

**UNIVERSIDADE
ESTADUAL DE
CAMPINAS**

doutorado

BC/52145

IB/ 81872

INSTITUTO DE BIOLOGIA



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO

Luciana Lasry Benchimol



**“SELEÇÃO ASSISTIDA POR MARCADORES
MOLECULARES MICROSSATÉLITES EM PROGRAMAS
DE RETROCRUZAMENTO EM MILHO
(*Zea mays* L.)”**

Este exemplar corresponde à redação final da tese defendida pelo(a) candidato (a) <i>Luciana Lasry</i> <i>Benchimol</i> e aprovada pela Comissão Julgadora.

**Tese apresentada ao Instituto de
Biologia para obtenção do título de
Doutor em Genética e Biologia
Molecular na área de Genética
Vegetal e Melhoramento.**

Orientador: Dr. Cláudio Lopes de Souza Júnior
Co-orientadora: Dra. Anete Pereira de Souza

Campinas, SP
2002

UNIDADE	IB/23872
Nº CHAMADA	UNICAMP B431A
V	EX
TOMBO BC/	52495
PROC.	IC-124/05
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	12/06/09
Nº CPD	

CH00179041-0

BIB ID 279904

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

B431s

Benchimol, Luciana Lasry

Seleção assistida por marcadores moleculares microssatélites em programas de retrocruzamento em milho (*Zea mays L.*) / Luciana Lasry Benchimol. -- Campinas, SP: [s.n.], 2002.

Orientador: Cláudio Lopes de Souza Júnior

Co-Orientadora: Anete Pereira de Souza

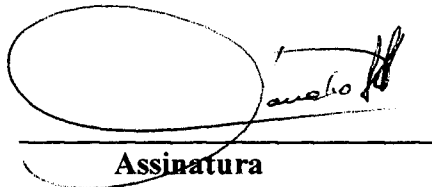
Tese (Doutorado) – Universidade Estadual de Campinas. Instituto de Biologia.

1. Seleção assistida. 2. Marcador molecular. 3. Milho. I. Souza Júnior, Cláudio Lopes de. II. Souza, Anete Pereira de. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Campinas, 25 de setembro de 2002

Banca Examinadora

Prof. Dr. Cláudio Lopes de Souza Júnior
(orientador)



Assinatura

Prof. Dr. Sérgio Furtado dos Reis



Assinatura

Prof. Dr. Carlos Eduardo Oliveira Camargo



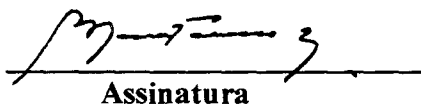
Assinatura

Prof. Dr. Sérgio Augusto Moraes Carbonell



Assinatura

Prof. Dr. Roland Vencovsky



Assinatura

Dedico...



Ao Senhor Meu DEUS,

*pela Obra que tem realizado em minha vida;
por ter-me guiado e sustentado até aqui.*

A Meus Filhos, Gabriel e Júlia,

*pela alegria constante;
por serem o maior tesouro que já recebi.*

A Meu Pai, Augusto,

pela infinita capacidade de amar.

"Tudo posso naquele que me fortalece."

(Filipenses, 4:13)

Agradeço...

Ao professor Dr. Cláudio Lopes de Souza Júnior, pela orientação, pelos "insights", sugestões e preocupação constante;

À professora Dra. Anete Pereira de Souza, pelo direcionamento, dedicação e co-orientação de grande valia à realização deste trabalho;

Aos professores da Pré-Banca e Banca de Tese, pelas suas sugestões e considerações, engrandecendo esta obra;

A Deirdre, pela revisão de inglês do artigo e ao professor Razera, pelo carinho e dedicação em revisar o português do manuscrito;

Aos colegas do esquema *power*: Sérgio T. Sibov, Alexandre F. Garcia, Adelmo R. da Silva, Claudete A. Mangolim, pelo apoio mútuo, interação, auxílio e companheirismo;

Aos demais colegas do "Laboratório de Análises Genéticas e Moleculares", pelos momentos de convivência no dia-a-dia do laboratório;

À Milena e Fernandinha, pela amizade e apoio, mesmo nas horas mais difíceis;

À Heloísa, Carolina, Andreza, Zildene, Aline, Marcus, Carlos, Juverlandi e Maria Zilda pelo trabalho conjunto e assistência técnica importantes ao sucesso do trabalho;

Ao colega Aurélio Aguiar, pelo auxílio nas análises de variâncias individuais e multivariadas;

Aos professores do Departamento de Genética/IB/UNICAMP e do Departamento de Genética da ESALQ/USP, pela minha formação acadêmica;

Às secretárias do CBMEG (Tânia, Sandra e Fabiana) e do Departamento de Genética da ESALQ/USP (Maydia, Neusinha e Márcia), pela prontidão em resolver as mais variadas questões e pelos auxílios prestados;

Aos técnicos agrícolas da ESALQ/USP, Jucelino, Ariberto, Miguel e Márcio, pela ajuda inestimável na condução dos experimentos de campo;

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro (PROC. 97/13397-1);

A meus irmãos-de-fé, pela amizade, carinho, assistência, orações e força, aliviando a caminhada neste deserto: Tiago, Adailton, Christiane, Bianca, Patrícia, Celso, Cristiana, Flávia, Beatriz, Geovana, Débora, Zenilda, Leo, Tereza, João Paulo, Paulo, D. Ordália, D. Dirce, Myriam, Pt. Edílson, Pt. Veimar, Pt. Toneguti e tantos outros.

À Teca, por me fazer acreditar;

À D. Zilda, por se fazer sempre presente;

A minha mãe e meus irmãos, pelo carinho com que os guardo no fundo de meu peito.

Índice

	Pg.
Lista de Figuras.....	ix
Lista de Tabelas.....	xii
Resumo.....	xiv
Abstract.....	xvii
1. Introdução.....	01
2. Revisão de Literatura.....	06
2.1. Melhoramento de Linhagens de Milho e o Método de Retrocruzamentos.....	07
2.2. Marcadores Moleculares	13
2.2.1. Marcadores Morfológicos e Bioquímicos.....	13
2.2.2. Marcadores Moleculares <i>RAPD</i> , <i>RFLP</i> e <i>AFLP</i>	14
2.2.3. Marcadores Moleculares Microssatélites ou SSR.....	17
2.3. Seleção Assistida por Marcadores Moleculares (SAM).....	23
3. Objetivos.....	33
4. Artigo.....	35
4.1. Marker-Assisted Selection Using Microsatellites Markers in Maize Backcross Breeding Programs.....	36
Abstract.....	37
Introduction.....	39
Material and Methods.....	42
<i>Plant Material and Backcrossing Generations</i>	42
<i>Molecular Selection Approach</i>	43

<i>DNA Extraction and Microsatellite Analysis</i>	44
<i>Amplification and Detection Conditions</i>	47
<i>Molecular Data Analysis</i>	47
<i>Experimental Evaluation and Statistical Analysis</i>	48
Results and Discussion.....	49
Acknowledgements.....	58
References.....	59
5. Conclusões.....	62
6. Perspectivas Futuras.....	64
7. Anexo.....	69
8. Referências Bibliográficas	85

Lista de Figuras

2. Revisão de Literatura

Figura I. Microssatélites: a variação no número de repetições de dinucleotídeos resulta na variação do tamanho dos fragmentos amplificados. Esta variação é visualizada em gel de acrilamida ou de agarose com alto poder de resolução. PPP = oligonucleotídeo adjacente às repetições em *tandem*, utilizando para a amplificação do microssatélite em diferentes indivíduos.....19

4. Artigo

Figure 1. Distribution of the recurrent parent genotype recovery with three selection criteria (SC_1 , SC_2 and SC_3) for three backcross generations (BC_1 , BC_2 , BC_3).....51

7. Anexo

Figura 1AN. Amplificação das plantas RC_1 com o microssatélite PHI037 (Bin 1.08). $\Phi X/HinfI$ foi usado como marcador de peso molecular. A é o alelo do genitor doador (L08-05F) e B é o alelo do genitor recorrente (L14-4B). H refere-se ao genótipo observado para o F_170

Figura 2AN. Microssatélites ancorados no cromossomo 01 e utilizados na amplificação das amostras RC_1 , RC_2 e RC_3 . CS_1 , CS_2 e CS_3 são os critérios de seleção;.....71

Figura 3AN. Microssatélites ancorados no cromossomo 02 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.72

Figura 4AN. Microssatélites ancorados no cromossomo 03 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.73

Figura 5AN. Microssatélites ancorados no cromossomo 04 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.74

Figura 6AN. Microssatélites ancorados no cromossomo 05 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.75

Figura 7AN. Microssatélites ancorados no cromossomo 06 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.76

Figura 8AN. Microsatélites ancorados no cromossomo 07 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.77

Figura 9AN. Microsatélites ancorados no cromossomo 08 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.78

Figura 10AN. Microsatélites ancorados no cromossomo 09 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.79

Figura 11AN. Microsatélites ancorados no cromossomo 10 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.80

Figura 12AN. Médias de recuperação do genótipo do genitor recorrente para cada um dos critérios (CS₁, CS₂ e CS₃) obtidas em cada geração de retrocruzamentos com seleção assistida por marcadores.81

Lista de Tabelas

2. Revisão de Literatura

Tabela I. Proporção esperada da recuperação do genoma do genitor recorrente (GR) durante os retrocruzamentos.....	11
--	----

4. Artigo

Table 1. Primers, their chromosome locations and respective Bins. <i>Sat</i> (saturated) notation refers to the extra-primer set used to amplify the 48 selected BC ₃ plants and saturate maize genome. SC is the selection criterion used.....	46
---	----

Table 2. Comparisons of BC ₃ lines percent of recurrent parent genotype for the expected mean of BC ₃ generation (93.75%) and for 100%, under non-saturated (non-sat) and saturated (sat) conditions.....	52
--	----

Table 3. Values and significances of mean squares (MS) from the analyses of variance for nine traits: ear weight (EW), number of kernel rows per ear (NKE), number of rows per ear (NRE), ear length (EL), ear diameter (ED), plant height (PH), ear height (EH), days to silk extrusion (DSE), and days to pollen shed (DPS).....	53
---	----

Table 4. Values and significances of F tests (Hotelling-Lawley Trace) from the multivariate analyses of variance for yield, height, and flowering.....	54
---	----

Table 5. Numbers and percentages (between parenthesis) of significant <i>Dunnett</i> tests for nine traits. BC ₃ lines were compared to the recurrent parent inbred line L-14-4B.....	55
---	----

7. Anexo

Tabela 1AN. Plantas RC ₁ selecionadas e o percentual de recuperação do genótipo do recorrente para os critérios de seleção avaliados (CS ₁ , CS ₂ e CS ₃).....	82
--	----

Tabela 2AN. Plantas RC ₂ selecionadas e o percentual de recuperação dos genótipos do genitor recorrente para cada critério de seleção avaliado (CS ₁ , CS ₂ e CS ₃).....	83
--	----

Tabela 3AN. Comparações entre RC _{3sel} ^{sat} (selecionada e saturada) x RC _{3sel} ^{non-sat} (selecionada e não-saturada) em relação a presença de alelos do genitor doador (A) $S_{\bar{x}}$ é o erro padrão da média.....	84
--	----

Tabela 4AN. Comparações entre os diferentes critérios das amostras RC ₃ totais (RC ₃ ^{all}), saturadas (RC _{3sel} ^{sat}) e não-saturadas (RC _{3sel} ^{non-sat}).....	85
--	----

Resumo

A aplicação da seleção assistida por marcadores moleculares em programas de retrocruzamento tem sido foco de grande interesse. Inúmeros trabalhos teóricos de simulação computacional despontaram a fim de averiguar os parâmetros necessários para que uma estratégia experimental eficiente seja posta em prática. Poucos estudos experimentais efetivamente relatam sucesso na introgressão de características de interesse com seleção assistida por marcadores moleculares.

Marcadores moleculares microssatélites foram utilizados para introgredir um alelo alvo de uma linhagem doadora (L-08-05) para uma linhagem recorrente (L-14-4B), provenientes de grupos heteróticos distintos. Uma marca molecular polimórfica e de fácil leitura, PHI037 (Bin 1.08), foi escolhida para ser transferida, simulando uma característica monogênica ou um traço quantitativo ligado a ela. Três gerações de retrocruzamentos foram realizadas. As plantas foram genotipadas, selecionadas apenas em nível molecular e ao mesmo tempo retrocruzadas. Três critérios de seleção foram utilizados para monitorar a recuperação do genótipo do genitor recorrente: a) duas marcas (CS₁); b) quatro marcas (CS₂); c) seis marcas (CS₃), distribuídas entre os dois braços de cada cromossomo. A cada ciclo, plantas foram selecionadas mediante a presença do alelo alvo cujo loco estava em estado heterozigoto e também pela maior recuperação do genótipo do recorrente. Apenas os microssatélites não-recuperados na geração anterior foram utilizados para genotipagem da outra geração. Em RC₁, 53 microssatélites polimórficos foram utilizados para genotipar 186 plantas e quatro plantas foram escolhidas de cada critério. Em RC₂, 60 plantas de cada

critério (180 plantas no total) foram analisadas e quatro plantas de cada progênie dentro de cada critério foram selecionadas. Em RC₃, 160 plantas por critério (480 plantas no total) foram genotipadas. As quarenta e oito plantas RC₃ selecionadas (16 de cada critério) foram genotipadas com marcas adicionais para saturar o genoma. Testes-*t* foram utilizados para comparar as médias entre linhagens RC₃, saturadas e não-saturadas, com a média esperada (93,75%) e com média saturada (100%). As linhagens RC₃ foram avaliadas com 09 caracteres fenotípicos, sendo que a linhagem L-14-4B foi incluída na avaliação. Análises de variâncias individuais, para os nove caracteres, e multivariadas, para três conjuntos de caracteres (produção, altura e florescimento) foram realizadas. As 48 linhagens RC₃ selecionadas foram comparadas com o genitor recorrente, para todos os caracteres, utilizando-se para isto o teste de *Dunnett*.

Para todos os critérios de seleção, com exceção das médias de recuperação para as linhagens selecionadas saturadas e não-saturadas com marcas do critério CS₁, as médias das linhagens RC₃ foram significativamente superiores ($P \leq 0,01$) ao valor esperado (93,75%) e significativamente inferiores ($P \leq 0,01$) ao valor máximo esperado de 100%. Evidenciou-se que, com o aumento do número de marcas (CS₃), houve a redução na diferença entre as médias obtidas para as linhagens selecionadas saturada e não-saturada e também redução no desvio-padrão. Com o aumento do número de marcas, também houve a redução da presença de alelos provenientes do genitor doador na geração RC₃ selecionada e saturada. Na análise de variância individual para os nove caracteres avaliados, apenas dois deles apresentaram baixa significância ($P \leq 0,05$). Pela análise multivariada, em nível de critérios de seleção, não houve qualquer significância. Pelo teste de *Dunnett*, para cada critério de seleção, foram realizados 120 testes dos quais 13, 14 e 8 foram significativos para CS₁, CS₂ e CS₃, respectivamente. Assim, com aumento de marcas em nível fenotípico, um maior

número de linhagens mostrou identidade em relação ao genitor recorrente. Pelos resultados moleculares e fenotípicos, pode-se concluir que o alelo PHI037 foi eficientemente transferido e que o critério CS₃, com aproximadamente seis marcas por cromossomo, se mostrou superior aos demais.

Abstract

Marker-assisted selection (MAS) has received great attention and many are the efforts to establish the best strategy for using marker-assisted selection as a routine in backcross breeding programs. Although much computer simulation has evaluated the parameters that influence marker-assisted selection efficiency, few experimental data report successful results of introgression of desirable traits and estimate the necessary numbers of markers needed for a complete recurrent genotype recovery.

Microsatellites were used to introgress a single-locus trait from a L-08-05 donor line to a L-14-4B recurrent divergent inbred line. A clear reading pattern and polymorphic microsatellite, PHI037 (Bin 1.08), was chosen to be transferred, simulating an monogenic trait or a QTL linked to it. Three generations of backcross were conducted. Plants were genotyped by molecular approach, selected, and at the same time, backcrossed. Three selection criteria were established to estimate the recoveries of the recurrent parent genotypes: a) two (SC_1) markers, b) four (SC_2) markers, c) six (SC_3) markers per chromosome. At each backcross cycle, plants were selected based on the target allele the *locus* of which was in a heterozygous state and on the greatest recovery of the recurrent genotype. Only the non-recovered microsatellites, from the previous cycle, were used to genotype BC samples in the next generation. In BC_1 , fifty-three polymorphic microsatellites were used to genotype the 186 plants and four plants were chosen based on each criterion. In BC_2 , 60 plants of each criterion (180 plants in total) were genotyped and four plants of each progeny within each criterion were selected. In BC_3 , 160 plants per criterion (480 plants in total) were genotyped. Forty-eight BC_3 lines (16 from each criterion) were selected from the BC_3 lines and genotyped with those microsatellites used only in the other criteria (for SC_1 and SC_2) and also with another additional 19 microsatellites distributed in Bins not assayed previously. This procedure was carried out with the

aim of saturating maize genome and comparing the estimates of the recurrent parent genotype recoveries, under saturated and non-saturated conditions, for each criterion. *t*-Tests for comparison between BC₃ line (saturated and non-saturated) means for the expected mean (93.75%) of this generation and for 100% recovery were computed. The 48 selected BC₃ lines were evaluated in a field experiment and the recurrent L-14-4B inbred line was also included. Individual analyses of variance for the nine traits and multivariate analyses of variance for three composite traits (yield, height and flowering) were performed. Comparison between the 48 selected BC₃ lines and the recurrent parent (L-14-4B) were performed using the *Dunnnett* test.

For all selection criteria, except from the recoveries for the selected lines for SC₁, under saturated and non-saturated conditions, BC₃ line averages were significantly superior ($P \leq 0.01$) to the expected value (93.75%) and significantly inferior ($P \leq 0.01$) to the maximum expected value (100%). The increase of the number of markers contributed towards a reduction in the donor fragment content. SC₃ criterion also showed smaller differences between saturated and non-saturated means and standard deviations for their selected lines. According to analyses of variance for the nine traits, only two of them presented low significance ($P \leq 0.05$). No significance was detected for selection criteria by multivariate analyses of variance. The *Dunnnett* test produced 120 tests for each selection criterion and 13, 14 and 8 were significant for SC₁, SC₂, SC₃, respectively. Phenotypically, the higher amount of markers produced a superior number of BC₃ lines that did not differ statistically from the recurrent parent. Gathering molecular and phenotypic results, it can be concluded that PHI037 allele was efficiently transferred and that the SC₃ criterion, with approximately 06 markers per chromosome, was superior to perform marker-assisted backcross selection.

1. INTRODUÇÃO

1. Introdução

Antes do século vinte, o melhoramento de plantas era artesanal, com pouco conhecimento de Genética. Entretanto, a redescoberta das “Leis de Mendel” fez com que o melhoramento de plantas evoluísse como arte e ciência. Na medida em que novas tecnologias se tornam mais abrangentes ganhos adicionais em produtividade agrícola são requeridos para suportar o aumento populacional. Oportunidades de usar novas formas de seleção genotípica ou métodos combinados de seleção fenotípica e genotípica visando ao aumento da produtividade agrícola têm emergido continuamente.

Vários métodos de seleção podem ser empregados no melhoramento genético de plantas (Allard, 1960). A escolha do método mais eficiente e adequado depende de vários fatores, dentre os quais, da herança da característica a ser melhorada. Entre os métodos utilizados no melhoramento de linhagens, destaca-se o método de retrocruzamentos cuja aplicação no melhoramento de plantas foi originalmente demonstrada por Harlan & Pope (1922).

O método de retrocruzamentos envolve uma série de cruzamentos da progênie de dois genitores (Borém, 2001). O genitor que participa apenas do cruzamento inicial é dito genitor doador ou não-recorrente e possui a característica alvo a ser transferida. O genitor que é utilizado em cruzamentos sucessivos, durante o programa, é dito genitor recorrente. Este método é geralmente utilizado para se melhorar uma característica qualitativa em um cultivar que é deficiente (Fehr, 1987). Ao final, o objetivo é se atingir a recuperação do genótipo do genitor recorrente, exceto em relação ao alelo em transferência. A proporção dos genes do genitor doador é reduzida à metade a cada geração de retrocruzamentos

(Borém, 2001). Na ausência de seleção e *linkage drag*, a recuperação média do genitor recorrente após seis gerações é de aproximadamente de 99%.

Com o advento dos marcadores moleculares e dos primeiros mapas moleculares genéticos, a seleção assistida por marcadores (SAM) passou a ser uma idéia atraente para os melhoristas. Em retrocruzamentos, a tecnologia dos marcadores moleculares pode ser aplicada em vários estágios. O *fingerprinting* de DNA de coleções de germoplasma pode ser usado para identificar os melhores genitores de um cruzamento; pode ser utilizado também para identificar a divergência entre os genitores, estimando-se o esforço que se terá para se re-estabelecer o genótipo genitor elite. Em cada ciclo de retrocruzamento, os marcadores podem ser utilizados para identificar linhagens portadoras do alelo alvo, contendo a menor quantidade de genótipo doador e para encontrar linhagens com menor quantidade de segmentos ao redor do alelo alvo (*linkage drag*). Para características recessivas, marcadores moleculares podem evitar a necessidade de uma geração adicional de autofecundação necessária para se identificar o alelo alvo (Langridge *et al.*, 2001).

Apesar de muitos trabalhos teóricos de simulação computacional, envolvendo algoritmos sofisticados (Hillel *et al.*, 1990; Hospital *et al.*, 1992; Frisch *et al.*, 1999a,b; Hospital, 2001; Frisch & Melchinger, 2000), poucos exemplos de sucessos práticos estão disponíveis (Weck *et al.*, 1991; Stuber, 1994; Openshaw *et al.*, 1994; Ragot *et al.*, 1995; Mesquita *et al.*, 2001). Isto parece ser devido à necessidade de desenvolvimento de novas estratégias que levem em conta a complexidade do genoma, a natureza e número dos marcadores moleculares disponíveis, a complexidade das características a serem introgridas, o número de plantas a serem avaliadas em cada geração e o número de populações que possam ser manipuladas (Ribaut & Hoisington, 1998). As magnitudes dos

benefícios esperados variam de acordo com o trabalho de simulação. Openshaw *et al.* (1994) relataram que o número de gerações de retrocruzamentos poderia diminuir de sete para três, usando um tamanho de amostra reduzido ($N < 100$) e também poucos marcadores ($N < 80$), em linhagens de milho. Hospital *et al.* (1992) concluíram que, com o uso da seleção assistida, seria possível uma redução do número de gerações de sete para duas. O resultado destas simulações não pode ser comparado, entretanto, pois as condições e pressupostos utilizados foram diferentes (Lee, 1995). O milho tem 10 cromossomos de aproximadamente 200 cM. Ao que parece, nos cromossomos não-portadores do alelo alvo, dois marcadores por cada 100 cM seriam adequados para a recuperação do genótipo do genitor recorrente (Hospital *et al.*, 1992).

