**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Variabilidade e divergência genética em cruzamentos
dialélicos de soja**

Rosa María Alvarez Parra

Tese apresentada para obtenção do título de Doutor
em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

**Piracicaba
2011**

Rosa María Alvarez Parra
Engenheiro Agrônomo

**Variabilidade e divergência genética em cruzamentos
dialélicos de soja**

versão revisada de acordo com a resolução CoPGr 5890 de 2010

Orientador:
Prof. Dr. **ISAIAS OLÍVIO GERALDI**

Tese apresentada para obtenção do título de Doutor em
Ciências. Área de concentração: Genética e
Melhoramento de Plantas

**Piracicaba
2011**

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/ USP

Alvarez Parra, Rosa María

Variabilidade e divergência genética em cruzamentos dialélicos de soja / Rosa María Alvarez Parra. - - versão revisada de acordo com a resolução CoPGr 5890 de 2010. - - Piracicaba, 2011.
98 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2011.

1. Cruzamento dialélico 2. Diversidade genética 3. Marcador molecular 4. Soja
5. Variação genética em plantas I. Título

CDD 633.34
A473v

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

À minha mãe **Rosa Parra Torrelles**,
Às minhas filhas **Dalia Alexandra e Rosa Alejandra**,
Ao meu esposo **Pedro José García**,
À minha avó **Barbara Rosa**,
Aos meus irmãos **Elvis, Carlos, Israel e Richard**.

DEDICO

À **Dalia Rosa Paez**,
...pelo amor e os sábios conselhos quando ainda estava conosco.

OFEREÇO

AGRADECIMENTOS

A Deus, pelas bênçãos recebidas na vida, assim como pela força, serenidade e esperança que me dá a cada dia.

Ao Instituto Nacional de Investigaciones Agrícolas (INIA), Venezuela, pelo financiamento oferecido.

À Fundación Nacional del Arroz (FUNDARROZ) pela ajuda financeira oferecida, em especial a seu presidente Ing. Pedro Luís Cordero.

À Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ/USP), pela oportunidade fornecida para realizar meus estudos.

Ao professor Isaias Olívio Geraldi, pela oportunidade de fazer parte de sua equipe de estudantes. Agradeço ao senhor pela aprendizagem recebida.

Ao Pedro José García, esposo e companheiro, pelo incentivo para voltar à casa de estudos e pelo apoio e amor fornecido.

Às minhas filhas, Dália Alexandra e Rosa Alejandra por existir, e ser o presente mais precioso que Deus me deu.

À minha família, pelo apoio incondicional para o alcance desta meta, especialmente minha mãe Rosa e minha irmã Elvis Nailander.

À Gheisa Marielys García pela ajuda proporcionada neste tempo, que tornou possível em grande parte o alcance da meta.

À minha amiga e colega Edicta Rosa Reyes, por todo o apoio oferecido.

Ao Edgar Torres pela amizade e auxílio oferecido nos momentos mais difíceis.

Ao Fredy Vegas, Pastora Querales e Carlos Vásquez, pelo afeto e colaboração.

Aos meus colegas de estudo Leandro Augusto Andrade Fumes, Fernando Hoshino Shirahige, e em especial a Guilherme José Farias e Larissa Pereira de Castro, pela ajuda, amizade e coleguismo.

Aos funcionários do laboratório, Gustavo Alexandre Perina e Fernandes de Araújo, pela colaboração no trabalho de campo.

A todos os que, em algum momento ofereceram ajuda para superar as dificuldades.

Grata para sempre, que Deus abençoe a todos e cada um de vocês.

Muitas coisas não ousamos empreender por parecerem difíceis; entretanto, são difíceis porque não ousamos empreendê-las.

Sêneca

SUMÁRIO

RESUMO.....	11
ABSTRACT	13
LISTA DE FIGURAS	15
LISTA DE TABELAS	17
1 INTRODUÇÃO	19
2 DESENVOLVIMENTO	21
2.1 Revisão bibliográfica	21
2.1.1 Características da soja	21
2.1.2 Origem e domesticação	22
2.1.3 Diversidade genética	23
2.1.4 Melhoramento genético da soja.....	26
2.1.5 Predição do potencial de cruzamentos para obtenção de linhagens superiores	28
2.1.5.1 Base genética da predição do potencial do cruzamento	29
2.1.5.2 Metodologias utilizadas para realizar predições de populações endogâmicas.....	31
2.1.5.3 Resultados das predições do comportamento de populações endogâmicas	36
2.2 Material e métodos.....	43
2.2.1 Material genético.....	43
2.2.2 Obtenção do material experimental	45
2.2.2.1 Obtenção das progênies F_1 e F_2	45
2.2.2.2 Obtenção das progênies $F_{2:3}$	45
2.2.3 Avaliação experimental das progênies $F_{2:3}$	45
2.2.4 Caracterização do ambiente experimental.....	46
2.2.5 Caracteres avaliados	46
2.2.6 Análises estatístico-genéticas.....	47
2.2.6.1 Análises de variância.....	47
2.2.6.2 Estimativas dos componentes de variância.....	48
2.2.6.3 Estimativas dos coeficientes de herdabilidade	50

2.2.6.4 Predição da resposta à seleção	50
2.2.7 Desdobramento das distâncias genéticas em DGG e DGE	53
2.2.8 Correlações.....	54
2.3 Resultados e discussão	55
2.3.1 Análise de variância das progênes $F_{2:3}$	55
2.3.1.1 Produção de grãos (PG)	55
2.3.1.2 Altura das plantas na maturação (AM).....	57
2.3.2 Análises das médias.....	59
2.3.3 Estimativas das variâncias genéticas entre progênes $F_{2:3}$	62
2.3.4 Estimativas dos coeficientes de herdabilidade	66
2.3.5 Resposta à seleção	68
2.3.6 Desdobramento das distâncias genéticas em distância genética geral e distância genética específica.....	71
2.3.7 Associações entre as estimativas de parâmetros e as distâncias genéticas.....	74
3 CONSIDERAÇÕES FINAIS	83
REFERÊNCIAS	87

RESUMO

Variabilidade e divergência genética em cruzamentos dialélicos de soja

As estimativas de divergência genética (DG) são pouco utilizadas pelos programas de melhoramento na seleção de genitores, pois os resultados das predições têm sido inconsistentes. O objetivo deste trabalho foi avaliar a possibilidade de prever a variabilidade genética (VG) de populações derivadas de cruzamentos biparentais de soja, utilizando dados de DG obtidos com marcadores AFLP (Bonato et al. 2006) e também avaliar a possibilidade de melhorar a associação entre DG e VG com o desdobramento das distâncias genéticas de acordo com um arranjo dialélico, isto é, em distância genética geral (DGG_i e DGG_j) e distância genética específica (DGE_{ij}). Para isso, utilizaram-se 10 populações oriundas de um dialélico com genitores adaptados e com diferentes níveis de divergência genética (DG): baixa ($DG \leq 0,2$), média ($DG \cong 0,4$) e alta ($DG \cong 0,6$). Cada população foi avaliada utilizando 100 progênies $F_{2:3}$ no delineamento experimental em látice simples 10×10 (duas repetições), durante o ano agrícola 2009/10. Foram avaliados os caracteres produção de grãos (PG) e altura das plantas na maturação (AM). Para cada cruzamento foram estimados: variância genética entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$), média das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), amplitude das médias das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), coeficiente de herdabilidade entre médias de progênies $F_{2:3}$ ($\hat{h}_x^2$) e respostas à seleção (Rs) com intensidades de 20% e 40%. Para PG, as correlações de Spearman de DG com $\hat{\sigma}_p^2$, $\hat{h}_x^2$ e Rs foram positivas, altas e significativas ($P < 0,05$), indicando um aumento na magnitude destas estimativas com o aumento da divergência genética (DG). Para AM as correlações foram altas e significativas somente entre DG e $\hat{h}_x^2$ ($P < 0,05$). Em geral, o desdobramento de DG não causou um aumento na associação com as estimativas de parâmetros, exceto para a correlação de DG com $\bar{F}_{2:3}$ e $\Delta\bar{F}_{2:3}$, em ambos os caracteres. Os resultados sugerem que as divergências genéticas entre os genitores determinadas pelos marcadores AFLP utilizadas neste trabalho refletem a real DG para o caráter PG, uma vez que os mesmos foram consistentes com os resultados obtidos em trabalho anterior, utilizando as mesmas estimativas de DG. Consequentemente, essas estimativas de DG poderiam auxiliar na escolha de genitores de alto potencial para a produção de grãos nos programas de melhoramento genético de soja.

Palavras-chaves: Soja; Divergência Genética; Marcadores AFLP; Predição de variabilidade.

ABSTRACT

Variability and genetic divergence in soybean diallel crosses

Estimates of genetic divergence (DG) have not been used for parental selection in breeding programs, because of the inconsistency of the results. The objective of this study was to evaluate the possibility of predicting the genetic variability (VG) of soybean populations derived from two-way crosses, using DG data from AFLP markers (Bonato et al., 2006), and also to evaluate the possibility of improving the relationship between DG and VG by partitioning the genetic distance (GD) in general genetic distance (DGG_i and DGG_j) and specific genetic distance (DGE_{ij}), according to a diallel scheme. Ten populations derived from a diallel composed by adapted parents with different levels of genetic divergence (GD) were used: low ($DG \leq 0,2$), medium ($DG \cong 0,4$) and high ($DG \cong 0,6$). One hundred $F_{2:3}$ progenies of each cross were evaluated in a simple lattice 10 x 10 design (two replications) during the 2009/10 growing season for the traits grain yield (PG) and plant height at maturity (AM). Genetic variance among $F_{2:3}$ progenies ($\hat{\sigma}_p^2$), $F_{2:3}$ progeny mean ($\bar{F}_{2:3}$), amplitude of individual $F_{2:3}$ progeny means ($\Delta\bar{F}_{2:3}$), heritability among $F_{2:3}$ progeny means ($\hat{h}_{\bar{x}}^2$), and responses to selection (Rs), with intensities of 20% and 40%, were estimated for each cross. For PG, the Spearman correlation between DG and $\hat{\sigma}_p^2$, $\hat{h}_{\bar{x}}^2$ and Rs (20% and 40%) were positive, high and significant ($P < 0.05$), indicating an increase of the parameter estimates following the genetic divergence (DG) increasing. For AM, the correlation were high and significant ($P < 0.05$) only for DG and $\hat{h}_{\bar{x}}^2$. In general, the partitioning of DG did not increase the association with the parameters estimates, except for the correlation between DG with $\bar{F}_{2:3}$ and $\Delta\bar{F}_{2:3}$ for both traits. General results suggest that AFLP genetic divergence between the parents considered in this study expresses the true DG for PG, since they were consistent with the results of a previous work, using the same estimates of DG. Thus, these estimates of genetic divergence could be used in order to select high potential parents for grain yield in soybean breeding programs.

Keywords: Soybean; Genetic Divergence; AFLP Markers; Predicting variability

LISTA DE FIGURAS

Figura 1 - Histogramas das médias individuais das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), dos 10 cruzamentos para produção de grãos (PG). Eixo Y: Frequências; Eixo X: Produção de grãos em g.m^{-2}

78

LISTA DE TABELAS

Tabela 1 - Cultivares utilizados e distâncias genéticas entre eles, obtidas através do marcador AFLP	43
Tabela 2 - Informação genealógica dos cultivares de soja.....	44
Tabela 3 - Esquema da análise de variância com as respectivas esperanças dos quadrados médios.....	48
Tabela 4 - Quadrados médios (QM) e respectivas significâncias (teste F) das análises de variância das progênes $F_{2:3}$ para o caráter produção de grãos (PG) de cada cruzamento, coeficientes de variação experimental (CV%) e médias das progênes $F_{2:3}$ ($\bar{F}_{2:3}$), em g.m ⁻²	56
Tabela 5 - Graus de liberdade das análises de variância das progênes $F_{2:3}$ para o caráter produção de grãos de cada cruzamento	56
Tabela 6 - Quadrados médios (QM) e respectivas significâncias (teste F) das análises de variância das progênes $F_{2:3}$ para o caráter altura das plantas na maturação (AM) de cada cruzamento, coeficientes de variação experimental (CV%) e médias ajustadas das progênes $F_{2:3}$ ($\bar{F}_{2:3}$), em cm	58
Tabela 7 - Graus de liberdade das análises de variância das progênes $F_{2:3}$ para o caráter altura das plantas na maturação de cada cruzamento.....	58
Tabela 8 - Médias dos genitores ($\bar{P}_i$ e $\bar{P}_j$), média das progênes $F_{2:3}$ ($\bar{F}_{2:3}$), média da progênie $F_{2:3}$ superior ($\bar{F}_{2:3(s)}$), média da progênie $F_{2:3}$ inferior ($\bar{F}_{2:3(i)}$) e amplitudes das médias individuais das progênes $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), de cada cruzamento para os caracteres produção de grãos (PG) e altura das plantas na maturação (AM)	60
Tabela 9 - Estimativas das variâncias genéticas entre progênes $F_{2:3}$ ($\hat{\sigma}_p^2$) de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança associados às estimativas, variâncias fenotípicas entre médias de progênes $F_{2:3}$ ($\hat{\sigma}_F^2$) e distâncias genéticas (DG), para o caráter produção de grãos (PG)	63

Tabela 10 - Estimativas das variâncias genéticas entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$) de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança associados às estimativas, variâncias fenotípicas entre médias de progênies $F_{2:3}$ ($\hat{\sigma}_F^2$) e distâncias genéticas (DG), para o caráter altura das plantas na maturação (AM)..... 65

Tabela 11 - Estimativas do coeficiente de herdabilidade, no sentido amplo, entre médias de progênies $F_{2:3}$ ($\hat{h}_x^2$) de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança e distâncias genéticas (DG), para todos os caracteres avaliados..... 67

Tabela 12 - Estimativas das médias dos genitores ($\bar{Y}_g$), médias das progênies $F_{2:3}$ originais ($\bar{Y}_0$), médias das progênies $F_{2:3}$ selecionadas ($\bar{Y}_s$), diferencial de seleção (ds), resposta esperada com a seleção (Rs), médias da população ao final do processo de endogamia F_n ($\bar{Y}_m$) e distâncias genéticas (DG), para a produção de grãos (PG), considerando intensidades de seleção de 20% e 40% 69

Tabela 13 - Distâncias genéticas (DG: acima da diagonal), distâncias genéticas geral (DGG), distância genética média ($\overline{DG}$), e distâncias genéticas específicas (DGE_{ij}: abaixo da diagonal), calculadas a partir de marcadores AFLP e desdobradas de acordo com o método de Griffing (1956)..... 72

Tabela 14 - Correlações de Spearman (r_s) entre as distâncias genéticas determinadas por marcadores AFLP (DG), e as estimativas dos demais parâmetros [$r(DG, Y)$]: estimativa da variância genética entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$); média das progênies $F_{2:3}$ ($\bar{F}_{2:3}$); amplitude das médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), estimativas do coeficiente de herdabilidade, no sentido amplo, entre médias de progênies $F_{2:3}$ ($\hat{h}_x^2$) e, estimativas de resposta à seleção (Rs) com intensidades de seleção de 20% e 40% para os caracteres produção de grãos (PG) e altura das plantas na maturação (AM)..... 75

1 INTRODUÇÃO

Os programas de melhoramento de plantas têm por objetivo gerar novos cultivares capazes de ter, em plantios comerciais, alta produção com menor custo. Para ter sucesso, é de vital importância contar com populações que possuam médias adequadas e com suficiente variabilidade genética a fim de identificar genótipos com as características desejadas (COX et al., 1985). Devido ao fato que o número de genes segregando para um determinado caráter define tanto a variância genética quanto a divergência genética entre genótipos, o conhecimento da distância genética entre os genitores constitui uma ferramenta que pode ser utilizada no planejamento de cruzamentos. O uso da distância genética baseada no conhecimento da genealogia (*pedigree*), tal como o coeficiente de parentesco, é uma opção; no entanto, seu uso tem gerado resultados contraditórios, e nem sempre estes estão disponíveis. Uma alternativa que tem sido proposta é o uso da distância genética baseada em algum tipo de marcador genético, principalmente os marcadores moleculares, os quais têm a vantagem de não serem influenciados pelo ambiente. Entretanto, as predições do comportamento de populações endogâmicas têm se mostrado poucos consistentes e, portanto, são pouco usadas pelos programas de melhoramento genético.

Se ponderar o custo e o tempo requeridos para o desenvolvimento de novos cultivares de soja, as pesquisas cuja finalidade é a procura da melhoria da capacidade preditiva são relevantes, devido ao impacto que poderiam ter no aumento da eficiência do processo, especialmente se o método permite a escolha *a priori*, que é mais facilmente aceito e posto em prática pelos melhoristas da cultura. Espera-se que a baixa utilidade observada até hoje seja revertida, à medida que prossigam as pesquisas nessa área e sejam superados os obstáculos que estão interferindo na consistência das predições, assim como no convencimento dos melhoristas sobre a utilidade dessa informação.

Considerando que na pesquisa realizada por Bonato et al. (2006) foram calculadas as distâncias genéticas entre 317 cultivares de soja liberados no Brasil entre 1962 e 1998, utilizando o marcador AFLP, e ainda que estes cultivares

certamente estão presentes no conjunto de genitores dos diferentes programas de melhoramento de soja no Brasil, há a possibilidade de que essas estimativas possam ser incorporadas no processo de seleção de genitores, se elas estiverem refletindo as divergências genéticas para os caracteres de interesse do programa de melhoramento. Caso essas estimativas possuam capacidade preditiva sobre as variâncias genéticas das populações endogâmicas, elas poderiam impactar positivamente na eficiência dos programas de melhoramento, poupando custo na obtenção de novos cultivares. Nesse contexto, Colombari Filho (2009) avaliou a associação entre distância genética e variância genética em seis cruzamentos biparentais de soja, utilizando as distâncias genéticas obtidas por Bonato et al. (2006). Os resultados mostraram a existência de alta correlação entre distância genética e variância genética em populações $F_{2:3}$, coeficiente de herdabilidade e amplitude das médias das progênies $F_{2:3}$, para os caracteres produção de grãos, dias para maturação e altura no florescimento. No entanto, esses resultados devem ser considerados com cautela, pois foi obtida pela análise de apenas seis populações, o que implica na necessidade de aprofundar a pesquisa.

Nesse sentido, apresenta-se este trabalho, que tem por objetivo geral, verificar a possibilidade de utilização dos dados de divergência genética com marcadores AFLP, obtidos por Bonato et al. (2006), para auxiliar na seleção de genitores em programas de cruzamentos de soja. Para isso, utilizaram-se 10 populações provenientes de um cruzamento dialélico, derivados de genitores com diferentes graus de divergência genética. Os objetivos específicos compreenderam: (i) estimar as correlações entre as distâncias genéticas dos genitores (DG) e a variabilidade genética gerada nos cruzamentos, médias, amplitude das médias, coeficiente de herdabilidade e respostas à seleção; e (ii) avaliar a viabilidade do uso dos componentes das distâncias genéticas (DG) desdobrados de acordo com um arranjo dialélico, isto é, distância genética geral (DGG_i e DGG_j) e distância genética específica (DGE_{ij}).

2 DESENVOLVIMENTO

2.1 Revisão bibliográfica

2.1.1 Características da soja

O gênero *Glycine* é um paleopoliplóide que pertence à divisão Magnoliophyta, classe Magnoliopsida, Sub-classe Rosidae, ordem Fabales, família Fabaceae e tribo Phaseolae (CAPELLARI JÚNIOR; RODRIGUEZ; SOUZA, 2007). Está dividido em dois subgêneros: o subgênero *Soja* no qual está a soja anual cultivada (*Glycine max* (L.) Merrill) junto a seu antecessor silvestre (*Glycine soja* Sieb. and Zucc.), e o subgênero *Glycine* com mais de 20 espécies de soja perenes. Com exceção de uma espécie do subgênero *Glycine*, a qual tem $2n = 38$, as espécies restantes de *Glycine* de ambos os subgêneros têm 20 pares de cromossomos ($2n = 40$). Devido ao fato de possuir um genoma com o dobro de cromossomos da maioria dos membros da tribo *Phaseoleae*, além de evidências de duplicações antigas do genoma, acredita-se que a soja é um tetraplóide, embora do ponto de vista citogenético se comporte como um diplóide e tenha herança dissômica para a maioria dos genes (DOYLE et al., 2004).

Os eventos de duplicação e diploidização do genoma da soja foram múltiplos e ocorreram, aproximadamente, há 14 e 45 milhões de anos. Esse processo gerou um genoma complexo caracterizado por duplicações regionais e rearranjos dos cromossomos. Sabe-se que os mais de 46.000 genes da soja cultivada estão agrupados em blocos de famílias, com uma quantidade média de 75 genes por bloco, sendo que 21% dos genes se encontram em zonas repetitivas e ricas em *transposons* nos centros dos cromossomos. A presença dos genes não é regular nos cromossomos; estima-se que 61,4% dos genes homólogos se encontram em blocos presentes em só dois cromossomos, 5,6% presentes em três cromossomos e 21,5% presentes em quatro cromossomos. Além disso, foi comprovada a existência de cromossomos altamente fragmentados com blocos presentes em outros cromossomos e o *splicing* alternativo em genes importantes, tais como aqueles

relacionados à produção de grãos (WALLING; PIRES; JACKSON, 2005; SCHMUTZ et al., 2010).

2.1.2 Origem e domesticação

A soja cultivada é uma leguminosa originária da China. Estudos baseados em análises com marcadores SSR e AFLP com raças locais de soja (*G. max*) e soja silvestre (*G. soja*), sugerem que áreas meridionais da China poderiam ser o centro de origem da soja, pelo fato de a divergência genética estar associada à latitude, e ser maior nos grupos de acessos provenientes de regiões de baixas latitudes (AN et al., 2009). A domesticação foi iniciada há cerca de cinco mil anos pelos chineses a partir da soja silvestre (*G. soja*), porém a área geográfica exata onde a domesticação ocorreu ainda não foi determinada com certeza. Atualmente, existem diversas teorias que propõem como possível lugar de domesticação a região Norte da China (HYMOWITZ, 1970), a região Nordeste da China (FUKUDA, 1933), o vale do Rio Amarelo (XU, 1986), o vale do Rio Yangtze e áreas do Sul (WANG, 1985), ou a área multicêntrica (LU, 1978; AN et al., 2009). O processo de domesticação começou com plantas rasteiras anuais e em sua evolução originaram os cultivares primitivos, os quais foram, em geral, geneticamente isolados em áreas geográficas relativamente pequenas (CHEN; NELSON, 2005). Isto produziu alguns efeitos de gargalo severos que causaram, na soja cultivada, a perda de mais de 80% dos alelos raros presentes na soja silvestre (HYTEN et al., 2006).

