*. **Chemosphere**, v. 92, p. 659-666. Disponível em: <http://dx.doi.org/10.1016/j.chemosphere.2013.01.082>.

RASKIN, I.; ENSLEY, B.D. Phytoremediation of Toxic Metals: Using Plants to Clean Up the Environment. **Wiley-Interscience**, 2000.

SHEORAN, V.; SHEORAN, A.; POONIA, P. Role of hyperaccumulators in phytoextraction of metals from contaminated mining sites: a review. **Criticals Review in Environmental Science and Technology**. v. 41, p. 168–214, 2011.

SINGH, A.; PRASSAD, S.M. Reduction of heavy metal load in food chain: technology assessment. **Rev. Environ. Sci. Biotechnol.** v. 10, p. 199-214, 2011.

SOUZA, E.P. de; FRANÇA, I. de, FERREIRA, L.E. Mecanismos de tolerância a estresses por metais pesados em plantas. **Revista Brasileira de Agrociência**, p.167–173, 2011.

VERBRUGGEN, N.; HERMANS, C.; SCHAT, H. Mechanisms to cope with arsenic or cadmium excess in plants. **Current Opinion in Plant Biology**, v.12, p. 364–372, 2009.

WUANA, R.A.; OKIEIMEN, F.E. Heavy metals in contaminated soils: a review of sources, chemistry, risks and best available strategies for remediation. **ISRN Ecology**, v. 1, p. 20, 2011.

Atualidades em Melhoramento Genético Florestal

Renan Furlan Gonsaga
Marcio José de Araújo
Tamara Mariel Valdovinos
João Vitor Oliveira de Souza
Rinaldo Cesar de Paula

1. Introdução

Os programas de melhoramento florestal proporcionaram grandes avanços na silvicultura brasileira nos últimos anos, o que pode ser confirmado principalmente pelo aumento na produtividade e na qualidade de nossas florestas. Hoje, o Brasil é referência mundial em Silvicultura de eucalipto, e parte deste reconhecimento deve-se aos trabalhos de melhoristas florestais.

As árvores mais plantadas no mundo são do gênero *Eucalyptus*, devido às características de rápido crescimento, boa produtividade, ampla adaptação e versatilidade para aplicação em diferentes usos. Espécies deste gênero também são as mais utilizadas para plantios de florestas comerciais no Brasil (Indústria Brasileira de Árvores, 2015). Estas florestas são formadas, em sua maioria, a partir de mudas multiplicadas vegetativamente, cujas técnicas vêm sendo aperfeiçoadas constantemente. Os primeiros passos objetivando a clonagem do eucalipto deram-se no século passado, quando um grupo de pesquisadores australianos e franceses multiplicaram, por enraizamento de estacas obtidas de mudas seminais, várias espécies do gênero *Eucalyptus* (ALFENAS et al., 2009).

Algumas técnicas alternativas surgiram para dar suporte para seleção de genótipos-elite em testes clonais, como, por exemplo, a proteômica. A proteômica consiste no estudo sistemático do conjunto ou complemento de proteínas expressas no genoma de um organismo, chamado de proteoma (PHIZICKY et al., 2003;

KLUG et al., 2010). Além de gerar dados inéditos, a proteômica tem contribuído para a validação, complementação e até a correção de informações obtidas por abordagens anteriores acerca da biologia das plantas. Na prática, as análises proteômicas recentemente se tornaram mais dinâmicas com o advento de métodos de análise rápida (SALVATO; CARVALHO, 2010), o que pode vir a favorecer sua aplicação nas pesquisas da área florestal. A transgenia é outro procedimento biotecnológico discutido atualmente nas pesquisas com o eucalipto, por exemplo, e muito se espera de seus resultados como uma ferramenta complementar para o melhoramento genético desse grupo de espécies (FONSECA et al., 2010). As espécies geneticamente modificadas, futuramente, poderão mudar de maneira bastante significativa a forma como as florestas são manejadas, além de permitir a colheita de uma matéria-prima aprimorada para a indústria.

2. Clonagem de espécies florestais

O principal gênero de espécies florestais plantadas no Brasil atualmente é o *Eucalyptus*. A área total de florestas plantadas no País é de 7,74 milhões de hectares, e deste montante, 5,56 milhões de hectares são de plantios de eucalipto (Indústria Brasileira de Árvores, 2015). A produção de mudas clonais de eucalipto encontra-se bem estabelecida e em estágio avançado, sendo praticada de forma intensiva (XAVIER et al., 2009).

A definição de “clone” fica bem estabelecida como sendo o grupo de células ou indivíduos geneticamente idênticos, derivados por meio da reprodução assexuada de um único antecessor comum. Portanto, pode-se referir a um clone como um ser oriundo de mitoses sucessivas ao longo de seu desenvolvimento. Assim, por meio da propagação vegetativa, a totipotencialidade das espécies de *Eucalyptus* é direcionada para obtenção de genótipos com características desejáveis, acrescentando-se a obtenção de ganhos genéticos rápidos e eficientes. A clonagem do eucalipto é realizada a partir do material juvenil coletado, pela técnica da estaquia. A estaquia é amplamente adotada e teve grande contribuição na silvicultura clonal intensiva. De princípios bem conhecidos, a estaquia tradicional foi aperfeiçoada, gerando a miniestaquia e a microestaquia, que surgiram como melhores alternativas para contornar dificuldades no enraizamento das

estacas, na formação das mudas e no desenvolvimento da árvore de clones e espécies florestais. Dentre as técnicas de clonagem, a miniestaquia tem sido a mais utilizada (XAVIER et al., 2009).

O processo de obtenção das miniestacas consiste na indução de brotação em minicepas (mudas) do material genético a ser propagado, com posterior coleta de brotos, cujo comprimento varia de 2 a 5 cm. Após a coleta, mantêm-se de duas a quatro folhas jovens, efetuando um corte na base dos propágulos e submetendo-os a recipientes contendo água, de modo que possam manter a turgescência das células. Em seguida, as miniestacas são acondicionadas em recipientes, com substrato adequado, para estimular o enraizamento em casa de vegetação. Nesta etapa, espécies de eucalipto permanecem por um período de 15 dias; logo após, são enviadas para casa de sombra por um período de aclimação durante igual período. Antes de serem direcionadas ao plantio no campo, as mudas passam por uma etapa de rustificação, que consiste na exposição às condições próximas àquelas submetidas no campo, com restrições hídricas e nutricionais. A etapa de rustificação conduz a redução de perda de qualidade e mortalidade de mudas no campo (FERRIANI et al., 2010).

A clonagem de *Pinus* spp. é limitada por uma combinação de fatores, tais como: i) baixo índice de enraizamento e baixa qualidade do sistema radical pelo método da estaquia; ii) o entendimento da relação entre as idades, fatores ambientais e épocas para a coleta das miniestacas, assim como a interação desses fatores no método da miniestaquia; iii) dificuldade para a reversão da juvenildade, de modo que possa atingir o potencial máximo de enraizamento (ALCANTARA et al., 2007). Porém, alternativas, como, por exemplo, a embriogênese somática pode ser utilizada para propagação de clones de *Pinus* com bom potencial produtivo para características de crescimento (DIAS, 2013).

Os programas de melhoramento florestal envolvem em sua fase final os testes clonais, aos quais são submetidos testes de competição em escalas-piloto ou até mesmo comerciais. Nesses testes, é recomendado realizar as avaliações genéticas a partir de dois anos de idade, de modo que se possa determinar o valor real do indivíduo no ambiente em que ele se encontra. Desta maneira, pode-se obter a otimização de ganhos genéticos com a seleção precoce, reduzindo tempo e esforços, gerando economia de recursos na condução do programa (ARAÚJO et al., 2015). A seleção é a etapa decisória de

sucesso ou fracasso de um programa de melhoramento genético. Atualmente, para a seleção de clones, os modelos lineares mistos vêm destacando-se dentre os programas de melhoramento genético florestal, permitindo ganhos genéticos substanciais (FONSECA et al., 2010), pois a seleção é realizada sobre o valor genotípico e não sobre a média fenotípica, como geralmente ocorre.

Na metodologia de modelos mistos, a estimação dos componentes de variância é realizada pelo método da máxima verossimilhança restrita (REML), e a predição dos valores genéticos, por meio do procedimento BLUP (melhor predição linear não viesado). O BLUP é o procedimento ótimo de seleção, que, por sua vez, dependendo da natureza do caráter, da estrutura genética da população de melhoramento e das características experimentais, permite estimar os efeitos genéticos aditivos (a), de dominância (d) e genotípicos (g) (Para mais detalhes, ver Resende, 2002).

A análise de dados via modelos mistos, além de corrigir simultaneamente os dados para os efeitos ambientais, estimar os parâmetros genéticos e prever os valores genéticos, possui outras vantagens, tais como: i) permite, de maneira adequada, a avaliação do comportamento e do desempenho dos indivíduos ao longo do tempo e dos espaços; ii) produz resultados não viesados e maximiza a acurácia seletiva; iii) maximiza os ganhos genéticos, melhorando a eficiência dos programas de melhoramento mesmo no caso de dados desbalanceados, gerando estimativas mais precisas; iv) possibilita trabalhar com conjunto de dados com características complexas, como, por exemplo, dados com medidas repetidas, vários locais, diferentes idades e diversos tipos de delineamentos experimentais, além do cômputo da interação genótipo x ambiente (PIRES et al., 2011).

De maneira geral, o desenvolvimento de um clone de eucalipto envolve várias técnicas e estratégias de melhoramento genético, tais como: i) a polinização de árvores matrizes, com características desejáveis; ii) produção e plantio de mudas oriundas desses cruzamentos em testes de progênies; iii) avaliação das progênies ao longo do tempo e posterior seleção entre e dentro das que apresentarem os melhores desempenhos; iv) clonagem dos indivíduos superiores e avaliação em testes clonais.

3. Proteômica no melhoramento florestal

A proteômica atualmente é considerada uma das ferramentas de grande importância para estudos da regulação da atividade gênica, ou seja, entender o comportamento dos seres vivos, como e quando eles expressam uma proteína, fundamental para seu ciclo de vida (FIALHO et al., 2008). O termo proteoma surgiu somente em 1994 (WASINGER et al., 1995), e mesmo dentro da genética, ainda há muito o que ser explorado, não só na área florestal, mas também em toda a área vegetal, animal e humana.

No ano de 2014, um importante acontecimento para a área florestal aconteceu: o sequenciamento do genoma do *Eucalyptus grandis*. Esse feito está contribuindo para que os estudos proteômicos sobre essa cultura avancem rapidamente (MYBURG et al., 2014). A proteômica auxilia na identificação de genótipos superiores, atuando efetivamente nas etapas inerentes ao melhoramento florestal, reduzindo o tempo gasto na seleção genética com os métodos tradicionais, fornece informações em nível molecular da proteína e toda sua variabilidade que é expressa pelo genoma (PENNINGTON; DUNN, 2001). Por meio do estudo das proteínas, busca-se compreender o comportamento das plantas quando submetidas a qualquer tipo de estresse. Além disso, compreender e identificar como e quais proteínas estão envolvidas nas rotas metabólicas pode adequar às etapas envolvidas num programa de melhoramento, aumentar a variabilidade genética e, conseqüentemente, os ganhos com a seleção (ZHANG et al., 2004; SOARES; MACHADO, 2007).

A área de análise proteômica tem-se tornado, também, uma ferramenta importante para auxiliar na relação entre o genótipo e a influência que este sofre pelo ambiente. O estudo do proteoma pode ser dividido em duas áreas: a proteômica funcional e a proteômica estrutural. A primeira é baseada na atividade biológica das proteínas, já a proteômica estrutural se baseia na variação das expressões em diferentes ambientes (ANDERSON et al., 2000). Com o surgimento da espectrometria de massas, o desenvolvimento da proteômica atual avançou muito, pois tornou-se viável a realização de experimentos em grande escala. São utilizadas plataformas de análise rápida, direta e sensível, podendo-se citar os géis 2D e o shotgun, que tornaram as análises ainda

mais dinâmicas (SALVATO; CARVALHO, 2010).

Uma das áreas de estudo da proteômica, considerada importante para entender o comportamento dos seres vivos sobre estresses e um dos principais temas discutidos atualmente no mundo, é a escassez dos recursos hídricos. Por isso, a busca por genótipos mais tolerantes ao estresse hídrico e ao térmico tem sido cada vez mais intensa (PIMENTEL, 2005). Neste sentido, a proteômica pode captar o comportamento das plantas submetidas ao estresse hídrico, térmico, etc. e, assim, traçar um panorama de como os genes são expressos nesta situação. São essas questões que trazem a proteômica para próximo do melhoramento florestal.

O estudo da proteômica, relacionado ao melhoramento florestal, surge para incrementar e complementar todo o conhecimento adquirido na área através dos anos, auxiliando numa maior compreensão no complexo sistema que é a biologia das plantas, e, conseqüentemente, num futuro próximo, auxiliar tanto o melhoramento, como as práticas agronômicas. Portanto, a proteômica abre novos caminhos no melhoramento florestal para os estudos com estresses fisiológicos, variabilidade genética e desenvolvimento de plantas.

4. Transgênese florestal

A transformação genética, também chamada transgênese ou transgenia, é a introdução controlada de um gene ou fragmento de DNA no genoma de uma célula receptora, de maneira que este seja expresso de forma estável (DIOUF, 2003). Com a posterior regeneração *in vitro* dessa célula inicialmente transformada, é possível gerar uma planta transgênica contendo nova informação genética. Técnicas de transgenia aliadas ao melhoramento genético clássico permitem aumentar os ganhos genéticos, tendo como vantagem a redução do tempo necessário para introdução de novas características em linhagens já estabelecidas por longos programas de cruzamento e de seleção (BORÉM, 2007). A transformação genética é uma ferramenta de grande potencial para espécies florestais, já que facilita a obtenção de genótipos superiores, ao incorporar genes de interesse em progênies e clones comerciais sem modificar a estrutura genética global da planta (STUDART-GUIMARÃES

et al., 2003). Assim, permite superar problemas inerentes ao melhoramento genético das árvores, tais como o longo tempo para a obtenção de novas gerações e a enorme diversidade existente entre e dentro das espécies (DICIERO; AMARAL, 2002).

Os métodos de transgênese mais efetivos para a geração de AGM (árvores geneticamente modificadas) são a transformação mediada por agrobactérias e a aceleração de partículas ou biobalística (BABU et al., 2003). No entanto, todo sistema eficiente de transformação vegetal necessita das seguintes etapas: identificação, isolamento e introdução do DNA exógeno na célula; seleção e crescimento das células transformadas; além do estabelecimento de um sistema simples e eficiente de regeneração das células transformadas (BRASILEIRO; DUSI, 1999). A obtenção de AGM tem como principal obstáculo a menor adaptabilidade à cultura de tecidos *in vitro* apresentada pelas plantas lenhosas. Embora várias AGMs de importância econômica tenham sido desenvolvidas por empresas privadas, seus protocolos de transformação são mantidos em segredo industrial (ROUT et al., 1999; BORÉM, 2007; SARTORETTO et al., 2008). As principais características de interesse que já foram aplicadas por meio da transgenia e conduzidas para avaliações em teste no campo são a produção de biomassa, a manipulação da composição de lignina, a resistência a insetos, o controle do florescimento, a tolerância a herbicidas e a estresses abióticos (HALPIN; BOERJAN, 2003; STUDART-GUIMARÃES et al., 2003). No Brasil, o primeiro eucalipto transgênico desenvolvido pela empresa FuturaGene contém um gene isolado da planta *Arabidopsisthaliana*, que codifica a proteína Cell1, que promove o alongamento da parede celular, aumentando o crescimento da planta e, conseqüentemente, gerando maior produtividade. Quando comparado com o eucalipto convencional, o eucalipto transgênico produz 20% mais celulose; além disso, o tempo de corte pode ser reduzido de 7 anos para 5,5 anos (DA SILVEIRA, 2013). Durante cerca de 5 anos, os plantios do híbrido transgênico (*E.grandis* x *E.urophylla*) foram monitorados para que se previníssem problemas ambientais e que se conhecesse melhor o comportamento deste transgênico no ambiente. Finalmente, a partir da decisão favorável da Comissão Técnica Nacional de Biossegurança (CTNBio), órgão responsável no Brasil pela avaliação de produtos transgênicos, em abril de 2015, o Brasil passou a ser o primeiro país a

liberar comercialmente um eucalipto transgênico (PASQUALI, 2015). Segundo um estudo do Centro para Avaliação do Risco Ambiental de Culturas Geneticamente Modificadas (Cera), calcula-se que, desde o final dos anos de 1980, já tenham sido realizados no mundo mais de 800 experimentos de campo com árvores geneticamente modificadas de cerca de 40 espécies (DA SILVEIRA, 2013). Devido ao enorme retorno econômico, além das possibilidades de adaptação dos plantios em áreas onde antes não era possível, a tendência mundial é que um número crescente de espécies florestais transgênicas seja liberado no mercado nos próximos anos (STUDART-GUIMARÃES et al., 2003; VALDETARO et al., 2011).

5. Referências bibliográficas

ALCANTARA, G. B. D.; RIBAS, L. L. F.; HIGA, A. R.; CHISTINA, K.; RIBAS, Z.; KOEHLER, H. S. Efeito da idade da muda e da estação do ano no enraizamento de miniestacas de *Pinus taeda* L. **Revista Árvore**, Viscosa, v. 31, n. 3, p. 399-404, 2007.

ALFENAS, A. C.; ZAUZA, E. A. V.; MAFIA, R. G.; ASSIS, T. F. **Clonagem e doenças do eucalipto**. 2. ed. Viçosa-MG: Ed. UFV, 2009, 500p.

ANDERSON, N. L.; MATHESON, A. D.; STEINER, S. Proteomics: applications in basic and applied biology. **Current Opinion in Biotechnology**, London, v. 11, n. 4, p. 408-412, 2000.

ARAÚJO, M. J.; DIAS, D. D. C.; SCARPINATI, E. A.; PAULA, R. C. Número de repetições, de plantas por parcela e de avaliações para testes clonais de eucalipto. **Pesquisa Agropecuária Brasileira**, Brasília, v. 50, n. 10, p. 923-931, 2015.

BABU, R. M.; SAJEENA, A.; SEETHAMAN, K.; REDDY, M. S. Advances in genetically engineered (transgenic) plants in pest management-an over view. **Crop Protection**, Willoughby, Ohio, v.22, p.1.071-1.086, 2003.

BORÉM, A. **Biotecnologia florestal**. Universidade Federal de Viçosa, 2007, 387 p.

BRASILEIRO, A. C. M.; DUSI, D. M. A. Transformação genética de plantas. In: TORRES, A. C.; CALDAS, L. S.; BUSO, J. A. (Eds.). **Cultura de tecidos e transformação genética de plantas**. Brasília: Embrapa-SPI/Embrapa-CNPq, 1999. p. 679-735.

DA SILVEIRA, E. Silvicultura: mais celulose por centímetro quadrado. **Revista Pesquisa FAPESP**, Ed. 204, 2013. Disponível Em: <<http://revista->

pesquisa.fapesp.br/2013/02/11/mais-celulose-por-centimetro-quadrado/>. Acesso em: jan. 2016.

DI CIERO, L.; AMARAL, W. Árvores geneticamente modificadas na silvicultura intensiva. **Biotecnologia, Ciência e Desenvolvimento**, Brasília, n.29, p.92-98, 2002.

DIAS, P. C. Avaliação genética de clones de *Pinus taeda* propagados via embriogênese somática. Viçosa: Universidade Federal de Viçosa, 2013. 89 f. **Tese** (Doutorado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2013.

DIOUF, D. Genetic transformation of forest trees. **African Journal of Biotechnology**, Nairobi, Kenya, v.2, n.10, p.328-333, 2003.

FERRIANI, A. P.; ZUFFELLATO-RIBAS, K. C.; WENDLING, I. Miniestaquia aplicada a espécies florestais. **Revista Agro@mbiente On-line**, Boa Vista, Roraima, v. 4, n. 2, p. 102-109, 2010.

FIALHO, E.; MORENO, F. S.; THOMAS, P. Nutrição pós-genoma: fundamentos e aplicações de ferramentas ômicas. **Revista Nutrição**, Campinas, v.21, n.6, p. 755-766, 2008.

FONSECA, S. M.; RESENDE, M. D. V.; ALFENAS, A. C.; GUIMARÃES, L. M. S.; ASSIS, T. F.; GRATTAPAGLIA, D. **Manual prático de melhoramento genético do eucalipto**. Viçosa-MG: Ed. UFV, DFP. 2010. 200p.

HALPIN, C.; BOERJAN, W. Stacking transgenes in forest trees. **Trends in Plant Science**, Kidlington, Oxford, v.8, n.8, p.363-365, 2003.

KLUG, W. S.; CUMMINGS, M. R.; SPENCER, C. A.; PALLADINO, M. A. **Conceitos de genética**. – 9. ed. – Porto Alegre : Artmed, 2010, 896 p.

MYBURG, A. A.; GRATTAPAGLIA, D.; TUSKAN, G. A.; HELLSTEN, U.; HAYES, R. D.; GRIMWOOD, J.; GOODSTEIN, D. M. The genome of *Eucalyptus grandis*. **Nature**. London, v. 510, n. 7505, p. 356-362, 2014.

PASQUALI, G. A liberação comercial de eucalipto transgênico. **Revista Opiniões**. 2015. Disponível em: <<http://florestal.revistaopinioes.com.br/revista/detalhes/3-liberacao-comercial-de-eucalipto-transgenico/>>. Acesso em: jan. 2016.

PENNINGTON, S. R.; DUNN. M. J. **Proteomics**: from protein sequence to function. New York: Springer-Verlag e BIOS Scientific Publishers, 2001, 313p.

PHIZICKY, E.; BASTIAENS, P. I.; ZHU, H.; SNYDER, M.; FIELDS, S. Protein analysis on a proteomic scale. **Nature**, London, v. 422, n. 6928, p. 208-215, 2003.

PIMENTEL, C. Respostas fisiológicas à falta d'água: limitação difusiva ou

metabólica In: NOGUEIRA, R.J.M.C.; ARAÚJO, E. de L.; WILLADINO, L.G.; CAVALCANTE, U.M.T.; (Ed.). **Estresses ambientais: danos e benefícios em plantas**. Recife: UFRPE, Impr. Universitária, 2005. p.13-21.

