

FRANCIELLE ALLINE MARTINS

**GENOTIPAGEM DE ENDOSPERMA COMO ESTRATÉGIA
AUXILIAR NO MAPEAMENTO E DETECÇÃO DE QTLs EM
POPULAÇÕES EXOGÂMICAS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2006

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

M386g
2006
2006.

Martins, Francielle Aline, 1982-

Genotipagem de endosperma como estratégia auxiliar
no mapeamento e detecção de QTLs em populações
exogâmicas / Francielle Aline Martins. – Viçosa : UFV,

xi, 73f. : il. ; 29cm.

Orientador: Pedro Crescêncio Souza Carneiro.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 68-73.

1. Genômica. 2. Locos de caracteres quantitativos.
3. Microssatélites (Genética). 4. Milho – Genética.
I. Universidade Federal de Viçosa. II. Título.

CDD 22.ed. 572.8633

FRANCIELLE ALLINE MARTINS

**GENOTIPAGEM DE ENDOSPERMA COMO ESTRATÉGIA
AUXILIAR NO MAPEAMENTO E DETECÇÃO DE QTLs EM
POPULAÇÕES EXOGÂMICAS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 29 de setembro de 2006.

Dr^a. Cláudia Teixeira Guimarães
(Co-orientadora)

Prof. Cosme Damião Cruz
(Co-orientador)

Dr^a. Eveline Teixeira Caixeta

Prof. Everaldo Gonçalves de Barros

Prof. Pedro Crescêncio Souza Carneiro
(Orientador)

Aos meus pais Gilsimar Francisco Martins e Warly Orlando Martins.
Ao meu companheiro Pedro Marcos de Almeida.
Ao meu filho Lucas.

Dedico

AGRADECIMENTOS

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realização do curso.

Ao Centro Nacional de Pesquisa Milho e Sorgo da EMBRAPA, pela oportunidade de realização deste trabalho.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro.

Ao Professor Pedro Crescêncio Souza Carneiro, pela orientação, pelos ensinamentos transmitidos, pelo incentivo, pela amizade, pela confiança e pela paciência.

Aos meus conselheiros Dr^a Cláudia Teixeira Guimarães, Professor Cosme Damião Cruz e Professor José Eustáquio de Souza Carneiro, pelo bom convívio e pelas contribuições para o desenvolvimento desta pesquisa.

Ao Professor Everaldo Gonçalves de Barros e à Dr^a Eveline Teixeira Caixeta, pelas sugestões para o desenvolvimento deste estudo.

Ao Dr. Jurandir Vieira de Magalhães, pelas sugestões e pela atenção.

Ao Dr. Sidney Netto Parentoni, pelo fornecimento das sementes para elaboração deste trabalho.

A todos os meus professores, pela contribuição ao meu crescimento profissional.

Aos funcionários do Núcleo de Biologia Aplicada, pelo convívio e pela atenção, em especial ao Ubiraci Gomes de Paula Lana.

Aos meus amigos que conquistei durante a minha estada em Sete Lagoas Lidiane, Éder, Maurício, Jane, Leo, Ana Paula, Renata e Marconi, pela convivência.

Aos meus eternos amigos Leonardo, Janaina e Maria Andréia, pelos momentos de agradável convívio e descontração.

Aos meus familiares, pelo carinho e incentivo.

Ao Pedro Marcos, pelo amor, pela dedicação, pela compreensão e pelo apoio; e à sua família, pelo carinho e incentivo.

Ao meu filhote Lucas, pelo sorriso sincero e pela compreensão da minha ausência enquanto estive em Sete Lagoas.

Àqueles que são a base desta conquista: meus pais Gilsimar e Warly, pelo amor incondicional, pelo incentivo, pela dedicação, pelo exemplo e pelos sábios conselhos e, ainda, por não medirem esforços para a minha formação.

A Deus, pelo dom da vida, por conceder-me força e possibilitar-me alcançar mais este objetivo.

BIOGRAFIA

FRANCIELLE ALLINE MARTINS, filha de Gilsimar Francisco Martins e Warly Orlando Martins, nasceu em Juiz de Fora, Minas Gerais, no dia 30 de abril de 1982.

Em 2000, iniciou o Curso de Ciências Biológicas na Universidade Federal de Viçosa (UFV), em Viçosa, MG, concluindo a Licenciatura em julho de 2004 e o Bacharelado em janeiro de 2005.

Durante o período de graduação foi bolsista do PIBIC/FAPEMIG no Departamento de Biologia Geral da UFV, onde desenvolveu atividades de pesquisa no Laboratório de Citogenética e Citometria Vegetal.

Em março de 2005, ingressou no Programa de Pós-Graduação, em nível de Mestrado, em Genética e Melhoramento da UFV, submetendo-se à defesa de tese em setembro de 2006.

SUMÁRIO

RESUMO	ix
ABSTRACT	xi
1. INTRODUÇÃO GERAL.....	1
2. REVISÃO DE LITERATURA.....	4
2.1 Mapas Genéticos.....	4
2.2 Marcadores Morfológicos, Bioquímicos e Moleculares.....	5
2.3 Mapeamento de QTLs.....	8
2.4 Populações de mapeamento.....	10
2.4.1 Populações endogâmicas.....	11
2.4.2 Populações exogâmicas.....	12
2.5 Detecção de QTLs em populações exogâmicas.....	12
2.5.1 Cruzamentos marcador-informativos.....	13
2.5.2 Frequência de genótipos não informativos.....	14
2.5.3 O endosperma como estratégia na detecção de QTL.....	18
CAPÍTULO 1.....	21
Germinação de sementes de milho submetidas à remoção parcial do endosperma e do pericarpo.....	21
1. INTRODUÇÃO.....	22
2. MATERIAL E MÉTODOS.....	24
3. RESULTADOS E DISCUSSÃO.....	26
CAPÍTULO 2.....	28
Adaptação da metodologia de PCR semiquantitativo e marcadores SSRs para diferenciação de amostras quanto ao número de cópias de um alelo.....	28
1. INTRODUÇÃO.....	29

2. MATERIAL E MÉTODOS	31
3. RESULTADOS	34
4. DISCUSSÃO.....	45
CAPÍTULO 3.....	47
Genotipagem do endosperma de milho utilizando marcadores SSR fluorescentes e PCR semiquantitativo	47
1. INTRODUÇÃO.....	48
2. MATERIAL E MÉTODOS	50
3. RESULTADOS	54
4. DISCUSSÃO.....	63
3. CONCLUSÕES GERAIS	66
4. REFERÊNCIAS.....	68

RESUMO

MARTINS, Francielle Aline, M.Sc., Universidade Federal de Viçosa, setembro de 2006. **Genotipagem de endosperma como estratégia auxiliar no mapeamento e detecção de QTLs em populações exogâmicas.** Orientador: Pedro Crescêncio Souza Carneiro. Co-Orientadores: Claudia Teixeira Guimarães, Cosme Damião Cruz e José Eustáquio de Souza Carneiro.

No mapeamento genético e detecção de QTLs em populações exogâmicas nem sempre é possível a determinação da fase de ligação dos alelos. Assim, indivíduos heterozigotos são descartados dessas análises por serem não informativos, uma vez que não é possível, por meio do seu genótipo, distinguir a origem de seus alelos em relação aos dois genitores. Dessa forma, o objetivo principal deste trabalho foi propor a genotipagem do endosperma para identificar a origem alélica dos indivíduos heterozigotos. Inicialmente, fragmentos do endosperma representando 10, 25 e 50% do peso das sementes de milho foram retirados, sendo as sementes submetidas ao teste de germinação. Observou-se que a remoção de até 50% do endosperma não afetou a taxa de germinação das sementes. A metodologia de PCR semiquantitativo foi otimizada para diferenciar doses dos alelos nas misturas de DNA foliar de duas linhagens de milho (L3 e L1113-01), representando as diferentes proporções alélicas observadas no tecido endospermático dos seus cruzamentos recíprocos, tendo como base a quantidade máxima de endosperma que podia ser utilizada na extração do DNA. Marcadores SSR foram gerados pela técnica de PCR semiquantitativo, e os fragmentos amplificados foram avaliados tanto em gel de agarose tratado com brometo de etídio quanto em gel de poliacrilamida, usando-se *primers* fluorescentes. A

resolução do gel de agarose não possibilitou a diferenciação das misturas dos DNAs parentais. No entanto, por meio da análise de regressão e da comparação da intensidade da banda correspondente a um mesmo alelo nas diferentes misturas, pôde-se inferir a concentração inicial de cada um dos alelos. A necessidade de um padrão de alelos limitou o uso dessa técnica nas análises de QTLs em populações nas quais pelo menos um dos genitores é conhecido. Já a resolução do gel de poliacrilamida utilizando marcadores fluorescentes foi mais eficiente na genotipagem de endospermas, uma vez que possibilitou a diferenciação da origem materna das sementes dos híbridos recíprocos. Assim, a estratégia de genotipagem do endosperma utilizando *primers* SSR fluorescentes amplificados pela técnica de PCR semiquantitativo possibilitou a determinação da origem dos alelos dos descendentes heterozigotos derivados de populações exogâmicas, permitindo a inclusão destes na detecção de QTLs e, conseqüentemente, aumentando a precisão das análises.

ABSTRACT

MARTINS, Francielle Aline, M.Sc., Universidade Federal de Viçosa, September de 2006. **Endosperm genotyping as assistant strategy in the QTLs' detection and mapping in outbred populations.** Adviser: Pedro Crescêncio Souza Carneiro. Co-Advisers: Cláudia Teixeira Guimarães, Cosme Damião Cruz and José Eustáquio de Souza Carneiro.

In the genetic mapping in outbred populations not always it is possible to determine the linkage phase of the alleles. Thus, heterozygous individuals are discarded from these analyses due to the lack of information, once it is not possible, through their genotype, to distinguish the origin of their parental alleles. In this way, the main objective of this work was to propose the endosperm genotyping as a strategy to identify the allelic origin of those heterozygotes individuals. Initially, fragments from the endosperm representing 10, 25 and 50% of the corn seeds weight were extracted and the seeds were submitted to the germination test. The results suggest that the elimination of up to 50% of the endosperm did not affected the seed germination. The methodology of semiquantitative PCR was optimized to differentiate doses of the alleles in the mixtures of DNA derived from leaves of two maize inbred lines (L3 and L1113- 01). It was represented different allelic proportions observed in the endosperm of their reciprocal crosses, based on the maximum amount of endosperm that could be used for DNA extraction. SSR markers were generated by semiquantitative PCR technique and the amplified fragments were evaluated in both agarose gels treated with ethidium bromide and polyacrylamide gels using fluorescently labeled primers. Gel resolution using agarose did not allow the differentiation of the mixtures of parental DNAs. However, through the regression analysis and comparison of the band intensity

corresponding to the same allele in the different mixtures, the initial concentration of each one of the alleles could be inferred. The requirement of an allelic pattern limited the use of this technique to QTL analysis in populations where at least one of the genitors is known. Although the resolution of polyacrylamide gels using fluorescent markers was more efficient in the endosperm genotyping, once it was allowed to differentiate the maternal origin of reciprocal hybrids seed's. So, the strategy of endosperm genotyping using fluorescent SSR primer amplified by semiquantitative PCR allowed the determination of allelic origin in the heterozygous offspring derived from outbred populations, including these individuals in the QTL detection, and consequently, increasing the precision of this analysis.

1. INTRODUÇÃO GERAL

Muitas características de importância econômica na agropecuária, como produção de grãos, volume de leite, altura, peso e teor de proteína, dentre outras, apresentam padrão contínuo de variação, ou seja, em dada população não há clara distinção entre as classes fenotípicas de indivíduos. A fundamentação genética dessas características foi elucidada logo após a redescoberta das leis de Mendel, quando foi proposto que essa variação contínua poderia ser o resultado da ação cumulativa de um número grande de genes. Isso deu origem à chamada Genética Quantitativa, que estuda as características governadas por uma grande quantidade de genes, em geral afetados por fatores ambientais (BEARZOTI, 2000).

Geldermann (1975), assim, propôs o termo *locos controladores de características quantitativas* (QTL) para designar um grupo de genes responsáveis pela expressão de determinada característica. O desenvolvimento dos marcadores moleculares, no entanto, principalmente a partir do final da década de 1980, estimularam o desenvolvimento de novas ferramentas de análise que procuraram, de modo preciso, detectar, localizar e caracterizar os locos que participam do controle da expressão desses caracteres (COELHO, 2000).

De modo análogo às metodologias de mapeamento genético, o princípio que fundamenta as análises de mapeamento de QTLs é a existência de desequilíbrio gamético de ligação decorrente da redução da frequência de recombinação entre genes situados em regiões próximas entre si ao longo de determinado cromossomo. Nesse sentido, diversos delineamentos genéticos podem ser utilizados para que esse desequilíbrio possa ser obtido com maior

poder de resolução possível, de modo a permitir o mapeamento genético dos locos de interesse (COELHO, 2000). Assim, a seleção da população de mapeamento, a escolha de genitores e a determinação do tipo de cruzamento são consideradas etapas críticas para o sucesso da construção do mapa (STAUB et al., 1996).

Em estudos envolvendo populações exogâmicas, a determinação da fase de ligação nem sempre é possível, pois muitas vezes um dos genitores é desconhecido (populações de meios-irmãos) ou até mesmo ambos (populações naturais). Além disso, a descendência pode ser resultado do cruzamento de ambos os genitores que apresentam o mesmo marcador heterozigoto, e, nesse caso, somente a progênie homozigota é informativa (LIU, 1998). Em todos esses casos, a descendência heterozigota M_1M_2 não possibilita, por meio de seu genótipo, distinguir a origem de seus alelos em relação aos dois genitores do cruzamento $M_1M_2 \times M_1M_2$. A falta de informatividade dos indivíduos heterozigotos faz que as avaliações destes sejam descartadas das análises, não contribuindo para a detecção do QTL. Assim, estratégias de identificação da origem dos alelos, se paternos ou maternos, são de grande interesse para uma detecção mais acurada dos QTLs.

Uma possível forma de incluir os indivíduos heterozigotos nas análises é por meio da genotipagem do seu endosperma. Espera-se que um indivíduo heterozigoto M_1M_2 apresente um endosperma $M_1M_1M_2$ ou $M_1M_2M_2$, quando os genitores femininos são doadores do alelo M_1 ou M_2 , respectivamente. A distinção quanto ao número de doses diferenciadas de cada alelo tem sido realizada com sucesso por meio da metodologia de PCR em tempo real. Apesar de sensível e confiável, esta metodologia é excessivamente cara, o que limita o seu uso rotineiro em vários laboratórios (MARIN et al., 2002).

Nesse sentido, a metodologia de PCR semiquantitativo (PCR-SQ) vem sendo adaptada ao longo dos últimos anos. Embora não permita uma quantificação absoluta, espera-se ser possível a inferência do número de moléculas amplificadas de cada genoma, se o produto da PCR-SQ corresponder à fase exponencial, em que o número de cópias é função do número inicial (GÉNÉTHON, 2001).

Devido ao potencial uso da PCR-SQ para inferir o número de moléculas amplificadas, o objetivo deste estudo foi adaptar uma metodologia de extração e genotipagem do DNA de endosperma de populações exogâmicas de milho para a determinação do número de cópias dos alelos nesse tecido e, dessa forma conhecer a origem deles para que indivíduos heterozigotos possam compor as análises utilizadas na detecção de QTLs.

2. REVISÃO DE LITERATURA

2.1. Mapas genéticos

Com a redescoberta dos trabalhos de Mendel, no final do século XIX, por Hugo de Vries, Carl Correns e Erick von Tschermak-Seysenegg, foram realizadas inúmeras pesquisas, a fim de esclarecer os mecanismos básicos da herança. Essas pesquisas culminaram na formulação da Teoria Cromossômica da Herança por W. Sutton e T. Boveri, em 1902, que estabeleceram a relação entre cromossomos e fatores mendelianos (genes), ou seja, os cromossomos são elementos celulares portadores dos genes (CARNEIRO; VIEIRA, 2002).

No mesmo ano, W. Bateson, E.R. Saunders e R.C. Punnett demonstraram, em ervilha-doce, que as segregações dos caracteres cor da flor e formato do pólen não ocorriam de forma independente (LANDER; WEINBERG, 2000). T.H. Morgan, C.B. Bridges, H.J. Muller e A.H. Sturtevant, em 1910, trabalhando com *Drosophila melanogaster*, também observaram distorções da Segunda Lei de Mendel. Morgan supôs, então, que alguns genes estariam situados no mesmo cromossomo e que durante a meiose, ocasionalmente, ocorreriam, entre os homólogos, trocas de segmentos denominadas *crossing-over*, ou permuta. A ocorrência de quiasma, que é a visualização citológica da permuta, dava suporte aos resultados genéticos encontrados por Morgan, que também percebeu que a proporção da progênie recombinante variava bastante, relatando que estas variações, de algum modo, refletiam a distância entre os genes no cromossomo (CARNEIRO; VIEIRA, 2002).

Em 1913, A.H. Sturtevant, interpretando dados oriundos da segregação de genes ligados, sugeriu o uso da porcentagem de recombinantes como indicador quantitativo da distância linear entre dois genes na construção de mapas genéticos, que evidenciaram que a posição dos genes correspondia à sua ordem linear nos cromossomos. Assim, o conceito de localização dos genes em uma ordem linear passou a ser incorporado à teoria cromossômica da herança (GARDNER; SNUSTAD, 1986; GRIFFITHS et al., 1998).

O mapeamento genético tornou-se, rapidamente, uma poderosa ferramenta para os geneticistas. Os primeiros mapas genéticos eram fundamentados em marcadores morfológicos e citológicos. Estes se originaram de *estoques cromossômicos* com aberrações (como aneuploidias, translocações, deleções e inversões), principalmente nas culturas do milho, tomate e ervilha (Coe et al., 1988; Rick; Yoder, 1988, citados por CARNEIRO; VIEIRA, 2002).

