

**INCORPORAÇÃO DA INFORMAÇÃO DE
PARENTESCO NO MÉTODO
GENEALÓGICO PELO ENFOQUE DE
MODELOS MISTOS**

JOSÉ AIRTON RODRIGUES NUNES

2006

JOSÉ AIRTON RODRIGUES NUNES

**INCORPORAÇÃO DA INFORMAÇÃO DE PARENTESCO
NO MÉTODO GENEALÓGICO PELO ENFOQUE DE
MODELOS MISTOS**

Tese apresentada à Universidade Federal de
Lavras como parte das exigências do Programa
de Pós-graduação em Agronomia, área de
concentração em Genética e Melhoramento de
Plantas, para a obtenção do título de "Doutor".

Orientador
Prof. Dr. Magno Antônio Patto Ramalho

LAVRAS
MINAS GERAIS - BRASIL
2006

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Nunes, José Airton Rodrigues

Incorporação da informação de parentesco no método genealógico pelo enfoque de modelos mistos / José Airton Rodrigues Nunes. – Lavras : UFLA, 2006.

113 p. : il.

Orientador: Magno Antônio Patto Ramalho.

Tese (Doutorado) – UFLA.

Bibliografia.

1. Pedigree. 2. BLUP. 3. Melhoramento genético vegetal. 4. Simulação. I. Universidade Federal de Lavras. II. Título.

CDD-631.523

JOSÉ AIRTON RODRIGUES NUNES

**INCORPORAÇÃO DA INFORMAÇÃO DE PARENTESCO
NO MÉTODO GENEALÓGICO PELO ENFOQUE DE
MODELOS MISTOS**

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-graduação em Agronomia, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de "Doutor".

APROVADA em 10 de fevereiro de 2006

Prof. Dr. Daniel Furtado Ferreira	UFLA
Dr. Francisco Rodrigues Freire Filho	EMBRAPA
Prof. Dr. Wagner Pereira Reis	UFLA
Prof. Dr. Júlio Sílvio de Sousa Bueno Filho	UFLA

Prof. Dr. Magno Antonio Patto Ramalho
UFLA
(Orientador)

LAVRAS
MINAS GERAIS – BRASIL
2006

A Deus,

*infinita bondade, misericórdia e amor, por me proporcionar
mais esta conquista na minha vida.*

OFEREÇO

*Aos meus maravilhosos pais, **Manoel Raimundo Nunes Rodrigues**
(in memorian) e **Maria Dedícia Rodrigues**, que sempre
estiveram ao meu lado com muito amor, confiança e
repassando essenciais ensinamentos de vida.*

*Ao meu querido avô, **João Rodrigues de Barros** (in
memorian) e à minha adorável avó, **Josefa Maria
da Conceição**, que sempre me incentivaram com
muito carinho, bondade e sabedoria.*

DEDICO

AGRADECIMENTOS

À Universidade Federal de Lavras, pela oportunidade de concluir o doutorado e ao Departamento de Biologia e Programa de Pós-graduação em Agronomia, Genética e Melhoramento de Plantas, pelo apoio irrestrito na condução e desenvolvimento da minha pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de pesquisa.

À Universidade Federal do Piauí, especialmente ao Centro de Ciências Agrárias desta IES, em nome de seu Diretor João Batista Lopes e ao Departamento de Planejamento e Política Agrícola, em nome de seu Chefe Fábio Coelho Gomes Nóbrega, pelo apoio e confiança ao permitirem meu afastamento num momento decisivo para a conclusão desta tese.

Ao Prof. Magno Antonio Patto Ramalho, exemplo de competência e dedicação como professor e pesquisador de uma universidade pública federal, pela excelência em orientação em todas as fases de elaboração desta tese, valiosos ensinamentos e, sobretudo, pela amizade e convívio.

Ao Prof. Eduardo Bearzoti, originalmente o mentor do projeto desta tese, pelas orientações fundamentais e ensinamentos, que muito contribuíram para a conclusão deste trabalho.

Ao Prof. Daniel Furtado Ferreira que prontamente se disponibilizou para a orientação deste trabalho, pelas valiosas sugestões, excelente convívio e amizade e à Dra. Ângela de Fátima Barbosa Abreu, pela imensa ajuda, apoio e gentileza durante toda a fase experimental desta tese.

Ao Dr. Marcos Deon Vilela de Resende e ao Prof. Júlio Sílvio de Sousa Bueno Filho, pelas importantes sugestões e críticas para o engrandecimento e elaboração deste trabalho.

Aos professores do Departamento de Biologia, Magno Ramalho, César Brasil, João Bosco, Elaine Aparecida, Lisete e Giovana, pelos ensinamentos fundamentais para minha formação profissional e humana.

Aos funcionários do Departamento de Biologia, Erondina, Léo, Lindolfo, Magda, Rafaela e Zélia, pela simpatia, ajuda e eficiência com que nos atenderam, e em especial à Elaine, pela gentileza e amizade.

Aos colegas do curso de Doutorado e do GEN, pelo ótimo convívio durante todo o curso. Também aos amigos, João Luis “Traquino”, Eduardo Lambert “Dudu”, Wilacildo “Wilão”, Kaesel, Mauroisrael, Regina, Alberto, Lidiana, José Waldemar e Elizângela, pela amizade e incentivo.

Ao professor Raimundo de Pontes Nunes, pelos ensinamentos fundamentais para a atividade científica, incentivo e amizade.

À minha noiva Andreia “Benzinho” que, em todos os momentos, esteve ao meu lado e com a qual divido todo o mérito desta conquista, pela cumplicidade, confiança e amor. Também a seus pais, Sr. Célio Soares Ferreira e sua esposa Sra. Célia Maria Nunes Ferreira, pelos quais tenho enorme respeito e admiração, pelo apoio e carinho.

Aos grandes amigos, Flávio, Carminha, Isaías e Luciana, pela sincera torcida e carinho e, em especial à Larissa, que com seu sorriso espontâneo nos trouxe muita inspiração e alegria.

Aos meus queridos avós, José Rodrigues e Júlia, pelos ensinamentos de vida, carinho e afeto e aos meus irmãos, Denise, Francisco, João e Luiz, os quais são para mim motivo de muita felicidade, pela motivação, união e carinho;

À minha sobrinha, Letícia, pela qual tenho enorme carinho, pela felicidade que nos proporciona e aos demais familiares, pelo incentivo.

SUMÁRIO

	Página
RESUMO.....	i
ABSTRACT	iii
1 Introdução	1
2 Referencial teórico.....	3
2.1 Métodos de melhoramento aplicáveis às plantas autógamas.....	3
2.1.1 Autogamia.....	3
2.1.2 Estrutura genética de populações em plantas autógamas	4
2.1.3 Métodos de condução de populações segregantes em plantas autógamas	8
2.1.3.1 Método genealógico ou <i>pedigree</i>	12
2.1.3.2 Método populacional ou <i>bulk</i>	18
2.1.3.3 Comparação de métodos de condução em plantas autógamas	22
2.2 Simulação computacional para comparação de métodos de condução em plantas autógamas	25
2.3 Emprego do <i>BLUP</i> no melhoramento de plantas autógamas.....	31
3 Material e métodos	45
3.1 Simulação computacional.....	45
3.1.1 Modelo genético e configurações estudadas	45
3.1.2 Avanço das populações segregantes e geração dos genótipos.....	45
3.1.2.1 Método genealógico ou <i>pedigree</i>	46
3.1.2.2 Método <i>bulk</i>	46
3.1.3 Geração dos fenótipos	47
3.1.4 Análise estatística	49
3.2 Experimentos de campo.....	53
3.2.1 Local	53
3.2.2 Cultivares.....	54

3.2.3 Obtenção da população segregante e avanço das gerações pelos métodos genealógico e <i>bulk</i>	54
3.2.4 Avaliação das progênies	55
3.2.4.1 Delineamento experimental e condução dos experimentos.....	55
3.2.4.2 Características avaliadas.....	55
3.2.4.3 Análise estatística	56
3.2.4.4 Estimação de parâmetros genéticos	58
3.2.4.5 Estimação das correlações de Spearman e Pearson, proporções de coincidências e ganhos realizados com a seleção.....	59
4 Resultados e discussão.....	60
4.1 Simulação computacional.....	60
4.2 Experimentos de campo.....	81
5 Conclusões.....	97
Referências bibliográficas	98
ANEXO	112

RESUMO

NUNES, José Airton Rodrigues. **Incorporação da informação de parentesco no método genealógico pelo enfoque de modelos mistos.** LAVRAS: UFLA, 2006. 113 p. (Tese – Doutorado em Agronomia/Genética e Melhoramento de Plantas)*

No método genealógico é, normalmente, despendido um trabalho adicional para a anotação da genealogia ao longo das gerações, à qual tem sido atribuída pouca utilidade. Com isso, o objetivo deste estudo foi avaliar a eficiência do método *pedigree* pela inclusão da informação de parentesco entre as progênes a partir da genealogia e, ainda, compará-lo com o método *bulk* via simulação computacional e em experimentos de campo. Nas simulações, assumiu-se um modelo genético de 20 locos segregantes, não ligados, de efeitos iguais e aditivos e com frequência alélica 0,5. Para as configurações avaliadas, foram considerados valores de herdabilidade no sentido amplo na média das progênes de 10%, 25%, 50% e 75% e grau médio de dominância de 0,0, 0,5 e 1,0. Foram simulados 1.000 experimentos, para cada configuração e método de condução, referentes à avaliação de 256 progênes $F_{4;5}$. Nos experimentos de campo, com a cultura do feijoeiro, foram avaliadas 256 progênes $F_{4;5}$ na safra ‘das águas’ 2004/05 e as correspondentes 256 progênes $F_{4;6}$ na safra ‘das secas’ 2005, sendo estas oriundas dos métodos genealógico e *bulk*. A população segregante foi obtida a partir do cruzamento entre os genitores BRS Talismã e BRS Valente. As análises dos dados simulados e de campo foram realizadas pela abordagem de modelos mistos. Os dados das progênes oriundas do método *pedigree* foram analisados de acordo com dois modelos: ignorando e considerando o parentesco genético aditivo entre as progênes. Para o método *bulk*, a análise foi realizada considerando as progênes como não relacionadas. Os critérios de seleção avaliados foram a média das progênes e o melhor preditor linear não tendencioso (*BLUP*) com a inclusão do parentesco, denotado por $BLUP_A$ e na ausência de parentesco, denotado por *BLUP*. A dominância teve pouca influência sobre os resultados das simulações para ambos os métodos de condução. O $BLUP_A$ foi mais eficiente que a média em todas as configurações, resultando no melhor ranqueamento das progênes. Nos experimentos de campo, foi evidenciado que a variação genética entre as progênes foi significativamente não nula para a produção de grãos e nota de porte, em ambos os métodos de

* Comitê Orientador: Magno Antônio Patto Ramalho - UFLA (Orientador), Daniel Furtado Ferreira - UFLA.

condução. Pode-se inferir que, utilizando-se o método *pedigree*, considerando a genealogia entre as progênes por meio do procedimento $BLUP_A$, resulta em ganhos seletivos superiores aos alcançados quando esta informação é ignorada, de forma a compensar o trabalho despendido na obtenção desses registros, sobretudo, para caracteres de herdabilidade baixa. Os métodos genealógico e *bulk*, nas condições simuladas, mostraram-se igualmente eficientes para a obtenção de genótipos superiores. Contudo, quando a informação da genealogia foi considerada, o método *pedigree* apresentou ganhos ligeiramente superiores.

ABSTRACT

NUNES, José Airton Rodrigues. **Using relationship information from pedigree method on mixed model estimation.** LAVRAS: UFLA, 2006. 113 p. (Thesis - Doctorate in Agronomy/Genetics and Plant Breeding)*

In the pedigree method it is usually spent an additional effort for recording the genealogies along the generations, that have been attributed little usefulness. This study aimed to evaluate the efficiency of the pedigree method considering the use of relationship among progenies and to compare its performance with the bulk method by computer simulation and in field experiments. It was assumed a genetic model with 20 segregating loci with allelic frequency equal to 0.5. All loci were assumed to have equal effect, no epistasis and no linkage. For the appraised configurations it was considered values of broad sense heritability at the progenies level of 10%, 25%, 50% and 75% and average level of dominance of 0.0, 0.5 and 1.0. It was evaluated 256 $F_{4:5}$ progenies in 1000 simulated experiments for each configuration and selection method. In the field experiments 256 $F_{4:5}$ progenies (in the rainy season of 2004/05) and 256 $F_{4:6}$ (in the dry season of 2005) from the pedigree and bulk methods were evaluated. The common bean segregating population was obtained from the crossing between parents BRS Talismã and BRS Valente. The analyses of the simulated and field data were accomplished by the mixed model approach. The pedigree's data were analyzed according of two models: ignoring and considering the additive genetic relationship among progenies. For the bulk method the analysis was accomplished considering the progenies as not related. The appraised selection criteria were the progenies average and the best linear unbiased prediction (*BLUP*) considering the relationship, denoted by $BLUP_A$ and in the relationship absence, denoted by *BLUP*. Dominance had little influence on the simulation results for both selection methods. The $BLUP_A$ was more efficient than the average in all studied configurations, resulting in the best ranking of the progenies. For the field experiments there was significantly no null genetic variation among the progenies for grain yield and plant stature in both methods of conduction of the segregating populations. It can be inferred that the pedigree selection take account the genetic similarity among progenies ($BLUP_A$) had selections gain greater than the same method ignoring this pedigree information (*BLUP*). This advantage may compensate the additional effort spent obtaining

* Guidance Committee: Magno Antônio Patto Ramalho - UFLA (Major Professor), Daniel Furtado Ferreira.

pedigree records, mainly for traits of low heritability. The pedigree and bulk methods, in the simulated conditions, were equally efficient for obtaining superior genotypes, however when the progenies relationship was considered the pedigree method was slightly superior to bulk method.

1 INTRODUÇÃO

O progresso genético nos programas de melhoramento de plantas, especialmente em autógamias, está intrinsecamente ligado à adoção de métodos de condução das populações segregantes que permitam manipular adequadamente a variabilidade genética liberada ao longo das gerações de avanço e, sobretudo, à utilização de eficientes procedimentos de seleção que possibilitem explorar ao máximo as informações disponíveis por ocasião da avaliação das progênies.

Dentre os métodos de condução das populações segregantes existentes, destaca-se, pelo freqüente uso nos programas de melhoramento, o método genealógico ou *pedigree*. Uma particularidade deste método, refere-se ao trabalho adicional normalmente despendido para as anotações dos registros genealógicos das progênies ao longo das gerações de autofecundação. Contudo, a esta informação da genealogia tem sido atribuída pouca utilidade para fins seletivos.

Nos vários estudos de comparação de métodos de condução já realizados, tanto por meio de experimentação convencional quanto via simulação computacional, o sistema de registros advindo da utilização do método genealógico foi sempre desconsiderado no momento da avaliação das progênies (Casali & Tiglechar, 1975; Raposo, 1999). Vale comentar que resultados provenientes de experimentos de campo são de difícil extrapolação devido, principalmente, à insuficiente amostragem. Assim, o uso de simulação computacional apresenta-se como ferramenta valiosa que pode contribuir efetivamente para a otimização dos programas de melhoramento (Ferreira, 2001; Fouilloux & Bannerot, 1988).

O objetivo dos melhoristas de plantas autógamas, na avaliação das progênies em experimentos com repetição, é selecionar aquelas que associem os melhores valores genéticos aditivos (*VGA*). Com isso, a eficiência da seleção, admitindo uma adequada experimentação, depende diretamente do emprego de métodos estatísticos que aproveitem ao máximo as informações disponíveis. Vários métodos têm sido propostos no contexto de modelos fixos e mistos, como o método dos quadrados mínimos e métodos baseados na verossimilhança, resultando em diferentes funções dos valores fenotípicos com a finalidade de prever os *VGA* das progênies, como, por exemplo, o *BLUE* (melhor estimador linear não tendencioso), o *BLP* (melhor preditor linear) e, mais recentemente, o *BLUP* (melhor preditor linear não tendencioso) (White & Hodge, 1989).

O enfoque de análise por modelos mistos, conforme proposto por Henderson em 1949, em que os *VGA* preditos correspondem a predições *BLUP*, tem a vantagem de permitir a inclusão da informação da similaridade genética entre as progênies, como acontece com o uso do método genealógico, diretamente no método de análise.

Diante do exposto, o presente trabalho teve por objetivo avaliar a eficiência da seleção mediante incorporação da informação de parentesco do método genealógico, determinada pela genealogia entre as progênies, pelo enfoque de modelos mistos e, adicionalmente, compará-lo ao método *bulk*, por meio de simulação computacional e em condições de campo.

2 REFERENCIAL TEÓRICO

2.1 Métodos de melhoramento aplicáveis às plantas autógamias

2.1.1 Autogamia

As espécies vegetais para as quais predominam a autofecundação são denominadas de autógamias. Inúmeras espécies cultivadas de importância econômica pertencem a este grupo: cereais, como o trigo, cevada, aveia e arroz; leguminosas, como a soja, ervilha, amendoim, feijoeiro comum e feijão caupi; olerícolas, como o tomate, quiabo, alface e beringela e perenes, como café arábica, pêsego e nectarina (Borém & Miranda, 2005 e Ramalho et al., 2001).

As espécies autógamias caracterizam-se por apresentarem mecanismos morfológicos e ou fisiológicos que favorecem ou conduzem à formação de sementes a partir da união de gametas oriundos da mesma planta (Fehr, 1987). Segundo Allard (1999), nas espécies silvestres predominam mecanismos que evitam a autogamia, o que sugere que as espécies autógamias tenham se originado a partir de espécies alógamas. Por exemplo, a soja cultivada (*Glycine max*) apresenta taxa de alogamia de 1%, enquanto que no seu ancestral selvagem (*Glycine soja*) esta taxa alcança 13% (Fujita et al., 1997).

Em geral, a taxa de fecundação cruzada natural em plantas autógamias é inferior a 5% (Allard, 1999). Contudo, em várias plantas autógamias, é relatada a ocorrência de variações nas quantidades relativas de polinização cruzada dentro da mesma espécie. Para o feijoeiro comum (*Phaseolus vulgaris* L.), por exemplo, Antunes et al. (1973), citados por Ramalho et al. (1993), encontraram taxas de alogamia de 9% em Pelotas, Rio Grande do Sul, enquanto Marques Junior e Ramalho (1995), em estudo realizado em diferentes épocas de semeadura em Lavras, Minas Gerais, evidenciaram que estas taxas variaram com

a época, sendo, em média, 1,39%. Este fato mostra que a taxa de fecundação cruzada, além de variar com o genótipo, é também dependente das condições ambientais durante o período de polinização e da disponibilidade de populações de insetos polinizadores (Ramalho et al., 1993; Vieira et al., 2005).

A partir do exposto, torna-se desejável que o melhorista determine adequadamente as taxas de alogamia que ocorrem nos genótipos que compõem a base de melhoramento da espécie autógama que está sendo trabalhada, nos seus respectivos ambientes (Allard, 1999). Este conhecimento é fundamental, sobretudo, em atenção às etapas finais de um programa de melhoramento de plantas autógamas em que se almeja manter a pureza genética das linhagens.

2.1.2 Estrutura genética de populações em plantas autógamas

O sistema de reprodução sexual interfere diretamente sobre a estrutura genética das populações, visto que, ao contrário do que ocorre em espécies de propagação vegetativa, os genótipos não são propriamente transmitidos para a próxima geração e, sim, os alelos pelos gametas formados, os quais originarão os genótipos da próxima geração (Wricke & Weber, 1986).

Para as espécies autógamas, conforme mostrado no item 2.1.1, a autofecundação é o sistema de acasalamento predominante. Este sistema é o que mais rapidamente conduz à endogamia (Miranda Filho, 2001).

O efeito da endogamia sobre a estrutura genética das populações constitui-se na principal diferença entre espécies autógamas e alógamas (Allard, 1999). Enquanto as plantas alógamas sofrem severa depressão por endogamia, as espécies autógamas quase não a exibem, devido, ao longo do processo evolutivo, a carga genética encoberta no heterozigoto ter sido gradativamente eliminada. Desse modo, os métodos de melhoramento apropriados para cada uma destas

espécies são bem diferentes. Especificamente para culturas autógamas, estes métodos devem, em geral, direcionarem-se para a condução e a manutenção da homozigose nas progênies e potenciais cultivares (Wricke & Weber, 1986). Porém, em algumas espécies naturalmente autógamas, como o arroz, tem-se evidenciado uma quantidade substancial de vigor híbrido, de forma que alguns programas de melhoramento para esta cultura têm sido direcionados para a obtenção de híbridos, cuja produção de grãos, têm superado às cultivares tradicionais em aproximadamente 20% (Virmani, 1999).

O fluxo alélico em plantas autógamas é reduzido devido às baixas taxas de polinização cruzada. Normalmente, as populações naturais consistem da mistura de muitas linhagens ou linhas puras (Allard, 1999; Wricke & Weber, 1986). Nestas espécies, os genótipos produzem gametas idênticos que se unem originando progênies genotipicamente idênticas (Wricke & Weber, 1986).

Quanto ao padrão de distribuição da variação genética em populações naturais de plantas autógamas, é esperado que a maior diversidade encontre-se entre as populações (Loveless & Hamrick, 1984). A respeito desta proposição, Allard (1988), por meio de resultados obtidos a partir de estudos de microevolução em cevada, comenta: “a autofecundação parece promover o desenvolvimento e manutenção da adaptabilidade dentro de populações e, ao mesmo tempo, facilitar o desenvolvimento de diferenciação espacial pelo retardo do fluxo gênico entre populações”.

Por meio do levantamento de trabalhos realizados durante o período de 1968 a 1988, que se reportavam à diversidade em populações de plantas usando aloenzimas, Hamrick & Goldt (1990) encontraram que, em média, a fração da variação genética total concentrada entre populações, em espécies autógamas, foi cinco vezes superior à encontrada em espécies que têm a polinização cruzada predominante. Estudos posteriores, realizados com populações de parentes selvagens de espécies autógamas cultivadas, têm obtido resultados concordantes.

Por exemplo, no estudo de diversidade e estrutura populacional de quatro populações de arroz selvagem (*Oryza glumaepatula*) da América do Sul utilizando marcadores izoenzimáticos e do tipo *RAPD*, foi constatado que a maior proporção da variação genética foi devido a diferenças entre as populações (Buso et al., 1998).

As populações de melhoramento de plantas autógamas são obtidas, em geral, a partir de hibridações artificiais entre duas linhagens. Deste cruzamento, obtém-se uma geração F_1 , homogênea, de indivíduos altamente heterozigóticos. Os indivíduos F_1 são, em seguida, autofecundados para liberação da variabilidade, obtendo-se a geração F_2 , a qual é heterogênea, constando de indivíduos homozigóticos e heterozigóticos. A população F_2 é considerada a população base em equilíbrio com indivíduos não endogâmicos, isto é, a endogamia somente será levada em conta a partir desta geração (Ramalho et al., 1993; Wricke & Weber, 1986).

Com a autofecundação, a consequência genética imediata é a redução da proporção de heterozigotos em 50% a cada geração, sem, contudo, promover alterações nas frequências alélicas (Bernardo, 2002; Fehr, 1987; Hedrick, 2000). Assumindo que os locos que controlam o caráter de interesse não estejam ligados, a população F_2 pode ser considerada como uma população de acasalamentos aleatórios em equilíbrio. A partir disso, o equilíbrio verdadeiro somente será alcançado ao final de ciclos sucessivos de autofecundação quando todos os genótipos estiverem em completa homozigose (Wricke & Weber, 1986).

Na presença de ligação gênica, a F_2 não está em equilíbrio e a homozigose será atingida mais lentamente com a endogamia, dependendo da intensidade da ligação. Assim, a probabilidade de encontrar os melhores genótipos homozigóticos contendo as combinações de alelos desejados pode ser reduzida (Wricke & Weber, 1986).

Um aspecto interessante com o decorrer das sucessivas gerações de autofecundação é a previsibilidade da constituição genética dessas populações. Tome-se a uma população F_2 heterozigótica para n locos não ligados. As proporções esperadas de homozigotos e heterozigotos para qualquer geração g podem ser expressas pela expansão do binômio de Newton $(a + b)^n$, em que: a é a frequência de locos em heterozigose obtida por $(1/2)^{g-1}$ e b é a frequência de locos em homozigose, correspondente a $1-(1/2)^{g-1}$. Substituindo a e b na expressão anterior e, realizando algumas operações matemáticas, obtém-se a seguinte expressão apresentada por Allard (1999): $[1+(2^{g-1}-1)]^n$. Por exemplo, a composição esperada da população F_6 considerando uma segregação para três locos com distribuição independente, é dada por:

$$\left[1 + \left(2^{6-1} - 1\right)\right]^3 = [1 + 31]^3 = (1)^3 + 3(1)^2(31) + 3(1)(31)^2 + (31)^3.$$

Assim, de um total de 32.768 indivíduos, nesta população são esperados:

1 indivíduo com todos os locos em heterozigose

93 indivíduos com dois locos em heterozigose e um loco em homozigose

2.883 indivíduos com um loco em heterozigose e dois locos em homozigose

29.791 indivíduos com todos os locos em homozigose.

A partir do exemplo, tomando-se a segregação para apenas três locos, observa-se que, na geração F_6 , a proporção de indivíduos homozigóticos é alta (91%). Contudo, apesar de se tratar de população avançada por cinco sucessivas autofecundações, o que geralmente tem sido apregoado para a obtenção de famílias em suposta homozigose, tem-se uma chance de 9% de extrair famílias que ainda estão segregando para pelo menos um loco.

De modo geral, as cultivares, em culturas autógamias, usualmente compreendem uma única linhagem ou uma mistura de várias linhagens fenotipicamente semelhantes, que podem ter sua constituição genética mantida indefinidamente (Fehr, 1987; Ramalho et al., 2001).

2.1.3 Métodos de condução de populações segregantes em plantas autógamas

O melhoramento em espécies autógamas tem como objetivo, em geral, a obtenção de linhagens que detenham alelos favoráveis no maior número de locos (Ramalho et al., 2001). Os métodos de melhoramento aplicados às espécies autógamas podem ser classificados em: introdução de cultivares, seleção de linhas puras e hibridação. A descrição detalhada desses métodos pode ser encontrada em vários livros texto (Borém & Miranda, 2005; Fehr, 1987; Ramalho et al., 2001).

Nesse referencial, o método da hibridação irá receber maior atenção. Esse método tem como finalidade reunir num único indivíduo – linhagem – os alelos desejáveis que se encontram em linhagens distintas (Fehr, 1987; Ramalho, 1997). É constituído basicamente de três etapas: a escolha dos genitores a serem cruzados, a obtenção da população segregante e a escolha do método de condução da população segregante (Ramalho et al., 2001).

A população segregante a partir da qual será realizada a extração das linhagens, como já comentado previamente, é obtida, normalmente, mediante cruzamento entre genitores superiores (Rasmusson & Phillips, 1997). Desse modo, capitalizando-se os efeitos benéficos da seleção anteriormente aplicada. A população base formada tem grande potencial de melhoramento, pois, caracteriza-se por apresentar média e variabilidade elevadas, aumentando-se a probabilidade de selecionar linhagens de performance superior (Fouilloux & Bannerot, 1988).

Após obtida a população base, o próximo passo consiste em proceder o avanço da endogamia até um nível adequado de homozigosidade (Fehr, 1987). Na literatura, são descritos vários métodos de condução da população segregante apropriados para espécies autógamas (Borém & Miranda, 2005; Fehr, 1987; Jensen, 1988; Ramalho et al., 2001). Conforme Fouilloux & Banerot (1988),

esses métodos podem ser classificados em duas categorias: a primeira, em que estão compreendidos os métodos que não separam as fases de endogamia e seleção, na qual estão inclusos a seleção massal e o método genealógico ou *pedigree*, e a segunda, em que constam os métodos que separam as fases de endogamia e seleção, na qual os mais importantes são os métodos do *bulk* ou populacional e descendente de uma única semente, este último mais comumente reportado pela sigla *SSD* (do inglês *single seed descent*).

Além dos métodos mencionados anteriormente, relata-se a existência de vários outros que se constituem em modificações e ou combinações destes. Estes métodos surgiram em virtude das desvantagens e ou limitações apresentadas pelos métodos originais (Bórem & Miranda, 2005; Fehr, 1987; Ramalho et al., 1993; Ramalho et al., 2001). Como exemplo, Wang et al. (2003) citam a evolução das estratégias de melhoramento utilizadas pelo Centro Internacional de Melhoramento do Milho e Trigo (*CIMMYT*) para o trigo, visando melhorar a eficiência do uso de recursos: o método genealógico, durante o período de 1944-1985; o método modificado *bulk-pedigree*, de 1985 até meados de 1998, o qual resultou em muitas das variedades de trigo cultivadas no mundo e, atualmente, o *selected bulk method*.

Em virtude do escopo deste trabalho, no presente referencial somente os métodos genealógico e populacional serão abordados. Maiores detalhes sobre os métodos de avanço da endogamia aplicáveis a espécies autógamas podem ser encontrados nos livros texto mencionados anteriormente.