PIRES, I. E.; RESENDE, M. D. V. D.; SILVA, R. L. D.; JR., M. F. R. D. R. **Genética florestal**. Viçosa-MG: Editora Arka, 2011. 318 p.

RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Embrapa Informação Tecnológica, Colombo: Embrapa Florestas, 2002.

ROUT, G. R.; SAMANTARAY, S.; MOTTLEY, J.; DAS, P. Biotechnology of the rose: a review of recent progress. **Scientia Horticulturae**, Amsterdam, v. 81, p. 201-228, 1999.

SALVATO, F.; CARVALHO, M. C. G. Métodos e estratégias em proteômica e suas aplicações na área vegetal. **Ciência Rural**, Santa Maria, Brasil, v. 40, n. 3, p. 727-734, 2010.

SARTORETTO, L. M.; SALDANHA, C. W.; CORDER, M. P. M. Transformação genética: estratégias e aplicações para o melhoramento genético de espécies florestais. **Ciência Rural**, Santa Maria, Brasil, v.38, n.3, p.861-871, 2008.

SOARES, A.M. S.; MACHADO, O.L.T. Defesa de plantas: Sinalização química e espécies reativas de oxigênio. **Revista Trópica – Ciências Agrárias e Biológicas**, Chapadinha, Maranhão, vol. 1, n. 1, p. 9-19, 2007.

STUDART-GUIMARÃES, C.; LACORTE, C.; BRASILEIRO, A. C. M. Transformação genética em espécies florestais. **Ciência Florestal**, Santa Maria, Brasil, v. 13, n. 1, p. 167-178, 2003.

VALDETARO, E. B.; CARVALHO, A. P. V.; GAMBA, O. S.; DEMOLINARI, M. D. S. M.; SILVA, E. Árvores geneticamente modificadas: técnicas, aplicações, riscos e os potenciais impactos associados a sua utilização. **Pesquisa Florestal Brasileira**, Colombo, Brasil, v. 31, n. 65, p. 51, 2011.

WASINGER, V.C.; CORDWELL, S.J.; CERPA-POLJAK, A.; YAN, J.X.; GOOLEY, A.A.; WILKINS, M.R.; DUNCAN, M.W.; HARRIS, R.; WILLIAMS, K.W.; HUM-PHERYSMITH, I. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. **Electrophoresis**, Weinheim, Germany, v. 16, p. 1.090-1.094, 1995.

XAVIER, A.; WENDLING, I.; SILVA, R. L. **Silvicultura clonal: princípios e técnicas**. Viçosa-MG: Ed. UFV, 2009, 272 p.

ZHANG, H. Y.; XIE, X. Z.; XU, Y. Z.; WU, N. H. Isolation and functional assessment of a tomato proteinase inhibitor II gene. **Plant Physiology and Biochemistry**, Amsterdam, v. 42, n. 5, p. 437-444, 2004.

Estudo da variabilidade genética aplicado à genômica animal

*Jorge Augusto Petroli Marchesi
Danielly Beraldo dos Santos Silva
Letícia Joaquim Borges
Nedenia Bonvino Stafuzza
Danísio Prado Munari*

1. Introdução

A variabilidade genética, expressa pela frequência dos diferentes alelos em um mesmo gene, além de apresentar fundamental importância para a evolução, é a base para o desenvolvimento de programas de melhoramento genético de animais domésticos. Eventos como mutação, transposição e recombinação geram variabilidade genética em nível de espécie, enquanto migração, deriva genética e seleção alteram as frequências alélicas em populações. No melhoramento genético, utiliza-se essa diversidade genética presente em populações de animais para aumentar a frequência de alelos favoráveis que estejam relacionados com características de interesse econômico, os quais, em sua grande maioria, apresentam natureza quantitativa. Mas afinal, como os alelos estão ligados a características de interesse? O que são marcadores genéticos? O que é seleção genômica? Como marcadores são utilizados em programas de melhoramento genético? Neste capítulo, serão abordados conceitos básicos sobre a genética de populações, quantitativa, molecular e das tecnologias genômicas aplicadas ao melhoramento genético que poderão responder a estas e a outras questões.

2. Conceitos básicos de genética de população

O melhoramento zootécnico tem como objetivo aprimorar animais domésticos, visando ao aumento da produtividade com qualidade, ao menor custo e demanda de tempo (PEREIRA, 2012). Para tanto, devem-se levar em consideração os efeitos ambientais e genéticos sobre as características em estudo. Desta forma, para descrever a constituição genética de um grupo de indivíduos da mesma espécie (população), seria necessário identificar os genótipos e alelos, e suas respectivas proporções (frequências genotípicas e frequências alélicas).

Em uma população, as propriedades genéticas são influenciadas por fatores que atuam no processo de transmissão dos genes de uma geração para outra. Um desses fatores é o tamanho da população, uma vez que a frequência alélica está sujeita a variações causadas pela amostragem entre gerações sucessivas, e quanto menor o número de pais, maior será a variação da amostragem (deriva genética). Em uma população grande, os gametas são amostras mais representativas da composição genética dos pais do que em uma população pequena (NEI, 1975).

Outro fator são as diferenças na fertilidade e na viabilidade. Durante o processo de transmissão gênica, os diferentes genótipos dos pais podem apresentar diferença na fertilidade. Assim, as frequências alélicas podem ser modificadas, contribuindo para a desigualdade do conjunto de gametas formadores da próxima geração. Além disso, os genótipos de zigotos recém-formados podem apresentar diferenças na taxa de sobrevivência. A probabilidade de sobrevivência e de reprodução de determinado fenótipo (genótipo) também pode ter influência do ambiente, responsável pelo valor adaptativo do indivíduo, ou seja, indivíduos mais adaptados a determinado ambiente têm maior chance de sobreviver e de deixar descendentes. Esse processo é denominado seleção natural (FALCONER; MACKAY, 1996).

As frequências alélicas de um gene, verificadas na progênie, também podem ser alteradas pela seleção artificial, que é a principal prática utilizada em programas de melhoramento genético animal, a qual tem por objetivo obter ganhos de desempenho dentro de populações. Nesse caso, somente os indivíduos com melhores características de interesse são utilizados como repro-

dutores (FALCONER; MACKAY, 1996).

A migração e a mutação também são fatores com forte influência na frequência alélica de uma população. O efeito da migração em uma população é dependente da taxa de imigração e da diferença na frequência alélica entre imigrantes e nativos. O efeito da mutação depende do tipo de evento mutacional: se raro (que pode ser considerado único), tem pouca importância na população por não produzir troca permanente; caso a mutação seja recorrente em uma grande população, a frequência de um alelo mutante nunca é tão baixa que possa ser “perdido” ao longo das gerações, então deve-se descobrir qual o efeito que essa mutação causa na frequência alélica da população (FALCONER; MACKAY, 1996).

Em uma população grande, sob acasalamento ao acaso (cada indivíduo tem igual possibilidade de acasalamento com qualquer outro indivíduo na população), na ausência de migração, mutação e seleção, as frequências alélicas e genotípicas permanecem constantes na população a cada geração (WEIR, 1990). Estas propriedades constituem a Lei de Hardy-Weinberg (1908). Pesquisadores utilizam esses parâmetros para compreender a diversidade e a estrutura genética de populações, as quais tradicionalmente têm sido estudadas por meio da abordagem frequentista de Wright (1921, 1931 e 1951) ou seus derivados. Atualmente, é feito um levantamento da diversidade genética em várias populações de determinada espécie por meio do uso de marcadores moleculares. Deste modo, é possível verificar como a variabilidade está distribuída entre as populações, permitindo fazer inferências quanto aos padrões de cruzamentos, dispersão, fluxo gênico e histórico biogeográfico da espécie (SCHNABEL et al., 1998) e assim contribuir para o avanço do conhecimento em genética de populações, tanto para planos de melhoramento como para projetos de conservação.

3. Conceitos básicos em genética quantitativa

Na genética quantitativa, um dos objetivos é estimar os valores genéticos dos indivíduos, para posterior seleção dos animais geneticamente superiores. Para isso, são utilizados registros fenotípicos em conjunto com dados provenientes do pedigree,

permitindo assim estimar o valor genético do animal para determinada característica. No entanto, esse valor calculado é apenas uma estimativa do que poderia ser o verdadeiro valor genético do animal, o qual depende, diretamente, das características avaliadas, que podem ser classificadas como características qualitativas ou quantitativas.

A maioria das características economicamente importantes na produção animal, como as características de peso, conversão alimentar e rendimento de carcaça, é controlada por vários pares de genes, cada um com efeito relativamente pequeno, apresentando variações contínuas dentro das populações. Sendo assim, cada indivíduo da população pode possuir diferentes alelos para cada um dos genes envolvidos na determinação do fenótipo de uma determinada característica. Essas características são denominadas de características poligênicas, de herança complexa ou quantitativa. Devido a esse tipo de herança, tais características são estudadas na população, por meio de parâmetros como média, variância e desvio-padrão.

Para verificar diferenças quantitativas entre características e determinar a ligação entre as frequências alélicas e as genotípicas mensuradas em indivíduos de determinada população é necessário considerar alguns conceitos de valor, expressos em unidades métricas pelas quais a característica é medida (FALCONER; MACKAY, 1996). Valor fenotípico (P) é o valor observado das características mensuradas no indivíduo e depende do valor genotípico (G), que é o conjunto particular de alelos que um indivíduo possui, somado ao efeito do ambiente (E). Assim, o modelo genético básico pode ser descrito como $P = \mu + G + E$, sendo μ a média populacional (BOURDON, 2000). O valor genotípico é composto pela soma dos efeitos aditivos (A), desvios da dominância (D) e efeitos de epistasia (I). Para um locus, a diferença entre o valor genotípico (G) e o valor genético (A) de um genótipo é o desvio de dominância (D). Desvio de dominância resulta da propriedade de dominância entre alelos em um locus. Na ausência de dominância, G e A apresentam o mesmo valor (FALCONER; MACKAY, 1996).

Pressupondo que não há influência do ambiente em uma população, o valor do efeito ambiental seria igual a zero, enquanto o valor fenotípico médio seria equivalente à média do valor genotípico (ELANDT-JOHNSON, 1971). Como na prática da genética

quantitativa clássica os valores dos genótipos não são mensuráveis, exceto quando se considera um único locus em que genótipos podem ser distinguidos fenotipicamente ou com genótipos de linhagens altamente endogâmicas, valores arbitrários devem ser atribuídos aos mesmos (FALCONER; MACKAY, 1996). Considerando um único locus com dois alelos, B_1 e B_2 , os valores genotípicos para os genótipos B_1B_1 , B_1B_2 e B_2B_2 serão, respectivamente, a , d e $-a$, assumindo que o alelo B_1 aumenta o valor genotípico. Considerando uma escala de valores genotípicos, o ponto médio entre os valores dos dois homozigotos será zero, e o valor d do heterozigoto dependerá do grau de dominância entre os alelos. Desta forma, $d = 0$ para ausência de dominância; $d > 0$ se B_1 for dominante sobre B_2 ; $d < 0$ se B_2 for dominante sobre B_1 ; $d = +a$ ou $d = -a$ para dominância completa; $d > +a$ ou $d < -a$ para sobredominância. A média fenotípica (μ) é importante parâmetro populacional, a qual, assumindo uma população com frequências alélicas p e q para os alelos B_1 e B_2 , respectivamente, considerando um único locus e equilíbrio de Hardy-Weinberg, é dada pela equação $\mu = m + a(p - q) + 2dpq$ (VAN VLECK et al., 1987; FALCONER; MACKAY, 1996).

A média dos valores genotípicos pode ser alterada por seleção, pois depende das frequências alélicas (VAN VLECK et al., 1987). Quando há efeito conjunto de vários loci, a média da população é a soma das contribuições de cada locus individualmente. O estudo da estrutura familiar de uma população depende da medida de valor dos genes, transmitida de pai para a progênie. Assim, o efeito médio é o valor associado aos genes de um indivíduo que são transmitidos para sua progênie, o qual depende dos valores genotípicos e das frequências alélicas. A soma dos efeitos médios dos genes de um indivíduo para todos os pares de alelos em cada locus e para todos os loci representa o valor genético do indivíduo (FALCONER; MACKAY, 1996). Segundo Van Vleck et al. (1987), a importância do valor genético está em sua representação, sendo a medida do desvio do desempenho esperado da progênie do indivíduo em relação à média da população. Como definido por Falconer e Mackay (1996), o valor genético é duas vezes o desvio médio da progênie em relação à média da população, pois a progênie contém apenas uma amostra da metade dos genes dos pais. O desvio da progênie em relação à média da população (metade do valor genético) representa a “habilidade de transmissão” do

pai (VAN VLECK et al., 1987).

O valor genético de um animal depende do número de características desejáveis que influenciam em seu desempenho. Assim, a seleção para determinada característica é importante não somente por sua resposta à seleção, mas também na resposta de outras características dependentes. O tamanho e o sentido das respostas correlacionadas são determinados principalmente pela correlação genética entre as características envolvidas. As correlações genéticas medem a probabilidade de duas características diferentes serem afetadas pelos mesmos genes. O processo pelo qual um mesmo gene pode afetar duas ou mais características, denominado pleiotropia, é a causa da existência das correlações genéticas, embora a ligação entre os genes que influenciam duas características também é causa da mesma. Além disso, a correlação observada entre duas características pode ocorrer devido ao meio comum entre as mesmas. Logo, a correlação observada na população, denominada correlação fenotípica, não expressa a magnitude de seus componentes genético e ambiental (PEREIRA, 2012).

4. Conceitos básicos em genética molecular

Marcador molecular pode ser definido como toda e qualquer característica herdável presente em uma sequência de DNA que diferencia dois ou mais indivíduos. Caracteriza-se por alterações na sequência de nucleotídeos, denominadas polimorfismos. Considerando que cerca de 97% do genoma de mamíferos é composto por sequências de DNA não codificantes, a maioria dos polimorfismos é estável e não acarreta mudanças fenotípicas. Tal fato favorece a transmissão para os descendentes e a fixação dos polimorfismos dentro de uma população.

Entre os marcadores moleculares mais utilizados atualmente, destacam-se os microssatélites e os polimorfismos de nucleotídeo único (SNPs). Os microssatélites são os marcadores mais polimórficos do genoma, os quais são compostos por sequências de 1 a 6 nucleotídeos repetidos em *tandem*. Cada microssatélite constitui um locus genético altamente variável, multialélico e de grande conteúdo informativo. Por essas razões, apresenta gran-

de relevância em diversos testes de DNA, os quais possibilitam a identificação de indivíduos. Esse tipo de marcador pode ser estudado pela técnica da reação em cadeia da polimerase (PCR) e sequenciamento.

Os SNPs são polimorfismos resultantes da alteração de um único nucleotídeo (substituição, inserção ou deleção). Para que essa alteração no genoma possa ser considerada um SNP, a mesma deve ocorrer em uma frequência alélica acima de 1% em determinada população. Esses marcadores são abundantes no genoma, porém a maioria são bialélicos, apresentando pouca informação em termos populacionais quando comparados com os marcadores microssatélites (multialélicos). Beuzen et al. (2000) destacaram as seguintes razões para o crescente interesse da utilização dos SNPs como marcadores para análise genômica: (a) por sua predominância no genoma, são os marcadores que apresentam maior potencial de estarem perto ou dentro de algum locus de característica de interesse econômico; (b) podem estar localizados em regiões codificantes, afetando diretamente a função proteica; (c) são herdados de maneira mais estável que os microssatélites, o que os torna mais apropriados à seleção de marcadores a longo prazo; (d) são mais adequados para análise genética em larga escala, por meio da tecnologia de chips de DNA. Esse tipo de marcador pode ser utilizado para estudos de associação, mapeamento genético, detecção de doenças genéticas, perfil de expressão gênica, análise de pequenos RNAs, estudos de metilação e identificação de indivíduos. Há diversas metodologias para a detecção de SNPs, das quais se destacam o sequenciamento e os painéis de DNA.

5. Seleção genômica e estudos genômicos

A possibilidade de detectar polimorfismos causadores de alterações fenotípicas ou alcançar alta precisão de previsão na seleção genômica tornaram-se cada vez mais fáceis com o progresso no sequenciamento de última geração e no desenvolvimento das tecnologias de genotipagem de SNPs de alta densidade. A incorporação dessas informações na avaliação genética de animais de criação apresenta grande potencial para aumentar a acurácia de seleção, acelerar os ganhos genéticos e reduzir os custos nos pro-

gramas de melhoramento, assim como para identificar os genes associados a determinadas características quantitativas (RESENDE et al., 2008).

Pensando nisso, Meuwissen et al. (2001) propuseram um novo método de seleção, denominado seleção genômica (GS, *Genome Selection*) ou seleção genômica ampla (GWS, *Genome Wide Selection*). Para isto, GWS enfatiza a predição simultânea (sem uso de testes de significância para marcadores individuais) dos efeitos genéticos de milhares de marcadores dispersos no genoma de um organismo, possibilitando capturar os efeitos da maioria dos loci (tanto de pequenos quanto de grandes efeitos) e explicar quase a totalidade da variação genética do caráter quantitativo. Este método pode ser aplicado em todas as populações sob avaliação nos programas de melhoramento genético de espécies alógamas, uma vez que apresenta alta acurácia para a seleção baseada exclusivamente em marcadores, além de não exigir prévio conhecimento das posições dos QTLs (*Quantitative Trait Loci*) (RESENDE et al., 2008).

Ainda com o desenvolvimento de painéis de SNPs de alta densidade, foi possível o surgimento de estudos de associação genômica ampla (GWAS, *Genome Wide Association Studies*) que possibilitaram a descoberta da base genética de diversas características quantitativas complexas. Inicialmente, esses estudos foram realizados em humanos (DONNELLY, 2008), mas logo foram aplicados em animais de produção (GODDARD; HAYES, 2009). Por exemplo, Hayes et al. (2010) utilizaram uma mistura de metodologias de modelos lineares para estimar a proporção da variância genética associada a três características quantitativas de bovinos de leite. Já Long et al. (2008) aplicaram o GWAS para a identificação de regiões genômicas ligadas à mortalidade precoce em frangos de corte, e Dorshorst et al. (2010) com a hiperpigmentação cutânea, polidactilia e outras características morfológicas em galinhas.

A utilização de dados massivos de polimorfismos tem auxiliado, significativamente, na seleção de animais em programas de melhoramento genético, e vem demonstrando ser muito eficiente na identificação de SNPs associados a características de interesse econômico, possibilitar a busca de genes candidatos e melhorar o entendimento dos mecanismos de regulação das características mais variadas.

5. Referências bibliográficas

BEUZEN, N. D.; STEAR, M. J.; CHANG, K. C. Molecular markers and their use in animal breeding. **The Veterinary Journal**, v. 160, n. 1, p. 42-52, 2000.

BOURDON, R. M. **Understanding Animal Breeding**. 2.ed. New Jersey: Prentice Hall, 2000. 538p.

DONNELLY, P. Progress and challenges in genome-wide association studies in humans. **Nature**, n. 456, p. 728-731, 2008.

DORSHORST, B.; OKIMOTO, R.; ASHWELL, C. Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the Silkie chicken. **Journal of Heredity**, v. 101, n. 3, p. 339-350, 2010.

ELANDT-JOHNSON, R. C. **Probability models and statistical methods in Genetics**. New York: John Wiley & Sons, Inc., 1971. 592p.

FALCONER, D. S.; MACKAY, F. C. **Introduction to quantitative genetics**. 4. ed. New York: Longman Group, 1996. 463 p.

GODDARD, M. E.; HAYES, B. J. Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, v.10, n.6, p. 381-391, 2009.

HAYES, B. J.; PRYCE, J.; CHAMBERLAIN, A. J.; BOWMAN, P. J.; GODDARD, M. E. Genetic Architecture of Complex Traits and Accuracy of Genomic Prediction: Coat Colour, Milk-Fat Percentage, and Type in Holstein Cattle as Contrasting Model Traits. **Plos Genetics**, v. 6, n. 9, 2010.

LONG, N.; GIANOLA, D.; ROSA, G. J.; WEIGEL, K. A.; AVENDAÑO, S. Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. **Jornal Animal Breed Genetics**. v. 124, n. 6, p. 377-389, 2007.

MEUWISSEN, T. H.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 4, n. 157, p. 1.819-1.829, 2001.

NEI, M. **Molecular population genetics and evolution**, 1975. 288p.

PEREIRA, J. C. C. **Melhoramento genético aplicado à produção animal**. 6. ed. Belo Horizonte: FEPMVZ, 2012. 758 p.

RESENDE, M. D. V.; LOPES, P. S.; SILVA, R. L.; PIRES, I. E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético

co. **Pesquisa Florestal Brasileira**, Colombo, n.56, p.63-77, 2008.

SCHNABEL, A.; BEERLI, P.; ESTOUP, A. et al. (1998) A guide to software packages for data analysis in molecular ecology. In Carvalho G, eds. **Advances in Molecular Ecology**. Amsterdam: IOS Press. 1998. p. 291-303.

VAN VLECK, L. D. **Selection index and introduction to mixed model methods**. Boca Raton: CRC Press, 1993. p. 481.

WEIR, B. S. **Genetic data analysis. Methods for discrete population genetic data**. Madison, 1990. p. 377.

WRIGHT, S. Systems of mating. **Genetics**, v.6, p.111-178, 1921.

WRIGHT, S. The genetical structure of populations. **Annals of Eugenics**, v.15, p.395-420, 1951.

WRIGHT, S. Evolution in Mendelian populations. **Genetics**, v.16, p.97-159, 1931.

Epidemiologia molecular e diversidade genética de microrganismos

*Andressa de Souza-Pollo
Luciana de Cenço Corrêa de Lacerda
Roberta Cristina Delphino Carboni
Luiz Augusto do Amaral*

1. Introdução

A epidemiologia molecular engloba a identificação e o monitoramento de espécies de microrganismos patogênicos a humanos, animais e plantas, por meio de técnicas de biologia molecular. A informação taxonômica correta de um microrganismo desconhecido é estritamente essencial para o conhecimento de sua epidemiologia, biodiversidade e relação com outros organismos no ecossistema. Com o advento da era genômica, estudos taxonômicos polifásicos e a aplicação de marcadores moleculares para o entendimento da diversidade genética das populações microbianas são abordagens indispensáveis para o desenvolvimento de pesquisas microbiológicas (PRAKASH et al., 2007; DAS et al., 2014).