A partir da década de 1980, com o advento dos marcadores de DNA, o mapeamento genético tornou-se, efetivamente, ilimitado a todas as espécies (BOSTEIN et al., 1980). Mapas genéticos que levaram décadas para serem desenvolvidos, em virtude da restrita disponibilidade de marcadores morfológicos e isoenzimáticos, foram amplamente saturados com o uso de marcadores de polimorfismo em nível de DNA. Várias espécies sobre as quais os estudos de herança eram restritos tiveram mapas genéticos rapidamente desenvolvidos (GRATTAPAGLIA; SEDEROFF, 1994).

2.2. Marcadores morfológicos, bioquímicos e moleculares

Caracteres de herança simples, apresentando variação discreta, cujas classes são facilmente distinguíveis, são bons marcadores morfológicos ou fenotípicos. Características como nanismo, deficiência de clorofila, cor das pétalas e morfologia foliar são usadas como marcadores fenotípicos, sendo de avaliação simples e de baixo custo. Entretanto, a disponibilidade de marcadores morfológicos é restrita a poucas espécies vegetais para as quais o número de informações genéticas é maior, por exemplo o milho, o tomate e a ervilha (FERREIRA; GRATTAPAGLIA, 1996; LIU, 1998). Outra desvantagem do uso desses marcadores é a necessidade de grande número de descritores

que serão identificados, em sua maioria, na planta inteira ou adulta; além do tempo gasto e da necessidade de espaço físico para a avaliação do material, os marcadores morfológicos podem ser influenciados pelo ambiente (PADILHA, 2002). Pode-se destacar, ainda, a dificuldade de obtenção de populações com grande número de marcadores morfológicos segregando e a desvantagem de freqüentemente possuírem herança dominante (BEARZOTI, 2000).

Após 1949, uma nova classe de marcadores foi descoberta a partir de polimorfismos bioquímicos, as isoenzimas. Essas são formas moleculares de uma mesma enzima que, embora desempenhando as mesmas funções catalíticas, podem apresentar propriedades cinéticas distintas, diferenciadas via eletroforese (FERREIRA; GRATTAPAGLIA, 1996). As isoenzimas passaram a ser usadas como marcadores bioquímicos, permitindo a construção de mapas genéticos, potencialmente, em todas as espécies de plantas (GARVIN; WEEDEN, 1994). Os marcadores isoenzimáticos apresentam um controle genético normalmente simples, em que a ação alélica é, na maioria dos casos, codominante (CARNEIRO; VIEIRA, 2002). Dessa forma, pode-se dizer que os marcadores bioquímicos apresentam um nível de polimorfismo, embora limitado, superior ao dos morfológicos, sendo menos influenciados pelo ambiente (PADILHA, 2002).

A primeira classe de marcadores de polimorfismo em nível de DNA desenvolvida foi o Restriction Fragment Length Polymorphism (RFLP). Essa técnica se baseia no comprimento dos fragmentos polimórficos obtidos a partir da reação do DNA com enzimas de restrição específicas para determinadas seqüências de nucleotídeos. O nível de polimorfismo obtido com o RFLP, embora maior que o de isoenzimas e consideravelmente alto em plantas alógamas como o milho, é relativamente baixo em algumas espécies autógamas (BEARZOTI, 2000).

Nos anos de 1990, marcadores com nível de polimorfismo ainda maior foram obtidos com o uso da chamada reação em cadeia da enzima polimerase (PCR). Essa técnica deu origem a diferentes classes de marcadores, com destaque para o RAPD, AFLP e SSR (microsatélites).

O Simple Sequence Repeat (SSR), ou microsatélite, é um marcador multialélico e codominante, apresentando elevado conteúdo de informação genética por loco. Tal técnica é caracterizada pela simplicidade, rapidez,

confiabilidade, precisão e reprodutibilidade em prover um perfil genético detalhado, além de apresentar a possibilidade de automação (PADILHA, 2002).

Os genomas eucariotos são densamente povoados por diferentes classes de seqüências repetidas. Os microssatélites ou SSR são seqüências curtas com um a seis nucleotídeos repetidos em *tandem* ao longo da molécula de DNA, as quais são flanqueadas por seqüências conservadas. Após sua descoberta, os microssatélites tornaram-se extensivamente usados em análises genéticas, pois são freqüentes, encontram-se distribuídos ao acaso no genoma e originam locos genéticos altamente polimórficos (ROIZÈS, 2000).

Cada microssatélite constitui um loco genético altamente variável, multialélico e com um elevado conteúdo informativo de polimorfismo (FERREIRA; GRATTAPAGLIA, 1996). O polimorfismo de um loco microssatélite pode ser constatado por meio de PCR, utilizando-se um par de *primers* específicos e complementares às seqüências únicas que flanqueiam a região repetida. Desse modo, cada segmento amplificado com diferente tamanho representa um alelo diferente do mesmo loco (QUEIROZ, 2004). Os SSRs apresentam elevado poder discriminatório e, normalmente, poucos locos garantem a completa diferenciação dos genótipos de interesse (PADILHA, 2002).

A identificação dos SSRs é um processo trabalhoso e de elevado custo, envolvendo a construção de bibliotecas genômicas, seleção de clones contendo os microssatélites, seqüenciamento desses clones, desenho dos *primers* específicos e sua confirmação por PCR. Para o milho, seqüências de 1855 pares de *primers* SSR encontram-se publicamente disponíveis no Maize Genetics and Genomics Database (<http://www.maizegdb.org/ssr.php>), estando bem caracterizados e mapeados ao longo de todo o genoma.

2.3. Mapeamento de QTLs

Há grande quantidade de mapas genéticos disponíveis, tanto para espécies animais quanto vegetais (CARNEIRO; VIEIRA, 2002). O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares. Com base nas informações fornecidas pelos mapas genéticos, procedimentos como a localização e o mapeamento de Quantitative Trait Loci (QTLs), estudos de sintenia e clonagem de genes com base em mapas genéticos passaram a representar complementos importantes a serem considerados em programas de melhoramento de diferentes espécies vegetais (LEE, 1995).

Uma das aplicações mais importantes dos mapas genéticos é a localização de genes que controlam características de importância econômica, como produção de grãos, altura da planta e teor de proteína que resultam da ação cumulativa de um conjunto de genes. Essas características são denominadas poligênicas, quantitativas ou de herança complexa, e os locos que as controlam são chamados de QTLs (TANKSLEY, 1993).

O mapeamento de QTLs possibilita mensurar o número de locos quantitativos envolvidos na herança complexa, bem como suas localizações cromossômicas e modo de ação gênica (aditividade, dominância, heterose e epistasia) (FERREIRA; GRATTAPAGLIA, 1996; AUSTIN; LEE, 1998). A capacidade de detecção de um QTL é 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, bem como da herdabilidade da característica (PATERSON et al., 1991; TANKSLEY, 1993).

Uma vez detectados os marcadores ligados a determinada característica de interesse, é possível selecionar os indivíduos com base no fenótipo do marcador, sem que haja a necessidade de avaliar o fenótipo da característica. Essa abordagem é conhecida como seleção assistida por marcadores (SAM), sendo uma estratégia importante, principalmente, nos casos em que a característica de interesse é de avaliação difícil e cara, nos estudos de espécies perenes de ciclo longo e para características fenotípicas de herdabilidade baixa, o que pode significar redução no tempo e nos custos de programas de melhoramento (SOUZA, 2001).

Dentre os métodos utilizados para a localização de QTLs, encontram-se o método Bulked Segregant Analysis (BSA), a detecção por marcas simples, o mapeamento por intervalo simples, o mapeamento por intervalo composto e o mapeamento por intervalo múltiplo ou multiponto.

A metodologia conhecida como BSA, ou análise de *bulks* segregantes, normalmente utilizada para detecção de genes que controlam características de herança mendeliana simples, é considerada processo mais direto de detecção de um marcador associado a um QTL. Trata-se de procedimento baseado em seleção prévia de marcadores, no sentido de otimizar o mapeamento de regiões flanqueando locos de caracteres quantitativos. O procedimento consiste em estabelecer dois grupos de indivíduos obtidos dos pontos extremos de uma população segregante para determinado caráter quantitativo e, então, avaliar a ocorrência de polimorfismo em locos de marcadores moleculares entre os grupos (SCHUSTER; CRUZ, 2004).

Segundo Tanksley (1993), os procedimentos mais simples para detecção de QTLs são os modelos lineares ou mapeamento por marcas simples que analisam, separadamente, a diferença entre os valores fenotípicos médios para cada marcador (EDWARDS et al., 1987), não sendo necessária a construção de mapas de ligação. O procedimento consiste na comparação estatística do valor médio do caráter para as distintas classes genotípicas do marcador. Um resultado significativo denotaria a existência de ligação entre o marcador e um QTL do caráter em questão, embora os efeitos do QTL e a sua distância do marcador não possam ser estimados (COELHO, 2000). Como limitação, há o fato de quanto maior for a distância entre o QTL e o marcador, menor será a probabilidade de detectá-lo em vista da possibilidade de recombinação entre o gene marcador e o QTL. Além disso, a magnitude do QTL, pelo mesmo motivo, é normalmente subestimada (LIU, 1998).

O método de mapeamento com base em intervalo simples foi proposto por Lander e Botstein (1989) e tem sido alternativa para aumentar o poder de detecção das associações, permitindo estimar o efeito e a posição dos QTLs. Esse método tem como base informações de segregação de pares de marcadores adjacentes como unidades de análise, utilizando o método de máxima verossimilhança para estimar a frequência de recombinantes e a magnitude do efeito do QTL no intervalo entre dois marcadores ligados

(LYNCH; WALSH, 1998). Uma desvantagem atribuída ao mapeamento fundamentado em intervalo é a de que os outros QTLs fora do intervalo são ignorados, o que pode resultar em duas conseqüências principais. Toda a variação genética devida a esses outros QTLs é residual, o que pode diminuir a precisão das estimativas e o poder dos testes. Além disso, eventuais QTLs que estejam ligados ao intervalo interferem no processo de estimação, levando, em casos extremos, a declarar a existência de um QTL no intervalo quando na realidade não há nenhum (BEARZOTI, 2000).

A fim de resolver esses problemas, Jansen (1993) e Zeng (1993; 1994), propuseram o mapeamento por intervalo composto, método em que se consideram os QTLs fora do intervalo em questão, por um modelo de regressão linear múltipla, proporcionando maior resolução no mapeamento (COELHO, 2000).

Outra abordagem mais recente é o mapeamento por intervalo múltiplo, proposto por Kao et al. (1999), que trouxe considerável aprimoramento no mapeamento de QTLs, pois incorpora parâmetros de epistasia ao modelo. Em presença de epistasia, esse método apresenta algumas vantagens potenciais como maior eficiência e precisão na identificação de QTLs; seus efeitos são estimados sem viés e, ao contrário de outros modelos, pode aumentar a eficiência da seleção assistida, considerando os efeitos epistáticos (BEARZOTI, 2000).

2.4. Populações de mapeamento

A seleção da população para mapeamento envolve a escolha de genitores e a determinação do tipo de cruzamento, sendo considerada uma etapa crítica para o sucesso da construção do mapa (STAUB et al., 1996). Para isso, duas condições básicas devem ser atendidas: o máximo de polimorfismo entre os genitores e que sejam produzidas gerações em que os locos estejam em desequilíbrio de ligação (PATERSON et al., 1991; TANKSLEY, 1993).

Os genitores escolhidos para gerar uma população de mapeamento devem ser contrastantes em relação aos genes que determinam a característica de interesse e para os possíveis locos marcadores. Em casos

cujo polimorfismo é baixo, podem-se realizar cruzamentos interespecíficos para obter maior número de locos informativos (TANKSLEY, 1993).

Gerações F_2 ou oriundas de retrocruzamentos, cruzamentos entre heterozigotos e linhagens duplo-haplóides são as populações mais usadas na construção de mapas de ligação (TANKSLEY, 1993; LYNCH; WALSH, 1998). Conforme salientou COELHO (2000), determinados genótipos (indivíduos F_2) são portadores de dois gametas informativos, enquanto em outras populações (retrocruzamento) cada indivíduo carrega somente um gameta informativo, sendo o outro complemento não informativo. Assim, de modo aproximado, um indivíduo F_2 fornece duas vezes mais informações que um indivíduo oriundo de uma geração de retrocruzamento. Essa relação é aproximada, porque nem todos os genótipos de F_2 são informativos (LIU, 1998).

A escolha do tipo de população mais apropriada ao mapeamento está também relacionada com o tipo de marcador. O máximo de informação genética é alcançado quando se usam uma população F_2 e marcadores codominantes (STAUB et al., 1996). Lynch e Walsh (1998) distinguiram duas populações de cruzamento: as linhagens de cruzamento endogâmico e as populações exogâmicas, sendo que esta última pode ser separada em populações de cruzamento controlado e populações naturais, ou seja, sem controle de cruzamento.

2.4.1. Populações endogâmicas

A forma mais simples e eficiente de detectar um QTL é por meio de cruzamento de populações endogâmicas, pois, cruzando duas linhagens ou populações endogâmicas, a progênie apresentará uma diferença fixa entre todos os marcadores e o loco para a característica. Dessa forma, todos os locos ligados na progênie (geração F_1) estarão em desequilíbrio de ligação, e esta pode ser utilizada para gerar as populações de mapeamento. Dentre essas, as mais comuns são as populações F_2 , que são geradas a partir do intercruzamento dos indivíduos F_1 , e as populações de retrocruzamento, que são obtidas com os cruzamentos entre os indivíduos F_1 e a população parental. Nas populações F_2 , todos os três genótipos possíveis estão representados,

enquanto nas populações de retrocruzamento apenas o genótipo heterozigoto e um dos homozigotos estão representados na progênie (LIU, 1998).

2.4.2. Populações exogâmicas

Algumas vezes é difícil fazer a clara distinção entre populações naturais e artificiais (experimentais). Populações obtidas a partir de cruzamentos controlados entre pais selecionados podem ser consideradas como populações experimentais. As populações produzidas por cruzamentos de ocorrência natural (sem controle artificial) são consideradas populações naturais.

Para algumas espécies, populações naturais podem ser geradas por acasalamentos preferenciais, até mesmo por completa autofecundação, ou por acasalamento ao acaso. Exemplos de populações naturais podem ser as famílias de meios-irmãos, mistura ao acaso e autopolinização. Populações geradas de uma única planta polinizada por muitas fontes de pólen desconhecidas ou parcialmente conhecidas são típicas famílias de meios-irmãos. Em termos genéticos, as fontes de pólen são subconjuntos ao acaso de populações. Outro tipo de população natural, comum no reino vegetal, é gerado pela combinação de polinização ao acaso e autopolinização como sementes de uma árvore (LIU, 1998).

2.5. Detecção de QTLs em populações exogâmicas

Em linhagens endogâmicas, vários delineamentos têm sido propostos para a detecção do desequilíbrio necessário para o mapeamento de QTL. A maior diferença entre cruzamentos endogâmicos e exogâmicos está no fato de os genitores, deste último, ser genótipos geralmente desconhecidos (SLATE, 2005).

Segundo Lynch e Walsh (1998), esta situação traz sérias conseqüências, pois somente uma fração dos descendentes é informativa em uma população exogâmica. Um genótipo informativo contém pelo menos um gameta informativo, o qual permite a identificação da origem dos seus alelos, mesmo que não seja possível determinar se esse gameta é recombinante ou não. A frequência de genótipos não informativos é característica de cada tipo

de população e está relacionada diretamente com a frequência alélica e com o tipo de cruzamento (DA; LEWIN, 1995).

2.5.1. Cruzamentos marcador-informativos

Três tipos de cruzamentos marcador-informativos foram descritos por Lynch e Walsh (1998) para o mapeamento de populações exogâmicas:

Cruzamentos completamente informativos ($M_1M_2 \times M_3 M_4$)

Os pais são diferentes heterozigotos. Toda a progênie é informativa, uma vez que é possível, por meio de seu genótipo, distinguir a origem de seus alelos em relação a ambos os genitores (DA; LEWIN, 1995).

Para a detecção do QTL, devem ser analisados os indivíduos que constituem a família. Nessa situação são identificadas quatro classes genotípicas, em relação ao loco M, das quais duas (M_1M_3 e M_1M_4) são representadas por indivíduos portadores do alelo M_1 do genitor P_1 e as outras duas, M_2M_3 e M_2M_4 , são representadas por indivíduos portadores do alelo M_2 do genitor P_1 (CRUZ, 2005, comunicação pessoal).

O estudo da ligação entre um loco Q e uma marca M pode ser feito com base na média dos indivíduos agrupados de acordo com a presença de diferentes alelos do loco M. A hipótese de associação entre os locos Q e M pode ser avaliada pela significância do contraste entre essas médias, através da estatística t (COELHO, 2000).

Retrocruzamento ($M_1M_2 \times M_3M_3$)

Um pai é heterozigoto para o marcador, e o outro é homozigoto, conseqüentemente toda a progênie é informativa, uma vez que é possível, por meio de seu genótipo, a distinção da origem de seus alelos em relação a cada um dos genitores (DA; LEWIN, 1995). Desse modo, em relação ao loco M, são identificadas duas classes genotípicas, das quais M_1M_3 é representada por indivíduos portadores do alelo M_1 do genitor P_1 e M_2M_3 é representada por indivíduos portadores do alelo M_2 do genitor P_1 . Assim, estabelece-se o

contraste entre as médias e utiliza-se a estatística t para avaliar a existência da ligação entre os locos M e Q (COELHO, 2000).

Cruzamentos ($M_1M_2 \times M_1M_2$)

Ambos os pais apresentam o mesmo marcador heterozigoto e somente a progênie homozigota é informativa, pois a descendência heterozigota não possibilita, por meio de seu genótipo, distinguir a origem de seus alelos em relação aos dois genitores (DA; LEWIN, 1995).