Independente do método de seleção escolhido, um questionamento constantemente realizado é sobre quantos indivíduos e ou progênies serão necessários em cada geração (Ramalho et al., 2001). Considerando uma população segregando para onze locos, com distribuição independente, Ramalho et al. (2001) discutem o número de plantas necessárias em cada geração, utilizando como referência a probabilidade de 95% de manter o genótipo

desejado. Estabelecendo que este evento seja a obtenção de genótipo em homozigose completa com os alelos favoráveis, a chance de ocorrência aumenta com o decorrer das gerações de autofecundação. No entanto, quando se considera o evento desejado como o genótipo com alelos favoráveis em homozigose ou heterozigose, esta probabilidade diminui. Com base nestas observações, o tamanho da população na geração inicial F_2 não necessita ser grande $(3/4)^{11}$, ou seja, com apenas 69 plantas, tem-se 95% de probabilidade de se obter pelo menos uma planta com o genótipo desejado (em homozigose ou heterozigose). Quando se considera o homozigoto na F_2 , esse número é superior a 12,5 milhões, evidenciando que, nas gerações sucessivas, o tamanho da população deve ser aumentado. Este resultado mostra a dificuldade de se identificar genótipos homozigóticos desejáveis, já nas gerações iniciais, sendo esta tarefa mais facilitada após a realização de algumas gerações de autofecundação, a qual promove um incremento concomitante na frequência de homozigotos na população com a conseqüente redução da proporção de heterozigotos (Ramalho et al., 2001; Wricke & Weber, 1986).

Outro aspecto relevante a ser considerado refere-se à constituição genética das populações ao longo das gerações de endogamia em termos dos componentes da variância genética. Este conhecimento tem grande importância para o estabelecimento de estratégias e tomada de decisões durante o ciclo de melhoramento, a fim de permitir explorar adequadamente a variabilidade genética existente. Para isso, Cockerhan (1963), citado por Ramalho & Vencovsky (1978), propôs um método prático para a obtenção dos componentes da variância genética em plantas autógamas, baseado no princípio estabelecido por Horner & Weber (1956), de que a covariância entre parentes é função da variância genética. Uma descrição criteriosa deste método e da notação utilizada foi realizada por Souza Junior (1989).

Considerando a população F_2 não endogâmica e com frequência alélica 0,5, por originar-se do cruzamento entre duas linhagens, a variância genética total nesta geração, na ausência de efeitos epistáticos, é dada por: $\sigma_{G_{F_2}}^2 = \sigma_A^2 + \sigma_D^2$, em que: σ_A^2 é a variância genética aditiva que, neste caso de frequência alélica 0,5, deve-se exclusivamente aos efeitos aditivos, os quais são efetivamente transmitidos de uma geração para outra e σ_D^2 é a variância genética de dominância (Ramalho et al., 1993; Souza Junior, 1989). Quanto à importância relativa destes componentes de variância, Bernardo (2002) comenta que a σ_A^2 é usualmente maior do que σ_D^2 em culturas autógamas.

A variância genética total nas gerações seguintes a F_2 ($\sigma_{G_{F_g}}^2$), desconsiderando os efeitos da seleção, pode ser obtida pela seguinte expressão:

$$\sigma_{G_{F_g}}^2 = (1 + I_g) \sigma_A^2 + (1 - I_g^2) \sigma_D^2, \quad (1)$$

em que: I_g corresponde ao coeficiente de endogamia da geração g , o qual pode ser obtido por $I_g = 1 - (1/2)^{g-2}$ (Ramalho & Vencovsky, 1978; Souza Junior, 1989). A partir da expressão (1), depreende-se que ocorre um aumento crescente da contribuição de σ_A^2 para a variância genética, enquanto que a contribuição de σ_D^2 diminui. Na F_∞ , a variância genética entre linhagens conterà $2 \sigma_A^2$ (Souza Junior, 1989).

Diante do exposto, torna-se necessário a escolha criteriosa do método de seleção que conduzirá a população segregante durante as gerações de endogamia de modo a manejá-la adequadamente, melhorando a eficiência do processo seletivo. Como exemplo, Wang et al. (2003) destacam que o sucesso do programa de melhoramento de trigo desenvolvido pelo *CIMMYT* é devido, dentre outros fatores, às estratégias de melhoramento adotadas por seus melhoristas.

Para a tomada de decisão sobre qual método de condução utilizar, o melhorista deve observar cuidadosamente qual é mais adequado a seus objetivos e ao tempo desejado para a obtenção das linhagens, além de avaliar outros fatores que têm influência direta, como a herdabilidade dos caracteres sob seleção, a adequabilidade dos ambientes para seleção e, principalmente, a quantidade de recursos e infra-estrutura disponíveis (Fehr, 1987; Fouilloux & Bannerot, 1988). Para Mallmann et al. (1994), a implementação dos métodos de condução constitui a fase mais crítica, laboriosa e onerosa de um programa de melhoramento. A alocação judiciosa dos recursos é, indubitavelmente, um fator *sine qua non* para atendimento do objetivo primordial do melhoramento de plantas, o qual, segundo Fehr (1987), é melhorar a performance genética de cultivares de uma dada espécie na maneira mais eficiente possível.

2.1.3.1 Método genealógico ou *pedigree*

O surgimento do método genealógico ou *pedigree* clássico é datado para a última década do século XIX, antes mesmo da redescoberta das leis de Mendel (Akerman et al., 1938, citado por Jensen, 1988). O termo *pedigree selection* foi primeiro empregado quando a seleção de plantas individuais foi usada para isolar linhas puras em populações heterogêneas de espécies autógamas (Fehr, 1987).

O método genealógico foi desenvolvido na *Swedish Seed Association*, em Svalof, Suécia, sendo sua proposição creditada a Hjalmar Nilsson, conforme descrição realizada por Newman, em 1912 (Fehr, 1987). Segundo esta descrição, a descoberta se deu a partir da coleta de plantas individuais de trigo e outras espécies, realizada durante a colheita de 1891. Posteriormente, estas plantas foram classificadas botânica e morfológicamente, sendo as sementes das plantas

de similar classe misturadas e uma amostra de cada plantada em linha na próxima geração. Da mesma maneira, as classes com apenas uma planta também tiveram suas sementes plantadas em linha. Durante a avaliação das linhas, Nilsson observou que somente as progênies das linhas provenientes de plantas individuais tinham características uniformes. A partir desta observação, Nilsson referiu-se ao uso da seleção de planta individual seguida da avaliação da progênie como “*system of pedigree*”. A primeira descrição completa e detalhada do método genealógico foi apresentada por Love, em 1927, que logo o introduziu em todos os programas de melhoramento por ele supervisionados na Universidade de Cornell (Jensen, 1988).

Desde então, o método *pedigree* foi evoluindo e sendo largamente empregado, originando muitas cultivares de sucesso em várias espécies autógamas na Europa, especialmente, após reconhecida a importância de sua utilização de forma combinada em populações base de melhoramento provenientes de hibridação artificial (Jensen, 1988). A partir disso, este método passou a ser freqüentemente utilizado para o desenvolvimento de linhagens em vários programas de melhoramento de espécies autógamas, como arroz (Castro et al., 2005), soja (Sedyama et al., 2005), feijão (Vieira et al., 2005), trigo (Federizzi et al., 2005) e café arábica (Sakiyama et al., 2005). Além disso, este método também se aplica às espécies alógamas, sendo, inclusive, o mais popular para a obtenção de linhagens endogâmicas de milho atualmente (Borém & Miranda, 2005).

No método genealógico, inicialmente, as plantas na geração F_2 são escolhidas com base em fenótipos predefinidos. As sementes destas plantas selecionadas são colhidas e semeadas separadamente em linhas, formando as progênies ou linhas $F_{2,3}$ que, conforme notação utilizada por Fehr (1987), designam as progênies F_3 provenientes de plantas F_2 . As melhores progênies $F_{2,3}$ são selecionadas e, ainda, os melhores indivíduos dentro dessas melhores

progênies com base na avaliação para os caracteres de interesse. O processo, então, se repete, até que, com sucessivas gerações de autofecundação, resultem em progênies com nível de homozigose desejado (F_6 ou F_7), cujas sementes serão colhidas em *bulk* para, logo depois, seguirem para a avaliação em experimentos com repetição em vários locais. As melhores linhagens identificadas irão participar dos ensaios regionais de competição de cultivares (Ramalho et al., 1993). A condução pelo método *pedigree* deve ser realizada em condições ambientais apropriadas para que ocorra a expressão das diferenças genéticas entre as plantas ou progênies para os caracteres de interesse, como a resistência a doenças, para a confiabilidade nos resultados da seleção (Fehr, 1987).

O princípio deste método é que a seleção individual de plantas na população segregante, com posterior avaliação das progênies e o conhecimento da genealogia dos tipos selecionados, permitam a maximização da eficiência da seleção (Borém & Miranda, 2005).

Na utilização do método genealógico, quatro aspectos importantes necessitam ser enfocados: a viabilidade da seleção dentro das progênies, a eficácia da seleção visual realizada durante as fases de endogamia, a eficiência da seleção precoce e, por último, as anotações dos registros genealógicos realizadas no decurso das gerações de endogamia.

Como comentado previamente, a condução da população segregante pelo método genealógico é efetuada com os indivíduos em cada geração estruturados em progênies, nas quais os indivíduos são relacionados geneticamente (Fehr, 1987; Fouilloux & Bannerot, 1988). Nesta condição, torna-se necessário conhecer, primeiramente, como a variância genética total (1) está distribuída entre e dentro das progênies nas sucessivas gerações (Ramalho et al., 2001). Conforme apresentado por Bernardo (2002) e Souza Júnior (1989), a variância

genética entre progênies da geração g oriundas de plantas da geração $(g-1)$ pode ser obtida pela fórmula geral:

$$\sigma_{G_{Entre F_{g-1:g}}}^2 = (I + I_{g-1}) \left[\sigma_A^2 + \frac{(I - I_g)^2}{I - I_{g-1}} \sigma_D^2 \right], \quad (2)$$

enquanto que a variância genética dentro média de progênies pode ser obtida subtraindo-se (1) de (2) resultando em:

$$\sigma_{G_{Dentro F_{g-1:g}}}^2 = (I - I_g) (\sigma_A^2 + \sigma_D^2), \quad (3)$$

Diante das expressões apresentadas, conclui-se que, a cada geração adicional de autofecundação, ocorre um aumento na variância aditiva entre progênies (2), contudo, com decréscimo desta dentro de progênies (3). Isso, conseqüentemente, acarreta na perda progressiva da efetividade da seleção dentro de progênies (Bernardo, 2002; Ramalho et al., 1993). Assim, a seleção dentro de progênies apenas é justificada até a geração F_5 , devendo, nas gerações seguintes, ser baseada no comportamento entre progênies (Ramalho et al., 2001).

É oportuno salientar que, pelo fato da seleção no método genealógico ser realizada entre e dentro de progênies, a variância genética aditiva total presente é explorada completamente naquela geração (Ramalho & Vencovsky, 1978). No entanto, uma característica intrínseca a este método, como já mencionado, é que esta seleção é efetuada, exclusivamente, com base em avaliações visuais. Uma discussão detalhada sobre o procedimento de seleção visual foi realizada por Jensen (1988), o qual comenta que a observação é uma parte inseparável do processo de melhoramento de plantas.

Uma das primeiras discussões sobre a seleção visual ocorreu no trabalho de Engledow e Wadham, em 1923, no qual os autores afirmam que o julgamento visual melhora com a experiência do melhorista (Jensen, 1988). Para Fehr (1987), a eficiência da seleção visual está relacionada com a habilidade do melhorista em discernir com consistência as diferenças genéticas verdadeiras.

No entanto, vários trabalhos têm demonstrado não haver essencialmente diferenças entre indivíduos experientes ou inexperientes quanto à competência em discriminar progênies superiores, sobretudo, para caracteres de herbabilidade baixa. Silva et al. (1994), avaliando o efeito deste tipo de seleção para o caráter produtividade de grãos de feijão, realizada por dez selecionadores, não constataram nenhuma tendência de associação entre a experiência do indivíduo com a cultura e a eficiência da seleção.

De acordo com Ramalho et al. (2001), o uso da seleção visual constitui uma limitação séria ao uso do método *pedigree* para caracteres quantitativos em virtude de sua baixa eficiência. Estudos realizados em várias oportunidades e com diferentes culturas têm apontado para esta ineficiência, como em feijão (Patiño & Singh, 1989; Silva et al., 1994), milho (Vargas, 1996) e arroz (Cutrim et al., 1997).

Outro ponto importante no método *pedigree*, que está ligado com a eficácia da seleção visual, é a sua utilização ainda nas gerações iniciais. A seleção precoce objetiva descartar progênies derivadas de F₂ ou F₃ com desempenho inferior e, com isso, permitir a alocação eficiente de recursos para testes e avaliações de progênies promissoras (Bernardo, 2002; Fehr, 1987). A pressuposição básica da seleção precoce, é que o desempenho de uma progênie na geração de endogamia inicial seja preditiva da performance média das linhagens derivadas, quando a homozigose for alcançada (Bernardo, 2003).

A efetividade da seleção precoce foi demonstrada por Bernardo (1991). O autor discute que a eficiência prática da seleção precoce é determinada pela correlação entre o valor fenotípico médio da progênie na geração de seleção (g) e o desempenho genotípico médio das linhagens descendentes em gerações avançadas $[r(\bar{p}_g, \bar{G}_\infty)]$, que é obtida por:

$$r(\bar{p}_g, \bar{G}_\infty) = h_g \cdot r(\bar{G}_g, \bar{G}_\infty), \quad (4)$$

em que: h_g é a raiz quadrada da herdabilidade na geração de seleção e $r_{(\bar{G}_g, \bar{G}_\infty)}$ é correlação genética entre as progênies na geração de seleção e as respectivas linhagens na geração F_∞ . Esta correlação genética, considerando a ausência de dominância, é função exclusivamente do coeficiente de endogamia na geração de seleção: $r_{(\bar{G}_g, \bar{G}_\infty)} = \sqrt{(I + I_g)/2}$. Dessa forma, as correlações genéticas entre progênies $F_{2:3}$, $F_{3:4}$, $F_{4:5}$ e $F_{5:6}$ e as linhagens são, respectivamente, iguais a 0,71, 0,87, 0,94 e 0,97. Portanto, com este resultado, observa-se que a efetividade da seleção precoce torna-se fundamentalmente limitada pela herdabilidade (Bernardo, 2002). Todavia, esta herdabilidade é geralmente baixa, devido ao número limitado de sementes para as avaliações das progênies nas gerações iniciais. Uma estratégia a ser utilizada, neste caso, é melhorar a precisão experimental na avaliação destas progênies.

Outra atividade peculiar ao método genealógico, que juntamente com a seleção visual torna sua aplicação laboriosa e dispendiosa, é a anotação da genealogia dos indivíduos selecionados. A razão de manter este sistema de registro durante o emprego deste método é maximizar a variabilidade genética entre progênies retidas durante a seleção (Fehr, 1987). Segundo Ramalho et al. (2001), outro argumento utilizado é o fato de que, com o conhecimento da genealogia das melhores linhagens identificadas, será possível para o melhorista voltar a estas progênies e repetir a seleção. Porém, segundo estes autores, este argumento é discutível, pois não implica em certeza de sucesso advindo da nova seleção pela obtenção de linhagens com performance igual ou superior às previamente selecionadas.

Em geral, à informação de parentesco gerada pelo trabalho adicional despendido em anotar os registros genealógicos, tem sido atribuída pouca utilidade prática para fins seletivos (Ramalho et al., 2001). No tocante a este aspecto, situa-se um dos objetivos do presente trabalho, o qual visa a utilização

dessa informação incorporada diretamente no método de análise baseado na teoria dos modelos lineares mistos (Henderson et al., 1959), com a finalidade de melhorar a eficiência da seleção de progênes superiores (itens 2.3 e 3.1.4).

2.1.3.2 Método populacional ou *bulk*

O método da população, também conhecido com o método *bulk*, foi desenvolvido inicialmente pelo melhorista sueco Nilsson-Ehle, em 1908, a partir de trabalhos realizados na *Swedish Seed Association (SSA)*, em Svalof, Suécia, com a cultura do trigo (Fehr, 1987; Jensen, 1988).

A descrição do método, conforme utilizado pela *SSA* para o desenvolvimento de cultivares de trigo adaptadas ao inverno, foi feita por Newman, em 1912 (Fehr, 1987). Todavia, segundo Jensen (1988), foi o trabalho clássico publicado por Florell, em 1929, que apresentou de forma clara e definitiva este método de seleção e mostrou seu potencial para competir eficientemente em resultados com outros métodos amplamente usados na época, como o método *pedigree*. Já no mencionado trabalho, aspectos sobre as bases teóricas para determinação do número de gerações necessárias para iniciar a seleção e o problema de amostragem foram discutidos.

A condução pelo método *bulk* é simples. Inicia-se a partir da geração F_2 , em que as plantas são colhidas massalmente e, então, uma amostra de sementes é retirada para a obtenção da população na geração F_3 . Este processo é repetido até a geração que apresente um nível adequado de homozigose, na qual será procedida a seleção de plantas individuais ou “abertura do *bulk*”. A partir disso, as progênes oriundas das plantas individuais colhidas serão avaliadas, em conformidade com o descrito anteriormente para o método *pedigree* (Borém & Miranda, 2005; Ramalho et al., 2001).

Três aspectos importantes inerentes ao emprego do método *bulk* precisam ser discutidos: a ação da seleção natural durante as gerações de autofecundação, o problema de amostragem e a determinação da época adequada para iniciar a seleção.

No método *bulk*, a seleção natural é permitida atuar sobre a população durante as gerações de endogamia. A premissa aceita é que esta seleção promova o aumento da frequência de genótipos agronomicamente desejáveis. Contudo, é inevitável questionar se isso realmente ocorrerá no sentido almejado pelo melhorista (Ramalho et al., 2001).

Este efeito da seleção natural por 50 gerações foi avaliado por Allard (1988). O autor tomou como referência a população de cevada sintetizada em 1928, por Harlan e Martini, constituída pelas sementes F_1 de todos os possíveis cruzamentos aos pares de 28 cultivares de cevada. Foram constatadas mudanças para vários caracteres. Para o rendimento de grãos, foi evidenciado um aumento de cerca de 1% a 2% ao ano, o que, segundo Ramalho et al. (2001), reflete ganho comparável ao da seleção artificial. Foi discutido que estas alterações no rendimento de grãos de cevada resultaram do aumento no número de sementes e não propriamente de um aumento no tamanho da semente.

Resultados experimentais semelhantes têm sido obtidos com populações de feijoeiro conduzidas pelo método *bulk*. Foi observado, pela avaliação de populações durante 15 gerações, que a ação da seleção natural contribuiu para um incremento médio do rendimento de grãos de 2,5% por geração (Corte et al., 2002). Em populações de feijão avançadas da geração F_2 a F_{13} , Gonçalves et al. (2001) observaram que os indivíduos com hábito de crescimento indeterminado e com menores sementes foram mantidos.

Os resultados anteriores evidenciam que a seleção natural atua no sentido de eliminar os indivíduos menos adaptados que, segundo Allard (1988), consistem em indivíduos com menor capacidade de sobreviver e reproduzir num

dado conjunto de ambientes. Ainda diante desses resultados, torna-se essencial, quando o objetivo for selecionar linhagens com determinados fenótipos, como hábito de crescimento determinado, semente com padrão carioca, florescimento precoce e sementes grandes, para o caso do feijão, que a seleção artificial seja praticada a partir da geração F₃ (Corte et al., 2002). Esta seleção tem por finalidade eliminar plantas com características indesejáveis e direcionar o *bulk* para o objetivo pretendido (Fehr, 1987; Florell, 1929, citado por Jensen, 1988). Esta prática reduz, ainda, as perdas não aleatórias oriundas da competição intergenotípica.

Ante o evidenciado, torna-se patente que a influência da seleção natural deverá ser levada em consideração quando da escolha dos ambientes onde as gerações pelo *bulk* serão avançadas. Estes ambientes deverão refletir as condições de cultivo e possibilitar a expressão das diferenças genéticas entre os indivíduos para os caracteres de interesse, de forma que a seleção natural atue no sentido desejado (Fehr, 1987). Este aspecto foi estudado por Pirola et al. (2002) usando populações de feijão avançadas pelo método *bulk* em diferentes ambientes. Os autores mostraram que a seleção natural atuou mantendo os indivíduos mais adaptados para aqueles ambientes onde as populações foram avançadas.

No método *bulk*, ocorre a impraticabilidade de avançar com todas as sementes colhidas a cada geração, tornando evidente a necessidade de aplicação de procedimento de amostragem. Segundo Fouilloux & Bannerot (1988), a amostragem constitui-se numa das limitações do método *bulk*, conduzindo, indubitavelmente, a uma perda substancial de progênies de uma geração para outra. Em estudo por simulação avaliando as conseqüências das perdas amostrais no método *bulk*, foi estimado que a taxa de extinção esperada de progênies após uma geração seja superior a 20% (Kervella & Fouilloux, 1992).

A partir de simulações, Muehlbauer et al. (1981) encontraram que a estimativa da variância genética aditiva em F_6 para o *bulk* foi inferior à obtida para o *SSD*. Este fato foi explicado em função da competição e fecundidade das plantas. Porém, segundo Fouilloux & Bannerot (1988), a substancial perda de progênies não conduz a uma correspondente redução na variabilidade genética, pelo fato de essa perda estar associada a um incremento da variabilidade dentro da progênie.

A fim de minimizar problemas com a amostragem, Kervella & Fouilloux (1992) demonstraram que o aumento do tamanho da população ao longo das gerações promove uma redução da taxa de extinção. Esta alternativa é exequível para o método *bulk*, em virtude de sua facilidade de condução (Ramalho et al., 2001).

Como a seleção é postergada até que um nível de homozigose seja atingido, uma questão relevante que surge é em que momento deverá ser “aberto o *bulk*” (Ramalho et al., 1993). Baseado nas alterações ocorridas nas frequências genotípicas e, por conseguinte, nos acréscimos cada vez menores nas frações da variância aditiva explorada ao longo das gerações de endogamia, conforme explicado anteriormente, não restam justificativas teóricas para protelar a etapa de seleção além das gerações F_4 ou F_5 (Ramalho et al., 1993). Todavia, Silva et al. (2004) revelaram, por meio de avaliação experimental de progênies de feijão extraídas nas gerações F_2 , F_8 e F_{24} conduzidas pelo método *bulk*, que a frequência de progênies superiores aumentou quando a extração da progênie foi adiada para gerações mais avançadas. Este resultado aponta para uma importante característica do método *bulk*, que consiste na sua flexibilidade em permitir o atraso para a extração das linhagens sem eventual perda para o melhorista (Silva et al., 2004).

2.1.3.3 Comparação de métodos de condução em plantas autógamas

Conforme descrito previamente, existem vários métodos de condução de populações segregantes disponíveis para espécies autógamas. Assim, o melhorista terá a tarefa de decidir sobre qual método utilizar no seu programa de melhoramento. Para auxiliar na tomada desta decisão, torna-se patente a necessidade de informações para que se proceda a escolha pelo método mais adequado e ou eficiente para atender aos objetivos do melhorista (Hallauer & Miranda Filho, 1988; Ramalho et al., 2001).

Para a comparação de métodos de condução de populações segregantes, as mesmas alternativas empregadas para comparar os diferentes métodos de seleção recorrente podem ser utilizadas (Fehr, 1987). Duas dessas são apontadas por Hallauer & Miranda Filho (1988): o uso de resultados obtidos por diversos pesquisadores e, outra, apoiada nas bases teóricas fornecidas pela genética quantitativa usando a mesma intensidade seletiva. A primeira alternativa apresentada é difícil de ser utilizada, em virtude da grande quantidade de variáveis envolvidas, como diferentes métodos, populações, ambientes, tamanhos amostrais, intensidades de seleção, além de diferenças relacionadas com os critérios de seleção dos diferentes melhoristas. Em contrapartida, quando opta-se pela segunda alternativa, as dificuldades acima não mais existem. Neste caso, a efetividade dos métodos é comparada por meio da predição do ganho genético. Porém, para sua aplicação, incorre-se na necessidade de estimativas acuradas dos parâmetros envolvidos nessas expressões teóricas do progresso genético esperado. Uma terceira alternativa e, indubitavelmente, a ideal, consiste em comparar diretamente os diferentes métodos numa mesma população segregante, em igualdade de condições, avaliando os mesmos caracteres (Ramalho, 1997; Weyhrich et al., 1998).

Vários trabalhos de pesquisa têm sido realizados visando confrontar os variados métodos de condução, sob a terceira alternativa, em diversas espécies autógamias utilizando variados critérios (Raposo, 1999). Mas, os resultados obtidos não são concordantes sobre qual método é mais eficiente. Contudo, estes trabalhos têm mostrado, em geral, que todos os métodos de condução avaliados são eficientes na obtenção de linhagens superiores.

Para a cultura da soja, diferentes métodos de condução foram comparados em várias circunstâncias (Boerma & Cooper, 1975; Martin & Geraldi, 2002; Voigt & Weber, 1960). Para o estudo conduzido por Boerma & Cooper (1975), por exemplo, foram utilizadas quatro populações segregantes. Os métodos *pedigree*, *bulk* dentro de F_2 e *SSD* foram comparados usando como critérios as médias de todas as progênies selecionadas, das cinco melhores e da melhor selecionada, além da variância genotípica entre progênies. Considerando as quatro populações, os autores não detectaram nenhuma diferença consistente entre os métodos para o caráter rendimento de grãos, entretanto, para o caráter precocidade, os métodos *SSD* e *pedigree* se mostraram mais eficientes. Quanto à variância genotípica, o método *pedigree* pareceu reter maior variabilidade entre as progênies para o rendimento de grãos.

Alguns trabalhos, objetivando comparar a eficiência dos métodos de condução, também são relatados para a cultura do arroz (Fahim et al., 1998; Mishaet al., 1994; Ntanos e Roupakias, 2001). Na investigação realizada por Fahim et al. (1998), foram comparados quatro métodos de condução: o *pedigree*, *pedigree* modificado, *bulk* e *SSD*, a partir de dois cruzamentos biparentais. Foram avaliadas as progênies F_6 de cada método para onze caracteres. O método *pedigree* modificado consistiu em colher as melhores plantas durante cada geração de endogamia e, em seguida, retirar de cada uma destas plantas uma semente. Após isso, as sementes foram misturadas e uma amostra retirada para formar a população da geração seguinte. Com base na performance das melhores

progênies selecionadas, não foram verificadas diferenças significativas entre os métodos para as características avaliadas. O método *SSD* foi tão eficiente quanto os demais, tendo a vantagem de ser rápido e menos oneroso, enquanto que os métodos *pedigree* e *pedigree* modificado foram os que apresentaram maiores custos.

Comparações neste contexto são referenciadas em pesquisas conduzidas com populações de feijão (Costa et al., 2002; Ranalli et al., 1996; Raposo et al., 2000; Urrea & Singh, 1994). Todavia, como já mencionado, as conclusões obtidas são, em determinados trabalhos, discordantes. Urrea & Singh (1994) constataram que a produtividade média das progênies derivadas do *SSD* foi menor quando comparada à alcançada com os métodos *bulk* e *bulk* dentro de F_2 , concluindo pela não recomendação deste método para condições tropicais. Já Ranalli et al. (1996), comparando os métodos *bulk* dentro de F_2 e *SSD* em duas populações de feijão, observaram que o método *SSD* resultou em progênies com desempenho superior para ambos os cruzamentos.

Os métodos *pedigree*, *bulk*, *bulk* dentro de progênies F_2 e F_3 e *SSD* foram comparados por Raposo et al. (2000). Utilizou-se uma população segregante resultante do cruzamento inter-racial entre os genitores ‘Carioca’, raça Mesoamericana e ‘Flor de Mayo’, raça Andina. As 64 progênies F_6 provenientes de cada método mais quatro testemunhas foram avaliadas em experimentos conduzidos em dois locais. A partir dos resultados, os autores verificaram que todos os métodos possibilitaram sucesso com a seleção. Constataram, ainda, que não houve diferenças marcantes entre os métodos na obtenção de progênies superiores. Entretanto, considerando as estimativas dos parâmetros genéticos e fenotípicos, a facilidade e a flexibilidade de condução, sugeriram os métodos *bulk* e *SSD* como os mais vantajosos.

Os estudos de comparação de métodos de condução fazendo uso de simulação computacional serão abordados no item 2.2.

2.2 Simulação computacional para comparação de métodos de condução em plantas autógamas

Com o advento da revolução computacional, o emprego de simulação tornou-se freqüente nas mais diversas áreas da ciência, fornecendo poderosas ferramentas para a pesquisa e ensino (Morgan, 1995).

A simulação, de acordo com Dachs (1988), consiste em imitar, por meio de programa de computador, o comportamento de um sistema real, para estudar seu funcionamento sob condições alternativas. Diante deste conceito, depreende-se que a simulação viabiliza a extração de informações importantes, com baixo custo e rapidez, mediante alterações de interesse nos fatores envolvidos no fenômeno em estudo (Ferreira, 2001).