As técnicas convencionais de microbiologia, comumente aplicadas para a identificação de microrganismos, não são suficientes para a classificação taxonômica segura de espécies. A evolução dos microrganismos em resposta à constante modificação ambiental que os envolve, torna pouco segura a identificação de espécies pela coloração e pela forma da colônia em meios de cultura específicos, forma e tamanho de esporos, propriedades de coloração, motilidade, gama de hospedeiros e patogenicidade (PRAKASH et al., 2007). Além disso, alguns microrganismos,

embora viáveis, não são cultiváveis e poderão ser identificados apenas pela aplicação de estudos moleculares. Tendo em vista a importância do conhecimento das várias técnicas disponíveis para estudos de epidemiologia molecular e diversidade genética de microrganismos, neste capítulo, serão abordados os principais métodos moleculares aplicados para este fim atualmente.

2. Identificação de espécies por PCR espécie-específico

A PCR espécie-específico permite a identificação de gêneros e espécies microbianas, assim como a presença de genes de interesse (MOLINA et al., 2015). A identificação, por meio de uma PCR básica, poderá ser realizada em isolados ou até mesmo diretamente do DNA total de amostras de planta, de excreções e tecidos animais, do solo, etc.

Para estudos de epidemiologia molecular, cujo objetivo é mapear a presença de espécies microbianas em uma grande quantidade de amostras, sem a necessidade de isolamento, a utilização de primers específicos pode constituir uma vantagem em termos de rapidez e custos. No entanto, como a análise de positividade da espécie de um determinado microrganismo, por PCR espécie-específico, é apenas baseada na identificação de uma banda com um tamanho específico de pares de bases, este tipo de identificação taxonômica não é confiável. Primers específicos para determinada espécie microbiana podem vir a amplificar uma banda de igual tamanho para outros microrganismos, resultando em dados errôneos de positividade. Dessa forma, ao usar a técnica de PCR espécie-específico, sugere-se a utilização de mais de um primer específico para o mesmo fim, preferencialmente desenhado a partir de regiões diferentes do genoma. Tendo em vista estes contratemplos, nas pesquisas microbiológicas aplicadas, é imprescindível o sequenciamento de genes específicos para a identificação taxonômica de microrganismos.

3. Abordagem polifásica para classificação microbiana

Nas últimas duas décadas, a identificação taxonômica de fungos e bactérias, cujo gênero era desconhecido, baseava-se, na

maioria das vezes, no sequenciamento da região ITS e 16S rDNA, respectivamente (WHITE et al., 1990; SRINIVASAN et al., 2015). Atualmente, soma-se ao sequenciamento de tais regiões ribossomais o sequenciamento de genes *housekeeping*, ou seja, de genes constitutivos essenciais para a manutenção de funções celulares básicas. A abordagem polifásica consiste no sequenciamento de sete a oito genes constitutivos e pode ser aplicada em estudos taxonômicos de fungos e bactérias (CROUS; HAWKSWORTH; WINGFIELD, 2015; WOO et al., 2008).

A abordagem polifásica para a identificação de fungos compreende o sequenciamento de genes codificadores das proteínas GAPDH (gliceraldeído-3-fosfato desidrogenase), *tub2* (beta-tubulina), *tef1* (Fator de alongação 1-alpha), *act* (actina) e *his3* (histona), suplementado a genes ribossomais específicos em nível de espécie. O sequenciamento de outras regiões conservadas como a LSU (Large subunit), SSU (small subunit) e RPB2 (RNA polimerase II) pode prover melhor discriminação em nível de gênero ou família (CROUS; HAWKSWORTH; WINGFIELD, 2015). Para a identificação de bactérias, a abordagem polifásica é também conhecida por MLST (*Multilocus Sequence Typing*). Neste caso, as regiões gênicas sequenciadas que se mostrarem únicas dentro de uma mesma espécie são caracterizadas por um número de alelo. Dessa forma, cada isolado bacteriano é caracterizado pelos alelos de cada um dos sete genes sequenciados, que, em conjunto, constituem o perfil alélico do isolado bacteriano ou a sua ST (*Sequence Type*) (PÉREZ-LOSADA et al., 2013; DAS et al., 2014). As árvores filogenéticas poderão ser obtidas pelos métodos de agrupamento de máxima parcimônia, máxima verossimilhança e/ou por análise Bayesiana, utilizando-se do modelo evolutivo que melhor se adapta aos dados (POSADA; BUCKLEY, 2004; PÉREZ-LOSADA et al., 2013; CROUS; HAWKSWORTH; WINGFIELD, 2015).

Além da abordagem polifásica pelo sequenciamento de Sanger, com o advento das técnicas de sequenciamento de segunda geração, ou NGS (*next generation sequencing*), pesquisadores têm encorajado os microbiologistas a considerarem a possibilidade da aplicação em larga escala do sequenciamento genômico completo, em substituição às técnicas atualmente aplicadas para classificação taxonômica (MAIDEN et al., 2013). O objetivo é criar bases de dados confiáveis, utilizadas mundialmente, e as-

sim evitar a nomeação de novas espécies de forma inapropriada (CROUS; HAWKSWORTH; WINGFIELD, 2015).

4. Diversidade genética: principais marcadores moleculares e suas aplicações

O sequenciamento genômico é imprescindível para a caracterização e identificação dos microrganismos; no entanto, quando realizado a partir de genes constitutivos e mesmo a partir de regiões ribossomais intergênicas, não fornece um conhecimento abrangente da diversidade genética existente em uma população microbiana da mesma espécie. Neste contexto, os marcadores moleculares podem ser utilizados complementarmente ao sequenciamento.

Os marcadores moleculares fornecem o perfil genético dos isolados microbianos e são aplicados em estudos de diversidade genética de populações da mesma espécie. A escolha apropriada do tipo de marcador molecular a ser aplicado depende significativamente do objetivo do trabalho e do contexto epidemiológico ao qual este se encontra inserido, além da espécie microbiana que se deseja analisar e o tempo disponível para o estudo (SABAT et al., 2013). Cada marcador molecular possui vantagens e desvantagens que devem ser levadas em consideração no momento da escolha. Assim, é importante conhecer a metodologia de cada marcador e os resultados que estes podem oferecer.

4.1. AFLP

Os marcadores AFLP (*Amplified fragment length polymorphism*) encontram-se no *ranking* dos marcadores mais acurados da atualidade, produzem impressões digitais particulares do DNA, específicas para cada espécie, e sua utilização não requer conhecimento prévio da sequência genômica (VOS et al., 1995; CHIAL, 2008). Comparativamente às demais técnicas, esta detecta maior número de loci, além de possibilitar ampla cobertura do genoma. Os marcadores AFLP, dentre outros, sobreviveram à era genômica e continuarão constituindo um método relevante mesmo para espécies com genoma conhecido, devido às suas

inúmeras funcionalidades. Tais marcadores podem ser utilizados para diferenciação entre espécies, identificação da variabilidade genética inter e intrapopulacional, mapeamento genético e isolamento de fragmentos do genoma com características particulares, cuja sequência é desconhecida (ZHANG; VAN PARIJS; XIAO, 2014). Além disso, tal técnica pode ser aplicada como auxílio à tecnologia NGS, que visa ao sequenciamento genômico completo (VAN POECKE et al., 2013).

A técnica AFLP é composta por quatro etapas. Primeiramente, o DNA genômico íntegro é digerido, concomitantemente, por duas enzimas de restrição, sendo uma de corte raro, e outra de corte frequente, por exemplo, a EcoRI e MseI, respectivamente. Aos fragmentos gerados são ligados adaptadores, cujas sequências são complementares às extremidades coesivas dos mesmos, por meio da enzima T4 DNA Ligase. A seguir, a PCR de pré-amplificação permite a seleção dos fragmentos clivados em uma extremidade pela EcoRI, e na outra, pela MseI. Para tanto, adicionalmente à sequência dos adaptadores, os primers da PCR pré-seletiva possuem um nucleotídeo complementar em sua extremidade 3'. Na etapa de amplificação seletiva, por sua vez, podem ser acrescentados mais um a dois nucleotídeos a esta extremidade dos primers, resultando assim na seleção refinada dos fragmentos. Nesta última etapa, os primers ligantes à extremidade EcoRI podem ser marcados com fluorescência.

Todas as etapas que compõem o AFLP devem ser realizadas com acurácia para que a eficiência das PCRs possibilite a correta detecção das bandas ao final do processo. Esta técnica é reprodutível e muito eficiente para analisar a diversidade genética de microrganismos (SOUZA et al., 2013).

4.2. Microssatélites

Os microssatélites tratam-se de uma parte repetitiva do DNA com cerca de 2-5 pares de bases que se repetem sequencialmente de 5 a 50 vezes no genoma de eucariotos e bactérias (RICHARD; KERREST; DUJON, 2008). Microssatélites e minissatélites podem ser chamados de VNTR (*variable number of tandem repeats*) pelos microbiologistas, STRs (*short tandem repeats*) pelos geneticistas forenses, ou simplesmente de SSRs (*simple sequence repeats*) pelos

melhoristas de plantas (SABAT et al., 2013; GONTHIER et al., 2015).

As mutações de microssatélites resultam no ganho ou na perda de uma unidade repetitiva inteira e, às vezes, de duas ou mais unidades repetitivas simultaneamente. Essas alterações estão relacionadas ao escorregamento da DNA polimerase na replicação, resultantes de incompatibilidades entre as fitas de DNA durante a meiose (VIGUERA; CANCEILL; EHRLICH, 2001). Os microssatélites são herdáveis e, portanto, são utilizados atualmente nos exames de teste de paternidade e em estudos de herdabilidade e diversidade genética de microrganismos (SABAT et al., 2013; GONTHIER et al., 2015). Muitos microssatélites estão localizados em regiões não codificadoras do DNA e são biologicamente silenciosos, já outros estão localizados em regiões codificadoras e, neste caso, essas mutações podem resultar em mudanças fenotípicas e doenças (MELISSA et al., 2016).

Há vários trabalhos na literatura que fornecem a sequência de primers que flanqueiam as regiões dos microssatélites especialmente desenhados para cada gênero microbiano, planta ou animal. No entanto, caso ainda não tenham sido realizados estudos de microssatélites para o organismo que se deseja estudar, a pesquisa de tais microssatélites, o desenho de primers e a otimização das reações constituem passos laboriosos. Uma vez possuindo os primers, os fragmentos de DNA são amplificados por meio de uma PCR básica. Outra variação desta técnica são os ISSRs (*inter-simple sequence repeat*); neste caso, as sequências complementares a dois microssatélites vizinhos são utilizadas para o desenho de primers, e a região variável entre estes é amplificada (GRAMAJE et al., 2014).

A técnica de microssatélites é reproduzível e muito utilizada atualmente no acesso à diversidade genética e, principalmente, em estudos de herdabilidade; no entanto, em caso de mutação do DNA que flanqueia os microssatélites, os primers podem passar a não amplificar corretamente os fragmentos contendo os microssatélites. Também, primers desenhados para uma determinada espécie podem funcionar em outras; entretanto, há o risco de que nenhum alelo seja de fato identificado e/ou regiões não relativas a microssatélites venham a ser amplificadas. Assim, é necessário atentar-se aos resultados obtidos para que eventuais modificações no protocolo possam ser realizadas.

4.3 REP-PCR

A Rep-PCR (*Repetitive-element PCR*) é uma técnica baseada na análise de perfis genéticos utilizada para classificar isolados bacterianos. O método utiliza primers que hibridizam sequências não codificadoras repetitivas que se encontram espalhadas pelo genoma. O fragmento de DNA flanqueado pelos primers específicos, localizado entre os elementos repetitivos, é amplificado por PCR de acordo com a distribuição desses elementos no genoma. Os padrões de bandejamento dos isolados, obtidos por eletroforese, são comparados, a fim de se determinar a relação genética entre os mesmos. Esta técnica não é reprodutível, podendo o perfil genético dos isolados mudar, dependendo do executor das reações de PCR. Mesmo diante deste fato, a técnica pode ser utilizada e produz resultados fidedignos quando executada por uma mesma pessoa (SABAT et al., 2013).

Os elementos repetitivos estão localizados em diferentes posições e orientações do genoma. A técnica Rep-PCR engloba a utilização dos marcadores moleculares REP (*Repetitive Extragenic Palindromic*), ERIC (*Enterobacterial Repetitive Intergenic Consensus*) e BOX (VERSALOVIC et al., 1994). O marcador REP corresponde a regiões repetitivas ao longo de elementos de 35-40 pb (STERN et al., 1984; VERSALOVIC et al., 1991). O marcador ERIC corresponde a elementos repetitivos de 124-127 pb (HULTON et al., 1991; VERSALOVIC et al., 1991) e o marcador BOX é dividido nas subunidades BOX A (57 pb), BOX B (43 pb) e BOX C (50 pb). As sequências da subunidade BOX A aparecem altamente conservadas em várias espécies de bactérias, já as sequências das subunidades BOX B e BOX C são encontradas especialmente em bactérias Gram-positivas, pois foram desenhadas com base nas subunidades repetitivas de *Streptococcus pneumoniae* (MARTIN et al., 1992; KOEUTH; VERSALOVIC; LUPSKI, 1995).

4.4 PFGE

A técnica de PFGE (*Pulsed Field Gel Electrophoresis*) é utilizada para separação de grandes moléculas de DNA por eletroforese em gel de agarose, em um campo elétrico que, periodicamente, muda de direção (SCHWARTZ; CANTOR, 1984). Esta técnica

é considerada como padrão-ouro em estudos epidemiológicos de bactérias patogênicas. Em tais estudos, o DNA genômico da bactéria é extraído dentro de um bloco de agarose, que é submetido à fragmentação com uma enzima de restrição específica para cada espécie. Os fragmentos obtidos são então separados em um equipamento de campo pulsado em gel de agarose. É uma metodologia reprodutível e recomendada na discriminação de isolados bacterianos obtidos de um surto de infecção (WERNER et al., 2015).

5. Abordagem metagenômica

Caso se deseje estudar a biodiversidade microbiana presente em um determinado ambiente, sem conhecimento prévio das espécies que o habitam, emprega-se o sequenciamento metagenômico, que, inclusive, abrange a identificação de microrganismos não cultiváveis (SHARPTON, 2014). Neste caso, obtém-se o DNA total de amostras de plantas, excreções animais, solo, água, dentre outras inúmeras possibilidades, e este é submetido à PCR para amplificação de regiões de interesse que, conjuntamente, serão sequenciadas aos milhares em sequenciadores de segunda geração.

A metagenômica pode auxiliar tanto na identificação de microrganismos quanto na prospecção de genes de interesse (SONI et al., 2016). Na maioria das vezes, para o estudo da biodiversidade de fungos, aplica-se o sequenciamento da região ITS rDNA (TONGE; PASHLEY; GANT, 2014), e para comunidades bacterianas, aplica-se o sequenciamento metagenômico da região 16S rDNA (TIKHONOV; LEACH; WINGREEN, 2015); no entanto, tais regiões podem variar de acordo com o objetivo do trabalho. Para a execução da metagenômica, existem diferentes metodologias, plataformas de equipamentos e rendimento de sequências que deverão ser escolhidos de acordo com o alvo pretendido, com o recurso financeiro e com profissionais qualificados disponíveis.

6. Referências bibliográficas

CHIAL, H. DNA fingerprinting using amplified fragment length polymor-

phisms (AFLP): No genome sequence required. **Nature Education**, London, v. 1, n. 1, p. 176, 2008.

CROUS, P. W.; HAWKSWORTH, D. L.; WINGFIELD, M. J. Identifying and Naming Plant-Pathogenic Fungi: Past, Present, and Future. **Annual Review of Phytopathology**, v. 53, p. 247-67, 2015.

DAS, S.; DASH, H.R.; MANGWANI, N.; CHAKRABORTY, J.; KUMARI, S. Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. **Journal of Microbiological Methods**, v. 103, p. 80-100, 2014.

DEL VECCHIO, V. G.; PETROZIELLO, J. M.; GRESS, M. J.; MCCLESKEY, F. K.; MELCHER, G. P.; CROUCH, H. K.; LUPSKI, J. R. Molecular genotyping of methicillin-resistant *Staphylococcus aureus* via fluorophore-enhanced repetitive-sequence PCR. **Journal of Clinical Microbiology**, v. 33, n. 8, p. 2.141-2.144, 1995.

GONTHIER, P.; SILLO, F.; LAGOSTINA, E.; ROCCOTELLI, A.; Cacciola, O. S.; Stenlid, J.; Garbelotto, M. Selection processes in simple sequence repeats suggest a correlation with their genomic location: insights from a fungal model system. **BMC Genomics**, v. 16, p. 1107, 2015.

GRAMAJE, D.; LEÓN, M.; SANTANA, M.; CROUS, P. W.; ARMENGOL, J. Multilocus ISSR Markers Reveal Two Major genetic groups in Spanish and South African populations of the grapevine fungal pathogen *Cadophora luteo-olivacea*. **PLoS One**, v. 9, n. 10, e110417, 2014.

HULTON, C.S.J.; HIGGINS, C.F.; SHARP, P.M. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. **Molecular Microbiology**, v.5, p.825-762, 1991.

KOEUTH, T.; VERSALOVIC, J.; LUPSKI, J. R. Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. **Genome Research, Cold Spring Harbor**, v. 5, p. 408-418, 1995.

MAIDEN, M. C. J.; MELISSA, J.; JANSEN VAN RENSBURG; James, E.; BRAY; SARAH, G.; EARLE; SUZANNE, A.; FORD; KEITH, A.; JOLLEY; Noel D. McCarthy. MLST revisited: the gene-by-gene approach to bacterial genomics. **Nature Review Microbiology**, v. 11, n.10, p. 728-736, 2013.

MARTIN, B.; HUMBERT, O.; CAMARA, M.; GUENZI, E.; WALKER, J.; MITCHELL, T.; ANDREW, P.; PRUDHOMME, M.; ALLOING, G.; HAKENBECK, R.; MORRISON, D.A.; BOULNOIS, G.J.; CLAVERYS, J.P. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. **Nucleic Acids Research**, Oxford v.20, p.3479–3483, 1992.

MELISSA, G.; WILLEMS, T.; GUILMATRE, A.; ZENG, H.; MARKUS, B.; GEORGIEV, S.; DALY, M. J.; PRICE, A. L.; PRITCHARD, J. K.; SHARP, A. J.; ERLICH, Y. Abundant contribution of short tandem repeats to gene expression variation in humans. **Nature Genetics**, v. 48, p. 22-29, 2016.

MOLINA, F.; LÓPEZ-ACEDO, E.; TABLA, R.; ROA, I.; GÓMEZ, A.; REBOLLO, J. E. Improved detection of *Escherichia coli* and coliform bacteria by multiplex PCR. **BMC Biotechnology**, v. 15, n. 48, 2015.

PÉREZ-LOSADA, M.; CABEZAS, P.; CASTRO-NALLAR, E.; CRANDALL, K. A. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. **Infection, Genetics and Evolution**, Volume 16, June 2013, p. 38-53.

POSADA, D.; BUCKLEY, T. R. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. **Systematic Biology**, v. 53, p. 793-808, 2004.

PRAKASH, O.; VERMA, M.; SHARMA, P.; KUMAR, M.; KUMARI, K.; SINGH, A.; KUMARI, H.; JIT, S.; GUPTA, S.K.; KHANNA, M.; Lal, R. Polyphasic approach of bacterial classification - an overview of recent advances. **Indian Journal of Microbiology**, v. 47, p. 98-108, 2007.

RICHARD, G.-F.; KERREST, A.; DUJON, B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. **Microbiology and Molecular Biology Reviews**, v. 72, n. 4, p. 686-727, 2008.

SABAT, A.J.; BUDIMIR, A.; NASHEV, D.; SÁ-LEÃO, R.; VAN DIJL, J.M.; LAURENT, F.; GRUNDMANN, H.; FRIEDRICH, A.W. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. **Euro Surveillance**, v. 18, n. 4, p. 1-14, 2013.

SCHWARTZ, D. C.; CANTOR, C. R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. **Cell**, v. 37, n. 1, p. 67-75, 1984.

SHARPTON, T. J. An introduction to the analysis of shotgun metagenomics data. **Frontiers in Plant Science**, v. 5, p. 1-14, 2014.

SONI, R.; SUYAL, D. C.; SAI, S.; GOEL, R. Exploration of nifH gene through soil metagenomes of the western Indian Himalayas. **3 Biotech**, v. 6, n. 25, p. 1-4, 2016.

SOUZA, A.; CARBONI, R. C. D.; WICKERT, E.; LEMOS, E. G. M.; Goes, A. Lack of host specificity of *Colletotrichum* spp. isolates associated with anthracnose symptoms on mango in Brazil. **Plant Pathology, Nottingham**, v. 62, p. 1038-1047, 2013.

SRINIVASAN, R.; KARAOZ, U.; VOLEGOVA, M.; MACKICHAN, J.; KATO-MAE-

DA, M.; MILLER, S.; NADARAJAN, R.; BRODIE, E. L.; LYNCH, S. V. Use of 16S rRNA Gene for Identification of a Broad Range of Clinically Relevant Bacterial Pathogens. **PLoS ONE**, v. 10, n. 2, e0117617, 2015.

STERN, M.J.; AMES, G.F.L.; SMITH, N.H.; ROBINSON, E.C.; HIGGINS, C.F. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. **Cell, Cambridge**, v.37, p.1015–1026, 1984.

TIKHONOV, M.; LEACH, R. W.; WINGREEN, N. S. Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. **The ISME Journal**, v. 9, p. 68-80, 2015.

TONGE, D. P.; PASHLEY, C. H.; GANT, T. W. Amplicon –Based metagenomic analysis of mixed fungal samples using proton release amplicon sequencing. **PLoS ONE**, v. 9, n. 4, e93849, 2014.

VAN POECKE, R. M.; MACCAFERRI, M.; TANG, J. F.; TRUONG, H. M.; JANSSEN, A.; VAN ORSOUW, N. J.; SALVI, S.; SANGUINETI, M. C.; TUBEROSA, R.; VAN DER VOSSSEN, E. A. Sequence-based SNP genotyping in durum wheat. **Plant Biotechnology Journal**, Malden, v. 11, p. 809–817, 2013.

VERSALOVIC, J.; KOEUTH, T.; LUPSKI, J.R. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. **Nucleic Acids Research**, Oxford, v.19, p.6.823-6.831, 1991.

