

# Parentesco na seleção para produtividade e teores de óleo e proteína em soja via modelos mistos

Larissa Correia de Melo Pinheiro<sup>(1)</sup>, Pedro Ivo Vieira Good God<sup>(1)</sup>, Vinícius Ribeiro Faria<sup>(1)</sup>, Ane Gabrielle Oliveira<sup>(1)</sup>, Aline Akemi Hasui<sup>(1)</sup>, Eduardo Henrique Guimarães Pinto<sup>(1)</sup>, Klever Márcio Antunes Arruda<sup>(2)</sup>, Newton Deniz Piovesan<sup>(3)</sup> e Maurilio Alves Moreira<sup>(3)</sup>

<sup>(1)</sup>Universidade Federal de Viçosa (UFV), Campus Rio Paranaíba, Caixa Postal 22, CEP 38810-000 Rio Paranaíba, MG. E-mail: larissa.correia@ufv.br, pedro.god@ufv.br, vinicius.faria@ufv.br, ane.oliveira@ufv.br, akemihassui@gmail.com, eduardo.pinto@ufv.br

<sup>(2)</sup>Instituto Agronômico do Paraná, Rodovia Celso Garcia Cid, Km 375, CEP 86047-902 Londrina, PR. E-mail: klever@iapar.br <sup>(3)</sup>UFV, Bioagro, Avenida P.H. Rolfs, s/nº, CEP 36570-000 Viçosa, MG. E-mail: piovesan@ufv.br, moreira@ufv.br

Resumo – O objetivo deste trabalho foi avaliar influência da informação de parentesco na seleção de progênies de soja quanto à produtividade e aos teores de óleo e proteína, com base no uso de modelos mistos de predição dos valores genéticos. Novecentas progênies  $F_{4,6}$  e 200 progênies  $F_{4,7}$  de soja foram avaliadas nas safras 2010/2011 e 2011/2012, respectivamente. As progênies foram obtidas de cruzamentos múltiplos a partir de 57 progenitores. Os dados foram analisados por meio de modelos aleatórios (quadrados mínimos) e mistos BLUP/REML (“best linear unbiased prediction/restricted maximum likelihood”). Os maiores valores de ganhos preditos foram obtidos com o BLUP/REML. Os valores genéticos preditos com o método BLUP/REML, sem informação de parentesco, apresentaram alta correlação com aqueles obtidos com o modelo aleatório, além de detectada alta coincidência das progênies selecionadas. A inclusão da matriz de parentesco resultou na seleção de progênies diferentes e em maior acurácia na predição dos valores genéticos.

Termos para indexação: *Glycine max*, BLUP/REML, ganhos de seleção, matriz de parentesco.

## Relationship in the selection for productivity and oil and protein contents in soybean using mixed models

Abstract – The objective of this work was to evaluate the influence of relationship information for selecting soybean progenies as to their productivity, and oil and protein contents, using mixed models for the prediction of breeding values. Nine hundred  $F_{4,6}$  and 200  $F_{4,7}$  soybean progenies were evaluated in the seasons 2010/2011 and 2011/2012, respectively. The progenies were obtained from multiple crosses from 57 parents. Data were analyzed using random models (least squares) and mixed models BLUP/REML (best linear unbiased prediction/restricted maximum likelihood). The highest values of predicted gains were obtained by BLUP/REML. The breeding values predicted with the use of BLUP/REML without relationship information were highly correlated with the ones obtained with the random model, and the selected progenies were rather coincident. The inclusion of the relationship matrix resulted in the selection of different progenies and in higher accuracy of breeding values.

Index terms: *Glycine max*, BLUP/REML, selection gain, relationship matrix.

## Introdução

A soja [*Glycine max* (L.) Merrill] é originária do continente asiático e destaca-se como uma das mais importantes oleaginosas cultivadas no Brasil e no mundo. Em razão de seu alto teor de óleo (19%) e proteína (38%), alta produtividade e versatilidade de uso, ela constitui uma das principais commodities no mercado mundial (Pires & Santos, 2013). No Brasil, a soja tem expressiva participação na pauta de exportações, na forma de farelo, óleo e grãos. Em termos mundiais, a soja contribui com 29,4% do

mercado de óleos vegetais, cuja produção é destinada principalmente ao consumo humano e como matéria-prima para a produção de biodiesel (Silva, 2013).

Nos programas de melhoramento de soja, a seleção de genótipos é feita com base nos valores fenotípicos, aos quais confundem-se os valores genotípicos e ambientais. O uso das equações de modelos mistos (EMM) (Henderson, 1975) pode prever o valor genotípico e aumentar a eficiência da seleção. Com base nas EMM, é possível obter o melhor preditor linear não-viesado (BLUP) dos valores genéticos. Para a

predição por meio de modelos mistos, é necessário obter as estimativas dos componentes de variância associados aos efeitos aleatórios do modelo. Essas estimativas são obtidas por meio do método da máxima verossimilhança restrita (REML) (Patterson & Thompson, 1971), simultaneamente à predição dos valores genéticos, o que consiste no método BLUP/REML.

