

MARCO ANTONIO MACHADO

INTEGRAÇÃO DE MAPAS, MAPEAMENTO DE QTLs E ANÁLISE DA DIVERSIDADE
GENÉTICA EM SOJA UTILIZANDO MARCADORES AFLP

**INTEGRAÇÃO DE MAPAS, MAPEAMENTO DE QTLs E ANÁLISE DA DIVERSIDADE
GENÉTICA EM SOJA UTILIZANDO MARCADORES AFLP**

Tese apresentada à Universidade Federal de
Viçosa, como parte das exigências do Curso
de Genética e Melhoramento, para obtenção
do título de *Doctor Scientiae*.

Dr. Van Schuler

Prof. Elza Fernandes de Araújo

BRASIL

VIÇOSA - MINAS GERAIS

SETEMBRO - 1999

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

T

M149i
1999

Machado, Marco Antonio, 1968-
Integração de mapas, mapeamento de QTLs e análise da
diversidade genética em soja utilizando marcadores AFLP /
Marco Antonio Machado. – Viçosa : UFV, 1999.
90p. : il.

Orientador: Maurílio Alves Moreira
Tese (doutorado) - Universidade Federal de Viçosa

1. Soja - Mapeamento genético. 2. Soja - Análise de
QTLs. 3. Soja - Diversidade genética. 4. Soja - Marcadores
AFLP. 5. Soja - Genética molecular. I. Universidade Federal
de Viçosa. II. Título.

CDD 19.ed. 633.343

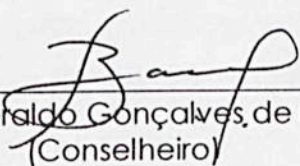
CDD 20.ed. 633.343

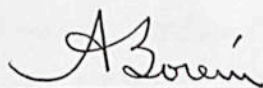
MARCO ANTONIO MACHADO


INTEGRAÇÃO DE MAPAS, MAPEAMENTO DE QTLs E ANÁLISE DA DIVERSIDADE
GENÉTICA EM SOJA UTILIZANDO MARCADORES AFLP

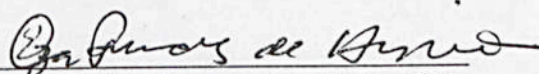
Tese apresentada à Universidade Federal de
Viçosa, como parte das exigências do Curso
de Genética e Melhoramento, para obtenção
do título de Doctor Scientiae.

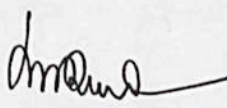
APROVADA: 28 de maio de 1999.


Prof. Everaldo Gonçalves de Barros
(Conselheiro)


Prof. Aluizio Borém de Oliveira
(Conselheiro)


Dr. Ivan Schuster


Prof. Elza Fernandes de Araújo


Prof. Maurílio Alves Moreira
(Orientador)

AGRADECIMENTO

À Universidade Federal de Viçosa, por toda a minha formação acadêmica e pela oportunidade oferecida para a realização do Curso.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de pesquisa. Aos meus pais Francisco e Coralli.

Ao Professor Maurício Alves, A Deus, pela orientação, pela amizade e pela confiança durante o Curso.

Ao Dr. Abramides do Val, pelo programa de colaboração técnico-científica entre a Monsanto Brasil e o Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO) da Universidade Federal de Viçosa, o que possibilitou o desenvolvimento de meus trabalhos de pesquisa na Monsanto de St. Louis, MO, EUA.

Ao Dr. Xavier Delannoy, pelo treinamento e pela realização de grande parte dos trabalhos de pesquisa em seu laboratório na Monsanto, em St. Louis, MO, EUA.

Ao Dr. Ivan Schuster, pela ajuda na análise dos dados, pela discussão e pela edição da tese.

Ao Professor Everilda Gonçalves de Sales, pelo aconselhamento durante a fase de pesquisa.

Aos Professores Aluizio Borém de Oliveira e Siza Fernandes de Araújo, pelas valiosas sugestões.

A Abelmar Gasteira e Inês Chammel, pela determinação dos ácidos graxos dos sementes de soja.

AGRADECIMENTO

À Universidade Federal de Viçosa, por toda a minha formação acadêmica e pela oportunidade oferecida para a realização do Curso.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de estudos.

Ao Professor Maurílio Alves Moreira, pela orientação, pela amizade e pela confiança durante o Curso.

Ao Dr. Abramides do Val, pelo programa de colaboração técnico-científico entre a Monsanto Brasil e o Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO) da Universidade Federal de Viçosa, o que possibilitou o desenvolvimento de meus trabalhos de pesquisa na Monsanto de St. Louis, MO, EUA.

Ao Dr. Xavier Delannay, pelo treinamento e pela realização de grande parte dos trabalhos de pesquisa em seu laboratório na Monsanto, em St. Louis, MO, EUA.

Ao Dr. Ivan Schuster, pela ajuda na análise dos dados, pela discussão e pela editoração da tese.

Ao Professor Everaldo Gonçalves de Barros, pelo aconselhamento durante a fase de pesquisa.

Aos Professores Alúzio Borém de Oliveira e Elza Fernandes de Araújo, pelas valiosas sugestões.

A Abelmon Gesteira e Inês Chammel, pela determinação dos ácidos graxos das sementes de soja.

Às minhas colegas de laboratório Kcal e Marta, pelo excelente convívio, pelo estímulo e pela grande amizade durante o Curso.

A todos que, de alguma forma, colaboraram para a conclusão do meu Curso, culminando com a defesa desta tese.

BIOGRAFIA

MARCO ANTONIO MACHADO, filho de Francisco Machado Filho e Elizabeth Coralli Sabioni Machado, nasceu em 15 de dezembro de 1968, em Austin, Texas, EUA.

Curso o primeiro e segundo graus em Viçosa, MG.

Em março de 1986, ingressou no Curso de Agronomia da Universidade Federal de Viçosa (UFV), em Viçosa, MG, diplomando-se em janeiro de 1991.

Em novembro de 1991, iniciou o Curso de Mestrado em Agroquímica na UFV cujo tese foi aprovada no dia 30 de junho de 1994.

Em agosto de 1994, iniciou o Curso de Doutorado em Genética e Melhoramento nessa mesma universidade, desenvolvendo, no período de dezembro de 1995 até janeiro de 1997, trabalhos de pesquisa na Monsanto Company, St. Louis, MO, EUA.

BIOGRAFIA

MARCO ANTONIO MACHADO, filho de Francisco Machado Filho e Carmélia Coralli Sabioni Machado, nasceu em 18 de dezembro de 1968, em Austin, Texas, EUA.	x
1. CURSOS	1
1.1. Curso de Agronomia da Universidade Federal de Viçosa (UFV), em Viçosa, MG, diplomando-se em janeiro de 1991.	4
1.2. Curso de Mestrado em Agroquímica na UFV, cuja tese foi aprovada no dia 30 de junho de 1994.	11
1.3. Curso de Doutorado em Genética e Melhoramento nessa mesma universidade, desenvolvendo, no período de dezembro de 1995 até janeiro de 1997, trabalhos de pesquisa na Monsanto Company, St. Louis, MO, EUA.	18
RESUMO	18
ABSTRACT	19
INTRODUÇÃO	20
MATERIAL E MÉTODOS	22
RESULTADOS E DISCUSSÃO	34
REFERÊNCIAS BIBLIOGRÁFICAS	49
IDENTIFICAÇÃO DE LOCOS DE CARACTERES QUANTITATIVOS (QTLs) PARA TIPOS DE ÁCIDOS GRAXOS EM CIMENTOS DE SOJA UTILIZANDO MARCADORES AFLP	50

RESUMO	72
ABSTRACT	73
INTRODUÇÃO	74
MATERIAL E MÉTODOS	74
RESULTADOS E DISCUSSÃO	74
REFERÊNCIAS BIBLIOGRÁFICAS	74
CONTEÚDO	
EXTRATO	viii
ABSTRACT	x
1. INTRODUÇÃO	1
1.1. Marcadores moleculares	1
1.2. Mapas de ligação	4
1.3. Mapeamento de QTLs	6
1.4. Diversidade genética	11
REFERÊNCIAS BIBLIOGRÁFICAS	14
INTEGRAÇÃO DE MAPAS DE LIGAÇÃO EM SOJA UTILIZANDO MARCADORES AFLP	18
RESUMO	18
ABSTRACT	19
INTRODUÇÃO	20
MATERIAL E MÉTODOS	22
RESULTADOS E DISCUSSÃO	34
REFERÊNCIAS BIBLIOGRÁFICAS	49
IDENTIFICAÇÃO DE LOCOS DE CARACTERES QUANTITATIVOS (QTLs) PARA TEORES DE ÁCIDOS GRAXOS EM SEMENTES DE SOJA UTILIZANDO MARCADORES AFLP	50

RESUMO	50
ABSTRACT	51
INTRODUÇÃO	52
MATERIAL E MÉTODOS	53
RESULTADOS E DISCUSSÃO	58
REFERÊNCIAS BIBLIOGRÁFICAS	70
DIVERSIDADE GENÉTICA EM ACESSOS DE <i>Glycine soja</i> E <i>Glycine max</i> UTILIZANDO MARCADORES AFLP	72
RESUMO	72
ABSTRACT	73
INTRODUÇÃO	74
MATERIAL E MÉTODOS	76
RESULTADOS E DISCUSSÃO	83
REFERÊNCIAS BIBLIOGRÁFICAS	88
2. RESUMO E CONCLUSÕES	89

EXTRATO

MACHADO, Marco Antonio, D.S., Universidade Federal de Viçosa, setembro de 1999. **Integração de mapas, mapeamento de QTLs e análise da diversidade genética em soja utilizando marcadores AFLP.** Orientador: Maurílio Alves Moreira. Conselheiros: Everaldo Gonçalves de Barros e Aluizio Borém de Oliveira.

Marcadores moleculares auxiliam no monitoramento de genes e possibilitam avaliar a variabilidade genética em programas de melhoramento. Marcadores AFLP são bastante polimórficos e apropriados para uso em culturas de base genética estreita como a soja. O primeiro objetivo deste trabalho foi a integração de um mapa de AFLP interespecífico de soja da Monsanto Company, EUA, com um mapa de domínio público desenvolvido pela Universidade de Utah, EUA. Este mapa foi gerado com uma população de 223 RILs (linhas recombinantes de autofecundação) do cruzamento de Minsoy X Noir. Esse mesmo mapa possui 468 marcadores (279 RFLP, 166 SSR e 23 clássicos), que cobrem 2.330 cM e contêm QTLs (loci de características quantitativas) identificados. DNA de 88 RILs foram amplificados, utilizando-se a técnica de AFLP, com 41 pares de *primers*, usados também no mapa da Monsanto. Dos 354 marcadores obtidos, 54 foram polimórficos para o mapa da Monsanto e serviram como pontos de ancoramento para a integração dos dois mapas. Essa integração vai permitir a transferência de QTLs do mapa Minsoy X Noir para o mapa da Monsanto. O segundo objetivo foi a identificação de QTLs para teores de ácidos graxos em soja, utilizando-se marcadores AFLP. Sementes de soja de 357 indivíduos F₂ do cruzamento da

linhagem de baixo teor de ácido linolênico, BARC-12, e a variedade CAC-1 foram analisadas quanto ao teor de ácidos graxos (palmítico, esteárico, oléico, linolênico e linoléico), por meio de cromatografia gasosa, usando-se um método não-destrutivo. Foram selecionados 89 indivíduos para análise pela técnica de AFLP. Trinta e cinco combinações de *primers* foram utilizadas, gerando 158 marcadores dominantes e 18 co-dominantes. Foram identificados QTLs para todos os ácidos graxos, sendo que um marcador co-dominante explicou 50% da variação do teor de ácido linolênico. Linhas recombinantes de autofecundação (RILs) estão sendo desenvolvidas para esse cruzamento, o que vai permitir realizar estudos mais detalhados sobre a herança do teor desses ácidos. O terceiro, e último, objetivo do presente trabalho foi estudar a variabilidade genética de 263 acessos de soja, sendo 31 *Glycine soja*, 10 *Glycine max* adaptados e 222 *Glycine max* não-adaptados, utilizando-se a técnica de AFLP. Os padrões de AFLP foram utilizados para gerar uma matriz de distâncias genéticas de Nei, que foi empregada na realização de uma análise multivariada. Os resultados confirmaram que a soja cultivada apresenta base genética muito estreita e que existe grande diversidade genética em materiais exóticos, que podem ser aproveitados em programas de melhoramento.

ABSTRACT

MACHADO, Marco Antonio, D.S., Universidade Federal de Viçosa, September, 1999. **Map integration, QTL mapping and genetic diversity in soybean using AFLP markers.** Adviser: Maurílio Alves Moreira. Committee Members: Everaldo Gonçalves de Barros and Aluizio Borém de Oliveira.

Molecular markers help monitoring genes and allow to access the genetic variability in germplasm. AFLP are very polymorphic markers and suitable to be used in crops with narrow genetic base like soybean. The first objective of this work was to integrate an interspecific AFLP map developed by Monsanto, USA with a public map developed by the University of Utah. This map was generated with 223 recombinant inbred lines from the cross of Minsoy and Noir. It has 468 markers (279 RFLP, 166 SSR and 23 Classical markers) covering 2,330 cM with many quantitative trait loci (QTLs) identified. DNA from 88 RILs from this population were analyzed by the AFLP technique with 41 primer pairs also used in the Monsanto's map. From the 354 AFLP markers generated 54 markers were also polymorphic with Monsanto map and were used as anchor points to merge the two maps. Map integration will allow the transfer of QTL data from the Minsoy X Noir map to Monsanto's map. The second objective of this work was to map quantitative trait loci (QTLs) for fatty acid content in soybean seeds. Soybean seeds of 357 F₂ individuals from a cross between the low linolenic line BARC-12 and the cultivar CAC-1 were analyzed for the content of palmitic, stearic, oleic, linoleic and linolenic acids by gas chromatography. A total of 89 plants were analyzed with the AFLP technique using 35 primer pairs that generated 158 dominant and 18 co-dominant

markers. It was possible to identify QTLs for all fatty acids and one co-dominant marker explained 50% of the variation for the content of linolenic acid. Recombinant Inbred Lines (RILs) are being developed for this cross to provide more detailed studies about the inheritance of fatty acid content in soybean seeds. The third objective of this work was to measure the genetic diversity present in unadapted and wild relatives of soybean. For that, 263 soybean accessions (31 *Glycine soja*, 10 adapted and 222 unadapted *Glycine max*) were evaluated using AFLP technique. AFLP patterns were used to generate a Nei & Li genetic distance matrix that was used to perform a multi-dimensional analysis. The results showed that cultivated soybean has a narrow genetic base and a great amount of genetic diversity is still present in exotic germplasm and could be used as a source of variability in breeding programs.

1.1.1. Marcadores moleculares

O avanço das técnicas moleculares permitiu o desenvolvimento de marcadores que podem ser usados para estudos genéticos. Os primeiros marcadores moleculares utilizados foram os marcadores que são variáveis de comprimento de fragmentos de DNA, conhecidos como RFLPs (Restriction Fragment Length Polymorphisms). Entretanto, o uso de marcadores RFLPs não é suficiente para estudos genéticos de alto resolução e o desenvolvimento de marcadores de alto resolução é necessário para o melhoramento de plantas (FRANKSLEY et al., 1989).

principalmente desenvolvida para mapeamento do genoma humano (BOTSTEIN et al., 1980). A técnica de sondagem de DNA de cópia única é utilizada em plantas de interesse e usada, como sonda, para acompanhar a regeneração de regiões homólogas do genoma de vários indivíduos numa população. Esta técnica baseia-se na digestão do DNA com vários sistemas de restrição, que cortam o DNA em regiões específicas denominadas sítios de restrição. O DNA digerido é submetido a um campo elétrico, em que os fragmentos de DNA são separados. Os estudos genéticos tradicionais são baseados na construção de mapas de ligação utilizando marcadores fenotípicos ou clássicos que, na sua maioria, são genes que determinam pigmentação, morfologia, estádios de desenvolvimento e resposta a pragas e doenças. Na cultura da soja, o mapeamento tradicional tem progredido muito vagarosamente, devido à falta de marcadores e à baixa variabilidade genética do germoplasma (SHOEMAKER, 1994). Aproximadamente, 250 marcadores morfológicos, de pigmentação e isoenzimáticos já foram identificados para a cultura da soja. No entanto, apenas 63 foram incluídos no mapa genético clássico, o que resultou numa média de apenas três marcadores em cada um dos 20 grupos de ligação da soja (PALMER e HEDGES, 1993).

1.1. Marcadores moleculares

O avanço das técnicas moleculares permitiu o desenvolvimento de marcadores que podem ser usados para estudos genéticos. Os primeiros marcadores moleculares utilizados foram as isoenzimas, que são variantes de uma enzima detectadas por diferenças na migração eletroforética em géis de amido. Esta técnica tem sido usada com sucesso em estudos genéticos, no entanto o número de marcadores obtidos não é suficiente para estudos detalhados de mapeamento no melhoramento de plantas (TANKSLEY et al., 1989).

Várias técnicas moleculares foram desenvolvidas para explorar a grande variabilidade genética existente no DNA, o que permitiu o início de uma nova era na construção de mapas de ligação.

A técnica de RFLP (*Restriction Fragment Length Polymorphism*), primeiramente desenvolvida para mapeamento do genoma humano (BOTSTEIN et al., 1980), teve grande aplicação no mapeamento de plantas. DNA de cópia única é clonado de uma espécie de interesse e usado, como sonda, para acompanhar a segregação de regiões homólogas do genoma de vários indivíduos numa população segregante. Esta técnica baseia-se na digestão do DNA com várias enzimas de restrição, que cortam o DNA em regiões específicas denominadas sítios de restrição. O DNA digerido é submetido a um campo elétrico, em que os fragmentos de DNA são separados pelo tamanho (eletroforese). Após a eletroforese, os fragmentos de DNA são transferidos para uma membrana. A diferença no tamanho dos fragmentos é detectada pelo uso de sondas marcadas que possuem homologia com regiões específicas de alguns fragmentos. O polimorfismo obtido é devido a inserções, deleções e mutações de ponto. Os fragmentos são herdados de modo co-dominante, permitindo a distinção entre indivíduos heterozigotos e homozigotos.

A técnica de RAPD (*Random Amplified Polymorphic DNA*) (WILLIAMS et al., 1990; WELSH e McCLELLAND, 1990) baseia-se numa modificação da técnica original da reação em cadeia da polimerase (PCR) que possibilita a síntese de DNA *in vitro*, a partir de um DNA-modelo. São utilizados *primers* de seqüência aleatória de 10 pares de base (pb), que vão se ligar a regiões homólogas no DNA-alvo. A amplificação ocorrerá quando dois *primers* estiverem flanqueando um segmento de DNA em direção oposta e distanciados de 500 a 2.000 pb, o que possibilita a síntese desse segmento pela *Taq* DNA polimerase. Repetidos ciclos de desnaturação do DNA-alvo, ligação dos *primers* e síntese da cadeia produzem os fragmentos de DNA, que são visualizados em gel de agarose. O polimorfismo ocorre quando: a distância entre os sítios de ligação dos *primers* é muito grande; o sítio de ligação dos *primers* não está presente (devido a inserções ou deleções entre os sítios); ou quando o sítio sofreu mutação de ponto. A técnica de RAPD amplifica vários locos em cada reação, dependendo da espécie estudada. Outra

característica é que a quantidade de DNA necessária é muito pequena. No entanto, a técnica não permite a distinção entre indivíduos heterozigotos e homozigotos, sendo, portanto, considerada dominante.

A técnica de SSR (*Simple Sequence Repeat*) ou microssatélites, introduzida em soja por AKKAYA et al. (1992), baseia-se na amplificação de regiões contendo seqüências repetidas simples, com o uso de *primers* especialmente construídos para ligar em regiões que flanqueiam os microssatélites. O polimorfismo baseia-se no número de seqüências repetidas presentes em cada indivíduo. Em soja, que possui base genética estreita, já foram detectados 23 alelos diferentes para um mesmo loco. Esse alto nível de alelismo possibilita a comparação de mapas de ligação entre diferentes variedades, podendo-se gerar um mapa-consenso. Essa técnica permite a detecção dos indivíduos heterozigotos, sendo, portanto, considerada co-dominante.

A técnica de AFLP (*Amplified Fragment Length Polymorphisms*) foi desenvolvida recentemente para mapeamento em larga escala (VOS et al., 1995). Os marcadores são fragmentos de restrição detectados após uma amplificação seletiva por PCR. O DNA genômico é digerido simultaneamente com duas enzimas de restrição, como *EcoRI* e *MseI*. A seguir, os fragmentos de DNA gerados são ligados a adaptadores nos sítios de *EcoRI* e de *MseI* e amplificados por PCR, utilizando-se *primers* que contêm seqüências complementares aos adaptadores. A seletividade é alcançada pela adição de nucleotídeos arbitrários à extremidade 3' do *primer*, causando aumento na seletividade do *primer* e resultando numa redução na complexidade dos produtos de amplificação por PCR. Quando três nucleotídeos são adicionados às extremidades dos dois *primers*, entre 50 e 100 locos são amplificados num genoma típico de planta (VOS et al., 1995). Os fragmentos de DNA amplificados são detectados pelo fracionamento em géis de seqüenciamento, o que permite detectar polimorfismos de até um par de bases. Esta técnica é extremamente poderosa na caracterização de genomas e no mapeamento genético. Em espécies, consideradas polimórficas, como o milho, pode-se obter grande número de marcadores com apenas um par de *primers*. O polimorfismo gerado por esta técnica é preferencialmente dominante, sendo possível, no entanto, identificar os indivíduos heterozigotos,

pela diferença de intensidade das bandas, se esta for realizada com bastante precisão.

1.2. Mapas de ligação

O grande avanço das técnicas moleculares permitiu a obtenção de grande número de marcadores e a elaboração de mapas de ligação detalhados para várias espécies. Na cultura da soja, os primeiros mapas de ligação foram desenvolvidos, utilizando-se marcadores RFLP. APUYA et al. (1988) desenvolveram uma biblioteca genômica de soja digerida com a enzima *Sau3AI* e clonada no fago M13 digerido com *BamHI*. Para diferenciar os clones de cópia única dos clones contendo DNA repetitivo, fez-se a hibridização dos clones com DNA genômico total de soja, marcado com radioatividade. Os clones de cópia única ou de poucas cópias foram identificados pela ausência de hibridização.

O DNA de plantas pode apresentar metilação da base citosina, e muitas enzimas de restrição não conseguem digerir o DNA metilado. Com base nesse fato, KEIM e SHOEMAKER (1988) utilizaram a enzima *PstI*, sensível à metilação, para construir uma biblioteca genômica de soja. Regiões transcritas apresentam menor metilação do DNA, podendo ser digeridas pela enzima *PstI*. Desse modo, os referidos autores conseguiram obter uma biblioteca de DNA de cópia única.

SHOEMAKER et al. (1992) identificaram marcadores RFLP polimórficos para grande número de genótipos de soja e demonstraram que um mapa detalhado de RFLP pode ser utilizado para entendimento da estrutura e da evolução do genoma, identificando-se quais ancestrais doaram os marcadores para dado genoma de uma variedade.

SHOEMAKER e OLSON (1993) desenvolveram um mapa saturado em soja numa população de 60 indivíduos F_2 do cruzamento da linhagem de *G. max* A81-356022 com o acesso de *G. soja* PI 468.916. Esse mapa com 490 marcadores contém marcadores RFLP de clones aleatórios e de genes conhecidos, marcadores RAPD, marcadores isoenzimáticos e alguns marcadores clássicos. Foram gerados 20 grupos de ligação, correspondendo a uma distância de recombinação de quase 3.000 cM.

LARK et al. (1993) desenvolveram um mapa de ligação utilizando o cruzamento intra-específico Minsoy (PI 27.890) X Noir 1 (PI 290.136). O mapa contém 133 marcadores RFLP, alguns marcadores clássicos e também isoenzimáticos. Trinta e um grupos de ligação foram gerados, correspondendo a uma distância de recombinação de 1.550 cM.

AKKAYA et al. (1995) mapearam 40 marcadores SSR em uma população F_2 de soja de 60 indivíduos de um cruzamento de duas linhagens oriundas das variedades Clark e Harosoy. Os marcadores SSR distribuíram-se aleatoriamente no genoma em relação aos marcadores RFLP e RAPD mapeados anteriormente. A inclusão dos marcadores SSR aumentou a abrangência do mapa, de 1.056 para 1.486 cM, em 29 grupos de ligação.

KEIM et al. (1997) construíram um mapa saturado de soja numa população de 300 linhagens recombinantes endogâmicas do cruzamento das variedades BSR-101 X PI 437.654. Primeiramente, foi construído um mapa-base com 165 marcadores RFLP, utilizando-se toda a população de 300 indivíduos. Quarenta e dois indivíduos foram utilizados para inclusão de 650 marcadores AFLP. O mapa final apresentou 840 marcadores, sendo 165 RFLP, 650 AFLP e 25 RAPD, abrangendo 3.441 cM em 28 grupos de ligação. Foi notada a ocorrência de aglomeração de marcadores AFLP em algumas regiões do mapa, possivelmente devido a uma redução de recombinação nessas regiões, podendo caracterizar regiões heterocromáticas, como centrômeros. Apesar dessa aglomeração, os referidos autores concluíram que os marcadores AFLP mapearam em todos os grupos de ligação e apresentaram boa distribuição em relação àquela obtida com marcadores RFLP.

É interessante que um mapa construído com marcadores moleculares apresente marcadores clássicos de fácil identificação fenotípica, pois se podem descobrir associações entre locos de caráter quantitativo (QTLs) de interesse e características morfológicas de fácil detecção. Infelizmente, os mapas de ligação, construídos com marcadores moleculares, apresentam poucos marcadores clássicos, pois as populações utilizadas para gerar esses mapas segregam para poucos locos clássicos. Com o objetivo de integrar os marcadores clássicos disponíveis aos marcadores moleculares, SHOEMAKER e SPECHT (1995) desenvolveram uma população F_2 de soja, segregando para 20 marcadores clássicos, oriunda do cruzamento de uma linhagem da variedade

Clark com uma linhagem da variedade Harosoy. Os referidos autores utilizaram 350 marcadores RFLP e oito marcadores RAPD, previamente mapeados num cruzamento interespecífico de *Glycine soja* X *Glycine max* (SHOEMAKER e OLSON, 1993). Dos 358 marcadores utilizados, 118 (33%) segregaram na população Clark X Harosoy. Foi utilizado o programa de computador JoinMap (STAM, 1993) para agrupar os dados de ambas as populações e construir grupos de ligações-consenso. Com este estudo, cerca da metade dos 19 grupos de ligação do mapa clássico foi associada a um grupo de ligação molecular.

Qi et al. (1996) construíram um mapa integrado em cevada, utilizando – se dados moleculares oriundos de quatro cruzamentos. Pelo menos 22% dos marcadores eram comuns entre dois mapas, o que permitiu a geração de um mapa integrado contendo 898 marcadores, cobrindo 1.060 cM e removendo regiões não saturadas existentes nos mapas individuais. Comparações feitas entre o mapa integrado e os mapas individuais evidenciaram que a ordem dos marcadores, na sua grande maioria, foi conservada.

A grande utilização de marcadores SSR no mapeamento genético da soja permitiu ao Dr. Perry Cregan, do USDA (Departamento de Agricultura dos EUA), realizar a integração de três mapas de soja desenvolvidos nos EUA: USDA/Universidade de Iowa, Universidade de Utah e Universidade de Nebraska (comunicação pessoal).

1.3. Mapeamento de QTLs

O estudo dos locos que controlam características quantitativas (QTLs) possibilita a dissecação de caracteres complexos nos seus componentes genéticos individuais, permitindo o seu tratamento como se fossem características de herança simples (TANKSLEY et al., 1989).

A utilização de marcadores moleculares, para mapear e caracterizar QTLs, está fundamentada no desequilíbrio de ligação existente numa população segregante, fazendo com que sejam criadas associações entre os marcadores e os QTLs. A população segregante utilizada pode ser de retrocruzamento, F₂ ou RILs (linhagens recombinantes endogâmicas), dentre outras. A população de retrocruzamento é pouco informativa, pois segrega

apenas para os alelos do progenitor doador. A população F_2 , por sua vez, é duas vezes mais informativa do que a de retrocruzamento e apresenta todas as combinações possíveis de alelos dos progenitores (AA, Aa, aa). A presença dos heterozigotos torna possível o estudo dos efeitos aditivos e dos desvios de dominância. Tais características fazem com que as populações F_2 sejam bastante utilizadas no mapeamento de plantas. A maior limitação das populações F_2 e de retrocruzamento é apresentar uma natureza não estática, ou seja, a autofecundação destas populações vai dar origem a populações diferentes, F_3 e RC_1F_2 , respectivamente, dificultando a análise de QTLs em diferentes localidades. As RILs constituem uma população permanente, visto que cada indivíduo é resultante de autofecundações sucessivas de uma planta F_2 . Os vários ciclos de meiose requeridos para alcançar a homozigose oferecem mais oportunidade para eventos de recombinação. Desse modo, os mapas obtidos por meio de RILs apresentam maior resolução, se comparados com os mapas de F_2 . Uma das principais características dessa população é que ela permite o estudo da interação QTLs x ambiente, uma vez que as mesmas linhas podem ser plantadas em diferentes ambientes. Essas características fazem com que esse tipo de população seja a mais utilizada atualmente no mapeamento de QTLs. A sua principal limitação é o tempo requerido para sua obtenção, cerca de sete ou mais gerações de autofecundação.

A detecção de QTLs pode ser feita por meio de dois procedimentos principais: modelos lineares e modelos utilizando máxima verossimilhança.

Os modelos lineares são a maneira mais simples para detecção de QTLs. A distribuição dos valores fenotípicos é analisada para cada marcador separadamente. Modelos lineares simples, como ANOVA e regressão linear, são métodos de fácil execução, bastante flexíveis e não necessitam de mapas de ligação. Por utilizarem apenas diferenças entre as médias dos marcadores não é possível estimar a magnitude do efeito do QTL e a sua posição no genoma (LYNCH e WALSH, 1998).

Os métodos de máxima verossimilhança são os mais utilizados no mapeamento de QTLs. Esse modelo leva em consideração toda a informação da distribuição marcador/fenótipo, tendo, portanto, mais poder de resolução. Esses métodos são mais complexos e requerem a utilização de programas de

computador específicos. LANDER e BOTSTEIN (1989) desenvolveram o método de mapeamento por intervalo, que se baseia nas informações da segregação de pares de marcadores adjacentes como unidade de análise, estimando-se a frequência de recombinação e a magnitude do efeito do QTL no intervalo entre dois marcadores ligados no mapa de ligação. O resultado da análise é mostrado por um mapa de LOD scores (verossimilhança das desigualdades) em função da posição do QTL no mapa. Em cada experimento, grande número de testes para associação marcador/fenótipo é realizado, dificultando o estabelecimento de níveis de significância para os testes de hipótese estatística. Isso faz com que possam ocorrer QTLs falso-positivos. Uma metodologia eficaz, desenvolvida para obtenção de níveis de significância geral, utiliza testes de permutação não-paramétricos, baseados na reamostragem dos dados originais (DOERGE e CHURCHILL,1996). Os dados fenotípicos são aleatoriamente reamostrados, gerando uma amostra contendo os dados moleculares originais, mas com os dados fenotípicos aleatoriamente correlacionados com os dados moleculares. O teste estatístico é, então, computado nesta amostra, cujo procedimento é repetido muitas vezes.