As primeiras citações sobre a soja aparecem no período entre 2.883 e 2.838 a.C., quando ela era considerada um grão sagrado, ao lado do arroz, do trigo, da cevada e do milheto. Apesar de ter sido conhecida e consumida pela civilização oriental por milhares de anos, a produção de soja ficou restrita à China até aproximadamente dois mil anos atrás, quando se expandiu para o sul da China, Coreia e Japão, passando a ser a base da alimentação desses povos (KIHARA, 1969). Na Europa, foi introduzida por volta do século XVII, como planta ornamental, e reintroduzida no século XVIII pelos missionários. Em 1765, a soja foi introduzida pela

primeira vez nos Estados Unidos, proveniente da China, e a partir dali, se dispersou pelo continente americano (TAVAUD-PIRRA et al., 2009).

As primeiras tentativas de produção comercial de soja na Rússia, Inglaterra e Alemanha fracassaram, provavelmente devido a fatores climáticos desfavoráveis, além da ausência de conhecimento sobre a cultura e suas exigências. Os norte-americanos conseguiram entre o fim do século XIX e início do século XX, melhorar a soja tornando possível sua produção comercial, criando novas variedades, com teor de óleo mais elevado, o que possibilitou a expansão da cultura pelo mundo (GIZLICE; CARTER; BURTON, 1993a).

2.1.3 Diversidade genética

Geralmente, para o melhorista, a definição de diversidade genética inclui apenas aqueles genes vitais para o melhoramento de caracteres agronômicos, baseando suas decisões principalmente nos cálculos de progresso esperado com a seleção (GIZLICE; CARTER; BURTON, 1993b). Em função disso, não é estranho, então, que a base do germoplasma dos programas de melhoramento de soja dos países produtores da cultura, fora de seu centro de origem, tenha sido formada por um pequeno número de linhagens ancestrais e que seja estreita (DELANNAY; RODGERSDMAN; PALMER, 1983; HIROMOTO; VELLO, 1986; GIZLICE; CARTER; BURTON, 1993a; 1994; GIZLICE et al., 1996; ABDELNOOR; BARROS; MOREIRA, 1995; KISHA; SNELLER; DIERS, 1997; MIRANDA et al., 2007). No entanto, existe outro conceito de diversidade genética utilizado em outros âmbitos, que implica a consideração de todos os genes que controlam os caracteres dentro de cada espécie. Nesse contexto, a variabilidade genética da cultura da soja para caracteres fisiológicos, morfológicos e agronômicos é considerada bastante ampla, embora pouco utilizada nos programas de melhoramento (GIZLICE; CARTER; BURTON, 1993a).

Na China, a diversidade genética dos acessos de soja é alta, sendo a única que varia de acordo com a latitude e a longitude (CHEN; NELSON, 2005; AN et al., 2009; TAVAUD-PIRRA et al., 2009). Estudos feitos com marcadores moleculares (RAPD,

SSR e AFLP) mostram que *G. soja* tem diversidade genética maior que *G. max* (AN et al., 2009) e que as raças locais (*G. max*) provenientes de Ásia têm maior diversidade genética do que aqueles provenientes de diferentes lugares do mundo. Concordando com a dispersão da espécie, a coleção de germoplasma preservado na Europa, apresenta uma diversidade genética intermediária entre o germoplasma asiático e o americano (TAVAUD-PIRRA et al., 2009).

Acredita-se que a base genética estreita do germoplasma de soja dos Estados Unidos, é devida ao pequeno número de ancestrais que formaram sua base e aos subsequentes cruzamentos entre as linhagens elites primárias, durante o processo de melhoramento (KISHA; SNELLER; DIERS, 1997). Nos últimos 50 anos, ocorreu uma redução da base genética original da cultura em aproximadamente um quarto (GIZLICE; CARTER; BURTON, 1993a). Estudos realizados com marcadores moleculares RFLP, AFLP e RAPD determinaram que o polimorfismo varia de 16 a 36% (MAUGHAN; SAGHAI-MAROOF; BUSS, 1996; THOMPSON; NELSON, 1998). Uma característica particular do germoplasma de soja dos Estados Unidos é que ele não é homogêneo, devido às diferenças contrastantes na base genética dos cultivares de soja da região Norte e da região Sul (GIZLICE et al., 1996). No entanto, o alto nível de relação genética entre cultivares foi evidenciado por Gizlice; Carter e Burton (1994) que observaram que 85% dos cultivares de soja liberados no mercado de 1947 a 1988 foram representados por aproximadamente 20 ancestrais. Da mesma forma, Diwan e Cregan (1997) constataram esse fato ao reunir 95% dos alelos presentes nas variedades de soja norte americanas em só 35 genótipos (10,1 alelos/loco).

Devido ao fato que a expansão da cultura de soja na América ocorreu a partir dos Estados Unidos, existe alto grau de parentesco entre variedades e linhagens utilizadas como genitores nos programas de melhoramento de soja dos Estados Unidos, Canadá, México e Brasil (GILL et al., 2006; ALCÂNTARA NETO, 2001). Em alguns casos, como no México, as semelhanças entre cultivares são tão altas que as características morfoagronômicas não discriminam as variedades locais e as introduzidas dos países antes mencionados (QUINTERO et al., 2005).

No Brasil, a formação do germoplasma de soja foi feita principalmente com plantas provenientes do sul dos Estados Unidos (HIROMOTO E VELLO, 1986), e

segundo estudos feitos por Yamanaka et al. (2007), esse germoplasma é completamente diferente do chinês e japonês. Vello, Hiromoto e Azevedo-Filho (1988) calcularam o coeficiente de parentesco de 74 cultivares comerciais de soja liberados no mercado brasileiro nos anos 1983-84, e obtiveram um valor de 16,4%, sendo que apenas 26 genótipos ancestrais eram responsáveis por 80% do acervo genético estudado. Coincidindo com esta tendência, Bonetti (1983) encontrou que aproximadamente 70% dos cultivares desenvolvidos para o estado do Rio Grande do Sul descendiam dos cultivares americanos Hill, Hood, ou de ambos. Abdelnoor, Barros e Moreira (1995), utilizando marcadores RAPD em 38 cultivares de soja, obtiveram coeficientes de similaridade genética de 69% a 100% com média de 82,6%. Bonato et al. (2006), avaliando 317 cultivares de soja brasileiras liberados no mercado durante 1962-1998 por meio de marcadores AFLP, obtiveram coeficientes de similaridade genética de 17% a 97% com média de 61%, sendo a maioria desses coeficientes maiores que 60%.

Contudo, alguns estudos detectam uma variabilidade genética mais ampla. Almeida, Peluzio e Afférri, (2011), estudando 12 cultivares de soja cultivados em várzeas irrigadas, por meio de métodos multivariados, encontraram uma dissimilaridade genética de Mahalanobis variando de 2,6 a 374,1. Peluzio et al. (2009), estudando 14 cultivares de soja cultivados no mesmo sistema, encontraram uma dissimilaridade genética de Mahalanobis variando de 3,5 a 440. Igualmente, Vieira et al. (2009), avaliando 53 cultivares de soja com marcadores microssatélites (SSR), obtiveram distâncias genéticas de 2% a 73%, com média de 47%, enquanto que Priolli et al. (2010) obtiveram coeficientes de similaridades genéticas de 0,01 a 0,9 com média de 0,25, quando avaliaram 168 cultivares de soja com marcadores SSR.

Considerando que a formação do germoplasma base de soja no Brasil foi derivada do germoplasma do Sul dos Estados Unidos, e que os objetivos de melhoramento genético tem variado através do tempo, é interessante observar que nesse germoplasma não se observa um efeito similar ao ocorrido no germoplasma norte-americano. Segundo estudos realizados por Priolli et al. (2004), a diversidade genética no germoplasma de diferentes programas de melhoramento brasileiros

manteve, em geral, um nível constante nos últimos 30 anos, além de relativa heterogeneidade dentro de alguns desses programas. Eles consideram que isso pode ser devido à introdução, ao longo desses anos, de diferentes fontes de germoplasma para resistência a patógenos.

Devido à importância econômica da cultura da soja, a preservação da diversidade genética é feita em bancos de germoplasma por diferentes países. Os maiores bancos de germoplasma de soja estão na China (22.600 acessos, onde 93% correspondem a cultivares primitivos), nos Estados Unidos (17.000 acessos), na Coreia (7.000 acessos), no Brasil (5.000 acessos) e na França (2.000 acessos) (CHEN; NELSON, 2005; TAVAUD-PIRRA et al., 2009; YOON et al., 2009; EMBRAPA, RECURSOS GENETICOS; ALMEIDA et al., 1999a). No Brasil, a coleção de germoplasma está sendo conservada em câmaras climatizadas no Banco Ativo de Germoplasma da Embrapa Soja, em Londrina, Paraná, e uma cópia dela é mantida, em condições de conservação de longo prazo, na Coleção Base da Embrapa Recursos Genéticos, em Brasília (DF). A maioria desses acessos foi introduzida da coleção de germoplasma norte-americana e é composta principalmente por genótipos procedentes da China, do Japão e de outros países onde ocorreu a diversificação da espécie (ALMEIDA et al., 1999a).

2.1.4 Melhoramento genético da soja

A obtenção de cultivares melhorados, com características de alta produtividade, ampla adaptação e boa resistência e/ou tolerância a fatores bióticos e abióticos adversos, é o objetivo geral de um programa de melhoramento genético. A soja foi introduzida no Brasil em 1882 no Estado da Bahia. Em São Paulo, começou a ser cultivada por imigrantes japoneses, por volta de 1908, e foi introduzida oficialmente no Rio Grande do Sul em 1914. Na década de 1940, ainda era cultivada em pequena escala por colonos japoneses e em instituições de pesquisa. Porém, na década 1960, devido ao impulso na produção de suínos e aves no Brasil, foi gerada uma demanda pelo farelo de soja, que tornou a produção da cultura uma ação estratégica. Em meados de 1970, ocorreu uma alta do preço da soja no mercado internacional, em

função da redução da safra de grãos na Rússia e China, e da redução da pesca da anchova no Peru, cuja farinha era utilizada como componente proteico na fabricação de rações para animais; devido a isso, os fabricantes do produto passaram a utilizar o farelo de soja. Essa situação despertou o interesse dos agricultores e do próprio governo brasileiro, ao perceber a vantagem competitiva em relação aos outros países produtores, já que a colheita de soja brasileira ocorre na entressafra americana, quando os preços atingem as maiores cotações. Desde então, o país passou a investir em tecnologia para adaptação da cultura às condições brasileiras, processo liderado pelo Instituto Agrônomo de Campinas (IAC) e a Empresa Brasileira de Pesquisa Agropecuária - Embrapa (EMBRAPA, 2004a; DOMIT et al.; 2007).

Como produto dos avanços do melhoramento da cultura realizados pelo IAC e a Embrapa, pela primeira vez na história, a soja foi plantada com sucesso em regiões de baixas latitudes, entre o trópico de capricórnio e a linha do equador. Essa conquista revolucionou a história mundial da soja e impactou o mercado a partir do final da década de 1980 e mais notoriamente em 1990, quando os preços dos grãos começaram a cair, devido aos incrementos na produção mundial. Atualmente, os líderes na produção mundial de soja são os Estados Unidos, Brasil, Argentina, China, Índia e Paraguai. No Brasil, a soja é cultivada nas regiões Sul e Sudeste, Centro-Oeste, Norte e Nordeste, sendo para isso utilizado um grande número de cultivares melhorados (ABDELNOOR; BARROS; MOREIRA, 1995; ALCÂNTARA NETO, 2001; PRIOLLI et al., 2002).

O tempo requerido para a produção de novos cultivares de soja é de aproximadamente 10 anos no Brasil (DOMIT et al.; 2007), e de 15 a 20 anos, nos Estados Unidos (MIGLIACCI, 2008). A estratégia dos programas de melhoramento de soja no Brasil baseou-se inicialmente em introduções de algumas linhagens do Sul dos EUA, com as quais foram desenvolvidos cultivares mais adaptados às regiões tradicionais de cultivo. Porém, para a região em expansão, seguiu-se a estratégia de desenvolvimento de linhagens adaptadas às baixas latitudes, por meio da incorporação do caráter período juvenil longo (PJL) (BONETTI, 1983; HIROMOTO; VELLO, 1986; PALUDZYSZYN FILHO; SOUZA; ALMEIDA, 1993; PRIOLLI et al., 2004). Como resultado desse trabalho, foram obtidos ganhos genéticos substanciais

através dos anos, em termos de produtividade e em outros caracteres (HIROMOTO; VELLO, 1986). O aumento da produtividade média anual devido aos ganhos genéticos foi estimado entre 1,1% e 1,8% (RUBIN; SANTOS, 1996; TOLEDO et al., 1990).

No âmbito mundial, a produção de novas tecnologias na soja, tanto em cultivares melhorados quanto em manejo agrônomo, impulsionou a produção, que cresceu mais de 200% ao longo dos últimos 30 anos (MIGLIACCI, 2008). No futuro, espera-se o desenvolvimento de novos cultivares com maior potencial de produção que possibilitem cobrir a demanda crescente de soja no mercado mundial. A identificação de genes envolvidos na expressão desse caráter e seu posterior uso nos trabalhos de melhoramento são fundamentais para atingir esse objetivo. Nesse contexto, estão sendo identificados QTLs (*Quantitative Trait Locus*) no genoma da soja, associados ao rendimento da cultura, indicando a existência de alelos com potencial para aumentar o rendimento tanto em germoplasma elite quanto em germoplasma introduzido (GUZMAN et al., 2007).

2.1.5 Predição do potencial de cruzamentos para obtenção de linhagens superiores

Mesmo que na atualidade se disponha de tecnologias que auxiliam o processo de melhoramento genético, o sucesso do programa de melhoramento em obter cultivares de impacto no mercado agrícola não está garantido; assim, hoje continua vigente o fator “arte”, em muitos programas de melhoramento, o que envolve a sorte ou a destreza do melhorista em reconhecer os genótipos superiores. O melhorista tenta aumentar a probabilidade de sucesso realizando o maior número de cruzamentos possíveis por ano, utilizando de maneira geral, genitores adaptados, divergentes, com médias similares e altas, sendo que o número de genitores e nível de caracterização varia entre os diversos programas de melhoramento. Essa situação faz com que a maior parte das populações formadas sejam rejeitadas no futuro, por não gerarem materiais genéticos competitivos. Como exemplo, tem-se que a Embrapa testa anualmente cerca de 4 mil linhagens de soja em ensaios preliminares

de 1º ano, em dois locais; e ao final do ciclo, são selecionadas apenas 15% a 20% das linhagens (ALMEIDA et al., 1999a), ou seja, 3.200 a 3.400 linhagens são descartadas anualmente depois de ter consumido tempo e recursos econômicos consideráveis. Se o melhorista pudesse contar com ferramentas práticas que permitam conhecer previamente o potencial das populações, poderia aumentar a eficiência do programa de melhoramento ao eliminar, nas primeiras etapas, aquelas populações com pouco potencial ou ainda evitar sua formação. Dessa forma, poderia concentrar seus esforços naquelas populações mais promissoras, poupando tempo e recursos financeiros no processo.

2.1.5.1 Base genética da predição do potencial do cruzamento

A obtenção contínua de cultivares que satisfaçam as exigências do mercado depende da disponibilidade de populações com potencial para produzir genótipos transgressivos com caracteres agronômicos desejáveis, capazes de serem competitivos na produção comercial. Segundo Cox et al. (1985), as populações com maiores potenciais são aquelas que têm médias altas e adequadas para os caracteres de interesse, além de alta variabilidade genética. Como os melhoristas, de maneira geral, escolhem cultivares comerciais ou linhagens elites como genitores, as diferentes populações geradas pelos programas de melhoramento em soja terão médias altas e similares, pelo fato que a média de um conjunto de linhagens endogâmicas derivadas de um cruzamento biparental, sem efeito de seleção, é igual à média dos genitores, se não houver epistasia, desequilíbrio de ligação ou efeito de adaptação (THORNE; FEHR, 1970; MOSER; LEE, 1994). Nesse caso, apenas a variância genética permitirá discriminar o potencial das populações (FALCONER; MACKAY, 1996). Segundo Cockerham (1954) em uma população derivada de um cruzamento biparental, a variância genética será composta pela eq. (1),

$$\sigma_G^2 = (1 + I_g)\sigma_A^2 + (1 - I_g^2)\sigma_D^2 \quad (1)$$

em que:

- σ_A^2 : variância genética aditiva;
- σ_D^2 : variância genética dominante; e
- I_g : corresponde ao coeficiente de endogamia na geração da descendência, sendo a F_2 a população de referência.

Durante o processo de homozigose, a variância devida à dominância vai sendo dissipada com a eliminação dos genótipos heterozigotos. Em função disso, a variância genética entre as linhas puras aleatórias (geração F_n) derivadas desse cruzamento é dada pelo dobro da variância aditiva existente na geração F_2 ($\sigma_{A(F_n)}^2 = 2\sigma_{A(F_2)}^2$). Em consequência, considerando que para um loco com dois alelos a variância aditiva está dada pela eq. (2),

$$\sigma_A^2 = 2pq[a + (q - p)d]^2 \quad (2)$$

em que:

- a: valor genotípico associado aos homozigotos; e
- d: valor genotípico associado aos heterozigotos.

Então, numa população de linhagens puras aleatórias (geração F_n) derivadas de um cruzamento biparental, com frequências alélicas $p=q=0,5$, a variância aditiva será $\sigma_A^2 = 0,5 a^2$; e a variância genética será: $\sigma_G^2 = a^2$ ou $\sigma_G^2 = \sum_1^n a^2$ (SOUZA JUNIOR, 1989). Isso implica que no caso de a variância aditiva apresentar valores mais altos que a epistática e os efeitos aditivos dos n genes que controlam o caráter de interesse forem semelhantes, então, a variância genética entre as linhas puras aleatórias (geração F_n) derivadas desse cruzamento biparental será determinada pelo número de genes segregando na população de linhagens endogâmicas (FALCONER; MACKAY, 1996).

Sabendo que o número de genes segregando determina tanto a variância genética quanto a divergência genética entre genótipos, então, na tentativa de prever o potencial das populações, usa-se a estimativa de divergência genética entre genitores como preditor da variância genética da população a ser formada pelo

cruzamento entre eles, esperando que genitores mais divergentes gerem as populações com maior potencial ao ter maior número de genes segregando.

2.1.5.2 Metodologias utilizadas para realizar predições de populações endogâmicas

Existem diferentes metodologias para fazer as predições, e visto que elas também determinam a maneira de fazer a escolha dos genitores, podem-se classificá-las segundo esse critério, em dois grupos. No primeiro grupo estão os métodos que estimam distâncias genéticas, sendo essas distâncias usadas posteriormente para prever a divergência entre genitores e, conseqüentemente, a variabilidade gerada do cruzamento; eles permitem fazer a escolha dos genitores *a priori* pelo fato da seleção estar baseada nas estimativas de divergência, além do desempenho *per se* dos genótipos candidatos a genitores. Os métodos do segundo grupo utilizam as médias populacionais de gerações precoces para fazer as predições, e apenas permitem fazer a escolha *a posteriori*, porque os pais são selecionados segundo o desempenho de suas progênes, com dados de F_1 , F_2 ou de gerações avançadas (DIAS et al., 2004).

Entre os métodos preditivos que permitem a escolha dos genitores *a priori*, estão o coeficiente de parentesco, a análise multivariada de caracteres quantitativos e os marcadores moleculares. O coeficiente de parentesco tem sido muito utilizado e geralmente é empregando o coeficiente de Malécot (MALÉCOT, 1948). Esse método determina a probabilidade de dois indivíduos apresentarem um alelo idêntico por descendência em um loco tomado ao acaso. Para o seu cálculo, é necessário o conhecimento da genealogia de cada genótipo, bem como as contribuições relativas de cada ancestral. O método tem três pressuposições: *i*) ancestrais originais dos genitores não são aparentados; *ii*) cada genitor de um dado cruzamento contribui com 50% de alelos para a sua progênie; e *iii*) os descendentes foram formados em ausência de seleção (natural ou artificial), mutação, migração ou deriva. (FALCONER; MACKAY, 1996). A capacidade preditiva do coeficiente de parentesco é baixa, pois

esse método apenas estima similaridade por descendência e, no caso da variância genética, ela é determinada pelo grau de similaridade por estado, ou seja, toda a similaridade existente entre dois indivíduos (HELMS et al., 1997). Além disso, pode haver falhas em algumas das pressuposições, sendo a mais comum aquela relativa ao item *iii*, porque geralmente usa-se como genitores cultivares melhorados, os quais foram obtidos a partir de populações segregantes submetidas a processos seletivos pelos melhoristas (BURKHAMER et al., 1998). Gizlice, Carter e Burton (1993a) mostraram que a primeira pressuposição também pode ser violada, ao demonstrar que alguns ancestrais dos cultivares de soja da América do Norte são aparentados, obtendo coeficientes de similaridade genética de até 0,88.

O método da análise multivariada de caracteres quantitativos envolve as análises de componentes principais, variáveis canônicas e métodos aglomerativos. Ele considera as diferenças morfológicas ou fisiológicas apresentadas pelos genitores na determinação da divergência, que é geralmente quantificada como uma medida de dissimilaridade. Seu princípio é que, quando se avalia divergência de uma população baseada em valores fenotípicos médios, também se avalia divergência entre valores genotípicos associados com a frequência alélica em diferentes unidades amostrais (BARROSO; HOFFMANN, 2003). Uma desvantagem do método é que a avaliação de muitos dados pode ser trabalhosa, além do fato que os valores fenotípicos poderem ser fortemente influenciados pelo ambiente, o que pode dificultar a obtenção de estimativas precisas da distância genética entre as linhagens avaliadas (THORMANN; OSBORN, 1992; MOSER; LEE, 1994). Contudo, os resultados de trabalhos de predição apontam que, em média, sejam preditores mais eficientes que o coeficiente de parentesco (BARROSO; HOFFMANN, 2003).

As estimativas de divergência genética determinadas por marcadores moleculares são as mais frequentes na literatura atual e tem como vantagem o fato de oferecer uma medida mais precisa desse parâmetro genético (FALCONER; MACKAY, 1996; THORMANN; OSBORN, 1992). A literatura é vasta em trabalhos sobre estudos de distâncias genéticas, utilizando diferentes tipos de marcadores moleculares (RFLP, RAPD, AFLP e SSR); porém, a utilização destas estimativas nos programas de

melhoramento é muito reduzida devido à inconsistência das predições (HELMS et al., 1997; BARROSO; HOFFMANN, 2003).