O uso do método BLUP/REML é comum no melhoramento genético de plantas perenes como eucalipto (Rosado et al., 2012), café (Ramalho et al., 2012) e frutíferas (Bruna et al., 2012). Entretanto, no melhoramento de espécies anuais, seu uso não é tão frequente, embora haja um número crescente de relatos, em diferentes espécies, como milho (Viana et al., 2010a, 2011, 2012), feijão (Baldissera et al., 2012), arroz (Borges et al., 2012) e trigo (Silva et al., 2011).

Há importantes relatos acerca da utilização de BLUP/REML no melhoramento da soja. Panter & Allen (1995a) estudaram a eficiência dos quadrados mínimos e do BLUP na identificação de linhagens superiores. Os autores simularam o balanceamento ou o desbalanceamento de dados, e observaram que as predições BLUP apresentaram menor erro-padrão e maior correlação dos valores preditos com o desempenho dos genótipos. Naoe et al. (2006), em um programa de retrocruzamentos, verificaram a influência da informação de parentesco sobre a estimação de parâmetros genéticos quanto ao teor de proteína em sementes de soja. A inclusão da análise de parentesco alterou as estimativas dos componentes de variância e da herdabilidade; no entanto, os resultados foram inconclusivos no que diz respeito ao aumento na acurácia da estimação. Carvalho et al. (2008) relataram que, quando não há incorporação da informação de parentesco, os procedimentos dos modelos mistos (BLUP/REML) para a predição e estimação de parâmetros e valores genéticos da soja não diferem dos utilizados com o modelo aleatório (quadrados mínimos).

O objetivo deste trabalho foi avaliar influência da informação de parentesco na seleção de progênies de soja quanto à produtividade e aos teores de óleo e proteína, com base no uso de modelos mistos de predição dos valores genéticos.

## Material e Métodos

As populações de soja deste trabalho foram obtidas a partir de 57 progenitores, representados por cultivares-elite e por linhagens do Programa de

Melhoramento da Qualidade da Soja, desenvolvido pela Universidade Federal de Viçosa. Entre os progenitores, foram incluídos genótipos com alto teor de óleo (>21%) e alto teor proteico (>40%).

Após a realização de todos os intercruzamentos, foram obtidos 450 tipos de híbridos óctuplos. Destes, 150 foram selecionados quanto à tolerância ao herbicida glifosato. Os híbridos foram autofecundados e as progênies F<sub>2</sub> foram conduzidas pelo método SPD (“single pod descent”), até a geração F<sub>4</sub>.

A partir da geração F<sub>4</sub>, todos os ensaios foram conduzidos na área experimental da Cooperativa Agropecuária do Alto Paranaíba (Coopadap), no Município de Rio Paranaíba, MG. Os teores de óleo e proteína foram determinados no Laboratório de Genética Molecular de Plantas/Bioagro/UFV, por meio de espectrometria do infravermelho próximo (FT-NIR). Os espectros foram obtidos a partir de amostras de grãos moídos.

A geração F<sub>4</sub> foi conduzida na safra 2008/2009. Foram selecionadas 1.300 plantas individuais para a composição da próxima geração, com base no desempenho agrônomo. A geração F<sub>4,5</sub> foi conduzida na safra 2009/2010, quando foram colhidas sementes F<sub>6</sub> de 1.243 progênies. Uma pequena amostra de cada parcela foi triturada e avaliada quanto aos teores de óleo e proteína. Foram selecionadas 487 (39%) famílias com maior teor de proteína e 479 (38%) famílias com maior teor de óleo. Como houve redundância em 77 famílias quanto ao alto teor de óleo e proteína, foram selecionadas ao todo 889 famílias F<sub>4,6</sub>.

Para avaliar a produtividade de grãos e os teores de óleo e proteína, as famílias previamente selecionadas, além de outras seis progênies com características agrônomicas desejáveis e 5 testemunhas (A 7002, Suprema, UFVTN 105AP, VX-044622 e NA 7255RR), foram semeadas na safra 2010/2011. Utilizou-se o delineamento em látice quadrado 30x30, com duas repetições, em parcelas de 2 m, com espaçamento de 0,5 m e densidade de semeadura de 13 a 15 sementes por metro.

A partir do ensaio anterior, 200 progênies F<sub>4,7</sub> foram selecionadas para compor o ensaio na safra agrícola 2011/2012. Como testemunhas, foram empregados genótipos de alto teor de óleo (A 7002, Msoy 6101 e Suprema), alto teor de proteína (VX-044622 e UFVTN 105AP) e alto rendimento de grãos (Conquista, CS 030 e CXA 2). Neste ensaio, utilizou-se o delineamento em blocos ao acaso com

duas repetições. Cada parcela foi composta de quatro fileiras de 5x0,50 m. A parcela útil para avaliação do rendimento de grãos e dos teores de óleo e proteína foi constituída pelas duas fileiras centrais. A densidade de semeadura utilizada foi de 13 a 15 sementes por metro.