MANSUR et al. (1993) identificaram QTLs para 15 características agronômicas, numa população de linhagens recombinantes endogâmicas de soja, oriundas do cruzamento entre as variedades Minsoy e Noir 1. Nessa população, esses autores encontraram segregação transgressiva para várias características, ou seja, a média da população apresenta faixa maior de variação, quando comparada aos valores dos pais. Foi utilizado mapeamento por intervalo para os marcadores ligados (mapeados) e ANOVA para os marcadores não-ligados. QTLs de grande expressão foram identificados pelo marcador de RFLP, R79, que explicou 31% da variação para maturidade, e G173, que explicou 53% da variação para altura de planta.

LEE et al. (1996a), trabalhando com uma população de linhas derivadas da geração F_4 , do cruzamento entre a variedade de soja Young e PI 416937, identificaram QTLs para altura de planta, acamamento e maturidade. Os dados de campo foram obtidos em quatro localidades, no intuito de verificar a consistência da detecção de QTLs em diferentes localidades. Utilizaram um mapa contendo 137 marcadores RFLP, abrangendo 1.600 cM. Dois

marcadores para altura de planta, Blt 043 e A063a, e um marcador para acamamento, A169, foram detectados em todas as localidades. Esses mesmos autores concluíram que o nível de consistência dos QTLs, nas diferentes localidades, depende da característica analisada.

LEE et al. (1996b) identificaram QTLs adicionais para altura de planta, acamamento e maturidade, numa população de soja do cruzamento de Coker 237 X PI 97100, segregando para hábito de crescimento. O marcador Dt1 explicou 67% da variação para altura de planta e 56,4% da variação para acamamento. Duas regiões, entre os marcadores R051 e N100 e entre os marcadores B032 e CpTI, explicaram 31,2 e 26,2% da variação para maturidade, respectivamente. Os QTLs foram consistentes nas diferentes localidades, e os referidos autores concluíram que isso pode ser devido ao fato de que foram identificados poucos QTLs de grande efeito.

MANSUR et al. (1996) utilizaram uma população de linhagens recombinantes endogâmicas de soja, do cruzamento de Minsoy X Noir 1, para mapear QTLs para várias características agrônômicas: dias para florescimento, dias para maturidade, período reprodutivo, altura de planta, acamamento, teor de óleo, teor de proteína, tamanho da semente, produtividade, número de sementes, largura da folha, comprimento da folha e área foliar. Foi construído um mapa contendo marcadores RFLP, SSR e marcadores clássicos, abrangendo aproximadamente 2.000 cM. Para muitas características, poucos QTLs explicaram grande proporção da variação observada. QTLs para a maioria das características foram mapeados em apenas três grupos de ligação e, freqüentemente, no mesmo loco genético. Devido ao nível de resolução do mapa, esses autores não conseguiram determinar se os QTLs apresentam efeitos pleiotrópicos ou se são vários genes ligados numa pequena região.

BRUMMER et al. (1997) avaliaram oito populações de soja do meio-oeste americano, para determinação de teores de óleo e proteína. Eles encontraram QTLs estáveis e instáveis à variação ambiental em todas as oito populações para os teores de óleo e proteína. Os QTLs encontrados foram sensíveis à variação ambiental e à composição genética de cada cruzamento, embora alguns QTLs foram estáveis em diferentes populações e em diferentes ambientes.

LANDER e BOTSTEIN (1989) e DARVASI e SOLLER (1992) descreveram uma metodologia prática, para detecção de QTLs, denominada genotipagem seletiva. A maioria dos caracteres quantitativos é mais fácil de ser medida do que a obtenção de marcadores moleculares. Portanto, é mais fácil avaliar a população quanto aos fenótipos desejados e analisar apenas uma parte selecionada (extremos da curva normal) da população com marcadores moleculares. Essa técnica se baseia no fato de que a maioria da informação sobre a ligação marcador-QTL vem dos indivíduos com fenótipos extremos.

Seguindo o mesmo fundamento da genotipagem seletiva, MICHELMORE et al. (1991) desenvolveram uma técnica denominada Análise da Segregação de Bulks (*Bulked Segregant Analysis* ou BSA). Os fenótipos extremos de cada lado da curva normal são selecionados para a formação de dois bulks de DNA dos indivíduos. Um bulk possui o alelo para o QTL e o outro bulk o alelo alternativo para este QTL. Marcadores não-relacionados com o QTL segregam aleatoriamente entre os dois bulks. Esse método requer populações grandes para que os bulks sejam bem contrastantes e identifica principalmente QTLs de maior efeito. No entanto, esse é um método que permite seleção rápida de marcadores ligados a QTLs. A análise de segregação de bulks também pode ser usada para o mapeamento preciso de QTLs, ou seja, saturar uma região associada a um QTL com marcadores adicionais. São construídos bulks baseados nos alelos alternativos dos marcadores ligados ao QTL, fazendo com que novos marcadores possam ser encontrados, saturando a região em questão (GIOVANNONI et al., 1991).

WANG e PATERSON (1994) concluíram que estratégias como genotipagem seletiva e BSA não são efetivas para detecção de QTLs de menor efeito, que normalmente influenciam uma característica complexa. EDWARDS e PAGE (1994) demonstraram, por meio de simulações em computador, que a seleção assistida por marcador (MAS) foi mais eficiente quando poucos QTLs explicam a maior parte da variação para a característica analisada. Genotipagem seletiva e BSA são técnicas que diminuem o trabalho de mapeamento necessário para localizar QTLs (MIKLAS et al., 1996; MANSUR et al., 1993) e se adaptam bem aos melhoristas que estão interessados na identificação de QTLs de maior efeito, objetivando uma seleção assistida por marcador (MAS).

1.4. Diversidade genética

O melhoramento da cultura da soja trouxe ganhos genéticos consideráveis. No Brasil, a média de está em torno de 2.300 kg/ha, com perspectivas de aumento (EMBRAPA, 1998). No entanto, todo esse ganho genético está sendo acompanhado de erosão genética.

HIROMOTO e VELLO (1986) constataram que 100% do conjunto gênico das variedades, recomendadas para o ano agrícola 1983/1984, eram derivadas de apenas 26 ancestrais. Nove desses ancestrais foram responsáveis por mais de 80% do conjunto gênico, e apenas quatro ancestrais contribuíram com cerca de 50% desse conjunto. Esses autores concluíram que existe necessidade de aumentar a base genética das variedades brasileiras, devido à vulnerabilidade genética do germoplasma, para aumentar a produção.

Nos Estados Unidos, situação semelhante foi encontrada. DELANNAY et al. (1983) constataram que 10 ancestrais contribuíram com 88% do conjunto gênico do germoplasma do norte e apenas sete ancestrais contribuíram com 70% do germoplasma do sul.

GIZLICE et al. (1993), estudando 14 ancestrais de soja e utilizando coeficientes de parentesco, concluíram que 25% da base genética foi erodida nos conjuntos gênicos do norte e sul dos Estados Unidos. Se a taxa de endogamia continuar constante, metade da diversidade genética da coleção de soja será perdida em 58 anos no conjunto gênico do norte e, em 89 anos, no conjunto gênico do sul dos Estados Unidos.

Marcadores moleculares vêm sendo bastante utilizados como ferramenta para estudo da variabilidade genética em plantas. ABDELNOOR et al. (1995) avaliaram a diversidade genética em 38 variedades brasileiras de soja, utilizando marcadores RAPD. Foram encontrados cinco subgrupos distintos A1, A2, B1, B2 e B3, no nível de 76% de distância genética relativa. Os resultados estão de acordo com o coeficiente de parentesco. As 14 variedades do grupo A apresentaram, pelo menos, 50% do conjunto gênico derivado da PI 60406 e Tanloxi, ao passo que as 24 variedades do grupo B não evidenciaram contribuição desses progenitores. No subgrupo B1, composto de 14 variedades, três ancestrais (Roanoke, Tokyo e PI 54610) contribuíram com pelo menos 37% do conjunto gênico desse subgrupo. Todas as quatro

variedades que compõem o subgrupo B2 apresentam a PI 240664 como ancestral comum. Quatro, das seis variedades que compõem o subgrupo B3, apresentam a PI 37335 como ancestral comum.

POWELL et al. (1996) utilizaram uma combinação de marcadores SSR nucleares e de cloroplasto para investigar os níveis e padrões de variabilidade detectados em genótipos de *G. max* e *G. soja*. Numa análise previamente realizada em 700 genótipos com 115 sondas de RFLP, foram detectados 12 acessos, que representam 92% da variabilidade desse conjunto gênico. Esses 12 genótipos, juntamente com outros 37 acessos de *G. max* e *G. soja*, foram avaliados com 11 SSRs nucleares, que detectaram 129 alelos. Os 12 genótipos foram responsáveis por 40% da variabilidade alélica detectada com os SSRs. Apesar de os SSRs serem, por natureza, multialélicos, os genótipos se agruparam de acordo com a espécie. Além dos 11 SSRs nucleares, dois SSRs de cloroplastos, maternalmente transmitidos, identificaram um total de sete haplótipos.

MAUGHAN et al. (1996) investigaram 23 acessos de soja cultivada e silvestre pela técnica de AFLP. Foram utilizados 15 pares de *primers*, que geraram 759 fragmentos, numa média de 51 fragmentos para cada par de *primer*. Do total de fragmentos, 274 (36%) foram polimórficos. Desses, 127 (17%) foram polimórficos em *G. max* e 237 (31%) em *G. soja*. O número de marcadores polimórficos detectados para cada par de *primer* variou de 9 a 27. A diversidade genética determinada por AFLP foi maior nos acessos silvestres do que na soja cultivada. A análise de agrupamento claramente separou os acessos de *G. soja* dos de *G. max*. Dentro do grupo de *G. max*, as variedades de soja cultivada se agruparam de forma mais densa, demonstrando a pequena diversidade genética da soja cultivada.

O primeiro objetivo deste trabalho foi promover a integração do mapa de AFLP de soja construído pela Monsanto (St. Louis, MO, EUA) com um mapa de domínio público desenvolvido pela University of Utah (Salt lake City, UT, EUA). O segundo objetivo foi a identificação de QTLs para teores de ácidos graxos em sementes de soja, utilizando-se marcadores AFLP. O terceiro objetivo foi avaliar a diversidade genética de linhagens não-adaptadas e de parentes silvestres da soja, por meio da técnica de AFLP.

Os artigos que compõem esta tese são apresentados nas normas da revista *Genetics and Molecular Biology*, publicada pela Sociedade Brasileira de Genética.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDELNOOR, I.V., BARROS, E.G., MOREIRA, M.A. Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA technology and comparative analysis with pedigree data. *Rev. Bras. Genet.*, v.18, n.2, p. 265-273, 1995.
- AKKAYA, M.S., BHAGWAT, A.A., CREGAN, P.B. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, v.132, n.3, p.1131-1137, 1992.
- AKKAYA, M.S., SHOEMAKER, R.C., SPECHT, J.E., BHAGWAT, A.A., CREGAN, P.B. Integration of simple sequence repeat DNA markers into soybean linkage map. *Crop Sci.*, v.35, n.4, p.1439-1445, 1995.
- APIYA, N.R., FRAZIER, B.L., KIM, P., ROTH, E.J., LARK, T.G. Restriction fragment length polymorphisms of gl-WAC markers in soybean. *Glycine max (L) Merr. Theor. Appl. Genet.*, v.75, n.2, p. 889-901, 1993.
- BOYSTEIN, D., WHITE, R.L., MCKINCK, M., DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, v.32, n.1, p. 314-331, 1980.
- BRUMMER, E.C., GRAFF, G.L., ORE, J., WELCH, J.B., SHOEMAKER, R.C. Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Sci.*, v.37, n.1, p. 370-378, 1997.
- DARVASI, A., SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theor. Appl. Genet.*, v.85, n.1, p. 353-357, 1992.
- DELABINAY, Y., RODGERS D.M., FALGOUT, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. *Crop Sci.*, v.23, n.2, p. 944-949, 1983.
- DOERGE, R.W., CRITCHFIELD, G.A. Population tests for multiple loci affecting a quantitative character. *Genetics*, v.142, n.1, p. 283-294, 1996.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDELNOOR, R.V., BARROS, E.G., MOREIRA, M.A. Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. **Rev. Brasil. Genet.**, v.18, n.2, p. 265-273, 1995.
- AKKAYA, M.S., BHAGWAT, A.A., CREGAN P.B. Length polymorphisms of simple sequence repeat DNA in soybean. **Genetics**, v.132, n.3, p.1131-1139, 1992.
- AKKAYA, M.S., SHOEMAKER, R.C., SPECHT, J.E., BHAGWAT, A.A., CREGAN, P.B. Integration of simple sequence repeat DNA markers into soybean linkage map. **Crop Sci.**, v.35, n.4, p.1439-1445, 1995.
- APUYA, N.R., FRAZIER, B.L., KEIM, P., ROTH, E.J., LARK, K.G. Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. **Theor. Appl. Genet.**, v.75, n.2, p. 889-901, 1988.
- BOTSTEIN, D., WHITE, R.L., SKOLNICK, M., DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **Am J. Hum. Genet.**, v.32, n.1, p. 314-331, 1980.
- BRUMMER, E.C., GRAEF, G.L., ORF, J., WILCOX, J.R., SHOEMAKER, R.C. Mapping QTL for seed protein and oil content in eight soybean populations. **Crop Sci.**, v.37, n.1, p. 370-378, 1997.
- DARVASI, A., SOLLER, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theor. Appl Genet.**, v.85, n.1, p. 353-359, 1992.
- DELANNAY, X., RODGERS D.M., PALMER, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Sci.**, v.23, n.2, p. 944-949, 1983.
- DOERGE, R.W., CHURCHILL, G.A. Permutation tests for multiple loci affecting a quantitative character. **Genetics**, v.142, n.1, p. 285-294, 1996.

- DON, R.H., COX P.T., WAINWRIGHT B.J., MATTICK J.S. "Touchdown" PCR to circumvent spurious priming during gene amplification. **Nuc. Acid. Res.**, v.19, n.4, p.4008, 1991.
- EDWARDS, M.D., PAGE, N.J. Evaluation of marker-assisted selection through computer simulation. **Theor. Appl. Genet.**, v.88, n.1, p. 376-382, 1994.
- EMBRAPA. Centro Nacional de Pesquisa de Soja (Londrina, PR). **Recomendações técnicas para a cultura da soja na região central do Brasil 1998/99**. Londrina, 1998. 164p. (Embrapa-Soja. Documentos, 98).
- GIOVANNONI, J.J., WING, R.A., GANL, M.W., TANKSLEY, S.D. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. **Nucl. Acids Res.**, v.19, n.5, p. 6553-6558, 1991.
- GIZLICE, Z., CARTER JR., T.E., BURTON, J.W. Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. **Crop Sci.**, v. 33, n.2, p. 614-620, 1993.
- HIROMOTO, D.M., VELLO, N.A. The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. **Rev. Brasil. Genet.**, v.9, n.1, p. 295-306, 1986.
- KEIM, P., SHOEMAKER, R.C. Construction of a random recombinant DNA library that is primarily single copy sequences. **Soy Genet. News**, v.15, n.1, p.147-148, 1988.
- KEIM, P., SCHUPP, J.M., TRAVIS, S.E., CLAYTON, K., FERREIRA, A., WEBB, D.M. A high-density soybean genetic map based upon AFLP markers. **Crop Sci.**, v.37, n.2, p. 537-543, 1997
- LANDER E.S., GREEN P., ABRAHAMSON J., BARLOW A., DALY M.J., LINCOLN S.E., NEWBURG, L. MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, v.1, n.1, p.174-181, 1987.
- LANDER, E.S., BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v.121, n.1, p.185-199, 1989.
- LARK, K.G., WEISEMANN, J.F., MATTHEWS, B.F., PALMER, R., CHASE, K., MACALMA, T. A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy and Noir 1'. **Theor. Appl. Genet.**, v.86, n.3, p. 901-906, 1993.
- LEE, S.H., BAILEY, M.A. MIAN, M.A. R., CARTER JR., T.E., ASHLEY, D.A., HUSSEY, R.S., PARROTT, W.A., BOERMA, H.R. Molecular markers associated with soybean plant height, lodging, and maturity across locations. **Crop Sci.**, v.36, n.2, p.728-735, 1996a.
- LEE, S.H., BAILEY, M.A. MIAN, M.A.R., SHIPE, E.R., ASHLEY, D.A., PARROTT, W.A., HUSSEY, R.S., BOERMA, H.R. Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. **Theor. Appl. Genet.**, v.92, n.2, p. 516-523, 1996b.

- LYNCH, M., WALSH, B. **Genetics and analysis of quantitative genetics**. Sinauer Associates. Sunderland, Massachusetts. 980 p. 1998.
- MANSUR, L.M., LARK, K.G., KROSS, H., OLIVEIRA, A.B. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). **Theor. Appl. Genet.**, v.86, n.3, p.907-913, 1993.
- MANSUR, L.M., ORF, J., LARK, K.G. Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean. **Theor. Appl. Genet.**, v.86, n.3, p.914-918, 1993.
- MANSUR, L.M., ORF, J.H., CHASE, K. JARVIK, T., CREGAN, P.B., LARK, K.G. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. **Crop Sci.**, v.36, n.3, p.1327-1336, 1996.
- MAUGHAN, P.J., SAGHAI MAROOF, M.A., BUSS G.R., HUESTIS, G.M. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic line analysis. **Theor. Appl. Genet.**, v.93, n.2, p.392-401, 1996.
- MICHELMORE, R.W, PARAN, I, KESSELI, R.V., Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genome regions by using segregating populations. **Proc. Natl. Acad. Sci. USA**, v.88, n.6, p.9828-9832, 1991.
- MIKLAS, P.N., JOHNSON, E., STONE, V., BEAVER, J.S., MONTOYA, C., ZAPATA, M. Selective mapping of QTL conditionig disease resistance in common bean. **Crop Sci.**, v.36, n.4, p.1344-1351, 1996.
- PALMER, R. G., HEDGES, B.R. **Linkage map of soybean (*Glycine max* L. Merr.)** In: Genetic Maps: locus maps of complex genomes. Cold Spring Harbour Laboratory Press. New York. 1993, p. 6139-6148.
- POWELL, W., MORGANTE, M., DOYLE, J.J., McNICOL, J.W., TINGEY, S.V., RAFALSKI, A.J. Genepool variation in genus *Glycine* sugenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. **Genetics**, v.144, n.2, p.793-803, 1996.
- QI, X., STAM, P., LINDHOUT, P. Comparison and integration of four barley genetic maps. **Genome**, v.39, n.1, p. 379-394, 1996.
- ROHLF, F.J. **NTSYS-PC numerical taxonomy and multivariate analyses system**. Version 1.8 Exeter Publ., Setauket, NY, EUA. 1993.
- SAS Institute. **SAS/STAT users guide**. 6.09 ed. SAS inst., Cary, NC, EUA. 1989.
- SHOEMAKER, R.C., OLSON, T.C. **Molecular linkage map of soybean (*Glycine max* L. Merr.)**. In: Genetic Maps: locus maps of complex genomes. Cold Spring Harbour Laboratory Press, New York. 1993, p. 6131-6138.
- SHOEMAKER, R.C., SPECHT, J.E. Integration of the soybean molecular and classical genetic linkage groups. **Crop Sci.**, v.35, n.2, p.436-446, 1995.

SHOEMAKER, R.C. **RFLP map of soybean**. In: DNA-Based markers in plants (R. L. Phillips e I. K. Vasil, Eds). Kluwer Academic Publishers. Netherlands. 1994, p. 299-309.

SHOEMAKER, R.C., GUFFY, R.D., LORENZEN, L.L., SPECHT, J.E. Molecular genetic mapping of soybean: map utilization. **Crop Sci.**, v.32, n.3, p.1091-1098, 1992.

STAM, P. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. **Plant J.**, v.3, n.3, p.739-744, 1993.

TANKSLEY, S.D., YOUNG N.D., PATERSON, A.H., BONIERBALE, M.W. RFLP mapping in plant breeding: new tools for an old science. **Bio/Technology**, v.7, n.1, p.257-265, 1989.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M., ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucl. Acids Res.**, v.23, n.5, p.4407-4414, 1995.

WANG, G.L., PATERSON, A.H. Assessment of DNA pooling strategies for mapping of QTLs. **Theor. Appl. Genet.**, v.88, n.2, p. 355-361, 1994.

WELSH, J., McCLELLAND, M. Fingerprinting genomes using PCR with arbitrary primers. **Nucl. Acids Res.**, v.18, n.7, p.7213-7218, 1990

WILLIAMS, J.G.K., KUBELICK, A.R., LIVAK, K.J., RAFALSKI, J.A., TINGEY, S.V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucl. Acids Res.**, v.18, n.6, p. 6531-6535, 1990.

RESUMO

A cultura da soja é amplamente estudada em todo o mundo e requer mapas genéticos que possam ser utilizados por vários grupos de pesquisa. A integração de mapas genéticos possibilita o intercâmbio de informações entre populações distintas, trazendo maior flexibilidade na utilização dos mapas genéticos no melhoramento. Este trabalho teve como objetivo a integração do mapa de AFLP de soja construído pela Monsanto Company (St. Louis, MO, EUA), desenvolvido a partir de uma população de 92 indivíduos F_2 do cruzamento interspecífico de HS-1 (*Glycine max*) X PI 326581 (*Glycine soja*), com um mapa de domínio público desenvolvido pela Universidade de Utah (Salt Lake City, UT, EUA). Este último mapa foi gerado com uma população de 223 RILs (linhas recombinantes de autofecundação) oriundas do cruzamento entre duas linhas não-adaptadas de soja, Minsoy e Noir 1. Ele possui 468 marcadores (279 RFLP, 166 SSR e 23 clássicos), que cobrem 2.330 cM e contêm QTLs (*loci* de características quantitativas) identificados para vários caracteres agrônômicos de interesse. DNA de 88 RILs de Minsoy X Noir 1 foram amplificados, utilizando-se a técnica de AFLP, com 41 pares de *primers* empregados na construção do mapa da Monsanto. Foram gerados 354 marcadores de AFLP, sendo 54 polimórficos também para o mapa da Monsanto e utilizados como pontos de ancoramento para a integração dos dois mapas. Foi encontrada uma boa colinearidade entre ambos os mapas, o que vai permitir a transferência de informações entre o mapa de domínio público, Minsoy X Noir 1, e o mapa da Monsanto, HS-1 X PI 326581.

Palavras-Chave: integração de mapas, mapa de ligação, AFLP, DNA e soja.

INTEGRATION OF TWO SOYBEAN LINKAGE MAPS USING AFLP TECHNOLOGY

ABSTRACT

Soybean is widely studied all over the world requiring flexible genetic maps to be used by a great number of research groups. Map integration allows exchange of information between distinct populations and provides a wider range of utilization for genetic maps in plant breeding. The objective of this study was to integrate an AFLP map developed by Monsanto Company (St. Louis MO, USA) of 92 F₂ individuals of a interspecific cross HS-1 (*Glycine max*) X PI 326581 (*Glycine soja*) with a public map developed by the University of Utah. This map was generated with 223 recombinant inbred lines from the cross of two unadapted *G.max* lines Minsoy and Noir 1. It has 468 markers (279 RFLP, 166 SSR and 23 Classical markers) covering 2,330 cM with many quantitative trait loci (QTLs) identified for many agronomic traits. DNA from 88 RILs from this population were analyzed by the AFLP technique with 41 primer pairs also used in the Monsanto map. Three hundred fifty four AFLP markers were generated with 54 markers being also polymorphic with Monsanto map serving as anchor points to merge the two maps. Good collinearity was found between the two maps what will allow the transfer of information between the public map and the Monsanto map.

Key Words: map integration, linkage map, AFLP, DNA and soybean.

INTRODUÇÃO

O avanço das técnicas moleculares permitiu a obtenção de grande número de marcadores genéticos que exploram a variabilidade existente no DNA, possibilitando a elaboração de mapas genéticos detalhados para várias espécies vegetais. Os mapas genéticos podem ser utilizados pelos melhoristas para o monitoramento de genes de interesse para o programa de melhoramento, por meio de seleção assistida por marcadores. A cultura da soja (*Glycine max* (L.) Merrill), que é amplamente estudada, requer mapas genéticos que possam ser utilizados por diferentes grupos de pesquisa, visando ao intercâmbio de informações.

Os mapas genéticos podem ser utilizados pelos melhoristas para monitoramento de genes de interesse, numa seleção assistida por marcadores, e também para clonagem de genes baseada em mapa, em pesquisas na área de genética molecular. Os mapas devem, também, possuir marcadores clássicos com fenótipos de fácil detecção e grande número de marcadores de DNA bem distribuídos por todo o genoma (Shoemaker e Specht, 1995). A flexibilidade de um mapa vai depender da utilização de marcadores que possam ser empregados em outras populações diferentes da população utilizada para gerar o mapa. Essa flexibilidade ocorre em razão da diversidade genética dos progenitores utilizados para gerar o mapa. Mapas com progenitores geneticamente bem distantes apresentam maior chance de utilização de seus marcadores em outras populações, se comparados com mapas gerados com progenitores menos divergentes entre si. Outra característica bastante importante na flexibilidade de utilização de marcadores em outras populações diz respeito à própria natureza do marcador. Este pode ser polimórfico entre dois progenitores divergentes e não apresentar polimorfismo em variedades-elite de interesse para o programa. O ideal seria identificar marcadores com maior probabilidade de serem polimórficos em uma grande extensão do germoplasma (Shoemaker *et al.*, 1992).

Vários mapas genéticos de ligação, utilizando-se marcadores moleculares, foram desenvolvidos para a cultura da soja. Shoemaker e Olson (1993) desenvolveram um mapa saturado numa população de 60 indivíduos F₂

do cruzamento da linhagem de *G. max* A81-356022 com o acesso de *G. soja* PI 468.916. Esse mapa contém 490 marcadores, como marcadores RFLP de clones aleatórios e de genes conhecidos, marcadores RAPD, marcadores isoenzimáticos e alguns marcadores clássicos. Foram gerados 20 grupos de ligação, que correspondem a uma distância de recombinação de quase 3.000 cM. Keim *et al.* (1997) construíram um mapa saturado de soja numa população de 300 linhagens recombinantes endogâmicas do cruzamento das variedades BSR-101 X PI 437.654. Primeiramente, foi construído um mapa-base com 165 marcadores RFLP, utilizando-se toda a população de 300 indivíduos. Quarenta e dois indivíduos foram empregados para a inclusão de 650 marcadores AFLP. O mapa final apresentou 840 marcadores, sendo 165 RFLP, 650 AFLP e 25 RAPD, abrangendo 3.441 cM em 28 grupos de ligação.

Para aumentar a utilização dos mapas genéticos, tem sido feito esforço no sentido de integrar os mapas existentes. A integração de mapas permite a troca de informações sobre a consistência dos grupos de ligação e possibilita a transferência de características de interesse agrônomo mapeadas entre os mapas. Shoemaker e Specht (1995) integraram o mapa genético clássico da soja com um mapa genético molecular desenvolvido por Shoemaker e Olson (1993), o qual apresentava 350 marcadores RFLP e oito marcadores RAPD. Perry Cregan (comunicação pessoal), utilizando marcadores SSR, realizou a integração de três mapas genéticos de ligação desenvolvidos nos EUA: USDA/Universidade de Iowa, Universidade de Utah e Universidade de Nebraska.

Este trabalho teve como objetivo a integração do mapa genético de soja desenvolvido pela Monsanto (St. Louis, MO, EUA) com um mapa de domínio público desenvolvido pela Universidade de Utah (Salt Lake City, UT, EUA), utilizando-se a tecnologia do AFLP (*Amplified Fragment Length Polymorphisms*) (Vos *et al.*, 1995).

MATERIAL E MÉTODOS

Material genético

A integração de mapas envolveu um mapa desenvolvido pela Monsanto (St. Louis, MO, EUA) com um mapa de domínio público desenvolvido pela Universidade de Utah (Salt Lake City, UT, EUA). O mapa da Monsanto é oriundo de um cruzamento interespecífico dos genótipos de soja HS-1 (*G. max*, variedade comercial da Hartz Seeds Co, EUA) com a PI 326581, que é um acesso de *G. soja*. Foram utilizados 92 indivíduos F_2 , com 58 combinações de *primer*, para gerar um mapa com 480 marcadores AFLP e 53 grupos de ligação, cobrindo uma extensão de 2.365 cM (Figura 1). O mapa de domínio público foi desenvolvido a partir de 223 RILs (*Recombinant Inbred Lines*) originadas de 15 autofecundações, a partir da F_2 , do cruzamento entre os genótipos Minsoy (PI 27890) X Noir 1 (PI 290136), ambos *G. max*. O mapa contém 468 marcadores (279 RFLP, 166 SSR e 23 Clássicos), com 33 grupos de ligação cobrindo uma extensão de 2.330 cM (Figura 2). Diversas instituições de pesquisa estão avaliando esta população quanto a várias características fenotípicas, como: dias para florescimento, dias para maturação, altura de planta, acamamento, produtividade, tamanho da semente e teor de óleo e proteína da semente, dentre outras. Todos esses dados fenotípicos gerados, aliados aos dados moleculares, vão possibilitar a identificação de grande número de locos de caracteres quantitativos (QTLs) nessa população, o que vai ser de grande utilidade para os melhoristas de soja. O interesse da Monsanto em integrar o seu mapa de AFLP com o mapa Minsoy X Noir 1 é possibilitar a transferência dos QTLs identificados.

Com base em valores fenotípicos (dados não-mostrados), 88 linhas foram escolhidas na população Minsoy X Noir 1, constituída de 223 RILs para serem analisadas pela técnica de AFLP.

Figura 1 – Mapa genético de ligação desenvolvido pela Monsanto (St. Louis, MO, EUA), derivado do cruzamento entre o HS-1 (genótipo comercial variedade comercial da Hartz Seeds Co, EUA) e a PI 326581 (genótipo acessão). Foram utilizados 92 indivíduos F_2 e 58 pares de primers para gerar um mapa com 480 marcadores AFLP e 53 grupos de ligação, cobrindo uma extensão de 2.365 cM.

