

## **Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas**

**República Federativa do Brasil**

*Luiz Inácio Lula da Silva*

Presidente

**Ministério da Agricultura, Pecuária e Abastecimento**

*Roberto Rodrigues*

Ministro

**Empresa Brasileira de Pesquisa Agropecuária – Embrapa**

**Conselho de Administração**

*José Amauri Dimárzio*

Presidente

*Clayton Campanhola*

Vice-Presidente

*Alexandre Kalil Pires*

*Hélio Tollini*

*Ernesto Paterniani*

*Luis Fernando Rigato Vasconcellos*

Membros

**Diretoria–Executiva da Embrapa**

*Clayton Campanhola*

Diretor-Presidente

*Mariza Marilena Tanajura Luz Barbosa*

*Gustavo Kauark Chianca*

*Herbert Cavalcante de Lima*

Diretores-Executivos

**Embrapa Roraima**

*Antonio Carlos Centeno Cordeiro*

Chefe Geral

*Oscar José Smiderle*

Chefe Adjunto de Pesquisa e Desenvolvimento

*Miguel Amador de Moura Neto*

Chefe Adjunto de Administração



*Empresa Brasileira de Pesquisa Agropecuária  
Centro de Pesquisa Agroflorestal de Roraima  
Ministério da Agricultura, Pecuária e Abastecimento*

*ISSN 0101 – 9805  
Julho, 2003*

# **Documentos** *01*

## **Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas**

Maria Aldete Justiniano da Fonseca Ferreira

Boa Vista, Roraima  
2003

4 Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas

Exemplares desta publicação podem ser obtidos na:

Embrapa Roraima

Rod. BR-174 Km 08 - Distrito Industrial Boa Vista-RR

Caixa Postal 133

69301-970 - Boa Vista - RR

Telefax: (095) 626.7018

e\_mail: [sac@cpafrr.embrapa.br](mailto:sac@cpafrr.embrapa.br)

**Comitê de Publicações:** Oscar José Smiderle - Presidente

Bernardo de Almeida Halfeld Vieira

Evandro Neves Muniz

Hélio Tonini

Moisés Cordeiro Mourão de Oliveira Júnior

Patrícia da Costa

Paulo Roberto Valle da Silva Pereira

Editoração: Maria Lucilene Dantas de Matos

**1ª edição**

1ª Impressão (2003): 300

Normalização Bibliográfica: Maria José Borges Padilha

FERREIRA, M. A. J. da F. Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas. Boa Vista: Embrapa Roraima, 2003. 63p. (Embrapa Roraima. Documentos, 1).

ISSN: 0101-9805

1. Planta – Genética – Melhoramento. 2. Marcadores moleculares. I. Embrapa Roraima. II. Título. III. Série

CDD: 581.3

## **Autora**

**Maria Aldete Justiniano da Fonseca Ferreira**

Dra. Genética e Melhoramento de Plantas,

Pesquisadora da Embrapa Roraima

Cx. Postal 133 Boa Vista-RR 626 7125

E-mail: [aldete@cpafrr.embrapa.br](mailto:aldete@cpafrr.embrapa.br)

## Sumário

1. Introdução.....	7
2. Marcadores Moleculares.....	8
2.1. Eletroforese.....	8
2.2. Polimorfismo no comprimento de fragmentos de restrição (RFLP).....	13
2.3. Marcadores baseados em reação de polimerase em cadeia (PCR).....	16
2.4. Polimorfismo de DNA amplificado ao acaso (RAPD).....	17
2.5. Polimorfismo de comprimento de fragmentos amplificados (AFLP).....	19
2.6. Minisatélites e Microsatélites.....	21
3. Marcadores moleculares na genética de populações.....	23
4. Marcadores moleculares na genética quantitativa.....	36
5. Marcadores moleculares no melhoramento de plantas .....	44
6. Considerações finais.....	52
7. Referências Bibliográficas.....	54

# **Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas**

---

Maria Aldete Justiniano da Fonseca Ferreira

## **1- INTRODUÇÃO**

A genética de populações e a genética quantitativa constituem alicerces de fundamental importância para o melhoramento de plantas. Para que um programa de melhoramento seja desenvolvido com sucesso são necessários alguns pré-requisitos, como, por exemplo, conhecer o sistema de acasalamento das espécies a serem melhoradas, assim como a estrutura genética das populações; detectar a existência ou não de variabilidade genética no germoplasma disponível com relação às características de interesse; conhecer a ação gênica destas características, entre outros fatores.

O melhoramento de plantas tem como meta principal reunir em um genótipo as características desejáveis, a nível comercial, de uma determinada espécie. A avaliação e seleção de indivíduos são realizadas, há muito tempo, com base em seus fenótipos. Entretanto, os caracteres de interesse, na sua grande maioria, são controlados por muitos genes, apresentam baixa herdabilidade e alta influência ambiental. Esses fatores minimizam o progresso esperado com os processos seletivos e, conseqüentemente, o êxito de programas de melhoramento.

Os marcadores moleculares, por outro lado, têm despontado como uma potente ferramenta, que pode auxiliar estes ramos da genética e, evidentemente, o melhoramento de plantas. Com a identificação direta dos genótipos, é possível determinar as frequências gênicas e com isso averiguar o sistema de acasalamento da espécie, a variabilidade genética das populações, o tipo de ação gênica predominante no controle de caracteres,

assim como superar a maior parte das limitações da análise fenotípica. Além disso, a duração de programas de melhoramento pode ser reduzida, resultando em liberação mais rápida de cultivares e retorno mais rápido do investimento aplicado.

É considerado marcador molecular qualquer fenótipo molecular oriundo de um gene expresso, como no caso de isoenzimas, ou de um segmento específico de DNA que corresponde a regiões expressas ou não do genoma. Um marcador molecular é definido como marcador genético, ao ser comprovado que seu comportamento está de acordo com as leis básicas da herança Mendeliana (Ferreira & Grattapaglia, 1995).

As técnicas de marcadores moleculares mais utilizadas, atualmente, são as isoenzimas, RFLP (Polimorfismo no Comprimento de Fragmentos de Restrição), RAPD (Polimorfismo de DNA Amplificado ao Acaso), AFLP (Polimorfismo de Comprimento de Fragmentos Amplificados), Minisatélites e Microsatélites. Estes marcadores podem ser aplicados com diversos fins, como, por exemplo, para estudar a estrutura genética de populações; análise de filogenias; detecção de ligação gênica com caracteres mono e poligênicos; identificação de variedades; avaliação de germoplasma; introgressão gênica; seleção indireta de características agrônômicas, dentre outros. Pesquisas nesta área estão sendo desenvolvidas, tendo-se alguns casos, de eficiência comprovada, do emprego das técnicas de marcadores moleculares na genética de populações e quantitativa e no melhoramento de plantas.

Nesta revisão, são descritas as técnicas de marcadores moleculares e suas aplicações tanto na área de genética de populações e quantitativa quanto no melhoramento de plantas.

## **2. Marcadores Moleculares**

### **2.1. Eletroforese**

Segundo Moraes (1992), a partir de 1926 a técnica de eletroforese teve um grande impulso, quando Summer conseguiu isolar e cristalizar a enzima urease, demonstrando que enzimas e proteínas são moléculas que podem ser purificadas. Beadle & Tatum, na década de 40, formularam a teoria de “um gene uma enzima”, dando origem à genética relacionada às isoenzimas, sendo que este termo só foi proposto no final dos anos 50 por



Markert & Moller, com a finalidade de designar as formas moleculares múltiplas nas quais as enzimas podem existir em um mesmo organismo. Em 1971, Gottlieb propôs o termo aloenzima para designar as diferentes formas de proteínas produzidas por diferentes alelos em um mesmo loco. Portanto, o termo isoenzima define um grupo de múltiplas formas moleculares da mesma enzima que ocorre em uma espécie, como resultado da presença de mais de um gene codificando cada uma das enzimas (Moss, 1982 citado por Ferreira & Grattaplagia, 1995).

Coolidge em 1939, tentou pela primeira vez caracterizar proteínas, através da mobilidade eletroforética em meios semi-porosos e Durrum em 1950 conseguiu separar proteínas, usando tiras de papel. O gel de amido como suporte para eletroforese foi empregado inicialmente por Smithies em 1955, o que proporcionou uma resolução superior ao meio de papel, separando as proteínas pelo tamanho e forma (Moraes, 1992).

De acordo com Ballvé (1992), a aplicação das isoenzimas como marcadores genéticos teve como base os trabalhos de Harris (1966) em humanos e de Lawontin & Hubby (1966) em *Drosophila*, os quais verificaram que as análises eletroforéticas revelavam polimorfismos genéticos que podiam ser utilizados para caracterizar populações e espécies. Desde então, houve aumento considerável nas pesquisas relativas à técnica de eletroforese e regulação gênica das isoenzimas (Shannon, 1968; Brewer & Singi, 1970; Nelson & Bun, 1973; Scandalios, 1974; Market, 1975; Shieds et al., 1983 citados por Ballvé, 1992).

Segundo Murphy et al. (1990) citado por Ferreira & Grattaplagia (1995), a premissa básica adotada ao se utilizar dados isoenzimáticos é que diferenças na mobilidade de isoenzimas em um campo elétrico são resultantes de diferenças ao nível de seqüências de DNA que codificam tais enzimas. Portanto, se os padrões de bandas de dois indivíduos diferem, assume-se que estas diferenças possuam base genética e sejam herdáveis.

A migração diferencial da proteína codificada por um alelo, baseia-se na diferente composição ou seqüência dos aminoácidos das cadeias polipeptídicas que são os produtos primários especificados por esses alelos e que compõem as proteínas. Mudanças de bases aminadas do DNA do gene estrutural podem resultar na substituição de um aminoácido e na formação de um polipeptídeo de carga ou configuração diferente.

A enzima resultante, também modificada, quando submetida à eletroforese, terá uma migração alterada, refletindo indiretamente a variabilidade no DNA correspondente ao gene estrutural. A maioria das alterações é discriminada, exceto as substituições neutras ou aquelas no interior da molécula que não produzem mudanças conformacionais (Ramshaw et al., 1979 citado por Ballvé, 1992).

A técnica de eletroforese é composta por três etapas. A primeira refere-se à extração de proteínas do tecido vegetal. Uma grande parte dos tecidos vegetais podem ser utilizados para análise isoenzimática, entretanto os mais empregados são as folhas cotiledonares. O tecido vegetal deve ser macerado com um tampão de extração, o qual deve possibilitar a extração das proteínas, a manutenção de sua atividade catalítica e a prevenção de oxidação de compostos fenólicos associados. Após a extração das proteínas, estas devem ser separadas através de eletroforese (2ª Etapa). Existem vários tipos de gel de eletroforese, como penetrose, amido, poliacrilamida. Nesta fase diversos sistemas tampão podem ser utilizados. Terminada a eletroforese ou “corrida do gel”, o mesmo é submetido à coloração histoquímica (3ª Etapa), sendo que os corantes utilizados são dependentes de cada sistema enzimático. As reações específicas, resultantes dos corantes, formam complexos coloridos e insolúveis no gel visualizados como bandas nas posições em que as enzimas se encontravam quando a eletroforese foi interrompida (Ferreira & Grattapaglia, 1995). É importante frisar que condições apropriadas da técnica de eletroforese, da metodologia de extração das proteínas e das soluções de coloração são requisitos fundamentais para uma resolução nítida das bandas.

Ao conjunto de bandas visualizadas no gel denomina-se zimograma, que refere-se ao fenótipo eletroforético, representando as formas multimoleculares das enzimas, ou seja, as isoenzimas. As aloenzimas correspondem às bandas resultantes de produtos de alelos codificando enzimas específicas. A interpretação dos zimogramas é importante para a caracterização do loco isoenzimático em termos genéticos (genótipos).

Considerando um dado loco, o número de bandas observadas no gel dependerá do estado alélico deste loco (homozigótico ou heterozigótico), assim como da estrutura quaternária da isoenzima codificada. Um loco homozigótico é representado por somente uma banda, de migração específica, correspondente ao único tipo de aloenzima codificada. Ao passo que os genótipos heterozigóticos são representados por um número variável de bandas de acordo com o número de polipeptídeos que forma a enzima ativa. Desta forma, para uma enzima monômera são vistas duas bandas, correspondentes às

duas aloenzimas monômeras codificadas (**A** e **B**). Para uma enzima dímera, três bandas, duas correspondentes aos homodímeros (**AA** e **BB**) e uma intermediária, o heterodímero (**AB**), composto dos dois polipeptídeos distintos, **A** e **B**, especificados por cada alelo parental (**a** e **b**). Quando a enzima tem estrutura tetramérica, o heterozigoto apresenta cinco bandas, dois homotetrâmeros (**AAAA** e **BBBB**) e três heterotetrâmeros (**AAAB**, **AABB** e **ABBB**) diferentes (Tabela 1) (Ballvé, 1992).

Segundo Weeden & Gottlieb (1979) citado por Ballvé (1992) a complexidade dos zimogramas aumenta com a expressão de mais de um loco e com as possíveis interações entre eles, sendo que em determinadas situações a interpretação dos fenótipos é facilitada com a análise dos zimogramas de grãos de pólen (haplóides).

**Tabela 1.** Zimogramas de genótipos homozigóticos e heterozigóticos e segregação da geração F<sub>2</sub> para um loco codificando uma enzima monômera, dímera e tetramérica.

Enzima	Genótipos			Segregação F <sub>2</sub>		
	Genitores		Híbrido	1	2	3
	Aa	bb	ab	aa	ab	bb
Monômera	— A	— B	— A — B	—	—	—
	— AA	— BB	— AA — AB — BB	—	—	—
Tetramera	— AAAA	— BBBB	— AAAA — AAAB — AABB — ABBB — BBBB	—	—	—

Ballvé (1992), de acordo com uma revisão bibliográfica, enumera uma série de situações excepcionais que podem complicar a interpretação dos zimogramas, como:

- Bandas heterodímeras que não migram exatamente na mediana entre as bandas homodímeras (Goodman & Stuber, 1983);
- Bandas adicionais que resultam do armazenamento, preparação das amostras ou mesmo do processo da eletroforese (Staub et al., 1985);

- c) Bandas múltiplas oriundas de modificações pós-tradução das enzimas, freqüentemente chamadas de sombras ou bandas fantasmas. Essas bandas múltiplas, de migração concomitante, podem se comportar geneticamente como alelos simples (Rick & Fobes, 1976; Kahler, 1983; Harry, 1986). Não se sabe ao certo se estas aloenzimas com padrão de bandas múltiplas são produzidas por modificações pós-tradução *in vivo* ou por mudanças conformacionais *in vitro*;
- d) Presença de “alelos nulos”, assim denominadas porque ou não são transcritos ou codificam para polipeptídeos defeituosos, que resultam em ausência de atividade enzimática, alterando assim o número de bandas normalmente observadas nos genótipos heterozigotos (Wendel & Weeden, 1989).

Segundo Weeden & Wendel (1989) citado por Ballvé (1992), as isoenzimas apresentam várias características que as tornam excelentes marcadores genéticos, pois exibem, geralmente, herança mendeliana simples, codominância, penetrância completa, ausência de pleiotropia e de interações epistáticas e estão disponíveis na população apresentando grande diversidade.

Medina Filho (1983) citado por Moraes (1992), também apresenta vários argumentos a favor da técnica de eletroforese, como a sua aplicabilidade a qualquer tipo de planta; revelar, de imediato, grande número de locos e alelos; o fato dos equipamentos empregados serem de fácil manipulação e econômicos; facilidade na utilização do gel de amido, devido a sua atoxicidade, assim como a vantagem de poder ser cortado em diversas camadas horizontais, possibilitando a revelação de sistemas enzimáticos diferentes e conseqüentemente proporcionando a análise rápida e simultânea de vários locos isoenzimáticos.