Evidências apontam que marcadores de DNA oferecem a vantagem de uma recuperação mais rápida do genótipo do genitor recorrente (Weck *et al.*, 1991; Openshaw *et al.*, 1994). A seleção assistida com marcadores é esperada ser mais eficiente que a seleção fenotípica quando a herdabilidade do caráter é baixa, quando existe uma ligação estreita entre o QTL a ser transferido e os marcadores de DNA (Dudley, 1993; Knapp, 1998) e quando se amplia o tamanho populacional (Moreau *et al.*, 1998).

Marcadores têm sido utilizados em programas de retrocruzamento para a introgressão de resistência a vírus em tomate (Young & Tanksley, 1989), em arroz (Ahmadi *et al.*, 2001) e para capacidade de regeneração em cultura de tecidos em milho (Armstrong *et al.*, 1992). A seleção assistida por marcadores pode gerar ganhos genéticos em muitas circunstâncias. Para introgridir alelos exóticos, por exemplo, para germoplasma adaptado, a seleção assistida por retrocruzamentos utilizando marcadores moleculares

poderia diminuir o número de gerações necessárias e facilitar a recuperação do genótipo recorrente (Lee, 1995).

De acordo com Mohan *et al.* (1997), a introgressão por seleção assistida por marcadores para características monogênicas, cuja posição do gene de interesse no cromossomo é conhecida com certeza, tem sido relatada por muitos autores. Existe, entretanto, menor número de relatos de trabalhos lidando com *QTLs* cujas posições ainda não foram determinadas (Bernacchi *et al.*, 1998; Van Berloo & Stam, 1998).

O presente trabalho se propôs a averiguar experimentalmente a utilização de marcadores moleculares do tipo microssatélites para assessorar a introgressão de uma característica monogênica arbitrária, um microssatélite polimórfico, em programas de retrocruzamentos com milho tropical. Entre os parâmetros avaliados, destaca-se o número de marcas por cromossomo e sua influência na estimativa de recuperação do genótipo do genitor recorrente. O procedimento adotado para a condução da seleção foi puramente molecular, tanto para os critérios de seleção para o alelo alvo a ser introgridido quanto para o monitoramento da recuperação do genótipo do genitor recorrente. Neste contexto, o presente trabalho pôde investigar a eficiência e viabilidade da seleção assistida por marcadores moleculares microssatélites em retrocruzamentos, quando uma marca molecular é transferida do genitor doador para o recorrente, em três ciclos de retrocruzamentos, simulando uma característica monogênica ou uma marca ligada a *QTL*. Teve-se como objetivo também avaliar o critério mais adequado para efetuar a introgressão do alelo alvo, com o maior percentual de recuperação do genótipo do genitor recorrente e a menor manutenção possível de alelos do genitor doador.

2. REVISÃO DE LITERATURA

2. Revisão de Literatura

2.1. Melhoramento de Linhagens de Milho e o Método de Retrocruzamentos

O milho (*Zea mays* L.) é uma espécie altamente politípica, com cerca de 300 raças identificadas (Paterniani, 1990) e milhares de variedades. Pela sua ampla capacidade adaptativa e grande importância econômica, é a espécie vegetal mais cosmopolita do mundo (Dowswell *et al.*, 1996). Há raças e variedades específicas adaptadas a condições ecológicas igualmente específicas (Paterniani, 1995), como a diferentes altitudes (0-1000m, 1000-2000m, acima de 2000m) e latitudes (Equador até 40° de latitude Norte e Sul). Há variabilidade quanto aos tipos de endospermas cultivados (dentado, duro ou cristalino, amiláceo, pipoca e doce) e estes são cultivados visando atender às necessidades específicas de consumo. No Brasil, o milho semidentado amarelo (híbrido resultante do cruzamento do dentado e cristalino) e o dentado amarelo são os mais populares.

Entre os cereais cultivados no mundo, o milho lidera o segundo lugar no *ranking* de produção, sendo apenas superado pelo trigo. É o milho uma espécie alógama, cuja forma de reprodução natural ocorre com o acasalamento entre plantas; assim, as plantas de uma geração qualquer são naturalmente oriundas de gametas femininos e masculinos de diferentes plantas (Souza Jr., 2001).

As quatro maiores atividades de melhoramento de linhagens e híbridos em milho são: (a) seleção de populações de base melhoradas, (b) desenvolvimento de novas linhagens, (c) melhoramento de linhagens estabelecidas e (d) avaliação dos genitores

quanto à capacidade de combinação geral e específica. Todos os quatro ramos de atividade são importantes e essenciais para um progresso genético contínuo. No caso, o melhoramento das linhagens estabelecidas envolve: (a) cruzamento com alguma fonte (genitor não-recorrente) portadora de características almejadas; (b) número variável de retrocruzamentos que podem ser feitos, especialmente, quando se almeja uma recuperação do tipo próximo à linhagem original e (c) testes das linhagens recuperadas em combinações híbridas (Bauman, 1977).

Para um programa de retrocruzamentos ser bem sucedido, certas condições básicas devem ser satisfeitas (Briggs & Allard, 1953): (a) um genitor recorrente satisfatório deve existir; (b) deve ser possível reter uma intensidade adequada do caráter que está sendo transferido através dos vários retrocruzamentos; (c) o genótipo do genitor recorrente deve ser reconstituído em um número razoável de gerações, com tamanho populacional de fácil manuseio.

Muitas características importantes economicamente já foram introduzidas em variedades cultivadas de plantas por meio de retrocruzamentos, como, por exemplo, genes de resistência a patógenos em tomate (Young & Tanksley, 1989), genes de tolerância a estresse e de qualidade de fruto (Rick, 1974; 1982), genes de resistência (Sharma & Gill, 1983) ou peso de sementes (Knott & Talukdar, 1971), em trigo. Em milho, um exemplo característico pode ser ilustrado pela variedade *Piranão*, portadora do genótipo br_2br_2 , sendo br_2 a designação do gene braquítico (Paterniani & Risi, 1976). Este gene reduz o porte da planta sem afetar o tamanho da espiga. A introdução deste gene foi feita por cruzamento com o genitor doador *Tuxpeño* que tem o mesmo germoplasma do *Piramex*. Nos EUA, os milhos braquíticos não tiveram muito sucesso porque, entre outras questões, o

número de retrocruzamentos (2 a 3) tem se mostrado insuficiente para uma recuperação satisfatória do genótipo do genitor recorrente.

O valor do potencial dos retrocruzamentos no melhoramento de plantas parece ter escapado à atenção dos melhoristas até 1922, quando Harlan & Pope assinalaram as possibilidades do método no melhoramento de cereais de grãos pequenos. Anteriormente, este método era apenas aplicado junto ao melhoramento animal, para fixar características raciais em animais domésticos.

O método de retrocruzamentos no melhoramento de plantas consiste numa maneira precisa de melhoria de linhagens ou variedades elites que possuam um grande número de características desejáveis, mas são deficientes em algum atributo (Allard, 1960). No final da série de cruzamentos, o(s) gene(s) que está(ao) sendo transferido(s) estará(ao) em condição heterozigótica, diferentemente dos demais. Depois do último retrocruzamento, uma autofecundação é conduzida para restaurar a homozigose para o gene introgridido. Este método pode ser utilizado tanto para plantas autógamas como para alógamas.

Normalmente, nesse processo, quando a linhagem elite é cruzada como doadora, os diversos retrocruzamentos realizados são acompanhados de intensa seleção para se incorporarem os caracteres desejáveis e para recuperação, pelo menos parcial, do desempenho da linhagem elite (Souza Jr., 2001). Espera-se que, em se praticando seleção, isto resultará numa linhagem com exatamente a mesma adaptação, produtividade e demais qualidades do genitor recorrente. Este método, ao contrário dos métodos de melhoramento genealógico e de população, dá ao melhorista um alto grau de controle genético do seu material (Allard, 1960). De acordo com Briggs & Allard (1953), este método é bem adequado para realizar pequenas substituições necessárias para aumentar a eficácia de

variedades bem sucedidas sem o risco de quebra de combinações gênicas existentes e favoráveis.

A transferência de alelos dominantes é feita mais diretamente que a de alelos recessivos, na qual a autofecundação torna-se essencial para manifestação do alelo. Além disto, características de alta herdabilidade são transferidas mais facilmente. O tamanho populacional necessário é tanto maior quanto maior o número de genes que se almeja transferir. Entretanto, o método se torna complexo para transferência de características quantitativas (Borém, 2001).

O número de retrocruzamentos depende do grau de divergência genética das linhagens genitoras; quanto mais diferentes em desempenho forem essas linhagens, maior deve ser o número de retrocruzamentos e vice-versa (Souza Jr, 2001). O mérito agrícola do genitor doador, a intensidade de seleção dos caracteres do genitor recorrente, da ligação gênica entre o gene que está sendo transferido e outros indesejáveis também influenciam. Para a transferência de apenas um alelo, no final do processo, deve-se obter uma linhagem original com apenas esse alelo transferido. Além disso, quando se trata de caracteres de alta herdabilidade, qualitativos ou quantitativos, procura-se recuperar a quase totalidade do genótipo da linhagem elite. O número de retrocruzamentos nestes programas é maior, chegando a seis ou sete gerações para garantir a recuperação da linhagem elite; o número de plantas a serem manipuladas também deve ser maior.

Entretanto, é sabido que a proporção de genes do genitor doador é reduzida à metade após cada geração de retrocruzamentos. Na ausência de seleção e *linkage drag*, a recuperação média do genitor recorrente (GR) pode ser estimada pela fórmula

$$GR\% = 100 \left[1 - \left(\frac{1}{2} \right)^{n+1} \right],$$

onde “n” é o número de retrocruzamentos (Tabela I). Após seis gerações, a recuperação esperada é de aproximadamente de 99%. Briggs & Allard (1953) afirmaram que, quando há

Tabela I. Proporção esperada da recuperação do genoma do genitor recorrente (GR) durante os retrocruzamentos.

Geração	GR %
F ₁	50,000
RC ₁	75,000
RC ₂	87,500
RC ₃	93,7500
RC ₄	96,8750
RC ₅	98,4375
RC ₆	99,2188
RC ₇	99,6094

rígida seleção nas primeiras gerações, o número de retrocruzamentos satisfatório encontra-se em torno de cinco.

É importante ressaltar que os valores da Tabela I são médias de recuperação, já que a cada geração há uma variação entre os indivíduos quanto ao número de genes do genitor recorrente que eles possuem.

Uma das maiores dificuldades durante um programa de retrocruzamentos é lidar com o efeito da ligação gênica, que promove um “arraste” de muitos outros genes e regiões indesejáveis junto com o gene alvo que está sendo transferido. Este fenômeno é dito

linkage drag. A estimativa de recuperação média do genitor recorrente $(1 - (1/2)^{n+1})$ expressa acima ignora o efeito do *linkage drag* (Borém, 2001).

Há vários estudos tentando prever ou estimar as porções indesejáveis ligadas e transferidas com este processo durante a introgressão gênica em retrocruzamentos. Hanson (1959), por exemplo, considerou o fenômeno de *linkage drag* e desenvolveu previsões de comprimento dos segmentos cromossômicos retidos em volta do loco mantido heterozigoto. Naviera & Barbadilla (1992) derivaram distribuições teóricas sobre os comprimentos dos segmentos de cromossomo doador. Stam & Zeven (1981) fizeram previsões sobre as porções de genoma doador na geração seguinte. De acordo com Stam & Zeven (1981) e Zeven *et al.* (1983), após 20 ciclos de retrocruzamentos, espera-se encontrar ainda uma porção considerável (10 cM) do cromossomo doador ligada à característica selecionada. Na maioria dos genomas, 10 cM é suficiente para conter centenas de genes. Young & Tanksley (1989) encontraram que os tamanhos de fragmentos introgrididos variavam de 4 a 51 cM. É importante ressaltar que a transferência de genes adicionais ao gene de interesse pode resultar uma nova linhagem modificada para outros caracteres que não sejam aqueles originalmente desejados. Isto é mais grave se o genitor doador for agronomicamente inferior, pois a probabilidade de que genes indesejáveis estejam ligados ao gene de interesse é grande.

2.2. Marcadores Moleculares

2.2.1. Marcadores Morfológicos e Bioquímicos

Mendel foi o primeiro a usar a avaliação fenotípica baseada em marcadores morfológicos para estudar a herança em ervilhas. Posteriormente, em *Drosophila*, também foram utilizados marcadores fenotípicos e demonstrou-se que alguns destes marcadores eram herdados de uma forma não-aleatória, estabelecendo-se a teoria de ligação gênica. Esta classe de marcadores possui características importantes como facilidade de medida, penetrância completa, além de natureza qualitativa. Até a década de 60, esta categoria de marcadores morfológicos foi utilizada em muitos estudos de genética e melhoramento relacionados com fenótipos de fácil identificação visual, como nanismo, deficiência clorótica, cor da pétala ou morfologia foliar (Ferreira & Grattapaglia, 1998).

Infelizmente, marcadores morfológicos exibem pouco ou nenhum polimorfismo na maioria dos organismos (Reiter, 2001). Isto reduz a probabilidade de se encontrarem associações significativas entre estes marcadores e genes de importância, limitando sua utilização em programas de melhoramento.

Uma categoria de marcadores “mais neutros” tornou-se necessária e o primeiro passo neste sentido veio com o desenvolvimento de métodos bioquímicos e com o estudo de proteínas. Iniciou-se o estudo da variabilidade das isoenzimas (Stuber & Goodman, 1983), grupo de múltiplas formas moleculares da mesma enzima que ocorre numa espécie. Com isto, o número de marcadores genéticos foi ampliado em pelo menos uma ordem de grandeza. Potencialmente, todas as espécies de plantas podem ser analisadas por meio de isoenzimas e variação alélica foi encontrada.

Isoenzimas têm sido utilizadas em diferentes estudos genéticos desde variabilidade genética (Allard *et al.*, 1992), melhoramento de espécies florestais (Adams, 1983), introgressão gênica e avaliação de germoplasma (Tanksley & Orton, 1983). Entretanto, esta classe de marcadores apresenta duas principais limitações que são o número total de locos que podem ser detectados no genoma e o número de alelos por loco ou polimorfismo genético detectável em cada loco. Há outras dificuldades quanto ao uso das isoenzimas que devem ser consideradas: modificações pós-traducionais das proteínas, especificidade de formas isoenzimáticas em tecidos vegetais, polimorfismo em resposta a condições ambientais, diferenças de atividade de expressão de acordo com o estágio de desenvolvimento da planta, dificuldade de interpretação dos isolocos (Murphy *et al.*, 1990).

2.2.2. Marcadores Moleculares *RAPD*, *RFLP* e *AFLP*

Com os avanços da Biologia Molecular, técnicas capazes de acessar diretamente o DNA surgiram. Apesar de técnicas de seqüenciamento já terem sido desenvolvidas, elas eram muito laboriosas e não muito convidativas a uma determinação genotípica rápida. Entretanto, com o estabelecimento do primeiro mapa genético usando polimorfismos no comprimento de fragmentos de restrição ou *RFLPs* (*Restriction Fragment Length Polymorphism*; Botstein *et al.*, 1980), houve uma guinada na área de análises genéticas. Os *RFLPs*, apesar de muito mais trabalhosos e custosos que as isoenzimas e difíceis de serem automatizados (Newbury & Ford Lloyd, 1993), mostraram-se poderosos pela abundância de locos polimórficos. A detecção do polimorfismo entre indivíduos é baseada na diferença do tamanho de fragmentos de restrição do DNA que podem ser causados por mutações ao nível do sítio de restrição, ou por rearranjos de maiores magnitudes, como translocações e

inversões (Beckman & Soller, 1986). O marcador *RFLP* apresenta as seguintes vantagens: grande número de marcadores disponíveis, reprodutibilidade e herança co-dominante. Tem sido utilizado em grande número de estudos, seja para a construção de mapas moleculares (O'Donoghue *et al.*, 1995), mapeamento de características de interesse (Millach *et al.*, 1997), caracterização de germoplasma (Oliveira *et al.*, 1996), avaliação de diversidade genética e correlação com dados de desempenho (Melchinger, 1999; Benchimol *et al.*, 2000).

Entretanto, com o desenvolvimento do *PCR* (*Polymerase Chain Reaction*; Saiki *et al.*, 1985), surgiram técnicas moleculares baseadas na amplificação do DNA. As metodologias atuais mais utilizadas são os *RAPDs* (*Random Amplified Polymorphic DNA*), fundamentados na amplificação de seqüências arbitrárias e os microssatélites (SSRs - *Single Sequence Repeats*), que requerem informações de seqüências estabelecidas *a priori*.

Os *RAPDs* foram descobertos por dois grupos simultâneos (Williams *et al.*, 1990; Welsh & McClelland, 1990). Caracterizam-se por serem marcadores dominantes, de extrema simplicidade técnica e de baixo custo. Além disto, requerem pequena quantidade de DNA e são de rápida execução (Waugh *et al.*, 1992). O nível de polimorfismo obtido com *RAPDs* varia grandemente com a espécie. Têm sido amplamente utilizados em estudos de divergência genética entre genótipos (Jain *et al.*, 1994; Santos *et al.*, 1994; Lanza *et al.*, 1997), mapeamento de características agronômicas de interesse (Wight *et al.*, 1994; Grattapaglia, 1994), mas apresentam muitas vezes dificuldades quanto à reprodutibilidade com o uso de *primers* aleatórios devido a anelamento com *mismatch* (Neale & Harry, 1994). Esta sensibilidade a condições experimentais dificulta a reprodução entre laboratórios. Pejic *et al.* (1998) empregaram diferentes marcadores moleculares comparativamente para acessar a diversidade genética entre linhagens de milho e

encontraram grande distorção do agrupamento gerado por *RAPD* em relação aos outros marcadores.

Uma outra técnica que evoluiu com a chegada do *PCR* foi a de *AFLP* (*Amplified Fragment Length Polymorphism*; Zabeau & Vos, 1993; Lin & Kuo, 1995), cujo polimorfismo observado por reação é muito maior do que o revelado por *RFLP* ou *RAPD*. Em uma única reação de *PCR*, podem-se detectar mais de 50 locos usando para isto pouca quantidade de DNA. Seu polimorfismo está baseado em diferenças entre genótipos, na distribuição dos sítios de restrição e na amplificação diferencial de fragmentos. Assim como o *RAPD*, o *AFLP* é dominante, entretanto, apesar de mais laborioso, é muito mais reprodutível. Esta metodologia oferece a oportunidade de realizar estudos genéticos detalhados em muitos organismos, tanto para a caracterização de cultivares (Powell *et al.*, 1996; Hongtrakul *et al.*, 1997) e mapeamento genético (Thomas *et al.*, 1995; Zimnoch-Guzowska *et al.*, 2000), quanto para análises de divergência genética (Lima, 2000; Barbosa *et al.*, 2001). Um grande número de mapas genéticos já foi construído em várias espécies de plantas usando *AFLP* (Vuylsteke *et al.*, 1997).

A técnica de *AFLP* tem custos iniciais mais baixos e é mais aplicável entre espécies. *AFLPs* têm sido amplamente aplicados em espécies como cana-de-açúcar, em que os microssatélites, outra classe de marcadores moleculares, estão ainda limitados em número e em uso (Cordeiro *et al.*, 2000).

2.2.3. Marcadores Moleculares Microssatélites ou SSR

Nos genomas de eucariotos, há diferentes classes de seqüências repetidas, umas mais complexas (minisatélites) e outras mais simples, os microssatélites (Hamada *et al.*, 1982; Tautz & Renz, 1984). Os microssatélites ou seqüências simples repetidas, ou ainda SSRs (*Simple Sequence Repeats*), consistem em pequenas seqüências com 2 a 6 nucleotídeos de comprimento, repetidas em *tandem*, de 10 a 60 vezes (Gupta *et al.*, 1996), centenas de vezes como no caso do motivo ATT de beterraba (Mörchen *et al.*, 1996), ou até milhares de vezes (Brown *et al.*, 1996). Estas seqüências estão dispersas pelo genoma. Geralmente, repetições compostas de dois ou mais motivos de repetição são encontradas (Morgante & Olivieri, 1993).

Os microssatélites ocorrem freqüentemente e ao acaso em todos os DNAs de eucariotos analisados (Gupta *et al.*, 1996). Em especial, em genomas vegetais, existe ampla variabilidade quanto à freqüência de ocorrência e a própria forma de distribuição (Ogliari, 1999). Para milho, a freqüência média de distribuição é de um loco a cada 50 mil pb (pares de bases; Chin *et al.*, 1996; Morgante & Olivieri, 1993). Wang *et al.* (1994) compararam mono e dicotiledôneas e verificaram que a freqüência média de ocorrência desses elementos repetitivos é de um loco a cada 21 mil e 65 mil pb, respectivamente.

As regiões repetidas são geralmente compostas por di, tri, tetranucleotídeos, ou ainda, às vezes, seqüências repetidas maiores (Tautz & Renz, 1984). Entretanto, seqüências em bancos de dados têm revelado que a abundância relativa dos vários motivos de microssatélites em plantas e animais difere consideravelmente. As repetições de tri e tetranucleotídeos são comuns em humanos. Nos genomas de humanos e ratos, as repetições (AC)_n são mais freqüentemente encontradas (Gupta *et al.*, 1996).

Em plantas, os marcadores microssatélites já foram descritos para inúmeras espécies vegetais, como arroz (Wu & Tanksley, 1993), milho (Chin *et al.* 1996; Holm *et al.* 1994; Senior & Heun, 1993), soja (Akkaya *et al.*, 1992; Morgante & Olivieri, 1993; Rongwen *et al.*, 1995), brássica (Lagercrantz *et al.*, 1993), videira (Browers *et al.*, 1996), entre outras. As repetições mais freqüentes são dinucleotídeos (AT)_n. Além disso, nestes genomas, as repetições (GA)_n são mais freqüentes do que as (CA)_n. Em trigo, um estudo de diversidade em 40 cultivares e linhagens utilizando 23 microssatélites, representou um total de 142 alelos, localizados em 15 cromossomos diferentes. O número de alelos em um loco variou de 3 a 6, com média de 6,2 alelos por loco (Gupta *et al.*, 1996). Em milho, Chin *et al.* (1996) identificaram três classes mais abundantes dentre as 576 seqüências analisadas: (AG)_n, (CCT)_n e (CCG)_n.

A explicação precisa para a variação encontrada no número de repetições em *tandem* de um microssatélite é desconhecida, mas acredita-se que esta esteja associada à ocorrência de *crossing-overs* desiguais ou por erros durante a replicação do DNA, processo descrito como *slippage* e seguido por falha no sistema de reparo que restaura a seqüência original (Chalesworth *et al.*, 1994; Levison & Gutman, 1987; Strand *et al.*, 1993; Souza, 2001).

