

**ÍSIS FERNANDA DE ALMEIDA**

**MELHOR PREDIÇÃO LINEAR NÃO VIESADA (BLUP) NA SELEÇÃO  
DENTRO DE PROGÊNIES NÃO ENDÓGAMAS DE ESPÉCIES ANUAIS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS – BRASIL  
2009

ÍSIS FERNANDA DE ALMEIDA

**MELHOR PREDIÇÃO LINEAR NÃO VIESADA (BLUP) NA SELEÇÃO  
DENTRO DE PROGÊNIES NÃO ENDÓGAMAS DE ESPÉCIES ANUAIS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

Aprovada: 19 de Fevereiro de 2009.

---

Dr. Marcos Deon Vilela de Resende

---

Prof. Adair José Regazzi

---

Dr. Paulo Bonomo

---

Prof. Glauco Vieira Miranda

---

Prof. José Marcelo Soriano Viana  
(Orientador)

Aos meus pais Iracy e Cely  
Ao meu irmão Ramon.

## **AGRADECIMENTOS**

A Deus por conceder o dom da vida e inspirar-me em todos instantes.

À Universidade Federal de Viçosa, pela oportunidade de realizar este curso.

Ao CNPq, pelo auxílio financeiro.

Ao meu orientador, professor José Marcelo Soriano Viana pela orientação e pelo apoio.

Aos amigos, Fábio, Magno, Keny, Admilson, pela colaboração e pela agradável convivência.

Às ex-companheiras de república, Josiane, Ana Paula e Vanessa pela amizade e pelos momentos de alegria compartilhados.

Aos professores e colegas do curso de pós-graduação em genética.

Às amigas Daniela, Adalgisa e Pâmela pelo carinho, apoio e companheirismo.

Ao amigo Vinícius por sua amizade, colaboração, disponibilidade e por suas valiosas sugestões.

Aos funcionários Vicente, Márcio e Antônio, pelos esforços despendidos nos trabalhos de campo.

Às funcionárias da secretaria, Rose e Rita, pela atenção e pela amizade.

Ao meu Irmão, Ramon, que me trouxe para Viçosa e que sem o qual, seria muito mais difícil estar onde estou hoje.

Aos membros da banca, Drs. Marcos Deon Vilela de Resende, Paulo Bonomo e professores Adair José Regazzi e Glauco Vieira Miranda pela disponibilidade e colaboração.

A todos os que colaboraram para realização deste trabalho, minha sincera gratidão.

## **BIOGRAFIA**

ÍSYS FERNANDA DE ALMEIDA, filha de Iracy de Almeida e Cely Aparecida de Almeida, nasceu na cidade de Volta Redonda - RJ, em 09 de novembro de 1984.

Realizou os estudos do primeiro e segundo graus, na Escola Estadual Getúlio Vargas e na Escola Técnica Pandiá Calógeras, respectivamente, até o ano de 2001.

Ingressou no curso de Agronomia no ano de 2002 na Universidade Federal de Lavras.

Em março de 2007, obteve o diploma de Engenheira Agrônomo pela Universidade Federal de Lavras.

Em 2007, iniciou no Programa de Mestrado em Genética e Melhoramento, pela Universidade Federal de Viçosa, com ênfase em Genética Quantitativa e Melhoramento de Milho-Pipoca, submetendo-se à defesa de tese em fevereiro de 2009.

## SUMÁRIO

|                                         |            |
|-----------------------------------------|------------|
| <b>RESUMO .....</b>                     | <b>vii</b> |
| <b>ABSTRACT .....</b>                   | <b>ix</b>  |
| <b>1. INTRODUÇÃO .....</b>              | <b>1</b>   |
| <b>2. MATERIAL E MÉTODOS .....</b>      | <b>3</b>   |
| <b>2.1. Aspectos teóricos .....</b>     | <b>3</b>   |
| <b>2.2. Dados Experimentais .....</b>   | <b>6</b>   |
| <b>2.3. Análises Estatísticas .....</b> | <b>7</b>   |
| <b>3. RESULTADOS .....</b>              | <b>9</b>   |
| <b>4. DISCUSSÃO .....</b>               | <b>11</b>  |
| <b>5. REFERÊNCIAS .....</b>             | <b>13</b>  |
| <b>6. ANEXO .....</b>                   | <b>19</b>  |

## RESUMO

ALMEIDA, Ísis Fernanda de M.Sc., Universidade Federal de Viçosa, fevereiro de 2009.  
**Melhor Predição Linear Imparcial (BLUP) no melhoramento vegetal: seleção dentro de famílias não endogâmicas.** Orientador: José Marcelo Soriano Viana. Co-orientadores: Fabyano Fonseca e Silva e Cosme Damião Cruz.