Por outro lado, quando a investigação científica exige cobertura mais ampla do genoma, como por exemplo no caso de mapeamento genético ou caracterização detalhada de germoplasma, a técnica isoenzimática não é recomendada, uma vez que o número total de locos que podem ser detectados no genoma e o número de alelos por loco, isto é, o nível de polimorfismo genético detectável em cada loco são limitados (Ferreira & Grattaplagia, 1995).

As isoenzimas apresentam outras limitações como marcadores moleculares, segundo Ferreira & Grattaplagia (1995), como:

- a) Ocorrência de modificações pós-tradução das enzimas, produzindo as chamadas “isoenzimas conformacionais” ou formas múltiplas do produto de um único gene, que diferem em estruturas secundárias e terciárias;
- b) Especificidade de formas isoenzimáticas em alguns tecidos vegetais;
- c) Influência das condições ambientais na atividade enzimática, resultando em polimorfismo isoenzimático em resposta a condições ambientais;
- d) Atividade isoenzimática diferenciada em relação a diferentes estádios de desenvolvimento;
- e) Dificuldades na interpretação de zimogramas quando isoenzimas com mobilidade eletroforéticas idênticas representam os produtos de dois locos distintos do mesmo sistema enzimático, tornando difícil ou impossível detectar quais alelos pertencem a cada loco (Murphy et al., 1990).

Apesar destas limitações, a técnica de isoenzimas continua sendo uma das classes de marcadores moleculares muito utilizada e sobretudo muito útil para análises genéticas que não necessitam de uma amostragem ampla do genoma.

## **2.2. Polimorfismo no comprimento de fragmentos de restrição (RFLP)**

Segundo Ferreira & Grattapaglia (1995) os marcadores RFLP foram estudados pela primeira vez por Grodzicker et al. (1974), em um experimento que visava a detecção de mutação em DNA de vírus. Outros pesquisadores, a exemplo de Botstein et al. (1980) e Wyman & White (1980), inspiraram-se nesta idéia e prepuseram a utilização da técnica em análises genômicas, alguns anos mais tarde. Desde então, esta técnica vem sendo utilizada extensivamente, tornando-se de grande utilidade e importância nos estudos biológicos.

O principal objetivo da técnica de RFLP é detectar diferenças na seqüência de DNA de dois indivíduos. Para um melhor entendimento, pode-se considerar o DNA extraído de dois indivíduos distintos, submetido a clivagem por enzima de restrição, sendo cortado em um grande número de pontos (sítios de restrição), originando  $10^4$  a  $10^7$  fragmentos de

DNA, a depender do genoma. A clivagem de fragmentos de tamanhos distintos é consequência de diferenças na seqüência de DNA dos dois indivíduos. Estes fragmentos são separados por meio de eletroforese em gel de agarose, porém ainda não é possível visualizar os fragmentos, porque o efeito produzido pela eletroforese é semelhante a um arraste contínuo no gel. Assim, para os marcadores serem detectados ou visualizados é preciso transferir os fragmentos separados no gel de agarose para uma membrana de nitrocelulose por capilaridade ou vácuo, processo este denominado de *Southern Blot*. A membrana de nitrocelulose pode ser reutilizada algumas vezes, o que contribui para diminuir os custos da técnica, bem como facilita o trabalho de geração de dados. A visualização dos fragmentos polimórficos entre todos fragmentos transferidos para a membrana de nitrocelulose, só é possível empregando-se “sondas”, que corresponde a hibridação de pequenos fragmentos clonados de DNA, que possuem seqüências homólogas do DNA imobilizado na membrana. Desta maneira, visualiza-se apenas os fragmentos que são homólogos ao DNA da sonda. Nas sondas são incorporados nucleotídeos que contem moléculas de fósforo radioativo ou nucleotídeos modificados nos fragmentos clonados do DNA. A membrana de nitrocelulose é finalmente exposta a um filme de raio X (autoradiografia) o que possibilita a revelação das bandas que são os marcadores RFLP. A emissão de partículas beta do fósforo radioativo ou fótons de luz provoca sensibilização do filme, produzindo bandas. As bandas somente serão visualizadas em posições diferentes na autoradiografia caso os dois indivíduos diferirem entre si em relação ao sítio de restrição enzimática na fita de DNA, ocasionando fragmentos de tamanhos distintos. Estas bandas, porém, são utilizadas como marcador genético, apenas se for verificada segregação Mendeliana entre elas, exigindo para tanto a utilização de populações segregantes.

O polimorfismo observado nesta técnica, portanto, ocorre devido ao DNA dos indivíduos diferirem na seqüência de nucleotídeos ao longo da fita. Assim quanto mais distinto um indivíduo do outro, maior será o polimorfismo, ou seja, maiores serão as diferenças em seqüências de nucleotídeos ao longo da fita de DNA.

Alguns procedimentos são fundamentais para o devido emprego desta técnica como marcador genético, tais como:

(1) Obtenção de uma biblioteca de clones

A obtenção dos clones que serão utilizados como sondas, pode ser feita de diferentes maneiras. Uma delas, é por meio da transcrição reversa de mRNA do indivíduo em estudo, o que proporciona a formação de uma biblioteca de moléculas de DNA complementar (*cDNA library*). Os fragmentos de DNA genômico podem, também, serem clonados ao acaso (*genomic library*). Outra forma é ampliar seqüências conhecidas, via PCR, utilizando *primers* específicos ou ainda, amplificar via PCR, bandas selecionadas e obtidas em gel de eletroforese.

## (2) Seleção de clones a serem utilizados como sondas

Os clones a serem utilizados como sondas não devem possuir seqüências de DNA repetitivo, pois isto provoca borrões não interpretáveis na autoradiografia ou padrões de bandas múltiplas que dificultam a distinção de locos, alelos e segregação Mendeliana. Estas conseqüências são devido ao fato dos clones que possuem DNA repetitivo hibridizarem com vários fragmentos da membrana de nitrocelulose.

Já as sondas que contem seqüências moderadamente repetitivas no genoma, são úteis quando a finalidade é obter impressão digital genética (*genetic fingerprint*) específica para cada indivíduo.

A principal vantagem dos marcadores RFLP, quando comparados as isoenzimas, é possibilitar a cobertura de todo o genoma dos indivíduos estudados, aumentando, conseqüentemente, a probabilidade de se detectar associações estatisticamente significativas entre genes que controlam caracteres de interesse e marcadores moleculares.

Por possuírem expressão codominante, como as isoenzimas, os marcadores RFLP permitem identificar genótipos heterozigotos e homozigotos, proporcionando mais informação genética e análise pormenorizada da ação gênica e interação alélica.

Outras vantagens destes marcadores em relação às isoenzimas é o número ilimitado de marcadores RFLP, resultando em maior polimorfismo alélico em cada loco e o maior tempo de conservação do DNA, devido a sua alta estabilidade, podendo as membranas de nitrocelulose serem reutilizadas 15 ou mais vezes.

O emprego de marcadores RFLP, por outro lado, pode ser limitado em decorrência de exigir mão-de-obra qualificada, instalações adequadas e dejetos de material radioativo, no

caso do uso de  $^{32}\text{P}$  e também por ser necessária a obtenção de biblioteca de sondas, o que requer alguns meses.

### **2.3. Marcadores baseados em reação de polimerase em cadeia (PCR)**

Em meados da década de 80, Kary Mullis, concebeu a tecnologia do PCR, causando uma revolução na biologia, o que levou a este pesquisador a ganhar o prêmio Nobel de medicina no início da década de 90 (Mullis & Falsone, 1987; Saiki et al., 1985 citados por Ferreira & Grattapaglia, 1995).

A enzima DNA polimerase de *E. coli* foi utilizada nos primeiros experimentos de PCR, entretanto a técnica se expandiu quando Saiki et al. (1988) isolaram uma DNA polimerase (Taq polimerase) de uma bactéria (*Thermus aquaticus*) que vive em fontes térmicas, polimeriza a 72°C, mantendo atividade por algumas horas a 95°C. Esta descoberta proporcionou a completa automatização do processo de PCR (Ferreira & Grattapaglia, 1995).

Esta técnica, segundo Gyllensten (1989) e White et al. (1989), é um método que tem revolucionado diversos campos da biologia molecular e genética, por permitir o isolamento e a caracterização de fragmentos específicos de DNA.

De acordo com White et al. (1989) a sensibilidade, rapidez e versatilidade da técnica de PCR tem apresentado um grande impacto sobre a biologia molecular, expandindo diversas áreas que antigamente eram inatingíveis, devido a carência de um método que apresentasse as vantagens do PCR. Estes pesquisadores acrescentam que em virtude da capacidade de adicionar, deletar ou alterar as informações dos *primers* feitos na PCR, este método é ideal para introduzir mutações específicas em posições específicas em um fragmento de DNA, facilitando o estudo das interações entre proteínas e DNA.

Ferreira & Grattapaglia (1995) citam três etapas que envolvem esta técnica: desnaturação, anelamento e extensão do DNA. A primeira etapa consiste na desnaturação da fita dupla de DNA alvo, por meio da elevação da temperatura para 92 a 95°C. A depender do tamanho e seqüência do *primer* utilizado, na etapa de anelamento, a temperatura é drasticamente reduzida para 35 a 60°C, a fim de que ocorra a hibridação DNA-DNA da cada *primer* com as seqüências complementares que flanqueiam a região alvo. Para que a enzima DNA polimerase realize a extensão a partir de cada terminal 3'



dos *primers*, a temperatura deve ser novamente elevada, neste caso para 72°C. A etapa de extensão compreende a adição de nucleotídeos empregando como molde a seqüência alvo, de forma que é formada uma cópia desta seqüência. Estas etapas são repetidas por inúmeras vezes, produzindo uma grande quantidade de DNA da seqüência alvo de interesse.

Um dos principais problemas associados ao PCR, segundo Gyllensten (1989), relaciona-se à rapidez na renaturação após a desnaturação dos produtos amplificados. A magnitude deste problema está diretamente relacionado a diferença de tamanho entre o *primer* e o produto do PCR. Para amenizar este problema diversos métodos têm sido propostos, os quais baseiam-se no princípio de que somente uma simples fita de DNA ou RNA pode ser avaliada por *primer* na reação.

#### **2.4. Polimorfismo de DNA amplificado ao acaso (RAPD)**

O desenvolvimento da técnica de RAPD foi possível graças a descoberta do PCR. O princípio desta técnica é o uso de *primers* mais curtos e de seqüência arbitrária durante a reação de amplificação, eliminando a necessidade do conhecimento prévio da seqüência.

A técnica foi desenvolvida independentemente por Williams et al. (1990), que patentearam a tecnologia com o nome mais comumente utilizado (RAPD - Polimorfismo de DNA Amplificado ao Acaso) e descreveram a técnica no contexto da análise Mendeliana, demonstrando a identificação de marcadores genéticos para mapeamento; e por Welsh & McClelland (1990), que propuseram a denominação de AP-PCR (*Arbitrarily Primed - Polymerase Chain Reaction*) para a técnica, sendo esta mais adequada, pois os *primers* possuem seqüências arbitrárias, mas a amplificação não ocorre ao acaso e sim em lugares específicos no genoma. Este grupo de pesquisadores desenvolveu *fingerprints* (impressões digitais) genômicos simples e reproduzíveis para a identificação de linhagens, utilizando géis de eletroforese em poliacrilamida de maior poder de resolução juntamente com *primers* um pouco mais longos.

Ferreira & Grattapaglia (1995) relatam que a técnica de RAPD é uma variação do protocolo de PCR, porém com duas características diferentes: a primeira é que utiliza um *primer* único, ao invés de um par de *primers*; e a segunda é que este *primer* possui

seqüência arbitrária, sendo desconhecida a seqüência alvo. É necessário que duas seqüências de DNA complementares ao *primer* arbitrário estejam adjacentes o suficiente e em orientação oposta, para que ocorra a amplificação exponencial de um segmento de DNA pela DNA polimerase. Este segmento pode ser visualizado, utilizando brometo de etídio, na forma de banda em um gel de eletroforese, normalmente de agarose ou em gel de poliacrilamida de alta resolução, neste caso as bandas são visualizadas por autoradiografia (AP-PCR) ou por meio de coloração com nitrato de prata. Cada *primer* arbitrário utilizado sintetiza vários segmentos de DNA em diversos pontos do genoma, produzindo várias bandas no gel, sendo, entretanto, dependente do comprimento do *primer* utilizado e do tamanho e complexidade do genoma analisado.

Os marcadores RAPD se comportam como marcadores genéticos dominantes, ou seja, um indivíduo diplóide homocigoto para um dado loco RAPD possui dois alelos RAPD idênticos (**AA**), a partir dos quais ocorre a amplificação; o indivíduo heterocigoto (**Aa**) para o mesmo loco RAPD possui um alelo (**A**) que é amplificado e o outro (**a**) que não é; já o genótipo homocigoto recessivo (**aa**) é identificado pela ausência de banda no gel (fenótipo nulo). Desta maneira os genótipos homocigotos dominantes e os heterocigotos são considerados juntos na mesma classe fenotípica (presença de banda no gel). O que acontece é que a técnica de RAPD não apresenta sensibilidade quantitativa suficiente para discriminar estes dois casos. Apesar da maioria dos marcadores RAPD apresentarem comportamento dominante, é possível amplificar marcadores codominantes (amplificar ambos os alelos de um loco) utilizando o mesmo *primer*. Porém isto só é possível em experimentos de análise de segregação de marcadores em progênies lado a lado com os parentais (Ferreira & Grattapaglia, 1995).

Em decorrência da técnica dos marcadores RFLP se basear na hibridação de DNA, enquanto a de RAPD ter como base a amplificação de DNA, as vantagens deste último sobre o primeiro são as seguintes:

- a) Obtenção mais rápida de dados;
- b) Não exige o desenvolvimento antecipado de uma biblioteca de sondas específicas para o organismo de interesse;
- c) Pode ser utilizado um conjunto único de *primers* arbitrários;

- d) Não necessita do uso de isótopos radioativos ou marcação não-radioativa;
- e) É preciso uma quantidade mínima de DNA para a análise genotípica de um indivíduo;
- f) Possibilita amostrar regiões de DNA repetitivo, pois os *primers* empregados para detecção de variação ao nível de DNA são arbitrários;
- g) Técnica mais sensível na detecção de polimorfismo ao nível de DNA;
- h) Apresenta custo mais baixo, em virtude dos gastos com mão-de-obra, reagentes e suplementos serem menores e não necessitar de instalações sofisticadas de laboratório.

Os marcadores RAPD, entretanto, apresentam algumas desvantagens, como o baixo conteúdo de informações genéticas por loco; a impossibilidade de discriminar genótipos heterozigotos de homozigotos; o desconhecimento prévio da base genética das bandas, que pode ser limitante em alguns casos.

Atualmente a técnica de RAPD é considerada mais barata que a do RFLP, porém mais cara que a de isoenzimas. Acredita-se que, nos próximos anos, ocorra uma redução no preço da enzima Taq Polimerase, responsável por 80% do custo total de um dado genotípico (*data point*) e conseqüentemente da técnica de RAPD.

## **2.5. Polimorfismo de comprimento de fragmentos amplificados (AFLP)**

A técnica de AFLP, mais recente que as anteriores, foi desenvolvida por Zabeau (1993) citado por Vantoi et al. (1996). Esta técnica combina a especificidade, resolução e poder de amostragem de digestão com enzimas de restrição com a velocidade e praticidade de detecção dos polimorfismos via PCR.

A análise de AFLP compreende quatro etapas:

- 1ª Etapa: o DNA genômico total do indivíduo é clivado com duas enzimas de restrição. A clivagem do DNA genômico origina três classes de fragmentos que diferem quanto às extremidades (fragmentos grandes, fragmentos pequenos

e fragmentos de tamanho intermediário), em decorrência de esta clivagem ser realizada com uma enzima de corte freqüente;

2ª Etapa: são ligados aos terminais dos fragmentos genômicos, originados da clivagem adaptadores específicos;

3ª Etapa: os fragmentos gerados são amplificados seletivamente via PCR, empregando-se *primers* que reconhecem as seqüências nos adaptadores;

4ª Etapa: é realizada a separação dos fragmentos amplificados em gel de alta resolução.