Primordialmente, para realizar simulações há a necessidade da formulação de modelos matemáticos que descrevam adequadamente o comportamento dos eventos reais em estudo (Mood et al., 1974). Estes modelos necessitam ser suficientemente simples para permitirem a operacionalização, na forma de algoritmos e interpretação de forma adequada (Cruz, 2001). Neste aspecto, surge, em muitos casos, uma desvantagem nos modelos simulados no que se refere à simplificação das pressuposições (Van Oeveren & Stam, 1992; Wang et al., 2003). Por outro lado, estes modelos permitem a variação nos valores assumidos pelos parâmetros, de tal forma que uma amplitude realística de fatores pode ser facilmente testada e, mais importante, o efeito relativo destes fatores sobre os resultados da seleção pode ser estudado (Van Oeveren & Stam, 1992). No campo da genética e melhoramento de plantas, os modelos são inerentemente probabilísticos por envolverem várias variáveis de natureza aleatória (Ferreira, 2001).

Os primeiros trabalhos utilizando simulação na genética foram realizados por Fraser, em 1957 (Gurgel, 2004). A partir disso, a simulação tem tido grande aplicabilidade nas pesquisas de genética e melhoramento de plantas,

contribuindo de forma substancial para a resolução e ou elucidação de problemas complexos. Como exemplos, citam-se estudos avaliando a eficiência da seleção recorrente (Fouilloux & Bannerot, 1988), determinação do número de famílias a serem avaliadas (Ferreira et al., 2000; Gurgel, 2004), eficiência de dialelos circulantes na escolha de genitores (Veiga et al., 2000), viabilidade da seleção assistida por marcadores moleculares (Bearzoti, 1997; Liu et al., 2004), dentre várias outras aplicações.

No melhoramento de plantas, dispõem-se de várias estratégias que podem ser adotadas pelos melhoristas em seus programas. Para o caso particular de espécies autógamas, as diferentes estratégias de melhoramento distinguem-se, principalmente, pela possibilidade de utilização de variados métodos de condução das populações segregantes formadas (item 2.1.3). Dessa forma, um grande interesse tem sido focalizado para a comparação destas estratégias em termos de sua eficiência, para assim decidir ou não pela sua adoção. Neste contexto, comparações experimentais são freqüentemente muito onerosas e as técnicas matemáticas somente levam em conta um limitado número de parâmetros (Fouilloux & Bannerot, 1988).

Van Oeveren & Stam (1992) acrescentam, ainda, que a experimentação convencional apresenta restrições também devido à insuficiente amostragem – experimentos de campo. Os resultados advindos desses experimentos de campo poderiam mascarar a situação real pela eventual presença de efeitos aleatórios ocasionais. Dessa forma, estes autores sugerem que o uso de simulação computacional pode contribuir efetivamente para a otimização de programas de melhoramento, por sobrepor estas limitações de ordem prática e por possibilitar que um grande número de repetições de experimentos seja utilizado, de forma que uma informação mais segura seja tomada por meio da média obtida com estes experimentos simulados.

O primeiro trabalho usando simulação computacional objetivando comparar métodos de condução em plantas autógamas foi realizado por Casali & Tigchelaar (1975). Na ocasião, foram comparados os métodos *pedigree*, *bulk*, *SSD* e métodos combinando o *pedigree* ou *bulk* até as gerações F_3 e F_4 seguidos pelo *SSD*. A população F_2 simulada constou de 400 indivíduos. A seleção foi baseada no valor fenotípico. Ao final, foram simuladas 32 progênies F_6 para os métodos *pedigree* e *bulk*, e 32, 64 e 320 para o *SSD* e métodos combinados. Utilizaram-se os valores genotípicos (*VG*) médios das melhores progênies F_6 para verificar o progresso com cada método.

Os *VG* foram gerados considerando um sistema genético assumindo 20 locos com distribuição independente, com dois alelos cada, de efeitos iguais e aditivos, admitindo a ausência de dominância. Para a obtenção dos valores fenotípicos foram acrescentados aos *VG* efeitos ambientais aleatórios normalmente distribuídos. As simulações foram realizadas para as herdabilidades em nível de indivíduo de 10%, 25%, 50% e 75%. Dois modelos genéticos foram empregados: Modelo I, assumindo um só caráter controlado por 20 locos; Modelo II, assumindo 5 caracteres controlados por 2, 3, 5, 5 e 5 locos com herdabilidades de 100%, 75%, 50%, 25% e 10%, respectivamente.

Os autores constataram que a escolha do método variou com a herdabilidade do caráter. O *pedigree* foi preferido para os casos de herdabilidades alta (75%) e moderada (50%), enquanto que, para herdabilidades de 10% e 25%, os métodos *bulk* e *SSD*, respectivamente, resultaram nas melhores progênies F_6 . Concluíram, ainda, que a seleção em gerações iniciais para caracteres altamente herdáveis, seguida pelo *SSD* até a geração F_6 e, posterior seleção das progênies para caracteres de baixa herdabilidade, seria mais efetiva.

Os métodos *bulk* e *SSD* foram comparados por simulação, após quatro gerações de endogamia, por Muehlbauer et al. (1981). Os autores assumiram um

modelo constando de sete cromossomos com seis locos cada e apresentando taxas de recombinação constante entre locos vizinhos de 0,0; 0,1; 0,2; 0,3 e 0,5. A taxa de multiplicação média (*TM*) por planta para o caso do método *bulk* foi de 50 sementes, com desvios de 25, 50 e 100 sementes por planta, ao passo que para o *SSD*, as taxas de sobrevivência (*TS*) das plantas de uma geração para outra foram de 0,707, 0,780, 0,841, 0,95 e 0,99. Toda a variação genética presente foi estritamente aditiva. Considerou-se na simulação uma população F_2 de 200 indivíduos para cada método. Os critérios utilizados para a comparação dos métodos foram a variância genética aditiva, os valores genéticos médios na geração F_6 , e a porcentagem de progênies F_2 representadas na F_6 .

Os autores observaram que a variância genética aditiva na F_6 no método *bulk* foi inferior à obtida no *SSD*, relatando que a *TM* afetou consideravelmente a variabilidade genética no *bulk*. Considerando-se um desvio na *TM* de 25, cerca de 75% das progênies F_2 não estavam representadas na F_6 . Para o método *SSD*, a *TS* foi importante. A perda de progênies na F_6 no *SSD* foi equiparada ao *bulk* quando a *TS* foi inferior a 75% a cada geração de avanço. Relataram, ainda, que a ligação teve pouca influência na variância genética aditiva.

Tendo em vista que a seleção pode ser praticada em diferentes estágios de um processo de melhoramento e com variadas intensidades, Van Oeveren & Stam (1992) realizaram um estudo via simulação computacional com o objetivo de avaliar a eficiência de dois distintos métodos: um método com teste em geração precoce (*TGP*), no caso em F_3 , estimando a probabilidade de obter linhagens na F_∞ que superem um determinado padrão e o método *SSD*, no qual a seleção é atrasada para gerações de endogamia mais avançadas (F_6).

Os métodos foram comparados para os casos de um e cinco cruzamentos sob seleção, sendo somente os cruzamentos mais promissores avançados para o *TGP*. Cada cruzamento teve diferente número de locos segregantes (no máximo 30). Admitiu-se, nas simulações, um modelo genético para o caráter quantitativo

com todos os locos de efeitos iguais e aditivos e sem a presença de ligação. Simularam-se as variáveis: número de plantas F_2 (25 a 400), herdabilidade individual (0,02 a 1,0) e o efeito de dominância ($d=0$, $d=1/2a$, $d=a$). As dez melhores progênies fenotipicamente superiores selecionadas de cada método foram avaliadas e comparadas. Os critérios utilizados foram a média, o máximo e o desvio padrão genotípico dessas progênies.

Os autores concluíram que a herdabilidade e o número de plantas na F_2 foram as duas variáveis mais influentes. Além disso, considerando somente um cruzamento, o *TGP* resultou em progênies superiores ao *SSD*. Porém, quando mais cruzamentos são envolvidos, os métodos não diferiram, à exceção quando a herdabilidade foi baixa, em que o método *SSD* foi até melhor devido à ineficiência do *TGP* em selecionar os melhores cruzamentos. Relatou-se, ainda, que a dominância teve pouca influência sobre ambos os métodos estudados.

Os métodos *bulk-pedigree (BP)* e *selected bulk method (SB)*, largamente utilizados pelo Programa de Melhoramento de Trigo do *CYMMYT*, foram comparados via simulação computacional (Wang et al., 2003). As simulações foram realizadas usando a plataforma de simulação para análise quantitativa de modelos genéticos *QU-GENE* (Podlich & Cooper, 1998) e o módulo de aplicação *QUCIM*, especificamente desenvolvido para simular o procedimento de melhoramento para a cultura do trigo, adotado pelo *CYMMYT*. Os métodos foram comparados em termos do ganho genético ajustado, do número de cruzamentos mantidos após um ciclo de melhoramento e da alocação de recursos. O ciclo de melhoramento tendo seu princípio com o cruzamento e término na geração em que as novas linhagens obtidas retornam à fase de cruzamento como genitores. Para este estudo, a geração considerada foi a F_{10} .

Os métodos avaliados neste estudo distinguiram-se apenas pela forma de condução das populações nas gerações segregantes F_3 a F_5 . Após o início, com a seleção das plantas na F_2 , ambos seguiram por três gerações de endogamia pelo

método *bulk*, com a diferença de que, no método *SB*, apenas as plantas selecionadas foram colhidas em *bulk*. Após isto, novamente praticou-se a seleção de plantas individuais na geração F_6 . Nas gerações seguintes (F_7 a F_{10}), as progênies foram conduzidas por meio de *bulk* dentro de progênies e avaliadas em experimentos. Para ambos os métodos, consideraram-se os mesmos 1.000 cruzamentos a partir de 200 genitores.

Para efeito de simulação, foram consideradas as características dos ambientes alvo para os quais o programa de melhoramento de trigo de primavera do *CYMMIT* está direcionado, ambientes irrigados e com baixa pluviosidade. Foram considerados nove caracteres do trigo, com diferentes números de locos segregantes baseado em estudos realizados. Para o caráter produtividade de grãos, devido ao desconhecimento do número de locos, considerou-se a segregação para 20 e 40 locos. Assumiram-se para os locos, efeitos iguais e aditivos, exceto para os genes que governaram a produtividade de grãos, em que se considerou a presença de epistasia. A dominância não foi considerada. Os efeitos dos genes pleiotrópicos foram assumidos presumivelmente causadores da correlação entre os caracteres. A herdabilidade individual para o rendimento de grãos foi fixada em 0,05. Foram simuladas doze configurações considerando o caráter produtividade de grãos, sendo dois níveis para o número de locos com frequência 0,5 (20 e 40), três níveis de epistasia (ausente, digênica e trigênica) e dois níveis de pleiotropia (ausente e presente).

Dos resultados, os autores inferiram que o ganho genético padronizado com o método *SB* foi 3,3% superior ao alcançado com o *BP*. Evidenciou-se que o aumento do número de locos e a presença concomitante de epistasia e pleiotropia dificultam a identificação dos melhores genótipos, independente da estratégia utilizada, reduzindo o ganho genético. Uma maior proporção de cruzamentos foi retida pelo método *SB*, comparado ao método *BP*. Quanto aos recursos alocados da F_1 a F_8 , o método *SB* ocupou cerca de 1/3 a menos de área

do que o método *BP*, além de produzir um menor número de progênies (cerca de 40% a menos) durante estas gerações. Assim, Wang et al. (2003) concluíram que o método *SB* foi mais efetivo quanto ao tempo, trabalho e custo envolvidos, além de resultar num ganho genético levemente superior ao método *BP*.

2.3 Emprego do *BLUP* no melhoramento de plantas autógamas

O objetivo básico dos programas de melhoramento de plantas autógamas é, em geral, selecionar genótipos superiores, ou seja, portadores de alelos favoráveis para a maioria dos locos (Ramalho et al., 2001). Como estes valores genotípicos não podem ser mensurados diretamente, passam, então, a ser preditos a partir dos correspondentes valores fenotípicos observados.

Em espécies autógamas, a eficiência da avaliação fenotípica em atestar o verdadeiro valor associado a um dado genótipo, ou mais especificamente sobre seu valor genético aditivo (*VGA*), está diretamente relacionada com a complexidade do controle genético do caráter em questão. Para caracteres quantitativos, os quais são governados por um grande número de locos e, por sofrerem forte influência de efeitos ambientais sobre a expressão fenotípica, tem-se que a herdabilidade é normalmente baixa. Dessa forma, o valor fenotípico observado constitui um pobre indicador do *VGA* real (Bernardo, 2002).

Dessa maneira, o sucesso do melhoramento genético de plantas autógamas para os principais caracteres agronômicos de interesse, os quais apresentam controle poligênico, depende, primeiramente, da criteriosa avaliação das progênies em experimentos bem delineados (Ramalho et al., 2001). De posse desses dados, o passo seguinte consiste em explorá-los adequadamente,

sobretudo pela adoção de acurados procedimentos de seleção das progênes que associem os maiores *VGA* (Mrode, 1996; Resende, 2002a).

A necessidade de uma experimentação adequada é condição primordial, dado que o processo seletivo é realizado por meio do ordenamento das progênes usando alguma função dos dados. Para que o progresso genético seja maximizado, é necessário que esta função dos dados resulte no ranqueamento dos *VGA* preditos o mais próximo possível do ranqueamento pelos verdadeiros *VGA* (White & Hodge, 1989).

Alguns métodos têm sido utilizados com a finalidade de predizer o *VGA* em plantas pelo uso de diferentes funções dos dados. Estes métodos diferem principalmente pelas pressuposições assumidas e pelas considerações realizadas sobre a natureza dos efeitos dos *VGA* (White & Hodge, 1989). Os procedimentos mais usuais têm sido o *BLUE* – *Best Linear Unbiased Estimator* (melhor estimador linear não tendencioso), o *BLP* – *Best Linear Predictor* (melhor preditor linear) e, mais recentemente, o *BLUP* – *Best Linear Unbiased Predictor* (melhor preditor linear não tendencioso). Abordagens detalhadas a respeito desses procedimentos são apresentadas em Bernardo (2002), Henderson et al. (1959), Henderson (1975), Henderson (1984), Lynch & Walsh (1998), Mrode (1996), Resende (2002a), Searle et al. (1992), White & Hodge (1989). Neste referencial será tratado mais especificamente sobre o *BLUP*.

Antes de qualquer consideração metodológica sobre o procedimento *BLUP*, duas questões que, apesar de bastante discutidas em artigos científicos e ou livros texto, necessitam ser abordadas: a natureza do efeito de genótipos no modelo estatístico e, a controvérsia sobre estimação ou predição de efeitos aleatórios.

A decisão de tratar um efeito de um fator como sendo fixo ou aleatório e as suas implicações na análise estatística são bastante referenciadas na literatura (Eisenhart, 1947; McCulloch & Searle, 2001, Robinson, 1991, Searle, 1971). A

classificação dos efeitos interfere diretamente na definição do modelo e, conseqüentemente, na utilização de diferentes funções dos dados para ranquear os genótipos (White & Hodge, 1989). Segundo Bearzoti (2002), o ordenamento de genótipos pode ser alterado, de acordo com a escolha do modelo estatístico, em situações nas quais houver informação de parentesco entre as unidades de seleção e ou, desbalanceamento nos dados. Em estudo teórico, Duarte & Vencovsky (2001) constataram que a adoção da abordagem de modelo misto com genótipos aleatórios comparativamente às análises tradicionais (genótipos fixos) leva à seleção de diferentes tratamentos genéticos, quando a variância genotípica for baixa em relação à variância do erro e os ensaios forem não ortogonais e desbalanceados.

Os conceitos sobre a natureza dos efeitos presentes num modelo têm sido ampliados, mas, de modo geral, apresentam, como ponto comum, a questão da amplitude da inferência a ser realizada. Para o caso de efeitos fixos, as conclusões se atêm aos resultados da avaliação e ou comparação dos efeitos dos níveis do fator em estudo, enquanto que, para efeitos aleatórios, a conclusão obtida se estenderá para além dos níveis estudados.

Para Eisenhart (1947), efeitos devem ser considerados fixos (constantes) quando o interesse do pesquisador centra-se nos efeitos dos níveis do fator em estudo; caso contrário, os efeitos são considerados aleatórios. Para Searle (1971), quando os níveis de um dado fator se constituem numa amostra aleatória de um população de níveis, estes efeitos devem ser tomados como aleatórios, ao passo que, tendo sido estes níveis previamente escolhidos, estes deverão ser considerados fixos.

De acordo com Bearzoti (2002), a definição dos efeitos de um modelo como aleatórios deveria se basear não apenas na alusão a alguma população de referência, mas também se há alguma informação a ser recuperada. Por exemplo, em delineamentos do tipo blocos incompletos (*DBI*) (como em látice), os blocos

retêm informações dos efeitos dos tratamentos a serem recuperadas, justificando, portanto, assumi-los como aleatórios (Resende, 2002b). Do mesmo modo, genótipos, quando geneticamente relacionados, devem ser tomados como aleatórios e, assim, o procedimento estatístico adequado será a denominada análise com recuperação da informação intergenotípica (Federer, 1998).

Para Robinson (1991), sentenciar se determinados efeitos são fixos ou aleatórios depende da resposta a um questionamento principal: os efeitos em estudo provêm de uma distribuição de probabilidade? Em caso de resposta positiva, os efeitos serão aleatórios; caso contrário, serão considerados fixos. Nota-se, neste conceito, que pressuposições adicionais ao modelo são requeridas, conforme será discutido posteriormente.

Contudo, apesar dos conceitos de Eisenhart (1947) e Searle (1971) não fazerem menção à distribuição dos efeitos aleatórios, esta será uma pressuposição requerida, invariavelmente, para fins de aplicação de testes de hipótese e ou, obtenção dos intervalos de confiança para as variâncias dos efeitos aleatórios ou componentes de variância (McCulloch & Searle, 2001).

Em síntese, observada a discussão prévia, fica perceptível que a decisão sobre a natureza do efeito pode, em algumas circunstâncias, não ser óbvia. Como exemplo, citam-se os efeitos de anos e locais, conforme já evidenciado pelo próprio Eisenhart (1947). Já outros efeitos, como o de blocos, são considerados pelo autor como inerentemente aleatórios, justificando que, na tentativa de repetições completas dos experimentos, novos arranjos experimentais no campo sejam obtidos. No entanto, uma prática comum na experimentação agrícola, sobretudo em experimentos fitotécnicos, nos quais o delineamento em blocos completos casualizados (*DBC*) é o mais utilizado, é considerar blocos como fixos (Piepho, 1994). Do mesmo modo, no melhoramento animal, todos os efeitos não genéticos presentes no modelo são classificados como fixos, inclusive o de blocos (Mrode, 1996). De maneira

genérica, Resende (2002a) comenta que os blocos, quando em *DBC*, devem ser tratados como fixos e, quando em *DBI*, devem ser considerados aleatórios, conforme já justificado anteriormente. Uma discussão detalhada sobre considerar blocos como fixos ou aleatórios é apresentada em Resende (2002b).

Em se tratando de efeitos genéticos no melhoramento vegetal, esta classificação parece já bem resolvida. Por exemplo, as progênies resultantes de uma população base de melhoramento, formada a partir do cruzamento de duas linhagens, ao final de sucessivas gerações de autofecundação, são facilmente assumidas como aleatórias; enquanto que cultivares comerciais em ensaios de recomendação são consideradas de efeitos fixos.

A explicação para o primeiro exemplo apoia-se na própria natureza estocástica das leis da genética. Os efeitos de progênies são, em parte, resultante da soma dos efeitos médios dos alelos componentes dos genes que controlam o caráter de interesse herdados dos seus parentais, ou seja, dos *VGA* (Falconer & Mackay, 1996), o qual é também denominado, no melhoramento animal, como valor reprodutivo (do inglês *breeding value*). Como apenas os alelos são passados para formar os genótipos dos descendentes, fica bastante evidente que somente metade dos *VGA* de cada genitor é transmitida para a descendência. Assim, fica perfeitamente evidenciado que os *VGA* são, conceitualmente, aleatórios. Logo, um conjunto de progênies obtidas a partir de uma população se constitui numa amostra aleatória de uma população infinitamente grande de possíveis progênies. Quanto ao exemplo da avaliação de cultivares, o interesse do melhorista está intrinsecamente voltado para os genótipos em questão. Assim é bastante razoável supor que os efeitos desses sejam assumidos como fixos.

De acordo com a pressuposição sobre a natureza dos efeitos presentes no modelo, Eisenhart (1947) distinguiu, fundamentalmente, duas classes de problemas resolvidos com o uso da análise de variância. Na primeira classe, sendo o efeito de tratamento fixo, o interesse estaria voltado para a detecção e

estimação de efeitos fixos, enquanto que, na segunda classe, quando o efeito de tratamento é aleatório, o foco seria a detecção e a estimação de componentes de variância. Torna-se evidente, por esta classificação, que o rigor da suposição sobre a natureza do efeito é apenas considerada para fins de estimação das variâncias associadas aos efeitos aleatórios e construção dos estimadores para os testes F (Duarte & Vencovsky, 2001; White & Hodge, 1989).

Todavia, em algumas situações práticas, como, por exemplo, no melhoramento vegetal, o interesse do melhorista baseia-se mais fortemente sobre o mérito genético dos indivíduos para fins de seleção, ou seja, sobre os valores realizados dos efeitos aleatórios. Neste contexto, estabelece-se um questionamento posterior sobre se este problema seria de estimação ou predição destes efeitos.

Uma denominação mais frequentemente aceita sobre o tipo de procedimento estatístico utilizado, de acordo com a natureza dos efeitos, é o uso dos termos estimação para efeitos fixos e predição para efeitos aleatórios. Esta distinção é a enfatizada na maioria dos livros texto que tratam do assunto, como em Bernardo (2002), Henderson (1984), Lynch & Walsh (1998), McCulloch & Searle (2001), Mrode (1996), Resende (2002a), White & Hodge (1989). Esta terminologia será adotada no presente trabalho.

A despeito desta ampla aceitação, Robinson (1991) teceu críticas ao uso desta terminologia. O autor comenta que o termo estimação pode ser usado indistintamente, tanto para efeitos fixos como para aleatórios. Realça, ainda, que o termo preditor se constitui apenas numa convenção para designar o estimador para efeitos aleatórios. Resende (2002a) explica que, embora o vocábulo predição faça alusão a eventos futuros, significa também a estimação do valor realizado de uma variável aleatória amostrada de uma população. Desse modo, esta discussão é puramente referente à terminologia, não tendo qualquer interferência sobre o método estatístico propriamente dito.

Conforme já comentado, existem vários procedimentos estatísticos que se destinam à predição dos efeitos aleatórios ou *VGA*, mais especificamente no contexto do melhoramento de plantas autógamas. Contudo, o procedimento *BLUP* tem sido o mais extensivamente aplicado no melhoramento animal (Mrode, 1996). No melhoramento vegetal, sua utilização tem sido mais restrita, devido à falta de vivência com o método, pequena divulgação (Bueno Filho, 1997) e, sobretudo, devido ao fato dos dados serem usualmente balanceados (Bernardo, 2002). Algumas aplicações do *BLUP* têm sido realizadas em culturas perenes (Bueno Filho, 1997; Durel et al., 1998; Resende, 2002a) e, também em anuais, como milho (Bernardo, 1994; Bernardo, 2002) e soja (Panter & Allen, 1995a; Panter & Allen, 1995b).

O procedimento *BLUP* foi desenvolvido por Henderson, em 1949, no contexto de modelos mistos, com a finalidade de prever valores genéticos de animais para a produção de leite sob condições de desbalanceamento nos dados (Mrode, 1996; White & Hodge, 1989). Porém, somente foi apresentado formalmente por Henderson, em 1973 (Resende, 2002a). Uma síntese histórica da evolução do procedimento *BLUP* é apresentada em Resende (2002a).

Devido à necessidade de realização de cálculos mais complicados, a utilização do *BLUP* como método rotineiro de análise apenas se tornou possível a partir do avanço no campo computacional, sobretudo, com o desenvolvimento de softwares estatísticos que traziam algoritmos especializados já implementados. O primeiro uso do *BLUP* foi em 1970, para a predição do valor genético de 1.200 touros para a produção de leite num programa de inseminação artificial (Schaeffer, 1991).

A sigla *BLUP*, como já explicado anteriormente, significa melhor predição linear não tendenciosa, melhor indicando que este procedimento maximiza a correlação entre o *VGA* predito, simbolizado por \hat{a} e o *VGA* verdadeiro (a), ou seja, minimiza a variância do erro de predição [$V(\hat{a} - a)$];

linear no sentido de que é uma função linear dos dados e não tendencioso por apresentar a esperança do *VGA* predito igual à esperança do *VGA* verdadeiro [$E(\hat{\mathbf{a}}) = E(\mathbf{a})$] (Mrode, 1996; White & Hodge, 1989). Em geral, para os efeitos genéticos ou *VGA* de progênies oriundas de uma mesma população base, o valor médio esperado é, por definição, assumido como zero [$E(\mathbf{a}) = \mathbf{0}$] (Bernardo, 2002). As provas matemáticas de derivação e das propriedades pertinentes ao *BLUP* são apresentadas em Resende (2002a).

Para apresentar o procedimento *BLUP*, será considerada, sem perda de generalidade, a avaliação de progênies provenientes de uma mesma população base de melhoramento sob a pressuposição de genes de ação e efeitos aditivos, considerando, ainda, a ausência de seleção. Os dados fenotípicos podem ser descritos a partir do modelo misto geral da forma (Henderson et al., 1959):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e} , \quad (5)$$

em que:

\mathbf{y} : vetor de observações ou fenótipos, de dimensões $n \times 1$;

\mathbf{X} : matriz do modelo referente aos efeitos fixos $\boldsymbol{\beta}$, de dimensões $n \times p$;

$\boldsymbol{\beta}$: vetor dos efeitos fixos, de dimensões $p \times 1$;

\mathbf{Z} : matriz do modelo referente aos efeitos aleatórios \mathbf{a} , de dimensões $n \times q$;

\mathbf{a} : vetor dos efeitos aleatórios ou valores genéticos aditivos (*VGA*) das progênies, de dimensões $q \times 1$,

\mathbf{e} : vetor de erros, de dimensões $n \times 1$, com $\mathbf{e} \sim N_n(\mathbf{0}, R)$, sendo $R = I\sigma_e^2$.

Para o modelo (5), assume-se que os efeitos aleatórios de progênies são normalmente distribuídos, ou seja, $\mathbf{a} \sim N_q(0, G)$, sendo $G = A\sigma_a^2$ é a matriz de covariâncias genéticas aditivas dos *VGA* das progênies. A matriz A corresponde à matriz de parentesco genético aditivo entre progênies, também denominada, no melhoramento animal, como *numerator relationship matrix*. O componente de variância σ_a^2 denota, neste exemplo, a variância genética aditiva na F_2 .

O elemento a_{ij} da matriz A corresponde ao parentesco genético aditivo entre as progênes i e j , equivalente ao coeficiente de parentesco de Wrigth (*relationship coefficient*), sendo igual a duas vezes o coeficiente de parentesco ou de coancestria de Malecot (r_{ij}) (Lynch & Walsh, 1998). O coeficiente de parentesco de Malecot entre i e j corresponde à probabilidade, de um alelo, tomado ao acaso em i ser idêntico por descendência a um outro alelo tomado ao acaso em j (Souza Júnior, 1989), o qual é, geralmente, determinado a partir da informação de genealogia decorrente do método de condução, como exemplo, no uso do método genealógico (item 2.1.3.1). Entretanto, na ausência desses registros, o parentesco pode também ser determinado por meio de informação gerada por marcadores moleculares (Bernardo, 1994; Bernardo, 2002; Lynch & Ritland, 1999). Bernardo (1994) comparou os resultados das predições *BLUP* obtidas por meio de informação de genealogia e com o uso de dados de marcadores moleculares *RFLP* de dois grupos de linhagens de milho. De acordo com o autor, não foram verificadas alterações significativas pelos dois procedimentos.

Considerando um esquema de autofecundações sucessivas para o caso de população oriunda de cruzamentos entre linhagens completamente endogâmicas ($I=1,0$), em que as frequências dos alelos segregantes são iguais a 0,5, tem-se que a covariância genética entre os genótipos nas gerações g e g' relacionados a um ancestral comum t é dada pela seguinte expressão (Bernardo, 2002; Souza Júnior, 1989):

$$Cov(t, g, g') = 2r_{gg'}\sigma_A^2 + \mu_{gg'}\sigma_D^2 + (2r_{gg'})^2\sigma_{AA}^2 + 2r_{gg'}\mu_{gg'}\sigma_{AD}^2 + \dots, \quad (6)$$

em que:

$2r_{gg'}$: coeficiente de parentesco genético aditivo entre g e g' , sendo $r_{gg'}$ o coeficiente de parentesco de Malecot entre g e g' ;

$\mu_{gg'}$: coeficiente de parentesco de dominância entre g e g' , também chamado de coeficiente de fraternidade (Lynch & Walsh, 1998). Este coeficiente refere-se à

probabilidade de g e g' serem heterozigotos para um dado loco e os dois alelos deste loco em g e g' serem idênticos por descendência (Souza Júnior, 1989);

σ_{AA}^2 : variância genética epistática do tipo aditiva x aditiva;

σ_{AD}^2 : variância genética epistática do tipo aditiva x dominante;

Os estimadores de $2r_{gg'}$ e $\mu_{gg'}$ são denotados pelas seguintes expressões (Souza Júnior, 1989):

$$2r_{gg'} = I + I_i;$$
$$\mu_{gg'} = (I + I_i)(I - I_g)(I - I_{g'}) / (I - I_i).$$

Portanto, por ocasião da análise, informações sobre a similaridade genética existentes podem ser incorporadas ao modelo (5), permitindo a obtenção de resultados com maior acurácia (pela consideração da informação de parentesco) e não viciados (pela consideração da correlação entre os valores genéticos na estimação dos efeitos fixos) (Resende, 2002a). Salienta-se que, pelo método dos quadrados mínimos, estas informações são desconsideradas. Neste âmbito, um estudo foi realizado por Bromley et al. (2000) com o intuito de investigar se a inclusão ou não da informação de parentesco entre linhagens de milho afetava as estimativas dos componentes de variância. Os autores observaram que houve um acréscimo no componente de variância referente a linhagens quando a informação foi considerada, indicando que o ganho potencial com a seleção mediante inclusão dessa informação seria maior do que o predito ignorando-a.

