

**AVALIAÇÃO DA DIVERSIDADE GENÉTICA ENTRE CULTIVARES  
BRASILEIRAS DE SOJA, ATRAVÉS DE MARCADORES AFLP**

**ANA LÍDIA VARIANI BONATO**  
Engenheira Agrônoma

Orientador: Prof. Dr. ISAIAS OLÍVIO GERALDI

Tese apresentada à Escola Superior de  
Agronomia "Luiz de Queiroz", Universidade  
de São Paulo, para obtenção do título de  
Doutor em Agronomia, Área de  
Concentração: Genética e Melhoramento  
de Plantas.

**PIRACICABA**  
Estado de São Paulo-Brasil  
Julho-2000

## ERRATA

<b>Página</b>	<b>Linha</b>	<b>Onde se lê</b>	<b>Leia-se</b>
13	23	<i>Glycine Max</i>	<i>Glycine max</i>
15	4	genótipos	genótipos
35	17	individuais por, com base...	individuais, com base...
46	2	coletado aproximadamente trinta folhas obtidas de plantas...	coletada uma folha por planta em aproximadamente trinta plantas...
36		Ignorar as equações de regressão em todas as figuras.	

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - Campus "Luiz de Queiroz"/USP**

Bonato, Ana Lídia Variani

Avaliação da diversidade genética entre cultivares brasileiras de soja, através de marcadores AFLP / Ana Lídia Variani Bonato. - - Piracicaba, 2000.

89 p. Acompanha CD-ROM

Tese (doutorado) - Escola Superior de Agricultura Luiz de Queiroz, 2000.

Bibliografia.

1. Diversidade genética 2. Marcador genético 3. Soja 4. Variabilidade genética 5. Variedade I. Título

CDD 633.34

**"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"**

**DEDICO**

***Aos meus pais, Emídio R. Bonato e Lídia V. Bonato***

***Compartilho o êxito deste trabalho com todos que junto comigo nele acreditaram...***

## AGRADECIMENTOS

À Deus.

À minha família, Emídio R. Bonato, Lídia V. Bonato e Cláudia V. Bonato pela contribuição, incentivo e apoio em todos os momentos.

Ao Prof. Dr. Isaias O. Geraldi pela orientação, ensinamentos e amizade.

Ao Dr. Eberson S. Calvo pela oportunidade de desenvolver este trabalho junto à sua equipe e pela inestimável contribuição.

À Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo e à Empresa Brasileira de Pesquisa Agropecuária - Soja pela oportunidade de realização do curso e dos experimentos relativos à tese, em suas instituições.

Aos professores do Curso de Pós-Graduação em Genética e Melhoramento de Plantas pelos ensinamentos.

Ao CNPq e ao Conselho do Curso de Pós-Graduação em Genética e Melhoramento de Plantas pela concessão da bolsa de estudos.

Ao Dr. Romeu A. de Souza Kiihl, Dr. Carlos A. Arrabal Arias, MSc. Ricardo V. Abdelnoor, Dr. Leones A. Almeida e Dr. José F. Ferraz de Toledo, pesquisadores da Embrapa Soja, pela contribuição e conhecimentos transmitidos no decorrer deste trabalho.

Ao Prof. Dr. Sinval Silveira Neto e Prof. Dra. Marinéia L. Haddad, do Departamento de Entomologia da ESALQ/USP e aos colegas Antônia Marlene M. Barbosa, Alexandre S. Guedes Coelho e Paulo A. Vianna Barroso pelo auxílio dispensado nas análises dos resultados da tese.

À Andrés E. L. Reyes, do Centro de Informática da ESALQ/USP, pelo fornecimento da versão atualizada do software PARENT.

Aos colegas de laboratório, Lilka Kawasaki, Carla Cripa Cispino, Solange M. Lugle, Fábio Sartori, Silvana R. Marin, Mayra K. Martins, Katiaires E. Delpin,

Lázara Caramori, Valdemar de P. Carvalho e Héber F. dos Reis pela colaboração no transcorrer das análises laboratoriais.

À Cândida V. de Oliveira e Adriana M. R. Ometto Bussi, secretárias do Departamento de Genética da ESALQ/USP e à Eliana M. Garcia e Silvana C.N. Oliveira, bibliotecárias da ESALQ/USP pelo auxílio prestado.

Aos colegas de curso, especialmente à Edina Regina Moresco, Ana Luiza Ramos, Andréa Mittelmann, Fábio Pandini, Betânia L. Rocha Cabral, Gustavo J. Meano Brito, Cantídio Fernando Gouvea, Paulo Roberto Martins, Adelmo Resende da Silva, Paulo Eduardo T. dos Santos, Solange R. Monteiro, Alexandre Sebben, Sandro R. Fuzatto e Vanoli Fronza pela demonstração de amizade e agradável convivência.

À Roseli F. Caseiro, Neusa P. Jorge, Márcia Schleder, Tânia Bittencourt, Célia Y. Takachi e Paulo Celso de F. de Mello pela fiel amizade.

À todos que contribuíram direta ou indiretamente na realização deste trabalho, o meu sincero agradecimento .

## SUMÁRIO

	Página
LISTA DE SIGLAS, ABREVIATURAS E SÍMBOLOS.....	viii
RESUMO.....	x
SUMMARY.....	xii
1 INTRODUÇÃO.....	1
2 REVISÃO DE LITERATURA.....	3
2.1 Origem da soja, introdução e expansão no Brasil.....	3
2.2 Métodos de determinação da diversidade genética.....	5
2.3 Correlações entre diversidade/similaridade genética obtidas através de coeficientes de parentesco e de marcadores moleculares.....	14
2.4 Predição da variância genética através das estimativas de diversidade genética.....	15
3 PREDIÇÃO DA VARIABILIDADE GENÉTICA EM POPULAÇÕES DE SOJA, ATRAVÉS DA DISTÂNCIA GENÉTICA BASEADA EM MARCADORES AFLP.....	17
RESUMO.....	17
SUMMARY.....	18
3.1 INTRODUÇÃO.....	20
3.2 MATERIAL E MÉTODOS.....	22
3.2.1 Material genético.....	22
3.2.2 Extração e quantificação do DNA.....	23
3.2.3 Análises com marcadores AFLP.....	24
3.2.4 Coleta e análise dos dados de AFLP.....	26
3.2.5 Cálculo do coeficiente de parentesco (f).....	26
3.2.6 Relações entre a variância genética aditiva (D) e as distâncias genéticas (DG e AFD).....	27

3.3 RESULTADOS E DISCUSSÃO.....	27
3.3.1 Análise de AFLP.....	27
3.3.2 Estimativas das distâncias genéticas.....	29
3.3.3 Correlações entre variância genética aditiva (D) e distâncias genéticas (AFD e DG).....	30
3.3.4 Eficiência de predição baseada em AFD.....	37
4 SIMILARIDADE GENÉTICA ENTRE CULTIVARES DE SOJA LIBERADAS NO BRASIL, AVALIADA ATRAVÉS DE MARCADORES AFLP.....	38
RESUMO.....	38
SUMMARY.....	39
4.1 INTRODUÇÃO.....	41
4.2 MATERIAL E MÉTODOS.....	43
4.2.1 Material genético.....	43
4.2.2 Extração e quantificação do DNA.....	46
4.2.3 Análises com marcadores AFLP.....	46
4.2.4 Análise dos dados obtidos por AFLP e cálculo de similaridade genética.....	46
4.2.5 Coeficiente de parentesco.....	47
4.3 RESULTADOS E DISCUSSÃO.....	48
4.3.1 Análise de AFLP.....	48
4.3.2 Estimativas da similaridade genética por AFLP (AFS).....	50
4.3.2.1 Comportamento da diversidade genética em diferentes épocas....	54
4.3.3 Agrupamento das 317 cultivares de soja baseado nas estimativas de similaridade genética via AFLP (AFS).....	56
4.3.4 Agrupamento das 63 cultivares de soja baseado nas estimativas de similaridade genética via AFLP (AFS).....	59
4.3.5 Fatores que podem influenciar as estimativas de similaridade genética obtidas por marcadores moleculares.....	62



4.3.6 Estimativas de similaridade genética baseadas no coeficiente de parentesco (f) e correlação entre a similaridade genética obtida por AFLP (AFS) e pelo coeficiente de parentesco.....	64
5 CONCLUSÕES.....	67
ANEXOS.....	68
REFERÊNCIAS BIBLIOGRÁFICAS.....	75

## LISTA DE SIGLAS, ABREVIATURAS E SÍMBOLOS

U - Unidades

M- Molar

mM- Milimolar

ul- Microlitros

ml- Mililitros

ng- Nanogramas

pmoles- Picomoles

p/v- Relação peso por volume

h- Horas

min- Minutos

s- Segundos

cm- Centímetros

°C- Graus Celsius

MgCl<sub>2</sub> - Cloreto de magnésio

KCl - Cloreto de potássio

TE- Tris-EDTA

Tris- Tris (hidroximetil) aminometano

EDTA- Ácido etilenodiaminotetra acético

ATP- Adenosina tri-fosfato

$\gamma$ -<sup>33</sup>P - Fósforo-gama

DNA- Ácido desoxirribonucleico

PCR- Polimerase Chain Reaction

RFLP- Restriction Fragment Length Polymorphism

RAPD- Random Amplified Polymorphic DNA

AFLP- Amplified Fragment Length Polymorphism

SSR- Single Sequence Repeat

f- Coeficiente de parentesco

DG- Distância Genética obtida pelo coeficiente f

**AFD- Distância Genética obtida pelos marcadores AFLP**

**AFS- Similaridade Genética obtida pelos marcadores AFLP**

**UPGMA- Unweighted Pair-Group Method Arithmetic Average**

# **AVALIAÇÃO DA DIVERSIDADE GENÉTICA ENTRE CULTIVARES BRASILEIRAS DE SOJA, ATRAVÉS DE MARCADORES AFLP**

Autora: ANA LÍDIA VARIANI BONATO

Orientador: Prof. Dr. ISAIAS OLÍVIO GERALDI

## **RESUMO**

Esta tese é composta por dois trabalhos. O objetivo do primeiro foi investigar o potencial de medidas de distâncias genéticas obtidas a partir de marcadores moleculares AFLP (AFD) e do coeficiente de parentesco (DG) na predição da variabilidade genética de populações segregantes de soja. A distância genética entre as seis combinações possíveis entre quatro genótipos foi baseada em 165 polimorfismos, obtidos com 21 combinações de *primers EcoRI/MseI*. A variância genética aditiva foi estimada para sete caracteres avaliados em quatro épocas de semeadura (setembro, outubro, novembro e dezembro) em quatro anos. Os resultados mostraram correlação negativa entre os coeficientes AFD e DG. A correlação entre a variância genética aditiva de produção de grãos e a distância genética (AFD), na média dos quatro anos e das quatro épocas de semeadura, foi baixa, mas significativa. Quando a estimativa foi feita separadamente por época, a correlação mais alta foi observada no mês de outubro. Obtiveram-se coeficientes de correlação maiores em épocas específicas para cada caráter, sendo estas estimativas maiores, também, em relação aos respectivos coeficientes gerais. Entre a variância genética aditiva e a distância genética (DG) houve correlação significativa

apenas para peso de 100 sementes em outubro e dezembro e para número de dias para floração em setembro. Esse estudo mostrou, portanto, que a distância genética avaliada por AFLP pode ser uma ferramenta útil na predição da variabilidade genética da soja e, conseqüentemente, na escolha de genitores para programas de melhoramento genético. O segundo trabalho teve como objetivo avaliar a similaridade genética com marcadores AFLP (AFS) entre 317 cultivares de soja, liberadas no Brasil no período de 1962 a 1998. Para isso, foram utilizadas seis combinações de *primers* *EcoRI/MseI* e estimados os coeficientes de similaridade genética pelo método de Nei & Li (1979). Os coeficientes de parentesco entre 100 cultivares, liberadas entre 1984 e 1998, foram estimados e correlacionados com os coeficientes de similaridade obtidos com os marcadores. A análise de AFLP revelou a ocorrência de aproximadamente 394 bandas, sendo 78 destas polimórficas. Esse número foi suficiente para avaliar a similaridade genética entre as cultivares, conforme revelou a análise *bootstrap*. A interpretação do dendrograma foi dificultada pelo grande número de cultivares utilizadas. Mesmo assim foi possível detectar agrupamentos de cultivares formados de acordo com o esperado através da genealogia dos mesmos. A utilização de um dendrograma formado por um número menor de cultivares permitiu uma melhor interpretação das similaridades entre as mesmas. Não foi encontrada correlação significativa entre os coeficientes de parentesco e os de similaridade genética (AFS). Os resultados mostraram que as cultivares de soja liberadas no Brasil no período entre 1962 e 1998 apresentam alta similaridade genética, a qual permaneceu constante no período estudado. Apesar disso, têm sido obtidos progressos substanciais com o melhoramento genético.

# EVALUATION OF GENETIC DIVERSITY AMONG BRAZILIAN SOYBEAN CULTIVARS USING AFLP MARKERS

Author: ANA LÍDIA VARIANI BONATO

Adviser: Prof. Dr. ISAIAS OLÍVIO GERALDI

## SUMMARY

This thesis is composed by two parts. The first carries out an investigation on the potential of genetic distance measurements obtained from AFLP molecular markers (AFD) and from the parentage coefficient (GD) to predict the genetic variability of segregant soybean populations. The genetic distance among the six possible combinations of four genotypes was based on 165 polymorphisms obtained from 21 *EcoRI/MseI* primer combinations. The additive genetic variance was estimated for seven traits in four sowing periods and four years. The results showed a negative correlation among AFLP and the parentage coefficient estimates of genetic divergence. The correlation between the additive variance for grain yield and the AFLP genetic distance in the average of the four years and four planting periods was low, but significant. Greater correlation coefficients were observed in specific periods for each trait, and these greater estimates were larger than the respective general coefficients. There was significant correlation among the additive variance and the genetic distance (GD) only for the of 100 seed weight trait in October and December and for days to flowering in September. This study showed that the genetic distance estimated by AFLP can be a useful tool for predicting genetic variability in soybean and for the choice of parents to be used in a plant breeding program.

The second study aimed to assess the genetic similarity measured by AFLP among 317 soybean cultivars released in Brazil between 1962 and 1998. Six combinations of *EcoRI/MseI* primers were used and the genetic similarity coefficients were estimated by the Nei and Li method (1979). The parentage coefficients among 100 cultivars, released between 1984 and 1998, were estimated and correlated with the similarity coefficients obtained by the markers. AFLP analysis showed 78 polymorphic bands out of a total of approximately 394. This number was sufficient to assess the genetic similarity among the cultivars in accordance with bootstrap analysis. The dendrogram interpretation was hindered by the large number of cultivars used, but it was possible to detect cultivar groups formed as expected from their genealogy. The use of the dendrogram involving a lower number of cultivars enabled a better interpretation of the similarities. No significant correlation was found among the parentage coefficient and AFLP similarity coefficients. The results showed that the soybean cultivars, released in Brazil between 1962 and 1998, had a high genetic similarity that remained constant throughout the period studied. In spite of this, substantial gains have been obtained in soybean breeding.

## 1 INTRODUÇÃO

A soja, *Glycine max* (L.) Merrill, é considerada a mais importante oleaginosa cultivada no mundo. Devido seu alto valor nutricional, proporcionando múltiplas utilizações, foi responsável pela formação de uma complexa estrutura de produção, de armazenamento, de processamento e de comercialização em todos os países onde é cultivada em larga escala.

Apesar de ser uma das espécies cultivadas há menos tempo no Brasil, ocupa posição de destaque na economia agrícola do país, conferindo-lhe, desde 1975/76, a posição de segundo produtor mundial. Na safra de 1998/99, alcançou uma produção de aproximadamente trinta e um milhões de toneladas de grãos e uma produtividade média de 2.400 kg/ha (CONAB, 2000).

O cultivo comercial iniciou no Rio Grande do Sul e em São Paulo no final dos anos 40. O seu estabelecimento e o crescimento inicial nesses Estados ocorreu, principalmente, devido à excelente adaptação das cultivares introduzidas do sul dos Estados Unidos. Estendeu-se para os demais Estados, estando hoje presente em todas as regiões situadas até a latitude de 5° Sul. A incorporação dessas áreas ao cultivo deveu-se, essencialmente, ao desenvolvimento de cultivares adaptadas às condições de menor latitude pelos programas brasileiros de melhoramento genético.

Em um programa de melhoramento de plantas, a escolha de genitores para a hibridação é uma das etapas mais importantes. De modo geral, o melhorista depara-se com duas situações básicas nesta etapa. A primeira é aquela em que deseja maximizar a recombinação. Nesse caso, deve escolher



genitores geneticamente mais distantes, além de possuírem as características de interesse, com o intuito de obter híbridos que produzam progênie segregantes com maior variabilidade genética e, desta maneira, aumentar a chance de encontrar segregantes transgressivos. Outra situação é a que deseja minimizar a recombinação, uma situação típica de retrocruzamentos. Nessa condição, deve escolher, entre aqueles que possuem a característica que deseja transferir, um genótipo que apresente maior similaridade com o genitor recorrente, e com isso, conseguir uma recuperação mais rápida das demais características deste genitor.

Deste modo, o conhecimento da diversidade genética entre genótipos pode contribuir para a escolha de genitores a serem utilizados nos programas de melhoramento. Nos últimos anos, novas técnicas, que auxiliam na avaliação das relações genéticas entre genótipos, foram desenvolvidas, acompanhando os rápidos e constantes avanços na área da genética molecular. Os marcadores moleculares, utilizados na investigação da diversidade genética de diferentes espécies, são um exemplo disso. Entre os marcadores moleculares empregados no Brasil, o AFLP (Amplified Fragment Length Polymorphism) é o mais recente, com utilização crescente em estudos de caracterização do genoma de plantas.

Esta tese consta de duas pesquisas. A primeira teve como objetivo investigar o potencial de medidas de distâncias genéticas, obtidas a partir de marcadores moleculares AFLP e do coeficiente de parentesco, na predição da variabilidade genética de populações segregantes de soja. Na segunda pesquisa, o objetivo foi avaliar a similaridade genética entre cultivares de soja liberadas no Brasil em diferentes épocas, com marcadores AFLP.

## **2 REVISÃO DE LITERATURA**

### **2.1 Origem da soja, introdução e expansão no Brasil**

A soja, *Glycine max* (L) Merrill, é originária do leste da Ásia, provavelmente da região nordeste da China, segundo Hymowitz (1970). A espécie foi introduzida no Brasil e começou a ser estudada no estado da Bahia, no final do século 19 (D'Utra, 1882). Posteriormente foi introduzida e estudada nos estados de São Paulo e Rio Grande do Sul (Bonato & Dall'Agnoll, 1985). No entanto, o seu cultivo comercial iniciou bem mais tarde, sendo que as primeiras estatísticas oficiais começaram a ser publicadas em 1941, e seu desenvolvimento no Brasil foi sustentado, entre outros fatores, pela fácil adaptação de cultivares introduzidas dos Estados Unidos (Bonato & Bonato, 1987).

A partir dos anos 60, a cultura se expandiu, inicialmente, em latitudes entre 30°S e 20°S, principalmente nos estados do Rio Grande do Sul, Santa Catarina, Paraná e São Paulo. Nesta época predominaram as introduções de cultivares e de linhagens americanas, muitas das quais tornaram-se importantes cultivares dessas regiões. Com a identificação das cultivares de melhor adaptação, os melhoristas passaram a combinar as características destas por meio de hibridações (Bonetti, 1983). Assim, uma série de novas cultivares provenientes de programas de melhoramento brasileiros contribuíram significativamente, já na década de 70, para a elevação dos índices de produtividade.

Nesta mesma época, a soja começou a ser cultivada na chamada região de expansão, compreendendo latitudes menores que 20°S, abrangendo os estados do Mato Grosso do Sul, Mato Grosso, Goiás, Minas Gerais, Distrito Federal, Bahia e Maranhão. Segundo Kiihl et al. (1986), essa expansão foi favorecida pelo desenvolvimento de cultivares melhoradas e adaptadas à zonas equatoriais. Para tanto, procurou-se manter o mesmo tipo de soja utilizada no sul do país, alterando, principalmente, a resposta ao fotoperíodo, através da introdução de genes que retardam o florescimento.

Desta forma, o desenvolvimento de cultivares mais produtivas, estáveis e melhor adaptadas às diferentes condições de cultivo tem representado uma importante contribuição no estabelecimento da soja como uma das principais culturas do país (Almeida et al., 1997). No entanto, estudos da base genética das cultivares brasileiras revelaram que o germoplasma utilizado em programas de melhoramento genético para esta espécie no país como um todo pode ser considerado como uma população de baixo tamanho efetivo (N=11 a 15), (Vello et al., 1988). Assim, apenas 11 ancestrais contribuíam com 89% do conjunto de genes do germoplasma de soja recomendado para cultivo em 1983/84 (Hiromoto & Vello, 1986). Nos Estados Unidos, uma situação semelhante foi encontrada, onde 25 ancestrais participavam de 90% do conjunto gênico das cultivares lançadas entre 1947 e 1988 (Gizlice et al., 1994). A estreita base genética, aliada à seleção das mesmas populações por um longo período de tempo, pode resultar na diminuição do progresso com seleção para produção de grãos e no aumento da vulnerabilidade genética a fatores bióticos e abióticos (Bhatnagar, 1999).

Apesar disso, tem-se obtido ganhos significativos com o melhoramento genético ao longo do tempo. Toledo et al. (1990) estimaram um ganho médio anual de 1,8% para um grupo de cultivares precoces e de 1,3% para o grupo de cultivares semi-precoces do programa de melhoramento de soja do Estado do Paraná, no período de 1981 a 1986. Mais recentemente, Rubin & Santos (1996), ao estudarem cultivares de soja que foram largamente cultivadas em

diferentes períodos no Rio Grande do Sul, concluíram que nos últimos 40 anos foi obtido um ganho médio de 19 kg/ha/ano, ou seja, de 1,1% ao ano. Esses autores comentaram, no entanto, que os ganhos vinham diminuindo com o decorrer do tempo, devido, entre outras razões, ao uso do mesmo germoplasma básico nas hibridações. Isso, no entanto, parece não estar se confirmando, segundo revelam recentes lançamentos de novas cultivares. Além de consideráveis ganhos no aprimoramento de características agrônomicas conseguidos pelas correções de defeitos controlados por caracteres qualitativos, são relatados também ganhos na produção de grãos. No Rio Grande do Sul, a nova cultivar BRS 153, liberada em 1998, tem apresentado produtividade 14% superior à da testemunha mais cultivada, BR-16 (Bonato et al., 1999). No Paraná, a cultivar BRS 133 superou em 8,5% a testemunha BR-16 (Almeida et al., 1999). No Estado de Mato Grosso do Sul, a nova cultivar MS/BRS-171 (Campo Grande), liberada em 1998, superou em 20% a cultivar FT-Cristalina (Zuffo et al., 1999). Em Mato Grosso, a cultivar MT/BR-50 (Parecis), liberada em 1996, foi 5% superior à cultivar padrão MT/BR-45 (Paiaguás), conforme Takeda et al. (1999).

Mesmo assim, a escolha de genitores geneticamente mais distantes é uma alternativa que pode ser utilizada para aumentar a variabilidade disponível nas populações de plantas, e, conseqüentemente, maximizar o ganho genético (Messmer et al., 1993).