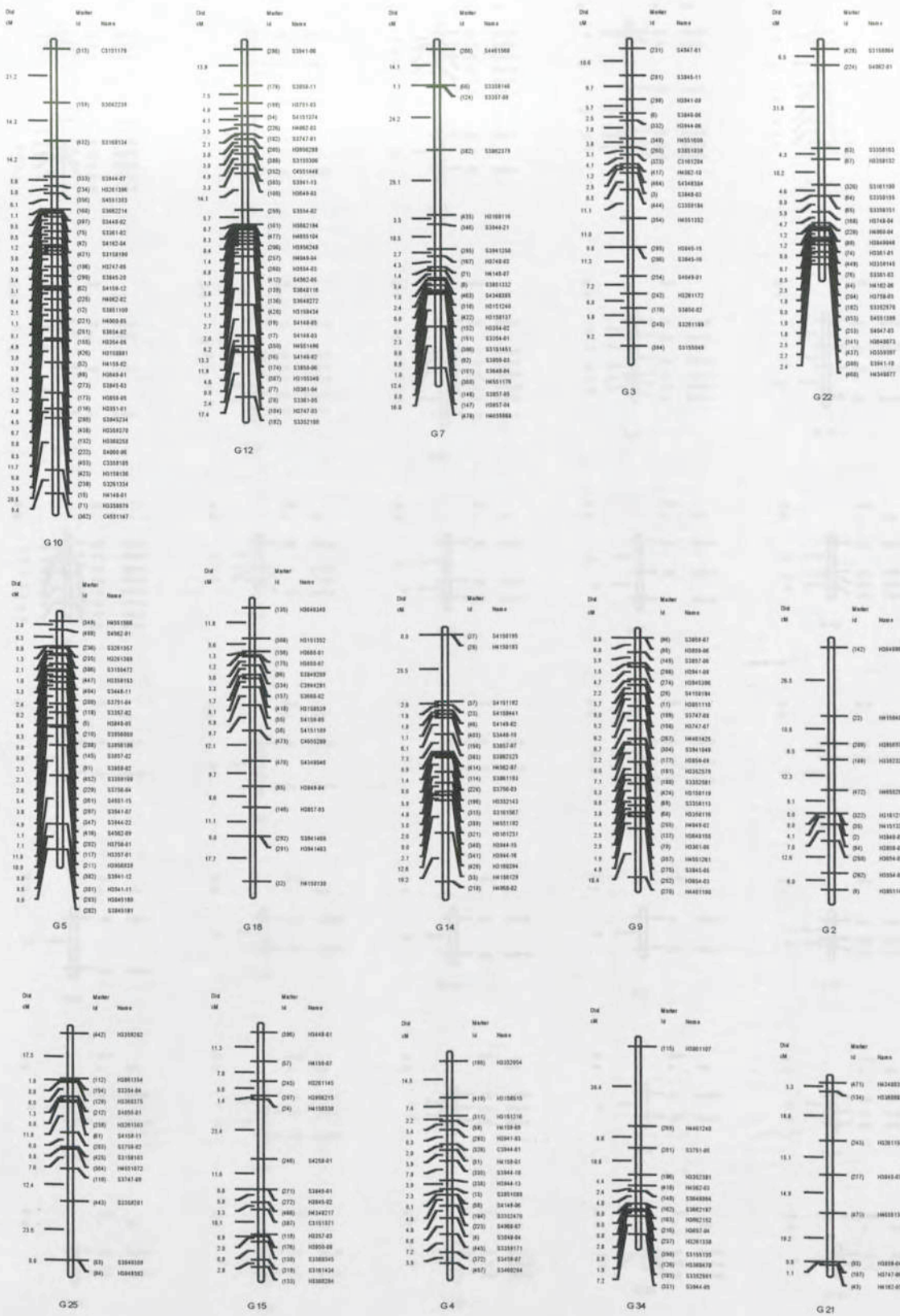


Figura 1 – Mapa genético de ligação desenvolvido pela Monsanto (St. Louis, MO, EUA), derivado do cruzamento entre a HS-1 (*Glycine max*, variedade comercial da Hartz Seeds Co, EUA) e a PI 326581 (*Glycine soja*). Foram utilizados 92 indivíduos F₂ e 58 pares de primers, que geraram um mapa com 480 marcadores AFLP e 52 grupos de ligação, cobrindo uma extensão de 2.365 cM.

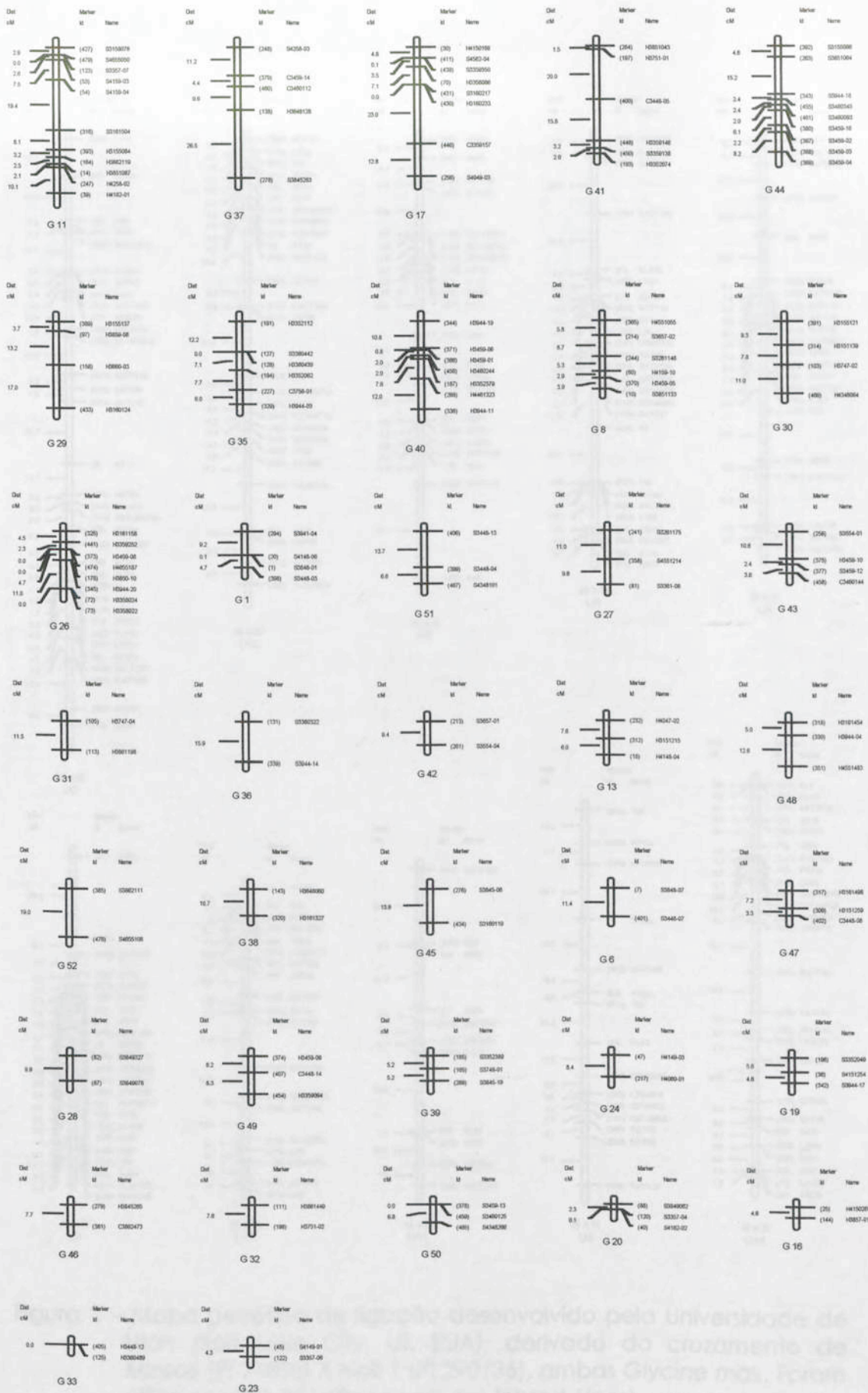


Figura 1, Cont.

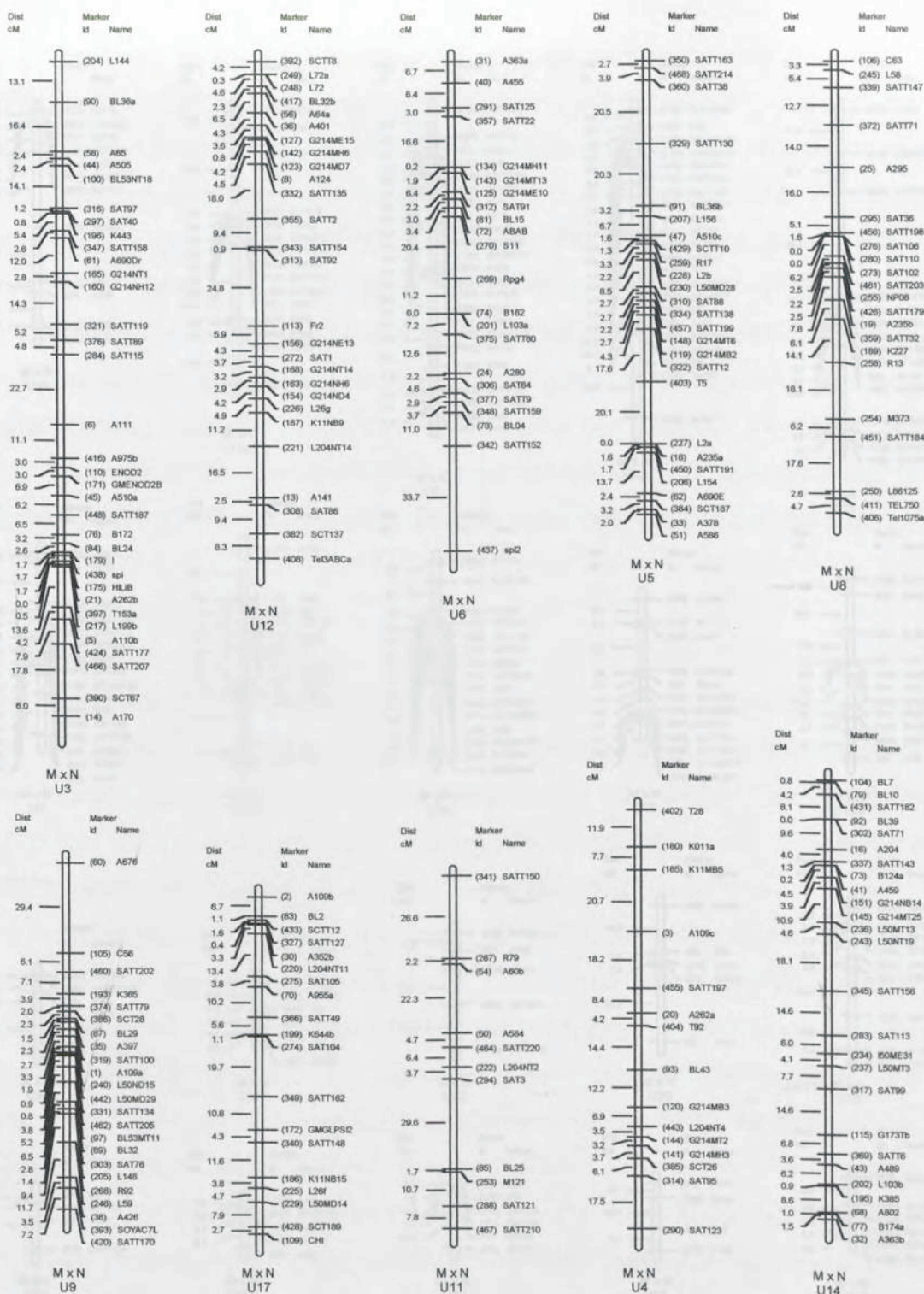


Figura 2 – Mapa genético de ligação desenvolvido pela Universidade de Utah (Salt Lake City, UT, EUA), derivado do cruzamento de Minsoy (PI 27890) X Noir 1 (PI 290136), ambas *Glycine max*. Foram utilizadas 223 RILs (*Recombinant Inbred Lines*), que geraram um mapa com 468 marcadores (279 RFLP, 166 SSR e 23 Clássicos) e 33 grupos de ligação, cobrindo uma extensão de 2.330 cM.

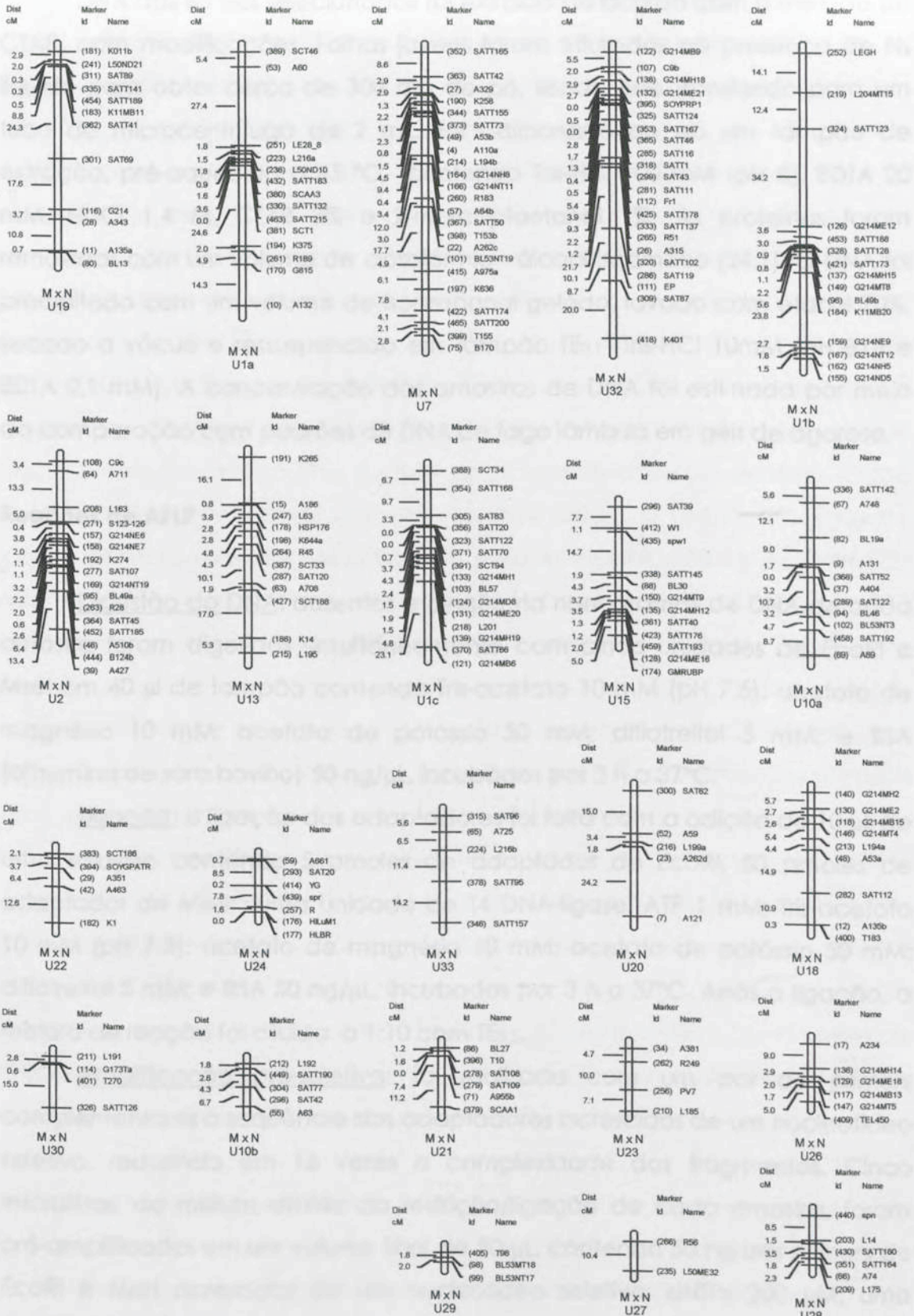


Figura 2, Cont.

Isolamento e quantificação de DNA

DNA das 88 RILs selecionadas foi extraído de acordo com o método do CTAB, com modificações. Folhas jovens foram trituradas na presença de N₂ líquido para obter cerca de 300 mg de pó, sendo este transferido para um tubo de microcentrífuga de 2 mL. Foi adicionado ao pó um tampão de extração, pré-aquecido a 65 °C, contendo Tris-HCl 100 mM (pH 8), EDTA 20 mM, NaCl 1,4 M, CTAB 2% e 2-mercaptoetanol 1%. As proteínas foram removidas com um volume de clorofórmio - álcool isoamílico (24:1). O DNA foi precipitado com um volume de isopropanol gelado, lavado com etanol 70%, secado a vácuo e ressuspendido em tampão TE_{0,1} (Tris-HCl 10mM (pH 8,0) e EDTA 0,1 mM). A concentração das amostras de DNA foi estimada por meio da comparação com padrões de DNA de fago lâmbda em géis de agarose.

Reações de AFLP

Digestão do DNA: duzentos e cinqüenta nanogramas de DNA de cada amostra foram digeridos simultaneamente com cinco unidades de *EcoRI* e *MseI* em 40 µl de tampão contendo Tris-acetato 10 mM (pH 7,5); acetato de magnésio 10 mM; acetato de potássio 50 mM; ditiotreititol 5 mM; e BSA (albumina de soro bovino) 50 ng/µL, incubados por 3 h a 37°C.

Ligação: a ligação dos adaptadores foi feita com a adição de 10 µL de uma solução contendo 5 pmoles de adaptador de *EcoRI*, 50 pmoles de adaptador de *MseI*; uma unidade de T4 DNA-ligase; ATP 1 mM; Tris-acetato 10 mM (pH 7,5); acetato de magnésio 10 mM; acetato de potássio 50 mM; ditiotreititol 5 mM; e BSA 50 ng/µL, incubados por 3 h a 37°C. Após a ligação, a mistura de reação foi diluída a 1:10 com TE_{0,1}.

Amplificação pré-seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de um nucleotídeo seletivo, reduzindo em 16 vezes a complexidade dos fragmentos. Cinco microlitros, da mistura diluída da restrição/ligação de cada amostra, foram pré-amplificados em um volume final de 50 µL, contendo 50 ng dos *primers* de *EcoRI* e *MseI* acrescidos de um nucleotídeo seletivo; dNTPs 200 µM; uma unidade de *Taq* polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl

50 mM. As reações de amplificação do DNA foram aplicadas no termociclador Perkin-Elmer 9600 (Norwalk, CT, EUA). Os ciclos foram os seguintes: 20 ciclos de 94°C por 30 segundos, 55°C por 30 segundos e 72°C por 60 segundos. Após a amplificação pré-seletiva, os produtos das reações foram diluídos a 1:20 com TE_{0,1}.

Amplificação seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de três nucleotídeos seletivos, reduzindo em 256 vezes a complexidade dos fragmentos. Para amplificação seletiva, o *primer* seletivo de *EcoRI* foi marcado na extremidade 5', utilizando-se [γ -³³P] ATP. A reação de marcação foi composta de 50 μ L, contendo 500 ng de *primer* de *EcoRI* com três nucleotídeos seletivos; [γ -³³P] ATP 100 μ Ci; 10 unidades de T4 polinucleotídeo cinase; Tris-HCl 25 mM (pH 7,5); MgCl₂ 10 mg; ditiotretitol 5 mM; e espermidina-3HCl 0,5 mM. Cinco microlitros da mistura de amplificação pré-seletiva, diluída a 1:20, foram amplificados num volume total de 20 μ L, contendo 5 ng do *primer* de *EcoRI*, marcado com três nucleotídeos seletivos; 30 ng de *primer* de *MseI* com três nucleotídeos seletivos; dNTPs 0,2 mM; 0,4 unidades de *Taq* polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl 50 mM. Os ciclos de amplificação seletiva foram os seguintes: 12 ciclos de 94°C por 30 segundos, 65°C (-0,7°C/ciclo) por 30 segundos e 72°C por 60 segundos, até atingir a temperatura ótima de ligação dos *primers* de 56°C - Touchdown Cycle Profile (Don *et al.*, 1991).

O DNA das 88 RILs de Minsoy X Noir 1 foi amplificado pela técnica de AFLP, com a utilização de 22 pares de *primer* também utilizados previamente para construção do mapa da Monsanto, visando à identificação de bandas polimórficas nas duas populações, servindo como pontos de ancoramento para integração dos dois mapas. Posteriormente, foram utilizados 19 pares de *primers* não comuns ao mapa da Monsanto, apenas com o objetivo de saturar o mapa Minsoy X Noir 1 com marcadores AFLP adicionais.

Eletroforese em gel de poliacrilamida

Os produtos da amplificação seletiva foram misturados com um volume igual (20 μ L) de tampão contendo formamida 98%; EDTA 10mM; xileno cianol

FF 1 mg/mL; e azul-de-bromofenol 1 mg/mL. As amostras foram, então, desnaturadas a 95°C por cinco minutos e mantidas em gelo até serem aplicadas no gel. Um microlitro e meio de cada amostra foi aplicado em um gel desnaturante de poliacrilamida 4,5% (acrilamida:bisacrilamida 19:1, uréia 7,5 M e Tris-borato EDTA (TBE) 0,5x), utilizando-se a cuba de eletroforese *SequiGen* de 38 x 50 cm (BioRad Laboratories Inc., Hercules, CA, EUA). Foram utilizados dois pentes, de 48 dentes, para cada gel. Em cada pente foram carregados os progenitores Minsoy e Noir 1 e um padrão 10 BL, que mostra uma banda de DNA a cada 10 pb entre 80 e 530 pb. A placa de vidro externa não recebeu nenhum tratamento, e a placa de vidro interna foi coberta com 3 mL de Sigma cote para a não-aderência do gel. A eletroforese foi realizada em TBE 1x, 130 W, 50°C por 1,5 h. Após a eletroforese, as placas foram separadas e o gel permaneceu ligado à placa de vidro externa, que não recebeu nenhum tratamento. O gel foi fixado em papel 3MM, envolto por um filme de plástico, secado a vácuo por 1 h a 70°C e exposto a um *phosphor screen* por 20 h. Os *screens* foram digitalizados, usando-se o *Storm System* (Molecular Dynamics, CA, EUA).

Scoring dos géis e análise dos dados

As imagens geradas pelo *Storm system* foram convertidas para o programa da Keygene (Wageningen, Holanda) para serem analisadas. Cada imagem foi convertida em duas, gerando-se o gel da esquerda e o gel da direita. Uma vez convertida a imagem, o primeiro procedimento foi a definição das faixas, seguindo-se as curvas do gel, para que o programa identificasse cada amostra. Cada faixa do gel foi nomeada de acordo com a amostra aplicada. A próxima etapa foi a definição das bandas comuns para definir a forma do gel. Uma vez definida a forma do gel, foi feito o *scoring* das bandas polimórficas, sendo (+) para a presença da banda, (-) para ausência da banda e (?) para falhas ou dúvida quanto ao *scoring*.

Foram utilizados dois pentes, de 48 dentes, para cada gel. Em cada pente foram aplicados os progenitores Minsoy e Noir 1 e um padrão 10 BL, que mostrou uma banda a cada 10 pb, entre 80 e 530 pb. No pente da esquerda, além dos progenitores e do padrão 10 BL, foi carregada uma mistura de DNA

de nove genótipos de soja, chamada de *pool* universal. Este *pool* universal foi utilizado no *scoring* dos géis, usando-se géis de referência, conforme descritos nos parágrafos subseqüentes.

Géis de referência: o mapa de AFLP construído pela Monsanto possui marcadores que são designados pelo par de *primers* utilizado, seguido da posição em que se encontra no gel em relação ao total de marcadores gerados, ex. M4150-08 (oitavo marcador gerado pelo par de *primers* E41 e M50). Este sistema de nomenclatura torna difícil a identificação de marcadores específicos em diferentes populações, impossibilitando a utilização deles em trabalhos de integração de mapas. Uma das maneiras de solucionar esse problema foi padronizar cada marcador pelo seu tamanho em pares de base, por meio do uso de géis de referência. Para isso, foi criado um *pool* de DNA de nove genótipos de soja, sendo oriundos do norte e sul dos EUA com a presença também de algumas PIs (introduções não-adaptadas). Esse *pool* é designado universal, pois gera grande parte das possíveis bandas da cultura da soja. As bandas geradas pelo *pool* universal foram caracterizadas pelo tamanho, em pares de base, utilizando-se para isso um DNA-padrão, 10 BL, que apresenta bandas a cada 10 pb, com resolução entre 80 e 530 pb. Na Figura 3, mostra-se um detalhe do *scoring* de um gel de referência. Géis de AFLP de 256 pares de *primers*, cobrindo todas as possíveis combinações, contendo DNA amplificado com os *primers* pré-seletivos EcoRI (+1)=A e MseI (+1)=C (Tabela 1), foram gerados contendo: DNA-padrão 10 BL, *pool* universal, HS1 e PI 326581 (progenitores do mapa da Monsanto). As bandas produzidas pelo *pool* universal foram caracterizadas pelo tamanho em pb por meio de um alinhamento da imagem do DNA padrão 10 BL do gel com uma imagem computadorizada desse padrão. As bandas polimórficas entre HS1 e PI 326581 foram, então, caracterizadas pelo tamanho em pb. Por meio da comparação dos géis de referência com os géis produzidos para gerar o mapa, todos os marcadores foram caracterizados pelo tamanho em pb, possibilitando a identificação deles em outras populações. Os géis de referência permitem a determinação do tamanho, em pb, de marcadores de qualquer população segregante de soja, por meio do alinhamento do DNA-padrão, 10 BL, do gel a ser analisado com o DNA-padrão do gel de referência, com o uso do

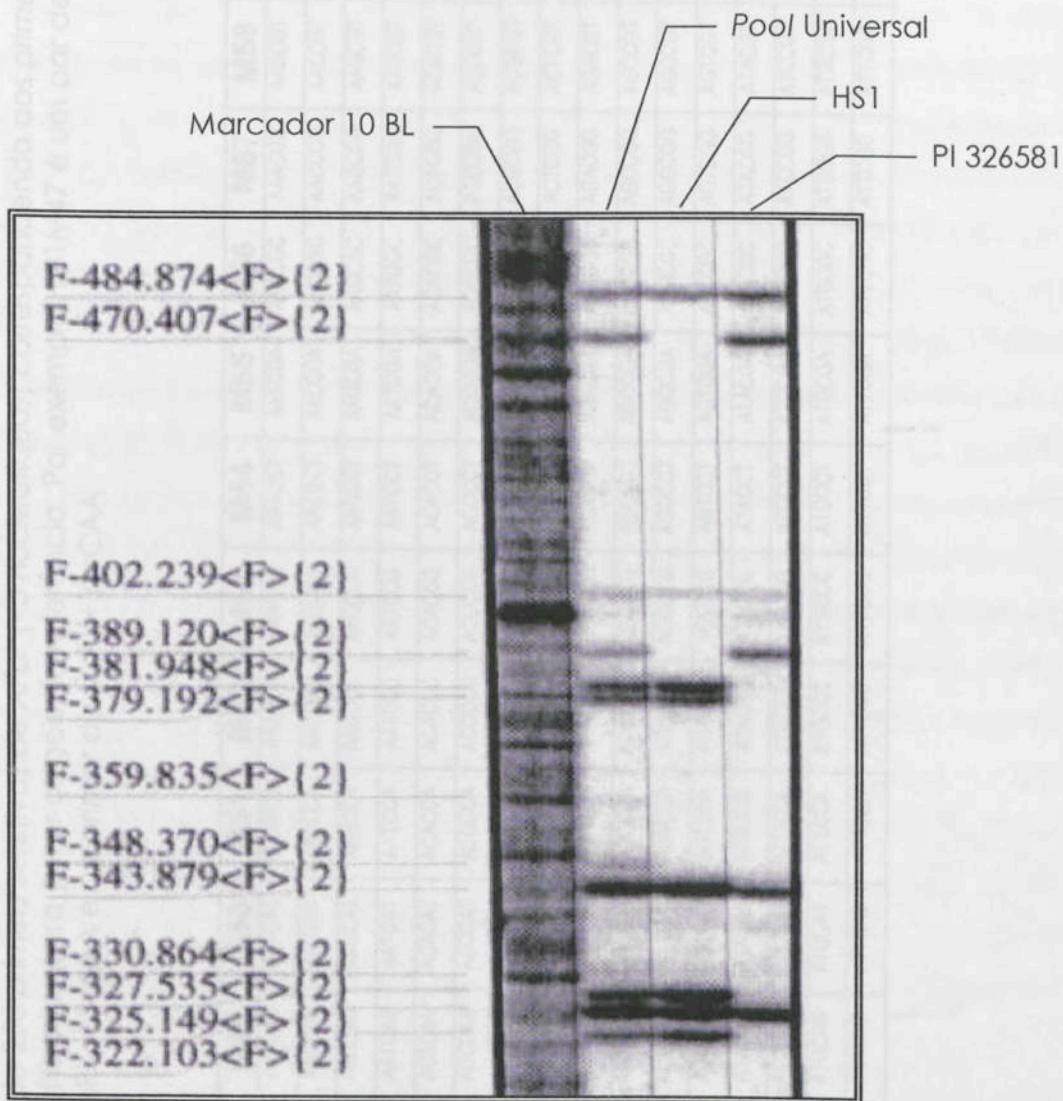


Figura 3 – Detalhe de um gel de referência mostrando uma região entre 320 e 510 bp. Foram carregados da esquerda para a direita: o marcador 10 BL, que mostra uma banda a cada 10 pb entre 80 e 530 pb; o *pool* universal, que é uma mistura de vários genótipos de soja; e os progenitores HS1 e PI 326581, utilizados no mapa da Monsanto. Foi feito o *scoring* de todas as bandas geradas de acordo com o tamanho, em pb.

Tabela 1 – Designação de 256 primers seletivos de AFLP (+3 nucleotídeos) correspondendo aos primers pré-seletivos EcoRI(+1)=A e MseI(+1)=C, utilizados para gerar os géis de referência. Por exemplo, E31M47 é um par de primers que corresponde ao primer de EcoRI(+3)=AA e ao primer de MseI(+3)=CAA

	M47	M48	M49	M50	M51	M52	M53	M54	M55	M56	M57	M58	M59	M60	M61	M62
E31	AAACAA	AAACAC	AAACAG	AAACAT	AAACCA	AAACCC	AAACCG	AAACCT	AAACGA	AAACGC	AAACGG	AAACGT	AAACTA	AAACTC	AAACTG	AAACTT
E32	AAGCAA	AACCAC	AACCAG	AACCAT	AACCCA	AACCCC	AACCCG	AACCCT	AACCGA	AACCGC	AACCGG	AACCGT	AACCTA	AACCTC	AACCTG	AACCTT
E33	AAGCAA	AAGCAC	AAGCAG	AAGCAT	AAGCCA	AAGCCC	AAGCCG	AAGCCT	AAGCGA	AAGCGC	AAGCGG	AAGCGT	AAGCTA	AAGCTC	AAGCTG	AAGCTT
E34	AATCAA	AATCAC	AATCAG	AATCAT	AATCCA	AATCCC	AATCCG	AATCCT	AATCGA	AATCGC	AATCGG	AATCGT	AATCTA	AATCTC	AATCTG	AATCTT
E35	ACACAA	ACACAC	ACACAG	ACACAT	ACACCA	ACACCC	ACACCG	ACACCT	ACACGA	ACACGC	ACACGG	ACACGT	ACACTA	ACACTC	ACACTG	ACACTT
E36	ACCCAA	ACCCAC	ACCCAG	ACCCAT	ACCCCA	ACCCCC	ACCCCG	ACCCCT	ACCCGA	ACCCGC	ACCCGG	ACCCGT	ACCCTA	ACCCTC	ACCCTG	ACCCTT
E37	ACGCAA	ACGCAC	ACGCAG	ACGCAT	ACGCCA	ACGCCC	ACGCCG	ACGCCCT	ACGCCGA	ACGCCGC	ACGCCGG	ACGCCGT	ACGCTA	ACGCTC	ACGCTG	ACGCTT
E38	ACTCAA	ACTCAC	ACTCAG	ACTCAT	ACTCCA	ACTCCC	ACTCCG	ACTCCT	ACTCGA	ACTCGC	ACTCGG	ACTCGT	ACTCTA	ACTCTC	ACTCTG	ACTCTT
E39	AGACAA	AGACAC	AGACAG	AGACAT	AGACCA	AGACCC	AGACCG	AGACCT	AGACGA	AGACGC	AGACGG	AGACGT	AGACTA	AGACTC	AGACTG	AGACTT
E40	AGCCAA	AGCCAC	AGCCAG	AGCCAT	AGCCCA	AGCCCC	AGCCCG	AGCCCT	AGCCGA	AGCCGC	AGCCGG	AGCCGT	AGCCTA	AGCCTC	AGCCTG	AGCCTT
E41	AGGCAA	AGGCAC	AGGCAG	AGGCAT	AGGCCA	AGGCCC	AGGCCG	AGGCCCT	AGGCCGA	AGGCCGC	AGGCCGG	AGGCCGT	AGGCTA	AGGCTC	AGGCTG	AGGCTT
E42	AGTCAA	AGTCAC	AGTCAG	AGTCAT	AGTCCA	AGTCCC	AGTCCG	AGTCCT	AGTCGA	AGTCGC	AGTCGG	AGTCGT	AGTCTA	AGTCTC	AGTCTG	AGTCTT
E43	ATACAA	ATACAC	ATACAG	ATACAT	ATACCA	ATACCC	ATACCG	ATACCT	ATACGA	ATACGC	ATACGG	ATACGT	ATACTA	ATACTC	ATACTG	ATACTT
E44	ATCCAA	ATCCAC	ATCCAG	ATCCAT	ATCCCA	ATCCCC	ATCCCG	ATCCCT	ATCCGA	ATCCGC	ATCCGG	ATCCGT	ATCCTA	ATCCTC	ATCCTG	ATCCTT
E45	ATGCAA	ATGCAC	ATGCAG	ATGCAT	ATGCCA	ATGCCC	ATGCCG	ATGCCCT	ATGCCGA	ATGCCGC	ATGCCGG	ATGCCGT	ATGCTA	ATGCTC	ATGCTG	ATGCTT
E46	ATTCAA	ATTCAC	ATTCAG	ATTCAT	ATTCCA	ATTCCC	ATTCCG	ATTCCCT	ATTCCGA	ATTCCGC	ATTCCGG	ATTCCGT	ATTCTA	ATTCTC	ATTCTG	ATTCTT

programa de computador Keygene. Uma vez alinhadas as imagens, os marcadores do gel de referência foram transferidos para o gel a ser analisado e os marcadores comuns aos dois géis foram identificados, ficando caracterizados por um nome universal oriundo do *pool* universal. Marcadores com nome universal em pb permitem a sua identificação em diferentes populações, possibilitando a integração de mapas e o estudo das informações fornecidas pelos mapas.