VERSALOVIC, J.; SCHNEIDER, M.; DE BRUIJN, F.J.; LUPSKI, J.R. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. **Methods in Molecular and Cell Biology**, v.5, p.25-40, 1994.

VIGUERA, E.; CANCEILL, D.; EHRLICH, S. D. Replication slippage involves DNA polymerase pausing and dissociation. **The EMBO Journal**, v. 20, n. 10, p. 2587-2595, 2001.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; LEE, T.; HORNES, M.; FRITERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, Oxford, v. 23, n. 21, p. 4407-4414, 1995.

WERNER, G.; FLEIGE, C.; NEUMANN, B.; BENDER, J. K.; LAYER, F.; KLARE, I. Evaluation of DiversiLab®, MLST and PFGE typing for discriminating clinical *Enterococcus faecium* isolates. **Journal of Microbiological Methods**, v. 118, p. 81-84, 2015.

WHITE, T. J.; BRUNS, T.; LEE, S.; TAYLOR, J. W. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: INNIS, M.A.; GELGARD, D.H.; SNINSKY, J.J.; WHITE, T.J. **PCR protocols: a guide to methods and applications**. San Diego: Academic Press, 1990. cap.38, p. 315-322.

WOO, P.C.Y.; LAU, S.K.P.; TENG, J.L.L.; TSE, H.; YUEN, K.-Y. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. **Clinical Microbiology and Infection**, v. 14, n. 10, p. 908-934, 2008.

ZHANG, Z.; VAN PARIJS, F. R. D.; XIAO, B. The status of AFLP in the genomics era and a pipeline for converting AFLPs into single-locus markers. **Molecular Breeding, Lleida**, v. 34, p. 1245-1260, 2014.

Genômica funcional: análise de transcriptoma animal por RNA-Seq

*Danielly Beraldo dos Santos Silva
Larissa Fernanda Simielli Fonseca
Jorge Augusto Petroli Marchesi
Daniele Fernanda Jovino Gimenez
Lucia Galvão de Albuquerque*

1. Introdução

No âmbito da genômica funcional animal, o estudo do transcriptoma avalia o padrão de expressão de genes em tecidos ou órgãos específicos, com funções importantes para o metabolismo animal (RAMAYO-CALDAS et al., 2012). A sequência completa dos RNAs de um organismo pode ser obtida por meio do sequenciamento de nova geração (NGS – Next Generation Sequencing), técnica conhecida como RNA-Seq, também chamado de High-throughput RNA sequencing (RAMAYO-CALDAS et al., 2012). O método de RNA-Seq, tipicamente, consiste no isolamento do RNA total, enriquecimento e depleção do rRNA, conversão do RNA-alvo em uma biblioteca de cDNA com adaptadores anexados em uma ou ambas as extremidades, sequenciamento em plataformas *high-throughput* e análise dos dados (GRIFFITH et al., 2015).

O desenvolvimento do NGS e o uso do RNA-Seq permitiram a quantificação de todos os genes expressos e suas isoformas (splicing alternativo) de maneira mais acurada, diferentemente de outros métodos tradicionais de análise de expressão gênica, como Microarranjos e Análise da Expressão Gênica em Série (SAGE) (TANG et al., 2011). Para obter uma análise de expressão diferencial de genes efetiva, é importante obter estimativas acuradas da expressão de cada amostra, além de contabilizar todas as possíveis fontes de variações, sejam técnicas ou biológicas.

Essa metodologia também permite identificar sítios de poliadenilação, RNAs não codificantes, polimorfismos de nucleotídeo único (SNPs) e elementos transponíveis. Além disso, propicia a descoberta de novos transcritos e RNAs longos não codificantes (RAMAYO-CALDAS et al., 2012; TRAPNELL et al., 2012).

Griffith et al. (2010) relataram que 85% de novos eventos de *splicing* e 88% da predição de genes diferencialmente expressos por RNA-Seq, são validados por abordagens “gold standard”, tais como a PCR Quantitativa em Tempo Real (RT-qPCR). O RNA-Seq permitiu descobertas importantes em relação aos mecanismos moleculares de animais de produção. Estas descobertas incluem expressão diferencial de genes relacionados com características de interesse econômico em diversas raças de bovinos de corte e leite (JIN et al., 2012; McCABE et al., 2012; CÉSAR et al., 2015; TIZIOTO et al., 2015). Deste modo, RNA-Seq é uma importante ferramenta para desvendar os mecanismos subjacentes de características complexas, promovendo melhor compreensão da regulação genética do fenótipo de interesse (TIZIOTO et al., 2015). Neste capítulo, exploramos os conceitos de biologia molecular inerente à técnica de RNA-Seq, análise e interpretação de dados.

2. Isolamento do RNA, preparação da biblioteca e estratégia de sequenciamento

A escolha dos parâmetros para RNA-Seq pode ter impactos significativos sobre estratégia de análise (VAN DIJK et al., 2014). Estes parâmetros incluem a possibilidade de enriquecer o RNA total com a calda poli (A) ou redução do RNA ribossomal no meio (depleção); métodos de construção de bibliotecas (*stranded* ou *unstranded*) e o uso de técnicas de normalização do cDNA (BOGDANOV et al., 2010; ARCHER et al., 2014). De igual modo, a escolha da plataforma de sequenciamento: ABI SOLiD, Illumina (GA I e II, HiSeq, MiSeq), Roche 454, IonTorrent, IonProton, Helicos, PacBio, entre outros; tamanho de *reads*; o uso de *single-end* (apenas uma fita) ou *paired-end* (ambas as fitas); tamanho da cobertura e profundidade de sequenciamento, a fim de obter transcritos raros ou variações (ADICONIS et al., 2013) e outros parâmetros podem influenciar nas etapas de análise e na inter-

pretação dos dados.

Como a maioria dos experimentos de RNA-Seq envolve comparações entre diferentes condições, é desejável que estes fatores sejam consistentes em todas as amostras e/ou réplicas biológicas do experimento. Para a construção de biblioteca de RNA-Seq, em geral, o RNA total de alta qualidade é extraído; após, o rRNA é depletado, é realizado o enriquecimento de subpopulações de RNAs (mRNAs ou pequenos RNAs), o RNA-alvo é fragmentado em tamanho uniforme e é transformado em cDNA. De acordo com a tecnologia de sequenciamento utilizada, são ligados adaptadores em uma ou ambas as extremidades; após, as amostras são amplificadas e sequenciadas. O sequenciamento gera sequências curtas, *reads* de 30-400 pb, podendo ser de modo *single-end* ou *pair-end* (HOLT; JONES, 2008; WANG et al., 2009).

Todas as tecnologias de NGS são capazes de gerar informação sobre milhões de pares de bases em uma única corrida (CARVALHO; SILVA, 2010). Dentre as tecnologias NGS, a mais utilizada no Brasil para RNA-Seq é a Illumina, que utiliza a tecnologia de sequenciamento por síntese. Cada uma das 8 *lanes* de corrida por placa da plataforma gera, em média, 37,5 Gb com leituras *single-end* de até 100 pb (ILLUMINA® HiSeq2500 System).

A cobertura utilizada no experimento é outro fator importante, o que tem implicações diretas no custo. Maior cobertura requer mais profundidade de sequenciamento, assim é possível detectar um transcrito raro ou uma variação. Normalmente, quanto maior o genoma, mais complexo será o transcriptoma e mais sequenciamento será necessário para que seja atingida a cobertura adequada. O cálculo da cobertura, em termos de transcriptomas, não é padronizado, já que o número e o nível exatos de isoformas são normalmente desconhecidos, também porque a atividade de transcrição varia muito ao longo do genoma (WANG et al., 2009).

Como outras tecnologias de alto rendimento de sequenciamento, a análise de RNA-Seq também enfrenta vários desafios de informática, tais como armazenamento, recuperação e processamento de grandes quantidades de dados. Os dados devem ser processados de tal modo que sejam removidas as *reads* de baixa qualidade, e os erros sejam reduzidos nas etapas de análises.

3. Processamento e análise de dados para obter o perfil da expressão diferencial de genes

Para analisar os dados gerados por RNA-seq, é preciso primeiramente conhecer os principais formatos de arquivos e modelos de dados subjacentes que os representam. Estes incluem o formato FASTA para o armazenamento de dados do genoma de referência (PEARSON et al., 1988); o formato de transferência de genes (GTF), para armazenar as informações de transcrição/ gene; o formato FASTQ, para armazenar dados de *reads* brutas (COCK et al., 2010); e o formato SAM/BAM, para o mapa de alinhamentos das *reads* (LI et al., 2009).

Os primeiros passos do *workflow* para RNA-Seq envolvem uma análise inicial de controle de qualidade (QC) dos dados brutos em arquivos FASTQ. O controle de qualidade, tipicamente, envolve a análise K-mer para identificar potenciais problemas, tais como a contaminação do adaptador, a remoção ineficiente de sequências ribossomais ou uma abundância de fragmentos mais curtos do que o comprimento das *reads*. Além disto, avaliações de QCs adicionais obtidas nesta fase podem identificar perfis de baixa qualidade, problemas que possam ter ocorrido durante o sequenciamento ou muitas bases ambíguas (indicado como ‘N’ em arquivos FASTQ). Dependendo da estratégia de construção da biblioteca de RNA-Seq, é necessário realizar a trimagem das sequências antes do alinhamento dos dados. Duas estratégias de trimagem comuns incluem “*adapter trimming*” e “*quality trimming*”. *Adapter trimming* envolve a remoção da sequência do adaptador, mascarando sequências específicas utilizadas durante a construção da biblioteca. *Quality trimming* geralmente remove as extremidades das leituras, onde os índices de qualidade de base são baixos, em que os erros de sequência e os *mismatches* resultantes impedem o alinhamento das *reads* (GRIFFITH et al., 2015).

Softwares, como o FastQC, PrinSeq, CutAdapt e Scythe, removem adaptadores, avaliam a qualidade e ajustam os dados de RNA-Seq brutos antes do alinhamento. Uma vez estabelecido o controle de qualidade, é necessário escolher a melhor estratégia (*softwares*) para avaliar os dados. Para a análise da expressão diferencial de genes, o protocolo Tuxedo vem sendo amplamente utilizado, e consiste na utilização de várias ferramentas que per-

mitem a análise de dados gerados a partir de RNA-Seq. O *pipeline* Tuxedo inclui o TopHat, que utiliza o Bowtie para mapear as *reads* no genoma de referência; em seguida, identifica as junções de união entre os transcritos. Inclui também Cufflinks *tools* (Cufflinks, Cuffmerge, Cuffquant, Cuffnorm, Cuffdiff), que permitem: 1) montar os transcritos; 2) fusionar os transcriptomas gerados em um transcriptoma de referência; 3) quantificar os níveis de expressão gênica; 4) calcular os valores normalizados para a expressão de genes e transcritos; 5) estimar o nível de expressão baseado em um índice denominado FPKM (*Fragments Per Kilo-base of transcript per Million fragments mapped*).

O Cufflinks é uma ferramenta de análise de expressão diferencial que compara os índices FPKM dos diferentes tratamentos experimentais e identifica alterações significativas no nível de expressão dos transcritos entre os tratamentos, por meio de um rigoroso teste estatístico. Uma vez que o Cufflinks não possibilita a visualização dos resultados, exportação de imagens e gráficos, foi desenvolvido o CummeRbund (implementado em R) para esta função (TRAPNELL et al., 2012; GRIFFITH et al., 2015). O DESeq2 é um *software*, implementado em R, que também é utilizado para o estudo de expressão de genes.

A maioria dos *softwares* para análise de RNA-Seq são projetados para Mac OS, Linux e Amazon Web Services (AWS) Elastic Compute (EC2) (GRIFFITH et al., 2015) e necessita de um conhecimento básico em programação e linha de comando. Para tanto, uma das alternativas é a utilização das plataformas como o Galaxy e iPlant Collaborative (baseadas em web), open source, utilizadas para a criação e execução de diversos *workflows* em bioinformática, dentre os quais, a análise de expressão diferencial de genes (GIARDINE et al., 2005; GOFF et al., 2011). Essas plataformas integram uma série de ferramentas de bioinformática, incluindo o *pipeline* Tuxedo (TRAPNELL et al., 2012), para a análise de RNA-seq, que inclui os programas Bowtie, Tophat e Cufflinks.

Muitos outros *softwares* são utilizados para análise de expressão diferencial, conforme descrito por Griffith et al. (2015). Para cada objetivo de análise, há exigências distintas e desafios inerentes aos processos biológicos. Como descrito acima, basicamente o *workflow* geral envolve a obtenção de dados brutos, pré-processamento e realização do controle de qualidade, alinhamento ou

montagem, estimação do nível de expressão e visualização dos resultados finais. Como resultado, é gerada uma lista de genes diferencialmente expressos, dos quais precisam ser identificadas suas funções biológicas.

3.1 Anotação funcional

O ponto final da análise de dados de expressão gênica por RNA-Seq é uma lista de genes diferencialmente expressos ou valores absolutos do nível de expressão dos genes de uma amostra, em que apenas esta lista pouco contribui para a compreensão de como a expressão do gene influencia nas alterações do fenótipo em estudo. Com isso, muitos métodos e ferramentas, incluindo alguns pacotes do R/Bioconductor (HUBER et al., 2015), foram desenvolvidos para a modelagem funcional dos dados de RNA-Seq, visando a novas descobertas biológicas e biomédicas.

Algumas destas ferramentas incluem a análise de vias de enriquecimento, como Gene Ontology (GO). O banco de dados GO fornece um vocabulário controlado que descreve o gene e os atributos do produto gênico em qualquer organismo. O GO inclui três ontologias que correspondem a três aspectos diferentes da biologia celular: a função molecular, o processo biológico e o componente ou a localização subcelular. A metodologia mais simples deste tipo de análise consiste no cálculo da porcentagem de genes de uma categoria GO determinada, que esteja sobre/sub-representado em um conjunto de genes, em relação a um conjunto de genes-referência. Embora GO inclua, principalmente, conceitos respeitantes ao nível celular e subcelular, também engloba os níveis acima, os órgãos e os sistemas do organismo correspondente (ASHBURNER et al., 2000).

Outra maneira de se trabalhar com estes tipos de dados é verificar a forma com que estes genes se relacionam. De modo geral, a função biológica da célula viva é um resultado de muitas moléculas que interagem, e ela não pode ser atribuída a apenas um único gene ou a uma única molécula. A atribuição funcional em KEGG (Kyoto Encyclopedia of Genes and Genomes) é um processo de ligação de um conjunto de genes no genoma com uma rede de moléculas que interagem na célula, tal como um caminho ou um complexo, representando uma função biológica de ordem

superior (KANEHISA et al., 2016). Desta maneira, a anotação funcional permite a classificação dos genes em classes funcionais, ou processos nos quais estes estão envolvidos, o que pode ser muito útil para compreender o significado fisiológico de grandes quantidades de genes e para avaliar as diferenças funcionais entre subgrupos de sequências.

4. Metodologias de validação da expressão de genes (qRT-PCR)

Para a validação dos resultados da análise de genes diferencialmente expressos, é utilizada qRT-PCR (Reação quantitativa em cadeia pela polimerase em tempo real). Na técnica qRT-PCR, diferentemente da técnica de PCR tradicional (*endpoint*), em que é necessária a análise do produto de PCR através de eletroforese em gel de agarose ou poliacrilamida após o último ciclo de amplificação (PELT-VERKUIL et al., 2008), os produtos da PCR em tempo real podem ser mensurados a cada ciclo, e este monitoramento é feito durante a fase exponencial de amplificação, e, com isso, pode-se determinar a quantidade inicial do alvo com grande precisão (AHMED, 2002). Para isso, a técnica utiliza corantes (sistemas de detecção) que se associam à amplificação do gene-alvo e emitem fluorescência. Esta fluorescência aumenta em proporção direta ao número de moléculas de produtos de PCR (amplicons) gerados (WONG; MEDRANO, 2005).

Os sistemas de detecção mais utilizados são o SYBR® Green e o TaqMan®. O sistema SYBR® Green possui capacidade de se ligar à fita dupla de DNA de forma covalente e emitir uma fluorescência que é medida e convertida numa quantidade de DNA (MACKAY et al., 2007), enquanto o sistema TaqMan® utiliza uma sonda fluorescente, que se liga à sequência-alvo, a qual tem total complementaridade, e que permite a detecção de um produto específico, conforme este se acumula durante os ciclos da PCR (PELT-VERKUIL et al., 2008). A principal diferença entre os dois sistemas é que, com o SYBR® Green, detecta-se todo o DNA fita dupla, inclusive produtos da reação inespecíficos, e, por isso, é necessário que as reações sejam otimizadas para a obtenção de resultados precisos. Por outro lado, o sistema TaqMan® é altamente específico, pois possui, além

do par de oligonucleotídeos iniciadores, mais uma sequência de nucleotídeos chamada de sonda. Desse modo, uma molécula de DNA somente será contabilizada se os oligonucleotídeos e a sonda se ligarem a ela, e, por isso, apresenta um custo muito maior em relação ao SYBR® Green.

A qRT -PCR pode ser realizada por meio de duas técnicas: quantificação absoluta ou quantificação relativa. O ensaio de quantificação absoluta descreve um experimento de qRT -PCR no qual amostras de quantidades conhecidas são diluídas em série e, em seguida, amplificadas para gerar uma curva-padrão. As amostras desconhecidas são então quantificadas por comparação com essa curva. Essa técnica permite que se saiba o número exato de cópias do RNA-alvo em uma determinada amostra biológica (PELT-VERKUIL et al., 2008).

A quantificação relativa é subdividida em duas metodologias, e a escolha da metodologia a ser utilizada depende da eficiência de amplificação dos oligonucleotídeos iniciadores utilizados nas reações, sendo que os resultados de expressão obtidos ao final do experimento devem ser iguais para os dois métodos. Quando a eficiência de amplificação dos iniciadores está dentro de um limiar aceitável, utiliza-se o método do Ct comparativo ($\Delta\Delta$ Ct); no entanto, se as eficiências não são satisfatórias, deve-se lançar mão do método da Curva Padrão para Quantificação Relativa (LIFE TECHNOLOGIES, 2016). Independentemente do método a ser seguido, em um experimento de quantificação relativa por qRT -PCR, a expressão de um gene de interesse em uma dada amostra é comparada com a expressão do mesmo gene em outra amostra (amostra calibradora).

Os resultados são expressos como um aumento ou uma diminuição da expressão do tratado em relação ao controle. Posteriormente, os valores obtidos são comparados aos resultados das mesmas amostras quantificadas em genes de referência. Este procedimento é utilizado como um controle na variabilidade experimental neste tipo de quantificação (LIFE TECHNOLOGIES, 2016). A escolha dos genes de referência deve ser feita de forma criteriosa, e não devem variar nos tecidos ou células sob investigação, ou em resposta ao tratamento experimental. Recomenda-se utilizar mais de um gene de referência, como forma de controlar possíveis valores discrepantes e grandes diferenças entre os genes (VANDESOMPELE et al., 2002).

5. Referências bibliográficas

- ADICONIS, X. et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. **Nature methods**. v.10, n.7, p.623–9, 2013.
- AHMED, F. E. Detection of genetically modified organisms in foods. **Trends in Biotechnology**, v. 20, n. 5, p. 215 – 233, 2002.
- ARCHER, S. K.; SHIROKIKH, N. E.; PREISS, T. Selective and flexible depletion of problematic sequences from RNA-seq libraries at the cDNA stage. **BMC genomics**. v.15, p.401, 2014.
- ASHBURNER, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. **Nature Genetics**, v. 25, p. 25-29, 2000.
- BOGDANOV, E. A. et al. Normalizing cDNA libraries. Current protocols in molecular biology / edited by Ausubel Frederick M et al., 2010, Chapter 5:Unit 5 12 1–27.
- CARVALHO, M. C. G.; SILVA, D. C. G. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas, **Ciência Rural**, v.40, n.3, p.735-744, 2010.
- CESAR, A.S.; REGITANO, L.C.; KOLTES, J.E.; FRITZ-WATERS, E.R.; LANNA, D.P.; et al. Putative regulatory factors associated with intramuscular fat content. **PLoS One**, v.10, n.6, p. 1-21, 2015.
- COCK, P. J.; FIELDS, C. J.; GOTO, N.; HEUER, M. L.; RICE, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. **Nucleic acids research**. v.38, n.6, p.1.767-1.771, 2010.
- GIARDINE, B. et al. Galaxy: a platform for interactive large-scale genome analysis. **Genome Research**, v. 15, n. 10, p. 1451-1455, Oct. 2005.
- GOFF, S. A. et al. The iPlant collaborative: cyberinfrastructure for plant biology. **Frontiers in Plant Science**, v.2, n.34, 2011.
- GRIFFITH, M. et al. Alternative expression analysis by RNA sequencing. **Nature methods**. v.7, n.10, p.843–847, 2010.
- GRIFFITH, M.; WALKER, J. R.; SPIES, N. C.; AINSCOUGH, B. J.; GRIFFITH, O. L. Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud. **PLOS one Computacional Biology**. v.6, p.1-20, 2015.
- HOLT, R. A.; JONES, S. J. The new paradigm of flow cell sequencing. **Genome Res**. v.18, p.839–846, 2008.

HUBER, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. **Nature Methods**.

JIN, W.; OLSON, E.N.; MOORE, S.S.; BASARAB, J.A.; BASU, U.; GUAN, L.L. Transcriptome analysis of subcutaneous adipose tissues in beef cattle using 3' digital gene expression-tag profiling. **J Anim Sci**, v.90, p.171–183, 2012.

KANEHISA, M.; SATO, Y.; KAWASHIMA, M.; FURUMICHI, M.; TANABE, M. KEGG as a reference resource for gene and protein annotation. **Nucleic Acids Research**, v. 44, p. 457–462, 2016

LI, H. et al., 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. **Bioinformatics**. v.25, m.16, p.2.078–2079, 2009.

LIFE TECHNOLOGIES, Real Time Handbook, Foster, 2012. Disponível em: <https://lifetech.postclickmarketing.com/Global/FileLib/qPCR/Real-TimePCR_Handboo_Update_FLR.pdf>. Acessado em: 15 fev. 2015.

MACKAY, I. M.; MACKAY, J.F.; NISSEN, M.D.; SLOOTS, T.P. Real-time PCR: History and Fluorogenic Chemistries in Mackay, p.1-40, I.M. (Ed.), Real-Time PCR in Microbiology, From Diagnosis to Characterization. **Press Norfolk**, UK, 2007.