## **2.2 Métodos de determinação da diversidade genética**

Diferentes métodos têm sido empregados no sentido de quantificar a divergência genética entre genótipos e prever o nível da variabilidade liberada nas gerações subsequentes a um cruzamento, auxiliando o melhorista na escolha de genitores. Uma das metodologias propostas é o conhecimento do grau de parentesco entre os possíveis genitores. O grau de parentesco entre diferentes cultivares ou linhagens pode ser estimado através do chamado

Coeficiente de Parentesco ou Coeficiente de Malécot ( $f$ ), (Malécot, 1947). Este coeficiente é definido como a probabilidade de duas cultivares ou linhagens possuírem, por descendência, o mesmo alelo em um dado loco escolhido ao acaso. Para seu cálculo, considera-se o número de ancestrais de cada um dos genótipos e suas respectivas contribuições genéticas relativas (Vello et al., 1988; Falconer & Mackay, 1996). Através desse coeficiente foi demonstrada a substancial relação de proximidade genética entre cultivares americanas de soja (Delannay et al., 1983; Specht & Williams, 1984; Carter et al., 1993). Estudos posteriores, através da genealogia, revelaram que cultivares do norte e do sul dos Estados Unidos apresentavam bases genéticas contrastantes (Gizlice, et al., 1993; 1996), e que a diversidade genética do germoplasma de soja norte americano estava sendo reduzido nos últimos 50 anos (Gizlice et al., 1994). No Brasil, Vello et al. (1988) estimaram, através desse coeficiente, o parentesco entre 69 cultivares de soja, recomendadas em 1983/84 e constataram a estreita base genética do germoplasma estudado.

Esta metodologia apresenta a vantagem de estimar a diversidade genética antes da realização do cruzamento. Entretanto, além da necessidade do conhecimento completo e preciso da genealogia, o cálculo do coeficiente de parentesco de Malécot assume pelo menos duas premissas básicas, que nem sempre são verificadas: (1) a ausência de parentesco ( $f=0$ ) entre os ancestrais que se situam na base da árvore genealógica e (2) em um dado cruzamento simples cada genitor contribui com 50% dos alelos para a sua progênie; portanto pressupõe a ausência de seleção (Vello et al., 1988; Carter et al., 1993). Por isso, podem ser mais úteis os métodos que possibilitem determinar mais realisticamente o grau de divergência entre genótipos, independentemente do conhecimento de suas respectivas genealogias.

Assim, diversidade genética também pode ser estimada através de marcadores genéticos, que são considerados como quaisquer características que evidenciem diferenças entre indivíduos, que sejam precisamente reproduzidas nas progênies e que possam ser utilizadas para correlacionar com

outras características de interesse (Almeida et al.,1997). Os marcadores genéticos classificam-se em morfológicos, bioquímicos e moleculares.

Em soja, o uso de marcadores morfológicos e bioquímicos (através de isoenzimas) para a estimação do grau de parentesco entre linhagens e/ou cultivares foi pioneiramente relatado por Cox et al. (1985). Os autores constataram que os índices de similaridade genética, calculados a partir dos dados desses marcadores, apresentaram baixa correlação, entre 0,15 e 0,49, com o coeficiente de parentesco de Malécot. A maior crítica a esse, bem como a outros trabalhos semelhantes, como, por exemplo os realizados com milho (Stubber & Moll,1972; Stubber et al. (1980) e com aveia (Hamrick & Allard, 1975), que tentaram usar dados bioquímicos e/ou morfológicos na estimação da diversidade genética em plantas, é que o número de locos amostrados foi muito pequeno, não representando satisfatoriamente o genoma da planta. Este problema pode ser mais crítico em espécies autógamas devido à ocorrência de baixa variabilidade genética aliada a condições de intensa seleção, tal como acontece com o germoplasma utilizado pelos melhoristas de soja no Brasil (Vello et al., 1988).

Na última década, os marcadores moleculares têm tido destaque em pesquisas como estimadores da divergência genética entre linhagens, cultivares, acessos ancestrais, PIs (Plant Introductions) e espécies de plantas. Eles têm sido considerados mais úteis e eficientes para este fim, pois fornecem medidas mais precisas da divergência, em razão do maior polimorfismo que apresentam, pela estabilidade ambiental e pela natureza genotípica (Santos, 1994). As técnicas de RFLP (Restriction Fragment Length Polymorphism; Botstein et al., 1980) e RAPD (Randomly Amplified Polimorphic DNA; Williams et al.,1990) constituem a primeira geração de marcadores moleculares (Vierling & Nguyen,1992; Demeke et al.,1993).

Na detecção do polimorfismo via RFLP o DNA genômico é digerido por uma enzima de restrição previamente determinada, separado por eletroforese em um gel de agarose, transferido para uma membrana e então hibridizado

com uma sonda marcada, radioativa ou não. As membranas são expostas a filmes e o resultado da hibridação é observado após a revelação deste. O polimorfismo é visualizado como bandas de diferentes tamanhos, entre os indivíduos que apresentarem diversidade genética para o loco analisado. Isto ocorre pois as enzimas de restrição cortam o DNA em sítios específicos, sendo que uma mutação pontual em um destes sítios resulta na perda ou na criação de um novo sítio, dando origem a um fragmento de restrição de tamanho diferente. Mutações causadas por inserção, deleção ou inversão de fragmentos de DNA também resultam em variação dos tamanhos dos fragmentos de restrição (Ferreira & Grattapaglia, 1996).

A técnica de RAPD baseia-se na amplificação de fragmentos não específicos, utilizando a reação em cadeia da polimerase (PCR; Polimerase Chain Reaction) e de oligonucleotídeos de bases arbitrárias como *primers* para amplificar o DNA genômico. Os produtos da PCR são produzidos em regiões ao acaso do genoma, flanqueadas por dois sítios que apresentam complementariedade aos *primers*, e são analisados por eletroforese em géis de agarose corados com brometo de etídio (Ferreira & Grattapaglia, 1996). Os polimorfismos são reconhecidos pela presença de um fragmento amplificado em um dos genótipos em relação à ausência desse mesmo fragmento em outro genótipo. Embora tenha sido amplificado por um *primer* arbitrário, a presença ou a ausência desse produto específico irá revelar a existência ou não, respectivamente, de sítios específicos complementares a esse *primer* no DNA em estudo (Williams et al., 1990).

Vários trabalhos vêm sendo desenvolvidos no sentido de analisar a diversidade genética diretamente dos dados obtidos de marcadores moleculares. D'Hont et al. (1995) identificaram eficientemente, com o auxílio de marcadores RFLP, a variabilidade genética entre híbridos interespecíficos mediante a presença de marcadores específicos de cana-de-açúcar e *Erianthus arundinaceus*, embora, esta última apresente um nível limitado de polimorfismo. O'Donoghue (2000), através de RFLP, verificou a similaridade genética entre

cultivares de aveia norte-americanas. Este análise possibilitou, também, separar claramente as cultivares de primavera das cultivares de inverno. Outros estudos mais recentes utilizando marcadores RAPD são relatados a seguir. Uma análise da diversidade genética em acessos de sorgo através desses marcadores possibilitou identificar e diferenciar raças e regiões onde havia maior divergência (Menkir et al., 1997). Johns et al. (1997) contribuíram para a classificaram variedades locais e cultivares de feijão em diferentes *pools* gênicos. De forma semelhante, Duarte et al. (1999b) classificaram cultivares dessa espécie de acordo com o centro de origem, a partir das estimativas de distância genética. Um trabalho mais completo de análise da diversidade entre genótipos de beterraba envolvendo 50 anos de pesquisa deste cultura nos EUA foi realizado com o objetivo de examinar o efeito do melhoramento, ao longo do tempo, no ganho ou na perda da diversidade genética. Os pesquisadores concluíram que a diversidade dentro do germoplasma analisado permaneceu constante (McGrath et al., 1999).

Em soja, Thompson & Nelson (1998) avaliaram PIs, mantidas na coleção da USDA, através do mesmo tipo de marcador e observaram que algumas linhagens experimentais, provenientes do cruzamento de cultivares modernas com PIs, além de mais divergentes geneticamente, apresentaram maior produtividade quando comparadas àquelas oriundas de genitores domesticados. Nesta linha de pesquisa, Thompson et al. (1998) estudaram a diversidade entre ancestrais e PIs de soja selecionadas da mesma coleção. Utilizando quatro métodos de agrupamento, sendo dois deles métodos hierárquicos (UPGMA e método da mínima variância de Ward) e outros dois não hierárquicos, obtiveram 87% de coincidência dos genótipos estarem no mesmo grupo, demonstrando a eficiência da técnica. Esses trabalhos demonstram a efetividade da aplicação dos marcadores moleculares na compreensão da relação entre genótipos através da diversidade.

Embora o emprego de marcadores RFLP e RAPD nesses estudos tenha sido encorajador, o uso dessas técnicas em programas de melhoramento



genético de plantas tem esbarrado em problemas técnicos intrínsecos à natureza do marcador. O marcador RFLP é tecnicamente trabalhoso, de difícil automação, e, no caso da soja, apresenta um agravante que é a baixa informatividade (Lorenzen et al., 1995). Comparado com o RFLP, o surgimento do RAPD representou um avanço técnico no que se refere à simplicidade e à perspectiva de automação, através da incorporação da técnica de PCR (Polimerase Chain Reaction), visto reduzir as quantidades de DNA necessárias, eliminar as etapas de transferência e de hibridização, e possibilitar amostrar-se maior número de locos num mesmo gel (1 a 10), segundo Ferreira & Grattapaglia (1996). Em contrapartida, o RAPD pode apresentar sérios problemas de reproducibilidade, o que dificulta bastante a sua aplicação em larga escala. Além disso, o RAPD não representou um avanço na informatividade.

Mais recentemente, outros tipos de marcadores vem sendo desenvolvidos como uma forma de contornar esses problemas: os microssatélites (Jacob et al., 1991) e o AFLP (Amplified Fragment Length Polymorphism; Vos et al., 1995). Os microssatélites são também conhecidos como SSR (Simple Sequence Repeats) e consistem em uma subclasse de DNA repetitivo formada por pequenas sequências (2 a 5 nucleotídeos) repetidas em tandem, tais como (AT) $n$ , (ATT) $n$ , por exemplo. A variação no número de repetições ( $n$ ) dessas sequências gera uma grande quantidade de polimorfismo, favorecendo sua utilização em estudos genéticos.

A ocorrência deste tipo de polimorfismo em plantas foi relatada, inicialmente em soja, por Akkaya et al. (1992). Entre outras aplicações, os microssatélites têm sido utilizados na caracterização da diversidade genética, na identificação de germoplasma e no estudo da dinâmica de populações (Thomas & Scott, 1993; Yang et al., 1994; Rongwen et al., 1995; Smith et al., 1997; Senior et al., 1998). Em soja, o alto nível de polimorfismos revelado pelos marcadores microssatélites foi explorado por Glazko et al. (1999), na caracterização e separação entre variedades locais e genótipos selvagens,

mediante a análise da diversidade genética. Xu et al. (1999) utilizaram os marcadores microssatélites para avaliarem a diversidade genética da soja na China, e esta análise permitiu a divisão das cultivares em sete grupos, com evidente associação de determinados grupos com resistência à fatores bióticos e abióticos. Diwan & Cregan (1997) ao analisarem a diversidade gênica dessa espécie, através de 20 locos por microssatélites, distinguiram cultivares norte-americanas modernas, que apresentam características morfológicas idênticas, entretanto, não conseguiram separar dois genótipos de soja, sendo que um deles era seleção do outro.

Na técnica de AFLP, o DNA genômico é digerido simultaneamente com duas enzimas de restrição. A seguir, os fragmentos gerados são ligados a adaptadores nos sítios das enzimas e amplificados por PCR, utilizando-se *primers* com sequências complementares aos adaptadores. Em seguida é feita uma amplificação seletiva por PCR com a adição de nucleotídeos arbitrários à extremidade 3' do *primer*, propiciando o aumento na seletividade do *primer* e a redução na complexidade dos produtos amplificados. Os fragmentos de DNA amplificados são detectados pelo fracionamento em géis de sequenciamento, permitindo detectar polimorfismos de até um par de bases (Vos et al., 1995). O marcador AFLP pode ser encarado como um RAPD bastante melhorado, onde até 100 locos (cerca de dez vezes mais que o RAPD) podem ser amostrados num único gel, com uma alta resolução e reproducibilidade (Meksem et al., 1995). O aumento da reproducibilidade deve-se, principalmente, ao fato do *primer* utilizado para a reação de PCR não ser de natureza arbitrária, como no RAPD, mas sim, específico para o adaptador. Isto permite o uso de condições mais estritas de anelamento do *primer*, aumentando assim a especificidade da reação (Vos et al., 1995). Por suas características, essa técnica tem sido utilizada com sucesso para estudos de diversidade genética em várias espécies. O número de trabalhos tem aumentado principalmente nos últimos dois anos. Alguns dos resultados serão apresentados a seguir.

Sharma et al. (1996) avaliaram a diversidade e as relações filogenéticas entre seis populações de lentilha cultivada e seus parentes selvagens e compararam os resultados com dados prévios, obtidos com os mesmos acessos pela técnica de RAPD. A quantidade de polimorfismo obtida por AFLP foi muito maior do que a observada por RAPD, permitindo detectar variações e afinidades entre as subespécies selvagens e os grupos de cultivares, ou mesmo, dentro das subespécies. Tohme et al. (1997) avaliaram a estrutura genética entre e dentro do conjunto de genes, com 114 genótipos selvagens de feijão, constatando distinções relevantes entre genótipos provenientes de diferentes regiões de origem. Em arroz, a diversidade entre acessos foi investigada por Zhu et al. (1998), obtendo marcadores grupo-específicos. A diversidade entre acessos de *Glycine max* e de *Glycine soja* também foi estudada por Machado (1999), que demonstrou existir baixa variabilidade dos genótipos adaptados em relação aos demais. A diversidade encontrada em coleções de germoplasma compostas por espécies selvagens, PIs e cultivares tem sido investigada também em cana-de-açúcar (Besse et al., 1998), em beterraba (Hansen et al., 1999) e em capim-bermuda (Zhang et al., 1999).

Visando a detecção da diversidade entre cultivares, Hartl & Seefelder (1998) empregaram marcadores AFLP na avaliação entre cultivares de lúpulo. A similaridade genética média foi de 0,89, demonstrando a efetividade da técnica na detecção de variação ao nível de sequências de DNA entre genótipos altamente relacionados. Em um estudo similar, Barrett & Kidwell (1998) compararam 56 cultivares de trigo de inverno e de primavera. Obtiveram um coeficiente de 0,58 de diversidade entre cultivares de inverno vs Primavera, 0,53 dentro das cultivares de inverno e de 0,49 dentro das cultivares de primavera. Em trabalho subsequente, compararam linhagens de trigo de inverno e de primavera quanto a diversidade, baseada na genealogia e em dados de AFLP e constataram que em 89% dos casos os valores da distância genética baseada na genealogia foram acima de 0,90; enquanto que os valores da distância genética medida por AFLP foram distribuídos normalmente, com

média 0,54, sugerindo que a diversidade foi melhor identificada por esse último método (Barret et al., 1998).

Estudos comparativos entre tipos de marcadores propiciaram informações sobre suas performances e limitações quanto a aplicabilidade na estimação da diversidade genética. Com esse objetivo, Pejic et al. (1998) avaliaram a similaridade genética entre linhagens de milho através das técnicas RFLP, RAPD, SSR e AFLP, e obtiveram uma maior eficiência com AFLP, em relação aos demais. Acessos de feijão Azuki, provenientes de diferentes regiões da Ásia, foram agrupados de acordo com a similaridade, onde o AFLP apresentou resultados mais elucidativos quando comparado ao RAPD (Yee et al., 1999). Chavarriaga-Aguirre et al. (1999) avaliaram a diversidade contida em uma *Core Collection* de mandioca através de isoenzimas, de SSR e de AFLP. Este último marcador possibilitou a identificação de acessos duplicados.

Em soja, Powell et al. (1996) compararam diferentes marcadores moleculares no estudo das relações genéticas entre genótipos (cultivares e ancestrais) de *Glycine max* e genótipos de *Glycine soja*. Os níveis mais elevados de polimorfismo foram detectados por marcadores microssatélites, enquanto que os marcadores AFLP apresentaram a mais efetiva proporção *multiplex*, isto é, o número de locos analisados simultaneamente em um gel. Comparando as matrizes de similaridades genéticas de *Glycine max* e *Glycine soja*, elaboradas a partir de dados obtidos por AFLP e por microssatélites, observaram que as matrizes eram altamente correlacionadas, indicando congruência entre os métodos, entretanto, analisando somente *Glycine Max*, a média de similaridade genética entre os genótipos foi menor quando utilizaram microssatélites (0,34), em comparação com o AFLP (0,65), refletindo, neste caso, um maior poder de resolução dos microssatélites.

### **2.3 Correlações entre diversidade/similaridade genética obtidas através de coeficientes de parentesco e de marcadores moleculares**

Estudos de correlações entre as estimativas da diversidade, ou da similaridade genética, obtidas pelo coeficiente de parentesco (Coeficiente de Malécot,  $f$ ) e por marcadores moleculares têm sido publicados para várias espécies, apresentando resultados distintos.

Em milho, a correlação entre o coeficiente de parentesco e as medidas com dados de RFLP revelou-se, na maioria dos casos, alta, variando de 0,57 a 0,98 (Smith et al., 1990; Melchinger et al., 1991; Bernardo, 1993; Brunklaus-Jung et al., 1993; Bernardo, 1994), em relação aos trabalhos com o mesmo objetivo, porém empregando-se a técnica de RAPD (0,48) (Hahn et al., 1995). Entretanto, Ajmone-Marsan et al. (1992) relataram valores mais baixos em dois grupos de linhagens (0,49 e 0,01, respectivamente), utilizando-se marcadores RFLP. Em aveia, tanto RFLP como RAPD revelaram baixas correlações com o coeficiente  $f$  (Tinker et al., 1993; Graner et al., 1994), enquanto que em arroz (Mackill, 1995) e em lúpulo (Sustar-Vozlic & Javornik, 1999) foi possível separar cultivares relacionadas através de RAPD, com suporte dos dados da genealogia.

Na cultura da soja, Keim et al. (1992), ao compararem as medidas de  $f$  e de similaridade genética, através de dados de RFLP, entre linhagens ancestrais e cultivares modernas, encontraram uma correlação geral de magnitude moderada ( $r=0,54$ ). Entretanto, quando somente as cultivares foram analisadas, a correlação encontrada foi expressivamente maior ( $r=0,73$ ). Segundo os autores, esses resultados refletem erros nas genealogias das linhagens ancestrais. Abdelnoor et al. (1995) avaliaram a diversidade genética entre 38 cultivares brasileiras de soja, através de RAPD. Embora os autores afirmem haver, de maneira geral, uma concordância entre os dados de RAPD e de análise de parentesco, nenhum valor de correlação foi apresentado. Mais importante, talvez, tenha sido o fato desses autores terem mostrado vários

casos de discrepância entre a análise de RAPD e da genealogia. Essa discrepância foi interpretada como evidência concreta da influência da seleção na contribuição do genoma de cada genitor no genoma de sua progênie. Há que se salientar que diferentes genótipos foram utilizados nesses estudos. Além disso, medidas de coeficiente de parentesco e de similaridade genética a partir de marcadores moleculares são medidas de natureza diferentes, baseadas em diferentes tipos de informações e, portanto, estão sujeitas a diferentes tipos de erro (Diwan & Cregan, 1997).

#### **2.4 Predição da variância genética através das estimativas da diversidade genética**

A utilização das estimativas de diversidade genética na predição da variabilidade genética resultante de um cruzamento tem sido investigada com o intuito de facilitar a escolha de genitores a serem incluídos em um programa de melhoramento.

Em relação ao emprego do coeficiente de parentesco na predição do potencial de um cruzamento, medido pela variância genética das progênies, em aveia, Cowen & Frey (1987) encontraram uma correlação positiva entre a distância genética dos genitores e a variância dos cruzamentos derivados. Ainda em aveia, Souza & Sorrels (1991) consideraram a distância genética, obtida pelo coeficiente  $f$ , um preditor significativo da variância para vários caracteres quantitativos. Entretanto, resultados inconsistentes entre o coeficiente de parentesco e a predição da performance de cultivares dessa espécie foram observados por Beer et al. (1995). Em soja, Toledo (1992) encontrou uma correlação de 0,59 entre o potencial de cruzamento, no caso, medido pelo número de linhagens com produção superior à produção de cultivares padrões, e o coeficiente  $f$ . Este autor concluiu que o método se mostrou eficaz quando associado à taxas medianas de seleção de

cruzamentos, com 78% de acerto quando foram escolhidos os 50% cruzamentos superiores.

Manjarrez-Sandoval et al. (1997) propuseram-se a investigar a utilidade das medidas de similaridade genética medida por RFLP e pelo coeficiente  $f$  na predição da variação genética para produção de grãos em soja, avaliando cinco populações conduzidas por *single seed descent*. Os autores obtiveram uma alta correlação ( $r=0,91$ ) entre os dados de  $f$  e os de RFLP. Entretanto, o coeficiente  $f$  foi mais eficiente ( $r=-0,81$ ) do que os marcadores RFLP ( $r=-0,58$ ) na predição da variação genética para produção de grãos. Com o mesmo propósito, Kisha et al. (1997) verificaram correlações positivas entre distância genética, baseada em dados de RFLP e  $f$ , e a variação genética para produção de grãos, maturidade e altura de planta, na maioria das populações avaliadas. Entretanto, Helms et al. (1997) além de não obterem aparente associação entre  $f$  e a distância genética baseada em marcadores RAPD, ao analisarem seis populações dessa espécie, não encontraram correlação significativa entre as estimativas de distâncias e a variância genética.

No que se refere à utilização de técnicas mais modernas na predição da variância genética, Burkhamer et al. (1998) utilizaram marcadores AFLP para estimar a distância genética entre genótipos de trigo e obtiveram correlações moderadas, embora não significativas, entre essa medida e a variância genética de vários caracteres de interesse. Da mesma forma, Bohn et al. (1999) utilizando marcadores AFLP e RFLP, não encontraram associações significativas com estimativas da variância genética, para sete caracteres analisados em trigo, entre eles altura de plantas na maturação, peso de 100 sementes e produção de grãos.

Esses resultados sobre a predição da variância genética através das distâncias genéticas evidenciaram-se, ainda, conflitantes, demonstrando que em alguns casos a aplicação das estimativas de diversidade genética para esta finalidade permanece incerta.

### **3 PREDIÇÃO DA VARIABILIDADE GENÉTICA EM POPULAÇÕES DE SOJA, ATRAVÉS DA DISTÂNCIA GENÉTICA BASEADA EM MARCADORES AFLP**

#### **RESUMO**

Este trabalho teve como objetivo investigar o potencial de medidas de distâncias genéticas obtidas a partir de marcadores moleculares AFLP (AFD) e do coeficiente de parentesco (DG) na predição da variabilidade genética de populações segregantes de soja. Para este propósito foi estimada a distância genética entre as seis combinações possíveis entre três cultivares e uma linhagem (Ocepar 8, BR-13, FT-2 e BR 85-29009). A análise de AFLP foi baseada em 165 polimorfismos, obtidos com 21 combinações de *primers* *EcoRI/MseI*. A variância genética aditiva foi estimada a partir de uma amostra aleatória de linhagens homozigóticas (F<sub>7</sub> a F<sub>10</sub>, dependendo do ano) para sete caracteres: produção de grãos, número de nós, número de dias para floração, altura de plantas na floração, número de dias para maturação, altura de plantas na maturação e peso de 100 sementes, em quatro épocas de semeadura (setembro, outubro, novembro e dezembro) e em quatro anos consecutivos. Os resultados mostraram uma correlação negativa entre AFD e DG ( $r=-0,39$ ). Foi encontrada uma baixa, mas significativa correlação geral ( $r=0,21^*$ ) entre a variância genética aditiva de produção de grãos e a distância genética (AFD) na média dos quatro anos e das quatro épocas de semeadura. Quando os coeficientes de correlação entre esses dois parâmetros foram estimados por época de semeadura, os maiores valores foram encontrados em outubro ( $r=0,41^*$ ) e novembro ( $r=0,37$ ). Na análise dos demais caracteres observou-se



coeficientes de correlação maiores em épocas específicas para cada caráter, sendo esses valores mais elevados, também, que os respectivos coeficientes gerais. Avaliando a relação entre a variância genética e as distâncias genéticas identificou-se um dos cruzamentos como sendo um *outlier*. Retirando este cruzamento da análise, a magnitude e o nível de significância da correlação geral e por épocas de semeadura aumentaram. No caso de produção de grãos, a correlação geral aumentou para 0,43\*\* e as correlações para as épocas normais de semeadura (outubro e novembro) aumentaram para 0,70\*\* e 0,52\*\*, respectivamente. Não encontrou-se correlação geral significativa entre a variância genética e a distância genética (DG) para todos os caracteres, com exceção de peso de 100 sementes ( $r=0,32^*$ ). A eficiência de predição, considerada como a probabilidade de identificar os três cruzamentos com maior variância genética através da seleção dos três cruzamentos com maiores coeficientes de distância genética (AFD), para uma taxa de 50% de seleção, variou entre 33,3% a 100%. Para a produção de grãos a eficiência de predição foi de 66,6%; 83,3%; 83,3% e 58,3% para setembro, outubro, novembro e dezembro, respectivamente. Concluiu-se que a distância genética medida por AFLP pode ser uma ferramenta útil na predição da variabilidade genética da soja e, conseqüentemente, na escolha de genitores para programas de melhoramento genético.