No modelo (5) é assumido, ainda, não existir covariância entre os desvios residuais (e) e efeitos genéticos (a), ou seja, $Cov(a, e) = \mathbf{0}$. A partir disso, Henderson (1950), citado por Henderson (1975), derivou o sistema de

equações de modelo misto (*EMM*) que permitiu obter simultaneamente o preditor para os efeitos aleatórios ($\hat{\mathbf{a}}$) e o estimador dos efeitos fixos ($\hat{\boldsymbol{\beta}}$):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\sigma_e^2/\sigma_a^2 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} X'\mathbf{y} \\ Z'\mathbf{y} \end{bmatrix}. \quad (7)$$

Henderson et al. (1959) mostraram que o estimador dos efeitos fixos ($\hat{\boldsymbol{\beta}}$), em (7), é o melhor estimador linear não tendencioso de $\boldsymbol{\beta}$, ou seja, *BLUE*($\hat{\boldsymbol{\beta}}$), sendo igual à solução de quadrados mínimos generalizados, dada por:

$$BLUE(\hat{\boldsymbol{\beta}}) = (X'V^{-1}X)^{-1}X'V^{-1}\mathbf{y}, \quad (8)$$

em que: $V = V(\mathbf{y}) = ZGZ' + R$.

Em seguida, Henderson (1963), citado por Henderson (1975), provou que o preditor ($\hat{\mathbf{a}}$), em (7), corresponde ao *BLUP* ($\hat{\mathbf{a}}$) dado por:

$$BLUP(\hat{\mathbf{a}}) = GZ'V^{-1}(\mathbf{y} - X\hat{\boldsymbol{\beta}}), \quad (9)$$

sendo: $\hat{\boldsymbol{\beta}}$ igual a (8) e $GZ' = Cov(\mathbf{y}, \mathbf{a})$.

Observa-se, pelo preditor (9), que as predições dos efeitos aleatórios ou *VGA* são ajustadas para os efeitos não genéticos presentes (efeitos fixos) e que o termo $GZ'V^{-1}$ promove o chamado efeito de encolhimento (do inglês *shrinkage*) destas predições, fazendo com que estas estejam menos dispersas em torno da $E(\mathbf{a})$ (Robinson, 1991). Sob condições de dados balanceados e na ausência de informação de parentesco, Kennedy & Sorensen (1988), citados por Bernardo (2002), mostraram que o *BLUP* ($\hat{\mathbf{a}}$) (7) é determinado por:

$$BLUP(\hat{\mathbf{a}}) = H(\mathbf{y} - X\hat{\boldsymbol{\beta}}), \quad (10)$$

em que H é uma matriz de ordem $q \times n$, cujos elementos não nulos, relativos à incidência dos níveis do fator aleatório, correspondem às contribuições das parcelas individuais na herdabilidade para seleção entre médias de progênies. Logo, é facilmente visto que, sob as condições supracitadas, o ranqueamento

realizado por meio do preditor (10) será o mesmo obtido com as médias fenotípicas.

A obtenção das soluções $\hat{\beta}$ e \hat{a} por meio de (7) é, em geral, computacionalmente mais vantajosa, pois, para obtê-las por (8) e (9), é necessário inverter a matriz V , de dimensões $n \times n$, enquanto que, por (7), a dimensão da matriz a ser invertida é normalmente grandemente reduzida, da ordem $(p+q) \times (p+q)$ (Henderson, 1975).

Uma questão importante para que (8) e (9) detenham a propriedade de mínimo erro quadrático médio é que os componentes de variância genéticos e não genéticos sejam conhecidos. Na prática, estes componentes são desconhecidos, tendo, portanto, de serem estimados por algum método de estimação (Robinson, 1991). Estudo realizado por Kackar & Harville (1981), citados por Lynch & Walsh (1998), mostrou que as predições *BLUP* permaneceram não viesadas quando estimativas dos componentes foram utilizadas em lugar dos valores reais. Porém, os autores comentam que o *BLUP* obtido pode não ser o melhor dentre os preditores lineares não viciados. Desse modo, para que a seleção pelo *BLUP* seja eficiente, é necessário, primeiramente, que boas estimativas dos componentes de variância estejam disponíveis (Bueno Filho, 1997). Nesta situação, tem-se que os termos *BLUE* e *BLUP* não seriam apropriados, sendo substituídos pelos termos *EBLUE* e *EBLUP* (*BLUE* e *BLUP* empíricos), para denotar o uso de estimativas para os componentes (Littell et al., 1996).

Existem vários métodos de estimação de componentes de variância que podem ser utilizados (Lynch & Walsh, 1998; Resende, 2002a; Searle et al., 1992). Um método que tem sido rotineiramente utilizado em modelos mistos é o método da máxima verossimilhança restrita (*REML*), desenvolvido por Patterson e Thompson, em 1971 (Resende, 2002a). Os estimadores *REML* apresentam

boas propriedades em relação aos demais procedimentos existentes (Harville, 1977; Lynch & Walsh, 1998; Resende, 2002a).

Variados algoritmos computacionais iterativos têm sido utilizados para a obtenção de estimadores *REML*, dentre os quais os mais usados são: o algoritmo *EM* (*Expectation-Maximization*), *DF-REML* (*Derivative-Free Restricted Maximum Likelihood*) (Resende, 2002a) e Newton-Raphson (Littel et al., 1996).

Considerando as *EMM* em (7), o algoritmo *EM* conduz aos seguintes estimadores de σ_a^2 e σ_e^2 (Resende, 2002a):

$$\hat{\sigma}_e^2 = \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \hat{\mathbf{a}}'\mathbf{Z}'\mathbf{y}}{n - r(\mathbf{X})}; \quad \hat{\sigma}_a^2 = \frac{\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} + \hat{\sigma}_e^2 \text{tr}(\mathbf{A}^{-1}\mathbf{C}^{22})}{q}, \quad (11)$$

em que:

$r(\mathbf{X})$: posto da matriz \mathbf{X} ;

$\text{tr}(\cdot)$: traço da matriz;

$$\mathbf{C}^{22} \text{ é dado por: } \mathbf{C} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\sigma_e^2/\sigma_a^2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C}^{11} & \mathbf{C}^{12} \\ \mathbf{C}^{21} & \mathbf{C}^{22} \end{bmatrix}$$

Assim, fica evidente que o algoritmo *EM* requer valores arbitrários iniciais para (11) para obter as soluções (8) e (9) por meio das *EMM* (7). Estas soluções são usadas para se obter novas estimativas de (11) e, assim por diante, até que a convergência seja alcançada.

Segundo Bernardo (2002), o procedimento *BLUP* é útil por duas razões: permite a análise de dados desbalanceados, retornando, nestas condições, predições mais confiáveis do que as obtidas pelo *MQM*; explora a informação entre indivíduos aparentados, conforme já comentado, maximizando o uso dos dados disponíveis na comparação entre os genótipos. Além disso, Bernardo (1994), Bernardo (1995) e Bernardo (1996) comentam sobre a utilidade do procedimento *BLUP* para a predição de valores genéticos de cruzamentos não realizados em delineamentos dialélicos. Para uma abordagem mais detalhada sobre este último aspecto, recomenda-se a leitura de Silva Filho (2004).

No que concerne ao melhoramento vegetal, a eficiência do *BLUP* tem sido comparada em relação a outros métodos, como o método dos quadrados mínimos ordinários (*MQM*). Panter & Allen (1995a) estudaram a eficiência dos *MQM* e *BLUP* na identificação de linhagens superiores em cruzamentos de linhagens de soja. Os autores simularam situações de balanceamento e desbalanceamento de dados. Em todas as situações, as previsões *BLUP* apresentaram menor erro padrão e maiores valores de correlação de Spearman entre os valores preditos e o desempenho dos cruzamentos avaliados em campo, bem como maior porcentagem de cruzamentos superiores dentre os realizados.

Ao comparar dois modelos de *BLUP* (com e sem a inclusão da informação de parentesco genético das linhagens utilizadas no cruzamento) com o *MQM* para a predição de cruzamentos em soja, Panter & Allen (1995b) observaram que os erros padrões das estimativas obtidas pelo *MQM* foram, em média, 6% maiores do que as obtidas com os modelos de *BLUP*. Contudo, não foram observadas diferenças marcantes entre os modelos *BLUP*, sendo o modelo com informação de parentesco ligeiramente superior.

Um estudo via simulação computacional visando comparar os métodos *MQM* e *BLUP*, considerando informações de similaridade genética por meio de marcadores moleculares em modelos dialélicos utilizando os métodos II e IV de Griffing, foi conduzido por André (1999). O autor concluiu pela superioridade do *BLUP* em relação ao método *MQM* quanto à predição de efeitos de cruzamentos e de capacidades de combinação, em diferentes condições de herdabilidade e de grau médio de dominância.

3 MATERIAL E MÉTODOS

3.1 Simulação computacional

Para a realização deste estudo foi implementado um programa computacional em ambiente Delphi 6.0 (Cantú, 2002). Foram simuladas situações comuns ao melhoramento de plantas autógamas, com ênfase à cultura do feijoeiro, com a aplicação de técnicas estatísticas pela abordagem de modelos lineares mistos.

3.1.1 Modelo genético e configurações estudadas

Assumiu-se um modelo genético simplificado para um caráter quantitativo qualquer controlado por 20 locos de segregação independente ($L=20$), com efeitos iguais e aditivos, e com frequência alélica 0,5. As configurações estudadas consideraram diferentes valores dos parâmetros herdabilidade no sentido amplo na média das progênies (h_a^2), de 10%, 25%, 50% e 75% e grau médio de dominância (g.m.d.), de 0,0, 0,5 e 1,0.

3.1.2 Avanço das populações segregantes e geração dos genótipos

Para cada configuração e método de condução avaliado na simulação foram geradas 1.000 populações F_2 com L locos segregantes. Cada população F_2 foi constituída por 64 plantas. O avanço dessas populações foi efetuado, independentemente, pelos métodos genealógico e *bulk*, considerando plantas

com igual taxa de multiplicação, ou seja, de cada planta foram gerados 40 descendentes, tendo como base o feijoeiro.

3.1.2.1 Método genealógico ou *pedigree*

As 64 plantas de cada população F_2 simulada deram origem a 64 progênes $F_{2:3}$ constituídas por 40 plantas cada. A partir de cada progênie $F_{2:3}$ foram tomadas aleatoriamente duas plantas, gerando 128 progênes $F_{3:4}$. O processo foi repetido na geração seguinte, de modo a obter 256 progênes $F_{4:5}$ contendo 40 plantas cada (Figura 1).

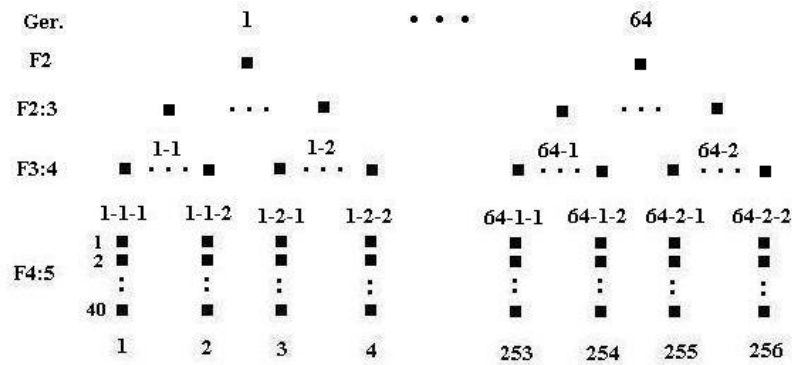


FIGURA 1 - Esquema de condução pelo método genealógico.

3.1.2.2 Método *bulk*

As 64 plantas F_2 simuladas geraram uma população F_3 de 2.560 plantas. Na geração F_3 foram tomadas aleatoriamente 128 plantas das quais foram considerados 40 descendentes, totalizando 5.120 plantas na geração F_4 . Por

ocasião dessa geração (F_4), foram retiradas, ao acaso, 256 plantas que deram origem às 256 progênies $F_{4:5}$, constando de 40 plantas cada (Figura 2).

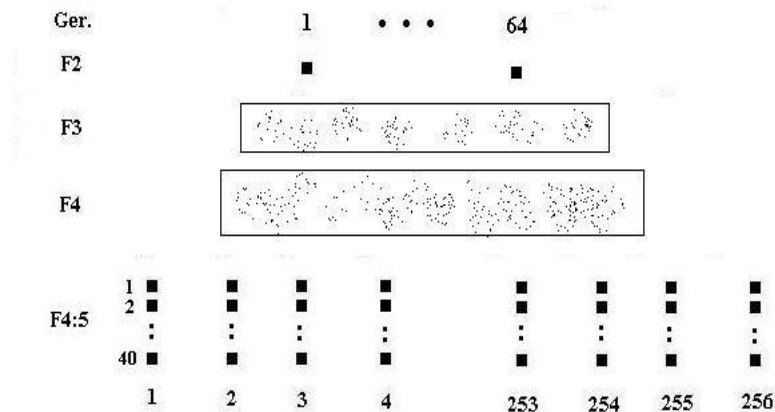


FIGURA 2 - Esquema de condução pelo método *bulk*.

3.1.3 Geração dos fenótipos

Os fenótipos das plantas foram obtidos adicionando-se erros aleatórios aos valores genotípicos (VG). Os VG foram atribuídos segundo o modelo genético descrito no item 3.1.1, partindo-se de uma constante 100. O efeito aditivo (a_l) do l -ésimo loco foi assumido igual a 1,0 e o efeito de dominância (d_l) foi determinado em função do g.m.d., ou seja, $d_l = g.m.d. \times a_l$, em que $l=1,2,\dots,L$. Assim, o VG da planta k (g_k) é dado por:

$$g_k = \sum_{l=1}^L \delta_l, \text{ em que } \delta_l = \begin{cases} 1 & \text{se } AA \\ d_l & \text{se } Aa \\ -1 & \text{se } aa \end{cases}$$

Simularam-se os valores fenotípicos para parcelas de vinte plantas ($n=20$), de forma que o número de repetições foi igual a dois ($r=2$). O valor fenotípico da planta k pertencente à progênie i na repetição j foi estabelecido pelo seguinte modelo:

$$y_{ijk} = \mu + g_{ijk} + e_{ij} + w_{ijk},$$

em que:

μ : constante, no caso 100;

g_{ijk} : valor genotípico da planta ijk ($i=1,2,\dots,256; j=1,2; k=1,2,\dots,20$);

e_{ij} : desvio ambiental entre parcelas, sendo $e_{ij} \sim N(0, \sigma_e^2)$;

w_{ijk} : desvio ambiental aleatório dentro da parcela, associado a y_{ijk} , sendo $w_{ijk} \sim N(0, \sigma_w^2)$.

Os valores de e_{ij} e w_{ijk} foram atribuídos aleatoriamente a partir da distribuição normal. Para estipular os valores de e_{ij} e w_{ijk} , foi considerada a relação σ_w^2/σ_e^2 fixada em oito ($c=8$), em que σ_e^2 é a variância ambiental entre parcelas e σ_w^2 é a variância ambiental dentro da parcela. Os erros variaram conforme valores assumidos para a herdabilidade h_a^2 .

A expressão da herdabilidade h_a^2 , nesse caso, é dada por:

$$h_a^2 = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2/r + \sigma_d^2/nr},$$

em que:

σ_p^2 : variância genética entre progênes F_{4;5}, ou seja, $\sigma_p^2 = 7/4 \sigma_A^2 + 7/64 \sigma_D^2$,

sendo que σ_A^2 é a variância aditiva na F₂, no caso $\sigma_A^2 = L/2$ e σ_D^2 é a variância

de dominância na F₂, no caso $\sigma_D^2 = [L(g.m.d. \times a)^2]/4$;

σ_d^2 : variância fenotípica dentro de parcela, isto é, $\sigma_d^2 = \sigma_{Gd}^2 + \sigma_w^2$, sendo que σ_{Gd}^2 é a variância genética dentro de progênie F_{4:5}, ou seja, $\sigma_{Gd}^2 = 1/8 \sigma_A^2 + 1/8 \sigma_D^2$ e σ_w^2 é a variância ambiental dentro da parcela, que pode ser obtida por:

$$\sigma_w^2 = [(1 - h_{F_2}^2) / h_{F_2}^2] \sigma_G^2,$$

sendo:

σ_G^2 : variância genética na F₂, ou seja, $\sigma_G^2 = \sigma_A^2 + \sigma_D^2$;

$h_{F_2}^2$: herdabilidade individual na geração F₂, que pode ser determinada em função dos valores pré-fixados da herdabilidade h_a^2 por:

$$h_{F_2}^2 = \sigma_G^2 (1 + n/c) / [(nr \sigma_p^2 (1 - h_a^2) / h_a^2) - \sigma_{Gd}^2 + \sigma_G^2 (1 + n/c)].$$

3.1.4 Análise estatística

Foram analisados 1.000 experimentos referentes à avaliação das 256 progênie F_{4:5} simuladas, provenientes de ambos os métodos de condução, para cada configuração estudada. A análise foi realizada com base nos dados fenotípicos médios das parcelas, considerando o delineamento experimental inteiramente casualizado, conforme item 3.1.3, com duas repetições.

A abordagem estatística das análises foi baseada na teoria de modelos mistos (Henderson et al., 1959), com estimação dos componentes de variância pelo método da máxima verossimilhança restrita (REML) (Patterson & Thompson, 1971).

Conforme descrição constante no item 3.1.2.1 para a condução pelo método genealógico, cada planta F₂ gerou, na F₅, quatro progênie F_{4:5} (Figura 1). Com base nesta genealogia detalhada, foi determinada a similaridade genética entre as progênie relacionadas a partir do cálculo dos coeficientes de parentesco de Malecot por meio do programa Rxy (Ferreira & Zambalde, 1997),

considerando a população F_2 como não endogâmica. Com isso, foram obtidos os coeficientes de parentesco aditivo entre as progênes, os quais correspondem a duas vezes o coeficiente de parentesco de Malecot (6). Logo, os dados fenotípicos das progênes oriundas do método genealógico foram analisados de acordo com dois modelos: o primeiro em que a informação de parentesco aditivo é ignorada – modelo G , e o segundo em que esta informação do parentesco aditivo é levada em consideração – modelo aditivo (G_A). Para o método *bulk*, por não se dispor da informação da genealogia entre as progênes $F_{4,5}$, a análise foi realizada de forma semelhante ao método genealógico, utilizando o modelo sem parentesco – modelo G .

Os dados fenotípicos médios das 256 progênes $F_{4,5}$ foram analisados a partir do modelo misto geral da forma:

$$y = X\beta + Za + e, \quad (12)$$

em que:

y : vetor de dados fenotípicos médios, de dimensões 512×1 ;

X : vetor do modelo referente ao efeito fixo da constante β , de dimensões 512×1 ;

β : efeito fixo da constante;

Z : matriz do modelo referente aos efeitos aleatórios de progênes a , de dimensões 512×256 ;

a : vetor dos efeitos aleatórios das progênes, de dimensões 256×1 , com $a \sim N(0, G)$, sendo $G = A\sigma_a^2$;

e : vetor de erros, de dimensões 512×1 , com $e \sim N(0, R)$, sendo $R = I\sigma_e^2$.

No enfoque de modelos mistos, G refere-se à matriz de covariâncias genéticas entre as progênes, sendo denotada por $A\sigma_a^2$, em que A recebeu diferentes estruturas dependendo do modelo adotado. No caso do modelo sem o

parentesco genético aditivo, a matriz G foi denotada por $I\sigma_p^2$, ou seja, $A=I$, indicando que as progênies foram assumidas como não relacionadas, sendo que σ_a^2 , nesse caso, equívale à variância genética entre progênies $F_{4;5}$ (σ_p^2). Já para o modelo considerando a inclusão do parentesco genético aditivo, a matriz G foi denotada por $A\sigma_A^2$, com $A \neq I$, ou seja, A contendo os coeficientes de parentesco aditivo entre as progênies e σ_a^2 refere-se à variância aditiva na F_2 (σ_A^2). Neste caso a matriz A foi dada por:

$$A = I_{64} \otimes \begin{bmatrix} 1,75 & 1,50 & 1,00 & 1,00 \\ 1,50 & 1,75 & 1,00 & 1,00 \\ 1,00 & 1,00 & 1,75 & 1,50 \\ 1,00 & 1,00 & 1,50 & 1,75 \end{bmatrix}, \quad (13)$$

em que \otimes é o produto de Kroneker.

As soluções para os efeitos aleatórios (\hat{a}) e fixos ($\hat{\beta}$) do modelo misto (12) foram obtidas pela resolução do seguinte sistema de equações de Henderson (1975):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1} \hat{\sigma}_e^2 / \hat{\sigma}_a^2 \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}. \quad (14)$$

Para a obtenção das soluções anteriores, os componentes de variância genéticos e não genéticos foram assumidos desconhecidos, como ocorre na prática, sendo estimados pelo método da máxima verossimilhança restrita (*REML*). Como o método *REML* é um processo iterativo, foi então utilizado o algoritmo numérico denominado de *EM* (por alternar passos de esperança e maximização), caracterizando o algoritmo como *EM-REML*. Os estimadores *EM-REML* dos componentes $\hat{\sigma}_a^2$ e $\hat{\sigma}_e^2$ são dados por:

$$\hat{\sigma}_e^2 = \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \hat{\mathbf{a}}'\mathbf{Z}'\mathbf{y}}{n - r(\mathbf{X})}; \hat{\sigma}_a^2 = \frac{\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} + \hat{\sigma}_e^2 \text{tr}(\mathbf{A}^{-1}\mathbf{C}^{22})}{q}, \quad (15)$$

em que:

n : número de observações, no caso 512;

q : número de progênies, no caso 256;

$r(\mathbf{X})$: posto da matriz \mathbf{X} ;

$\text{tr}(\cdot)$: operador traço da matriz;

\mathbf{C}^{22} é uma submatriz da matriz \mathbf{C} definida por:

$$\mathbf{C} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\sigma_e^2/\sigma_a^2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C}^{11} & \mathbf{C}^{12} \\ \mathbf{C}^{21} & \mathbf{C}^{22} \end{bmatrix}.$$

A partir de valores arbitrários iniciais para $\hat{\sigma}_a^2$ e $\hat{\sigma}_e^2$ em (14) são obtidas as soluções $\hat{\boldsymbol{\beta}}$ e $\hat{\mathbf{a}}$. Estas soluções são usadas para se obter novas estimativas de $\hat{\sigma}_a^2$ e $\hat{\sigma}_e^2$ (15) e, assim por diante, até que a convergência seja alcançada. Adicionalmente foi implementado na rotina de análise o acelerador de convergência de Aitken, conforme descrito em Conte (1971).

De acordo com o modelo adotado (G ou G_A), diferentes significados são apregeados às predições ($\hat{\mathbf{a}}$) dos efeitos aleatórios (\mathbf{a}), visto que o melhor preditor linear não-viesado ($BLUP$) é especificado por:

$$BLUP(\hat{\mathbf{a}}) = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}),$$

ou seja, depende da matriz $\mathbf{G} = \mathbf{A}\sigma_a^2$. A matriz \mathbf{V} refere-se à matriz de covariâncias dos dados, sendo denotada por $\mathbf{V}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$. Assim, concernente ao modelo sem parentesco, o vetor $\hat{\mathbf{a}}$ corresponde às predições dos valores genotípicos das progênies ($BLUP$), enquanto que, para o modelo considerando o parentesco genético aditivo, o vetor $\hat{\mathbf{a}}$ corresponde às predições dos valores genéticos aditivos das progênies ($BLUP_A$). Vale salientar que, adicionalmente para cada experimento simulado, em ambos os métodos de condução, foram também obtidas as médias fenotípicas (M).

Foram estimadas, para cada experimento simulado, as herdabilidades no sentido amplo na média das progênies (\hat{h}_a^2) no modelo sem parentesco em ambos os métodos de condução, e a herdabilidade no sentido restrito na média das progênies (\hat{h}_r^2) no modelo com parentesco aditivo, apenas no método genealógico, como já mencionado.

Os procedimentos de seleção das progênies F_{4,5} foram avaliados e comparados a partir do conhecimento dos valores genotípicos verdadeiros correspondentes. Foram estimadas, para cada experimento simulado, em ambos os métodos de condução, as correlações de Spearman, as proporções de coincidências nas frações selecionadas de 5%, 10% e 25%, em ambos os extremos e os valores genotípicos médios das progênies selecionadas pelos diferentes procedimentos nas proporções selecionadas de 0,4% (melhor progênie), 5%, 10% e 25%.

Foram calculados os valores médios das estimativas dos parâmetros anteriormente citados, como também os desvios padrões associados a essas estimativas ao longo dos 1.000 experimentos simulados para cada configuração estudada.

3.2 Experimentos de campo

3.2.1 Local

Os experimentos foram conduzidos no município de Lavras, na área experimental do Departamento de Biologia da Universidade Federal de Lavras, localizada na região Sul do estado de Minas Gerais, a 918 metros de altitude, 21° 14' S de latitude e 40° 17' W de longitude.

3.2.2 Cultivares

Os genitores utilizados na obtenção da população segregante de feijoeiro foram as cultivares BRS Talismã, origem na UFLA, com hábito de crescimento indeterminado tipo III, tipo de grão carioca, com semente de coloração creme com rajadas marrom-claras e resistente à antracnose (patótipos 65, 67, 73, 75, 81, 83, 89, 95, 117), ao mosaico comum e resistência intermediária à mancha angular; e BRS Valente, origem na Embrapa, com hábito de crescimento indeterminado tipo II, tipo de grão preto, resistente à antracnose (patótipos 89, 95, 453 e 585), mosaico comum e resistência intermediária à mancha angular.

3.2.3 Obtenção da população segregante e avanço das gerações pelos métodos genealógico e *bulk*

As sementes da geração F_1 foram obtidas a partir do cruzamento artificial entre os genitores ‘BRS Talismã’ e ‘BRS Valente’ (item 3.2.1), em casa de vegetação, de forma análoga à descrita por Ramalho et al. (1993). As plantas da geração F_1 e F_2 foram obtidas sob condições de campo. O avanço das demais gerações até a obtenção das progênies $F_{4,5}$ foi efetuado segundo preconizado para cada método de condução, de maneira semelhante à descrita no item 3.1.2.1 para o método genealógico (Figura 1) e item 3.1.2.2 para o método *bulk* (Figura 2). Em adição, as sementes de cada progênie $F_{4,5}$, em ambos os métodos de condução, foram colhidas e misturadas, tomando-se uma amostra ao acaso para gerar as progênies $F_{4,6}$.

Adicionalmente, durante o processo de condução da população segregante pelo método *pedigree*, foi realizada a anotação dos registros genealógicos para fins de determinação do parentesco entre as progênies nas gerações $F_{4,5}$ e $F_{4,6}$, considerando a geração F_2 como não endogâmica ($I_2=0$).

3.2.4 Avaliação das progênies

3.2.4.1 Delineamento experimental e condução dos experimentos

A avaliação das 256 progênies $F_{4,5}$ oriundas dos métodos genealógico e *bulk* foi realizada em experimentos contíguos na safra ‘das águas’ de 2004/05. O delineamento experimental empregado foi o látice simples 16 x 16. A parcela experimental foi constituída por uma linha de um metro de comprimento, espaçadas em 50 cm, com densidade de plantio de 20 sementes por metro.

Os experimentos foram implantados sob área de plantio direto, sendo a adubação de plantio equivalente a 300 kg/ha do formulado 8-28-16 de N, P_2O_5 e K_2O . Foi procedida a adubação de cobertura com 150 kg/ha de sulfato de amônia aos 25 dias após a semeadura. O controle das plantas invasoras foi efetuado por meio de herbicida de pós-emergência. Irrigação complementar por aspersão foi realizada quando necessário.

De forma análoga, foram conduzidos os experimentos referentes à avaliação das 256 progênies $F_{4,6}$ oriundas dos métodos genealógico e *bulk* na safra ‘das secas’ de 2005. O delineamento experimental empregado foi também o látice simples 16 x 16, entretanto, a parcela experimental, nesse caso, foi constituída por uma linha de dois metros de comprimento, espaçadas em 50 cm, com densidade de plantio de 15 sementes por metro. O manejo da cultura nesses experimentos da geração $F_{4,6}$ foi semelhante ao adotado para a geração $F_{4,5}$.