MCCABE, M. et al. RNA-seq analysis of differential gene expression in liver from lactating dairy cows divergent in negative energy balance. **BMC Genomics**, v.13, p.193. 2012.

PEARSON, W. R.; LIPMAN, D. J. Improved tools for biological sequence comparison. **Proceedings of the National Academy of Sciences of the United States of America**. v.85, n.8, p.2.444–2.448, 1998.

PELT-VERKUIL, E.; VAN BELKUM, A.; HAYS, J.P. Principles and technical aspects of PCR amplification, Springer, 332 p. 2008.

RAMAYO-CALDAS, Y. et al. Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics**, v.13, n.547, p.2-18, 2012.

TANG, F.; LAO, K.; SURANI, M. A. Development and applications of single-cell transcriptome analysis. **Nature methods**, 2011.

TIZIOTO, P.C. et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n.242, p.1-14, 2015.

TRAPNELL, C. et al. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. **Nature Protocols**.

v.7, n.3, p.562-578, Mar. 2012.

VAN DIJK, E. L.; JASZCZYSZYN, Y.; THERMES, C. Library preparation methods for next-generation sequencing: tone down the bias. **Experimental cell research**. v.322, n.1, p.12–20, 2014.

VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biology**, 2002.v.12, p.115–121. 2015.

WANG, Z.; GERSTEIN, M.; SNYDER, M.; RNA-Seq: a revolutionary tool for transcriptomics. **Nat Rev Genet**. v.10, n.1, p.57–63. 2009.

WONG, M.L.; MEDRANO, J. RealTime PCR for mRNA quantification. **Bio-techniques**, v. 39, p. 75 - 85, 2005.

Principais marcadores moleculares utilizados no melhoramento de plantas

*Kian Eghrari Moraes
Carlos Henrique Caprio
Camila Baptista do Amaral
Marcela Bonafin Marconato
Gustavo Vitti Môro*

1. Introdução

Diante da crescente demanda por alimento, o melhoramento genético de plantas contribui para suprir tal necessidade, buscando garantir o abastecimento por meio do aumento da produtividade das culturas, realizado através da seleção de caracteres de interesse dentro da variabilidade genética de cada espécie.

A partir disso, os marcadores moleculares surgiram como uma ferramenta auxiliar para o melhoramento genético, agilizando processos comumente morosos, identificando genes de interesse com maior facilidade e gerando informações relevantes para o desenvolvimento das culturas. Os marcadores baseiam-se no polimorfismo existente no DNA de cada espécie, de maneira a identificar as diferenças no genoma de indivíduos da mesma ou de diferentes populações, visando a elucidar o funcionamento dos genes quando submetidos a diversas condições ambientais, entre outras aplicações

Neste capítulo, será apresentado um breve resumo dos principais marcadores moleculares utilizados no melhoramento genético de plantas. Os aspectos relevantes de cada marcador envolvendo sua origem, seu funcionamento, suas aplicações, suas vantagens e limitações serão aqui apresentados.

2. Marcadores Baseados em Hibridização

2.1 RFLP (*Restriction Fragment Length Polymorphisms*)

Logo depois de Edward Southern, em 1975, desenvolver o Southern Blot, que constitui na transferência de fragmentos de DNA do gel de agarose para uma membrana de nitrocelulose, sendo hibridizada com uma fita simples de ácido nucleico marcado radioativamente, que permite a visualização por autorradiografia ou fluorografia (SOUTHERN, 1975), David Bolstein et al. (1980) propuseram a primeira técnica de marcadores moleculares chamada RFLP, capaz de detectar polimorfismos na sequência de DNA através da hibridização de uma sonda, a qual possui uma sequência de DNA conhecida, com fragmentos de DNA, digeridos por uma ou mais enzimas de restrição.

A tradução livre de RFLP significa polimorfismos do comprimento dos fragmentos de restrição e é um marcador molecular baseado em hibridização, ou seja, a sequência de DNA de interesse é encontrada pelo pareamento da mesma com um fragmento de DNA de fita simples (sonda), marcado radioativamente (BOLSTEIN et al., 1980).

Os passos para a execução do marcador baseiam-se na utilização de uma ou mais enzimas de restrição, as quais têm o papel de reconhecer e clivar o DNA em regiões específicas; na separação desses fragmentos de DNA por eletroforese em gel de agarose; na transferência dos fragmentos para uma membrana de náilon ou nitrocelulose; e na hibridização dos fragmentos clivados com uma sonda de sequência de DNA conhecida (BOLSTEIN et al., 1980; BORÉM; FRITSCHÉ-NETO, 2014).

Marcadores robustos para análises genéticas e uso em programas de melhoramento; marcador codominante, fácil reprodutibilidade e alta especificidade da sonda com a sequência de interesse são algumas das vantagens do RFLP. Entre as limitações estão o alto custo, a utilização de sondas radioativas, a pequena quantidade de dados gerada, as dificuldades técnicas e operacionais, e a necessidade de possuir muita habilidade e conhecimento laboratorial para seu desenvolvimento (BORÉM; FRITSCHÉ-NETO, 2014).

3. Marcadores Baseados em PCR

3.1 RAPD (*Random Amplified Polymorphic DNA*)

A análise que utiliza marcadores do tipo RAPD foi citada pela primeira vez na década de 90 por Williams et al. (1990), Welsh & McClelland. (1990) e Caetano-Anollés et al. (1991), com denominações diferentes, mas todas baseiam-se na amplificação de fragmentos aleatórios de DNA genômico obtidos a partir de um único *primer* aleatório, composto por 10 pares de bases. É basicamente uma variação da técnica de PCR em si, utilizada para genotipagem. A técnica é fundamentada na premissa de que um *primer* aleatório se liga a diversas sequências do DNA-alvo, possibilitando assim a amplificação desses fragmentos por meio da reação com a enzima DNA polimerase. Os fragmentos amplificados representam diferentes locos, e podem ser separados por tamanho pela eletroforese. Desta forma, os marcadores RAPD podem ser utilizados para a obtenção de impressões digitais genômicas de indivíduos, variedades ou populações, análise da estrutura e diversidade genética em populações, estabelecimento de relações filogenéticas interespecíficas e construção de mapas genéticos de alta cobertura genômica para localização de genes de interesse econômico.

Devido à plasticidade desse tipo de marcador, sua utilização engloba tanto espécies cultivadas quanto espécies geneticamente desconhecidas e, no âmbito das culturas agrícolas, esse tipo de marcador também pode ser adaptado para identificação de cultivares com diferentes graus de eficiência entre as espécies. Dentre as culturas já pesquisadas para esse fim, estão o algodão, cereais, soja, trigo e rosa. Todavia, este uso da técnica ainda não é comum, uma vez que os *primers* e as bandas geradas na eletroforese são muito variáveis, o que dificulta a reprodutibilidade dos resultados, sendo então necessário um estudo criterioso para estabelecer os protocolos de identificação de cultivares.

Os marcadores RAPD são dominantes, ou seja, os alelos de um mesmo loco são identificados pela presença ou ausência da banda, resultante da amplificação de um fragmento. Todavia, não é possível determinar se este loco está em homozigose ou heterozigose, apenas que há presença de um alelo dominante naquele loco. Outra limitação diz respeito à baixa reprodutibilidade des-

se tipo de marcador (BORÉM; FRITSCHÉ-NETO, 2014), o que dificulta sua aceitação por parte de alguns pesquisadores.

A principal vantagem está na simplicidade da técnica. Por tratar-se de uma adaptação da PCR, não é necessária grande quantidade de material genético e, por utilizar *primers* aleatórios, também não é necessário conhecimento prévio sobre a estrutura genética da espécie estudada (LACERDA et al., 2002). É uma técnica barata, com poucas etapas e da qual provém uma expressiva quantidade de informações.

3.2 AFLP (*Amplified Fragment Length Polymorphism*)

A análise de AFLP foi descrita por Zabeau e Vos (1993) e baseia-se na clivagem do DNA genômico por duas enzimas de restrição, uma de corte raro e outra de corte frequente, gerando fragmentos polimórficos. Uma vez que grande parte das enzimas de restrição produz extremidades com uma só cadeia, chamadas de “caudas”, após a clivagem, é necessário adicionar adaptadores, que são sequências de DNA conhecidas, com aproximadamente 25 pares de base em cada extremidade. Em seguida, são utilizados *primers* complementares aos adaptadores para a amplificação dos fragmentos, por meio da reação de PCR. O último passo da análise é a eletroforese, que permite a visualização dos fragmentos ou bandas no gel. O uso de duas enzimas de restrição faz com que sejam produzidos três tipos de fragmento: os cortados em ambas as extremidades pela enzima de corte frequente, que correspondem a 90% das bandas; os cortados em ambas as extremidades pela enzima de corte raro, e os cortados pelas duas, e o perfil dos fragmentos é analisado para gerar as informações.

A técnica de AFLP pode ser utilizada para o monitoramento de herança de características agrônômicas, análise de *pedigree* e o *screening* de marcadores de DNA ligados a determinadas características genéticas.

Assim como a técnica que utiliza marcadores RAPD, os marcadores AFLP são dominantes, o que impede a distinção de locos em homozigose e heterozigose, e caracteriza a principal desvantagem desta análise. Além disso, é mais onerosa do que o RAPD, uma vez que possui mais etapas e são necessários reagentes adicionais, como os adaptadores e as enzimas de restrição.

Dentre as vantagens, pode-se citar a necessidade de pouca quantidade de amostra de DNA, eficiência devido à capacidade de revelar grande número de bandas polimórficas, rapidez e reprodutibilidade, além de dispensar o desenho dos *primers* específicos (usam-se *primers* complementares aos adaptadores) e o conhecimento prévio da estrutura genética do organismo estudado.

3.3 SSR (*Simple Sequence Repeats*)

Os marcadores de sequência simples repetidas, posteriormente denominados também de microssatélites, foram gerados pela pesquisa de Litt e Luty (1989), são uma classe de sequências de DNA repetitivo que ocorre em todos os organismos, incluindo os eucariontes e os procariontes. Consistem em repetições de sequências, geralmente de um a seis nucleotídeos, ocupando uma extensão de até 100 pares de bases. Os microssatélites são classificados com base em seu tamanho, tipo de unidade de repetição e localização no genoma (GUIMARÃES et al., 2009).

De modo geral, as sequências mais repetidas em plantas são (AT) n (GT) n e (AG) n , e nos animais, a (AC) n . Essas sequências estão distribuídas por todo o genoma e são flanqueadas por sequências altamente conservadas entre os indivíduos de uma espécie. Com isso, permite-se seleção de *primers* específicos que amplificam, via metodologia PCR, os fragmentos contendo o DNA repetitivo em todos os genótipos. As sequências de microssatélites ocorrem tanto nas regiões que codificam quanto naquelas que não codificam, porém suas frequências são maiores nas sequências transcritas (BORÉM; FRITSCHÉ-NETO, 2014).

Acredita-se que as sequências repetidas em *tandem* nos organismos ocorreram devido aos erros de duplicação do DNA, adicionando conjuntos extras de sequências repetidas. Com o passar do tempo, essas sequências variaram em número entre uma cultivar e outra. Como consequência disso, há a produção de números variados das unidades repetidas, fazendo com que cada fragmento de microssatélite assuma comprimentos diferentes que podem ser identificados em eletroforese. Cada um desses fragmentos de um dado locus é equivalente a um alelo de um gene, tornando assim um dos marcadores com maior variabilidade alélica (GUIMARÃES et al., 2009).

Os microssatélites são utilizados em diversas espécies de plantas, com aplicações variadas, como na construção de mapas genéticos, em estudos de diversidade, análise filogenética, *finger-print* e no melhoramento assistido (CAIXETA et al., 2009).

Esses marcadores são amplamente usados, pois são muito mais frequentes e distribuídos ao acaso, permitindo a mais completa cobertura do genoma dos eucariotos e, também, por serem codominantes e multialélicos, fornecendo grande conteúdo informativo genético por locus. Além disso, a técnica é de fácil execução e possui elevada taxa de reprodutibilidade, facilitando a troca de informações entre pesquisadores (CAIXETA et al., 2009).

Uma das desvantagens da técnica é a obtenção dos *primers* específicos que flanqueiam os microssatélites. É laborioso, uma vez que é necessária a construção de bibliotecas genômicas enriquecidas, além do desenho e do sequenciamento dos *primers*. Essa limitação tem sido resolvida com o sequenciamento em larga escala, tecnologia possível pelos avanços nas técnicas moleculares.

O possível acesso a uma grande quantidade de informação, somada à especificidade e à rapidez da tecnologia PCR fazem dos microssatélites uma ferramenta eficiente para estudos em várias espécies de plantas cultivadas.

4. Marcadores Baseados em Sequenciamento

4.1 SNP (*Single Nucleotide Polymorphism*)

O marcador molecular conhecido como SNP, ou Polimorfismo do Nucleotídeo Único, em tradução livre, pertence à classe de marcadores mais utilizados atualmente na comunidade científica. Esta técnica é baseada na alteração na sequência de nucleotídeos que compõem o DNA, os chamados polimorfismos, os quais, neste caso, ocorrem a nível de nucleotídeo único em regiões específicas do genoma em estudo.

Dado que em sequências de DNA, cada unidade de nucleotídeo é composta por um grupo fosfato, uma pentose (ribose ou desoxirribose) e uma base nitrogenada (Adenina, Guanina, Citosina e Tirosina), as alterações detectadas por marcadores do

tipo SNP dão-se nas bases nitrogenadas. Este rearranjo é causado por mutações (inserções ou deleções) entre bases nitrogenadas do mesmo grupo (purinas/purina), processo mais frequente e conhecido como transição, ou mesmo entre bases de grupos distintos (purinas/pirimidinas), este, denominado de transversão.

Inicialmente desenvolvido visando a verificar a precisão do marcador SNP ao mapear um gene de defesa em resposta à presença de um fungo na espécie *Arabidopsis thaliana*, hoje esta técnica é amplamente utilizada em processos de genotipagem, mapeamento genético, resposta à presença de patógenos, filogenia e identificação de diferenças genéticas intra e interespecíficas (CHO et. al., 1999).

O uso da técnica requer o conhecimento prévio da sequência do gene em estudo, de maneira a ser possível desenhar primers e sondas relativos ao mesmo. Em seguida, o processo caminha para a amplificação, clonagem e sequenciamento dos fragmentos de um ou mais genes provenientes de organismos distintos.

Entre as principais vantagens da utilização de SNPs, está sua ampla distribuição no genoma, sendo assim possível identificar diferenças entre os alelos do mesmo gene, entre indivíduos de uma população. Porém, a necessidade de se conhecer previamente a sequência do gene de interesse, assim como o desenho de primers e sondas, demanda tempo, alto custo e mão de obra especializada.

5. Considerações Finais

Atualmente, a integração entre as técnicas convencionais de melhoramento de plantas e a biotecnologia vegetal têm-se tornado imprescindíveis para o avanço da agricultura, pois permitem explorar de maneira mais eficiente a variabilidade genética das espécies, a identificação de genes de interesse agrônômico e a geração de novas cultivares comerciais, em menor espaço de tempo.

Neste sentido, este capítulo abordou, brevemente, os principais marcadores moleculares que podem ser empregados em programas de melhoramento, suas vantagens, aplicações e limitações. No entanto, é interessante ressaltar que as técnicas convencionais aplicadas em programas de melhoramento são inerentes a qualquer outro avanço de tecnologia, sendo assim insubstituíveis para o sucesso do programa.

6. Referências bibliográficas

BOREM, A.; CAIXETA, E. T. **Marcadores moleculares**. 2. ed. Editora Folha de Viçosa, Viçosa-MG, 2009, p. 532.

BOREM, A.; FRITSCH-NETO, R. **Biotechnology and Plant Breeding**. Applications and Approaches for Developing Improved Cultivars. Elsevier, 257 p. 2014.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **Am. J. Hum. Genet.** v. 32, p. 314–331, 1980.

CAETANO-ANOLLÉS, G. et al. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. **Nat Biotech**, v. 9, p. 553–557, 1991.

CHO, R.; MINDRINOS, M.; RICHARDS, D.; SAPOLSKY, R.; ANDERSON, M.; DRENKARD, E.; DEWDNEY, J.; REUBER, T.; STAMMERS, M.; FEDERSPIEL, N., 1999. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. **Nature Genetics** 23, p. 203–207.

GUIMARÃES, C. T.; MAGALHÃES, J. V. de; LANZA, M. A.; SCHUSTER, I. Marcadores moleculares e suas aplicações no melhoramento genético. **Informe Agropecuário**, Belo Horizonte, v. 30, n.253, p.34-33, nov./dez. 2009.

LACERDA, D. R. et al. A técnica de RAPD: uma ferramenta molecular em estudos de conservação de plantas. **Lundiana**, v. 3, n. 2, p. 87-92, 2002.

LITT, M.; LUTY, J.A., 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American Journal of Human Genetics** 44, p. 397–401.

SOUTHERN, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. **J. Mol. Biol.** v. 98, p. 503–517, 1975.

WELSH, J.; McCLELLAND, M. Fingerprinting genomes using PCR with arbitrary primers. **Nucleic Acids Research**, Oxford, v.18, n.24, p.7.213-7.218, 1990.

WILLIAMS, J.G.K. et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, Oxford, v.18, n.22, p.6.531-6.535, 1990.

ZABEAU, M.; VOS, P. Selective restriction fragment amplification: fingerprinting. US 6045994 A. **Google Patents**, 1993. Disponível em <<http://www.google.com/patents/US6045994>>.

Melhoramento Clássico de Alógamas

*Sophia Mangussi Franchi Dutra
Lucas Tadeu Mazza Revolti
Gustavo Hugo Ferreira de Oliveira
Flávia Alves Marques da Silva
Gustavo Vitti Môro*

1. Introdução

Espécies alógamas são aquelas que apresentam mais de 95% de fecundação cruzada (Tabela 1). O melhoramento genético de alógamas pode ter dois objetivos principais: o melhoramento populacional e a obtenção de híbridos. Diante desse cenário, empregam-se métodos a fim de desenvolver e selecionar o genótipo desejado, garantindo sua reprodução para que o agricultor possa cultivá-lo.

Tabela 1. Exemplos das principais culturas alógamas de importância agronômica.

Cereais	Frutífera	Florestal	Agroenergéticas
Azevém	Abacate	Araucária	Cana-de-açúcar
Centeio	Maçã	Aroeira	Canola
Milheto	Mamão	Eucalipto	Girassol
Milho	Uva	Seringueira	Mamona

A primeira etapa de qualquer programa de melhoramento é a formação da população-base, sendo esta de suma importância, pois poderá definir o sucesso ou o fracasso do programa, já que os principais requisitos para que a seleção seja eficiente são a existência de variabilidade genética e a média elevada para as características sob seleção. No melhoramento populacional, o tamanho efetivo da população e a presença da variabilidade genética são de suma importância para a realização da seleção massal

ou da seleção recorrente. Nas plantas alógamas, o híbrido é um genótipo já existente na população onde foi fixado, permitindo, assim, sua reprodução. Desta forma, para que ocorram genótipos superiores na população e para que não se esgote a variabilidade, é necessária a adoção de métodos de melhoramento populacional. Assim, os métodos de melhoramento populacional visam ao aumento da frequência de alelos favoráveis na população.

1.1 Equilíbrio de Hardy-Weinberg

Para o melhor entendimento dos diferentes métodos de melhoramento populacional, principalmente em espécies alógamas, faz-se necessário ter o conhecimento da colaboração das pesquisas feitas de forma independente por G.H. Hardy e W. Weinberg. Conhecido como “Equilíbrio de Hardy-Weinberg” (EHW), o fenômeno elucidado por esses pesquisadores ocorre em populações grandes, com acasalamento ao acaso, e desde que não haja seleção, migração, mutação nem deriva genética, chamando a atenção para o fato de que as frequências alélicas permanecem as mesmas a cada geração de acasalamento ao acaso. Em outras palavras, no campo de produção, uma população em EHW, por manter as mesmas frequências gênicas ao longo das gerações de acasalamento ao acaso, não sofre alteração na média de produtividade, desde que seja mantido o tamanho populacional e que fatores como seleção, mutação e migração, e outros fatores citados anteriormente não atuem.

Fato interessante é que o equilíbrio pode estar presente para cada locus separadamente; por exemplo, se tivermos duas características controladas por um conjunto de loci diferentes e que a mudança em uma característica não afete a outra, a seleção feita para a primeira característica será retirada do equilíbrio, enquanto a segunda característica se manterá no EHW. Assim, de maneira geral, uma população mendeliana está em EHW quando a frequência dos diferentes genótipos, em cada geração, através das gerações de acasalamento ao acaso, estiver de acordo com a expressão $(p+q)^2=p^2+2pq+q^2=1$.

Muitas das teorias e metodologias no melhoramento genético de plantas têm sido formuladas a partir da premissa de que as populações de referência estão em EHW.

1.2 Significado da seleção no melhoramento de espécies alógamas

O principal objetivo do melhoramento populacional em plantas alógamas é o aumento das frequências de alelos favoráveis na população. A taxa de elevação das frequências alélicas pode ser influenciada por diversos fatores como: a influência do ambiente, a variabilidade genética da população, as ligações gênicas, as correlações, tanto fenotípica quanto genotípica, dentre outros. Porém, a seleção é um dos processos que ocorrem em todos os métodos de melhoramento populacional para obtenção de genótipos superiores (Figura 1).

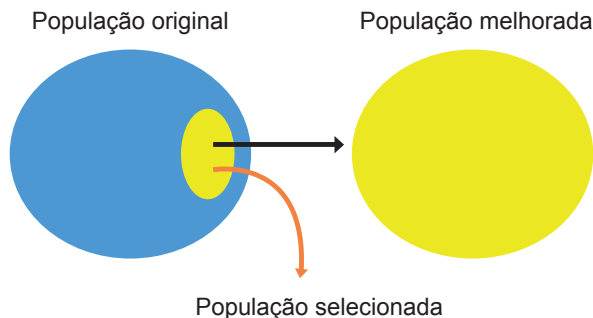


Figura 1. Significado da seleção em espécies alógamas.

Para que haja o melhoramento, os indivíduos selecionados devem dar origem a descendentes superiores, independentemente do método utilizado. Assim, a seleção é feita sobre a população com média alta e ampla variabilidade, com o objetivo principal de obter uma população melhorada (alta frequência de alelos favoráveis para a variável em estudo), superior à população original.