Enquanto o número de repetições em *tandem* de um microssatélite geralmente varia, a seqüência de bases adjacente ao microssatélite pode ser única no genoma e conservada (no mesmo loco), entre diferentes indivíduos da mesma espécie (Souza, 2001). Assim, pode-se desenhar um oligonucleotídeo específico para as seqüências adjacentes a um dado microssatélite e através de uma reação de *PCR* será possível amplificar este loco em diferentes genótipos. A detecção dos locos de microssatélites pode ser feita por amplificação individual através de *PCR*, utilizando um par de *primers* específicos (de 20 a

30 bases) complementares a seqüências únicas que flanqueiam os microssatélites. Para que esta reação aconteça, é necessária pouca quantidade de DNA (50 a 100 ng), sendo que os produtos gerados têm geralmente de 50 a 300 pb. Cada loco de microssatélites é analisado individualmente ao se utilizar o par de *primers* construído especificamente para sua amplificação. A detecção das seqüências de microssatélites via *PCR* é feita em gel de eletroforese utilizando-se poliacrilamida ou agarose especial de alta resolução, já que é necessário um gel adequado para a separação de segmentos que diferem por poucos pares de bases. A visualização é feita por brometo de etídeo, nitrato de prata ou por autoradiograma, quando se utiliza oligos marcados com radioisótopos (Ferreira & Grattapaglia, 1998).

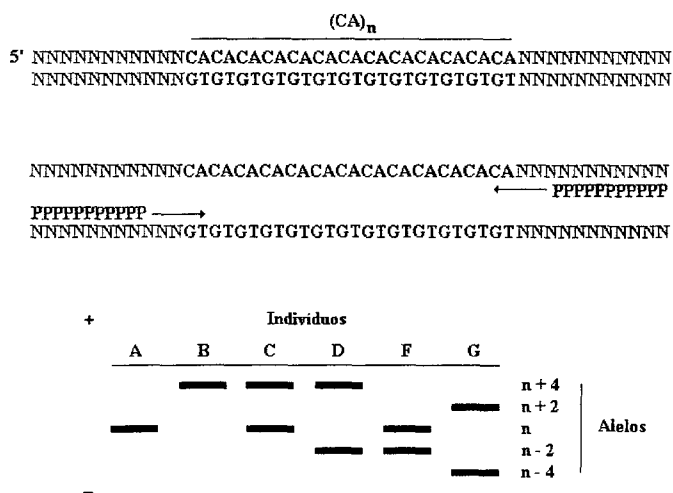


Figura 1. Microssatélites: a variação no número de repetições de dinucleotídeos resulta na variação do tamanho dos fragmentos amplificados. Esta variação é visualizada em gel de acrilamida ou de agarose com alto poder de resolução. PPP = oligonucleotídeo adjacente às repetições em *tandem*, utilizando para a amplificação do microssatélite em diferentes indivíduos.

Os locos de microssatélites são altamente polimórficos entre espécies e também entre indivíduos dentro de uma mesma espécie ou população. A variação do número de unidades de repetição é a base do polimorfismo observado; quanto maior o número de repetições, maior o número de alelos detectados (Gupta *et al.*, 1996, Figura I).

O polimorfismo dos locos microssatélites é detectado pela diferença nos fragmentos amplificados na reação de *PCR* com *primers* específicos. Senior & Heun (1993) encontraram, para microssatélites em milho, polimorfismo semelhante ao encontrado com marcador *RFLP*.

Esta classe de marcadores revelou os seguintes atributos (Gupta *et al.*, 1996): (a) são co-dominantes como os *RFLPs*, (b) existem muitos alelos numa população e o nível de heterozigosidade é extremamente alto e (c) são marcadores herdáveis de maneira Mendeliana e podem ser utilizados para análise de ligação e mapeamento. Outras qualidades também foram apontadas por Burr (2001), como o fato de serem baseados em amplificação por *PCR* e poderem ser resolvidos por eletroforese e separados em géis de agarose de alta densidade. Esta classe de marcadores é ainda a que possui o mais elevado conteúdo de informação de polimorfismo (*PIC* ou *Polimorphism Information Content*). De acordo com Garcia *et al.* (2001), que avaliaram comparativamente as quatro classes de marcadores (*RAPDs*, *RFLPs*, *AFLPs* e *SSRs*) para estudos de divergência em linhagens tropicais de milho, os marcadores dominantes retiveram níveis de polimorfismo menores que os codominantes e, após *RFLPs*, os microssatélites foram os marcadores moleculares que se revelaram como muito promissores em conteúdo de informação e polimorfismo. Neste trabalho, o tamanho amostral avaliado para os marcadores microssatélites foi pequeno (68 locos para *SSR* versus 185 locos para *RFLP*), produzindo menor informatividade em relação aos *RFLPs*. Entretanto, se a saturação de microssatélites fosse ampliada, provavelmente sua informatividade se equivaleria ou suplantaria a dos *RFLPs*.

A maior limitação do método talvez seja o conhecimento e obtenção dos pares de *primers* que flanqueiam as repetições *em tandem*, sendo este processo muitas vezes de alto investimento (Chin *et al.*, 1996). Estas seqüências podem ser obtidas inicialmente através

de busca em banco de dados de seqüências de bases, para algumas espécies. Entretanto, somente através de um programa de desenvolvimento de microssatélites é que se é capaz de se obter um conjunto completo de *primers* para a espécie. Este processo envolve a geração de uma biblioteca genômica, utilizando-se fragmentos de 200 a 600 pares de bases, já selecionados previamente quanto à presença, neles, de repetições *em tandem*. A Biblioteca será “enriquecida” em clones capazes de gerar microssatélites. Os clones são hibridizados em condições de alta estringência, com sondas constituídas por *primers*, apresentando repetições encontradas nos microssatélites [(GA)₁₅, (GT)₁₅ ou (CT)₁₅]. Os clones positivos são inicialmente analisados através de amplificação seletiva para se deduzir a localização da repetição em relação às extremidades do fragmento clonado (Rafalski *et al.*, 1996). Os clones detectados devem ser seqüenciados. Este método para obtenção de microssatélites requer altos investimentos e infra-estrutura especializada de laboratório (Chin *et al.*, 1996; Souza, 2001), o que pode se configurar limitante para muitas culturas.

Em milho, há grande número de seqüências disponíveis no *site* do *Maize Data Bank* (<http://www.agron.missouri.edu>), passíveis de serem sintetizadas e utilizadas por laboratórios de todo o mundo, uma vez que esta é uma base de dados pública. Até julho de 2002, 1.855 microssatélites para milho já estavam disponíveis, dos quais 1.797 destes mapearam em locos diferentes. Este número tem sido expandido rapidamente. Isto facilita bastante o trabalho em relação a outras culturas onde se requerem investimentos iniciais para desenvolvimento dos microssatélites. O genoma de milho está ainda dissecado em 100 Bins, espaçados aproximadamente a cada 20 cM ao longo dos 10 cromossomos (Gardiner *et al.*, 1993).

A aplicação dos microssatélites em plantas é recente e muitas espécies já contam com *primers* desenvolvidos e mapeados, como soja (Rongwen *et al.*, 1995), arroz (Zhao &

Rochert, 1993), milho (Senior & Heun, 1993; Taramino & Tingey, 1996), feijão (Yu *et al.*, 2000), trigo (Röder *et al.*, 1998), cevada (Becker & Heun, 1995) e sorgo (Taramino *et al.*, 1997).

Apesar de muitas vantagens, os marcadores microssatélites apresentam ainda algumas limitações quanto a sua aplicação. Além da questão referente aos custos e dificuldades técnicas relacionadas acima quanto ao seu desenvolvimento, muitos, mesmo quando disponíveis, não podem ser utilizados. Sibov (2002) relataram que de 859 microssatélites escolhidos para amplificar suas amostras de DNA de milho, apenas 213 (25%) mostraram um padrão adequado de amplificação, havendo polimorfismo entre os genitores. Do restante, 251 (29%) não foram polimórficos, 113 (13%) falharam na amplificação mesmo depois de terem suas temperaturas de anelamento otimizadas e terem sido reamplificados. Outros 281 (33%) não tiveram resolução em agarose. Ainda, dos 213 microssatélites polimórficos, 119 (56%) puderam ser utilizados no mapeamento, pois segregaram de forma Mendeliana (1:2:1); 35 (16%) não apresentaram bandas claramente distinguíveis durante a genotipagem; 40 (19%) apresentaram distorção depois da genotipagem; e 19 (9%) apresentaram padrão não-Mendeliano de leitura com ausência de uma das bandas dos genitores em uma ou mais plantas F_1 testadas. Os autores relataram que cerca de 5,57 microssatélites foram testados para encontrar um marcador polimórfico sem problemas. Apesar de todo este esforço, algumas regiões não tiveram cobertura completa, criando “buracos” sobretudo nos cromossomos 4, 5, 9 e 10. Outros trabalhos também relatam dificuldades associadas com o uso dos microssatélites como desvios de segregação, ausência de alelos dos genitores da descendência genotipada, baixa resolução em agarose ou, até mesmo, ausência de ligação a qualquer grupo de ligação (Mangolim, 2002). Neste

trabalho de mapeamento de *QTLs* associados com o teor de óleo em linhagens tropicais de milho, alguns cromossomos apresentaram distâncias entre suas marcas superiores à 70cM (cromossomo 2 e 4).

2.3. Seleção Assistida por Marcadores Moleculares (SAM)

A seleção assistida por marcadores moleculares (SAM) foi originalmente proposta por Sax (1923) e Thoday (1961) e tem merecido grande destaque e entusiasmo nos programas de melhoramento atuais. Young (1999) verificou mais de 400 artigos listados pelo *Current Contents* de 1995 a 1999 e que portavam como palavra-chave o termo *marker-assisted breeding* ou *marker-assisted selection*. Poucos, entretanto, descreviam a liberação de variedades ou germoplasma tendo usado técnicas de seleção assistida por marcadores moleculares.

Em programas de retrocruzamento, muitas possibilidades surgiram com o advento dos marcadores moleculares para assessorar a introgressão gênica. Marcadores de DNA podem, por exemplo, ser utilizados na seleção indireta do alelo alvo do genitor doador a ser transferido; selecionar regiões não-ligadas à região introgridida de forma a acelerar o retorno ao genótipo do genitor recorrente e também selecionar regiões ligadas ao alelo alvo, contribuindo para a redução do *linkage drag* (Lee, 1995). A SAM oferece a possibilidade de analisar plantas em estágio precoce e mesmo de genotipar caracteres epistáticos.

Quando os marcadores moleculares são utilizados para monitorar a presença do alelo introgridido, denomina-se esta seleção de *foreground selection* (Tanksley, 1983; Melchinger, 1990; Hospital & Charcosset, 1997). Isto é particularmente interessante quando a avaliação fenotípica não é possível, ou é muito cara, ou ainda pode apenas ser

realizada tardiamente. Marcadores também podem ser utilizados para acelerar a recuperação do genótipo do recorrente em outros locos que não referentes àquele introgridido e isto é chamado de *background selection* (Hillel *et al.*, 1990, 1993; Hospital *et al.*, 1992). Estudos preliminares teóricos assumiram que o gene introgridido podia ser detectado sem ambigüidade, o que nem sempre é verdade, principalmente quando se trata de um *QTL*, cujos heterozigotos não podem ser distinguidos distintamente dos homozigotos.

Inúmeros trabalhos de simulação computacional têm surgido avaliando e sugerindo parâmetros para se realizar a introgressão assistida por marcadores. Estas simulações têm sido bem diversas. Por exemplo, SAM tem sido testada para combinar dados fenotípicos e genotípicos em um índice de seleção (Knapp, 1994; Xie & Xue, 1998) e considerando diferentes gerações de melhoramento (Edwards & Page, 1994; Xie & Xue, 1998). A eficiência de SAM tem sido avaliada considerando a herdabilidade do caráter alvo (Hospital *et al.*, 1997; Knapp, 1998), o efeito genético do loco alvo (Van Berloo & Stam, 1998) e o monitoramento de regiões genômicas, uma a uma (Hospital *et al.*, 1997). A maioria dos artigos teóricos relacionados a SAM apresentam modelos matemáticos complexos, tornando-se muitas vezes difícil delinear, na prática, um experimento relacionado à seleção assistida. Difícilmente são levadas em consideração, nestes artigos teóricos, as implicações do uso de diferentes estratégias de laboratório necessárias a se atingir a seleção assistida bem sucedida, como o tipo de marcador de DNA a ser utilizado, suas vantagens riscos e dificuldades (Ribault *et al.*, 2002).

Entre estes trabalhos teóricos, Hillel *et al.* (1990, 1993) propuseram que os *fingerprinting* de DNAs pudessem ser utilizados para introgressão de alelos em populações de retrocruzamentos selecionando-os a favor de *background* genético ou contra ele.

Usaram-se minissatélites, que não são adequados para estudos de introgressão, pois os seus padrões não podem ser mapeados e possuem uma não-aleatoriedade de distribuição no genoma. Estes autores também desconsideraram a recombinação ao redor dos marcadores.

Hospital *et al.* (1992) concluíram que, se marcadores moleculares forem usados, o genoma do genitor recorrente poderia ser recuperado em aproximadamente duas gerações de retrocruzamentos. Openshaw *et al.* (1994) determinaram o tamanho populacional e a densidade de marcador necessária para efetuar a recuperação do genótipo do genitor recorrente, utilizando o milho como modelo. Eles recomendaram o uso de quatro marcadores por cromossomo (de 200 cM de comprimento) e a seleção para recombinantes próximos ao alelo alvo.

Jarboe *et al.* (1994), em milho, também relataram redução no número de gerações de retrocruzamentos necessárias para a obtenção de indivíduos recuperados através de marcadores moleculares. Desta forma, por meio de simulação computacional, obtiveram, com três gerações de retrocruzamentos e 80 marcadores, uma recuperação de 99% do genitor recorrente.

Outro trabalho investigou por meio de simulações o ganho relativo em um programa de retrocruzamentos cuja seleção fosse feita apenas por marcadores, por características fenotípicas ou por um índice combinando marcadores e características fenotípicas (Visscher *et al.*, 1996). Os autores relataram que marcadores moleculares seriam eficientes em programas de retrocruzamento para simultaneamente introgridir um alelo e selecionar para o genótipo desejado. Marcadores espaçados de 10 a 20 cM geraram ganho de uma a duas gerações de retrocruzamentos, abreviando o total de gerações para completa recuperação do genótipo do genitor recorrente, em relação à seleção apenas fenotípica ou a ausência de seleção.

Diferentes estratégias de seleção assistida em programas de retrocruzamentos foram testadas, por meio de simulações, e demonstraram que, com o uso de marcadores moleculares, o grau de recuperação do genitor recorrente em RC₃ equivaleria ao que seria alcançado em RC₇ sem seleção para transferir um único alelo alvo (Frisch *et al.*, 1999b). Os autores ainda ressaltaram a importância de se elaborarem mapas de ligação plenamente saturados para obtenção de valores máximos de recuperação do genótipo do recorrente nas gerações RC₂ e RC₃. A maior eficácia de SAM, em relação ao melhoramento convencional, dependente da saturação do mapa de ligação, já havia sido ressaltada por Stuber & Edwards (1986).

Embora os resultados teóricos demonstrem que marcadores moleculares possam assistir programas de retrocruzamentos, muito poucos trabalhos experimentais relatam sucesso na introgressão de características de interesse por seleção assistida por marcadores moleculares. Ragot *et al.* (1995) utilizaram *RFLPs* para monitoramento da introgressão de um construto gênico (com genes para proteínas de resistência à fosfinotricina e à inseticida - gene *Bt*) em linhagens de milho através de retrocruzamentos e relataram grande progresso do genitor recorrente em RC₃. A linhagem doadora (Lancaster) recebeu o construto por bombardeamento com microprojétil e houve uma única inserção no cromossomo 01. O controle com inseticida foi feito a cada geração para eliminar as plantas que não portavam o gene incorporado no genoma. Na geração RC₁, 61 combinações sonda-enzimas foram utilizadas. Os autores calcularam a porcentagem de recuperação levando em conta todo o genoma, inclusive o gene *Bt* introgridido, e obtiveram para a primeira geração 50% de recuperação, abaixo da média esperada. Entretanto, para as outras gerações, encontraram médias acima da esperada, se nenhuma seleção fosse praticada.

Weck *et al.* (1991) analisaram experimentalmente 125 plantas RC₁ com sondas de *RFLP* para introgressão de um único alelo do gene de interesse. As 57 plantas heterozigotas para o caráter foram analisadas com 38 sondas. Três plantas com maior percentual de recuperação do genótipo do genitor recorrente foram escolhidas e retrocruzadas. As 58 plantas RC₂ heterozigotas para o caráter foram analisadas com as sondas que foram heterozigotas em RC₁. Concomitante às análises de laboratório, a geração RC₂ foi retrocruzada para gerar RC₃. A média populacional das plantas RC₂ selecionadas por *RFLP* foi próxima à média esperada para a população RC₃. Os autores afirmaram que duas gerações de retrocruzamentos assistidos por *RFLP* são capazes de recuperar o genótipo do genitor recorrente.

Utilizando marcadores ligados a *QTLs*, foi possível introgridir regiões específicas do genoma que conferem características quantitativas específicas desejáveis a uma variedade elite (Tanksley *et al.*, 1989; Paterson *et al.*, 1991; Dudley, 1993). Em tomate (*Lycopersicon esculentum* Mill.), linhagens têm sido criadas contendo *QTLs* da espécie selvagem *Lycopersicum hirsutum* Hub. & Bonpl. Estas linhagens ultrapassam a variedade elite original em produção, conteúdo de sólidos solúveis e coloração do fruto (Tanksley & McCouch, 1997). Este resultado foi obtido com o uso da metodologia desenvolvida por Tanksley & Nelson (1996), de *advanced backcross QTL analysis*, ou método de retrocruzamentos avançados, no qual a população é conduzida através de retrocruzamentos, mapeamento e seleção assistida concomitantemente.

Em um trabalho prévio, *QTLs* presentes em linhagens de milho oriundas de grupos heteróticos distintos foram detectados e mapeados (Stuber *et al.*, 1992). Dando continuidade, em um outro trabalho, estes fragmentos foram transferidos para duas linhagens alvo formadoras de híbridos elites. Isto foi feito em três gerações de

retrocruzamentos e as linhagens RC₃ resultantes dos dois programas de retrocruzamentos foram cruzadas. Pelo menos um marcador foi identificado em cada braço cromossômico, além de dois marcadores adicionais, flanqueando o segmento doador a ser transferido. Apenas os marcadores não-fixados foram avaliados na geração subsequente. A avaliação dos híbridos melhorados mostrou que houve variação para produtividade de grãos, tendo sido possível selecionar híbridos superiores. Interessante é que, neste trabalho, uma característica quantitativa (rendimento de grãos) foi tratada e manipulada tal qual se fosse um caráter monogênico. Entretanto, detalhes de como foi feita esta introgressão, como o número de marcas utilizadas ou porcentagem de recuperação do genótipo do recorrente, não foram fornecidos (Stuber *et al.*, 1994).

O gene *sucr* foi introgridido do tomate *Lycopersicon chmielewskii* para o genótipo do tomate *L. esculentum* cv. 'Hunt 100' por seleção assistida por marcadores, usando marcadores *RFLPs* proximais e também um cDNA para enzima ácido-invertase de tomate, como sondas do gene *sucr* (Chetelat *et al.*, 1995). Na geração RC₁F₂, o segmento contendo o gene *sucr* foi composto de 43,2 cM (1/3 do comprimento do cromossomo 3). Após seleção para recombinações entre o gene alvo e marcadores adjacentes, o fragmento introgridido foi reduzido a 0,8-7,1 cM, na geração RC₅.

Outro exemplo bem sucedido de seleção assistida por marcadores moleculares é o de introdução do gene *Xa21* de resistência à BB (*Bacterial Blight*) em uma linhagem restauradora de arroz, a "Minghui 63", o qual foi feito com sucesso através de três gerações de retrocruzamentos assistidos, seguido de uma geração de autofecundação (Chen *et al.*, 2000). Um sistema baseado em *PCR* foi estabelecido consistindo em um marcador contendo uma parte do gene *Xa21*, uma marca localizada a 0,8 cM do loco *Xa21* de um lado e uma outra marca distando 3,0 cM do gene do outro lado. Um total de 128 locos de

RFLPs, distribuídos ao longo dos 12 cromossomos, foram analisados e utilizados para selecionar a recuperação do genótipo do genitor recorrente. A “Minghui 63 (*Xa21*)” obtida foi exatamente idêntica à original, exceto pela presença de um fragmento menor que 3,8 cM (<0,21%) de comprimento envolvendo o loco *Xa21* do genitor doador, mantendo o resto do genótipo do genitor recorrente intacto. A linhagem recuperada “Minghui 63 (*Xa21*)” e o seu híbrido F₁, “Shanyou 63 (*Xa21*)”, produziram o mesmo espectro de resistência do genitor doador. Também em relação ao desempenho, a linhagem recuperada e o seu híbrido mostraram desempenho idêntico ao da linhagem original “Minghui 63” e ao híbrido “Shanyou 63” e, sob condições de doença, atingiu um potencial produtivo superior aos originais.

Plantas RC₁, RC₂ e RC₃ selecionadas para resistência ao fungo *Colletotrichum lindemuthianum* foram avaliadas por meio de marcadores *RAPD* (Hagiwara *et al.*, 2001). A diversidade genética entre os genótipos foi avaliada bem como o parentesco das populações RC₁F₄ e RC₂F₁ para com os genitores, de acordo com a genealogia delas. Os autores confirmaram a utilidade desta classe de marcadores para auxiliar na seleção de plantas mais semelhantes ao genitor recorrente, nas populações segregantes, mesmo em RC₁.

A redução do *linkage drag* é um dos maiores objetivos da seleção assistida por retrocruzamentos (Tanksley *et al.*, 1989). Estudos teóricos mostraram que porções ligadas ao alelo alvo permanecem em gerações avançadas de retrocruzamentos, sem seleção assistida por marcadores. No caso de alelos alvo introgrididos a partir de germoplasma não adaptado, o *linkage drag* é muitas vezes a causa principal das diferenças entre a linhagem recorrente e convertida (Frisch *et al.*, 1999a).

Para um único loco, as frequências esperadas para os indivíduos serem homozigotos para o genótipo do genitor recorrente seriam 0,5, 0,75 e 0,875, em RC₁, RC₂ e RC₃,

respectivamente. Estas são probabilidades de os locos serem homozigóticos para alelos do genitor recorrente. As variâncias destas proporções, em cada geração, seriam $0,5 (1-0,5)/n$, $0,75 (1-0,75)/n$, $0,875 (1-0,875)/n$, respectivamente, onde n é o número de unidades de recombinação independentes no genoma. A variância presente em RC_1 é muito maior do que aquela presente nas gerações subsequentes, o que indica uma freqüência de distribuição em RC_1 mais ampla dos genótipos do genitor recorrente do que em gerações mais tardias. Sendo assim, espera-se que um indivíduo com recombinação entre o gene alvo e os marcadores flangeadores ocorra em uma freqüência muito maior na geração RC_1 do que em RC_2 , sendo possível de se realizar a recuperação do genótipo do genitor recorrente em RC_1 (Chen *et al.*, 2000).

Uma das maiores limitações da aplicação da seleção assistida é seu alto custo, muitas vezes associado ao tipo de marcador e delineamento utilizados. Com o desenvolvimento de marcadores baseados em *PCR*, é possível se mapear e localizar qualquer caráter de maneira mais eficiente e menos custosa (Kumar *et al.*, 1999).