Assim como nos marcadores RAPD, os marcadores AFLP apresentam dominância, não sendo possível a detecção de heterozigotos. Entretanto estão sendo desenvolvidos sistemas de leitura quantitativa das bandas, proporcionando a diferenciação dos genótipos homozigotos dos heterozigotos.

Os marcadores AFLP apresentam três vantagens importantes sobre os demais marcadores. A primeira grande vantagem é o grande número de fragmentos que são gerados e visualizados em um único gel, proporcionando uma amostragem ampla e simultânea de um genoma. Esta técnica apresenta também grande poder de detecção de variabilidade genética, pois explora o polimorfismo de presença e ausência de sítios de restrição como no ensaio de RFLP e a ocorrência ou não de amplificação a partir de seqüências arbitrárias como na técnica de RAPD. Além destas vantagens, apresenta maior robustez quando comparado com RAPD, uma vez que *primers* mais longos são utilizados na etapa de PCR, aumentando a especificidade da amplificação e evitando, conseqüentemente, a competição que ocorre durante a PCR na técnica de RAPD.

A principal limitação desta técnica, assim como a de RAPD, é a impossibilidade de diferenciar genótipos heterozigotos dos homozigotos. Esta técnica também apresenta custo superior ao de RAPD, pois exige uma maior quantidade de reagentes, maior número de equipamentos de biologia molecular, protocolos de extração de DNA mais elaborados, além de um maior número de etapas, como as etapas de digestão enzimática e ligação de adaptadores. O sucesso desta técnica está diretamente relacionada a qualidade do DNA, ou seja, um DNA de alto nível de pureza é necessário para que ocorra digestão completa pelas enzimas de restrição.

## 2.6. Minisatélites e Microsatélites

Os marcadores baseados em locos hipervariáveis de minisatélites e aqueles baseados na amplificação de microsatélites, se difundiram após a demonstração, no início dos anos 80, que os genomas eucariotos são densamente povoados por diferentes classes de seqüências repetidas, algumas mais complexas (minisatélites) e outras mais simples (microsatélites) (Tautz & Renz, 1984).

Wyman & Colute (1980) citados por Ferreira & Grattaplagia (1995) isolaram a primeira região hipervariável a partir de uma biblioteca genômica humana. Nos últimos anos, diversas seqüências hipervariáveis têm sido identificadas e clonadas em outros organismos, inclusive em plantas (Wu & Wu, 1989; Schmidt, 1991; Thomas et al., 1993 citados por Ferreira & Grattaplagia 1995).

Um loco hipervariável ou minisatélite é constituído de um número variável de seqüências idênticas repetidas lado a lado. Estas seqüências possuem de 15 a 100 pares de bases e são repetidas até 50 vezes em cada loco hipervariável. Esta classe de marcadores também é denominada como VNTR ou *Variable Number of Tandem Repeats*.

A obtenção e detecção destes tipos de marcadores têm princípio semelhante ao utilizado para RFLP. Da mesma forma que na técnica de RFLP, o DNA dos indivíduos é clivado com enzimas de restrição, separado por meio de eletroforese, imobilizado em membrana e detectado através de hibridação com sondas. A única diferença entre os marcadores RFLP e VNTR é o tipo de sonda empregada na detecção do polimorfismo de DNA. As sondas são constituídas de seqüências núcleo homólogas às seqüências repetidas dos minisatélites, de forma que são detectados todos os locos hipervariáveis simultaneamente, sendo obtido um perfil complexo de bandas múltiplas na autoradiografia, ao invés de um padrão simples contendo uma banda para genótipos homozigotos ou duas para heterozigotos como na técnica de RFLP.

As vantagens e desvantagens dos marcadores baseados em locos hipervariáveis são semelhantes àquelas enumeradas para RFLP. Porém, a grande vantagem destes marcadores em relação ao RFLP, é a amostragem de diversas regiões do genoma simultaneamente, pois explora o polimorfismo de comprimento dos fragmentos de restrição (RFLP) em cada loco hipervariável e o polimorfismo no número e distribuição destes locos ao longo do genoma. Isto resulta em um complexo padrão de bandas,

fazendo com que todo indivíduo possua um perfil de bandas diferentes, semelhante a um código de barras de embalagens de produtos industrializados, surgindo daí o conceito de impressão digital genética ou *genetic fingerprint*.

Esta vantagem apresentada por este marcador, torna-se uma desvantagem quando vista por outro ângulo, pois é impossível identificar as bandas (alelos) correspondentes a cada loco, o que compromete a determinação de genótipos e o teste de hipóteses de segregação Mendeliana.

As seqüências simples repetidas ou microsátélites constituem pequenas seqüências com 1 a 4 nucleotídeos de comprimento, repetidas em *tandem*. Estas seqüências simples são mais freqüentes, melhor distribuídas ao acaso e formam locos genéticos mais polimórficos do que os locos hipervariáveis constituídos por minisátélites. Os sítios de microsátélites marcados por seqüência, ou seja, locos altamente polimórficos amplificados via PCR, denominados de STMS - *Sequence Tagged Microsatellite Sites*, representam a classe mais polimórfica de marcadores moleculares atualmente.

A detecção de marcadores microsátélites baseia-se na amplificação individual através de PCR de regiões contendo seqüências simples repetidas, utilizando um par de *primers* específicos e complementares as seqüências únicas que flanqueiam o microsátélite. Cada segmento amplificado de tamanho diferente equivale a um alelo diferente do mesmo loco.

Como apresentam expressão co-dominante e multialelismo, esta classe de marcadores é a que apresenta o mais elevado conteúdo de informações de polimorfismo, favorecendo a mais completa cobertura do genoma.

Esta técnica, no entanto, apresenta uma grande desvantagem que é a necessidade do desenvolvimento prévio dos marcadores, o que exige muito trabalho, mão-de-obra especializada, equipamento sofisticado para sequenciamento automático e elevado custo de implantação.

### 3. Marcadores moleculares na genética de populações

Neste ramo da genética, os marcadores moleculares têm sido empregados freqüentemente no estudo da estrutura genética das populações.

A estrutura genética das populações refere-se a distribuição da variabilidade genética entre e dentro de populações. A formação da estrutura é resultante de fatores como sistema de acasalamento, níveis de endogamia, seleção natural ou artificial, fluxo gênico e deriva genética entre e dentro das populações.

O número de alelos por loco, a porcentagem de locos polimórficos e a heterozigidade têm sido utilizados como índices de diversidade para caracterizar e comparar os níveis de variação genética nas populações.

O número de alelos por loco, consiste do número total de alelos detectados na população divididos pelo número de locos polimórficos.

A porcentagem de locos polimórficos é uma medida da variação genética que depende do tamanho da amostra, uma vez que se baseia na freqüência alélica. Ela é quantificada pela razão entre o número de locos polimórficos e o número total de locos da amostra. Para se determinar quando um loco é polimórfico ou não, é preciso a adoção de critérios quantitativos arbitrários, que classificam as freqüências alélicas de acordo com sua maior ou menor ocorrência na população, ou seja, se o alelo é mais ou menos comum. São usados dois critérios para definir se um loco é polimórfico ou não. O primeiro considera que a freqüência do alelo mais comum não deve ser maior que 0,99, e o outro considera que a freqüência do alelo mais comum não deve ser superior a 0,95 (Torggler et al., 1995).

A freqüência de heterozigotos, segundo Weir (1990), é um importante indicador da diversidade genética, pois cada heterozigoto carrega alelos diferentes, representando melhor a variação existente. Nei (1987) citado por Reis (1996) define a heterozigidade em termos de freqüência gênica. Para uma população de acasalamento ao acaso, a heterozigidade para um loco é definida como

$$h = 1 - \sum_{i=1}^m x_i^2$$

onde  $X_i$  é a frequência do  $i$ -ésimo alelo para um determinado loco e  $m$  é o número de alelos. A heterozigosidade média ( $H$ ) é a média das quantidades  $h$  para todos os locos, ou seja

$$H = \sum_{k=1}^n \frac{h_k}{r}$$

onde  $h_k$  é a estimativa da heterozigosidade do  $k$ -ésimo loco e  $r$  é o número de locos examinados.

Em indivíduos haplóides ou em diplóides, nos quais as frequências genotípicas desviam das proporções do equilíbrio, a melhor estimativa é a diversidade genética.

A diversidade genética de uma população subdividida pode ser analisada, segundo o método de Nei (1973), em diversidade entre e dentro de subpopulações. Este método pode ser aplicado em qualquer população e desconsidera o número de alelos por loco, as forças evolucionárias como mutação, seleção e migração, o sistema de acasalamento e o tipo de organismo em estudo. O método tem como base a identidade de duas amostras de genes dentro e entre populações. Considerando um loco, a frequência do alelo  $k$  na população é  $x_k$ . A probabilidade de identidade de dois genes é

$$J = \sum_k x_k^2$$

e a probabilidade de não identidade, que é a heterozigosidade ou diversidade é

$$H = 1 - J$$

Quando a população é subdividida em  $s$  subpopulações, é possível estabelecer a relação

$$J_T = J_S - D_{ST}$$

em que  $J_T$  e  $J_S$  são as identidades genéticas médias total e dentro das subpopulações, respectivamente; e  $D_{ST}$  é a diversidade genética média entre subpopulações.

A diversidade genética média na população total pode ser subdividida nas diversidades médias dentro e entre subpopulações

$$H_T = H_S + D_{ST}$$

onde  $H_T = 1 - J_T$  e  $H_S = 1 - J_S$ .

A diversidade entre subpopulações relativa à população total é obtida por

$$G_{ST} = \frac{D_{ST}}{H_T}$$

De acordo com as expressões apresentadas é possível obter a relação

$$1 - J_S = (1 - G_{ST})(1 - J_T)$$

Esta expressão é análoga a apresentada por Wright (1965) para as estatísticas  $F$ , sendo diferentes apenas nas interpretações. O índice de fixação de Wright, também chamado



coeficiente de endogamia, é empregado para quantificar os efeitos da endogamia em termos de redução na heterozigosidade. Na expressão de Wright (1965)

$$1 - F_{IT} = (1 - F_{IS})(1 - F_{ST})$$

$F_{IT}$  representa a correlação entre os gametas que se unem para produzir os indivíduos em relação aos gametas da população total;  $F_{IS}$  representa a média de correlações, sendo cada uma delas proveniente de gametas que se unem em cada subpopulação em relação aos gametas desta subpopulação e  $F_{ST}$  representa a correlação entre os gametas ao acaso dentro das subpopulações em relação aos gametas da população total.

Nei (1977) citado por Reis (1996) demonstra que as estatísticas  $F$  não precisam ser definidas como correlação entre gametas. Elas podem ser definidas como uma função de heterozigosidade esperada ( $H_S$ ) e observada ( $H_O$ )

$$F_{IS} = \frac{H_S - H_O}{H_S}, \quad F_{IT} = \frac{H_T - H_O}{H_T}, \quad F_{ST} = \frac{H_T - H_S}{H_T} \quad e \quad H_T = 1 - \sum p_k^{-2}$$

O índice  $F_{ST}$  de Wright é semelhante ao coeficiente de diferenciação genética  $G_{ST}$  de Nei.

Cockerham (1969) baseando-se na análise de variância das freqüências gênicas em populações experimentais ou naturais, utiliza as correlações como medidas de probabilidade de identidade por descendência. Considerando o modelo linear

$$x_{kij} = p + a_k + b_{ki} + m_{kij}$$

onde:

$x_{kij}$  é o valor observado do gene  $j$ , do indivíduo  $i$ , pertencente ao grupo  $k$ , que vale 1 se o alelo é **A** e vale 0 se o alelo é **a**;

$p$  é uma constante associada a todas observações;

$a_k$  é o efeito do grupo  $k$ ;

$b_{ki}$  é o efeito do indivíduo  $i$  dentro do grupo  $k$ ;

$w_{kij}$  é o efeito do gene  $j$  dentro de cada indivíduo.

Considerando que os parâmetros  $a_k, b_{ki}$  e  $w_{kij}$  são aleatórios, não correlacionados e com variâncias, respectivamente  $\sigma_a^2, \sigma_b^2$  e  $\sigma_w^2$ , as esperanças matemáticas dos quadrados dos termos são

$$\begin{aligned} E(x_{kij}, x_{k'i'j'}) &= p^2 + \sigma^2 && \text{se } k = k', i = i' \text{ e } j = j' \\ E(x_{kij}, x_{k'i'j'}) &= p^2 + Cov_{ab} && \text{se } k = k', i = i' \text{ e } j \neq j' \\ E(x_{kij}, x_{k'i'j'}) &= p^2 + Cov_a && \text{se } k = k', i \neq i' \\ E(x_{kij}, x_{k'i'j'}) &= p^2 + Cov_b && \text{se } k \neq k' \end{aligned}$$

Como os grupos são não correlacionados então a  $Cov_g = 0$  e em termos de correlações tem-se

$$\sigma^2 = p(1-p), Cov_{ab} = \rho_{ab} p(1-p) \text{ e } Cov_a = \rho_a p(1-p)$$

As correlações acima estão associadas aos componentes de variância, desta forma

$$\begin{aligned} (1 - \rho_{ab}) p(1-p) &= \sigma_w^2, \\ (\rho_{ab} - \rho_a) p(1-p) &= \sigma_b^2, \\ \rho_{ab} p(1-p) &= \sigma_a^2 \text{ e} \\ \sigma^2 &= \sigma_w^2 + \sigma_b^2 + \sigma_a^2 = p(1-p) \end{aligned}$$

Verifica-se que as correlações intraclasse, as covariâncias e os componentes de variância representam parametrizações equivalentes.

Estes resultados podem ser expressos em termos de  $F$  e  $\bar{\theta}$ , se a variação é originada devido à deriva e ao sistema de acasalamento. A correlação entre freqüências gênicas de diferentes indivíduos no mesmo grupo ( $\rho_a$ ), entre genes dentro de indivíduos de diferentes grupos ( $\rho_{ab}$ ) e entre genes dentro de indivíduos dentro de grupos ( $\rho_b$ ), são respectivamente:

$$\rho_a = \frac{\sigma_a^2}{\sigma_w^2 + \sigma_b^2 + \sigma_a^2} = \bar{\theta}$$

$$\rho_{ab} = \frac{\sigma_b^2 + \sigma_a^2}{\sigma_w^2 + \sigma_b^2 + \sigma_a^2} = F$$

$$\rho_b = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2} = \frac{F - \bar{\theta}}{1 - \bar{\theta}}$$

Verifica-se equivalência entre as correlações propostas por Cockerham (1969) e os índices de fixação de Wright (1965), ou seja

$$\rho_a = \bar{\theta} = F_{ST}, \rho_{ab} = F = F_{IT} \quad \text{e} \quad \rho_b = f = F_{IS}$$

Vencovsky (1992) descreve a estrutura genética de populações, tendo como base a técnica de análise de variância com frequências gênicas. Os dados de marcadores moleculares devem ser obtidos de acordo com uma estrutura hierárquica, onde devem ser consideradas uma ou mais populações para o estudo; devem ser colhidas sementes de plantas mãe, que são mantidas separadamente constituindo progênies e os marcadores são aplicados planta a planta, permitindo um número arbitrário de locos e de alelos por loco.