Uma vez feito o *scoring* dos géis das 88 RILs de Minsoy X Noir 1, utilizando-se os géis de referência, os dados foram exportados pelo programa Keygene, que gerou uma matriz +/- . Os marcadores que apresentavam a origem da banda de Minsoy receberam a letra M e os que exibiam a origem da banda de Noir 1, a letra N. Após a letra M ou N, vem o número do *primer* de *EcoRI*, seguido pelo *primer* de *MseI* e pelo tamanho do fragmento, em pb, fornecido pelos géis de referência. Exemplificando, o marcador M4348266 significa uma banda originada de Minsoy, gerada pelo par de *primers* seletivos E43 e M48 (Tabela 1), com o tamanho de 266 pb. A matriz +/- foi convertida para uma matriz de A e B para ser utilizada pelos programas de mapeamento. Os grupos de ligação foram determinados pelo programa MapMaker 3.0 (Lander *et al.*, 1987), utilizando-se LOD 5,0. A ordem dos marcadores em cada grupo de ligação foi determinada pelo programa JoinMap (Stam, 1993). Os grupos de ligação foram gerados graficamente pelo programa MapMaker para Macintosh, usando-se a distância de Kosambi.

A comparação de um dos 88 RILs de Minsoy X Noir com os 22 pares de *primers* comuns ao mapa da Maranta, produziu um total de 151 marcadores, e sua representação uma matriz de 431 bandas polimórficas para cada par de *primers* utilizado. Desses 151 marcadores, 34 foram comuns aos dois mapas, sendo utilizados como pontos de ancoramento (Tabela 2).

RESULTADOS E DISCUSSÃO

Na Figura 4, mostra-se o produto de 12 dos 256 pares de *primers* utilizados nos géis de referência. Dos 256 pares de *primers* usados, foi realizado o *scoring* de 41 daqueles que foram empregados na amplificação das linhagens recombinantes endogâmicas (RILs) de Minsoy X Noir 1. Uma média de 90 bandas foi gerada para cada par de *primer*, com uma variação de 53 até 156 bandas. Todas as bandas detectadas foram nomeadas de acordo com o par de *primer* utilizado e o tamanho em pb. A inclusão do *pool* universal permitiu o *scoring* de inúmeras bandas não presentes nos progenitores HS1 e PI 326581, demonstrando que a técnica de AFLP foi muito sensível para a detecção de bandas oriundas de baixas concentrações de DNA, visto que o *pool* universal foi realizado com uma mistura do DNA de nove genótipos (Figura 3). É importante que os géis de referência contenham maior número de bandas possível, visto que eles serão utilizados para realizar o *scoring* de géis de outros experimentos, envolvendo outras populações.

Para identificação de marcadores comuns entre o mapa da Monsanto e o mapa Minsoy X Noir 1, foi necessário renomear os marcadores oriundos do mapa da Monsanto, visto que apresentavam nomenclatura designada pelo par de *primer* utilizado, seguido da posição em que se encontrava no gel. Desse modo, não se podia saber ao certo se dois marcadores eram realmente comuns aos dois mapas. As bandas polimórficas do mapa da Monsanto foram caracterizadas pelo tamanho em pb, por meio do alinhamento dos géis produzidos para gerar o mapa com os géis de referência de 22 pares de *primers* utilizados no mapa Minsoy x Noir 1. Na Figura 5, mostra-se o padrão de AFLP obtido com 88 RILs de Minsoy X Noir 1. Os marcadores oriundos desses 22 pares de *primers* foram renomeados de acordo com o par de *primer* utilizado e o tamanho em pb, possibilitando a identificação de marcadores comuns entre o mapa da Monsanto e o mapa Minsoy X Noir 1.

A amplificação do DNA das 88 RILs de Minsoy X Noir com os 22 pares de *primers*, comuns ao mapa da Monsanto, produziu um total de 151 marcadores, o que representa uma média de 6,61 bandas polimórficas para cada par de *primer* utilizado. Desses 151 marcadores, 54 foram comuns aos dois mapas, sendo utilizados como pontos de ancoramento (Tabela 2).

Faixas

- 1 = 10BL
- 2 = UP
- 3 = HS-1
- 4 = PI 326581
- 5 = 10BL
- 6 = UP
- 7 = HS-1
- 8 = PI 326581
- 9 = 10BL
- 10 = UP
- 11 = HS-1
- 12 = PI 326581
- 13 = 10BL
- 14 = UP
- 15 = HS-1
- 16 = PI 326581
- 17 = 10BL
- 18 = UP
- 19 = HS-1
- 20 = PI 326581
- 21 = 10BL
- 22 = UP
- 23 = HS-1
- 24 = PI 326581
- 25 = 10BL
- 26 = UP
- 27 = HS-1
- 28 = PI 326581
- 29 = 10BL
- 30 = UP
- 31 = HS-1
- 32 = PI 326581
- 33 = 10BL
- 34 = UP
- 35 = HS-1
- 36 = PI 326581
- 37 = 10BL
- 38 = UP
- 39 = HS-1
- 40 = PI 326581
- 41 = 10BL
- 42 = UP
- 43 = HS-1
- 44 = PI 326581
- 45 = 10BL
- 46 = UP
- 47 = HS-1
- 48 = PI 326581

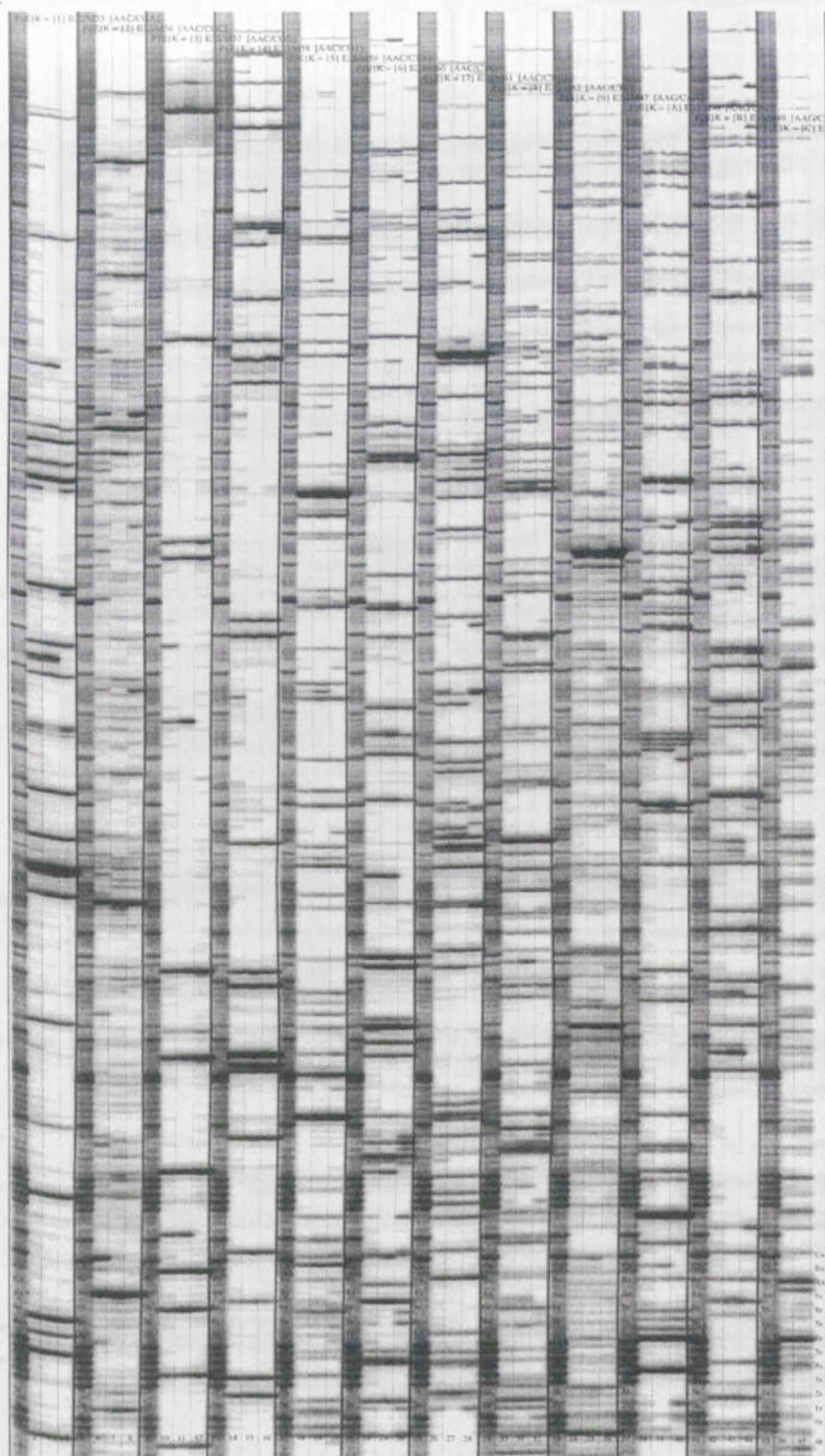


Figura 4 – Gel de referência mostrando 12 dos 256 pares de primers de AFLP utilizados (Tabela 1). Para cada par de primer, foi carregado um padrão 10 BL, que mostra uma banda a cada 10 pb entre 80 e 530 pb, e foram amplificadas o DNA do pool universal e o dos progenitores do mapa da Monsanto (Figura 1) HS1 e PI 326581.

MONSANTO
ST. LOUIS

RIL4348R N [..]
Minsoy X Noir RILs

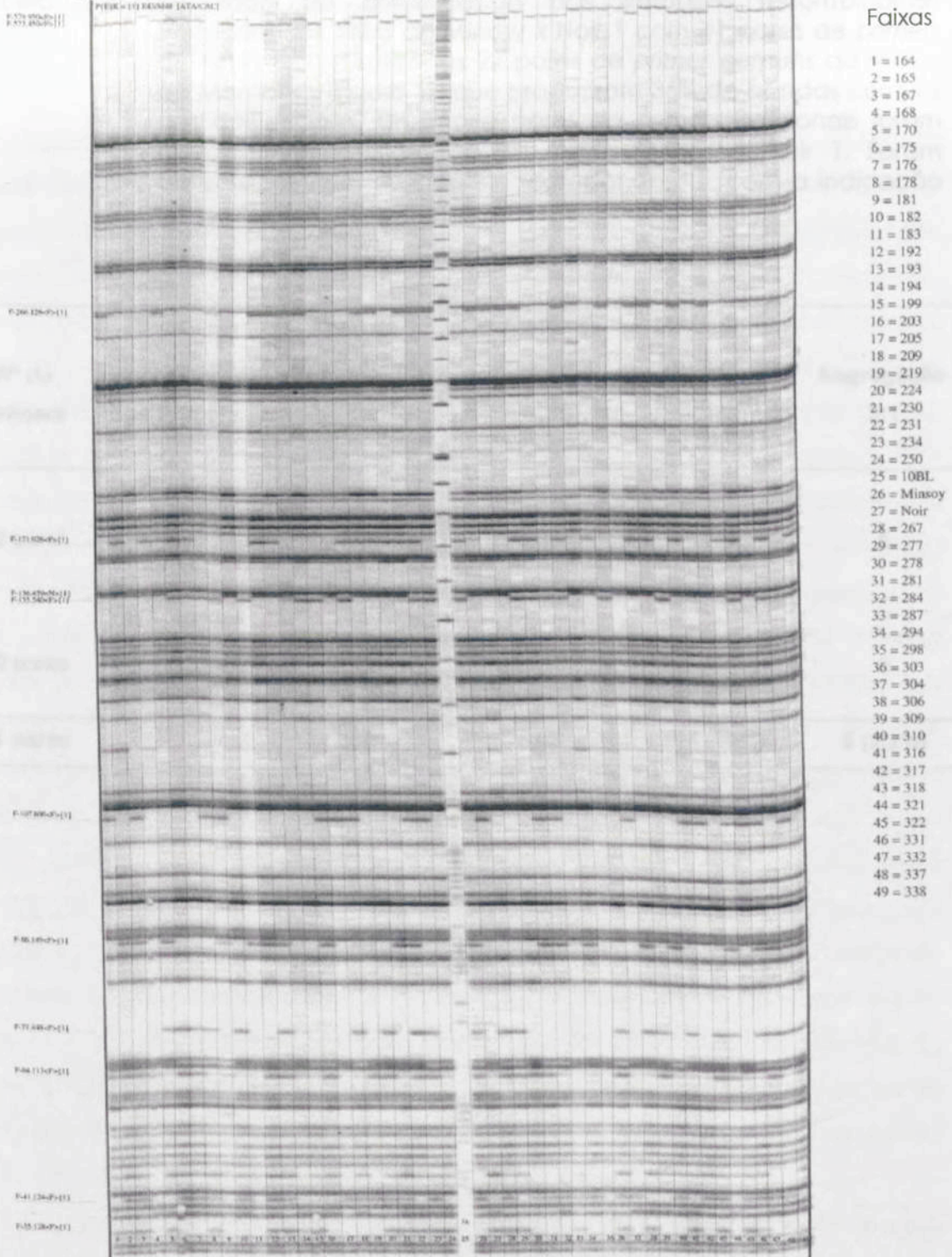


Figura 5 – Padrão de amplificação de DNA obtido pela técnica de AFLP utilizando o par de *primers* E43M48 (ATA/CAC), para as linhagens recombinantes endogâmicas (RILs) de Minsoy x Noir 1. As faixas do gel estão caracterizadas conforme a legenda à direita. O padrão 10 BL, Minsoy e Noir 1 foram carregados nas faixas 25, 26 e 27, respectivamente. Este par de *primers* produziu 12 bandas polimórficas, e o *scoring* foi feito, considerando-se a presença da banda = + e a ausência da banda = - .

Tabela 2 – Resultado da amplificação das linhagens recombinantes endogâmicas (RILs) de Minsoy x Noir 1 com 41 pares de *primers* de AFLP. Foram utilizados 22 pares de *primer* comuns ao mapa da Monsanto (Figura 1), que produziram 36% de bandas comuns aos dois mapas. Dezenove pares de *primers* adicionais foram utilizados para saturação do mapa Minsoy X Noir 1. Foram detectados 2,2% de bandas segregando 3:1, com a indicação de possíveis regiões duplicadas no genoma

Nº de Primers	Especificação dos Primers	Total de Marcadores Obtidos	Nº Médio de Marcadores para cada Par de Primer	Marcadores Comuns ao Mapa HS1 x PI 326581	Segregação 3:1
22 pares	Usados no mapa HS1 x PI 326581	151	6,86	54 (36%)	5
19 pares	Extra, apenas para saturação do mapa	211	11,10	-	3
41 pares		362	8,82	54 (36%)	8 (2,2%)

em populações semelhantes oriundas de genótipos de soja adaptados.

Além dos 22 pares de *primers* utilizados na integração dos mapas, o DNA das 88 RILs de Minsoy X Noir 1 foi amplificado com mais 19 pares de *primers*, com o objetivo de saturar o mapa Minsoy X Noir 1. Os 19 pares de *primers* adicionais produziram 211 marcadores, o que representa uma média de 11,1 bandas polimórficas para cada par de *primer* utilizado (Tabela 2). No total, o DNA das 88 RILs de Minsoy X Noir 1 foi amplificado com 41 pares de *primers* de AFLP, que produziram 362 marcadores, representando uma média de 8,82 marcadores para cada par de *primer*.

Foram detectados oito marcadores (2,2%) com segregação de tipo 3:1, que não foram incluídos no mapa de Minsoy X Noir 1. Considerando-se que a população de RILs de Minsoy X Noir 1 sofreu 10 autofecundações, o nível de heterozigose seria $1 - 0,99^{10} = 9,1\%$. Portanto, esses 2,2% de marcadores com segregação 3:1 podem estar detectando regiões duplicadas no genoma que, por sua vez, também podem estar associadas à possível origem tetraploide da soja (Cicchini et al., 1994).

A identificação de 54 marcadores comuns, num total de 151 marcadores, representa 36% de marcadores comuns aos dois mapas. Considerando que o mapa da Monsanto é interespecífico (*G. max* X *G. soja*) e o mapa Minsoy X Noir 1 é intra-específico (*G. max* X *G. max*), este resultado está de acordo com a integração de mapas realizada por Shoemaker e Specht (1995), que encontraram 32% de marcadores de RFLP polimórficos entre uma população de *G. soja* X *G. max* e uma população de *G. max* X *G. max*. Akkaya *et al.* (1995) comentaram que apenas 32% dos marcadores foram polimórficos, segundo o trabalho de Shoemaker e Specht (1995), e acreditavam que esse polimorfismo será muito menor se a integração de mapas envolver duas populações de *G. max* adaptadas, que apresentam base genética estreita segundo Delannay *et al.* (1983). Uma alternativa para contornar esse baixo nível de polimorfismo seria o uso de marcadores microssatélites ou SSRs. As seqüências de DNA que flanqueiam os microssatélites são, geralmente, conservadas entre indivíduos da mesma espécie, permitindo-se a seleção de *primers* que irão gerar fragmentos em todos os genótipos. O alto nível de alelismo dos microssatélites permite que dado loco seja polimórfico entre duas populações bem distintas ou até mesmo em populações semelhantes oriundas de genótipos de soja adaptados.

Além dos 22 pares de *primers* utilizados na integração dos mapas, o DNA das 88 RILs de Minsoy X Noir 1 foi amplificado com mais 19 pares de *primers*, com o objetivo de saturar o mapa Minsoy X Noir 1. Os 19 pares de *primers* adicionais produziram 211 marcadores, o que representa uma média de 11,1 bandas polimórficas para cada par de *primer* utilizado (Tabela 2). No total, o DNA das 88 RILs de Minsoy X Noir 1 foi amplificado com 41 pares de *primers* de AFLP, que produziram 362 marcadores, representando uma média de 8,82 marcadores para cada par de *primer*.

Foram detectados oito marcadores (2,2%), com segregação do tipo 3:1, que não foram incluídos no mapa de Minsoy X Noir 1. Considerando-se que a população de RILs de Minsoy X Noir 1 sofreu 10 autofecundações, o nível de heterozigose seria $1 - (\frac{1}{2})^{10} \approx 0,1\%$. Portanto, esses 2,2% de marcadores com segregação 3:1 podem estar detectando regiões duplicadas no genoma que, por sua vez, também podem estar associadas à possível origem tetraplóide da soja (Shoemaker *et al.*, 1996)

O mapa original de Minsoy X Noir 1 (Figura 2) contém 468 marcadores, sendo 279 RFLP, 166 SSR e 23 clássicos, distribuídos em 33 grupos de ligação, que cobrem uma extensão de 2.330 cM, com um nível de saturação médio de um marcador para cada 5 cM. A inclusão de 354 marcadores de AFLP provocou mudança substancial no mapa original de Minsoy X Noir 1. O novo mapa (Figura 6) contém 822 marcadores, cobrindo uma extensão de 2.651 cM, com saturação média de um marcador para cada 3,2 cM. Esse novo mapa apresentou uma melhor saturação do que o mapa descrito por Keim *et al.* (1997), o qual exibiu um marcador para cada 4 cM. Os marcadores de AFLP incluídos no mapa Minsoy X Noir 1 apresentaram boa distribuição em relação aos marcadores SSR e RFLP, em todos os grupos de ligação. Foi notada, no entanto, certa aglomeração de marcadores AFLP em algumas regiões dos grupos de ligação. Uma possível explicação para essa distribuição desuniforme pode ser devida ao fato de que tais regiões exibem baixa frequência de recombinação, como centrômeros. Desse modo, a inclusão dos marcadores AFLP provavelmente tenha permitido detalhamento físico dos grupos de ligação do mapa Minsoy X Noir 1, com a possível identificação de centrômeros.

Os 54 marcadores de AFLP polimórficos entre a população F₂ de HS1 e PI 326581, utilizada na construção do mapa da Monsanto, e a população de RILs de Minsoy X Noir 1, possibilitaram a integração dos dois mapas. Vinte e cinco grupos de ligação do mapa Minsoy X Noir 1 foram ligados a 27 grupos de ligação do mapa da Monsanto (Figura 7). Ocorreram fusões de grupos de ligação em ambos os mapas, formando três novos grupos de ligação no mapa HS1 X PI 326581 e um novo grupo de ligação no mapa Minsoy X Noir 1. Na Figura 7A, mostra-se o novo grupo G21 de HS1 X PI 326581 formado pela fusão dos grupos G29 e G50 ao G21 antigo, apresentando cinco pontos de ancoramento com M X N U19. Na Figura 7R, observa-se o novo grupo G11 de HS1 X PI 326581 formado pela fusão dos grupos G11 e G41, apresentando dois pontos de ancoramento com M X N U1a. Na Figura 7U, mostra-se o novo grupo G15 de HS1 X PI 326581 formado pela fusão dos grupos G15 e G27, apresentando dois pontos de ancoramento com M X N U11. Ocorreu uma fusão de grupos de ligação no mapa Minsoy X Noir 1. Na Figura 7P, mostra-se o novo grupo M X N U28/U10b com a fusão do grupo U28 com o grupo U10b,

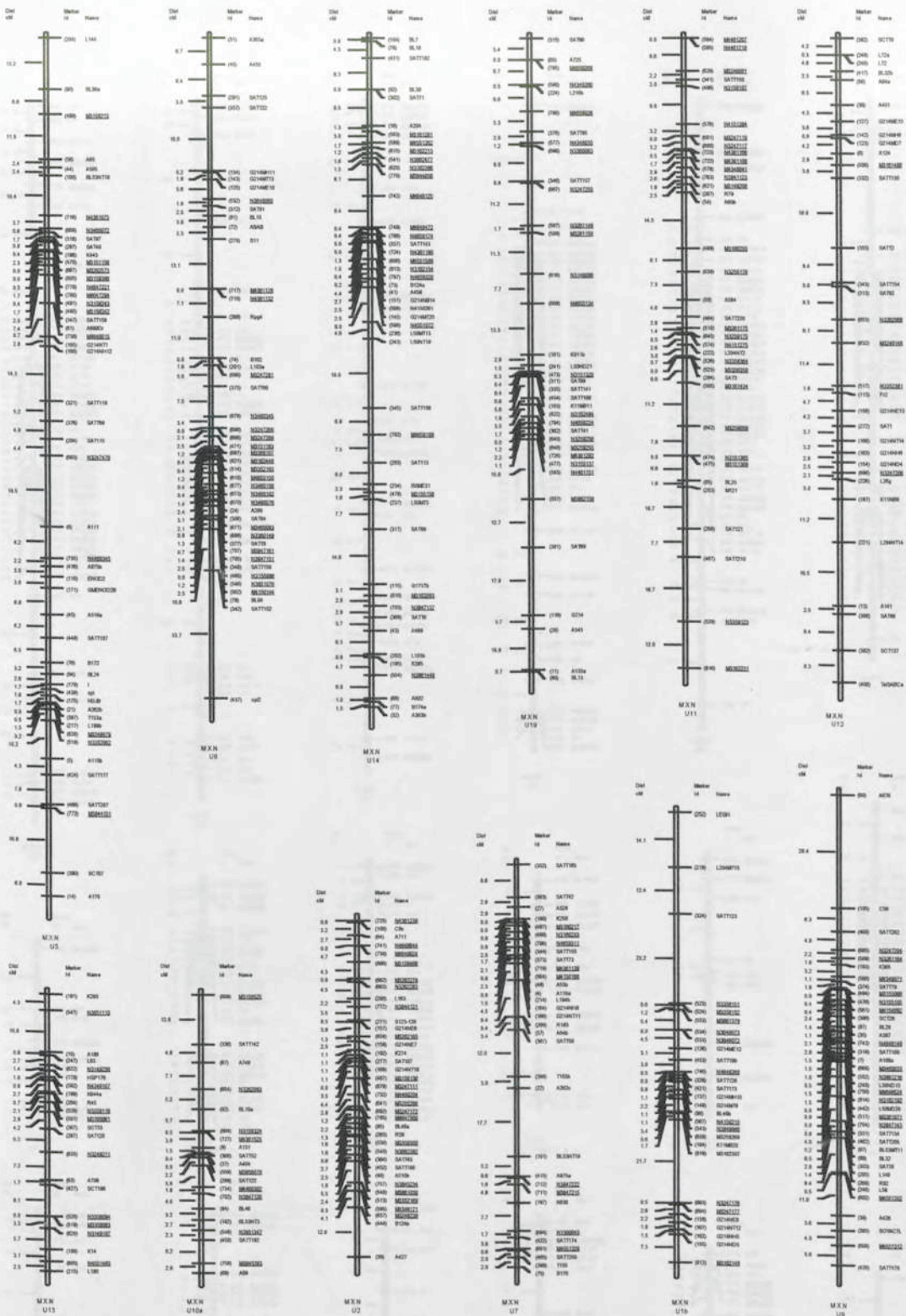


Figura 6 – Novo mapa genético de ligação de Minsoy (PI 27890) X Noir 1 (PI 290136). Foram acrescentados 354 marcadores AFLP (sublinhado) aos 468 marcadores (279 RFLP, 166 SSR e 23 clássicos) existentes. Esse novo mapa possui 32 grupos de ligação, com a fusão do grupo U10b com U28 cobrindo uma extensão de 2.651 cM, apresentando saturação de um marcador para cada 3,2 cM.

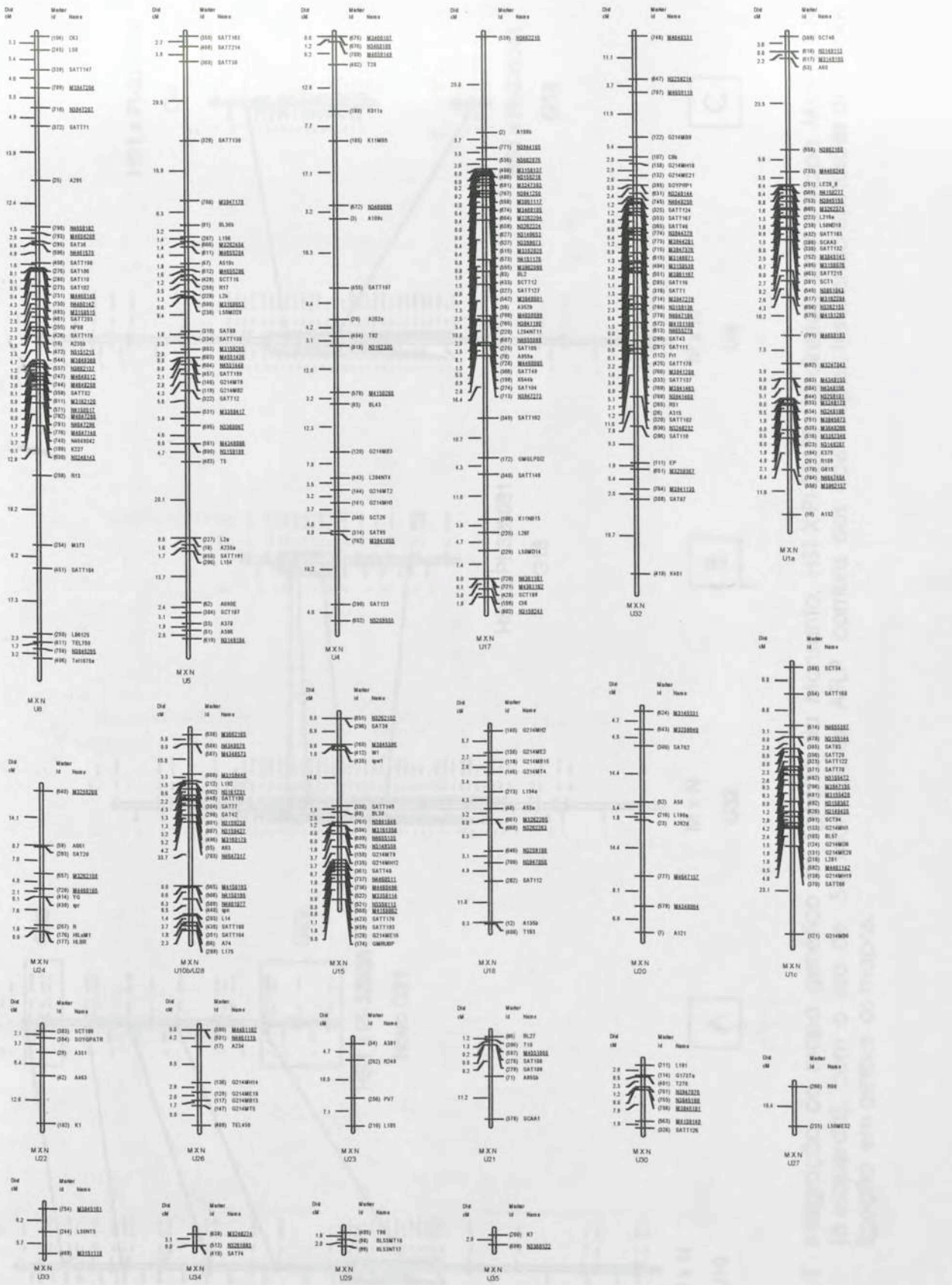


Figura 6, Cont.