## SUMMARY

The potential of genetic distance measurements obtained from AFLP molecular markers (AFD) and from the parentage coefficient (GD) to predict the genetic variability in segregant soybean populations was investigated. The genetic distance among the six possible combinations between three cultivars and a breeding line (Ocepar 8, BR-13, FT-2 and BR 85-29009) was estimated. The AFLP analysis was based on 165 polymorphisms, obtained with 21 *EcoRI/MseI* primer combinations. The additive genetic variance was estimated

from inbred lines data ( $F_7$  to  $F_{10}$ , depending on the year) for seven traits: grain yield, number of nodes, days to flowering, plant height at flowering, days to maturity, plant height at maturity and 100 seed weight, in four sowing periods (September, October, November, and December) in four consecutive years. A negative correlation between AFD e GD ( $r=-0.39$ ) was obtained. A low but significant general correlation was found ( $r=0.21^*$ ) between the additive genetic variance of grain yield and the genetic distance by AFLP in the mean of the four years and the four sowing periods. When the correlation coefficients between these two parameters were estimated by sowing period, the largest value were obtained in October ( $r=0.41^*$ ) and in November( $r=0.37$ ). The analysis of the other traits revealed the presence of greater correlation coefficients in specific periods, and these values were also higher than the respective general coefficients. One of the crosses was identified as an outlier when the relationship between the additive genetic variance and the distances was assessed. The removal of this cross from the analysis implied in an increase of the magnitude and level of significance of the general and by sowing period correlations. The overall correlation for grain yield increased ( $r=0.43^{**}$ ). Moreover, correlations for the normal sowing dates (October and November) increased to  $0.70^{**}$  and  $0.52^{**}$ , respectively. The general correlations among additive genetic variance and genetic distance GD were, in general, non-significant for any trait. The only exception was 100 seed weight ( $r=0.32^*$ ). The prediction efficiency considered as the probability of identifying the three crosses with greatest variance by selecting the three crosses with the greatest AFLP genetic distance coefficients, meaning a 50% selection proportion, varied between 33.3% and 100%. Prediction efficiency for grain yield was 66.6%, 83.3%, 83.3%, and 58.3% for September, October, November, and December, respectively. It was concluded that AFLP genetic distance estimates can be a useful tool in predicting soybean genetic variability, and in the choice of parents to be used in a plant breeding program.

### 3.1 INTRODUÇÃO

A escolha de genitores para cruzamentos é uma das etapas fundamentais num programa de melhoramento genético de plantas. Considerando genótipos com performances agronômicas similares, os cruzamentos entre pares geneticamente mais distantes deverão ter um maior número de alelos segregantes e, conseqüentemente, uma maior variância genética nas progênies (Cox et al., 1985; Messmer et al., 1993). Desta forma, a sustentabilidade do ganho genético com seleção pode ser alcançada com a escolha de genitores adaptados, com boa produtividade e que apresentem maior divergência genética. O conhecimento *a priori* da variabilidade genética oriunda de um cruzamento e a conseqüente predição do potencial dos cruzamentos são ferramentas propostas para auxiliar o melhorista na escolha de genitores.

O Coeficiente de Parentesco ( $f$ ), ou Coeficiente de Malécot (Malécot, 1947), definido como a probabilidade de dois genótipos possuírem, por descendência, o mesmo alelo num dado loco escolhido ao acaso, é um dos métodos empregados para se medir a divergência genética ou o grau de parentesco entre cultivares e/ou linhagens. Essa metodologia apresenta a vantagem de permitir a estimativa da diversidade genética antes da realização do cruzamento. Entretanto, além da necessidade do conhecimento completo e preciso da genealogia dos genótipos genitores, o cálculo do coeficiente de Malécot assume, pelo menos, duas premissas básicas: 1) ausência de parentesco entre as linhagens ancestrais que se situam na base da árvore genealógica e 2) em um dado cruzamento simples, cada genitor contribui com 50% dos alelos para sua progênie (Vello et al., 1988).

Esta medida tem sido utilizada em estudos de avaliação da base genética de germoplasma de soja. Vello et al. (1988), estimaram o coeficiente  $f$  entre 69 cultivares de soja recomendadas no Brasil e caracterizaram a base

genética desta espécie no país. Nos Estados Unidos, foi demonstrada a substancial relação de proximidade genética entre cultivares nacionais, através desse coeficiente (Delannay et al., 1983; Specht & Williams, 1984; Carter et al., 1993). Da mesma forma, Gizlice et al. (1994) analisaram as cultivares dos EUA e do Canadá. A viabilidade de seu emprego na predição do potencial de cruzamentos ainda foi pouco estudada. Em aveia, Cowen & Frey (1987) encontraram uma correlação positiva entre a distância genética ( $1 - f$ ) entre genitores, medida a partir de dados de genealogia, e a variância genética dos cruzamentos derivados. Souza & Sorrels (1991), consideraram a distância genética estimada pelo coeficiente  $f$  um bom preditor da variância genética de vários caracteres quantitativos para essa espécie.

Na última década, os marcadores moleculares têm recebido destaque em pesquisas visando o emprego dos mesmos como estimadores da divergência genética entre linhagens, cultivares e espécies de plantas. Ao contrário das análises baseadas em dados genealógicos, os marcadores moleculares fornecem uma medida direta ao nível de DNA. Considerando a utilidade da variância genética na previsão de cruzamentos biparentais superiores (Toledo, 1987) e que a variância genética é dependente, em parte, do grau de divergência genética entre genótipos (Moser & Lee, 1994), a associação entre a variância genética e a distância genética por marcadores em soja tem sido relatada. Em relação à correlação das medidas de divergência entre genótipos a partir de marcadores com as variâncias genéticas de características de interesse em soja, Manjarrez-Sandoval et al. (1997) observaram que o coeficiente  $f$  foi mais eficiente que os marcadores RFLP na predição da variância em cinco populações. No entanto, essas medidas ( $f$  e RFLP) não forneceram uma previsão precisa da variância de cruzamentos individuais, mas foram capazes de estimar diferenças médias significativas de variabilidade entre grupos de cruzamentos (Kisha et al., 1997). Da mesma forma, Helms et al. (1997), embora não tenham obtido associação significativa entre a distância genética através do coeficiente de parentesco e

RAPD com a magnitude da variância genética entre populações individuais, conseguiram relacionar um grupo de populações mais distantes geneticamente com uma maior variância e outro grupo de populações mais próximas com uma menor variância.

Embora o emprego de marcadores RFLP e RAPD em estudos relativos à divergência genética tenha sido encorajador, o uso dessas técnicas esbarra em problemas como a morosidade e a difícil automação do RFLP e a falta de reproducibilidade do RAPD. Uma alternativa para se contornar esses problemas seria o uso da técnica de AFLP (Vos et al., 1995) devido à sua alta informatividade associada à maior resolução e reproducibilidade (Meksem et al., 1995). Esta técnica já tem sido empregada em estudos da predição da variância em progênies de trigo (Burkhamer et al., 1998; Bohn et al., 1999).

Este trabalho tem por objetivo investigar o potencial de medidas de distâncias genéticas obtidas a partir de marcadores moleculares AFLP (AFD) e do coeficiente de parentesco (DG) na predição da variabilidade genética de populações segregantes de soja.

## **3.2 MATERIAL E MÉTODOS**

### **3.2.1 Material genético**

Os genótipos utilizados neste trabalho foram: BR 85-29009, FT-2, BR-13, e Ocepar 8. Esses genótipos foram previamente cruzados em um esquema dialélico, originando seis populações distintas (Triller & Toledo, 1996; Tabela 1). As variâncias entre famílias das gerações avançadas  $F_7$ ,  $F_8$ ,  $F_9$  e  $F_{10}$  foram previamente estimadas a partir de dados experimentais obtidos nos anos agrícolas de 1991/92, 1992/93, 1993/94 e 1994/95, respectivamente. Nos anos 1991/92 e 1992/93 as épocas de semeadura avaliadas foram outubro, novembro e dezembro. Nos anos seguintes foi incluído o mês de setembro. Um

total de 100 famílias, com 4 plantas por família, foram avaliadas em cada geração e os experimentos foram conduzidos na Embrapa Soja, localizada em Londrina, PR, Brasil, conforme descrito em Triller & Toledo (1996).

As estimativas da variância genética aditiva em gerações avançadas de endogamia (D) foram obtidas pelo método dos quadrados mínimos ponderados, de acordo com Cavalli e Hayman, citados por Mather & Jinks (1984), através do programa Genfit (Toledo, 1991). Os caracteres incluídos neste trabalho foram: produção de grãos (g/planta), número de dias para floração, altura de plantas na floração (cm), número de dias para maturação, altura de plantas na maturação (cm), número de nós e peso de 100 sementes (g).

Tabela 1. Populações de soja avaliadas.

<b>População</b>	<b>Cruzamento</b>
1	BR 85-29009 x FT-2
2	BR 85-29009 x BR-13
3	BR 85-29009 x OCEPAR 8
4	FT-2 x BR-13
5	FT-2 x OCEPAR 8
6	BR-13 x OCEPAR 8

### 3.2.2 Extração e quantificação de DNA

Para extração do DNA, foram coletadas folhas jovens (V2) de plantas de cada genótipo, cultivadas em casa de vegetação e conservadas a  $-80^{\circ}\text{C}$ . O DNA dos quatro genótipos foi extraído a partir do tecido foliar congelado, conforme metodologia descrita por Saghai-Marroof et al. (1984).

A concentração de DNA foi estimada em géis de agarose 0,8% (p/v) através da comparação visual da intensidade das bandas do DNA extraído com as bandas do DNA do fago lambda de concentrações conhecidas. As amostras de DNA foram diluídas para uma concentração de aproximadamente 30 ng/ $\mu$ l e mantidas a -20°C.

### 3.2.3 Análises com marcadores AFLP

Esta etapa foi realizada com o auxílio do “AFLP Analyses Kit” (Gibco-BRL), seguindo as normas do fabricante. O DNA genômico (300 ng) foi digerido em um volume final de 25  $\mu$ l com 1,25 U de cada uma das enzimas de restrição *EcoRI* e *MseI* em tampão contendo 10 mM Tris-HCl pH 7,5, 10 mM acetato de magnésio, 50 mM acetato de potássio, a 37°C por 2 h. As enzimas foram inativadas a 70°C por 15 min. Adaptadores *EcoRI* e *MseI* foram ligados ao DNA digerido através da adição de 25  $\mu$ l de uma solução contendo 0,4 mM ATP, 10 mM Tris-HCl pH 7,5, 10 mM acetato de magnésio, 50 mM acetato de potássio e 1 U de DNA ligase T4. A reação foi incubada por 2 h a 20°C e em seguida o DNA ligado aos adaptadores foi diluído em TE (10:1). Uma alíquota de 5  $\mu$ l foi utilizada em uma reação de PCR contendo 10 mM Tris-HCl pH 8,3, 1,5 mM MgCl<sub>2</sub>, 50 mM KCl, 1 U de Taq polimerase e 50 pmoles de *primers* de pré-amplificação. A reação foi submetida a 20 ciclos de PCR com 30 s a 94°C para desnaturação, 60 s a 56°C para anelamento, e 60 s a 72°C para extensão por ciclo. Essa reação foi diluída 50 vezes em TE e armazenada a -20°C para uso posterior. Antes da amplificação final e seletiva, realizou-se a fosforilação do *primer* seletivo *EcoRI* numa reação contendo 0,5 ng de *primer*, 50 mM Tris-HCl pH 7,6, 10 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM 2-mercaptoetanol, 10  $\mu$ l de [ $\gamma$ -<sup>33</sup>P]ATP e 20 U de T4 polinucleotídeo kinase. A reação foi incubada a 37°C por 1 h e em seguida foi interrompida por inativação térmica da enzima a 70°C por

15 minutos. A amplificação final seletiva foi feita a partir de 5 µl do DNA diluído em solução contendo 10 mM Tris-HCl pH 8,3, 1,5 mM MgCl<sub>2</sub>, 50 mM KCl, 5 ng de *primer* seletivo *EcoRI* marcado com [ $\gamma$ -<sup>33</sup>P]ATP, 30 ng de *primer* seletivo *Msel*, 0,1 U de Taq polimerase. A reação foi submetida a 1 ciclo de PCR com 30 s a 94°C para denaturação, 30 s a 65°C para anelamento, e 60 s a 72°C para extensão, seguida por mais 12 ciclos de "touch down" PCR (onde a temperatura de anelamento dos *primers* foi reduzida de 0,7°C por ciclo) e 23 ciclos de PCR com 30 s a 94°C para denaturação, 30 s a 56°C para anelamento, e 60 s a 72°C para extensão. Ao produto da amplificação adicionou-se o mesmo volume de um tampão de amostra (98% formamida, 0,25% azul de bromofenol e 10mM EDTA). Para detecção dos produtos de PCR, uma alíquota de 2,0 ml de cada reação, descrita acima, foi fracionada em gel de sequenciamento (5% de poliacrilamida, 8 M em uréia, 90 mM Tris-borato, 2 mM EDTA). As amplificações foram conduzidas no termociclador Perkin-Elmer Gene Amp 9600 (Perkin-Elmer Corp., Norwalk, CT). Após secagem, o gel foi exposto a um filme de autoradiografia (Kodak X-OMAT-AR). Foram utilizadas 21 combinações de *primers* para amplificação (Tabela 2).

Tabela 2. Combinações de *primers EcoRI/Msel* utilizados para as análises de AFLP

<i>EcoRI</i>	<i>Msel</i>	<i>EcoRI</i>	<i>Msel</i>	<i>EcoRI</i>	<i>Msel</i>
AAC	CTG	AAG	CTG	ACT	CTG
AAC	CTC	AAG	CTC	ACT	CTA
AAC	CTA	AAG	CTA	ACT	CTT
AAC	CAC	AAG	CTT	ACT	CAC
AAC	CAG	AAG	CAC	ACT	CAG
AAC	CAA	AAG	CAA	ACT	CTC
AAC	CAT	AAG	CAT	ACT	CAT



### 3.2.4 Coleta e análise dos dados de AFLP

A interpretação das autoradiografias foi feita usando-se um sistema binário, ou seja, presença ou ausência de uma mesma banda numa determinada combinação de genótipos.

A similaridade genética (AFS) entre os pares de cultivares foi calculada a partir da matriz de dados de AFLP e baseada no coeficiente de Nei & Li (1979), segundo a equação:

$$AFS_{ij} = [2N_{i,j} / (2N_{i,j} + N_i + N_j)]$$

em que  $AFS_{ij}$  representa a medida da similaridade, a partir de dados de AFLP, entre os genótipos  $i$  e  $j$ ,  $N_{i,j}$  é o número total de bandas comuns a  $i$  e  $j$ , e  $N_i$  e  $N_j$  correspondem ao número de bandas encontradas nos genótipos  $i$  e  $j$ , respectivamente.

A distância genética por AFLP (AFD) foi obtida pela diferença:

$$AFD = 1 - AFS$$

### 3.2.5 Cálculo do coeficiente de Parentesco (f)

O coeficiente de parentesco ( $f$ ; Malecót, 1947) foi calculado a partir da genealogia dos genitores, conforme Vello et al. (1988). A distância genética, calculada através do coeficiente de parentesco (DG), foi estimada como  $1 - f$ .

### **3.2.6 Relações entre a variância genética aditiva (D) e as distâncias genéticas (DG e AFD)**

As relações entre D, DG e AFD foram avaliadas baseando-se na correlação entre as mesmas, para cada caráter, através do programa estatístico SAS (SAS, 1987). Foi obtida uma correlação geral, considerando a média dos anos e das épocas de semeadura avaliados, e uma correlação para cada época separadamente.

A eficiência de predição foi considerada como a probabilidade de se identificar os cruzamentos com maior variabilidade genética, através das distâncias genéticas (AFD). Foram selecionados os três cruzamentos com maiores valores de AFD, o que corresponde a uma taxa de seleção de 50%. A eficiência de predição foi avaliada, em percentagem, comparando-se esses valores de AFD com os três maiores valores da variância genética aditiva encontrada para cada caráter e para cada época.

## **3.3 RESULTADOS E DISCUSSÃO**

### **3.3.1 Análise de AFLP**

Os 21 pares de combinações de *primers* geraram um total de 963 bandas, sendo 165 polimórficas, ou seja, uma média de 7,86 bandas polimórficas por par entre os quatro genótipos analisados (Tabela 3). Através de marcadores AFLP, Hartl & Seefelder (1998) encontraram uma média de 18 bandas polimórficas por combinação de *EcoRI/MseI* entre cultivares de lúpulo com alta similaridade genética. Com a mesma técnica, Barret & Kidwel (1998) obtiveram uma média de 14,3 bandas polimórficas por combinação de *primers*,

utilizando 16 *primers* *Pst*I/*Mse*I e *Eco*RI/*Mse*I, entre linhagens adaptadas de trigo. Maughan et al. (1996) estudando a variação entre acessos de *Glycine max* e *Glycine soja*, através de AFLP, detectaram uma média de 18 bandas polimórficas por combinação de *primer*. Comparado com outros tipos de marcadores moleculares em espécies autógamas, o AFLP tem demonstrado uma maior eficiência em termos de efeito *multiplex*, ou seja, o número de locos analisados simultaneamente em um gel (Powell et al., 1996). A menor taxa de polimorfismo encontrada no presente estudo pode ser explicada pelo fato de que foram utilizados um grupo pequeno (quatro genótipos) de soja cultivada.

Tabela 3. Número total de bandas e número de bandas polimórficas obtidas com os 21 pares de *primers*.

<i>Eco</i> RI	<i>Mse</i> I	Nº total de bandas	Bandas polimórficas
AAC	CTG	50	16
AAC	CTC	37	5
AAC	CTA	62	14
AAC	CAC	42	4
AAC	CAG	34	3
AAC	CAA	34	4
AAC	CAT	86	9
AAG	CTG	45	12
AAG	CTC	48	10
AAG	CTA	65	7
AAG	CTT	72	13
AAG	CAC	40	5
AAG	CAA	34	8
AAG	CAT	42	5
ACT	CTG	32	8
ACT	CTA	44	9
ACT	CTT	37	5
ACT	CAC	27	5
ACT	CAG	41	6
ACT	CTC	26	4
ACT	CAT	65	13
<b>Total</b>		<b>963</b>	<b>165</b>

As maiores taxas de polimorfismos foram detectadas pelas combinações: *EcoRI-AAC/Msel-CTG*, *EcoRI-AAC/Msel-CTA*, *EcoRI-AAG/Msel-CTT* e *EcoRI-ACT/Msel-CAT*, isto é, 16, 14, 13 e 13, respectivamente (Tabela 3). Considerando o pequeno número de genótipos envolvidos nesta análise estas combinações de *primers* podem ser sugeridas para trabalhos futuros de estudos de diversidade genética em soja.

### 3.3.2 Estimativas das distâncias genéticas

As distâncias genéticas, estimadas através dos marcadores AFLP (AFD) revelaram uma boa amplitude de variação (0,37 a 0,72), apesar do pequeno número de populações avaliadas (Tabela 4). Utilizando RFLP, Kisha et al. (1997) obtiveram distâncias variando de 0,06 a 0,54, mas com amplitudes de 0,30 entre os genótipos de cada um dos três diferentes grupos de populações de soja analisados. Em outro trabalho, as distâncias medidas por RFLP em cinco populações de soja variaram de 0,25 a 0,43. A amplitude obtida foi considerada, pelos autores, adequada para se comparar as distâncias genéticas com os valores encontrados da variância genética (Manjarrez-Sandoval et al., 1997). Os resultados de AFD, neste trabalho, indicaram que os cruzamentos BR 85-29009 x Ocepar 8 e BR-13 x Ocepar 8 como os mais divergentes. As distâncias genéticas (DG) entre os genitores, calculadas a partir do coeficiente de parentesco ( $f$ ), variaram de 0,50 a 0,90 (Tabela 4). No entanto, a maior DG foi encontrada no cruzamento BR 85-29009 x BR-13, distintamente dos valores encontrados por marcadores, sugerindo uma baixa correlação entre essas medidas.

Tabela 4. Distâncias genéticas por marcadores AFLP (AFD) e pelo coeficiente de parentesco (DG) para as seis populações de soja.

População	AFD	DG
1	0,49	0,70
2	0,37	0,90
3	0,72	0,70
4	0,51	0,89
5	0,60	0,50
6	0,62	0,89

### 3.3.3 Correlações entre variância genética aditiva (D) e distâncias genéticas (AFD e DG)

A correlação entre AFD e DG foi negativa ( $r=-0,39$ ). Outros trabalhos com soja, embora salientando que diferentes genótipos foram utilizados em cada experimento, não evidenciaram relação consistente entre DG e distâncias obtidas por marcadores (Helms et al., 1997; Abdelnoor et al., 1995). Por outro lado, Kisha et al. (1998) obtiveram valores significativos para as correlações com DG entre e dentro de grupos de cultivares ancestrais e elites dessa espécie. Keim et al. (1992), comparando essas medidas entre linhagens exóticas e cultivares modernas, encontraram uma correlação considerada moderada ( $r=0,54$ ), pois não se tem conhecimento dos genitores das linhagens exóticas; no entanto, esses valores aumentaram ( $r=0,73$ ) quando foram comparadas apenas cultivares. Embora o número pequeno de populações utilizado no presente trabalho possa ter tido um papel importante na definição da natureza negativa da correlação, recentemente, 100 cultivares, componentes do germoplasma de soja brasileiro, foram avaliados com marcadores AFLP e também foi encontrada uma baixa correlação ( $r=0,12$ ) entre AFS e f (Bonato et al., 2000a).

Tabela 5. Estimativas da variância genética aditiva (D: média de quatro anos) para sete caracteres em quatro épocas de semeadura de seis populações de soja<sup>1</sup>.

Época	População	Produção de grãos	Número de nós	Dias para floração	Altura da planta na floração	Dias para maturação	Altura da planta na maturação	Peso 100 sementes
Setembro	1	49,91	0,91	6,96	13,69	30,08	35,86	2,81
	2	153,32	1,49	20,52	26,13	102,98	63,07	2,68
	3	146,82	1,78	16,04	27,24	121,59	79,11	1,07
	4	32,24	0,63	10,35	9,88	60,10	22,88	2,41
	5	67,21	0,85	8,63	9,58	80,15	28,87	1,54
	6	126,24	1,07	21,85	26,33	123,68	62,90	2,26
Outubro	1	69,09	1,08	9,05	38,43	35,32	79,56	2,62
	2	211,13	1,89	25,64	61,60	60,53	127,59	3,65
	3	390,25	2,06	41,97	116,08	105,48	159,94	2,21
	4	97,44	1,33	10,26	27,86	65,04	77,13	3,57
	5	187,04	1,21	11,23	46,51	87,91	68,51	1,92
	6	216,84	2,17	16,46	50,14	87,34	130,05	3,22
Novembro	1	66,44	0,62	3,44	40,86	23,24	67,15	1,96
	2	76,54	0,83	12,17	59,46	37,23	107,88	2,17
	3	118,27	2,30	21,92	87,56	60,20	140,34	2,31
	4	43,39	0,78	9,63	38,25	49,43	52,68	1,77
	5	85,51	1,17	11,22	27,32	57,03	72,70	2,10
	6	90,66	1,32	13,59	31,38	52,11	76,43	3,04
Dezembro	1	40,78	0,64	2,10	28,10	12,91	43,62	1,99
	2	58,21	0,86	7,42	47,03	24,13	74,58	2,10
	3	60,15	1,73	12,98	60,87	45,61	84,52	1,59
	4	26,40	1,09	7,77	29,56	21,71	60,41	2,17
	5	44,96	1,25	9,24	40,12	31,72	60,47	1,46
	6	53,41	1,34	9,50	33,38	31,35	68,04	2,84

<sup>1</sup> Significativo ao nível de 1% de probabilidade, pelo teste t.