O procedimento GLM do programa SAS versão 9.2 (SAS Institute, Cary, NC, EUA) foi utilizado para a análise em modelos aleatórios, com base no método dos quadrados mínimos. O procedimento Mixed foi utilizado para análises em modelos mistos. As rotinas utilizadas nas análises de modelos mistos estão apresentadas em Viana et al. (2011). Os dados da safra 2010/2011 foram analisados com base no modelo em látice quadrado. O modelo estatístico adequado foi obtido com a equação  $y_{ijk} = \mu + t_i + r_j + b_{k(j)} + \varepsilon_{ijk}$ , em que:  $y_{ijk}$  é o valor fenotípico;  $\mu$  é a constante do modelo;  $t_i$  é o efeito da progênie  $i$ ;  $r_j$  é o efeito da repetição  $j$ ;  $b_{k(j)}$  é o efeito do bloco  $k$ , dentro da repetição  $j$ ; e  $\varepsilon_{ijk}$  é o erro experimental. Os dados da safra 2011/2012 foram analisados conforme o delineamento de blocos ao acaso, por meio da equação  $y_{ij} = \mu + t_i + b_j + \varepsilon_{ij}$ , em que:  $y_{ij}$  é o valor fenotípico;  $\mu$  é a constante do modelo;  $t_i$  é o efeito da progênie  $i$ ;  $b_j$  é o efeito do bloco; e  $\varepsilon_{ij}$  é o erro experimental.

O procedimento BLUP/REML foi utilizado de duas formas. Na primeira, não foram incluídas as informações de parentesco das progênies em seleção. Na segunda, foi incluída a informação de parentesco, por meio da matriz de parentesco genético, calculada com base na genealogia das progênies.

O modelo misto aditivo univariado foi utilizado e é dado matricialmente por  $y = X\beta + Z_1a + Z_2b + \varepsilon$ , em que:  $y$  é o vetor de observações (valores fenotípicos);  $X$  é a matriz de incidência dos efeitos fixos;  $\beta$  é o vetor de efeitos fixos;  $Z_1$  e  $Z_2$  são as matrizes de incidência dos efeitos aleatórios;  $a$  é o vetor de valores genéticos aditivos dos pais;  $b$  é vetor de efeitos de blocos; e  $\varepsilon$  é o vetor de erros.

A matriz de covariâncias do vetor  $a$  é dada por

$$\text{Var}(a) = G_1 = \sigma_A^2 \begin{bmatrix} 2F & 2r_{12} & \cdots & 2r_{1t} \\ 2r_{12} & 2F & \cdots & 2r_{2t} \\ \cdots & \cdots & \ddots & \cdots \\ 2r_{1t} & 2r_{2t} & \cdots & 2F \end{bmatrix} = \sigma_A^2 A$$

em que  $r_{jj}$  é o coeficiente de parentesco entre os pais das progênies  $j$  e  $j'$ , e  $F$  é o coeficiente de endogamia. A matriz de parentesco  $A$  foi obtida por meio do

procedimento Proc Inbred, conforme Viana et al. (2011). No cenário sem informação de parentesco, considerou-se a matriz de parentesco como uma matriz identidade, o que faz com que a matriz de covariâncias seja dada por  $I\sigma_A^2$ .

A matriz de covariâncias do vetor  $b$  é dada por  $\text{Var}(b) = G_2 = I\sigma_b^2$ .

A predição dos efeitos aleatórios e estimação dos efeitos fixos são feitas por meio das equações de modelo misto, descritas por Henderson (1984) como:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_1 & X'R^{-1}Z_2 \\ Z_1'R^{-1}X & Z_1'R^{-1}Z_1 + G_1^{-1} & Z_1'R^{-1}Z_2 \\ Z_2'R^{-1}X & Z_2'R^{-1}Z_1 & Z_2'R^{-1}Z_2 + G_2^{-1} \end{bmatrix} \begin{bmatrix} \beta^0 \\ \tilde{a} \\ \tilde{b} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z_1'R^{-1}y \\ Z_2'R^{-1}y \end{bmatrix}$$

Para a obtenção das soluções do sistema de EMM, os componentes de variância genéticos e não genéticos foram considerados desconhecidos, como ocorre na prática, tendo sido estimados pela máxima verossimilhança restrita (REML) (Patterson & Thompson, 1971).

A comparação dos métodos e cenários analíticos foi feita pela determinação da acurácia dos valores genéticos preditos e pela correlação entre os valores genéticos preditos pela análise de modelos mistos e as médias fenotípicas, utilizadas para seleção no modelo aleatório. Verificou-se, também, a coincidência entre as progênies selecionadas pelos diferentes métodos.