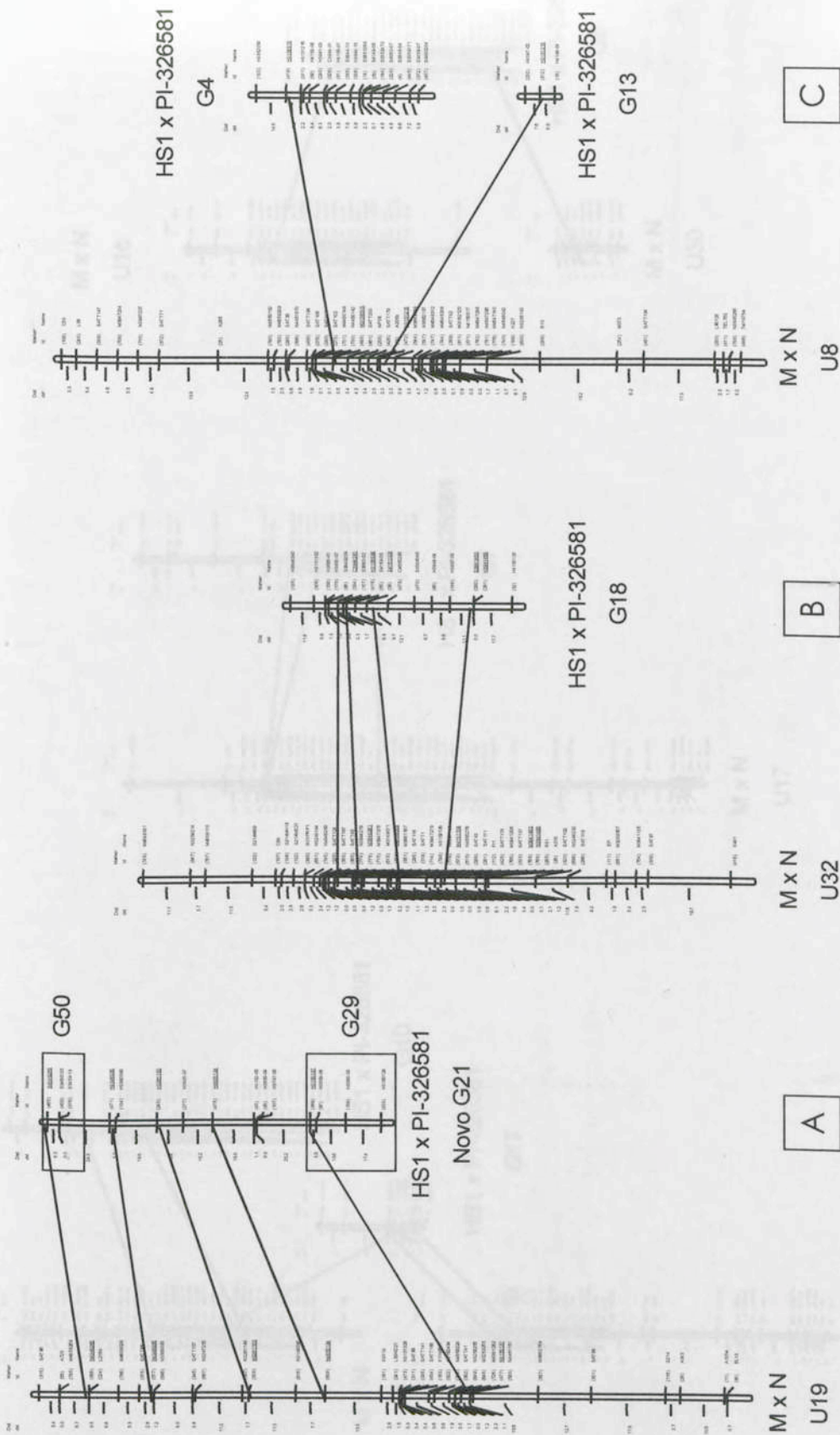


Figura 7 - Integração do mapa genético de ligação da Monsanto, HS1 X PI-326581 (à direita), com o mapa Minsoy X Noir 1 (à esquerda), com o uso de 54 marcadores AFLP comuns aos dois mapas, resultando em fusões de grupos de ligação em ambos os mapas.

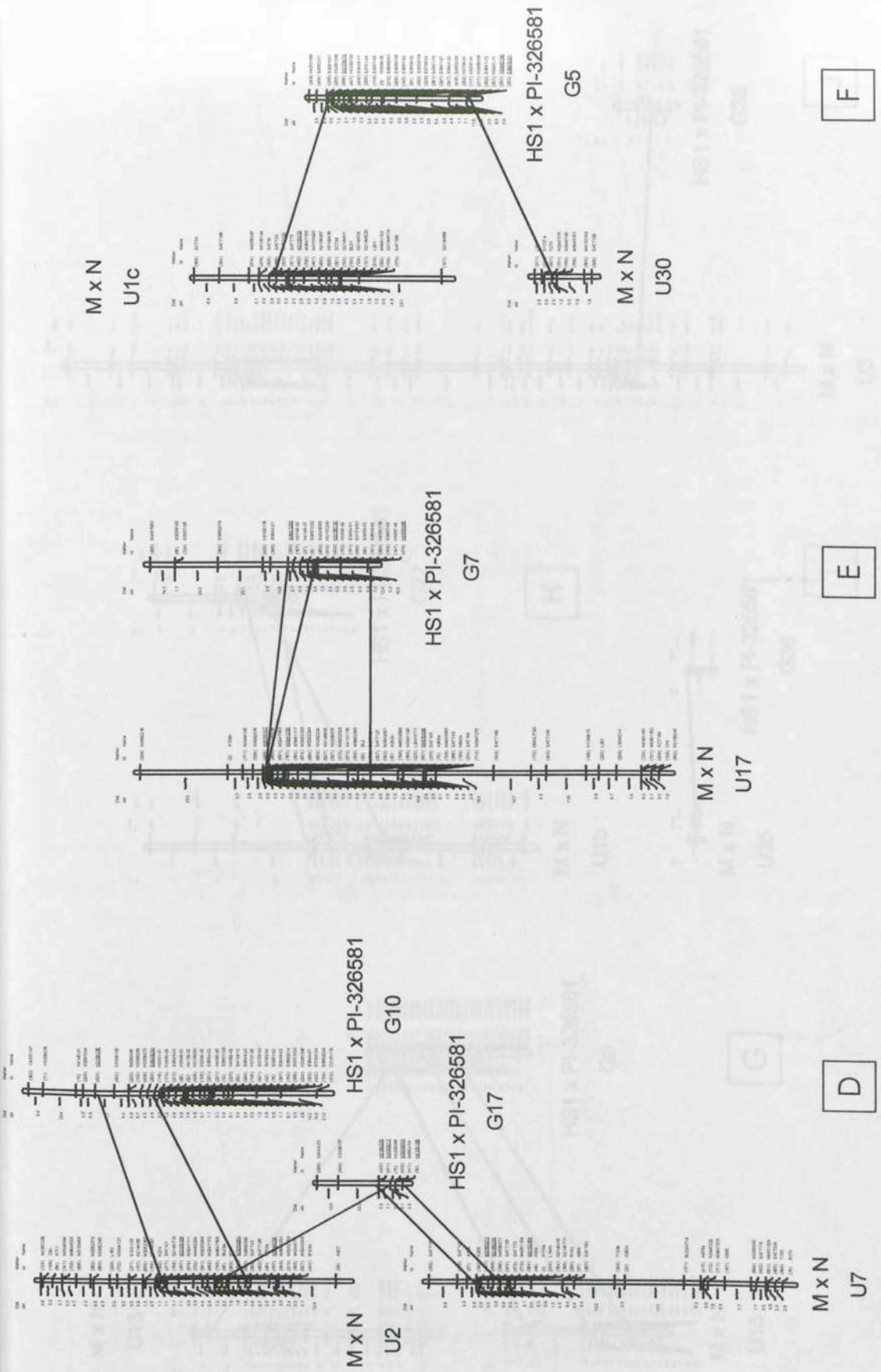


Figura 7, Cont.

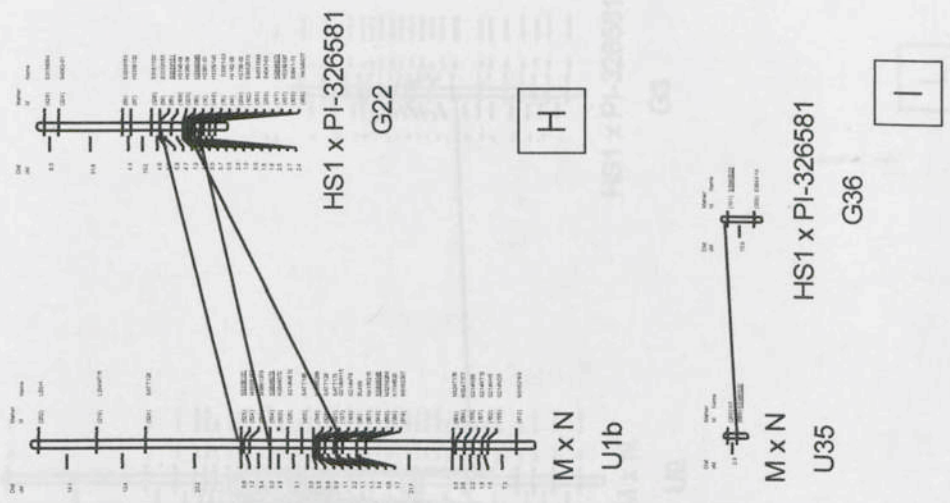
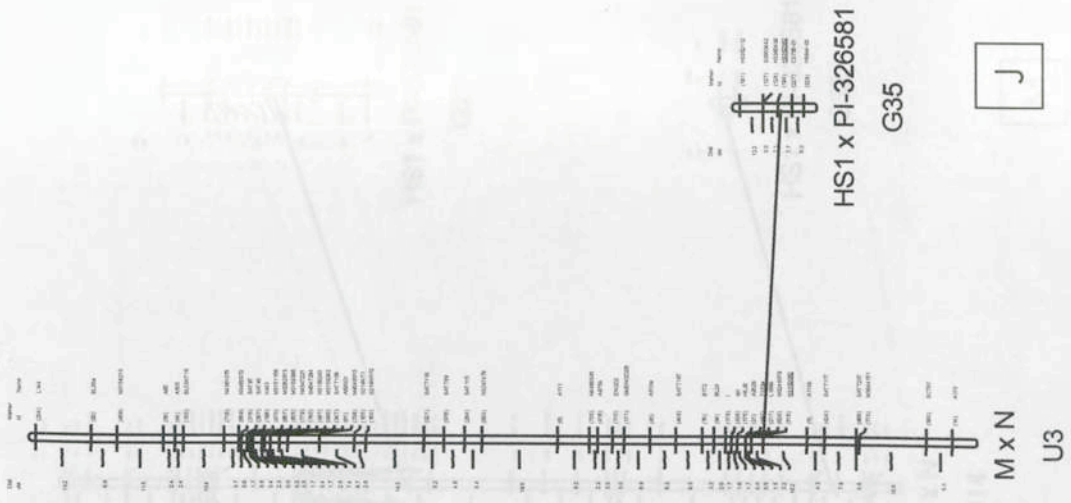


Figura 7, Cont.

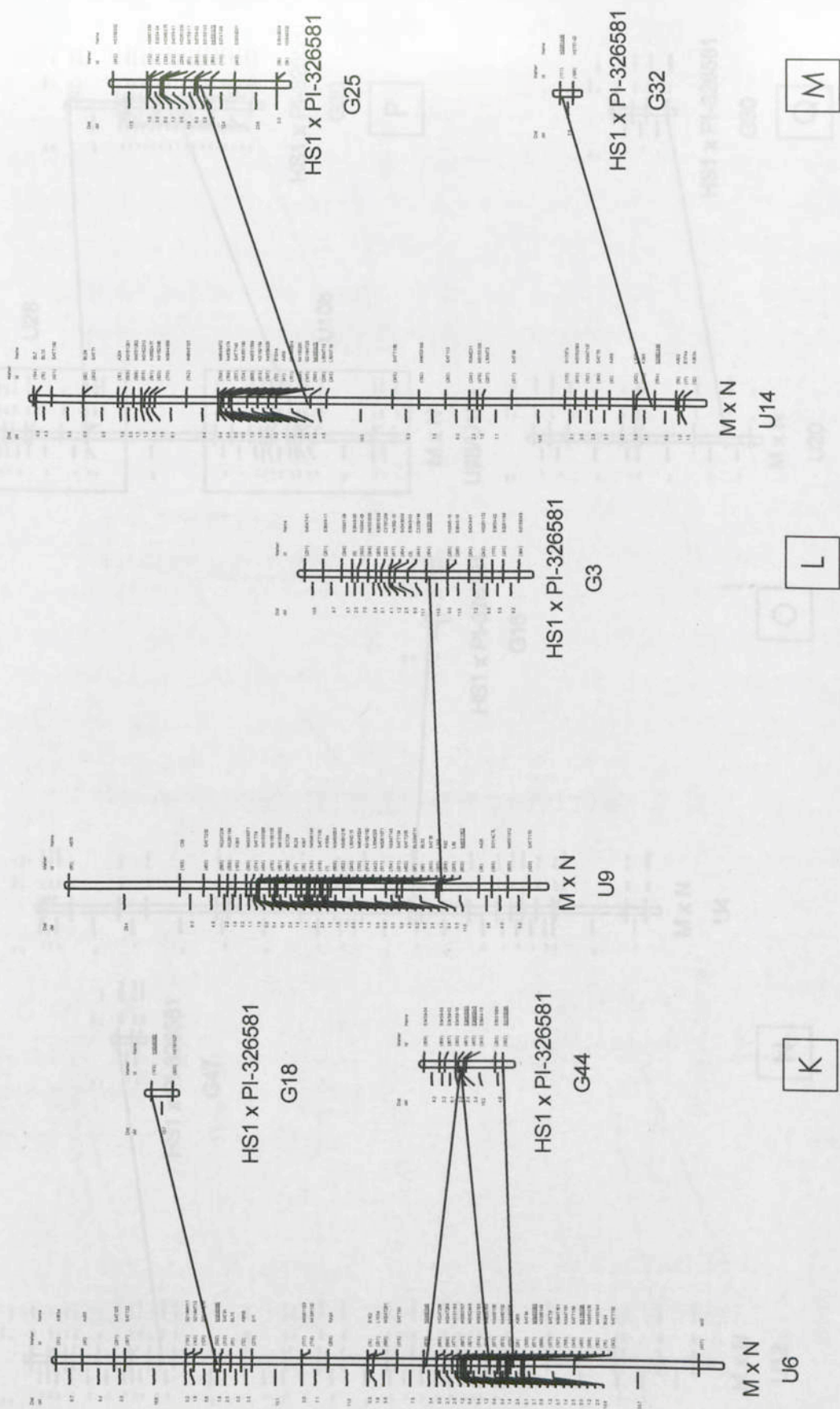


Figura 7, Cont.

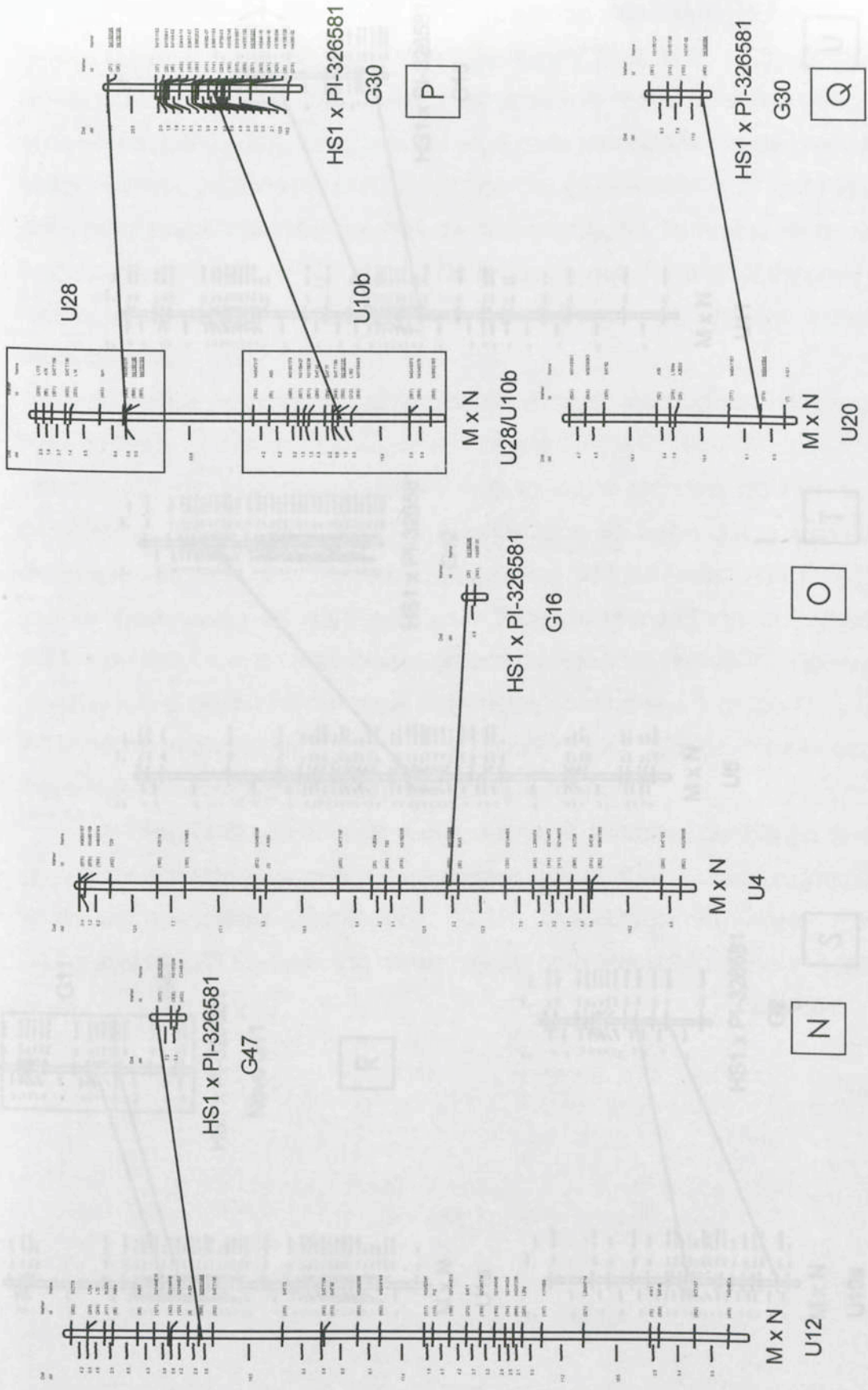
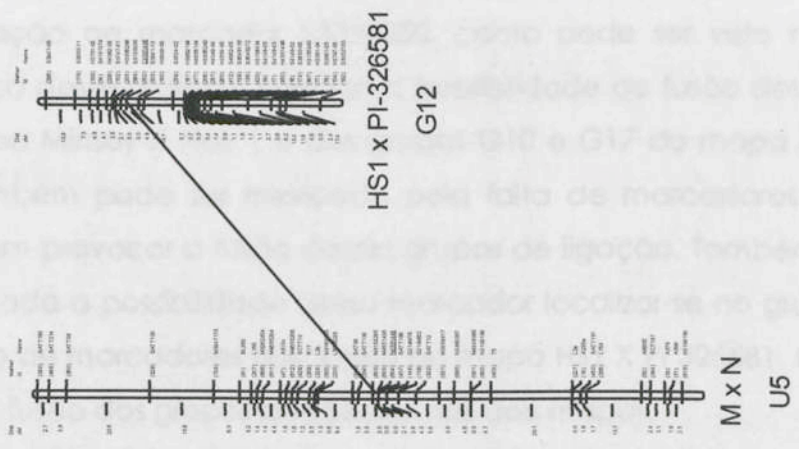
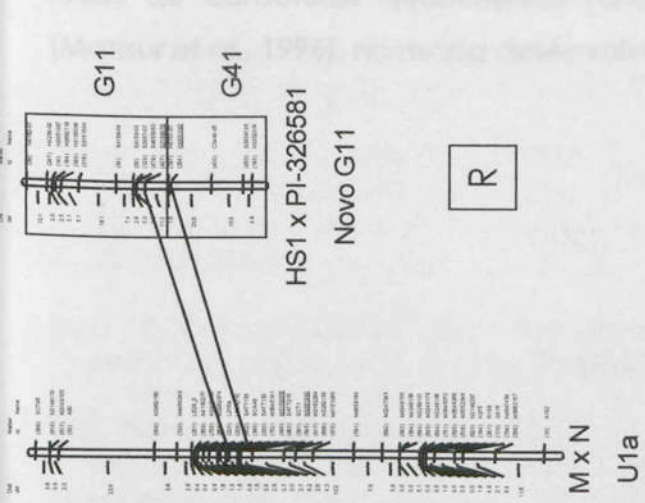


Figura 7, Cont.



apresentando dois pontos de ancoramento com o grupo G30 de HS1 X PI 326581. No mapa HS1 X PI 326581 foram realizadas tentativas de fusão entre os grupos G4 e G13 (Figura 7C), G18 e G44 (Figura 7K) e G25 e G32 (Figura 7M). No entanto, a fusão dos grupos excedeu o limite da frequência de recombinação de 50%. Isso pode ser explicado pela ausência de marcadores adicionais que poderiam sustentar a fusão dos grupos dentro do limite imposto, pelo programa MapMaker, de 50% de recombinação. Também, foi tentada a fusão dos grupos U1c e U30 (Figura 7F) e dos grupos U13 e U 15 (Figura 7G) do mapa Minsoy X Noir 1, no entanto a fusão dos grupos também excedeu o limite de 50%.

No geral, houve boa colinearidade entre os marcadores nos grupos de ligação entre os dois mapas. Ocorreu apenas pequena falta de colinearidade em relação ao marcador S3359350, como pode ser visto na Figura 7D. A presença desse marcador indica a possibilidade de fusão dos grupos U2 e U7 do mapa Minsoy X Noir 1 e dos grupos G10 e G17 do mapa HS1 X PI 326581. Isso também pode ser explicado pela falta de marcadores adicionais que poderiam provocar a fusão desses grupos de ligação. Também, não pode ser descartada a possibilidade desse marcador localizar-se no grupo G10, com a inclusão de marcadores adicionais ao mapa HS1 X PI 326581. Nesse caso, não haveria fusão dos grupos de ligação nos dois mapas.

A integração de mapas realizada neste trabalho permitiu a fusão de grupos de ligação em ambos os mapas e vai permitir o posicionamento dos locos de caracteres quantitativos (QTLs), mapeados em Minsoy X Noir 1 (Mansur *et al.*, 1996), no mapa desenvolvido pela Monsanto, HS1 X PI 326581.

Endersby, J.C., Polak, K., Labate, J., Specht, J., Brummer, E.C., Odeh, I., Young, M., Condeelis, V., Wilson, J., Jamieson, J.P., Kochert, G. & Boerma, H.R. Genome duplication in soybean (*Glycine subgenus soja*). *Genetics*, 166: 229-238, 1994.

Shan, P. Construction of integrated genetic linkage maps by means of a new computer package. *Join Map, Plant J.*, 3:739-744, 1993.

Vos, P., Hogert, R., Bentzen, M., Peleman, M., Lee, T., James, M., Rijbes, A., Pol, J., Falkman, J., Kubler, M. & Izopet, M. AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res.*, 22: 467-474, 1995.

REFERÊNCIAS BIBLIOGRÁFICAS

- Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A. e Cregan, P.B. Integration of simple sequence repeat DNA markers into soybean linkage map. **Crop Sci.**, **35**: 1439-1445, 1995.
- Delannay, X., Rodgers D.M. e Palmer, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Sci.**, **23**: 944-949, 1983.
- Don, R.H., Cox, P.T., Wainwright B.J. e Mattick J.S. "Touchdown" PCR to circumvent spurious priming during gene amplification. **Nuc. Acid. Res.**, **19**:4008, 1991.
- Keim, P., Schupp, J.M., Travis, S.E., Clayton, K., Ferreira, A. e Webb, D.M. A high-density soybean genetic map based upon AFLP markers. **Crop Sci.**, **37**:537-543, 1997.
- Lander E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S.E. e Newburg, L. MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, **1** :174-181, 1987.
- Mansur, L.M., Orf, J.H., Chase, K., Jarvik, T., Cregan, P.B. e Lark, K.G. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. **Crop Sci.**, **36**: 1327-1336, 1996.
- Shoemaker, R.C. e Olson, T.C. Molecular linkage map of soybean (*Glycine max* L. Merr.). In: **Genetic Maps: locus maps of complex genomes**. Cold Spring Harbour Laboratory Press, New York. 1993, p. 6.131-6.138.
- Shoemaker, R.C. e Specht, J.E. Integration of the soybean molecular and classical genetic linkage groups. **Crop Sci.**, **35**: 436-446, 1995.
- Shoemaker, R.C., Guffy, R.D., Lorenzen, L.L. e Specht, J.E. Molecular genetic mapping of soybean: map utilization. **Crop Sci.**, **32**: 1091-1098, 1992.
- Shoemaker, R.C., Polzin, K., Labate, J., Specht, J., Brummer, E.C. Olson, T., Young, N., Concibido, V., Wilcox, J., Tamulonis, J.P., Kochert, G. e Boerma, H.R. Genome duplication in soybean (*Glycine* subgenus *soja*). **Genetics**, **144**: 329-338, 1996.
- Stam, P. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. **Plant J.**, **3**: 739-744, 1993.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. e Zabeau, M. AFLP: a new technique for DNA fingerprinting. **Nucl. Acids Res.**, **23**: 4407-4414, 1995.

IDENTIFICAÇÃO DE LOCOS DE CARACTERES QUANTITATIVOS (QTLs) PARA TEORES DE ÁCIDOS GRAXOS EM SEMENTES DE SOJA UTILIZANDO MARCADORES AFLP

RESUMO

A fração óleo da semente de soja é composta dos ácidos graxos: palmítico, esteárico, oléico, linolênico e linoléico. É sabido que os teores desses ácidos graxos apresentam herança poligênica, o que dificulta a manipulação genética dessas características. Nesse caso, o uso de marcadores moleculares pode auxiliar no entendimento da herança dos teores desses ácidos graxos. Este trabalho teve como objetivo o mapeamento de QTLs para teores de ácidos graxos em soja. Sementes de soja de 357 indivíduos de uma população F_2 do cruzamento entre uma linhagem de baixo teor de ácido linolênico, BARC-12, e a variedade CAC-1 foram analisadas quanto ao teor dos cinco ácidos graxos (palmítico, esteárico, oléico, linolênico e linoléico) por meio de cromatografia gasosa, usando-se um método não-destrutivo. Quinze sementes com maiores e menores teores de cada ácido graxo foram selecionadas e plantadas em casa de vegetação, sendo utilizados trifolíolos não-expandidos para a extração de DNA. Um total de 89 plantas foram analisadas pela técnica de AFLP. Trinta e cinco pares de *primers* foram utilizados, gerando-se 158 marcadores dominantes e 18 co-dominantes. Foram identificados QTLs para todos os ácidos graxos, ressaltando-se que um marcador co-dominante explicou 50% da variação do teor de ácido linolênico. Linhas recombinantes de autofecundação (RILs) estão sendo desenvolvidas para esse cruzamento, o que vai permitir realizar estudos mais detalhados sobre a herança do teor de ácidos graxos em sementes de soja.

Palavras-Chave: ácido graxo, QTLs, AFLP, DNA e soja.

MAPPING QUANTITATIVE TRAIT LOCI (QTLs) FOR FATTY ACID COMPOSITION IN SOYBEAN SEEDS USING AFLP MARKERS

ABSTRACT

Soybean seed oil fraction is composed by five fatty acids: palmitic, stearic, oleic, linoleic and linolenic. The fatty acid content show a polygenic inheritance making it difficult to manipulate these traits in breeding programs. Molecular markers help to understand the inheritance of fatty acids content in soybean seeds. The objective of this work was to map quantitative trait loci (QTLs) for fatty acid composition in soybean seeds. Soybean seeds of 357 F₂ individuals from a cross between the low linolenic line BARC-12 and the cultivar CAC-1 were analyzed for the composition of palmitic, stearic, oleic, linoleic and linolenic acids by the means of gas chromatography using a non destructive method. Fifteen seeds with low and high-content for each fatty acid were selected and planted in green house to generate a leaflet for DNA extraction. A total of 89 plants were analyzed with the AFLP technique using 35 primer pairs that produced 158 dominant and 18 co-dominant markers. It was possible to identify QTLs for all fatty acids and one co-dominant marker explained 50% of the variation for the content of linolenic. Recombinant inbred lines (RILs) are being developed for this cross to allow more detailed studies about the inheritance of fatty acid content in soybean seeds.

Key Words: fatty acid, QTLs, AFLP, DNA and soybean.

INTRODUÇÃO

O mercado de soja está cada vez mais especializado, requerendo variedades modificadas para várias características, como teor de óleo e proteína, composição de ácidos graxos e teor de açúcares solúveis, dentre outras (Embrapa, 1998). Grande parte dos caracteres de interesse dos programas de melhoramento é controlada por vários genes, sendo considerada poligênica ou quantitativa. Tais caracteres são altamente influenciados pelo ambiente e de difícil manipulação e compreensão (Mansur *et al.*, 1996). Para a maioria dos caracteres quantitativos, pouco é conhecido a respeito do número, da localização e das interações de genes que controlam a sua expressão. A qualidade do óleo de soja é uma característica agrônômica de grande interesse atualmente nos programas de melhoramento. Cinco ácidos graxos compõem a fração óleo do grão de soja: palmítico, esteárico, oléico, linolênico e linoléico, cujos teores também apresentam herança poligênica. O melhoramento para características quantitativas é dificultado, uma vez que os efeitos dos genes individuais não são facilmente identificados (Dudley, 1993). O estudo dos locos que controlam características quantitativas (QTLs), por meio da associação destes a marcadores moleculares, possibilita a dissecação de caracteres complexos nos seus componentes genéticos individuais, possibilitando também lidar com eles, com a eficácia de características simples. O uso de mapas genéticos obtidos com marcadores moleculares oferece a oportunidade de identificar, mapear e quantificar o efeito dos QTLs. A detecção de QTLs vai depender da magnitude do efeito do QTL, do tamanho da população estudada e da frequência de recombinação entre o QTL e o marcador que depende da saturação do mapa (Tanksley *et al.*, 1989). Este trabalho teve como objetivo o mapeamento de QTLs para teores de ácidos graxos no cruzamento entre a linhagem de baixo teor de ácido linolênico, BARC-12, e a variedade CAC-1, utilizando-se a técnica de AFLP (Vos *et al.*, 1995).

MATERIAL E MÉTODOS

Material genético e determinação de ácidos graxos

O mapeamento de QTLs para teores de ácidos graxos em sementes de soja foi realizado numa população F_2 do cruzamento de uma linhagem de soja com baixo teor de linolênico (3,0%), BARC-12 (USDA-ARS at Beltsville, MD, EUA), com a variedade CAC-1 (Cooperativa Agrícola de Cotia, Cotia, SP), que apresenta teores normais de linolênico (8,5%). Toda a população F_2 de 357 indivíduos foi analisada quanto aos teores dos ácidos graxos que compõem a fração óleo do grão de soja: palmítico, esteárico, oléico, linoléico e linolênico. Por meio de um método não-destrutivo, 15 mg de cada semente F_2 foram utilizados para extração lipídica seguida de transesterificação com 1 M de metóxido de sódio. A fase orgânica foi extraída com hexano, e 1 μ L de cada amostra foi utilizada para cromatografia gasosa. Foi utilizada uma coluna carbowax de 30 m de comprimento e 0,32 cm de diâmetro interno, no equipamento SHIMADZU 17A. Foi usado nitrogênio (1,1 mL/min) para fazer o arraste dos ácidos graxos.

A metodologia da genotipagem seletiva (Lander e Botstein, 1989; Darvasi e Soller, 1992), com modificações, foi utilizada para selecionar indivíduos a serem avaliados pela técnica de AFLP. Esta técnica baseia-se no fato de que a maioria da informação sobre a ligação marcador-QTL vem dos indivíduos com fenótipos extremos. Portanto, apenas os extremos da curva normal são avaliados com marcadores moleculares. Foram selecionadas 15 sementes com os maiores teores e 15 sementes com os menores teores para cada ácido graxo. Um total de 89 indivíduos foi selecionado, devido à correlação fenotípica entre os teores dos ácidos graxos. Os indivíduos selecionados foram analisados pela técnica de AFLP.

Isolamento e quantificação de DNA

As 89 sementes selecionadas foram plantadas em casa de vegetação, sendo coletadas folhas jovens de cada planta para extração de DNA, segundo o método do CTAB, com modificações. Folhas jovens foram trituradas

na presença de N₂ líquido para obter cerca de 300 mg de pó, sendo este transferido para um tubo de microcentrífuga de 2 mL. Foi adicionado ao pó um tampão de extração, preaquecido a 65°C, contendo Tris-HCl 100 mM (pH 8), EDTA 20 mM, NaCl 1,4 M, CTAB 2% e 2-mercaptoetanol 1%. As proteínas foram removidas com um volume de clorofórmio - álcool isoamílico (24:1). O DNA foi precipitado com um volume de isopropanol gelado, lavado com etanol 70%, secado a vácuo, ressuspendido em tampão TE_{0,1} (Tris-HCl 10mM (pH 8,0) e EDTA 0,1 mM). A concentração das amostras de DNA foi estimada por meio da comparação com padrões de DNA de fago lãmbda em géis de agarose.