Fonte: Dados originais obtidos por J.F.F. Toledo (Embrapa Soja: dados não publicados)

As estimativas da variância genética aditiva para os sete caracteres, na média dos anos de 1991 a 1995, em quatro épocas de semeadura, são apresentadas na Tabela 5 e as estimativas individuais por ano estão apresentadas no Anexo A. As correlações entre DG ou AFD e D, considerando a média dos diferentes anos, são mostradas na Tabela 6. Observou-se que DG e D, de maneira geral, não apresentaram correlações significativas. Com exceção do carácter peso de 100 sementes nos meses de outubro e dezembro e número de dias para floração em setembro. Desta forma, DG não se revelou um bom preditor da variância genética aditiva das populações. Provavelmente

a baixa correlação está condicionada ao fato dos genótipos em questão não estarem de acordo com as premissas para que o coeficiente de parentesco seja eficientemente estimado, ou seja, ausência de parentesco entre as linhagens ancestrais que se situam na base da árvore genealógica e, em um dado cruzamento simples, cada genitor contribui com 50% dos alelos para sua progênie (Carter et al., 1993; Vello et al., 1988). Ou ainda, devido ao pequeno número de genitores utilizados, um possível erro na genealogia pode ocasionar uma distorção considerável nos resultados de DG.

Em contrapartida, houve correlações significativas e mais altas entre AFD e D para quase todos os caracteres estudados neste trabalho na média dos quatro anos e épocas. O caráter que apresentou maior correlação geral entre AFD e D foi o de número de nós ( $r=0,50^{**}$ ). Os caracteres peso de 100 sementes e altura de plantas na floração não apresentaram valores significativos (Tabela 6). Ao se analisar desta forma, tida como geral, os resultados da correlação encontrados foram de moderados a baixos. As estimativas por marcadores AFLP são consideradas uma amostra representativa, porém aleatória do genoma. Deste forma, não se sabe quais as associações de ligação existentes entre marcadores e genes que controlam características agronômicas (Kisha et al., 1997). Principalmente para o caráter produção de grãos, muitos genes, possivelmente, encontram-se esparsos pelo genoma e há possivelmente muitos marcadores deste estudo que não estão ligados a estes genes de produção, influenciando as respostas da correlação.

Kisha et al. (1997) encontraram correlações significativas entre marcadores RFLP e as variâncias das populações de soja para altura de plantas e número de dias para maturação em diferentes grupos de populações, mas não obtiveram resultados significativos em outras populações avaliadas em diferentes anos. Manjarrez-Sandoval et al. (1997), avaliando cinco populações de soja, conseguiram identificar eficientemente a população com maiores variâncias através de marcadores RFLP, embora estes tenham

tido um menor poder preditivo que a DG. Segundo os autores este fato pode estar relacionado com o baixo número de polimorfismos detectados por esta técnica em genótipos com estreita base genética. Com o intuito de minimizar este problema, Burkhamer et al. (1998) utilizaram marcadores AFLP para estimar as distâncias genéticas entre genótipos de trigo e obtiveram correlações moderadas embora não significativas, entre essa medida e a variância genética de vários caracteres de interesse. Da mesma forma, Bohn et al. (1999), utilizando a mesma técnica, além de RFLP, não encontraram associações significativas com estimativas da variância genética em sete caracteres analisados em trigo, dentre eles altura das plantas na maturação, peso de 100 sementes e produção de grãos. Um dos motivos citados para a falta de associação entre as medidas seria o alto erro padrão das estimativas da variância genética, provavelmente relacionado com o baixo número de linhagens avaliadas (22). Neste trabalho, foi utilizado um número bem maior (100), proporcionando, provavelmente, uma melhor estimativa de D.

Correlações mais expressivas foram observadas para épocas específicas. A análise das correlações por épocas sugere a influência deste fator na resposta diferencial das plantas em relação à variância genética aditiva. Para produção de grãos, altura de plantas na floração e número de dias para maturação, a maior correlação ocorreu no mês de outubro, enquanto que, para número de nós, a correlação atingiu índices significativos em novembro e em dezembro. Para número de dias para floração, a correlação foi crescente de setembro a dezembro, enquanto que, para altura de plantas na maturação, as correlações mantiveram-se próximas entre os meses, com um pequeno acréscimo em setembro. Essas respostas diferenciais encontradas por épocas, provavelmente, estão relacionadas com a expressão dos genes ligados a cada caráter. Desta forma, sugere-se que as melhores respostas encontram-se nos meses em que há melhor expressão desses genes específicos. Entretanto, não existem trabalhos que comparam separadamente este fator.



Tabela 6. Correlações entre distância genética (DG) e variância genética aditiva (D) e entre distância genética (AFD) e a variância genética aditiva (D), com e sem 'outliers', para sete caracteres, em quatro épocas de semeadura, entre seis populações de soja.

Caráter	Época	DG x D	AFD X D com outliers	AFD X D sem outliers
Produção de grãos	geral	0,01	0,21*	0,43**
	setembro	0,25	0,20	0,74*
	outubro	-0,02	0,41*	0,70**
	novembro	-0,19	0,37	0,52**
	dezembro	0,03	0,11	0,29
Número de nós	geral	-0,04	0,50**	0,64**
	setembro	0,11	0,10	0,44
	outubro	0,35	0,35	0,62**
	novembro	-0,22	0,65**	0,72**
	dezembro	-0,22	0,70**	0,75**
Dias para floração	geral	0,14	0,35**	0,56**
	setembro	0,69**	0,20	0,79**
	outubro	0,22	0,29	0,49*
	novembro	0,06	0,54**	0,84**
	dezembro	-0,16	0,69**	0,75**
Altura da planta na floração	geral	0,09	0,16	0,33**
	setembro	0,47	0,15	0,56
	outubro	0,04	0,24	0,35
	novembro	0,10	0,08	0,27
	dezembro	0,07	0,19	0,59**
Dias para maturação	geral	-0,05	0,42**	0,58**
	setembro	0,23	0,56*	0,93**
	outubro	-0,16	0,68**	0,76**
	novembro	-0,15	0,55**	0,60**
	dezembro	-0,08	0,52**	0,73**
Altura da planta na maturação	geral	0,17	0,22*	0,41**
	setembro	0,21	0,35	0,49
	outubro	0,30	0,21	0,43*
	novembro	0,08	0,26	0,53**
	dezembro	0,17	0,30	0,57**
Peso 100 sementes	geral	0,32**	-0,11	-0,03
	setembro	0,24	-0,44	-0,44
	outubro	0,56**	-0,38	-0,26
	novembro	0,11	0,33	0,53*
	dezembro	0,49**	-0,12	-0,02

\* significativo ao nível de 5% de probabilidade.

\*\* significativo ao nível de 1% de probabilidade.

Notadamente verificou-se, em todas os caracteres, que o cruzamento BR 85-29009 x BR-13 influenciou significativamente as magnitudes de correlação entre AFD e D. Foi possível detectar que esse cruzamento se inseriu desproporcionalmente aos demais em relação às duas variáveis em questão, como demonstrado na Figura 1, para produção de grãos, com um símbolo ◆, sugerindo na realidade uma super-estimação da variância em relação à baixa distância genética obtida por marcadores. Esses valores discrepantes se mantiveram em todas as épocas, sendo o cruzamento considerado, neste caso, um “outlier”, mas houve uma tendência desse efeito ser mais pronunciado em setembro. Esses fatores sugerem a possibilidade da linhagem BR 85-29009 apresentar genes maiores, relacionados ao período juvenil longo. Esta suposição vem sendo investigada pelos geneticistas (Embrapa Soja, Comunicação pessoal). Aliado a isto, tem-se o fato da cultivar BR-13 caracterizar-se por ser muito sensível ao fotoperíodo, diferenciando-se extremamente quanto a este fator da cultivar BR 85-29009. Provavelmente, esses fatores causaram uma alta e discrepante variância nesse cruzamento.

A correlação geral e as correlações individuais por, com base na média dos anos, aumentaram para todas as épocas com a eliminação desse cruzamento “outlier” (Tabela 6). Observou-se nitidamente o acréscimo expressivo em setembro, onde o efeito da discrepância era maior. O efeito da resposta diferencial em alguns anos, provavelmente esteja relacionado a fatores ambientais não controlados, como a temperatura e a umidade. Kisha et al. (1997) também encontraram uma variância genética para produção de grãos desproporcional às demais, em uma população de soja e a consideraram um “outlier”, por ocorrer o mesmo efeito em dois anos consecutivos. Retirando esta população das análises, as correlações entre a variância genética e a distância genética por RFLP, aumentaram nos dois anos testados. Segundo os autores pode ser esperado que determinados genitores de cruzamentos específicos difiram em muitos genes afetando um caráter e que isto não seja detectado pelos marcadores genéticos.

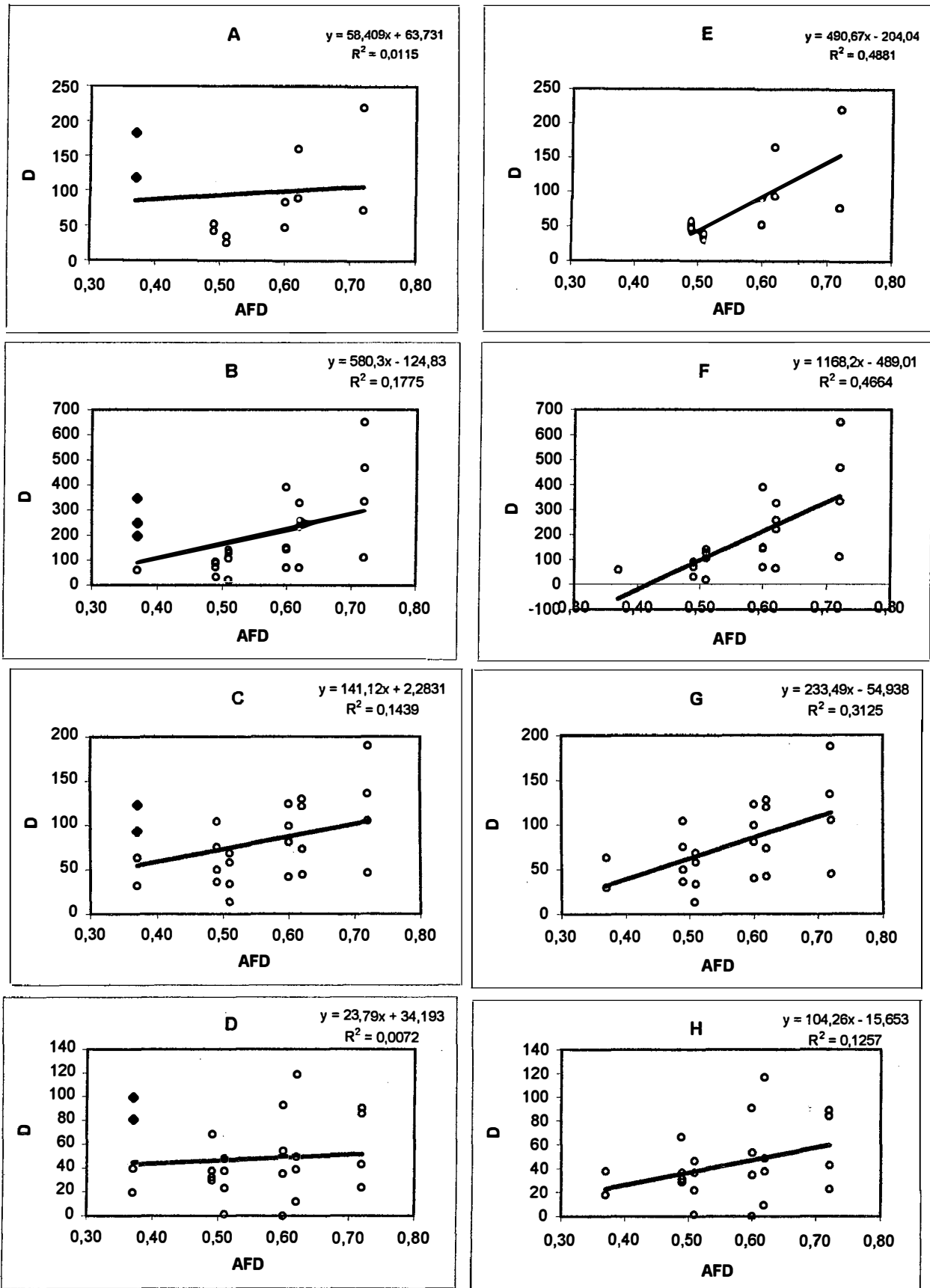


Figura 1. Relação entre variância genética aditiva (D) e distância genética obtida por AFLP (AFD) para o caráter produção de grãos em quatro anos, em quatro épocas de semeadura: A, B, C e D referem-se à setembro, outubro, novembro e dezembro, respectivamente, com os dados "outliers" (◆). E, F, G e H referem-se às mesmas épocas, sem os "outliers".

### 3.3.4 Eficiência de predição baseada em AFD

Considerando todas as populações avaliadas, a predição apresentou em quase todas as situações, em relação a caracteres e épocas, uma eficiência superior a 50%, com exceção de peso de 100 sementes (33%), atingindo valores máximos em número de dias para maturação (100%), como mostra a Tabela 7. Para produção de grãos, os mais altos índices de eficiência foram encontrados nos meses de outubro e de novembro (83,3%), sendo esses os meses mais recomendados para semeadura de soja no país. Desta forma, se um melhorista utilizar esta informação de AFD para reduzir o número de cruzamentos e, conseqüentemente, o número de populações a serem avançadas em um programa de melhoramento, ele teria, a partir desses resultados, uma taxa de sucesso de 83% durante as épocas normais de semeadura. Esses resultados estão de acordo com a conclusão de Kisha et al. (1997) de que a estimativa da distância genética pode ser útil na identificação de cruzamentos que resultem em populações com maiores variâncias.

Este trabalho demonstra a viabilidade da utilização de marcadores AFLP na predição da variância genética aditiva em programas de melhoramento de soja, como também, pode substituir com vantagens não somente o coeficiente de parentesco, como as análises com os marcadores RFLP e RAPD.

Tabela 7. Eficiência de predição (%), baseada em AFD, para sete caracteres de soja em quatro épocas de semeadura.

<b>Caráter</b>	<b>Setembro</b>	<b>Outubro</b>	<b>Novembro</b>	<b>Dezembro</b>
Produção de grãos	66,6	83,3	83,3	58,3
Número de nós	66,6	66,6	74,9	91,6
Dias para floração	66,6	58,3	83,3	83,3
Altura da planta na floração	66,6	66,6	66,6	58,3
Dias para maturação	66,6	100,0	83,3	100,0
Altura da planta na maturação	66,6	66,6	66,6	66,6
Peso de 100 sementes	33,3	41,6	58,3	33,3

## 4 SIMILARIDADE GENÉTICA ENTRE CULTIVARES DE SOJA LIBERADAS NO BRASIL, AVALIADA ATRAVÉS DE MARCADORES AFLP

### RESUMO

Com o objetivo de avaliar a similaridade genética entre cultivares de soja foram analisadas 317 cultivares liberadas no Brasil no período de 1962 a 1998, utilizando a técnica de AFLP, com seis combinações de *primers EcoRI/MseI*. Em seguida foram estimados os coeficientes de similaridade genética de acordo com o método de Nei & Li (1979), bem como feito o agrupamento das cultivares pelo método UPGMA. Foram calculados também os coeficientes de parentesco (f) entre 100 cultivares, liberadas entre 1984 e 1998, os quais foram correlacionados com os coeficientes de similaridade obtidos pelos marcadores. A análise de AFLP revelou a ocorrência de aproximadamente 394 bandas, sendo 78 delas polimórficas, com um número médio de 13 bandas por combinação de *primer*. Os coeficientes de similaridade genética (AFS) variaram de 0,17 a 0,97, com média de 0,61. Observou-se uma frequência maior (56,8%) de coeficientes acima de 0,60, enquanto que apenas 9,7% foram iguais ou menores que 0,50, refletindo o alto grau de similaridade genética entre as cultivares. Apesar disso, os coeficientes de similaridade (AFS) permaneceram constantes ao longo de três décadas, indicando que o grau de similaridade genética entre as cultivares não tem aumentado. A análise *bootstrap* indicou que as 78 marcas polimórficas foram suficientes para avaliar a similaridade genética entre as 317 cultivares de soja, devido ao CV relativamente baixo (7,7%). A interpretação do dendrograma foi dificultada pelo grande número de

cultivares utilizadas. Mesmo assim foi possível detectar agrupamentos de cultivares formados de acordo com o esperado através da genealogia dos mesmos. No entanto, verificou-se em alguns casos que cultivares com altos coeficientes de similaridade genética alocaram-se em diferentes grupos. Um novo dendrograma, composto por apenas 63 cultivares, permitiu uma melhor interpretação dos agrupamentos. A correlação cofenética foi de 0,60 e 0,70, respectivamente, nos dendrogramas compostos por 317 e 63 cultivares. Os coeficientes de parentesco entre 100 cultivares variaram de 0 a 1, com média de 0,21, sendo esta muito próxima de outros dados relatados para cultivares brasileiras e norte-americanas, e considerada alta. Porém, não detectou-se correlação significativa entre os coeficientes  $f$  e AFS ( $r=0,12$ ). Os marcadores AFLP evidenciaram, portanto, a alta similaridade genética entre as cultivares de soja avaliadas. Apesar disso, têm sido obtidos progressos substanciais com o melhoramento genético.

## SUMMARY

The genetic similarity among soybean cultivars was studied using the AFLP technique with six combinations of *EcoRI/MseI* primers on 317 cultivars released in Brazil from 1962 to 1998. The genetic similarity coefficients was estimated according to Nei and Li (1979) and the cultivars were clustered by the UPGMA method. The parentage coefficients ( $f$ ) among 100 cultivars released between 1984 and 1998 were also calculated and correlated with the similarity coefficients obtained by the markers. AFLP analysis showed approximately 394 bands, 78 of which were polymorphic, with a mean number of 13 bands per primer combination. The genetic similarity coefficients (AFS) varied from 0.17 to 0.97, with a mean of 0.61. Coefficients above 0.60 predominated (56.8% of the cases), while only 9.7% were equal or less than 0.50, reflecting the high degree of genetic similarity among the cultivars. The similarity coefficients (AFS), however, have remained constant throughout the last three decades, indicating

that the degree of genetic similarity among the cultivars has not increased. Bootstrap analysis indicated that the 78 polymorphic markers were sufficient to assess the genetic similarity among the 317 soybean cultivars, due to the relatively low coefficient of variation (7.7%). The dendrogram interpretation was hindered by the large number of cultivars used, but it was possible to detect groups of cultivars formed as expected by their genealogy. However, some cases were found where cultivars with high genetic similarity coefficients were allocated in different groups. Another dendrogram, composed of only 63 cultivars, allowed a better interpretation of the groups. The cophenetic correlation was 0.60 and 0.70, respectively, in the 317 and 63 cultivar dendrograms. The parentage coefficients among 100 cultivars varied from 0 to 1, with a mean of 0.21, which is very close to other reported data for Brazilian and North American cultivars, and was considered high. However, no significant correlation was detected among the  $f$  and AFS coefficients ( $r=0.12$ ). The AFLP markers indicated the high genetic relationship among the cultivars assessed, but in spite this substantial gains have been obtained in soybean breeding.

## 4.1 INTRODUÇÃO

O conhecimento do grau de similaridade entre os diferentes genótipos é um passo de fundamental importância na eficiência dos programas de melhoramento genético de plantas. Essas informações são úteis na organização de coleções de germoplasma, na identificação de grupos heteróticos e na seleção de parentais para cruzamento.

O coeficiente de parentesco (Malécot, 1947), baseado nas informações sobre a genealogia entre os genótipos, tem sido empregado para a obtenção das estimativas de similaridade genética e para o estudo da estrutura genética do germoplasma cultivado de soja. As análises de genealogia do germoplasma americano demonstraram uma substancial relação de proximidade genética entre as mesmas (Delannay et al., 1983; Specht & Williams, 1984; Carter et al., 1993). Estudos mais recentes revelaram que as cultivares do norte e do sul dos Estados Unidos apresentam bases genéticas contrastantes (Gizlice et al., 1994; 1996). Esses estudos também evidenciaram que a diversidade genética do germoplasma de soja norte americano, como um todo, estava sendo reduzido nos últimos 50 anos (Gizlice et al., 1993). As análises de genealogia também revelaram uma base genética estreita no germoplasma brasileiro de soja (Vello et al., 1988). No entanto, a utilização do coeficiente de Malécot depende da disponibilidade e precisão das informações acerca da genealogia dos genótipos.

A diversidade genética entre indivíduos pode ser estimada diretamente através de marcadores bioquímicos e moleculares. Em soja, a utilização de marcadores bioquímicos, tais como as isoenzimas, esbarrou no baixo grau de polimorfismo (Cox et al., 1985). Este problema vem sendo contornado pelo emprego de marcadores moleculares. Assim, Sneller et al. (1997) separaram claramente as cultivares elite americanas do norte e do sul do país empregando marcadores RFLP. Essa mesma técnica foi utilizada no estudo de germoplasma



exótico de soja e permitiu a identificação de diferentes *pools* gênicos (Keim et al., 1989, 1992; Kisha et al., 1998). Trabalhos similares foram conduzidos empregando-se outros tipos de marcadores moleculares tais como o RAPD (Abdelnoor et al., 1995; Thompson & Nelson, 1998), microssatélites (Akkaya et al., 1992; Rongwen et al., 1994; Powell et al., 1996; Diwan & Cregan, 1997) e AFLP (Maughan et al., 1996; Machado, 1999; Zhu et al., 1999).

Um estudo comparativo da performance dos diferentes tipos de marcadores em análise genética de soja revelou que os marcadores microssatélites apresentam maior grau de polimorfismo para a discriminação entre genótipos, enquanto que os marcadores AFLP apresentam maior grau *multiplex*, ou seja, o número de locos analisados simultaneamente em um gel (Powell et al., 1996), além de ser uma eficiente ferramenta na distinção de genótipos altamente relacionados (Hartl & Seefelder, 1998).

Apesar desses estudos de diversidade com marcadores moleculares terem sido conduzidos com diferentes tipos de marcadores e empregado diferentes genótipos, os resultados têm confirmado, como era de se esperar, a presença de maior diversidade genética no germoplasma exótico (Maughan et al., 1996; Machado, 1999; Zhu et al., 1999). No entanto, a correlação entre estimativas de diversidade genética, obtida por marcadores e pelo coeficiente de parentesco em material cultivado, tem apresentado resultados distintos. Em trabalhos com RFLP, a magnitude dessas correlações variaram de 0,54 a 0,91 (Keim et al., 1992; Manjarrez-Sandoval et al., 1997; Kisha et al., 1998). Através de marcadores RAPD, Helms et al. (1997) não obtiveram aparente relação entre os dois tipos de estimativas. Abdelnoor et al. (1995) indicaram alguns casos de discrepância entre a distância genética entre cultivares brasileiras, analisadas por RAPD e pelo coeficiente de parentesco, embora, de maneira geral, tenham encontrado concordância entre os dados. É possível que essas discrepâncias estejam refletindo os erros relacionados com as medidas baseadas na genealogia, uma vez que as medidas através de marcadores moleculares se constituem em uma medida direta da distância genética.