3.2.4.2 Características avaliadas

O caráter produção de grãos, em gramas por parcela, foi avaliado nas progênies provenientes de ambos os métodos de condução, enquanto o porte das

plantas foi somente avaliado nas progênies oriundas do método genealógico. A avaliação do porte foi feita na média da parcela, por meio uma escala de notas (variando a 1 a 5) utilizada por Collicchio (1995): nota 1 – planta de hábito de crescimento II, ereta, com uma haste e com inserção alta das primeiras vagens; nota 2 - planta de hábito de crescimento II, ereta e com algumas ramificações; nota 3 - planta de hábito de crescimento II ou III, ereta, com muitas ramificações e tendência a prostrar-se; nota 4 - planta de hábito de crescimento III, semi-ereta e medianamente prostada; nota 5 - planta de hábito de crescimento III, com entrenós longos e muito prostrada.

3.2.4.3 Análise estatística

Os dados fenotípicos dos experimentos de avaliação das progênies F_{4:5} e F_{4:6}, em ambos os métodos de condução, foram analisados pela abordagem de modelos mistos à semelhança do descrito no item 3.1.4, conforme o seguinte modelo estatístico:

$$y_{ijk} = \mu + r_j + b_{k(j)} + a_i + e_{ijk} \quad (16)$$

em que:

y_{ijk} : observação da parcela que recebeu a progênie i no bloco k dentro da repetição j ($i=1,2,\dots,256; j=1,2; k=1,2,\dots,16$);

μ : constante associada a todas as observações;

r_j : efeito fixo da repetição j ;

$b_{k(j)}$: efeito aleatório do bloco k dentro da repetição j ;

a_i : efeito aleatório da progênie i ;

e_{ijk} : erro experimental aleatório associado à observação y_{ijk} .

O modelo estatístico misto (16) pode ser escrito na forma matricial por:

$$y = X\beta + Z_1b + Z_2a + e, \quad (17)$$

em que:

y : vetor de dados, de dimensões 512×1 ;

X : matriz do modelo referente aos efeitos fixos β , de dimensões 512×3 ;

β : vetor dos efeitos fixos, de dimensões 3×1 ;

Z_1 : matriz do modelo referente aos efeitos aleatórios de blocos dentro de repetições b , de dimensões 512×32 ;

b : vetor dos efeitos aleatórios de blocos dentro de repetições, de dimensões 32×1 , com $b \sim N(0, I\sigma_b^2)$;

Z_2 : matriz do modelo referente aos efeitos aleatórios de progênies a , de dimensões 512×256 ;

a : vetor dos efeitos aleatórios das progênies, de dimensões 256×1 , com $a \sim N(0, G)$, sendo $G = A\sigma_a^2$;

e : vetor de erros, de dimensões 512×1 , com $e \sim N(0, R)$, sendo $R = I\sigma_e^2$.

As análises dos dados fenotípicos das progênies $F_{4,5}$, provenientes dos métodos genealógico e *bulk*, foram realizadas de forma análoga à descrita no item 3.1.4. Para os dados da geração $F_{4,6}$, o modelo com a inclusão do parentesco genético aditivo (G_A) considerou a seguinte matriz A :

$$A = I_{64} \otimes \begin{bmatrix} 1,875 & 1,50 & 1,00 & 1,00 \\ 1,50 & 1,875 & 1,00 & 1,00 \\ 1,00 & 1,00 & 1,875 & 1,50 \\ 1,00 & 1,00 & 1,50 & 1,875 \end{bmatrix}.$$

Adicionalmente, foram realizadas as análises dos dados experimentais considerando apenas a recuperação da informação interblocos (blocos aleatórios e progênies fixas).

As análises estatísticas pela abordagem de modelos mistos foram realizadas por meio do procedimento *Mixed* do programa SAS 8.0 (Little et al., 1996; SAS, 1999). A rotina para a análise com recuperação da informação

interblocos e intergenotípica para o modelo estatístico (16), considerando o parentesco genético aditivo entre as progênes, adaptada de Balzarini & Milligan (2003), é apresentada em Anexo.

3.2.4.4 Estimação de parâmetros genéticos

O método especificado para a estimação dos componentes de variância foi o da máxima verossimilhança restrita (*REML*). As estimativas dos componentes de variância foram testadas por meio do teste *Z* de Wald, a 5% de probabilidade, sendo a estatística do teste dada por:

$$Z = \hat{\sigma}_i^2 / s(\hat{\sigma}_i^2),$$

em que: $s(\hat{\sigma}_i^2)$ é a estimativa do erro padrão assintótico associado ao componente $\hat{\sigma}_i^2$ obtida a partir da inversa da matriz de informação de Fisher (Searle et al., 1992, p. 284).

Os intervalos de confiança (*IC*) associados às estimativas dos componentes de variância foram estimados a partir da expressão (Ramalho et al., 2005):

$$IC(\sigma^2) : P\left[v\hat{\sigma}_i^2 / \chi_{(v;1-\alpha/2)}^2 < \sigma_i^2 < v\hat{\sigma}_i^2 / \chi_{(v;\alpha/2)}^2\right] = 1 - \alpha,$$

em que:

α : nível de significância pré-estabelecido, no caso $\alpha=0,05$;

v : graus de liberdade associados ao componente $\hat{\sigma}_i^2$ obtidos pela aproximação de Satterthwaite (1946), conforme expressão denotada por: $v = 2Z^2$ (SAS, 1999);

$\chi_{(v;1-\alpha/2)}^2$; $\chi_{(v;\alpha/2)}^2$: quantis da distribuição de qui-quadrado χ^2 para v graus de liberdade;

Foram estimadas, para cada experimento, as herdabilidades no sentido amplo na média das progênes (\hat{h}_a^2) para o modelo sem parentesco em ambos os métodos de condução pelo seguinte estimador:

$$\hat{h}_a^2 = \hat{\sigma}_p^2 / (\hat{\sigma}_p^2 + \hat{\sigma}_e^2 / r).$$

A herdabilidade no sentido restrito na média das progênes (\hat{h}_r^2) no modelo com parentesco aditivo (G_A), apenas, no método genealógico, como já mencionado, foi obtida pela seguinte expressão:

$$\hat{h}_r^2 = 7/4 \hat{\sigma}_A^2 / (7/4 \hat{\sigma}_A^2 + \hat{\sigma}^2 / r).$$

3.2.4.5 Estimação das correlações de Spearman e Pearson, proporções de coincidências e ganhos realizados com a seleção

Foram estimadas para cada experimento, em ambos os métodos de condução, as correlações de Spearman e Pearson, e proporções de coincidências nas frações selecionadas de 1%, 5%, 10% e 25%, em ambos os extremos, para os procedimentos de seleção: média ajustada (*MAj*), *BLUP* sem parentesco e *BLUP* com parentesco (*BLUP_A*).

Em adição, foram também calculados, para cada caráter, em ambos os métodos de condução, os ganhos percentuais realizados (*GR*) na geração $F_{4:6}$, com base na seleção das progênes $F_{4:5}$, pelos diferentes procedimentos nas proporções selecionadas de 1%, 5%, 10% e 25%, pela seguinte expressão:

$$GR(\%) = \left[\frac{(MAj F_{4:6}(s) - MG F_{4:6})}{MG F_{4:6}} \right] \times 100,$$

em que:

MAjF_{4:6}(s): valor médio das correspondentes progênes $F_{4:6}$ selecionadas na $F_{4:5}$ com base num dos procedimentos considerando a proporção selecionada *s*;

MGF_{4:6}: média geral das progênes $F_{4:6}$.

4 RESULTADOS E DISCUSSÃO

4.1 Simulação computacional

Não há muitos relatos sobre o grau médio de dominância (g.m.d.) de caracteres em espécies cultivadas, especialmente autógamas. Os existentes preconizam que este não é elevado, normalmente inferior a 1,0, como por exemplo, em feijão (Souza & Ramalho, 1995) e trigo (Novoselovic et al., 2004), não sendo, portanto, relatada a ocorrência de sobredominância no controle genético da maioria dos caracteres. Por essa razão, optou-se, no presente estudo de simulação, por utilizar o g.m.d. variando de 0,0 a 1,0, o que certamente representa a quase totalidade do controle genético das características em plantas cultivadas (Crow, 2000).

Na questão da amplitude de variação para as herdabilidades no sentido amplo na média das progênies pré-fixadas (h_a^2) nas simulações, também procuraram-se valores que estivessem próximos do normalmente observado em experimentos de avaliação de progênies. Tomando-se como referência a cultura do feijoeiro, em experimentos conduzidos no estado de Minas Gerais, tem-se que a maioria das estimativas desse parâmetro (h_a^2) para o caráter produção de grãos situou-se entre 10% e 75% (Moreto, 2005). O mesmo fato tem sido constatado para outras espécies cultivadas, como, por exemplo, no milho (Hallauer & Miranda Filho, 1988). Assim, os valores paramétricos da herdabilidade h_a^2 utilizados nas simulações representam o que é comumente observado sob condições de campo no Brasil.

Quando da realização de simulação, a primeira preocupação que surge é verificar se o programa implementado foi eficaz e, mais ainda, se os resultados refletem o que se espera que ocorra em condições reais (Ferreira, 2001). Os

valores médios obtidos para as estimativas das herdabilidades h_a^2 (\hat{h}_a^2), para ambos os métodos de condução e configurações estudadas, foram próximos dos valores populacionais pré-definidos. Por exemplo, considerando o g.m.d. de 0,0, os valores médios das estimativas \hat{h}_a^2 variaram de 10,8% no *bulk* a 11,12% no genealógico, quando a herdabilidade h_a^2 fixada foi de 10% (Tabela 1).

Outro aspecto que realça a fidedignidade da rotina implementada foi que os desvios padrões associados às estimativas \hat{h}_a^2 , para os métodos de condução avaliados, foram continuamente menores para maiores valores de h_a^2 , em todos os valores de g.m.d. estudados. Tomando-se, como exemplo, o g.m.d. de 0,0, observa-se que, para o caso de herdabilidade h_a^2 baixa (10%), o desvio padrão correspondeu, em média, a 82,91% da estimativa \hat{h}_a^2 no método genealógico e de 82,89% no *bulk*, enquanto que, quando considerada a alta herdabilidade h_a^2 (75%), estas proporções foram de apenas 5,02% no genealógico e de 5,01% no *bulk* (Tabela 1).

Outra observação importante diz respeito aos resultados obtidos para os valores médios das estimativas da variância genética entre progênies ($\hat{\sigma}_p^2$) para os métodos genealógico e *bulk*. Afere-se que as estimativas médias $\hat{\sigma}_p^2$, para ambos os métodos, foram semelhantes (Tabela 1). Neste ponto, convém mencionar que para o método *bulk*, assumiu-se que as plantas simuladas tinham igual taxa de multiplicação (item 3.1.2). Isso implica que a ação da seleção natural não foi considerada durante o processo de condução. Há inúmeros relatos da ação da seleção natural durante a condução de populações segregantes por esse método (Allard, 1988; Gonçalves et al., 2001). Vale salientar que o efeito da seleção natural, nestes trabalhos, foi avaliado por inúmeras gerações. Provavelmente, com duas gerações apenas, como foi o caso dessa simulação

TABELA 1 Valores genotípicos médios das 256 progênes F_{4:5} (VGM) e valores médios estimados para a variância genética entre progênes F_{4:5} ($\hat{\sigma}_p^2$), variância genética aditiva na F₂ ($\hat{\sigma}_A^2$) e herdabilidades no sentido amplo (\hat{h}_a^2) e restrito (\hat{h}_r^2) na média de progênes F_{4:5} para os métodos genealógico (com e sem informação de parentesco) e *bulk*, variando o grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênes (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1000. Valores entre parenteses correspondem aos desvios padrões associados.

g.m.d.	h_a^2 (%)	Genealógico					<i>Bulk</i>		
		Sem parentesco			Com parentesco		VGM	$\hat{\sigma}_p^2$	\hat{h}_a^2
		VGM	$\hat{\sigma}_p^2$	\hat{h}_a^2 (%)	$\hat{\sigma}_A^2$	\hat{h}_r^2 (%)			
0,0	10	100,0(0,5)	20,48(17,90)	11,12(9,22)	10,31(7,17)	10,02(6,58)	100,0(0,5)	19,84(17,34)	10,81(8,96)
0,0	25	100,0(0,4)	17,31(7,68)	24,23(9,43)	10,05(3,37)	24,78(6,85)	100,0(0,5)	17,21(7,99)	24,11(9,79)
0,0	50	100,0(0,4)	17,35(3,88)	49,25(6,96)	9,97(2,12)	49,46(6,30)	99,98(0,5)	17,33(3,98)	49,21(6,93)
0,0	75	100,0(0,5)	17,48(2,80)	74,65(3,75)	9,97(1,50)	74,72(3,60)	100,0(0,5)	17,38(2,68)	74,59(3,74)
0,5	10	100,6(0,5)	20,31(17,51)	11,00(8,99)	10,58(7,08)	10,24(6,45)	100,6(0,5)	19,98(17,85)	10,81(9,17)
0,5	25	100,6(0,4)	17,60(7,82)	24,30(9,43)	10,17(3,50)	24,74(7,04)	100,6(0,5)	17,26(8,04)	23,93(9,71)
0,5	50	100,6(0,5)	17,55(3,89)	49,45(6,83)	10,05(2,12)	49,51(6,18)	100,6(0,5)	17,62(4,02)	49,38(6,95)
0,5	75	100,6(0,4)	17,47(2,88)	74,38(3,81)	9,99(1,57)	74,76(3,70)	100,6(0,5)	17,51(2,84)	74,48(3,58)
1,0	10	101,2(0,5)	20,24(17,25)	10,68(8,66)	10,50(6,98)	9,90(6,24)	101,3(0,5)	20,06(17,82)	10,63(9,02)
1,0	25	101,3(0,5)	17,96(7,98)	24,31(9,48)	10,13(3,55)	24,21(7,10)	101,3(0,5)	18,06(8,45)	24,41(10,00)
1,0	50	101,2(0,5)	17,96(4,02)	49,52(7,04)	10,25(2,17)	49,42(6,29)	101,3(0,5)	17,87(3,93)	49,33(6,71)
1,0	75	101,3(0,5)	17,84(2,72)	74,44(3,57)	10,24(1,52)	74,46(3,55)	101,3(0,5)	17,77(2,87)	74,43(3,75)

(item 3.1.2.2), o efeito da seleção natural não seja pronunciado, o que não invalida os resultados obtidos com a simulação realizada.

Do mesmo modo, foi desconsiderado o possível efeito da seleção visual, como preconizado pelo método genealógico (Borém & Miranda, 2005), na tomada de plantas dentro das progênes nas diferentes gerações (item 3.1.2.1). É oportuno enfatizar que há inúmeros resultados evidenciando a ineficiência da seleção visual para caracteres de herdabilidade baixa, inferiores a 50%, como ocorre na maioria dos caracteres de importância econômica (Cutrim et al., 1997; Silva et al., 1994). Infere-se, assim, que a tomada aleatória das duas plantas para gerar as progênes da geração subsequente, provavelmente, não teve repercussão expressiva sobre os resultados, sobretudo para herdabilidades h_a^2 menores que 50%. Observação semelhante é válida para a seleção entre progênes, a qual é, originalmente, visual no método genealógico. Neste caso, esta não foi efetuada visando à exploração da maior variabilidade possível gerada pelo cruzamento. Tal fato fica evidente, como já mencionado, pelas estimativas $\hat{\sigma}_p^2$ nos dois métodos que foram muito semelhantes. Um outro aspecto a ser ressaltado, é que a tomada de apenas dois indivíduos para representar a progênie não teve maiores conseqüências em termos da variância liberada entre as progênes, em comparação com o método *bulk* (Tabela 1).

No que concerne ao método genealógico, uma de suas principais particularidades é o conhecimento da genealogia existente entre as progênes obtidas e, para a qual, tem sido apregoada pouca utilidade (Ramalho et al., 2001). Uma aplicação para justificar o tempo despendido na anotação dos registros genealógicos seria sua utilização para determinar os parentescos entre os indivíduos para fins de análise pela abordagem de modelos mistos. Esse foi, inclusive, o principal objetivo deste trabalho. Por isso, foi realizada, adicionalmente, outra análise pelo procedimento *BLUP* (Henderson et al., 1959) com a inclusão desta informação (*BLUP_A*). Salienta-se que o fato de assumir esta

similaridade genética entre as progênes implica na recuperação da informação intergenotípica, ou seja, que, na predição realizada pelo procedimento $BLUP_A$ de uma dada progênie, é levado em conta não somente o seu desempenho *per se*, mas também, o desempenho das progênes a ela relacionadas (Bernardo, 2002; Henderson, 1984; Panter e Allen, 1995b; Resende, 2002a).

Para esta análise, utilizou-se apenas o parentesco genético aditivo entre as progênes (item 3.1.4) à semelhança do que já é rotineiramente feito no melhoramento animal – modelo animal - sob a pressuposição de genes de ação aditiva (g.m.d.=0,0) (Lynch & Walsh, 1998; Mrode, 1996).

Em se tratando da situação em que o grau médio de dominância é nulo, verifica-se que as estimativas de h_a^2 para o método genealógico, desconsiderando o parentesco (\hat{h}_a^2) e levando-se em conta o parentesco (\hat{h}_r^2), foram equivalentes (Tabela 1). Contudo, observa-se que as estimativas da herdabilidade h_a^2 , quando essa informação é considerada, tiveram menores desvios padrões associados.

Este resultado demonstra que levar em conta a genealogia, uma vez disponível, para fins de análise, como ocorre freqüentemente na condução pelo método *pedigree*, é vantajoso. Para os demais casos em que existe variância genética de dominância (g.m.d.=0,5 e 1,0), as estimativas \hat{h}_r^2 permaneceram próximas dos valores reais, o que pode ser explicado em função da reduzida contribuição da dominância nesta geração avançada de endogamia (F_5). Observando-se a relação \hat{h}_r^2/\hat{h}_a^2 e considerando-se que o denominador em ambos os estimadores é o mesmo, nota-se que a diferença se deve apenas ao numerador, isto é, no que está contido na variância genética entre as progênes, no caso $F_{4,5}$, ou seja, $7/4\sigma_A^2/(7/4\sigma_A^2 + 7/64\sigma_D^2)$. Considerando o modelo genético simulado de locos de segregação independente com frequência alélica 0,5, tem-se que a relação anterior equívale a 0,99 no g.m.d. de 0,5 e 0,97 no g.m.d. de 1,0,

representando, dessa forma, uma contribuição da dominância de, apenas, um ponto percentual no g.m.d de 0,5 e três pontos percentuais no g.m.d. de 1,0. Logo, o valor esperado da \hat{h}_r^2 (h_r^2) corresponde praticamente ao valor da herdabilidade h_a^2 , como constatado (Tabela 1).

Outro resultado importante da análise com o modelo, levando-se em consideração o parentesco aditivo, foi que os valores médios das estimativas do componente de variância genética associado a progênies reportaram diretamente à variância genética aditiva na F_2 (σ_A^2) em todas as configurações simuladas. Percebe-se que as estimativas $\hat{\sigma}_A^2$ foram próximas de 10, o verdadeiro valor populacional segundo o modelo genético simulado com 20 locos segregantes (item 3.1.1). Nesse caso, convém destacar que o modelo analisado, como já comentado anteriormente, supõe que o controle genético seja puramente aditivo, de modo que o seu emprego em presença de dominância pode não ser adequado (Lynch & Walsh, 1998). Contudo, pôde-se verificar que, mesmo com g.m.d. iguais a 0,5 ou 1,0, o modelo utilizado se mostrou bastante robusto, não acarretando em maiores problemas na estimação dos componentes de variância, sobretudo, σ_A^2 , mostrando que, naquela geração avaliada, a dominância parece ter pouca influência sobre o modelo ajustado (Tabela 1).

Uma questão de interesse para os melhoristas na avaliação das progênies refere-se à seleção daquelas que apresentam melhores valores genéticos associados, dado que o objetivo final, no caso de autógamias, é a obtenção de linhagens superiores. Nesse contexto, procurou-se avaliar, por meio da simulação, alguns procedimentos de seleção quanto ao ordenamento correto das progênies $F_{4:5}$ provenientes dos métodos genealógico e *bulk*, tendo como referência o conhecimento dos verdadeiros valores genotípicos (*VG*) associados. Foram avaliadas as classificações pela média fenotípica (*M*) e pelo preditor *BLUP* na condição de ausência de parentesco, para ambos os métodos.

Adicionalmente, para o método genealógico, foi também avaliada a classificação pelo preditor $BLUP$, considerando a informação do parentesco aditivo ($BLUP_A$) existente entre as progênes (Tabela 2). Vale comentar que, nesse contexto de avaliação de progênes oriundas de uma mesma população de melhoramento, as predições pelo procedimento $BLUP_A$ remetem aos valores genéticos aditivos (VGA) preditos das progênes, ou seja, os *breeding value*, de forma que a sua utilização leva, possivelmente, a maximizar a chance de realizar o ranqueamento verdadeiro das progênes (Mrode, 1996; Panter e Allen, 1995b; Resende, 2002a ; White & Hodge, 1989).

Inicialmente, é preciso salientar que os valores médios das estimativas das correlações de Spearman (r_S) entre a média M e $BLUP$ sem parentesco, para todas as configurações simuladas, foram iguais a 1,0, evidenciando não haver qualquer vantagem seletiva entre os procedimentos. A justificativa para este resultado está nas condições de balanceamento em que a simulação foi efetuada, com o uso de delineamento experimental ortogonal e, ainda, sem a existência de qualquer desbalanceamento por perda de dados (Bearzoti, 2002; Bernardo, 2002; Duarte & Vencovsky, 2001; Panter & Allen, 1995b). Assim, são mostrados os resultados advindos com a utilização da média M , subentendendo sua equidade com o $BLUP$ sem parentesco nas condições comentadas anteriormente.

Antes de qualquer comentário, pode-se observar que a dominância não exerceu qualquer influência sobre as estimativas das r_S entre os procedimentos em avaliação. Atendo-se à situação de herdabilidade h_a^2 igual a 25% e às correlações r_S entre $BLUP_A$ e VG [$r_{S(BLUP_A, VG)}$], por exemplo, verifica-se que as estimativas das $r_{S(BLUP_A, VG)}$ foram iguais a 0,62, em todas as situações de g.m.d estudadas (Tabela 2). Este resultado ratifica, mais uma vez, que a dominância parece ter pouca importância e, no tocante ao preditor $BLUP_A$, novamente mostrando que o modelo simplificado, levando-se em conta apenas o parentesco aditivo, parece ser suficiente para explicar a variação existente. Este resultado

está de acordo com o obtido por Van Oeveren & Stam (1992) que, por meio de simulação computacional, também verificaram que a dominância tem pouca importância no melhoramento de plantas autógamas.

De modo geral, as estimativas médias das correlações r_s entre os procedimentos avaliados foram diretamente proporcionais aos valores da herdabilidade h_a^2 . Por exemplo, observando-se a estimativa da correlação r_s entre a média M e VG [$r_{S(M, VG)}$] na situação de baixa herdabilidade h_a^2 (10%), percebe-se que ela foi de apenas 0,30, enquanto que, quando a herdabilidade h_a^2 passou a ser de 50%, esta correlação $r_{S(M, VG)}$ subiu para 0,69 (Tabela 2). A explicação para este resultado baseia-se no que conceitualmente entende-se pelo

TABELA 2 Valores médios das correlações de Spearman (r_s) e desvios padrões associados (valores entre parenteses) entre valores genotípicos paramétricos (VG), médias fenotípicas (M) e $BLUP$ considerando o parentesco aditivo ($BLUP_A$) de progênies $F_{4:5}$, conduzidas pelos métodos genealógico e *bulk*, variando o grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênies (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1000.

g.m.d.	h_a^2 (%)	Método de condução			
		Genealógico			<i>Bulk</i>
		$r_{S(M, VG)}$	$r_{S(BLUP_A, VG)}$	$r_{S(BLUP_A, M)}$	$r_{S(M, VG)}$
0,0	10	0,30(0,06)	0,43(0,09)	0,69(0,04)	0,30(0,06)
0,0	25	0,48(0,05)	0,62(0,06)	0,76(0,04)	0,48(0,05)
0,0	50	0,69(0,04)	0,79(0,04)	0,87(0,03)	0,69(0,04)
0,0	75	0,85(0,02)	0,89(0,02)	0,95(0,01)	0,85(0,02)
0,5	10	0,30(0,06)	0,43(0,08)	0,69(0,04)	0,30(0,06)
0,5	25	0,48(0,05)	0,62(0,06)	0,76(0,04)	0,48(0,05)
0,5	50	0,69(0,04)	0,79(0,04)	0,87(0,03)	0,69(0,04)
0,5	75	0,85(0,02)	0,89(0,02)	0,95(0,01)	0,85(0,02)
1,0	10	0,30(0,06)	0,43(0,08)	0,69(0,04)	0,30(0,06)
1,0	25	0,48(0,05)	0,62(0,06)	0,76(0,04)	0,48(0,05)
1,0	50	0,69(0,04)	0,78(0,04)	0,87(0,03)	0,69(0,04)
1,0	75	0,85(0,02)	0,89(0,02)	0,95(0,01)	0,85(0,02)

parâmetro herdabilidade (h_a^2), o qual exprime o quanto da variação fenotípica observada está sendo explicada pela variação genotípica (Falconer & Mackay, 1996). Assim, considerando caracteres que possuem baixa herdabilidade h_a^2 , isto implica que existe, proporcionalmente, maior quantidade da variação fenotípica não devido a causas genéticas e, sim, advindas da influência de fatores ambientais, atestando que a média M pode não refletir fielmente o VG presente. Ao contrário, para caracteres com alta herdabilidade h_a^2 , o erro da média M é grandemente reduzido, sendo, neste caso, mais representativo do VG real.

Pelo exposto anteriormente, depreende-se que a herdabilidade h_a^2 representa um coeficiente de determinação entre a média M e VG , de modo que os valores médios das estimativas da correlação $r_{S(M), VG}$ podem ser utilizados para verificar a qualidade das simulações, por se tratarem, diretamente, de estimadores de $\sqrt{h_a^2}$ (Falconer & Mackay, 1996). Evidencia-se que os valores da correlação $r_{S(M), VG}$ foram próximos dos valores esperados ($\sqrt{h_a^2}$) para todas as configurações avaliadas. Como exemplo, observe que, para uma herdabilidade h_a^2 de 25%, a estimativa média da correlação $r_{S(M), VG}$ foi de 0,48, valor próximo ao valor populacional de 0,5 (Tabela 2).

É oportuno comentar que tanto a média M quanto o $BLUP_A$ constituem-se em funções dos dados fenotípicos e que objetivam prever, igualmente, os valores genéticos aditivos (VGA) associados às progênies, os quais são melhor entendidos como realizações de variáveis aleatórias não observáveis, cujos valores reais nunca são conhecidos (White & Hodge, 1989). Assim, o melhor preditor será aquele que resultar nos VGA ranqueados o mais próximo possível do ranqueamento pelos VGA reais. Com esse objetivo, foram comparados os valores médios das correlações $r_{S(M), VG}$ com as correlações r_S entre $BLUP_A$ e VG [$r_{S(BLUP_A), VG}$]. Observa-se que, para todos os valores estudados de herdabilidade

h_a^2 , o $BLUP_A$ se mostrou superior à média M , ou seja, $r_{S(M, VG)} < r_{S(BLUP_A, VG)}$ (Tabela 2). Esta relação demonstra uma maior eficiência do $BLUP_A$ para a correta classificação das progênes com a incorporação do parentesco. Bernardo (2002, pg. 234) afirma que o procedimento $BLUP$, mesmo com a inclusão do parentesco, não teria qualquer vantagem prática para propósitos de ranqueamento e seleção, quando da avaliação de linhagens provenientes de uma mesma população de melhoramento. Não há nenhuma contradição nesses resultados. O referido autor considerou, como exemplo, quatro linhagens tomadas ao acaso na geração F_∞ , sem nenhuma informação acerca da genealogia. Já no presente caso, a genealogia foi obtida. Assim, a vantagem obtida nas simulações pelo $BLUP_A$ pode ser explicada observando-se a matriz de parentesco (A) considerada (13). Observa-se que a matriz A contém elementos fora da diagonal principal de diferentes magnitudes, denotando que uma quantidade desigual de informação é explorada para a obtenção das predições de cada progênie, diferentemente do exemplificado por Bernardo (2002).

Pelas simulações, observa-se que a eficiência relativa com o $BLUP_A$ foi mais acentuada para menores valores de herdabilidade h_a^2 (<50%), ratificada pela diminuição na correlação r_S entre $BLUP_A$ e média M [$r_{S(M, BLUP_A)}$]. Desse modo, consoante a herdabilidade h_a^2 diminui, a informação de parentesco passa a ter maior importância, ao passo que, com elevada herdabilidade h_a^2 (>50%), os valores genotípicos já estão bem determinados pelos próprios valores fenotípicos (M), tendo o parentesco nesta circunstância menor peso (Duarte & Vencovsky, 2001; Resende, 2002a), corroborando com os altos valores obtidos para as correlações $r_{S(M, BLUP_A)}$. Por exemplo, considerando uma herdabilidade h_a^2 de 10%, a eficiência relativa na correlação $r_{S(BLUP_A, VG)}$ em relação à correlação $r_{S(M, VG)}$ foi de 43,33%, enquanto que, para a herdabilidade h_a^2 de 50%, esta

eficiência diminuiu para apenas 14,5%, confirmado pelo alto valor da correlação $r_{S(M, BLUP_A)}$ observada, de 0,87 (Tabela 2).