Assim, em relação ao loco M, são identificadas três classes genóticas, mas apenas duas são informativas, das quais M_1M_1 é representada por indivíduos portadores do alelo M_1 do genitor P_1 e M_2M_2 o é por aqueles portadores do alelo M_2 do genitor P_1 . Desse modo, o contraste entre as médias pode ser estabelecido e sua significância, avaliada pela estatística t (COELHO, 2000). Portanto, observa-se que, nesse caso, a detecção do QTL realizada por meio deste contraste envolve apenas os indivíduos informativos.

2.5.2. Freqüência de genótipos não informativos

A falta de informatividade dos heterozigotos faz que as avaliações destes sejam descartadas, pois não é possível distinguir qual genitor foi o doador da marca, uma vez que ambos apresentam-na.

A freqüência de indivíduos não informativos varia entre os tipos de cruzamentos e depende do conhecimento do genótipo dos pais, como pode ser observado na Tabela 1, adaptada de Da e Lewin (1995) e Guo e Elston (1999).

Tabela 1 - Frequência de genótipos não informativos para um loco nos diferentes tipos de cruzamentos em uma população de acasalamento ao acaso

Genótipo do Genitor Feminino	Genótipo do Genitor Masculino	Frequência do Cruzamento	Genótipo da Descendência	Frequência de Genótipos Não Informativos		
				I^*_2	I^*_1	I^*_0
M_1M_1	M_1M_1	p^4	1 M_1M_1	0	0	0
	M_1M_2	$2p^3q$	$\frac{1}{2} M_1M_1 + \frac{1}{2} M_1M_2$	0	0	$\frac{1}{2}$
	M_2M_2	p^2q^2	1 M_1M_2	0	0	1
M_1M_2	M_1M_1	$2p^3q$	$\frac{1}{2} M_1M_1 + \frac{1}{2} M_1M_2$	0	$\frac{1}{2}$	$\frac{1}{2}$
	M_1M_2	$4p^2q^2$	$\frac{1}{4} M_1M_1 + \frac{1}{2} M_1M_2 + \frac{1}{4} M_2M_2$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	M_2M_2	$2pq^3$	$\frac{1}{2} M_1M_2 + \frac{1}{2} M_2M_2$	0	$\frac{1}{2}$	$\frac{1}{2}$
M_2M_2	M_1M_1	p^2q^2	1 M_1M_2	0	0	1
	M_1M_2	$2pq^3$	$\frac{1}{2} M_1M_2 + \frac{1}{2} M_2M_2$	0	0	$\frac{1}{2}$
	M_2M_2	q^4	1 M_2M_2	0	0	0
Soma		1,0		i_2	i_1	i_0

* i_2 é a frequência quando se conhece o genótipo de ambos os genitores; i_1 , quando se conhece apenas o genótipo de um genitor e i_0 , quando se desconhecem ambos os genitores.

Considerando-se o conhecimento do genótipo de ambos os pais, por exemplo no delineamento de irmãos-completos, parte dos descendentes do cruzamento $M_1M_2 \times M_1M_2$ é não informativa. Isso corresponde a uma porcentagem de informatividade de 50%, o que resulta na perda de metade dos investimentos em experimentos que envolvem tal cruzamento.

Nas populações em que é possível o conhecimento de apenas um dos pais, por exemplo, nas populações de meios-irmãos observa-se que a frequência de indivíduos não informativos é maior que nos cruzamentos em que ambos os pais são conhecidos. Nesse caso, descendentes heterozigotos M_1M_2 originados a partir do cruzamento de um genitor conhecido também heterozigoto e um genitor homozigoto desconhecido são também não informativos, ao contrário dos cruzamentos em que os dois pais são conhecidos. Quando os genótipos de ambos os pais são desconhecidos, o número de indivíduos não informativos é ainda maior, como pode ser observado na Tabela 1.

A partir da Tabela 1, pode-se estimar a freqüência de genótipos não informativos em uma população de acasalamento ao acaso, nas três diferentes situações. Quando os genótipos de ambos os pais são conhecidos (i_2), essa freqüência pode ser calculada por meio da expressão (1); se apenas o genitor feminino é conhecido (i_1), o cálculo pode ser realizado por meio da expressão (2); e se o genótipo de ambos os pais são desconhecidos (i_0), a freqüência pode ser estimada por meio da expressão (3).

$$i_2 = \frac{1}{2} 4p^2q^2 = 2p^2q^2 \quad (1)$$

$$i_1 = \frac{1}{2} 2p^3q + \frac{1}{2} 4p^2q^2 + \frac{1}{2} 2pq^3 = pq \quad (2)$$

$$i_0 = \frac{1}{2} 2p^3q + p^2q^2 + \frac{1}{2} p^3q + \frac{1}{2} 4p^2q^2 + \frac{1}{2} 2pq^3 + p^2q^2 + \frac{1}{2} 2pq^3 = 2pq \quad (3)$$

Pelas equações anteriores, observa-se que a determinação da freqüência de genótipos não informativos é função da freqüência alélica. Na Figura 1 estão plotados os gráficos da porcentagem de indivíduos não informativos obtidos a partir das expressões citadas, para diferentes freqüências de p e q.

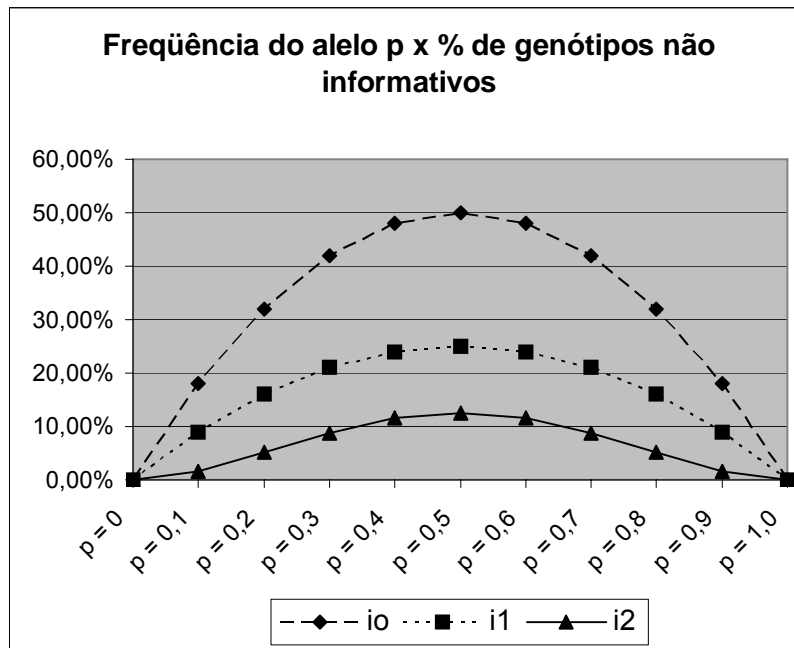


Figura 1 - Gráfico da distribuição da frequência de genótipo não informativos em populações de acasalamento ao acaso, em diferentes frequências de p e q, quando não se conhece o genótipo dos dois pais (i_0), quando apenas um dos pais tem genótipo conhecido (i_1) e quando ambos os genótipos paternos são conhecidos (i_2).

Nos estudos de populações de acasalamento ao acaso, a frequência de genótipos não informativos máxima é de 50%, quando as frequências de p e q são iguais a 0,5 e ambos os genitores são desconhecidos.

Nas populações em que o genótipo de todos os genitores são conhecidos, ou de pelo menos um deles, os valores máximos de descendentes não informativos foram 12,5 e 25%, respectivamente, quando em ambas as situações p e q são também iguais a 0,5.

Em populações F_2 de mapeamento, os pais devem ser contrastantes genotipicamente para cada gene em questão. Assim, espera-se que as frequências de p e q sejam iguais a 0,5 nessas populações e que a porcentagem de indivíduos não informativos seja de 12,5% quando os genótipos dos dois pais são conhecidos, 25% quando apenas um pai é conhecido e 50% quando nenhum pai é conhecido. Já nas populações de retrocruzamento, espera-se uma frequência de p = 0,75 e q = 0,25, o que corresponde a uma frequência de indivíduos não informativos de 7,03%, também quando os genótipos de ambos os pais são conhecidos, de 18,75%

quando apenas um é conhecido e de 37,5% quando os dois são desconhecidos. Assim, faz-se necessária a busca por estratégias que aumentem o índice de informatividade em cada um desses cruzamentos.

2.5.3. O endosperma como estratégia na detecção de QTL

O endosperma é um tecido de reserva que garante a nutrição do embrião em desenvolvimento. Nas Angiospermas, é formado pela fusão de um gameta masculino com os dois núcleos polares do gametófito feminino. Essa fusão dá origem ao núcleo triplóide do endosperma, que se desenvolve por divisões mitóticas e apresenta, no seu conjunto genômico, duas cópias dos alelos maternos e uma cópia do alelo paterno (WU et al., 2002).

Uma das dificuldades na detecção de QTLs em populações exogâmicas está na falta de informatividade dos descendentes heterozigotos M_1M_2 . Uma possível forma de incluí-los nas análises é por meio da genotipagem do seu endosperma. Espera-se que um indivíduo heterozigoto M_1M_2 apresente um endosperma $M_1M_1M_2$ ou $M_1M_2M_2$, quando os genitores femininos são doadores do alelo M_1 ou M_2 , respectivamente. Assim, independentemente do conhecimento dos genótipos parentais, pode-se determinar a origem de ambos os alelos nos descendentes heterozigotos, de forma que esses passam a compor as médias no contraste para avaliação da hipótese de associação entre o loco M e o QTL Q.

Como exemplo, pode-se considerar a constituição de uma família de irmãos-completos a partir de um cruzamento entre um genitor feminino ($M_1Q_1 // M_2Q_2$) e um genitor masculino ($M_1Q_3 // M_2Q_4$) (Tabela 2).

Tabela 2 - Genótipo do endosperma (END) e do embrião (EMB) dos descendentes de uma família de irmãos-completos derivada do cruzamento $M_1Q_1//M_2Q_2$ (genitor feminino) x $M_1Q_3//M_2Q_4$ (genitor masculino)

Gametas ♂	$M_1Q_3(1-r)$		$M_1Q_4(r)$		$M_2Q_3(r)$		$M_2Q_4(1-r)$	
Gametas ♀	END	EMB	END	EMB	END	EMB	END	EMB
$M_1Q_1(1-r)$	$M_1M_1M_1$	M_1M_1	$M_1M_1M_1$	M_1M_1	$M_1M_1M_2$	$M_1M_2^*$	$M_1M_1M_2$	$M_1M_2^*$
$M_1Q_2(r)$	$M_1M_1M_1$	M_1M_1	$M_1M_1M_1$	M_1M_1	$M_1M_1M_2$	$M_1M_2^*$	$M_1M_1M_2$	$M_1M_2^*$
$M_2Q_1(r)$	$M_2M_2M_1$	$M_1M_2^*$	$M_2M_2M_1$	$M_1M_2^*$	$M_2M_2M_2$	M_2M_2	$M_2M_2M_2$	M_2M_2
$M_2Q_2(1-r)$	$M_2M_2M_1$	$M_1M_2^*$	$M_2M_2M_1$	$M_1M_2^*$	$M_2M_2M_2$	M_2M_2	$M_2M_2M_2$	M_2M_2

* Indivíduos heterozigotos, até então não informativos, pela incapacidade de distinção da origem de seus alelos em relação aos seus genitores.

Assim, em relação ao loco M, poderiam ser identificadas na descendência quatro classes genóticas e não apenas duas. $M_1M_1M_1$ e $M_1M_1M_2$ representam indivíduos portadores do alelo M_1 e $M_2M_2M_2$ e $M_2M_2M_1$, representam indivíduos portadores do alelo M_2 , recebidos dos genitores femininos. As médias fenotípicas esperadas de cada uma das classes genóticas estão representadas na Tabela 3.

Tabela 3 - Médias fenotípicas esperadas de cada uma das classes genóticas

Genótipo	Média	Esperado
$M_1 M_1 M_1$	μ_{11}	$(1-r)^2 \mu_{Q_1Q_3} + r(1-r) \mu_{Q_1Q_4} + r(1-r) \mu_{Q_2Q_3} + r^2 \mu_{Q_2Q_4}$
$M_1 M_1 M_2$	μ_{12}	$r(1-r) \mu_{Q_1Q_3} + (1-r)^2 \mu_{Q_1Q_4} + r^2 \mu_{Q_2Q_3} + r(1-r) \mu_{Q_2Q_4}$
$M_2 M_2 M_1$	μ_{21}	$r(1-r) \mu_{Q_1Q_3} + r^2 \mu_{Q_1Q_4} + (1-r)^2 \mu_{Q_2Q_3} + r(1-r) \mu_{Q_2Q_4}$
$M_2 M_2 M_2$	μ_{22}	$r^2 \mu_{Q_1Q_3} + r(1-r) \mu_{Q_1Q_4} + r(1-r) \mu_{Q_2Q_3} + (1-r)^2 \mu_{Q_2Q_4}$

Uma das maneiras de estudar a ligação entre os locos Q e M é agrupar as médias dos indivíduos portadores de determinado alelo do loco M. Assim, pode-se estabelecer:

$$\mu_1 = \frac{\mu_{11} + \mu_{12}}{2} = (1-r) \left(\frac{\mu_{Q_1Q_3} + \mu_{Q_1Q_4}}{2} \right) + r \left(\frac{\mu_{Q_2Q_3} + \mu_{Q_2Q_4}}{2} \right) e$$

$$\mu_2 = \frac{\mu_{21} + \mu_{22}}{2} = r \left(\frac{\mu_{Q1Q3} + \mu_{Q1Q4}}{2} \right) + (1-r) \left(\frac{\mu_{Q2Q3} + \mu_{Q2Q4}}{2} \right)$$

O contraste é, então, expresso por meio de:

$$\delta = \mu_1 - \mu_2 = (1-2r) \left(\frac{\mu_{Q1Q3} + \mu_{Q1Q4}}{2} - \frac{\mu_{Q2Q3} + \mu_{Q2Q4}}{2} \right)$$

Expressando as médias dos genótipos relativos ao loco Q em função dos efeitos aditivos e da dominância, tem-se:

$$\delta = (1-2r) \left[(\alpha_1 - \alpha_2) + \frac{\delta_{13} + \delta_{14}}{2} - \frac{\delta_{23} + \delta_{24}}{2} \right]$$

Assim, a hipótese de associação entre os locos Q e M pode ser avaliada pela significância do contraste entre as médias, utilizando-se a estatística t, para avaliar $H_0: \delta = 0$.

Na ausência de ligação entre os locos M e Q, o contraste δ é nulo e nenhum QTL é detectado.

Com a inclusão dos descendentes heterozigotos no cálculo das médias que irão compor o contraste, espera-se maior acurácia na utilização da estatística t, como forma de detecção dos QTLs. Assim, faz-se necessária a adaptação de uma metodologia de genotipagem de endosperma capaz de determinar a origem dos alelos a partir de um mínimo de DNA, para que a sua extração não afete a germinação da semente e o desenvolvimento da planta.

CAPÍTULO 1

Germinação de sementes de milho submetidas à remoção parcial do endosperma e do pericarpo

1. INTRODUÇÃO

A semente de milho, botanicamente definida como cariópse, é composta, ordinariamente, por embrião, endosperma e pericarpo (GALINAT, 1979). O embrião representa a futura planta em seu estado primitivo, enquanto, endosperma e pericarpo se constituem nos acessórios encarregados de nutrir e proteger, inicialmente, o embrião durante a germinação (CARVALHO; NAKAGAWA, 2000). Nesse processo, contudo, as funções desempenhadas pelo endosperma e pelo pericarpo vão perdendo a importância à medida que a plântula passa a desenvolver os mecanismos autotróficos da vida (GOMES; KARAZAWA, 1982).

A presença de resíduos endospermáticos, envoltos pelo pericarpo, acoplados às plantas capazes de realizar fotossíntese e, portanto, em condições de subsidiar as suas necessidades nutricionais, evidencia a existência de reservas superiores às necessárias para emergência em condições ambientais adequadas. Assim, a retirada, em pré-semeadura, da parte endospermática disponível em excesso para o estabelecimento da plântula parece biologicamente viável (Da SILVA et al., 1994). Esses autores verificaram a interferência da excisão de partes da semente de milho no desempenho biológico, a partir de tratamentos representados por remoções aproximadas de 0, 14, 28 e 37%, em peso, sem que o embrião fosse diretamente afetado. A avaliação, conduzida em laboratório e em campo, permitiu concluir que há possibilidade biológica de excisar parte ($\pm 14\%$ em peso) da semente de milho, em pré-semeadura e com tratamento fungicida, em lotes de alta qualidade fisiológica.

Nas Angiospermas, o endosperma é formado pela fusão de um gameta masculino com os dois núcleos polares do gametófito feminino. Essa fusão dá origem ao núcleo triplóide do endosperma, que se desenvolve por divisões mitóticas e apresenta, no seu conjunto genômico, duas cópias dos alelos maternos e uma cópia do alelo paterno (WU et al., 2002). Essa característica torna-se interessante, sobretudo, nos estudos de mapeamento de populações em que os indivíduos heterozigotos são não informativos (LIU, 1998) e suas avaliações, até então, são descartadas, pois não se pode distinguir qual o genitor é o doador de determinada marca M_1 ou M_2 . Assim, estratégias de identificação da origem dos alelos se paternos ou maternos são de grande interesse.

Uma possível forma de incluir os indivíduos heterozigotos nas análises seria por meio da genotipagem do seu endosperma, na qual se espera que um indivíduo heterozigoto M_1M_2 apresente um endosperma $M_1M_1M_2$ ou $M_1M_2M_2$, quando os genitores femininos são doadores do alelo M_1 ou M_2 , respectivamente.

Dessa forma, o objetivo deste trabalho foi determinar a quantidade mínima necessária de endosperma à germinação da semente de milho. Para isso, foram realizados tratamentos representados por remoções aproximadas de 10, 25 e 50%, em peso, sem que o embrião fosse diretamente afetado. A preservação do embrião é necessária para o desenvolvimento da planta, a qual será utilizada posteriormente para o mapeamento genético.