A Melhor Predição Linear Não Viesada (BLUP) tornou-se a metodologia mais empregada na avaliação genética animal e de espécies perenes, sendo potencialmente relevante para espécies anuais. O objetivo desse trabalho foi avaliar a metodologia na seleção dentro de famílias não endogamas, em programas de seleção recorrente. Ajustou-se o modelo animal. Os dados foram valores fenotípicos de plantas nos lotes de recombinação de dois a três ciclos de seleção na população de milho-pipoca Viçosa, com famílias não endogamas. Em cada planta foram mensuradas a produção e a capacidade de expansão. Nas análises foi utilizado o programa SAS, por meio dos procedimentos *inbreed* e *mixed*. Houve redução da variabilidade genética relativa à CE e produção com os ciclos de seleção. Os valores aditivos preditos por BLUP/REML foram mais acurados que os valores fenotípicos. A análise multi-gerações foi mais acurada que a análise por geração apenas em relação à CE. Esses dois métodos proporcionaram valores aditivos altamente correlacionados. Contudo, a porcentagem

de coincidência entre os indivíduos selecionados revela diferença entre os métodos BLUP. Na população estruturada em famílias de meios-irmãos o método de melhoramento mais eficiente foi seleção massal. Na população estruturada em famílias de irmãos completos destacou-se a seleção dentro de progênies.

## ABSTRACT

ALMEIDA, Ísis Fernanda de M. Sc., Universidade Federal de Viçosa, February of 2009. **Best Unbiased Linear Prediction (BLUP) in plant breeding: selection inside non-endogamic families.** Adviser: José Marcelo Soriano Viana. Co-advisers: Fabyano Fonseca e Silva and Cosme Damião Cruz.

The *Best Linear Unbiased Prediction* (BLUP) became the most employed methodology in the genetic evaluation of animals and perennial species, with special relevance for annual species. The objective of this work was to evaluate this methodology in the selection inside non-endogamous families, in recurring selection programs. The animal model was adjusted. The data were phenotypical values of plants in the recombination lots of two to three selection cycles in the population of the popcorn maize Viçosa. The production and expansion capacity were measured in each plant. The SAS software system was used in the analyses by means of the *inbreed* and *mixed procedures*. There was a reduction in the genetic variability related to the CE and the production with the selection cycles. The additive values predicted by BLUP/REML were more accurate than the phenotypical values. The multi-generation analysis was more accurate than the analysis by generation only in relation to the CE. These two methods provided highly correlated additive values. However, the

percentage of coincidence among the individuals selected reveals a difference between the BLUP methods. In the population structured into families of half-brothers, the most efficient breeding method was the massal selection. In the population structured into families with full brothers, the selection inside progenies stood out.

## **Introdução**

A seleção recorrente visa o acréscimo contínuo na frequência dos genes favoráveis. Sua eficiência, qualquer que seja a unidade de seleção, depende da correta identificação dos indivíduos com melhor valor genético aditivo. Uma metodologia adequada para a predição de valores genéticos aditivos de indivíduos sob seleção, de uso comum no melhoramento animal e de espécies perenes, é a Melhor Predição Linear Não Viesada, proposta por Henderson (1974). Idealmente a predição é feita levando em conta o parentesco entre os indivíduos, usando estimativas dos componentes de variância do modelo (genéticos e não genéticos), com correção para os efeitos fixos.

Com relação à utilização de BLUP em culturas anuais, a aplicação mais relevante até recentemente era a predição do desempenho de híbridos não avaliados, proposta por Bernardo (1996), a qual é possível apenas quando suficientes informações de parentesco entre as linhagens estão disponíveis. Em seu trabalho com 4099 híbridos de 16 combinações entre nove grupos heteróticos de milho, avaliados em até 15 locais, ele encontrou estimativas de correlação entre o valor predito e o desempenho fenotípico, para produção de grãos, variando de 0,49 a 0,70, com uma tendência de estimativas mais elevadas para os grupos com maior número de híbridos avaliados. Segundo o autor, estes valores podem ser considerados moderadamente elevados e sugerem que o BLUP pode ser eficiente para identificação de cruzamentos não realizados, com desempenho superior.

Ao longo dos últimos cinco anos outras aplicações relevantes, de uso comum no melhoramento animal e de espécies perenes, foram incorporadas aos programas de melhoramento de anuais alógamas e autógamas. Piepho et al. (2008) analisaram o uso da metodologia BLUP no melhoramento de plantas, demonstrando sua relevância, mesmo sem o uso da informação de pedigree. Dentre as aplicações destacam-se a avaliação genética de indivíduos (predição de valores genéticos aditivos), e de linhas puras e linhagens (predição de

valores genotípicos e de efeitos de capacidade geral de combinação), com e sem o uso das informações de pedigree, em análises uni e multicaracterística. Flachenecker (2006a, b) empregaram BLUP na predição dos valores genéticos aditivos médios dos pais de famílias de irmãos completos de uma população de milho, os quais permitiram satisfatória predição do desempenho das progênies do ciclo subsequente, em relação à umidade dos grãos e/ou ao caráter índice de seleção, calculado como função da produção de grãos e do conteúdo de matéria-seca. A partir da análise de genótipos de trigo em vários ambientes, Crossa et al. (2006) concluíram que a predição de valores genéticos por BLUP com informação de parentesco foi mais precisa, comparativamente às análises tradicionais, assumindo efeitos fixos e efeitos aleatórios. Reis et al. (2005) empregaram BLUP na análise dialélica de linhagens de milho, em combinação com a técnica de bootstrap. Eles verificaram, em relação à produção, que a correlação entre as médias preditas e observadas dos híbridos diminuiu com a redução do número de genitores, destacando que a redução foi menos acentuada pelo emprego do método BLUP, comparativamente ao método BLUE (modelo fixo).