Os métodos preditivos que permitem a escolha dos genitores *a posteriori*, são aqueles que envolvem a avaliação em gerações precoces; entre eles estão a análise dialélica, o método de Jinks e Pooni (1976) para estimar proporção de linhagens superiores e as estimativas de resposta à seleção. Esses métodos são mais consistentes nas predições, porém, requerem maior trabalho pelo fato de ser necessário formar e avaliar as populações (BARROSO; HOFFMANN, 2003). Na análise dialélica, o sistema de cruzamentos dialélico corresponde ao intercruzamento de n genitores, gerando n^2 combinações possíveis que correspondem aos n genitores, sendo $n(n-1)/2$ o número de híbridos simples possíveis e $n(n-1)/2$ o número de recíprocos dos híbridos simples. Para estimar os componentes de variâncias genéticas com informações geradas de cruzamentos dialélicos, o efeito de genitores deve ser aleatório, por isso, os genitores devem ser uma amostra representativa de uma população de referência (BAKER, 1978; RAMALHO; SANTOS; ZIMMERMANN, 1993); uma vez cumprida esta restrição, os componentes da variância genética podem ser estimados por meio de duas metodologias, a de Jinks e Hayman (1953) ou a de Griffing (1956). No caso específico da metodologia de Griffing (1956), ela estima a capacidade geral de combinação (CGC) e a capacidade específica de combinação (CEC). O termo CGC é utilizado para designar o comportamento médio de um genitor numa série de combinações híbridas, e está associado aos efeitos aditivos dos alelos e às ações epistáticas do tipo aditivas; ao passo que a CEC expressa a superioridade ou inferioridade de certas combinações híbridas, em relação ao que seria esperado segundo as CGC de seus genitores; ela está associada aos efeitos de desvio de dominância dos genes e à epistasia envolvendo dominância (CRUZ; VENCOSKY, 1989; RAMALHO; SANTOS; ZIMMERMANN, 1993). Para que a CGC possa ser preditora do comportamento de populações endogâmicas, é necessário que não exista interação entre a CGC e as gerações, ou seja, que as magnitudes dos efeitos epistáticos do tipo aditivo x aditivo e de ordens superiores, não sejam suficientes para alterar a classificação das populações realizadas pela CGC (STHUTHMAN; STUCKER, 1975; RAMALHO; SANTOS; PEREIRA FILHO, 1988).

O método de Jinks e Pooni (1976), por sua vez, possibilita a predição da proporção de linhagens superiores derivadas de um cruzamento biparental. A consistência de sua predição depende de três pressuposições: i) a média das linhagens da população deve ser igual à média de seus genitores; ii) os efeitos epistáticos devem ser pequenos ou ausentes a fim de que não interfiram com a simetria da distribuição das linhagens endogâmicas em relação à média genitor; e iii) a estimativa da σ_A^2 deve ser precisa. O cálculo da predição da proporção de linhagens superiores se obtém por meio da distribuição normal padrão, através do cálculo dos valores “z” pela fórmula de Kearsey e Pooni (1996), eq. (3),

$$z = \frac{PD - \bar{P}}{\sqrt{2\hat{\sigma}_A^2}} \quad (3)$$

em que:

- z: é o valor da variável z (abscissa) de uma distribuição normal padrão;
- PD: é o padrão adotado como limite para obtenção de linhagens superiores, que é determinado pelo melhorista, podendo ser a média do cruzamento, média do genitor superior ou média de cultivares comerciais;
- $\bar{P}$: é a média dos genitores do cruzamento, que corresponde à média da população após atingir a homozigose; e
- $\hat{\sigma}_A^2$: é a variância genética aditiva da população na geração F_2 .

Deve-se salientar que a baixa produção de sementes híbridas da maioria das espécies autógamas limita o uso dos métodos preditivos que permitem a escolha dos genitores *a posteriori*. Jinks e Pooni (1980); Toledo (1987) e Kearsey e Pooni (1996) sugerem a estimação da variância aditiva com a utilização de progênies $F_{2:3}$, isto é, progênies F_3 derivadas de plantas F_2 . Segundo Cockerham (1954), a variância genética entre médias de progênies F_g derivadas de plantas F_2 provenientes de cruzamentos biparentais corresponde à eq. (4),

$$\sigma_{\bar{F}_g:F_2}^2 = \sigma_A^2 + (1 - I_g)^2 \sigma_D^2 \quad (4)$$

em que:

- σ_A^2 : é a variância genética aditiva;
- σ_D^2 : é a variância genética de dominância; e
- I_g : é o coeficiente de endogamia na geração da descendência, sendo a F_2 a população de referência.

Assim, na geração $F_{2:3}$, onde o coeficiente de endogamia $I_{F_3} = 1/2$, a variância genética entre progênies é eq. (5),

$$\sigma_{p(F_{2:3})}^2 = \sigma_A^2 + \frac{1}{4} \sigma_D^2 \quad (5)$$

Como a σ_D^2 está dividida por quatro, sugere-se ignorar esse componente para obter uma estimativa aproximada da variância genética aditiva, isto é $\sigma_{p(F_{2:3})}^2 \cong \sigma_A^2$. Desta maneira é possível o uso desta metodologia nos programas de melhoramento genético nas espécies autógamas, podendo avaliar o potencial de cada cruzamento na primeira geração de abertura de progênies e, conseqüentemente, eliminar aqueles com pouco potencial, concentrando-se os esforços e recursos nos cruzamentos promissores. A consistência das predições de porcentagem de progênies superiores (JINKS; POONI, 1976), utilizando a estimativa de variância aditiva derivada de avaliações de progênies $F_{2:3}$ é alta, e foi testada em culturas como fumo (JINKS; POONI, 1980); soja (TOLEDO, 1987; TRILLER; TOLEDO, 1996; PULCINELLI, 1997) e aveia (POWELL et al., 1985; CALIGARI; POWELL; JINKS, 1985).

Finalmente, Hallauer e Miranda Filho (1988) enfatizam que a estimativa da resposta á seleção constitui uma das mais importantes contribuições da genética quantitativa para o melhoramento de plantas. Por meio dela, pode-se avaliar se uma dada população é adequada aos propósitos de melhoramento, calculando a possibilidade de se conseguir alterar, positivamente, a média da população, mediante a utilização de técnicas de seleção apropriadas. A resposta à seleção ou ganho

genético pode ser calculada pelo produto do diferencial de seleção (ds) pela herdabilidade, sendo que a principal função da herdabilidade, segundo Falconer (1987), é seu papel preditivo, que expressa a confiabilidade do valor fenotípico como estimador do valor genotípico, de tal forma que quanto maior a herdabilidade maior o ganho genético por seleção. É provável que nos casos onde a divergência genética seja preditora da variância genética, também seja preditora da resposta à seleção, devido ao fato que a seleção está associada à existência de variabilidade genética, à seleção natural e/ou artificial e ao ajuste dos genótipos aos ambientes existentes (REIS et al., 2004).

2.1.5.3 Resultados das predições do comportamento de populações endogâmicas

Existem inúmeros trabalhos na área de predição do comportamento de populações endogâmicas, utilizando diferentes métodos preditivos. Os resultados mostram que as predições mais consistentes são as que tendem a predizer a média das populações endogâmicas e as menos satisfatórias são as relativas à predição da variância genética (BARROSO; HOFFMANN, 2003). As causas disso podem estar relacionadas ao fato de que as médias das gerações utilizadas podem ser obtidas com maior precisão que as variâncias, já que apresentam um pequeno erro associado à sua estimativa, em comparação ao erro associado às estimativas das variâncias (BURKHAMER et al., 1998). No entanto, existem outras possíveis causas da inconsistência na predição das variâncias, uma delas está relacionada às médias dos genitores, visto que se observa que à medida que as médias dos genitores vão se tornando diferentes, a divergência genética perde seu poder de predição, fazendo com que o comportamento das populações endogâmicas seja independente da variância genética (THORNE; FEHR, 1970; SCHOENER; FEHR, 1979; VELLO; FEHR; BAHRENFUS, 1984; ININDA et al., 1996). Outra possível causa da inconsistência está relacionada à correlação entre a divergência genética dos genitores e a variância genética do caráter de interesse na população sob estudo,

pois a capacidade de predizer o potencial das populações usando a estimativa de divergência genética vai depender dessa correlação (GIZLICE; CARTER; BURTON, 1993b). Sabe-se que a correlação pode ser devida aos fatores genéticos tais como a pleiotropia e a ligação gênica ou devida aos fatores ambientais (FALCONER, 1987).

Os trabalhos publicados com aveia e trigo mostram que as correlações entre a distância genética estimada por coeficiente de parentesco e a variância genética são nulas ou de baixa magnitude (COWEN; FREY, 1987; SOUZA; SORRELS, 1991; MOSER; LEE, 1994; BARBOSA-NETO; SORRELLS; CISAR, 1996; FABRIZIUS et al., 1998; BURKHAMER et al., 1998). Na cultura da soja, embora correlações altas tenham sido verificadas por Manjarrez-Sandoval et al. (1997b), os resultados de outros trabalhos mostram que esse método não apresenta a precisão necessária para ser um bom preditor (MANJARREZ-SANDOVAL et al., 1997a; KISHA; SNELLER; DIERS, 1997; HELMS et al., 1997).

As correlações entre distância genética estimada por métodos multivariados e a variância genética variam entre espécies e entre conjuntos de populações da mesma espécie. Entretanto, os resultados apontam que, em média, sejam preditores mais eficientes que o coeficiente de parentesco (BARROSO; HOFMANN, 2003).

As estimativas de divergência genética obtidas por marcadores moleculares não são capazes de predizer de modo consistente a variância genética das populações endogâmicas. Em trigo, Burkhamer et al. (1998) verificaram que as distâncias genéticas estimadas a partir dos marcadores AFLP e STS-PCR (*Sequence Tagged Sites*) não foram bons preditores da variância genética e do número de segregantes transgressivos em populações $F_{3:5}$. Resultados similares foram obtidos por Helms et al. (1997) em soja e Gumber et al. (1999) em *Vicia faba*, com marcadores RAPD; assim como Moser e Lee (1994) em aveia e Kisha; Sneller e Diers (1997) em soja, utilizando marcadores RFLP. Em aveia, Moser e Lee (1994) utilizaram simultaneamente distâncias genéticas baseadas em marcadores moleculares, coeficiente de parentesco e desempenho dos genitores em caracteres quantitativos, e encontraram que a variância genética para caracteres associados ao rendimento se mostrou independente ou muito pouco correlacionada com a distância genética. No entanto, alguns trabalhos realizados com marcadores moleculares, obtiveram

associação entre ambos os parâmetros genéticos. Nesse sentido, Manjarrez-Sandoval et al. (1997b) observaram uma relação quadrática entre a distância genética baseada em marcadores RFLP e a variância genética da produção de grãos em populações endogâmicas de soja; porém, eles acrescentam que essa associação deve ser considerada com cautela, pois foi obtida pela análise de apenas cinco populações. De igual maneira, Colombari Filho (2009) obteve alta correlação entre as distâncias genéticas baseada em marcadores AFLP e variância genética em seis populações de soja para os caracteres produção de grãos, dias para maturação e altura no florescimento; embora altas, as associações foram não significativas, fato que o autor sugere que possa ser devido ao pequeno número de graus de liberdade associados ao teste t.

Manjarrez-Sandoval et al. (1997b) demonstraram que é possível obter estimativas de DG com capacidades preditivas em relação à Rs. Nesse trabalho, foram utilizadas estimativas de DG por marcadores RFLP e linhagens $F_{6:8}$ e $F_{8:10}$ provenientes de cinco populações pertencentes ao programa de melhoramento de soja da USDA-ARS. Os autores validaram os resultados obtidos da resposta à seleção (Rs) fazendo seguimento no tempo das linhagens selecionadas de cada uma das cinco populações e constataram que as predições de Rs concordaram com as seleções realizadas pelo programa de melhoramento, sendo que duas das três populações com maiores Rs geraram as linhagens presentes nos testes finais do programa de melhoramento. Os autores também verificaram que a média predita da linhagem mais produtiva, estimada pela média observada da população acrescentada pelo ganho genético esperado, foi mais precisa que a Rs na identificação do potencial das populações, devido ao fato que a população que obteve a segunda melhor média predita foi a que gerou um novo cultivar de soja (cultivar Brim).

Entendendo que a correlação entre ambos os parâmetros genéticos é imprescindível para que exista a capacidade preditiva da divergência genética, então, é importante a identificação das causas das falhas na correlação. Segundo Gumber et al. (1999), essas falhas podem ser atribuídas: i) à baixa repetibilidade da variância genética estimada, devido ao elevado erro padrão; ii) às estimativas de diversidade genética viesadas; ou iii) ao fato da correlação entre a variância genética real e

diversidade genética real ser baixa. No primeiro caso, têm que ser considerados vários fatores, entre os quais estão: o método usado para calcular a estimativa, o número de populações desenvolvidas (ao menos 10 populações diferentes), a obtenção das progênies livres do efeito da deriva génica, o tamanho da amostra avaliada (superior a 100 linhagens por cruzamento em 3 ou 4 repetições) e o número de ambientes usados nas avaliações (GUMBER et al., 1999; GIZLICE; CARTER; BURTON, 1993b).

No segundo caso, a estimativa de diversidade genética pode estar viesada com respeito à verdadeira diversidade genética dos genitores, devido a erros no processo de obtenção da estimativa ou devido ao método usado para sua determinação (BARROSO; HOFFMANN, 2003). Em relação aos métodos, observa-se que de maneira geral, os valores da estimativa de divergência genética variam segundo o método utilizado. Nesse sentido, ao comparar divergência genética calculada por coeficiente de parentesco e RAPD, Helms (1997) não encontrou correlação entre elas, entanto que Abdelnoor, Barros e Moreira (1995) observaram que ambas as técnicas agrupavam os materiais de maneira distinta. De maneira similar, baixo valores de correlação foram encontrados entre as divergências genéticas estimadas por coeficiente de parentesco e AFLP (BONATO et al., 2006); assim como entre as divergências genéticas estimadas por coeficiente de parentesco e SSR (PRIOLLI et al., 2010). Entretanto, correlações de até 0,91 têm sido obtidas em estudos em que foi calculada a divergência genética por coeficiente de parentesco e RFLP, (MAJARREZ-SANDOVAL et al., 1997b; KISHA et al., 1998). Aparentemente, as diferenças nas estimativas da divergência genética determinadas por coeficiente de parentesco, análise multivariada e por marcadores moleculares são mais acentuadas à medida que os genitores estão mais relacionados (MIRANDA et al., 2001). As causas da variação das estimativas de divergência genética geradas pelos diferentes métodos podem estar presentes em função das pressuposições assumidas, que causam viés na estimativa, tanto superestimando a divergência genética, ao ignorar possíveis relações entre ancestrais, quanto subestimando ao assumir, por exemplo, que duas bandas com a mesma mobilidade em um gel são idênticas quando elas podem ser diferentes (MOSER; LEE, 1994). Priolli et al. (2010) estabelecem que a inconsistência

observada nos diferentes estudos é esperada, já que os marcadores moleculares identificam alelos idênticos por estado, o que não é considerado na estimação da divergência genética calculada por coeficiente de parentesco, já que esse identifica alelos idênticos por descendência.

A terceira causa nas falhas da correlação citadas por Gumber et al. (1999), faz referência aos casos em que realmente a correlação entre a variância genética e diversidade genética é baixa; geralmente isso ocorre em cruzamentos entre genitores adaptados e não adaptados. Em trabalhos realizados em aveia (MOSER; LEE, 1994), feijão (GUMBER, 1999) e soja (BARROSO et al., 2003), encontrou-se que as DG foram geralmente maiores em cruzamentos entre genitores não adaptados do que entre genitores adaptados. No entanto, as variâncias genéticas foram geralmente maiores em cruzamentos entre genitores adaptados do que entre genitores não adaptados; nesses casos, a adaptação constitui outro fator que influencia a variância genética liberada em um cruzamento, por afetar a expressão gênica dos QTLs que determinam a expressão de um caráter. Nos casos em que os genitores adaptados apresentam uma estreita relação genética é provável que as estimativas de DG possam ser preditoras das médias das populações endogâmicas (BARROSO et al., 2003).

Outras causas das falhas de correlação entre a DG e as variâncias genéticas foram discutidas por Gumber, et al. (1999), que consideraram que as estimativas de divergência genética obtidas por marcadores moleculares podem representar apenas uma diversidade genética geral dos genitores e não uma diversidade genética específica para o caráter de interesse. Neste sentido, observa-se que as magnitudes das correlações entre as distâncias genéticas e as variâncias genéticas dependem da distribuição dos QTLs pelo genoma, da magnitude dos efeitos dos QTLs sobre o caráter, do grupo de populações e das condições ambientais em que as variâncias genéticas são estimadas (HELMS et al., 1997; KISHA; SNELLER; DIERS, 1997; MANJARREZ-SANDOVAL et al., 1997b; BARROSO; HOFFMANN, 2003); no entanto, a maioria dos trabalhos sobre divergência genética são realizados desconhecendo a posição dos QTLs que controlam o caráter de interesse.

O uso dos diferentes marcadores se faz assumindo dois aspectos: ampla cobertura do genoma pelos marcadores e distribuição homogênea pelo genoma dos QTLs que controlam o caráter alvo, esperando-se que a divergência genética estimada pelos marcadores reflita a divergência genética para o caráter de interesse. No entanto, existem diferentes fatores que impedem a correlação entre ambas as divergências genéticas, tais como: *i*) distribuição dos marcadores pelo genoma, a qual depende da técnica usada e, ainda, dentro da região acessada pela técnica, pois pode ser que a distribuição dos marcadores não seja uniforme, devido a diferentes causas como, por exemplo: número insuficiente de marcadores utilizados, assim como tipo de sequência preferencialmente detectada pelo marcador. Isso implica que parte dos genes amostrados pelos marcadores podem não estar ligada aos QTLs que influenciam o caráter de interesse (BERNARDO, 1992; BARROSO; HOFFMANN, 2003); *ii*) diferentes QTLs que afetam o caráter de interesse podem apresentar efeitos diferentes ou estar ligados. Em ambos os casos, as regiões do genoma vão possuir diferentes efeitos sobre o caráter e, se essas regiões não forem densamente amostradas pelos marcadores, não será possível medir a verdadeira divergência genética; e *iii*) pode existir mobilidade de segmentos de DNA dentro dos cromossomos. Nesse sentido, sabe-se que na cultura de milho QTLs importantes para rendimento podem estar localizados em certos segmentos do cromossomo e que esses segmentos podem estar em diferentes partes do cromossomo em populações derivadas de diferentes cruzamentos (STUBER; EDWARDS; WENDEL, 1987; STUBER, 1989); isso faz com que os cálculos de divergência genética com marcadores moleculares distribuídos uniformemente pelo genoma não possam ser úteis para determinar a verdadeira divergência genética para o caráter de interesse.

O nível de relacionamento genético dos genitores pode aumentar a possibilidade de estar se estimando uma diversidade genética geral, visto que quando os genitores estão relacionados geneticamente, os alelos do marcador são idênticos por descendência e devem ser informativos sobre o loco e regiões ligadas; no caso em que os genitores não estão relacionados, os alelos dos marcadores só podem ser iguais em estado e dão pouca informação sobre áreas adjacentes (MOSER; LEE, 1994; SARAWAT; STODDARD; MARSHALL, 1994b; GUMBER et al., 1999).

Melchinger et al. (1990a; 1990b; 1992) e Charcosset, Lefort-Buson e Gallais (1991) indicaram que para garantir a obtenção de estimativas de divergência genética com capacidade preditiva, deve-se abranger o conhecimento da posição dos QTLs que controlam o caráter de interesse, ter desequilíbrio de ligação entre os marcadores moleculares e esses QTLs, e que a fase de ligação entre os marcadores moleculares e os QTLs se mantenham invariáveis nos diferentes grupos de germoplasma. Além disso, Melchinger et al. (1990a) propôs uma alternativa com vistas a melhorar a associação entre divergência genética e outros parâmetros genéticos, que se baseia na decomposição da distância genética de acordo com o método de Griffing (1956). Segundo esta proposta, as distâncias genéticas para um grupo de genótipos podem ser arranjadas na forma de um dialelo (distâncias obtidas com todas as combinações duas a duas a partir de um grupo de genótipos) e decompostas de acordo o método de Griffing (1956) para cruzamentos dialélicos, isto é: eq. (6),

$$DG = \overline{DG} + DGG_i + DGG_j + DGS_{ij} \quad (6)$$

em que:

- DG é a distância genética entre dois genitores;
- $\overline{DG}$ é a distância genética média do grupo;
- DGG_i é a distância genética geral associada ao genitor “i”;
- DGG_j é a distância genética geral associada ao genitor “j”; e
- DGS_{ij} é a distância genética específica associada à combinação “ij”.

Poucos trabalhos têm utilizado essa alternativa, e os resultados mostram que a decomposição da distância genética apenas melhorou em pequenas magnitudes a associação entre divergência genética e variância genética. Segundo os autores, os fatores que podem ter gerado essa situação são: estimativas de divergência genética com viés e o tipo de germoplasma utilizado (MELCHINGER et al., 1990a, 1990b, 1992; BARROSO et al., 2003).

2.2 Material e métodos

2.2.1 Material genético

O material genético utilizado no presente trabalho compreende 10 populações $F_{2:3}$ formadas a partir dos cruzamentos dialélicos de cinco linhagens genitoras, excetuando-se os recíprocos. Os genitores foram escolhidos com base nas distâncias genéticas calculadas por Bonato et al. (2006), usando marcadores moleculares do tipo AFLP, seguido da adaptabilidade dos genitores ao Estado de São Paulo, ciclos semelhantes e alta produção. No estudo realizado por Bonato et al. (2006) foram estabelecidos os coeficientes de similaridades genéticas (SG) entre 317 cultivares de soja liberados no Brasil de 1962 até 1998, usando seis combinações de *primers* *EcoRI/MseI* (E-AAC/M-CAT, E-AAC/MCTA, E-AAC/M-CTC, E-AAC/M-CTG, E-AAG/M-CTT, e E-ACT/M-CAT) e o coeficiente de Nei e Li (1979).

As estimativas de distâncias genéticas (DG) dos cinco cultivares compreenderam um amplo intervalo (Tabela 1). Assim, tem-se a existência de diferentes graus de DG, isto é, cultivares com alta DG (em torno de 0,6), cultivares com DG intermediária (em torno de 0,4) e cultivares com baixa DG (menor que 0,2). As sementes dos genitores utilizados foram obtidas do Banco Ativo de Germoplasma da Embrapa Soja (Londrina-PR).

Os cultivares utilizados são adaptados e recomendados para o plantio comercial no Estado de São Paulo (Tabela 2), tendo ciclos que vão desde semiprecoce até médios. A genealogia dos cultivares mostra que alguns deles têm alto grau de relacionamento

Tabela 1 - Cultivares utilizados e distâncias genéticas entre eles, obtidas através do marcador AFLP

	Cultivares	Embrapa-60	IAC-12	IAC-100	FT-10	FT-14
1	Embrapa-60	-				
2	IAC-12	0,367	-			
3	IAC-100	0,464	0,128	-		
4	FT-10	0,586	0,385	0,451	-	
5	FT-14	0,509	0,334	0,407	0,161	-

Fonte: Bonato et al.(2006).

Tabela 2 - Informação genealógica dos cultivares de soja

	Cultivar	GM	Instituição responsável	Genealogia	Origem dos genitores
1	Embrapa-60	M	Embrapa	FT-Abyara x BR83-147	União x Sant' Ana União x BR-1
2	IAC-12	S	IAC	Paraná x IAC 72-231	Hill x D52-810
3	IAC-100	S	IAC	IAC-12 x IAC 78-2318	Paraná x IAC 72-231
4	FT-10 (Princesa)	S	FT- Pesquisa e Sementes	FT9510 x Sant' Ana	Seleção em Andrews D51-5427 x D49-2491
5	FT-14 (Piracema)	S	FT- Pesquisa e Sementes	FT9510 x Sant' Ana	Seleção em Andrews D51-5427 x D49-2491

GM: Grupo de maturação; M: médio (131-140 dias); S: semiprecoce (121 a 130 dias).