Esse processo de obtenção da população original (população-base), seleção e obtenção da população melhorada pode ser realizado de várias formas, que são conhecidas como métodos de melhoramento populacional. Os diferentes métodos a serem utilizados dependem do tipo de população melhorada que o melhorista objetiva desenvolver, seja ela linhagens, população de polinização aberta, seja mesmo híbrido.

2. Métodos de melhoramento populacional

2.1 Seleção Intrapopulacional

Os métodos intrapopulacionais baseiam-se na seleção de pro-
gênes ou indivíduos em uma única população.

2.1.1 Seleção massal

A seleção massal é um método que visa à escolha das plantas, por meio da avaliação fenotípica, com as melhores características agrônômicas dentro de uma população, sendo recomendada para seleção de características de alta herdabilidade. As sementes oriundas das plantas selecionadas são misturadas e semeadas, produzindo a população melhorada. Tomando-se como exemplo a cultura do milho, as espigas das melhores plantas são selecionadas, e suas sementes são misturadas, sendo utilizada uma amostra dessas sementes para a obtenção da população melhorada.

É importante que se tenha alto grau de conhecimento da cultura que se está trabalhando, para obter-se bons resultados com a seleção massal. Ainda levando em consideração a cultura do milho, as respostas das plantas de uma geração para a outra são reduzidas, sendo necessários vários ciclos de seleção para alcançar resultados almejados.

O controle parental e ambiental é reduzido nesse método de seleção, sendo que geralmente tem-se apenas o controle do progenitor feminino, uma vez que o pólen provém de toda a população. Com relação ao ambiente, não há controle da heterogeneidade do solo e, conseqüentemente, nos locais mais férteis do campo podem ocorrer as melhores plantas.

Buscando-se aumentar a eficiência da seleção massal, para obter-se maior controle ambiental, foram propostas algumas variações para a seleção massal original, denominadas seleção massal estratificada e seleção massal estratificada geneticamente.

Seleção massal estratificada

Este sistema de seleção constitui-se em dividir a área em parcelas ou estratos homogêneos, realizando a colheita e a seleção

em cada estrato separadamente dos demais.

Utilizando a cultura do milho como exemplo, semeia-se um lote isolado com espaçamento recomendado para a cultura. As plantas que apresentam as características de interesse em cada estrato são selecionadas, tomando-se a mesma quantidade de sementes por espiga para obter-se a população melhorada. As plantas que estão ao lado de falhas são descartadas, portanto somente plantas competitivas devem ser escolhidas. A intensidade de seleção varia em função do objetivo do programa de melhoramento, ficando geralmente entre 5 e 20%.

Seleção massal estratificada geneticamente

Neste esquema, o campo também é dividido em estratos como na seleção massal estratificada, porém é semeado um genótipo constante, geralmente um híbrido simples, entre as plantas da população, obtendo-se assim um controle ambiental maior.

O plantio sistemático de um genótipo constante de performance conhecida serve de referência comparativa para a seleção fenotípica das plantas da população. Dessa forma, é importante que o genótipo constante seja macho-estéril, despendoado e/ou possua um gene marcador, para que as plantas selecionadas não sejam polinizadas por este, ou seja, possível identificar as sementes polinizadas.

2.1.2 Seleção com testes de progênies

Neste tipo de seleção, avalia-se o genótipo dos genitores com base no fenótipo de seus descendentes. Este método tem apresentado maior sucesso em caracteres de alta herdabilidade. É mais eficiente que a seleção massal, pois as progênies podem ser avaliadas em ensaios com repetição e em mais de um ambiente, propiciando maior exatidão das médias das plantas selecionadas.

O método consiste na seleção das melhores plantas dentro de uma população de polinização livre. Uma parte das sementes é semeada em fileira, uma linha para cada planta selecionada, para a avaliação das progênies. A outra parte das sementes é guardada, sendo denominado de sementes remanescentes. Após a avaliação das fileiras, as sementes remanescentes que deram

origem às melhores fileiras são misturadas na mesma proporção e semeadas em campo isolado no ano seguinte.

2.1.3 Seleção recorrente intrapopulacional

A seleção recorrente é um método de melhoramento cíclico que visa a aumentar de forma contínua e progressiva a frequência de alelos favoráveis nas populações, mantendo a variabilidade em níveis para permitir ganhos subsequentes com a seleção. A cada ciclo, são obtidos genótipos superiores aos existentes no ciclo anterior, melhorando geneticamente as populações, sem que haja a redução drástica da variabilidade. Os genótipos resultantes podem ser usados como novas variedades comerciais e também como material básico para programas de melhoramento em geral.

O método é composto por quatro fases: obtenção das progênies, avaliação das progênies, seleção de progênies superiores e recombinação das progênies selecionadas. Estas etapas são repetidas sucessivamente, obtendo-se n ciclos seletivos. A recombinação tem por objetivo a manutenção da variabilidade de geração para geração.

A seleção recorrente intrapopulacional visa ao melhoramento da população *per se*, sendo um método importante para espécies alógamas de propagação vegetativa e para gerar novas variedades. Ao final do processo de seleção, obtém-se uma progênie intrapopulacional recombinada e que passa a ser avaliada. Tipo de seleção interessante para caracteres quantitativos por possibilitar maior ocorrência de indivíduos desejados na população com maior efetividade na seleção para os efeitos gênicos aditivos.

2.2 Seleção Interpopulacional

2.2.1 Seleção Recorrente Recíproca (SRR)

A seleção recorrente interpopulacional ou recíproca (SRR) tem por objetivo melhorar a capacidade de combinação entre duas populações, aumentando a resposta heterótica entre elas. Este método utiliza os efeitos genéticos de dominância, dominância parcial e sobredominância. A SRR inicia-se com a utilização de duas po-

pulações básicas. As progênies intrapopulacionais são utilizadas para recombinação, e a interpopulacional é avaliada em ensaios de competição. Em seguida, ocorre a seleção de plantas de cada população antes do florescimento, e essas plantas serão recombinadas para obtenção de progênies interpopulacionais. As plantas selecionadas de cada população devem ser autofecundadas para a obtenção das sementes S_1 que serão utilizadas para recombinação. As progênies inter devem ser avaliadas em experimentos com repetição, e as melhores devem ser selecionadas para produtividade e qualidade. A recombinação das populações, em campo isolado das duas populações, utilizando sementes S_1 , deve ser feita em seguida para a obtenção da população melhorada.

3. Métodos de obtenção de híbridos

Após o estabelecimento da população-base, as plantas são conduzidas à homozigose, dando origem às linhagens que serão combinadas entre si para formar os híbridos, que são heterozigotos. Os genótipos das linhagens são facilmente reproduzidos por meio de autofecundação, e desta forma pode-se reproduzir o híbrido indefinidamente por meio do cruzamento das linhagens parentais.

3.1 Obtenção de linhagens endogâmicas

As linhagens são obtidas por meio de autofecundações sucessivas, de forma artificial. Para isso, protegem-se os órgãos femininos das plantas antes da emissão de estilo-estigma e, no momento apropriado, realiza-se a autopolinização das plantas. Os dois métodos principais, utilizados para obtenção de linhagens, são apresentados a seguir:

Método genealógico ou padrão

O método genealógico, também conhecido como padrão, é o mais utilizado para a obtenção de linhagens com características desejáveis. Primeiramente, realiza-se a autofecundação de uma população-base, e a partir desta etapa são obtidas as sementes

S_1 . Entre as plantas S_1 são selecionadas aquelas com os melhores padrões agronômicos, principalmente para as características de alta herdabilidade. As sementes das plantas selecionadas são semeadas em linhas e autofecundadas e, em seguida, é realizada uma seleção entre e dentro das linhagens. Esse processo continua até as gerações S_6 - S_8 , quando se obtêm genótipos em homozigose.

A vantagem do método genealógico é que se obtêm linhagens com bom padrão agronômico, uma vez que a seleção inicia em S_1 , e o número de linhagens é menor do que o número de plantas inicial. A desvantagem do método é que, dependendo da espécie e da região, como é feita a seleção, pode-se realizar apenas um ciclo por ano, pois as condições de cultivo têm de ser ideais.

Single seed descendent (SSD)

Realiza-se a autofecundação de uma população inicial da qual se obtêm as sementes S_1 . A partir de S_2 , cultiva-se uma semente por linhagem a cada geração de autofecundação, sem que haja seleção artificial, de modo que a homozigose ocorra ao acaso. O método pode sofrer variações, por exemplo, quando ocorre semeadura de uma cova por planta na geração seguinte, para garantir que todas as plantas da população inicial estejam representadas quando a homozigose for atingida.

O método SSD pode gerar maior variabilidade; contudo, com o alcance da homozigose ao acaso, as linhagens finais têm um padrão, que pode não ser o desejado. O método ainda reduz o tempo de obtenção das linhagens, pois não requer condições ideais de cultivo devido à não seleção artificial, podendo-se realizar vários ciclos por ano desde que se garanta a sobrevivência da espécie.

Cruzamento das linhagens

Após o desenvolvimento das linhagens, é necessária a avaliação das mesmas em cruzamentos, para avaliação e seleção do híbrido superior. Os cruzamentos em dialelo e *topcross* são os mais comuns e indicados, porém devido à grande quantidade de cruzamentos possíveis, a partir de poucas linhagens, o melhorista pode optar pelo *topcross*, cruzando todas as linhagens com um testador

comum. Além das informações dos próprios híbridos *topcross*, este tipo de cruzamento permite obter informações sobre a capacidade de combinação das linhagens. Além disso, as linhagens podem ser separadas em grupos heteróticos distintos. A formação de grupos heteróticos é um método de agrupamento baseado nas capacidades de combinação geral e específica dos indivíduos, e definido com base na heterose produzida nas combinações híbridas, de forma que se maximize a heterose entre grupos e aumente a frequência de alelos favoráveis dentro dos grupos.

Com as informações dos *topcross*, realizam-se apenas os cruzamentos entre as linhagens que apresentam as melhores capacidades de combinação e que estão em grupos heteróticos diferentes, viabilizando a obtenção e a avaliação dos híbridos. Nesta etapa, podem ser formados diferentes tipos de híbridos, sendo os principais:

a) Híbrido simples (HS): obtido pelo cruzamento de duas linhagens endogâmicas. É mais produtivo que os demais, apresentando uniformidade de plantas e espigas; no entanto, as sementes possuem custo mais elevado em relação aos outros tipos de híbridos, pois a quantidade de sementes produzida por área é reduzida.

b) Híbrido triplo (HT): resultado do cruzamento de um híbrido simples ($A \times B$) com uma terceira linhagem (C). Possui menor custo de produção que o HS, porém apresenta menor uniformidade e produção de grãos que ambos.

c) Híbrido duplo (HD): resultante do cruzamento de dois híbridos simples ($A \times B$) $\times$ ($C \times D$). É menos produtivo, apresenta maior desuniformidade, melhor estabilidade e exige tecnologia frente aos demais.

Após a obtenção e a avaliação dos híbridos, o melhorista seleciona os superiores, que seguirão para ensaios regionais e de VCU, e aqueles que forem realmente superiores, após essas avaliações, serão lançados comercialmente.

4. Referências bibliográficas

ALLARD, R. W. **Princípios de Melhoramento Genético das Plantas**. Ed. Edgard Blucher. São Paulo. 1971. 381p.

BORÉM, A. **Melhoramento de espécies cultivadas**. -- 2.ed. – Viçosa: Ed. UFV, 2005. 969p.

BORÉM, A.; MIRANDA, G.V. **Melhoramento de Plantas**. 5. ed. ver. e ampl. –Viçosa: Ed. UFV, 2009, 529 p.

BUENO, L. C. S.; MENDES, A. N. G.; CARVALHO, S. P. **Melhoramento genético de plantas**: princípios e procedimentos. -- 2.ed. – Lavras: Ed. UFLA, 2006. 319p.

PATERNIANI, E. Métodos tradicionais de melhoramento de milho. In: BULL, L. T.; CANTARELLA, H. (Ed.). **Cultura do milho**: fatores que afetam a produtividade. Piracicaba: Potafos, 1993. p. 22-43.

RAMALHO, M. A. P.; SANTOS, J. B.; PINTO, C. A. B. P.; SOUZA, E. A.; GONÇALVES, F. M. A.; SOUZA, J. C. **Genética na agropecuária**. -- 5.ed. – Lavras: Ed. UFLA, 2012. 566p.

Princípios e aplicações da PCR em Tempo Real Quantitativa e LAMP (Loop Mediated Isothermal Amplification)

*Prof. Dr. Marcos Rogério André
Profa. Dra. Rosângela Zacarias Machado
Doutorando Otávio Luiz Fidelis Junior
Doutorando Paulo Henrique Sampaio
Mestrando Luiz Ricardo Gonçalves
Mestrando Jyan Lucas Benevenuto*

1. Bases da PCR em Tempo Real Quantitativa

A PCR em tempo real quantitativa é uma ferramenta utilizada em muitos laboratórios, em aplicações voltadas à investigação científica básica e no diagnóstico direto de material genético. Esta tecnologia une o princípio da Reação em Cadeia pela Polimerase (PCR) com o monitoramento de emissão de fluorescência emitido por moléculas adicionadas nas reações, permitindo a visualização da polimerização de produtos de PCR, gerados durante a reação (NAVARRO et al., 2015; PABINGER et al., 2014).

A cada ciclo de amplificação, ocorre um acréscimo linear na emissão de fluorescência, conforme produtos de PCR são produzidos. A correlação entre a quantidade de produtos de PCR gerados a cada ciclo e a intensidade de fluorescência emitida é utilizada para calcular a quantidade de DNA presente no início da reação, levando em consideração o número de ciclos necessários para que estes produtos sejam detectados pela primeira vez (BUSTIN e NOLAN, 2004).

Os níveis de fluorescência emitidos em cada amostra podem variar devido a alterações no meio em que a reação ocorre, criando um “sinal de fundo” (*background signal*). Este sinal de fundo

é mais evidente durante os primeiros ciclos das reações, antes que ocorra acúmulo significativo do produto-alvo. A intensidade de “sinal de fundo” emitido é utilizada para a determinação de um limiar de fluorescência. Este limiar, ou linha basal (*Baseline Threshold*), é calculado através de uma plataforma algorítmica específica, determinando o ponto em que a fluorescência emitida por amostra excede a linha basal. O ciclo em que este ponto ocorre é chamado de Ciclo de Quantificação (*Cq*), que é inversamente relacionado ao logaritmo do número inicial de cópias. O valor de *Cq* reflete, portanto, o ponto em que ocorre acúmulo de produtos de PCR, suficiente para a detecção do DNA-alvo (BUSTIN; NOLAN, 2004).

2. Sistemas de Detecção

2.1 Intercalantes de DNA

As moléculas mais comumente utilizadas para este método são o SYBR Green I e Eva Green, que se ligam de forma reversível, porém forte, às moléculas de DNA de cadeia dupla por alocação do cromóforo entre os pares de bases empilhados. Estas moléculas possibilitam o monitoramento da amplificação de produtos de PCR a cada ciclo, através da leitura da absorbância de energia luminosa (fótons). Esta energia, emitida por fluoróforos, dá-se através da excitação destes por absorção de comprimentos de ondas específicos, emitidos por *lasers* ou lâmpadas halógenas acopladas ao termociclador, bem como um fluorímetro para leitura da emissão de fluorescência. Desta forma, o acréscimo de fluorescência é medido ao final de cada ciclo de PCR, quando ocorre a formação de DNA de cadeia dupla (BUSTIN; NOLAN, 2004).

As principais vantagens deste método são a incorporação destes fluoróforos em protocolos previamente estabelecidos, com oligonucleotídeos iniciadores comuns. Outra vantagem é o menor custo em relação ao uso de sondas de hidrólise, porém apresenta menor especificidade quando comparado a esta. Esta inespecificidade ocorre devido à possibilidade de ligação dos intercalantes em cadeias simples de DNA e também de ligação in-

discriminada em qualquer outra cadeia dupla gerada durante a reação, por exemplo: a emissão de fluorescência inespecífica gerada por ligação aos dímeros de *primers* (BUSTIN; NOLAN, 2004).

Uma forma de aumentar a especificidade das reações é a utilização de curva de dissociação ou curva de “*melting*”. Este processo tem em vista a diferenciação de produtos específicos dos não específicos, gerados após o término da reação, através de uma etapa adicional em que ocorre desnaturação gradativa das moléculas de DNA de cadeia dupla, conforme o aumento gradativo na temperatura (50°C a 95°C).

Quando ocorre a desnaturação do DNA, ocorre a dissociação dos intercalantes no meio, resultando em uma queda na emissão de fluorescência, até que permaneçam apenas moléculas de DNA de cadeia simples na reação, gerando curvas de dissociação para cada amostra. Estas curvas de dissociação individuais são então convertidas em picos distintos, representando graficamente a primeira derivada negativa da fluorescência em função da temperatura.

Diferentes curvas de dissociação são observadas quando ocorre a amplificação de fragmentos inespecíficos. Isto ocorre devido à maior dificuldade de rompimento das três pontes de hidrogênio geradas entre as bases Guanina e Citosina, quando comparadas à dupla ligação entre Adenina e Timina. Portanto, levando em consideração a especificidade da sequência de nucleotídeos em cada cadeia, o tempo necessário para a desnaturação também é diferente, gerando picos distintos (BUSTIN; NOLAN, 2004).

2.2 Sondas de Hidrólise (Sistema TaqMan®)

O sistema de detecção, popularmente conhecido como *TaqMan*, utiliza a atividade exonucleásica da enzima *Taq* polimerase como chave para o monitoramento dos produtos de PCR gerados a cada ciclo. Esta técnica tem como base a utilização de oligonucleotídeos, ou sondas de hidrólise, duplamente marcadas com: um fluoróforo “repórter” na extremidade 5’ e um *quencher*, ou abafador, na extremidade 3’(sendo TAMRA o mais comumente utilizado). Semelhante aos fluoróforos intercalantes de DNA, a emissão de fluorescência do repórter dá-se através da excitação deste por absorção de comprimentos de ondas específicos e

leitura por fluorímetro, acoplado ao equipamento termociclador (KIM et al., 2013).

Em uma sonda *TaqMan* intacta, a fluorescência emitida pelo repórter acoplado à extremidade 5' é absorvida pelo *quencher* localizado na extremidade 3'. Isto ocorre através de transferência de energia de ressonância de fluorescência (FRET), fazendo com que a intensidade de fluorescência emitida permaneça abaixo da linha basal (KIM et al., 2013; BUSTIN; NOLAN, 2004).

Mais recentemente, *quenchers* escuros estão substituindo os fluorescentes em reações de PCR em Tempo Real, liberando, assim, o TAMRA como repórter nas análises multiplex, reações estas em que são utilizadas mais de uma sonda de hidrólise, com fluoróforos que emitem comprimentos de onda distintos, para detecção de diferentes genes (BUSTIN; NOLAN, 2004).

Durante a etapa de polimerização, todas as ligações de nucleotídeos à cadeia-base serão quebradas. Isto ocorre devido à atividade exonucleásica da enzima *Taq* polimerase. Desta forma, a sonda de hidrólise é clivada gradativamente a partir da extremidade 5', separando o fluoróforo de seu abafador, permitindo que esta fluorescência possa então ser detectada pelo equipamento acoplado ao termociclador. Assim, a geração de produtos de PCR é detectada pelo monitoramento da fluorescência, gerado a cada ciclo, conforme ocorre a clivagem das sondas de hidrólise, utilizadas na reação (Bustin; Nolan, 2004).

Este sistema apresenta maior especificidade quando comparado ao sistema de intercalantes de DNA, devido à utilização de três oligonucleotídeos que deverão ser complementares ao DNA molde. Entretanto, a necessidade de utilização de uma sonda complica o processo de concepção dos oligonucleotídeos que serão utilizados em conjunto, devendo ser específicos para a sequência de interesse. É necessário, ainda, salientar que o comprimento do gene-alvo para as reações de qPCR é limitado, devendo conter menos que 500 pares de bases. Tais dificuldades podem levar à concepção e à utilização de conjuntos, iniciador e sonda, falhos, o que pode levar à produção de resultados distorcidos (BUSTIN; NOLAN, 2004).

Portanto, a escolha correta da sequência de nucleotídeos utilizados na sonda de hidrólise (compatível com a temperatura de anelamento dos iniciadores) e de fluoróforos resulta em um per-

feito anelamento e, conseqüentemente, em eficiente clivagem e liberação do repórter durante cada ciclo, possibilitando a confiabilidade dos resultados obtidos por este sistema de detecção e a criação de protocolos que possibilitem a detecção, com eficiência, de produtos de PCR distintos, como no caso de reações multiplex (BUSTIN; NOLAN, 2004; NAVARRO et al., 2015).

PCR em tempo real quantitativa relativa

A qPCR pode ser dividida em dois métodos de quantificação:

- **Quantificação absoluta** → feita através da comparação dos valores do Cq das amostras testadas com uma curva-padrão. O resultado da análise é quantidade de ácido nucleico (número de cópias, µg) por dada quantidade de amostra.
- **Quantificação relativa** → o resultado da análise é uma relação: a quantidade relativa (diferença vezes) do gene-alvo, comparada à quantidade de um gene de referência (altamente preservado). O gene de referência deve ter sua expressão constante dentre as amostras testadas e não deve ser influenciado pelo tratamento experimental em estudo.

Neste capítulo, abranger-se-ão apenas os métodos de quantificação relativa. Para facilitar o entendimento de cada método, utilizou-se de um exemplo prático. Seu experimento visa a verificar se existe diferença entre o nível de expressão do fator de necrose tumoral alfa (TNFα) entre células ovarianas cancerosas e normais, e, caso sim, o quanto. Esta comparação pode ser feita de duas formas:

- (1) Pode-se extrair o RNA de 1.000 células cancerígenas e 1.000 células normais e determinar a quantidade de TNFα das amostras, realizando uma RT-qPCR específica para TNFα. Descobrir-se-á, então, a diferença relativa de TNFα, através da razão entre o número equivalente de células ovarianas cancerosas e normais.
- (2) Ou então, por exemplo, verificar a expressão de um gene de referência altamente conservado entre as amostras, por exemplo, o gliceraldeído-3-fosfato desidrogenase (GAPDH), cuja expressão deve ser semelhante em células cancerígenas e normais. Neste caso, amostras de RNA podem ser

extraídas a partir de quantidades aproximadas de células cancerosas e normais, sem a necessidade de se saber o número exato de células. Após a extração, os níveis de expressão de TNF α e de GAPDH das amostras são determinados através da RT-qPCR. A expressão relativa de TNF α de células cancerosas VS. normais é determinada tomando a razão entre a expressão de TNF α *normalizado* pelo GAPDH nos dois tipos de células.