O mérito da inclusão das informações de pedigree foi investigado por Nunes et al. (2008), empregando simulação. Eles concluíram que a inclusão das informações de parentesco proporcionaram maiores ganhos preditos. Bauer et al. (2006), que utilizaram informação de marcas moleculares para o cálculo da matriz de parentesco, concluíram que a metodologia BLUP pode ser empregada no melhoramento de alógamas, na seleção de linhas puras. O mesmo foi verificado por Sebolai et al. (2005), a partir de simulação, na seleção precoce de linhas puras. Em outro estudo de simulação, Bauer e Léon (2008) analisaram a eficiência da seleção de linhas superiores pelo BLUP multicaracterística. Os resultados evidenciaram que essa análise foi superior à unicaracterística, quanto à identificação das linhas superiores e acurácia das predições. Kruuk (2004) empregou o modelo animal na estimação de parâmetros genéticos de populações naturais, concluindo que a metodologia

BLUP permitiu o uso mais eficiente dos dados disponíveis, além de proporcionar maior poder estatístico dos testes.

Visando contribuir para que essa metodologia passe a ser regularmente empregada pelos melhoristas de espécies anuais, autógamas e alógamas, o objetivo desse trabalho foi avaliar seu emprego na seleção de plantas em famílias não endógamas, em programas de melhoramento populacional recorrente.

## **Material e Métodos**

### *Aspectos Teóricos*

A eficiência da seleção em programas de melhoramento populacional recorrente é proporcional à eficiência de identificação dos genitores de melhor valor genético aditivo, uma vez que os demais componentes do valor genotípico não têm correlação com o número de genes favoráveis presentes. Embora valor genotípico (G) e valor genético aditivo (A) sejam correlacionados com o número de genes favoráveis (N), em geral a correlação de A com N é maior ou igual (na ausência de dominância) à correlação de G com N (Viana 2005).

Admitindo ausência de epistasia tem-se

$$G = M + A + D$$

sendo M a média populacional e D o valor genético devido à dominância.

Então, o valor fenotípico de um indivíduo é

$$P = M + A + D + \sum_k (\text{efeito fixo})_k + \text{erro}$$

Em termos matriciais

$$y = X\beta + Zu_1 + Zu_2 + \varepsilon$$

em que y é o vetor de valores fenotípicos, X é a matriz de incidência dos efeitos fixos,  $\beta$  é o vetor dos níveis de cada fator fixo (incluindo a média populacional), Z é a matriz de incidência dos efeitos aleatórios,  $u_1$  é o vetor dos valores genéticos aditivos,  $u_2$  é o vetor dos valores genéticos devidos à dominância e  $\varepsilon$  é o vetor de resíduos. Esse modelo é conhecido

como modelo animal. O critério de obtenção do melhor preditor linear não-viesado de um vetor aleatório é a maximização da função densidade de probabilidade conjunta de  $y$  e dos vetores aleatórios, ou da função densidade de probabilidade conjunta dos vetores aleatórios e  $\varepsilon$ , obtendo, sob normalidade, as equações de modelo misto (Henderson 1974).

Admitindo população não endógama ( $\text{Cov}(u_1, u_2) = \phi$ ) e que não há correlação entre os erros ( $\text{Var}(\varepsilon) = R = I\sigma^2$ ), as equações de modelo misto são

$$\begin{bmatrix} X'X & X'Z & X'Z \\ Z'X & ZZ + \sigma^2 G_1^{-1} & Z'Z \\ Z'X & Z'Z & Z'Z + \sigma^2 G_2^{-1} \end{bmatrix} \begin{bmatrix} \beta^0 \\ \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ Z'y \end{bmatrix}$$

em que  $G_1 = \sigma_A^2 A_{(n)}$  e  $G_2 = \sigma_D^2 D_{(n)}$  são as matrizes de variâncias dos vetores de valores genéticos aditivos e devidos à dominância.  $A_{(n)}$  e  $D_{(n)}$  são as matrizes de parentesco genético aditivo e de parentesco genético devido à dominância, dadas por

$$A_{(n)} = \begin{bmatrix} 2r_{11} & 2r_{12} & \cdots & 2r_{1n} \\ 2r_{12} & 2r_{22} & \cdots & 2r_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ 2r_{1n} & 2r_{2n} & \cdots & 2r_{nn} \end{bmatrix} \text{ e } D_{(n)} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{12} & u_{22} & \cdots & u_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{bmatrix}$$

em que  $r_{ij}$  é o coeficiente de parentesco entre os indivíduos  $i$  e  $j$  e  $u_{ij} = u_{(mf)(m'f')} = r_{mm'}r_{ff'} + r_{mf'}r_{fm'}$  é a probabilidade dos indivíduos  $i$  e  $j$  terem genótipos idênticos por descendência, sendo  $m$  e  $f$  os pais de  $i$ , e  $m'$  e  $f'$  os pais de  $j$  (Lynch e Walsh 1988).