Fonte: Embrapa 2004b; IICA, 1987; Almeida et al., 1999b; Lourenção et al., 2005.

genético, tal como FT-10 e FT-14, as quais são derivadas do mesmo cruzamento, assim como IAC-12 e IAC-100, em que uma é genitora da outra.

Ao comparar esta informação com as DG resultantes do trabalho de Bonato et al. (2006), pode-se observar coincidência, pois as menores DG correspondem aos cultivares com maior grau de relacionamento genético tal como IAC-12 e IAC-100, seguidos de FT-10 e FT-14. Entre os genitores mais divergentes, destacam-se Embrapa-60 com os cultivares FT-10 e FT-14. Entretanto, é interessante observar que Embrapa-60 apresenta maior semelhança genética com IAC-12 e IAC-100 que com FT-10 e FT-14, ainda que com estes compartilhem um ancestral comum (Sant' Ana).

2.2.2 Obtenção do material experimental

2.2.2.1 Obtenção das progêneses F_1 e F_2

No ano de 2008 foram realizados todos os cruzamentos entre os cultivares em casa de vegetação. Os genitores foram cruzados em um dialelo completo, sem recíprocos, originando 10 cruzamentos F_1 , os quais foram autofecundados para obtenção da geração F_2 no inverno de 2008. Nos cruzamentos em que os genitores diferiam na cor das flores, utilizou-se esse marcador morfológico (cor da flor) para distinguir possíveis autofecundações.

2.2.2.2 Obtenção das progêneses $F_{2:3}$

No ciclo de verão (2008/9), foram semeadas em bandejas de isopor 110 sementes descendentes de uma planta de cada cruzamento, em que os pais tiveram flores de cores diferentes, e 110 sementes descendentes de cada uma de duas plantas, naqueles cruzamentos em que os pais tiveram flores da mesma cor. As plântulas foram transplantadas aos 15 dias após a semeadura nas bandejas, na Estação Experimental Anhembi do Departamento de Genética da ESALQ/USP, localizada no distrito de Anhumas, para o avanço de uma geração de autofecundação. Durante o florescimento, foi verificada a presença de segregação em todas as descendências plantadas, em seguida, foi selecionada apenas uma descendência por cruzamento. Na maturação, as plantas foram colhidas individualmente dentro de cada cruzamento e foram formadas as populações $F_{2:3}$, selecionando ao acaso 100 plantas individuais de cada cruzamento.

2.2.3 Avaliação experimental das progêneses $F_{2:3}$

As 1.000 progêneses $F_{2:3}$ provenientes das 10 populações foram semeadas no dia 04/11/2009 na E. E. Anhembi, em 10 experimentos com delineamento experimental em látice simples 10 x 10, com duas repetições. Cada experimento foi constituído de 100 progêneses $F_{2:3}$ provenientes de um cruzamento particular. As parcelas foram

constituídas por uma linha de 2,0 m de comprimento com espaçamento de 0,50 m entre linhas, contendo 30 plantas após o desbaste, que foi feito 26 dias depois do plantio. No início, meio e fim de cada repetição foram semeados os genitores correspondentes a cada população, em parcelas iguais às progênies. Os tratos culturais e cuidados necessários foram aplicados, visando à obtenção de um experimento com boa precisão experimental, e, para isso, o local conta com infraestrutura necessária, incluindo irrigação.

2.2.4 Caracterização do ambiente experimental

Os cruzamentos e avanços de gerações de autofecundação até a F_2 foram realizados em casa de vegetação, nas dependências do Departamento de Genética da Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo (ESALQ/USP), localizada no município Piracicaba, SP. A obtenção das progênies $F_{2:3}$ e avaliação de ensaios foram realizados na Estação Experimental Anhembi, localizada a uma distância de 42 km da sede da ESALQ, próxima ao distrito de Anhumas, com relevo plano, solo do tipo latossolo amarelo aluvial distrófico, textura médio-arenosa, com acidez e toxicidade de alumínio neutralizado pela aplicação de calcário dolomítico em anos anteriores.

2.2.5 Caracteres avaliados

Foram avaliados os seguintes caracteres

- a) **Estande (ST):** número de plantas da parcela, avaliadas logo após do desbaste.
- b) **Altura das plantas na maturação (AM):** distância medida em centímetros, do nível do solo ao ápice da haste principal, na época de maturação (estádio R_8) avaliada em cinco plantas por parcela.

- c) **Produção de grãos (PG):** produção total de grãos da parcela, em g.m^{-2} , e com umidade corrigida para 13%.

Para o caráter PG, as observações das progênes $F_{2:3}$ foram corrigidas para o estande. Além disso, foram eliminadas observações com estandes menores que 11 plantas por parcela, para garantir mais de cinco plantas por metro linear; assim, oito observações pertencentes a quatro populações foram eliminadas para as análises. Estandes mínimos por parcela semelhantes aos utilizados neste trabalho foram considerados como adequados por Manjarrez-Sandoval et al. (1997b). As observações foram ajustadas para umidade de 13% e, posteriormente, as análises dos dados foram feitas utilizando apenas o estande como covariável.

2.2.6 Análises estatístico-genéticas

2.2.6.1 Análises de variância

As análises estatísticas foram realizadas com o PROC GLM do SAS (SAS INSTITUTE, 2001), separadamente para cada experimento. Antes de proceder com as análises, foi testada a normalidade dos resíduos pelo teste de Shapiro-Wilk (SAS INSTITUTE, 2001). As análises de variâncias foram realizadas com base no modelo matemático apresentado a seguir, em que todos os efeitos, exceto a média (μ), são aleatórios eq. (7),

$$Y_{ijk} = \mu + t_i + r_j + b_{k(j)} + \varepsilon_{ijk} \quad (7)$$

em que:

- Y_{ijk} : é o valor observado da progênie i na repetição j , no bloco k ;
- μ : é a média geral (constante inerente a todas as observações);
- t_i : é o efeito aleatório da progênie i ($i = 1, 2, \dots, I$);
- r_j : é o efeito aleatório da repetição j ($j = 1, 2, \dots, J$);

- $b_{k(j)}$: é o efeito aleatório do bloco k ($k = 1, 2, \dots, K$); dentro de repetição j ; e
- ε_{ijk} : é o erro experimental associado à observação Y_{ijk} .

Com base nesse modelo, tem-se o seguinte esquema da análise de variância com as respectivas esperanças dos quadrados médios, para cada experimento (Tabela 3).

Tabela 3 - Esquema da análise de variância com as respectivas esperanças dos quadrados médios

FV	GL	QM	E(QM)	F
Repetições (R)	(J-1)	QM_R		
Blocos / R	J (K-1)	QM_B		
Progênies Ajustadas (P)	(K ² -1)	QM_P	$\sigma^2 + J\sigma_p^2$	QM_P / QM_E
Erro Intra-Blocos	(K-1)(JK-K-1)	QM_E	σ^2	
Total	(JK ² -1)			

Para a obtenção das esperanças matemáticas dos quadrados médios foram adotadas as seguintes pressuposições:

- I) $t_i \sim \text{NID}(0, \sigma_p^2)$;
- II) $r_j \sim \text{NID}(0, \sigma_r^2)$;
- III) $b_{k(j)} \sim \text{NID}(0, \sigma_b^2)$;
- IV) $\varepsilon_{ijk} \sim \text{NID}(0, \sigma^2)$; e
- V) t_i , r_j , $b_{k(j)}$ e ε_{ijk} são independentes.

2.2.6.2 Estimativas dos componentes de variância

Com base nas esperanças dos quadrados médios foram estimados os componentes de variância, utilizando as seguintes expressões:

Variância do resíduo entre parcelas, eq. (8),

$$\hat{\sigma}^2 = QM_E \quad (8)$$

Variância genética entre progênes F_{2:3}, eq. (9),

$$\hat{\sigma}_p^2 = \frac{QM_P - QM_E}{J} \quad (9)$$

Variância fenotípica entre médias de progênes F_{2:3}, eq. (10),

$$\hat{\sigma}_F^2 = \frac{QM_P}{J} \quad (10)$$

em que:

- QM_P é o quadrado médio de progênes;
- QM_E é o quadrado médio do erro experimental; e
- J é o número de repetições.

Segundo Cockerham (1963), a variância genética entre médias de progênes F_{2:3} ($\hat{\sigma}_p^2$), ignorando os efeitos epistáticos, é composta pela eq. (11),

$$\sigma_{G(F_{2:3})}^2 = \sigma_A^2 + \frac{1}{4} \sigma_D^2 \quad (11)$$

Como o componente de variância atribuído aos efeitos de dominância está dividido por quatro, é possível a utilização da estimativa de variância genética entre médias de progênes F_{2:3} como uma estimativa de variância genética aditiva, isto é, $\hat{\sigma}_{p(F_{2:3})}^2 \cong \hat{\sigma}_A^2$, sem afetar significativamente as predições efetuadas (JINKS; POONI, 1980; POWELL et al., 1985; CALIGARI; POWELL; JINKS, 1985; TOLEDO, 1987; TRILLER; TOLEDO, 1996; PULCINELLI, 1997). Por isso, neste trabalho, as estimativas das variâncias genéticas aditivas para todos os cruzamentos foram obtidas dessa maneira.

2.2.6.3 Estimativas dos coeficientes de herdabilidade

Com base nas estimativas dos componentes de variâncias, calculou-se o coeficiente de herdabilidade no sentido amplo, entre médias de progênie F_{2:3} ($\hat{h}_x^2$), através da expressão, eq. (12),

$$\hat{h}_x^2 = \frac{\hat{\sigma}_p^2}{\sigma_F^2} \cdot 100 \quad (12)$$

2.2.6.4 Predição da resposta à seleção

Com base nas médias das 100 progênie F_{2:3} de cada cruzamento, foi aplicado um processo de seleção das 20 e 40 progênie F_{2:3} superiores para produções de grãos, ou seja, uma intensidade de seleção de 20 e 40%, respectivamente. Com as médias das progênie selecionadas e as estimativas de herdabilidade no sentido amplo, entre médias de progênie F_{2:3} ($\hat{h}_x^2$), foi possível estimar a resposta à seleção (Rs) para produção de grãos. Segundo Souza Junior (1989) a Rs utilizando progênie F_{2:3} está dada pela eq. (13),

$$Rs = ds \frac{Cov_{2,3,\infty}}{\hat{\sigma}_F^2} \quad (13)$$

em que:

- ds é o diferencial de seleção com intensidade de seleção de 20 e 40%;
- $Cov_{2,3,\infty}$ é covariância entre médias de progênie F₃ e progênie F_n, tendo todas elas um ancestral comum em F₂; e
- $\hat{\sigma}_F^2$ é variância fenotípica entre médias de progênie F_{2:3}.

Por se tratar de covariâncias entre parentes em gerações diferentes, a $Cov_{2,3,\infty}$ é de naturaliza genética, então, segundo Cockerham (1954) ela é dada pela eq. (14),

$$\text{Cov}_{2,3,\infty} = \sigma_A^2 + (1 - I_g)(1 - I_{g'})\sigma_D^2 \quad (14)$$

em que, nesse caso:

- I_g : é o coeficiente de endogamia na geração F_3 ; e
- $I_{g'}$: é o coeficiente de endogamia na geração das linhagens puras F_n

Assim, pelo fato de que $I_{F3} = 1/2$ e $I_{Fn}=1$, a $\text{Cov}_{2,3,\infty}$ será dada por toda a variância aditiva contida na geração F_2 , o seja, $\text{Cov}_{2,3,\infty} = \sigma_A^2$. Devido que a $\sigma_A^2/\sigma_F^2 = h^2$, então a R_s foi calculada pela eq. (15)

$$R_s = ds \cdot h^2 \quad (15)$$

em que:

- ds é o diferencial de seleção; e
- h^2 é o coeficiente de herdabilidade, entre médias de progênie ($\hat{h}_x^2$).

Foram calculadas os intervalos de confiança (IC) para as estimativas da variância genética entre médias de progênie $F_{2:3}$ ($\hat{\sigma}_p^2$), herdabilidades no sentido amplo, entre médias de progênie $F_{2:3}$ ($\hat{h}_x^2$), de acordo com o processo descrito por Burdick e Graybill (1992), isto é:

Intervalos de confiança (IC) para as estimativas da variância genética, eq. (16),

$$\text{IC}(\hat{\sigma}_p^2)_{0,95} = \left[\frac{n_t \hat{\sigma}_p^2}{\chi_{n_t;0,975}^2} \leq \sigma_p^2 \leq \frac{n_t \hat{\sigma}_p^2}{\chi_{n_t;0,025}^2} \right] \quad (16)$$

sendo:

- n_t é o número de graus de liberdade associado à estimativa do componente de variância genética entre progênie $F_{2:3}$ ($\hat{\sigma}_p^2$);

- $\chi^2_{n_t;0,975}$ é o valor tabelado de χ^2 com n_t graus de liberdade e $\alpha = 97,5\%$ de probabilidade; e
- $\chi^2_{n_t;0,025}$ é o valor tabelado de χ^2 com n_t graus de liberdade e $\alpha = 2,5\%$ de probabilidade.

O número de graus de liberdade (n_t) foi calculado segundo o método descrito por Satterthwaite (1946) eq. (17),

$$n_t = \frac{(QM_P - QM_E)^2}{\frac{(QM_P)^2}{f_1} + \frac{(QM_E)^2}{f_2}} \quad (17)$$

em que:

- QM_P é o quadrado médio de progênies;
- QM_E é o quadrado médio do erro experimental;
- f_1 é o número de graus de liberdade de progênies; e
- f_2 é o número de graus de liberdade do erro experimental.

Intervalos de confiança (IC) para as estimativas das herdabilidades, eq. (18),

$$IC(\hat{h}_x^2)_{0,95} = \{1 - [(\frac{QM_P}{QM_E})F_{0,975:f_1,f_2}]^{-1} \leq h_x^2 \leq 1 - [(\frac{QM_P}{QM_E})F_{0,025:f_1,f_2}]^{-1}\} \quad (18)$$

em que:

- $F_{0,975:f_1,f_2}$ é o valor da Tabela de F com $\alpha = 97,5\%$ de probabilidade associado a f_1 graus de liberdade de progênies e f_2 graus de liberdade do erro experimental; e
- $F_{0,025:f_1,f_2}$ é o valor da Tabela de F com $\alpha = 2,5\%$ de probabilidade associado a f_1 graus de liberdade de progênies e f_2 graus de liberdade do erro experimental.

2.2.7 Desdobramento das distâncias genéticas em DGG e DGE

As distâncias genéticas (DG) utilizadas neste trabalho foram calculadas a partir dos dados de similaridade genéticas (SG_{ij}) obtidos por Bonato et al. (2006), utilizando 78 marcadores AFLP em germoplasma preservado pela Embrapa Soja. A similaridade genética foi estimada pelo método de Nei e Li (1979), através da fórmula da eq. (19),

$$SG_{ij} = \frac{2N_{ij}}{(N_i + N_j)} \quad (19)$$

em que:

- SG_{ij} é o valor da similaridade estimada entre os genótipos i e j , baseados em dados AFLP;
- N_{ij} é o total de bandas comuns em ambos dos genitores; e
- N_i e N_j são os números de bandas encontradas nos genitores i e j , respectivamente.

A partir destas foram obtidas as distâncias genéticas, isto é, eq. (20),

$$DG_{ij} = 1 - SG_{ij} \quad (20)$$

No presente estudo, as distâncias genéticas entre os cinco genitores foram decompostas de acordo com o Modelo I, Método 4 da análise dialélica de Griffing (1956), conforme sugerido por Melchinger et al. (1990a, 1990b), com a finalidade de detalhar a associação entre as distâncias genéticas dos genitores e os diferentes parâmetros genéticos das populações geradas por eles, utilizando a seguinte expressão eq. (21),

$$DG_{ij} = \overline{DG} + DGG_i + DGG_j + DGE_{ij} \quad (21)$$

em que:

- DG_{ij} é a distância genética entre dois genitores;
- $\overline{DG}$ é a distância genética média do grupo de genitores;
- DGG_i é a distância genética geral associada ao genitor “ i ”;
- DGG_j é a distância genética geral associada ao genitor “ j ”; e
- DGE_{ij} é a distância genética específica associada à combinação “ ij ”.

2.2.8 Correlações

A associação entre as distâncias genéticas e as estimativas de variância genética entre progênies ($\hat{\sigma}_{G(F_{2:3})}^2$), médias das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), amplitude de médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$) e resposta esperada com a seleção, foram realizadas através do coeficiente de correlação de Spearman (STEEL; TORRIE, 1997). Da mesma maneira, foram avaliadas as associações entre a soma das distâncias genéticas gerais ($DGG_i + DGG_j$) e distâncias genéticas específicas (DGE_{ij}) e as outras estimativas genéticas através do coeficiente de correlação de Spearman.

2.3 Resultados e discussão

2.3.1 Análise de variância das progênes F_{2:3}

2.3.1.1 Produção de grãos (PG)

Os valores dos quadrados médios e significâncias detectados pelo teste de F na análise de variância, coeficientes de variação experimental, médias ajustadas das progênes F_{2:3} ($\bar{F}_{2:3}$) e graus de liberdade para o caráter produção de grãos, em g.m⁻² estão apresentados nas Tabelas 4 e 5.

As análises de variância detectaram diferenças significativas ($P < 0,05$) ou altamente significativas ($P < 0,01$) para estande (ST) em seis das dez populações avaliadas (Tabela 4), implicando que a correção dos dados de produção de grãos pelo estande através da análise de covariância foi apropriada.

Foram detectadas diferenças significativas ($P < 0,05$) ou altamente significativas ($P < 0,01$) para progênes em todos os cruzamentos avaliados, indicando a ocorrência de variabilidade para o caráter entre as progênes F_{2:3}. Apenas no cruzamento 2x3 (IAC-12 x IAC-100), a análise de variância não detectou diferenças entre as progênes, indicando pouca variabilidade para o caráter PG nessa população. Esse resultado pode ser devido ao alto grau de relacionamento genético entre os genitores desse cruzamento. Os coeficientes de variação (CV) oscilaram entre 14,9% e 20,5%, estando essas magnitudes dentro dos valores relatados na literatura para o caráter PG utilizando o mesmo tipo de parcela, o que implica em uma boa precisão experimental (SILVA, 2005; BARONA; COLOMBARI FILHO; GERALDI, 2009; CARVALHO, 2008; FARIAS, 2008; COLOMBARI FILHO, 2009).

Tabela 4 - Quadrados médios (QM) e respectivas significâncias (teste F) das análises de variância das progênies $F_{2,3}$ para o caráter produção de grãos (PG) de cada cruzamento, coeficientes de variação experimental (CV%) e médias das progênies $F_{2,3}$ ($\bar{F}_{2,3}$), em g.m⁻²

FV	QM				
	1x2	1x3	1x4	1x5	2x3
Repetições	234,2 ^{ns}	16.592,0 ^{ns}	6.212,3 ^{ns}	471,2 ^{ns}	56.560,8 [*]
Blocos/R	7.060,2 ^{**}	5.424,8 ^{**}	2.599,3 ^{ns}	3.351,4 ^{**}	6.975,1 [*]
Progênies	3.635,0 [*]	5.508,6 ^{**}	5.549,2 ^{**}	4.750,1 ^{**}	3.282,8 ^{ns}
ST	10.823,8 [*]	276,1 ^{ns}	12.331,0 [*]	10.509,0 ^{**}	1.554,0 ^{ns}
Erro Intrablocos	2.399,4	2.409,5	2.288,5	1.191,0	3.596,3
CV%	15,7	15,1	14,9	15,1	19,5
$\bar{F}_{2,3}$	312,5	325,4	321,0	228,6	308,4

FV	QM				
	2x4	2x5	3x4	3x5	4x5
Repetições	27,1 ^{ns}	42,2 ^{ns}	8.903,4 ^{ns}	180,1 ^{ns}	3.370,8 ^{ns}
Blocos/R	3.751,7 ^{ns}	2.576,1 ^{ns}	6.764,1 ^{**}	6.503,0 ^{**}	3.839,7 ^{ns}
Progênies	5.124,1 ^{**}	4.951,0 ^{**}	7.414,4 ^{**}	6.921,6 ^{**}	4.825,2 [*]
ST	14.016,4 [*]	7.253,6 [*]	9.195,0 ^{ns}	16.636,1 [*]	2.770,5 ^{ns}
Erro Intrablocos	2.738,6	1.772,1	2.608,5	2.472,1	3.014,5
CV%	20,1	18,4	16,4	20,5	20,3
$\bar{F}_{2,3}$	261,0	229,4	311,2	242,5	270,4

^{*}, ^{**} e ^{ns}: teste F significativo a 5%, 1% e não significativo, respectivamente.

Genitores: 1:Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Tabela 5 - Graus de liberdade das análises de variância das progênies $F_{2,3}$ para o caráter produção de grãos de cada cruzamento

GL	Cruzamentos									
	1x2	1x3	1x4	1x5	2x3	2x4	2x5	3x4	3x5	4x5
Repetições	1	1	1	1	1	1	1	1	1	1
Blocos/R	18	18	18	18	18	18	18	18	18	18
Progênies	98	98	98	98	99	98	98	98	95	99
ST	1	1	1	1	1	1	1	1	1	1
Erro	78	79	78	74	79	77	75	78	71	79

Genitores: 1:Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

As médias ajustadas das progênies $F_{2:3}$ ($\bar{F}_{2:3}$) oscilaram entre 228,6 g.m⁻², para o cruzamento 1x5 (Embrapa-60 x FT-14), a 325,4 g.m⁻², para o cruzamento 1x3 (Embrapa-60 x IAC-100), que correspondem, respectivamente, a 2,29 t.ha⁻¹ e 3,25 t.ha⁻¹. Considerando a produtividade da soja comercial para o ano agrícola 2009-2010 (BRASIL- CONAB), tanto em nível nacional (2,93 t.ha⁻¹), quanto para o Estado de São Paulo (2,77 t.ha⁻¹), pode-se observar que as produções médias obtidas pelas diferentes populações estiveram próximas destes valores. Portanto, considerando que as populações foram derivadas de cruzamento entre cultivares comerciais adaptados, e que as condições experimentais garantem um melhor desenvolvimento da planta em comparação aos plantios comerciais, pelo qual é de esperar que a produtividade nesse sistema seja superior às obtidas pelos agricultores, podem ser consideradas como baixas, as médias ajustadas das progênies $F_{2:3}$. Essa observação é corroborada pelas produtividades obtidas por Colombari Filho (2009) que, usando alguns cruzamentos similares e avaliando na mesma localidade (E. E. Anhembi), obteve médias ajustadas das progênies $F_{2:3}$ superiores às obtidas neste trabalho, reportando produção de grãos de 394,7 g.m⁻² para o cruzamento IAC-12 x IAC-100 e 490,8 g.m⁻² para o cruzamento FT-10 x FT-14. As causas das baixas produtividades obtidas provavelmente sejam as precipitações fora do normal, no ano de avaliação (ano agrícola 2009/10), que foi relativamente atípico, o que poderia ter afetado o desenvolvimento normal das plantas, incidindo negativamente na produção de grãos.