A frequência da variável indicadora  $x$  vale um, quando um dado alelo, por exemplo  $A_1$ , de um loco está na planta e vale zero quando este alelo está ausente. Considerando  $r$  populações,  $m$  progênies por população,  $n$  plantas por progênies e dois genes de um dado loco por planta, o modelo hierárquico para cada frequência alélica é

$$y_{ijkl} = \mu + p_i + f_{j(i)} + i_{k(ij)} + g_{l(ijk)}$$

onde:

$y_{ijkl}$  é a frequência do gene  $l$ , dentro do indivíduo  $k$ , da família  $j$ , pertencente à população  $i$ ;

$\mu$  é uma constante associada a todas observações;

$p_i$  é o efeito da população  $i$ , com  $i = 1, 2, \dots, r$ ;

$f_j(i)$  é o efeito da família  $j$  dentro da população  $i$ , com  $j = 1, 2, \dots, m$ ;

$i_k(ij)$  é o efeito do indivíduo  $k$  dentro da família  $j$  e da população  $i$ , com  $k = 1, 2, \dots, n$ ;

$g_l(ijk)$  o efeito do gene  $l$  dentro do indivíduo  $k$ , dentro da família  $j$  e da população  $i$ , com  $l = 1, 2$ .

Este modelo pressupõe que as entidades da estrutura hierárquica são obtidas por amostragem e admite que a diversidade entre as populações foi provocada por deriva genética ocorrida nas gerações anteriores. O esquema da análise de variância esta esquematizada na Tabela 2.

**Tabela 2.** Análise de variância das freqüências de um alelo de um loco.

Fontes de Variação	G.L.	S.Q.	Q.M.	E (Q.M.)
Populações (P)	$r - 1$	$S_P$	$Q_P$	$\sigma_G^2 + 2\sigma_I^2 + 2n\sigma_F^2 + 2nm\sigma_P^2$
Família (F)/P	$r(m - 1)$	$S_F$	$Q_F$	$\sigma_G^2 + 2\sigma_I^2 + 2n\sigma_F^2$
Indivíduos (I)/F	$rm(n - 1)$	$S_I$	$Q_I$	$\sigma_G^2 + 2\sigma_I^2$
Genes (G)/I	$rmn$	$S_G$	$Q_G$	$\sigma_G^2$
Total	$2rmn - 1$	$S_T$		

A variância total da freqüência alélica é

$$\sigma_T^2 = \sigma_P^2 + \sigma_F^2 + \sigma_I^2 + \sigma_G^2 = p(1 - p)$$

sendo:

$\sigma_P^2$  a variância entre populações;

$\sigma_F^2$  a variância entre famílias;

$\sigma_I^2$  a variância entre indivíduos;

$\sigma_G^2$  a variância entre genes.

Existem as seguintes relações paramétricas:

$$\begin{aligned}\sigma_P^2 &= p(1-p)\theta_2, \\ \sigma_F^2 &= p(1-p)(\theta_1 - \theta_2), \\ \sigma_I^2 &= p(1-p)(F - \theta_1), \\ \sigma_G^2 &= p(1-p)(1 - F), \\ f &= \frac{F - \theta_2}{1 - \theta_2}\end{aligned}$$

onde:

$p$  é a frequência de um gene em um determinado loco;

$F = F_{IT}$  é o coeficiente médio de endogamia de todas as plantas nas populações analisadas, ao nível de espécie;

$f = F_{IS}$  é o coeficiente médio de endogamia dentro das populações;

$\theta_1$  é o coeficiente de parentesco das plantas dentro das famílias;

$\theta_2 = F_{ST}$  é a medida da distância genética entre as populações estudadas, correspondendo também ao parentesco entre as plantas de famílias diferentes dentro das populações.

As taxas aparentes de autofecundação ( $s$ ) ou de fecundação cruzada ( $t$ ) relacionam-se com  $f$ , quando as populações estão em equilíbrio com endocruzamento e também quando os marcadores moleculares não sofrem seleção

$$f = \frac{s}{2-s} = \frac{1-t}{1+t}$$

O método dos momentos pode ser empregado para obter as estimativas dos componentes de variância e dos parâmetros genéticos, a partir da análise de variância apresentada na Tabela 2, obtendo-se as expressões

$$\hat{\sigma}_P^2 = \frac{1}{2mn}(Q_P - Q_F),$$

$$\hat{\sigma}_I^2 = \frac{1}{2}(Q_I - Q_G),$$

$$\hat{\sigma}_G^2 = Q_G,$$

$$\hat{\theta}_2 = \frac{\hat{\sigma}_P^2}{\hat{\sigma}_I^2},$$

$$\hat{\theta}_1 - \hat{\theta}_2 = \frac{\hat{\sigma}_F^2}{\hat{\sigma}_I^2},$$

$$\hat{F} - \hat{\theta}_1 = \frac{\hat{\sigma}_I^2}{\hat{\sigma}_I^2} \quad e$$

$$1 - \hat{F} = \frac{\hat{\sigma}_G^2}{\hat{\sigma}_I^2},$$

A estimativa da taxa aparente de fecundação cruzada é

$$\hat{t} = \frac{1 - \hat{f}}{1 + \hat{f}}$$

O sistema de acasalamento caracteriza o modo de transmissão dos genes de uma geração para outra. Sua determinação proporciona o conhecimento dos padrões de parentesco entre os indivíduos que se cruzam dentro de uma mesma população.

Segundo Ritland (1983) e Shaw & Allard (1982) citados por Torggler et al. (1995), a determinação das taxas de autofecundação em populações de espécies florestais e em algumas plantas cultivadas é fundamental, porque a mesma pode ter efeitos deletérios ou causar depressão por endogamia em características de importância econômica ou adaptativa. Eles enfatizam, ainda, a importância de se detectar a endogamia nos experimentos designados a estimar componentes de variância, nos quais normalmente se assume que os indivíduos dentro das progênes de polinização aberta são meios-irmãos. Porém, quando ocorre endogamia, o grau de parentesco entre os indivíduos pode ser maior que aquele esperado para meios-irmãos, resultando em estimativas enganosas dos parâmetros genéticos da população.

A principal consequência da endogamia em termos genéticos é o aumento da probabilidade de indivíduos que se cruzam carregarem alelos semelhantes, em decorrência da descendência de um ancestral comum, ocasionando uma redução da variabilidade genética dentro de famílias e um aumento da variabilidade genética entre as famílias (Allard et al., 1968 citado por Torggler et al., 1995).

A determinação do sistema de acasalamento em plantas, empregando dados de marcadores moleculares é feita de acordo com duas metodologias. A primeira se baseia no uso do coeficiente de endogamia ( $f$ ) de Wright

$$t = \frac{1 - f}{1 + f}$$

semelhante ao obtido por Vencovsky (1992).

A outra metodologia consiste no emprego de modelos multilocos. Segundo Reis (1996) o emprego de modelos multilocos permite a obtenção de estimativas mais adequadas da taxa de cruzamento, porque leva em consideração as combinações genóticas envolvendo todos os locos. De acordo com este autor, o modelo proposto por Ritland & Jain (1981) tem sido o mais utilizado nos trabalhos recentes. Este modelo considera que as progênes de um genótipo materno representam um conjunto de genótipos derivados de óvulos, que se cruzam em uma probabilidade  $t$ , com um conjunto de pólen com frequência alélicas  $p$  e se autofecundam em uma probabilidade  $(1 - t)$ . O processo de obtenção das estimativas tem implícito um equilíbrio de endogamia, equivalente ao equilíbrio de Wright, com  $f$  dependente de  $t$ .

Segundo Torggler et al. (1995) outra metodologia utilizada é a baseada na estrutura de famílias. A determinação dos locos marcadores polimórficos é feita através da distribuição dos genótipos entre progênes originadas de um parental materno. A obtenção das informações é baseada na hipótese de que a progênie derivada de autofecundação seja menos heterozigota que a derivada de fecundação cruzada. Este método fornece maior precisão estatística e está livre da suposição de equilíbrio de endogamia, pois o valor  $t$  (taxa de cruzamento) é dado em função das proporções observadas dos genótipos das progênes de fecundação cruzada, ou seja, progênie do tipo  $AA = (1 - t) + t_p$  e progênie  $Aa = t(1 - p)$ , sendo variável, uma vez que se aumentar a taxa de cruzamento  $t$ , aumenta-se a produção de progênes heterozigotas  $Aa$ .

Conforme o exposto, é possível verificar que a metodologia desenvolvida por Nei complementa as desenvolvidas por Wright, Cockerham e Vencovsky. Por meio da metodologia desenvolvida por Nei é possível detectar quanto da diversidade genética é devida a variação dentro e entre as populações ou subpopulações, enquanto as outras metodologias, também fornecem a taxa de autofecundação e de fecundação cruzada das espécies. Portanto, é aconselhável, que nas análises estatísticas de trabalhos científicos sejam utilizadas tanto a metodologia de Nei quanto as outras, a fim de que se tenha uma visão mais ampla da estrutura genética das populações estudadas.

Diversos trabalhos científicos têm sido desenvolvidos utilizando estas metodologias.

Hamrick & Godt (1990) citados por Reis (1996), analisaram 449 espécies vegetais de 165 gêneros, objetivando associar níveis de diversidade a aspectos da história de vida e ecologia de espécies vegetais. Os autores detectaram que, em média, 50,5% dos locos foram polimórficos e a heterozigosidade média foi de 0,113. Concluíram que o sistema reprodutivo e a distribuição geográfica das espécies foram os fatores que mais contribuíram para a variação dos dados. Ao utilizarem as estatísticas de Nei, constataram que em espécies perenes de vida longa, com fecundação cruzada e de fase final de sucessão, ocorreu maior variação dentro do que entre populações.

Moraes (1992) estudou a variabilidade e diversidade genética em populações naturais de aroeira (*Myracrodruon urundeuva*), utilizando a divergência genética entre duas populações classificadas hierarquicamente, através da técnica de eletroforese e por caracteres quantitativos. Os resultados indicaram alta taxa de autofecundação e baixa taxa aparente de fertilização cruzada nas populações, indicando a predominância de cruzamento entre indivíduos aparentados. Detectou, ainda, que a maior parte da variação genética está dentro das populações e a menor entre as populações. Concluiu que a utilização dos métodos quantitativos e de eletroforese mostram-se coerentes e complementares para o estudo genético de populações naturais de aroeira.

Paiva (1992) fez inferências sobre a estrutura genética de populações naturais de seringueira (*Hevea brasiliensis*), a fim de orientar o processo de amostragem na coleta e conservação de germoplasma. A variabilidade e a diversidade genética das populações foi determinada de acordo com uma análise de variância, considerando modelo desbalanceado em classificação hierárquica com três níveis de hierarquia, para cada alelo



identificado. De acordo com os resultados obtidos foi possível verificar que as populações apresentaram distâncias genéticas bem semelhantes.

Mori (1993) estudou a variabilidade genética em uma população de eucalipto (*Eucalyptus grandis*), com diferentes intensidades de seleção, empregando análise de variância semelhante à utilizada por Paiva (1992) e Moraes (1992). Detectou pequena distância genética entre as subpopulações e fixação de alelos em locos polimórficos na subpopulação composta por menor número de clones.

Muniz (1994), com o objetivo de estudar a análise de variância de freqüências gênicas que visam caracterizar a estrutura genética de populações de indivíduos diplóides, estabeleceu as propriedades dos estimadores dos parâmetros genéticos obtidos pelo método dos momentos, assim como a distribuição do quociente entre quadrados médios tentando propor um critério para testar a nulidade do coeficiente de endogamia. Detectou que os estimadores dos coeficientes de endogamia e da taxa de fecundação cruzada foram tendenciosos, sendo que as expressões das tendências apresentaram-se em função de  $1/n$ , tornando-se desprezíveis com o aumento do tamanho da amostra ( $n$ ). As fórmulas propostas para estimação da variância do coeficiente de endogamia e do estimador da taxa de fecundação cruzada apresentaram resultados satisfatórios quando o coeficiente de endogamia da população foi inferior a 0,5; a freqüência gênica estava entre 0,3 e 0,7 e o tamanho da amostra superior a 30 indivíduos. A estimação do coeficiente de endogamia em populações com alelos múltiplos, usando a análise conjunta com todos os alelos, foi menos tendenciosa que a estimativa através da média das estimativas nas análises de cada alelo, embora apresentassem a mesma variância. Foram apresentados resultados também para estimação destes coeficientes no caso de amostras de indivíduos de populações diferentes.

Marmey et al. (1994) avaliaram três espécies de mandioca (dois acessos de *Manihot glaziovii*, quatro de *M. esculenta* e dois de *M. caerulescens* Pohl), objetivando estudar a diversidade entre estas espécies. Foi detectado alto nível de polimorfismo tanto dentro quanto entre as espécies.

Simonsen & Heneen (1995) estudaram a variabilidade genética dentro e entre acessos de *Brassica campestris* e *Brassica oleracea* de diferentes origens. A heterozigosidade média para todos os acessos de *B. campestris* foi 0,21, para as raças chinesas foi 0,22, para as

cultivares chinesas foi 0,19, para as cultivares suecas foi 0,16 e para as variedades indianas foi 0,11. Os acessos de *B. oleracea* apresentaram menor nível de heterozigosidade (0,07), do que *B. campestris*. De acordo com a estatística *F* de Wright foi constatado que a diferenciação em *B. campestris* foi devida tanto a variabilidade dentro quanto entre acessos.

Nesbitt et al. (1995) ao analisarem quatro subespécies (*globulus*, *bicostata*, *pseudoglobulus* e *maidenii*) de *Eucalyptus globulus* de diferentes origens, detectaram maior variação dentro de locais. Alguns locais identificados previamente como intermediários entre as subespécies, com base em dados morfológicos, foram não intermediários com base em dados moleculares.

Isabel et al. (1995) ao compararem a diversidade genética em populações naturais de *Picea mariana* (Mill.), verificaram que há importantes níveis de diversidade genética e que esta é predominante dentro das populações.

Skroch & Nienhuis (1995) desenvolveram pesquisa, no qual um dos objetivos foi obter estimativas da diversidade genética e grau de parentesco entre 10 genótipos de feijão (*Phaseolus vulgaris*). A diversidade genética média foi 0,388; significando que, em média, o marcador RAPD detectou neste estudo somente 38,8% de segregantes entre todos os cruzamentos entre os 10 genótipos.

Reis (1996), realizou um trabalho visando gerar informações a respeito da distribuição e dinâmica dos alelos em populações naturais de palmito (*Euterpe edulis*), caracterizando os níveis e a distribuição da variabilidade genética, a estrutura genética das populações, o sistema reprodutivo e o fluxo gênico das mesmas. Os resultados revelaram níveis elevados de heterozigosidade tanto a partir das progênes quanto a partir de indivíduos adultos. Isto apresenta relevância, uma vez que as gerações posteriores poderão sempre apresentar novos recombinantes ou mais recombinantes, permitindo adaptação a microambientes e a manutenção da dinâmica populacional. A estrutura genética foi caracterizada a partir das estimativas das estatísticas *F* de Wright, dos coeficientes de coancestralidade ( $\theta$ ) de Cocherham e da análise de populações subdivididas de Nei. Os resultados indicaram menor divergência entre populações contíguas, sugerindo a existência de efeitos de deriva e/ou seleção entre regiões e locais, sendo as mesmas de menor efeito ou contrapostas pelo fluxo gênico entre as populações contíguas. Concluiu

que a forma básica de organização da variabilidade genética na espécie é o modelo de isolamento por distância. O sistema reprodutivo foi caracterizado pela avaliação da existência de panmixia e endogamia e pela estimativa da taxa de cruzamento. Os resultados evidenciaram que a cada ciclo reprodutivo predominam cruzamentos não aleatórios entre os indivíduos, produzindo progênies que não apresentam equilíbrio de panmixia ou endogamia, mesmo quando os indivíduos adultos estão em equilíbrio de panmixia. O fluxo gênico estimado foi elevado, o que leva à recomendação da manutenção de pelo menos 60 indivíduos reprodutivos por hectare nas populações sob manejo. A coerência dos resultados obtidos e a estratégia de amostragem empregada, proporcionaram a formação de um referencial, em relação aos níveis e distribuição da variação genética, para a coleta de germoplasma, a produção de sementes e o melhoramento da espécie, assim como para o monitoramento das populações naturais sob manejo sustentado.