Uma questão importante, que se torna indispensável ressaltar, é sobre o significado de não-viesado para o $BLUP$, mais especificamente para o $BLUP_A$. Esta preocupação procede, visto a falta de vivência de melhoristas com estes métodos, aliada à sua pequena divulgação (Bueno Filho, 1997). Como já comentado, o $BLUP_A$, no presente contexto, é um preditor dos VGA das progênes (\mathbf{a}), cuja esperança é igual a zero, ou seja, $E(\mathbf{a}) = \mathbf{0}$ (Falconer & Mackey, 1996). O $BLUP_A$ é não-tendencioso no sentido de que $E(\hat{\mathbf{a}}) = E(\mathbf{a})$ (Bernardo, 2002; Robinson, 1991), em que $\hat{\mathbf{a}}$ corresponde aos VGA preditos. Depreende-se, portanto, que, diferentemente do conceito de não-viesado para estimador de efeitos fixos, a não-tendenciosidade do $BLUP$ não está no âmbito das predições dos efeitos aleatórios individuais, ou seja, $E(\hat{\mathbf{a}}) = \mathbf{a}$, mas, referenciado à população destes efeitos. Em síntese, quando a $h_a^2 \rightarrow 100\%$, $\hat{\mathbf{a}} = E(\mathbf{a}/\mathbf{y}) \rightarrow \mathbf{a}$, enquanto que, com $h_a^2 \rightarrow 0\%$ tem-se que $\hat{\mathbf{a}} = E(\mathbf{a}/\mathbf{y}) \rightarrow \mathbf{0}$, demonstrando que o efeito *shrinkage* nas predições $BLUP$ se dá mais acentuadamente diante de baixos valores de h_a^2 , resultando na menor concordância entre a média M e $BLUP_A$, isto é, em menores estimativas da $r_{S(M, BLUP_A)}$, conforme mostrado na Tabela 2.

De modo geral, os desvios padrões associados às estimativas das correlações r_S foram baixos. Verifica-se que estes desvios decresceram, coerentemente, como já explicado, à medida que a herdabilidade h_a^2 aumenta (Tabela 2).

Outra questão de maior relevância para os melhoristas, que sobrepuja o aspecto da classificação de todas as progênes avaliadas, é identificar as progênes selecionadas, ou seja, aquelas situadas nos extremos. Para isso, foram estimadas as proporções de coincidência das progênes selecionadas pelo $BLUP_A$

e média M em comparação com as efetivamente selecionadas com base nos VG reais (Tabelas 3, 4 e 5).

Nota-se que, independente da proporção selecionada (s), a dominância não afetou os resultados das proporções de coincidências nas diferentes condições em que foram efetuadas as simulações. Observa-se, ainda, que, fixado o valor de s , as correspondentes proporções de coincidências estimadas, em ambos os extremos selecionados e métodos de condução, foram idênticas, o que, presumivelmente, resulta das restrições distribucionais assumidas, no caso a distribuição normal para o caráter e fatores envolvidos no modelo simulado (Tabelas 3, 4 e 5). Adicionalmente, vê-se que, quanto maior o valor de s , maiores foram as coincidências entre a média M e $BLUP_A$ [$C_{(M, BLUP_A)}$], mostrando que, sob intensidade seletiva branda, ambos se mostram bons preditores dos VGA das progênies.

Para o método *pedigree*, visualiza-se que as proporções de coincidências entre $BLUP_A$ e VG [$C_{(BLUP_A, VG)}$] foram superiores às observadas entre a média M e VG [$C_{(M, VG)}$], para todas as configurações e valores de s . Percebe-se, então, que o $BLUP_A$ proporcionou uma superioridade seletiva em relação à média M , isto é, $C_{(M, VG)} < C_{(BLUP_A, VG)}$, corroborando com o que foi visto anteriormente com as estimativas das correlações r_s (Tabelas 3, 4, 5 e 2). Outra vez a eficiência relativa do $BLUP_A$ em relação à média M nas coincidências com os VG foi proporcionalmente maior com o decréscimo nos valores da herdabilidade h_a^2 . Considerando, por exemplo, uma herdabilidade h_a^2 de 10%, g.m.d. de 0,0 e proporção selecionada s de 5%, a $C_{(BLUP_A, VG)}$ foi de 0,21, ao passo que a $C_{(M, VG)}$ foi de 0,15, isto é, uma eficiência relativa de 40%. Já para a proporção selecionada s de 10%, nas mesmas condições anteriores, o valor desta eficiência foi de 31,8%, e com uma proporção selecionada s igual a 25%, de 15,4%. Quando se considerou a herdabilidade h_a^2 de 50%, ainda nas mesmas

TABELA 3 Valores médios das proporções de coincidências (C) e desvios padrões associados (valores entre parênteses) para 5% das progênies $F_{4:5}$ superiores e inferiores ranqueadas pelos valores genotípicos paramétricos (VG), médias fenotípicas (M) e $BLUP$ considerando o parentesco aditivo ($BLUP_A$) para os métodos genealógico e *bulk*. As configurações simuladas variaram em função do grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênies (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1.000.

g.m.d.	h_a^2 (%)	5% superiores				5% inferiores			
		Genealógico			<i>Bulk</i>	Genealógico			<i>Bulk</i>
		$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$	$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$
0,0	10	0,15(0,09)	0,21(0,13)	0,40(0,11)	0,15(0,09)	0,15(0,09)	0,21(0,14)	0,40(0,11)	0,15(0,09)
0,0	25	0,24(0,11)	0,33(0,14)	0,48(0,11)	0,23(0,11)	0,24(0,11)	0,34(0,15)	0,47(0,11)	0,23(0,11)
0,0	50	0,38(0,12)	0,48(0,14)	0,60(0,10)	0,38(0,12)	0,39(0,12)	0,49(0,14)	0,60(0,11)	0,39(0,12)
0,0	75	0,57(0,12)	0,63(0,13)	0,75(0,09)	0,57(0,12)	0,57(0,12)	0,63(0,13)	0,75(0,09)	0,56(0,11)
0,5	10	0,14(0,09)	0,20(0,14)	0,40(0,11)	0,14(0,09)	0,15(0,09)	0,21(0,14)	0,40(0,10)	0,15(0,10)
0,5	25	0,23(0,11)	0,32(0,15)	0,47(0,12)	0,23(0,11)	0,24(0,11)	0,34(0,15)	0,47(0,12)	0,24(0,11)
0,5	50	0,38(0,12)	0,48(0,14)	0,60(0,11)	0,38(0,13)	0,39(0,12)	0,48(0,14)	0,60(0,11)	0,39(0,12)
0,5	75	0,56(0,11)	0,63(0,12)	0,75(0,09)	0,56(0,12)	0,57(0,12)	0,63(0,12)	0,76(0,09)	0,57(0,11)
1,0	10	0,15(0,10)	0,21(0,14)	0,40(0,11)	0,15(0,09)	0,15(0,10)	0,21(0,14)	0,40(0,11)	0,15(0,09)
1,0	25	0,24(0,11)	0,33(0,15)	0,47(0,11)	0,23(0,11)	0,24(0,11)	0,33(0,15)	0,46(0,11)	0,24(0,11)
1,0	50	0,38(0,12)	0,47(0,14)	0,60(0,11)	0,39(0,11)	0,40(0,12)	0,49(0,14)	0,61(0,11)	0,39(0,12)
1,0	75	0,55(0,12)	0,61(0,12)	0,75(0,09)	0,56(0,11)	0,57(0,12)	0,63(0,12)	0,75(0,09)	0,57(0,11)

TABELA 4 Valores médios das proporções de coincidências (C) e desvios padrões associados (valores entre parênteses) para 10% das progênies $F_{4,5}$ superiores e inferiores ranqueadas pelos valores genotípicos paramétricos (VG), médias fenotípicas (M) e $BLUP$ considerando o parentesco aditivo ($BLUP_A$) para os métodos genealógico e *bulk*. As configurações simuladas variaram em função do grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênies (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1.000.

g.m.d.	h_a^2 (%)	10% superiores				10% inferiores			
		Genealógico			<i>Bulk</i>	Genealógico			<i>Bulk</i>
		$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$	$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$
0,0	10	0,22(0,07)	0,29(0,11)	0,47(0,08)	0,22(0,07)	0,22(0,07)	0,29(0,11)	0,48(0,08)	0,22(0,07)
0,0	25	0,33(0,08)	0,42(0,11)	0,54(0,07)	0,33(0,08)	0,32(0,08)	0,42(0,11)	0,55(0,07)	0,32(0,08)
0,0	50	0,46(0,09)	0,56(0,10)	0,66(0,07)	0,47(0,08)	0,47(0,09)	0,56(0,10)	0,66(0,07)	0,47(0,08)
0,0	75	0,63(0,08)	0,69(0,08)	0,79(0,06)	0,64(0,08)	0,64(0,09)	0,69(0,08)	0,79(0,06)	0,63(0,08)
0,5	10	0,22(0,07)	0,29(0,11)	0,47(0,07)	0,22(0,07)	0,22(0,08)	0,29(0,11)	0,48(0,08)	0,22(0,07)
0,5	25	0,32(0,08)	0,41(0,11)	0,54(0,08)	0,32(0,08)	0,33(0,08)	0,42(0,11)	0,54(0,08)	0,33(0,08)
0,5	50	0,47(0,09)	0,56(0,10)	0,66(0,07)	0,47(0,08)	0,47(0,08)	0,56(0,10)	0,66(0,07)	0,48(0,08)
0,5	75	0,63(0,08)	0,69(0,08)	0,79(0,06)	0,63(0,08)	0,64(0,08)	0,69(0,08)	0,79(0,06)	0,64(0,08)
1,0	10	0,23(0,07)	0,29(0,11)	0,48(0,08)	0,23(0,07)	0,23(0,08)	0,29(0,11)	0,48(0,08)	0,22(0,07)
1,0	25	0,32(0,08)	0,41(0,11)	0,54(0,07)	0,32(0,08)	0,32(0,08)	0,42(0,11)	0,54(0,08)	0,32(0,08)
1,0	50	0,47(0,09)	0,56(0,10)	0,66(0,07)	0,47(0,08)	0,48(0,08)	0,56(0,10)	0,66(0,07)	0,47(0,09)
1,0	75	0,63(0,08)	0,68(0,08)	0,79(0,06)	0,62(0,08)	0,64(0,08)	0,69(0,08)	0,79(0,06)	0,64(0,07)

TABELA 5 Valores médios das proporções de coincidências (C) e desvios padrões associados (valores entre parênteses) para 25% das progênies $F_{4,5}$ superiores e inferiores ranqueadas pelos valores genotípicos paramétricos (VG), médias fenotípicas (M) e $BLUP$ considerando o parentesco aditivo ($BLUP_A$) para os métodos genealógico e *bulk*. As configurações simuladas variaram em função do grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênies (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1.000.

g.m.d.	h_a^2 (%)	25% superiores				25% inferiores			
		Genealógico			<i>Bulk</i>	Genealógico			<i>Bulk</i>
		$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$	$C_{(M, VG)}$	$C_{(BLUP_A, VG)}$	$C_{(BLUP_A, M)}$	$C_{(M, VG)}$
0,0	10	0,39(0,05)	0,45(0,07)	0,60(0,05)	0,39(0,05)	0,39(0,05)	0,45(0,07)	0,60(0,05)	0,39(0,05)
0,0	25	0,48(0,05)	0,56(0,06)	0,66(0,04)	0,48(0,05)	0,48(0,05)	0,56(0,06)	0,66(0,04)	0,48(0,05)
0,0	50	0,60(0,05)	0,68(0,06)	0,75(0,04)	0,60(0,05)	0,61(0,05)	0,68(0,06)	0,75(0,04)	0,61(0,05)
0,0	75	0,74(0,04)	0,78(0,04)	0,85(0,03)	0,73(0,04)	0,74(0,04)	0,78(0,04)	0,85(0,03)	0,73(0,05)
0,5	10	0,39(0,05)	0,45(0,07)	0,61(0,04)	0,39(0,05)	0,39(0,05)	0,46(0,07)	0,60(0,05)	0,39(0,05)
0,5	25	0,48(0,05)	0,56(0,07)	0,66(0,06)	0,48(0,05)	0,48(0,05)	0,56(0,08)	0,66(0,06)	0,48(0,05)
0,5	50	0,60(0,05)	0,68(0,06)	0,75(0,05)	0,60(0,05)	0,61(0,05)	0,68(0,06)	0,75(0,05)	0,61(0,05)
0,5	75	0,74(0,04)	0,78(0,04)	0,85(0,04)	0,73(0,04)	0,74(0,04)	0,78(0,05)	0,85(0,04)	0,73(0,04)
1,0	10	0,39(0,05)	0,45(0,07)	0,61(0,05)	0,39(0,05)	0,39(0,05)	0,45(0,07)	0,60(0,05)	0,39(0,05)
1,0	25	0,48(0,05)	0,56(0,06)	0,66(0,04)	0,48(0,05)	0,48(0,05)	0,56(0,07)	0,66(0,05)	0,48(0,05)
1,0	50	0,61(0,05)	0,68(0,05)	0,75(0,04)	0,60(0,05)	0,61(0,05)	0,68(0,05)	0,75(0,04)	0,61(0,05)
1,0	75	0,74(0,04)	0,77(0,04)	0,85(0,03)	0,73(0,04)	0,74(0,04)	0,78(0,04)	0,85(0,04)	0,74(0,04)

circunstâncias anteriores, os valores das eficiências relativas passaram a ser de 26,3% com proporção selecionada s de 5%, 21,7% com proporção selecionada s de 10% e 13,3% com proporção selecionada s de 25% (Tabelas 3, 4 e 5). Evidencia-se, assim, que a eficiência do $BLUP_A$ possivelmente será maior quando o melhorista estiver trabalhando com um caráter de herabilidade baixa, como mencionado e mediante a aplicação de alta intensidade seletiva.

Contudo, faz-se necessário evidenciar que as estimativas das proporções de coincidência tiveram desvios padrões proporcionalmente elevados, sendo mais acentuados para menores valores da herdabilidade h_a^2 em ambos os métodos de condução. Percebe-se, ainda, que os erros associados às estimativas das $C_{(BLUP_A, VG)}$ foram, proporcionalmente, maiores quando comparados com as $C_{(M, VG)}$ nas condições de baixa herdabilidade h_a^2 . Como exemplo, considerando-se uma alta intensidade seletiva, ou seja, com proporção selecionada s igual a 5% na herdabilidade h_a^2 de 10%, o desvio padrão associado a $C_{(M, VG)}$ correspondeu a 60% da estimativa e, para $C_{(BLUP_A, VG)}$, esta correspondência foi 61,9%, ao passo que, com herdabilidade h_a^2 de 75%, estas proporções foram de 21,05% para $C_{(M, VG)}$ e de 20,63% para $C_{(BLUP_A, VG)}$ (Tabela 3).

Atentando-se agora para uma intensidade seletiva mais branda, ou seja, para a proporção selecionada s de 25%, as proporções relativas dos desvios padrões foram de 12,82% para $C_{(M, VG)}$ e 15,56% para $C_{(BLUP_A, VG)}$ na herdabilidade h_a^2 de 10%; ao mesmo tempo, na herdabilidade h_a^2 de 75%, estas proporções relativas foram de 5,41% para $C_{(M, VG)}$ e 5,13% para $C_{(BLUP_A, VG)}$ (Tabela 5). Assumindo-se que as proporções de coincidência seguem distribuição binomial, tem-se que o estimador do desvio padrão populacional de uma proporção p é dado por $\sqrt{\hat{p}(1-\hat{p})/n}$, em que n é o tamanho da amostra (Ferreira, 2005). Assim, tomando-se como exemplo a proporção selecionada s de 5%, o que equiivale a um valor de n de 13, por se tratar da avaliação de 256

progênies, na configuração com herdabilidade h_a^2 de 10% e g.m.d. de 0,0, verifica-se, para o método *pedigree*, que a estimativa de $C_{(M, VG)}$ foi de 0,15, isto é, $\hat{p} = 0,15$ e que o desvio padrão médio associado foi de 0,09 (Tabela 3), ou seja, próximo do valor esperado de 0,10.

Em realidade, os melhoristas almejam é que as progênies selecionadas associem os maiores valores genéticos possíveis que, em última análise, refletem o ganho realizado com a seleção, desconsiderando a interação. Pelos resultados das Tabelas 6 e 7, verifica-se que os *VG* médios das progênies $F_{4:5}$ superiores selecionadas pelos procedimentos em estudo nas diferentes proporções selecionadas s , em ambos os métodos de condução, foram mais afetados pela herdabilidade h_a^2 , obtendo um contínuo incremento com o aumento da herdabilidade h_a^2 , como era esperado (Falconer & Mackay, 1996). Ademais, foram observados ligeiros aumentos nos *VG*, com a presença do efeito de dominância (g.m.d. = 0,5 e 1,0), em ambos os métodos de condução. Considerando a proporção selecionada s de 0,4%, o que equivale a identificar a progênie de melhor desempenho, por exemplo, e fixando a herdabilidade h_a^2 em 10%, observa-se que os *VG* médios das melhores progênies, do método *pedigree*, selecionadas pelo $BLUP_A$, foram de 104,7 no g.m.d. de 0,0, 105,3 no g.m.d. de 0,5 e de 106,2 no g.m.d. de 1,0. Essa mesma tendência foi observada para as demais herdabilidades h_a^2 e proporções selecionadas s (Tabelas 6 e 7).

Examinando-se os *VG* médios obtidos das progênies selecionadas pelo $BLUP_A$ relativo à média M para o método *pedigree*, nas proporções selecionadas s avaliadas, pode-se verificar que o procedimento $BLUP_A$ apresenta vantagem, porém, com ganhos relativos de menor expressão, em todas as configurações estudadas e, sobretudo, com o aumento da herdabilidade h_a^2 . Como exemplo, escolhendo uma herdabilidade h_a^2 de 10% e g.m.d. de 0,0, verifica-se que os

TABELA 6 Valores genotípicos médios e desvios padrões associados (valores entre parênteses) da progênie F_{4:5} de melhor desempenho e das 5% superiores ranqueadas pela média e *BLUP* considerando o parentesco aditivo (*BLUP_A*) para os métodos genealógico e *bulk* variando o grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênie (*h_a²*) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1.000.

g.m.d.	<i>h_a²</i> (%)	Melhor Progênie			5% superiores		
		Genealógico		<i>Bulk</i>	Genealógico		<i>Bulk</i>
		<i>Média</i>	<i>BLUP_A</i>	<i>Média</i>	<i>Média</i>	<i>BLUP_A</i>	<i>Média</i>
0,0	10	103,9(3,8)	104,7(3,7)	103,6(3,7)	102,6(1,2)	103,7(1,5)	102,7(1,2)
0,0	25	105,5(3,4)	106,9(3,3)	105,9(3,6)	104,2(1,2)	105,4(1,4)	104,2(1,2)
0,0	50	108,0(2,9)	108,7(2,7)	107,9(3,0)	105,9(1,1)	106,7(1,2)	105,9(1,1)
0,0	75	109,6(2,4)	109,9(2,3)	109,6(2,5)	107,3(1,1)	107,6(1,1)	107,3(1,0)
0,5	10	104,4(3,8)	105,3(3,8)	103,2(4,9)	103,3(1,2)	104,4(1,6)	102,5(2,3)
0,5	25	106,2(3,7)	107,3(3,3)	105,7(4,1)	104,8(1,2)	105,9(1,6)	104,5(1,9)
0,5	50	108,6(2,9)	109,1(2,7)	108,5(3,1)	106,6(1,1)	107,3(1,3)	106,5(1,3)
0,5	75	110,1(2,3)	110,4(2,1)	110,2(2,5)	107,9(1,0)	108,2(1,0)	107,9(1,0)
1,0	10	105,2(3,8)	106,2(3,8)	104,9(3,9)	104,0(1,2)	105,1(1,6)	104,0(1,3)
1,0	25	107,1(3,6)	108,1(3,2)	106,8(3,5)	105,5(1,2)	106,7(1,5)	105,5(1,2)
1,0	50	109,1(3,0)	109,8(2,6)	109,1(3,1)	107,2(1,1)	108,0(1,2)	107,2(1,1)
1,0	75	110,7(2,3)	110,9(2,2)	110,8(2,5)	108,5(1,0)	108,8(1,0)	108,5(1,1)

TABELA 7 Valores genotípicos médios e desvios padrões associados (valores entre parênteses) de 10% e 25% das progênes $F_{4,5}$ superiores ranqueadas pela média e $BLUP$ considerando o parentesco aditivo ($BLUP_A$) para os métodos genealógico e *bulk* variando o grau médio de dominância (g.m.d.) de 0,0, 0,5 e 1,0 e herdabilidade no sentido amplo na média das progênes (h_a^2) de 10%, 25%, 50% e 75%. O número de experimentos simulados para cada configuração foi de 1.000.

g.m.d.	h_a^2 (%)	10% superiores			25% superiores		
		Genealógico		<i>Bulk</i>	Genealógico		<i>Bulk</i>
		<i>Média</i>	$BLUP_A$	<i>Média</i>	<i>Média</i>	$BLUP_A$	<i>Média</i>
0,0	10	102,3(0,9)	103,2(1,1)	102,3(0,9)	101,7(0,6)	102,3(0,8)	101,7(0,7)
0,0	25	103,6(0,9)	104,7(1,1)	103,6(0,9)	102,6(0,7)	103,4(0,7)	102,7(0,7)
0,0	50	105,1(0,9)	105,8(1,0)	105,1(0,9)	103,7(0,7)	104,2(0,7)	103,7(0,7)
0,0	75	106,2(0,9)	106,5(0,9)	106,3(0,8)	104,6(0,6)	104,8(0,6)	104,6(0,6)
0,5	10	102,9(0,9)	103,9(1,1)	102,9(0,9)	102,3(0,6)	103,0(0,8)	102,3(0,7)
0,5	25	104,2(0,9)	105,2(1,2)	104,0(1,6)	103,3(0,6)	104,0(0,8)	103,1(1,2)
0,5	50	105,7(0,9)	106,4(1,0)	105,7(1,0)	104,4(0,6)	104,9(0,7)	104,4(0,8)
0,5	75	106,9(0,8)	107,1(0,8)	106,9(0,8)	105,2(0,6)	105,4(0,6)	105,2(0,6)
1,0	10	103,6(0,9)	104,5(1,2)	103,6(0,9)	103,0(0,6)	103,6(0,8)	102,9(0,7)
1,0	25	104,9(0,9)	105,9(1,1)	104,9(0,9)	103,9(0,6)	104,6(0,7)	103,9(0,6)
1,0	50	106,4(0,8)	107,1(0,9)	106,4(0,9)	105,1(0,6)	105,6(0,7)	105,0(0,7)
1,0	75	107,5(0,8)	107,8(0,8)	107,5(0,9)	105,9(0,6)	106,1(0,6)	105,8(0,7)

acrécimos relativos nos VG das progênes selecionadas pelo $BLUP_A$ em relação à média M foram de apenas 0,77% para s de 0,4%, 1,07% para s de 5%, 0,88% para s de 10% e de 0,59% para s de 25%. Com maiores valores de herdabilidade h_a^2 , por exemplo 50%, estes ganhos relativos foram ainda menores, ou seja, de 0,65% para s de 0,4%, 0,76% para s de 5%, 0,67% para s de 10% e de 0,48% para s de 25% (Tabelas 6 e 7, Figura 3). Com isso, pode-se inferir que quando o parentesco entre as progênes conduzidas pelo método genealógico é levado em consideração, os ganhos genotípicos com as progênes selecionadas pelo $BLUP_A$ foram ligeiramente superiores aos obtidos pela média M , ou seja, quando esta informação foi ignorada.

Um outro enfoque dessa simulação, era o de verificar a eficiência relativa do método *pedigree* com o *bulk*. Levando-se em consideração as ressalvas já mencionadas no que tange à condução de ambos os métodos, observa-se que o VG médio das progênes selecionadas, empregando-se a média M , em ambos os métodos, foram semelhantes em todas as configurações estudadas (Tabelas 6 e 7). Já no estudo via simulação realizado por Casali & Tigchelaar (1975), utilizando de modelo genético semelhante, diferindo apenas por admitir genes de ação aditiva, os autores obtiveram que o método *pedigree* foi considerado mais efetivo que o *bulk* para herdabilidades superiores a 50%, usando como critério de seleção o valor fenotípico. Contudo, ressalta-se que, durante a condução pelo método genealógico, por estes autores, foi efetuada a seleção fenotípica, enquanto que, para o *bulk*, parecem não ter considerado a ação da seleção natural, o que resultou, coerentemente, em maior vantagem para o método *pedigree*.

Inferese, assim, com base nos resultados da presente simulação, que os dois métodos, em termos de ganhos proporcionados, foram igualmente eficientes. Esses resultados são compatíveis com os relatados para algumas culturas autógamas em condições de campo, como, por exemplo, o trigo (Singh

et al., 1998) e o feijão (Raposo et al., 2000). Caberá, então, aos melhoristas, optar pelo melhor método em função dos recursos disponíveis. Vale salientar que em aplicando-se o método genealógico e identificando-se a genealogia, a utilização do procedimento $BLUP_A$ justificaria o trabalho adicional despendido na anotação dos registros genealógicos e, adicionalmente, proporcionaria ganhos superiores ao *bulk*.

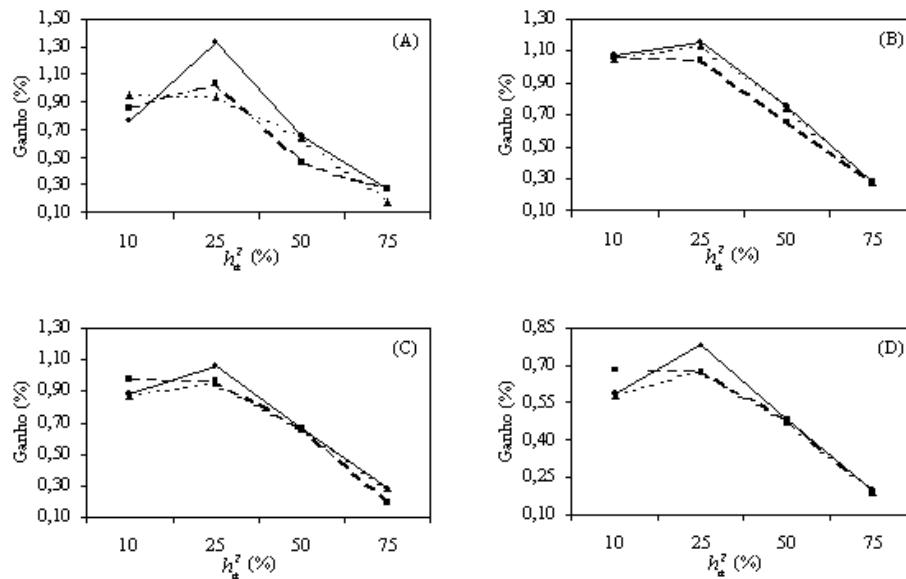


FIGURA 3 Ganho, em porcentagem, nos valores genotípicos médios das progênes $F_{4.5}$ selecionadas pelo $BLUP$ considerando a informação do parentesco aditivo ($BLUP_A$) em relação à média (M), em função da herdabilidade no sentido amplo na média das progênes (h_a^2) nas proporções selecionadas (s) de 0,4% (A), 5% (B), 10% (C) e 25% (D) e grau médio de dominância de 0,0 (---○---), 0,5 (---■---) e 1,0 (---▲---).

4.2 Experimentos de campo

O experimento de campo foi realizado com o intuito de aferir os resultados da simulação em um exemplo real. No caso do método genealógico, novamente, foram considerados dois modelos sob o enfoque de modelos mistos: considerando a informação da genealogia pela inclusão do parentesco genético aditivo entre as progênes, modelo G_A e ignorando essa informação, modelo G . É oportuno enfatizar que o modelo aditivo-dominante não foi incluído porque os resultados das simulações evidenciaram a inexpressiva contribuição da dominância, fato que já havia sido observado em outra oportunidade (Van Oeveren & Stam, 1992).

Inicialmente, é importante comentar que ocorreram desbalanceamentos nos dados de produção de grãos devido a perdas ocasionais de parcelas experimentais, sendo seis na safra ‘das águas’ de 2004/05 e três na ‘safra das secas’ de 2005 para o método genealógico, enquanto que, para o método *bulk*, estas perdas foram de doze parcelas na safra ‘das águas’ de 2004/05 e de duas na safra ‘das secas’ de 2005. As análises para esse caráter, portanto, foram realizadas levando-se em conta essas perdas.