O ajuste do modelo, por meio de um processo iterativo, proporciona a estimação dos componentes de variância, a predição de cada vetor aleatório e uma solução para o vetor de efeitos fixos. O processo mais empregado para a estimação dos componentes de variância é o da máxima verossimilhança restrita/residual (REML), desenvolvido por Patterson e Thompson (1971). Alternativamente podem ser empregados os métodos da máxima

verossimilhança (ML), da estimação quadrática não viesada de variância mínima (MIVQUE) e o método III de Henderson (Littell et al. 2006).

Embora o modelo descrito sugira ser sempre possível a predição dos valores genéticos aditivos e devidos à dominância, isso infelizmente não é verdade. Por exemplo, no caso de espécies anuais alógamas, como o milho, se a população no lote de recombinação está estruturada em famílias de meios-irmãos, não é possível o ajuste do modelo aditivo-dominante, pois em cada ciclo a matriz de parentesco genético de dominância é uma identidade. Então, como  $G_2 = \sigma_D^2 I$  e  $R = \sigma^2 I$  não há convergência no processo REML. Isso não é necessariamente verdade no caso de animais e espécies perenes. Se a população está estruturada em famílias de irmãos completos, mesmo que a matriz de parentesco de dominância não seja uma identidade, não está garantido ser possível o ajuste do modelo aditivo-dominante. Em geral, para que o modelo aditivo-dominante possa ser ajustado é necessário existir pelo menos uma covariância entre parentes que seja função da variância aditiva e uma covariância entre parentes que seja função da variância aditiva e da variância de dominância. Então, em relação aos indivíduos no lote de recombinação do primeiro ciclo de seleção com famílias de irmãos completos, como só existe uma covariância entre parentes, a de irmãos completos, também não é possível ajustar o modelo aditivo-dominante. Para ajustar o modelo aditivo-dominante e aumentar a acurácia das predições relativas aos indivíduos do segundo ciclo, por exemplo, pode-se considerar em conjunto os indivíduos dos lotes de recombinação dos ciclos 1 e 2 (análise multi-gerações).

A acurácia de predição dos valores genéticos aditivos é

$$r_{A\tilde{A}} = \sqrt{1 - \left( \frac{PEV}{\tilde{\sigma}_A^2} \right)}$$

sendo PEV a variância do erro de predição e  $\tilde{\sigma}_A^2$  o estimador de máxima verossimilhança restrita da variância aditiva (Mrode 2005).

Em cada ciclo foram consideradas duas estratégias de seleção: seleção de 10 plantas em cada uma das vinte famílias selecionadas com base no teste e seleção de 200 plantas. Os ganhos genéticos diretos e indiretos foram calculados com base no controle parental ( $p = 1/2$ , pois a seleção ocorreu após o florescimento) e nos valores genéticos aditivos preditos pela metodologia BLUP/REML, pela função

$$\Delta M = (1/2) \left( \frac{1}{200} \sum_i \tilde{A}_i \right)$$

### *Dados Experimentais*

Foram analisados os valores fenotípicos de capacidade de expansão (EV; mL/g) e de produção de grãos (g/0,18 m<sup>2</sup>), de plantas nos lotes de recombinação dos ciclos 1, 2 e 3, do programa de seleção com famílias de meios-irmãos, e dos ciclos 1 e 2, do programa de seleção com famílias de irmãos completos, ambos com a população de milho-pipoca Viçosa. Os lotes com famílias de meios-irmãos foram conduzidos na estação experimental da Universidade Federal de Viçosa (UFV), em Ponte Nova-MG, Brazil, nos anos agrícolas de 2000/01, 2002/03 e 2004/05. Em cada lote foram plantadas 20 famílias selecionadas com base no teste de 196 progênies, intercalando uma fileira de macho a cada quatro fileiras de fêmeas. Também foram plantadas duas fileiras do híbrido simples IAC 112 (testemunha). Cada fileira tinha 20 m, com até 100 plantas. Por ocasião do florescimento as fileiras de fêmeas e da testemunha foram despendoadas. Os lotes com progênies de irmãos completos foram conduzidos em campo experimental na UFV, em Viçosa-MG, nos anos agrícolas de 2003/04 e 2005/06. Em cada lote também foram plantadas 20 famílias selecionadas com base no teste de 196 progênies. Cada fileira tinha 10 m, com até 50 plantas. Por ocasião do florescimento fez-se controle de cruzamentos, de forma a serem obtidas pelo menos 19 famílias de irmãos completos em cada progênie (um macho de cada uma das demais 19 progênies. Infelizmente as plantas usadas como machos não foram identificadas nem mensuradas.

Nos lotes com famílias de meios-irmãos foram mensuradas 20 plantas em cada progênie e 10 plantas em cada fileira do IAC 112. Nos lotes com famílias de irmãos completos foram mensuradas 324 e 390 plantas. A EV foi avaliada em pipoqueira de ar quente (1,200 w), utilizando amostras de 10 g.