O coeficiente de parentesco (CP) e estimativas da similaridade genética por RFLP (RFLP-GS) tem sido propostos para estimar a distância genética e estudar a diversidade dentro de espécies cultivadas. RFLP-GS é definido como a proporção de bandas RFLP idênticas, comparadas com um número total de bandas entre duas cultivares. Em soja, RFLPs têm sido eficientes em detectar diferenças genéticas no *pedigree*. Esta pesquisa foi realizada para investigar a relação de CP e RFLP-GS entre parentais e a variância genética entre suas progênies endogâmicas. Os resultados indicaram que o CP foi mais eficiente para predição da variância genética entre linhas endogâmicas do que RFLP-GS. A correlação entre CP e RFLP-GS para as cinco populações estudadas foi 0,45, não significativa estatisticamente. Esta baixa correlação foi devido a uma população que apresentou alto valor de CP e baixa estimativa de RFLP-GS. Os autores utilizaram os resultados obtidos para calcular também o ganho esperado com a seleção para as cinco populações. Os resultados indicam que as populações com baixo CP entre os parentais apresentam vantagens na seleção, ou seja, contribuem com uma maior porcentagem de linhagens superiores para produção. Os resultados validam a aplicação prática do coeficiente de parentesco na seleção de parentais para o desenvolvimento de populações com alta produtividade (Manjarrez-Sandoval et al., 1996b).

#### 4. Marcadores moleculares na genética quantitativa

Em plantas, características quantitativas, são aquelas controladas por vários genes de pequeno efeito, que agem conjuntamente, geralmente de forma aditiva, causando um tipo de variação fenotípica contínua, caracterizada pela presença de um gradiente de fenótipos, onde os intermediários são mais freqüentes que os extremos (superiores e inferiores) (Allard, 1971). São poucas as informações sobre o número, posição cromossômica, magnitude do efeito e interações dos locos que controlam a expressão destes caracteres. Os locos que controlam a expressão dos caracteres quantitativos são denominados QTL - *Quantitative Trait Loci* (locos controladores de características quantitativas).

Os estudos de herança poligênica e o progresso genético no melhoramento de plantas são estudados e estimados graças à teoria e as técnicas da Genética Quantitativa.

Entretanto, os genes que controlam estas características não podem ser estudados separadamente por meio das técnicas clássicas da Genética Quantitativa, pois, a mesma se baseia em modelos biométricos, nos quais os fatores (genéticos e ambientais) controladores destes caracteres são considerados em conjunto. A expressão fenotípica de um determinado caráter quantitativo é:  $F = m + g + e + ge$ , onde  $m$  é a média do valor numérico fenotípico na população estudada,  $g$  é o efeito genético,  $e$  é o efeito ambiental e  $ge$  o efeito da interação genótipo com o ambiente (Vencovsky, 1992).

No entanto, com o uso de marcadores moleculares é possível identificar e investigar os genes que controlam as características quantitativas; obter informações sobre a organização dos genomas, sobre a contribuição relativa dos genes maiores e menores (com efeitos grandes e pequenos) para a variação contínua e sobre aspectos como a epistasia, pleiotropia e a base genética da heterose (Edwards et al., 1987 citados por Torggler et al., 1995).

Segundo Ferreira & Grattapaglia (1995) a determinação da ligação genética entre marcadores e QTLs, depende da existência de desequilíbrio de ligação entre alelos no loco marcador e alelos do QTL. O desequilíbrio de ligação gera efeitos quantitativos associados ao marcador que podem ser detectados e estimados por meio de análises estatísticas apropriadas.

O desequilíbrio de ligação é decorrente de efeitos de ligação ou epistáticos entre os locos, resultando em desvios das freqüências gaméticas em relação as freqüências de equilíbrio. As probabilidades de ocorrência dos alelos nos gametas são:

$$\begin{aligned}p(ab) &= p(a) p(b) + D \\p(AB) &= p(A) p(B) + D \\p(aB) &= p(a) p(B) - D \\p(Ab) &= p(A) p(b) - D \\D &= p(ab) p(AB) - p(aB) p(Ab)\end{aligned}$$

Nestas expressões, temos que:

$p(ab)$  e  $p(AB)$  são as probabilidades de ocorrência de dois alelos em associação nos gametas (alelos recessivos e dominantes dos dois locos juntos no mesmo cromossomo);

$p(aB)$  e  $p(Ab)$  são as probabilidades de ocorrência de dois alelos em repulsão nos gametas (alelos recessivos de um loco e dominantes do outro loco juntos no mesmo cromossomo);

$p(a)$ ,  $p(b)$ ,  $p(A)$  e  $p(B)$  são as probabilidades de ocorrência dos alelos recessivos e dominantes dos dois locos, e

$D$  mede o desequilíbrio de ligação.

Segundo Lewontin (1974) citado por Torggler et al. (1995), quando:

$D = 0$  os locos estão em equilíbrio de ligação (não estão ligados) e as mudanças nas freqüências gênicas de um loco são independentes das que ocorrem no outro loco;

$D > 0$  existe excesso de gametas em associação ( $AB, ab$ );

$D < 0$  existe excesso de gametas em repulsão ( $Ab, aB$ )

e, nestes dois últimos casos, os locos são considerados ligados, não independentes e, portanto, em desequilíbrio de ligação.

Por meio do mapeamento de genes que controlam características quantitativas (QTLs) ou características quantitativas de importância econômica (ETL – *Economic Trait Loci*), é possível identificar e estudar os efeitos destes genes separadamente.

O mapeamento de QTLs e ETLs é realizado tendo como base populações segregantes, que pode ser uma população  $F_2$ , de retrocruzamento, linhagens puras recombinantes ou linhagens dihaplóides, obtidas a partir da geração  $F_1$ , resultante do cruzamento entre duas linhagens puras, geneticamente divergentes e fenotipicamente extremas para as características quantitativas de interesse. Centenas de indivíduos destas populações são avaliadas para as características e genotipadas para algumas dezenas de marcadores moleculares distribuídos a intervalos regulares (10 a 30 centiMorgans) ao longo do genoma. Este tipo de pedigree maximiza a quantidade de desequilíbrio de ligação entre marcadores e QTLs e/ou ETLs, sendo isto importante, uma vez que a detecção de um QTL está em função da magnitude do seu efeito sobre a característica, do tamanho da população segregante avaliada, da frequência de recombinação entre o marcador e o QTL e da herdabilidade do caráter (Ferreira & Grattapaglia, 1995).

Já foi proposta uma série de métodos de identificação de ligação ou associação entre alelos marcadores e QTLs. Estes são divididos em dois tipos: (1) Análise de um só ponto e (2) Análise por intervalo. Este primeiro método é o mais simples, pois são analisados os dados de um único marcador a cada vez. Contudo, quanto mais distante o QTL estiver do marcador, menor é a chance de se detectar associações estatisticamente significativas, em decorrência das recombinações entre os poligenes e o marcador serem mais frequentes, o que pode ocasionar falhas de classificação. Além disto, a magnitude do efeito de um QTL detectado é geralmente subestimada, devido também à recombinação entre o marcador e o poligene. Estes problemas podem ser minimizados aumentando-se o número de marcadores analisados, com distâncias menores a 15 cM (Tanksley, 1993 citado por Ferreira, 1996).

A análise por intervalo consiste na análise simultânea de grupos de marcadores ligados com relação aos seus efeitos na característica quantitativa. Com isto é possível compensar a ocorrência de recombinação entre o marcador com o QTL, aumentando a probabilidade de detectar o QTL e também de obter estimativas não tendenciosas dos seus efeitos genotípicos (Lander & Botstein, 1986, 1989 citados por Ferreira, 1996).

Segundo Ferreira & Grattapaglia (1995) a análise por intervalo, por minimizar o efeito da distância genética sobre a diferença quantitativa associada ao marcador, permite: (1) localizar com precisão o QTL no segmento; (2) estimar a magnitude do efeito fenotípico do QTL na forma de proporção da variância fenotípica na característica quantitativa explicada pelo QTL e (3) aumentar o poder estatístico de detecção e ao mesmo tempo limitar a ocorrência de falsos positivos (falsos QTLs).

Dentre as metodologias estatísticas propostas para detectar a ligação entre marcadores e QTLs, utilizando as estratégias de um só ponto ou por intervalo, os métodos mais comuns na literatura, segundo Ferreira (1996) são: (1) Análise de Variância; (2) Estimativas de máxima verossimilhança e (3) Regressão linear e não linear múltipla.

Segundo Dudley (1993) a forma mais simples de detectar ligação ou associação entre marcadores e QTLs é através da análise de variância. Os efeitos da ligação sobre a identificação de associações entre QTLs e marcadores pode ser ilustrado da seguinte forma: considerando o cruzamento  $M_1M_1T_1T_1 \times m_1m_1t_1t_1$ , onde  $M_1$  e  $m_1$  são alelos alternativos dos loco 1 marcador e  $T_1$  e  $t_1$  são alelos alternativos do QTL 1, e que  $M_1$  e  $T_1$  estão ligados com um valor de recombinação  $r$ . Então, na geração  $F_2$  deste cruzamento, os parentais homozigotos podem ser recuperados com a frequência de  $[0,5(1-r)]^2$  cada. Para identificar associações entre marcadores e QTLs compara-se as médias genotípicas dos homozigotos de cada classe de marcadores. Se a diferença ( $M_1M_1 - m_1m_1$ ) é significativa, então o marcador está ligado ao QTL. Assumindo que  $T_1T_1$  tem um valor genotípico  $a$ ,  $T_1t_1$  valor  $d$  e  $t_1t_1$  valor  $-a$ , a média genotípica do marcador na geração  $F_2$  será:

$$M_1M_1 = [a(1-2r) + 2r(1-r)d]$$

$$M_1m_1 = [d(1-2r + 2r^2)]$$

$$m_1m_1 = [-a(1-2r) + 2r(1-r)d]$$

$$M_1M_1 - m_1m_1 = 2a(1-2r)$$

A magnitude desta diferença ou contraste e, conseqüentemente, da capacidade de se detectar associações entre marcadores e QTLs está em função do valor de recombinação ( $r$ ) e da diferença genotípica ( $2a$ ) entre os homozigotos do QTL.

Lander & Botstein (1989) citados por Ferreira (1996) citam algumas falhas deste método: (1) se  $0 < r < 0,5$  o efeito genotípico é subestimado por um fator de  $(1 - 2r)$ ; (2) é exigido um grande número de progênies e (3) o método não define a posição do QTL. Além disto, não se pode distinguir entre uma forte ligação a um QTL com pequeno efeito e uma ligação em menor grau com um QTL de grande efeito.

O método da máxima verossimilhança consiste em formar uma função (máxima verossimilhança) para os parâmetros do modelo e com base na distribuição da variável (geralmente é a distribuição normal) que se deseja estimar, buscam-se as estimativas que a maximizem. A resolução analítica para os valores dos parâmetros exige métodos iterativos, os quais consomem tempo, principalmente quando se dispõe de grandes conjuntos de dados e numerosos parâmetros (Ferreira, 1996).

A detecção de ligação marcador-QTL por meio de regressão é realizada tendo como base mapas de ligação dos locos marcadores. São realizadas regressões lineares múltiplas para cada grupo de ligação. Moreno-Gonzalez (1992a, b; 1993) citado por Ferreira (1996) apresenta métodos de regressão linear múltipla para estimar a associação de marcadores com QTLs. Ele constatou que o método é mais eficiente para detectar QTLs não ligados e que para QTLs ligados é preciso um maior número de progênies.

Ferreira (1996) apresenta uma extensiva revisão bibliográfica sobre estes métodos, inclusive comparando-os entre si. Além disto, cita outros métodos desenvolvidos para detecção de associações entre marcadores e QTLs, como por exemplo o desenvolvido por Neumann (1990), que utilizou o teorema de Bayes. Pinheiro & Carneiro (2000) também apresentam uma consistente revisão sobre a análise de QTL no melhoramento de plantas.

Outras metodologias estatísticas têm sido estudadas e desenvolvidas com o objetivo de investigar os efeitos da epistasia e pleiotropia (Gibson, 1996), o desequilíbrio de ligação de locos microsátélites (Pritchard & Feldman, 1996), assim como sobre o comportamento do desequilíbrio de ligação sobre flutuações ambientais em populações panmíticas



infinitas com não sobreposição de gerações e regimes de seleção em diplóides e haplóides (Kirzhner et al., 1995).

De acordo com Ferreira & Grattaplgia (1995), apesar do número limitado de estudos de mapeamento de QTLs publicados, um número significativo de associações entre marcadores e QTLs tem sido detectado, sendo que alguns QTLs com grande efeito foram identificados, os quais explicam até 70% da variação fenotípica observada, assim como outros QTLs de menor efeito também foram identificados.

Com o objetivo de determinar se diferentes genótipos enzimáticos estão associados a diferentes características morfológicas, foram mensurados 20 caracteres quantitativos em indivíduos de duas populações de milho, sendo que uma foi selecionada massalmente para prolificidade e a outra para produção de grãos. Genótipos para locos enzimáticos foram determinados nos mesmos indivíduos. Análise discriminante foi empregada para discriminar genótipos enzimáticos em um determinado locus usando como variáveis discriminantes os caracteres morfológicos. Quando detectada discriminação, foi considerada a existência de associação entre genótipos enzimáticos e caracteres morfológicos. Funções discriminantes foram associados com produção, maturidade e fosfatase ácida 1 (Acp1) em ambas as populações. A produção total de grãos em uma das populações apresentou correlação com a primeira função discriminante, a qual foi responsável por 61,1% da variação. Na outra população, a segunda função discriminante apresentou correlação com maturidade e tamanho da planta (Pollak et al., 1984).

Empregando análise de variância, Garvey & Hewitt (1992) verificaram que seis isoenzimas apresentaram associações significativas com teor de sólidos solúveis ( $^{\circ}$ Brix) em tomate. Efeitos epistáticos entre QTLs que condicionam brix não foram detectados. Neste estudo foram utilizadas três espécies de tomate: *Lycopersicon cheesmanii* (apresenta alto Brix) foi cruzada com *L. pennellii* (apresenta baixo Brix) e as plantas  $F_1$  foram cruzadas com *L. esculentum* (tomate cultivado, que apresenta Brix baixo). Já Eshed & Zamir (1994) observaram três regiões de QTLs nos cromossomos 1, 5 e 7 de tomate controlando o teor de sólidos solúveis totais.

Prince et al. (1993) obtiveram 19 grupos de ligação com 192 marcadores moleculares RFLP e isoenzimas, ao contruírem o mapa molecular de pimenta (*Capsicum* spp.) utilizando população  $F_2$  do cruzamento interespecífico entre *C. annum* cv. NuMex e *C.*

*chinense*. Observaram dois QTLs controlando número de flores por nó no grupo de ligação 10.

Lindhout et al. (1994) identificaram três locos marcadores RFLP associados com precocidade em tomate. Um deles estava associado principalmente com tempo de florescimento, outro com frutificação e o outro com tempo para colheita. Dois destes três locos estão também associados com peso de frutos.

Geralmente resistência à doenças é uma característica monogênica, ou seja, controlada por um gene. Entretanto, em ervilha (*Pisum sativum*), Dirlewanger et al. (1994) detectou QTLs para resistência a raça C de *Ascochyta pisi*. No cromossomo 4 foi detectado o QTL com maior efeito.