2. MATERIAL E MÉTODOS

Material vegetal

Foram utilizadas sementes da variedade de milho BR201, de alta qualidade fisiológica, fornecidas pela Embrapa Milho e Sorgo.

Metodologia

A retirada parcial do endosperma e do pericarpo foi realizada por meio de cortes dirigidos, sem que o embrião fosse atingido. Para tanto, as sementes foram pesadas individualmente e, em seguida, submetidas aos cortes, com o auxílio de uma lâmina de barbear e bisturi, que representaram remoções aproximadas de 10, 25 e 50%, em peso (Figura 1). Esses cortes constituíram, respectivamente, os tratamentos 1, 2 e 3, além do controle que constitui a semente íntegra. Após a remoção, as sementes foram germinadas em rolos de papéis germitest embebidos em água, mantidas em germinador com temperatura controlada entre 20 e 30° C e fotoperíodo de oito horas. As avaliações foram realizadas no quinto e sétimo dias após o início do experimento. Os testes de germinação foram realizados de acordo com as Regras de Análise de Sementes-RAS (BRASIL, 1992), fornecendo as taxas de plântulas normais, anormais e de sementes mortas.

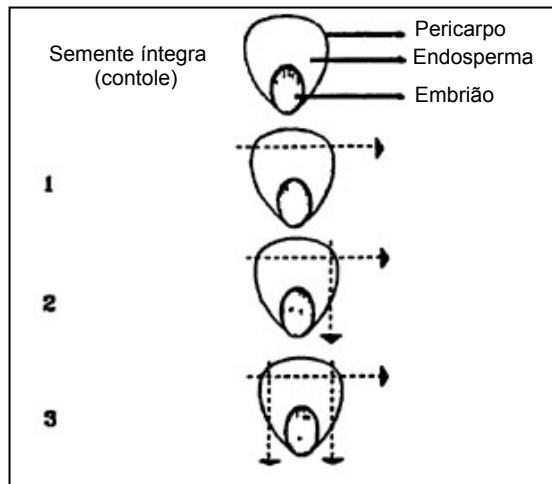


Figura 1 - Cortes realizados na semente de milho 1, 2 e 3 representam remoções aproximadas de 10, 25 e 50%, em peso.

Foram utilizadas três repetições em delineamento inteiramente casualizado, sendo cada repetição composta por 30 sementes, totalizando 90 sementes por tratamento. Os dados coletados foram submetidos à análise de variância, segundo o modelo apresentado a seguir:

$$y_{ij} = \mu + t_i + \varepsilon_{ij}.$$

Em que:

y_{ij} = observação referente ao i -ésimo tratamento e j -ésima repetição ($i = 1, 2, 3$ e $j = 1, 2$ e 3);

μ = média geral;

t_i = efeito do i -ésimo tratamento e

ε_{ij} = erro aleatório.

3. RESULTADOS E DISCUSSÃO

Os dados relativos à distribuição porcentual das frações obtidas após a excisão de parte do endosperma nas sementes estão apresentados na Tabela 1. Assim, pôde-se determinar a quantidade de endosperma, em miligramas, extraído, em média, nos cortes 1, 2 e 3, correspondentes às porcentagens 10, 25 e 50%, respectivamente.

Tabela 1 - Distribuição porcentual da parte excisada (Ex) e da porção contendo o embrião (Em). Quantidade média de endosperma, em miligramas, de cada porção após a aplicação dos cortes 1, 2 e 3

Cortes	Fração	% em Massa	Massa Média (mg)
controle	Ex	0,0	0,0
	Em	100,0	350,0
1	Ex	10,0	34,0
	Em	90,0	316,0
2	Ex	25,0	85,0
	Em	75,0	265,0
3	Ex	50,0	175,0
	Em	50,0	175,0

As 90 sementes germinadas avaliadas para cada tratamento foram classificadas em três categorias, de acordo com a RAS: plantas normais, anormais e sementes mortas. Em seguida, as médias dos tratamentos para cada classe foram submetidos à análise de variância, em nível de 5% de probabilidade (TABELA 2).

Tabela 2 - Dados médios, obtidos no teste de germinação, das taxas de plântulas normais (PN), anormais (PA) e sementes mortas (SM). ANOVA a 5% de probabilidade. As médias de cada tratamento foram calculadas a partir de três repetições, com 30 sementes cada

Classe	Médias dos Tratamentos				Probabilidade
	Controle	1	2	3	
PN	28,33	26,67	28,33	28,00	35,1359%
PA	1,00	2,67	1,67	1,67	30,6703%
SM	0,67	0,67	0,00	0,33	36,3687%

Apesar das diferenças aparentes, a análise de variância dos dados indicou não existir diferença entre as médias dos tratamentos e do controle tanto para plântulas normais e anormais quanto para sementes mortas.

O aumento não significativo do número de plântulas anormais nos tratamentos pode ser explicado pelas dificuldades de remoção do endosperma sem afetar o embrião. Segundo Da Silva et al. (1994), a taxa de plântulas anormais nas sementes cortadas pode ser reduzida com o uso do fungicida Captan 75%, na dosagem de 1,5 g/kg de sementes.

No geral, a retirada de parte do endosperma e pericarpo não afetou a germinação do embrião. Assim, os resultados indicaram que até 50% da massa do endosperma pode ser removida sem prejuízo na germinação, o que correspondeu, em média, a 175 mg, podendo variar de acordo com o tipo de grão. Estabelecida a quantidade de endosperma que pode ser retirada, sem prejuízo para o desenvolvimento da plântula, pôde-se seguir com análises moleculares com a extração de DNA do fragmento excisado para genotipagem e identificação da origem dos alelos.

CAPÍTULO 2

Adaptação da metodologia de PCR semiquantitativo utilizando marcadores SSRs para diferenciação de amostras quanto ao número de cópias de um alelo

1. INTRODUÇÃO

Grande impulso para a biologia molecular foi dado em meados da década de 1980, quando Karl Mullis divulgou a técnica denominada *Polymerase Chain Reaction* (PCR). Essa técnica envolve a síntese *in vitro* de um segmento específico de DNA na presença da DNA polimerase. A reação baseia-se no anelamento e extensão enzimática de um par de oligonucleotídeos utilizados como iniciadores (*primers*), que delimitam a seqüência de DNA-alvo a ser amplificada. A partir daí, várias foram as técnicas derivadas da PCR, permitindo avanços significativos nas áreas da Genética e da Biologia Molecular (PADILHA, 2002).

Dentre as metodologias mais recentes, desenvolvidas a partir da PCR para a quantificação de doses diferenciadas de um alelo, destaca-se a técnica de PCR em tempo real ou PCR quantitativo. Nessa técnica, utilizaram-se sondas marcadas com fluoresceína, para determinar, comparando-se com uma curva-padrão, a quantidade de fluorescência emitida e, assim, determinar a quantidade de moléculas amplificadas. Apesar de sensível e confiável, essa metodologia ainda é cara e requer equipamentos específicos, limitando o seu uso rotineiro em vários laboratórios (MARIN et al., 2002). Assim, faz-se necessária a adaptação de metodologias mais simples e menos onerosas para a quantificação de alelos.

Nesse sentido, a metodologia de PCR semiquantitativo (PCR-SQ) vem sendo adaptada com sucesso para o diagnóstico e determinação da carga viral em humanos infectados pelo citomegalovírus (MARIN et al., 2004). No entanto, é na análise de expressão gênica que a RT-PCR-SQ tem sido utilizada com maior freqüência, a exemplo de estudos com abelhas para quantificação dos

valores relativos de mRNA da enzima metiltransferase em rainhas e operárias de diferentes idades, no conhecimento da reprodução e organização social em *Melipona scutellaris* (VIEIRA et al., 2004).

Embora o PCR semiquantitativo não permita uma quantificação absoluta, espera-se que seja possível inferir quanto ao número de moléculas amplificadas em cada genoma se o seu produto for avaliado na fase exponencial, em que o número de cópias é função da quantidade inicial (GÉNÉTHON, 2001). Assim, ao contrário do PCR em tempo real, o PCR semiquantitativo possui a necessidade do conhecimento prévio da cinética da reação, uma vez que é preciso determinar o tempo de reação ideal no qual os produtos do PCR correspondem à fase exponencial, em que as estimativas podem ser mais acuradas.

A falta de informatividade dos indivíduos heterozigotos é uma das dificuldades do mapeamento de QTLs em populações exogâmicas (LIU, 1998), uma vez que não é possível identificar a origem de cada um dos seus alelos. Com isso, tais indivíduos são excluídos das análises do contraste entre médias, não contribuindo para a detecção do QTL. Uma possibilidade de incluí-los nas análises seria por meio da genotipagem do seu endosperma, em que se espera que um indivíduo heterozigoto M_1M_2 apresente um endosperma $M_1M_1M_2$ ou $M_1M_2M_2$, quando os genitores femininos são doadores do alelo M_1 ou M_2 , respectivamente.

Devido ao potencial uso do PCR-SQ para inferir o número de moléculas amplificadas, o objetivo deste estudo foi adaptar essa metodologia para a determinação do número de cópias dos alelos em mistura de DNA que simulam as condições observadas no tecido endospermático.

2. MATERIAL E MÉTODOS

Material vegetal

Foram utilizadas folhas das linhagens endogâmicas de milho L3 e L1113-01 cedidas pelo Programa de Melhoramento de Milho da Embrapa Milho e Sorgo. Os procedimentos moleculares foram realizados no Laboratório de Biologia Molecular do Núcleo de Biologia Aplicada da Embrapa Milho e Sorgo, em Sete Lagoas, MG.

Extração de DNA foliar

Devido à pouca quantidade de DNA esperada das extrações de endosperma, inicialmente optou-se pela utilização do DNA genômico foliar na otimização da técnica de PCR semiquantitativo. Dessa forma, extraiu-se o DNA das linhagens parentais, segundo o método descrito por Saghai-Marroof et al. (1984), em que um volume de 5 mL de tampão CTAB (CTAB 2%; Tris-HCl 0,2 M; NaCl 1,4 M; EDTA 0,02 M e β -Mercaptoetanol 2%) foi adicionado a aproximadamente 5 mg do material vegetal moído em N₂ líquido até a obtenção de um pó fino. A mistura foi mantida a 65 °C durante 1 h, com homogeneizações a cada 15 min. Após esse período, foi realizada uma lavagem com igual volume de clorofórmio-octanol (24:1), com homogeneização constante por 10 min. O material foi centrifugado a 15.000 *xg* por 10 min e o sobrenadante, transferido para outro tubo, onde foi repetida a lavagem. Após a centrifugação, o DNA foi precipitado com isopropanol, ressuspendido em 3 mL de TE (Tris-HCl 10 mM; EDTA 1 mM; pH 8,0) e tratado com 20 μ L RNase A (10

mg/mL) durante 1 h a 37 °C. O DNA foi novamente precipitado com etanol 95%, lavado com etanol 70% e ressuspendido em TE.

A quantificação do DNA foi realizada em gel de agarose 0,8%, utilizando-se um padrão de DNA de concentração conhecida. Após a realização da eletroforese a 100 V por 1 h, o gel foi tratado com brometo de etídeo (1 mg/mL) e visualizado sob luz ultravioleta (UV) no Eagle Eye II (Stratagene).

Após a quantificação, o DNA-estoque foi mantido a -20 °C e o DNA, diluído na concentração de trabalho, sendo armazenado a 4 °C. Nos procedimentos posteriores foram utilizados os DNAs dos parentais individuais, bem como misturas em diferentes proporções destes, na tentativa de simular a constituição do endosperma (Tabela 1).

Tabela 1 - Misturas em diferentes proporções do DNA dos parentais, utilizadas nas reações para simulação da situação real no endosperma

Identificação	Qt de DNA L3	Qt de DNA L1113-01
1. L3	30 ng	-
2. L1113-01	-	30 ng
3. Mistura 1:1	15 ng	15 ng
4. Mistura 1:2	10 ng	20 ng
5. Mistura 2:1	20 ng	10 ng

Otimização da técnica de PCR semiquantitativo em agarose

Inicialmente, três *primers* SSRs, polimórficos entre as linhagens parentais, foram selecionados baseando-se no estudo de informatividade e padrão de amplificação realizado por Padilha (2002) (Tabela 2). Devido às limitações de resolução no gel de agarose, foram utilizados apenas *primers* cujo polimorfismo foi superior a 18 pb.

Tabela 2 - *Primers* polimórficos entre as linhagens parentais L3 e L1113-01. O Bin refere-se à posição no cromossomo, as seqüências destacadas são os núcleos de repetição característicos de cada *primer* e os valores em pb indicam o tamanho do alelo presente em cada uma das linhagens

<i>Primer</i>	Bin	Núcleo de Repetição	Alelo L1113 (pb)	Alelo L3 (pb)
<i>umc1653</i>	6.07	(GAAA) ₂₄	93	112
<i>bnlg1006</i>	5.00	(AG) ₂₀	224	171
<i>bnlg182</i>	1.03	?	400	112

As reações de amplificação foram realizadas em um volume final de 10 µL contendo 30 ng de DNA, Tris-HCl 1 mM, KCl 50 mM, MgCl₂ 2 mM, 125 µM de cada um dos dNTPs, 0,6 µM de cada um dos *primers* e 1 U da enzima Taq polimerase. Os ciclos da reação de PCR consistiram de uma desnaturação inicial a 95 °C por 2 min, nove ciclos de 94 °C por 20 s, 68 °C por 20 s com a redução de 1 °C a cada ciclo e 72 °C por 20 s, seguidos de 15 a 25 ciclos de amplificação a 94 °C por 20 s, 60 °C por 20 s e 72 °C por 20 s, no estudo da Cinética de Amplificação.

Os fragmentos amplificados foram observados em géis de agarose 3% sob eletroforese a 80 V, por 2 h. O gel foi tratado com brometo de etídeo (1 mg/mL) e a imagem, capturada pelo Eagle-Eye II, sob luz UV. As imagens foram analisadas pelo programa ImageQuANT[®] (Molecular Dynamics), para quantificação das intensidades das bandas.

Após a determinação do número de ciclos ideal para a visualização das diversas proporções de DNA das linhagens L3 e L1113-01 nas misturas, novas amplificações foram realizadas para os *primers umc1653, bnlg1006 e bnlg182*, na tentativa de propor uma metodologia aplicável a diferentes *primers* SSRs. Foram realizadas três repetições para cada *primer* e o número de ciclos de amplificação total foi fixado em 30, sendo mantidas as demais condições de reação.

3. RESULTADOS

Otimização da técnica de PCR semiquantitativo em agarose

Cinética da reação

O estudo da cinética da reação com os *primers bnlg182, bnlg1006 e umc1653* permitiu a determinação do melhor tempo de reação, no qual foi possível a visualização das diferentes proporções nas misturas.

Os fragmentos amplificados, observados em agarose 3%, e tratados com brometo de etídio sob luz UV, foram observados a partir de 28 ciclos de amplificação. As bandas correspondentes aos fragmentos de menor tamanho se apresentaram mais difusas, em comparação com as bandas dos fragmentos maiores, dificultando a análise da intensidade relativa entre as bandas das diferentes misturas (Figura 1).

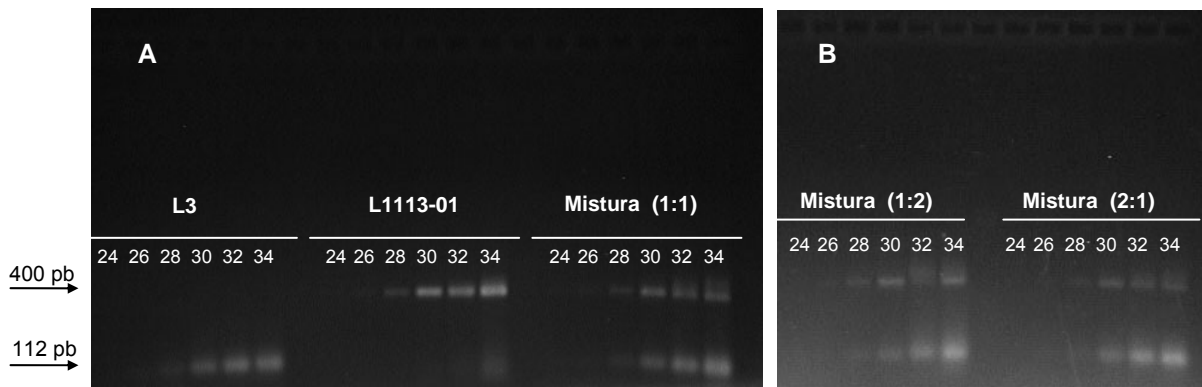


Figura 1 - Eletroforese em gel de agarose 3%, tratado com brometo de etídio, dos produtos amplificados com o *primer* SSR *bnlg182* das amostras de DNA foliar e das respectivas misturas. Gel A: produtos da reação com amostras de DNA das linhagens L3, L1113-01 e da mistura 1:1. Gel B: produtos da reação com as misturas (1:2) e (2:1). O número de ciclos de amplificação variou de 24 a 34, com intervalo de dois ciclos.

Após a análise da intensidade das bandas pelo programa ImageQuANT[®], os valores correspondentes a cada leitura foram plotados em gráficos, como apresentado nas Figuras 2 e 3.