#### *Análises Estatísticas*

Para as análises foram utilizados os softwares SAS v9.1, por meio dos procedimentos *inbred* e *mixed*, e ASReml (Gilmour et al. 2006). Como o *proc mixed* da versão 9.1 do SAS não permite a entrada da matriz de parentesco de dominância, o SAS serve atualmente apenas para a análise do modelo aditivo. O *proc inbred* calcula a matriz de parentesco genético aditivo e o *proc mixed* a utiliza para análise pelo método BLUP/REML. O SAS program utilizado na análise dos dados dos lotes com famílias de meios-irmãos foi

```
data file1;
input individual male female;
datalines;
pedigree file
;
proc inbred data= file1 covar noprint outcov= file2;
var individual male female;
run;
data cAmatrix; set file2(firstobs= n1); keep coln1-coln2;
run;
data Amatrix; set cAmatrix; rename coln1-coln2 = coln3-coln4;
parm= 1;
row= _n_;
run;

data file3;
input population plant EV yield;
datalines;
data file
;
proc mixed data= file3 method= REML covtest cl;
class population plant;
model one dependent variable (EV or yield) = population / solution ddfm= kr;
```

```
random plant / solution type= lin(1) ldata= Amatrix;
run;
```

Informações sobre os comandos devem ser obtidas em Littell et al. (2006). Os termos n1 e n2 referem-se ao primeiro e ao último indivíduos considerados na análise. Como os dados do IAC 112 foram utilizados nas análises da população estruturada em famílias de meios-irmãos, os números foram 197 e 597 na análise dos dados do lote do ciclo 1, recodificando o IAC 112 como indivíduo 597 e eliminando os demais indivíduos (ver pedigrees completos no Anexo 1). Na análise dos dados do lote do ciclo 2 os números foram 597 e 997, agora recodificando o IAC 112 como indivíduo 997 e eliminando os demais indivíduos. Em relação ao ciclo 3 os números foram 997 e 1397, recodificando o IAC 112 como indivíduo 1397. Na análise combinando os ciclos 1 e 2, por exemplo, os números foram 197 e 997, recodificando o IAC 112 como indivíduo 997. Em todos os casos o IAC 112 foi inserido no pedigree como um indivíduo de pais desconhecidos. Os termos n3 e n4 são os códigos dos indivíduos no arquivo de dados. Na análise dos dados dos ciclos 1, 2 e 3 os valores foram 1 e 401, 1 e 801 e 1 e 1201, recodificando as plantas do IAC 112 como planta 401, 801 e 1201, respectivamente, e eliminando os demais valores fenotípicos. Na análise combinando os ciclos 1 e 2 os números foram 1 e 801, recodificando as plantas do IAC 112 como planta 801 e eliminando os demais valores fenotípicos. A inclusão da testemunha implicou na necessidade de inserir o efeito fixo de população ( $p_1$  e  $p_2$ ), sendo utilizados os valores 1 para a população Viçosa e 2 para o IAC 112. Então, as médias populacionais de Viçosa e do IAC 112 foram estimadas pelas seguintes funções previsíveis

$$\tilde{M} = M^0 + p_1^0 + (1/400) \sum_i \tilde{A}_i \text{ e } \tilde{M}_{\text{IAC112}} = M^0 + p_2^0 + \tilde{A}_{\text{IAC112}}$$

sendo  $M^0$ ,  $p_1^0$  e  $p_2^0$ , os elementos da solução para o vetor de efeitos fixos. A média dos selecionados é dada por

$$\tilde{M}_s = M^0 + p_1^0 + (1/196) \sum_i \tilde{A}_{i_s}$$

As análises dos dados dos lotes de recombinação com famílias de irmãos completos foram feitas com o SAS (modelo aditivo) e com o ASReml (modelo aditivo-dominante). Contudo, certamente pela falta dos valores fenotípicos dos machos utilizados nos cruzamentos, os quais também não foram identificados (para permitir o ajuste do modelo aditivo-dominante, foi admitido que os machos eram indivíduos no lote de recombinação não utilizados como fêmeas; códigos 717 a 912; ver Anexo 1), não houve convergência na estimação dos componentes de variância, quando se ajustou o modelo aditivo-dominante.

### **Resultados**

Os testes de nulidade da variância aditiva (estatística  $z$ ) revelaram, independente da estruturação da população em famílias de meios-irmãos ou irmãos completos, a existência de variabilidade genética em relação à CE e produção de grãos (Tabelas 1 e 2). Dois testes, contudo, estão associados a valores de probabilidade maiores que 5% (26.42%, em relação à CE no ciclo 3 da população estruturada em progênes de meios-irmãos, e 14.56%, também em relação à CE no ciclo 1 da população estruturada em progênes de irmãos completos). Levando em conta o valor de probabilidade do teste e a magnitude da variância aditiva, tem-se evidência de redução da variabilidade com seleção entre e dentro de famílias de meios-irmãos.

As estimativas de correlação genética entre CE e produção só foram consistentes nas duas populações em relação ao ciclo 2 (aproximadamente -0.1) (Tabelas 1 e 2). Em relação ao primeiro ciclo obteve-se estimativa negativa de pequena magnitude na população estruturada em progênes de meios-irmãos (-0.15) e correlação próxima de zero na população estruturada em progênes de irmãos completos. Os resultados obtidos revelam que a seleção com base em qualidade não deve implicar em alteração relevante na produção de grãos.