Velboom & Lee (1994) estudaram uma população de milho (*Zea mays* L.) composta por 150 linhas F<sub>2:3</sub> obtidas do cruzamento entre duas linhagens endogâmicas do grupo heterótico Lancaster, objetivando identificar QTLs associados com produção de grãos e todos seus componentes por meio de marcadores RFLP. Foram detectados QTLs para produção de grãos e seus componentes, sendo que apenas um QTL foi responsável por 35% da variação. Foram detectados, ainda, algumas regiões com pleiotropismo ou efeitos de ligação sobre alguns componentes da produção. A ação gênica foi de dominância parcial a sobredominância para todos os caracteres. Alguns QTLs exibiram sobredominância ou “pseudo-sobredominância”, indicando possível ligação de dois ou mais genes em repulsão.

Um cruzamento entre uma linhagem (GY 14) e um germoplasma (PI 432860) de pepino (*Cucumis sativus* var. *sativus*) foi realizado e marcadores moleculares foram utilizados para determinar o número, magnitudes de efeitos e variação total descritos pelos genes condicionados por herança quantitativa em relação aos caracteres comprimento, diâmetro, tamanho da cavidade de sementes, cor, relação comprimento/diâmetro e relação tamanho da cavidade de sementes/diâmetro. QTLs foram detectados por meio do programa MAPMAKER/QTL, usando 100 linhas F<sub>3</sub> avaliadas em experimentos realizados em um local por dois anos. Foram construídos modelos multilocus. Geralmente o mesmo loci estava associado com QTLs. Herdabilidades de QTLs individuais foram confirmados pela análise de retrocruzamento (67 linhas BC<sub>1</sub>P<sub>1</sub> e 68 linhas BC<sub>1</sub>P<sub>2</sub>). A maioria dos QTLs foram confirmados em pelo menos uma população de retrocruzamento. Foram calculados

as variâncias aditivas totais e herdabilidades usando o delineamento III da Carolina do Norte (delineamento NCIII) e as estimativas foram comparadas com as variâncias totais estimadas por meio dos marcadores. Os resultados indicaram que os marcadores foram responsáveis por uma grande variação, mas não toda, da variância aditiva estimada pela covariância entre parentes. A utilização de retrocruzamentos para confirmar QTLs individuais e a variância total descrita pelos QTLs é recomendada para evitar falsos positivos ou superestimativas de efeitos. O número de QTLs e/ou a proporção da variância fenotípica descrita pelos marcadores e pelos delineamentos estão de acordo com relatos anteriores de herdabilidades empregadas em germoplasma similar (Kennard & Havey, 1995).

Por meio de análise com marcadores RFLP, setenta e três associações diferentes foram detectadas entre marcadores e QTLs para peso de fruto, sólidos solúveis e peso de sementes em um cruzamento interespecífico entre *Lycopersicon esculentum* e *L. cheesmanii*. No cromossomo 6 foi verificado um QTL de grande efeito para estas três características (Goldman et al., 1995).

Foram identificados três QTLs afetando tolerância ao frio em tomate, utilizando dados isoenzimáticos. Um dos QTLs apresentou efeito negativo para tolerância ao frio, ao passo que os outros dois apresentaram efeito positivo (Vallejos & Tanksley, 1983 citado por Cardoso, 1996).

As técnicas multivariadas são, também, muito utilizadas com dados de marcadores moleculares. Segundo Wilches (1995) citado por Duarte (1997) os procedimentos estatísticos multivariados mais empregados nestes estudos são as distâncias genéticas, índices de similaridade e dendogramas e as escalas multidimensionais. Estes estudos objetivam, em sua grande maioria, avaliar a variabilidade genética, agrupar germoplasmas, classificar e investigar a estrutura genética de populações, etc. Em relação aos trabalhos realizados na área de genética de populações, os mesmos estão relatados neste ítem desta revisão. Outros exemplos serão citados no próximo ítem, pois estão relacionados ao melhoramento de plantas.

## **5. Marcadores moleculares no melhoramento de plantas**

Os marcadores moleculares têm sido empregados no melhoramento de plantas para diversos fins, constituindo-se uma ferramenta muito importante e que, no futuro, será essencial em programas de melhoramento, em virtude das vantagens que apresentam, como por exemplo, maior rapidez na obtenção de resultados e maior confiabilidade dos dados obtidos.

Dentre as atividades ou áreas de pesquisas inseridas em programas de melhoramento que podem ser auxiliadas pelos marcadores moleculares, segundo Ferreira & Grattapaglia (1995), destacam-se:

- (1) Identificação e proteção de variedades e/ou clones patenteados;
- (2) Certificação de pureza genética de linhagens e híbridos;
- (3) Construção de coleções nucleares ("core collection") em bancos de germoplasmas;
- (4) Atribuição de linhagens a grupos heteróticos;
- (5) Mapeamento genético de QTLs e/ou ETLs;
- (6) Introgessão de características via retrocruzamentos assistido por marcadores;
- (7) Seleção e recombinação dirigida de genótipos superiores;
- (8) Seleção durante o desenvolvimento de linhagens endogâmicas;
- (9) Predição de fenótipos esperados;
- (10) Seleção indireta para características de difícil avaliação (resistência a fatores bióticos e abióticos);
- (11) Seleção precoce em cultivares perenes;
- (12) Monitoramento de fecundação cruzada e autofecundação em pomares de sementes florestais.

Diante da enorme quantidade de trabalhos realizados de acordo com revisão realizada no CAB-ABSTRACTS, apenas algumas exemplificações serão dadas para algumas destas áreas de pesquisa, a fim de que seja possível visualizar a eficiência do uso dos marcadores moleculares no melhoramento genético de plantas.

Com o advento do patenteamento de cultivares melhoradas, os marcadores moleculares ocupam lugar de destaque na área de identificação e proteção de variedades. É indiscutível a eficiência da caracterização de variedades, linhagens e híbridos, ao nível molecular. A verificação da pureza genética de linhagens ou híbridos é fundamental, também, no controle de qualidade por ocasião da produção de sementes comerciais.

Em espécies autógamas, por exemplo, a produção de híbridos é feita com emasculação e polinizações manuais, quando não se dispõe de fontes de macho-esterilidade. Estes procedimentos aumentam a probabilidade de contaminação, devido a autofecundação da planta mãe em decorrência da emasculação incompleta ou tardiamente realizada. Além disto, a ocorrência de cruzamentos naturais; a mistura física das sementes durante a colheita dos frutos; extração, beneficiamento e empacotamento das sementes, ocasionam contaminação genética dos lotes de sementes híbridas.

Segundo Bailey (1983) citado por Ferreira & Grattapaglia (1995) ao menos quatro critérios devem ser satisfeitos para a eficiente utilização dos marcadores moleculares na identificação de variedades: (1) variação inter-varietal distinta; (2) mínima variação intra-varietal, (3) estabilidade ambiental e (4) reproducibilidade experimental.

Smith & Smith (1991) desenvolveram um trabalho que tinha como objetivo investigar a capacidade de marcadores RFLP discriminar 78 híbridos de milho. A correlação entre as distâncias estimadas com base no pedigree e aquelas estimadas com base nos dados moleculares foi 0,92, indicando que marcadores RFLP são um meio prático de diferenciação de híbridos.

Marcadores RAPD foram utilizados com o objetivo de verificar sua eficiência na discriminação de cultivares de *Brassica napus*. Um grupo de 100 primers foram testados, destes 70 produziram bandas e 22 mostraram polimorfismo. Foi verificado que 6 primers são suficientes para diferenciar as cultivares e que RAPDs são eficientes na identificação e discriminação das cultivares (Mailer et al., 1994).

Rom et al. (1995) estudaram a possibilidade de se avaliar a pureza genética dos lotes de sementes híbridas por meio de marcadores RAPD. Verificaram que isto é possível, uma vez que primers detectaram polimorfismo nos genitores e nos híbridos testados. Paran et al. (1995) conseguiram detectar plantas contaminantes em quatro híbridos de tomate e, também, discriminar os híbridos entre si.

Híbridos de pimentão também foram identificados e discriminados por marcadores RFLP (Tanksley & Young, 1989) e por PCR (Liveneh et al., 1992), sendo que o último método é mais simples e pode ser utilizado em escala comercial no controle de qualidade durante a produção de híbridos de pimentão.

Marcadores RAPD-PCR foram empregados na identificação de híbridos de soja e de feijoeiro, derivados de cruzamentos entre genitores que não constatavam para características facilmente monitoráveis. Os marcadores OP-F12 e OP-CO3 geraram bandas polimórficas de DNA, presentes no genitor masculino e ausentes no genitor feminino, de feijão e soja, respectivamente. Cerca de 30% das plantas analisadas não apresentaram as bandas esperadas, portanto, devem ter sido resultantes de autofecundação. Desta forma, estes dois marcadores podem ser utilizados eficientemente na identificação e discriminação de híbridos de feijão e soja (Alzate-Marin et al., 1996).

Marcadores RAPD foram usados para determinar a frequência de polimorfismo de DNA em 39 germoplasmas de melancia [*Citrullus lanatus* (Thunb.) Mansf.]. A fração de bandas (F) comuns entre os genótipos foi estimada pela estatística de Nei e a dissimilaridade foi calculada subtraindo o valor de F por 1. O método UPGMA (Unweighted Pair Group Method with Arithmetic Mean) foi utilizado para determinar o parentesco genético entre os genótipos. Foram determinados quatro grupos. Foi determinado, também, que oito marcadores RAPD podem ser utilizados para discriminação varietal de melancia. Concluíram que RAPDs podem ser empregados para identificar e discriminar genótipos, assim como para detectar QTLs em melancia (Lee et al., 1996).

A divergência genética é avaliada para diversos fins, em programas de melhoramento, como por exemplo identificar amostras duplicadas em bancos de germoplasmas; eleger descritores essenciais para caracterização de acessos de germoplasmas; estabelecer coleções-núcleo ("core collection"); orientar a escolha de parentais a serem intercruzados

em programas de melhoramento, visando obter maior efeito heterótico nas populações híbridas.

Objetivando verificar a eficiência de marcadores RFLP para elucidar a heterose entre linhas endogâmicas de milho e a similaridade genética entre linhagens aparentadas e não aparentadas, Melchinger et al. (1991) analisaram 32 linhas endogâmicas de milho. As distâncias entre as linhagens, estimadas pela distância de Roger's (RD), revelaram considerável diversidade entre as linhagens dos grupos heteróticos BSSS (Iowa Stiff Stalk Synthetic), RYD (Reid Yellow Dent) e LSC (Lancaster Sure Crop). Os autores concluíram que dados de RFLP podem ser empregados para determinação de grupos heteróticos de linhagens e para quantificação da similaridade genética entre linhagens aparentadas.

Ajmone-Marsan et al. (1992) analisaram 31 linhagens endogâmicas de milho com 149 marcadores RFLP, com o objetivo de determinar a eficiência de RFLPs na estimação de similaridade genética entre 16 linhagens dos grupos heteróticos BSSS (Iowa Stiff Stalk Synthetic) e LSC (Lancaster Sure Crop) e comparar as similaridades genéticas baseadas em marcadores moleculares com as baseadas em informações do pedigree. Foram estimados coeficientes de similaridade genética (GS) e coancestralidade (f), entre pares de linhagens do mesmo grupo heterótico, com os dados de RFLPs e pedigrees, respectivamente. Para as linhagens do grupo heterótico BSSS, as análises de agrupamento baseados em dados de RFLPs e pedigrees foram similares (correlação entre GS e f de 0,70). Entretanto, para as linhagens do grupo heterótico LSC, consideráveis discrepâncias entre GS e f foram detectadas (correlação foi 0,07).

Pela análise de dados de marcadores RFLP, Nienhuis et al. (1993) estimaram a distância genética entre genótipos de *Brassica oleracea* (brócoli, couve-flor e repolho). A similaridade genética detectada entre os genótipos de cada grupo foram coincidentes com as informações de suas genealogias. Os autores argumentam que esta informação é importante em programas de melhoramento por permitir a realização de amostragem mais eficiente dos genótipos e realizar cruzamentos entre genótipos mais divergentes com a finalidade de ampliar a base genética, assim como para produzir híbridos com boas performances.

Objetivando determinar quais seriam as melhores combinações para formação de híbridos de *Brassica juncea*, Jain et al (1994) mensurou com RAPDs o distanciamento

genético entre 12 genótipos indianos e 11 exóticos. Verificaram baixo polimorfismo nos genótipos indianos e alto nos exóticos. Os genótipos foram agrupados em dois grupos, sendo um formado por apenas genótipos indianos e o outro pelos genótipos exóticos e quatro indianos. Foi observada elevada heterose nos cruzamentos entre os genótipos indianos e exóticos.

A principal causa de sabores indesejáveis nos produtos derivados de soja, acredita-se que sejam as isozimas lipoxigenases (LOX) das sementes. Foram identificados três mutantes (semLOX1, semLOX2 e semLOX3) deficientes dessas isozimas; e os mesmos foram eleitos para introduzir os alelos controladores dessas isozimas na variedade brasileira Cristalina. A técnica de RAPD foi empregada para caracterizar os genitores e algumas outras linhagens. Os resultados obtidos permitiram estabelecer “fingerprint” de cada genótipo e determinar a distância genética entre os mesmos. A distância genética relativa foi de 76% e foram definidos quatro grupos distintos (VilarinhoS et al., 1994).

Burstin et al. (1994) analisaram 21 linhas endogâmicas de milho usando isoenzimas, RFLPs e 2-D PAGE (two-dimensional electrophoresis of denatured proteins), com o objetivo de investigar o parentesco genético entre as linhas que são potenciais para produção de híbridos para silagem. O grupo de linhagens foi geneticamente divergente. O polimorfismo foi menor a nível de proteínas do que a nível de DNA.

Ali et al. (1995) observaram correlação alta entre distância genética e heterose em canola (*Brassica napus*) para diversas características importantes, dentre elas produtividade de sementes, número de vagens por planta e número de sementes por vagem.

Sekhon & Gupta (1995) estudaram o uso da distância genética entre parentais, estimada por dados isoenzimáticos, na predição da heterose em mostarda. Os resultados sugerem que a heterose é mais responsiva em cruzamentos derivados de parentais relacionados do que entre aqueles resultantes de parentais não relacionados, não estando de acordo com o que seria esperado.

Com base em dados de marcadores RAPD, Tomita et al. (1996) avaliaram acessos de limão verdadeiro e híbridos relacionados e verificaram que as principais variedades, com denominações distintas, constituem apenas um único clone ou acessos idênticos que sofreram mutações somáticas de difícil detecção por estes marcadores. Para medir a similaridade genética empregaram o coeficiente de Jaccard, o qual variou de 0,76 a 1,00.



Utilizando também o coeficiente de Jaccard para construir a matriz de similaridade, Coletta Filho et al. (1996) detectaram similaridade completa entre 14 acessos de tangerina Ponkan. Os resultados indicam tratar-se de um único clone, com denominações locais diferentes ou, também, mutações não detectadas pelos marcadores RAPD.

Amaral Jr. (1996) realizou o agrupamento de cultivares de tomate utilizando o método de agrupamento de Tocher, a partir de dados de marcadores RAPD e dados de marcadores morfoagronômicos, objetivando avaliar a divergência genética. Verificou que os agrupamentos a partir de dados de RAPD diferiram daqueles baseados em características morfoagronômicas e da qualidade de frutos, atribuindo isto à impossibilidade discriminatória dos marcadores RAPD utilizados, ou seja, os fragmentos heteromórficos devem ter sido amplificados em regiões nucleicas não envolvidas na expressão das características avaliadas, não permitindo, portanto, a descrição apropriada dos genomas.

Lanza et al. (1996a, 1996b) a partir de dados de marcadores RAPD, estimaram as distâncias genéticas entre linhagens obtidas das populações BR 105 e BR 106 de milho, as quais foram avaliadas a nível de campo para produção de grãos. Utilizando as estimativas das distâncias genéticas estimadas pelo coeficiente de similaridade de Jaccard, construíram um dendograma, empregando o critério UPGMA, obtendo três grupos. Constataram, desta forma, que os marcadores RAPD são eficientes na obtenção de grupos heteróticos e na predição de cruzamentos entre linhagens de milho.