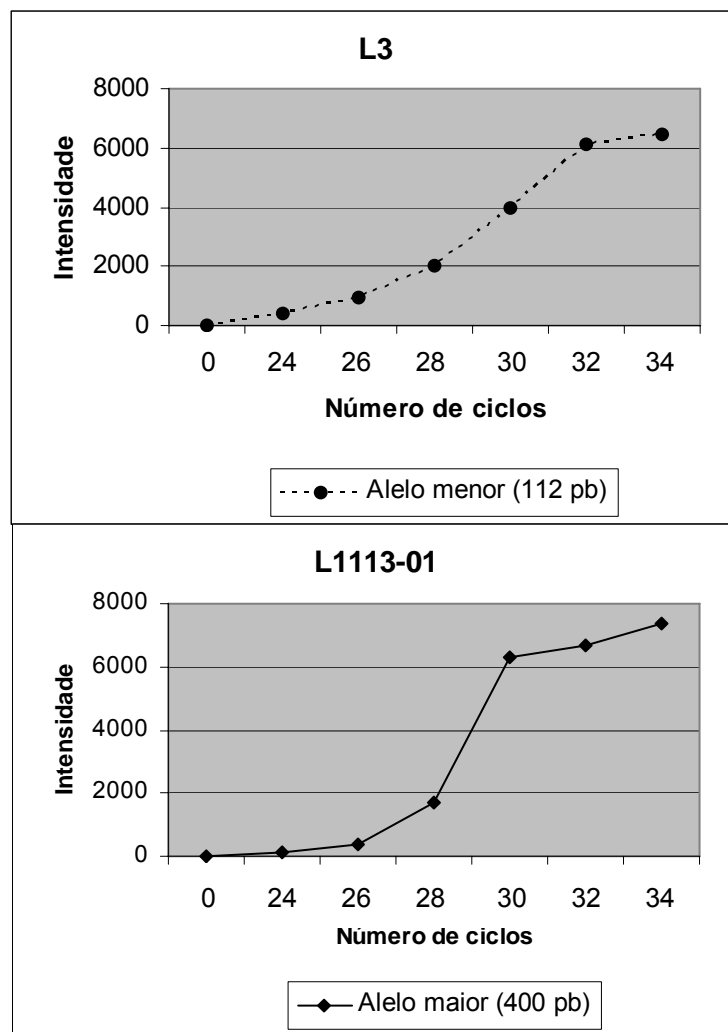


Figura 2 - Intensidade dos fragmentos amplificados de DNA foliar das linhagens L3 e L1113-01. As curvas obtidas correspondem à cinética da reação de amplificação pelo *primer bnlg182*, que variou entre 24 e 34 ciclos.

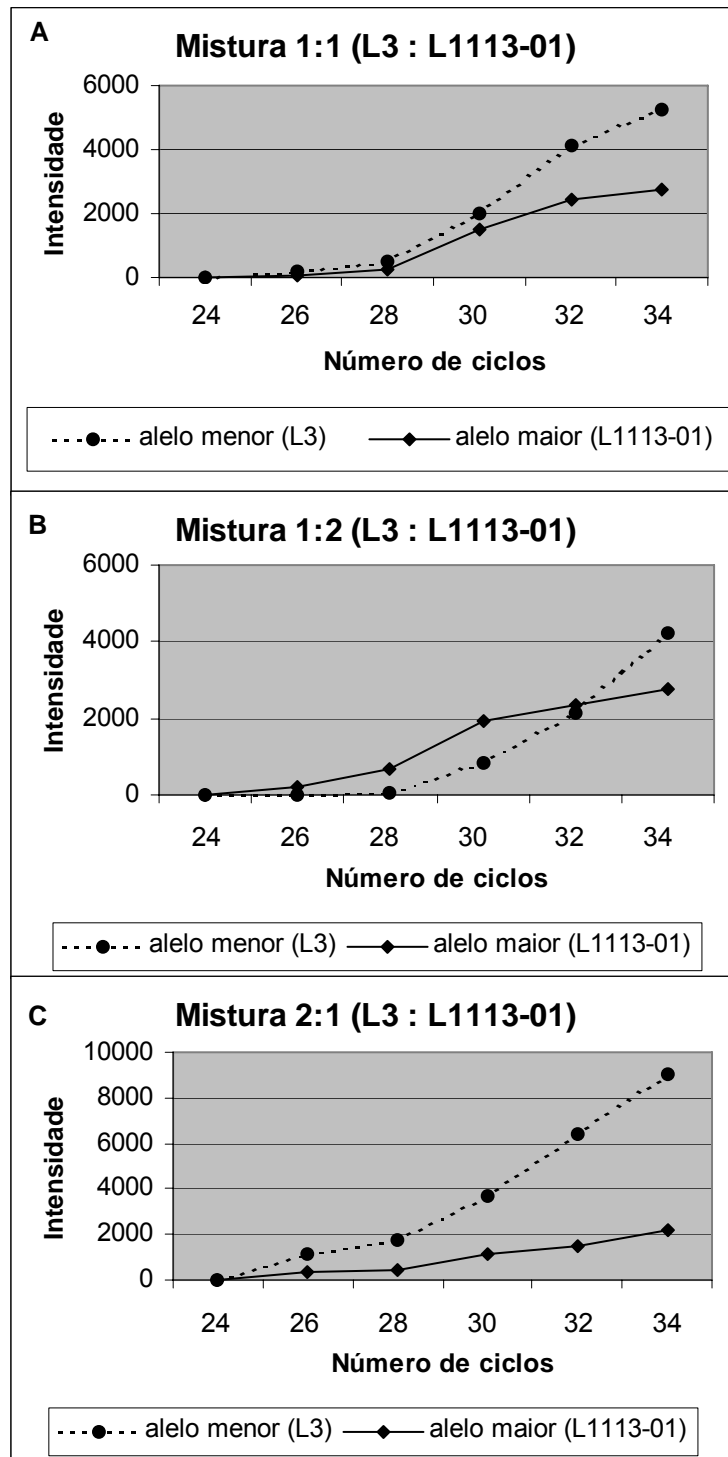


Figura 3 - Intensidade dos fragmentos amplificados das misturas de DNA foliar, nas diferentes proporções das linhagens L3 e L1113-01. As curvas obtidas correspondem à cinética da reação de amplificação pelo *primer bnlg182*, que variou entre 24 e 34 ciclos. A: produtos da reação de amplificação da mistura 1:1 (15 ng do DNA L3 e 15 ng do DNA L1113-01). B: produtos da reação de amplificação da mistura 1:2 (10 ng do DNA L3 e 20 ng do DNA L1113-01). C: produtos da reação de amplificação da mistura 2:1 (20 ng de DNA L3 e 10 ng de DNA L1113-01).

A partir das análises dos gráficos, observou-se que a reação de amplificação com 30 ciclos foi a que melhor representou a fase exponencial, tanto no DNA genômico individual quanto nas diferentes misturas, utilizando o *primer* SSR *bnlg182*. Nos demais *primers* estudados, embora não tenha sido possível a observação gráfica da proporção 2:1 (L3 : L1113-01) em nenhum momento da curva, considerou-se também o número de ciclos ideal igual a 30. Entre os pontos analisados, esse se apresentou ligeiramente anterior ao início da fase de saturação, correspondendo, ainda, a fase exponencial, em que o número de amplificações é função da quantidade inicial de cada alelo.

O estudo da cinética da reação de amplificação do *primer* *bnlg182* apontou que, nas misturas do DNA, a amplificação do fragmento de maior tamanho atinge a fase de saturação com menor número de ciclos, e a intensidade do alelo de menor tamanho foi superior, exceto em alguns ciclos onde a proporção do DNA original foi a metade. Tal perfil de amplificação indica que o alelo de menor tamanho foi preferencialmente amplificado na reação de PCR. No entanto, tal padrão de amplificação não foi constante entre os diferentes *primers* SSR, uma vez que a cinética da reação de amplificação com o *primer* *bnlg1006* indicou uma preferência pela amplificação do alelo de maior tamanho.

Pôde-se observar ainda que, quanto menor for a diferença de tamanho entre os alelos polimórficos, menor será a diferença de intensidade entre as bandas quando o DNA das linhagens é misturado e que o número de ciclos de amplificação para atingir a fase de saturação tende a ser mais semelhante. Um exemplo foi verificado com o estudo da cinética da reação de amplificação do *primer* *umc1653*, que apresenta 18 pb de polimorfismo (Figura 4).

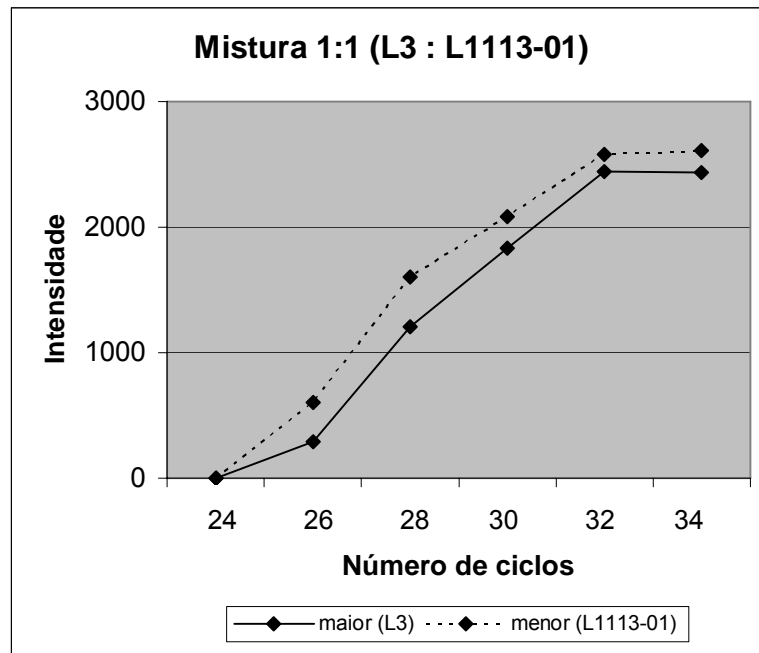


Figura 4 - Intensidade dos fragmentos amplificados das misturas de DNA foliar, na proporção 1:1 das linhagens L3 e L1113-01. As curvas obtidas correspondem à cinética da reação de amplificação pelo *primer umc1653*, que variou entre 24 e 34 ciclos. Observar que ambas as curvas atingem a saturação a partir de 32 ciclos.

Ajuste de um modelo de regressão para a reação de amplificação

As análises visuais dos fragmentos produzidos utilizando 30 ciclos de amplificação permitiram a diferenciação de quantidades de DNA dos parentais em cada uma das misturas, embora não tenham permitido a observação das proporções 1:2 e 2:1, nas misturas 3 e 4, respectivamente (Figura 5).

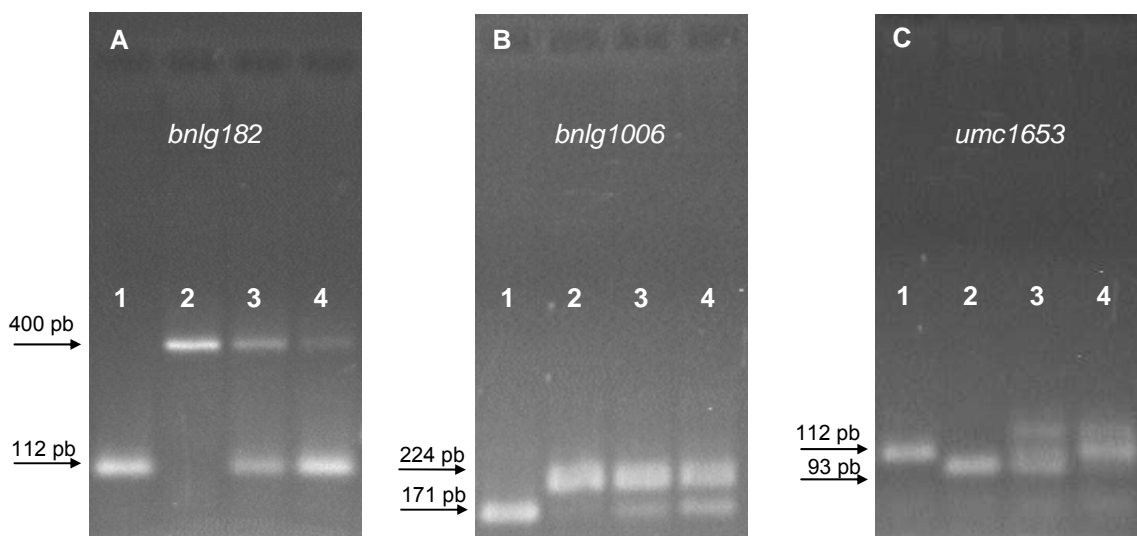


Figura 5 - Reação de amplificação com 30 ciclos completos para os *primers bnlG182* (gel A), *bnlg1006* (gel B) e *umc1653* (gel C). As amostras aplicadas no gel correspondem à reação de amplificação contendo: **1.** 30 ng do DNA da L3; **2.** 30 ng do DNA da L1113-01; **3.** Mistura 1(L3):2(L1113-01), 10 ng do DNA da L3 e 20 ng do DNA da L1113-01 e **4.** Mistura 2(L3):1(L1113-01), 20 ng de DNA L3 e 10 ng de DNA L1113-01.

A resolução dos fragmentos amplificados pelo *primer umc1653* não foi suficiente para realizar a análise de intensidade das bandas pelo programa ImageQuaNT[®]. Os valores de intensidade observados nas três repetições para os *primers bnlG182* e *bnlg1006* foram convertidos para um intervalo entre 0 e 100, sendo a intensidade das bandas das amostras parentais consideradas 100 e as demais, recalculadas por meio de regra de três simples. Os dados convertidos foram submetidos à análise de variância para o teste da hipótese $H_0: Y = \beta_0 + \beta_1 X$ (Tabela 3).

Os valores calculados de F a 5% foram não-significativos em três das análises de variância, indicando a não rejeição da hipótese H_0 . Apenas a

análise dos dados relativos à intensidade das bandas do menor alelo do *primer bnlg1006* apresentou um valor de F significativo. No entanto, a escolha do modelo linear foi mantida, uma vez que o efeito biológico do número de doses diferentes é linear.

Tabela 3 - Análise de variância dos valores de intensidade das bandas em cada amplificação, utilizando-se as três repetições. A – *Primer bnlg182*, alelo presente na linhagem L3 (112 pb). B – *Primer bnlg182*, alelo presente na linhagem L1113-01 (400 pb). C – *Primer bnlg1006*, alelo presente na linhagem L1113-01 (224 pb). D – *Primer bnlg1006*, alelo presente na linhagem L3 (171 pb).

A	FV	GL	SQ	QM	Probabilidade
Tratamento		2	1607,8567		
Regressão		1	1604,3028		
Desvio da regressão		1	3,5539	3,5539	21,1278%
Resíduo		6	10,8871	1,8145	
Total		8	1618,7438		R ² = 99,78%

B	FV	GL	SQ	QM	Probabilidade
Tratamento		2	132,8049		
Regressão		1	132,2706		
Desvio da regressão		1	0,5343	0,5343	100%
Resíduo		6	40,0933	6,6822	
Total		8	172,8982		R ² = 99,60%

C	FV	GL	SQ	QM	Probabilidade
Tratamento		2	207,1947		
Regressão		1	206,8086		
Desvio da regressão		1	0,3861	0,3861	100%
Resíduo		6	2,5034	0,4172	
Total		8	209,6981		R ² = 99,81%

D	FV	GL	SQ	QM	Probabilidade
Tratamento		2	1387,2975		
Regressão		1	1285,1942		
Desvio da regressão		1	102,1032	102,1032	1,2668%
Resíduo		6	49,7224	8,2871	
Total		8	1437,0199		R ² = 92,64%

A estimativa do coeficiente de determinação foi realizada pela razão entre as somas de quadrados da regressão e tratamento. Os valores observados variaram entre 92,64 e 99,81%, indicando uma resposta linear da intensidade da banda em função da quantidade de DNA, em nanogramas.

Os parâmetros da regressão foram calculados com o auxílio do Programa Genes (CRUZ, 2001) e também estabelecidos os intervalos de confiança, a 1% de probabilidade, para cada uma das variáveis independentes nas diferentes reações (Figuras 6 e 7).

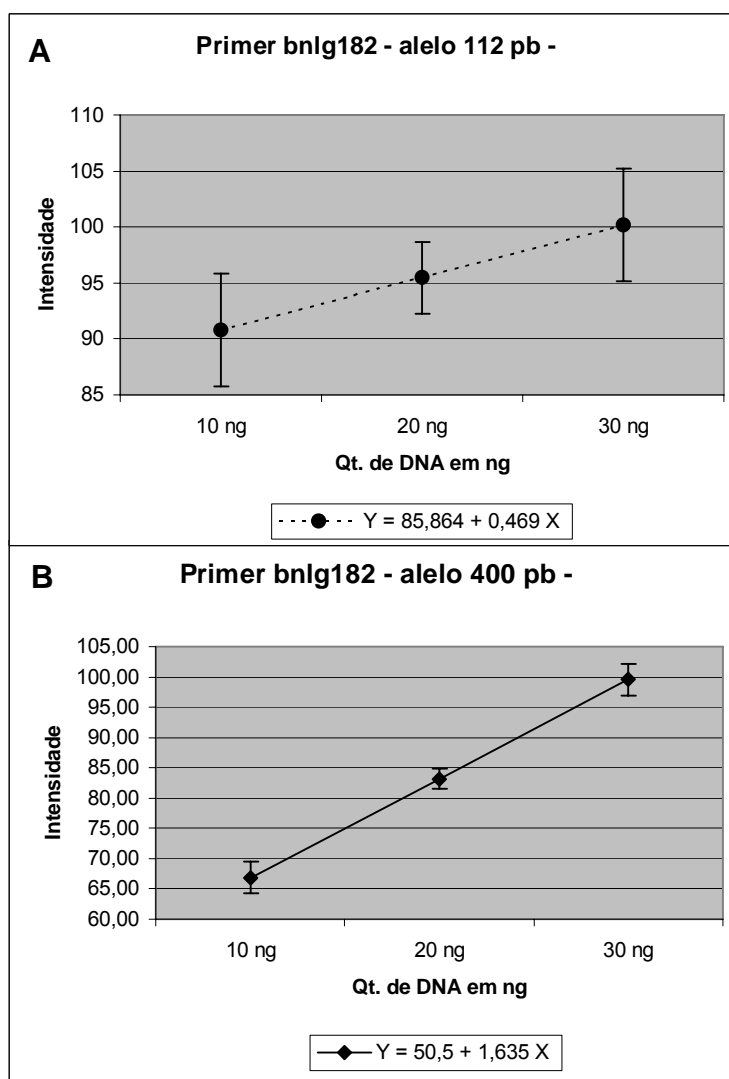


Figura 6 - Modelo de regressão linear da intensidade dos fragmentos amplificados pelo *primer bnlg182* vs quantidade de DNA em cada amostra. A – Intensidade da amplificação do alelo de 112 pb presente em L3. B – Intensidade da amplificação do alelo de 400 pb presente em L1113-01. A barra vertical indica o intervalo de confiança calculado para cada conjunto de dados, a 1% de probabilidade.