A análise de eficiência de identificação dos indivíduos de melhor valor genético aditivo revelou que os valores preditos por BLUP/REML, nas análises uni e multi-gerações,

são mais acurados que os valores fenotípicos (Tabela 3), principalmente na população estruturada em famílias de irmãos completos. A acurácia dos valores aditivos preditos por BLUP/REML superou a acurácia dos valores fenotípicos (dada pela raiz quadrada da herdabilidade) em 12% a 42% para CE e em 19% a 67% para produção, na população estruturada em progênies de meios-irmãos, e em 82% a 208% para CE e em 98% a 200% para produção, na população estruturada em progênies de irmãos completos.

Independente da estruturação da população em famílias de meios-irmãos ou irmãos completos, observou-se aumento da acurácia pelo método BLUP/REML multi-gerações, em relação à análise por geração, apenas em relação à CE (Tabela 3). Quanto à produção, pode-se considerar que não houve incremento da acurácia pela combinação de duas a três gerações.

Tanto para CE como para produção, nos ciclos 2 e 3 da população estruturada em famílias de meios-irmãos a correlação entre os valores aditivos preditos pelos métodos BLUP uni e multi-gerações foi elevada, com estimativas próximas a 0.9, ocorrendo uma ligeira queda do segundo para o terceiro ciclo. Na população estruturada em família de irmãos completos houve uma queda menos expressiva do primeiro para o segundo ciclo, indo de 0.97 para 0.96.

Selecionando em cada geração com base nos valores aditivos preditos a partir da análise BLUP/REML, em relação à CE, verificou-se maior eficiência de melhoramento na população estruturada em progênies de meios-irmãos com seleção massal, e maior eficiência na população estruturada em progênies de irmãos completos com seleção dentro (Tabelas 1 e 2). Na primeira o incremento no ganho predito em CE com seleção massal, em comparação à seleção dentro, variou entre 24% a 262%. Na segunda, a seleção dentro superou a seleção massal em 229% a 540%. Os ganhos indiretos preditos em produção foram, como esperados em razão da reduzida correlação, irrelevantes, variando de -1.5% a 3.4%. Os ganhos preditos em CE são de baixa magnitude, variando de 0.3% a 4.7%. Na população estruturada em

famílias de meios-irmãos os ganhos preditos diminuíram com o passar das gerações de seleção. Resultados equivalentes foram obtidos por Matta e Viana (2003), no ciclo 1 da população de milho-pipoca Beija-Flor estruturada em famílias de meios-irmãos, empregando BLUE/Método III de Henderson. Os ganhos preditos em CE, usando diferencial de seleção, foram de 0.53 mL/g e 0.35 mL/g, com seleção massal e seleção dentro, respectivamente.

Na população estruturada em famílias de meios-irmãos a coincidência envolvendo os 196 indivíduos selecionados por seleção massal pelas análises BLUP/REML e por seleção fenotípica decresceram ao longo dos ciclos. Nas análises individuais, 85,7%; 77,5% e 53,1% dos indivíduos selecionados no primeiro, segundo e terceiros ciclos, também foram indicados como superiores pela seleção fenotípica. Portanto, à medida que se adicionou a informação de parentesco nas análises individuais por BLUP/REML houve redução na coincidência entre os métodos. Entretanto, ao se realizar a seleção com base nas predições da análise conjunta, houve valores um pouco mais elevados de coincidência com a seleção fenotípica, que foram de 83,7% e 58,2% para o segundo e terceiro ciclos, respectivamente. Na população estruturada em família de irmãos completos a coincidência entre seleção dentro de famílias pelas análises individuais BLUP/REML e por seleção fenotípica decresceu ao longo dos ciclos, onde as coincidências foram de 85,7% e 82,6 %, para o primeiro e segundo ciclos, respectivamente. De maneira oposta, verificou-se um acréscimo da coincidência entre a seleção fenotípica e a seleção com base na análise conjunta BLUP/REML, que foi de 70,0% para 98,0% ao longo dos dois ciclos de seleção.

### **Discussão**

O aspecto mais importante que deveria ser discutido quando se emprega o método BLUP por meio do modelo animal, entre outros, como o modelo de reprodutor e o modelo de repetibilidade, é o ajuste do modelo aditivo, mesmo quando sabe-se que a expressão do caráter depende de efeitos de interação entre genes alélicos (dominância), como a produção de

grãos de plantas de milho, entre muitas outras importantes características de espécies animais e vegetais. Isso, contudo, nunca é levado em conta nos estudos aplicados e de simulação, provavelmente por que há um consenso entre os melhoristas de que a metodologia proporciona adequada predição dos verdadeiros valores genéticos aditivos, em geral mais acuradas que as obtidas por outros métodos (Kumar et al. 2008). Os estudos de simulação realizados por Leclerc et al. (2008) e Stock et al. (2007), envolvendo caracteres contínuos e categóricos, bem como dados moleculares na análise de QTL, demonstraram a grande eficiência e utilidade da metodologia BLUP, inclusive em análises multi-trait e de regressão aleatória.

Também há consenso de que a técnica também oferece adequada predição de valores genéticos aditivos quando há sobreposição de gerações, falta parcial ou total de informações de pedigree e ação da seleção, direcionada pelo homem ou natural. Kruuk (2004) considera que a metodologia BLUP/REML é poderosa para a estimação de componentes de variância e predição de valores genéticos aditivos em populações naturais, Contudo, em seu estudo de simulação, Cantet et al. (2000) verificaram subestimação da herdabilidade em população sob seleção, em razão de disponibilidade apenas parcial das informações de parentesco. Adicionalmente, verificaram também que a existência de seleção torna a herdabilidade viesada.