## Resultados e Discussão

Os teores de óleo e proteína, bem como a produtividade de grãos, apresentaram variabilidade genética significativa nos experimentos conduzidos nas safras 2010/2011 e 2011/2012 (Tabela 1). Os coeficientes de variação (CV) relativos aos teores de óleo estiveram entre 3,5 e 3,6% e os de proteína, entre 2,2 e 3,0%. Rodrigues et al. (2010) mapearam locos QTL em soja e obtiveram valores de CV de 4,06 e 3,8%, para os teores de proteína e óleo, respectivamente. Valores fenotípicos do teor de proteína e óleo são estimados, em geral, com alta precisão experimental.

Quanto à produtividade, o coeficiente de variação ficou entre 20 e 30,5% (Tabela 1). Storck et al. (2010) verificaram que 76,4% de 216 ensaios de competição de cultivares de soja tiveram alta precisão, avaliada por meio da acurácia seletiva. Apenas 6,9% desses ensaios tiveram baixa precisão, e o coeficiente de variação médio nos ensaios foi de 13,3%, com valor máximo de

33,9%. No presente trabalho, deve-se considerar que o esperado é que haja maior variabilidade intraclasse, por se tratar de ensaios com testes de progênies. O alto nível descritivo do teste ( $p < 0,0001$ ) para a variabilidade genética é uma indicação de que o estudo da característica é viável.

Os valores de herdabilidade para produtividade variaram de 22 a 36%, os de proteína, de 38 a 87%, e os de óleo, de 41 a 84% (Tabelas 1 e 2). Miranda (2006), ao avaliar 207 famílias nas gerações  $F_2$  e  $F_3$ , observou herdabilidades de 62,36%, para o teor de proteína; 71,5%, para o de óleo; e 56,5%, para produção de grãos. Chung et al. (2003) relataram herdabilidades entre 65 e 98% para os três caracteres, em populações de linhagens endogâmicas recombinantes.

A inclusão da informação de parentesco diminuiu as estimativas das variâncias aditivas para todas as características e, portanto, resultou em menores valores de herdabilidade. Embora as magnitudes de herdabilidade tenham variado, os ganhos preditos pelas duas análises foram semelhantes, uma vez que o cálculo de ganhos é realizado com base no valor genético predito das progênies selecionadas e não com base em funções que levam em consideração a estimativa de herdabilidade.

As correlações genóticas, obtidas para teor de óleo x proteína, teor de óleo x produtividade e teor de proteína x produtividade, foram de -0,57, 0,49 e 0, respectivamente. Portanto, a seleção simultânea de genótipos com alta concentração de óleo e proteína deve apresentar baixa eficiência, e ganhos para concentração de óleo podem ser obtidos de forma indireta, pelo aumento da produtividade (Clemente & Cahoon, 2009). A dificuldade para obtenção de ganhos simultâneos quanto aos teores de óleo e proteína é bem conhecida. No entanto, em estudo recente, verificou-se a existência de locos QTL e de interações epistáticas relacionadas ao conteúdo de óleo que apresentam efeitos positivos sobre a concentração de proteína e produtividade (Eskandari et al., 2013).

Os maiores ganhos preditos foram obtidos quando se utilizou o método BLUP/REML (Tabelas 1 e 2). No modelo aleatório, o ganho de seleção relativo ao teor de proteína variou de 2,9 a 4,6%. Ao empregar o método BLUP/REML, os ganhos obtidos foram de 5,2 a 9,8%, a depender da inclusão do parentesco na análise. Os ganhos quanto ao teor de óleo, com base no modelo aleatório, foram de 3,7 e 5,4% e, na análise BLUP/REML, esses valores atingiram 11,9%. Quanto à produtividade de grãos, as diferenças entre os métodos

**Tabela 1.** Parâmetros genéticos estimados com o método de quadrados mínimos, para seleção de progênies de soja com altos teores de proteína, óleo e produtividade.