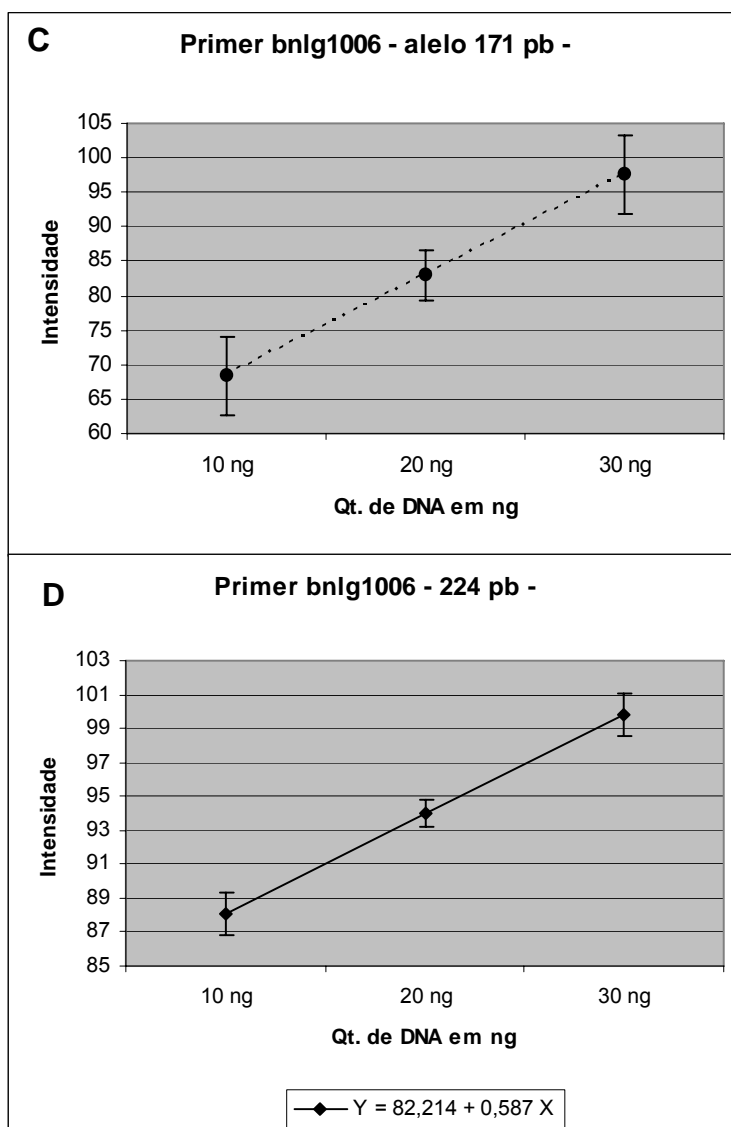


Figura 7 - Modelo de regressão linear da intensidade dos fragmentos amplificados pelo *primer bnlg1006* vs quantidade de DNA em cada amostra. C – Intensidade de amplificação do alelo de 171 pb presente em L3. D – Intensidade de amplificação do alelo de 224 pb presente em L1113-01. A barra vertical indica o intervalo de confiança calculado para cada conjunto de dados, a 1% de probabilidade.

No gráfico A da Figura 6, pode-se observar a sobreposição dos intervalos de confiança, mostrando que existem valores de intensidade para a banda correspondente ao alelo de tamanho menor (112 pb), que podem ter sido amplificados a partir de 10, 20 ou 30 ng de DNA. Análises realizadas com base em tal regressão não são informativas, pois as amostras que apresentam valores de intensidades pertencentes ao intervalo em sobreposição não podem ser classificadas quanto à dose de alelo. Dentre os dados observados, 67%

dos valores de intensidade das bandas menores que compõem as misturas encontram-se nessa região de sobreposição, dessa forma, mais da metade dos híbridos permaneceu não informativa, pois não foi possível determinar a quantidade de DNA originária de cada pai nas misturas, por meio da regressão.

No gráfico B da Figura 6, não se observou sobreposição entre os intervalos de confiança. No entanto, um dos valores de intensidade observados durante as análises não pertence a nenhum dos intervalos, sendo este um dado perdido. Entre os nove valores de intensidade que constituíram a análise de regressão, apenas um não foi explicado pelo modelo, ou seja, 11% dos dados foram perdidos.

Os gráficos da Figura 7 também não apresentaram sobreposições. A análise de regressão dos valores de intensidade das bandas correspondentes ao alelo de menor tamanho (171 pb) explicou todos os valores observados. No entanto, 11% das observações não puderam ser explicadas pela análise de regressão dos valores de intensidade das bandas do alelo de 224 pb, correspondendo, dessa forma, à mesma porcentagem de dados perdidos na análise da regressão do alelo maior do loco *bnlg182*.

4. DISCUSSÃO

Segundo Généthon (2001), o conhecimento da cinética da reação é necessário nesse tipo de estudo, principalmente para se conhecer a fase exponencial da reação de amplificação, em que o número de cópias de um fragmento é função da sua quantidade inicial.

Neste estudo, preocupou-se com a padronização da metodologia, com o intuito de diferenciar a quantidade de DNA parental nas misturas em diferentes proporções. Dessa forma, a determinação da intensidade dos fragmentos amplificados a partir do DNA parental e de suas misturas em diferentes proporções foi o ponto crítico para a determinação das variações quanto ao número de cópias de um loco no genoma, por meio da técnica de PCR-SQ. Para tal, foram padronizadas e mantidas constantes as condições eletroforéticas (tempo e voltagem), concentração de agarose, da solução-tampão e da solução de brometo de etídio. Os parâmetros para captura e análise das imagens dos géis foram padronizados para minimizar qualquer efeito que não fosse ao acaso. Condições semelhantes também foram utilizadas por Marin et al. (2002).

As diferenças quanto à intensidade das bandas no mesmo gel foram atribuídas ao número de produtos amplificados. Para o *primer bnlg1006* houve uma amplificação preferencial do alelo presente na linhagem L1113-01 (224 pb). Independentemente das diferentes concentrações iniciais, o alelo de maior tamanho apresentou maior intensidade em todas as amostras avaliadas, indicando maior quantidade de moléculas amplificadas. No entanto, os locos *bnlg182* e *umc1653* apresentaram ligeira preferência pela amplificação do alelo de menor tamanho.

Neste trabalho, observou-se que a cinética da reação foi específica para cada *primer* e, em especial, para cada alelo. Deve-se ressaltar ainda que a diferença de tamanho entre os alelos polimórficos interfere diretamente na cinética da reação, de forma que os *primers* que apresentam menores diferenças de tamanho entre os alelos amplificados são mais adequados para estudos de PCR-SQ, em que se pretende diferenciar as doses dos alelos.

O gel de agarose na concentração de 3,0% não apresentou resolução suficiente para avaliar diferença menor que 20 pb, e, ainda, a observação de bandas difusas foi constante, dificultando a delimitação da região para a quantificação precisa da sua intensidade, sobretudo nos fragmentos de menor tamanho. Ferreira e Grattapaglia (1996) recomendaram o uso de matriz de alta resolução para a visualização desses fragmentos, como géis de poliacrilamida corados com brometo de etídio ou prata. Nesta última coloração, um grande cuidado deve ser observado quanto à sua padronização.

As análises de regressão da intensidade dos fragmentos gerados após 30 ciclos de amplificação com os *primers bnlg182* e *bnlg1006*, utilizando-se misturas com diferentes proporções dos DNAs parentais, resolveram, em parte, a falta de informatividade dos indivíduos heterozigotos. Analisando os valores de intensidade dos fragmentos amplificados, pelo modelo de regressão linear foi possível concluir quanto às quantidades dos DNAs nas misturas em 17% das amostras, no *primer bnlg182*. Já no *primer bnlg1006* as análises foram conclusivas em 100% das misturas.

Assim, pode-se concluir que a metodologia de PCR-SQ, seguida da análise de regressão, pode ser aplicada no mapeamento de QTL com o objetivo de diminuir o número de indivíduos não informativos de populações exogâmicas. No entanto, ressalta-se que essa metodologia necessita ser previamente padronizada com o número de doses dos alelos conhecidos, limitando-a aos cruzamentos em que pelo menos o genótipo de um dos genitores é conhecido, como é o caso das famílias de meios-irmãos. Dessa forma, ainda é necessária a padronização de uma metodologia capaz de identificar o número de cópias de um alelo em descendentes de cruzamentos em que ambos os genitores são desconhecidos.

CAPÍTULO 3

Genotipagem do endosperma de milho utilizando marcadores SSR fluorescentes e PCR semiquantitativo

1. INTRODUÇÃO

Uma das aplicações mais importantes dos mapas genéticos é a localização de genes que controlam características de importância econômica, como produção de grãos, altura da planta e teor de proteína, entre outras, as quais resultam da ação cumulativa de um conjunto de genes. A idéia de se utilizar em marcadores genéticos para localizar locos que controlam a expressão de um caráter quantitativo (QTL) não é nova. Segundo Tanksley (1993), a análise dos dados utilizando marcadores é o procedimento mais simples para se detectar um QTL. A metodologia de análise baseada em um único loco consiste na comparação estatística do valor médio do caráter para as distintas classes genotípicas do marcador. Um resultado significativo denotaria a existência de ligação entre o marcador e um QTL do caráter em questão (COELHO, 2000).

A seleção da população para mapeamento envolve a escolha de genitores e a determinação do tipo de cruzamento, sendo considerada uma etapa crítica para o sucesso da construção do mapa (STAUB et al., 1996). No entanto, independentemente dessa escolha, duas condições devem ser atendidas: máximo de polimorfismo entre os genitores e que os locos estejam em desequilíbrio de ligação (PATERSON et al., 1991; TANKSLEY, 1993).

O desequilíbrio de ligação entre os locos segregantes é atribuído à ligação física entre os locos e originado do desequilíbrio gamético de ligação, decorrente, por sua vez, da redução na frequência de recombinação entre genes situados em regiões próximas ao longo de determinado cromossomo (COELHO, 2000). O desequilíbrio de ligação, decorrente da ligação física entre os locos, atinge seu ponto máximo nas populações derivadas de cruzamentos

controlados e, como conseqüência, a capacidade de detectar a ligação também é máxima (TANKSLEY, 1993; LIU, 1998).

Em populações exogâmicas, o conhecimento prévio da fase de ligação nem sempre é possível, pois parte das vezes se desconhece um dos genitores (populações de meios-irmãos) ou até mesmo ambos (populações naturais). A descendência pode ser também resultado de um intercruzamento em que ambos os genitores apresentam o mesmo marcador heterozigoto e somente a progênie homozigota é informativa. Em todos esses casos, a descendência heterozigota não possibilita, por meio de seu genótipo, distinguir a origem de seus alelos em relação aos dois genitores (LIU, 1998). A falta de informatividade dos indivíduos heterozigotos faz com que as avaliações destes sejam descartadas das análises de contraste entre médias, não contribuindo para a detecção do QTL. Assim, estratégias de identificação da origem dos alelos, se paternos ou maternos, são de grande interesse para se obter uma estimativa mais acurada na identificação de QTLs.

Uma possível forma de incluí-los nas análises é por meio da genotipagem do seu endosperma, em que se espera que um indivíduo heterozigoto, descendente do cruzamento $M_1M_2 \times M_1M_2$, apresentasse um endosperma $M_1M_1M_2$ ou $M_1M_2M_2$ quando recebeu do genitor feminino o alelo M_1 ou M_2 , respectivamente. Dessa forma, o objetivo deste estudo foi desenvolver uma metodologia de genotipagem de endosperma de milho, para utilizá-la na identificação da origem dos alelos em indivíduos heterozigotos até então não-informativos.

2. MATERIAL E MÉTODOS

Material vegetal

Foram utilizadas folhas e sementes das linhagens endogâmicas de milho L3 e L1113-01 cedidas pelo Programa de Melhoramento de Milho da Embrapa Milho e Sorgo, bem como sementes originadas do cruzamento recíproco e de seu entre essas linhagens. Os procedimentos moleculares foram realizados no Laboratório de Biologia Molecular do Núcleo de Biologia Aplicada da Embrapa Milho e Sorgo, de Sete Lagoas, MG.

Sementes oriundas do cruzamento entre as linhagens L3 e L1113-01, cujo genitor feminino foi a linhagem L3, foram nomeadas como híbrido L3 X L1113-01, enquanto as sementes do cruzamento recíproco, cujo genitor feminino foi L1113-01, receberam a denominação de híbrido L1113-01 X L3.

Extração de DNA do endosperma

O DNA do endosperma foi extraído como descrito por Leiva (2003). Após a remoção do pericarpo com o auxílio de uma lâmina de barbear, aproximadamente 100 mg de endosperma, cerca de 1/3 do peso da semente, foram moídos até a obtenção de um pó fino. A quantidade de endosperma removido foi determinada de acordo com o rendimento das diferentes quantidades extraídas e com os resultados do teste de germinação de sementes realizado no Capítulo 1, onde a remoção de até 50% do endosperma em peso não comprometeu a viabilidade da semente.

Ao endosperma moído foram adicionados 600 μ L de tampão de extração (Sarcosil 1%; Tris-HCl 100 mM, pH 8,0; NaCl 100 mM; EDTA 200 mM, pH 8,0) sob agitação no vórtex por 20 segundos, seguido de homogeneização à temperatura ambiente por 15 min. Após esse período, uma lavagem com igual volume de fenol e de uma mistura de clorofórmio-octanol na proporção (24:1) foi realizada por 20 min. O material foi centrifugado a 15.000 xg e o sobrenadante transferido para outro tubo, onde foi realizada uma nova lavagem com clorofórmio-octanol (24:1), por 5 min. Após a centrifugação a 15.000 xg por 10 min, o sobrenadante foi removido para um novo tubo e tratado com 5 μ L de RNase A (10 mg/mL) a 37 °C, por uma hora. Em seguida, o DNA foi precipitado com 450 μ L isopropanol a -80 °C por 30 min e centrifugado a 15.000 xg por 10 min. O sedimento foi lavado com etanol 70% por 5 min e secado à temperatura ambiente por 24 h.

A quantificação do DNA foi realizada em gel de agarose 0,8%, utilizando-se um padrão de DNA de fago lâmbda com concentração conhecida. Após a eletroforese realizada a 60 V durante 30 min, o gel foi tratado com brometo de etídeo (1 mg/mL) e visualizado sob luz ultravioleta, no Eagle Eye II (Stratagene).

Extração de DNA foliar

Devido à pouca quantidade de DNA extraída do endosperma, inicialmente optou-se pela utilização do DNA genômico foliar na etapa de otimização da técnica de PCR semiquantitativo. Dessa forma, extraiu-se o DNA das linhagens parentais, segundo o método descrito por Saghai-Marroof et al., 1984.

A quantificação e as misturas do DNA das linhagens parentais foram realizadas conforme descrito no Capítulo 2.

Otimização da metodologia de PCR semiquantitativo com marcadores SSR fluorescentes em poliacrilamida a partir das misturas de DNA foliar

Foram utilizados nove *primers* SSRs marcados com as fluorescências: azul (6-FAM), verde (HEX) ou amarelo (NED), na extremidade 5' do *primer forward*. Os *primers* foram previamente selecionados no banco de dados da Embrapa Milho e Sorgo como sendo polimórficos entre as linhagens parentais (PADILHA, 2002). A Tabela 1 descreve as características dos *primers* SSR selecionados.

O material inicialmente amplificado foram as misturas de DNA foliar dos parentais apresentadas no Capítulo 2.

Tabela 1 - *Primers* SSR polimórficos selecionados. Bin corresponde a localização no genoma e as seqüências, aos núcleos de repetição específicos para cada *primer*. Os valores em pb correspondem ao tamanho dos alelos presentes em cada uma das linhagens parentais (L3 e L1113-01), e na coluna à direita estão os valores de fluorescência correspondentes

Primer	Bin	Seqüência	L1113-01 (pb)	L3 (pb)	Fluorescência
<i>umc2025</i>	1.05	(AGCT) ₄	130	123	FAM
<i>bnlg1208</i>	5.03	(AG) ₁₀	104	97	FAM
<i>mmc041</i>	1.08	(CA) ₁₂	164	174	HEX
<i>dup024</i>	2.08	(GA) ₁₆	106	118	NED
<i>umc1016</i>	7.02	(CT) ₂₅	116	102	FAM
<i>umc1653</i>	6.07	(GAAA) ₂₄	93	112	HEX
<i>phi053</i>	3.05	(ATAC)	195	169	NED
<i>bnlg1006</i>	5.00	(AG) ₂₀	224	171	HEX
<i>bnlg182</i>	1.03	?	400	112	FAM

As reações de amplificação foram realizadas isoladamente em cada fluorescência, seguindo-se os parâmetros descritos por Padilha (2002). Foram utilizados 30 ciclos de amplificação, conforme previamente estabelecido para as condições em gel de agarose no Capítulo 2.

Os produtos finais da reação de PCR de cada uma das três fluorescências foram diluídos, separadamente, nas proporções de 1:50, 1:20 e 1:15, respectivamente para 6-FAM, HEX e NED. Um volume de 3 μ L da diluição foi misturado com 1,0 μ L de formamida e 0,4 μ L do corante *blue dextran*. As amostras foram desnaturadas a 95 °C durante 3 min e mantidas no gelo até à sua aplicação no gel desnaturante de poliacrilamida 5%. Foi aplicado 1,5 μ L da amostra desnaturada, e os produtos amplificados foram separados por eletroforese a 2000 V durante 3 h, no seqüenciador ABI Prism 377 (Applied Biosystems).