Reações de AFLP

Digestão do DNA: duzentos e cinqüenta nanogramas de DNA de cada amostra foram digeridos simultaneamente com cinco unidades de *EcoRI* e *MseI* em 40 µl de tampão contendo Tris-acetato 10 mM (pH 7,5); acetato de magnésio 10 mM; acetato de potássio 50 mM; ditiotreitól 5 mM; e BSA (albumina de soro bovino) 50 ng/µL, incubados por 3 h a 37°C.

Ligação: a ligação dos adaptadores foi feita com a adição de 10 µL de uma solução contendo 5 pmoles de adaptador de *EcoRI*; 50 pmoles de adaptador de *MseI*; uma unidade de T4 DNA-ligase; ATP 1 mM; Tris-acetato 10 mM (pH 7,5); acetato de magnésio 10 mM; acetato de potássio 50 mM; ditiotreitól 5 mM; e BSA 50 ng/µL, incubados por 3 h a 37°C. Após a ligação, a mistura de reação foi diluída a 1:10 com TE_{0,1}.

Amplificação pré-seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de um nucleotídeo seletivo, reduzindo em 16 vezes a complexidade dos fragmentos. Cinco microlitros, da mistura diluída da restrição/ligação de cada amostra, foram pré-amplificados em um volume final de 50 µL, contendo 50 ng dos *primers* de *EcoRI* e *MseI* acrescidos de um nucleotídeo seletivo; dNTPs 200 µM; uma unidade de *Taq* polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl 50 mM. As reações de amplificação do DNA foram aplicadas no termociclador Perkin-Elmer 9600 (Norwalk, CT, EUA). Os ciclos foram os seguintes: 20 ciclos de 94°C por 30 segundos, 55°C por 30 segundos e 72°C por 60 segundos. Após a

amplificação pré-seletiva, os produtos das reações foram diluídos a 1:20 com TE_{0,1}.

Amplificação seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de três nucleotídeos seletivos, reduzindo-se em 256 vezes a complexidade dos fragmentos. Para amplificação seletiva, o *primer* seletivo de EcoRI foi marcado na extremidade 5', utilizando-se [γ -³³P] ATP. A reação de marcação foi composta de 50 μ L, contendo 500 ng de *primer* de EcoRI com três nucleotídeos seletivos; [γ -³³P] ATP 100 μ Ci; 10 unidades de T4 polinucleotídeo cinase; Tris-HCl 25 mM (pH 7,5); MgCl₂ 10 mg; ditiotretitol 5 mM; e espermidina-3HCl 0,5 mM. Cinco microlitros da mistura de amplificação pré-seletiva, diluída a 1:20, foram amplificados num volume total de 20 μ L, contendo 5 ng do *primer* de EcoRI marcado com três nucleotídeos seletivos; 30 ng de *primer* de MseI com três nucleotídeos seletivos; dNTPs 0,2 mM; 0,4 unidades de Taq polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl 50 mM. Os ciclos de amplificação seletiva foram os seguintes: 12 ciclos de 94°C por 30 segundos, 65°C (-0,7°C/ciclo) por 30 segundos e 72°C por 60 segundos, até atingir a temperatura ótima de ligação dos *primers* de 56°C - Touchdown Cycle Profile (Don et al., 1991).

Primeiramente, foi realizada uma avaliação do polimorfismo entre os dois progenitores, BARC-12 e CAC-1, com 64 pares de *primers*. Destes, foram selecionados 35 pares de *primers* mais polimórficos, que foram utilizados para amplificação dos 89 indivíduos selecionados.

Eletroforese em gel de poli-acrilamida

Os produtos da amplificação seletiva foram misturados com um volume igual (20 μ L) de tampão contendo formamida 98%, EDTA 10mM, xileno cianol FF 1 mg/mL e azul-de-bromofenol 1 mg/mL. As amostras foram, então, desnaturadas a 95°C por cinco minutos e mantidas em gelo até serem aplicadas no gel. Um microlitro e meio de cada amostra foi aplicado em um gel desnaturante de poli-acrilamida 4,5% (acrilamida:bisacrilamida 19:1, uréia 7,5 M e Tris-borato EDTA (TBE) 0,5x), utilizando-se a cuba de eletroforese SequiGen de 38 x 50 cm (BioRad Laboratories Inc., Hercules, CA, EUA). Foram

utilizados dois pentes, de 48 dentes, para cada gel. Em cada pente foram carregados os progenitores BARC-12 e CAC-1 e um padrão 10 BL que mostra uma banda de DNA a cada 10 pb, entre 80 e 530 pb. A placa de vidro externa não recebeu nenhum tratamento, e a placa de vidro interna foi coberta com 3 mL de Sigma cote para a não-aderência do gel. A eletroforese foi realizada em TBE 1x, 130 W, 50°C por 1,5 h. Após a eletroforese, as placas foram separadas e o gel permaneceu ligado à placa de vidro externa, que não recebeu nenhum tratamento. O gel foi fixado em papel 3 MM, envolto por um filme de plástico, secado a vácuo por 1 h a 70°C e exposto a um *phosphor screen* por 20 h. Os *screens* foram digitalizados, usando-se o *Storm System* (Molecular Dynamics, CA, EUA).

Scoring dos géis e análise dos dados

As imagens geradas pelo *Storm System* foram convertidas para o programa da Keygene (Wageningen, Holanda) para serem analisadas. Cada imagem foi convertida em duas, gerando-se o gel da esquerda e o gel da direita. Uma vez convertida a imagem, o primeiro procedimento foi a definição das faixas, seguindo-se as curvas do gel para que o programa identificasse cada amostra. Cada faixa do gel foi nomeada de acordo com a amostra carregada. A próxima etapa foi a definição das bandas comuns para definir a forma do gel. Uma vez definida a forma do gel, foi feito o *scoring* das bandas polimórficas, sendo (+) para a presença da banda, (-) para a ausência da banda e (?) para falhas ou dúvida quanto ao *scoring*.

Feito o *scoring* dos géis, os dados foram exportados pelo programa Keygene, gerando-se uma matriz +/- . Os marcadores que apresentavam a origem da banda de BARC-12 receberam a letra B e os que apresentavam a origem da banda de CAC-1, a letra C. Após a letra B ou C, vinham os três nucleotídeos seletivos do *primer* de *EcoRI*, seguidos dos três nucleotídeos seletivos do *primer* de *MseI* e do tamanho do fragmento em pb. Exemplificando, o marcador BACCCTA234 caracteriza uma banda originada de BARC-12, gerada pelo par de *primers* seletivos *EcoRI* ACC e *MseI* CTA, com o tamanho de 234 pb. A matriz +/- foi convertida para uma matriz de A, B, C, D e H para ser utilizada pelos programas de mapeamento. As designações

foram as seguintes: A= ausência de banda para BARC-12, C= presença de banda para BARC-12, B= ausência de banda para CAC-1 e D= presença de banda para CAC-1. Os marcadores co-dominantes receberam a seguinte denominação: A= homozigoto para CAC-1, B= homozigoto para BARC-12 e H= heterozigoto. Todos os marcadores co-dominantes receberam a letra H antes do par de *primers* seletivos. Os grupos de ligação foram determinados pelo programa MapMaker 3.0 (Lander *et al.*, 1987), utilizando-se LOD 3,0. A ordem dos marcadores em cada grupo de ligação foi determinada pelo programa JoinMap (Stam, 1993). Os grupos de ligação foram gerados graficamente pelo programa MapMaker para Macintosh, utilizando-se a distância de Kosambi para conversão das freqüências de recombinação em centimorgans (cM).

A identificação de marcadores associados aos teores de ácido graxo foi realizada por meio de uma análise de variância simples entre as médias dos teores dos ácidos graxos e os dados moleculares. A proporção da variância fenotípica explicada pela segregação de um marcador ligado a um QTL foi determinada por uma análise de regressão. A localização dos QTLs nos grupos de ligação foi determinada por meio do mapeamento por intervalo (Lander e Botstein, 1989), utilizando-se o pico do valor de F como estimativa da posição mais provável do QTL. Os valores de significância ($F_{0,05}$ e $F_{0,01}$) foram calculados para cada ácido graxo e para cada grupo de ligação, por meio de permutações (Doerge e Churchill, 1996). Os efeitos aditivo (a) e dos desvios de dominância (d) foram calculados para os marcadores co-dominantes, sendo corrigidos pela freqüência de recombinação (r) entre o marcador e o QTL (Edwards *et al.*, 1987). O grau médio de dominância foi determinado pela razão d/a. Todas as análises envolvendo QTLs foram feitas através do programa Qgene, desenvolvido pela Universidade de Cornell (Ithaca, NY, EUA).

em conexão com CAC-1, que apresenta uma banda normal (B+C). Poucos indivíduos da população B exibiram valores que decediram de dois progenitores BARC-12 e CAC-1.

No total QTL, apresentaram os coeficientes de correlação negativa dos teores dos cinco ácidos graxos na população B do cruzamento de BARC-12 X CAC-1. Os valores obtidos como de 0,145 são significativos ao nível de 0,01 de probabilidade. Os valores de CAC-1 e B exibiram semelhanças com o QTL de probabilidade de 0,001. A análise de regressão linear de correlação de -0,287 e a de regressão linear de correlação de -0,305, de

RESULTADOS E DISCUSSÃO

Segregação dos teores de ácido graxo

A população F_2 de 357 indivíduos foi analisada quanto aos teores dos ácidos graxos que compõem a fração óleo do grão de soja: palmítico, esteárico, oléico, linoléico e linolênico. Na Figura 1, mostra-se a distribuição dos fenótipos em relação aos teores desses ácidos, os quais apresentaram distribuição normal, indicando que vários genes estão envolvidos, o que caracteriza uma herança quantitativa. O teor de oléico, no entanto, apresentou distribuição que diferiu um pouco da distribuição normal. Os progenitores BARC-12 e CAC-1 exibiram teores semelhantes para palmítico (9,9 e 10,3%, respectivamente) e esteárico (3,4 e 3,7%, respectivamente). A grande maioria dos indivíduos da população F_2 excedeu os progenitores BARC-12 e CAC-1, quanto aos teores de palmítico e esteárico, caracterizando-se a ocorrência de segregação transgressiva. Isso indica que os dois genótipos parentais apresentaram teores similares de palmítico e esteárico, por meio de combinações gênicas distintas. A ocorrência de segregação transgressiva também foi notada nos trabalhos de Mansur *et al.* (1993), Mansur *et al.* (1996), Hnetkovsky *et al.* (1996) e Xiao *et al.* (1996). BARC-12 e CAC-1 apresentaram teores mediantemente discrepantes para oléico (36,3 e 20,2%, respectivamente) e linoléico (47,0 e 57,1%, respectivamente). Alguns indivíduos da população F_2 apresentaram teores que excediam aos dos progenitores BARC-12 e CAC-1, quanto aos teores de oléico e linoléico. Em relação ao linolênico, os progenitores BARC-12 e CAC-1 mostraram teores bastante contrastantes, visto que BARC-12 é uma linhagem desenvolvida para baixo teor de linolênico (3,4%), em contraste com CAC-1, que apresenta teores normais (8,4%). Poucos indivíduos da população F_2 exibiram valores que excediam ao dos progenitores BARC-12 e CAC-1.

Na Tabela 1, apresentam-se os coeficientes de correlação fenotípica dos teores dos cinco ácidos graxos na população F_2 do cruzamento de BARC-12 X CAC-1. Os valores absolutos acima de 0,148 são significativos no nível de 0,01 de probabilidade. Os teores de oléico e esteárico apresentaram correlação de -0,2471 e os de oléico e linolênico, correlação de -0,3530. Os

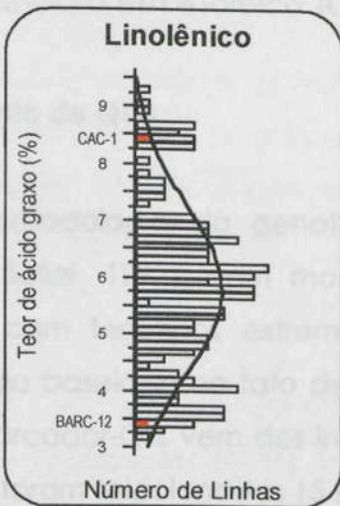
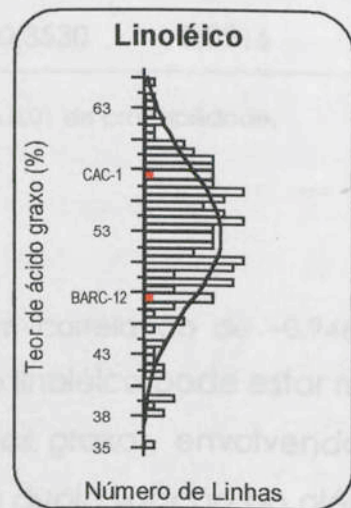
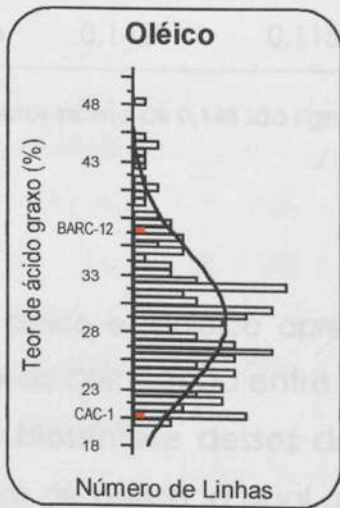
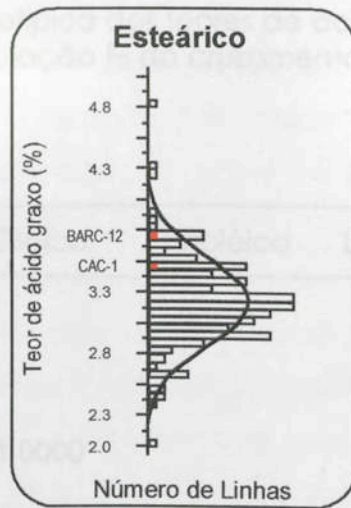
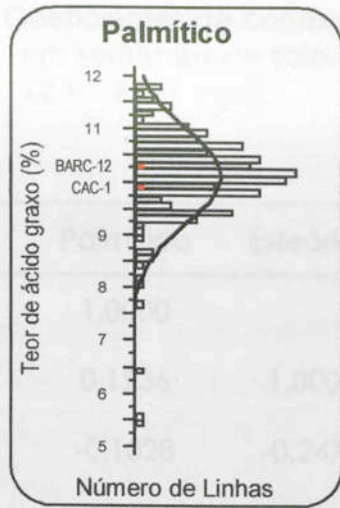


Figura 1 - Distribuição dos teores de cinco ácidos graxos da população F₂ do cruzamento dos genótipos de soja BARC-12 X CAC-1. Para cada ácido graxo são mostrados os valores dos progenitores BARC-12 e CAC-1.

Tabela 1 – Coeficientes de correlação fenotípica dos teores de ácidos graxos em sementes de soja, na população F₂ do cruzamento de BARC-12 X CAC-1

	Palmítico	Estearico	Oléico	Linoléico	Linolênico
Palmítico	1,0000				
Estearico	0,1236	1,0000			
Oléico	-0,1028	-0,2471	1,0000		
Linoléico	-0,0850	0,1550	-0,9460	1,0000	
Linolênico	0,1421	0,1156	-0,3530	0,0916	1,0000

Valores absolutos acima de 0,148 são significativos a 0,01 de probabilidade.

teores de oléico e linoléico apresentaram correlação de $-0,9460$. Esse alto coeficiente de correlação entre oléico e linoléico pode estar relacionado à rota de biossíntese desses dois ácidos graxos, envolvendo a enzima dessaturase de oléico, a qual adiciona dupla ligação ao oléico (C18:1), transformando-o em linoléico (C18:2).

Mapeamento de QTLs

A metodologia da genotipagem seletiva (Lander e Botstein, 1989; Darvasi e Soller, 1992), com modificações, foi utilizada para selecionar os indivíduos com fenótipos extremos da distribuição normal (Figura 1). Essa metodologia baseia-se no fato de que a maior parte da informação sobre a ligação marcador-QTL vem dos indivíduos extremos. Da população F₂, de 357 indivíduos, foram selecionadas 15 sementes com os maiores teores e 15 com os menores teores de cada ácido graxo. Os indivíduos, num total de 89, foram selecionados, devido à correlação fenotípica entre os teores dos ácidos graxos, e analisados pela técnica de AFLP.

Para determinar quais os pares de *primers* de AFLP eram mais polimórficos para a população, foi realizada uma avaliação prévia incluindo apenas os progenitores BARC-12 e CAC-1, mediante o emprego de 64 pares de *primers* (Figura 2). Destes, foram selecionados 35 pares de *primers* mais polimórficos, para avaliação dos 89 indivíduos da população F₂ (Figura 3). Os 35 pares de *primers* selecionados produziram um total de 194 bandas polimórficas, sendo que apenas 11% delas apresentaram desvio da segregação do tipo 3:1 (3 presenças : 1 ausência da banda), confirmada pelo teste do qui-quadrado. A técnica de AFLP é essencialmente dominante, pois, na grande maioria das vezes, o outro alelo não está presente no mesmo gel. No entanto, quando o outro alelo é identificado no mesmo gel, é possível realizar o *scoring* do marcador como co-dominante. O *scoring* das bandas permitiu identificar, entre os 194 marcadores obtidos, 36 marcadores, que puderam ser transformados em 18 marcadores co-dominantes. Portanto, foram identificados 158 marcadores dominantes e 18 co-dominantes, totalizando-se 176 marcadores gerados pelos 35 pares de *primers*, o que representa uma média de cinco marcadores para cada par de *primer*. A média de cinco marcadores/par de *primer* ficou bem abaixo do valor encontrado por Keim *et al.* (1997), que obtiveram uma média de nove marcadores/par de *primer* num cruzamento de BSR101 X PI-437654. Isso, provavelmente, ocorreu pelo fato de esses autores terem utilizado uma PI, que por ser não-adaptada gerou uma população com maior diversidade e, conseqüentemente, maior número de marcadores. BARC-12 e CAC-1 são genótipos adaptados que apresentam base genética mais reduzida (Delannay *et al.*, 1983) e, portanto, com menor polimorfismo.

Os 176 marcadores de AFLP foram utilizados para construir um mapa genético de ligação para essa população, o qual cobriu uma distância de recombinação de 1.476 cM, em 26 grupos de ligação. Um total de 21 marcadores permaneceu não ligado, e a distância média obtida entre dois marcadores adjacentes foi de 9,5 cM. Essa baixa saturação foi devida ao pequeno número de marcadores utilizado para gerar o mapa.

Na Tabela 2, mostram-se os marcadores que apresentaram associação significativa com um QTL (P<0,01), correspondendo a um limite de F=4,68. Foram encontradas associações marcador-QTLs para todos os ácido graxos.

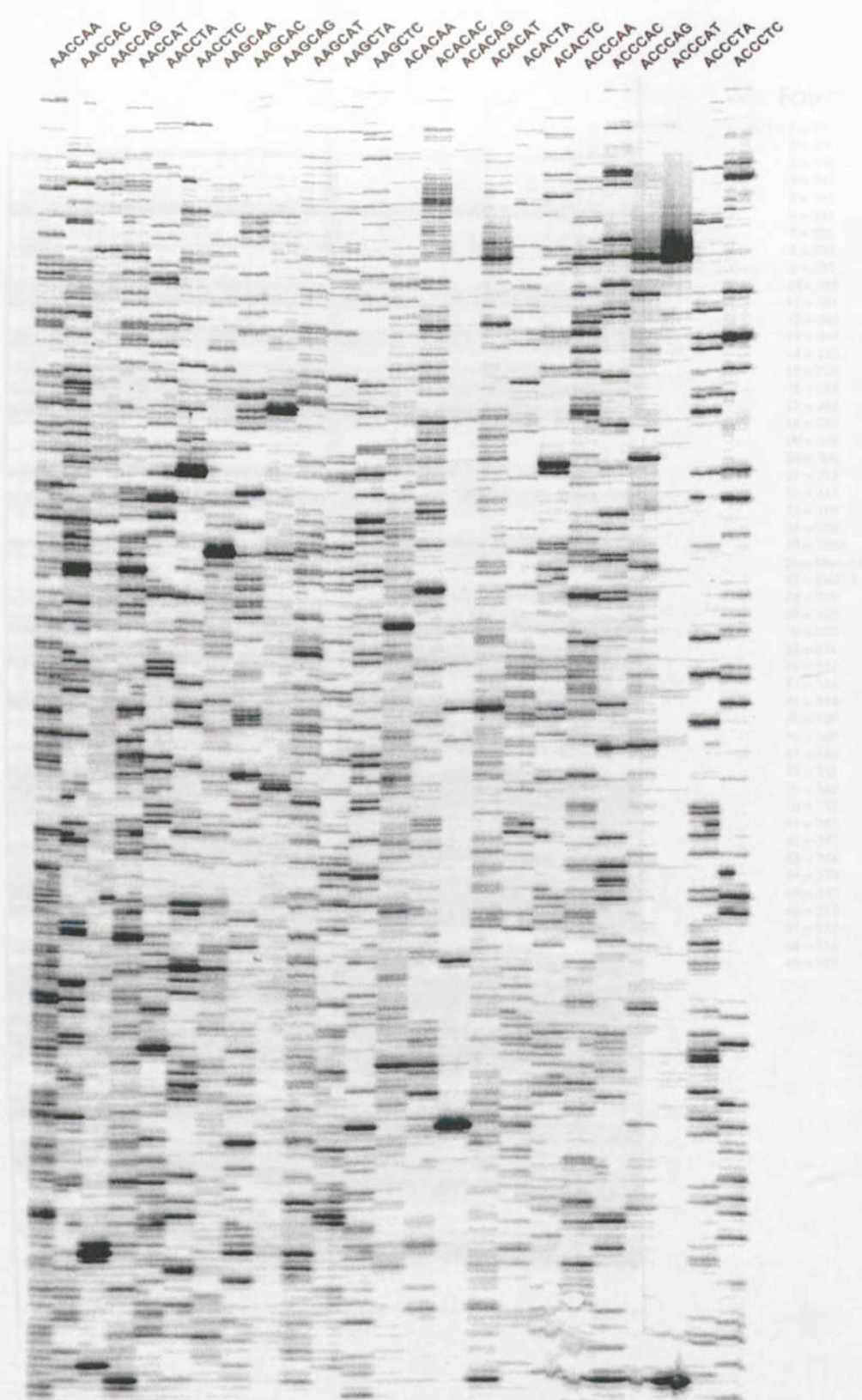


Figura 3 – Perfil de amplificação de DNA obtido pela técnica de AFLP utilizando-se o par de primers ACC/CTA na população F₂ do cruzamento de BARC12 x CACT. As fotos da gel estão invertidas.

Figura 2 – Avaliação do polimorfismo entre os progenitores BARC-12 (à esquerda) e CAC-1 (à direita), utilizando-se 24 pares de primers de AFLP. As três primeiras letras representam o primer seletivo (+3) de EcoRI, seguido pelo primer seletivo (+3) de MseI. Os pares de primers que produziram número maior de bandas polimórficas, entre os dois progenitores, foram utilizados para avaliação da população F₂.

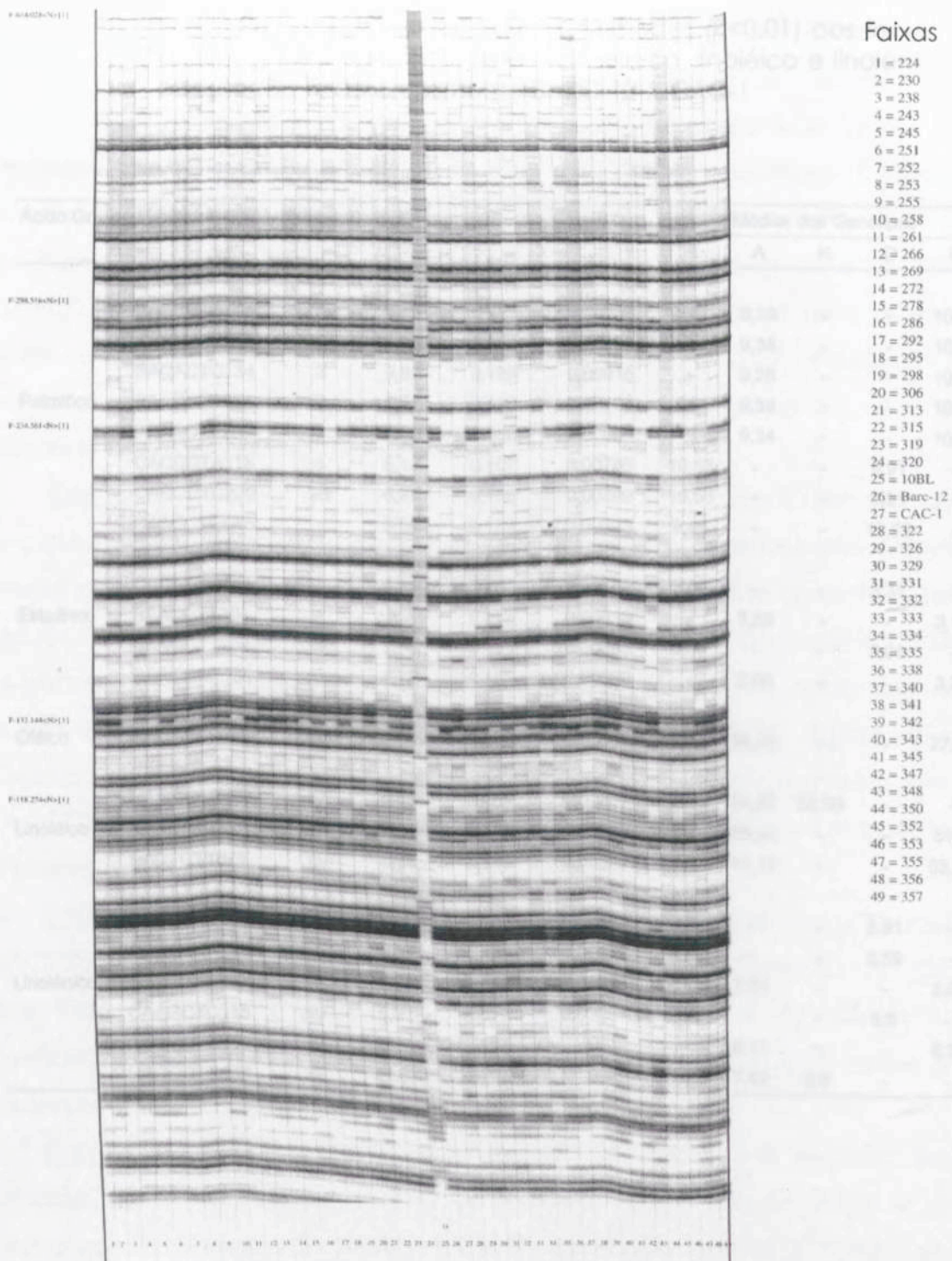


Figura 3 – Padrão de amplificação de DNA obtido pela técnica de AFLP, utilizando-se o par de *primers* ACC/CTA na população F_2 do cruzamento de BARC-12 X CAC-1. As faixas do gel estão caracterizadas conforme a legenda à direita. O padrão 10 BL, que mostra uma banda a cada 10 pb entre 80 e 530 pb; BARC-12; e CAC-1 foram carregados nas faixas 25, 26 e 27, respectivamente. Esse par de *primers* produziu cinco bandas polimórficas, e o *scoring* foi feito, considerando-se a presença da banda = + e a ausência da banda = -

Tabela 2 – Marcadores significativamente associados ($P < 0,01$) aos teores dos ácidos graxos palmítico, esteárico, oléico, linoléico e linolênico na população F_2 do cruzamento BARC-12 X CAC-1

Ácido Graxo	Marcador	Grupo	F	R ²	P	Médias dos Genótipos				
						B	A	H	D	C
Palmítico	BAACCTA144	3	6,49	0,131	0,00237	-	9,39	-	-	10,21
	BACTCAT209	3	7,76	0,154	0,00080	-	9,35	-	-	10,23
	BACACTC134	3	9,81	0,186	0,00015	-	9,28	-	-	10,25
	BAACCTA054	3	6,87	0,139	0,00172	-	9,39	-	-	10,23
	BAACCTT263	3	7,63	0,152	0,00090	-	9,34	-	-	10,23
	CACCCTA118	4	5,13	0,107	0,00786	10,57	-	-	9,91	-
	CACCCTC290	20	6,23	0,128	0,00299	10,65	-	-	9,91	-
CAACCTG085	-	6,56	0,132	0,00223	9,44	-	-	10,22	-	
Esteárico	CACACAT614	4	6,76	0,136	0,00188	3,06	-	-	3,31	-
	BAACCTT305	4	5,4	0,112	0,00618	-	3,35	-	-	3,11
	CAACCTT294	5	7,58	0,154	0,00095	3,35	-	-	3,08	-
	BAACCAC135	8	7,73	0,152	0,00082	-	2,96	-	-	3,26
Oléico	BACTCTT072	25	5,04	0,105	0,00852	-	31,99	-	-	27,84
Linoléico	HAGCCTC480	17	5,03	0,108	0,00867	48,49	54,67	52,69	-	-
	BAACCTT166	17	7,51	0,150	0,00099	-	55,45	-	-	51,1
	BACTCTT072	25	5,08	0,106	0,00822	-	49,17	-	-	53,07
Linolênico	CACCTT194	4	7,1	0,142	0,00140	6,82	-	-	5,61	-
	CACCCAA415	4	6,87	0,139	0,00172	6,79	-	-	5,59	-
	BACTCTT110	16	8,15	0,161	0,00058	-	7,01	-	-	5,61
	CAAGCTC115	20	8,91	0,175	0,00031	7,03	-	-	5,6	-
	BAGCCTG155	22	18,39	0,302	0	-	6,71	-	-	5,23
	HAGCCTG116	22	41,52	0,497	0	4,08	7,42	5,9	-	-

Palmítico: a análise de variância (Tabela 2) indicou oito marcadores associados significativamente ao teor de palmítico, sendo cinco marcadores no grupo 3, um marcador nos grupos 4 e 20 e um marcador não ligado. O marcador mais fortemente associado foi BACTCT134, que explicou 18,6% da variação do teor de palmítico. O mapeamento por intervalo (Figura 4A) identificou uma região no grupo 3, situada entre os marcadores BAACCTA144 e BAACCTT263, significativa a $F_{0,01}$. No grupo 4, foi detectada uma região situada entre os marcadores CACCCTA118 e CAACCTC257, significativa apenas a $F_{0,05}$. O marcador CACCCTC290, localizado no grupo 20, não foi detectado no mapeamento por intervalo.

Esteárico: a análise de variância (Tabela 2) indicou quatro marcadores associados significativamente ao teor de esteárico, sendo dois marcadores no grupo 4 e um nos grupos 5 e 8. O marcador mais fortemente associado foi BAACCAC135, que explicou 15,2% da variação do teor de esteárico. O mapeamento por intervalo (Figura 4B) identificou uma região no grupo 4, situada entre os marcadores CACACAT614 e BAAGCAT150, significativa apenas a $F_{0,05}$. O marcador BAACCAC135 localizado no grupo 8 e o marcador CAACCTT294 localizado no grupo 5 não foram detectados no mapeamento por intervalo.