Até o presente momento, nenhum estudo abrangente de análise do germoplasma brasileiro de soja com uso de marcadores moleculares foi conduzido. Com o objetivo de avaliar a metodologia, Abdelnoor et al. (1995) avaliaram a diversidade genética entre 38 cultivares brasileiras, através de RAPD, e encontraram cinco subgrupos distintos. Entretanto, a aplicação desses resultados no programa de melhoramento genético esbarra no número limitado de genótipos empregados no estudo.

O objetivo desse trabalho foi avaliar a similaridade genética entre as cultivares de soja liberadas no Brasil no período de 1962 a 1998, utilizando marcadores AFLP.

## **4.2 MATERIAL E MÉTODOS**

### **4.2.1 Material genético**

O presente trabalho foi realizado no Centro Nacional de Pesquisa de Soja (Embrapa Soja), localizado em Londrina, Estado do Paraná, Brasil. Foram analisadas 317 cultivares de soja, liberadas no Brasil no período de 1962 a 1998. As sementes foram obtidas no Banco Ativo de Germoplasma da Embrapa Soja. As cultivares estão listadas na Tabela 1.

Tabela 1- Cultivares de soja liberadas no Brasil no período de 1962 a 1998 e utilizadas para análise com marcadores AFLP.

nº	Cultivar	nº	Cultivar	nº	Cultivar
1	BR-36 *	55	FT-11 (Alvorada) *	109	CAC-1
2	BR-37	56	Cep 12 (Cambará) *	110	IAC-13
3	GO/BR-26 (Tocantins)	57	FT-6 (Veneza)	111	RS 6 (Guassupi) *
4	MS/BR-19 (Pequi)	58	FT-7 (Tarobá)	112	Coker 136
5	BR-30 *	59	FT-8 (Araucária) *	113	UFV-2
6	MS/BR-34 (EMPAER 10) *	60	FT-2	114	IAC-3
7	BR-23 *	61	FT-3	115	São Carlos *
8	BR-38	62	FT-1	116	Ivorá
9	BR-28 (Seridó)	63	Emgopa-308 *	117	IAC-7
10	MS/BR-20 (Ipê)	64	BR/Emgopa-312 (Potiguar)	118	IPAGRO-21
11	Emgopa-301	65	Emgopa-311	119	IAC-12
12	MS/BR-21 (Buriti) *	66	Emgopa-313 (Ananguera)	120	Embrapa 20 (Doko RC) *
13	MS/BR-17 (São Gabriel)	67	Emgopa-309	121	IAS 4
14	BA/BR-31	68	Emgopa-310	122	UFV-4
15	MS/BR-18 (Guavira) *	69	FT-15	123	Ocepar 2 (Iapó)
16	MG/BR-22 (Garimpo) *	70	FT-Manaca	124	Ocepar 7 (Brilhante) *
17	BR-5	71	FT-16	125	Paraná
18	BR-14 (Modelo) *	72	FT-10 (Princesa)	126	Ocepar 14 *
19	BR-10 (Teresina)	73	FT-Eureka	127	Vila Rica
20	BR-4	74	FT-Guaíra *	128	Ocepar 11 *
21	BR-12	75	FT-14 (Piracema)	129	UFV-9 (Sucupira)
22	BR-24 *	76	FT-Jatoba	130	Ocepar 10 *
23	BR-15 (Mato Grosso) *	77	FT-Canarana *	131	IAS 2
24	BR-1	78	FT-Estrela	132	Embrapa 3 *
25	BR-13 (Maravilha) *	79	FT-9 (Inae)	133	MG/BR-42 (Kage) *
26	BR-2 (Vagem clara)	80	FT-18 (Xavante)	134	Ocepar 5 (Piquiri)
27	BR-11 (Carajás)	81	FT-17 (Bandeirantes) *	135	Lancer
28	BR-6 (Nova Bragg)	82	Cristalina	136	FT-440
29	Bossier	83	FT-19 (Macacha)	137	Pampeira *
30	BR-3	84	FT-Abyara *	138	Embrapa 1 (IAS 5 RC)
31	Andrews	85	FT-13 (Aliança)	139	Missões
32	BR-8 (Pelotas)	86	FT-20 (Jaú)	140	RS 5 (Esmeralda) *
33	Emgopa-305 (Caraíba) *	87	FT-Bahia *	141	Prata *
34	BR-29 (Londrina)	88	CAC/BR-43 *	142	IAC-16
35	BR-16	89	IAC-5	143	Ocepar 9 *
36	BR-9 (Savana)	90	FT-Maracaju	144	Embrapa 19
37	SP/BR-41	91	IAC-4	145	Embrapa 2
38	Doko	92	Davis	146	IPAGRO-20
39	Dourados	93	IAC-2	147	Nova IAC-7 *
40	BR-7	94	IAC-6	148	UFV-8 (Monte Rico)
41	Emgopa-302	95	FT-Seriema	149	IAC-11 *
42	Cep 20 (Guajuvira) *	96	IAS 1	150	Paranaíba
43	MS/BR-39 (Chapadão) *	97	IAC-Foscarin 31	151	Hood 75
44	Emgopa-304 (Campeira)	98	Pérola	152	IAS 5 (Vagem Escura)
45	Campos Gerais	99	Hardee	153	Cajeme
46	GO/BR-33 (Javaes)	100	BR-32 *	154	UFV-14
47	Cep 10	101	UFV-5	155	Santa Rosa
48	Cobb	102	UFV-3	156	UFV/ITM-1
49	Bragg	103	UFV-Araguaia	157	Viçoja
50	Cep 16 (Timbó) *	104	IAC-100	158	UFV-15 (Uberlândia) *
51	FT-12 (Nissei)	105	Industrial	159	IAC-9
52	BR-40 (Itiquira) *	106	Mineira	160	Ocepar 3 (Primavera)
53	FT-4	107	IAS 3 (Delta)	161	UFV-10 (Uberaba) *
54	FT-5 (Formosa)	108	IAC-14	162	IAC-8

Tabela 1- Cultivares de soja liberadas no Brasil no período de 1962 a 1998 e utilizadas para análise com marcadores AFLP (Continuação).

nº	Cultivar	nº	Cultivar	nº	Cultivar
163	Embrapa 5	218	Soberana	273	IAC-19
164	MT/BR-45 (Paiaguás) *	219	Ocepar 15 (Paracatu)	274	BRAS-60 (Celeste) *
165	Paranagoiana	220	Embrapa 62 *	275	BRS 157
166	Ocepar 8 *	221	KI-S 702	276	BRS 158 (Milena) *
167	IAC-1	222	Nobre *	277	MA/BR-64 (Parnaíba)
168	Pelicano	223	Embrapa 58 *	278	MG/BR-56 (Confiança) *
169	Sertaneja *	224	FT-Saray	279	MA/BR-65 (Sambaíba)
170	Ivaí	225	Cep 26 (Umbú)	280	BRS 137 *
171	Planalto	226	FT- Morena *	281	GO/BRS-160 (Goiatuba) *
172	Hill	227	Embrapa 46	282	Emgopa-316 (Rio Verde)
173	UFV-6 (Rio Doce) *	228	Ocepar 16	283	Embrapa 64 (Ponta Porã)
174	RS 7 (Jacui) *	229	Embrapa 59	284	BRS 133 *
175	IAC-10	230	MT/BR-50 (Parecis)	285	Ocepar 19 (Cotia)
176	J-200	231	MT/BR-47(Canário)	286	GO/BRS-161 (Catalão)
177	Tiaraju	232	FT- Iramaia	287	MG/BR-58 (Segurança)
178	IAS 5 (Vagem Clara)	233	FT- Líder *	288	BRS 134 *
179	IAC-15	234	MT/BR-52 (Curió)	289	Coodetec-203
180	UFV-7 (Juparana) *	235	Embrapa 26	290	MS/BR-57 (Lambari) *
181	União	236	Ocepar 12 *	291	MS/BRS-168 (Piapara)
182	Ocepar 13	237	Embrapa 32 (Itaqui) *	292	MT/BR-55 (Uirapuru)
183	UFV-1	238	Ocepar 6 *	293	MS/BRS-171 (C. Grande) *
184	Teresina RCH	239	Rainha	294	Cariri RCH
185	Ocepar 4 (Iguaçu)	240	MS/BRS-172 (Tuiuiú) *	295	MS/BRS-169 (Bacuri)
186	São Luiz	241	Embrapa 60 *	296	BRS 136
187	BR/IAC-21	242	KI-S 801	297	Coodetec-202 *
188	Vitória	243	BRS 154	298	BRS 135 *
189	Tropical	244	MS/BRS-167 (Carandá) *	299	GO/BRS-162 (Bela Vista) *
190	Embrapa 48 *	245	MG/BRS-66 (Liderança) *	300	BRS 156
191	Embrapa 9 (Bays) *	246	FT-45263	301	Serido RCH
192	Embrapa 47	247	Embrapa 30	302	Emgopa-315 (Rio Vermelho) *
193	MT/BR-53 (Tucano)	248	Ocepar 17	303	MT/BRS- 63 (Pintado)
194	Timbira	249	BRS 153	304	Emgopa-307
195	BR-35 (Rio Balsas) *	250	MS/BRS-173 (Piraputanga)	305	KI-S 602
196	IAC-18	251	FT-Cristalina RCH *	306	FT-Cometa
197	Invicta *	252	KI-S 602 RCH	307	Emgopa-306 (Chapada)
198	Década	253	Embrapa 33 (Cariri RC) *	308	MS/BR-61 (Surubi) *
199	FT-102	254	BRS 155 *	309	BR-27 (Cariri)
200	FT-100	255	Embrapa 31 (Mina)	310	KI-S 601
201	SPS-1	256	FT-106	311	MG/BR-54 (Renascença)
202	BRS 65 ( Itapoty) *	257	Coodetec-201	312	MT/BR-49 (Pioneira) *
203	FT-103	258	MS/BR-59 (Mandi)	313	GO/BRS-163 (Jataí)
204	FT-104	259	BR/Emgopa-314 *	314	MG/BR-46 (Conquista) *
205	Embrapa 61 *	260	FT-2002	315	Emgopa-303 *
206	Embrapa 25 *	261	Embrapa 34 (Teresina RC)	316	GO/BR-25 (Aruaná) *
207	MT/BR-51 (Xingu)	262	BRS 138	317	Numbaira
208	Cristalina Antiga	263	MT/BRS-159 (Crixás)		
209	FT- Cristal *	264	MG/BRS-68 (Vencedora)		
210	IAC-17	265	Pirapó 78		
211	BRS 66 *	266	BRS 132		
212	FT-101 *	267	IAC-20		
213	RS 9 (Itaubá) *	268	Embrapa 63 (Mirador) *		
214	Embrapa 4 (BR-4 RC)	269	Ocepar 18		
215	Fepagro-RS 10 *	270	MS/BRS-166 (Apaiari)		
216	FT- Cristalina *	271	BRAS-62 (Carla)		
217	MG/BR-48 (Garimpo RCH)	272	MS/BRS-170 (Taquari)		

\* Cultivares utilizadas para cálculo do coeficiente de parentesco.

#### 4.2.2 Extração e quantificação do DNA

O DNA foi extraído a partir de uma amostra composta de tecido foliar. Para cada cultivar foi coletado aproximadamente trinta folhas obtidas de plantas individuais em estágio V2, cultivadas em casa de vegetação. As amostras foram imediatamente congeladas em nitrogênio líquido e armazenadas a -80°C. Posteriormente, foram utilizadas para a extração do DNA genômico, conforme Saghai-Marooof et al. (1984).

A concentração de DNA foi estimada em géis de agarose 0,8% (p/v) através da comparação visual da intensidade das bandas do DNA extraído com as bandas do DNA do fago lambda de concentrações conhecidas. As amostras de DNA foram então diluídas para uma concentração de aproximadamente 30 ng/μl e mantidas a -20°C.

#### 4.2.3 Análises com marcadores AFLP

Para a análise de AFLP foi utilizado o "AFLP Analyses Kit" (GIBCO-BRL), seguindo-se o protocolo descrito por Bonato et al. (2000b). Para as reações de amplificação foram selecionadas seis combinações de *primers EcoRI/MseI*, baseado na taxa de polimorfismos previamente observadas (Bonato et al., 2000b): E-AAC/M-CAT, E-AAC/M-CTA, E-AAC/M-CTC, E-AAC/M-CTG, E-AAG/M-CTT e E-ACT/M-CAT.

#### 4.2.4 Análise dos dados obtidos por AFLP e cálculo da similaridade genética

A interpretação das autoradiografias foi baseada na observação visual da presença ou ausência de uma mesma marca (banda) polimórfica para cada

genótipo analisado. A partir desses dados foi estimada a similaridade genética entre todos os possíveis pares de genótipos através do coeficiente de Nei & Li (1979), segundo a equação:

$$AFS_{ij} = [ 2N_{i,j} / ( 2N_{i,j} + N_i + N_j ) ],$$

em que  $AFS_{ij}$  representa a medida da similaridade a partir de dados de AFLP entre os genótipos  $i$  e  $j$ ,  $N_{i,j}$  é o número total de bandas comuns a  $i$  e  $j$  e  $N_i$  e  $N_j$  correspondem ao número de bandas encontradas nos genótipos  $i$  e  $j$ , respectivamente.

A matriz gerada com essas estimativas foi utilizada para a análise de agrupamento em forma de dendrograma pelo método UPGMA (Unweighted pair group method using arithmetic averages) (Rolf, 1997). A correlação cofenética entre os dados da matriz de similaridade genética e os valores cofenéticos do dendrograma foi estimada para indicar a representatividade do dendrograma em relação às estimativas originais de similaridade. Todas as análises a partir da matriz de dados binários foram realizadas com o auxílio do software NTSYS 2.02 ( Rolf, 1997).

A análise *bootstrap* (Tivang et al., 1994) foi empregada para verificar se o número de marcadores foi suficiente para caracterizar as cultivares quanto à similaridade genética. Os procedimentos desta análise de reamostragem estão descritos em Barroso (2000).

#### 4.2.5 Coeficiente de Parentesco

As estimativas do coeficiente de parentesco,  $f$  (Malécot, 1947), entre 100 cultivares de soja, liberadas no período de 1984 a 1998, foram baseadas nas respectivas genealogias e obtidas com o auxílio do programa PARENT (CIAGRI-ESALQ/USP). Foram incluídas somente 100 cultivares devido à capacidade limitada do programa PARENT, aquelas cuja genealogia completa

foi possível recuperar e, além disso, as cultivares liberadas após 1984, para não repetir o estudo realizado por Vello et al. (1988). Estas estão indicadas com asterisco (\*) na Tabela 1. Posteriormente, essas estimativas foram utilizadas no cálculo de correlação entre os coeficientes de parentesco e de similaridade genética (AFS), entre os respectivos pares de cultivares.

### 4.3 RESULTADOS E DISCUSSÃO

#### 4.3.1 Análise de AFLP

De acordo com os resultados apresentados na Tabela 2, as seis combinações de *primers* utilizadas para analisar as 317 cultivares geraram aproximadamente 394 bandas, sendo que 78 (19,8%) destas foram polimórficas e possíveis de serem analisadas com segurança para todos os genótipos. O número médio de marcadores polimórficos foi 13 por combinação de *primer*, variando entre 6 e 25, e as combinações *EcoRI-AAC/MseI-CTC* e *EcoRI-AAC/MseI-CAT* geraram os maiores níveis de polimorfismo, isto é, 18 (34,6%) e 25 (30,1%), respectivamente.

Maughan et al. (1996) detectaram, através de AFLP, 274 (36%) bandas polimórficas entre o número total de bandas observadas entre acessos de *Glycine max* e *Glycine soja*. A maioria dos polimorfismos (31%), foi encontrada entre os acessos de *Glycine soja*, enquanto que apenas 17% foram constatados em *Glycine max*. Resultado similar foi verificado por Machado (1999), que através da mesma técnica molecular, encontrou uma proporção superior de polimorfismos entre genótipos de *Glycine soja* (dados não apresentados). Em relação ao número de bandas polimórficas por combinação de *primer*, Maughan et al. (1996) obtiveram um número médio de 18, considerando acessos das duas espécies, enquanto que Machado (1999) encontrou uma média muito superior (61).

Tabela 2. Combinações de *primers* *EcoRI/Msel* para análise de AFLP e número de bandas obtidas.

<b>Primers</b>		<b>Número de bandas</b>	
<b><i>EcoRI</i></b>	<b><i>Msel</i></b>	<b>Total</b>	<b>Polimórficas</b>
AAC	CTC	52	18
AAC	CTA	68	11
AAC	CTG	55	8
AAC	CAT	83	25
AAG	ATT	86	10
ACT	CAT	50	6
<b>Total</b>		<b>394</b>	<b>78</b>

Considerando que no presente estudo foram utilizados apenas cultivares adaptadas de *Glycine max*, os níveis de polimorfismo encontrados podem ser considerados bons, pois possibilitaram a distinção entre os genótipos com base na similaridade genética. Esses resultados confirmam os resultados anteriores de que o AFLP é uma técnica molecular que detecta polimorfismos em múltiplos locos, gerando uma vasto número de marcadores reproduzíveis em um curto período de tempo (Maheswaran et al., 1997), e que se constitui uma poderosa ferramenta na distinção entre genótipos altamente relacionados (Hartl & Seefelder, 1998).

De acordo com o trabalho de Zhu et al. (1999), as combinações de *primers* que geraram mais polimorfismos foram *EcoRI-AAC/Msel-CTC* (53%) e *EcoRI-AAC/Msel-CAT* (50%). Embora este estudo tenha encontrado uma frequência maior de polimorfismo, provavelmente por se tratar de acessos adaptados e não adaptados de *Glycine max* e de *Glycine soja*, os resultados concordam com os obtidos no presente trabalho, e confirmam essas duas



combinações de *primers* como altamente informativas para o germoplasma brasileiro de soja.

#### 4.3.2 Estimativas da similaridade genética por AFLP (AFS)

O cálculo da similaridade genética entre os 317 genótipos resultou em uma matriz com 50.086 coeficientes de similaridade de Nei & Li (1979), apresentada em anexo (Anexo B). A distribuição de frequência dos referidos coeficientes encontra-se na Figura 1.

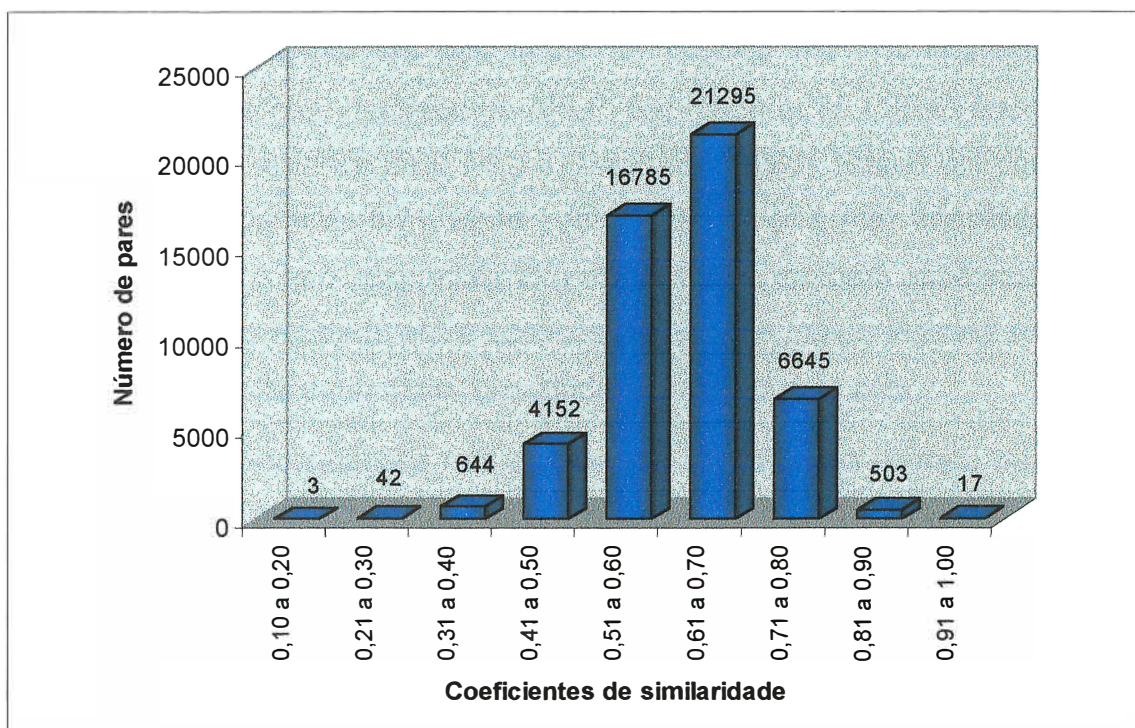


Figura 1- Distribuição de frequência de 50.086 coeficientes de similaridade genética, baseados em AFLP, entre 317 cultivares de soja.

O coeficiente médio, entre todos os genótipos, foi 0,61, variando de 0,17 a 0,97 entre os pares de cultivares Nobre / Bossier e FT-Cristal / FT-Cristalina, respectivamente. Constatou-se que 56,8% dos coeficientes estimados tiveram

valores superiores a 0,60, refletindo o alto grau de similaridade genética existente entre as cultivares incluídas neste estudo. Entretanto, 9,7% dos coeficientes foram iguais ou menores que 0,50 e podem ser explorados no caso de escolha de genitores altamente divergentes.

A alta similaridade genética entre cultivares de soja brasileiras também foi evidenciada por Abdelnoor et al. (1995), através de análises com marcadores RAPD, os quais obtiveram um coeficiente médio de 0,82, com variação de 0,69 a 1. O par mais distante geneticamente correspondeu às cultivares Tropical e UFV-6 e o mais próximo às cultivares Ocepar-9 e Paranagoiana. Outros trabalhos também demonstraram a condição de alta similaridade entre genótipos adaptados de *Glycine max*. A exemplo disso, através da análise com marcadores AFLP, Maughan et al. (1996) encontraram uma similaridade variando de 0,74 a 1. Machado (1999) obteve um coeficiente médio de similaridade de 0,71, entre acessos de *Glycine max* adaptados, e de 0,57, entre acessos não adaptados. Zhu et al. (1999), também observaram altos coeficientes de similaridade entre acessos de *Glycine max* e *Glycine soja* (0,60 a 0,94), mas ressaltaram a maior proximidade para *Glycine max*.

Uma das preocupações nas estimativas de AFS é o número de marcas necessárias para amostrar o genoma. Através da análise *bootstrap* obteve-se um coeficiente da variação (CV) médio de 7,7% com os 78 marcadores utilizados na análise de AFLP (Figura 2). Observou-se que o CV diminuiu com o aumento do tamanho amostral, indicando que a precisão das estimativas de similaridade genética aumentou com o aumento do número de locos polimórficos. A transformação do CV médio e do tamanho da amostra na escala logarítmica tornou a relação entre as duas variáveis linear. Utilizando a equação de regressão, apresentada na Figura 3, estimou-se que com 100 locos polimórficos obter-se-ia um CV médio de 6,7%, portanto um aumento pouco expressivo na precisão das estimativas.

Entre as considerações relacionadas com o número de marcadores necessários para amostrar corretamente o genoma em estudo foi sugerido que

em milho um mínimo de 100 locos devem ser utilizados (Messmer et al., 1991). Mais recentemente, Pejic et al. (1998), em uma análise comparativa da similaridade genética em milho, envolvendo RFLP, RAPD, AFLP e microssatélites, obtiveram coeficientes de variação de 5% a 10%, para os quatro tipos de marcadores, utilizando 150 marcas polimórficas. Os autores consideraram ser esse número de marcas suficiente para estimar a similaridade com boa precisão.