| Parâmetro              | Proteína      | Óleo                  |  | Produtividade             |
|------------------------|---------------|-----------------------|--|---------------------------|
|                        |               | Experimento 2010/2011 |  |                           |
| Variância aditiva      | 2,3990*       | 1,2330*               |  | 69.833,83*                |
| Intervalo de confiança | 2,1895–2,6402 | 1,1253–1,3569         |  | 63.722,46–76.873,26       |
| Variância residual     | 1,1090        | 0,5379                |  | 245.799                   |
| Acurácia               | 0,8270        | 0,8344                |  | 0,4704                    |
| Herdabilidade          | 0,6839        | 0,6963                |  | 0,2213                    |
| Ganho direto           | 1,1381        | 0,7589                |  | 96,1678                   |
| Ganho de seleção (%)   | 2,90          | 3,70                  |  | 5,40                      |
| Média da população     | 38,09%        | 20,52%                |  | 1.779 kg ha <sup>-1</sup> |
| CV (%)                 | 2,80          | 3,60                  |  | 27,80                     |
|                        |               | Experimento 2011/2012 |  |                           |
| Variância aditiva      | 5,4682*       | 2,7555*               |  | 224940,9*                 |
| Intervalo de confiança | 4,5506–6,6959 | 2,2931–3,3741         |  | 187.195,57–275.442,14     |
| Variância residual     | 0,8048        | 0,5276                |  | 392.094,3                 |
| Acurácia               | 0,9336        | 0,9161                |  | 0,6038                    |
| Herdabilidade          | 0,8717        | 0,8393                |  | 0,3645                    |
| Ganho direto           | 1,8501        | 1,1123                |  | 201,3962                  |
| Ganho de seleção (%)   | 4,6           | 5,4                   |  | 6,4                       |
| Média da população     | 40,28%        | 20,44%                |  | 3.119 kg ha <sup>-1</sup> |
| CV (%)                 | 2,2           | 3,5                   |  | 20,0                      |

\* Significativo a 1% de probabilidade, pelo teste F.

analíticos foram maiores, com 6,4%, no modelo aleatório, e 18,6% na análise BLUP/REML. Carvalho et al. (2008) avaliaram a resposta da produtividade de grãos a diferentes intensidades de seleção, por meio do método dos quadrados mínimos e pela predição BLUP/REML. Os autores observaram que as diferenças entre os ganhos preditos pelos dois métodos analíticos foram menores em maiores intensidades de seleção, e que os maiores ganhos foram obtidos com o uso do BLUP/REML, embora diferenças práticas não tenham sido relatadas.

A inclusão da informação de parentesco aumentou a acurácia dos valores genéticos preditos, quanto à produtividade (Tabela 2). Pequenas alterações nos valores de acurácia foram verificadas quanto aos teores de óleo e proteína. Os modelos aleatório e BLUP/REML coincidiram quanto às progênes selecionadas e apresentaram alta correlação entre os valores genéticos preditos, quando não foi considerada informação de parentesco (Tabela 3). Carvalho et al. (2008) obtiveram resultados semelhantes, o que apoia a constatação de equivalência entre os métodos no

**Tabela 2.** Parâmetros genéticos estimados com o método BLUP/REML, com e sem a informação de parentesco, para seleção de progênes de soja com altos teores de proteína, óleo e produtividade.

| Parâmetro              | Proteína       |                | Óleo           |                | Produtividade                |                |
|------------------------|----------------|----------------|----------------|----------------|------------------------------|----------------|
|                        | Sem parentesco | Com parentesco | Sem Parentesco | Com parentesco | Sem parentesco               | Com parentesco |
| Experimento 2010/2011  |                |                |                |                |                              |                |
| Varição aditiva        | 2,4930*        | 1,2892*        | 1,2997*        | 0,6748*        | 72.946*                      | 29.261*        |
| Intervalo de confiança | 2,2096–2,8349  | 0,8–1,7784     | 1,1544–1,4744  | 0,4130–0,9365  | 53.397–105.663               | 17.331–41.190  |
| Varição residual       | 1,1225         | 2,0479         | 0,5413         | 0,9679         | 244.535                      | 273.178        |
| Acurácia               | 0,9374         | 0,9605         | 0,9408         | 0,9626         | 0,8071                       | 0,9191         |
| Herdabilidade          | 0,6895         | 0,3863         | 0,7060         | 0,4108         | 0,2298                       | 0,0968         |
| Ganho direto           | 2,5629         | 2,0007         | 1,7074         | 1,6001         | 288,4174                     | 311,2129       |
| Ganho de seleção (%)   | 6,70           | 5,20           | 8,30           | 7,79           | 16,20                        | 17,50          |
| Média da população     | 38,08%         |                | 20,51%         |                | 1.779,42 kg ha <sup>-1</sup> |                |
| Experimento 2011/2012  |                |                |                |                |                              |                |
| Varição aditiva        | 5,4975*        | 5,3805*        | 2,7658*        | 1,8113*        | 223.898*                     | 65.728*        |
| Intervalo de confiança | 4,5132–6,8448  | 3,5001–7,2610  | 2,2611–3,4616  | 1,0857–2,5368  | 155.784–349.181              | 23.706–107.751 |
| Varição residual       | 0,8011         | 0,9805         | 0,5276         | 0,7146         | 390.829                      | 494.396        |
| Acurácia               | 0,9804         | 0,9633         | 0,9754         | 0,9653         | 0,8679                       | 0,9226         |
| Herdabilidade          | 0,8728         | 0,8458         | 0,8398         | 0,7171         | 0,3642                       | 0,1173         |
| Ganho direto           | 3,9583         | 3,8352         | 2,4006         | 2,4464         | 581,5165                     | 442,65         |
| Ganho de seleção (%)   | 9,80           | 9,50           | 11,70          | 11,90          | 18,60                        | 14,20          |
| Média da população     | 40,27%         |                | 20,44%         |                | 3.119,12 kg ha <sup>-1</sup> |                |

\*Significativo a 1% de probabilidade, pelo teste da razão de verossimilhança.