Luby & Shaw (2001) avaliaram a eficiência e custos comparativos de SAM para caracteres monogênicos e poligênicos em relação a esquemas de melhoramento convencionais, com seleção fenotípica, para programas de melhoramento de frutíferas, e concluíram que SAM pode representar uma ferramenta de seleção de indivíduos substancialmente superiores comparados com o melhoramento convencional. Esta visão otimista também é evidenciada por Young (1999), que atentou para o fato de a tecnologia de seleção assistida necessitar de um melhor planejamento e execução no que tange a genotipagem dos indivíduos, ao tamanho populacional, a uma análise genética quantitativa apropriada e a conteúdos genéticos diferenciados. Uma questão interessante a ser abordada é que a maioria dos caracteres agronomicamente importantes são quantitativos e

influenciados por numerosos locos espalhados no genoma, com efeitos pequenos e individuais (Lande, 1981). Esforços crescentes vêm sendo aplicados para maximizar a eficiência do mapeamento de *QTL* conjuntamente com o processo de SAM. Uma vez mapeados, os *QTLs* podem ser manipulados e transferidos. Entretanto, as teorias de genética quantitativa aplicada a SAM têm sugerido que a eficiência da seleção assistida por marcadores seja inversamente proporcional a herdabilidade do caráter (Lande & Thompson, 1990; Paterson *et al.*, 1991). Este pressuposto também foi comprovado por Schneider *et al.* (1997), utilizando seleção assistida por RAPD para melhorar a resistência à seca em feijão comum (*Phaseolus vulgaris* L.).

Em relação a germoplasma exótico, há várias características interessantes (qualitativas e quantitativas) que poderiam ser incorporadas em híbridos elites e transferidas por SAM. Com os avanços da técnica de *PCR*, novas tecnologias, tais como novos tipos de marcadores moleculares, como *SNPs* (*Single Nucleotide Polychips de DNA* - Gilles *et al.*, 1999), ou o uso de *microarrays* (Lemieux *et al.*, 1998; Winzeler *et al.*, 1998), ou ainda o perfil de expressão de DNA, podem ser adotadas. Os melhoristas precisam apenas reconhecer a importância de mudanças no paradigma e usar as melhores ferramentas possíveis e, desta integração do convencional e molecular, novos ganhos genéticos poderão ser atingidos.

Desta forma, a integração de estratégias genômicas ou de bioinformática e os avanços na tecnologia dos marcadores de DNA levarão a uma profunda revolução no melhoramento de plantas de forma a minimizar os custos e viabilizar a aplicação de SAM a projetos específicos, a qual tem se mostrado superior ao melhoramento convencional, quando este é baseado apenas no fenótipo. Não se espera que marcadores moleculares

suplantem o melhoramento convencional, mas que ampliem a eficiência da seleção, promovendo aumento dos ganhos genéticos.

3. OBJETIVOS

3. OBJETIVOS

- a) Investigar a viabilidade do uso de marcadores moleculares do tipo microssatélites para assessorar a introgressão de uma marca molecular, através de três ciclos de retrocruzamentos;
- b) Selecionar o percentual de recuperação do genótipo do genitor recorrente com base em três critérios de seleção molecular: (a) duas marcas por cromossomo (CS_1); (b) quatro marcas por cromossomo (CS_2) e (c) todas as marcas polimórficas disponíveis, o que representou aproximadamente seis marcas por cromossomo (CS_3);
- c) Comparar as médias de recuperação obtidas para cada geração com seleção puramente molecular, dentro de cada critério de seleção, com a média de recuperação esperada para três ciclos de retrocruzamentos na ausência de seleção e de *linkage drag*, e também com a média de 100% de recuperação;
- d) Verificar, nas plantas RC_3 selecionadas, como estão as demais regiões não aferidas por microssatélites durante o processo de seleção assistida.
- e) Comparar fenotipicamente as linhagens RC_3 obtidas entre si e também com o genitor recorrente.

4. ARTIGO

4.1. Marker-Assisted Selection Using Microsatellites In Maize Backcross Breeding Programs

L.L. Benchimol^{1,2*}, C.L. Souza Jr.³ and A.P. Souza^{1,2}.

¹Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas (UNICAMP), P.O. 6010, 13083-970, Campinas, SP, Brazil

²Departamento de Genética e Evolução – Instituto de Biologia (IB), UNICAMP, 13083-970, Campinas, SP, Brazil

³Departamento de Genética, Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), 13400-970, P.O. 83, Piracicaba, SP, Brazil

*To whom correspondence should be addressed: Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas (CBMEG/UNICAMP), Cidade Universitária “Zeferino Vaz”, Cep 13083-970, C.P. 6010, Campinas, SP, Brazil.

E-mail: llasry@unicamp.br

FAX: (55-19) 3788-1089

Phone: (55-19) 3788-1132

Abstract

Marker-assisted selection (MAS) has received great attention and many are the efforts to establish the best strategy for using marker-assisted selection as a routine in backcross breeding programs. Although much computer simulation has evaluated the parameters that influence marker-assisted selection efficiency, few experimental data report successful results of introgression of desirable traits and estimate the necessary numbers of markers needed for a complete recurrent genotype recovery.

Microsatellites were used to introgress a single-locus trait from a L-08-05 donor line to a L-14-4B recurrent divergent inbred line. A clear reading pattern and polymorphic microsatellite, PHI037 (Bin 1.08), was chosen to be transferred, simulating an oligogenic trait or a QTL linked to it. Three generations of backcross were conducted. Plants were genotyped by molecular approach, selected, and at the same time, backcrossed. Three selection criteria were established to estimate the recoveries of the recurrent parent genotypes: a) 02 (SC_1) markers, b) 04 (SC_2) markers, c) 06 (SC_3) markers per chromosome. At each backcross cycle, plants were selected based on the target allele whose *locus* was in a heterozygous state, and on the greatest recovery of the recurrent genotype. Only the non-recovered microsatellites, from the previous cycle, were used to genotype BC samples in the next generation. In BC_1 , fifty-three polymorphic microsatellites were used to genotype the 186 plants and four plants were chosen based on each criterion. In BC_2 , 60 plants of each criterion (180 plants in total) were genotyped and four plants of each progeny within each criterion were selected. In BC_3 , 160 plants per criterion (480 plants in total) were genotyped. Forty-eight BC_3 lines (16 from each criterion) were selected from the BC_3 lines and genotyped with those microsatellites used only in the other criteria (for SC_1 and SC_2) and also with another additional 19 microsatellites distributed in Bins not assayed previously. This procedure was carried out with the aim of saturating maize genome and comparing the estimates of the recurrent parent genotype recoveries, under

saturated and non-saturated conditions, for each criterion. *t*-Tests for comparison between BC₃ line (saturated and non-saturated) means for the expected mean (93.75%) of this generation and for 100% recovery were computed. The 48 selected BC₃ lines were evaluated in a field experiment and the recurrent L-14-4B inbred line was also included. Individual analyses of variance for the nine traits and multivariate analyses of variance for three composite traits (yield, height and flowering) were performed. Comparison between the 48 selected BC₃ lines and the recurrent parent (L-14-4B) were performed using the *Dunnnett* test.

For all selection criteria, except from the recoveries for the selected lines for SC₁, under saturated and non-saturated conditions, BC₃ line averages were significantly superior ($P \leq 0.01$) to the expected value (93.75%) and significantly inferior ($P \leq 0.01$) to the maximum expected value (100%). The increase of the number of markers contributed towards a reduction in the donor fragment content. SC₃ criterion also showed smaller differences between saturated and non-saturated means and standard deviations for their selected lines. According to analyses of variance for the nine traits, only two of them presented low significance ($P \leq 0.05$). No significance was detected for selection criteria by multivariate analyses of variance. The *Dunnnett* test produced 120 tests for each selection criterion and 13, 14 and 8 were significant for SC₁, SC₂, SC₃, respectively. Phenotypically, the higher amount of markers produced a superior number of BC₃ lines that did not differ statistically from the recurrent parent. Gathering molecular and phenotypic results, it can be concluded that PHI037 allele was efficiently transferred and that the SC₃ criterion, with approximately 06 markers per chromosome, was superior to perform marker-assisted backcross selection.

Key words: MAS, marker-assisted backcrossing, marker-assisted selection, microsatellites, backcross

Introduction

Important economic traits have been introduced into elite maize inbred lines by backcross breeding. This method is used to improve a characteristic of a cultivar which is deficient (Fehr 1987). Usually, the trait to be transferred is controlled by a single gene, however more complexly inherited traits have been transferred successfully by this method (Rinke and Sentz 1961; Shaver 1976). A donor inbred line (DP) carrying a trait of interest is crossed with the recurrent parent (RP), an elite line that is lacking the trait (Openshaw 1994). Backcrossing of the selected plants with the RP can be repeated during each cycle until a version of the RP including the introgressed allele has been obtained.

Backcrossing approach has been usually successful but presents certain limitations such as the number of generations to achieve a complete recovery of the genetic background of the RP, which depending on the crop, may be time consuming. The expected fraction of the genome from the recurrent parent in the n^{th} backcross generation is calculated as $1 - (\frac{1}{2})^{n+1}$. However, this expression does not take into account the occurrence of recombination, which leads to a variation in this average value. Thus, in a backcross population the amount of the genome of the recurrent parent can vary considerably among plants. Another problem in backcross breeding programs is the presence of the donor's genetic portions surrounding the allele that is being introgressed, which is termed "linkage drag".

With the development of molecular markers, high-density linkage maps have been available in most crop species, opening many possibilities to backcross introgression. Molecular markers could be used to assess the presence of the introgressed gene, i. e., the

“foreground selection” (Tanksley 1983; Melchinger 1990; Hospital and Charcosset 1997) and/or to accelerate the recovery of the recurrent parent genotype at other *loci*, i. e., the “background selection” (Tanksley 1983; Young and Tanksley 1989). Tanksley et al. (1989) have proposed an approach for the selection of individuals that carry the target gene and are also homozygous for the recurrent parent alleles, in a large proportion of markers. This background selection would be employed with the aim of accelerating recovery of the recurrent parent genome and reducing the number of backcross generations required. DNA markers can aid in the selection of the maximum similarity for the recurrent inbred line and the minimum similarity for the donor inbred line (Hillel et al. 1990).

Many theoretical studies (Hospital et al. 1992; Visscher et al. 1996; Frisch and Mechinger 2001; Hospital 2001) have evaluated different parameters involved; however there are still few experimental and practical examples evaluating and testing the necessary parameters for an efficient marker-assisted selection. Planning a marker-assisted backcross program, includes considering the time advantage to assist backcrossing, the number of markers needed and the number of genotypes to be evaluated (Openshaw et al. 1994). Computer simulations suggest that recurrent parent genotype can be recovered in three generations by molecular markers, for background selection (Tanksley et al. 1989). Simulations carried out by Hospital et al. (1992), for instance, have indicated that four markers per 200 cM chromosome would be adequate to greatly increase the effectiveness of selection in BC₁. Openshaw et al. (1994) supported these results. They also recommended that preference should be given to the selection for proximal recombinants to the allele of interest without forgetting the selection of recovery of the recurrent parent elsewhere in the genome.

Weck et al. (1991) have stated that RFLP-assisted selection could increase the gain over phenotypic selection; using 125 plants in BC₁ and a single gene of interest, and after two backcross generations, 97.4% of recovery of the recurrent parent genotype could be achieved. Stuber (1994) used marker-facilitated introgression to transfer identified QTLs into the target divergent lines, in three backcross generations. After the third backcross, the resulting lines were crossed, generating enhanced hybrids which were evaluated under field conditions and compared to a check hybrid derived from the original parental lines. Both crosses produced a proportion of hybrids that yielded more than the checks, demonstrating that marker-facilitated backcrossing could be successfully employed to manipulate complexly inherited traits in maize, such as grain yield.

The *sucr* gene was introgressed from *Lycopersicon chmielewskii* into a cultivated tomato, *L. esculentum* cv. 'Hunt 100', by marker-assisted selection using tightly linked RFLP markers and a tomato acid invertase cDNA as probes for *sucr* (Chetelat et al. 1995). Sanchez et al. (2000) have used sequence tagged site marker-assisted selection via backcrossing to improve the resistance of three rice lines in *Xanthomonas oryzae* pv. *oryzae*. Three bacterial blight resistance genes were successfully transferred to these lines by MAS, demonstrating the usefulness of this strategy for pyramiding in terms of resistance recessive genes. Chen et al. (2000) introgressed a gene that conferred resistance to a broad spectrum of pathogen races into an elite restorer line of a hybrid rice using a PCR-based MAS system. A total of 128 RFLPs were used to recover the genetic background of the hybrid rice.

Marker assisted selection with microsatellites was implemented to assess the feasibility of using it in maize breeding programs. Thus, three objectives of this research were: (1) to evaluate the number of markers per chromosome to be used for background

selection, (2) to compare the selected lines with the recurrent parent inbred line after three generations of backcross marker-assisted selection.

Material and Methods

Plant Material and Backcrossing Generations

The L-08-05 ($S_7 - P_1$) inbred line, derived from the IG-1 population was used as donor parent and crossed with the L-14-4B ($S_5 - P_2$) inbred line, used as a recurrent parent, and derived from the BR-106 population. L-08-05 is early-maturing with orange flint kernels and L-14-4B is early maturing with yellow dent kernels. These populations were allocated into different heterotic groups (personal communication).

A unique F_1 ear was chosen from $P_1 \times P_2$ cross. BC_1 generation resulted from the backcrossing of F_1 with the recurrent parent. A hundred and eighty six BC_1 plants were genotyped and four plants were selected from each selection criterion (SC_1 , SC_2 , SC_3), totalizing 12 progenies. BC_2 generation was derived from the cross of BC_1 selected plants to P_1 . Fifteen plants per progeny/selection criterion were genotyped (a total of 60 plants/selection criterion and 180 plants genotyped). Four plants were selected from each criterion/progeny, performing a total of 16 plants per selection criterion and 48 BC_3 progenies selected.

The cross of $BC_2 \times P_1$ generated the BC_3 generation. Ten plants from each progeny/selection criterion were genotyped making up 160 plants per selection criterion and 480 plants in general. Following the same strategy described for the previous cycles, four plants were selected per progeny/selection criterion, making up a total of 16 plants per selection criterion, and 48 lines selected.

Molecular Selection Approach

Molecular approach for selecting plants in each backcross generation was based on a three selection criteria: (a) two (SC₁) markers per chromosome, (b) four (SC₂) markers per chromosome and (c) six (SC₃) markers per chromosome, approximately (53 polymorphic microsatellites available in BC₁). Those markers were chosen for their clear polymorphic profile and distribution along the chromosomes. Thus, i. e., for SC₁, one marker per chromosome arm was chosen, maintaining a certain distance from the centromere to increase recombination and decrease the “linkage drag”; for SC₂, two markers per chromosome arm were used, and for SC₃, approximately three markers per chromosome arm (Table1). However, this average in SC₃ criteria was not achieved for all chromosomes due to the difficulties in obtaining polymorphic markers. At each generation, selection was focused on plants that gathered favorable genetic recurrent inbred line background and maintained the specific target locus in a heterozygous state. The comparison of the percent of recovery of the recurrent parent genotype and the effectiveness of the increase in the number of markers per chromosome used during the selection process were done based on the molecular criteria.

A clear reading polymorphic microsatellite, **PHI037** (Bin 1.08), was elected and used as a target locus to be introgressed. Its recovery was monitored throughout all the three backcross generations. Apart from its reliable pattern, PHI037 also presented clear polymorphic flanking microsatellites (PHI011, at Bin 1.09 for SC₁/SC₂; BNLG615, at Bin 1.07 for SC₃) that could contribute to diminish the “linkage drag” around the locus. Every selected plant in any backcross generation would need to have, in order of importance, the following conditions: (1) the introgressed marker (PHI037) in a heterozygous state; (2) the

highest percent of recovery of the recurrent parent genotype, (3) no donor allele present in the genotyping of the individual plant. A plant that did not have PHI037 in a heterozygous state would not be selected, even if it had achieved a superior recovery of the recurrent parent genotype. This condition was generated by the combined need for introgression of the target locus and maximum recovery of the genetic background of the recurrent parent.

DNA Extraction and Microsatellite Analysis

Leaf material from parents, BC₁, BC₂ and BC₃ generations were collected, freeze-dried, ground to powder and stored at -20°C in individually labeled vials. DNA extraction followed the procedures described in Hoisington et al. (1994). Fifty-three polymorphic microsatellites were selected among 392 primer pairs to amplify maize microsatellites and were obtained from the Maize Data Bank (2001). Gibco BRL synthesized them (São Paulo, SP). When available, tetranucleotides repeats or even a greater number of repeats, were preferably selected to relate to di- or trinucleotide ones. Independently from the repeat type, microsatellite *loci* with higher repeat counts and perfect motif were the first choice.

By means of agarose gels, the primer pairs were screened against the parents and the F₁ plants. Microsatellites showing the clearest polymorphism among the parental lines were chosen for genotyping backcrossing generations. *Locus* designation, chromosome location, primer sequence information, description of microsatellite motif, and size of PCR product amplified in reference lines, can be found in the Maize Data Bank (2001).

Each amplification was carried out containing the DNA from donor, recurrent and F₁, apart from the backcrossed samples. Only the non-recovered microsatellites at BC₁ were used to screen BC₂. This approach was adopted up to the final BC₃ generation. The 48

(BC_{3sel}) BC₃ selected plants were analyzed with extra primers not used in the previous generation of selection with the aim of saturating the genome up to 07 markers, in average, per chromosome (19 additional primers). Apart from SC₃, SC₁ and SC₂ plants were also screened with primers used in genotyping of selected plants from other criterion. By this procedure, an amount of 72 microsatellites distributed over the ten chromosomes was reached (Table 1). Recurrent parent recovery was recalculated for this generation (BC_{3sel}^{sat}). Although the same genetic background constituted BC_{3sel}^{sat} and BC_{3sel}^{non-sat}, the estimates of the genetic content associated with the recurrent parent recoveries were different, regarding genome saturation.

Table 1. Primers used in each selection criterion and their respective Bins. *Sat* (saturated) notation refers to the extra-primer set used to amplify the 48 selected BC₃ plants and saturate maize genome. SC is the selection criterion used.

Primers	SC	Bin	Primers	SC	Bin
BNLG 1627	SC ₁ /SC ₂ /SC ₃	1.02	PHI 126	SC ₁ /SC ₂ /SC ₃	6.00
BNLG 2238	SC ₂ /SC ₃	1.04	PHI 077	SC ₂ /SC ₃	6.01
BNLG 1598	SC ₃	1.06	BNLG 1371	SC ₂ /SC ₃	6.02
BNLG 615	SC ₃	1.07	UMC 1614	<i>Sat</i>	6.04
PHI 037*	<i>All</i>	1.08	NC013	SC ₁ /SC ₂ /SC ₃	6.05
PHI 011	SC ₁ /SC ₂ /SC ₃	1.09	UMC1520	<i>Sat</i>	6.06
PHI 120	SC ₂ /SC ₃	1.11	BNLG 1759	SC ₃	6.07
BNLG 125	SC ₂ /SC ₃	2.02	UMC 1426	<i>Sat</i>	7.00
BNLG 2248	SC ₃	2.03	UMC 1632	<i>Sat</i>	7.01
BNLG 166	SC ₁ /SC ₂ /SC ₃	2.04	PHI 114	SC ₂ /SC ₃	7.03
DUPSSR 21	SC ₃	2.05	BNLG 434	SC ₁ /SC ₂ /SC ₃	7.03
BNLG 1396	SC ₃	2.06	DUPSSR 13	SC ₂ /SC ₃	7.04
UMC 1560	<i>Sat</i>	2.07	UMC 1407	<i>Sat</i>	7.05
BNLG1520	SC ₁ /SC ₂ /SC ₃	2.09	PHI 116	SC ₂ /SC ₃	7.06
UMC1394	<i>Sat</i>	3.01	UMC 1139	<i>Sat</i>	8.01
BNLG 1144	SC ₁ /SC ₂ /SC ₃	3.03	PHI 119	SC ₁ /SC ₂ /SC ₃	8.02
BNLG 602	SC ₃	3.04	PHI 115	SC ₃	8.03
MMC 0022	SC ₂ /SC ₃	3.05	BNLG 1863	SC ₂ /SC ₃	8.04
BNLG 420	SC ₂ /SC ₃	3.05	BNLG 1176	SC ₃	8.05
UMC 1659	<i>Sat</i>	3.07	BNLG 1823	SC ₁ /SC ₂ /SC ₃	8.07
BNLG 1536	<i>Sat</i>	3.09	DUPSSR14	SC ₂ /SC ₃	8.09
BNLG 1098	SC ₁ /SC ₂ /SC ₃	3.10	DUPSSR 06	SC ₁ /SC ₂ /SC ₃	9.02
UMC 1276	<i>Sat</i>	4.01	BNLG 430	SC ₂ /SC ₃	9.03
UMC 1652	<i>Sat</i>	4.04	UMC 1107	<i>Sat</i>	9.04
PHI 026	SC ₁ /SC ₂ /SC ₃	4.05	BNLG 1012	SC ₂ /SC ₃	9.05
BNLG 252	SC ₃	4.06	BNLG 292	SC ₃	9.06
DUPSSR 34	SC ₂ /SC ₃	4.07	BNLG 128	SC ₁ /SC ₂ /SC ₃	9.07
BNLG 2244	<i>Sat</i>	4.08	BNLG 619	<i>Sat</i>	9.07-9.08
BNLG 1917	SC ₁ /SC ₂ /SC ₃	4.10	PHI 117	<i>Sat</i>	10.00
BNLG 589	SC ₂ /SC ₃	4.11	UMC 1318	<i>Sat</i>	10.01
BNLG 1006	SC ₁ /SC ₂ /SC ₃	5.00	BNLG1451	SC ₁ /SC ₂ /SC ₃	10.02
PHI 113	SC ₂ /SC ₃	5.03-5.04	BNLG 640	SC ₂ /SC ₃	10.03
DUPSSR 10	SC ₂ /SC ₃	5.04	BNLG 1526	SC ₂ /SC ₃	10.04
BNLG 278	SC ₃	5.05-5.06	UMC 1930	<i>Sat</i>	10.05
UMC 1019	SC ₁ /SC ₂ /SC ₃	5.06	BNLG 1839	SC ₁ /SC ₂ /SC ₃	10.07
PHI 128	SC ₃	5.07			
UMC 1153	<i>Sat</i>	5.09			

*PHI037 was transferred as a microsatellite allele for all selection criteria.

Amplification and Detection Conditions

The 20µl reactions contained 10 ng of genomic DNA, 0.2µM of each primer, 2 mM of magnesium chloride, 100 µM of each dNTP, 0.5U of *Taq* DNA polymerase (Gibco BRL) and 1x of Gibco-BRL buffer. Reactions were carried out in 96-well microtitier plates using a PTC-100 Thermal Cycler (MJ Research, Watertown, MA), and an adaptation of the PCR protocol called Touchdown (Don et al. 1991; Chin et al. 1996). When the pairs of primers failed to amplify using the “touchdown” program, the amplifications were performed with low stringency annealing temperatures.

Products were separated by electrophoresis on a 4% metaphor:agarose gel, using TBE buffer 1x (5x: 89mM Tris-HCL pH 8.0, 89 mM Boric Acid, 2mM EDTA dissodic), over a 2:15 hour-run under 100V. The gel was stained with ethidium bromide and visualized under UV. The visual score of the amplifications was done directly over the UV apparatus, always comparing with P₁, P₂ and F₁ patterns. Gel was photographed and the photo kept as a register.