As precisões experimentais dos experimentos envolvendo as progênes $F_{4.5}$ e $F_{4.6}$ para ambos os métodos de condução foram realizadas por meio das estimativas da herdabilidade. Nessa condição, verifica-se que os experimentos referentes à avaliação das progênes $F_{4.5}$ foram menos precisos para o caráter produção de grãos em ambos os métodos de condução (Tabelas 8 e 9). Um fator que, provavelmente, ocasionou essa perda na precisão, foi o menor tamanho da parcela experimental utilizado na avaliação das progênes $F_{4.5}$ (item 3.2.4.1). Na cultura do feijoeiro, tem sido relatada a redução da precisão experimental quando da utilização de parcelas experimentais menores, como de uma linha de um metro, como foi o caso em $F_{4.5}$ (Bertolucci et al., 1991). Essa é uma condição

inevitável, visto que, nesta ocasião, a progênie é gerada a partir de uma única planta, constituindo-se a quantidade de sementes um fator limitante para se aumentar o tamanho da parcela ou número de repetições (Fouilloux & Bannerot, 1988). Outra explicação pode ser devido ao efeito da safra de avaliação, a qual foi na denominada safra ‘das águas’. Levantamentos de experimentos realizados na região Sul de Minas Gerais mostram que perdas de precisão são,

TABELA 8 Estimativas, testes de hipótese e intervalos de confiança (IC) de parâmetros genéticos e ambientais (σ_b^2 : variância entre blocos; σ_p^2 : variância genética entre progênies; σ_A^2 : variância genética aditiva na F_2 ; σ_e^2, σ^2 : variâncias residuais; h_a^2 : herdabilidade no sentido amplo na média das progênies; h_r^2 : herdabilidade no sentido restrito na média das progênies) para o caráter produção de grãos (g/parcela) referente à avaliação das progênies $F_{4,5}$ conduzidas pelos métodos genealógico, considerando o parentesco aditivo (G_A) e ausência de parentesco (G), e *bulk* (B).

Método	Parâmetro	Estimativa	Erro padrão	Pr > Z	IC (95%)	
					LI	LS
G	σ_b^2	377,00	236,13	0,0552	147,91	2212,64
	σ_p^2	1655,14	575,23	0,0020	925,96	3768,24
	σ_e^2	7027,46	649,05	<0,0001	5910,59	8495,58
	h_a^2	30,32%				
G_A	σ_b^2	328,60	238,38	0,0838	84,07	2131,72
	σ_A^2	1002,38	309,19	0,0006	593,28	2047,28
	σ^2	7049,43	549,74	<0,0001	6084,64	8264,68
	h_r^2	32,13%				
B	σ_b^2	1056,98	399,99	0,0041	566,19	2632,57
	σ_p^2	2602,46	595,50	<0,0001	1739,79	4316,18
	σ_e^2	5866,46	562,87	<0,0001	4902,54	7147,15
	h_a^2	43,20%				

em geral, observadas nesta safra, sobretudo para o caráter produção de grãos (Matos, 2005). A queda da precisão experimental reflete diretamente redução na qualidade e quantidade de informação a ser retirada do experimento (Steel et al., 1997).

Como o delineamento experimental empregado foi o de blocos incompletos parcialmente balanceados (*PBIB*) ou látice, considerando o efeito de blocos dentro de repetições como aleatórios, foi estimado, desse modo, o

TABELA 9 Estimativas, testes de hipótese e intervalos de confiança (IC) de parâmetros genéticos e ambientais (σ_b^2 : variância de blocos; σ_p^2 : variância genética entre progênies; σ_A^2 : variância genética aditiva na F_2 ; σ_e^2, σ^2 : variâncias residuais; h_a^2 : herdabilidade no sentido amplo na média das progênies; h_r^2 : herdabilidade no sentido restrito na média das progênies) para o caráter produção de grãos (g/parcela) referente à avaliação das progênies $F_{4,6}$ conduzidas pelos métodos genealógico, considerando o parentesco aditivo (G_A) e ausência de parentesco (G), e *bulk* (B).

Método	Parâmetro	Estimativa	Erro padrão	Pr > Z	IC (95%)	
					LI	LS
G	σ_b^2	174,60	110,36	0,0568	68,06	1048,53
	σ_p^2	1388,17	297,45	<0,0001	949,74	2220,90
	σ_e^2	3006,79	275,66	<0,0001	2532,00	3629,64
	h_a^2	46,24%				
G_A	σ_b^2	122,41	107,53	0,1275	37,06	2345,09
	σ_A^2	548,18	176,05	<0,0009	318,38	1160,04
	σ^2	3495,61	288,55	<0,0001	2992,29	4138,26
	h_r^2	31,96%				
B	σ_b^2	228,59	122,94	0,0315	99,53	958,20
	σ_p^2	1320,69	286,23	<0,0001	900,03	2125,72
	σ_e^2	2928,84	274,08	<0,0001	2457,99	3550,03
	h_a^2	46,67%				

componente de variância entre blocos (σ_b^2). Verifica-se que, nos modelos analisados com o método genealógico (G, G_A), as estimativas de σ_b^2 foram não significativamente diferentes de zero ($\text{Pr} > 0,05$) para o caráter produção de grãos. Já para o experimento das progênes provenientes do *bulk*, o componente σ_b^2 foi expressivo ($\text{Pr} < 0,05$) (Tabelas 8 e 9), assim como para o caráter nota de porte avaliado nas progênes conduzidas pelo método genealógico (Tabela 10). Os intervalos de confiança $IC(95\%)$ para o componente σ_b^2 foram amplos para os casos em que o teste Z foi não significativo ($\text{Pr} > 0,05$). Vale salientar que os intervalos de confiança para componentes de variância são apenas assintóticos sob ausência de normalidade e que os graus de liberdade constituem-se em aproximações obtidas por Satterthwaite (1946).

A não significância do teste de hipótese Z ($\text{Pr} > 0,05$) para o componente σ_b^2 implica que o delineamento experimental, látice, foi pouco eficiente para os experimentos concernentes ao método genealógico ou, de outro modo, que nestes experimentos pouca informação interblocos foi recuperada (Duarte & Vencovsky, 2001). Segundo Malheiros (1982), a baixa eficiência da análise com recuperação da informação interblocos é verificada quando a seguinte relação é aceita: $\sigma_b^2/\sigma_e^2 \geq 1/k$, sendo k o tamanho do bloco. Com relação ao caráter produção de grãos, observa-se que as estimativas $\hat{\sigma}_b^2/\hat{\sigma}_e^2$ para o método genealógico com base no modelo G foram de 0,05 na safra ‘das águas’ e 0,06 na safra ‘das secas’, ou seja, praticamente iguais a $1/k$, ou seja, $1/16$. Já para o *bulk*, as estimativas $\hat{\sigma}_b^2/\hat{\sigma}_e^2$ foram de 0,18 na safra ‘das águas’ e de 0,078 na safra ‘das secas’, ou seja, superiores a $1/16$ (Tabelas 8 e 9).

Outro ponto relacionado às estimativas de σ_b^2 para o método genealógico, foi a sua redução com a utilização do modelo G_A (Tabelas 8, 9 e 10). Uma explicação provável se deve ao plano experimental utilizado, pois, por

ocasião do sorteio utilizado, em ambos os experimentos, progênies relacionadas foram alocadas dentro do mesmo bloco num dos arranjos ortogonais, resultando num confundimento entre os efeitos de blocos e de

TABELA 10 Estimativas, testes de hipótese e intervalos de confiança (IC) de parâmetros genéticos e ambientais (σ_b^2 : variância de blocos; σ_p^2 : variância genética entre progênies; σ_A^2 : variância genética aditiva na F_2 ; σ_e^2, σ^2 : variâncias residuais; h_a^2 : herdabilidade no sentido amplo na média das progênies; h_r^2 : herdabilidade no sentido restrito na média das progênies) para o caráter nota de porte referente à avaliação das progênies $F_{4:5}$ e $F_{4:6}$ conduzidas pelos método genealógico, considerando o parentesco aditivo (G_A) e ausência de parentesco (G).

Geração	Método	Parâmetro	Estimativa	Erro padrão	Pr > Z	IC (95%)	
						LI	LS
$F_{4:5}$	G	σ_b^2	0,1057	0,0361	0,0017	0,0596	0,2364
		σ_p^2	0,1189	0,0384	0,0010	0,0689	0,2527
		σ_e^2	0,4294	0,0399	<0,0001	0,3607	0,5199
		h_a^2	30,92%				
	G_A	σ_b^2	0,0790	0,0292	0,0034	0,0429	0,1914
		σ_A^2	0,0946	0,0217	<0,0001	0,0631	0,1573
		σ^2	0,3962	0,0304	<0,0001	0,3428	0,4632
		h_r^2	43,06%				
$F_{4:6}$	G	σ_b^2	0,1157	0,0344	0,0004	0,0697	0,2290
		σ_p^2	0,0625	0,0202	0,0010	0,0362	0,1326
		σ_e^2	0,2231	0,0210	<0,0001	0,1870	0,2708
		h_a^2	30,88%				
	G_A	σ_b^2	0,1010	0,0308	0,0005	0,0601	0,2042
		σ_A^2	0,0477	0,0107	<0,0001	0,0321	0,0784
		σ^2	0,2015	0,0160	<0,0001	0,1736	0,2368
		h_r^2	41,24%				

progênies, de forma que com a recuperação da informação interprogênies, possivelmente, parte do efeito de blocos pode já ter sido recuperada (Bueno Filho & Gilmour, 2003).

Embora os genitores utilizados no cruzamento sejam do *pool* gênico Mesoamericano, estes são originários de programas distintos, sem parentesco entre si e com divergência para hábito de crescimento e vários outros caracteres (item 3.2.2). Essa divergência foi comprovada, pois, em ambos os métodos de condução e características avaliadas, o componente de variância genética associado a progênies (σ_p^2), nas duas gerações avaliadas, foi significativamente não nulo ($Pr < 0,001$) (Tabelas 8, 9 e 10). A amplitude da variação genética das progênies $F_{4,5}$ e $F_{4,6}$ para a produção de grãos e nota de porte também é evidenciada pelas distribuições de frequência mostradas nas Figuras 4, 5 e 6, elaboradas com base nas médias ajustadas obtidas da análise com recuperação da informação interblocos (análise não mostrada). Essa é uma condição indispensável para o propósito desse trabalho, ou seja, comparar métodos de condução (Fouilloux & Bannerot, 1988).

É oportuno salientar que as estimativas de σ_p^2 para o caráter produção de grãos nos dois métodos de condução, *pedigree* e *bulk*, não diferiram significativamente, sobretudo para os experimentos da safra 'das secas' (Tabela 9). Este resultado concorda com o obtido nas simulações, subentendendo-se que a amostragem realizada durante a condução pelo método genealógico - duas plantas por progênie a cada geração - garantiu que a variação genética fosse adequadamente explorada.

No que se refere ao método genealógico, as análises pelo modelo G_A para as progênies $F_{4,5}$ e $F_{4,6}$ resultaram em estimativas altamente significativas ($Pr < 0,001$) para a variância genética aditiva na F_2 (σ_A^2) para as características avaliadas (Tabelas 8, 9 e 10). Teoricamente, sob a mesma influência de fatores

ambientais e ausência de interação safras por progênies, as estimativas de σ_A^2 para estes caracteres nas duas gerações deveriam ser semelhantes. No entanto, ocorreram diferenças nas magnitudes destas estimativas nas diferentes gerações avaliadas. Novamente, a explicação baseia-se nas diferenças de precisões dos experimentos, resultantes das peculiaridades da atuação de fatores ambientais em cada safra, sem mencionar o efeito da interação progênies por safras, o qual é muito comum em experimentos de avaliação de progênies de feijoeiro conduzidos no Sul de Minas Gerais (Abreu et al. 1990; Rosal et al., 2000; Takeda et al., 1991).

É esperado que, com o modelo G_A , a estimativa da variância genética aditiva entre as progênies seja superior àquela obtida sem o parentesco. Esse fato tem sido, inclusive, relatado na literatura (Bromley et al., 2000; Durel et al., 1998). No presente trabalho, o modelo G estima σ_p^2 , isto é, a variância genética entre as progênies. Na geração $F_{4,5}$, por exemplo, σ_p^2 contém $7/4\sigma_A^2 + 7/64\sigma_D^2$. Sob hipótese de ausência de dominância, a estimativa de σ_A^2 do modelo G_A pode ser comparada com a estimativa de σ_p^2 do modelo G . Tendo como referência o caráter produção de grãos na geração $F_{4,5}$ (Tabela 8), observa-se que a estimativa de σ_p^2 , no modelo G , foi de 1.655,14. Já a estimativa de σ_A^2 , no modelo G_A , foi de 1.002,38. Considerando, como já mencionado, que a variância genética entre progênies $F_{4,5}$ explora $7/4\sigma_A^2$, tem-se que esta estimativa para o modelo G_A equivaleria a 1.754,165. Depreende-se, desconsiderando o erro inerente a esta estimativa, que esse valor é superior à estimativa de σ_p^2 no modelo G , o que está de acordo com o mostrado nas simulações. Esse fato só não foi constatado para a produção de grãos na geração $F_{4,6}$.

Normalmente, a seleção praticada pelos melhoristas é efetuada com base na média das progênies. Nesse caso, como os experimentos foram conduzidos

em látice, foram obtidas, adicionalmente, as médias ajustadas (MA_j) das progênie para as características avaliadas a partir da análise com apenas a recuperação da informação interblocos (análise não mostrada). Na Tabela 11 estão dispostas as estimativas das correlações de Spearman (r_s) e de Pearson (r) entre os critérios de seleção MA_j e $BLUP$, considerando o parentesco genético aditivo ($BLUP_A$) e a ausência de parentesco ($BLUP$). Primeiramente,

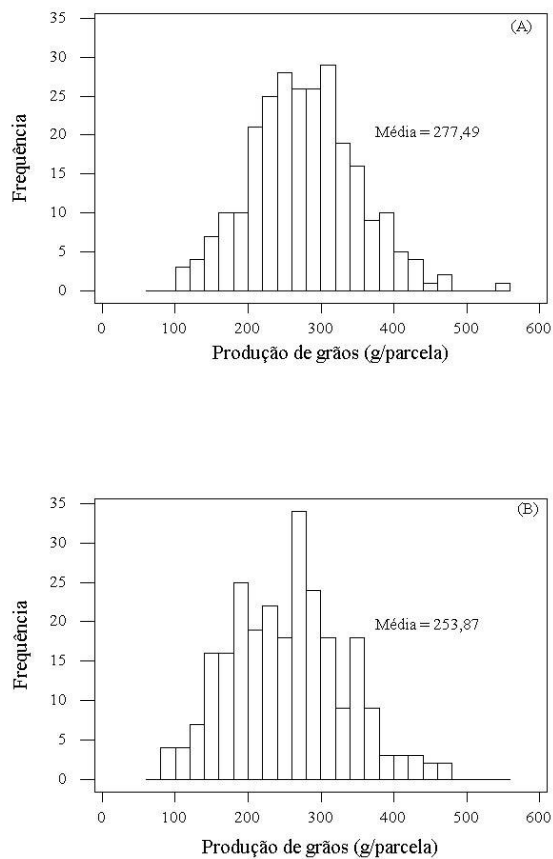


FIGURA 4 Distribuição de frequência para a produção de grãos (g/parcela) das 256 progênie $F_{4.5}$ conduzidas pelos métodos genealógico (A) e *bulk* (B).

observa-se que as estimativas das correlações de Pearson e Spearman foram semelhantes. Este resultado pode ser justificado pela boa aproximação para a distribuição normal apresentada pelos caracteres (Figuras 4, 5 e 6).

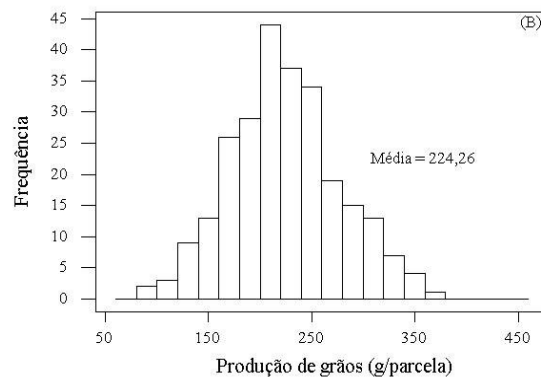
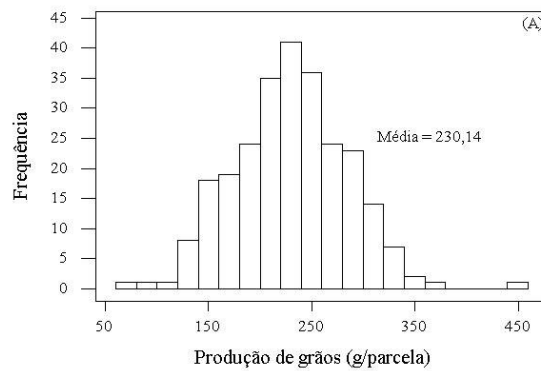


FIGURA 5 Distribuição de frequência para a produção de grãos (g/parcela) das 256 progênies $F_{4:6}$ conduzidas pelos métodos genealógico (A) e *bulk* (B).

Verifica-se que, para o caráter produção, em ambas as gerações e métodos de condução, as estimativas das $r_{S(MAj, BLUP)}$ foram sempre elevadas, enquanto que as $r_{S(MAj, BLUP)}$ e $r_{S(BLUP, BLUP)}$ foram de menores magnitudes, porém, ainda expressivas. Para o caráter nota de porte, estas considerações também são válidas. Estes resultados permitem inferir que a inclusão

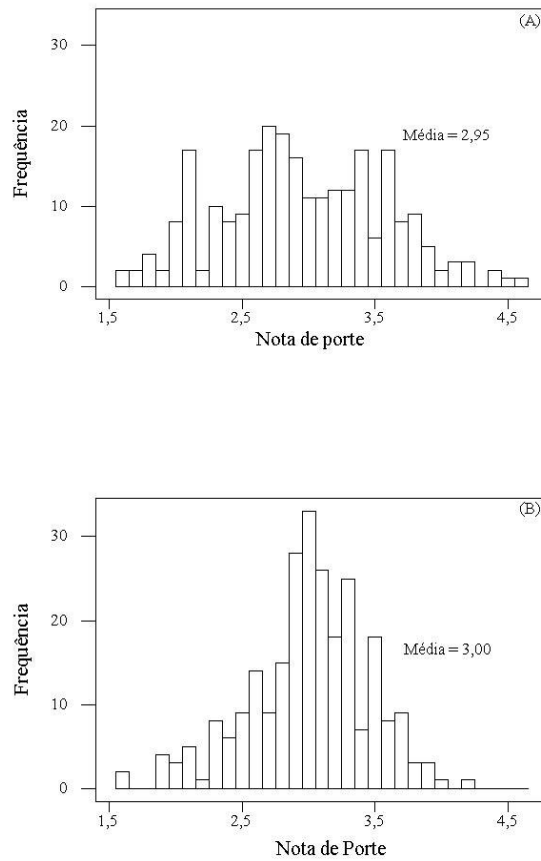


FIGURA 6 Distribuição de frequência para a nota de porte das 256 progênies $F_{4:5}$ (A) e $F_{4:6}$ (B) conduzidas pelo método genealógico.

TABELA 11 Estimativas e testes de *t*-Student das correlações (*Corr*) de Spearman (r_s) e Pearson (r) entre médias ajustadas (*MAj*) e *BLUP*, considerando o parentesco aditivo ($BLUP_A$) e ausência de parentesco (*BLUP*) das progênies nas gerações $F_{4;5}$ e $F_{4;6}$ conduzidas pelos métodos genealógico e *bulk*, para as características produção de grãos (g/parcela) e nota de porte.

Correlação	Característica	Método	<i>Corr</i> (. , .)	Progênies $F_{4;5}$	Progênies $F_{4;6}$
Spearman	Produção de grãos (g/parcela)	Genealógico	$r_{S(MAj, BLUP)}$	0,996**	0,997**
			$r_{S(MAj, BLUP_A)}$	0,802**	0,822**
			$r_{S(BLUP, BLUP_A)}$	0,798**	0,811**
		<i>Bulk</i>	$r_{S(MAj, BLUP)}$	0,992**	0,997**
	Nota de porte	Genealógico	$r_{S(MAj, BLUP)}$	0,985**	0,977**
			$r_{S(MAj, BLUP_A)}$	0,880**	0,880**
$r_{S(BLUP, BLUP_A)}$			0,848**	0,855**	
Pearson	Produção de grãos (g/parcela)	Genealógico	$r_{(MAj, BLUP)}$	0,993**	0,997**
			$r_{(MAj, BLUP_A)}$	0,810**	0,825**
			$r_{(BLUP, BLUP_A)}$	0,809**	0,815**
		<i>Bulk</i>	$r_{(MAj, BLUP)}$	0,990**	0,998**
	Nota de porte	Genealógico	$r_{(MAj, BLUP)}$	0,987**	0,985**
			$r_{(MAj, BLUP_A)}$	0,884**	0,903**
$r_{(BLUP, BLUP_A)}$			0,853**	0,879**	

da informação de parentesco ocasionou maiores mudanças no ordenamento das progênies do que simplesmente o desbalanceamento presente, ocasionado pelo delineamento experimental utilizado e ou perda de parcelas, como aconteceu para a produção de grãos (Panter e Allen, 1995a; Panter e Allen, 1995b; Resende, 2002a).

As proporções de coincidências entre os procedimentos de seleção, nas progênies $F_{4;5}$ e $F_{4;6}$, considerando diferentes proporções selecionadas (*s*), nos extremos, são apresentadas nas Tabelas 12 e 13. As coincidências entre *MAj* e *BLUP* [$C_{(MAj, BLUP)}$] para a produção de grãos nos dois métodos de condução avaliados foram, em geral, elevadas, corroborando com as estimativas das $r_{S(MAj, BLUP)}$ ($r > 0,95$) apresentadas na Tabela 11. Este resultado ratifica,

TABELA 12 Proporções de coincidências (C) para 1% e 5% das progêneses $F_{4:5}$ e $F_{4:6}$ superiores e inferiores conduzidas pelos métodos genealógico e *bulk*, selecionadas com base na média ajustada (MA_j) e $BLUP$, considerando o parentesco aditivo ($BLUP_A$) e ausência de parentesco ($BLUP$), para as características produção de grãos (g/parcela) e nota de porte.

s (%)	Característica	Método	$C(., .)$	Progêneses	Progêneses
				$F_{4:5}$	$F_{4:6}$
1% (+)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,33	1,00
			$C_{(MA_j, BLUP_A)}$	0,00	0,67
			$C_{(BLUP, BLUP_A)}$	0,00	0,67
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,67	0,67
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,67	0,67
			$C_{(MA_j, BLUP_A)}$	0,33	0,67
$C_{(BLUP, BLUP_A)}$			0,33	0,33	
1% (-)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	1,00	1,00
			$C_{(MA_j, BLUP_A)}$	0,33	0,67
			$C_{(BLUP, BLUP_A)}$	0,33	0,67
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	1,00	1,00
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,67	1,00
			$C_{(MA_j, BLUP_A)}$	0,33	0,67
$C_{(BLUP, BLUP_A)}$			0,00	0,67	
5% (+)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,85	0,85
			$C_{(MA_j, BLUP_A)}$	0,62	0,38
			$C_{(BLUP, BLUP_A)}$	0,62	0,46
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,85	0,92
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,85	0,92
			$C_{(MA_j, BLUP_A)}$	0,62	0,62
$C_{(BLUP, BLUP_A)}$			0,62	0,54	
5% (-)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,92	0,92
			$C_{(MA_j, BLUP_A)}$	0,62	0,46
			$C_{(BLUP, BLUP_A)}$	0,69	0,46
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,77	0,92
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,77	1,00
			$C_{(MA_j, BLUP_A)}$	0,54	0,54
$C_{(BLUP, BLUP_A)}$			0,54	0,54	

TABELA 13 Proporções de coincidências (C) para 10% e 25% das progênies $F_{4:5}$ e $F_{4:6}$ superiores e inferiores conduzidas pelos métodos genealógico e *bulk*, selecionadas com base na média ajustada (MA_j) e $BLUP$, considerando o parentesco aditivo ($BLUP_A$) e ausência de parentesco ($BLUP$), para as características produção de grãos (g/parcela) e nota de porte.

s (%)	Característica	Método	$C(. , .)$	Progênies	Progênies
				$F_{4:5}$	$F_{4:6}$
10% (+)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,96	1,00
			$C_{(MA_j, BLUP_A)}$	0,58	0,62
			$C_{(BLUP, BLUP_A)}$	0,58	0,62
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,92	1,00
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,92	0,88
			$C_{(MA_j, BLUP_A)}$	0,58	0,73
$C_{(BLUP, BLUP_A)}$			0,58	0,65	
10% (-)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,96	1,00
			$C_{(MA_j, BLUP_A)}$	0,62	0,54
			$C_{(BLUP, BLUP_A)}$	0,62	0,54
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,77	0,92
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,73	0,85
			$C_{(MA_j, BLUP_A)}$	0,58	0,69
$C_{(BLUP, BLUP_A)}$			0,58	0,69	
25% (+)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,97	0,98
			$C_{(MA_j, BLUP_A)}$	0,72	0,70
			$C_{(BLUP, BLUP_A)}$	0,73	0,69
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,95	0,95
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,94	0,89
			$C_{(MA_j, BLUP_A)}$	0,67	0,81
$C_{(BLUP, BLUP_A)}$			0,69	0,77	
25% (-)	Produção de grãos (g/parcela)	Genealógico	$C_{(MA_j, BLUP)}$	0,95	0,97
			$C_{(MA_j, BLUP_A)}$	0,66	0,67
			$C_{(BLUP, BLUP_A)}$	0,69	0,66
		<i>Bulk</i>	$C_{(MA_i, BLUP)}$	0,89	0,97
	Nota de porte	Genealógico	$C_{(MA_j, BLUP)}$	0,89	0,94
			$C_{(MA_j, BLUP_A)}$	0,77	0,80
$C_{(BLUP, BLUP_A)}$			0,73	0,78	

novamente, que as condições de desbalanceamento não causaram maiores problemas para fins de seleção. Contudo, verificou-se que as proporções de coincidências entre o $BLUP_A$ e MAj [$C_{(MAj,BLUP_A)}$] e entre o $BLUP_A$ e $BLUP$ [$C_{(BLUP,BLUP_A)}$] foram de menor magnitude. Ademais, verificou-se que as proporções de coincidências entre os diferentes procedimentos foram mais variáveis sob baixos valores de s , indicando, como já realçado, que o efeito *shrinkage* do $BLUP$, sobretudo com parentesco ($BLUP_A$), é mais pronunciado sobre valores extremos. Vale, ainda, mencionar que as coincidências nos dois extremos tiveram comportamento semelhante, sobretudo para maiores valores de s .

Em plantas autógamas, o interesse principal está voltado para o ganho realizado na geração F_∞ , quando se atinge a homozigose completa. Contudo, no presente caso, somente foi possível estimar o ganho realizado na geração $F_{4:6}$ pela seleção efetuada na geração $F_{4:5}$, utilizando os diferentes critérios em diferentes intensidades seletivas nos dois extremos (Tabela 14). Observa-se que no $BLUP_A$ com s de 1% o ganho foi de 2,85% e de 2,31% com s de 5%, enquanto que com a MAj esses ganhos foram de 13,78% e 13,96%, respectivamente. Com maiores valores de proporções selecionadas s , os ganhos praticamente não diferiram entre os procedimentos de seleção. Já para os ganhos obtidos para os extremos inferiores, o $BLUP_A$ se mostrou mais eficiente, especialmente para valores de s de 1% e 5%. Para a nota de porte, o comportamento foi similar.

Em princípio, os resultados anteriores não estão concordantes com os relatados nas simulações, como também com as estimativas dos parâmetros genéticos obtidas com os experimentos de campo. Há alguns fatores que, indubitavelmente, contribuíram para essa discrepância. O primeiro deles é a interação genótipos por ambientes, que não foi considerada na obtenção das estimativas anteriores e que, normalmente, é expressiva com a cultura do

TABELA 14 Ganho realizado, em porcentagem, para as características produção de grãos (g/parcela) e nota de porte na geração F_{4:6} dos métodos genealógico e *bulk*, com base na seleção na geração F_{4:5}, utilizando como procedimentos a média ajustada (*MAj*) e *BLUP*, considerando o parentesco aditivo (*BLUP_A*) e ausência de parentesco (*BLUP*), nas proporções selecionadas (*s*) de 1%, 5%, 10% e 25%.

Característica	<i>s</i> (%)	Genealógico			<i>Bulk</i>	
		<i>MAj</i>	<i>BLUP</i>	<i>BLUP_A</i>	<i>MAj</i>	<i>BLUP</i>
Produção de grãos (g/parcela)	1	13,78	5,82	2,85	7,56	11,89
	5	13,96	9,20	2,31	8,53	14,47
	10	6,95	4,00	7,50	9,19	9,48
	25	9,63	9,24	7,40	10,52	9,70
	-25	-6,04	-4,83	-7,27	-8,37	-10,73
	-10	-10,52	-10,85	-2,37	-10,34	-11,79
	-5	-3,21	-2,13	-10,18	-16,37	-11,33
	-1	2,12	2,12	-16,88	-3,96	-3,97
	-	-	-	-	-	-
Nota de porte	1	7,11	8,55	0,68	-	-
	5	7,05	9,89	9,94	-	-
	10	11,04	7,00	11,66	-	-
	25	7,57	7,81	8,40	-	-
	-25	-8,43	-8,20	-7,41	-	-
	-10	-13,10	-11,13	-13,23	-	-
	-5	-12,11	-13,57	-16,42	-	-
	-1	-15,95	-5,43	-34,35	-	-
	-	-	-	-	-	-

feijoeiro, como já mencionado, quando o efeito ambiental envolve época de semeadura ou safra, como ocorreu nesse caso. A presença do efeito de interação pode ser realçado pela correlação de Spearman das progênieis entre as gerações F_{4:5} e F_{4:6}. Observou-se que estas correlações para os diferentes critérios foram de baixa magnitude ($r_{S(MAjF_{4:5}, MAjF_{4:6})}=0,22$; $r_{S(BLUPF_{4:5}, BLUPF_{4:6})}=0,20$; $r_{S(BLUPAF_{4:5},$

$BLUP_{AF4:6}=0,33$). Essa pequena magnitude da correlação evidencia a ocorrência de interação e, inclusive, de natureza complexa (Cruz et al., 2004; Vencovsky e BARRIGA, 1992).