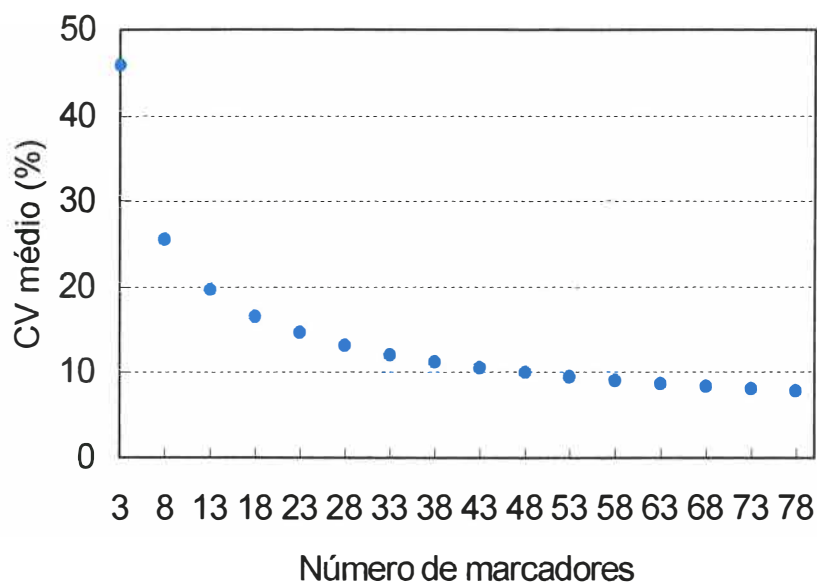


Figura 2. Coeficiente de variação médio da similaridade genética entre as 317 cultivares, de acordo com o número de locos AFLP reamostrados utilizando a técnica *bootstrap*.

$$\text{Log}(\text{CV médio})=1,898435-0,53413\times\text{log}(\text{número de marcadores})$$

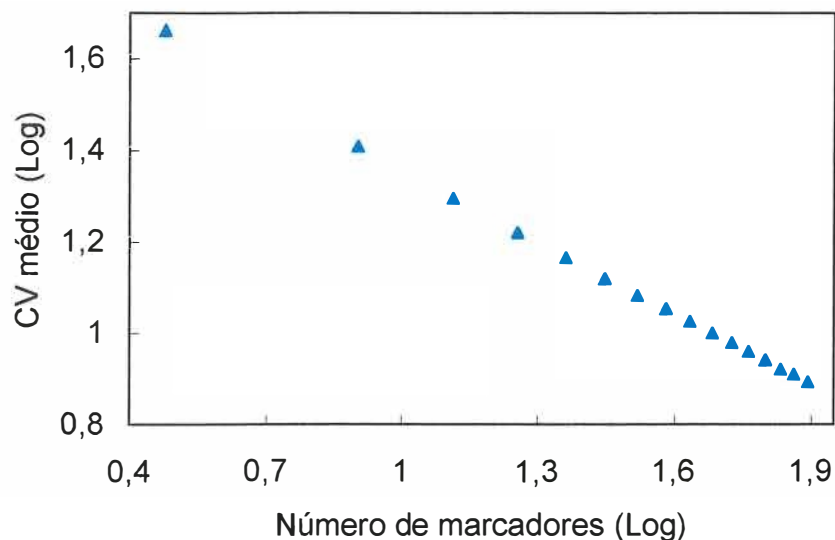


Figura 3. Equação de regressão entre o logaritmo natural do número de locos AFLP reamostrados e o logaritmo natural do coeficiente de variação médio das similaridades genéticas, utilizando a técnica *bootstrap*.

Moser & Lee (1994) sugeriram que espécies com menor grau de polimorfismo que o milho apresentam erro padrão dessas estimativas menor e, portanto, um menor número de locos é necessário. Corroborando esta hipótese, Keim et al. (1992) constataram que, em soja, o número ideal de locos que devem ser amostrados no genoma oscila entre 65 e 90 para marcadores RFLP, dependendo do número de alelos que cada combinação de genótipos tenha em comum. Diante do resultado obtido através do *bootstrap* podemos considerar que o número de marcadores AFLP utilizado neste trabalho foi suficiente para caracterizar as cultivares de soja quanto à similaridade genética.

#### 4.3.2.1 Comportamento da diversidade genética em diferentes épocas

Com o intuito de avaliar o comportamento da diversidade genética da soja ao longo do tempo em que o melhoramento genético foi realizado no Brasil, investigou-se os coeficientes (AFS) obtidos, através do agrupamento das cultivares em três épocas, de acordo com sua liberação para o mercado. Considerou-se na primeira época, 48 cultivares liberadas entre 1962 e 1980, na segunda época, 122 cultivares liberadas entre 1981 a 1990, e, na terceira época, 121 cultivares liberadas entre 1991 a 1998. Vinte e seis cultivares não foram incluídas neste estudo devido a falta de informações quanto aos respectivos anos de liberação. Os resultados obtidos (Tabela 3) mostram que os valores médios permaneceram praticamente constantes para as três épocas, sugerindo que, em termos gerais, as cultivares de soja desenvolvidas pelos programas brasileiros de melhoramento têm mantido um nível de similaridade genética semelhante ao longo do tempo. Em trabalho semelhante, com marcadores RFLP, Kisha et al. (1998), verificaram que a diversidade entre cultivares norte-americanas estava sendo reduzida com o decorrer do tempo devido aos efeitos do melhoramento.

Tabela 3. Valores médios, máximos e mínimos dos coeficientes de similaridade obtidos por marcadores AFLP entre cultivares de soja liberadas em diferentes épocas no Brasil

Coeficiente	Época de liberação das cultivares		
	1962 a 1980	1981 a 1990	1991 a 1998
Médio	0,62	0,62	0,63
Máximo	0,86	0,97	0,92
Mínimo	0,23	0,34	0,23
nº de cultivares	48	122	121

Apesar da estreita base genética encontrada na soja cultivada no Brasil (Hiromoto & Vello, 1986) e da relativa alta similaridade entre as cultivares, têm sido obtidos ganhos substanciais para produção de grãos e outros caracteres. Fatos semelhantes têm ocorrido nos Estados Unidos da América (Fehr, 1987). Neste sentido, Toledo et al. (1990) estimaram um ganho médio anual de 1,8% para um grupo de cultivares precoces e de 1,3% para o grupo de cultivares semi-precoces do programa de melhoramento de soja do Estado do Paraná, no período de 1981 a 1986. Mais recentemente, Rubin & Santos (1996), ao estudarem cultivares de soja que foram largamente cultivadas em diferentes períodos no Rio Grande do Sul, concluíram que nos últimos 40 anos foi obtido um ganho médio de 19 kg/ha/ano, ou seja, de 1,1% ao ano. Esses autores comentaram, no entanto, que os ganhos vinham diminuindo com o decorrer do tempo, devido, entre outras razões, ao uso do mesmo germoplasma básico nas hibridações. Isso, no entanto, parece não estar se confirmando, segundo revelam recentes lançamentos de novas cultivares. Além de consideráveis ganhos no aprimoramento de características agronômicas conseguidos pelas correções de defeitos controlados por caracteres qualitativos, são relatados também ganhos na produção de grãos. No Rio Grande do Sul, a nova cultivar BRS 153, liberada em 1998, tem apresentado produtividade 14% superior à da testemunha mais cultivada, BR-16 (Bonato et al., 1999). No Paraná, a cultivar BRS 133 superou em 8,5% a testemunha BR-16 (Almeida et al., 1999). No Estado de Mato Grosso do Sul, a nova cultivar MS/BRS-171 (Campo Grande), liberada em 1998, superou em 20% a cultivar FT-Cristalina (Zuffo et al., 1999). Em Mato Grosso, a cultivar MT/BR-50 (Parecis), liberada em 1996, foi 5% superior à cultivar padrão MT/BR-45 (Paiaguás), conforme Takeda et al. (1999).

### **4.3.3 Agrupamento das 317 cultivares de soja baseado nas estimativas de similaridade genética via AFLP (AFS)**

Para permitir a visualização das similaridades genéticas (AFS) entre os 317 cultivares estudadas, realizou-se o agrupamento pelo método UPGMA, que está apresentado na forma de um dendrograma (Anexo C). Apesar da interpretação ter sido dificultada pelo grande número de genótipos utilizados no trabalho, tornando-se impraticável a avaliação das similaridades de cada cultivar em relação às demais, foi possível detectar grupos formados de acordo com o esperado segundo a genealogia. Esses resultados corroboram a indicação da eficiência na utilização do AFLP como base para estimação da similaridade genética entre genótipos. O conhecimento da genealogia foi essencial para o auxílio e complementação da interpretação do dendrograma, o que concorda com outros trabalhos (Abdelnoor et al., 1995; Maughan et al., 1996 ; Diwan & Cregan, 1997).

Os coeficientes de similaridade (AFS) e o dendrograma mostraram que as cultivares originadas por mutações naturais apresentaram, como era esperado, coeficientes de similaridade genética elevados com as cultivares de origem. Em relação à cultivar Paraná, os mutantes Paranagoiana e Ocepar 9-SS1 apresentaram coeficientes de similaridade de 0,95. A cultivar mutante São Carlos revelou um coeficiente de similaridade genética de 0,83 com a cultivar original Davis, enquanto que a cultivar mutante UFV-1 mostrou coeficiente de 0,67 com a cultivar original Viçosa. As diferenças de magnitudes entre os coeficientes de similaridade genética observados para diferentes mutantes parecem indicar que nem todas as mutações que originaram os genótipos considerados foram pontuais, caso realmente se confirme que as cultivares em questão tenham se originado de mutações. No caso de Paranagoiana e de Ocepar 9-SS1, Derbyshire et al. (1990) comprovaram, através de estudos por eletroforese de isoenzimas, tratar-se realmente de mutantes da cultivar Paraná.

Os coeficientes de similaridade genética estimados entre cultivares que se originaram de outras cultivares também evidenciaram casos de alta similaridade. Exemplos disso foram constatados com as cultivares FT-Cristal e FT-Bahia, que apresentaram coeficientes (AFS) de 0,97 e 0,87, respectivamente, com FT-Cristalina, de onde foram selecionadas. Da mesma forma, a cultivar Ocepar-8 teve sua alta similaridade confirmada (AFS=0,90) com a cultivar Paraná, da qual foi selecionada. Entretanto, em alguns casos verificou-se uma menor similaridade, como foi constatado na cultivar IAS 5. Esta foi separada em dois tipos, de acordo com a cor da vagem, isto é, IAS 5 (vagem escura) e IAS 5 (vagem clara), na Embrapa Soja (Embrapa Soja: Comunicação pessoal). A análise por AFLP indicou que os dois tipos devem divergir em outros genes, além dos que determinam a cor da vagem, visto apresentarem entre si um coeficiente de similaridade de apenas 0,63. A cultivar FT-2, que se originou de seleção em IAS 5, evidenciou estar mais próxima da IAS-5 de vagem clara (AFS=0,78) do que da IAS-5 de vagem escura (AFS=0,65). Esses dados indicam uma refinação da técnica de AFLP na distinção, mesmo entre genótipos muito próximos. Diwan e Cregan (1997), através da análise de 20 microssatélites, não conseguiram separar dois genótipos de soja, sendo que um deles (Ilini) era seleção do outro (AK, Harrow).

Entre cultivares irmãs, ou seja, provenientes do mesmo cruzamento, foram observados vários casos de concordância entre os valores dos coeficientes de similaridade e suas alocações em um mesmo grupo no dendrograma. As cultivares MT/BR-50 (Parecis), MT/BR-51 (Xingu), MT/BR-52 (Curió) e MT/BR-53 (Tucano), oriundas do cruzamento BR 83-9520-1(2) x FT-Estrela, apresentaram entre si coeficientes de similaridade genética maiores de 0,80. Situação semelhante ocorreu entre as cultivares UFV-2, UFV-3, UFV-4 e UFV-Araguaia, descendentes do cruzamento Hardee x IAC-2, cujos coeficientes foram maiores de 0,83. Entre as cultivares FT-5, FT-10, FT-14 e FT-15, originadas do cruzamento entre FT 9510 x Sant'Ana, os coeficientes variaram de 0,71 a 0,93. As cultivares Embrapa 59, Embrapa 60, Embrapa 61 e Embrapa



66, descendentes do cruzamento FT-Abyara x BR 83-147, apresentaram entre si coeficientes maiores que 0,75. Por outro lado, as cultivares BRS 133, BRS 135, BRS 158, MS/BR-57 (Lambari) e MS/BRS-171 (Campo Grande), também selecionadas do mesmo cruzamento, apresentaram coeficientes menores que 0,65 com as demais cultivares irmãs acima citadas. Essas diferenças de similaridade, observadas entre as cultivares irmãs, podem ser atribuídas aos efeitos da seleção. Abdelnoor et al. (1995) também constataram similaridades em variados níveis entre cultivares provenientes de um mesmo cruzamento.

As cultivares desenvolvidas a partir de retrocruzamentos revelaram coeficientes variados de similaridade genética com seus genitores recorrentes. Assim, por exemplo, as cultivares BR-6 (Nova Bragg) e BR-13 (Maravilha), obtidas a partir de, respectivamente, três e quatro retrocruzamentos para a cultivar Bragg, apresentaram coeficientes de similaridade genética maiores que 0,75 com esta última. A cultivar Embrapa 1, obtida de seis retrocruzamentos para IAS 5, apresentou coeficiente de similaridade genética de 0,68 com IAS 5 (vagem escura) e de 0,54 com IAS 5 (vagem clara). A cultivar Embrapa 4, originária de seis retrocruzamentos, mostrou coeficiente de similaridade de apenas 0,61 com o genitor recorrente BR-4. A menor similaridade genética que a esperada, detectada neste estudo entre produtos de retrocruzamentos e seus respectivos genitores recorrentes, pode ser explicada, segundo Muehlbauer et al. (1988), pela introgressão de outros marcadores existentes no mesmo grupo de ligação do gene que foi transferido. Outra possibilidade seria a menor pressão de seleção aplicada para recuperar as características genéticas do genitor recorrente.

A análise do dendrograma não possibilitou a separação das cultivares em grupos baseados na distribuição geográfica para onde foram liberadas ou recomendadas. Em contrapartida, análises com RFLP mostraram uma clara separação entre cultivares de soja do norte e do sul dos EUA, havendo maior similaridade entre os genótipos do sul (Keim et al., 1992; Kisha et al., 1998). Este resultado está relacionado com o fato das cultivares de cada região serem

provenientes de distintos grupos de ancestrais. No Brasil, provavelmente, não ocorrem diferenças devido ao fato de que as cultivares desenvolvidas em diferentes locais originaram-se do mesmo grupo de ancestrais.

No agrupamento das cultivares pelo dendrograma foi possível constatar que na maioria dos casos discutidos anteriormente, os coeficientes estimados foram maiores que os visualizados através do dendrograma. Observou-se, também, que algumas cultivares que apresentaram coeficientes indicativos de grande similaridade genética, localizaram-se em grupos distintos no dendrograma. Isso ocorreu, por exemplo, com as cultivares Pirapó 78 e Nova IAC-7 que, mesmo apresentando coeficientes de 0,82, localizaram-se em grupos distintos das respectivas cultivares Paraná e IAC-7, das quais se originaram. Possíveis causas dessa distorção seriam o grande número de genótipos geneticamente muito próximos analisados e a baixa correlação cofenética obtida entre os coeficientes originais e os estimados pelo agrupamento, que foi de  $r=0,60$ . Usando marcadores AFLP, Powell et al. (1996) obtiveram uma correlação cofenética de 0,78 entre os acessos de *Glycine max*, enquanto que considerando acessos desta espécie juntamente com acessos de *Glycine soja*, o valor se elevou para 0,96. A menor correlação cofenética, observada entre acessos de *Glycine max*, pode ter sido decorrente da maior similaridade genética dos genótipos dessa espécie. Esta explicação pode ser estendida aos resultados encontrados no presente trabalho, onde as cultivares apresentaram coeficientes muito próximos entre si, o que pode ter interferido nos resultados estimados no agrupamento.

#### **4.3.4 Agrupamento das 63 cultivares de soja baseado nas estimativas de similaridade genética via AFLP (AFS)**

Com o intuito de facilitar a interpretação e, com isso, aferir a validade das considerações mencionadas anteriormente, elaborou-se um novo dendrograma

com apenas 63 cultivares, dentre as quais foram incluídas cultivares que apresentaram resultados considerados discrepantes pelo dendrograma que continha todos os genótipos. Também foram incluídas cultivares com coeficientes de similaridade de diferentes magnitudes (Figura 4). Neste caso, observou-se, além da maior facilidade de visualização dos grupos, um pequeno aumento nos valores de correlação cofenética, de 0,60 para 0,70. Em termos práticos, esses resultados sugerem que a seleção de genitores, feita com base na diversidade genética, pode se tornar mais efetiva se os melhoristas realizarem agrupamentos, a partir dos coeficientes de similaridade estimados nesse trabalho, com número menor de genótipos, previamente eleitos a partir de suas características qualitativas e quantitativas desejáveis.

Entre as cultivares consideradas discrepantes no dendrograma anterior (Anexo C) houve pares de cultivares que se realocaram dentro de um mesmo grupo, evidenciando concordância com os respectivos coeficientes originais. Exemplo disso foi o grupo formado pelas cultivares Embrapa 32 (Itaqui), Embrapa 33 e Cariri RCH, provenientes de retrocruzamentos para a cultivar BR-27 (Cariri). Situação semelhante ocorreu para a cultivar MS/BRS-172 (Tuiuiú) e sua genitora recorrente (FT-Cristalina), em quatro retrocruzamentos (AFS=0,77). A cultivar MT/BR-45 (Paiaguás), resultante do cruzamento entre Doko x IAC-7, que apresentou coeficiente de similaridade (AFS) de 0,79 com a cultivar IAC-7, também se agrupou junto com esta e com a Nova IAC-7. Um outro exemplo de melhoria no agrupamento ocorreu entre as cultivares irmãs MT/BR-55 (Uirapuru) e MT/BRS-159 (Crixás), em relação à MT/BR-50, com coeficientes de similaridade (AFS) acima de 0,70.

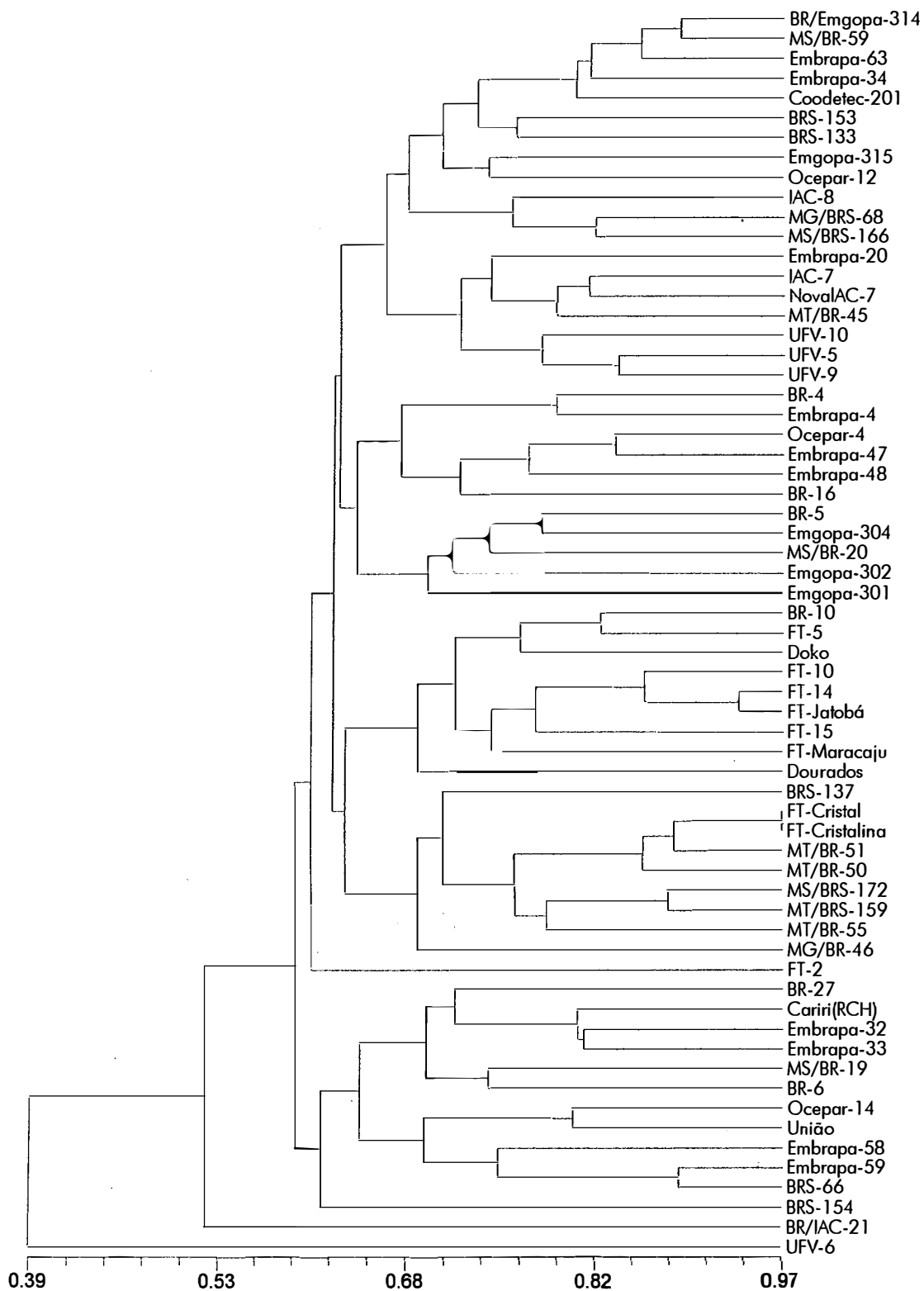


Figura 4. Dendrograma gerado por UPGMA baseado em coeficientes de similaridade obtidos através de AFLP, entre 63 cultivares de soja.

Algumas cultivares, no entanto, mantiveram-se em diferentes grupos, apesar de exibirem altos coeficientes de similaridade entre elas. Isso, provavelmente, seja reflexo de deficiências da metodologia de agrupamento, evidenciadas pela moderada correlação cofenética. Tais fatos puderam ser constatados nos agrupamentos da cultivar BR/Emgopa-314, obtida por seis retrocruzamentos para a cultivar Emgopa 301, cujo coeficiente de similaridade foi de 0,70; da cultivar BR/IAC-21, desenvolvida a partir de cinco retrocruzamentos para a cultivar IAC-8 e com coeficiente de 0,74; da cultivar Coodetec 201, produto de cinco retrocruzamentos para a Ocepar 4-Iguaçu (AFS=0,75) e entre as cultivares irmãs BR-4 e BR-5 (AFS=0,75) e BRS 153 e BRS 154 (AFS=0,71).

Finalmente, algumas cultivares situaram-se em grupos distintos devido a baixa similaridade, constatada pelos marcadores, apesar de serem aparentadas, como foram os casos das cultivares BRS-137 e o genitor recorrente Dourados, com coeficiente de similaridade (AFS) de 0,65, e de Embrapa-20 (Doko RC), oriunda de quatro retrocruzamentos para a cultivar Doko, com um coeficiente de similaridade (AFS) de 0,63. Fato semelhante foi também observado entre a cultivar UVF-6 e suas irmãs UFV-5, UFV-9 e UFV-10.

#### **4.3.5 Fatores que podem influenciar as estimativas de similaridade genética obtidas por marcadores moleculares**

Resultados contraditórios às expectativas teóricas, baseadas na genealogia de cada cultivar, são difíceis de serem explicados com precisão. No entanto, diversas fontes de erro podem existir em estudos dessa natureza, tais como: impurezas genéticas das sementes, qualidade dos reagentes e produtos utilizados, imprecisão de procedimentos relacionados à análise de AFLP e leitura das marcas polimórficas. Dentre esses fatores, o mais relevante para o

presente estudo seria provavelmente a impureza genética da sementes, uma vez que não foram utilizadas amostras foliares obtidas de apenas um planta. Por outro lado, o uso de amostra obtida de uma única planta poderia incorrer em uma outra fonte de erro, em casos em que a contaminação possa ser mais elevada, ou mesmo no caso de cultivares formadas por mistura de linhas irmãs.