Molecular Data Analysis

Results were scored as A for P₁ allele (GD), B for the P₂ allele (GR) and H for the F₁ pattern, composed by P₁ and P₂ alleles. The percentage of the recurrent parent genome present in each genotyped plant was estimated by the following expression: $GR\% = [B + (0.5H) / (B + H + A)] \cdot 100$. This percentage corresponds to the number of B alleles present in the genotype of each genotyped plant. The introgressed locus, PHI037, was taken out from the computation of the total recoveries of the recurrent parent genotypes.

BC₁ and BC₂ percent recovery means were compared to the expected averages for each backcross generation. The 48 selected BC₃ plants were analyzed considering saturating (BC_{3sel}^{sat}) and non-saturating (BC_{3sel}^{non-sat}) conditions. Data were transformed by arcsin function in order to compare the percentages of recovery between the selected plants. *t*-Tests of comparisons were performed for the percent recoveries of the recurrent parent for BC₃ lines, under saturated and non-saturated conditions, to the expected average of 93.75%, and the saturated average of 100%, for each selection criterion.

Experimental Evaluation and Statistical Analysis

The 48 BC₃ selected lines were evaluated in the Experimental Station at the Department of Genetics, Piracicaba, in the state of São Paulo. Plots were 6m long, spaced 0.80 m between rows, and 0.20 m between plants within plots, with 30 plants per plot. Ten plots containing the recurrent parent (L-14-04B) were included among the BC₃ lines plots. In the BC₃ lines plots, phenotypic data was recorded from ten competitive plants. In the L-14-04B plots, data from 05 competitive plants was taken from each plot. Data was recorded on the following phenotypic traits: ear weight (EW), number of kernel-rows per ear (NKE), number of rows per ear (NRE), ear length (EL), ear diameter (ED), plant height (PH), ear height (EH), days to silk extrusion (DSE), days to pollen shed (DPS). DSE and DPS were taken as one record per plot each. DSE refers to the days after planting date until 50% of the plants in a plot had visible silks extruded and DPS refers to the days after the planting date until 50% of the plants in a plot were shedding pollen.

Hierarchical analyses of variance were performed, and the sources of variation were partitioned into two main effects: Selection Criteria (SC) and BC₃ Lines (L)/SC. This latter

level was partitioned into L/SC₁, L/SC₂ and L/SC₃. Multivariate analyses of variance were performed for three composite traits: yield, height and flowering. Yield was composed by ear weight, number of kernel rows per ear, number of rows per ear, ear length and ear diameter; height was composed by plant and ear heights; flowering was composed by days to silk extrusion and days to pollen shed, with the sources of variations partitioned as for the individual analyses of variance, and significances were obtained by Hotelling-Lawley Trace *F*-test. *Dunnett* bilateral test was applied to compare the 48 BC₃ lines to the recurrent parent (L-14-4B). All these analyses were performed using SAS software (SAS Institute 2001). For seven traits (EW, NKE, NRE, EL, ED, PH, EH), the 48 selected lines were compared; but, for two traits (DSE and DPS), as they had just one record per plot each, they were grouped by their BC₂ origin and 04 tests were performed for each criterion.

Results and Discussion

The three molecular criteria showed an average of 74.32% of recovery of the recurrent parent genotype for the first backcross generation. This average did not differ from the expected one (75%), little variation was detected among the three criteria, and the maximum value was 87.50%. The average recovered value of the four selected plants in BC₁ generation was 83.75% for SC₁, 83.13% for SC₂ and 82.69% for SC₃ (Figure 1a, 1b and 1c). In the BC₂ generation, the three criteria showed an average of 91.16% and maximum value of 98.08%. The mean expected value for this generation, without selection, was 87.5%. The average recovered value of the 16 selected plants in the BC₂ generation was 94.38% for SC₁, 92.50% for SC₂, 92.05% for SC₃ (Figure 1d, 1e and 1f). In the BC₃ generation, the three criteria showed an average of 96.61% which was superior to the

expected mean for this generation (93.75%). The maximum value was 100% (without considering the PHI037 allele). The average recovered value of the 16 selected plants in the BC₃ generation was 97.25% for SC₁, 96.23% for SC₂ and 96.35% for SC₃ (Figure 1g, 1h and 1i). These averages were compatible to the ones expected in BC₄ or BC₅, indicating a genetic gain in recurrent parent recovery due to marker-assisted selection.

Hospital et al. (1992), using computer simulations, described a shortening in the number of backcross generations, from seven to two, to achieve a complete recovery of the recurrent genotype using MAS. Weck et al. (1991) reported a complete recovery of the recurrent genotype in only two RFLP-assisted selections.

The 16 selected BC₃ lines in the three selection criteria were saturated with more markers, and the mean values of the recurrent parent recovery showed that the averages for the selected saturated BC₃ lines were significantly lower ($P \leq 0.01$) for SC₁ and SC₂, but for SC₃, there were no significant differences. Indeed, there was a great decrease in the recurrent parent genotype recovery for SC₁ (99.84% to 92.40%) and for SC₂ (99.06% to 95.50%) with genome saturation. The increase in the number of markers also reduced significantly ($P \leq 0.01$) the difference between the averages and standard deviations for BC_{3sel}^{non-sat} and BC_{3sel}^{sat}. All selection criteria, except from BC_{3sel}^{non-sat} and BC_{3sel}^{sat} from SC₁, were significantly superior ($P \leq 0.01$) to the expected mean value (93.75%), and significantly inferior ($P \leq 0.01$) to the maximum value (100%). For SC₁, BC_{3sel}^{non-sat} did not differ from the maximum value (100%) whereas BC_{3sel}^{sat} was significantly inferior to the expected mean value (93.75%), for this generation. It should be emphasized that CS₃ criterion differed significantly from CS₂ and CS₁ in BC_{3sel}^{sat} (Table 2).

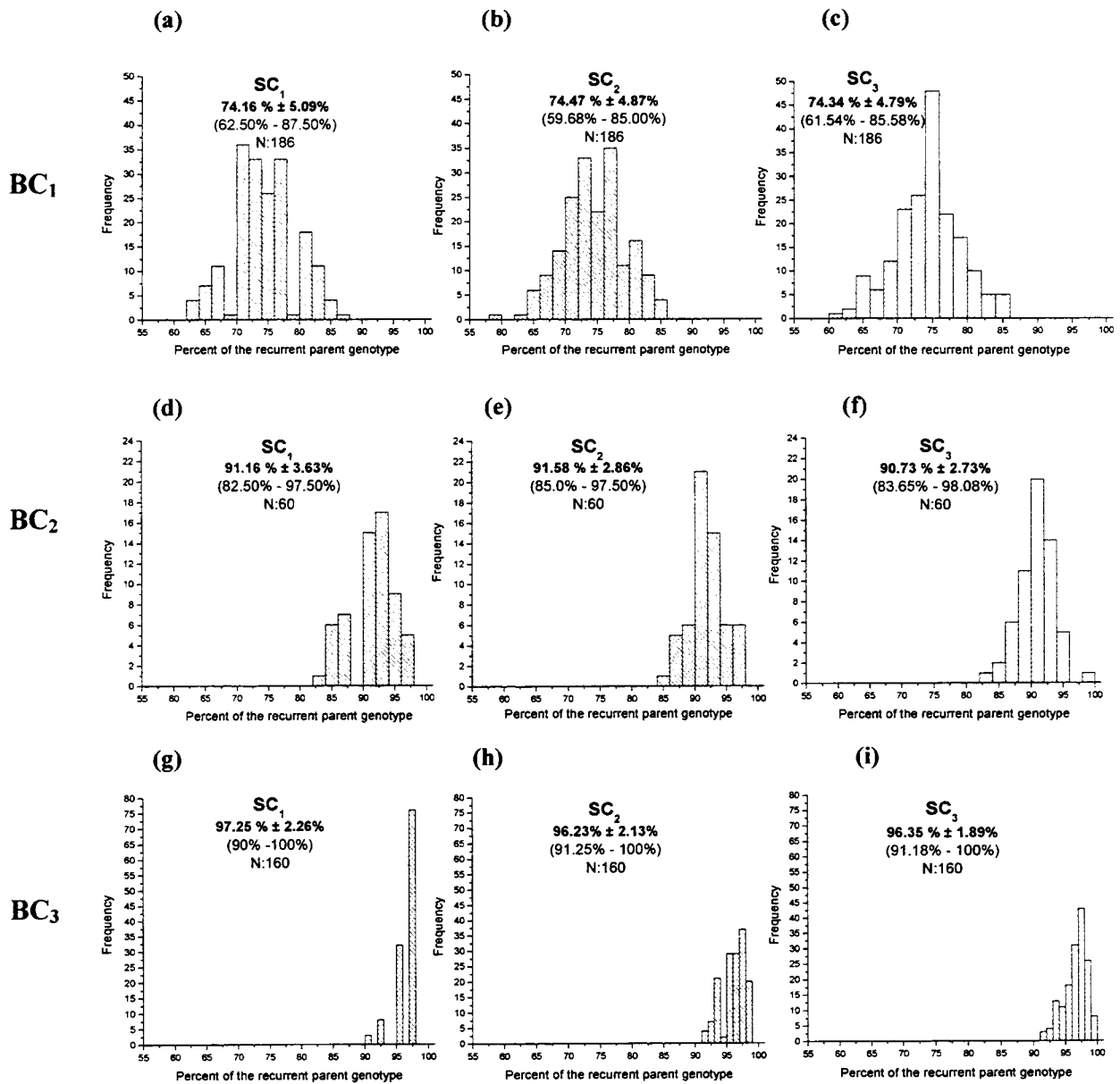


Figure 1. Distribution of the recurrent parent genotype recovery with three selection criteria (SC₁, SC₂ and SC₃) for three backcross generations (BC₁, BC₂ and BC₃).

The presence of alleles derived from the donor parent was not detected in the non-saturated condition. However, in the selected saturated BC₃ lines some donor alleles were detected in Bins that had not been assayed previously for selection (*sat* primers, Table 1). It should be emphasized that the average number of donor alleles reduced with the increment of markers per chromosome used for selection, i. e. 0.75 ± 0.25 for SC₁, 0.44 ± 0.16 for SC₂, and 0.13 ± 0.09 for SC₃. These results showed that only few markers per chromosome could not efficiently monitor the recovery of the recurrent parent genotype for background selection.

Table 2. Comparisons of BC₃ lines percent of recurrent parent genotype for the expected mean of BC₃ generation (93.75%) and for 100%, under non-saturated (non-sat) and saturated (sat) conditions.

Selection Criteria											
SC ₁ [†]				SC ₂ [†]				SC ₃ [†]			
Mean	Sd	N	R	Mean	Sd	N	R	Mean	Sd	N	R
BC₃^{all}				BC₃^{all}				BC₃^{all}			
97.25**	2.26	160	10	96.23**	2.13	160	8.75	96.35**	1.89	160	8.82
BC_{3sel}^{non-sat}				BC_{3sel}^{non-sat}				BC_{3sel}^{non-sat}			
99.84 [‡]	0.63	16	2.50	99.06**	0.72	16	2.50	97.47**	0.46	16	1.05
BC_{3sel}^{sat}				BC_{3sel}^{sat}				BC_{3sel}^{sat}			
92.40 [§]	1.90	16	5.90	95.50**	1.50	16	4.30	96.90**	1.20	16	3.50

[†] Sd: standard deviation; N: number of plants analysed; R: range between maximum and minimum values.

*, ** Significant at 0.05 and 0.01 probability levels, respectively.

[‡] Significant at 0.01 probability level for 93.75%, and non-significant at 100%.

[§] Significant at 0.05 probability level for 93.75%, and significant 0.01 probability level for 100%.

In the individual analyses of variance for the nine phenotypic traits no significance was detected between different selection criteria, except from the number of kernel rows per ear (NKE) and ear diameter (ED) that showed low significance ($P \leq 0.05$). This could be

explained by the great similarity found among the averages for the three selection criteria. For BC₃ lines, significance ($P \leq 0.01$) was detected for most of the traits (Table 3). In the multivariate analyses of variance, for the three types of traits evaluated (yield, height, flowering), no significances were detected for selection criteria and the BC₃ lines differed significantly ($P \leq 0.01$) within each criterion (Table 4). Thus, the analysis of the phenotypic data did not detect differences among the three selection criteria, but there were significant differences between the BC₃ lines within each selection criteria.

Table 3. Values and significances of mean squares (MS) from the analyses of variance for nine traits: ear weight (EW), number of kernel rows per ear (NKE), number of rows per ear (NRE), ear length (EL), ear diameter (ED), plant height (PH), ear height (EH), days to silk extrusion (DSE), and days to pollen shed (DPS).

Sources of Variation	DF	Mean Squares								
		EW (g plant ⁻¹)	NKE (number)	NRE (number)	EL (cm ear ⁻¹)	ED [†] (cm ear ⁻¹)	PH (cm plant ⁻¹)	EH (cm ear ⁻¹)	DSE [‡] (number)	DPS [‡] (number)
Selection Criteria (SC)	2	272.31	195.31*	1.11	7.14	120.06*	1466.56	259.38	1.69	1.65
BC ₃ Lines (L)/SC	45	470.09**	33.62*	1.47*	4.86**	26.44**	553.82**	303.96**	1.57	1.60
L/SC ₁	15	422.71**	35.81	1.33	4.55**	30.98**	536.15**	447.31**	2.11	1.85
L/SC ₂	15	488.65**	53.52**	2.05**	2.89**	13.56**	622.67**	267.25**	2.03	1.36
L/SC ₃	15	498.91**	11.53	1.04	7.14**	34.77**	502.65**	197.33**	1.30	1.60
Sampling Error	432	129.21	20.55	0.89	1.04	6.73	143.02	81.13	-	-
General Mean		57.73	30.97	8.68	12.89	3.34	190.75	111.50	58.50	59.73
SC ₁ mean		57.27	32.24	8.70	12.65	3.24	189.81	110.56	58.13	59.38
SC ₂ mean		57.77	30.45	8.75	13.02	3.40	191.00	112.13	58.69	59.81
SC ₃ mean		58.15	30.22	8.60	12.98	3.38	191.44	111.80	58.69	60.00
CV%		19.69	14.64	10.89	7.90	7.77	6.27	8.08	1.99	1.81

*, ** Significant at 0.05 and 0.01 probability levels, respectively; [†] mean squares multiplied by 10⁻². [‡]For DSE and DPS, there was only one record per plot each, and it was not possible to have an estimate for error value.

Table 4. Values and significances of F tests (Hotelling-Lawley Trace) from the multivariate analyses of variance for yield, height, and flowering.

Sources of Variation	F Test		
	Yield [†]	Height [‡]	Flowering [§]
Selection Criteria (SC)	1.27	1.75	0.39
BC ₃ Lines (L)/SC	2.10**	2.29**	1,28
L/SC ₁	1.99**	3.35**	1,15
L/SC ₂	2.10**	2.00**	0,62
L/SC ₃	2.00**	1.70**	0,35
Sampling Error	---	---	---

**Significant at 0.01 probability level; [†]Yield refers to the following traits: ear weight, number of kernel rows per ear, number of rows per ear, ear length, and ear diameter. [‡]Height refers to plant and ear heights. [§]Flowering refers to days to silk extrusion and days to pollen shed; there was only one record per plot, and consequently, it was not possible to compute the matrix for the error term.

The 48 selected BC₃ lines were compared with the parental inbred recurrent L-14-4B line, using the *Dunnnett* test. For each selection criterion, 120 tests were performed and 13 (10.83%), 14 (11.67%) and 8 (6.67%) of them were significant for SC₁, SC₂ and SC₃ respectively. Plant height presented the highest number of significant tests (10); 4, 4 and 2 of them for SC₁, SC₂ and SC₃, respectively. For number of rows per ear (NKE) and flowering traits (DSE and DPS), the BC₃ lines did not present differences in relation to the parental recurrent line (Table 5). Based on these phenotypic results, the SC₃ criterion showed an inferior number of significant tests, indicating that the increase in the number of markers during marker-assisted selection resulted in plants that were closer to the recurrent parent (Table 5). *Dunnnett* test performed was conservative since Bonferroni's correction for the probability level (Zar 1999) was not used; if it were applied, no BC₃ lines were significantly different for any of the traits in relation to the parental inbred line (L-14-4B). This result showed that even though the analyses of variance had shown the existence of genetic variability among the BC₃ lines, this variability was not enough to generate great

differences in relation to the recurrent parent line. However, the SC₃ (approximately six markers per chromosome) criterion showed itself to be superior to any other, as it presented the fewest number of significant tests and lines with significant tests, for more than one trait simultaneously. Therefore, in the phenotypic level, the superior number of markers per chromosome increased the efficiency of marker-assisted selection in a way that the final SC₃/BC₃ selected lines were closer to the parental type than the other lines selected by SC₁ and SC₂.

Table 5. Numbers and percentages (between parenthesis) of significant *Dunnnett* tests for nine traits. BC₃ lines were compared to the recurrent parent inbred line L-14-4B.

Trait [†]	Number of significant tests [‡]							
	SC ₁ (%)		SC ₂ (%)		SC ₃ (%)		Total (%)	
EW	3	(18.75)	3	(18.75)	3	(18.75)	9	(18.75)
NKE	2	(12.50)	3	(18.75)	0	(0.00)	5	(10.42)
NRE	0	(0.00)	0	(0.00)	0	(0.00)	0	(0.00)
EL	1	(6.25)	3	(18.75)	1	(6.25)	5	(10.42)
ED	0	(0.00)	0	(0.00)	1	(6.25)	1	(2.08)
PH	4	(25.00)	4	(25.00)	2	(12.50)	10	(20.83)
EH	3	(18.75)	1	(6.25)	1	(6.25)	5	(10.42)
DSE	0	(0.00)	0	(0.00)	0	(0.00)	0	(0.00)
DPS	0	(0.00)	0	(0.00)	0	(0.00)	0	(0.00)
Total	13	(10.83)	14	(11.67)	8	(6.67)	35	(9.72)
Total of tests	120		120		120		360	

[†]EW: ear weight, NKE: number of kernel rows per ear, NRE: number of rows per ear, EL: ear length, ED: ear diameter, PH: plant height, EH: ear height, DSE: days to silk extrusion and DPS: days to pollen shed. [‡] SC₁, SC₂, SC₃: selection criteria.

Backcrossing has been mainly used to introgress single Mendelian traits such as disease resistance or quality factors into elite germplasm. It allows target modifications to be performed without disrupting the existing overall genetic balance of the recurrent parent.

Nevertheless, reaching a near isogenic line through classical backcrossing procedures is a lengthy process. In the absence of linkage drag, seven classical backcross generations are required to recover more than 99% of the recurrent parent genotype. Maize is a crop with a fast turnover of elite cultivars, and therefore, the attractiveness of reducing the number of backcross generations through molecular markers is immense (Tanksley et al. 1989; Hospital et al. 1992; Openshaw et al. 1994).

Many are the efforts to establish the best strategy for using marker-assisted selection as a routine in breeding programs. A great number of papers focusing on computer simulations and investigating the parameters involved have turned up (Hospital et al. 1992, 1997; Tanksley and Nelson 1996; Visscher et al. 1996; Reyes-Valdés 2000), but few are the experimental results validating the efficiency on marker-assisted methodologies in different crop species (Stuber 1994; Barone et al. 2001; Chen et al. 2000). This paper has the aim of introgressing a polymorphic microsatellite simulating an oligogenic trait or a QTL linked to it, by marker-assisted selection, with no field selection steps in-between. The idea was to observe recurrent parent genotype recovery when different numbers of markers per chromosome are carried out to assist selection, and whether efficiency is supported or not by genome sampling. According to our molecular and phenotypic results, the PHI037 allele could be efficiently transferred by marker-assisted selection and SC₃ criteria showed itself to be the most efficient.

Before microsatellites availability, marker-assisted selection schemes had to be designed within the constraints imposed by RFLPs, a very reliable marker system, but rather laborious. RFLPs have practical limits on the number of individuals that can be genotyped efficiently (Ribaut et al. 1997). Microsatellites are codominant markers and can assess a great number of individuals. They are reliable, PCR-based markers with a high

level of heterozygosity, inherited in a Mendelian pattern. However, microsatellites sometimes have drawbacks related to their use; some genomic regions present large intervals among markers. Sibov (2002) reported that 5.57 microsatellites on average were tested to find a polymorphic reliable marker. Despite the efforts, some regions could not show complete coverage, and gaps remained in chromosomes 4, 5, 9 and 10. There were intervals up to 40 cM on chromosomes 4 and 5 related to Bins without polymorphic markers. Another study reported difficulties of the same nature (Mangolim 2002). Some features were observed with these microsatellites, such as segregation distortion as some microsatellites segregated in a non-Mendelian fashion, absence of one parental allele in the descendant generation and low resolution in agarose gels. In that study, QTLs associated to oil content were mapped in tropical maize inbred lines, but some chromosomes presented distances up to 70cM between their markers. In our research, some gaps between the markers still remained, even after genome saturation, due to the difficulties in finding polymorphic microsatellites for these Bins. Others molecular markers such as RFLPs mapped on this interval could solve this question.

Hospital et al. (1992) studied the effects of population size, number and position of selected markers in introgression breeding programs considering the expected proportion of the recipient genome. After analyzing the behavior of the recurrent parent markers in chromosomes not carrying the gene of interest, they reported that increasing the number of markers to more than three per chromosome would not be efficient. Our study, however, has showed that a higher number of markers would enhance the confidence of the recurrent parent estimates. This difference might be due to what Reyes-Valdés (2000) had attempted in relation to Hospital et al. (1992), their assumption of independence among all *loci* when calculating the expected proportion of recipient genome in non-carrier chromosomes.

Marker-assisted selection using microsatellites appears to be particularly promising. As a result from this research, the number of markers per chromosome per generation contributes to increasing the effectiveness of MAS. Compared to phenotypic selection, MAS proved to generate consistent results and high genetic gains in only three generations of backcrosses. The strategy used could be adjusted to a single gene trait with a tightly molecular marker linked to it. The results herewith demonstrated the usefulness of MAS in allele introgression and the importance of considering the correct parameters to design and develop a marker-assisted selection program. Usually, when a breeding backcross program is developed in the field, phenotypic selection is performed and plants are compared with the recurrent inbred line. In this research, marker-assisted selection showed itself to be efficient without phenotypic selection; but a 100% of recovery of the recurrent parent genotype was not achieved and the BC₃ lines differed from the parental line as in the Dunnett test. Practically, MAS could be associated to phenotypic selection and this partnership will enhance parental genotype recovery.