2.3.1.2 Altura das plantas na maturação (AM)

Os valores dos quadrados médios e significâncias detectados pelo teste de F na análise de variância, coeficientes de variação experimental, médias ajustadas das progênies $F_{2:3}$ ($\bar{F}_{2:3}$) e graus de liberdade para o caráter altura das plantas na maturação (AM), em centímetros, estão apresentados nas Tabelas 6 e 7.

Tabela 6 - Quadrados médios (QM) e respectivas significâncias (teste F) das análises de variância das progênies $F_{2:3}$ para o caráter altura das plantas na maturação (AM) de cada cruzamento, coeficientes de variação experimental (CV%) e médias ajustadas das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), em cm

FV	QM				
	1x2	1x3	1x4	1x5	2x3
Repetições	8,9 ^{ns}	165,1 ^{ns}	145,5 [*]	882,6 ^{**}	820,1 ^{**}
Blocos/R	45,9 ^{**}	106,8 ^{**}	26,7 ^{**}	41,6 ^{**}	75,3 ^{**}
Progênies	52,5 ^{**}	93,0 ^{**}	56,6 ^{**}	40,9 ^{**}	69,3 ^{**}
Erro Intrablocos	11,6	10,4	12,0	9,9	22,2
CV%	7,8	7,0	7,7	7,4	10,2
$\bar{F}_{2:3}$	43,7	46,1	45,0	42,6	46,2

FV	QM				
	2x4	2x5	3x4	3x5	4x5
Repetições	174,4 [*]	64,7 ^{ns}	327,3 ^{**}	19,9 ^{ns}	123,1 ^{ns}
Blocos/R	31,0 ^{**}	35,0 [*]	26,5 ^{ns}	52,5 ^{**}	43,2 ^{**}
Progênies	56,5 ^{**}	43,2 ^{**}	115,3 ^{**}	59,8 ^{**}	44,7 ^{**}
Erro Intrablocos	13,0	18,0	22,2	12,3	14,5
CV%	8,5	9,9	10,3	7,6	7,9
$\bar{F}_{2:3}$	42,3	42,7	46,0	46,3	48,4

^{*}, ^{**} e ^{ns}: teste F significativo a 5%, 1% e não significativo, respectivamente.

Genitores: 1:Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Tabela 7 - Graus de liberdade das análises de variância das progênies $F_{2:3}$ para o caráter altura das plantas na maturação de cada cruzamento

GL	Cruzamentos									
	1x2	1x3	1x4	1x5	2x3	2x4	2x5	3x4	3x5	4x5
Repetições	1	1	1	1	1	1	1	1	1	1
Blocos/R	18	18	18	18	18	18	18	18	18	18
Progênies	99	99	99	99	99	99	99	99	99	99
Erro	80	78	80	80	81	79	79	80	76	75

Genitores: 1:Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

As análises de variância mostraram diferenças significativas ($P < 0,05$) ou altamente significativas ($P < 0,01$) para as fontes de variação de repetições e blocos/repetições em cinco e nove experimentos, respectivamente. Neste sentido, o delineamento experimental utilizado foi vantajoso, pois possibilitou a redução do erro experimental, tornando o teste mais sensível. O coeficiente de variação (CV%) oscilou entre 7,0% e 10,3%, indicando uma boa precisão experimental. As análises de variância mostraram diferenças altamente significativas ($P < 0,01$) para a fonte de variação de progênies $F_{2:3}$ em todos os cruzamentos avaliados, evidenciando variabilidade genética para este caráter nas diferentes populações. As médias das progênies $F_{2:3}$ ($\bar{F}_{2:3}$) variaram de 42,3 cm para o cruzamento 2x4 (IAC-12 x FT-10) até 48,4 cm para o cruzamento 4X5 (FT-10 x FT-14), valores que podem ser considerados relativamente baixos. Isso provavelmente ocorreu devido à condição atípica do ano em que os experimentos foram avaliados, que influenciou negativamente o desenvolvimento normal dos genótipos, fato refletido nas médias das progênies $F_{2:3}$ observadas tanto para AM quanto para PG.

2.3.2 Análises das médias

Na Tabela 8 estão apresentadas, para cada cruzamento, as médias dos genitores ($\bar{P}_1$ e $\bar{P}_2$); média das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), média da progênie $F_{2:3}$ superior ($\bar{F}_{2:3(s)}$); média da progênie $F_{2:3}$ inferior ($\bar{F}_{2:3(i)}$); e amplitudes das médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), para os caracteres peso de grãos e altura das plantas na maturação.

Para o caráter produção de grãos (PG), observa-se que as médias das progênies $F_{2:3}$ das diferentes populações não foram homogêneas, sendo que os valores extremos foram 228,6 g.m⁻², para o cruzamento 1x5 (Embrapa-60 x FT-14), e 325,4 g.m⁻², para o cruzamento 1x3 (Embrapa-60 x IAC-100). A maioria dos cruzamentos envolvendo o genitor Embrapa-60 expressaram médias de progênies

Tabela 8 - Médias dos genitores ($\bar{P}_i$ e $\bar{P}_j$), média das progênes $F_{2:3}$ ($\bar{F}_{2:3}$), média da progênie $F_{2:3}$ superior ($\bar{F}_{2:3(s)}$), média da progênie $F_{2:3}$ inferior ($\bar{F}_{2:3(i)}$) e amplitudes das médias individuais das progênes $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), de cada cruzamento para os caracteres produção de grãos (PG) e altura das plantas na maturação (AM)

Cruzamento	PG					
	$\bar{P}_i$	$\bar{P}_j$	$\bar{F}_{2:3}$	$\bar{F}_{2:3(s)}$	$\bar{F}_{2:3(i)}$	$\Delta\bar{F}_{2:3}$
1x2	329,2	254,5	312,5	455,5	162,2	293,4
1x3	297,5	212,9	325,4	439,9	162,2	277,7
1x4	323,5	269,9	321,0	446,0	187,2	258,8
1x5	333,0	181,7	228,6	379,4	105,1	274,2
2x3	251,7	192,2	308,4	466,9	212,2	298,1
2x4	232,7	243,1	261,0	412,9	129,5	283,4
2x5	258,3	226,9	229,4	403,5	104,1	299,4
3x4	274,0	289,8	311,2	480,1	142,5	337,6
3x5	218,2	247,9	241,8	406,7	119,2	287,5
4x5	335,5	205,1	270,4	403,3	168,7	234,6

Cruzamento	AM					
	$\bar{P}_i$	$\bar{P}_j$	$\bar{F}_{2:3}$	$\bar{F}_{2:3(s)}$	$\bar{F}_{2:3(i)}$	$\Delta\bar{F}_{2:3}$
1x2	43,3	44,8	43,7	59,1	29,7	29,4
1x3	47,0	46,7	46,1	64,3	27,6	36,7
1x4	42,9	39,7	45,0	61,2	33,7	27,5
1x5	42,3	43,0	42,6	54,4	31,8	22,6
2x3	41,5	38,0	46,2	64,8	30,0	34,8
2x4	39,2	37,0	42,3	60,6	29,8	30,8
2x5	42,7	47,2	42,7	52,9	29,1	23,8
3x4	43,5	39,4	46,0	65,0	30,9	34,1
3x5	42,5	49,0	46,3	59,7	32,5	27,2
4x5	46,0	50,0	48,4	63,2	33,2	30,0

Genitores: 1: Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14; i= 1,2,3,4; j=2,3,4,5.

$F_{2:3}$ altas. Os valores da média da progênie $F_{2:3}$ superior ($\bar{F}_{2:3(s)}$) e da progênie $F_{2:3}$ inferior ($\bar{F}_{2:3(i)}$) indicam que todos os cruzamentos geraram linhagens transgressivas. As amplitudes de médias individuais de progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$) oscilaram entre 234,6 g.m⁻² para o cruzamento entre os cultivares irmãos 4x5 (FT-10 x FT-14) até 337,6 g.m⁻², para o cruzamento 3x4 (IAC-100 x FT-10), com estimativas de DG de 0,161 e 0,451, respectivamente.

Com relação ao caráter altura das plantas na maturação (AM), tanto as populações quanto os genitores expressaram plantas de porte baixo (Tabela 8). As médias ($\bar{F}_{2:3}$) para este caráter oscilaram de 42,3 cm, para o cruzamento 2x4 (IAC-12 x FT-10), até 48,4 cm, para o cruzamento 4x5 (FT-10 x FT-14).

Todas as populações apresentaram altura das plantas na maturação inferior ao limite máximo estabelecido para seleção nos programa de melhoramento, uma vez que o valor mais alto da progênie $F_{2:3}$ superior foi de 65 cm ($\bar{F}_{2:3(s)}$), e segundo Miranda (1998 apud PELUZIO et al., 2009), o melhoramento genético em soja busca obter cultivares com altura de plantas de 60 até 80 cm, e altura de inserção de primeira vagem de 10 até 15 cm, já que existe uma tendência das plantas mais altas e com menor altura de primeira vagem apresentar um maior número de vagens. Porém, a seleção de plantas muito altas (> 80 cm) e com baixa altura de inserção da primeira vagem (<10 cm) pode acarretar perdas na colheita mecanizada. Considerando que os genitores são cultivares comerciais e que o caráter altura das plantas na maturação está de acordo com o que é considerado aceitável pelo mercado, as alturas expressas por esses genitores neste trabalho, assim como as apresentadas pelas médias das suas respectivas progênies $F_{2:3}$ (todas inferiores a 60 cm), podem ser o resultado de condições ambientais atípicas ocorridas durante o período de desenvolvimento da cultura, especialmente no que se refere ao suprimento de água.

As populações 1x5 (Embrapa-60 x FT-14) e 1x3 (Embrapa-60 x IAC-100) foram as que apresentaram amplitudes de médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$) mais extremas, com valores de 22,6 cm e 36,7 cm, respectivamente, sendo que essas populações também tiveram valores DG de 0,509 e 0,464, respectivamente.

Dessa forma, observa-se que as estimativas de DG de ambas as populações aparentemente não refletem o nível de variabilidade apresentado pelo caráter.

2.3.3 Estimativas das variâncias genéticas entre progênies $F_{2:3}$

As estimativas da variância genética entre as progênies $F_{2:3}$ ($\hat{\sigma}_p^2$) de cada população, bem como seus intervalos de confiança, estão apresentadas nas Tabelas 9 e 10. As estimativas da variância genética que apresentaram valores negativos foram consideradas como nulas. Todas as estimativas das variâncias genéticas para o caráter produção de grãos foram significativamente diferentes de zero ($\alpha=0,05$) (Tabela 9), com exceção da variância do cruzamento 2x3 (IAC-12 x IAC-100). Os cruzamentos expressaram variâncias genéticas que oscilaram entre nulas, no cruzamento 2x3 (IAC-12 x IAC-100), e $2.403,0 \text{ (g.m}^{-2}\text{)}^2$, no cruzamento 3x4 (IAC-100 x FT-10), indicando a ocorrência de diferenças de variabilidade genética entre populações.

A estimativa de variância genética nula do cruzamento 2x3 (IAC-12 x IAC-100) era esperada, já que foram observadas diferenças não significativas para o efeito de progênies na análise de variância para esse cruzamento. Dessa forma, a variabilidade observada para o caráter PG, de acordo com a amplitude das médias individuais das progênies $F_{2:3}$ do cruzamento 2x3 ($298,1 \text{ g.m}^{-2}$, Tabela 8), possivelmente, foi resultante exclusivamente de efeitos ambientais ou de um componente genético, porém de magnitude tão baixa, que não foi possível ser detectado na análise de variância.

Tabela 9 - Estimativas das variâncias genéticas entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$) de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança associados às estimativas, variâncias fenotípicas entre médias de progênies $F_{2:3}$ ($\hat{\sigma}_F^2$) e distâncias genéticas (DG), para o caráter produção de grãos (PG)

	Cruzamentos				
	1x2	1x3	1x4	1x5	2x3
LS	2.675,2	2.960,8	3.118,9	2.793,2	
$\hat{\sigma}_p^2$	617,8	1.549,5	1.630,3	1.779,5	-156,8
LI	282,3	955,7	1.052,3	1.265,5	
$\hat{\sigma}_F^2$	1.817,5	2.754,3	2.774,6	2.375,0	1.641,4
DG	0,367	0,464	0,586	0,509	0,128

	Cruzamentos				
	2x4	2x5	3x4	3x5	4x5
LS	2.967,1	2.777,3	4.162,8	3.917,3	3.139,0
$\hat{\sigma}_p^2$	1.192,8	1.589,4	2.403,0	2.224,8	905,4
LI	675,9	1.058,5	1.609,4	1.470,9	445,6
$\hat{\sigma}_F^2$	2.562,1	2.475,5	3.707,2	3.460,8	2.412,6
DG	0,385	0,334	0,451	0,407	0,161

PG: produção de grãos [(g.m⁻²)²]. Genitores: 1: Embrapa-60; 2: IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Os intervalos de confiança associados às estimativas de variâncias genéticas foram satisfatórios em todos os cruzamentos, o que sugere uma boa precisão. Essas estimativas poderiam ter uma elevada repetibilidade, uma vez que foram seguidos aspectos considerados importantes para alcançar esse objetivo, dentre os quais se destacam o uso do tamanho amostral adequado (100 progênies por cruzamento) (GUMBER et al., 1999; GIZLICE; CARTER; BURTON, 1993b). Em relação ao número de ambientes utilizados, é necessário acrescentar que este trabalho foi desenvolvido com base nos resultados da avaliação de apenas um ambiente, de forma que não foi

possível estimar o componente da variação genética sem a interferência da interação genótipo x ambiente (GARDNER, 1963); porém, futuramente serão conduzidas outras avaliações, procurando obter as estimativas livres desse efeito. Entretanto, a concordância das estimativas de variâncias genéticas obtidas neste trabalho com as relatadas por Colombari Filho (2009), em cruzamentos comuns tais como (IAC-12 x IAC-100) e (FT-10 x FT-14) indicam que elas podem não estar viesadas e que o tamanho amostral utilizado nas populações foi apropriado para obter estimativas confiáveis.

Colombari Filho (2009) relata estimativas de variâncias genéticas para PG nulas e de $1.769,96 \text{ (g.m}^{-2}\text{)}^2$, para os cruzamentos (IAC-12 x IAC-100) e (FT-10 x FT-14), respectivamente, sendo que a variância genética do cruzamento (FT-10 x FT-14) foi equivalente a 36,4% da variância máxima obtida nas populações por eles avaliadas. No presente trabalho, esses mesmos cruzamentos expressaram variância genética para PG nula e $905,4 \text{ (g.m}^{-2}\text{)}^2$, respectivamente (Tabela 9), sendo que a magnitude da variância genética apresentada pelo cruzamento 4x5 (FT-10 x FT-14) foi equivalente a 37,7% da variância máxima obtida nas populações avaliadas. Os cruzamentos 3x4 (IAC-100 x FT-10), 3x5 (IAC-100 x FT-14), 1x5 (Embrapa-60 x FT-14) e 1x4 (Embrapa-60 x FT-10) foram os que expressaram as maiores variâncias genéticas, enquanto que os cruzamentos 2x3 (IAC-12 x IAC-100), 1x2 (Embrapa-60 x IAC-12) e 4x5 (FT-10 x FT-14), expressaram as menores variâncias genéticas. Analisando os dados de distâncias genéticas (DG) e as estimativas de variâncias genéticas, pode-se observar que para PG, aparentemente existe uma associação entre as estimativas de DG e de variâncias genéticas. Observe-se que os quatro cruzamentos que geraram as maiores estimativas de variâncias genéticas para PG (3x4, 3x5, 1x5, e 1x4) estão contidos dentro dos cinco cruzamentos com as maiores estimativas de DG (1x4, 1x5, 1x3, 3x4 e 3x5); enquanto que os cruzamentos que apresentaram as menores estimativas de variâncias genéticas (2x3, 1x2 e 4x5) formam parte dos quatro cruzamentos com as menores estimativas de DG (2x3, 4x5, 2x5 e 1x2).

Para altura das plantas na maturação (AM), os intervalos de confiança associados às estimativas indicam que estas têm precisão satisfatória (Tabela 10).

Tabela 10 - Estimativas das variâncias genéticas entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$) de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança associados às estimativas, variâncias fenotípicas entre médias de progênies $F_{2:3}$ ($\hat{\sigma}_F^2$) e distâncias genéticas (DG), para o caráter altura das plantas na maturação (AM)

	Cruzamentos				
	1x2	1x3	1x4	1x5	2x3
LS	31,1	59,0	33,6	23,7	39,1
$\hat{\sigma}_p^2$	20,4	41,3	22,3	15,5	23,6
LI	14,7	31,1	16,1	11,0	16,1
$\hat{\sigma}_F^2$	26,2	46,5	28,3	20,4	34,7
DG	0,367	0,464	0,586	0,509	0,128

	Cruzamentos				
	2x4	2x5	3x4	3x5	4x5
LS	32,9	24,1	69,6	35,6	25,3
$\hat{\sigma}_p^2$	21,8	12,6	46,6	23,8	15,1
LI	15,5	8,1	34,0	17,2	10,3
$\hat{\sigma}_F^2$	28,2	21,6	57,7	29,9	22,3
DG	0,385	0,334	0,451	0,407	0,161

AM: altura das plantas na maturação [(cm)²] e Genitores: 1: Embrapa-60; 2: IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Todas as estimativas das variâncias genéticas para o caráter AM foram significativamente diferentes de zero ($\alpha=0,05$), variando de 12,6 cm², no cruzamento 2x5 (IAC-12 x FT-14), até 46,6 cm², no o cruzamento 3x4 (IAC-100 x FT-10), indicando a ocorrência de diferenças significativas em relação à variabilidade genética para o caráter nas diferentes populações. Os cruzamentos 3x4 e 1x3 apresentaram as maiores estimativas de variâncias genéticas para o caráter AM, enquanto as populações geradas pelos cruzamentos 2x5, 4x5 e 1x5 foram as que apresentaram as estimativas mais baixas; os demais cruzamentos apresentaram valores intermediários.

Diferentemente ao observado para PG, não foi observada uma clara associação entre as distâncias genéticas (DG) e as estimativas de variâncias genéticas para o caráter AM nas populações avaliadas. Desse modo, pode-se notar que os cruzamentos que apresentaram as maiores variâncias genéticas para AM (3x4 e 1x3), têm estimativas de DG médias (0,451 e 0,464) (Tabela 1), enquanto os cruzamentos que expressaram as menores variâncias genéticas (2x5, 4x5 e 1x5) têm estimativas de DG média, baixa e alta, respectivamente (0,334, 0,161 e 0,509). É importante salientar a presença do genitor 5 (FT-14) nos cruzamentos que apresentaram as menores variâncias genéticas para AM, sugerindo que para altura de planta a cultivar FT-14 apresenta pouca divergência com esse grupo de cultivares (IAC-12, FT-10 e Embrapa-60), fato que não foi detectado em sua totalidade pelos marcadores AFLP.

2.3.4 Estimativas dos coeficientes de herdabilidade

As estimativas dos coeficientes de herdabilidade no sentido amplo entre médias de progênies $F_{2:3}$ ($\hat{h}^2_x$) de cada cruzamento, com seus respectivos intervalos de confiança, para os caracteres PG e AM, estão apresentadas na Tabela 11.

Para o caráter PG, todas as estimativas dos coeficientes de herdabilidade no sentido amplo entre médias de progênies $F_{2:3}$ foram significativamente diferentes de zero ($\alpha=0,05$) em todas as populações, exceto para os cruzamentos 2x3 (IAC-12 x IAC-100) e 1x2 (Embrapa-60 x IAC-12), cujos intervalos de confiança incluíram o valor de zero. As estimativas de coeficiente de herdabilidade oscilaram entre nulas e 74,9%, o que sugere a existência de potenciais diferentes para a seleção de genótipos superiores nas diferentes populações. Os cruzamentos que apresentaram os valores mais altos de herdabilidade, em ordem decrescente, foram 1x5 (Embrapa-60 x FT-14), 3x4 (IAC-100 x FT-10), 3x5 (IAC-100 x FT-14), 2x5 (IAC-12 x FT-14), 1x4 (Embrapa-60 x FT-10) e 1x3 (Embrapa-60 x IAC-100). De maneira geral, os valores de herdabilidade, entre médias de progênies $F_{2:3}$ para PG cresceram em magnitude à medida que as populações foram geradas entre genitores com maiores estimativas de

DG. Nota-se que as populações que apresentaram valores mais altos de herdabilidade foram derivadas de cruzamentos com valores médios ou altos de DG.

Tabela 11 - Estimativas do coeficiente de herdabilidade, no sentido amplo, entre médias de progênes $F_{2:3}(\hat{h}_x^2)$ de cada cruzamento, com os respectivos limites superiores (LS) e inferiores (LI) dos intervalos de confiança e distâncias genéticas (DG), para todos os caracteres avaliados

Caráter	Parâmetro	Cruzamentos									
		1x2	1x3	1x4	1x5	2x3	2x4	2x5	3x4	3x5	4x5
PG	LS	57,0	71,5	73,1	83,8	28,4	65,2	76,8	77,1	77,1	59,2
	$\hat{h}_x^2$	34,0	56,3	58,8	74,9	-9,6	46,6	64,2	64,8	64,3	37,5
	LI	-0,3	33,6	37,3	61,7	-66,0	18,7	45,4	46,6	45,1	5,3
AM	LS	85,5	92,7	86,2	84,2	79,0	85,0	72,9	87,4	86,7	79,0
	$\hat{h}_x^2$	77,9	88,8	78,9	75,8	67,9	77,0	58,5	80,7	79,5	67,5
	LI	66,5	83,0	68,0	63,4	51,5	65,2	37,0	70,8	68,8	50,5
DG		0,367	0,464	0,586	0,509	0,128	0,385	0,334	0,451	0,407	0,161

PG: produção de grãos [(g.m⁻²)²]; AM: altura das plantas na maturação [(cm)²]; Genitores: 1: Embrapa-60; 2: IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Para o caráter altura das plantas na maturação (AM), todas as estimativas do coeficiente de herdabilidade, no sentido amplo, entre médias de progênes $F_{2:3}(\hat{h}_x^2)$ diferiram significativamente de zero ($\alpha=0,05$), oscilando desde 58,5%, no cruzamento 2x5 (IAC-12 x FT-14), até 88,8% no cruzamento 1x3 (Embrapa-60 x IAC-100) (Tabela 11). As populações geradas pelos cruzamentos 1x3, 3x4, 3x5, 1x4, 1x2, 2x4 e 1x5 apresentaram as maiores estimativas de herdabilidade ($\hat{h}_x^2$), de magnitudes semelhantes. Por outro lado, as populações geradas pelos cruzamentos 2x5, 4x5 e 2x3 apresentaram as menores estimativas de herdabilidade ($\hat{h}_x^2$), não muito diferentes entre si. Observa-se que o primeiro grupo de populações tem estimativas de DG intermediárias e altas, enquanto que o segundo grupo de populações, com os menores valores de $\hat{h}_x^2$, apresenta estimativas de DG intermediárias e baixas.