Além disso, diversos outros fatores podem influenciar as estimativas de similaridade genética e devem ser considerados em trabalhos desta natureza. Um deles é o número de marcadores utilizados, que pode afetar a variância da estimativa da similaridade, ao se considerar que um marcador representa uma amostra independente do genoma (Keim et al., 1992; Powell et al., 1996). Neste trabalho, porém, os 78 marcadores polimórficos obtidos foi considerado um número adequado na análise das 317 cultivares, como já comentado anteriormente.

A distribuição dos marcadores no genoma, também é um fator considerado importante para estudos de diversidade, uma vez que quanto maior for a cobertura do genoma, mais eficiente será sua representação e comparação entre indivíduos. A princípio, considera-se que os marcadores são distribuídos ao acaso no genoma (Williams et al., 1990; Becker et al., 1995), havendo evidências que o AFLP tem ampla cobertura sobre o genoma das plantas (Maheswaran et al., 1997; Zhu et al., 1998). No entanto, alguns trabalhos têm indicado que as combinações de *primers* utilizados na análise de AFLP atuam na distribuição dos marcadores no genoma, sendo que as combinações de *primers EcoRI/MseI* estão relacionadas com regiões hipermetiladas do genoma, enquanto que as combinações *PstI/MseI* com regiões hipometiladas, sugerindo não haver uma distribuição ao acaso no genoma (Barret & Kidwell, 1998; Young et al., 1999).

Outro fator que pode ser definido previamente e pode influenciar os resultados de similaridade é o coeficiente de similaridade genética utilizado. Comparando os coeficientes mais empregados com dados de AFLP, o coeficiente de Nei & Li (1979) distingue-se do coeficiente de Distância Comum,

ou Simple Matching (Sokal & Mechener, 1958), por não considerar a ausência de bandas como evidência de similaridade entre os indivíduos. Esta consideração do coeficiente de Distância Comum pode conduzir a uma superestimação da similaridade genética, uma vez que a ausência de amplificação de uma determinada banda entre dois genótipos que ocorre em marcadores dominantes, não representa necessariamente similaridade genética entre eles. Em relação ao coeficiente de Jaccard, o coeficiente de Nei & Li difere apenas pelo peso duplo nas ocorrências de bandas nos dois genótipos em questão. Desta forma, o coeficiente de Nei & Li é indicado para este tipo de análise (Duarte et al., 1999a).

As modificações genéticas, tidas como variações *de novo*, vêm sendo investigadas e consideradas como fontes de variabilidade, a princípio não esperada entre indivíduos. Entre os mecanismos que podem gerar essas variações estão: recombinação intragênica, *crossing over* desigual, atividade de transposons, e metilação do DNA (Phillips, 1999).

#### **4.3.6 Estimativas de similaridade genética baseadas no coeficiente de parentesco (f) e correlação entre a similaridade genética obtida por AFLP (AFS) e pelo coeficiente de parentesco**

Os 4.950 coeficientes de parentesco entre 100 cultivares de soja, liberadas a partir de 1984, variaram de 0 a 1, com média de 0,21 (Anexo D). Os valores de  $f=0$  representam, por esta metodologia, ausência de parentesco entre os pares de cultivares. Foram obtidas 294 (5,9%) estimativas com este valor. O máximo de parentesco, representado por  $f=1$ , foi obtido em 14 (0,3%) pares de cultivares. Vello et al. (1988) estimaram o coeficiente de parentesco entre 69 cultivares recomendadas em 1983/84 e observaram um coeficiente médio de 0,16. Nos EUA, Gizlice et al. (1993) obtiveram um coeficiente de parentesco médio de 0,18 para as cultivares do norte e de 0,23 para as

cultivares do sul dos Estados Unidos, liberadas entre 1947 e 1988, coeficientes estes altos. Sneller et al. (1994) relataram resultados similares quando compararam cultivares elite do norte e do sul dos EUA. Entretanto, encontraram coeficientes de parentesco surpreendentemente baixos ( $f=0,10$ ) entre as regiões norte e sul, sugerindo que os melhoristas tem mantido *pools* distintos de genes. Por outro lado, a análise de cultivares de soja chinesas revelou um coeficiente de parentesco muito baixo ( $f=0,02$ ), demonstrando o alto grau de diversidade genética desse germoplasma (Cui et al, 1999).

O coeficiente de correlação entre a similaridade genética obtida via AFLP (AFS) e o coeficiente de parentesco ( $f$ ) foi de 0,12. Diversos são os fatores que podem justificar a baixa correlação entre esses coeficientes encontrados neste trabalho. Inicialmente, os dois tipos de coeficientes não são congruentes no tipo de similaridade genética avaliada, ou seja, o coeficiente de parentesco é uma determinação matemática baseada em probabilidades, enquanto que os marcadores moleculares detectam similaridades diretamente ao nível do DNA. As suposições relacionadas ao cálculo do coeficiente  $f$  são alicerçadas em condições nem sempre realistas, em se tratando do germoplasma utilizado no melhoramento de plantas. A violação dessas suposições parece ser crítica para o emprego deste coeficiente e pode ter como consequência a baixa correlação com os dados de marcadores genéticos. O cálculo do  $f$  assume a premissa que cada genitor contribui com 50% do genoma para sua progênie, não considerando o impacto da pressão de seleção e da deriva genética (Barret et al., 1998). Nos casos de cultivares derivadas de seleção ou de mutação a partir de outras cultivares, o coeficiente de parentesco assume  $f=1$ , desconsiderando a possibilidade de possuírem genes diferentes. A suposição de que  $f=0$ , quando os ancestrais não tem parentesco em comum, nem sempre é adequada e pode levar a uma sub-estimação da relação entre dois genótipos (Messmer et al., 1993). Como exemplo, Gizlice et al. (1993), através de análise multivariada, encontraram coeficientes de similaridade variando de 0 a 0,88 entre ancestrais de soja norte-americanas. Além disso, a falta de conhecimento preciso da



genealogia pode interferir negativamente nas estimativas do coeficiente  $f$ , o que não ocorre com os marcadores moleculares que não exigem o conhecimento prévio da genealogia para o cálculo da similaridade genética (Prabhu et al., 1997).

A correlação entre a similaridade genética obtida por marcadores e pelo coeficiente de parentesco tem sido investigada em muitos trabalhos, tendo sido obtidos resultados variados, atingindo desde valores extremamente baixos ( $r=-0,33$ ) em trigo, com RFLP (Barbosa-Neto et al., 1996), a altos ( $r=0,97$ ) em milho, com o mesmo tipo de marcador (Smith et al., 1991). Com AFLP, Barret & Kidwell (1998), trabalhando com trigo, encontraram uma correlação de 0,42, e as justificativas para a moderada correlação, utilizadas pelos autores, concordam com os fatores citados anteriormente.

Concluindo, neste trabalho foi relatado a primeira e ampla análise da similaridade genética do germoplasma brasileiro de soja. Os marcadores AFLP provaram ser uma efetiva e rápida ferramenta para este tipo de análise. Além de, demonstrar a capacidade da técnica, os resultados apresentados poderão auxiliar na seleção de genitores para programas de melhoramento genético.

## **5 CONCLUSÕES**

As análises dos dados e a interpretação dos resultados nesse trabalho permitiram apontar as seguintes conclusões:

- Existe correlação entre a variância genética aditiva e a distância genética obtida com marcadores AFLP (AFD) para os caracteres produção de grãos, número de nós, número de dias para floração, número de dias para maturação e altura das plantas na maturação.
- Não existe correlação entre a variância genética aditiva e a distância genética estimada pelo coeficiente de parentesco (DG) para os caracteres produção de grãos, número de nós, número de dias para floração, altura de plantas na floração, número de dias para maturação e altura das plantas na maturação.
- O marcador AFLP é uma ferramenta útil na predição da variabilidade genética e na estimativa da similaridade genética entre cultivares de soja.
- Existe alto grau de similaridade genética entre as cultivares de soja, liberadas no Brasil no período de 1962 a 1998.
- O grau médio de similaridade genética entre as cultivares de soja liberadas nas últimas três décadas tem permanecido constante.
- Não existe correlação entre os coeficientes de parentesco e de similaridade genética estimados por AFLP entre as cultivares de soja.

## **ANEXOS**

Anexo A. Estimativas da variância genética aditiva para sete caracteres em quatro anos e quatro épocas de semeadura em seis populações de soja<sup>1</sup>.

ANO	EPOCA	POPULAÇÃO	PRODUÇÃO DE GRÃOS	NÚMERO DE NÓS	DIAS PARA FLORAÇÃO	ALTURA DA PLANTA NA FLORAÇÃO	DIAS PARA MATURAÇÃO	ALTURA DA PLANTA NA MATURAÇÃO	PESO 100 SEMENTES
1991/92	Out.	1	86,72	1,06	9,86	21,19	35,91	75,33	2,20
		2	195,34	1,34	23,84	36,76	41,18	104,10	4,54
		3	649,52	2,34	67,51	92,24	71,52	185,46	2,87
		4	127,42	0,72	12,42	12,96	37,25	51,63	3,16
		5	388,38	1,08	19,07	43,66	58,99	61,02	0,77
		6	324,38	2,39	18,41	39,55	61,31	140,78	1,81
1992/93	Nov.	1	50,00	0,53	3,37	20,88	35,55	50,03	2,20
		2	63,32	0,75	9,77	29,63	48,07	77,07	1,76
		3	134,77	2,74	15,56	44,82	83,03	178,63	2,10
		4	33,71	0,44	9,27	18,01	61,16	39,35	2,02
		5	122,54	1,70	10,22	31,66	75,39	85,58	1,93
		6	127,76	1,42	10,08	37,27	59,17	80,58	3,31
1992/93	Dez.	1	36,36	0,28	3,09	26,83	11,85	44,23	2,04
		2	78,86	1,04	8,76	55,12	32,99	92,78	2,58
		3	90,17	2,18	19,44	94,50	60,00	115,19	2,16
		4	36,48	0,91	9,36	25,64	30,54	75,30	2,96
		5	91,89	1,15	12,01	70,78	36,29	67,59	1,73
		6	117,41	1,45	8,94	32,92	47,64	88,36	3,69
1992/93	Out.	1	30,57	1,00	10,42	24,15	35,21	55,65	1,39
		2	58,44	1,53	31,95	42,06	75,72	79,77	1,90
		3	110,44	1,88	31,38	35,84	126,61	74,74	1,79
		4	18,14	0,70	15,93	21,74	85,43	34,64	2,99
		5	68,03	1,30	11,14	21,64	117,16	42,37	1,18
		6	64,34	1,57	28,58	36,13	111,12	75,92	3,00
1992/93	Nov.	1	36,29	0,73	3,46	20,99	20,56	44,99	1,22
		2	29,67	0,79	12,57	26,86	42,16	82,79	2,12
		3	44,88	2,41	16,72	35,83	67,50	103,42	2,93
		4	13,75	1,00	9,57	19,43	60,35	50,69	1,96
		5	39,80	0,94	11,56	19,01	73,21	71,08	2,42
		6	42,29	2,05	16,64	25,76	65,69	85,68	2,60
1992/93	Dez.	1	28,65	0,71	2,07	22,16	15,69	41,64	1,03
		2	18,12	0,34	7,15	34,15	31,69	64,76	1,53
		3	22,61	2,06	13,85	38,27	62,90	62,88	1,07
		4	1,22	0,68	8,04	28,21	29,04	46,77	1,65
		5	0,15	1,38	8,98	26,74	43,79	66,49	0,97
		6	10,63	0,82	11,44	26,42	38,52	36,13	2,05

Anexo A. Estimativas da variância genética aditiva para sete caracteres em quatro anos e quatro épocas de semeadura em seis populações de soja<sup>1</sup> (Continuação).

ANO	ÉPOCA	POPULAÇÃO	PRODUÇÃO DE GRÃOS	NÚMERO DE SEMENTES	DIAS PARA FLOREÇÃO	ALTURA DA PLANTA NA FLOREÇÃO	DIAS PARA MATURAÇÃO	ALTURA DA PLANTA NA MATURAÇÃO	PESO 100 SEMENTES
1993/94	Set.	1	54,77	0,79	6,78	13,60	23,77	36,49	2,50
		2	185,79	1,27	22,19	27,09	90,34	63,60	3,38
		3	220,51	1,45	16,50	24,34	118,26	74,75	2,13
		4	36,82	0,51	10,43	10,46	59,58	18,40	3,56
		5	84,93	0,59	7,66	40,38	70,27	25,46	3,07
		6	161,81	0,50	23,05	34,35	109,68	70,97	2,89
1993/94	Out.	1	69,21	0,77	7,41	64,20	35,27	84,99	3,21
		2	246,06	1,94	23,07	112,79	66,40	179,75	3,36
		3	332,82	2,42	34,85	190,16	125,34	235,08	1,84
		4	139,39	2,37	12,70	47,92	65,87	131,65	4,53
		5	142,73	1,36	14,70	87,93	94,14	105,57	2,93
		6	221,23	2,73	18,83	78,66	93,11	161,82	2,80
1993/94	Nov.	1	104,22	0,73	4,20	44,31	12,46	73,39	2,05
		2	119,97	0,96	17,89	75,30	24,77	149,93	2,42
		3	188,49	2,50	39,21	141,12	40,37	162,84	2,05
		4	68,27	0,72	11,24	36,38	32,34	40,73	1,21
		5	80,65	1,12	14,43	58,61	30,64	64,76	1,90
		6	119,47	1,60	18,56	62,48	42,29	81,29	2,83
1993/94	Dez.	1	66,96	0,75	1,59	35,80	11,37	39,41	2,35
		2	97,89	1,06	7,51	56,26	15,97	64,64	2,36
		3	85,33	1,16	7,58	62,80	29,15	81,75	1,46
		4	43,16	1,89	7,43	39,53	7,42	66,30	2,24
		5	53,24	1,08	9,49	31,76	19,16	39,27	1,42
		6	48,26	1,79	9,15	45,21	16,14	80,99	2,79
1994/95	Set.	1	45,04	1,02	7,13	13,77	36,38	35,22	3,11
		2	120,84	1,70	18,84	24,36	115,61	62,54	1,97
		3	73,12	2,10	15,58	30,14	124,92	83,46	0
		4	27,66	0,74	10,27	9,30	60,62	27,35	1,25
		5	49,48	1,10	9,59	8,78	90,02	32,27	0
		6	90,67	1,64	20,64	18,30	137,67	54,82	1,63
1994/95	Out.	1	89,84	1,48	8,49	44,18	34,89	102,25	3,69
		2	344,68	2,74	23,71	54,81	58,83	146,73	4,78
		3	468,22	1,61	36,12	146,07	98,46	144,48	2,34
		4	104,82	1,51	14,27	28,80	71,60	90,61	3,58
		5	149,02	1,09	11,35	32,80	81,33	65,07	2,79
		6	257,41	1,98	23,13	46,21	83,80	141,68	5,26

Anexo A. Estimativas da variância genética aditiva para sete caracteres em quatro anos e quatro épocas de semeadura em seis populações de soja<sup>1</sup> (Continuação).

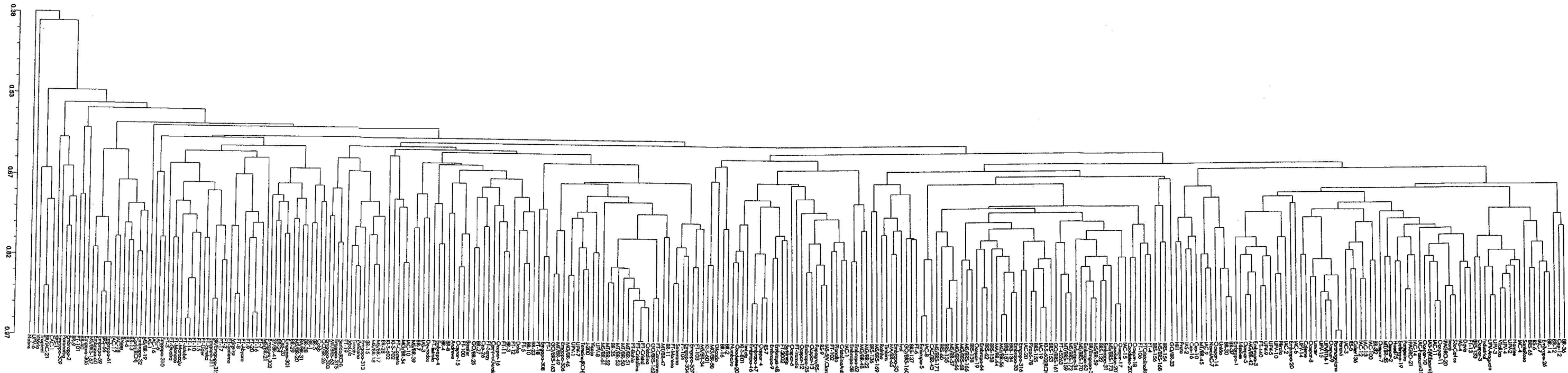
ANO	ÉPOCA	POPULAÇÃO	PRODUÇÃO DE GRÃOS	NÚMERO DE NÓS	DIAS PARA FLORAÇÃO	ALTURA DA PLANTA NA FLORAÇÃO	DIAS PARA MATURAÇÃO	ALTURA DA PLANTA NA MATURAÇÃO	PESO 100 SEMENTES
1994/95	Nov.	1	75,23	0,49	2,71	77,25	24,39	100,19	2,36
		2	93,21	0,82	8,46	106,03	33,91	121,73	2,37
		3	104,93	1,54	16,19	128,46	49,91	116,48	2,15
		4	57,82	0,96	8,42	79,18	43,86	79,96	1,87
		5	99,03	0,92	8,67	73,88	48,86	69,38	2,14
		6	73,11	0,19	9,07	64,47	41,29	58,18	3,42
Dez.		1	31,15	0,80	1,63	27,61	12,73	49,19	2,55
		2	37,98	1,00	6,25	42,58	15,85	76,12	1,92
		3	42,48	1,51	11,03	47,92	30,39	78,25	1,68
		4	21,72	0,86	6,25	24,87	19,83	53,25	1,83
		5	34,54	1,37	6,46	31,20	27,62	68,52	1,72
		6	37,35	1,29	8,45	28,97	23,10	66,67	2,83

<sup>1</sup> Significativo ao nível de 1% de probabilidade, pelo teste t.

Fonte: Dados originais obtidos por J.F.F. Toledo (Embrapa Soja: dados não publicados)

**Anexo B. Coeficientes de similaridade genética entre 317 cultivares de soja, liberadas no Brasil, obtidos através de AFLP.**

- ♦ Apresentação em compact-disc



Anexo C. Dendrograma gerado por UPGMA baseado em coeficientes de similaridade obtidos através de AFLP, entre 317 cultivares de soja.



**Anexo D. Coeficientes de parentesco e coeficientes de similaridade genética por AFLP entre 100 cultivares de soja.**

- ♦ Apresentação em compact-disc

## 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. **Revista Brasileira de Genética**, v.18, n.2, p.265-273, 1995.
- AJMONE-MARSAN, P.; LIVINI, C.; MESSMER, M.M.; MELCHINGER, A.E.; MOTTO, M. Cluster analyses of RFLP data from related maize inbred lines of the BSSS and LSC heterotic groups and comparison with pedigree data. **Euphytica**, v.60, p.139-148, 1992.
- AKKAYA, M.S.; BHAGWAT, A.A.; CREAGAN, P.B. Length polymorphisms of simple sequence repeat DNA in soybean. **Genetics**, v.132, p.1131-1139, 1992.
- ALMEIDA, L.A.; KIIHL, R.A.S.; ABDELNOOR, R.V. Melhoramento de soja. In: Abreu, A.F.B.; Gonçalves, F.M.A.; Marques Jr, O.G.; Ribeiro, P.H.E. (Eds.) Simpósio sobre atualização em genética e melhoramento de plantas. Lavras: Editora UFLA/GEN, 1997. 290 p.
- ALMEIDA, L.A.; KIIHL, R.A.S.; MIRANDA, L.C.; PÍPOLO, A.E.; YORINORI, J.T.; KASTER, M.; DOMIT, L.A. Cultivar de soja BRS 133. In: CONGRESSO BRASILEIRO DE SOJA, 1, Londrina, 1999. **Anais**. Londrina: Embrapa Soja, 1999. p.483. (Embrapa Soja. Documentos, 124).
- BARBOSA-NETO, J.F.; SORRELS, M.E.; CISAR, G. Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. **Genome**, v.39, p.1142-1149, 1996.

- BARRETT, B.A.; KIDWELL, K.K. AFLP-based genetic diversity among wheat cultivars from the Pacific Northwest. **Crop Science**, v.38, p.1261-1271, 1998.
- BARRETT, B.A.; KIDWELL, K.K.; FOX, P.N. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. **Crop Science**, v.38, p.1271-1278, 1998.
- BARROSO, P.A.V. Comparações entre métodos quantitativos e métodos baseados em marcadores RAPD para a predição do comportamento de populações de soja. Piracicaba, 2000. 162p. Tese (Doutorado) - Escola Superior de Agronomia "Luiz de Queiroz", Universidade de São Paulo.
- BECKER, J.; VOS, P.; KUIPER, M.; SALAMINI, F.; HEUN, M. Combined mapping of AFLP and RFLP markers in barley. **Molecular General Genetics**, v.249, p.65-73, 1995.
- BEER, S.C.; SOUZA, E.; SORRELS, M.E. Prediction of genotype performance from ancestral relationship in oat. **Crop Science**, v.35, p.69-73, 1995.
- BERNARDO, R. Estimation of coefficient of coancestry using molecular markers in maize. **Theoretical and Applied Genetics**, v. 85, p.1055-1062, 1993.
- BERNARDO, R. Prediction of maize single-cross performance using RFLPs and information on related hybrids. **Crop Science**, v.34, p. 20-25, 1994.
- BESSE, P.; TAYLOR, G.; CARROL, B.; BERDING, N.; BURNER, D.; McINTYRE, C.L. Assessing genetic diversity in a sugarcane germoplasm collection using an automated AFLP analysis. **Genetica** v.104, p.143-153, 1998.
- BOHN, M.; UTZ, H.F.; MELCHINGER, A.E. Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. **Crop Science**, v.39, p.228- 237, 1999.

- BONATO, A.L.V.; CALVO, E.S.; GERALDI, I.O.; ARIAS, C.A.A. Genetic similarity of soybean (*Glycine max* (L) Merrill) cultivars released in Brazil, using AFLP markers. 2000a (Submetido).
- BONATO, A.L.V.; CALVO, E.S.; ARIAS, C.A.A.; TOLEDO, J.F.F.; GERALDI, I.O. Prediction of genetic variability through AFLP-based measure of genetic distance in soybean (*Glycine max* (L). Merrill). 2000b (Submetido).
- BONATO, E.R.; BONATO, A.L.V. **A soja no Brasil: história e estatística.** Londrina: Embrapa, CNPSo, 1987. 61p. (Embrapa CNPSo. Documentos, 21).
- BONATO, E.R. ; DALL'AGNOLL, A. Soybean in Brazil: Production and Research. In: WORLD SOYBEAN RESEARCH CONFERENCE, 3, Ames, 1984. Proceedings. Boulder: Westview Press, 1985. p.1248-1256.
- BONATO, E.R.; BERTAGNOLLI, P.F.; LINHARES, A.G.; COSTAMILAN, L.M.; KIIHL, R.A.S.; ALMEIDA, L.A. Caracterização da nova cultivar de soja BRS 153. In: CONGRESSO BRASILEIRO DE SOJA, 1, Londrina, 1999. **Anais.** Londrina: Embrapa Soja, 1999. p.478. (Embrapa Soja. Documentos, 124).
- BONETTI, L.P. Cultivares e seu melhoramento genético. In: VERNETTI, F. de J. (Coord.) **Soja, genética e melhoramento.** Campinas: Fundação Cargill, 1983. v.2, p.741-800.
- BOTSTEIN, D.; WHITE, R. I.; SKOLNIC, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetic**, v.32, p. 314-331, 1980.
- BRATNAGAR, P.S. Importance and future prospects for genetic diversity for yield improvement. In: WORLD SOYBEAN RESEARCH CONFERENCE, 6, Chicago, USA, 1999. **Proceedings.** p. 185.
- BRUNKLAUS-JUNG, E.; LUDWING, W. F.; BOPPENMAIER, J.; MELCHINGER A. E.; HERRMANN, R. G. RFLP analyses of early maturing European maize germplasm: Correlation between RFLP and pedigree data. **Maydica**, v.38, p.239-244, 1993.