**Tabela 3.** Correlação (diagonal superior) e coincidência (diagonal inferior) na seleção de progênes de soja com o modelo aleatório ou com o método BLUP/REML (com e sem parentesco), para as características teores de óleo, proteína e produtividade.

| Modelo                   | Proteína         |                |                | Óleo             |                |                | Produtividade    |                |                |
|--------------------------|------------------|----------------|----------------|------------------|----------------|----------------|------------------|----------------|----------------|
|                          | Modelo aleatório | Sem parentesco | Com parentesco | Modelo aleatório | Sem parentesco | Com parentesco | Modelo aleatório | Sem parentesco | Com parentesco |
| Experimento 2010/2011    |                  |                |                |                  |                |                |                  |                |                |
| Modelo aleatório         | -                | 0,9773         | 0,7728         | -                | 0,9876         | 0,8007         | -                | 0,9621         | 0,6103         |
| BLUP/REML sem parentesco | 0,86             | -              | 0,7633         | 0,84             | -              | 0,7884         | 0,77             | -              | 0,5584         |
| BLUP/REML com parentesco | 0,52             | 0,52           | -              | 0,60             | 0,57           | -              | 0,41             | 0,42           | -              |
| Experimento 2011/2012    |                  |                |                |                  |                |                |                  |                |                |
| Modelo aleatório         | -                | 0,999          | 0,9842         | -                | 0,9993         | 0,9650         | -                | 0,9932         | 0,7351         |
| BLUP/REML sem parentesco | 1                | -              | 0,9845         | 1                | -              | 0,9667         | 1                | -              | 0,7402         |
| BLUP/REML com parentesco | 0,85             | 0,85           | -              | 0,87             | 0,87           | -              | 0,50             | 0,50           | -              |



processo de seleção. Esta equivalência é esperada, uma vez que, em condições balanceadas, tanto o método dos mínimos quadrados quanto o BLUP/REML produzem estimativas idênticas (Anderson, 1979).

No entanto, baixos valores de correlação entre os valores genéticos preditos pelas análises BLUP/REML com e sem informação de parentesco foram detectados (Tabela 3). A baixa correlação ocorreu em todas as características avaliadas no experimento de 2010/2011, e na produtividade em 2011/2012. A coincidência entre as progênies selecionadas também foi baixa nessas situações. A redução da coincidência de progênies selecionadas é esperada quando se utiliza informação de parentesco, principalmente para a seleção fenotípica (Vianna et al., 2010a). Portanto, a informação de parentesco influencia a seleção de progênies para compor as próximas gerações.

Nas situações em que a coincidência foi menor, foram obtidos maiores valores de acurácia para os valores genéticos preditos quando a informação de parentesco foi utilizada. Isto fica mais evidente quando se analisa a produtividade, cujos valores genéticos apresentaram correlação de 0,74 na safra 2011/2012, com apenas 50% de coincidência nas progênies selecionadas. Valores semelhantes foram verificados no experimento 2010/2011. Em estudo com dados simulados, verificou-se que a inclusão da informação de parentesco permite maior acurácia na obtenção de valores genéticos aditivos (Ramalho et al., 2012). Portanto, a inclusão de parentesco pode aumentar a eficiência de seleção, principalmente nos casos de menor precisão experimental.

A aplicação do BLUP no melhoramento de várias espécies anuais foi descrita por Piepho et al. (2008). De acordo com os autores, a inclusão da matriz de parentesco pode reduzir o viés da estimação de valores genéticos em populações sob seleção. Outros estudos mostraram que a inclusão de parentesco, na análise de modelos mistos, aumenta a acurácia dos valores e dos parâmetros genéticos (Nunes et al., 2008a, 2008b). Viana et al. (2010a) descreveram a eficiência da utilização do método BLUP juntamente com a utilização das informações de parentesco, em culturas anuais. Os autores concluíram que a matriz de parentesco determina maior eficiência de seleção, com base nos valores genéticos preditos pela análise BLUP. Em um estudo de predição de cruzamentos em soja, por meio do método BLUP, Panter & Allen

(1995b) observaram que o modelo com informação de parentesco foi ligeiramente superior.

Portanto, quando se consideram condições experimentais limitantes, ou na seleção de multicaudalísticas, o emprego de modelos mistos com informação de parentesco pode ser o mais recomendado (Viana et al., 2010b, 2012). Em termos práticos, a manutenção do registro de genealogias para progênies em teste pode demandar maior esforço, o que influencia a opção do uso destas informações na seleção. No entanto, de acordo com Heslot et al. (2012) e Segura et al. (2012), a utilização de modelos mistos no desenvolvimento de novas cultivares deve ser cada vez mais frequente. Este fato se deve principalmente ao avanço no uso de informações de marcadores moleculares e na implantação de sistemas de seleção assistida ou na seleção genômica, o que deve potencializar o uso de registros de genealogia.