A diferença entre os dois métodos de análise descritos acima é o tipo de normalizador utilizado. No primeiro cenário, o normalizador é o número de células. No segundo, o normalizador é o nível de expressão de um gene de referência (no caso, GAPDH). As manipulações matemáticas para os dois cenários serão apresentadas a seguir. Em ambos os casos, o resultado final é uma descrição quantitativa de uma determinada propriedade de uma amostra em relação à de outra amostra.

3.1 Quantificação relativa *normalizada pela unidade de massa* (Exemplo 1)

A vantagem da utilização de uma unidade de massa (tal como o número de células ou de μg de ácido nucleico), ao invés de um gene de referência como normalizador, é que o desenho experimental é conceitualmente simples, assim como o tratamento matemático.

O método requer a quantificação precisa do material de partida, independentemente se o número de células ou μg de ácido nucleico for utilizado como normalizador. Quando se comparam amostras múltiplas através da quantificação relativa, uma amostra é geralmente escolhida como calibrador (amostra-controle), e a expressão do gene-alvo (ex. TNF α), em todas as outras amostras, é expressa como aumentado ou diminuído em relação ao calibrador. Usualmente, a amostra do grupo não tratado ou sadio é escolhida como calibrador. No caso do exemplo, as células ovarianas normais seriam escolhidas como calibrador, e as cancerosas, como amostra-teste.

Para determinar a expressão relativa de TNF α em células cancerosas de ovário VS. normais, amostras de RNA total foram extraídas a partir de um número igual de células ovarianas normais e

cancerosas. Os valores de Cq para a amostra-teste (células cancerosas) e a amostra-calibrador (células normais) foram então usados para calcular a relação entre os dois, através da seguinte equação:

$$\text{Relação (teste/calibrador)} = E^{\text{Cq(calibrador)} - \text{Cq(teste)}}$$

E → Eficiência do teste – Assumindo-se que a eficiência da reação seja 100%, o E = 2, então:

$$\text{Relação (teste/calibrador)} = 2^{\text{Cq(calibrador)} - \text{Cq(teste)}} \text{ ou}$$

$$\text{Relação (teste/calibrador)} = 2^{\Delta\text{Cq}} \text{ em que, } \Delta\text{Cq} = \text{Cq(calibrador)} - \text{Cq(teste)}$$

3.2 Quantificação relativa *normalizada pelo gene de referência* (Exemplo 2)

Este tipo de normalização tem a vantagem de não necessitar de uma quantificação exata do material de partida, sendo vantajoso em casos em que o material de partida é limitado. A desvantagem desta técnica deve-se ao fato de requerer um gene de referência com expressão constante em todas as amostras testadas e cuja expressão não é alterada pelo tratamento. Recentemente, foi proposto que, na maioria dos casos, a utilização de múltiplos genes de referência é necessária para a quantificação, a fim de se verificar qual deles ou a associação entre eles melhor se adapta ao experimento. A verificação de qual ou de quais genes de referência devem ser utilizados, pode ser realizada através de softwares específicos (geNorm, NormFinder e/ou BestKeeper).

Para determinar a expressão relativa de um gene-alvo (ex. TNFα) da amostra-teste e amostra calibrador usando gene(s) de referência como o normalizador(es), os níveis de expressão do(s) gene(s) de referência e do alvo devem ser determinados utilizando-se, para tanto, da RT-qPCR. Em suma, é preciso determinar os valores do Cq, como mostra a Tabela 1.

Tabela 1. Valores Cq necessários para a quantificação relativa

	Teste	Calibrador (cal)
Gene-alvo	Cq (alvo, teste)	Cq (alvo, cal)
Gene de referência (ref)	Cq (ref, teste)	Cq (ref, cal)

Após os Cqs serem obtidos, existem três diferentes métodos que podem ser utilizados para determinar o nível de expressão do gene-alvo da amostra-teste em relação à amostra-calibrador: 1) **Método de Livak**, também conhecido como **método $2^{-\Delta\Delta Cq}$** ; 2) **Método de ΔCq** ; e 3) **Método de Pfaffl**.

Método de Livak (Método $2^{-\Delta\Delta Cq}$)

Este método é amplamente utilizado e de fácil execução. Ele assume que tanto os genes-alvo quanto os de referência são amplificados com eficiências próximas de $100\% \pm 5\%$. Após estabelecer que os genes-alvo e de referência possuem eficiências de amplificação semelhantes e próximas de 100%, pode-se determinar a diferença relativa do gene-alvo em diferentes amostras, utilizando os seguintes passos:

Primeiramente, normalizar o Cq do gene-alvo através do gene de referência (ref), tanto para a amostra-teste quanto para a amostra-calibrador:

$$\begin{aligned}\Delta Cq(\text{teste}) &= Cq(\text{alvo, teste}) - Cq(\text{ref, teste}) \\ \Delta Cq(\text{calibrador}) &= Cq(\text{alvo, calibrador}) - Cq(\text{ref, calibrador})\end{aligned}$$

Posteriormente, normalizar o ΔCq da amostra-teste pelo ΔCq da amostra-calibrador:

$$\Delta\Delta Cq = \Delta Cq(\text{teste}) - \Delta Cq(\text{calibrador})$$

Por último, a relação entre ambos:

$$2^{-\Delta\Delta Cq} = \text{Relação normalizada}$$

O resultado obtido nos remete a quantas vezes o gene-alvo está aumentado (ou diminuído) na amostra-teste em relação ao calibrador, sendo normalizado pela expressão do gene de referência. Se os genes de referência e alvo não possuem eficiências de amplificação semelhantes, pode-se otimizar ou redesenhar os ensaios ou, então, utilizar o método Pfaffl. Por outro lado, se os genes de referência e alvo apresentam eficiência de amplificação idênticas, mas a eficiência não é igual a 2 (100%), uma forma modificada do método $2^{-\Delta\Delta Cq}$ pode ser utilizada, substituindo o 2 na equação pela eficiência de amplificação real.

Método do ΔCq

Este método é uma variação do método de Livak, sendo mais simples de calcular e apresentando o mesmo resultado. Este método utiliza a diferença de valores de Cq entre os genes de referência e alvo para cada amostra. Os valores de expressão resultantes obtidos neste método são divididos pelo valor do calibrador escolhido. Os resultados deste cálculo são exatamente os mesmos que os obtidos com o método $2^{-\Delta\Delta Cq}$:

$$\text{Expressão da amostra-calibrador (referência/alvo)} = 2^{Cq(\text{referência}) - Cq(\text{alvo})}$$

$$\text{Expressão da amostra-teste (referência/alvo)} = 2^{Cq(\text{referência}) - Cq(\text{alvo})}$$

$$\text{Relação normalizada} = \text{Expressão da amostra-teste} / \text{Expressão da amostra-calibrador}$$

Os pressupostos matemáticos para esta abordagem são os mesmos que aqueles para o método $2^{-\Delta\Delta Cq}$.

Método de Pfaffl

O método $2^{-\Delta\Delta Cq}$ (Livak) para calcular a expressão gênica relativa só é válido quando as eficiências de amplificação dos genes-alvo e de referência forem semelhantes. Caso os dois fragmentos amplificados apresentem eficiências diferentes, uma fórmula alternativa deve ser utilizada. Para determinar a taxa de expressão entre a amostra-teste e o calibrador, deve-se utilizar a seguinte fórmula:

$$\text{Relação normalizada} = \frac{(E_{\text{teste}})^{\Delta Cq, \text{teste (referência - alvo)}}}{(E_{\text{calibrador}})^{\Delta Cq, \text{calibrador (referência - alvo)}}$$

Na equação acima, a E_{teste} e $E_{\text{calibrador}}$ são as eficiências de amplificação das amostras-teste e controle, respectivamente. $\Delta Cq, \text{teste (calibrador - teste)}$ = Cq do gene de referência menos o Cq do gene-alvo na amostra-teste; e o $\Delta Cq, \text{calibrador (calibrador - teste)}$ é o Cq do gene de referência menos o Cq do gene-alvo da amostra-calibrador. A equação acima assume que cada gene

(alvo e referência) apresente a mesma eficiência de amplificação nas amostras, mas não necessariamente que as amostras-teste e calibrador possuem a mesma eficiência de amplificação.

4. PCR em tempo real quantitativa absoluta

Ao contrário da PCR convencional (cPCR), em que o produto amplificado é submetido à eletroforese em gel de agarose para visualização do resultado, a qPCR pode mensurar a amplificação de DNA durante o progresso da reação, ou seja, em tempo real. A detecção em tempo real só é possível pela adição de uma molécula fluorescente que reporta o aumento na quantidade de DNA por meio do aumento proporcional do sinal de fluorescência. Sendo assim, a principal vantagem da qPCR sobre a cPCR é sua capacidade de poder determinar a quantidade de *template* no início da reação, promovendo, com alta acurácia, precisão e sensibilidade à quantificação do número de cópias de DNA-alvo.

A simplicidade conceitual e prática da qPCR, aliada a sua rapidez, alta sensibilidade e especificidade, faz com que essa ferramenta seja a técnica “padrão-ouro” para a quantificação de ácidos nucleicos. Essa popularidade é refletida em um alto número de publicações, reportando os dados de qPCR. Porém, a falta de consenso na realização, na análise e na interpretação dos dados quantitativos obtidos por experimentos, utilizando a qPCR, pode levar a erros graves, dificultando e, às vezes, impossibilitando a análise da qualidade dos trabalhos.

Nesse sentido, Bustin et al. (2009), em seu artigo “*The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments*”, descrevem importantes informações com o objetivo de aumentar a confiabilidade dos resultados, ajudando a garantir a integridade da literatura científica, promovendo a consistência entre laboratórios e o aumento da transparência experimental.

Em princípio, a quantificação de ácidos nucleicos por meio da qPCR não mostra dificuldades. Quanto maior o número de cópias de ácidos nucleicos no início da reação, menor é a quantidade de ciclos requeridos para gerar um número específico de produto. Portanto, o número de ciclos requeridos para a fluorescência chegar a um específico limiar de detecção (*threshold*) é inversa-

mente correlacionado com a quantidade de ácidos nucleicos na amostra. A conversão do ciclo de quantificação (Cq) para a quantidade de ácidos nucleicos na amostra é obtida pela interpolação dos valores de Cq da amostra-teste, versus os valores de Cq da curva-padrão obtidos por meio de uma diluição seriada de DNA, para a qual a quantidade de *template* inicial é previamente conhecida.

Um dos principais entraves na análise de dados quantitativos é a obtenção de dados imprecisos ou pouco acurados. É necessário ter em mente que: (1) um resultado pode ser preciso, mas, ao mesmo tempo, precisamente errado; (2) uma quantificação precisa não é a mesma coisa que uma quantificação acurada; (3) dados altamente acurados são difíceis de serem obtidos. Enquanto a precisão se refere à reprodutibilidade e à repetibilidade, a acurácia é a medida de concordância entre os valores experimentalmente obtidos e o valor real, o qual provavelmente é desconhecido.

A qPCR possui uma amplificação exponencial de moléculas-alvo. Assumindo que a eficiência de amplificação (E) é de 100%, cada molécula de DNA será duplicada a cada ciclo. Desta forma, é possível extrapolar o número de moléculas que havia no início da reação, utilizando a fórmula ($N_n = N_0 \times 2^n$). Entretanto, na vida real, uma E de 100% dificilmente é obtida, e um modelo matemático mais apropriado e que leva em conta a E pode ser baseado na equação: $N_n = N_0(1+E)$.

Outra propriedade importante da qPCR é a influência do número de cópias de DNA no início da reação. A obtenção de dados quantitativos a partir de amostras contendo 1.000 cópias de DNA-alvo no início da reação é muito acurada. Por outro lado, amostras que contenham número de cópias inicial menor que 1.000, fazem com que a estimativa do número de moléculas-alvo seja dificultada. Uma molécula de DNA, ao final de um ciclo de amplificação, pode ser duplicada ou, ainda, permanecer como uma única molécula. De fato, se a E é de 0,8, há a probabilidade de 20% que, após um ciclo de amplificação, ela ainda permaneça como uma simples cópia. Essa incerteza de amplificação é correlacionada com a E : Uma amostra contendo 100 cópias de DNA-alvo no início da reação possui a incerteza de 10% se a E for de 0,9, mas essa incerteza aumenta dramaticamente para 25% se a

E for de 0,5.

Quando amostras contendo baixo número de cópias de DNA *template* no início da reação são mesuradas por qPCR, pode ser observado alta diferença no número de cópias entre as réplicas. Essa diferença no número de cópias entre as réplicas pode ser explicada pelo “Efeito de Monte Carlo”, uma inerente limitação da qPCR quando amostras contendo baixa concentração de DNA são analisadas.

O cerne da quantificação de ácidos nucleicos por qPCR está pautado na eficiência de amplificação. De forma simplificada, a quantificação absoluta refere-se em comparar uma amostra desconhecida com um padrão externo (curva-padrão), cuja concentração de DNA é previamente conhecida. A curva-padrão pode ser preparada usando DNA genômico, Gblocks, DNA plasmidial, RNAs, etc., cada um com suas vantagens e desvantagens. Além de ser um passo crucial para a realização da qPCR, a curva-padrão fornece informações vitais que asseguram a qualidade da reação. Dentre essas, os resultados da eficiência de amplificação (E), o quadrado do coeficiente de regressão (R_2), intercepto do eixo Y (y -intercepto) e o *Slope*, mostram o quão acurado e sensível é seu teste. Os parâmetros para cada um dos indicadores ($E = >90$ ou $<105\%$; [$R_2 = >0.980$]; [$slope = -3.32$]) devem ser verificados e respeitados, a fim de se obter resultados acurados.

Sendo assim, é possível, por exemplo, determinar o número de partículas virais na amostra de um paciente infectado ou avaliar o potencial de diferentes princípios ativos frente à infecção por bactérias, protozoários, helmintos, etc., entre pacientes humanos e animais. De posse dessas informações, pode-se avaliar a carga parasitária durante o curso da infecção e identificar qual princípio ativo possui melhor desempenho no tratamento de diferentes enfermidades.

5. LAMP (Loop Mediated Isothermal Amplification)

A amplificação circular isotérmica de DNA (LAMP) é um método desenvolvido para amplificar DNA com alta especificidade, eficiência e rapidez em condições isotérmicas, fato que possibilita a não utilização de termocicladores para sua execução (NOTOMI et al., 2000).

Inicialmente, este método utilizava um conjunto de quatro oligonucleotídeos iniciadores, os quais reconhecem seis regiões distintas do DNA-alvo, sendo estes: os externos sense e antissense, e os internos sense e antissense. Estes últimos são, na verdade, oligonucleotídeos complementares a duas regiões distintas do DNA-alvo (NOTOMI et al., 2000). Posteriormente, foi verificado que a adição de dois novos oligonucleotídeos iniciadores poderia acelerar a reação, os quais foram denominados *loop primers* (NAGAMINE et al., 2002).

A enzima utilizada neste método, originalmente isolada do *Bacillus stearothermophilus* e denominada BST polimerase, possui atividade 5' → 3' polimerase, porém não possui atividade 5' → 3' e 3' → 5' exonuclease, o que permite a utilização da mesma em uma reação isothermal (NEW ENGLAND, BIOLABS, 2016). Portanto, não é necessária a utilização de três temperaturas distintas, como classicamente empregada nos protocolos de cPCR, a citar: temperaturas de desnaturação, anelamento e extensão.

A LAMP possui alta eficiência em amplificar o DNA-alvo. É capaz de detectar baixo número de cópias de uma amostra de DNA-alvo, além de mostrar-se altamente específica (NOTOMI et al., 2000). Este método sofre menor interferência de inibidores de reação de amplificação de DNA quando comparada ao cPCR (KANEKO et al., 2007). Para sua realização, não é necessária a utilização de equipamentos sofisticados. O custo por reação é muito baixo (aproximadamente 1 dólar) (GRAB et al., 2011).

A visualização dos resultados da LAMP pode ser realizada por diferentes métodos. Inicialmente, a visualização dos resultados era baseada na eletroforese em gel de agarose contendo intercalantes de DNA (NOTOMI et al., 2000). Logo em seguida, perceberam que a turbidez formada pela presença de pirofosfato de magnésio, composto este formado durante a amplificação de DNA, poderia ser um método de visualização dos resultados (MORI et al., 2001). Ainda, a aquisição dos resultados pode ser realizada em aparelhos de PCR em tempo real equipado com turbidômetro ou a olho nu, por meio da observação de turbidez nas amostras positivas, logo após a reação ser finalizada.

Posteriormente, outros métodos de revelação foram desenvolvidos, tais como a utilização de calceína, a qual inicialmente se liga aos íons de manganês (Mn^{+2}), ficando “abafada” sem emitir

fluorescência; durante a reação, devido à formação de pirofosfato, a calceína perde o Mn^{+2} e emite fluorescência, que por sua vez é amplificada pela entrada de Mg^{+2} no lugar antes ocupado pelo Mn^{+2} (TOMITA et al., 2008). Neste mesmo sentido, propuseram a utilização de probes marcadas com fluoróforos, para visualizar o resultado da reação a olho nu (MORI et al., 2006). Adicionalmente, pode-se adicionar SYBR Green I na amostra logo após o término da reação. Nesta condição, as amostras positivas mudam a cor para um tom de verde, possibilitando a visualização a olho nu, sem necessidade de equipamentos (NJIRU et al., 2008). Ainda, pode-se utilizar azul de hidroxinaftol para a revelação da reação. Neste caso, o produto da reação adquire cor azulada nas amostras positivas, ao passo que amostras negativas adquirem coloração roxa. Nesta última variação de revelação da reação, não se faz necessária a utilização de equipamentos sofisticados (GOTO et al., 2009), possibilitando a realização desta técnica em locais sem grandes recursos financeiros.

Inicialmente, a LAMP foi criada visando à detecção de RNA, requisitando para tal a adição de transcriptase reversa (NOTOMI et al., 2000), sendo descritos diversos protocolos de RT-LAMP (BUATES et al., 2010). Alguns anos após a criação deste método de amplificação de DNA, alguns pesquisadores adaptaram a técnica para a detecção de proteínas específicas com alta eficiência, mostrando novos rumos à utilização deste método de diagnóstico (GRAB et al., 2005; BURBULIS et al., 2015).

6. Referências bibliográficas

BIO-RAD LABORATORIES. **Real-Time PCR Applications Guide**. 2006, 105p.

BUATES, S.; BANTUCHAI, S.; SATTABONGKOT, J.; HAN, E. T.; TSUBOI, T.; UDOMSANGPETCH, R.; SIRICHAISINTHOP, J.; TAN-ARIYA, P. Development of a reverse transcription-loop-mediated isothermal amplification (RT-LAMP) for clinical detection of *Plasmodium falciparum* gametocytes. **Parasitology International**, Amsterdam, v. 59, n. 3, p. 414-420, 2010.

BURBULIS, I. E.; YAMAGUCHI, K.; NIKOLSKAIA, O. V.; PRIGGE, S. T.; MAGEZ, S.; BISSER, S.; RELLER, M. E.; GRAB, D. J. Detection of Pathogen-Specific Antibodies by Loop-Mediated Isothermal Amplification. **Clinical**

and Vaccine Immunology, Washington, v. 22, n. 4, p. 374-380, 2015.

BUSTIN S. A. **Why the need for qPCR publication guidelines?—The case for MIQE**. *Methods*. n.50, p. 217-226, 2010.

BUSTIN, S. A.; BENES, V.; GARSON, J. A.; HELLEMANS, J.; HUGGETT, J.; UBISTA, M.; MUELLER, R.; NOLAN, T.; PFAFFL, M. W.; SHIPLEY, G. L.; VANDESOMPELE, J.O.; WITTEWER, CARL T. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. **Clinical Chemistry**, v.44, n.4, p. 611-622, 2009.

BUSTIN, S. A.; NOLAN, T. In: BUSTIN, S. A. **A-Z Quantitative PCR**. 2004, 1. ed. p. 218-268.

GOTO, M.; HONDA, E.; OGURA, A.; NOMOTO, A.; HANAKI, K. I. Colorimetric detection of loop-mediated isothermal amplification reaction by using hydroxy naphthol blue. **Biotechniques**, London, v. 46, n. 3, p. 167-172, 2009.

GRAB, D. J.; LONSDALE-ECCLES, J.; INOUE, N. LAMP for tadpoles. **Nature Methods**, New York, v. 2, n. 9, p. 635-635, 2005.

GRAB, D. J.; NIKOLSKAIA, O. V.; INOUE, N.; THEKISOE, O. M.; MORRISON, L. J.; GIBSON, W.; DUMLER, J. S. Using detergent to enhance detection sensitivity of African trypanosomes in human CSF and blood by loop-mediated isothermal amplification (LAMP). **PLoS Neglected Tropical Diseases**, San Francisco, v. 5, n. 8, p. e1.249, 2011.

KANEKO, H.; KAWANA, T.; FUKUSHIMA, E.; SUZUTANI, T. Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. **Journal of Biochemical and Biophysical Methods**, Amsterdam, v. 70, n. 3, p. 499-501, 2007.

KIM, J.; LIM, J.; LEE, C. Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. **Biotechnology Advances**, v.31, p. 1.358-1.373, 2013.

MORI, Y.; HIRANO, T.; NOTOMI, T. Sequence specific visual detection of LAMP reactions by addition of cationic polymers. **BMC Biotechnology**, London, v. 6, n. 1, p. 3, 2006.

MORI, Y.; NAGAMINE, K.; TOMITA, N.; NOTOMI, T. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. **Biochemical and Biophysical Research Communications**, New York, v. 289, n. 1, p. 150-154, 2001.

NAGAMINE, K.; HASE, T.; NOTOMI, T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. **Molecular and cel-**

lular probes, London, v. 16, n. 3, p. 223-229, 2002.

NAVARRO, E.; SERRANO-HERAS, G. A.; CASTAÑO, M. J. A.; SOLERA, J. B. Real-time PCR detection chemistry. **Clinica Chimica Acta**, v.439, p. 231-250, 2015.