Não há dúvida, contudo, que o ajuste do modelo aditivo-dominante é importante, quando possível, nos estudos de capacidade de combinação, principalmente quando há interesse na predição da performance de cruzamentos não realizados (Bernardo 1996). No trabalho de Flachenecker et al. (2006a, b), em que foi ajustado o modelo aditivo, a predição do comportamento das famílias de irmãos completos pelo valor genético aditivo médio dos pais não foi satisfatória para produção, mas sim para uma característica de importância menor, a umidade dos grãos, provavelmente por ela não depender acentuadamente de efeitos de

dominância, e para uma segunda, denominada índice de seleção, a qual é função da produção e do conteúdo de matéria-seca.

Embora os resultados obtidos nesse estudo e em muitos outros teóricos e aplicados revelem sistematicamente superioridade do método BLUP/REML em relação ao método BLUE/Método III de Henderson, em razão da maior acurácia dos valores preditos em relação aos valores fenotípicos, nem sempre a discriminação entre distintos métodos de seleção é acentuada (Kumar et al. 2008) , principalmente quando os dados são balanceados e de apenas uma geração (Knott et al. 1995).

### **Agradecimentos**

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, pelos auxílios financeiros.

### **Referências**

- Arnason T, Jonmundsson, JV (2008) Multiple trait genetic evaluation of ewe traits in Icelandic sheep. *Journal of Animal Breeding and Genetics* 125:390-396.
- Bauer AM, Léon J (2008) Multiple-trait breeding values for parental selection in self-pollinating crops. *Theoretical and Applied Genetics* 116:235-242.
- Bernardo R (1996) Best linear unbiased prediction maize single-cross performance. *Crop Science* 36:50-56.
- Cantet RJC, Birchmeier AN, Santos-Cristal MG, de Avila VS (2000) Comparison of restricted maximum likelihood and Method R for estimating heritability and predicting breeding value under selection. *Journal of Animal Science* 78:2554-2560.
- Flachenecker C, Frisch M, Falke KC, Melchinger AE (2006a) Modified full-sib selection and best linear unbiased prediction of progeny performance in a European F2 maize population. *Plant Breeding* 125:248-253.

- Flachenecker C, Frisch M, Falke KC, Melchinger AE (2006b) Trends in population parameters and best linear unbiased prediction of progeny performance in a European F2 maize population under modified recurrent full-sib selection. *Theoretical and Applied Genetics* 112:483–491.
- Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead.
- Henderson CR (1974) General flexibility of linear model techniques for sire evaluation. *Journal of Dairy Science* 57:963-972.
- Knott SA, Sibly RM, Smith RH, Moller H (1995) Maximum likelihood estimation of genetic parameters in life-history studies using the ‘Animal Model’. *Functional Ecology* 9: 122-126.
- Kruuk LEB (2004) Estimating genetic parameters in natural populations using the “animal model”. *Philos Trans R Soc Lond B Biol Sci* 359: 873–890.
- Kumar A, Gandh RS, Singh A, Haile A (2008) Comparison of animal model with other conventional methods of sire evaluation for milk production in Karan Fries cattle. *Indian Journal of Animal Science* 78:1393-1396.
- Leclerc H, Wensch-Dorendorf M, Wensch J, Ducrocq V, Swalve HH (2008) A general method to validate breeding value prediction software. *Journal of Dairy Science* 91:3179-3183.
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) SAS for mixed models. 2<sup>nd</sup> ed. SAS Institute Inc., Cary.
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinuauer, Sunderland.
- Mrode, RA (2005) Linear models for the prediction of animal breeding values. 2.ed. CABI Publishing, Oxfordshire.

Nunes, JAR, Ramalho MAP, Ferreira DF (2008) Inclusion of genetic relationship information in the pedigree selection method using mixed models. *Genetics and Molecular Biology*, 31:73-78.

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545–554.

Piepho HP, Mohring J, Melchinger AE, Buchse A. (2008) BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161:209-228.

Stock KF, Hoeschele I, Distl O (2007) Estimation of genetic parameters and prediction of breeding values for multivariate threshold and continuous data in a simulated horse population using Gibbs sampling and residual maximum likelihood. *Journal of Animal Breeding and Genetics* 124:308-319.

Viana JMS (2005) Dominance, epistasis, heritabilities and expected genetic gains *Genetics and Molecular Biology* 28:67-74.