Acknowledgements

We would like to thank S.T. Sibov, C.A. Mangolim, A.R. Silva, A.F. Garcia, H.C. Acosta, A.C. Avancini, A. Santos, Z. Gonçalves, A.P. Alves, M.Z. Silva, J. Lugli for technical support during the development of the research; A.M. Aguiar for helping in the data analysis; A.J. Desidério, A.S. Oliveira, M.R.C. Filho, M.A. Silva for assisting with field plot management. This work was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) 99/12143-1 and 99/11479-6 by means of a grant to C.L. de Souza Jr and A.P. de Souza, respectively. L.L. Benchimol received a fellowship from FAPESP (contract 97/13397-1); C.L. de Souza Jr and A.P. de Souza are recipients of a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

- Barone A, Sebastiano A, Carputo D, della Rocca F (2001) Molecular marker-assisted introgression of wild *Solanum commersonii* genome into *S. tuberosum* gene pool. *Theor Appl Genet* 102:900-907
- Chen S, Lin XH, Xu CG, Zhang Q (2000) Improvement of bacterial blight resistance of "Minghui63", an elite restorer line of hybrid rice, by molecular marker-assisted selection. *Crop Sci* 40:239-244
- Chen S, Xu CG, Lin XH, Zhang Q (2001) Improving bacterial blight resistance of '6078', an elite restorer line of hybrid rice, by molecular marker-assisted selection. *Plant Breeding* 120: 133-137
- Chetelat RT, De Verna JW, Bennett (1995) Introgression into tomato (*Lycopersicon esculentum*) of the *L. chmielewskii* sucrose accumulator gene (*sucr*) controlling fruit sugar composition. *Theor Appl Genet* 91: 327-333
- Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* 19 (14)
- Fehr WF (1987) Principles of cultivar development: v.1. Theory and technique. Macmillan, New York
- Frisch M, Melchinger AE (2001) The length of the intact donor chromosome segment around a target gene in marker-assisted backcrossing
- Hillel J, Schaap T, Haberland A, Jeffreys J, Plotzky Y, Cahaner A, Lavi U (1990) DNA fingerprints applied to gene introgression in breeding programs. *Genetics* 124:783-789
- Hoisington D, Khairallah M, Gonzalez-de-Lion D (1994) Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. Second Edition, Mexico DF: CIMMYT
- Hospital F (2001) Size of donor segments around introgressed *loci* and reduction of linkage drag in marker-assisted backcross programs. *Genetics* 158: 1363-1379
- Hospital F, Charcosset A (1997) Marker-assisted introgression of quantitative trait *loci*. *Genetics* 147: 1469-1485
- Hospital F, Chevalet C, Mulsant P (1992) Using markers in gene introgression breeding programs. *Genetics* 132: 1199-1210
- Hospital, F.; Moreau, L.; Lacoudre, F.; Charcosset A, Gallais, A. (1997). More on the efficiency of marker-assisted selection. *Theor. Appl. Genet* 95: 1181-1889

Mangolim CA (2002) Mapeamento de QTLs associados com o teor de óleo em linhagens tropicais de milho (*Zea mays* L.) Dissertação de Doutorado, UNICAMP, IB, Campinas

Melchinger AE, Lee M, Lamkey KR, Hallauer AR, Woodman WL (1990) Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. *Theor Appl Genet* 80: 488-496

Openshaw SJ, Jarboe SG, Beavis WD (1994) Marker-assisted selection in backcross breeding. Pp. 41-43. In: Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon

Reyes-Valdés MH (2000) A model for marker-based selection in gene introgression breeding programs. *Crop Sci* 40:91-98

Ribaut JM, Jiang C, González-de-Léon D, Edmeades GO, Hoisington DA (1997) Identification of quantitative trait *loci* under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. *Theor Appl Genet* 94: 887-896

Rinke EH, Sentz JC (1961) Moving corn-belt germplasm northward. *Ann. Hybrid Corn Industry Conf.* 16:53-56

Sanchez AC, Brar DS, Huang N, Li Z, Khush GS (2000) Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. *Crop Sci* 40:792-797

SAS institute 2001 SAS/SAT user's guide v6.03. SAS Institute. Cary, N.C.

Shaver DL (1976) Conversions for earliness in maize inbreds. *Maize Genet Coop Newsltr* 50:20-23

Sibov ST (2002) Mapeamento de locos quantitativos (QTLs) associados à produção de milho tropical através de marcadores moleculares. Dissertação de Doutorado, UNICAMP, IB, Campinas

Stuber CW (1994) Enhancement of grain yield in maize using marker-facilitated introgression of QTLs. Pp. 44-46. In: Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon

Tanksley SD (1983) Molecular markers in plant breeding. *Plant Mol Biol Rep* 1:3-8

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor Appl Genet* 92: 191-203

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. *Bio/Technology* 7: 257-264

Visscher PMS, Haley CS, Thompson R (1996) Marker assisted introgression in backcross breeding programs. *Genetics* 144:1923-1932

Weck E; Bergstedt M, Farrar D, Kieler M, Krumm C, Lundell M, Mead D, Sunderland H, Taylor R, Wagen CA (1991) RFLPs push BC₂ selections ahead. *Maize Genet Newsltr* 65:101-102

Young ND, Tanksley SD (1989) Restriction Fragment length polymorphism maps and the concept of graphical genotypes. *Theor Appl Genet* 77:95-101

Zar JH (1999). *Biostatistical analysis*. 4th Edition Prentice Hall, Inc. Upper Saddle River, New Jersey

5. CONCLUSÕES

5. Conclusões

- a) Marcadores microssatélites se mostraram adequados para serem utilizados na introgressão de uma característica monogênica, para monitoração e recuperação rápida do genótipo do genitor recorrente, em três gerações de retrocruzamentos assistidos;
- b) O alelo PHI037 foi transferido da linhagem doadora para a linhagem recorrente e a eficiência desta transferência mostrou-se superior com a utilização do critério CS₃, com aproximadamente seis marcas por cromossomo;
- c) Em nível molecular e fenotípico, o critério CS₃ foi o mais eficiente para praticar a seleção assistida por microssatélites. Mostrou maior precisão para as estimativas de recuperação do genótipo do recorrente sob diferentes condições de saturação do genoma, menor proporção de alelos remanescentes do doador nas plantas RC₃ selecionadas e maior identidade das plantas selecionadas por este critério com o genitor recorrente;
- d) O espaçamento médio de 25 cM para os 72 microssatélites utilizados, ao longo dos 10 cromossomos, foi adequado para realizar seleção assistida por marcadores.

6. PERSPECTIVAS FUTURAS

6. Perspectivas Futuras

O uso de seleção assistida por marcadores moleculares em programas de retrocruzamentos ainda é um desafio. Este trabalho acrescenta uma nova dimensão teórico-prática de que a seleção assistida por marcadores moleculares do tipo microssatélites é viável e pode trazer uma recuperação mais rápida do genótipo do genitor recorrente, em três gerações de retrocruzamentos assistidos, comparada ao melhoramento convencional, quando baseado apenas no fenótipo. Entretanto, o delineamento experimental a ser utilizado num protocolo de SAM consiste em um passo determinante para o sucesso do método, principalmente no que tange ao número de marcadores a serem utilizados para o monitoramento da introgressão da característica de interesse e para a recuperação eficaz do genótipo do genitor recorrente. De acordo com os nossos resultados, evidenciou-se a necessidade de se utilizar seis marcas por cromossomo no mínimo para um monitoramento fiel da introgressão do alelo alvo e da recuperação da característica de interesse.

O emprego de SAM deve requerer o estabelecimento de prioridades e estratégias adequadas a depender do perfil da característica, do número de características que se almeja selecionar e da espécie em questão, de forma a se obter uma probabilidade razoável de recuperação do genótipo com a menor alocação de recursos e esforços. Torna-se complicada a utilização de SAM, em se tratando de características quantitativas, quando nenhum gene de efeito maior for detectado (Grattapaglia, 2000).

Considerando-se que as principais características de importância agrônômica são de herança quantitativa, o principal objetivo dos trabalhos envolvendo mapeamento de *QTLs* é a sua aplicação para a seleção assistida por marcadores em programas de melhoramento de

várias espécies. Quanto maior a dificuldade relativa à seleção fenotípica de uma dada característica, mais útil deverá ser a aplicação de SAM (Lande & Thompson, 1990).

Paralelamente, quanto maior a complexidade de manejo de uma espécie, mais frutos podem ser colhidos pela utilização de técnicas de seleção assistida por marcadores moleculares. Em espécies florestais como *Eucalyptus*, a seleção assistida deve ser útil para características de baixa herdabilidade (Grattapaglia, 2000). Entretanto, bom senso é necessário, pois muitas vezes a implementação de SAM ainda pode ser um desafio.

Com o avanço das técnicas de DNA Recombinante, a seleção assistida por marcadores moleculares pode ser um veículo útil para a introgressão de construtos gênicos já previamente incorporados e transmitidos de forma estável em uma dada linhagem (Ragot *et al.*, 1995). Apesar dos riscos já conhecidos, não se pode negar que os transgênicos ampliem o espectro de manipulação genética no melhoramento de muitas espécies.

Deve-se considerar que após uma dada característica ser introgridida para uma linhagem elite, torna-se necessário também utilizar a seleção assistida via retrocruzamentos para transferir esta característica para outras linhagens elites. No caso de características monogênicas, o sucesso da aplicação de SAM é muito mais garantido. Entretanto, deve-se avaliar o custo-benefício do método, deixando-se a cargo da seleção assistida aquelas características cuja detecção e monitoramento fenotípico não são triviais por metodologias clássicas. Para estas características, pode-se pensar e sugerir uma metodologia alternativa à desenvolvida neste trabalho. Por exemplo, ao invés de se fazer o monitoramento com todas as marcas, desde o primeiro ciclo de retrocruzamentos, esta averiguação poderia ser feita, nos ciclos iniciais, apenas amplificando-se a marca ligada ao alelo alvo a ser transferido (*foreground selection*). Combinadamente, nestes ciclos iniciais, seleção fenotípica seria

praticada. Na terceira geração de retrocruzamentos, entretanto, todos os marcadores moleculares polimórficos e disponíveis seriam utilizados para a averiguação do percentual de recuperação do genótipo do parental recorrente. Isto demandaria um tamanho populacional maior do que o utilizado neste trabalho para a terceira geração de retrocruzamentos, a fim de se atingir o máximo de saturação possível do genoma (*background selection*). Desta forma, o monitoramento da recuperação do genótipo do genitor recorrente seria totalmente feito neste ciclo. Para saber qual a viabilidade desta metodologia, as estimativas de recuperação do genitor recorrente a serem alcançadas teriam que ser avaliadas.

Torna-se importante ressaltar que marcadores microsatélites constituem uma ferramenta poderosa para se praticar a seleção assistida possibilitando a clara identificação rápida e eficaz dos alelos dos genitores além da visualização do heterozigoto. Esta é uma técnica de fácil manipulação baseada em amplificação por *PCR* e leitura das bandas em géis de agarose de alta resolução, o que faz com que a obtenção dos resultados seja rápida e eficaz, uma vez que já se tenham os pares de *primers* desenvolvidos para a espécie. Possibilita a manipulação de grandes quantidades de genótipos ao mesmo tempo, de forma não destrutiva. No caso de milho, já existem muitos microsatélites disponíveis no *Maize Data Bank* para serem sintetizados e utilizados. Entretanto, em muitas culturas, como é no caso da cultura de feijoeiro, não existem microsatélites suficientes desenvolvidos nem tão pouco mapeados. Mesmo para aquelas espécies em que estes já são abundantes, ainda há dificuldades técnicas quanto à detecção de polimorfismo, desvios de segregação ou até inconstância de um dos alelos presentes nos genitores (Sibov, 2002; Mangolim, 2002).

Num curto período de tempo, empregando-se a metodologia correta para se ter uma boa avaliação genotípica durante a seleção, SAM deve ganhar difusão entre muitas espécies de importância comercial. A expectativa de sua utilização é imensa e tem gerado grande atenção. Há resistência principalmente por desconhecimento da melhor metodologia a ser utilizada. Neste contexto, espera-se que o presente trabalho sirva de contribuição quanto à definição de alguns parâmetros reais e por prover resultados práticos e bem sucedidos da aplicação da seleção assistida por marcadores moleculares microsatélites em programas de retrocruzamentos em milho.

7. ANEXO

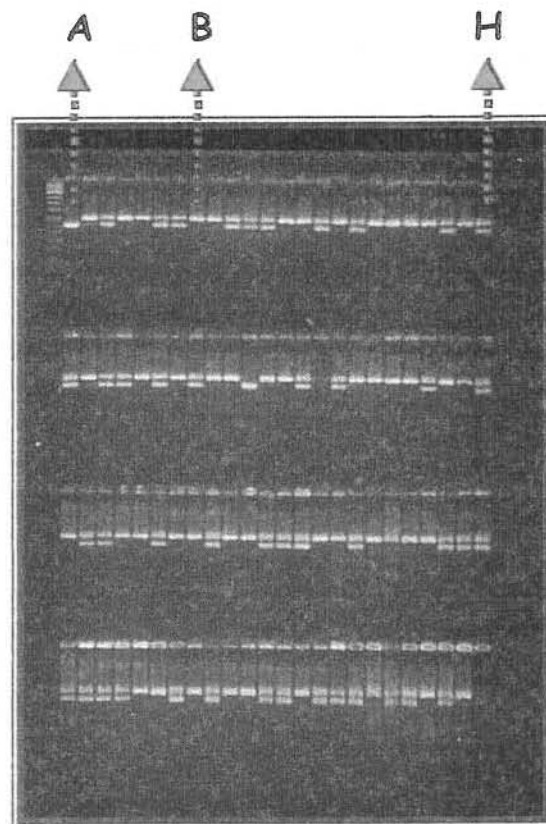


Figura 1AN. Amplificação das plantas RC_1 com o microsatélite PHI 037 (Bin 1.08). $\Phi X/HinfI$ foi usado como marcador de peso molecular. A é o alelo do genitor doador (L08-05F) e B é o alelo do genitor recorrente (L14-4B). H refere-se ao genótipo observado para o F_1 .

Marcas Passadas no Cromossomo 01

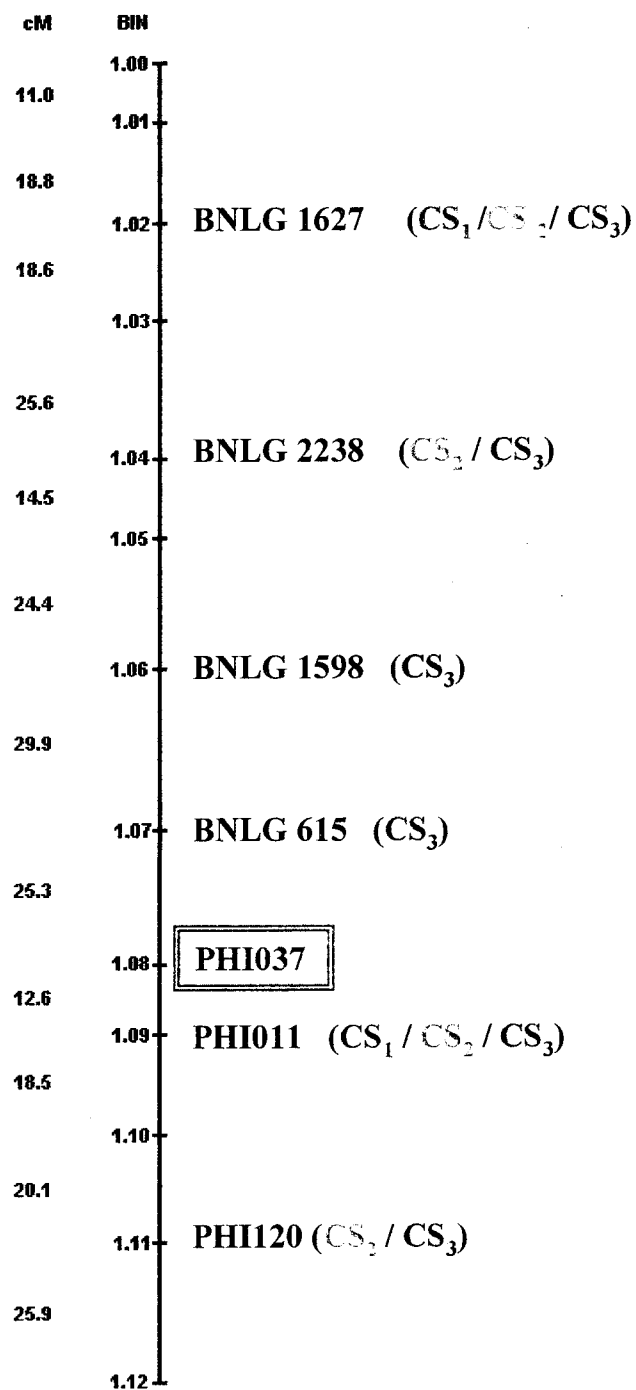


Figura 2AN. Microssatélites ancorados no cromossomo 01 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção.

Marcas Passadas no Cromossomo 02

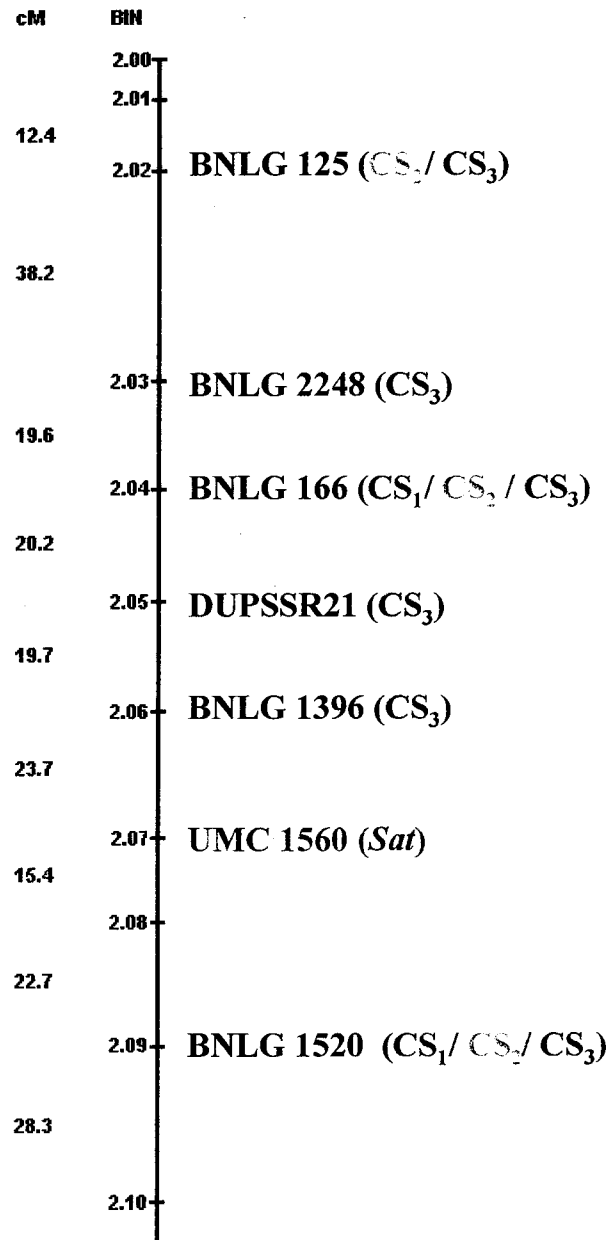


Figura 3AN. Microssatélites ancorados no cromossomo 02 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.

Marcas Passadas no Cromossomo 03

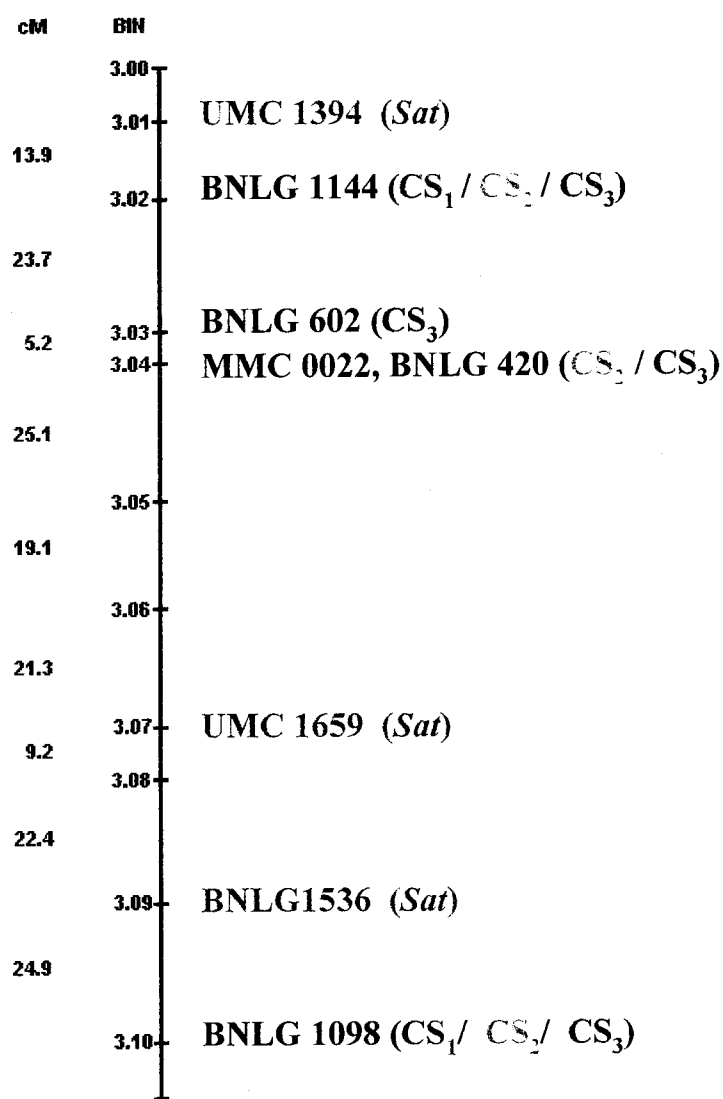


Figura 4AN. Microsatélites ancorados no cromossomo 03 e utilizados na amplificação das amostras RC_1 , RC_2 e RC_3 . CS_1 , CS_2 e CS_3 são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas $RC_{3sel.}$

Marcas Passadas no Cromossomo 04

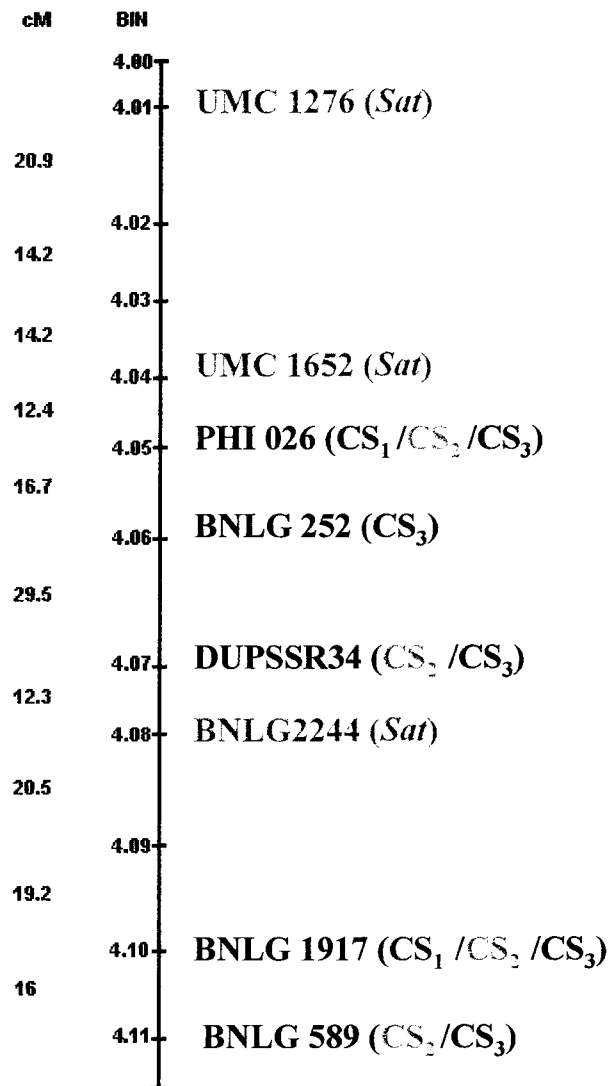


Figura 5AN. Microssatélites ancorados no cromossomo 04 e utilizados na amplificação das amostras RC_1 , RC_2 e RC_3 . CS_1 , CS_2 e CS_3 são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas $RC_{3sel.}$