Mesmo considerando que as estimativas de coeficientes de herdabilidade com base em médias de progênies são difíceis de comparar com os resultados relatados na literatura, pelo fato dos valores serem dependentes, entre outros fatores, do tipo de progênie, do número de repetições e de ambientes utilizados para a estimação do parâmetro (HALLAUER; MIRANDA FILHO, 1988), os valores de coeficientes de herdabilidade obtidos neste trabalho são condizentes com os obtidos por outros trabalhos avaliando progênies $F_{2:3}$ (COLOMBARI FILHO, 2009; CARVALHO, 2008).

2.3.5 Resposta à seleção

Na Tabela 12, estão apresentadas as estimativas das médias dos genitores ($\bar{Y}_g$), médias das progênies $F_{2:3}$ ($\bar{Y}_0$), médias das progênies $F_{2:3}$ selecionadas ($\bar{Y}_s$), diferencial de seleção (d_s), resposta esperada com a seleção (R_s) considerando intensidades de seleção de 20% e 40% e média esperada da população ao final do processo de endogamia ($\bar{Y}_m$), para o caráter produção de grãos (PG).

A resposta à seleção das 10 populações foram calculadas a partir das estimativas de variância genética, assumindo uma intensidade de seleção de 20% e 40%, para cada uma das populações. As médias das populações ao final do processo de endogamia ($\bar{Y}_m$) foram estimadas utilizando a média das progênies $F_{2:3}$, isto é, $F_n = \bar{F}_{2:3} + R_s$, devido à inconsistência observada nas médias dos genitores. As causas da inconsistência nas médias dos genitores podem ter sido geradas pelo número reduzido de observações, já que cada experimento contou com apenas seis repetições de cada genitor. As médias das progênies $F_{2:3}$ foram consideradas mais precisas, pois foram utilizadas 100 observações com duas repetições para seu cálculo.

Tabela 12 - Estimativas das médias dos genitores ($\bar{Y}_g$), médias das progênes $F_{2:3}$ originais ($\bar{Y}_0$), médias das progênes $F_{2:3}$ selecionadas ($\bar{Y}_s$), diferencial de seleção (ds), resposta esperada com a seleção (Rs), médias da população ao final do processo de endogamia F_n ($\bar{Y}_m$) e distâncias genéticas (DG), para a produção de grãos (PG), considerando intensidades de seleção de 20% e 40%

Parâmetro	Estimativas com intensidade de seleção de 20%									
	Cruzamentos									
	1x2	1x3	1x4	1x5	2x3	2x4	2x5	3x4	3x5	4x5
$\bar{Y}_g$	291,9	255,2	296,7	257,4	222,0	237,9	242,6	281,9	233,1	270,3
$\bar{Y}_0$	312,5	325,4	321,0	228,6	308,4	261,0	229,4	311,2	241,8	270,4
$\bar{Y}_s$	384,3	400,5	405,1	310,3	369,3	335,3	306,8	395,3	336,8	345,4
ds	71,8	75,1	84,0	81,7	60,9	74,4	77,5	84,1	95,0	75,0
Rs	24,4	42,2	49,4	61,2	0	34,6	49,7	54,5	61,1	28,1
Rs(%)	7,8	13,0	15,4	26,8	0	13,3	21,7	17,5	25,3	10,4
$\bar{Y}_m$	336,9	367,6	370,4	289,8	302,5	295,6	279,1	365,7	302,9	298,5
DG	0,367	0,464	0,586	0,509	0,128	0,385	0,334	0,451	0,407	0,161

Parâmetro	Estimativas com intensidade de seleção de 40%									
	Cruzamentos									
	1x2	1x3	1x4	1x5	2x3	2x4	2x5	3x4	3x5	4x5
$\bar{Y}_g$	291,9	255,2	296,7	257,4	222,0	237,9	242,6	281,9	233,1	270,3
$\bar{Y}_0$	312,5	325,4	321,0	228,6	308,4	261,0	229,4	311,2	241,8	270,4
$\bar{Y}_s$	359,6	380,4	377,0	279,1	348,6	310,6	281,0	369,0	300,6	323,3
ds	47,1	55,0	55,9	50,6	40,2	49,6	51,6	57,8	58,8	52,9
Rs	16,0	30,9	32,9	37,9	0	23,1	33,1	37,4	37,8	19,8
Rs(%)	5,1	9,5	10,2	16,6	0	8,8	14,5	12,0	15,6	7,3
$\bar{Y}_m$	328,5	356,3	353,9	266,4	304,5	284,1	262,5	348,6	279,6	290,2
DG	0,367	0,464	0,586	0,509	0,128	0,385	0,334	0,451	0,407	0,161

Genitores: 1: Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

Sabe-se que, na ausência de epistasia, o comportamento médio das progênes segregantes derivadas do um cruzamento biparental sem efeito de seleção, dependerá do comportamento médio dos genitores e dos efeitos de dominância (MATHER; JINKS, 1982). Assim, a média da geração $F_{2:3}$ é composta de (eq. 22),

$$\mu_{F_{2:3}(P_1 \times P_2)} = \mu_{\bar{P}(P_1 \times P_2)} + \frac{1}{4}d \quad (22)$$

em que:

- $\mu_{\bar{P}(P_1 \times P_2)}$ é a média dos genitores.
- d é a soma dos efeitos de dominância no cruzamento dos genitores $P_1 \times P_2$.

A existência de dominância e/ou epistasia envolvendo dominância faz com que a divergência genética ao nível dos genes com dominância entre os genitores gere diferenças entre as médias das populações e as dos genitores. Em função da baixa magnitude desse viés na geração $F_{2:3}$, que está dividida por quatro, e do alto número de dados utilizados no cálculo das médias entre progênie $F_{2:3}$, a utilização dessas médias foi assumida como a melhor alternativa para estimar $\bar{Y}_m$, no presente trabalho.

As estimativas de resposta à seleção para produção de grãos (Tabela 12) detectaram diferenças na capacidade de ganho entre as populações avaliadas. A resposta seleção foi maior nas populações 1x5 (Embrapa-60 x FT-14), 3x5 (IAC-100 x FT-14) e 3x4 (IAC-100 x FT-10), seguidas das populações 2x5 (IAC-12 x FT-14) e 1x4 (Embrapa-60 x FT-10), que são as populações que apresentaram maior variância genética para o caráter PG. O ganho representado como porcentagens da média das progênie $F_{2:3}$, oscilou de zero, para o cruzamento 2x3 (IAC-12 x IAC-100), em ambas as intensidades de seleção, a 26,8% e 16,6% para o cruzamento 1x5 (Embrapa-60 x FT-14), quando se aplicaram intensidades de seleção de 20% e 40%, respectivamente. As populações com as menores resposta à seleção foram aquelas derivadas dos cruzamentos 2x3 (IAC-12 x IAC-100) e 1x2 (Embrapa-60 x IAC-12).

As estimativas de médias esperadas das linhagens endogâmicas na geração F_n ($\bar{Y}_m$), com intensidades de seleção de 20% e 40%, mostraram que a população com menor $\bar{Y}_m$ foi derivada do cruzamento 2x5 (IAC-12 x FT-14), com 279,1 g.m⁻² e 262,5 g.m⁻², respectivamente; enquanto que a população com maior $\bar{Y}_m$ foi derivada do cruzamento 1x4 (Embrapa-60 x FT-10) com 370,4 g.m⁻² e 353,9 g.m⁻²,

respectivamente, seguida do cruzamento 1x3 (Embrapa-60 x IAC-100) com estimativas muito próximas.

As quatro populações com maiores $\bar{Y}_m$, com ambas as intensidades de seleção, correspondem às aquelas derivadas dos cruzamentos 1x4 (Embrapa-60 x FT-10), 1x3 (Embrapa-60 x IAC-100), 3x4 (IAC-100 x FT-10) e 1x2 (Embrapa-60 x IAC-12) (Tabela 12). Comparando as médias $\bar{Y}_m$ com a produtividade média comercial de soja registrada pela (BRASIL- CONAB), para o ano agrícola 2009-2010, de 2,93 t.ha⁻¹ em nível nacional e 2,77 t.ha⁻¹ para o Estado de São Paulo, observa-se que essas populações apresentam médias $\bar{Y}_m$ superiores às médias de produtividade comercial da cultura, tanto em nível nacional quanto regional.

Os resultados obtidos de reposta à seleção (Rs) devem ser vistos com cautela, visto que existe a possibilidade das estimativas de Rs estarem sub ou superestimadas em razão das variâncias genéticas conterem um provável viés oriundo da interação progênie x ambientes. Carvalho (2008) acrescenta ainda que a seleção realizada em gerações precoces tem maior probabilidade de selecionar progênie cuja superioridade não seja consistente nas gerações seguintes. No entanto, é interessante observar as diferenças entre as populações em relação ao potencial de ganho, ficando evidente que aquelas derivadas de genitores aparentados, tais como 2x3 (IAC-12 x IAC-100) e 4x5 (FT-10 X FT-14), foram as que expressaram os menores potenciais.

2.3.6 Desdobramento das distâncias genéticas em distância genética geral e distância genética específica

As distâncias genéticas (DG) dos cinco genitores utilizados neste trabalho, já haviam sido previamente calculadas por Bonato et al. (2006), com base em marcadores moleculares AFLP e a metodologia proposta por Nei e Li (1979). No presente estudo, essas estimativas foram desdobradas de acordo ao Modelo I Método 4 de Griffing (1956) para análises dialélicas, de acordo à proposta de Melchinger et al. (1990a) e estão apresentadas na Tabela 13.

Mesmo quando a partição da divergência genética dos genitores em seus componentes geral e específico objetivou verificar as mudanças na associação entre esse parâmetro e os outros parâmetros genéticos estimados nas populações $F_{2:3}$ derivadas desses genitores, especialmente com as variâncias genéticas, esses dados aportam informação genética adicional interessantes para serem analisados. Neste

Tabela 13 - Distâncias genéticas (DG: acima da diagonal), distâncias genéticas geral (DGG), distância genética média ($\overline{DG}$), e distâncias genéticas específicas (DGE_{ij} : abaixo da diagonal), calculadas a partir de marcadores AFLP e desdobradas de acordo com o método de Griffing (1956)

Genitores	Genitores					DGG
	1	2	3	4	5	
1		0,3670	0,4640	0,5860	0,5090	0,1364
2	-0,0477		0,1280	0,3850	0,3340	-0,1009
3	-0,0293	-0,1280		0,4510	0,4070	-0,0223
4	0,0483	0,0847	0,0720		0,1610	0,0221
5	0,0287	0,0910	0,0853	-0,2050		-0,0353
					$\overline{DG}$	0,3792

Genitores: 1: Embrapa-60; 2:IAC-12; 3: IAC-100; 4: FT-10 e 5: FT-14.

sentido, tem-se que a partição da divergência genética dos genitores em seus componentes geral e específico mostrou que, para os cruzamentos avaliados, as estimativas de distâncias genéticas específicas (DGE_{ij}) foram geralmente maiores que as estimativas de distância genética geral (DGG). O uso de germoplasma homogêneo pode ter causado esse efeito, pois de acordo com Freire e Filho (1988), quando o conjunto de genitores é heterogêneo espera-se que a capacidade geral de combinação seja superior à capacidade específica. Esse fato tem sido comprovado por trabalhos realizados em milho e soja. Melchinger et al. (1990a, 1992), avaliando populações derivadas de cruzamentos entre linhagens de milho dentado e duro, populações derivadas de cruzamentos entre diferentes tipos de linhagens (oriundas de variedades de polinização aberta ou cruzamentos dirigidos), e linhagens selecionadas pelo seu comportamento agrônomo, observaram que as DGE_{ij} foram

superiores às estimativas de DGG. Quando avaliaram populações de milho derivadas do cruzamento entre linhagens liberadas em épocas diferentes (antes de 1960 e durante a década de 1970), observaram que tanto a DGG quanto a DGE_{ij} explicaram 50% da variação dos valores da DG (MELCHINGER et al (1990b). Adicionalmente, Barroso et al. (2003), avaliando populações derivadas de cruzamento entre genitores adaptados e não adaptados em soja, observaram que as estimativas de DGG foram, geralmente, maiores que as estimativas DGE_{ij} .

As estimativas de distâncias genéticas geral (DGG) apresentadas pelos genitores 3, 4 e 5 (IAC-100, FT-10 e FT-14), próximas de zero (Tabela 13), indicam que os seus respectivos valores médios de distâncias genéticas (DG) não diferem muito da distância genética média (0,379). Por apresentar valores negativos de DGG, os genitores com maior número de alelos em comum foram: 2 (IAC-12), 3 (IAC-100) e 5 (FT-14), sendo que o primeiro deles é o menos divergente. Em contraste, o genitor 1 (Embrapa-60) apresenta o valor de DGG mais alto e positivo, e é o genitor com a maior quantidade de alelos que ocorrem em baixa frequência no grupo restante dos genitores estudados.

Com relação às estimativas de distância genética específica (DGE_{ij}), observa-se que os cruzamentos 1x2 (Embrapa-60 x IAC-12), 1x3 (Embrapa-60 x IAC-100), 2x3 (IAC-12 x IAC-100) e 4x5 (FT-10 x FT-14), apresentaram estimativas negativas de DGE_{ij} , sendo que os genitores dos cruzamentos 2x3 e 4x5 são aparentados. As estimativas negativas de DGE_{ij} indicam uma menor capacidade de complementação, do ponto de vista dos marcadores moleculares AFLP, de forma que esses cruzamentos são os de maior similaridade genética do grupo e, conseqüentemente, devem gerar populações com baixa variabilidade genética.

Segundo Almeida et al. (2011), cruzamentos com maiores valores de DGE_{ij} , envolvendo genitores contrastantes em DGG, podem gerar segregação transgressiva se os genes com efeitos aditivos forem complementares e atuarem na mesma direção de maximizar a expressão do caráter. Os resultados deste trabalho mostraram que a obtenção de segregação transgressiva foi independente dos valores de DGE_{ij} e DGG, pois todas as populações expressaram esse tipo de segregação (Tabela 8). As maiores estimativas de variâncias genéticas para PG foram obtidas por populações

derivadas de cruzamentos envolvendo diferentes combinações de valores DGE_{ij} e DGG ; como exemplo tem-se que as populações 3x4 (IAC-100 x FT-10) e 3x5 (IAC-100 x FT-14) com variâncias genéticas de $2.403,0 (g.m^{-2})^2$ e $2.224,8 (g.m^{-2})^2$, respectivamente, apresentaram altos valores positivos de DGE_{ij} e valores baixos, positivos ou negativos, de DGG ; enquanto que os cruzamentos 1x5 (Embrapa-60 x FT-14) e 1x4 (Embrapa-60 x FT-10), com variâncias genéticas $1.779,5 (g.m^{-2})^2$ e $1.630,3 (g.m^{-2})^2$, respectivamente, apresentaram baixos valores de DGE_{ij} e a combinação de um genitor com alto valor positivo de DGG , (Embrapa-60), com genitores com valores de DGG intermediários.

É importante acrescentar o fato de que mesmo quando as populações com maiores potenciais apresentaram combinações diferentes de DGE_{ij} e DGG , as populações com menores potenciais coincidiram em apresentar valores negativos de DGE_{ij} . Neste sentido, as populações 2x3 (IAC-12 x IAC-100), 1x2 (Embrapa-60 x IAC-12) e 4x5 (FT-10 X FT-14), com valores negativos de DGE_{ij} , expressaram, para o caráter PG, estimativas de variâncias genéticas e de resposta à seleção de magnitudes baixas ou nulas.

2.3.7 Associações entre as estimativas de parâmetros e as distâncias genéticas

As correlações de Spearman $[r(DG, Y)]$ entre as distâncias genéticas determinadas por AFLP (DG), e as estimativas das variâncias genéticas entre progênes $F_{2:3} (\hat{\sigma}_p^2)$, médias das progênes $F_{2:3} (\bar{F}_{2:3})$, amplitudes das médias das progênes $F_{2:3} (\Delta \bar{F}_{2:3})$, coeficientes de herdabilidade no sentido amplo entre médias de progênes $F_{2:3} (\hat{h}_x^2)$ e respostas à seleção, com intensidades de seleção de 20% e 40%, para os caracteres PG e AM, estão apresentadas na Tabela 14.

Para o caráter produção de grãos, as análises de correlações detectaram associação entre as medidas de divergência genética (DG) e as estimativas das variâncias genéticas entre progênes $F_{2:3}$, coeficientes de herdabilidade no sentido

Tabela 14 - Correlações de Spearman (r_s) entre as distâncias genéticas determinadas por marcadores AFLP (DG), e as estimativas dos demais parâmetros [$r(\text{DG}, Y)$]: estimativa da variância genética entre progênies $F_{2:3}$ ($\hat{\sigma}_p^2$); média das progênies $F_{2:3}$ ($\bar{F}_{2:3}$); amplitude das médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), estimativas do coeficiente de herdabilidade, no sentido amplo, entre médias de progênies $F_{2:3}$ ($\hat{h}_x^2$) e, estimativas de resposta à seleção (Rs) com intensidades de seleção de 20% e 40% para os caracteres produção de grãos (PG) e altura das plantas na maturação (AM)

r_s	PG	AM
$[r(\text{DG}, \hat{\sigma}_p^2)]$	0,68 *	0,36 ^{ns}
$[r(\text{DGG}_i + \text{DGG}_j, \hat{\sigma}_p^2)]$	0,26 ^{ns}	0,27 ^{ns}
$[r(\text{DGE}_{ij}, \hat{\sigma}_p^2)]$	0,65 *	0,05 ^{ns}
$[r(\text{DG}, \bar{F}_{2:3})]$	0,21 ^{ns}	-0,29 ^{ns}
$[r(\text{DGG}_i + \text{DGG}_j, \bar{F}_{2:3})]$	0,56 ^{ns}	-0,02 ^{ns}
$[r(\text{DGE}_{ij}, \bar{F}_{2:3})]$	-0,42 ^{ns}	-0,43 ^{ns}
$[r(\text{DG}, \Delta\bar{F}_{2:3})]$	0,10 ^{ns}	-0,21 ^{ns}
$[r(\text{DGG}_i + \text{DGG}_j, \Delta\bar{F}_{2:3})]$	-0,19 ^{ns}	0,02 ^{ns}
$[r(\text{DGE}_{ij}, \Delta\bar{F}_{2:3})]$	0,67 *	-0,43 ^{ns}
$[r(\text{DG}, \hat{h}_x^2)]$	0,66 *	0,65 *
$[r(\text{DGG}_i + \text{DGG}_j, \hat{h}_x^2)]$	0,22 ^{ns}	0,60 ^{ns}
$[r(\text{DGE}_{ij}, \hat{h}_x^2)]$	0,62 *	0,13 ^{ns}
$[r(\text{DG}, \text{Rs20})]$	0,65 *	
$[r(\text{DGG}_i + \text{DGG}_j, \text{Rs20})]$	0,20 ^{ns}	
$[r(\text{DGE}_{ij}, \text{Rs20})]$	0,65 *	
$[r(\text{DG}, \text{Rs40})]$	0,65 *	
$[r(\text{DGG}_i + \text{DGG}_j, \text{Rs40})]$	0,20 ^{ns}	
$[r(\text{DGE}_{ij}, \text{Rs40})]$	0,65 *	

* e ^{ns}: teste t significativo a 5% e não significativo, respectivamente.

amplo, bem como com as estimativas de resposta à seleção, tanto com intensidades de seleção de 20% como 40%, sendo todas essas correlações positivas, de magnitudes relativamente altas e significativas ($P < 0,05$). A associação entre a divergência genética e a média das progênes $F_{2:3}$ ($\bar{F}_{2:3}$) com a amplitude das médias individuais das progênes $F_{2:3}$ ($\Delta \bar{F}_{2:3}$) foram positivas, porém de baixas magnitudes e não significativas.

As correlações entre DG e $\hat{\sigma}_p^2$, $\hat{h}_x^2$ e Rs, de magnitudes semelhantes, possuem valores preditivos, pelo qual são úteis para predizer esses parâmetros genéticos das populações. Esses resultados sugerem que valores relativamente altos de DG entre genitores produzam populações de maior potencial, resultando em maior variância genética, herdabilidade e resposta à seleção para o caráter PG. Neste sentido, observa-se que, dos cinco cruzamentos com estimativas de DG superiores (1x4, 1x5, 1x3, 3x4 e 3x5), quatro deles apresentaram as maiores variâncias genéticas (3x4, 3x5, 1x5 e 1x4), com altas $\hat{h}_x^2$ (superiores a 58%) e as maiores respostas à seleção, as quais variaram de 15,4 a 26,8%, e de 10,2 a 16,6%, quando se aplicam intensidades de seleção de 20% e 40%, respectivamente. A presença de associação entre DG e variância genética também foi observada por Colombari Filho (2009) utilizando as mesmas estimativas de DG obtidas por Bonato et al. (2006), assim como por Manjarrez-Sandoval et al. (1997b) utilizando marcadores RFLP. Estes resultados mostram que é possível utilizar as medidas de distâncias genéticas (DG) geradas por marcadores moleculares, para auxiliar na escolha de genitores, ainda que muitos trabalhos na área de predição mostrem o contrario (HELMS et al., 1997; KISHA; SNELLER; DIERS, 1997; GUMBER et al., 1999; MOSER; LEE, 1994; BURKHAMER et al., 1998).

A ausência de associação entre DG e média das progênes $F_{2:3}$ ($\bar{F}_{2:3}$) era esperada. Segundo Barroso et al. (2003), as correlações entre DG e as médias das populações podem existir quando as populações são derivadas de genitores adaptados e não adaptados, condição ausente nos genitores do presente trabalho. No entanto, a associação entre DG e $\Delta \bar{F}_{2:3}$ era esperada, visto que esse parâmetro

reflete a variabilidade do caráter. A ausência de correlação entre DG e $\Delta\bar{F}_{2:3}$ para o caráter PG no presente trabalho, pode ter sido causada pelo estresse sofrido pelas populações, acarretando uma baixa produção de grãos. Na Figura 1 estão apresentados os histogramas das médias individuais das progênies $F_{2:3}$ ($\bar{F}_{2:3}$), para o caráter produção de grãos (PG). Foram utilizadas classes de intervalos de 25 g.m^{-2} , correspondentes a cerca de $(0,25)\sigma_{\bar{F}}$, em que $\sigma_{\bar{F}}$ é o desvio padrão fenotípico médio entre progênies $F_{2:3}$ dos 10 cruzamentos. Pode-se notar que as distribuições das médias individuais das progênies $F_{2:3}$ ($\bar{F}_{2:3}$) nas diferentes populações não apresentam uma distribuição normal perfeita, existindo dados discrepantes extremos como os observados nas populações derivadas dos cruzamentos 1x2, 2x3, 2x5, 1x5 e 3x4, os quais, mesmo sendo gerados por poucas observações, têm efeito na determinação das amplitudes de médias das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), podendo minimizar as diferenças entre as populações. Observam-se poucas diferenças entre as médias individuais das progênies superiores ($\bar{F}_{2:3(s)}$) e inferiores ($\bar{F}_{2:3(i)}$), refletidas na baixa amplitude de médias das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), as quais oscilaram de $234,6 \text{ g.m}^{-2}$ a $337,6 \text{ g.m}^{-2}$ (Tabela 8). Consequentemente, as $\Delta\bar{F}_{2:3}$ não tiveram relação tanto com as estimativas de DG, quanto com as variâncias genéticas, visto que as populações com maiores $\Delta\bar{F}_{2:3}$ (3x4, 2x5, 2x3, 1x2 e 3x5) eram oriundas de genitores com DG baixas e intermediárias, e $\hat{\sigma}_p^2$ de nulas a altas.