A intensidade dos fragmentos amplificados foi analisada pelo programa Genescan 2.1 (Applied Biosystems) e quantificada em função da fluorescência emitida em cada pico dos eletroferogramas. Os *primers* que apresentaram o padrão de amplificação com bandas *stutter* foram desconsiderados devido às dificuldades de mensuração na intensidade dos picos.

Genotipagem de endosperma pela metodologia de PCR semiquantitativo

A reação de amplificação do DNA de endosperma foi realizada, assim como a amplificação do DNA foliar no Capítulo 2, exceto a quantidade de DNA, que foi reduzida para 20 ng de DNA endospermático.

Para avaliar a repetibilidade dos valores de intensidade encontrados, foram calculadas as razões de intensidade de fluorescência entre os picos correspondentes ao alelo recebido do genitor L3 e L1113-01. Em cada híbrido L3 x L1113-01 e L1113-01 x L3, extraiu-se DNA de três sementes diferentes que constituíram as réplicas biológicas, e foram realizadas três reações de amplificação para cada amostra (réplicas experimentais). Assim, foram feitas 18 reações para cada par de *primers*, as quais foram analisadas segundo o delineamento inteiramente casualizado com três tratamentos, que constituíram as amostras de um mesmo genótipo e três repetições. O produto da amplificação do DNA dos parentais foi utilizado como controle positivo ou negativo durante as análises.

3. RESULTADOS

Extração do DNA de endosperma

A adaptação do protocolo descrito por Leiva (2003) para o isolamento do DNA de endosperma permitiu a extração de DNA das linhagens e dos híbridos suficientes para a realização das análises. Após extrações em diferentes quantidades de endosperma e respeitando a quantidade mínima estabelecida no Capítulo 1, para a não-destruição da viabilidade da semente observou-se maior rendimento de DNA nas amostras de 100 mg de endosperma, o que correspondeu a um terço do peso total. A eficiência de extração foi de 71,15%. Embora algumas extrações apresentassem precipitado após a lavagem com etanol, a quantidade de DNA não foi suficiente para ser quantificada em gel de agarose 0,8% corado com brometo de etídeo.

O rendimento de DNA por semente foi, em média, de 250 ng, o que permitiria realizar aproximadamente 10 reações de amplificação. Assim, optou-se por realizar as otimizações das reações de PCR semiquantitativo utilizando as mesmas misturas de DNA foliar das linhagens mencionadas no Capítulo 2.

Otimização da metodologia de PCR semiquantitativo com marcadores SSRs fluorescentes em poliacrilamida a partir das misturas de DNA foliar

A análise dos produtos do PCR semiquantitativo indicou que tanto as condições de reação previamente estabelecidas quanto o número de ciclos igual a 30 se aplicam às reações com marcadores fluorescentes e eletroforese em poliacrilamida, permitindo a observação das diferentes proporções de DNA parental nas misturas.

O cálculo da intensidade das bandas, resultantes da reação de amplificação do *primer umc1653*, realizado pelo programa Genescan 2.1, permitiu diferenciar a mistura 1:2 da mistura 2:1. A razão entre os valores de intensidade de fluorescência (RIF) dos produtos de amplificação da mistura 1:2 se aproximou a 2,16, indicando que o número inicial de alelos da linhagem L1113-01 é o dobro em relação ao número de alelos da linhagem L3, nessa mistura. Enquanto a RIF na análise da mistura 2:1 se aproximou de 0,53, indicando que o número inicial de alelos da linhagem L1113-01 é a metade em relação ao número de alelos da linhagem L3.

Os valores de intensidade observados tanto nas linhagens parentais quanto nas misturas foram plotados em eletroferogramas representados na Figura 1.

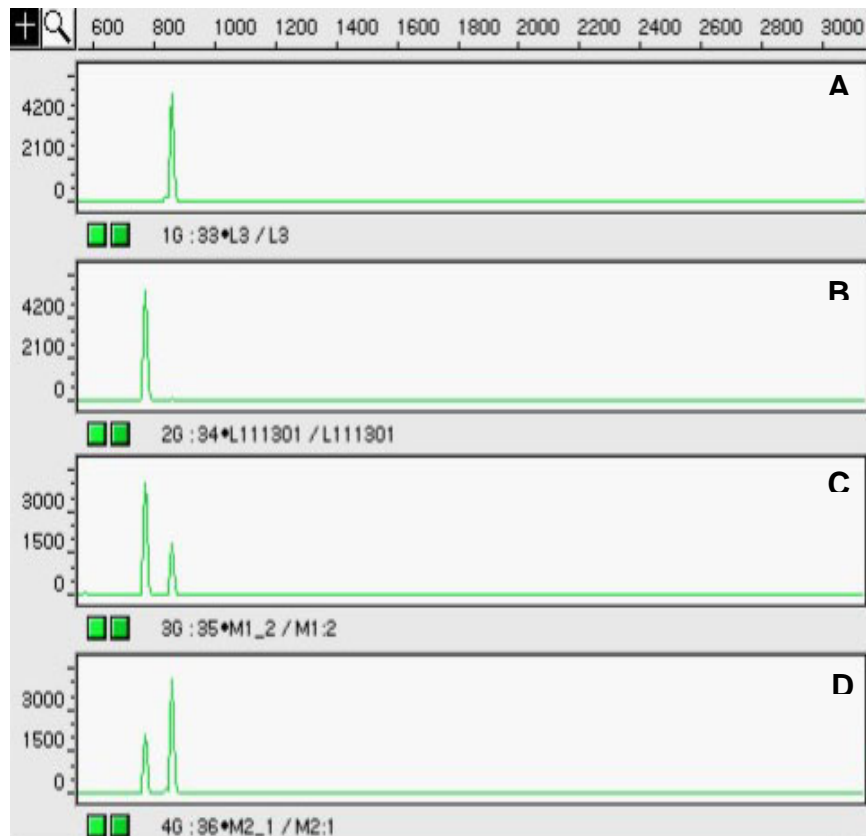


Figura 1 - Eletroferograma dos valores de intensidade observados nos produtos de amplificação do *primer umc1653*: A – produto da reação de amplificação realizada com DNA da linhagem L3. B – produto da reação de amplificação realizada com DNA da linhagem L1113-01. C – produtos da reação de amplificação realizada com a mistura 1:2 (10 ng do DNA L3 e 20 ng do DNA L1113-01). D – produtos da reação de amplificação realizada com a mistura 2:1 (20 ng do DNA L3 e 10 ng do DNA L1113-01). O eixo horizontal corresponde ao tamanho em pares de bases do fragmento detectado, enquanto no eixo vertical estão os valores de intensidade dos picos.

As reações de PCR-SQ realizadas com os outros *primers* apresentaram comportamento semelhante. Pôde-se observar, ainda, que as razões de intensidade de fluorescência se aproximam mais dos valores de 0,5 ou 2,0 quando as amostras aplicadas no gel foram mais diluídas (1:50).

Um total de nove *primers* foi analisado, sendo que três foram descartados por apresentarem padrão de bandas *stutter*, o que dificultou a determinação da intensidade das bandas. Dentre os seis *primers* que não apresentaram bandas *stutter*, apenas os *primers bnlg1208* e *umc2025* foram os

escolhidos para o estudo da amplificação do DNA endospermático, devido à limitação das quantidades de DNA e da necessidade de repetições.

Genotipagem de endosperma utilizando PCR semiquantitativo

Os produtos de amplificação dos *primers bnlg1208* e *umc2025* foram plotados em eletroferogramas (Figuras 2 e 3).

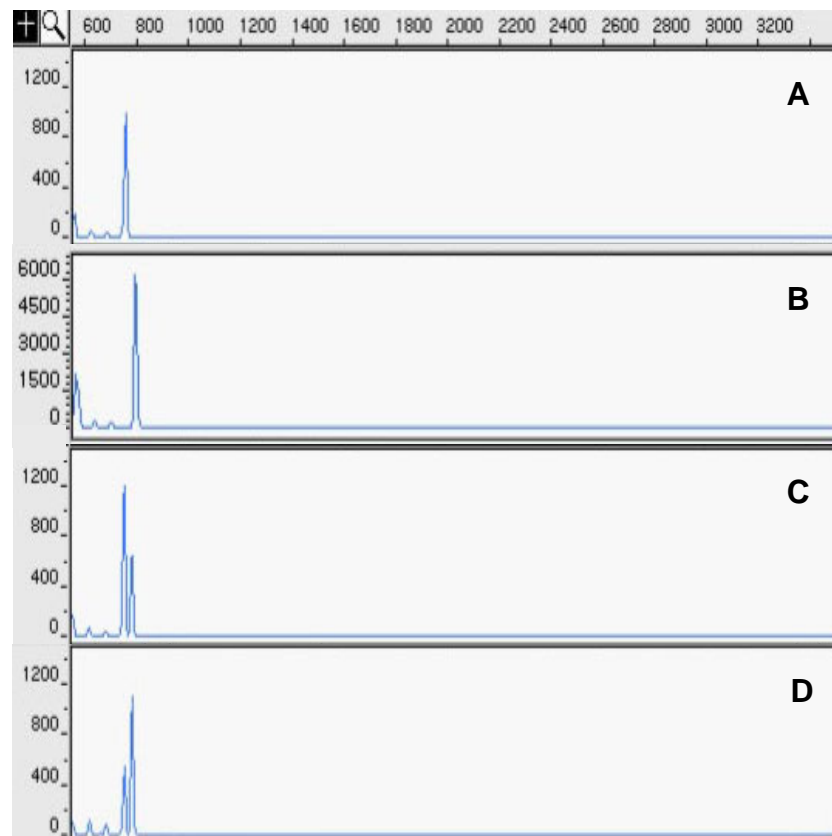


Figura 2 - Eletroferograma dos valores de intensidade observados nos produtos de amplificação do *primer bnlg1208*: A – produto da reação de amplificação realizada com DNA do genitor L3. B – produto da reação de amplificação realizada com DNA do genitor L1113-01. C – produtos da reação de amplificação realizada com DNA do híbrido (L3 x L1113-01). D – produtos da reação de amplificação realizada com DNA do híbrido (L1113-01 x L3). O eixo horizontal corresponde ao tamanho em pares de bases do fragmento detectado, enquanto no eixo vertical estão os valores de intensidade dos picos.

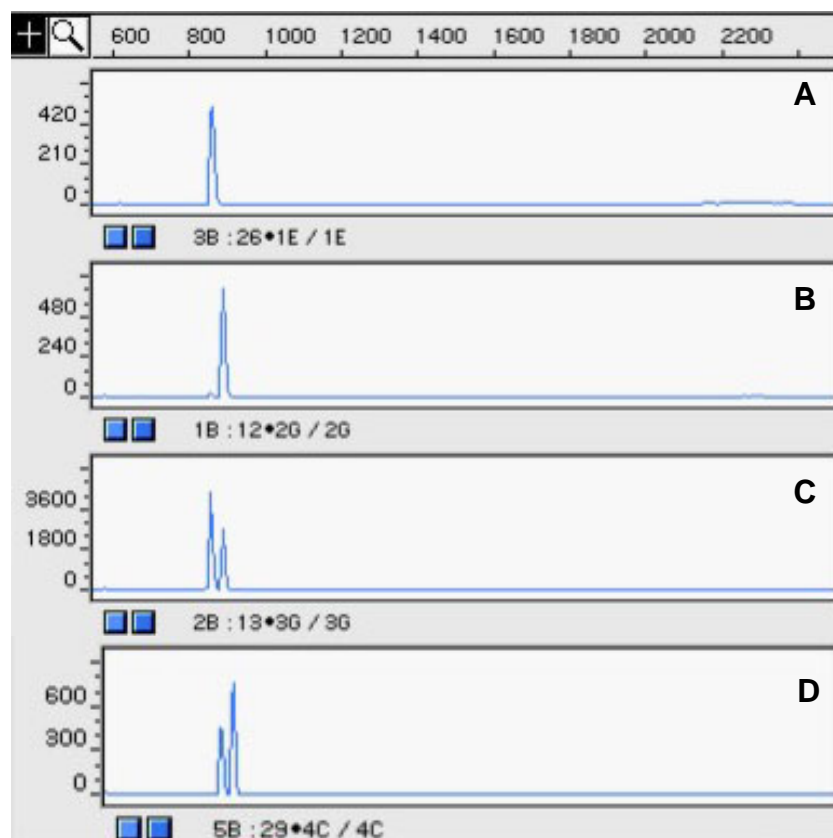


Figura - Eletroferograma dos valores de intensidade observados nos produtos de amplificação do *primer umc2025*: A – produto da reação de amplificação realizada com DNA do genitor L3. B – produto da reação de amplificação realizada com DNA do genitor L1113-01. C – produtos da reação de amplificação realizada com DNA do híbrido L3 x L1113-01. D – produtos da reação de amplificação realizada com DNA do híbrido L1113-01 x L3. O eixo horizontal corresponde ao tamanho em pares de bases do fragmento detectado, enquanto no eixo vertical estão os valores de intensidade dos picos.

A partir da análise de intensidade dos picos pelo programa Genescan 2.1, foram calculadas as razões de intensidade de fluorescência (RIF) entre os valores de intensidade das bandas do alelo recebido do genitor L3 e L1113-01 (Tabela 2).

Tabela 2 - Razão de intensidade de fluorescência entre os valores de intensidade das bandas correspondentes aos alelos doados pelo genitor L3 e L1113-01, amplificados pelos *primers bnlg1208* e *umc2025*. Os tratamentos A e C representam repetições biológicas do híbrido cujo genitor materno é L3, enquanto os tratamentos B e D são repetições biológicas do híbrido cujo genitor materno é L1113-01. R1, R2 e R3 são repetições experimentais

Primer <i>bnlg1208</i>				
Híbrido L3 x L1113-01		R1	R2	R3
Tratamento A	A1	2,0570	1,8407	1,6222
	A2	1,7460	1,6227	1,9388
	A3	1,9633	1,6746	1,6920
Híbrido L1113-01 x L3				
Tratamento B	B1	0,5072	0,5498	0,5010
	B2	0,5420	0,6776	0,4821
	B3	0,6151	0,6100	0,4986
Primer <i>umc2025</i>				
Híbrido L3 x L1113-01		R1	R2	R3
Tratamento C	C1	1,6601	1,9566	1,7852
	C2	1,4786	1,5518	2,0393
	C3	1,5600	2,2060	1,6456
Híbrido L1113-01 x L3				
Tratamento D	D1	0,5931	0,4923	0,5960
	D2	0,4485	0,4837	0,5561
	D3	0,5223	0,5272	0,5593

Por meio da análise de variância dos dados obtidos pela razão de intensidade de fluorescência referente aos híbridos L3 x L1113-01 e L1113-01 x L3, sob a hipótese $H_0: m_1 = m_2 = m_3$, verificou-se que não existiram diferenças significativas a 5% de probabilidade, entre as médias das réplicas biológicas, em nenhum dos tratamentos (Tabela 3).

Tabela 3 - Resumo da análise de variância dos dados obtidos pela razão de intensidade de fluorescência sob a hipótese $H_0: m_1 = m_2 = m_3$, a 5% de probabilidade. CV (%) é o coeficiente de variação de cada conjunto de dados

Primer bnlg 1208					
FV	GL	SQ	QM	F	CV (%)
Tratamento A	2	0,0091	0,0045	100%	10,1116
Resíduo	6	0,1977	0,0330	NRH ₀	
Total	8	0,2068			

FV	GL	SQ	QM	F	CV (%)
Tratamento B	2	0,0054	0,0027	100%	12,8015
Resíduo	6	0,0301	0,0050	NRH ₀	
Total	8	0,0355			

Primer umc 2025					
FV	GL	SQ	QM	F	CV (%)
Tratamento C	2	0,0252	0,0126	100%	15,9669
Resíduo	6	0,4764	0,0794	NRH ₀	
Total	8	0,5017			

FV	GL	SQ	QM	F	CV (%)
Tratamento D	2	0,0063	0,0032	32,1858%	9,0343
Resíduo	6	0,0138	0,0023	NRH ₀	
Total	8	0,0201			

De forma semelhante, por meio da análise de variância dos dados obtidos pela razão de intensidade de fluorescência verificou-se que não houve diferenças significativas entre as médias das repetições experimentais. Em nenhum dos tratamentos os valores de probabilidade calculados foram inferiores a 5% de probabilidade.

Uma vez não rejeitadas as hipóteses de nulidade das análises de variância, considerou-se nas análises posteriores que, para cada tratamento, existiram nove repetições.

Para a diferenciação quanto às razões entre os valores de intensidade das bandas nos híbridos L3 x L1113-01 e L1113-01 x L3, realizou-se o teste t a 5% de probabilidade, para análise do contraste entre as médias dos

tratamentos A e B, C e D, que correspondem às ampliações pelos *primers* *bnlg1208* e *umc2025*.

Inicialmente, as variâncias dos dois conjuntos de dados para cada *primer* foram comparadas pelo teste $F_{\text{máx}}$, de Hartley (HARTLEY, 1950) (Tabela 4).