Marcas Passadas no Cromossomo 05

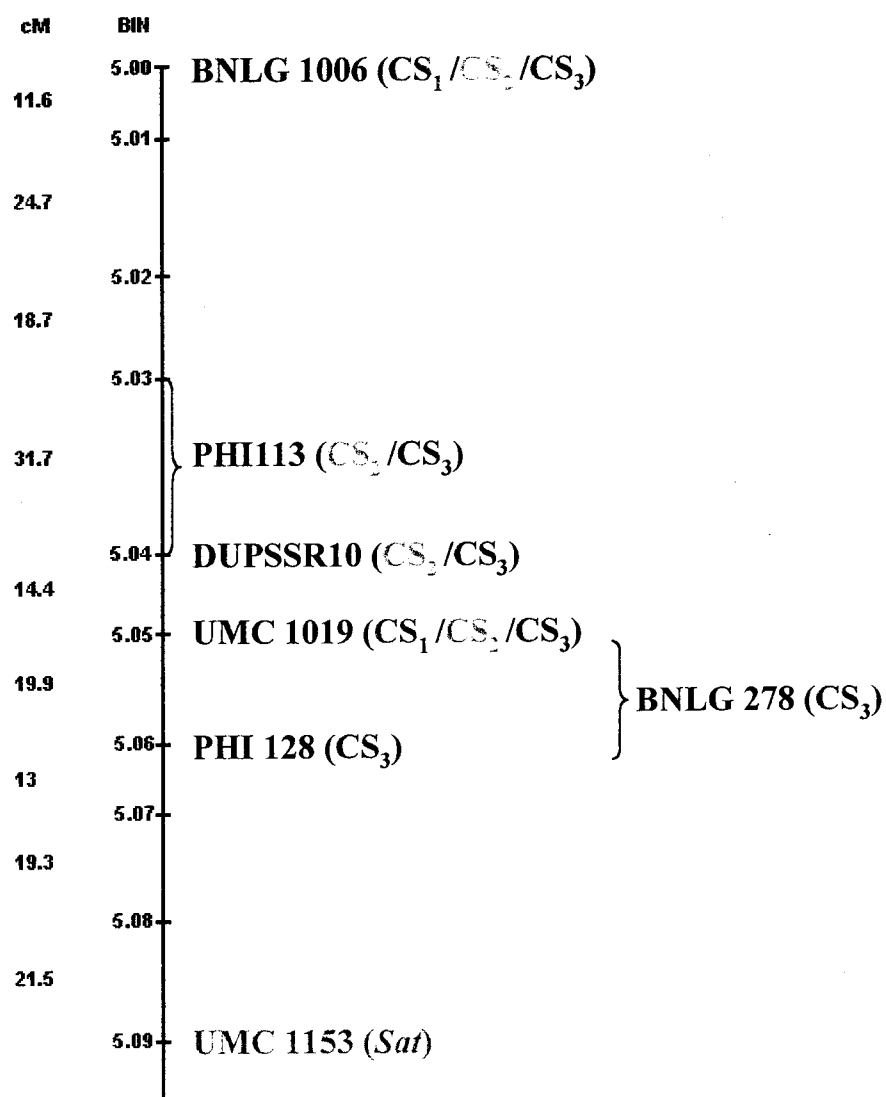


Figura 6AN. Microsatélites ancorados no cromossomo 05 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel.}.

Marcas Passadas no Cromossomo 06

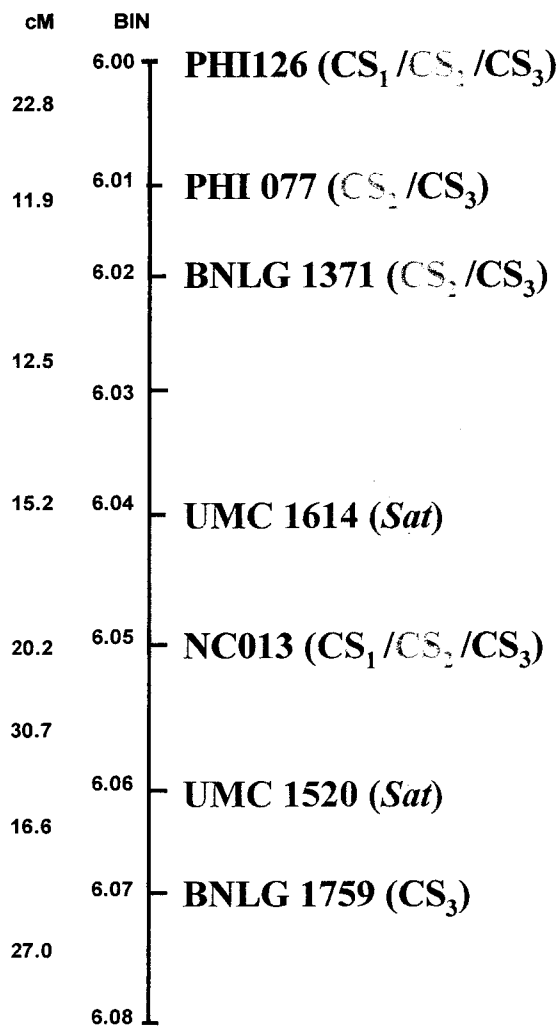


Figura 7AN. Microsatélites ancorados no cromossomo 06 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.

Marcas Passadas no Cromossomo 07

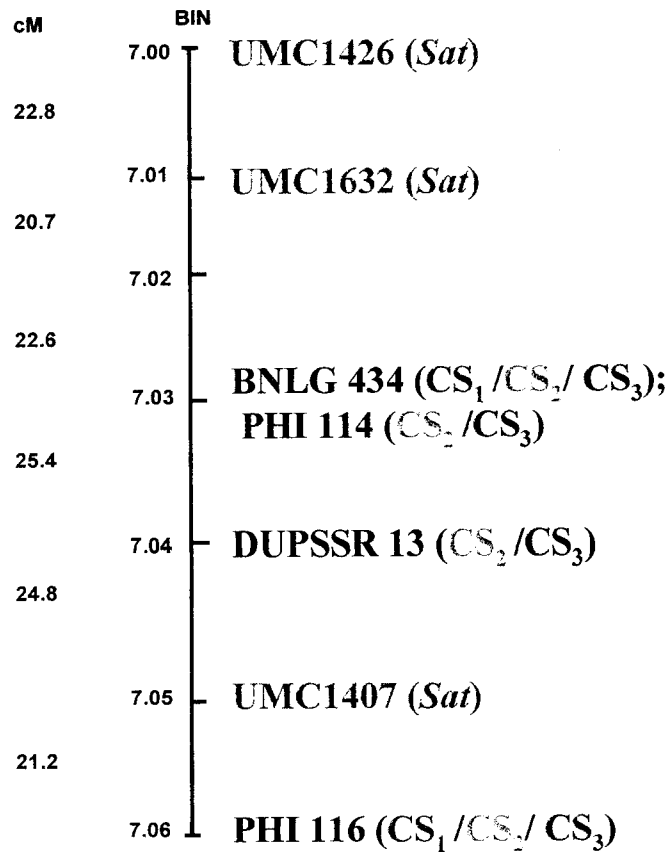


Figura 8AN. Microsatélites ancorados no cromossomo 07 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.

Marcas Passadas no Cromossomo 08

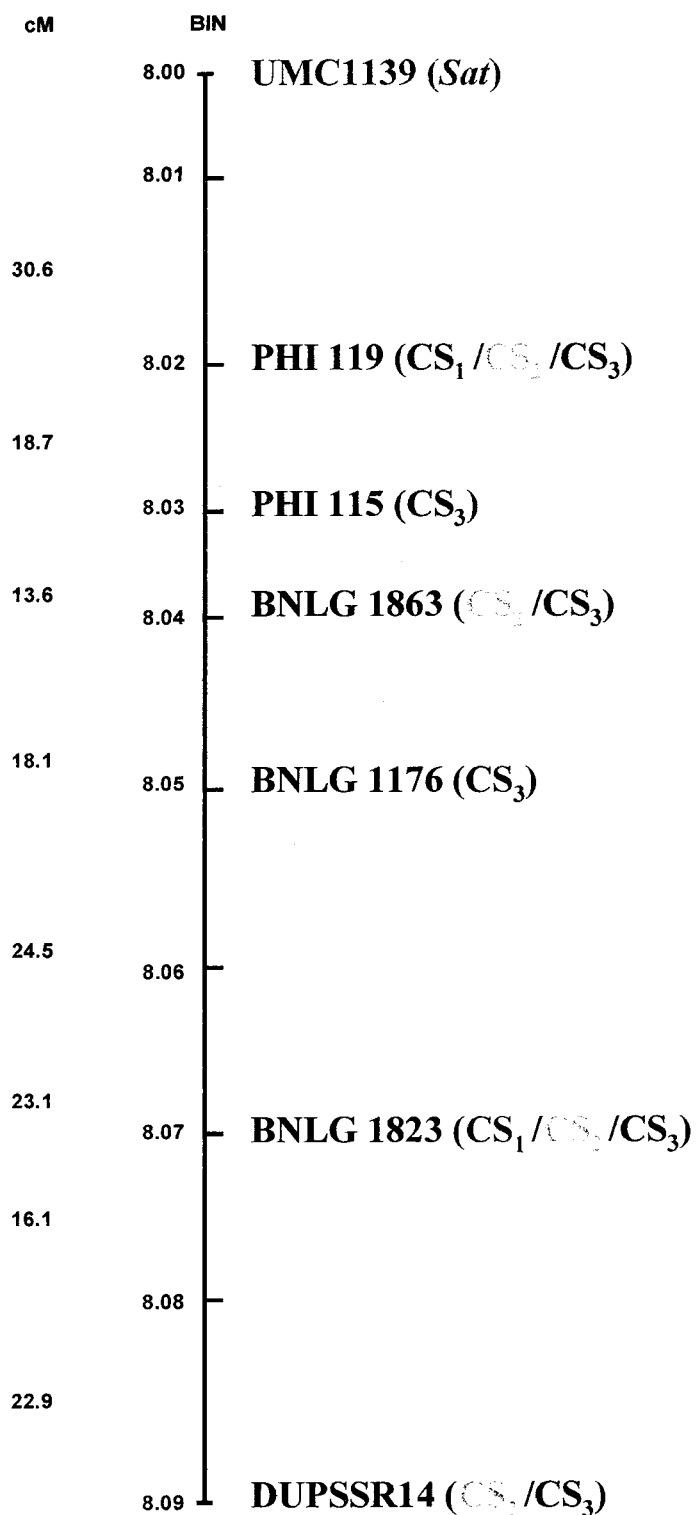


Figura 9AN. Microsatélites ancorados no cromossomo 08 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel}.

Marcas Passadas no Cromossomo 09

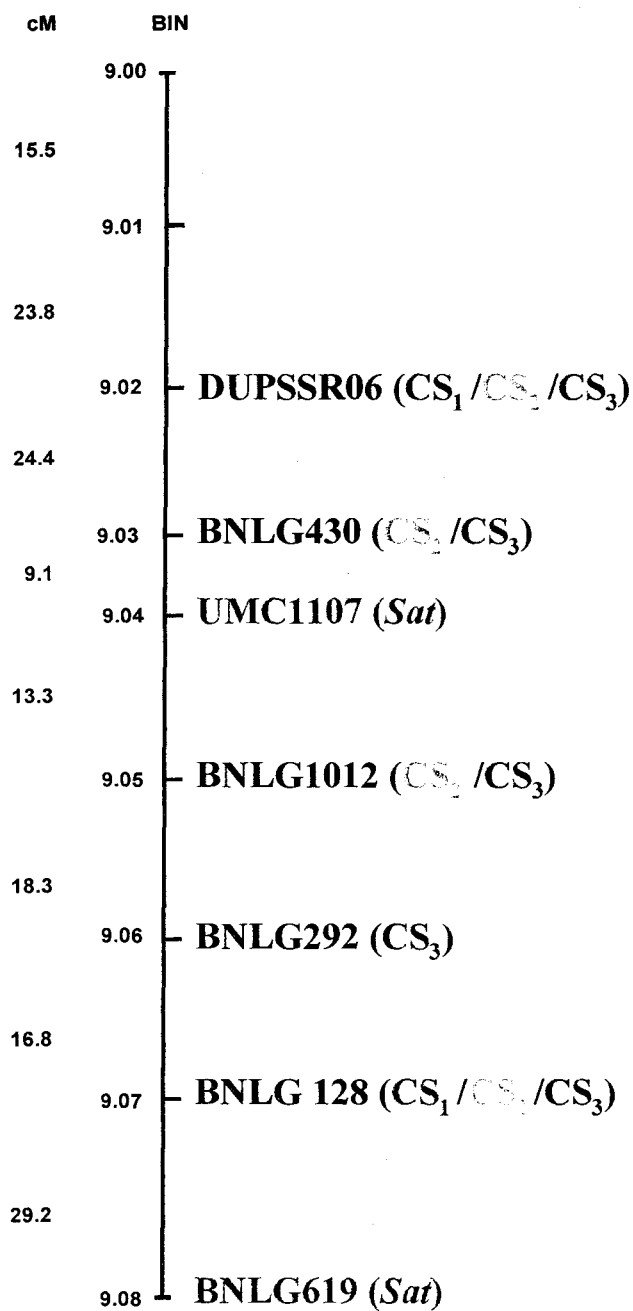


Figura 10AN. Microsatélites ancorados no cromossomo 09 e utilizados na amplificação das amostras RC_1 , RC_2 e RC_3 . CS_1 , CS_2 e CS_3 são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas $RC_{3sel.}$

Marcas Passadas no Cromossomo 10

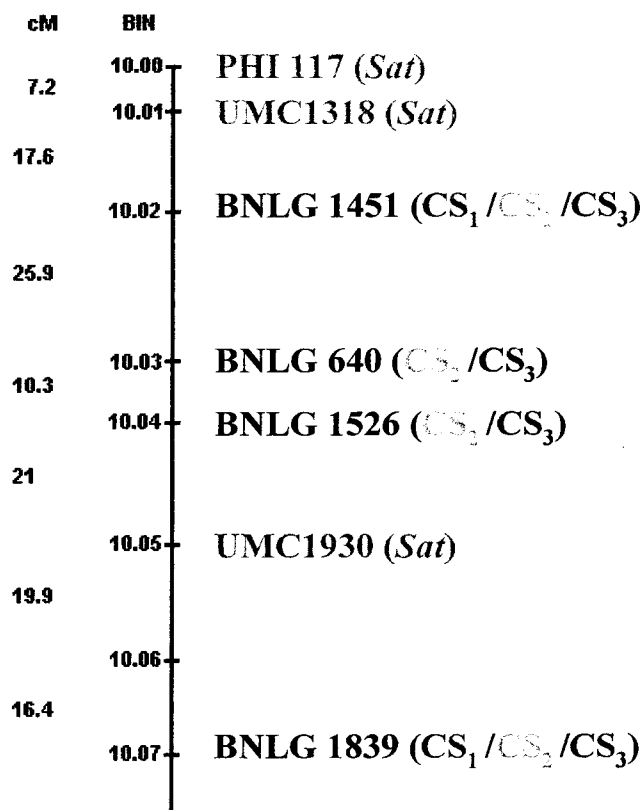


Figura 11AN. Microsatélites ancorados no cromossomo 10 e utilizados na amplificação das amostras RC₁, RC₂ e RC₃. CS₁, CS₂ e CS₃ são os critérios de seleção; *Sat* é a notação para designar o conjunto de pares de *primers* complementares utilizados para saturar o genoma das 48 plantas RC_{3sel.}.

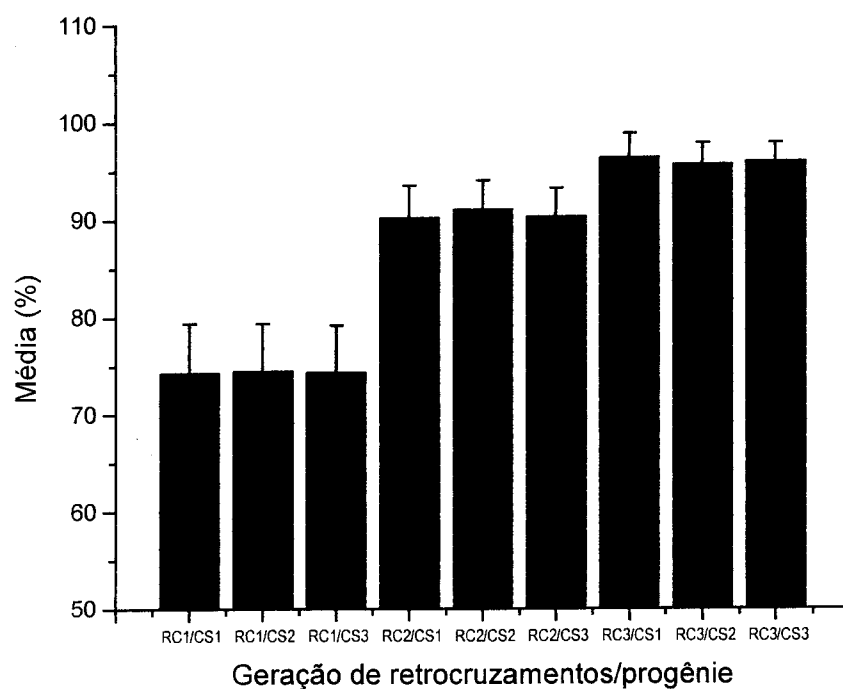


Figura 12AN. Médias de recuperação do genótipo do genitor recorrente para cada um dos critérios (CS_1 , CS_2 e CS_3) obtidas em cada geração de retrocruzamentos com seleção assistida por microssatélites.

Tabela 1AN. Plantas RC₁ selecionadas e o percentual do genótipo do genitor recorrente para os critérios de seleção avaliados (CS₁, CS₂ e CS₃).

CS ₁			CS ₂			CS ₃		
Progênes	Rec. (%)	A	Progênes	Rec. (%)	A	Progênes	Rec. (%)	A
8	82,50	0	12A	83,75	0	12B	82,69	0
93	82,50	0	201	81,25	0	91	82,69	0
204A	82,50	0	204B	82,50	0	232	80,77	0
247	87,50	0	300	85,00	0	300	84,62	0
<i>Média</i> _{CS1}	83,75		<i>Média</i> _{CS2}	83,13		<i>Média</i> _{CS3}	82,69	
<i>Média</i> _{RC1} [*]	74,16		<i>Média</i> _{RC1} [*]	74,47		<i>Média</i> _{RC1} [*]	74,34	
<i>Desvio</i> _{RC1} [*]	5,09		<i>Desvio</i> _{RC1} [*]	4,87		<i>Desvio</i> _{RC1} [*]	4,79	
<i>Min</i>	62,50		<i>Min</i>	59,68		<i>Min</i>	61,54	
<i>Max</i>	87,50		<i>Max</i>	85,00		<i>Max</i>	85,58	
<i>Amplitude</i>	25		<i>Amplitude</i>	25,35		<i>Amplitude</i>	24,04	
<i>N</i> [†]	186		<i>N</i> [†]	186		<i>N</i> [†]	186	

^{*} As médias e os desvios-padrões foram calculados para as quatro plantas selecionadas dentro de cada critério.

[†] N: é o número de observações em cada critério de seleção

Tabela 2AN. Plantas RC₂ selecionadas e o percentual de recuperação dos genótipos do genitor recorrente para cada critério de seleção avaliado (CS₁, CS₂ e CS₃).

CS ₁				CS ₂				CS ₃			
Progênes	Plantas	A	Rec. (%)	Progênes	Plantas	A	Rec. (%)	Progênes	Plantas	A	Rec. (%)
8	12	0	97,50	12A	1	0	92,50	12B	5	0	88,46
	26	0	97,50		14	0	93,75		23	0	91,35
	29	0	95,00		17	0	92,50		26	0	91,35
	34	0	95,00		22	0	96,25		31	0	88,46
93	11	0	90,00	201	11	0	90,00	91	25	0	92,31
	15	0	92,50		29	0	91,25		32	0	92,10
	16	0	92,50		30	0	90,00		35	0	92,31
	34	0	95,00		34	0	90,00		36	0	94,23
204A	10	0	95,00	204B	3	0	90,00	232	2	0	91,35
	16	0	90,00		10	0	92,50		5	0	91,35
	24	0	92,50		21	0	90,00		12	0	95,19
	28	0	90,00		25	0	90,00		18	0	90,38
247	14	0	97,50	300A	6	0	93,75	300	9	0	93,27
	21	0	97,50		8	0	95,00		11	0	95,19
	25	0	95,00		33	0	96,25		17	0	90,38
	31	0	97,50		34	0	96,25		40	0	95,19
<i>Média</i> CS ₁			94,38	<i>Média</i> CS ₂			92,50	<i>Média</i> CS ₃			92,05
<i>Média</i> RC ² CS ₁			91,16**	<i>Média</i> RC ² CS ₂			91,58**	<i>Média</i> RC ² CS ₃			90,73**
<i>Desvio</i> RC ² CS ₁			3,63	<i>Desvio</i> RC ² CS ₂			2,86	<i>Desvio</i> RC ² CS ₃			2,73
<i>Min</i>			82,50	<i>Min</i>			85,00	<i>Min</i>			83,65
<i>Max</i>			97,50	<i>Max</i>			97,50	<i>Max</i>			98,08
<i>Amplitude</i>			15	<i>Amplitude</i>			12,50	<i>Amplitude</i>			14,43
<i>N</i> [†]			60	<i>N</i> [†]			60	<i>N</i> [†]			60

** Comparações com o valor esperado de 87,5% foram significativas ao nível de probabilidade de 0,01.

[†] N: é o número de observações em cada critério de seleção.

Tabela 3AN. Comparações entre RC_{3sel}^{sat} (selecionada e saturada) x $RC_{3sel}^{non-sat}$ (selecionada e não-saturada), em relação a presença de alelos do genitor doador (A). $S_{\bar{x}}$ é o erro padrão da média.

Critérios	$RC_{3sel}^{non-sat}$	RC_{3sel}^{sat}
	Média	Média ($S_{\bar{x}}$)
CS_1	0	0,75 (0,25)
CS_2	0	0,44 (0,16)
CS_3	0	0,13 (0,09)

Tabela 4 AN. Comparações entre os diferentes critérios das amostras RC_3 totais (RC_3^{all}), saturadas (RC_{3sel}^{sat}) e não-saturadas ($RC_{3sel}^{non-sat}$).