Colombari Filho (2009) obteve correlações de Spearman entre DG e $\Delta\bar{F}_{2:3}$ de magnitudes máximas ($r=1,00$) e altamente significativas ($P<0,01$), concluindo que as estimativas de DG derivadas de marcadores moleculares AFLP (BONATO et al, 2006) foram preditoras de $\Delta\bar{F}_{2:3}$ para o caráter PG. Nesse trabalho, as distribuições das médias individuais das progênies $F_{2:3}$ ($\bar{F}_{2:3}$) nas diferentes populações foram mais homogêneas que no presente trabalho, apresentando dados com valores extremos discrepantes em apenas duas populações, sendo essas populações geradas por cruzamentos entre genitores com altas distâncias genéticas. Além disso, as $\Delta\bar{F}_{2:3}$

Figura 1 - Histogramas das médias individuais das progênes $F_{2:3}$ ($\bar{F}_{2:3}$), dos 10 cruzamentos para produção de grãos (PG). Eixo Y: Frequências; Eixo X: Produção de grãos em $g.m^{-2}$

foram mais acentuadas, variando de 304,4 g.m⁻² a 690,4 g.m⁻², sendo que as médias individuais das progênies superiores ($\bar{F}_{2:3(s)}$) e inferiores ($\bar{F}_{2:3(i)}$) oscilaram entre 121,8 g.m⁻² e 815,40 g.m⁻², com grandes diferenças entre as populações, dependendo da DG de seus genitores.

Deve-se ressaltar que neste trabalho, mesmo quando houve condições ambientais que afetaram a expressão do caráter PG nas populações, fazendo com que as variâncias fenotípicas não refletissem diferenças razoáveis entre elas, o elevado número de progênies avaliadas (100) tornou possível a detecção do componente genético de maneira precisa, tal que a associação entre DG e $\hat{\sigma}_p^2$, não foi prejudicada. Aparentemente o poder preditivo das distâncias genéticas (DG) sobre as amplitudes de médias das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$) é mais sensível ao normal desenvolvimento das progênies avaliadas.

Para o caráter altura das plantas na maturação, as análises de correlações detectaram associação apenas entre divergência genética e herdabilidade, no sentido amplo, entre médias de progênies $F_{2:3}$, com magnitude alta e significativa ($P < 0,05$). Para o resto dos parâmetros, as correlações foram não significativas e de magnitudes baixas. Os resultados sugerem que, de forma independente, as estimativas de variância fenotípica e genética para altura de planta na maturação entre as populações, sob as condições avaliadas, aumentaram o grau de correspondência entre o valor fenotípico e valor genético à medida que a estimativa de DG aumentava. A ausência de correlação entre a DG e os demais parâmetros genéticos para o caráter AM coincidiram com os obtidos por Colombari Filho (2009), usando marcadores AFLP, assim como os obtidos por Barroso et al. (2003), usando marcadores RADP em soja. Kisha; Sneller e Diers (1997) utilizando marcadores moleculares RFLP detectaram associação entre a DG e a variância genética para AM em soja, porém essa associação não foi consistente por ter variado entre o conjunto de populações e ambientes utilizados.

A concordância dos resultados com aqueles obtidos por Colombari Filho (2009), para o caráter altura das plantas na maturação, indica que as estimativas de DG determinadas pelos marcadores moleculares AFLP, obtidas por Bonato et al. (2006),

não são preditoras para AM. Isso implica que, mesmo quando esses marcadores AFLP estiveram dispersos pelo genoma, pois foram capazes de refletir a divergência genética para o caráter produção de grãos, não puderam refletir a DG existente para AM. As causas disso podem ser atribuídas à ausência de desequilíbrio de ligação entre os marcadores moleculares e os QTLs determinantes da AM, por efeito da existência de QTLs ligados ou QTLs com efeitos desiguais sobre o caráter. Outra possibilidade seria que os marcadores AFLP estejam ligados aos QTLs que controlam o caráter AM, porém em fase de ligação variável nos genitores (HELMS et al., 1997; KISHA; SNELLER; DIERS, 1997; BARROSO; HOFFMANN, 2003). Além disso, existe a possibilidade que a seleção de genitores utilizados neste trabalho possa ter incidido neste resultado, devido ao fato que a seleção foi feita por critérios pré-estabelecidos tendo como base principalmente as distâncias genéticas pré-existentes por AFLP, seguidos da adaptabilidade dos genitores às condições edafoclimáticas da área experimental, dos ciclos e dos rendimentos produtivos, não sendo levado em conta o caráter AM. Entretanto, considerando que no processo de seleção nos programas de melhoramento, a AM é fixada dentro de limites próximos, existindo um intervalo de variação permitido para esse caráter de apenas 20 cm (MIRANDA, 1998 apud PELUZIO et al., 2009), é plausível que o efeito deste tipo de seleção gere a existência de diferentes graus de divergência genética entre dois cultivares dependendo da região do genoma, podendo existir regiões do genoma que tenham uma divergência genética que esteja em função da origem dos cultivares e regiões do genoma com divergência genética que estejam em função do efeito de seleção aplicado à cultura. Neste sentido, pode ser que a divergência genética para os QTLs determinantes da AM entre os cultivares comerciais de soja seja suficientemente baixa, que não reflitam as diferenças dos cultivares pela sua origem. Isso implica que a utilização de distâncias genéticas geradas por marcadores moleculares distribuídos aleatoriamente no genoma, não seja útil na predição deste caráter em soja.

Observe-se que neste trabalho, mesmo quando a expressão da AM foi afetada pelas condições ambientais atípicas ocorridas durante o período de desenvolvimento da cultura, a diferença entre as médias das progênes $F_{2:3}$ das populações avaliadas

foi pequena, alcançando apenas 6,1cm. Além disso, as variâncias genéticas para o caráter AM foram independentes das estimativas DG, sendo que os cruzamentos que expressaram as menores variâncias genéticas (2x5, 4x5 e 1x5) têm DG média, baixa e alta, respectivamente (0,334, 0,161 e 0,509), ao passo que os cruzamentos que apresentaram as maiores variâncias genéticas para AM (3x4 e 1x3), têm DG médias (0,451 e 0,464).

O desdobramento da DG para o caráter PG, segundo a metodologia de Griffing (1956), mostrou em todos os casos, exceto para $\bar{F}_{2:3}$, que o componente correspondente à DGE_{ij} foi o que esteve mais associado com as mesmas estimativas (Tabela 14). Em geral, o desdobramento não causou um aumento na associação entre os caracteres avaliados, pois as magnitudes das associações entre os parâmetros genéticos e a DGE_{ij} foram semelhantes à associação entre estes parâmetros e a estimativa de DG. No entanto, para a correlação entre DG com $\bar{F}_{2:3}$ e $\Delta\bar{F}_{2:3}$, o desdobramento da DG foi efetivo, ao detectar um aumento na associação, sendo que a maior delas ocorreu para o caso específico de $\Delta\bar{F}_{2:3}$ com a DGE_{ij} onde a associação foi de magnitude alta, positiva e significativa ($P < 0,05$).

Para o caráter AM, o desdobramento da DG mostrou que tanto para variância genética, quanto para herdabilidade, no sentido amplo, entre médias de progênies $F_{2:3}$, o componente correspondente à soma das distâncias genéticas geral ($DGG_i + DGG_j$) foi o que esteve mais associado às estimativas, enquanto que para $\bar{F}_{2:3}$ e amplitude das médias individuais das progênies $F_{2:3}$ ($\Delta\bar{F}_{2:3}$), foi o componente correspondente a DGE_{ij} que mais esteve associado. De forma similar à PG, o desdobramento aumentou a associação entre $\bar{F}_{2:3}$ e DGE_{ij} , assim como entre $\Delta\bar{F}_{2:3}$ e DGE_{ij} . Entretanto a maioria das correlações foi baixa e não significativa, não tendo valor preditivo.

Estes resultados corroboram com os obtidos por outros autores, no sentido de que o desdobramento das estimativas da DG apenas aumenta a associação entre essa estimativa e os outros parâmetros genéticos de interesse, em magnitudes pequenas. Melchinger et al. (1990a, 1990b e 1992), trabalhando com milho, avaliou a associação entre a distância genética determinada a partir de marcadores

moleculares RFLP, a capacidade específica de combinação (GRIFFING, 1956) e a heterose específica (GARDNER; EBERHART, 1966). Os autores observaram, de maneira geral, aumentos na associação, especialmente com o componente de DGE_{ij} , mesmo quando as magnitudes das correlações não alcançaram valores preditivos. Barroso et al. (2003), utilizando populações em diferentes níveis de endogamia ($F_{2:8}$ até $F_{2:11}$) provenientes de cruzamentos entre genitores adaptados e não adaptados de soja, observaram que a decomposição da distância genética não melhorou a associação entre a DG e a variância genética tanto para PG quanto para AM.

3 CONSIDERAÇÕES FINAIS

Os resultados obtidos no presente trabalho sugerem que as estimativas de distâncias genéticas determinadas pelos marcadores moleculares AFLP determinadas por Bonato et al. (2006) apresentaram capacidade preditora para o comportamento das populações $F_{2:3}$, uma vez que foram observadas correlações altas e significativas entre as estimativas de DG e a variância genética entre progênies $F_{2:3}$, o coeficiente de herdabilidade no sentido amplo, bem como com as estimativas de resposta à seleção tanto com intensidades de seleção de 20% quanto com 40% para o caráter produção de grãos (PG). Para o caráter altura das plantas na maturação (AM), essas estimativas não tiveram capacidade preditiva. Em geral, o desdobramento das distâncias genéticas (DG) segundo o método de Griffing (1956) não causou um aumento na associação entre os caracteres avaliados.

Segundo os resultados obtidos em varias culturas, nenhuma estimativa de diversidade genética obtida por marcadores moleculares é capaz de predizer de modo consistente a variância genética das populações (MOSER, LEE, 1994; HELMS et al., 1997; KISHA; SNELLER; DIERS, 1997; BURKHAMER et al., 1998; GUMBER et al., 1999). Portanto, é oportuna a sugestão de Manjarrez-Sandoval et al. (1997b) de considerar com cautela a associação entre as estimativas de DG e variância genética, quando elas são detectadas pela análise de pequeno número de cruzamentos. No entanto, é necessário destacar que, neste trabalho, foi utilizado um tamanho amostral satisfatório de acordo com o sugerido para estimação da variância genética com alta repetibilidade (GUMBER et al., 1999; GIZLICE; CARTER.; BURTON, 1993b). Além disso, a concordância na associação entre as estimativas de DG e variância genética para os caracteres produção de grãos e altura das plantas na maturação obtidos neste trabalho e as apresentadas por Colombari Filho (2009), utilizando as mesmas estimativas de DG geradas por Bonato et al. (2006), sugerem que, mesmo quando as populações foram avaliadas em apenas um ambiente, houve boa precisão das estimativas genéticas.

Considerando que os genitores utilizados neste trabalho apresentaram estimativas de DG não contrastantes, sendo alguns deles altamente relacionados, os

resultados obtidos são relevantes, pelo fato de que esses tipos de genitores podem gerar estimativas de DG viesadas (MIRANDA et al., 2001). Observa-se que os marcadores AFLP detectaram maior similaridade genética entre os cultivares IAC-12 e IAC-100, que entre as linhagens irmãs FT-10 e FT-14, o que não está de acordo com as similaridades genéticas esperadas, considerando a informação de origem dos genitores (Tabela 2). Os resultados obtidos neste trabalho confirmam que para PG, as estimativas de DG com marcadores AFLP foram precisas, já que a população 4x5 (FT-10 x FT-14) apresentou variância genética maior que a população 2x3 (IAC-12 x IAC-100), cuja estimativa não deferiu estatisticamente de zero. Quando são comparados estes resultados com os obtidos por Colombari Filho (2009), para esses mesmos cruzamentos, considerando o caráter PG, observa-se que são altamente similares. Em ambos os trabalhos a população gerada pelo cruzamento 2x3 não gerou variância genética, enquanto que a população gerada pelo cruzamento 4x5 sim. Esta variância genética foi equivalente a 37,7% da variância máxima obtida nas populações avaliadas no presente trabalho, em comparação a 36,4% da variância máxima obtida nas populações avaliadas por Colombari Filho (2009). Estes resultados sugerem que o nível de similaridade genética determinada pelos marcadores AFLP nos genitores considerados neste trabalho refletem a real DG para o caráter PG. Por isso, pode-se considerar que essas estimativas de diversidade genética têm capacidade preditiva sobre o comportamento das populações endogâmicas de soja para o caráter PG.

A concordância dos resultados com aqueles obtidos por Colombari Filho (2009), para o caráter altura das plantas na maturação, indica que, mesmo quando os marcadores AFLP estiverem dispersos pelo genoma, eles não predizem a DG existente para esse caráter. As causas disso podem ocorrer em função da ausência de desequilíbrio de ligação entre os marcadores moleculares e os QTLs determinantes da AM, por efeito da existência de QTLs ligados ou QTLs com efeitos desiguais sobre o caráter. Outra possibilidade seria que os marcadores AFLP estejam ligados aos QTLs que controlam o caráter, porém em fase de ligação variável nos genitores (HELMS et al., 1997; KISHA; SNELLER; DIERS, 1997; BARROSO; HOFFMANN, 2003).

No desdobramento das distâncias genéticas (DG), mesmo que não tenha melhorado a associação entre DG e os parâmetros genéticos utilizados, as estimativas de DGE_{ij} identificaram as populações com menores potenciais. Portanto, mesmo quando aparentemente seu uso não tenha o efeito desejado, a utilização do desdobramento das DG poderia ser uma ferramenta útil para o processo de seleção de genitores, pois poderia ajudar no processo de identificação de cruzamentos com tendência de gerar populações com baixo potencial, evitando sua execução. Com isso, as melhores combinações de genitores poderiam ser feitas de maneira indireta, levando em conta tanto as distâncias genéticas, quanto a capacidade de produção e adaptação geral (HELMS et al., 1997; SARAWAT et al., 1994a).

Entretanto, é necessário progredir as pesquisas na área de predição, considerando que para alcançar o aumento da eficiência dos programas de melhoramento são necessários resultados de predições consistentes, assim como a incorporação de estimativas de DG com poder preditivo na rotina dos melhoristas. Neste sentido, é preciso realizar outras avaliações das populações utilizadas neste trabalho, visto que os resultados foram obtidos com a avaliação em apenas um ambiente. Mesmo assim, devido ao considerável número de cultivares de soja utilizadas por Bonato et al. (2006), novas pesquisas envolvendo outros ambientes e grupos de genitores serão necessárias para identificar se esse poder preditivo é uniforme para todo esse grupo de cultivares.

Finalmente, é importante acrescentar que a desconsideração da localização dos QTLs associados aos caracteres de interesse, nos trabalhos de predição, seguirá gerando resultados inconsistentes, uma vez que a associação entre marcadores moleculares e esses QTLs depende do acaso. Observa-se, como nos trabalhos de predição, a obtenção de associação entre DG e variância genética dos principais caracteres de interesse é baixa, e mesmo quando se detecta, ela não é igual para todos os caracteres avaliados. Os resultados do presente trabalho apontam que os marcadores AFLP usados por Bonato et al. (2006) estão associados apenas ao caráter PG. Nesse sentido, mesmo quando implique maior trabalho, é acertada a sugestão feita por Melchinger (1990a; 1990b; 1992), em relação à utilização prioritária de marcadores específicos ligados às regiões do genoma que contribuam para a

expressão do caráter de interesse, ligados a estratégias que maximizem a associação entre caracteres. No entanto, visto que nem todas as pesquisas na área de predição em soja poderiam realizar esse trabalho adicional e que PG é o caráter de maior importância nos programas de melhoramento, poderia ser estratégico que, na escolha de marcadores moleculares, os futuros trabalhos levem em consideração os marcadores usados por Bonato, et al. (2006), por serem possíveis peças chaves para a obtenção de resultados consistentes nesse caráter.

REFERÊNCIAS

ABDELNOOR, R.V.; BARROS, E.G.; MOREIRA, M.A. Determination of genetic diversity within Brazilian soybean germplams using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. **Revista Brasileira de Genética**, Ribeirão Preto, v. 18, n. 2, p. 265-273, 1995.

ALCÂNTARA NETO, F. **Marcadores microssatélites na identificação de cultivares de soja**. 2001. 46p. Tese (Mestrado em Fitotecnia) - Universidade Federal de Viçosa, Viçosa, 2001.

ALMEIDA, A; KIIH, R; MIRANDA, M; CAMPELO, G. **Melhoramento da soja para regiões de baixas latitudes - Recursos Genéticos e Melhoramento de Plantas para o Nordeste Brasileiro**. Londrina: EMBRAPA-CNPSO, 1999a. 15 p. Disponível em: <<http://www.cpatsa.embrapa.br:8080/catalogo/livrorg/sojamelhoramento.pdf>> Acesso em: 10 dez. 2010.

ALMEIDA, R.D.de; PELUZIO, J.M.; AFFÉRI, F.S. Divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do Estado Tocantins. **Revista Ciência Agronômica**, Fortaleza, v. 42, n. 1, p. 108-115, 2011.

ALMEIDA, L.A.; KIIHL, R.A.S.; MIRANDA, L.C.; YORINORI, J.T.; DOMIT, L.A.; PIPOLO, A.E.; KASTER, M. Cultivar de soja Embrapa 60. In: CONGRESSO BRASILEIRO DE SOJA, 1999, Londrina. **Anais**. Londrina: Embrapa Soja, 1999b. p.486. Disponível em: <http://orton.catie.ac.cr/cgi-bin/wxis.exe/?IisScript=ACERVO.xis&method=post&formato=2&cantidad=1&expresion=mfn=018073>. Acesso em: 10 dez. 2010.

AN, W.; ZHAO, H.; YINGSHANDONG; WANG, Y.; LI, Q.; ZHUANG, B.; GONG, L.; LIU, B. Genetic diversity in annual wild soybean (*Glycine Soja* Sieb. Et Zucc.) and cultivated soybean (*G. Max*. Merr.) from different latitudes in China. **Pakistan Journal of Botany**, Karachi, v. 41, n. 5, p. 2229-2242, 2009.

BAKER, R.J. Issues in diallel analysis. **Crop Science**, Madison, v.18, n.4, p.533-536, 1978.

BARBOSA-NETO, J.F.; SORRELLS, M.E.; CISAR, G. Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. **Genome**, Ottawa, v. 39, p. 1142-1149, 1996.

BARONA, M.A.A; COLOMBARI FILHO, J. M.; GERALDI, I.O. Utilización del diseño TTC para detectar epistasis en soja. **Agronomía Tropical**, Maracay, v. 59, n. 3, p. 323-330, 2009.

BARROSO, P.A.V.; GERALDI, I.O.; VIEIRA, M.L.C.; PULCINELLI C.E.; VENCOVSKY, R. Predicting performance of soybean populations using genetic distances estimated with RAPD markers. **Genetics and Molecular Biology**, Ribeirão Preto, v. 26, p. 43-348, 2003.

BARROSO, P.A.V.; HOFFMANN, V.L. **Métodos de predição do comportamento de populações em melhoramentos**. Campina Grande: Embrapa Algodão, 2003. 38p.

BERNARDO, R. Relationship between single-cross performance and molecular marker heterozygosity. **Theoretical and Applied Genetics**, Berlin, v. 83, p. 628-634, 1992.

BONATO, A.L.V.; CALVO, E.S.; GERALDI I.O.; ARIAS C.A.A. Genetic similarity among soybean (*Glycine max* (L) Merrill) cultivars released in Brazil, using AFLP markers. **Genetics and Molecular Biology**, Ribeirão Preto, v. 29, p. 692-704, 2006.

BONETTI, L. Cultivares e seu melhoramento genético. In: VERNETTI, F. J. (Coord.). **Soja Genética e Melhoramento**. Campinas: Fundação Cargill, 1983. v. 2, p. 741-800.

BRASIL. Ministério de Agricultura, Pecuária e Abastecimento. Companhia Nacional de Abastecimento CONAB. Disponível em: >http://www.conab.gov.br/conteudos.php?a=1252&t=2&Pagina_objcmsconteudos=3#A_objcmsconteudos>. Acesso em: 29 ago. 2011.

BURDICK, R. K.; GRAYBILL, F.A. **Confidence Intervals on variance components**. New York: Marcel Dekker, 1992. v. 127, 211 p.

BURKHAMER, R.; LANNING, S.; MARTENS, R.; MARTIN, J.; TALBERT, L. Predicting progeny variance from parental divergence in hard red spring wheat. **Crop Science**, Madison, v. 38, p. 243-248, 1998.

CALIGARI, P.D.S.; POWELL, W.; JINKS, J.L. The use of doubled haploids in barley breeding. 2: An assessment of univariate cross prediction methods. **Heredity**, London, v. 54, p. 353-358, 1985.

CAPELLARI JUNIOR, L.; RODRIGUES, R. R.; SOUZA, V.C. **Botânica sistemática aplicada aos cursos de engenharia agrônômica e engenharia florestal**. Piracicaba: ESALQ, LCB, 2007. 103 p.

CARVALHO, F.A.D. de. **Avaliação da eficiência do teste precoce no melhoramento genético de soja**. 2008. 102p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2008.

CHARCOSSET, A., LEFORT-BUSON, M.; GALLAIS, A. Relationship between heterosis and heterozygosity at marker loci, a theoretical computation. **Theoretical and Applied Genetics**, Berlin, v. 81, p. 571-575. 1991.

CHEN, Y.W.; NELSON, R.L. Genetic variation and relationships among cultivated, wild, and semiwild soybean. **Crop Science**, Madison, v. 44, p. 316–325, 2004.

_____. Relationship between Origin and Genetic Diversity in Chinese Soybean Germplasm. **Crop Science**, Madison, v. 45, n. 4, p.1645-1652, 2005.

COCKERHAM, C.C. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. **Genetics**, Austin, v. 39, p. 859-882, 1954.

_____. Estimation of genetic variances. In: HANSON, W.D.; ROBINSON, H.F. (Ed.). **Statistical genetics and plant breeding**. Washington: National Academy of Science, 1963. p.53-93.

COLOMBARI FILHO, J.M. **Seleção de genitores para cruzamentos com base em distâncias genéticas moleculares e perspectivas para o melhoramento de soja**. 2009. 98 p. Tese (Doutorado em Genética e Melhoramento de Plantas). - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2009.