- BURKHAMER, R.L.; LANNING, S.P.; MARTENS, R.J.; MARTIN, J.M.; TALBERT, L.E. Predicting progeny variance from parental divergence in hard red spring wheat. **Crop Science**, v.38, p. 234-248, 1998.
- CARTER JÚNIOR, T. E.; GIZLICE, Z.; BURTON, J. W. **Coefficient-of-parentage and genetic-similarity estimates for 258 North American soybean cultivars released by public agencies during 1945-1988.** Washington:USDA, 1993. ( USDA. Technical Bulletin, 1814)
- CHATFIELD, C.; COLLINS, A.J. **Introduction to multivariate analysis.** 3.ed. London: Chapman and Hall, 1986. 246p.
- CHAVARRIAGA-AGUIRRE, P.; MAYA, M.M.; TOHME, J.; DUQUE, M.C.; IGLESIAS, C.; BONIERBALE, M.W.; KRESOVICH, S.; KOCHERT, G. Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. **Molecular Breeding**, v.5, p. 263-273, 1999.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Indicadores da Agropecuária.** CONAB, 2000. p.11
- COWEN, N. M.; FREY, K. J. Relationship between genealogical distance and breeding behavior in oats (*Avena sativa* L.) **Euphytica**, v.36, p.413-424, 1987.
- COX, T.S.; KIANG, Y.T.; GORMAN, M.B.; RODGERS, D.M. Relationship between coefficient of parentage and genetic similarity indices in the soybean. **Crop Science**, v.25, p.529-532, 1985.
- CUI, Z.; CARTER JÚNIOR, Jr.T.E.; BURTON, J.W. Genetic diversity patterns in chinese soybean cultivars based on coefficient of parentage In: WORLD SOYBEAN RESEARCH CONFERENCE, 6 Chicago, 1999. **Proceedings.** p. 507.

- DELANNAY, X.; RODGERS, D.M.; PALMER, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Science**, v.23, p. 944-949, 1983.
- DEMEKE, T.; ADAMS, R.P.; CHIBBAR, R. Potential use of random amplified polymorphic DNA (RAPD): a case study in *Brassica*. **Theoretical and Applied Genetics**, v.84, p. 990-994, 1993.
- DERBYSHIRE, E.; CARVALHO, M.T.V.; BONATO, E.R. Comparison of natural variants of the soybean cultivar Paraná by isoenzyme analysis. **Brasilian Journal of Genetics**, v.13, p.83-87, 1990.
- D'HONT, A.; RAO, P.S.; FELDMANN, P.; GUVIT, L.; ISLAM-FARIDI, N.; TAYLOR, P.; GLASZMANN, J.C. Identification and characterization of sugarcane intergeneric hybrids, *Saccharum officinarum* x *Erianthus arundinaceus*, with molecular markers and DNA *in situ* hybridisation. **Theoretical and Applied Genetics**, v. 91, n.2, p. 320-326, 1995.
- DIWAN, N.; CREGAN, P.B. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. **Theoretical and Applied Genetics**, v.95, p. 723-733, 1997.
- DUARTE, J.M.; SANTOS, J.B. dos; MELO, L.C. Comparison of similarity coefficients based on RAPD markers in the common bean. **Genetics and Molecular Biology**, v.22, n.3, p. 427-432, 1999a.
- DUARTE, J.M.; SANTOS, J.B. dos; MELO, L.C. Genetic divergence among common bean cultivars from different races based on RAPD markers. **Genetics and Molecular Biology**, v.22, n.3, p. 419-426, 1999b.
- DUBREUIL, P.; CHARCOSSET, A. Relationships among maize inbred lines and populations from European and North American origins as estimated using RFLP markers. **Theoretical and Applied Genetics**, v.99, p. 473-480, 1999.
- D'UTRA, G. Soja. **Jornal do Agricultor**, v.4, n.168, p. 185-188, 1982.

- FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics** . 4.ed. Essex: Longman, 1996 . 464p.
- FEHR, W.R. **Principles of cultivar development**. New York: Macmillan,, 1987. V. 2: Crop species, p.533-576: Soybean.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores RAPD e RFLP em análise genética**. Brasília: EMBRAPA, CENARGEN, 1996. 220p.
- GIZLICE, Z.; CARTER JÚNIOR, T.E.; BURTON, J. W. Genetic basic for North American public soybean cultivars released between 1947 and 1988. **Crop Science**, v.34, p. 1143-1151, 1994.
- GIZLICE, Z.; CARTER JÚNIOR, T.E.; BURTON, J.W. Genetic diversity in north American soybean: I. Mutivariate analysis of founding stock and relation to coefficient of parentage. **Crop Science**, v.33, p.614-620, 1993.
- GIZLICE, Z.; CARTER JÚNIOR,T.E.; GERIG, T.M.; BURTON, J.W. Genetic diversity patternes in north American public soybean cultivars based on coefficient of parentage. **Crop Science**, v.33, p.753-765, 1996.
- GLAZCO, V.I.; DUBIN, A.V.; GLAZCO G.V. Genetic interrelations between soybean varieties, evaluated with the use of ISSR markers. In: WORLD SOYBEAN RESEARCH CONFERENCE, 6, Chicago, 1999. **Proceedings**. p. 515-516.
- GRANER, A.; LUDWIG, W.F.; MELCHINGER, A.E. Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. **Crop Science**, v.34, p. 1199-1205, 1994.
- HAHN, V.; BLANKENHORN, K.; SCHAWALL, M.; MELCHINGER, A.E. Relationships among early european maize inbreds: III. genetic diversity revealed with RAPD markers and comparison with RFLP and pedigree data. **Maydica**, v.40, p. 299-310, 1995.

- HAMRICK, J.L.; ALLARD, R.W. Correlation between quantitative characters and enzyme genotypes in *Avena barbata*. **Evolution**, v.29, p. 438-443, 1975.
- HANSEN, M.; KRAFT, T.; CHISTIANSOMN, M.; NILSSON, N.O. Evaluation of AFLP in *Beta*. **Theoretical and Applied Genetics**, v.98, p. 845-852, 1999.
- HARTL, L.; SEEFELDER, S. Diversity of selected hop cultivars detected by fluorescent AFLPs. **Theoretical and Applied Genetics**, v.96, p.112-116, 1998.
- HELMS, T.; ORF, J.; VALLAD, G.; McCLEAN, P. Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. **Theoretical and Applied Genetics**, v.94, p. 20-26, 1997.
- HIROMOTO, D. M.; VELLO, N. A. The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. **Revista Brasileira de Genética**, v.9, p. 295-306, 1986.
- HYMOWITZ, T. On the domestication of the soybean. **Economy Botany**, v.24, p. 408- 421, 1970.
- JACCARD, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura **Bull. Soc. Vaudoise Sci. Nat.** v.37, p.547-579, 1901.
- JACOB, H.J.; LINDPAINTER, K.; LINCOLN, S.E.; KUSUMI, K.; BUNKER, R.K.; YI-PEI MAO; GANTEN, D.; DZAU, V.J.; LANDER, E.S. Genetic mapping of a gene using hypertension in the stroke-prone spontaneously hipertensive rat. **Cell**, v.167, p. 213-224, 1991.
- JOHNS, M. A .; SKROCH, P.W.; NIENHUIS, J.; HINRICHSEN, P.; BASCUR, G.; MUÑOZ-SCHICK, C. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. **Crop Science**; v.37, p.605-613, 1997.
- KEIM, P.; BEAVIS, W.; SCHUPP, J.; FREESTONE, R. Evaluation of soybean RFLP marker diversity in adapted germplasm. **Theoretical and Applied Genetics**, v.85, p. 205-212, 1992.



- KEIM, P.; SHOEMAKER, R.C.; PALMER, R.G. Restriction fragment length polymorphism in soybean. **Theoretical and Applied Genetics**, v.77, p.786-792, 1989.
- KIIHL, R.A.S.; BAYS, I.A.; ALMEIDA, L.A. 1986 Soybean breeding for the brazilian tropics. In: **Soybean in Tropical and Subtropical Cropping Systems**, Tsukuba, Japan. 26 sept-oct. 1st, 1983. Proceedings. p. 141-144.
- KISHA, T.J.; SNELLER, C.H.; DIERS, B.W. Relationship between genetic distance among parents and genetic variance in populations of soybean. **Crop Science**, v.37, p. 1317-1325, 1997.
- KISHA, T.J.; DIERS, B.W.; HOYT, J.M.; SNELLER, C.H. Genetic diversity among soybean plant introductions and North American germplasm. **Crop Science**, v.38, p.1669-1680, 1998.
- LORENZEN, L. L.; BOUTIN, S.; YOUNG, N.; SPECHT, J. E.; SHOEMAKER, R. C. Soybean pedigree analysis using map-based molecular markers: I. tracking RFLP markers in cultivars. **Crop Science**, v.35, p.1326-1336, 1995.
- MACHADO, M.A. Diversidade genética e mapeamento da soja por meio de marcadores moleculares AFLP. Viçosa, 1999. 88p. Tese (Doutorado)-Universidade Federal de Viçosa.
- MACKILL, D.J. Classifying japonica rice cultivars with RAPD markers. **Crop Science**, v.35, p. 889-894, 1995.
- MAHESWARAN, M.; SUBUDHI, P.K.; NANDI, S.; XU, J.C.; PARCO, A.; YANG, D.C.; HUANG, N. Polymorphism, distribution and segregation of AFLP markers in a doubled haploid rice population. **Theoretical and Applied Genetics**, v.94, p. 39-45. 1997.
- MALÉCOT, G. **Les mathématiques de l'hérédité**. Paris: Masson et Cie, 1947.

- MANJARREZ- SANDOVAL, P.; CARTER JÚNIOR, T. E.; WEBB, D.M.; BURTON, J. W. RFLP Genetic Similarity Estimates and Coefficient of Parentage as Genetic Variance Predictors for Soybean Yield. **Crop Science**, v.37, p. 698-703, 1997.
- MAUGHAN, P.J.; SAGHAI-MAROOF, M.A.; BUSS, G.R. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance and near-isogenic line analysis. **Theoretical and Applied Genetics**, v.93, p.392-401, 1996.
- McGRATH, J.M.; DERRICO, C.A. ; YU, Y. Genetic diversity in selected, historical US sugarbeet germoplasm and *Beta vulgaris* ssp. *maritima* . **Theoretical and Applied Genetics**, v.98, p. 968-967, 1999.
- MEKSEM, K.; LEISTER, D.; PELEMAN, J.; ZABEAU, M.; SALAMIMI, F.; GEBHARDT, C. A high-resolution map of the vicinity of the R1 locus on the chromosome V of potato based on RFLP and AFLP markers. **Molecular and General Genetics**, v.249, p. 74-81, 1995.
- MELCHINGER, A.E.; MESSMER, M.M.; LEE, M.; WOODMAN, W.L.; LAMKEY, K.R. Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. **Crop Science**, v.31, p. 669-678, 1991.
- MENKIR, A. ; GOLDSBROUGH, P.; EJETA, G. RAPD based assessment of genetic diversity in cultivated races of sorghum. **Crop Science**, v.37, p. 564-569, 1997.
- MESSMER, M. M.; MELCHINGER, A.E.; HERMANN, R.G.; BOPPENMAIER, J. Relationships among early European maize inbreds: II. comparison of pedigree and RFLP data. **Crop Science**, v.3, p. 944-950, 1993.
- MESSMER M. M., MELCHINGER A. E., LEE M., LAMKEY K. R. Genetic diversity among progenitors and elite lines from the Iowa stiff stalk synthetic (BSSS) maize population: comparison of allozyme and RFLP data **Theoretical and Applied Genetics**, v.83, p.97-107, 1991.

- MOSER, H.; LEE, M. RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. **Theoretical and Applied Genetics**, v. 87, p.947-956, 1994.
- MUEHLBAUER, G.J.; SPECHT, J.E.; THOMAS-COMPTON, M.A.; STASWICK, P.E.; BERNARD, R.L. Near-isogenic lines - A potential resource in the integration of conventional and molecular marker linkage maps. **Crop Science**, v.28, n.5, p.729-735, 1988.
- NEI, M.; LI, W.H. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proceedings of the National Academy of Sciences of the United States of America**, v.76, p. 5269-5273, 1979.
- O'DONOUGHUE, L.S. Molecular markers as tools to characterize and use genetic resources in oat. In: REUNIÃO DA COMISSÃO BRASILEIRA DE PESQUISA DE AVEIA, 15, Pelotas, 2000. **Anais**. Pelotas: Universidade Federal de Pelotas, 2000. p. 2-15.
- PEJIC, I.; AJMONE-MARSAN, P.; MORGANTE, M.; KOZUMPLICK, V.; CASTIGLIONI, P.; TARAMINO, G.; MOTTO, M. Comparative analysis of genetic similarity among inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. **Theoretical and Applied Genetics**, v.97, p. 1248-1255, 1998.
- PHILLIPS, R.L. Unconventional sources of genetic diversity: *De Novo* variation and elevated epistasis. In: Borém, A. (Ed.) **Plant breeding in the turn of the millenium**. Viçosa. 1999. 379 p.
- POWELL, W.; MORGANTE, M.; ANDRÉ, C.; HANAFEY, M.; VOGEL, J.; TINGEY, S.; RAFALSKI, A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. **Molecular Breeding**, v.2, p. 225-238, 1996.
- PRABHU, R.R.; WEBB, D.; JESSEN, H.; LUK, S.; SMITH, S.; GRESSHOFF, P.M. Genetic relatedness among soybean genotypes using DNA amplification fingerprinting (DAF), RFLP and pedigree. **Crop Science**, v.37, p. 1590-1595, 1997.

- ROHLF, F.J. **NTSYS-pc: Numerical taxonomy and multivariate analysis system.** New York: EXETER Software, 1997.
- RONGWEN, J.; AKKAYA, M.S.; BHAGWAT, A.A.; LAVI, U.; CREGAN, P.B. The use of microsatellite DNA markers for soybean genotype identification. **Theoretical and Applied Genetics**, v.90, p.43-48, 1995.
- RUBIN, S. A. L.; SANTOS, O.S. Progresso do melhoramento genético da soja no Estado do Rio Grande do Sul: I. Rendimento de grãos. **Pesquisa Agropecuária Gaúcha**, v.2, n.2, p.139-147, 1996.
- SAGHAI-MAROOF, M.A .; SLOMAN, K.M.; JORGENSEN, R.A .; ALLARD, R.W. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. **Proceedings of the National Academy of Sciences of the United States of America**, v.81, p. 8014-8018, 1984.
- SANTOS, J.B. Emprego de marcadores moleculares no melhoramento de plantas. **Horticultura Brasileira**, v. 12, n.2, p. 282-286, 1994.
- SAS INSTITUTE. **SAS/STAT user's guide for personal computers**, version 6, Cary, 1987.
- SENIOR, M.L.; MURPHY, J.P.; GOODMAN, M.M.; STUBER, C.W. Utility of SSRs for determining genetic similarities and relationships in maize using na agarose gel system. **Crop Science**, v.38, p. 1088-1098, 1998.
- SHARMA, S.K.; KNOX, M.R.; ELLIS, T.H.N. AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. **Theoretical and Applied Genetics**, v.93, p.751-758, 1996.
- SMITH, J.S.C.; CHIN, E.C.L.; SHU, H.; SMITH, O.S.; WALL, S.J.; SENIOR, M.L.; MITCHELL, S.E.; KRESOVITCH, S.; ZIEGLE, J. An evaluation of the utility of SSR loci as molecular markers in maize: Comparisons with data from RFLPs and pedigree. **Theoretical and Applied Genetics**, v.95, p. 163-173, 1997.

- SMITH, J.S.C.; SMITH, O.S. Restriction fragment length polymorphisms can differentiate among U.S. maize hybrids. **Crop Science**, v.31, p.893-899, 1991.
- SMITH, O.S.; SMITH, J.S.C.; BOWEN, S.L.; TENBORG, R.A., WALL, S.J. Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yields, heterosis, and RFLPs. **Theoretical and Applied Genetics**, v.80, p. 833-840, 1990.
- SNELLER, C.H. Pedigree analysis of elite soybean lines. **Crop Science**, v.34, p.515-1522, 1994.
- SNELLER, C.H.; MILES, J.; HOYT, J.M. Agronomic performance of soybean plant introduction and their genetic similarity to elite lines. **Crop Science**, v.37, p.1595-1600, 1997.
- SOKAL, R.R.; MICHENER, C.D. A statistical method for evaluating systematic relationships. **University Kansas Science Bulletin**, v.38, p.1409-1438, 1958.
- SOUZA, E.; SORRELLS, M.E. Prediction of progeny variation in oat from parental genetic relationships. **Theoretical and Applied Genetics**, v.82, p.233-241, 1991.
- SPECHT, J.E.; WILLIAMS, J.H. Contribution of genetic technology to soybean productivity – retrospect and prospect. In: FEHR, W.R. (Ed.) **Genetic contributions to yield gains in five major crop plants**. Madison: Crop Science Society of America, 1984. p.49-74. (Special Publication, 7).
- STUBER, C.W.; MOLL, R.H. Frequency changes of isozyme alleles in a selection experiment for grain yield in maize (*Zea mays* L.) **Crop Science**, v.12, p. 337-340, 1972.
- STUBER, C.W.; MOLL, R.H., GOODMAN M.M., SCHAFFER H.E., WEIR B.S. Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea mays* L.) **Genetics**, v.95, p. 225-236, 1980.

- SUSTAR-VOZLIC, J.; JAVORNIK, B. Genetic relationships in cultivars of hop, *Humulus lupulus* L., determined by RAPD analysis. **Plant Breeding**, v.118, p. 175-181, 1999.
- TAKEDA, C.; HIROMOTO, D.M.; ALMEIDA, L.A.; KIIHL, R.A.S.; YORINORY, J.T. Cultivar de soja MT/BR-50 (Parecis) In: CONGRESSO BRASILEIRO DE SOJA, 1, Londrina,1999. **Anais**. Londrina: Embrapa Soja, 1999. p.471. (Embrapa Soja. Documentos, 124).
- THOMAS, M.R.; SCOTT, W.S. Microsatellite repeats in grapevine reveal DNA polymorphisms when analyzed as sequence-tagged sites. **Theoretical and Applied Genetics**, v.86, p. 985-990, 1993.
- THOMPSON, J.A.; NELSON, R.L. Utilization of diverse germplasm for soybean yield improvement. **Crop Science**, v.38, p.1362-1368, 1998.
- THOMPSON, J.A.; NELSON, R.L.; VODKIN, L.O. Identification of diverse soybean germplasm using RAPD markers. **Crop Science**, v.38, p.1348-1355, 1998.
- TINKER, N.A.; FORTIN, M.G.; MATHER, D.E. Random amplified polymorphic DNA and pedigree relationships in spring barley. **Theoretical and Applied Genetics**, v.85, p. 976-984, 1993
- TIVANG, J.G.; NIENHUIS, J.; SMITH, O.S. Estimation of sampling variance of molecular marker data using bootstrap procedure. **Theoretical and Applied Genetics**, v.89, p. 259-264, 1994.
- TOHME, J.; GONZALEZ, D.O.; BEEBE, S.; DUQUE, M.C. AFLP analysis of gene pools of a wild bean core collection. **Crop Science**, v.36, p.1375-1384, 1996.
- TOLEDO, J.F.F. Programa de computador para estimar parâmetros genéticos, componentes de médias e variâncias pelo método dos quadrados mínimos ponderados. **Pesquisa Agropecuária Brasileira**, v. 26, n.7, p.1023-1029, 1991.

TOLEDO, J.F.F. Mid parent and coefficient of parentage as predictors for screening among single crosses for their inbreeding potential. **Revista Brasileira de Genética**, v.15, p. 429-437, 1992.

TOLEDO, J.F.F. Predicting the inbreeding and the outcrossing potential of soybean varieties. **Revista Brasileira de Genética**, v.10, p.543-558, 1987.

TOLEDO, J.F.F.; ALMEIDA, L.A .; KIIHL, R.A.S. ;MENOSSO, O.G. Ganho genético em soja no Estado do Paraná, via melhoramento. **Pesquisa Agropecuária Brasileira**, Brasília, v.25, n.1, p.89-94, 1990.

TRILLER,C.; TOLEDO, J.F.F. Using the F<sub>3</sub> generation for predicting the breeding potential of soybean crosses. **Revista Brasileira de Genética**, v.19, n.2, p.289-294, 1996.

VELLO, N.A.; HIROMOTO, D.M.; AZEVEDO-FILHO, A.J.B.V. Coefficient of parentage and breeding of Brazilian soybean germplasm. **Revista Brasileira de Genética**, v.11, p. 679-697, 1998.

VIERLING, R.A.; NGUYEN, H. T. Use of RAPD markers to determine the genetic diversity of diploid, wheat genotypes. **Theoretical and Applied Genetics**, v.84, p. 835-838, 1992.

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

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

XU, Z.; QIU, L.; CHANG, R.; LI, X.; GUO, P.; ZHENG, C. Using SSR markers to evaluate genetic diversity of soybean in China In: **WORLD SOYBEAN RESEARCH CONFERENCE**, 6, Chicago, 1999. **Proceedings**. p. 510-511.

- YANG, G.; SAGHAI-MAROOF, M.A.; XU, C.G.; ZHANG, Q.; BIYASHEV, R.M. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. **Molecular and General Genetics**, v.245, p. 187-194, 1994.
- YEE, E.; KIDWELL, K.K.; SILLS, G.R.; LUMPKIN, T.A. Diversity among selected *Vigna angularis* (Azuki) accessions on the basis of RAPD and AFLP markers. **Crop Science**, v.39, p. 268-275, 1999.
- YOUNG, W.P.; SCHUPP, J.M.; KEIM, P. DNA methylation and AFLP marker distribution in the soybean genome. **Theoretical and Applied Genetics**, v.99, p. 785-790, 1999.
- ZHANG, L.H.; OZIAS-AKINS, P.; KOCHERT, G.; KRESOVICH, S.; DEAN, R.; HANNA, W. Differentiation of bermudagrass (*Cynodon* spp.) genotypes by AFLP analyses. **Theoretical and Applied Genetics**, v.98, p. 895-902, 1999.
- ZHU, J.; GALE, M.D.; QUARRIE, S.; JACKSON, M.T.; BRYÁN, G.J. AFLP markers for the study of rice biodiversity. **Theoretical and Applied Genetics**, v.96, p. 602-611, 1998.
- ZHU, S.L.; MONTI, L.M.; AVITABILE, A.; RAO, R. Evaluation of genetic diversity in Chinese soyabean germplasm by AFLP. **Plant Genetic Resources Newsletter**, n.119, p.10-14, 1999.
- ZUFFO, N.L.; ALMEIDA, L.A.; OTUBO, T.; KIIHL, R.A.S.; MORCELI, A.A. Cultivar de soja BRS/MS (Campo Grande). In: CONGRESSO BRASILEIRO DE SOJA, 1, Londrina, 1999. **Anais**. Londrina: Embrapa Soja, 1999. p.467. (Embrapa Soja. Documentos, 124).