Outro fator foi a baixa precisão na avaliação das progênies $F_{4:5}$, pelas razões já comentadas, além da questão referente ao planejamento experimental dos experimentos em ambas as safras, o qual resultou no confundimento dos efeitos de bloco e tratamentos genéticos. Vale salientar que os melhoristas de feijão normalmente aplicam intensidade seletiva fraca ($s \cong 50\%$), apenas com o objetivo de eliminar as progênies com algum defeito ou de desempenho bem inferior às demais nessa primeira geração, em que é realizada a avaliação das progênies em experimentos com repetição em parcelas menores. Assim, as comparações anteriores serão mais pertinentes para maiores magnitudes das proporções selecionadas. Seria importante também continuar avaliando estas progênies para realizar comparações a partir de experimentos mais bem planejados, levando em conta a similaridade genética entre as progênies, e com maior número de repetições e tamanho de parcela.

5 CONCLUSÕES

A inclusão da informação da similaridade genética entre as progênies do método genealógico usando o procedimento $BLUP_A$ resultou em ganhos seletivos superiores aos alcançados quando esta genealogia foi ignorada, de forma a compensar o trabalho adicional despendido na obtenção desses registros, sobretudo, para caracteres de herdabilidade baixa.

Os métodos genealógico e *bulk*, nas condições simuladas, mostraram-se igualmente eficientes para a obtenção de combinações genotípicas com performances superiores, contudo, quando a informação da genealogia foi considerada, o método genealógico apresentou ganhos superiores.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, A. de F. B. RAMALHO, M. A. P.; SANTOS, J. B. dos; PEREIRA FILHO, I. A. Effects genotype x environment interaction on estimations of genetic and phenotypic parameters of common bean. **Revista Brasileira de Genética**, Ribeirão Preto, v. 13, n. 1, p. 75-82, mar. 1990.
- ALLARD, R. W. Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. **Journal of Heredity**, Cary, v. 79, n. 4, p. 225-238, July/Aug. 1988.
- ALLARD, R. W. **Principles of plant breeding**. 2. ed. New York: John Willey, 1999. 254 p.
- ANDRÉ, C. M. G. **Avaliação da melhor predição linear não tendenciosa (BLUP) associada ao uso de marcadores moleculares na análise dialélica**. 1999. 101 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.
- BALZARINI, M.; MILLIGAN, S. Best linear unbiased prediction (BLUP) for genotype performance. In: KANG, M. **Handbook of formulas and software for plant geneticists and breeders**. New York: The Haworth Press, 2003, p. 181-191.
- BEARZOTI, E. Emprego de marcadores em modelos mistos. In: SIMPÓSIO SOBRE ATUALIZAÇÃO EM GENÉTICA E MELHORAMENTO DE PLANTAS, 4., 2002, Lavras. **Anais...** Lavras: UFLA, 2002. p. 21-43.
- BEARZOTI, E. **Simulação de seleção recorrente assistida por marcadores moleculares em espécies autógamas**. 1997. 230 p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.
- BERNARDO, R. Best linear unbiased prediction of maize single cross performance. **Crop Science**, Madison, v. 36, n. 1, p. 50-56, Jan./Feb. 1996.

BERNARDO, R. **Breeding for quantitative traits in plants**. Woodbury: Stemma Press, 2002. 359 p.

BERNARDO, R. Correlation between tests cross performance of lines early and last selfing generation. **Theoretical and Applied Genetics**, Berlin, v. 82, n. 1, p. 17-21, 1991.

BERNARDO, R. Genetic models for predicting maize performance in unbalanced yield trial data. **Crop Science**, Madison, v. 35, n.1, p. 141-147, Jan./Feb. 1995.

BERNARDO, R. On the effectiveness of early generation selection in self-pollinated crops. **Crop Science**, Madison, v. 43, n. 4, p. 1558-1560, jul./ago. 2003.

BERNARDO, R. Prediction of maize single-cross performance using RFLPs and information from related hybrids. **Crop Science**, Madison, v. 34, n.1, p. 20-25, Jan./Feb. 1994.

BERTOLUCCI, F. L. G.; RAMALHO, M. A. P.; DUARTE, G. S. Alternativas de tamanho e forma da parcela para avaliação de progênies de feijoeiro. **Ciência e Prática**, Lavras, v. 15, n. 3, p. 295-305, jul./set. 1991.

BOERMA, H. H.; COOPER, R. L. Comparison of three selection procedures for yield in soybeans. **Crop Science**, Madison, v. 15, n. 2, p. 225-229, Mar./Apr. 1975.

BORÉM, A.; MIRANDA, G. V. **Melhoramento de plantas**. 4. ed. rev. amp. Viçosa: UFV, 2005. 525 p.

BROMLEY, C. M.; VAN VLECK, L. D.; JOHNSON, B. E.; SMITH, O. S. Estimation of genetic variance in corn from F₁ performance with and without pedigree relationships among inbred lines. **Crop Science**, Madison, v. 40, n. 3, p. 651-655, May/June 2000.

BUENO FILHO, J. S. de S. **Modelos mistos na predição de valores genéticos aditivos em testes de progênies florestais**. 1997. 118 p. Tese (Doutorado em

Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, São Paulo, SP.

BUENO FILHO, J. S. de S.; GILMOUR, S. G. Planning incomplete block experiments when treatments are genetically related. **Biometrics**, Washington, v. 59, n. 2, p. 375-381, June 2003.

BUSO, G. S. C.; RANGEL, P. H.; FERREIRA, M. E. Analysis of genetic variability of South America wild rice populations (*Oryza glumaepatula*) with isoenzymes and RAPD markers. **Molecular Ecology**, Oxford, v. 7, n. 1, p. 107-117, Jan. 1998.

CANTÚ, M. **Dominando o Delphi 6: A Bíblia**. São Paulo: MAKRON Books, 2002. 934 p.

CASALI, V. W. D.; TIGCHELAAR, E. C. Computer simulation studies comparing pedigree, bulk, and single seed descent selection in self pollinated populations. **Journal of American Society Horticultural Science**, Alexandria, v. 100, n. 4, p. 364-367, July 1975.

CASTRO, E. M. de; BRESEGHELLO, F.; RANGEL, P. H. N.; MORAES, O. P. de. Melhoramento do arroz. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. p. 103-140.

COLLICCHIO, E. **Associação entre o porte da planta do feijoeiro e o tamanho dos grãos**. 1995. 98 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

CONTE, S. D. **Elementos de análise numérica**. 3. ed. Porto Alegre: Globo, 1971. 331 p.

CORTE, H. R.; RAMALHO, M. A. P.; GONÇALVES, F. M. A.; ABREU, A. de F. B. A. Natural selection for grain yield in dry bean populations bred by the bulk method. **Euphytica**, Wageningen, v. 123, n. 3, p. 287-393, 2002.

COSTA, J. G. C. da; RAVA, C. A.; ZIMMERMANN, F. J. P. Comparação da eficiência de métodos de seleção em gerações segregantes de feijoeiro-comum

considerando a resistência à antracnose e o rendimento. **Ciência e Agrotecnologia**, Lavras, v. 26, n. 2, p. 244-251, mar./abr. 2002.

CROW, J. F. The rise and fall of overdominance. In: JANICK, J. **Plant breeding reviews**. New York: John Wiley, 2000. v. 17, p. 226-258.

CRUZ, C. D. A informática no melhoramento genético. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. de; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento – plantas**. Rondonópolis: Fundação MT, 2001. p. 1085-1118.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 3. ed. Viçosa: UFV, 2004. v. 1, 480 p.

CUTRIM, V. dos A.; RAMALHO, M. A. P.; CARVALHO, A. M. Eficiência da seleção visual na produtividade de grãos de arroz (*Oryza sativa* L.) irrigado. **Pesquisa Agropecuária Brasileira**, Brasília, v. 32, n. 6, p. 601-606, jun. 1997.

DACHS, J. N. W. **Estatística computacional: uma introdução ao Turbo Pascal**. Rio de Janeiro: LTC, 1988. 236 p.

DUARTE, J. B.; VENCOVSKY, R. Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agrícola**, Piracicaba, v. 58, n. 1, p. 109-117, jan./mar. 2001.

DUREL, C. E.; LAURENS, F.; FOUILLET, A.; LESPINASSE, Y. Utilization of pedigree information to estimate genetic parameters from large unbalanced data sets in apple. **Theoretical and Applied Genetics**, Berlin, v. 96, n. 8, p. 1077-1085, June 1998.

EISENHART, C. The assumptions behind the analysis of variance. **Biometrics**, Raleigh, v. 3, n. 1, p. 1-21, 1947.

FAHIM, M.; DHANAPALA, M. P.; SENADHIRA, D.; LAWRENCE, M. J. Quantitative genetics of rice II. A comparison of the efficiency of four breeding methods. **Field Crop Research**, Amsterdam, v. 55, n. 3, p.257-266, Feb. 1998.

FALCONER, D. S.; MACKEY, T. F. C. **Introduction to quantitative genetics**. 4. ed. London: Longman, 1996. 463 p.

FEDERER, W. T. Recovery of interblock, intergradient, and intervariatal information in incomplete block and lattice rectangle designed experiments. **Biometrics**, Washington, v. 54, n. 2, p. 471-481, June 1998.

FEDERIZZI, L. C.; SCHEEREN, P. L.; BARBOSA NETO, J. F.; MILACH, S. C. K.; PACHECO, M. T. Melhoramento do trigo. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. p. 659-697.

FEHR, W. R. **Principles of cultivar development: theory and technique**. New York: Macmillan, 1987. v. 1, 525 p.

FERREIRA, D. F. **Estatística básica**. Lavras: UFLA, 2005. 664 p.

FERREIRA, D. F. Simulação no melhoramento. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. de; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento – plantas**. Rondonópolis: Fundação MT, 2001. p. 1119-1141.

FERREIRA, W. D.; RAMALHO, M. A. P.; FERREIRA, D. F.; SOUZA, M. A. de. Family number in common bean selection. **Genetics and Molecular Biology**, Ribeirão Preto, v. 23, n. 2, p. 403-409, June 2000.

FERREIRA, D. F.; ZAMBALDE, A. L. Simplificação das análises de algumas técnicas especiais da experimentação agropecuária no Mapgen e softwares correlatos. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE INFORMÁTICA APLICADA À AGROPECUÁRIA E AGROINDÚSTRIA, 1., 1997, Belo Horizonte. **Anais...** Belo Horizonte, 1997. p. 285-291.

FOIULLOUX, G.; BANNEROT, H. Selection methods in the common bean (*Phaseolus vulgaris* L.). In: GEPS, P. **Genetic resources of phaseolus beans**. Dordrecht: Kluwer Academic, 1988. p. 503-541.

FUJITA, R.; OKASAKI, K.; SHIMAMOTO, Y. The extent of natural cross-pollination in wild soybean. **Journal of Heredity**, Cary, v. 88, n. 2, p. 124-128, Mar./Apr. 1997.

GONÇALVES, F. M. A.; RAMALHO, M. A. P.; ABREU, A. de F. B. Natural selection in four common bean traits. **Crop Breeding and Applied Biotechnology**, Londrina, v. 1, p. 213-220, 2001.

GURGEL, F. de L. **Simulação computacional no melhoramento genético de plantas**. 2004. 174 p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

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

HAMRICK, J. L.; GODT, J. W. Allozyme diversity in plant species. *In*: BROWN, A. W. D.; CLEGG, M. T.; KALHER, A. L.; WEIR, B. S. **Plant population genetics, breeding, and genetic resources**. Massachusetts: Sinauer Associates, 1990. p. 43-63.

HARVILLE, D. A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, Washington, v. 72, n. 358, p. 320-328, June 1977.

HEDRICK, P. W. **Genetics of populations**. 2. ed. London: Jones and Bartlett Publishers, 2000. 539 p.

HENDERSON, C. R. **Applications of liner models in animal breeding**. Guelph: University of Guelph Press, 1984. 462 p.

HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, Raleigh, v. 31, n. 2, p. 423-447, June 1975.

HENDERSON, C. R.; KEMPTHORNE, O.; SEARLE, S. R.; VON KROSIGK, C. M. The estimation of environmental and genetic trends from records subject to culling. **Biometrics**, Raleigh, v. 13, n. 1, p. 192-218, June 1959.

HORNER, T. W.; WEBER, C. R. Theoretical and experimental study of self fertilized populations. **Biometrics**, Raleigh, v. 12, n. 4, p. 404-414, Dec. 1956.

JENSEN, N. F. **Plant breeding methodology**. New York, 1988. 676 p.

KERVELLA, J.; FOUILLOUX, G. A theoretical study of the bulk breeding method 1. Importance and consequences of losses due to sampling. **Euphytica**, Wageningen, v. 60, n. 3, p. 185-195, Apr. 1992.

LITTLEL, R. C.; MILLEKEN, G. A.; STROUP, W. W.; WOLFINGER, R. D. **SAS[®] System for mixed models**. Cary: SAS Institute, 1996. 633 p.

LIU, P.; ZHU, J.; LU, Y. Marker-assisted selection in segregating generations in self-fertilizing crops. **Theoretical and Applied Genetics**, Berlin, v. 109, n. 2, p. 370-376, July 2004.

LOVELESS, M. D.; HAMRICK, J. L. Ecological determinants of genetic structure in plant populations. **Annual Review of Ecology and Systematics**, v. Palo alto, 15, p. 65-95, 1984.

LYNCH, M.; RITLAND, K. Estimation of pairwise relatedness with molecular markers. **Genetics**, Baltimore, v. 152, n. 4, p. 1753-1766, Aug. 1999.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, Inc., 1998. 948 p.

MALHEIROS, E. B. **Efeitos da recuperação da informação interblocos na inferência estatística em ensaios em blocos incompletos equilibrados**. 1982. 110 p. Tese (Doutorado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP.

MALLMANN, J. L.; BARBOSA NETO, J. F.; CARVALHO, F. I. F. De; FEDERIZZI, L. C. Mecanismos de seleção aplicados sobre o caráter tamanho de grãos em populações segregantes de trigo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 29, n. 3, p. 427-437, mar. 1994.

MARQUES JUNIOR, O. G.; RAMALHO, M. A. P. Determinação da taxa de fecundação cruzada de feijoeiro (*Phaseolus vulgaris* L.) nas diferentes épocas de semeadura em Lavras-MG. **Ciência e Prática**, Lavras, v. 19, n. 3, p. 339-341, jul./set. 1995.

MARTIN, S. K. St.; GERALDI, I. O. Comparison of three procedures for early generation testing of soybean. **Crop Science**, Madison, v. 42, n. 3, p. 405-709, May 2002.

MATOS, J. W. de. **Análise crítica do programa de melhoramento genético do feijoeiro da UFPA no período de 1974 a 2004**. 2005. 116 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Universidade Federal de Lavras, Lavras, MG.

McCULLOCH, C. E.; SEARLE, S. R. **Generalized, linear, and mixed models**. New York: John Wiley, 2001. 320 p.

MIRANDA FILHO, J. B. Endogamia ou consaguinidade. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. de; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento – plantas**. Rondonópolis: Fundação MT, 2001. p. 629-648.

MISHA, D. K.; SINGH, C. B.; RAO, S. K. Effectiveness of different selection methods in segregating population of rice (*O. sativa* L.) in ACR 10372 x IR 36 in different environment. **Indian Journal Genetics**, New Delhi, v. 54, n. 4, p. 402-408, Dec. 1994.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. Tokyo: McGraw-Hill, 1974. p. 556.

MORETO, A. L. **Componentes da variância fenotípica em feijoeiro utilizando o método genealógico**. 2005. 75 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

MORGAN, B. J. T. **Elements of simulation**. London: Chapman & Hall, 1995. 351 p.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. Guildford: Biddles, 1996. 184 p.

MUEHLBAUER, F. J.; BURNELL, D. G.; BOGYO, T. P.; BOGYO, M. T. Simulated comparisons of single seed descent and bulk population breeding methods. **Crop Science**, Madison, v. 21, n. 4, p. 572-577, jul./ago. 1981.

NOVOSELOVIC, D.; BARIC, M.; DREZNER, G.; GUNJACA, J.; LALIC, A. Quantitative inheritance of some wheat plant traits. **Genetics and Molecular Biology**, Ribeirão Preto, v. 27, n. 1, p. 92-98, Mar. 2004.

NTANOS, D. A.; ROUPAKIAS, D. G. Comparative efficiency of two breeding methods for yield and quality in rice. **Crop Science**, Madison, v. 41, n. 2, p. 345-350, Mar./Apr. 2001.

PANTER, D. M.; ALLEN, F. L. Using best linear unbiased predictions to enhance breeding for yield in soybean: I choosing parents. **Crop Science**, Madison, v. 35, n. 2, p. 397-405, Mar./Apr. 1995a.

PANTER, D. M.; ALLEN, F. L. Using best linear unbiased predictions to enhance breeding for yield in soybean: II Selection of superior crosses from a limited number of yield trials. **Crop Science**, Madison, v. 35, n. 2, p. 405-410, Mar./Apr. 1995b.

PATIÑO, H.; SINGH, S. P. Visual selection for seed yield in the F₂ and F₃ generations of nine common bean crosses. **Annual Report Bean Improvement Cooperative**, New York, v. 32, p. 79-80, 1989.

PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, London, v. 58, n. 3, p. 545-554, 1971.

PIEPHO, H. P. Best linear unbiased prediction (*BLUP*) for regional yield trials: a comparison to additive main effects and multiplicative interaction (*AMMI*) analysis. **Theoretical and Applied Genetics**, Berlin, v. 89, n. 5, p. 647-654, Nov. 1994.

PIROLA, L. H.; RAMALHO, M. A. P.; CARNEIRO, J. E. de S.; ABREU, A. de F. B. Natural selection and family x location interaction in the common (dry) bean plant. **Genetics and Molecular Biology**, Ribeirão Preto, v. 25, n. 3, p. 343-347, Sept. 2002.

PODLICH, D. W.; COOPER, M. *QU-GENE*: a platform for quantitative analysis of genetic models. **Bioinformatics**, Oxford, v. 14, n. 7, p. 632-653, 1998.

RAMALHO, M. A. P. Melhoramento do feijoeiro. In: SIMPÓSIO SOBRE ATUALIZAÇÃO EM GENÉTICA E MELHORAMENTO DE PLANTAS, 1., 1997, Lavras. **Anais...** Lavras: UFLA, 1997. p. 167-196.

RAMALHO, M. A. P.; ABREU, A. F. B.; SANTOS, J. B. dos. Melhoramento de espécies autógamas. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S. de; VALADARES-INGLIS, M. C. **Recursos genéticos e melhoramento – plantas**. Rondonópolis: Fundação MT, 2001. p. 201-230.

RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. **Experimentação em genética e melhoramento de plantas**. 2. ed. Lavras: UFLA, 2005. 326 p.

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

RAMALHO, M. A. P.; VENCOSKY, R. Estimação de componentes da variância genética em plantas autógamas. **Ciência e Prática**, Lavras, v. 2, n. 2, p. 117-140, jul./dez. 1978.

RANALLI, P.; RUARO, G.; DEL RE, P.; FAETI, V. Comparison of early generation yield testing and a single seed descent procedure in two bean (*Phaseolus vulgaris* L.) crosses. **Journal of Genetics and Breeding**, Rome, v. 50, n. 2, p. 103-108, June 1996.

RAPOSO, F. C.; RAMALHO, M. A. P.; ABREU, A. de F. B. Comparação de métodos de condução de populações segregantes do feijoeiro. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 10, p. 1991-1997, out. 2000.

RAPOSO, F. V. **Comparação de métodos de condução de populações segregantes de feijoeiro**. 1999. 72 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

RASMUSSEN, D. C.; PHILLIPS, R. L. Plant breeding progress and genetic diversity from *de novo* variation and elevated epistasis. **Crop Science**, Madison, v. 37, n. 2, p. 303-310, mar./abr. 1997.

RESENDE, M. D. V. de. **Efeitos fixos ou aleatórios de repetições no contexto dos modelos mistos no melhoramento de plantas perenes**. Colombo: Embrapa Florestas, 2002b. 15 p. (Embrapa Florestas. Documentos, 68).

RESENDE, M. D. V. de. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: EMBRAPA Informação tecnológica, 2002a. 975 p.

ROBINSON, G. K. That BLUP is a good thing: the estimation of random effects. **Statistical Science**, Hayward, v. 6, n. 1, p. 15-51, Feb. 1991.

ROSAL, C. J. de S.; RAMALHO, M. A. P.; GONÇALVES, F. M. A.; ABREU, A. de F. B. Seleção precoce para a produtividade de grãos no feijoeiro. **Bragantia**, Campinas, v. 59, n. 2, p. 189-195, 2000.

SAKIYAMA, N. S.; PEREIRA, A. A.; MOURA, W. M.; ZAMBOLIM, L. Melhoramento do café arábica. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. p. 203-223.

SAS/STAT® 8.0 User'Guide. Cary, NC, USA, 1999.

SATTERTHWAITE, F. E. An approximate distribution of estimates of variance components. **Biometrics Bulletin**, Washington, v. 2, p. 110-114, 1946.

SCHAEFFER, L. R. C. R. Henderson: Contributions to predicting genetic merit. **Journal of Dairy Science**, Champaign, v. 74, n. 11, p. 4052-4066, Nov. 1991.

SEARLE, S. R. **Linear models**. New York: John Wiley, 1971. 532 p.

SEARLE, S. R.; CASELA, G.; McCULLOCH, C. E. **Variance components**. New York: John Wiley, 1992. 492 p.

SEDIYAMA, T.; TEIXEIRA, R. de C.; REIS, M. S. Melhoramento da soja. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. p. 553-603.

SILVA FILHO, J. L. da. **Estratégias de predição de cruzamentos não realizados em análise dialética, utilizando informações de marcadores moleculares**. 2004. 70 p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

SILVA, H. D.; RAMALHO, M. A. P.; ABREU, A. de F. B.; MARTINS, L. A. Efeito da seleção visual para produtividade de grãos em populações segregantes do feijoeiro. II. Seleção entre famílias. **Ciência e Prática**, Lavras, v. 18, n. 2, p. 181-185, abr./jun. 1994.

SILVA, N. O.; RAMALHO, M. A. P.; ABREU, A. de F. B.; CARNEIRO, J. E. de S. Performance of common bean families after different generations under natural selection. **Genetics and Molecular Biology**, Ribeirão Preto, v. 27, n. 4, p. 574-578, Dec. 2004.

SINGH, R. P.; RAJARAM, S.; MIRANDA, A.; HUERTA ESPINO, J.; AUTRIQUE, E. Comparison of two crossing and four selection schemes for yield, yield traits and slow rusting resistance to leaf rust in wheat. **Euphytica**, Wageningen, v. 100, n. 2, p. 35-43, 1998.

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

SOUZA, G. A.; RAMALHO, M. A. P. Estimates of genetic and phenotypic variance of some traits of dry bean using a segregating population from the cross Jalo x Small White. **Revista Brasileira de Genética**, Ribeirão Preto, v. 18, n. 1, p. 87-91, mar. 1995.

STEEL, R. G. D.; TORRIE, J. H.; DICKEY, D. A. **Principles and procedures of statistics: A biometrical approach**. 3. ed. New York: McGraw Hill, 1997. 656 p.

TAKEDA, C.; SANTOS, J. B. dos; RAMALHO, M. A. P. Choice of parental lines for common bean (*Phaseolus vulgaris* L.) breeding. II. Reaction of cultivars and of their segregating populations to variations in different

environments. **Revista Brasileira de Genética**, Ribeirão Preto, v. 14, n. 2, p. 455-465, jun. 1991.

URREA, C. A.; SINGH, S. P. Comparison of mass, F_2 – derived family, and single-seed-descent selection method in a interracial population of common bean. **Canadian Journal of Plant Science**, Ottawa, v. 74, n. 3, p. 461-464, July 1994.

VAN OEVEREN, A. J.; STAM, P. Comparative simulation studies on the effects of selection for quantitative traits in autogamous crop: early selection versus single seed descent. **Heredity**, Cary, v. 69, n. 4, p. 342-351, Oct. 1992.

VARGAS, M. A. **Eficiência da seleção visual e determinação do tamanho de parcela e número de repetições para avaliação da produtividade de grãos em família S_1 , S_2 de milho (*Zea mays* L.)**. 1996. 92 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, MG.

VEIGA, R. D.; FERREIRA, D. F.; RAMALHO, M. A. P. Eficiência dos dialelos circulantes na escolha de genitores. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 7, p. 1395-1406, jul. 2000.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496 p.

VIEIRA, C.; BORÉM, A.; RAMALHO, M. A. P.; CARNEIRO, J. E. de S. Melhoramento do feijão. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Viçosa: UFV, 2005. p. 301-392.

VIRMANI, S. S. Exploitation of heterosis for shifting the yield frontier in rice. In: COORS, J. G.; PANDEY, S. **The genetics and exploitation of heterosis in crops**. Madison: American Society of Agronomy/Crop Science Society, 1999. p. 423-438.

VOIGT, R. L.; WEBER, C. R. Effectiveness of selection methods for yield in soybean crosses. **Agronomy Journal**, Madison, v. 52, n. 9, p. 527-530, Sept. 1960.

WANG, J.; VAN GINKEL, M.; PODLICH, D.; YE, G.; TRETOWAN, R.; PFEIFFER, W.; DELACY, I. H.; COOPER, M.; RAJARAM, S. Comparison of two breeding strategies by computer simulation. **Crop Science**, Madison, v. 43, n. 6, p. 1764-1773, Nov./Dec. 2003.

WEYHRICH, R. A.; LAMKEY, K. R.; HALLAUER, A. R. Responses to seven methods of recurrent selection in the BS11 maize population. **Crop Science**, Madison, v. 38, n. 2, p. 308-314, Mar/Abr. 1998.

WHITE, T. L.; HODGE, G. R. **Predicting breeding values with applications in forest tree improvement**. Dordrecht: Kluwer Academic Publishers, 1989. 363 p.

WRICKE, G.; WEBER, W. E. **Quantitative genetics and selection in plant breeding**. Berlin: Walter de Gruyter, 1986. 395 p.

ANEXO

1 Rotina implementada no programa SAS 8.0 usando o *Proc Mixed* para análise de experimentos em látice (16) com recuperação da informação interblocos e interprogênes, considerando a informação do parentesco genético aditivo entre progênes (Adaptado de Balzarini e Milligan (2003))

EXEMPLO: Experimento em látice simples 16 x 16 referente à avaliação de progênes F_{4:5} conduzidas pelo método genealógico, conforme descrito no item 3.1.2.1. Os coeficientes de parentesco de Malecot entre progenies foram determinados a partir da informação da genealogia, assumindo a geração F₂ como não endogâmica.

```
options nodate nonumber;
data dados;
input repeticao bloco progenie producao;
datalines;
1      1      1      260.00
1      1      2      200.00
:      :      :      :
:      :      :      :
2      15     255     330.00
2      16     256     180.00
;
run;quit;

proc iml;

    b=32; * entrar com o número de blocos;
    p=256; * entrar com o número de progênes avaliadas;

    *****;
    ** Entrar com a matriz dos parentescos de Malecot (A) entre progenies **;
    *****;
    CoefMalecot={
        0.875  0.750  0.500  0.500  ,
        0.750  0.875  0.500  0.500  ,
        0.500  0.500  0.875  0.750  ,
        0.500  0.500  0.750  0.875
    };
```

```

*****
***** Matriz dos parentescos genéticos aditivos (A) entre progenies *****
*****
      A=2#I(64)@ CoefMalecot;

*****
***** Matriz associada aos efeitos aleatórios de blocos *****
*****
      Ab=I(b);

*****
***** Criar Data Set - Ldata *****
**Contém a matriz dos coeficientes da covariância dos efeitos aleatorios**
*****
      Id=1:b+p;
      matriz=j(b+p,b+p+2,0);
      matriz[,1]=j(b,1,1)//j(p,1,2);
      matriz[,2]=t(Id);
      matriz[1:b,3:b+2]=Ab;
      matriz[b+1:b+p,b+3:b+p+2]=A;
      create novo from matriz;
      append from matriz;

data ldata;
set novo;
rename col1=Parm;
rename col2=row;
rename col3-col290=col1-col288;
run;quit;

proc mixed data=dados covtest cl;
title1 'Ajuste de Modelos Mistos: Proc Mixed SAS';
title2 'Análise com informação de parentesco - Progenies F4:5';
class repeticao bloco progenie;
model producao = repeticao;
random bloco(repeticao) progenie //ldata=ldata type=lin(2) solution;
run;quit;

```