RC_3	Crítérios	Valores de t
RC_3^{all}	$CS_1 \times CS_2$	4,59**
	$CS_2 \times CS_3$	0,13
	$CS_3 \times CS_1$	5,10**
RC_{3sel}^{sat}	$CS_1 \times CS_2$	7,31**
	$CS_2 \times CS_3$	2,82*
	$CS_3 \times CS_1$	8,65**
$RC_{3sel}^{non-sat}$	$CS_1 \times CS_2$	4,04**
	$CS_2 \times CS_3$	5,37**
	$CS_3 \times CS_1$	14,05**

*, ** Significativos a níveis de probabilidade de 0,05 e 0,01, respectivamente.

8. REFERÊNCIAS BIBLIOGRÁFICAS

8. Referências Bibliográficas

- Adams WT (1983).** Application of isozymes in tree breeding. In: Isozymes in plant genetics and breeding. Part A. Tanskey SD & Orton TJ (Eds.). Elsevier Science Publishers, Amsterdam pp 60-64.
- Allard RW (1960).** Principles of plant breeding. John Wiley & Sons, Inc. New York.
- Allard RW, Zhang Q, Saghai Maroof MA, Muona OM (1992).** The evolution of multilocus genetic structure in an experimental barley population. *Genetics*, 131:957-969.
- Ahmadi N, Albar L, Pressoir G, Pinel A, Fargette D, Ghesquière A (2001).** Genetic basis and mapping of the resistance to *Rice yellow mottle virus*. III. Analysis of QTL efficiency in introgressed progenies confirmed the hypothesis of complementary epistasis between two resistance QTLs. *Theor. Appl. Genet.*, 103: 1084-1092.
- Armstrong CL, Romero-Severson J, Hodges TK (1992).** Improved tissue culture response of an elite maize inbred through backcross breeding, and identification of chromosomal regions important for regeneration by RFLP analysis. *Theor. Appl. Genet.*, 84: 755-762.
- Barbosa AMM, Geraldi IO, Benchimol LL, Garcia AAF, Souza Jr CL, Souza AP (2002).** Relationship of intra- and interpopulation tropical maize single cross hybrid performance and genetic distances computed from AFLP and SSR markers. Aceito para publicação na *Euphytica* (aceite em Set/2002).
- Bauman LF (1977).** Improvement of established maize inbreds. *Maydica* 4: 213-222.

- Becker J, Heun M (1995).** Allele variation and mapping. *Plant Mol. Biol.*, 27:835-845.
- Beckman JS, Soller M (1986).** Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. *Euphytica*, 35:111-124.
- Beckman JS, Weber JL (1992).** Survey of human and rat microsatellites. *Genomics*, 12: 627-631.
- Benchimol LL, Souza Jr CL, Garcia AAF, Kono, PMS, Mangolim CA, Barbosa AMM, Coelho ASG, De Souza AP (2000).** Genetic diversity in tropical maize inbred lines: heterotic group assignment and hybrid performance determined by RFLP marker. *Plant Bred.*, 119(6): 491-496
- Bernacchi D, Beck-Bunn EY, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley S (1998).** Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.*, 97: 381-397.
- Borém A (2001).** Melhoramento de plantas. Viçosa: UFV, 3. ed., 500p.
- Botstein D, White RL, Skolnick M, Davis RW (1980).** Construction of a genetic linkage map in a man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, 32: 314-331.
- Briggs FN, Allard RW (1953).** The current status of the backcross method of plant breeding. *Agron. J.*, 131-138.

- Burr B (2001).** Some concepts and new methods for molecular mapping in plants. In: Phillips RL & Vasil IK (Eds.). DNA-Based Markers in Plants, pp. 1-8. Kluwer Academic Publishers, Netherlands.
- Chen S, Lin XH, Xu CG, Zhang Q (2000).** Improvement of bacterial blight resistance of “Minghui63”, an elite restorer line of hybrid rice, by molecular marker-assisted selection. *Crop Sci.*, 40:239-244.
- Chetelat RT, De Verna JW, Bennett (1995).** Introgression into tomato (*Lycopersicon esculentum*) of the *L. chmielewskii* sucrose accumulator gene (*sucr*) controlling fruit sugar composition. *Theor. Appl. Genet.*, 91: 327-333.
- Chin ECL, Senior ML, Shu H, Smith JSC (1996).** Maize simple repetitive DNA sequences: abundance and allele variation. *Genome*, 39: 866-873.
- Cordeiro GM, Taylor GO, Henry RJ (2000).** Characterization of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. *Plant Science*, 155:161-168
- Dowswell CR, Paliwal RL, Cantrell RP (1996).** Maize in the World Economy. In: Maize in The Third World. Breth SA (Ed.). Winrock development-oriented literature series, Westview Press, Oxford. Pp. 268.
- Dudley JW (1993).** Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. *Crop Sci.*, 27: 944-947.
- Edwards MD, Page NJ (1994).** Evaluation of marker-assisted selection through computer simulation. *Theor. Appl. Genet.*, 88: 376-382.

- Fehr WR (1987).** Principles of cultivar development. Volume 1: Theory and Thecnique. McGraW-Hill, Inc., New York, NY.
- Ferreira ME, Grattapaglia D (1998).** Classes de marcadores moleculares para análise genética. In: Introdução ao uso de marcadores moleculares em análise genética. 3^a ed. EMBRAPA_CENARGEN, Pp. 220, Brasília.
- Frisch M, Bohn M, Melchinger AE (1999a).** Minimum sample size and optimum positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. *Crop Sci.*, 39:967-975.
- Frisch M, Bohn M, Melchinger AE (1999b).** Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Sci.*, 39: 1295-1301.
- Frisch M, Melchinger AE (2000).** The length of the intact donor chromosome segment around a target gene in marker-assisted backcrossing. *Genetics*, 157: 1343-1356.
- Garcia AAF, Benchimol LL, Barbosa AM, Geraldi IO, Souza Jr CL, Souza AP (2002).** Comparing of RAPDs, RFLPs, AFLPs, and SSRs for computation of genetic divergence in tropical maize inbred lines. A ser Submetido à *Euphytica*.
- Gardiner J, Coe EH, Melia-Hancock S, Hoisongton DA, Chao S (1993).** Development of a core RFLP map in maize using an immortalized-F2 population. *Genetics*, 134:917-930.
- Gilles PN, Wu DJ, Foster CB, Dillon PJ, Chanock SJ (1999).** Single nucleotide polymorphic discrimination by an electronic dot blot assay on semiconductor microchips. *Nat. Biotechnol.*, 17:365-370.
- Gupta PK, Balyan HS, Sharma PC, Ramesh B (1994).** Microsatellites in plants: a new class of molecular markers. *Current Science*, 70(1): 45-54.

- Grattapaglia D (1994).** Genetic mapping of quantitatively inherited economically important traits in Eucalyptus. Ph.D. Dissertation, North Carolina State University, 289 pp.
- Grattapaglia D (2000).** Além dos QTL's: desafios para a incorporação de informações de QTL's no melhoramento florestal. In: Análise de QTL no melhoramento de plantas. Pinheiro J.B., Carneiro I.F., Goiânia, FUNAPE.
- Hagiwara WE, dos Santos, JB, do Carmo, SLM (2001).** Use of RAPD to aid selection in common bean backcross breeding programs. *Crop Breeding and Applied Biotechnology*, 1 (4):355-362.
- Hamada H, Petrino MG, Kakunaga T (1982).** A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukariotic genomes. *Proc. Natl. Acad. Sci. USA*, 79: 6465 – 6469.
- Hanson WD (1959).** Early generation analysis of lengths of heterozygous with backcrossing or selfing. *Genetics*, 44:833-837.
- Harlan HV, Pope MN (1922).** The use and value of backcross in small grain breeding. *Jour. Hered.*, 13:319-322.
- Hillel J, Schaap T, Haberfeld A, Jeffreys J, Plotzky Y, Cahaner A, Lavi U (1990).** DNA Fingerprints applied to gene introgression in breeding programs. *Genetics*, 124:783-789.
- Hillel J, Verrinder Gibbins M, Etches RJ, Mcq Shaver D (1993).** Strategies for rapid introgression of a specific gene modification into a commercial pouldry flock from a single carrier. *Pouldry Sci.*, 72: 1197-1211.

- Hongtrakul V, Huetis GM, Knapp SJ (1997).** Amplified fragment length polymorphisms as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theor. Appl. Genet.*, 95:400-407.
- Hospital F (2001).** Size of donor segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. *Genetics*, 158: 1363-1379.
- Hospital F, Charcosset A (1997).** Marker-assisted introgression of quantitative trait loci. *Genetics*, 147: 1469-1485.
- Hospital F, Chevalet C, Mulsant P (1992).** Using markers in gene introgression breeding programs. *Genetics*, 132: 1199-1210.
- Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997).** More on the efficiency of marker-assisted selection. *Theor. Appl. Genet.*, 95: 1181-1189.
- Jain A, Bhatai S, Banga SS, Prakash S, Lakhmikumaran M (1994).** Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. *Theor. Appl. Genet.*, 88: 116-122.
- Jorboe SG, Beavis WD, Openshaw SJ (1994).** Prediction of responses to selection in marker-assisted backcross programs by computer simulation. In: Abstracts of the second international conference on plant genome. Scherago International Inc. 38.
- Knott DR, Taludar B (1971).** Increasing seed weight and its effects on yield, yield components, and quality. *Crop Sci.*, 11:280-283.

- Knapp SJ (1994).** Selection using molecular markers indexes. p. 1-11. In: Supplement, Joint Plant Breeding Symposia Series, American Society for Horticultural Science, Corvallis OR. 5-6 August 1994. CSSA, Madison, WI.
- Knapp SJ (1998).** Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. *Crop Sci.*, 38: 1164-1174.
- Kumar LS (1999).** DNA markers in plant improvement: An overview. *Biotechnology Advances*, 17:143-182.
- Lande R (1981).** The minimum number of genes contributing to quantitative variation between and within populations. *Genetics*, 99: 541-553.
- Lande R, Thompson R (1990).** Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 124:743-756.
- Langridge P, Lagudah ES, Holton TA, Appels R, Sharp PJ, Chalmers KJ (2001).** Trends in genetic and genome analyses in wheat: a review. *Aust. J. Agric. Res.*, 52:1043-1077.
- Lanza LLB, Souza Jr CL, Ottoboni LMM, Vieira MLC, Souza AP (1997):** Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. *Theor. Appl. Genet.*, 94, 1023-1030.
- Lee M (1995).** DNA markers and plant breeding programs. *Advances in Agronomy*, 55: 265-344.
- Lee M, Godshalk EB, Lamkey KR, Woodman WL (1989).** Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. *Crop Sci.*, 29:1067-1071.
- Lemieux B, Aharoni A, Schena M (1998).** DNA chip technology. *Mol. Breed.*, 4: 277-289.

- Lin JJ, Kuo J (1995).** AFLP: a novell PCR-based assay for plant and bacterial DNA fingerprinting. *Focus*, 17:66-71.
- Lima MLA (2000).** Análise da similaridade genética entre genótipos de cana-de açúcar (*Saccharum spp.*) detectada por AFLP e coeficiente de parentesco. Dissertação de Mestrado, Pp. 82, IB.
- Luby JJ, Shaw DV (2001).** Does Marker-assisted selection make dollars and sense in a fruit breeding programs? *HortScience*, 36(5): 872-879.
- Mangolim CA (2002).** Mapeamento de QTIs associados com o teor de óleo em linhagens tropicais de milho (*Zea mays* L.). Dissertação de Doutorado, UNICAMP, IB, Campinas.
- Melchinger AE (1990).** Use of molecular markers in breeding for oligogenic disease resistance. *Plant Breed.*, 104:1-19.
- Melchinger AE (1999).** Genetic diversity and heterosis. In: J.G. Coors and S. Pandey (eds.), Genetic and exploitation of heterosis in crops. ASA – CSSA. – SSSA, 99-118. Madison, WI, U.S.A.
- Mesquita AGG, Guimarães CT, Netto SP, Paiva E (2001).** Programa de retrocruzamento assistido por marcadores SSR para acelerar a recuperação do genótipo recorrente em milho. IV Latin-American Meeting On Plant Biotechnology - REDBIO, CD, Goiânia, Brazil
- Milach SCK, Rines HW, Phillips RL (1997).** Molecular genetic mapping of dwarfing genes in oat. *Theor. Appl. Genet.*, 95:783-790.
- Mohan M, Bair S, Bhagwat A, Krishna TG, Yano M (1997).** Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*, 3:87-103.

- Mörchen M, Cuguen J, Michaelis G, Hänni C, Saumitou-Laprade P (1996). Abundance and length polymorphism of microsatellite repeats in *Beta vulgaris* L. *Theor. Appl. Genet*, 92: 326-333.
- Moreau L, Charcosset A, Hospital F, Gallais A. (1998). Marker-assisted selection efficiency in populations of finite size. *Genetics*, 148: 1353-1365.
- Morgante M, Olivieri AM (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*, 3: 175-182.
- Murphy RW, Sites JW Jr., Buth DG, Haufler CH (1990). Proteins I: isozyme electrophoresis. In: Molecular Systematics. Hillis, DM & Moritz C (Eds.). Sinauer Associates, Sunderland MA. Pp. 45-126.
- Navieira H, Barbadilla A. (1992). The theoretical distribution of lengths of intact chromosome segments around a locus held heterozygous with backcrossing in a diploid species. *Genetics*, 130:205-209.
- Neale DB, Harry DE (1994). Genetic mapping in forest trees: RFLPs, RAPDs and beyond. *AgBiotech News Inf.*, 6: 107N-114N
- Newbury HJ, Ford-Lloyd BV (1993). The use of RAPD for assessing variation in plants. *Plant Growth Regul.*, 12: 43-51.
- O'Donoghue LS, Kianian SF, Rayapati PJ, Penner GA, Sorrells ME, Tanskley SD, Phillips RL, Rines HW, Lee M, Fedak G, Molnar SJ, Hoffman D, Salas CA, Wu B, Autrique E, Van Deynze A (1995). A molecular linkage map of cultivated oat. *Genome*, 38:368-380.
- Oliveira AC, Ritcher T, Bennetzen J (1996). Regional and racial specificities in sorghum germplasm assessed with DNA markers. *Genome*, 39:579-587.

- Openshaw SJ, Jarboe SG, Beavis WD (1994).** Marker assisted selection in backcross breeding. Pp. 41-43. In: Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon.
- Paterniani E, Risi R (1976).** Seleção entre e dentro de famílias de meios-irmãos nas populações Piranão A e Piranão B. Rel. Cient. Depto. De Genética, ESALQ, Piracicaba, 167-173.
- Paterniani E (1990).** Maize breeding in the tropics. In: *Critical Reviews in Plant Scienc.* 9:125-154.
- Paterniani E (1995).** Importância do milho na agroindústria. In: Produção e melhoramento do milho. Osuna, J. A. e Moro, J. R. (Eds.). Jaboticabal, FUNEP, 176p.
- Paterson AH, Damon S, Hewit JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991).** Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics*, 127: 181-197.
- Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Catiglioni P, Taramino G, Motto M (1998).** Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor. Appl. Genet.*, 97:1248-1255
- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996).** The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for gemplasm analysis. *Molecular Breeding*, 2: 225:238
- Rafalski JA, Morgante M, Powell W, Vogel JM, Tingey SV (1996).** Generating and using DNA markers in plants. In Birren B and Lai E (eds.). Analysis of non-mammalian genomes – a practical guide, pp. 75-134. Academic Press, NY.

- Ragot M, Biosioli M, Delbut MF, Dell'Orco A, Malgarini L, Thevenin P, Vernoy J, Vivant J, Zimmermann R, Gray G (1995).** Marker-assisted backcrossing: a practical example. In: *Techniques et utilisations des marqueurs moléculaires*. Montpellier, France. 29-31 March 1994. INRA, Paris.
- Ribaut J-M, Hoisington D (1998).** Marker-assisted selection: new tools and strategies. *Trends in Plant Science*, 3(6): 236-239.
- Ribaut J-M, Hoisington D (2002).** Simulation experiments on efficiencies of gene introgression by backcrossing. *Crop Sci.*, 42:557-565.
- Ribaut JM, Jiang C, González-de-Léon D, Edmeades GO, Hoisington DA (1997).** Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. *Theor. Appl. Genet.*, 94: 887-896.
- Rick CM (1974).** High soluble-solids content in large-fruited tomato lines derived from a wild green-fruited species. *Hilgardia*, 42: 493-510.
- Rick CM (1982).** The potential of exotic germplasm for tomato improvement and somatic cell genetics. Academic Press, New York, pp 1-28.
- Reiter R (2001).** PCR-based marker systems. In: Phillips RL & Vasil IK (Eds.), *DNA-Based Markers in Plants*, 9-29. Kluwer Academic Publishers, Netherlands.
- Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Granal MW (1995).** Abundance, variability and chromosomal location of microsatellites in wheat. *Mol. Gen. Genet.*, 246: 327-333.
- Röder MS, Korzum V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganai MW (1998).** A microsatellite map of wheat. *Genetics*, 149: 2007-2023.

- Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995).** The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet.*, 90:43-48.
- Saiki RK, Scharf SJ, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985).** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230: 1350-1354.
- Santos JB dos, Nienhuis J, Skroch P, Tivang J, Slocum MK (1994).** Comparison of RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. *Theor. Appl. Genet.*, 87:909-915.
- Sax K (1923).** The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, 8:552-560.
- Schneider KA, Brothers ME, Kelly JD (1997).** Marker-assisted selection to improve drought resistance in common bean. *Crop Sci.*, 37:51-60.
- Senior ML, Heun M (1993).** Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome*, 36: 884-889.
- Sharma HC, Gill BS (1983).** Current status of wide hybridization in wheat. *Euphytica*, 32: 17-31.
- Sibov ST (2002).** Mapeamento de locos quantitativos (QTLs) associados à produção de milho tropical através de marcadores moleculares. Dissertação de Doutorado, UNICAMP, IB, Campinas.
- Souza Jr CL (2001).** Melhoramento de espécies alógamas. In: Recursos genéticos e melhoramento – plantas. Nass LL, Valois ACC, Melo IS, Valadares-Inglis MC (Eds). Fundação MT, Pp. 1183.

- Stam P, Zeven AC (1981).** The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica*, 30: 227-238.
- Stuber CW (1994).** Enhancement of grain yield in maize using marker-facilitated introgression of QTLs. Pp. 44-46. In: Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon.
- Stuber CW, Edwards MD (1986).** Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. P. 70-83. In: Proc. 41st Annual Corn and Sorghum Industry Res. Conf. Chicago, IL. 10-11 Dec. 1985. American Seed Trade Association, Washington, DC
- Stuber CW, Goodman MM (1983).** Allozyme genotypes for popular and historically important inbred lines of corn, *Zea mays* L. USDA Agric. Res. Results, Southern Ser., No. 16.
- Tanskley SD (1983).** Molecular markers in plant breeding. *Plant. Mol. Biol. Rep.*, 1:3-8.
- Tanskley SD, McCough SR (1997).** Seed banks and molecular maps: unlocking genetic potential from wild. *Science*, 277: 1063-1066.
- Tanskley SD, Nelson JC (1996).** Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.*, 92: 191-203.
- Tanskley SD, Orton TJ (1983).** Isozymes in plant genetic and breeding (Part A & B). Elsevier Science Publishers, Amsterdam.

- Tanskley SD, Young ND, Paterson AH, Bonierbale MW (1989).** RFLP mapping in plant breeding: new tools for an old science. *Bio/Technology*, 7: 257-264.
- Taramino G, Tarchini R, Ferrario S, Lee M, Pe' ME (1997).** Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theor. Appl. Genet.*, 95:66-72.
- Taramino G, Tingey S (1996).** Simple sequence repeats for germoplasm analysis and mapping in maize. *Genome*, 39: 277-287.
- Tautz D, Renz M (1984).** Simple sequences are ubiquitous repetitive components of eukariotic genomes. *Nucleic Acids Res.*, 12: 4127-4138.
- Thoday JM (1961).** Location of polygenes. *Nature*, 191: 368-370.
- Thomas CM, Vos P, Zabeau M, Jones DA, Norcott KA, Charwick BP, Jones JDG (1995).** Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato *Cf-9* gene for resistance to *Cladosporium*. *Plant J.* 8: 785-794.
- Van Berloo R, Stam P. (1998).** Marker-assisted selection in autogamous RIL populations: a simulation study. *Theor. Appl. Genet.*, 96: 147-154.
- Visscher PM, Haley CS, Thompson R (1996).** Marker-assisted introgression in backcross breeding programs. *Genetics*, 144: 1923-1932.
- Vuylsteke M, Anoise R, Bastiaans E, Señor M, Stuber C, Kuiper M (1997).** A density AFLP linkage map of *Zea mays* L. Final Programs and Abstracts Guide, Plant and Animal genome V, Abs. P206, P104.

- Young ND (1999).** A cautiously optimistic vision for marker-assisted breeding. *Mol. Breed.*, 5:505-510.
- Young ND, Tanksley SD (1989).** Restriction Fragment length polymorphism maps and the concept of graphical genotypes. *Theor. Appl. Genet.*, 77:95-101.
- Yu K, Park SJ, Poysa V, Gepts P (2000).** Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *The Journal of Heredity*, 91: 429-434.
- Waugh R, Baird E, Powell W (1992).** The use of RAPD markers for the detection of gene introgression in potato. *Plant Cell Rep.*, 11: 466-469.
- Weck E, Bergstedt M, Farrar D, Kieler M, Krumm C, Lundell M, Mead D, Sunderland H, Taylor R, Wagen CA (1991).** RFLPs push BC2 selections ahead. *Maize Genet. Newsltr.*, 65:101-102.
- Welsh J, McClelland M (1990).** Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.*, 18: 7213-7218.
- Wight CP, Penner GA, O'Donoghue LS, Burrows VD, Molnar SJ, Fedak G (1994).** The identification of random amplified polymorphic DNA (RAPD) markers for day length insensitive in oat. *Genome*, 37: 910-914.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990).** DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, 18: 6531-6535.

- Winzeler FA, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, McCusker JH, Stevens DA, Wodicka I, Lockhart DJ, Davis RW (1998).** Direct allelic variation scanning of the yeast genome. *Science*, 281: 1194-1197.
- Zabeau M, Vos P. (1993).** Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application. Number: 9240229.7. Publication number 0534858 A 1.
- Zeven AC, Knott DR, Johson R (1983).** Investigation of linkage drag in near isogenic lines of wheat by testing for seedling reaction to races of stem rust, leaf rust and yellow rust. *Euphytica*, 32: 319-327.
- Zhao X, Kochert G (1993).** Phylogenetic distribution and genetic mapping of a (GGC)_n microsatellite form rice (*Oryza sativa* L.). *Plant Molec. Biol.*, 21: 607-614.
- Zimnoch-Guzowska E, Marczewski W, Lebecka R, Flis B, Schäfer-Pregl R, Salamini F, Gebhardt C (2000).** QTL analysis of new sources of resistance to *Erwinia carotovora* ssp *atroseptica* in potato done by AFLP, RFLP, and resistance-gene-like markers. *Crop Sci*, 40:1156-1167
- Xie C, Xu S (1998).** Strategies of marker-assisted aided recurrent selection. *Crop Sci.*, 38: 1526-1535.