NEW ENGLAND, BIOLABS. Bst DNA Polymerase, Large Fragment. Disponível em: < <https://www.neb.com/products/m0275-bst-dna-polymerase-large-fragment> >. Acesso em: 27 jan. 2016, 2016.

NJIRU, Z. K.; MIKOSZA, A. S.; ARMSTRONG, T.; ENYARU, J. C.; NDUNG'U, J. M.; THOMPSON, A. R. Loop-mediated isothermal amplification (LAMP) method for rapid detection of *Trypanosoma brucei rhodesiense*. **PLoS Neglected Tropical Diseases**, San Francisco, v. 2, n. 2, p. e147, 2008.

NOTOMI, T.; OKAYAMA, H.; MASUBUCHI, H.; YONEKAWA, T.; WATANABE, K.; AMINO, N.; HASE, T. Loop-mediated isothermal amplification of DNA. **Nucleic Acids Research**, London, v. 28, n. 12, p. e63, 2000.

PABINGER, S.; RÖDIGER, S.; KRIEGNER, A.; VIERLINGER, K.; WEINHÄUSEL, A. A survey of tools for the analysis of quantitative PCR (qPCR) data. **Biomolecular Detection and Quantification**, v.1, p. 23-33, 2014.

TOMITA, N.; MORI, Y.; KANDA, H.; NOTOMI, T. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. **Nature Protocols**, London, v. 3, n. 5, p. 877-882, 2008.

Melhoramento Genético de Plantas Autógamas

*Bruno Henrique Pedroso Val
Cleber Vinícius Giarretta Azevedo
Eder Licieri Groli
Eduardo Henrique Bizari
Sandra Helena Unêda-Trevisoli*

1. Introdução

O melhoramento genético de plantas constitui-se em um processo que visa a modificar a constituição genética das plantas, a fim de atender às necessidades da população humana (BUENO et al., 2006). As plantas autógamas possuem o mecanismo de autotofecundação como principal sistema de reprodução. Sendo assim, para atingir os objetivos de um programa de melhoramento genético de plantas autógamas com sucesso, é necessário seguir pelo menos cinco etapas principais: 1º) definição das características de interesse com estudo detalhado da diversidade genética e escolha dos melhores genitores para compor os processos de hibridação, etapa conhecida como pré-melhoramento; 2º) hibridação artificial entre os genitores escolhidos e formação da população sobre a qual será realizada a seleção; 3º) condução das populações segregantes e seleção dos genótipos promissores; 4º) ensaios finais com os genótipos selecionados, em experimentos com repetições, conduzidos em diversos locais e anos, e 5º) registro da nova cultivar desenvolvida.

2. Pré-Melhoramento

O principal objetivo de um programa de melhoramento genético de autógamias, assim como para a maioria das espécies de importância econômica, é desenvolver novas cultivares que atendam tanto aos interesses dos agricultores quanto aos da sociedade. No entanto, para alcançar estes objetivos, é necessário, além de conhecimento profundo nas áreas de genética, melhoramento vegetal e estatística, conhecimento em outras áreas correlacionadas, como fisiologia vegetal, bioquímica, fitopatologia, botânica, entomologia e sistema de produção, que envolvem o manejo e o cultivo da cultura com a qual se está trabalhando.

A primeira etapa é definir qual ou quais serão os objetivos principais a serem considerados no programa de melhoramento e, a partir deste momento, delinear quais serão os próximos passos para a obtenção de novas cultivares, sempre lembrando que cada linha de pesquisa não deve envolver muitos caracteres, pois, geralmente, à medida que é incrementado o número de caracteres de interesse, o ganho com a seleção tende a diminuir (DE CARVALHO et al., 2008).

Em seguida, deve-se realizar um estudo detalhado da diversidade genética da cultura para a escolha dos genitores que poderão proporcionar as melhores combinações, expondo a máxima variabilidade genética possível e, conseqüentemente, possibilitar a obtenção das melhores progênies sobre as quais será realizada a seleção. Por este motivo, as hibridações artificiais sempre devem ser feitas entre os genitores de melhor desempenho, que apresentem complementariedade de caracteres e apresentem diversidade genética adequada para a obtenção de variabilidade genética nas populações segregantes. Atualmente, a principal ferramenta utilizada para identificar a diversidade genética são os marcadores moleculares, os quais possibilitam avaliar um elevado número de genótipos em um curto espaço de tempo.

Vale ressaltar que a escolha dos genitores se caracteriza como uma das principais etapas do processo de melhoramento genético, pois um equívoco nesta etapa pode comprometer todo o trabalho que será realizado posteriormente. Após a escolha acurada dos parentais, a próxima etapa do programa é a realização das hibridações artificiais.

3. Hibridações Artificiais em Plantas Autógamas

A hibridação artificial de plantas é a base do programa de melhoramento genético, pois a partir dela, formam-se as populações sobre as quais será realizada a seleção. Para realizar este processo, sempre que possível, devem-se utilizar genitores que apresentem características agronômicas desejáveis, de acordo com o objetivo do programa de melhoramento.

A principal contribuição das hibridações artificiais está relacionada ao incremento da variabilidade genética, a qual é indispensável para que o processo de seleção seja eficiente e possibilite a obtenção dos resultados almejados, pois o objetivo principal das hibridações é tentar unir, na progênie, alelos de interesse agronômico que anteriormente se encontravam separados em ambos os genitores.

Para exemplificar como é feito o processo de hibridação em autógamas, será utilizada como exemplo a cultura da soja (*Glycine max* L.), visto que, para cada cultura, existem particularidades.

A hibridação em soja geralmente é feita de forma manual, com o auxílio de uma pinça de ponta reta, sem ranhuras no interior e com a extremidade fina. O primeiro passo é realizar a emasculação do genitor feminino. Esta etapa deverá ser realizada antes da abertura do botão floral, pois devido ao fato de a soja apresentar cleistogamia, a autofecundação do estigma ocorre antes da abertura da flor (BORÉM et al., 2009).

O botão floral utilizado como genitor feminino deverá ser aquele que abrirá no dia seguinte à emasculação. Neste estágio de desenvolvimento, é possível ver o ápice das pétalas surgindo entre as sépalas. A emasculação é feita segurando-se o botão floral com o auxílio do dedo indicador e do polegar, então com a pinça retira-se cada uma das sépalas com um movimento para baixo e, em seguida, devem-se remover as pétalas. O operador que possui experiência, consegue, com um único movimento, remover todas as pétalas e as anteras, deixando a flor emasculada. Caso as pétalas e os grãos de pólen não saiam, os mesmos devem ser removidos um a um, sempre tendo o máximo de cuidado para não danificar o ovário da flor, pois se isto ocorrer, possivelmente, ocorrerá o abortamento da mesma.

A flor a ser utilizada como genitor masculino, deve ser, de pre-

ferência, uma flor do dia, com as pétalas bem vistosas. Após a escolha da flor, deve-se retirar cuidadosamente a sépala que fica em frente à quilha e, então, com o auxílio de uma pinça, deve-se abrir a quilha, retirar o ovário com as anteras e derriçar o pólen, gentilmente, sobre o estigma da flor feminina emasculada. No caso da soja, quando se trabalha em casa de vegetação, não é necessário fazer a proteção da flor após a polinização.

Após o processo de hibridação, o cruzamento deve ser identificado. Para isto, em uma etiqueta (modelo relojoeiro), deve-se escrever o nome do genitor feminino (é padrão sempre escrever o nome do genitor feminino no início), do genitor masculino, a data em que foi realizada a hibridação e o nome do melhorista que realizou o processo.

Sempre que possível, devem-se utilizar, como genitor feminino, plantas que apresentem flores brancas e, como genitor masculino, plantas que apresentem flores roxas, visto que a cor da flor roxa apresenta dominância sobre a branca (marcador morfológico). Sendo assim, se todas as plantas F_1 apresentarem flor de coloração roxa, indica que elas são provenientes de hibridação; no entanto, se as plantas F_1 apresentarem flores de coloração branca, ocorreu autofecundação. Além disso, o caráter cor da flor apresenta pleiotropia com o caráter cor do hipocótilo e, assim sendo, a identificação das plantas provenientes de hibridação em F_1 pode ser realizada logo após a emergência das plântulas (BORÉM et al., 2009).

4. Métodos de Condução de Populações Segregantes

Após a escolha dos genitores e a realização das hibridações, o processo de seleção tem por finalidade identificar os melhores indivíduos dentro da população segregante. Para o sucesso da seleção, a experiência do melhorista e a qualidade de condução das populações são de extrema importância, pois a mesma é feita principalmente por meio de avaliações visuais. Os métodos de condução de populações segregantes mais utilizados no melhoramento de autógamas são: Genealógico (*Pedigree*), Populacional (*Bulk*), Retrocruzamento e SSD (*Single Seed Descent*).

5. Método Genealógico (*Pedigree*)

O método genealógico possui como característica principal o início do processo da seleção em F_2 , com sucessivas seleções até que os genótipos apresentem elevado grau de homozigose. As decisões devem ser tomadas entre centenas ou milhares de indivíduos da população, os quais estão segregando para as características de interesse. Em todas as etapas do método, é realizada uma série de anotações que proporcionam as relações entre as famílias que estão sendo avaliadas. Desta forma, ao atingir as gerações de elevada homozigose (F_6 em diante), o melhorista consegue saber exatamente qual foi a planta F_2 que deu origem àquele genótipo.

Na geração F_1 , não é realizada seleção, porém devem-se fornecer condições ideais para as plantas produzirem o máximo de sementes possível, pois, desta forma, o número de recombinações genéticas na geração F_2 será maior. A seleção inicia-se na geração F_2 , onde começa a segregação. Em F_2 , as plantas devem ser semeadas de forma espaçada, visando a permitir que as mesmas expressem todo seu potencial genético e que o trabalho de avaliação do melhorista seja facilitado. Na geração F_3 , cada planta selecionada em F_2 dará origem a uma linha no campo, na qual será realizada nova seleção. Nesta etapa, inicia-se a seleção entre e dentro das linhas (famílias), sempre buscando selecionar as plantas com as melhores características. As gerações F_4 e F_5 são conduzidas da mesma forma que a geração F_3 , porém acentua-se a seleção entre famílias. A partir de F_6 , são conduzidos os ensaios com repetições e não ocorre mais a seleção dentro da linha, somente entre as mesmas. As famílias que se destacarem em todos os anos e locais são multiplicadas e poderão ser disponibilizadas aos agricultores (Figura 1A).

A principal vantagem do método genealógico consiste no fato de o melhorista possuir um conhecimento detalhado de todas as famílias que estão sendo avaliadas. No entanto, devido à grande necessidade de anotações, este método torna-se bastante trabalhoso, demandando muita mão de obra. Além disso, como é realizada seleção em todas as gerações, as populações segregantes não podem ser conduzidas fora da estação de cultivo, tornando o processo de obtenção de novas cultivares mais demorado, quando comparado com outros métodos.

6. Método Populacional (*Bulk*)

O método populacional, também denominado de *Bulk*, é conhecido por ser um método de condução mais simples quando comparado ao método genealógico. Este método baseia-se na seleção natural para a identificação dos genótipos superiores.

A partir da geração F_2 , as plantas são colhidas em conjunto e apenas uma amostra é reservada para a geração F_3 . Esse processo é repetido até que o nível de homozigose desejado seja alcançado. Em algumas espécies, exige-se que as populações sejam conduzidas até gerações F_7 ou F_8 , porém em alguns casos, a partir da geração F_5 , já se torna suficiente.

Na geração em que ocorrerá a seleção, o plantio deverá ser mais espaçado para melhor observação das plantas pelo melhorista. Os genótipos desejáveis são selecionados, e as plantas são trilhadas separadamente. Cada planta irá gerar uma linha na geração seguinte. As fileiras que se apresentarem mais promissoras serão colhidas em “Bulk” ou de forma conjunta, sendo avaliadas em ensaios preliminares de linhagens. Após a coleta e a análise dos dados, as melhores linhagens serão destinadas para os ensaios intermediários de linhagens e, posteriormente, para os ensaios finais de linhagens (Figura 1B).

Este método possui como vantagens sua simplicidade e baixo custo, permitindo o uso de mecanização na colheita e por ser facilmente associado a outros métodos por meio de modificações. As desvantagens consistem em: parte da geração F_2 não é representada nas gerações posteriores, além da necessidade de se testar um grande número de linhagens após a homozigose.

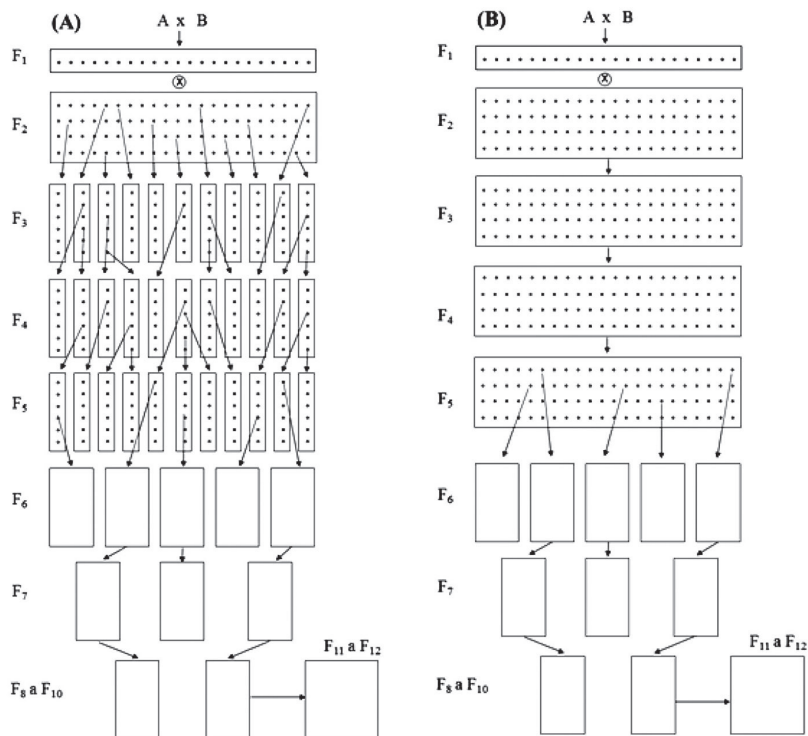


Figura 1. Esquema do Método Genealógico (A) e Método da População (B).

7. Método SSD (*Single Seed Descent*)

O método SSD tem como característica a redução do tempo requerido para a obtenção das linhagens homozigóticas. Nas fases iniciais desse método, não é necessário ter condições ideais de cultivo, podendo ser realizada mais de uma geração por ano.

O método consiste na colheita de uma única semente por planta, que irá originar uma nova planta na geração seguinte. Esse processo é realizado de F_2 até F_5 . A partir da geração F_5 , as condições devem ser as recomendadas, pois será iniciada a seleção. Da geração F_6 em diante, as progêies com desempenho superior são colhidas em *Bulk* e avaliadas em ensaios preliminares

de linhagens (EPL). A partir dessa etapa, os procedimentos são comuns aos outros métodos (Figura 2).

Apesar de ser um método muito utilizado, apresenta alguns problemas, como a perda de indivíduos por causas ambientais, diminuindo a representatividade da população, além da forte influência da seleção natural. Para corrigir esse problema e alguns inconvenientes surgidos, algumas variações do método foram propostas. Dentre essas variações, podem-se mencionar: SPD (*Single Pod Descent*), SPDS (*Single Pod Descent with Selection*), SHD (*Single Hill Descent*) e SHDT (*Single Hill Descent Thinned*).

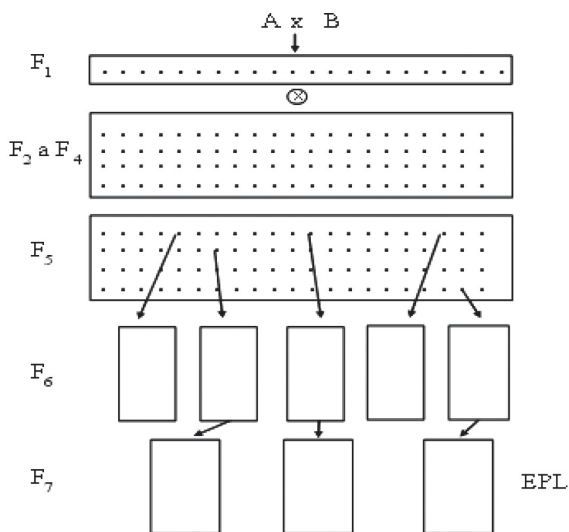


Figura 2. Esquema do Método SSD (*Single Seed Descent*).

8. Método do Retrocruzamento

O método do retrocruzamento é uma técnica que permite a transferência de um ou de poucos genes de interesse de um genitor denominado doador (não recorrente) para outro denominado receptor (recorrente).

De acordo com Fehr (1987), este método é utilizado para melhorar uma característica pontual de um genótipo que já possua

bons atributos agronômicos. Desta forma, o genitor recorrente, geralmente, é uma cultivar de interesse comercial, porém com limitações em uma ou poucas características, as quais o melhorista busca solucionar com a transferência de genes de um genitor doador.

Na prática, como demonstrado na Figura 3, este método consiste na realização do cruzamento entre um genitor doador (D) para a característica de interesse, com um genitor recorrente (R) a ser melhorado, obtendo-se a F_1 . Em seguida, as progênie são retrocruzadas com o genitor recorrente repetidas vezes, selecionando-se para a característica a ser incorporada, recuperando-se, assim, as características do genitor recorrente em conjunto com o novo caráter de interesse. Normalmente, são necessários entre cinco a seis retrocruzamentos para se obter alto grau de recuperação do genitor recorrente.

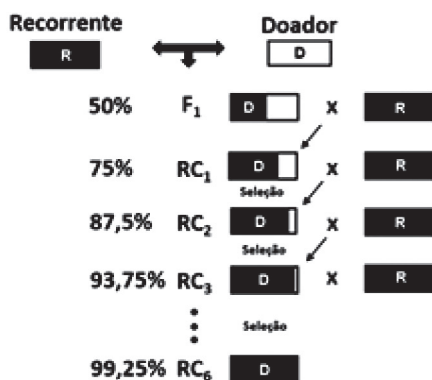


Figura 3: Esquema básico do Método do Retrocruzamento

Por meio deste método, características controladas por um ou poucos genes são facilmente transferidas, sendo que a seleção para alelos dominantes, geralmente, é realizada de forma mais fácil que para os alelos recessivos. No caso da transferência de genes recessivos, é necessária a autofecundação nas gerações ímpares para a realização da seleção da característica. Detalhes sobre estes procedimentos podem ser encontrados em Fehr (1987), Destro e Montalván (1999).

Este método possui como vantagem a utilização de um genótipo já conhecido, com alta previsibilidade do resultado, além de poder ser conduzido fora da região onde o novo genótipo será cultivado; por este motivo, após a obtenção da nova cultivar, não é necessária a condução de testes finais para comprovar seu desempenho. Como desvantagens, cita-se a dificuldade na realização dos diversos retrocruzamentos e o elevado tempo gasto na obtenção da cultivar, o que pode tornar obsoleto o novo genótipo obtido.

Visando-se a reduzir o tempo gasto para a recuperação do genitor recorrente, marcadores moleculares têm sido utilizados como ferramentas para auxiliar na seleção dos genótipos com a presença do gene de interesse, além de permitir, também, a identificação de genótipos mais semelhantes ao genitor recorrente. De acordo com Faleiro et al. (2004), por meio desta estratégia, o número de retrocruzamentos necessários para recuperar o genitor recorrente pode ser reduzido para três.

9. Ensaios de VCU

Em programas de melhoramento genético de espécies autógamas, as populações segregantes são geralmente avançadas e selecionadas em apenas um ambiente, devido, principalmente, ao elevado número de progênies em testes e à reduzida quantidade de sementes obtidas. Desta forma, as progênies são selecionadas sob uma condição específica. Entretanto, na prática, ao final do processo de seleção, os genótipos selecionados serão cultivados em diferentes condições de cultivo.

Por este motivo, no final do processo de seleção, é necessário que todos os genótipos selecionados sejam avaliados em diversos ambientes, pois é possível que o melhor genótipo em um determinado ambiente não apresente o mesmo desempenho em outro. Este comportamento diferenciado dos genótipos nos diferentes ambientes de cultivo ocorre devido ao efeito da interação genótipos por ambientes (CRUZ et al., 2004). Então, para contornar este problema, faz-se necessária a avaliação dos genótipos selecionados em diferentes locais e anos, representativos da região para a qual se pretende recomendar a nova cultivar.

Estas avaliações são denominadas de ‘Ensaios de VCU’ (Valor

de Cultivo e Uso), sendo obrigatórias para o registro de novas cultivares junto ao Ministério da Agricultura Pecuária e Abastecimento. Estes ensaios devem ser realizados de acordo com critérios mínimos estabelecidos para cada cultura e têm como finalidade fornecer subsídios que possibilitem a indicação dos genótipos mais estáveis e adaptados à região de cultivo indicada.

Após esta etapa, geralmente, os genótipos selecionados já satisfazem aos requisitos mínimos necessários para registro e indicação como novas cultivares comerciais superiores.

10. Referências bibliográficas

BORÉM, A.; ALMEIDA, L.A.; KIIHL, R.A.S. Hibridação em Soja. In: BORÉM, A. (Ed.). **Hibridação Artificial de plantas**. – 2. ed. – Viçosa: UFV, 2009. 625p.

BUENO, L. C. de S.; MENDES, A. N. G.; DE CARVALHO, S. M. **Melhoramento genético de plantas**: princípios e procedimentos. – 2. ed. - Lavras: UFLA, 2006. 319p.

CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2004.

DE CARVALHO, F. I. F.; LORENCETTI, C.; MARCHIORO, V. S.; SILVA, S. A. **Condução de populações no melhoramento genético de plantas**. – 2. ed.- Pelotas: UFPel, 2008. 288p.

DESTRO, D.; MONTALVÁN, R. **Melhoramento genético de plantas**. 1. ed. Londrina: UEL, 1999. 245p.

FALEIRO, F. G.; RAGAGNIN, V. A.; MOREIRA, M. A.; DE BARROS, E. G. Use of molecular markers to accelerate the breeding of common bean lines resistant to rust and anthracnose. **Euphytica**, v. 138, n. 3, p. 213-218, 2004.

FEHR, W. R. **Principles of cultivar development**: theory and technique. New York: MacMillan, 1987. 536 p.