A seleção assistida por marcadores moleculares é uma área que vem se expandindo nos últimos anos. Estudos que visam a identificação de associações entre marcadores moleculares e genes que conferem resistência em plantas são, talvez, os mais realizados e de grande importância para programas de melhoramento, uma vez que possibilitam a seleção de genótipos resistentes na fase de plântulas, diminuindo os custos dos experimentos, bem como acelerando os programas de melhoramento. Além disto, permitem o “screening” de um grande número de genótipos simultaneamente.

Reuveni & Bothma (1985) ao investigarem em melão (*Cucumis melo*) o quanto a atividade de peroxidases em plantas infectadas está associada a resistência ao oídio (*Sphaeroteca fuliginea*), observaram que a atividade foi superior em folhas de plantas resistentes, e que com a infecção esta atividade aumenta com o tempo, principalmente em folhas infectadas de cultivares suscetíveis. Isto indica que as plantas suscetíveis ativam sua peroxidase

após a infecção, sendo que esta reação é tardia para evitar o desenvolvimento da doença. Diante disto, pode-se concluir que a atividade de peroxidases pode ser empregada como marcador na seleção de genótipos resistentes de melão.

Williamson et al. (1994) detectaram a ligação entre o marcador de PCR (REX 1) com o gene de resistência a nematóide (Mi) em tomate (*L. peruvianum*). Este marcador foi testado até a geração F<sub>2</sub>, possibilitando a distinção de homozigotos de heterozigotos.

Foram identificados, também, dois marcadores RFLP ligados ao gene TM-2<sup>a</sup> que confere resistência ao TMV (vírus do mosaico do fumo) em tomate (Young et al., 1988). Porém, DAX et al. (1994) detectaram marcadores RAPD associados com a resistência a este mesmo vírus, utilizando o primer OPA 12. Como esta última técnica é mais barata e rápida que a primeira, é recomendável, portanto, o uso do marcador OPA 12 em programas de melhoramento que visem a obtenção de cultivares resistentes a este vírus.

Outros genes que conferem resistência em tomate já foram identificados, como por exemplo: marcadores RFLP (CT71 e CT220) ligados ao gene Sw-5 que confere resistência ao vírus TSWV (Stevens et al., 1995); localizado o gene Ty-1 para tolerância ao vírus TYLCV (Michelson et al., 1994; Zamir et al., 1994); marcadores RAPD associados ao gene Pto que condiciona resistência a *Pseudomonas syringae* pv *tomato* (MARTIN et al, 1991); marcadores RFLP associados aos genes Cf-9 e CF-4 que conferem resistência a *Clasporium fulvum* (Balint-Kurti et al., 1994); marcador isoenzimático Aat2 ligado ao gene I-3 que confere resistência a raça 3 de *Fusarium oxysporum* f. sp. *lycopersici* (Bournival et al, 1989); localização de um gene (Lv) para resistência a oídio (*Leveillula taurica*) por meio de marcadores RAPD e RFLP (Chunwongse et al, 1994).

Denny et al. (1996) apresentam resultados de experimentos com o objetivo de verificar a eficiência da seleção assistida por marcadores RFLP e RAPD na seleção de genótipos de soja resistentes a nematóides. Os autores discutem estes dois tipos de marcadores, inclusive relacionando a eficiência com o custo das técnicas.

Em melancia foi identificado o marcador isoenzimático associado ao gene que controla resistência ao vírus ZYMV (vírus do mosaio amarelo da abóbora) (Provvidenti, 1991) e em abóbora foi detectado o marcador isoenzimático (Aldo-p) associado ao gene de resistência ao vírus WMV-2 (vírus do mosaio da melancia-2) (Weeden et al., 1984).

Em outras hortaliças, também já foram localizados genes que conferem resistência a algumas doenças, como em alface (Farrara et al., 1987; Hulbert & Michelmore, 1985; Cole et al. 1991; Witsenboer et al., 1995); em cenoura (Ronfort et al., 1995), em ervilha (Dirlewanger et al., 1994) entre outras.

Navot et al. (1990) detectaram que a cor vermelha da polpa do fruto de melancia é determinado por um gene recessivo ligado ao marcador Gdh2 e que o sabor amargo é governado por um gene dominante ligado ao marcador Pgm1. Estes, portanto, podem ser utilizados em programas de melhoramento para seleção assistida de genótipos.

A macho esterilidade é um fenômeno muito importante e utilizado em programas de melhoramento, por facilitar a obtenção de híbridos em espécies onde cruzamentos manuais são de difícil execução. Lorenz et al. (1994) submeteram o DNA nuclear, mitocondrial (mt) e do cloroplasto (cp) de duas linhas quase isogênicas de beterraba (*Beta vulgaris*), sendo uma macho estéril e a outra fértil, à análises com RAPD.

Detectaram diferenças nos padrões de “fingerprinting” do mtDNA entre plantas macho estéreis e férteis, evidenciando os resultados experimentais que verificaram rearranjos no genoma mitocôndrial como fonte primária de macho esterilidade. Entretanto, eles aconselham o uso de cpDNA na identificação de linhas macho estéreis ou férteis.

Por outro lado, Steinborn et al. (1995) ao estudarem o DNA das mitocôndrias e cloroplastos de cenoura (*Daucus carota* subsp. *sativus*), por meio de marcadores RFLP, detectaram que mtDNA apresentou herança materna, ao passo que cpDNA apresentou também herança paterna.

O emprego de marcadores moleculares em programas de introgressão via retrocruzamentos é, segundo Ferreira & Grattapaglia (1995), a aplicação mais consistente da tecnologia de marcadores moleculares no melhoramento.

Neste contexto, os marcadores moleculares devem estar associados aos genes que se deseja introgridir, tornando possível monitorá-los no decorrer das gerações de retrocruzamento. Eles podem ser utilizados, ainda, na seleção de indivíduos que possuem além do gene em introgressão, a maior proporção do genoma recorrente. Estes procedimentos reduzem o número de gerações de retrocruzamento.

Vantoi et al. (1996) monitoraram ciclos de seleção recorrente utilizando marcadores AFLP em soja. A frequência de cada banda polimórfica entre os ciclos de seleção foi

determinada como:  $y = \frac{(D + H)}{N}$ , onde  $D$  e  $H$  são, respectivamente, o número de homozigotos e heterozigotos selecionados para a banda, e  $N$  o número total de selecionados. As contribuições  $(p_1, p_2, \dots, p_6)$  de cada linha parental para os ciclos de seleção, foi estimada pela equação:  $y = X\beta + \epsilon$ , onde  $y$  é o vetor das frequências de cada banda na seleção,  $\beta$  é o vetor dos valores  $p_j$  a serem estimados e  $\epsilon$  é o vetor do erro. Foi possível verificar que a contribuição dos dois parentais adaptados, após dois ciclos de seleção recorrente não mudou e que 25% do pool gênico das populações C0 e C1 foram devidos aos genótipos introduzidos (PI).

## 6. Considerações finais

É possível verificar por meio desta revisão que estudos sobre a aplicação de técnicas de marcadores moleculares na genética de populações, quantitativa e no melhoramento de plantas, vêm sendo realizados intensivamente, sendo que alguns resultados consistentes têm sido obtidos, proporcionando o emprego destas técnicas em programas de melhoramento.

De acordo com os conhecimentos adquiridos com a realização desta revisão, foi possível constatar que em relação ao estudo da estrutura genética de populações, é recomendável o emprego da metodologia desenvolvida por Nei, juntamente com uma das outras metodologias desenvolvidas por Wright, Cockerham e Vencovsky.

A associação entre marcadores moleculares e caracteres quantitativos (QTLs) foi detectada em uma série de pesquisas já realizadas, porém, é preciso refletir sobre os caracteres que apresentam baixas herdabilidades. Para detecção de associação ou ligação entre marcadores moleculares e QTLs, os dados morfoagronômicos são correlacionados com dados moleculares, então os caracteres são avaliados em ensaios de campo. Se apresentam baixas herdabilidades, logicamente, são altamente influenciados pelo ambiente, não sendo coerente a generalização dos resultados, ou

melhor, o uso de marcadores moleculares, em ambientes muito distintos daqueles onde foram realizados os experimentos, pois provavelmente não serão detectados. É preciso atentar, também, para a consistência da associação entre marcadores moleculares e QTLs, pois se o valor de recombinação for muito alto, a associação ou ligação pode ser desfeita ou quebrada após algumas gerações, comprometendo a eficiência da seleção assistida por marcadores. A construção de mapas genéticos da espécie a ser melhorada, nos ambientes de interesse, com alto nível de saturação, deve ser, portanto, a primeira meta de programas de melhoramento que pretendam utilizar marcadores moleculares na seleção de genótipos.

Acredita-se que o emprego de marcadores moleculares é eficiente na avaliação e caracterização molecular de germoplasma, na formação de coleções-núcleo, na identificação e diferenciação de cultivares, na certificação de pureza genética de lotes de sementes comerciais e proteção de cultivares, na verificação da pureza genética de linhagens e híbridos, uma vez que não é fundamental para algumas destas áreas que os marcadores moleculares estejam associados a caracteres agrônômicos de interesse, e para outras, eles são utilizados quando já estão fixados os genótipos, ou seja, as características de interesse.

Entretanto, para estudos de divergência genética, que visam a determinação de grupos heteróticos, a fim de orientar a escolha de parentais para obtenção de híbridos, assim como em programas de retrocruzamento e seleção de genótipos, é preciso lembrar dos caracteres que apresentam baixas herdabilidades, bem como assegurar, com alto nível de precisão, que os marcadores moleculares empregados estão realmente e proximamente associados às características de interesse.

Por outro lado, a seleção assistida para caracteres que apresentam altas herdabilidades, como por exemplo, resistência a doenças, é eficiente e extremamente proveitosa em programas de melhoramento, pela rapidez e confiabilidade dos resultados obtidos.

Em decorrência da grande abrangência do tema desta revisão, os tópicos foram apresentados de uma forma geral, não sendo possível o aprofundamento em cada um deles, pois, como é perceptível, cada um destes tópicos podem, separadamente, originar novas revisões. Porém, é bom frisar que esta revisão contribuiu muito para a iniciação na busca de conhecimentos sobre esta área de pesquisa. É desejado que a mesma seja útil,

também, a estudantes e/ou pesquisadores que tenham interesse em adquirir conhecimentos gerais sobre a área.

## 7. Referências Bibliográficas

AJMONE-MARSAN, P.; LIVINI, C.; MESSMER, M.M.; MELCHINGER, A.E.; MOTTO, M. Cluster analysis of RFLP data from related maize inbred lines of the BSSS and LSC heterotic groups and comparison with pedigree data. **Euphytica**, v.60, p.139-148, 1992.

ALI, M.; COPELAND, L.O.; ELIAS, S.G.; KELLY, J.D. Relationship between genetic distance and heterosis for yield and morphological traits in winter canola (*Brassica napus* L.). **Theoretical and Applied Genetics**, v.91, n.1, p.118-121, 1995.

ALLARD, R.W. 1971. Princípios de melhoramento genético de plantas. São Paulo, Edgard Blucher, 381 p.

ALZATE-MARIN, A.L.; BAÍA, G.S.; MARTINS FILHO, S.; PAULA JUNIOR, T.J. de; SEDIYAMA, C.S.; BARROS, E.G. de; MOREIRA, M.A. Use of RAPD-PCR to identify true hybrid plants from crosses between closely related progenitors. **Brazilian Journal of Genetics**, v.19, n.4, p.621-623, 1996.

AMARAL Jr., A.T. do. **Análise dialélica de betacaroteno, vitamina c, sólidos solúveis e produção e variabilidade em cultivares de tomateiro (*Lycopersicon esculentum* Mill.) via marcadores RAPD**. 1996. 198 p. (Doutorado - Universidade Federal de Viçosa).

BALINT-KURTI, P.J.; DIXON, M.S.; JONES, D.A.; NORCOTT, K.A.; JONES, J.D.G. RFLP linkage analysis of the Cf-4 and Cf-9 genes for resistance to *Cladosporium fulvum* in tomato. **Theoretical and Applied Genetics**, v.88, n.6/7, p.691-700, 1994.

BALLVÉ, R.M.L. **Marcadores isoenzimáticos e biologia da reprodução de *Stevia rebaudiana* (Bert.) Bertoni**. Campinas, 1992. 106p. (Doutorado - Universidade Estadual de Campinas / UNICAMP).

BOURNIVAL, B.L.; SCOTT, J.W.; VALLEJOS, C.E. An isozyme maker for resistance to race 3 of *Fusarium oxysporum* f. sp. *lycopersici* in tomato. **Theoretical and Applied Genetics**, v.78, n.4, p.489-494, 1989.

BURSTIN, J.; de VIENNE, D.; DUBREUIL, P.; DAMERVAL, C. Molecular markers and protein quantities as genetic descriptors in maize. I. Genetic diversity among 21 inbred lines. **Theoretical and Applied Genetics**, v.89, p.943-950, 1994.

CARDOSO, A.I.I. **Uso de marcadores moleculares em hortaliças. Piracicaba, 1996. 69p.** (Monografia - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

CHUNWONGSE, J.; BUNN, T.B.; CROSSMAN, C.; JIANG, J.; TANKSLEY, S.D. Chromosomal localization and molecular marker tagging of the powdery mildew resistance gene (Lv) in tomato. **Theoretical and Applied Genetics**, v.89, n.1, p.76-79, 1994.

COCKERHAM, C.C. Variance of gene frequency. **Evolution**, v.23, p.72-84. 1969.

COLE, R.A.; SUTHERLAND, R.A.; RIGGALL, W..E. The use of polyacrilamide gel electrophoresis to identify variation in isozymes as markers for *Lactuca* species and resistance to the lettuce root aphid *Pemphigus bursarius*. **Euphytica**, v.56, n.3, p.237-242, 1991.

COLETTA FILHO, H.D., TARGON, M.L.P., PONPEU JR., J., MACHADO, M.A. Genetic variability of Ponkan mandarins (*Citrus reticulata* Blanco). In: Congresso Nacional de Genética, 42. Caxambú-Brasil: SBG, **Brazilian Journal of Genetics**, v. 19, n.3, p.212, 1996.

DAX, E.; LIVNEH, O.; EDELBAUM, O.; KEDAR, N.; GAVISH, N.; KARCHI, H.; MILO, J.; SELA, I.; RABINOWITCH, H.D. A random amplified polymorphic DNA (RAPD) molecular marker for the Tm-2<sup>a</sup> gene in tomato. **Euphytica**, v.74, n.1/2, p.159-163, 1994.

DENNY, R.; MUDGE, J.; LANGE, D.; ORF, J.; PENUELA, S.; YOUNG, N. Marker-assisted selection for soybean cyst nematode resistance. **Soybean Genetics Newsletter**, v.23, p.179-182, 1996.

DIRLEWANGER, E.; ISAAC, P.G.; RANDES, S.; BELAJOUZA, M.; COUSIN, R.; de VIENNE, D. Restriction fragment length polymorphism analysis of loci associated with disease resistance genes and developmental traits in *Pisum sativum* L. **Theoretical and Applied Genetics**, v.88, n.1, p.17-27, 1994.

DUARTE, J. B. Princípios e utilização de técnicas multivariadas no melhoramento de plantas, 1996. Monografia. ESALQ/USP. 69 p.

DUDLEY, J.W. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. **Crop Science**, v.33, p.660-668, 1993.

ESHED, Y.; ZAMIR, D. Introgressions from *Lycopersicon pennellii* can improve the soluble solids yield of tomato hybrids. **Theoretical and Applied Genetics**, v.88, n.6/7, p.891-897, 1994.