Oléico: a análise de variância (Tabela 2) detectou o marcador BACTCTT072, localizado no grupo 25, associado significativamente ao teor de oléico. Esse marcador explicou 10,5% da variação do teor de oléico. O mapeamento por intervalo não detectou regiões significativamente associadas ao teor deste ácido.

Linoléico: a análise de variância (Tabela 2) indicou três marcadores associados significativamente ao teor de linoléico, sendo dois no grupo 17 e um no grupo 25. O marcador mais fortemente associado foi BAACCTT166, que explicou 15% da variação do teor de linoléico. O mapeamento por intervalo (Figura 4C) identificou uma região no grupo 17, situada entre os marcadores HAGCCTC480 e BAACCTT166, significativa para $F_{0,01}$. O marcador BACTCTT072, localizado no grupo 25, não foi detectado no mapeamento por intervalo.

Linolênico: a análise de variância (Tabela 2) mostrou seis marcadores associados significativamente ao teor de linolênico, sendo dois nos grupos 4 e 22 e um nos grupos 16 e 20. O marcador mais fortemente associado foi

Palmitico

Esteárico

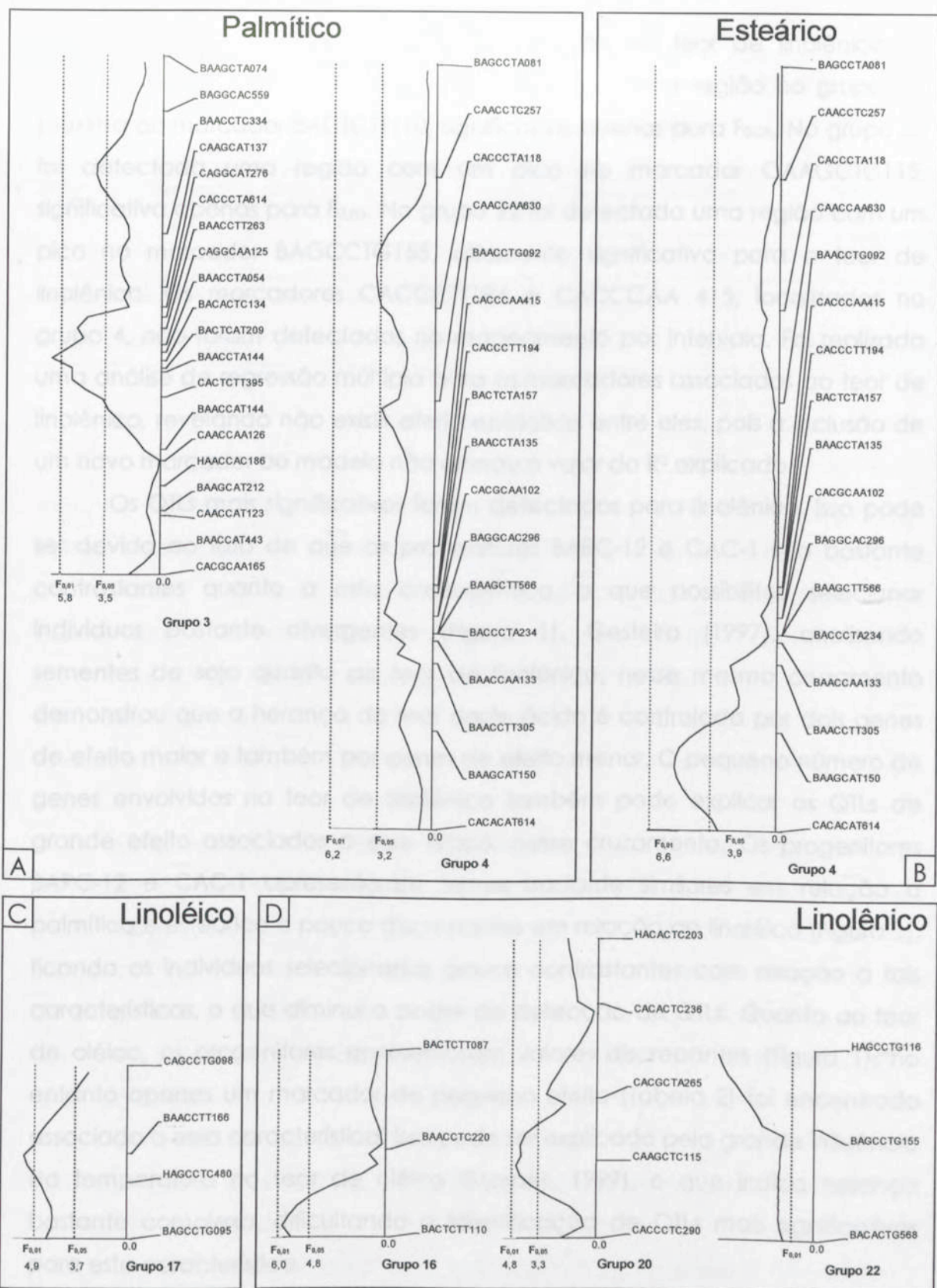


Figura 4 – Mapeamento por intervalo para os teores de palmítico, esteárico, linoléico e linolênico. O pico do valor de F indica a provável localização do QTL. F_{0,01} e F_{0,05} foram calculados por meio de permutações (Doerge e Churchill, 1996).

HAGCCTG116, que explicou 49,7% da variação do teor de linolênico. O mapeamento por intervalo (Figura 4D) identificou uma região no grupo 16, próxima ao marcador BACTCT110, significativa apenas para $F_{0,05}$. No grupo 20 foi detectada uma região com um pico no marcador CAAGCTC115, significativa apenas para $F_{0,05}$. No grupo 22 foi detectada uma região com um pico no marcador BAGCCTG155, altamente significativa para o teor de linolênico. Os marcadores CACCCT194 e CACCCTA 415, localizados no grupo 4, não foram detectados no mapeamento por intervalo. Foi realizada uma análise de regressão múltipla para os marcadores associados ao teor de linolênico, revelando não existir efeito epistático entre eles, pois a inclusão de um novo marcador ao modelo não alterou o valor do R^2 explicado.

Os QTLs mais significativos foram detectados para linolênico. Isso pode ser devido ao fato de que os progenitores BARC-12 e CAC-1 são bastante contrastantes quanto a esta característica, o que possibilitou selecionar indivíduos bastante divergentes (Figura 1). Gesteira (1997), analisando sementes de soja quanto ao teor de linolênico, nesse mesmo cruzamento demonstrou que a herança do teor deste ácido é controlada por dois genes de efeito maior e também por genes de efeito menor. O pequeno número de genes envolvidos no teor de linolênico também pode explicar os QTLs de grande efeito associados a esse ácido, nesse cruzamento. Os progenitores BARC-12 e CAC-1 apresentaram teores bastante similares em relação a palmítico e esteárico e pouco discrepantes em relação ao linoléico (Figura 1), ficando os indivíduos selecionados pouco contrastantes com relação a tais características, o que diminui o poder de detecção de QTLs. Quanto ao teor de oléico, os progenitores apresentaram valores discrepantes (Figura 1), no entanto apenas um marcador de pequeno efeito (Tabela 2) foi encontrado associado a essa característica. Isso pode ser explicado pela grande influência da temperatura no teor de oléico (Moraes, 1999), o que indica herança bastante complexa, dificultando a identificação de QTLs mais significativos para esta característica.

Os efeitos aditivo (a) e dos desvios de dominância (d) foram calculados para os marcadores co-dominantes (Tabela 3). O marcador HAGCCTG116 possui efeito aditivo bastante acentuado para o teor de linolênico, visto que a razão d/a foi de apenas 0,09, indicando haver ausência de dominância neste

Tabela 3 - Efeitos genéticos de QTLs associados aos ácidos graxos linoléico e linolênico na população F₂ do cruzamento de BARC-12 X CAC-1

Ácido Graxo	Marcador	Grupo	F	R ²	P	Médias dos Genótipos			Efeitos Genéticos		
						B	A	H	a	d	d/a
Linoléico	HAGCCTC480	17	5,03	0,108	0,00867	48,49	54,67	52,69	3,67	1,57	0,43
Linolênico	HAGCCTG116	22	41,52	0,497	0	4,08	7,42	5,9	1,67	0,15	0,09

loco. Os valores corrigidos de d e a para este marcador foram os próprios valores obtidos das médias, pois o QTL se encontrava exatamente na posição do marcador (Figura 4D). O marcador HAGCCTC480 mostrou dominância parcial quanto ao teor de linoléico, sendo a razão d/a de 0,43, visto que os alelos dominantes tendem a aumentar o teor de linoléico. Os valores de d e a para este marcador foram corrigidos, pois o pico da atividade de QTL se encontrava no marcador BAACCTT166, situado a 8,1 cM do mesmo (Figura 4C).

A segregação transgressiva foi confirmada pela ocorrência de herança de alelos complementares para QTLs, oriundos dos dois progenitores. Os marcadores CACCCTT194, CACCCAA415 e CAAGCTC115, oriundos do progenitor de alto teor de linolênico (CAC-1), apresentaram alelos que diminuía o teor de linolênico (Tabela 2). Isso indica que o progenitor CAC-1, apesar de possuir alto teor de linolênico, possui também alelos para baixo teor, que não são detectados fenotipicamente. O mesmo foi também encontrado para os marcadores CAACCTG085, quanto ao teor de palmítico; CACACAT614 e BAACCTT305, quanto ao teor de esteárico; e BACTCTT072, quanto aos teores de oléico e linoléico. A ocorrência de complementação alélica em QTLs também foi notada por Xiao *et al.* (1996) e Lee *et al.* (1996).

A metodologia da genotipagem seletiva pode superestimar a variação explicada pelo QTL, se comparada com a genotipagem da população inteira (Darvasi e Soller, 1992). Neste trabalho foi selecionado um total de 30 indivíduos com fenótipos extremos para cada ácido graxo (15 indivíduos com

maior teor e 15 indivíduos com menor teor), totalizando 89 indivíduos genotipados pela técnica de AFLP. Portanto, cada ácido graxo foi representado por 30 indivíduos com fenótipos extremos e 59 indivíduos aleatórios, considerando-se a baixa correlação geral entre os caracteres (Tabela 1). Desse modo, espera-se que os marcadores obtidos não apresentem desvios de segregação, o que foi confirmado pelo teste de qui-quadrado, pois apenas 11% dos marcadores apresentaram desvios de segregação do tipo 3:1 para os marcadores dominantes e 1:2:1 para os marcadores co-dominantes. A genotipagem seletiva, embora tenha sido utilizada para selecionar os indivíduos extremos para cada ácido graxo, não deve estar superestimando a variação explicada pelos QTLs detectados, visto que os marcadores obtidos exibiram pequeno desvio de segregação.

Gesteira (1997), trabalhando com esse mesmo cruzamento de BARC-12 X CAC-1, analisou a variância dos progenitores quanto ao teor de linolênico em relação à distribuição da população F₂. Ele relatou que o controle do teor de ácido linolênico é feito por dois genes maiores de efeito aditivo e também por genes de efeito menor. Os resultados com marcadores AFLP obtidos neste trabalho estão, de certa forma, em concordância com os do referido autor, pois indicam a existência de um gene maior de efeito aditivo identificado pelo marcador HAGCCTG116, que explica 49,7% da variação do teor de linolênico. Foram também detectados genes, com menor efeito, identificados nos grupos de ligação 16 e 20 (Figura 4 D). O presente trabalho identificou QTLs para o teor de linolênico em grupos de ligação com poucos marcadores (Figura 4 D), o que dificultou determinar, com precisão, quantos locos estavam envolvidos nesta característica, visto que esses pequenos grupos de ligação podem-se fundir entre si ou em grupos maiores, se mais marcadores forem adicionados a esta população. Apesar de terem sido identificados QTLs para todos os ácidos graxos na referida população, um maior número de marcadores seria necessário para gerar um mapa de ligação mais saturado e possibilitar uma análise mais precisa da herança dos teores dos ácidos graxos. Linhas recombinantes de autofecundação (RILs) estão sendo desenvolvidas para esse cruzamento, o que vai permitir avaliar a consistência dos QTLs encontrados em diferentes ambientes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Darvasi, A. e Soller, M. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. **Theor. Appl. Genet.**, **85**: 353-359, 1992.
- Delannay, X., RODGERS D.M. e PALMER, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Sci.**, **23**: 944-949, 1983.
- Doerge, R.W. e Churchill, G.A. Permutation tests for multiple loci affecting a quantitative character. **Genetics**, **142**: 285-294, 1996.
- Don, R.H., Cox P.T., Wainwright B.J. e Mattick J.S. "Touchdown" PCR to circumvent spurious priming during gene amplification. **Nuc. Acid. Res.**, **19**:4008, 1991.
- Dudley, J.W. Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. **Crop Sci.**, **33**: 660-668, 1993.
- Edwards, M.D., Stuber, C.W. e Wendel, J.F. Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I: numbers, genomic distribution, and types of gene action. **Genetics**, **116**: 113-125, 1987.
- Embrapa. Centro Nacional de Pesquisa de Soja (Londrina, PR). **Recomendações técnicas para a cultura da soja na região central do Brasil 1998/99**. Londrina, 1998. 164p. (Embrapa-Soja. Documentos, 98).
- Gesteira, A.S. **Estudo da herança e de marcadores moleculares para teor de ácido linolênico em soja**. Viçosa, MG, 1997. 50p. (Tese MS)
- Hnetkovsky, N., Chang, S.J.C., Doubler, T.W., Gibson, P.T. e Lightfoot D.A. Genetic mapping of loci underlying field resistance to soybean sudden death syndrome (SDS). **Crop Sci.**, **36**: 393-400, 1996.
- Keim, P., Schupp, J.M., Travis, S.E., Clayton, K., Ferreira, A. e Webb, D.M. A high-density soybean genetic map based upon AFLP markers. **Crop Sci.**, **37**:537-543, 1997.
- Lander E.S., Green P., Abrahamson J., Barlow A., Daly M. J., Lincoln S.E. e Newburg L. MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, **1** :174-181, 1987.
- Lander, E.S. e Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, **121**:185-199, 1989.
- Lee, S.H., Bailey, M.A., Mian, M.A.R., Shipe, E.R., Ashley, D.A., Parrott, W.A., Hussey, R. S. e Boerma, H. R. Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. **Theor. Appl. Genet.**, **92**: 516-523, 1996.

- Mansur, L.M., Lark, K.G., Kross, H. e Oliveira, A.B. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). **Theor. Appl. Genet.**, **86**: 907-913, 1993.
- Mansur, L.M., Orf, J.H., Chase, K. Jarvik, T., Cregan, P.B. e Lark, K.G. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. **Crop Sci.**, **36**: 1327-1336, 1996.
- Moraes, R.M.A. **Eficiência de seleção para teores dos ácidos oléico e linolênico em soja cultivada em duas temperaturas.** Viçosa, MG, 1999. 36p. (Tese MS).
- Stam, P. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. **Plant J.**, **3**: 739-744, 1993.
- TANKSLEY, S.D., YOUNG N.D., PATERSON, A.H., BONIERBALE, M.W. RFLP mapping in plant breeding: new tools for an old science. **Bio/Technology**, v.7, n.1, p.257-265, 1989.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. e Zabeau, M. AFLP: a new technique for DNA fingerprinting. **Nucl. Acids Res.**, **23**: 4407-4414, 1995.
- Xiao, J., Li, J., Yuan, L. e Tanksley, S.D. Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a specific rice cross. **Theor. Appl. Genet.**, **92**: 230-244, 1996.
- Palavras-Chave: diversidade genética, parentesco, AFLP e soja.

DIVERSIDADE GENÉTICA EM ACESSOS DE *Glycine soja* E *Glycine max* UTILIZANDO MARCADORES AFLP

RESUMO

A soja cultivada apresenta base genética estreita, devido ao fato de que poucos ancestrais contribuíram para o germoplasma utilizado nos programas de melhoramento. No entanto, o germoplasma da soja é muito extenso e ainda não explorado. O objetivo deste trabalho foi avaliar a diversidade genética presente em linhagens não-adaptadas e em parentes silvestres da soja. Foram analisados 263 acessos de soja, sendo 31 *Glycine soja*, 10 *Glycine max* adaptados e 222 *Glycine max* não-adaptados. Foram utilizados três pares de *primers*, que geraram 202 marcadores, demonstrando a grande resolução da técnica de AFLP em estudos genéticos. Somente dois marcadores foram monomórficos em todos os acessos, devido à grande variabilidade genética presente nos acessos de *G. soja*. Os padrões de AFLP foram utilizados para gerar uma matriz de distâncias genéticas de Nei e Li, que foi empregada para realizar uma análise multivariada. Os resultados indicaram que a soja cultivada apresenta uma genética muito estreita e que existe uma grande diversidade genética em materiais exóticos, que podem ser aproveitados em programas de melhoramento.

Palavras-Chave: diversidade genética, parentes silvestres, AFLP e soja.

GENETIC DIVERSITY OF WILD AND CULTIVATED SOYBEAN USING AFLP TECHNOLOGY

ABSTRACT

Cultivated soybean has a narrow genetic base because few ancestors contributed for most of the breeding programs germplasm. Nevertheless, soybean has an extensive germplasm not yet explored (Delannay et al., 1983). In order to measure the genetic diversity present in non adapted and wild relatives, 263 soybean accessions (31 *Glycine soja*, 10 adapted and 222 unadapted *Glycine max*) were evaluated using AFLP technique. Three primers pairs generated 202 polymorphic markers showing a great discriminatory power of AFLP in genetic investigations. Only two markers were monomorphic among all accessions due to the great genetic diversity of the *G. soja* lines. AFLP patterns were used to generate a Nei & Li genetic distance matrix what was used to perform a multi-dimensional analysis. The results showed that cultivated soybean has a narrow genetic base and a great amount of genetic diversity is still present in exotic germplasm wich could be used as a source of variability in breeding programs.

Key Words: genetic diversity, wild relatives, AFLP, DNA and soybean.

INTRODUÇÃO

A soja atualmente cultivada possui base genética estreita devido ao pequeno número de ancestrais utilizados na formação de seu germoplasma-núcleo. Nos Estados Unidos da América, 10 ancestrais contribuíram com 80% do germoplasma do norte e sete ancestrais com 70% do germoplasma do sul (Delannay *et al.*, 1983). Gizlice *et al.* (1993), utilizando coeficientes de parentesco, concluíram que 25% da base genética da soja foi erodida dos conjuntos gênicos do norte e sul dos Estados Unidos. No Brasil, situação semelhante foi observada: nove linhagens contribuíram com mais de 80% do conjunto gênico (Hiromoto e Vello, 1986).

Apesar da pequena base genética da soja atualmente cultivada, existe ainda grande variabilidade genética nas populações silvestres. O conhecimento da estrutura genética do germoplasma auxilia na escolha de linhagens que possam ser utilizadas nos programas de melhoramento desta cultura. Marcadores moleculares se adequam perfeitamente ao estudo da variabilidade genética existente no germoplasma e auxiliam na seleção de genótipos contendo as características desejadas dos programas de melhoramento da soja. Abdelnoor *et al.* (1995) avaliaram a diversidade genética em 38 variedades brasileiras de soja, utilizando marcadores RAPD, com o objetivo de auxiliar os melhoristas na escolha de progenitores em programas de melhoramento. Powell *et al.* (1996) empregaram uma combinação de marcadores SSR nucleares e de cloroplasto para investigar os níveis e padrões de variabilidade detectados em genótipos de *Glycine max* e *Glycine soja*. Apesar de os SSRs serem, por natureza, multialélicos, os genótipos se agruparam de acordo com a espécie. Maughan *et al.* (1996) investigaram 23 acessos de soja cultivada e silvestre pela técnica de AFLP. A diversidade genética determinada por AFLP foi maior nos acessos silvestres do que na soja cultivada. A análise de agrupamento claramente separou os acessos de *G. soja* dos acessos de *Glycine max*. Dentro do grupo de *G. max*, as variedades de soja cultivada se agruparam mais densamente, demonstrando a pequena diversidade genética existente na soja cultivada.

O objetivo deste trabalho foi avaliar a diversidade genética presente em 263 acessos de soja, sendo 31 *G. soja*, 10 *G. max* adaptados e 222 *G. max* não-adaptados, utilizando-se a técnica de AFLP (Vos et al., 1995).

Foi utilizada um total de 263 acessos de soja, envolvendo *G. max* adaptados, *G. max* não-adaptados e *G. soja* (Tabela 1). Foram analisados 10 *G. max* adaptados, que eram linhagens ou variedades comerciais, representando a soja atualmente cultivada; 222 *G. max* (introduções) de *G. max*, para retratar a variabilidade existente dentro da espécie da soja cultivada; e 31 *G. soja*, para representar a diversidade genética de uma espécie silvestre próxima.

Isolamento e quantificação de DNA

Foram germinadas 30 sementes de cada acesso de soja em papel germinador. As sementes de *G. soja* foram escarificadas para permitir a hidratação. Cada papel germinador, correspondendo a cada acesso, foi enrolado e colocado verticalmente dentro de um béquer, com uma lâmina de água de 3 cm. Cada béquer, contendo 10 raízes, foi coberto com saco plástico para evitar a evaporação de água. Após 10 dias, um kitólo não expandido de 200 mL foi transferido para um microtubo de 1,5 mL, previamente refrigerado em gelo seco e triturado com um pestle de madeira. O kitólo triturado foi homogeneizado com um mixer, na presença de 200 µL de tampão de extração contendo de 10-HCl (pH 8,0) 125 mM; EDTA 45 mM; NaCl 0,83 M; CTAB 0,83 %; ureia 2,0 M; mercapto-*o*-etilanol 0,83 % e β-mercapto-*o*-etilanol 0,16 %. Outros 500 µL de tampão de extração foram adicionados, sendo a mistura incubada por 1 hora em 60 °C. As proteínas foram removidas com um volume de cloroformio - dicloro-metano (1:1). O DNA foi precipitado com um volume de etanol gelado, lavado com etanol 70%, secado a vácuo, e ressuspenso em tampão TE (10 mM Tris-HCl, pH 8,0 e 0,1 mM EDTA). A concentração das amostras de DNA foi determinada por meio da comparação com padrões de DNA de fogo-linha de um DNA-seqüenciador.

MATERIAL E MÉTODOS

Material genético

Foi utilizado um total de 263 acessos de soja, envolvendo *G. max* adaptados, *G. max* não-adaptados e *G. soja* (Tabela 1). Foram analisados 10 *G. max* adaptados, que eram linhagens ou variedades comerciais, representando a soja atualmente cultivada; 222 PIs (introduções) de *G. max*, para retratar a variabilidade existente dentro da espécie da soja cultivada; e 31 *G. soja*, para representar a diversidade genética de uma espécie silvestre próxima.

Isolamento e quantificação de DNA

Foram germinadas 10 sementes de cada acesso de soja em papel germitest. As sementes de *G. soja* foram escarificadas para permitir hidratação. Cada papel germitest, correspondendo a cada acesso, foi enrolado e colocado verticalmente dentro de um béquer, com uma lâmina de água de 3 cm. Cada béquer, contendo 10 rolos, foi coberto com saco plástico para evitar evaporação de água. Após 10 dias, um trifolíolo, não expandido de cada acesso, foi transferido para um microtubo de 1,5 mL, previamente resfriado em gelo seco e triturado com um palito de madeira. O trifolíolo triturado foi homogenizado com um mixer, na presença de 200 μ L de tampão de extração composto de Tris-HCl (pH 8,0) 125 mM; EDTA 45 mM; NaCl 0,83 M; CTAB 0,83 %; sorbitol 0,15 M; sarcosil 0,83 %; e bissulfito de sódio 0,16 %. Outros 500 μ L de tampão de extração foram adicionados, sendo a mistura incubada por 1 h a 65°C. As proteínas foram removidas com um volume de clorofórmio - álcool isoamílico (24:1). O DNA foi precipitado com um volume de isopropanol gelado, lavado com etanol 70%, secado a vácuo, e ressuspendido em tampão TE_{0,1}: Tris-HCl (pH 8,0) 10mM e EDTA 0,1 mM. A concentração das amostras de DNA foi estimada por meio da comparação com padrões de DNA de fago lâmbda em géis de agarose.

Tabela 1 – Lista dos 263 acessos de soja utilizados na análise da diversidade genética e suas classificações quanto à espécie e à sua adaptação. (s)=*Glycine soja*, (mn)=*Glycine max* não-adaptado e (ma)=*Glycine max* adaptado

Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma
PI 65549	x			PI 5461	x			PI 123440	x			PI 171454			x
PI 135624	x			PI 54854	x			PI 132201	x			PI 171652			x
PI 163453	x			PI 64698	x			PI 132	x			PI 174854			x
PI 245331	x			PI 65342	x			PI 133226	x			PI 174855			x
PI 339735B	x			PI 68679	x			PI 153263	x			PI 174859			x
PI 339871B	x			PI 79739	x			PI 153271	x			PI 174861			x
PI 342618A	x			PI 79760	x			PI 153280	x			PI 174867			x
PI 342621C	x			PI 80476	x			PI 153285	x			PI 175174			x
PI 366121	x			PI 81029	x			PI 153292	x			PI 175177			x
PI 378692	x			PI 81041	x			PI 153	x			PI 175178			x
PI 407020	x			PI 82264	x			PI 154193	x			PI 175179			x
PI 407042	x			PI 82291	x			PI 157404	x			PI 175185			x
PI 407160	x			PI 82312	x			PI 157421	x			PI 175187			x
PI 407	x			PI 82326	x			PI 157435	x			PI 175192			x
PI 407197	x			PI 84611	x			PI 157440	x			PI 175196			x
PI 407203	x			PI 84751	x			PI 157462	x			PI 175199			x
PI 407211	x			PI 85505	x			PI 157	x			PI 179935			x
PI 407225	x			PI 86006	x			PI 157492	x			PI 189885			x
PI 407240	x			PI 86023	x			PI 159764	x			PI 189946			x
PI 407246	x			PI 86091-1	x			PI 159926	x			PI 192871			x
PI 407	x			PI 86103	x			PI 161431A	x			PI 192873			x
PI 407282	x			PI 86490	x			PI 165563	x			PI 196164			x
PI 423988	x			PI 87059	x			PI 165583	x			PI 196168			x
PI 423992	x			PI 87631-1	x			PI 165914	x			PI 200508			x
PI 424003	x			PI 88289	x			PI 165926	x			PI 204337			x
PI 424004A	x			PI 88788	x			PI 165947	x			PI 205085			x
PI 424033	x			PI 89008	x			PI 165989	x			PI 208429			x
PI 464927A	x			PI 89772	x			PI 166028	x			PI 208433			x
PI 468916	x			PI 90763	x			PI 166032	x			PI 208438			x
FC 03981	x			PI 91102	x			PI 166048	x			PI 208439			x
FC 31715	x			PI 91732-1	x			PI 167277	x			PI 209331			x
PI 32176		x		PI 103080	x			PI 171429	x			PI 209332			x
PI 54608-1		x		PI 103091	x			PI 171438	x			PI 209578			x

Tabela 1, Cont.

Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma	Nome do Acesso	s	mn	ma
PI 209833		x		PI 291325		x		PI 374186		x		PI 416868A			x
PI 212604		x		PI 291331		x		PI 377574		x		PI 417020			x
PI 212605		x		PI 295947		x		PI 378663		x		PI 417091			x
PI 212716		x		PI 295948		x		PI 378670A		x		PI 417138			x
PI 215811		x		PI 303652		x		PI 379559A		x		PI 417140			x
PI 219698		x		PI 307836		x		PI 379559B		x		PI 417290			x
PI 219732		x		PI 307881		x		PI 379559C		x		PI 417428			x
PI 227327		x		PI 309658		x		PI 379561		x		PI 423809			x
PI 238109		x		PI 319525		x		PI 381682		x		PI 423833B			x
PI 243516		x		PI 323276		x		PI 391577		x		PI 424024			x
PI 243538		x		PI 323550		x		PI 398643		x		PI 424078			x
PI 243550		x		PI 323551		x		PI 398652		x		PI 424182B			x
PI 247679		x		PI 323555		x		PI 398720		x		PI 424326			x
PI 248397		x		PI 323557		x		PI 399077		x		PI 424503			x
PI 248400		x		PI 323561		x		PI 401418		x		PI 427241			x
PI 248409		x		PI 323564		x		PI 404151		x		PI 428691			x
PI 253656A		x		PI 323569		x		PI 404156		x		PI 428692			x
PI 274453		x		PI 323574		x		PI 404165		x		PI 430457			x
PI 274454		x		PI 323575		x		PI 404168		x		PI 458072A			x
PI 274456		x		PI 323579		x		PI 404179B		x		Peking			x
PI 279648		x		PI 339734		x		PI 404185		x		Prolina			x
PI 284873		x		PI 342438		x		PI 404189		x		H5050			x
PI 290116B		x		PI 346306		x		PI 404194		x		HS-1			x
PI 290120		x		PI 360844		x		PI 404195		x		HS-2			x
PI 290123A		x		PI 361080		x		PI 407653		x		Minsoy			x
PI 290136		x		PI 365426		x		PI 407658C		x		Archer			x
PI 290148		x		PI 371611		x		PI 407729		x		H5545RR			x
PI 291272		x		PI 374154		x		PI 407730		x		Williams			x
PI 291275		x		PI 374159		x		PI 407810		x		CX375			x
PI 291278		x		PI 374165		x		PI 407960A		x		CX445			x
PI 291309C		x		PI 374170		x		PI 408251		x		K111			x
PI 291312		x		PI 374176		x		PI 408260B		x		Resnick			x
PI 291318		x		PI 374181		x		PI 416762		x					

Reações de AFLP

Digestão do DNA: duzentos e cinqüenta nanogramas de DNA de cada amostra foram digeridos simultaneamente com cinco unidades de *EcoRI* e *MseI* em 40 μL de tampão contendo Tris-acetato 10 mM (pH 7,5), acetato de magnésio 10 mM, acetato de potássio 50 mM, ditioneitol 5 mM e BSA (albumina de soro bovino) 50 ng/ μL , incubados por 3 h a 37°C.

Ligação: a ligação dos adaptadores foi feita com a adição de 10 μL de uma solução contendo 5 pmoles de adaptador de *EcoRI*, 50 pmoles de adaptador de *MseI*, uma unidade de T4 DNA-ligase, ATP 1 mM, Tris-acetato 10 mM (pH 7,5), acetato de magnésio 10 mM, acetato de potássio 50 mM, ditioneitol 5 mM e BSA 50 ng/ μL , incubados por 3 h a 37°C. Após a ligação, a mistura de reação foi diluída a 1:10 com TE_{0,1}.

Amplificação pré-seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de um nucleotídeo seletivo, reduzindo-se em 16 vezes a complexidade dos fragmentos. Cinco microlitros da mistura diluída da restrição/ligação de cada amostra foram pré-amplificados em um volume final de 50 μL , contendo 50 ng dos *primers* de *EcoRI* e *MseI*, acrescidos de um nucleotídeo seletivo; dNTPs 200 μM ; uma unidade de *Taq* polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl 50 mM. As reações de amplificação do DNA foram aplicadas no termociclador Perkin-Elmer 9600 (Norwalk, CT, EUA). Os ciclos foram os seguintes: 20 ciclos de 94°C por 30 segundos, 55°C por 30 segundos e 72°C por 60 segundos. Após a amplificação pré-seletiva, os produtos das reações foram diluídos a 1:20, com TE_{0,1}.