Tabela 4 - Teste $F_{\text{máx}}$, de Hartley, a 5% de probabilidade

Primer <i>bnlg1208</i>		
$H_0: \sigma^2_A = \sigma^2_B$ e $H_a: \sigma^2_A > \sigma^2_B$		
σ^2		σ^2_A / σ^2_B
σ^2_A	0,0259	2,4228%
σ^2_B	0,0058	RH ₀
Primer <i>umc2025</i>		
$H_0: \sigma^2_C = \sigma^2_D$ e $H_a: \sigma^2_C > \sigma^2_D$		
σ^2		σ^2_C / σ^2_D
σ^2_C	0,0627	7,0818. 10 ⁻³ %
σ^2_D	0,0025	RH ₀

Rejeitadas as hipóteses de nulidade, o valor de t foi calculado por meio da equação 1, que segue aproximadamente a distribuição *t*, de *Student* com n^* graus de liberdade, em que n^* pode ser calculado por meio da equação 2.

$$(1) \quad t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} \quad e$$

$$(2) \quad n^* = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{n_x - 1} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{n_y - 1}}$$

A hipótese de nulidade $H_0: m_A = m_B$ foi rejeitada, com a probabilidade de apenas $7,5868 \cdot 10^{-7}$ e 11 graus de liberdade, indicando a existência de uma diferença significativa entre os conjuntos de médias que compuseram os tratamentos. Assim, a partir das amplificações com o *primer bnlg1208* foi possível diferenciar os dois grupos de híbridos do cruzamento entre L3 e L1113-01, quanto ao valor da razão de intensidade de fluorescência que reflete o número de cópias de cada alelo correspondente no genoma do endosperma.

Semelhantemente, a hipótese de nulidade $H_0: m_C = m_D$ foi rejeitada com probabilidade de $8,0417 \cdot 10^{-5}$ com nove graus de liberdade, indicando a existência de uma diferença significativa entre os conjuntos de médias que compuseram cada tratamento. Assim como as amplificações com o *primer bnlg1208*, as reações com o *primer umc2025* permitiram diferenciar os dois grupos de híbridos do cruzamento entre L3 e L1113-01, quanto ao valor da razão de intensidade de fluorescência das bandas correspondentes aos alelos doados pelo genitor L3 e L1113-01, que reflete o número de cópias de cada alelo correspondente no genoma do endosperma.

Em seguida, foi estabelecido um intervalo de confiança, a 1% de probabilidade, para cada conjunto de dados (Tabela 5).

Tabela 5 - Intervalo de confiança dos valores da razão de intensidade de fluorescência (RIF), a 1% de probabilidade e oito graus de liberdade, para cada tratamento

Tratamento	Intervalo de Confiança
A	$1,615 \leq \text{RIF} \leq 1,975$
B	$0,466 \leq \text{RIF} \leq 0,636$
C	$1,484 \leq \text{RIF} \leq 2,045$
D	$0,475 \leq \text{RIF} \leq 0,587$

4. DISCUSSÃO

Extração do DNA de endosperma e foliar

As diferenças observadas nas concentrações de DNA extraído do tecido foliar e endospermático podem ser explicadas por variações específicas de cada tecido, como o número e tamanho das células, a taxa de divisão celular e a quantidade de DNA extracelular, ou, mesmo, por diferenças na composição química e estrutural desses tecidos (RAMOS et al., 2006). A baixa eficiência de extração do DNA de endosperma, quando comparada com DNA foliar, pode ser atribuída à perda de DNA pelo tecido endospermático durante a fase de acúmulo de reservas nas sementes. Quando o amido é formado, ocorrem rupturas em organelas e o conteúdo celular é modificado, com redução expressiva na quantidade de ácidos nucleicos, culminando com a morte da célula e total preenchimento pelas reservas de amido (McDONALD et al., 1995).

As sementes são estruturas de multiplicação vegetal, com tecido de reserva rico em carboidratos, lipídios e proteínas, responsáveis pelos suprimentos do embrião durante o processo de germinação (CARVALHO; NAKAGAWA, 2000). Essas substâncias de reserva dificultam o isolamento do DNA, exigindo maior número de lavagens com solvente orgânico (clorofórmio-álcool isolamílico), durante o protocolo de extração. Esse procedimento implica, também, redução na quantidade total de DNA isolado, pois parte dessas moléculas não se separa de outros compostos orgânicos e é, conseqüentemente, descartada (RAMOS et al., 2006).

O baixo rendimento nas extrações realizadas limitou o estudo. Dessa forma, optou-se por utilizar misturas em diferentes proporções do DNA foliar, na tentativa de simular a ocorrência no endosperma e validar a metodologia e as condições de reações preestabelecidas em eletroforese em agarose. A utilização das misturas em diferentes proporções foi uma alternativa viável, uma vez que refletiu bem a situação real observada no endosperma, no qual a quantidade do DNA de origem materna é o dobro da quantidade do DNA paterno.

Otimização da metodologia de PCR semiquantitativo com marcadores SSRs fluorescentes em poliacrilamida e genotipagem do endosperma

Os parâmetros da metodologia de PCR semiquantitativo anteriormente determinados se aplicaram a PCR-SQ com *primers* fluorescentes e eletroforese em gel de poliacrilamida, pois as análises das misturas de DNA dos parentais L3 e L1113-01 permitiram diferenciar a quantidade inicial do DNA parental em cada uma das misturas.

A eficiência da reação de amplificação do DNA de endosperma foi apenas de 29,37%, possivelmente a qualidade das extrações foi prejudicada devido à quantidade elevada de impurezas associadas ao DNA isolado a partir de sementes, o que pode ter impedido o acesso da Taq polimerase ao DNA-molde durante a reação. Tal fato também explica a necessidade de aplicação, no gel de amostras, das reações com DNA endospermático menos diluídas para melhor observação, pois, mesmo que a reação tenha ocorrido, ela pode ter sido prejudicada pela menor disponibilidade de DNA-molde a ser amplificado (RAMOS et al., 2006).

A utilização da razão entre os valores de intensidade de fluorescência dos alelos amplificados mostrou ser uma boa alternativa diante da discrepância dos valores observados de intensidade nas diferentes repetições de cada amostra. Utilizando os *primers bnlg1208* e *umc2025*, os valores das razões observadas puderam ser agrupados em apenas dois grupos, os híbridos do cruzamento L3 x L1113-01 e os do cruzamento L1113-01 x L3, uma vez que a

análise de variância não mostrou diferenças significativas entre as médias que compunham cada grupo.

O contraste entre as médias das razões dos dois grupos de híbridos mostrou-se significativo. Dessa forma, foi possível distingui-los a 5% de probabilidade quanto à razão entre os valores de intensidade dos alelos derivados dos dois parentais. A utilização de marcadores microssatélites fluorescentes e a adaptação da metodologia de PCR semiquantitativo foram eficientes em discriminar a origem do endosperma.

A determinação do intervalo de confiança no estudo de cada *primer*, em particular para cada híbrido, permitiu agrupar um indivíduo quanto à razão dos valores de intensidade dos alelos amplificados. Assim, foi possível discriminar a origem destes. Valores da razão de intensidade de fluorescência que não foram contemplados nos respectivos intervalos podem ser considerados dados perdidos. Dessa forma, dos 36 valores de razão de intensidade de fluorescência deste trabalho, 33% seriam perdidos se não houvesse repetições experimentais.

O estabelecimento da origem de cada alelo faz que indivíduos heterozigotos, até então não informativos e por isso não computados nas análises de contraste para a determinação de um QTL, possam ter suas análises consideradas, permitindo maior acurácia no mapeamento de populações exogâmicas.

Ressalta-se que a metodologia de PCR semiquantitativo, aqui apresentada, tem como vantagem o baixo custo, quando comparada com a metodologia de PCR em tempo real. Ainda, não necessita de uma amostra-padrão de DNA com quantidades conhecidas de cada alelo, podendo ser aplicada a qualquer tipo de população, da qual seja possível extrair DNA de endosperma suficiente para as análises, diferentemente da metodologia descrita no capítulo anterior.

3. CONCLUSÕES GERAIS

A genotipagem de endosperma utilizando marcadores SSRs e PCR semiquantitativos é uma alternativa eficiente para determinar a origem alélica nos descendentes heterozigotos.

A retirada de parte do endosperma e pericarpo não afetou a germinação do embrião. Até 50% da massa do endosperma pode ser removida sem prejuízo da germinação, o que correspondeu, em média, a 175 mg.

As baixas eficiências da extração de DNA endospermático e de amplificação representam uma limitação à utilização da genotipagem de endosperma como ferramenta para diminuir o número de descendentes não informativos. Dessa forma, recomenda-se um trabalho de otimização desses processos para aumento da eficiência.

A metodologia de PCR semiquantitativo em gel de agarose, seguida da análise de regressão, pode ser utilizada como ferramenta no mapeamento de QTL, com o objetivo de diminuir o número de indivíduos não informativos em populações exogâmicas onde, pelo menos, o genótipo de um dos genitores é conhecido, como é o caso das famílias de meios-irmãos.

A metodologia de PCR semiquantitativo com marcadores SSRs fluorescentes possibilitou distinguir dois grupos de híbridos do cruzamento entre L3 e L1113-01, quanto ao número de cópias dos diferentes alelos no endosperma, identificando a origem desses, independentemente do conhecimento prévio do genótipo de um dos genitores, podendo ser aplicada a qualquer tipo de população e, conseqüentemente, aumentando a acurácia na detecção de QTLs.

Ambas as metodologias aqui apresentadas têm como vantagem o baixo custo, em comparação com a metodologia de PCR em tempo real.

4. REFERÊNCIAS

AUSTIN, D. F.; LEE, M. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. **Crop Science**, v. 38, p. 1296-1308, 1998.

BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J.B.; CARNEIRO, I.F. (Eds.). **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE, 2000. p. 63-223.

BOSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal Human Genetics**, v. 32, n. 3, p. 314-331, 1980.

BRASIL, Ministério da Agricultura e Reforma Agrária. **Regras de Análise de Sementes**. Brasília, 1992. 365 p.

CARNEIRO, S. M.; VIEIRA, M. L. C. Mapas genéticos em plantas. **Bragantia**, v. 61, n. 2, p. 89-100, 2002.

CARVALHO, N.M.; NAKAGAWA, J. **Sementes: ciência, tecnologia e produção**. 4. ed. Jaboticabal; SP: FUNEP, 2000. 588 p.

COELHO, A.S.G. Considerações gerais sobre a análise de QTL's. In: PINHEIRO, J.B.; CARNEIRO, I.F. (Eds.). **Análise de QTL no melhoramento de plantas**. Goiânia: FUNAPE 2000. p. 1-36.

CRUZ, C. D.; **Programa Genes** – Versão Windows: aplicativo computacional em genética e estatística. Viçosa, MG: Editora UFV, 2001. 648 p.

Da SILVA, W. R.; CHAMMA, H. M. C. P.; NOVEMBRE, A. D. Da L.C.; De MORAES, M. H. D. Avaliação do desempenho de sementes de milho submetidas à remoção parcial do endosperma e do pericarpo. **Sci. Agric.** Piracicaba, v. 51, n.1, p. 138-152, 1994.

Da, Y.; LEWIN, H. A. Linkage information content and efficiency of full-sib and half-sib designs for gene mapping. **Theoretical and Applied Genetics**, v. 90, p. 699-706, 1995.

EDWARDS, M.D.; STUBER, C.W.; WENDEL, J.F. Molecular-marker-facilitated investigations of quantitative trait loci in maize. Numbers, genomic distribution and types of gene action. **Genetics**, v. 116, p. 13-125, 1987.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. Brasília: EMBRAPA/CENARGEM, 1996. 220 p.

GALINAT, W.C. Botany and origin of maize. In: **Maize**, New York, v. 105, n. 6, p. 12, 1979.

GARDNER, E. J.; SNUSTAD, P. **Genética**. Rio de Janeiro: Interamericana, 1986. 497 p.

GARVIN, D.F.; WEEDEN, N.F. Genetic-linkage between isozyme, morphological, and DNA markers in tepary bean. **Journal of Heredity**, v. 85, n. 4, p. 273-278, 1994.

GELDERMANN, H. Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. **Theoretical and Applied Genetics**, v. 46, p. 319-330, 1975.

GÉNÉTHON, La PCR quantitative en temps reels ou la "Taqman". 2001. Disponible em: <<http://www.genethon.fr>>. Acesso em: 11 nov 2005.

GOMES, J.; KARAZAWA, M. Como a planta de milho se desenvolve. In: **O milho no Paraná**. Curitiba: Fundação Instituto Agrônômico do Paraná, 1982. p. 35-50.

GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-tescross mapping strategy and RAPD markers. **Genetics**, v. 137, p. 1121-1137, 1994.

GRIFFITHS, A.J.F.; MILLER, J.H.; SUZUKI, D.T.; LEWONTIN, R.C.; GELBART, W.M. **Introdução à genética**. Rio de Janeiro: Guanabara, 1998. 856 p.

GUO, X.; ELSTON, R. Linkage information content of polymorphic genetics markers. **Hum Hered**, v. 49, p. 112-118, 1999.

HARTLEY, H. O. The use of rang in analysis of variance. **Biometrika**, v. 37, p. 271-280, 1950.

JANSEN, R. C. Interval mapping of multiple quantitative trait loci. **Genetics**, v. 135, p. 205-211, 1993.

LANDER, E. S.; BOTSTEIN, D. Mapping mendelianfactors underlying quantitative traits using RFLP linkage maps. **Genetics**, v. 121, n. 1, p. 185-199, 1989.

LANDER, E.S.; WEINBERG, R.A. Genomics: journey to the century of biology. **Science**, v. 287, n. 5459, p. 1777-1782, 2000.

LEE, M. DNA markers and plant breeding programs. **Advances in Agronomy**, v. 55, p. 265-344, 1995.

LEIVA, J. DNA isolation from mature endosperm, 2003. Disponível em: <<http://ag.arizona.edu/research/larkinslab/>>.

LIU, B.H. **Statistical genomics**. New York: CRC, 1998. 610 p.

LYNCH & WALSH. **Genetics and analysis of quantitative traits**. 1º ed. Sunderland, MA: Ed. Sinauer Associates, Inc., 1998. 980 p.

Maize Genetics and Genomics Database. Disponível em : <<http://www.maizegdb.org/ssr.php>>. Acesso em : 20 agosto 2006.

MARIN, L.J.; CUNHA, A. A.; AQUINO, V. H.; FIGUEIREDO, L. T. M. Desenvolvimento de uma metodologia de PCR semiquantitativa utilizando plasmídeo clonado com parte do gene gb de citomegalovírus. **Medicina, Ribeirão Preto**, v. 35, p. 85-94, 2002.

MARIN, L.J.; CUNHA, A. A.; AQUINO, V. H.; FIGUEIREDO, L. T. M. The use of qualitative and quantitative polymerase chain reactions for diagnosis of cytomegalovirus infections in bone marrow and kidney transplant recipients. **Revista Brasileira de Medicina tropical**, v. 37, n. 2, p. 158-164, 2004.

McDONALD, M.D. Genetic purity: from eletrophoresis to RAPDs. In: **American Seed Trading Association**. Washington, D.C: Proceeding annual corn and soeghum industry research conference 50, 1995. p. 256-271.

PADILHA, L. **Marcadores moleculares semi-automatizados na caracterização e determinação da diversidade genética entre linhagens de milho tropical**. Tese (Doutorado em Agronomia-Fitotecnia) – Universidade Federal de Lavras, Lavras, MG: UFLA, 2002. 85.

PATERSON, A.H.; TANKSLEY, S.D.; SORRELLS, M.E. DNA markers in plant improvement. **Advances in Agronomy**, v. 46, p. 39-89, 1991.

QUEIROZ, V. T. **Obtenção de primers microssatélites e desenvolvimento, validação e mapeamento de marcadores scar em feijoeiro-comum**. Tese

(Doutorado em Bioquímica Agrícola) – Universidade Federal de Viçosa, Viçosa, MG: UFV, 2004. 155.

RAMOS, N. P.; BRUNELLI, K. R.; CAMARGO, L. E. A.; FILHO, J. M. Sensibilidade dos microssatélites para determinar a pureza varietal em sementes de milho. **Revista Brasileira de Sementes**, v. 28, n. 1, p. 99-105, 2006.

ROIZÈS, G. Identification of microsatellite markers: screening for repeat sequences and mapping polymorphisms. In: HAJEER, A.; WORTHINGTON, J.; JOHN, S (Eds.). **SNP and micrisatellite genotyping: marker for genetic analysis**. Westborough: Eaton Publishing, 2000. p. 35-48.

SAGHAI-MAROOF, M. A.; SOLIMAN, K. M.; JORGENSEN, R. A.; ALLARD, R. W. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. **Proc. Natl. Acad. Sci.**, USA, v. 81, p. 8014-8018, 1984.

SAX, K. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. **Genetics**, v. 8, p. 552-560, 1923.

SCHUSTER, I.; CRUZ, C. D. **Estatística genômica**. Viçosa, MG: Editora UFV. 2004. 568 p.

SLATE, J. Quantitative trait loci mapping in natural populations: progress, caveats and future directions. **Molecular Ecology**, v. 14, p. 363-379, 2005.

SOUZA, A. P. Biologia molecular aplicada ao melhoramento. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S.; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento**. [S.l.] Fundação MT, 2001. p. 939-965.

STAUB, J.E.; SERQUEN, F.C.; GUPTA, M. Genetic markers, map construction, and their application in plant breeding. **HortScience**, v. 31, n. 5, p. 729-740, 1996.

TANKSLEY, S.D. Mapping polygenes. **Annual Review of Genetics**, v. 27, p. 205-233, 1993.

THODAY, J. M. Location of polygenes. **Nature**, v. 191, p. 368-370, 1961.

VIEIRA, C. U.; COSTA, M. C. R.; SIQUIEROLI, A. C. S.; MARANHÃO, A. Q.; COSTA, C. S.; KERR, W. E.; BAFFI, M. A.; BONETTI, A. M. Expressão do gene hormônio juvenil metiltransferase em abelhas *Melipona scutellaris* (Hymenoptera, Apidae, Meliponini). **Arq. Inst. Biol.**, São Paulo, v. 71 (supl.), p. 379-380, 2004.

WU, R.; MA, CHANG-XING.; GALLO-MEAGHER, M; LITTELL, R. C.; CASELLA, G. Statistical Methods for dissecting triploid endosperm traits using molecular markers: an autogamous model. **Genetics**, v. 162, p. 875-892, 2002.

ZENG. Z. B. Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. **Proc. Natl. Acad. Sci.**, USA, v. 90, p. 10972-10976, 1993.

ZENG. Z. B. Precision mapping of quantitative trait loci. **Genetics**, v. 136, p. 1457-1468, 1994.