FARRARA, B.F.; ILOTT, T.W.; MICHELMORE, R.W. Genetic analysis of factors for resistance to downy mildew (*Bremia lactucae*) in species of lettuce (*Lactuca sativa* and *L. serriola*). **Plant Pathology**, v.36, p.499-514, 1987.

FERREIRA, D.F. **Eficiência de métodos de mapeamento de locos quantitativos (QTLs) e de seleção assistida por marcadores moleculares através de simulação**. Piracicaba, 1996. 210p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

FERREIRA, M.E.; GRATTAPLAGIA, D. **Introdução ao uso de marcadores RAPD e RFLP em análise genética**. EMBRAPA - CENARGEN, Brasília, 1995. 220p.

GARVEY, T.C.; HEWITT, J.D. Use of molecular markers to locate quantitative trait linked to high soluble solids content in a hybrid of *Lycopersicon cheesmanii*. **Journal of American Society of Horticultural Science**, v.117, n.3, p.497-499, 1992.

GIBSON, G. Epistasis and pleiotropy as natural properties of transcriptional regulation. **Theoretical Population Biology**, v.49, p.58-89, 1996.

GOLDMAN, I.L.; PARAN, I.; ZAMIR, D. Quantitative trait locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculentum* X *Lycopersicon cheesmanii* cross. **Theoretical and Applied Genetics**, v.90, p.925-932, 1995.

GYLLENSTEN, U.B. PCR and DNA sequencing. **BioTechniques**, v.7, n.7, p.700-705, 1989.



HULBERT, S.H.; MICHELMORE, R.W. Linkage analysis of genes for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). **Theoretical and Applied Genetics**, v.70, n.5, p.520-528, 1985.

ISABEL, N.; BEAVLIEV, J.; BOUSQUET, J. Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. **Proc. Natl. Acad. Sci.**, v.92, p.6369-6373, 1995.

JAIN, A.; BHATIA, S.; BANGA, S.S.; PRAKASH, S.; LAKSHMIKUMARAN, M. Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. **Theoretical and Applied Genetics**, v.88, n.1, p.116-122, 1994.

KENNARD, W.C.; HAVEY, M.J. Quantitative trait analysis of fruit quality in cucumber: QTL detection, confirmation, and comparison with mating-design variation. **Theoretical and Applied Genetics**, v.91, p.53-661, 1995.

KIRZHNER, V.M.; KOROL, A.B.; RONIN, Y.I. The dynamics of linkage disequilibrium under temporal environmental fluctuation: two-locus selection. **Theoretical Population Biology**, v.47, p.257-276, 1995.

LANZA, L.L.B., SOUZA Jr., C.L., OTTOBONI, L.M.M., VIEIRA, M.L.C., SOUZA, A.P. Estimativa das distâncias genéticas para linhagens endogâmicas de milho e predição da performance dos respectivos híbridos simples através de marcadores moleculares do tipo RAPD. In: **Congresso Nacional de Milho e Sorgo, XXI. Resumos**, Londrina: ABMS/IAPAR, p. 109, 1996a.

LANZA, L.L.B., SOUZA Jr., C.L., VIEIRA, M.L.C., OTTOBONI, L.M.M., SOUZA, A.P. Distâncias genéticas entre linhagens endogâmicas de milho (*Zea mays* L.) e predição de híbridos simples através de marcadores moleculares do tipo RAPD. In: Congresso Nacional de Genética, 42. Caxambú-Brasil: SBG, **Brazilian Journal of Genetics**, v. 19, n.3, p.222, 1996b.

LEE, S.J.; SHIN, J.S.; PARK, K.W.; HONG, Y.P. Detection of genetic diversity using RAPD-PCR and sugar analysis in watermelon [*Citrullus lanatus* (Thunb.) Mansf.] germoplasm. **Theoretical and Applied Genetics**, v.92, p.719-725, 1996.

LINDHOUT, P.; VAN HEUSDEN, S.; PET, G.; VAN OOIJEN, J.W.; SANDBRINK, H.; VERKER, R.; VRIELINK, R.; ZABEL, P. Perspectives of molecular marker assisted breeding for earliness in tomato. **Euphytica**, v.79, p.279-286, 1994.

LIVNEH, O.; VARDI, E.; STRAM, Y.; EDELBAUM, O.; SELA, I. The conversion of a RFLP assay into PCR for the determination of purity in a hybrid pepper cultivar. **Euphytica**, v.62, n.2, p.97-102, 1992.

LORENZ, M.; WEIHE, A.; BORNER, T. DNA fragments of organellar origin in random amplified polymorphic DNA (RAPD) patterns of sugar beet (*Beta vulgaris* L.). **Theoretical and Applied Genetics**, v.88, n.6/7, p.775-779, 1994.

MAILER, R.J.; SCARTH, R.; FRISTENSKY, B. Discrimination among cultivars of rapeseed (*Brassica napus* L.) using DNA polymorphism amplified from arbitrary primers. **Theoretical and Applied Genetics**, v.87, p.697-704, 1994.

MANJARREZ-SANDOVAL, P.; CARTER Jr., T.E.; WEBB, D.M.; BURTON, J.W. Can heterosis be predicted by genetic similarity measures in soybean? **Soybean Genetics Newsletter**, v.23, p.192-196, 1996a.

MANJARREZ-SANDOVAL, P.; CARTER Jr., T.E.; WEBB, D.M.; BURTON, J.W. Coefficient of parentage and RFLP markers: Are they useful in predicting genetic variance in soybean populations? **Soybean Genetics Newsletter**, v.23, p.197-202, 1996b.

MARMEY, P.; BEECHING, J.R.; HAMON, S.; CHARRIER, A. Evaluation of cassava (*Manihot esculenta* Crantz) germplasm collections using RAPD markers. **Euphytica**, v.74, p.203-209, 1994.

MARTIN, G.B.; WILLIAMS, G.K.; TANKSLEY, S.D. Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato using random primers and near isogenic lines. **Proceedings National Academy Science USA**, v.88, p.2336-2340, 1991.

MELCHINGER, A.E.; MESSMER, M.M.; LEE, M.; WOODMAN, W.L.; LAMKEY, K.R. Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. **Crop Science**, v.31, p.669-678, 1991.

MICHELSON, I.; ZAMIR, D.; CZOSNECK, H. Accumulations and translocation of tomato yellow leaf curl virus (TYLCV) in a *Lycopersicon esculentum* breeding line containing the *L. chilense* TYLCV tolerance gene Ty-1. **Phytopathology**, v.84, n.9, p.928-933, 1994.

MORAES, M.L.T. **Variabilidade genética por isoenzimas e caracteres quantitativos em duas populações naturais de aroeira**. Piracicaba, 1992. 139p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

MORI, E.S. **Variabilidade genética isoenzimática em uma população de *Eucalyptus grandis* Hill ex. Maiden submetida a diferentes intensidades de seleção**. Piracicaba, 1993. 119p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

MUNIZ, J.A. **Inferência sobre parâmetros relativos à estrutura genética de populações com dados de frequência s gênicas**. Piracicaba, 1994. 224p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

NAVOT, N.; SARFATTI, M.; ZAMIR, D. Linkage relationship of genes affecting bitterness and flesh color in watermelons. **Journal of Heredity**, v.81, n.2, p.162-165, 1990.

NEI, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences**, Washington, v.70, n.12, p.3321-3323, 1973.

NESBITT, K.A.; POTTS, B.M.; VAILLAN COURT, R.E.; WEST, A.K.; REID, J.B. Portioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). **Heredity**, v.74, p.628-637, 1995.

NIENHUIS, J.; SLOCUM, M.K.; DeVOS, D.A.; MUREN, R. Genetic similarity among *Brassica oleracea* L. genotypes as measured by restriction fragment length polymorphism. **Journal of American Society for Horticultural Science**, v.118, n.2, p.298-303, 1993.

PAIVA, J.R. **Variabilidade enzimática em populações naturais de seringueira (*Hevea brasiliensis* (WILLD. EX. ADR. E JUSS.) MUELL. ARG.)**. Piracicaba, 1992. 145p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

PARAN, I.; HOROWITZ, M.; ZAMIR, D.; WOLF, S. Random amplified polymorphic DNA markers are useful for purity determination of tomato hybrids. **HortScience**, v.30, n.2, p.377, 1995.

POLLAK, L.M.; GARDNER, C.O.; PARKHURST, A.M. Relationships between enzyme marker loci and morphological traits in two mass selected maize populations. **Crop Science**, v.24, p.1174-1179, 1984.

PRINCE, J.P.; POCHARD, E.; TANKSLEY, S.D. Construction of a molecular linkage map of a pepper and a comparison of synteny with tomato. **Genome**, v.36, n.3, p.404-417, 1993.

PRITCHARD, J.; FELDMAN, M.W. Statistics for microsatellite variation based on coalescence. **Theoretical Population Biology**, v.50, p.325-344, 1996.

PROVVIDENTI, R. Inheritance of resistance to the Florida strain of Zucchini Yellow Virus in watermelon. **HortScience**, v.26, n.24, p.407-408, 1991.

REIS, M.S. dos. **Distribuição e dinâmica de variabilidade genética em populações naturais de palmitero (*Euterpe edulis* MARTIUS)**. Piracicaba, 1996. 209p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz"/USP).

REUVENI, R.; BOTHMA, G.A. The relationship between peroxidases activity and resistance to *Sphaeroteca fuliginea* in melons. **Phytopathology**, v.114, n.3, p.260-267, 1985.

ROM, M.; BAR, M.; ROM, A.; PILOWSKY, M.; GIDONI, D. Purity control of F<sub>1</sub>-hybrid tomato cultivars by RAPD markers. **Plant Breeding**, v.114, n.2, p.188-190, 1995.

RONFORT, J.; SAUMITOU-LAPRADE, P.; GUGUEN, J.; COUVET, D. Mitochondrial DNA diversity and male sterility in natural populations of *Daucus carota* spp. *carota*. **Theoretical and Applied Genetics**, v.91, n.1, p.150-159, 1995.

SEKHON, M.S.; GUPTA, V.P. Genetic distance and heterosis in indian mustard: developmental isozymes as indicators of genetic relationships. **Theoretical and Applied Genetics**, v.91, p.1148-1152, 1995.

SIMONSEN, V.; HENEEN, W.K. Genetics variation within and among different cultivars and landraces of *Brassica campestris* L. and *B. oleracea* L. based on isozymes. **Theoretical and Applied Genetics**, v.91, p.346-352, 1995.

SKROCH, P.W.; NIENHUIS, J. Qualitative and quantitative characterization of RAPD variation among snap bean (*Phaseolus vulgaris*) genotypes. **Theoretical and Applied Genetics**, v.91, p.1078-1085, 1995.

SMITH, J.S.C.; SMITH, O.S. Restriction fragment length polymorphism can differentiate among U.S. maize hybrids. **Crop Science**, v.31, p.893-899, 1991.

STEINBORN, R.; LINKE, B.; NOTHNAGEL, T.; BORNER, T. Inheritance of chloroplast and mitochondrial DNA in alloplasmic forms of the genus *Daucus*. **Theoretical and Applied Genetics**, v.91, n.4, p.632-638, 1995.

STEVENS, M.R.; LAMB, E.M.; RHOADS, D.D. Mapping the Sw-5 locus for tomato wilt virus resistance in tomatoes using RAPD and RFLP analyses. **Theoretical and Applied Genetics**, v.90, n.3/4, p.451-456, 1995.

TANKSLEY, S.D.; YOUNG, N..D. RFLP mapping in plant breeding: new tools for an old science. **Biotechnology**, v.7, n.2, p.257-264, 1989.

TAUTZ, D.; RENZ, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res*, v. 12, p. 4127-4138, 1984.

TOMITA, R.N., FIGUEIREDO, J.O. de, MACHADO, M.A. Análise de similaridade genética entre variedades de limão (*Citrus limon* Burm.). In: Congresso Nacional de Genética, 42. Caxambú-Brasil: SBG, **Brazilian Journal of Genetics**, v. 19, n.3, p.212, 1996.

TORGGLER, M.G.F.; CONTEL, E.P.B.; TORGGLER, S.P. **Isoenzimas - Variabilidade Genética em Plantas**. Sociedade Brasileira de Genética - Série Monográfica, 1, Ribeirão Preto, 1995.

VANTOAI, T.T.; PENG, J.; MARTIN, S.S. Using AFLP markers to determine the contribution of parental genomes during recurrent selection. **Soybean Genetics Newsletter**, v.23, p.214-216, 1996.

VILARINHOS, A..D.; BARROS, E.G.; PAIVA, E.; SEDIYAMA, C.S.; MOREIRA, M.A. Use of the random amplified polymorphic DNA technique to characterize soybean (*Glycine max* (L.) Merrill) genotypes. **Brazilian Journal of Genetics**, v.17, n.3, p.287-290, 1994.

VELDBOOM, L.R.; LEE, M. Molecular-marker-facilitated studies of morphological traits in maize. II: Determination of QTLs for grain yield and yield components. **Theoretical and Applied Genetics**, v.89, p.451-458, 1994.

VENCOVSKY, R. Análise de variância de freqüências alélicas. In: Congresso Latino Americano de Genética, 10, Rio de Janeiro, 1992. Proceedings, **Brazilian Journal of Genetics**. Suplemento 1, Ribeirão Preto, v.15, n.1, p.53-60, 1992.

WEEDEN, N.F.; ROBINSON, R.W.; IGNART, F. Linkage between an isozyme locus and one of the genes controlling resistance to watermelon mosaic virus 2 in *Cucurbita ecuadorensis*. **Cucurbit Genetics Cooperative Reporter**, v.7, p.86-87, 1984.

WEIR, B.S. Genetic data analysis. Methods for discrete population genetic data. Sinauer Associates Inc. Publishers, Sunderland Massachusetts. 1990. 377p.

WELSH, J.; McCLELLAND, M. Fingerprinting genome using PCR with arbitrary primers. **Nucleic Acids Research**, v.18, p.7213-7218, 1990.

WHITE, T.J.; ARNHEIM, N.; ERLICH, H.A. The polymerase chain reaction. **Trends in Genetics**, v.5, n.6, p.185-189, 1989.

WILLAMS, J.G.K.; KUBELIK, A.R.; LIVAK, K.J.; RAFALSKI, J.A.; TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, v.18, n.22, p.6531-6535, 1990.

WILLIAMSON, V.M.; HO, J.Y.; WU, F.F.; MILLER, N.; KALOSHIAN, I. A PCR-based marker tightly linked to the nematode resistance gene, Mi, in tomato. **Theoretical and Applied Genetics**, v.87, n.7, p.757-763, 1994.

WITSENBOER, H.; KESSELI, R.V.; FORTIN, M.G.; STANGHELLINI, M.; MICHELMORE, R.W. Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. **Theoretical and Applied Genetics**, v.91, n.1, p.178-188, 1995.

WRIGHT, S. The interpretation of population structure by F-statistics with special regard to system of mating. **Evolution**, v.19, p.395-420, 1965.

YOUNG, N.D.; ZAMIR, D.; GANAL, M.W.; TANKSLEY, S.D. Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the Tm-2<sup>a</sup> gene in tomato. **Genetics**, v.120, n.2, p.579-585, 1988.

ZAMIR, D.; MICHELSON, I.E.; ZAKAY, Y.; NAVOT, N.; ZEIDAN, M.; SARFATTI, M.; ESHED, Y.; HAREL, E.; PLEBAN, T.; Van OSS, H.; KEDAR, N.; RABINOWITCH, H.D.; CZOSNEK, H. Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, TY-1. **Theoretical and Applied Genetics**, v.88, n.2, p.141-146, 1994.

64 Utilização das técnicas de marcadores moleculares na genética de populações, na genética quantitativa e no melhoramento de plantas

**Embrapa**

---

*Roraima*

MINISTÉRIO DA AGRICULTURA,  
PECUÁRIA E ABASTECIMENTO