Amplificação seletiva: foi realizada com um par de *primers* complementares à seqüência dos adaptadores acrescidos de três nucleotídeos seletivos, reduzindo-se em 256 vezes a complexidade dos fragmentos. Para amplificação seletiva, o *primer* seletivo de *EcoRI* foi marcado na extremidade 5', utilizando-se [γ -³³P] ATP. A reação de marcação foi composta de 50 μL , contendo 500 ng de *primer* de *EcoRI* com três nucleotídeos seletivos; [γ -³³P] ATP 100 μCi ; 10 unidades de T4 polinucleotídeo cinase; Tris-HCl 25 mM (pH 7,5); MgCl₂ 10 mg; ditioneitol 5 mM; e espermidina-3HCl 0,5 mM. Cinco microlitros da mistura de amplificação pré-seletiva diluídos

a 1:20 foram amplificados num volume total de 20 μ L, contendo 5 ng do *primer* de *EcoRI* marcado com três nucleotídeos seletivos; 30 ng de *primer* de *MseI* com três nucleotídeos seletivos; dNTPs 0,2 mM; 0,4 unidades de *Taq* polimerase; Tris-HCl 10 mM (pH 8,0); MgCl₂ 1,5 mM; e KCl 50 mM. Os ciclos de amplificação seletiva foram os seguintes: 12 ciclos de 94°C por 30 segundos, 65°C (-0,7°C/ciclo) por 30 segundos e 72°C por 60 segundos, até atingir a temperatura ótima de ligação dos *primers* de 56°C - Touchdown Cycle Profile (Don et al., 1991).

Os 263 acessos de soja foram analisados pela técnica de AFLP, utilizando-se três pares de *primers* seletivos: *EcoRI/MseI* - AAC/CGA, AAG/CTC e AAT/CTC.

Eletroforese em gel de poli-acrilamida

Os produtos da amplificação seletiva foram misturados com um volume igual (20 μ L) de tampão contendo formamida 98%, EDTA 10 mM, xileno cianol FF 1 mg/mL e azul-de-bromofenol 1 mg/mL. As amostras foram, então, desnaturadas a 95°C por cinco minutos e mantidas em gelo até serem aplicadas no gel. Um microlitro e meio de cada amostra foi aplicado em um gel desnaturante de poli-acrilamida 4,5% (acrilamida:bisacrilamida 19:1, uréia 7,5 M e Tris-borato EDTA (TBE) 0,5x), utilizando-se a cuba de eletroforese *SequiGen* de 38 x 50 cm (BioRad Laboratories Inc., Hercules, CA, EUA). Os produtos da amplificação dos 263 acessos foram carregados em três géis, designados GSA, GSB e GSC. Cada gel foi carregado com dois pentes, de 48 dentes cada, acrescentando-se uma amostra extra na lateral do pente. Em cada pente foi carregado um padrão 10 BL, que mostrava uma banda a cada 10 pb, entre 80 e 530 pb, na faixa 25 e um genótipo designado P1 na faixa 26. A placa de vidro externa não recebeu nenhum tratamento, e a placa de vidro interna foi coberta com 3 mL de Sigma cote para a não-aderência do gel. A eletroforese foi realizada em TBE 1x, 130 W e 50°C por 1,5 h. Após a eletroforese, as placas foram separadas e o gel permaneceu ligado à placa de vidro externa, que não recebeu nenhum tratamento. O gel foi fixado em papel 3 MM, envolto por um filme de plástico, secado a vácuo por 1 h a 70°C

e exposto a um *phosphor screen* por 20 h. Os *screens* foram digitalizados, usando-se o *Storm System* (Molecular Dynamics, CA, EUA).

Scoring dos géis e análise dos dados

As imagens geradas pelo *storm system* foram convertidas para o programa da Keygene (Wageningen, Holanda) para serem analisadas. Cada imagem é convertida em duas, gerando-se o gel da esquerda e o gel da direita. Portanto, seis imagens para cada par de *primer* foram geradas: GSA (L e R), GSB (L e R) e GSC (L e R). Uma vez convertida a imagem, o primeiro procedimento foi a definição das faixas, seguindo-se as curvas do gel, para que o programa identificasse cada amostra. Cada faixa do gel foi nomeada de acordo com a amostra carregada. A próxima etapa foi a definição das bandas comuns para definir a forma do gel. Uma vez definida a forma do gel, foi feito o *scoring* das bandas polimórficas, sendo (+) para a presença da banda, (-) para a ausência da banda e (?) para falhas ou dúvida quanto ao *scoring*. A primeira imagem ao ser feito o *scoring* foi a GSA-L. A imagem GSA-R foi, então, alinhada com a imagem GSA-L, por meio do padrão 10 BL carregado na faixa 25 e do P1 na faixa 26. Com o alinhamento das imagens, foi possível copiar os marcadores da imagem GSA-L para a imagem GSA-R, para que o mesmo loco pudesse ser analisado com as amostras do gel GSA-R. Seguindo-se o mesmo procedimento, a imagem GSB-L foi alinhada com a imagem GSA-L para copiar os marcadores oriundos desta. Tal procedimento foi realizado até a última imagem GSC-R, e dessa forma todas as 263 amostras foram analisadas para os mesmos locos. A presença dos acessos de *G. soja* na imagem GSA-L minimizou bastante o aparecimento de novos marcadores nas outras imagens, evitando-se que eles tivessem que ser copiados de volta para todas as outras imagens.

Uma vez feito o *scoring* dos géis, os dados foram exportados pelo programa Keygene, gerando-se uma matriz +/-, que foi convertida numa matriz 0/1. O programa NTSYS 1.8 (Rohlf, 1993) foi utilizado para gerar uma matriz de distâncias genéticas, segundo a definição de similaridade $S_{ij} = 2a / (2a + b + c)$ proposta por Nei e Li (1979). S_{ij} é a similaridade entre dois indivíduos i e j , sendo a o número de bandas presentes em i e j ; b , o número de

bandas presentes em i e ausentes em j ; e c , o número de bandas presentes em j e ausentes em i . Essa definição de similaridade não considera as bandas ausentes em ambos os indivíduos, visto que ausência mútua não pode ser atribuída a uma mesma causa. O peso 2 para presença da banda em ambos os indivíduos permite uma melhor diferenciação entre indivíduos com um pequeno índice de similaridade. Foi gerada uma matriz de dissimilaridade de acordo com a fórmula $D_{ij}=1-S_{ij}$, que foi utilizada com o procedimento CORRESP do programa SAS, para visualizar a dispersão dos indivíduos em relação a três principais eixos de variação (SAS Institute, 1989).

As três primers relativas à β -AFLP, que geraram 202 bandas polimórficas, representando uma média de 67 bandas polimórficas para cada primer. Na Figura 1, mostra-se o padrão de amplificação obtido com o par de primers ACC/CAC, que produziu 72 bandas polimórficas. Para analisar as 263 acessos, foi necessário coregar três géis, gerando-se seis imagens a serem analisadas para cada par de primers. As 31 G. soja foram incluídas na primeira metade do primeiro gel (GSA-U), já prevendo a alta variabilidade genética nesses acessos. Essa grande variabilidade foi comprovada pela detecção de um alto nível de polimorfismo entre os acessos de G. soja. A maioria das bandas polimórficas utilizadas na avaliação dos 263 acessos foi gerada nos acessos de G. soja. Uma vez feito o scoring da primeira imagem (GSA-U), as bandas obtidas foram copiadas para as demais imagens, para que todas as 263 acessos fossem avaliadas.

Foi gerada uma matriz de $d \times j$ (dados não mostrados) com 263 colunas, representando o número de acessos e 202 linhas, representando o número de bandas polimórficas obtidas. Essa matriz foi utilizada para obtenção dos índices de dissimilaridade de Nei e Li entre os 263 acessos avaliados. A definição de similaridade de Nei e Li foi escolhida, devido ao alto nível de polimorfismo obtido com a técnica de AFLP. A fórmula de Nei e Li supervaloriza a presença da banda para ambos os indivíduos e desconsidera a ausência mútua de bandas. Deste modo, esta função permite uma melhor diferenciação entre indivíduos com pequeno índice de similaridade. No Tabela 2, mostram-se as médias e os desvios da média dos índices de dissimilaridade de Nei e Li para os 263 acessos avaliados. Os acessos de Glycine soja apresentaram o maior valor médio de dissimilaridade (0,58).

RESULTADOS E DISCUSSÃO

A técnica do AFLP foi utilizada para amplificar o DNA de 263 acessos de soja, sendo 10 *G. max* adaptados, que são linhagens ou variedades comerciais, representando a soja atualmente cultivada; 222 PIs (introduções) de *G. max*, para retratar a variabilidade dentro da espécie da soja cultivada; e 31 *G. soja*, para avaliar a diversidade genética de uma espécie silvestre próxima. A técnica de AFLP demonstrou-se bastante poderosa na detecção da alta variabilidade entre os acessos. Foram utilizados apenas três *primers* seletivos (+3) de AFLP, que geraram 202 bandas polimórficas, representando uma média de 67 bandas polimórficas para cada *primer*. Na Figura 1, mostra-se o padrão de amplificação obtido com o par de *primers* ACC/CAC, que produziu 72 bandas polimórficas. Para analisar os 263 acessos, foi necessário carregar três géis, gerando-se seis imagens a serem analisadas para cada par de *primers*. As 31 *G. soja* foram incluídas na primeira metade do primeiro gel (GSA-L), já prevendo a alta variabilidade genética nesses acessos. Essa grande variabilidade foi comprovada pela detecção de um alto nível de polimorfismo entre os acessos de *G. soja*. A maioria das bandas polimórficas utilizadas na avaliação dos 263 acessos foi originada dos acessos de *G. soja*. Uma vez feito o *scoring* da primeira imagem (GSA-L), as bandas obtidas foram copiadas para as demais imagens, para que todos os 263 acessos fossem avaliados.

Foi gerada uma matriz de 0 e 1 (dado não mostrado) com 263 colunas, representando o número de acessos e 202 linhas, representando o número de bandas polimórficas obtidas. Essa matriz foi utilizada para obtenção dos índices de dissimilaridade de Nei e Li entre os 263 acessos avaliados. A definição de similaridade de Nei e Li foi escolhida, devido ao alto nível de polimorfismo obtido com a técnica de AFLP. A fórmula de Nei e Li supervaloriza a presença da banda para ambos os indivíduos e desconsidera a ausência mútua de bandas. Desse modo, essa função permite uma melhor diferenciação entre indivíduos com pequeno índice de similaridade. Na Tabela 2, mostram-se as médias e os desvios da média dos índices de dissimilaridade de Nei e Li, para os 263 acessos avaliados. Os acessos de *Glycine soja* apresentaram o maior valor médio de dissimilaridade (0,58),

F-492.572<NS>[1]
 F-464.572<NS>[1]
 F-455.411<NS>[1]
 F-444.850<NS>[1]
 F-433.307<NS>[1]
 F-418.856<NS>[1]
 F-404.697<NS>[1]
 F-393.184<NS>[1]
 F-388.180<NS>[1]
 F-380.970<NS>[1]
 F-374.766<NS>[1]
 F-367.961<NS>[1]
 F-361.838<NS>[1]
 F-350.088<NS>[1]
 F-348.247<NS>[1]
 F-334.424<NS>[1]
 F-328.247<NS>[1]
 F-327.796<NS>[1]
 F-318.391<NS>[1]
 F-311.028<NS>[1]
 F-306.997<NS>[1]
 F-304.499<NS>[1]
 F-302.462<NS>[1]
 F-298.738<NS>[1]
 F-294.234<NS>[1]
 F-287.734<NS>[1]
 F-282.861<NS>[1]
 F-277.842<NS>[1]
 F-272.588<NS>[1]
 F-267.160<NS>[1]
 F-261.988<NS>[1]
 F-248.659<NS>[1]
 F-253.860<NS>[1]
 F-248.538<NS>[1]
 F-240.015<NS>[1]
 F-231.967<NS>[1]
 F-230.441<NS>[1]
 F-227.773<NS>[1]
 F-222.064<NS>[1]
 F-218.188<NS>[1]
 F-214.338<NS>[1]
 F-210.738<NS>[1]
 F-206.346<NS>[1]
 F-199.127<NS>[1]
 F-196.470<NS>[1]
 F-193.704<NS>[1]
 F-189.131<NS>[1]
 F-185.191<NS>[1]
 F-179.827<NS>[1]
 F-177.737<NS>[1]
 F-172.001<NS>[1]
 F-165.729<NS>[1]
 F-159.738<NS>[1]
 F-157.997<NS>[1]
 F-154.976<NS>[1]
 F-153.600<NS>[1]
 F-148.217<NS>[1]
 F-147.217<NS>[1]
 F-144.898<NS>[1]
 F-141.899<NS>[1]
 F-137.386<NS>[1]
 F-134.348<NS>[1]
 F-132.509<NS>[1]
 F-131.710<NS>[1]
 F-128.622<NS>[1]
 F-125.736<NS>[1]
 F-121.961<NS>[1]
 F-116.611<NS>[1]
 F-114.098<NS>[1]
 F-108.060<NS>[1]
 F-105.316<NS>[1]
 F-97.922<NS>[1]
 F-87.921<NS>[1]

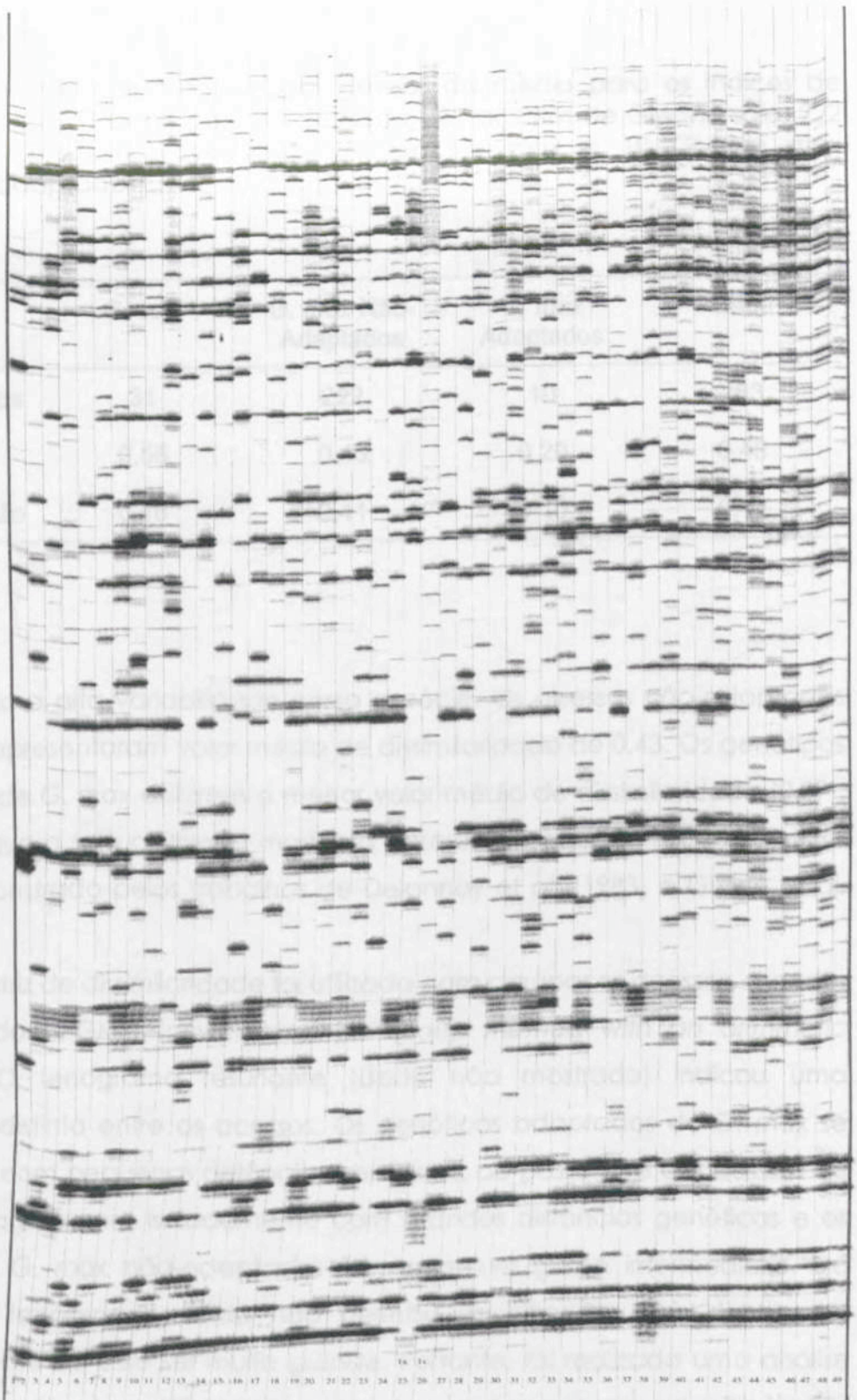


Figura 1 – Padrão de amplificação de DNA obtido pela técnica de AFLP, mostrando-se 48 dos 263 acessos de soja analisados. O padrão 10 BL, que mostra uma banda a cada 10 pb entre 80 e 530 pb, e P1 foram carregados nas faixas 25 e 26, respectivamente. Esse par de primers, ACC/CAC, gerou 73 bandas polimórficas entre os acessos, e o scoring foi feito, considerando-se a presença da banda = + e a ausência da banda = -.

Tabela 2 – Valores de média e dos desvios da média para os índices de dissimilaridade de Nei e Li para os 31 acessos de *Glycine soja*, 222 acessos não-adaptados de *Glycine max* e 10 *Glycine max* adaptados

	<i>G. soja</i>	<i>G. max</i> Não-Adaptados	<i>G. max</i> Adaptados	Geral
Nº de acessos	31	222	10	263
Média	0,58	0,43	0,29	0,45
Desvio-padrão	0,16	0,11	0,10	0,12

demonstrando a alta variabilidade nessa espécie. Os acessos não-adaptados de *G. max* apresentaram valor médio de dissimilaridade de 0,43. Os genótipos adaptados de *G. max* exibiram o menor valor médio de dissimilaridade (0,29), indicando que a soja cultivada moderna apresenta pequena base genética, como demonstrado pelos trabalhos de Delannay *et al.* (1983) e Gizlice *et al.* (1993).

A matriz de dissimilaridade foi utilizada para agrupar os acessos, usando-se o método UPGMA (*unweighted pair-group method with an arithmetic average*). O fenograma resultante (dado não mostrado) indicou uma separação distinta entre os acessos. Os genótipos adaptados de *G. max* se agruparam com pequenas distâncias genéticas, ao passo que os acessos de *G. soja* se agruparam isoladamente com grandes distâncias genéticas e os acessos de *G. max* não-adaptados formaram um grupo intermediário. No entanto, o fenograma obtido não permitiu visualização mais detalhada, devido ao fato de esse ser muito grande. Portanto, foi realizada uma análise de correspondência múltipla, que organiza os acessos em relação a três principais eixos de variação. Na Figura 2, mostra-se a dispersão tridimensional da variabilidade encontrada nos acessos de soja avaliados. Os genótipos adaptados de *G. max* formaram uma pequena aglomeração, estando cercados por uma camada de *G. max* não-adaptados, que por sua vez estão circundados bem externamente pelos acessos de *Glycine soja*. A dispersão

soja cultivada não adaptados, principalmente por serem de G. max. Isso indica que este germoplasma exótico pode fornecer variabilidade genética para os programas de melhoramento de soja. Os resultados encontrados neste trabalho estão de acordo com o trabalho de Maughan et al. (1996), que demonstraram a pequena base genética da soja cultivada em relação a acessos não-adaptados de G. max e G. soja.

Fazendo em vista a grande erosão genética resultante de vários anos de melhoramento e objetivando utilizar a grande variabilidade em materiais exóticos, Tanky et al. Nelson (1994) produziram uma mel de soja denominada análise de GTUs em relação ao melhoramento genético para introdução de transferir GTUs de variedades germoplasmas exóticos para melhoramento. Tanky et al. Nelson (1994) produziram uma mel de soja denominada análise de GTUs em relação ao melhoramento genético para introdução de variedades oriundas de G. soja e G. max. Essa nova mel de soja foi desenvolvida para proporcionar aumento de produtividade e resistência a pragas e doenças comerciais. Essa nova mel de soja foi desenvolvida para proporcionar aumento de produtividade e resistência a pragas e doenças comerciais. Essa nova mel de soja foi desenvolvida para proporcionar aumento de produtividade e resistência a pragas e doenças comerciais.

O gráfico tridimensional derivado da análise de correspondência múltipla dos 263 acessos de soja avaliados pela técnica de AFLP. Os resultados obtidos neste trabalho mostram a pequena base genética da soja cultivada em relação a acessos não-adaptados de G. max e G. soja. Isso indica que este germoplasma exótico pode fornecer variabilidade genética para os programas de melhoramento de soja. Os resultados encontrados neste trabalho estão de acordo com o trabalho de Maughan et al. (1996), que demonstraram a pequena base genética da soja cultivada em relação a acessos não-adaptados de G. max e G. soja.

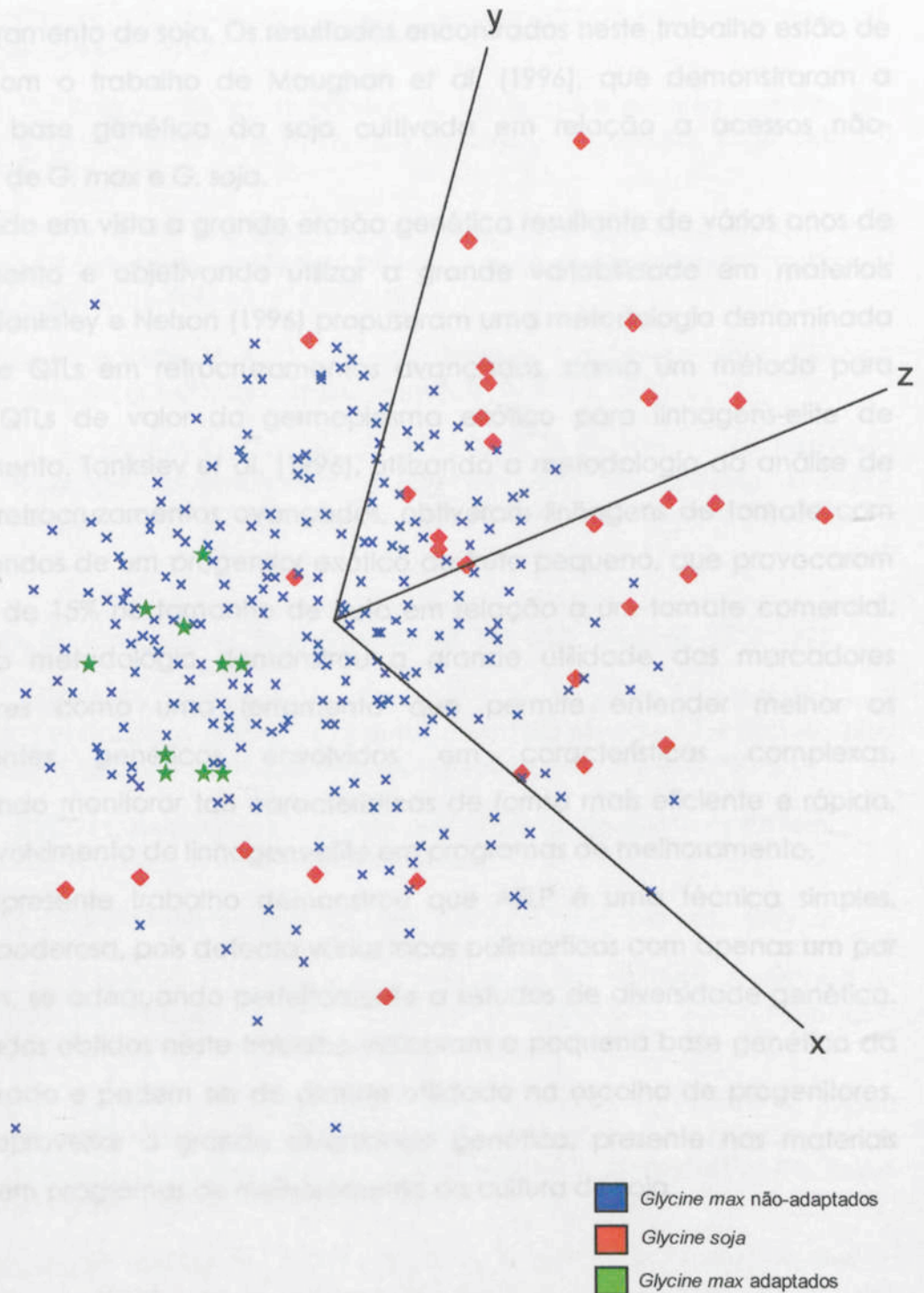


Figura 2 – Gráfico tridimensional derivado da análise de correspondência múltipla dos 263 acessos de soja avaliados pela técnica de AFLP.

tridimensional mostra claramente a pequena base genética da soja cultivada atualmente em relação à grande variabilidade ainda presente em acessos não adaptados, principalmente nos acessos de *G. soja*. Isso indica que este germoplasma exótico pode fornecer variabilidade genética para os programas de melhoramento de soja. Os resultados encontrados neste trabalho estão de acordo com o trabalho de Maughan *et al.* (1996), que demonstraram a pequena base genética da soja cultivada em relação a acessos não-adaptados de *G. max* e *G. soja*.

Tendo em vista a grande erosão genética resultante de vários anos de melhoramento e objetivando utilizar a grande variabilidade em materiais exóticos, Tanksley e Nelson (1996) propuseram uma metodologia denominada análise de QTLs em retrocruzamentos avançados, como um método para transferir QTLs de valor do germoplasma exótico para linhagens-elite de melhoramento. Tanksley *et al.* (1996), utilizando a metodologia da análise de QTLs em retrocruzamentos avançados, obtiveram linhagens de tomate com alelos oriundos de um progenitor exótico de fruto pequeno, que provocaram aumento de 15% no tamanho de fruto em relação a um tomate comercial. Essa nova metodologia demonstrou a grande utilidade dos marcadores moleculares como uma ferramenta que permite entender melhor os componentes genéticos envolvidos em características complexas, possibilitando monitorar tais características de forma mais eficiente e rápida, no desenvolvimento de linhagens-elite em programas de melhoramento.

O presente trabalho demonstrou que AFLP é uma técnica simples, rápida e poderosa, pois detecta vários locos polimórficos com apenas um par de *primers*, se adequando perfeitamente a estudos de diversidade genética. Os resultados obtidos neste trabalho indicaram a pequena base genética da soja cultivada e podem ser de grande utilidade na escolha de progenitores, visando aproveitar a grande diversidade genética, presente nos materiais exóticos, em programas de melhoramento da cultura da soja.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abdelnoor, R. V., Barros, E. G. e Moreira, M. A. Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. **Rev. Brasil. Genet.**, **18** (2): 265-273, 1995.
- Delannay, X., Rodgers D. M. e Palmer, R. G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Sci.**, **23**: 944-949, 1983.
- Gizlice, Z., Carter JR., T. E. e Burton, J. W. Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. **Crop Sci.**, **33**: 614-620, 1993.
- Hiramoto, D. M. e Vello, N. A. The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. **Rev. Brasil. Genet.**, **9**: 295-306, 1986.
- Maughan, P. J., Saghai Maroof, M. A., Buss G. R. e Huestis, G. M. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic line analysis. **Theor. Appl. Genet.**, **93**: 392-401, 1996.
- Nei, M., e Li, W. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proc. Natl. Acad. Sci. USA** , **76**:5269-5273, 1979.
- Powell, W., Morgante, M., Doyle, J. J., Mcnicol, J. W., Tingey, S. V. e Rafalski, A. J. Genepool variation in genus *Glycine* sugenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. **Genetics**, **144**: 793-803, 1996.
- Rohlf, F.J. **NTSYS-PC numerical taxonomy and multivariate analyses system. Version 1.8.** Setauket, NY, EUA. Exeter Publ., 1993.
- SAS Institute. **SAS/STAT users guide.** 6.09 ed. Cary, NC, EUA SAS inst., 1989.
- Tanksley, S. D. e Nelson, J. C. 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, 1996.
- Tanksley, S. D., Grandillo, S., Fulton, T. M., Zamir, D., Eshed, Y., Petiard, V, Lopez, J. e Beck-Bunn, T. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. **Theor. Appl. Genet.**, **92**: 213-224, 1996.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. e Zabeau, M. AFLP: a new technique for DNA fingerprinting. **Nucl. Acids Res.**, **23**: 4407-4414, 1995.

2. RESUMO E CONCLUSÕES

A utilização de marcadores moleculares em programas de melhoramento de culturas, como a soja, auxilia no monitoramento de genes de interesse e possibilita avaliar a variabilidade genética nas várias etapas do programa de melhoramento. Marcadores AFLP são bastante polimórficos e apropriados para uso em culturas de base genética estreita como a soja.

O primeiro objetivo deste trabalho foi a integração do mapa de AFLP de soja construído pela Monsanto Company (St. Louis, MO, EUA), a partir de uma população de 92 indivíduos F_2 do cruzamento de HS-1 X PI-326581 (*Glycine soja*), com um mapa de domínio público desenvolvido pela Universidade de Utah (Salt lake City, UT, EUA). Esse mapa foi gerado com uma população de 223 RILs (linhas recombinantes de autofecundação) oriundas do cruzamento de duas linhas não-adaptadas de soja, Minsoy e Noir 1. O mapa possui 468 marcadores (279 RFLP, 166 SSR e 23 clássicos), que cobrem 2.330 cM e contêm QTLs (*loci* de características quantitativas) identificados para vários caracteres agronômicos de interesse. DNA de 88 RILs foi amplificado, utilizando-se a técnica de AFLP, com 41 pares de *primers*, utilizados também no mapa da Monsanto.

Foram gerados 354 marcadores de AFLP, sendo 54 polimórficos também para o mapa da Monsanto e utilizados como pontos de ancoramento para a integração dos dois mapas. Foi encontrada uma colinearidade entre os dois mapas, permitindo-se a transferência de informações do mapa de domínio público, Minsoy X Noir 1, para o mapa da Monsanto, HS-1 X PI-326581.

O segundo objetivo do presente trabalho foi a identificação de QTLs para teores de ácidos graxos em soja, utilizando-se marcadores AFLP. Sementes de soja de 382 indivíduos de uma população F₂ do cruzamento da linhagem de baixo teor de ácido linolênico, BARC-12, com a variedade CAC-1 foram analisadas para o teor dos ácidos graxos palmítico, esteárico, oléico, linolênico e linoléico por meio de cromatografia gasosa, usando-se um método não destrutivo.

Quinze sementes com maiores e menores teores de cada ácido graxo foram selecionadas e plantadas em casa de vegetação, sendo utilizados trifolíolos não expandidos para a extração de DNA. Um total de 89 plantas foi analisado pela técnica de AFLP. Trinta e cinco combinações de *primers* foram utilizadas, gerando-se 158 marcadores dominantes e 18 co-dominantes. Foram identificados QTLs para todos os ácidos graxos, ressaltando-se que um marcador co-dominante explicou 50% da variação do teor de ácido linolênico. Linhas recombinantes de autofecundação (RILs) estão sendo desenvolvidas para esse cruzamento, o que vai permitir realizar estudos mais detalhados sobre a herança do teor de ácidos graxos em sementes de soja.

O terceiro objetivo deste trabalho foi estudar a variabilidade genética em 263 acessos de soja, sendo 31 *Glycine soja*, 10 *Glycine max* adaptados e 222 *Glycine max* não-adaptados, utilizando-se a técnica de AFLP.

Foram utilizados três pares de *primers*, que geraram 202 marcadores, demonstrando o grande polimorfismo da técnica de AFLP para estudos genéticos. Somente dois marcadores foram monomórficos em todos os acessos, devido à grande variabilidade genética presente nos acessos de *G. soja*. Os padrões de AFLP foram utilizados para gerar uma matriz de distâncias genéticas de Nei e Li, que foi utilizada para realizar uma análise multivariada.

Os resultados indicaram que a soja cultivada apresenta base genética muito estreita e que existe grande diversidade genética em materiais exóticos, que podem ser aproveitados em programas de melhoramento.