## Conclusões

1. O procedimento BLUP/REML é ferramenta eficaz na seleção de genótipos de soja.

2. A inclusão da matriz de parentesco permite maior acurácia na estimação dos valores genéticos das populações e aumenta a eficiência de seleção, principalmente quando se tem menor precisão experimental.

## Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e à Fundação de Amparo a Pesquisa do Estado de Minas Gerais (Fapemig), por auxílio financeiro e concessão de bolsas; à Cooperativa Agropecuária do Alto Paranaíba (Coopadap), pelo apoio no desenvolvimento dos ensaios em campo.

## Referências

- ANDERSON, R.D. Estimating variance components from balanced data: optimum properties of REML solutions and MIVQUE estimators. In: VLECK, L.D. van; SEARLE, S.R. (Ed.). **Variance components and animal breeding**. Ithaca: Cornell University, 1979. p.205-215.
- BALDISSERA, J.N. da C.; BERTOLDO, J.G.; VALENTINI, G.; COAN, M.M.D.; ROZETO, D.S.; GUIDOLIN, A.F.; COIMBRA, J.L.M. Uso do melhor preditor linear não viesado (BLUP) na

- predição de híbridos em feijão. **Bioscience Journal**, v.28, p.395-403, 2012.
- BORGES, V.; SOARES, A.A.; RESENDE, M.D.V. de; REIS, M. de S.; CORNELIO, V.M. de O.; LEITE, N.A.; SOARES, P.C.; JÚNIOR, G.T.C. Value for cultivation and use of upland rice cultivars tested in multienvironments. **Crop Breeding and Applied Biotechnology**, v.12, p.25-33, 2012. DOI: 10.1590/S198470332012000100004.
- BRUNA, E.D.; MORETO, A.L.; DALBÓ, M.A. Uso do coeficiente de repetibilidade na seleção de clones de pessegueiro para o litoral sul de Santa Catarina. **Revista Brasileira de Fruticultura**, v.34, p.206-215, 2012. DOI: 10.1590/S010029452012000100028.
- CARVALHO, A.D.F. de; FRITSCH NETO, R.F.; GERALDI, I.O. Estimation and prediction of parameters and breeding values in soybean using REML/BLUP and least squares. **Crop Breeding and Applied Biotechnology**, v.8, p.219-224, 2008.
- CHUNG, J.; BABKA, H.L.; GRAEF, G.L.; STASWICK, P.E.; LEE, D.J.; CREGAN, P.B.; SHOEMAKER, R.C.; SPEECH, J.E. The seed protein, oil, and yield QTL on soybean linkage group I. **Crop science**, v.43, p.1053-1067, 2003. DOI: 10.2135/cropsci2003.1053.
- CLEMENTE, T.E.; CAHOON, E.B. Soybean oil: genetic approaches for modification of functionality and total content. **Plant Physiology**, v.151, p.1030-1040, 2009. DOI: 10.1104/pp.109.146282.
- ESKANDARI, M.; COBER, E.R.; RAJCAN, I. Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. **Theoretical and Applied Genetics**, v.126, p.1677-1687, 2013.
- HENDERSON, C.R. **Applications of linear models in animal breeding**. Guelph: University of Guelph, 1984. 462p.
- HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, v.31, p.423-447, 1975. DOI: 10.2307/2529430.
- HESLOT, N.; YANG, H.P.; SORRELLS, M.E.; JANNINCK, J.L. Genomic selection in plant breeding: a comparison of models. **Crop Science**, v.52, p.146-160, 2012. DOI: 10.2135/cropsci2011.06.0297.
- MIRANDA, F.D. de. **Produção, conteúdo de proteína e óleo no grão da soja: herdabilidades, correlações e seleção de genótipos superiores**. 2006. 76p. Tese (Doutorado) Universidade Federal de Viçosa, Viçosa.
- NAOE, L.K.; SEDIYAMA, C.S.; PIOVESAN, N.D.; LOPES, P.S.; COIMBRA, R.R.; BARROS, E.G. de; MOREIRA, M.A. Estimação de componentes de variância e parâmetros genéticos, considerando o parentesco em população de retrocruzamento de soja. **Revista Ciência Agroambiental**, v.1, p.21-26, 2006.
- NUNES, J.A.R.; MORETO, A.L.; RAMALHO, M.A.P. Using genealogy to improve selection efficiency of pedigree method. **Scientia Agrícola**, v.65, p.25-30, 2008a. DOI: 10.1590/S010390162008000100004.
- NUNES, J.A.R.; RAMALHO, M.A.P.; FERREIRA, D.F. Inclusion of genetic relationship information in the pedigree selection method using mixed models. **Genetics and Molecular Biology**, v.31, p.73-78, 2008b. DOI: 10.1590/S141547572008000100015.
- PANTER, D.M.; ALLEN, F.L. Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. **Crop Science**, v.35, p.397-405, 1995a. DOI: 10.2135/cropsci1995.0011183X003500020020x.
- 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**, v.35, p.405-410, 1995b. DOI: 10.2135/cropsci1995.0011183X003500020021x.
- PATTERSON, H.D.; THOMPSON, R. Recovery of interblock information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971. DOI: 10.1093/biomet/58.3.545.
- PIEPHO, H.P.; MÖHRING, J.; MELCHINGER, A.E.; BÜCHSE, A. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica**, v.161, p.209-228, 2008. DOI: 10.1007/s1068100794498.
- PIRES, M.J. de S.; SANTOS, G.R. dos. **Modelo agroexportador, política macroeconômica e a supremacia do mercado: uma visão do modelo brasileiro de exportação**. Brasília: Instituto de Pesquisa Econômica Aplicada, 2013. 29p. (IPEA. Texto para discussão, 1817).
- RAMALHO, M.A.P.; CARVALHO, B.L.; NUNES, J.A.R. Perspectives for the use of quantitative genetics in breeding of autogamous plants. **ISRN Genetics**, v.2013, p.16, 2012. DOI: 10.5402/2013/718127.
- RODRIGUES, J.I. da S.; MIRANDA, F.D. de; FERREIRA, A.; BORGES, L.L.; FERREIRA, M.F. da S.; GOODGOD, P.I.V.; PIOVESAN, N.D.; BARROS, E.G. de; CRUZ, C.D.; MOREIRA, M.A. Mapeamento de QTL para conteúdos de proteína e óleo em soja. **Pesquisa Agropecuária Brasileira**, v.45, p.472-480, 2010. DOI: 10.1590/S0100204X2010000500006.
- ROSADO, A.M.; ROSADO, T.B.; ALVES, A.A.; LAVIOLA, B.G.; BHERING, L.L. Seleção simultânea de clones de eucalipto de acordo com produtividade, estabilidade e adaptabilidade. **Pesquisa Agropecuária Brasileira**, v.47, p.964-971, 2012. DOI: 10.1590/S0100204X2012000700013.
- SEGURA, V.; VILHJALMSSON, B.J.; PLLATT, A.; KORTE, A.; SEREN, U.; LONG, Q.; NORDBORG, N. An efficient multilocus mixedmodel approach for genomewide association studies in structured populations. **Nature Genetics**, v.44, p.825-830, 2012. DOI: 10.1038/ng.2314.
- SILVA, J.A. da. Avaliação do Programa Nacional de Produção e Uso do Biodiesel no Brasil – PNPB. **Revista de Política Agrícola**, v.22, p.18-31, 2013.
- SILVA, R.R.; BENIN, G.; SILVA, G.O. da; MARCHIORO, V.S.; ALMEIDA, J.L. de; MATEI, G. Adaptabilidade e estabilidade de cultivares de trigo em diferentes épocas de semeadura, no Paraná. **Pesquisa Agropecuária Brasileira**, v.46, p.1439-1447, 2011. DOI: 10.1590/S0100204X2011001100004.
- STORCK, L.; CARGNELUTTI FILHO, A.; LÚCIO, A.D.C.; MISSIO, E.L.; RUBIN, S. de A.L. Avaliação da precisão experimental em ensaios de competição de cultivares de soja. **Ciência e Agrotecnologia**, v.34, p.572-578, 2010. DOI: 10.1590/S141370542010000300007.