COWEN, N.M.; FREY, K.J. Relationships between genealogical distance and breeding behaviour in oats (*Avena sativa* L.) **Euphytica**, Wageningen, 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**, Madison, v. 25, p. 529-532, 1985.

CRUZ, C. D; VENKOVSKY, R. Comparação de alguns métodos de análise dialélica. **Revista Brasileira de Genética**, Ribeirão Preto, v. 12, n. 2, p. 425-438, 1989.

DELANNAY, X.; RODGERSDMAN, D.; PALMER, R.G. Relative genetic contribution among ancestral lines to North American soybean cultivars. **Crop Science**, Madison, v. 23, p. 944-949, 1983.

DIAS, L.; PICOLI, E.; ROCHA, R.; ALFENAS, A. A. Priori choice of hybrid parents in plants. **Genetics and Molecular Research**, Ribeirão Preto, v. 3, n. 3, p.356-368, 2004.

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**, Berlin, v. 95, p. 723-733, 1997.

DOMIT, L.; PÍPOLO, A.; MIRANDA, L.; GUIMARÃES, M. Transferência de tecnologia para cultivares de soja desenvolvidas pela Embrapa soja para o Paraná. **Revista Brasileira de Sementes**, Brasília, v. 29, n. 2, p. 1-9, 2007.

DOYLE, J.J.; DOYLE, J.L.; RAUSCHER, J.T.; BROWN, A.H.D. Diploid and Polyploid Reticulate Evolution Throughout the History of the Perennial Soybeans (*Glycine* Subgenus *Glycine*). **New Phytologist**, Cambridge, v. 161, n. 1, p. 121 -132, 2004.

EMBRAPA. **Embrapa sistema de produção 1, tecnologias de produção de soja Região Central do Brasil**. Londrina: EMBRAPA-CNPSo, 2004a. Disponível em: <<http://www.cnpso.embrapa.br/producaosoja/SojanoBrasil.htm>>. Acesso em: 15 jan. 2011.

EMBRAPA. **Embrapa sistema de produção 1, tecnologias de produção de soja Região Central do Brasil**. Londrina: EMBRAPA-CNPSo, 2004b. Disponível em: <http://www.cnpso.embrapa.br/producaosoja/tab/tab_5_2.htm>. Acesso em: 15 jan. 2011.

EMBRAPA. **Plataforma Nacional de Recursos Genéticos, PA15 - Banco Ativo de Germoplasma de Soja**. Brasília: EMBRAPA. Disponível em: <<http://plataformargen.cenargen.embrapa.br/pnrg/rede-vegetal/projetos-componentes/pc3-bancos-ativos-de-germoplasma-de-especies-leguminosas-oleaginosas-e-fibrosas/planos-de-acao/pa15-banco-ativo-de-germoplasma-de-soja>>. Acesso em: 14 nov. 2010.

FABRIZIUS, M.A.; BUSCH, R.H.; KHAN, K.; HUCKLE, L. Genetic diversity and heterosis of spring wheat crosses. **Crop Science**, Madison, v.38, p.1108-1112, 1998.

FALCONER, D.S. **Introdução à genética quantitativa**. Viçosa: UFV, 1987. 279 p.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to Quantitative Genetics**. 4th ed. London: Longman, 1996. 464 p.

FARIAS, G.J. **Avaliação de progênes F_{2:4} de uma população de soja e perspectivas de melhoramento**. 2008. 61 p. Tese (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2008.

FREIRE FILHO, F.R. **Análise genética de um dialelo entre genótipos precoces de soja (*Glycine max* (L.) Merrill)**. 1988. 224p. Tese (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1988.

FUKUDA, Y. Cytogenetical studies on the wild and cultivated Manchurian soybeans (*Glycine* L.). **Japanese journal of botany**, Tokyo, v.6, p. 489-506, 1933.

GARDNER, C.O. Estimates of genetic parameters in cross fertilizing plants and their implications in plant breeding. In: HANSON, W.D.; ROBINSON, H.F. (Eds). **Statistical genetics and plant breeding**. Washington: National Academy of Science, 1963. p. 225-252.

GARDNER, C.O.; EBERHART, S.A. Analysis and interpretation of the variety cross diallel and related populations. **Biometrics**, Washington, v. 22, p. 439-452, 1966.

GILL, L.H.R.; PECINA, Q.V.; MALDONADO, M.N.; HERNÁNDEZ, D.S.; MAYEK, P.N. Caracterización morfoagronómica y genética de germoplasma mejorado de soja. **Agricultura técnica en México**, México DF, v. 32, n. 3, p 281-294, 2006.

GIZLICE, Z.; CARTER, T.E.; BURTON, J.W Genetic diversity in North American soybean: Multivariate analysis of founding stock and relation to coefficient of parentage. **Crop Science**, Madison, v. 33, p. 614-620, 1993a.

_____. Genetic diversity in North American soybean: II. Prediction of heterosis in F2 populations of southern founding stock using genetic similarity measures. **Crop Science**, Madison, v.33, p.620-626, 1993b.

_____. Genetic base for the North American public soybean cultivars released between 1947 and 1988. **Crop Science**, Madison, v. 34, p. 1143–1151, 1994.

GIZLICE Z.; CARTER JR T.; GERIG T.; BURTON J. Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. **Crop Science**, Madison, v. 36, p. 753-765, 1996.

GRIFFING, J.B. Concept of general and specific combining ability in relation to diallel systems. **Australian Journal of Biological Science**, Melbourne, v. 9, p. 463-493, 1956.

GUMBER, R.K.; SCHILL, B.; LINK, W.; MELCHINGER, A.E. Mean, genetic variance, and usefulness of selfing progenies from intra- and inter- pool crosses in faba beans, and their prediction from parental parameters. **Theoretical and Applied Genetics**, Berlin, v. 98, p. 569-580, 1999.

GUZMAN, P.; DIERS, B.; NEECE, D.; SAINT MARTIN, S.K.; LEROY, A.; GRAU, C.; HUGHES, T.; NELSON, R. QTL Associated with Yield in Three Backcross-Derived Populations of Soybean. **Crop Science**, Madison, v. 47. p. 111-122, 2007.

HALLAUER, A.R.; MIRANDA FILHO, J.B. **Quantitative genetics in maize breeding**. 2nd ed. Ames: Iowa State University Press, 1988. 468 p.

HELMS, T.; VALLAD, G.; McCLEAN, P.; ORF, J. Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. **Theoretical and Applied Genetics**, Berlin, 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**, Ribeirão Preto, v. 9, p. 295-306, 1986.

HYMOWITZ, T. On the domestication of the soybean. **Economic Botany**, Bronx, v. 23, p. 408-421, 1970.

HYTEN, D.; SONG, Q.; ZHU, Y.; CHOI, I.; NELSON, R.; COSTA, J.; SPECHT, J.; SHOEMAKER, R.; CREGAN, P. Impact of genetic bottlenecks on soybean genome diversity. **Proceedings of the National Academy of Sciences of the USA**, Washington, v. 103, p. 16666-16671, 2006.

ININDA, J.; FEHR, W.R.; SILVIA, R.; SCHNEBLY, S.R. Genetic gain in soybean populations with different percentages of plant introduction parentage. **Crop Science**, Madison, v. 36, p. 1470-1472, 1996.

INSTITUTO INTERAMERICANO DE COOPERACIÓN PARA LA AGRICULTURA (IICA). Manejo del cultivo, control de plagas y enfermedades de la soja. In: SEMINÁRIO SOBRE MANEJO DAS CULTURAS DE SOJA E GIRASSOL, 21.; REUNIÃO SOBRE O CONTROLE DE ENFERMIDADES DE SOJA E GIRASSOL, 1987, Montevideo. **Diálogo...** Montevideo: IICA/BID/PROCISUR, 1987. 179 p. Disponível em: <http://books.google.co.ve/books?id=0ukNAQAAIAAJ&pg=PA74&lpg=PA74&dq=FT+-+Pesquisa+e+Sementes,+FT+10&source=bl&ots=KCm9-yGARr&sig=6J87xZfxVOLSLXcr-RNI4CE1yrg&hl=es-419&ei=e8zzTZDRBoXEgAf_2aTKCw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBcQ6AEwAA#v=onepage&q=FT%20-%20Pesquisa%20e%20Sementes%2C%20FT%2010&f=false>. Acesso em: 09 jul. 2010.

JINKS, J.L.; HAYMAN, B.I. The analysis of diallel crosses. **Maize Genetics News Letter**, Columbia, v. 27, p. 48-54, 1953.

JINKS, J.L.; POONI, H.S. Predicting the properties of recombinant inbred lines derived by single seed descent. **Heredity**, London, v. 36, p. 253-266, 1976.

_____. Comparing predictions of mean performance and environmental sensitivity of recombinant inbred lines based upon F_3 and triple test cross families. **Heredity**, London, v. 45, p. 305-312, 1980.

KEARSEY, M.J.; POONI, H.S. **The genetical analysis of quantitative traits**. London: Chapman & Hall, 1996. 381 p.

KIHARA, H. History of biology and other sciences in Japan in retrospect. In: Proceeding of the International Congress of Genetics, 12, 1968, Tokyo. **International Congress of Genetics...** Tokyo: Science Council of Japan, 1969. p. 49-70.

KISHA, T.J.; SNELLER, C.H.; DIERS, B.W. Relationship between genetic distance among parents and genetic variance in populations of soybean. **Crop Science**, Madison, 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**, Madison, v. 38, p. 1669-1680, 1998.

LOURENÇÃO, A.L.; BRAGA, N.R.; MIRANDA M.C. de; RECO, P.C., FUGI C.G.; PEREIRA, J.C. avaliação de danos de insetos e de severidade de oídio em genótipos de soja . **Bragantia**, Campinas, v. 64, n.3, p. 423-433, 2005.

LU, S.L. Discuss about origin of cultivated soybean in China. **China Agricultural Science**, Beijing, v. 4, p. 90-94, 1978.

MALÉCOT, G. **Les mathématiques de l'hérédité**. París: Masson, 1948. 63 p.

MANJARREZ-SANDOVAL, P. CARTER, T.E.; WEBB, D.M.; BURTON, J.W. Heterosis in soybean and its prediction by genetic similarity measures. **Crop Science**, Madison, v. 37, p. 1443-1452, 1997a.

_____. RFLP Genetic similarity and coefficient of parentage as genetic variance predictors for soybean yield . **Crop Science**, Madison, v. 37, p. 698-703, 1997b.

MATHER, K.; JINKS, J.L. **Biometrical genetics, the study of continuous variation**, 3rd ed. London: New York Chapman and Hall, 1982. p. 396.

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**, Berlin, v. 93, p. 392-401, 1996.

MELCHINGER, A.E.; BOPPENMAIER, J.; DHILLON, B.S.; POLLMER, W.G.; HERRMANN, R.G. Genetic diversity for RFLPs in European maize inbreds: II. Relation to performance of hybrids within versus heterotic groups for forage traits **Theoretical and Applied Genetics**, Berlin, v. 84, p.672-681, 1992.

MELCHINGER, A.E.; GUMBER, R.K.; LEIPERT, R.B.; VUYLSTEKE, M.; KUIPER, M. Prediction of testcross means and variances among F_3 progenies of F_1 crosses from testcross means and genetic distances of their parents in maize. **Theoretical and Applied Genetics**, Berlin, v. 96, p. 503 - 512, 1998.

MELCHINGER, A.E.; LEE, M.; LAMKEY, K.R.; WOODMAN, W.L. Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. **Crop Science**, Madison, v. 30, p.1033-1040, 1990a.

MELCHINGER, A.E.; LEE, M.; LAMKEY, K.R.; HALLAUER, A.R.; WOODMAN, W.L. Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. **Theoretical and Applied Genetics**, Berlin, v. 80, p.488-496, 1990b.

MIGLIACCI, P. **Cientistas conseguem sequenciar genoma da soja.** Porto Alegre: Terra Networks Brasil S.A, 2008. Disponível em: <<http://noticias.terra.com.br/ciencia/interna/0,,OI3416136-EI1434,00.html>>. Acesso em: 04 jun. 2010.

MIRANDA, G. V.; SEDIYAMA, C. S.; REIS, M. S.; CRUZ, C. D. Genetic Diversity among Elite Brazilian Soybean Cultivars with Narrow Genetic Base. **Crop Breeding and Applied Biotechnology**, Londrina, v. 1, n. 2, p. 115-123, 2001.

MIRANDA, S.Z.F.; ARIAS, A.C.A.; PRETE, C.C.E.; SOUZA K.R.A. de; ALMEIDA, A.L. de; TOLEDO, F.J.F. de; DESTRO, D. Genetic characterization of ninety elite soybean cultivars using coefficient of parentage. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, p. 363-396, 2007.

MOSER, H.; LEE, M. RFLP variation and genealogical distance, multivariate distance, heterosis and genetic variance in oats. **Theoretical and Applied Genetics**, Berlin, v. 87, n. 8, p. 947-956, 1994.

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**, Washington, v. 76, p. 5269-5273, 1979.

PALUDZYSZYN FILHO, E.; SOUZA K.R.A. de; ALMEIDA, A.L. de. Desenvolvimento de cultivares de soja na região Norte e Nordeste do Brasil. In: ARANTES, N.E.; SOUZA, P.I. de (Ed.). **Cultura da soja nos cerrados**. Piracicaba: Potafos, 1993. p. 255-266.

PELUZIO, J.M.; VAZ-DE-MELO, A.; AFFÉRI, F.S.; SILVA, R.R.; BARROS, H.B.; NASCIMENTO, I.R.; FIDELIS R.R. Variabilidade genética entre cultivares de soja, sob diferentes variações edafoclimáticas, na região centro-sul do estado do Tocantins. **Pesquisa Aplicada & Agrotecnologia**, Guarapuava, v. 2, n. 3, p. 55-64, 2009.

POWELL, W.; CALIGARI, P.D.S.; SWANSTON J.S.; JINKS, J.L. Genetic investigations into B-glucan content in barley. **Theoretical and Applied Genetics**, Berlin, v. 71, p. 461-466, 1985.

PRIOLLI, G.R.; MENDES-JUNIOR, T.C.; ARANTES, N.E.; CONTEL, B.E. Characterization of Brazilian soybean cultivars using microsatellite markers. **Genetics and molecular biology**, Ribeirão Preto, v. 25, p. 185-193, 2002.

PRIOLLI, G.R.; MENDES-JUNIOR, T.C.; SOUSA, B.S.; SOUSA, A.N.; CONTEL, B.E. Diversidade genética da soja entre períodos e entre programas de melhoramento no Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 39, n. 10, p. 967-975, 2004.

PRIOLLI, G.R.; PINHEIRO, J.B.; ZUCCHI, M.I.; BAJAY, M.M.; VELLO, N.A. Genetic Diversity among Brazilian Soybean Cultivars Based on SSR Loci and Pedigree Data. **Brazilian archives of biology and technology**, Curitiba, v. 53, n. 3, p. 519-531, 2010.

PULCINELLI, C.E. **Avaliação de cruzamentos dialélicos de soja em gerações avançadas de endogamia**. 1997. 167 p. Tese (Doutorado em Genética e Melhoramento de Plantas). – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1997.

QUINTERO, V.; HERNÁNDEZ, L.; MORENO, L.; SIMPSON, J.; DE LA VEGA, O.; VEGA, K. Diversidad genética en soja del trópico húmedo de México determinada con marcadores AFLP. **Revista Fitotecnia Mexicana**, Chapingo, v. 28, n. 1, p 63-69, 2005.

RAMALHO, M.A.P.; FERREIRA, F.D.; OLIVEIRA, A.C. de. **Experimentação em genética e melhoramento de plantas**. 2. ed. Lavras: UFLA, 2005. 322p.

RAMALHO, M.A.P.; SANTOS, J.B.; PEREIRA FILHO, I. Choice os parents for dry bean (*Phaseolus vulgaris* L.) breeding: I. Interaction of mean components by generation and by location. **Revista Brasileira de Genética**, Ribeirão Preto, v.11, n.2, p.391-400, 1988.

RAMALHO, M.A.P.; SANTOS, J.B.; ZIMMERMANN, M.J.O. **Genética quantitativa em plantas autógamas: aplicações ao feijoeiro**. Goiânia: UFG, 1993. 271 p.

REIS, F.E. dos; REIS, M.S.; CRUZ, C.D.; SEDIYAMA, T. Comparação de procedimentos de seleção para produção de grãos em populações de soja. **Ciência Rural**, Santa Maria, v.34, n.3, 2004.

RUBIN, S.; SANTOS, P. Progresso do melhoramento genético da soja no Estado do Rio Grande do Sul: I. Rendimento de grãos. **Pesquisa Agropecuária Gaucha**, Porto Alegre, v. 2, p. 139-147, 1996.

SARAWAT, P.; STODDARD, F.L.; MARSHALL, D.R. Derivation of superior F₅ lines from heterotic hybrids in pea. **Euphytica**, Wageningen, v. 73, n. 3, p. 265-272, 1994a.

_____. Genetic distance and its association with heterosis in peas. **Euphytica**, Wageningen, v. 73, n. 3, p. 255-264, 1994b.

SAS INSTITUTE, **SAS system for windows. Version 8.2**. Cary, 2001.

SATTERTHWAITE, F. E. An aproximade distribuition of estimates of variance components. **Biometrics**, Washington, v. 2, n. 6, p. 110-114, 1946.

SCHMUTZ, J.; CANNON, S.; SCHLUETER, J.; MA, J.; MITROS, T.; NELSON, W.; HYTEN, D.; SONG, Q.; THELEN, J.; CHENG, J.; XU, D.; HELLSTEN, U.; MAY, G.; YU, Y.; SAKURAI, T.; UMEZAWA, T.; BHATTACHARYYA, M.; SANDHU, D.; VALLIYODAN, B.; LINDQUIST, E.; PETO, M.; GRANT, D.; SHU, S.; GOODSTEIN, D.; BARRY, K.; FUTRELL-GRIGGS, M.; ABERNATHY, B.; DU, J.; TIAN, Z.; ZHU, L.; GILL, N.; JOSHI, T.; LIBAULT, M.; SETHURAMAN, A.; ZHANG, X.; SHINOZAKI, K.; NGUYEN, H.; WING, R.; CREGAN, P.; SPECHT, J.; GRIMWOOD, J.; ROKHSAR, D.; STACEY, G.; SHOEMAKER, R.; JACKSON, S. Genome sequence of the palaeopolyploid soybean. **Nature**, London, v. 463, p. 178-183, 2010.

SCHOENER, C.S.; FEHR, W.R. Utilization of plant introductions in soybean breeding populations. **Crop Science**, Madison, v. 19, p. 185-188, 1979.

SILVA, V.S. **Seleção de pré-cultivares de soja baseada em índices**. 2005. 104 p. Tese (Doutorado em Genética e Melhoramento de Plantas)-Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2005.

SOUZA JUNIOR, C. L. de. **Componentes da variância genética e suas implicações no melhoramento vegetal**. Piracicaba: FEALQ, 1989. 134 p.

SOUZA, E.; SORRELLS, M.E. Prediction of progeny variation in oat from parental genetic relationships . **Theoretical and Applied Genetics**, Berlin, v. 82, p. 233-241, 1991.

STEEL, R.G.D.; TORRIE, J.H. **Principles and procedures of statistics: a Biometrical Approach**. 3rd ed. New York: McGraw-Hill, 1997. 666 p.

STHUTHMAN, D.D.; STUCKER, R.E. Combining ability analysis of near homozygous lines derived from a 12-parent diallel cross in oats. **Crop Science**, Madison, v. 15, n. 6, p. 800–803, 1975.

STUBER, C.W. Molecular marker in the manipulation of quantitative characters. In: BROWN, A.; CLEGG, M.; KANKR, A.; WEIR, B (Ed.). **Plant Population Genetics, Breeding and Genetic Resource**. Sunderland: Sinauer Associate, 1989. p. 334-350.

STUBER, C.W; EDWARDS, M.D; WENDEL, J.F. Molecular-marker facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. **Crop Science**, Madison, v. 27, n. 4, p. 639-648, 1987.

TAVAUD-PIRRA, M.; SATRE, P.; NELSON, R.; SANTONI, S.; TEXIER, N.; ROUMET, P. Genetic Diversity in a Soybean Collection. **Crop Science**, Madison, v. 49, p. 895-902, 2009

THOMPSON, J.A.; NELSON, R.L. Utilization of diverse germplasm for soybean yield improvement. **Crop Science**, Madison, v. 38, n. 5, p.1362- 1368, 1998.

THORMANN, C.E.; OSBORN, T.C. Use of RAPD and RFLP markers for germplasm evaluation. *In* Applications of RAPD technology to plant breeding. **Crop Science**, Madison, p. 9-11, 1992.

THORNE, J.C.; FEHR, W.R. Exotic germplasm for yield improvement in 2-way and 3-way soybean crosses. **Crop Science**, Madison, v. 10, p. 677-678, 1970.

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

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

TRILLER, C.; TOLEDO, J.F.F. Using the F₃ generation for predicting the breeding potential of soybean. **Revista Brasileira de Genética**, Ribeirão Preto, v.19, p.289-294, 1996.

VELLO, N.A.; FEHR, W.R.; BAHRENFUS, J.B. Genetic variability and agronomic performance of soybean populations developed from plant introductions. **Crop Science**, Madison, v. 24, p. 511-514, 1984.

VELLO, N.A.; HIROMOTO, D.M.; AZEVEDO-FILHO, A.J. Coefficient of parentage and breeding of Brazilian soybean germplasm. **Brazilian journal of genetics**, Ribeirão Preto, v. 11, p. 679-697, 1988.

VIEIRA, E.; SCHUSTER, I.; SILVA, R. da; OLIVEIRA, M. Variabilidade genética em cultivares de soja determinada com marcadores microssatélites em gel de agarose **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 11, p.1460-1466, 2009.

WALLING, J.G.; PIRES, J.C.; JACKSON, S.A. Preparation of samples for comparative studies of plant chromosomes using in situ hybridization methods. **Methods in Enzymology**, New York, v. 395, p. 443–460, 2005.

WANG, L.Z. Origin and diffusion of soybean. **Soybean science**, Beijing v. 4, p. 1-7, 1985.

XU, B.. Three New evidences for origin of soybean. **Soybean science**, Beijing v.5, p.123-130, 1986.

YAMANAKA, N.; SATO, H.; YANG, Z.; XU, D.H.; CATELLI, L.L.; BINNECK, E.; ARIAS, C.A.A.; ABDELNOOR, R.V. Genetic relationships between Chinese, Japanese and Brazilian soybean gene pools revealed by simple sequence repeat (SSR) markers, **Genetics and Molecular Biology**, Ribeirão Preto, n. 30, v. 1, p. 85-88, 2007.

YOON, M.; LEE, J.; KIM, C.; KANG, J.; CHO, E.; BAEK, H. DNA profiling and genetic diversity of Korean soybean (*Glycine max* (L.) Merrill) landraces by SSR markers. **Euphytica**, Wageningen, v. 165, n. 1, p. 69–77, 2009.