**Tabela 1** - Parâmetros estimados por BLUP/REML, médias da população e IAC 112, ganho direto em CE (mL/g) e ganho indireto em produção (g/0,18m<sup>2</sup>), relativo à população estruturada em família de meios irmãos.

| Parâmetros           | Ciclo 1                 |                          | Ciclo 2                 |                          | Ciclo 3                 |                          |
|----------------------|-------------------------|--------------------------|-------------------------|--------------------------|-------------------------|--------------------------|
|                      | CE                      | Produção                 | CE                      | Produção                 | CE                      | Produção                 |
| Var. Aditiva         | 19.12 <sup>0.0001</sup> | 646.19 <sup>0.0060</sup> | 12.29 <sup>0.0420</sup> | 519.60 <sup>0.0001</sup> | 41.05 <sup>0.2642</sup> | 357.50 <sup>0.0500</sup> |
| (intervalo conf.)    | (11.76 - 26.49)         | (253.87 - 1038.52)       | (0 - 26.27)             | (247.41 - 791.78)        | (0 - 168.71)            | (0 - 784.33)             |
| Var. Residual        | 8.17                    | 327.57                   | 14.87                   | 281.51                   | 350.48                  | 689.34                   |
| Corr. Genética       | -0.148                  |                          | -0.077                  |                          | -                       |                          |
| Média da Pop.        | 29.52                   | 89.67                    | 34.79                   | 84.30                    | 35.31                   | 91.03                    |
| Ganho c/ sel. massal | 1.40                    | 0.48                     | 0.98                    | 0.85                     | 0.58                    | 0.87                     |
| Ganho c/ sel. dentro | 1.13                    | -0.04                    | 0.74                    | 0.83                     | 0.16                    | -0.24                    |
| Média do IAC 112     | 38.00                   | 113.68                   | 37.90                   | 89.60                    | 41.90                   | 66.16                    |

**Tabela 2** - Parâmetros estimados por BLUP/REML, médias da população e IAC 112, ganho direto em CE (mL/g) e ganho indireto em produção (g/0,18m<sup>2</sup>), relativo à população estruturada em família de irmãos completos.

| Parâmetros                        | Ciclo 1                               |                                              | Ciclo 2                              |                                            |
|-----------------------------------|---------------------------------------|----------------------------------------------|--------------------------------------|--------------------------------------------|
|                                   | CE                                    | Produção                                     | CE                                   | Produção                                   |
| Var. Aditiva<br>(intervalo conf.) | 3.75 <sup>0.1456</sup><br>(0 - 13.63) | 128.30 <sup>0.0401</sup><br>(14.48 - 282.46) | 2.54 <sup>0.0553</sup><br>(0 - 5.68) | 61.45 <sup>0.0150</sup><br>(5.93 - 116.97) |
| Var. Residual                     | 57.51                                 | 740.96                                       | 22.46                                | 200.63                                     |
| Corr. Genética                    | 0.0387                                |                                              | -0.1410                              |                                            |
| Média da Pop.                     | 30.49                                 | 42.48                                        | 31.85                                | 36.41                                      |
| Ganho c/ sel. massal              | 0.1015                                | -0.1997                                      | 0.4459                               | -0.2409                                    |
| Ganho c/ sel. dentro              | 0.6421                                | 1.4610                                       | 1.4793                               | -0.5426                                    |

**Tabela 3** - Acurácias do BLUP/REML e BLUE/ Método III de Henderson.

| Análises                      | Família de meios irmãos |         |         |          |         |         | Família de irmãos completos |         |          |         |
|-------------------------------|-------------------------|---------|---------|----------|---------|---------|-----------------------------|---------|----------|---------|
|                               | CE                      |         |         | Produção |         |         | CE                          |         | Produção |         |
|                               | Ciclo 1                 | Ciclo 2 | Ciclo 3 | Ciclo 1  | Ciclo 2 | Ciclo 3 | Ciclo 1                     | Ciclo 2 | Ciclo 1  | Ciclo 2 |
| BLUP Uni-ger.                 | 0.93                    | 0.87    | -       | 0.98     | 0.98    | 0.97    | 0.77                        | 0.69    | 0.96     | 0.95    |
| BLUP Multi-ger. C1+C2         |                         | 0.95    | -       |          | 0.95    | -       |                             | 0.83    |          | 0.95    |
| BLUP Multi-ger. C1+C2+C3      |                         |         | 0.98    |          |         | 0.96    |                             |         |          |         |
| BLUE/ Método III de Henderson | 0.83                    | 0.67    | -       | 0.81     | 0.80    | 0.58    | 0.25                        | 0.38    | 0.32     | 0.48    |

## **ANEXO**

**Anexo 1** - Pedigrees utilizados nas análises.

| Seleção de meios irmãos     |       |       |
|-----------------------------|-------|-------|
| Indivíduo                   | macho | fêmea |
| 1                           | .     | .     |
| ...                         |       |       |
| 196                         | .     | .     |
| 197                         | .     | 180   |
| ...                         |       |       |
| 596                         | .     | 192   |
| 597                         | .     | 524   |
| ...                         |       |       |
| 996                         | .     | 455   |
| 997                         | .     | 891   |
| ...                         |       |       |
| 1396                        | .     | 955   |
| Seleção de irmãos completos |       |       |
| 1                           | .     | .     |
| 2                           | .     | .     |
| ...                         |       |       |
| 196                         | .     | .     |
| 197                         | .     | .     |
| ...                         |       |       |
| 392                         | .     | .     |
| 393                         | 249   | 53    |
| ...                         |       |       |
| 716                         | 208   | 12    |
| 717                         | .     | .     |
| ...                         |       |       |
| 911                         | 307   | 111   |
| 912                         | .     | .     |
| 913                         | 798   | 422   |
| ...                         |       |       |
| 1302                        | 725   | 701   |