VIANA, J.M.S.; ALMEIDA, I.F. de; RESENDE, M.D.V. de; FARIA, V.R.; SILVA, F.F. BLUP for genetic evaluation of plants in noninbred families of annual crops. **Euphytica**, v.174, p.31-39, 2010a. DOI: 10.1007/s1068100900942.

VIANA, J.M.S.; DELIMA, R.O.; FARIA, V.R.; MUNDIM, G.B.; RESENDE, M.D.V. de; SILVA, F.F. Relevance of pedigree, historical data, dominance, and data unbalance for selection efficiency. **Agronomy Journal**, v.104, p.722-728, 2012. DOI: 10.2134/agronj2011.0358.

VIANA, J.M.S.; FARIA, V.R.; SILVA, F.F. e; RESENDE, M.D.V. de. Best linear unbiased prediction and family selection in crop species. **Crop Science**, v.51, p.2371-2381, 2011. DOI: 10.2135/cropsci2011.03.0153.

VIANA, J.M.S.; SOBREIRA, F.M.; RESENDE, M.D.V. de; FARIA, V.R. Multitrait BLUP in halfsib selection of annual crops. **Plant Breeding**, v.129, p.599-604, 2010b. DOI: 10.1111/j.14390523.2009.01745.x.

---

Recebido em 4 de março de 2013 e aprovado em 2 de agosto de 2013