Genetics and Molecular Biology, 31, 1, 73-78 (2008)

www.sbg.org.br

Research Article

Inclusion of genetic relationship information in the pedigree selection method using mixed models

José Airton Rodrigues Nunes[1], Magno Antonio Patto Ramalho[2] and Daniel Furtado Ferreira[3]

[1]*Departamento de Planejamento e Política Agrícola, Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brazil.*
[2]*Departamento de Biologia, Universidade Federal de Lavras, Lavras, MG, Brazil.*
[3]*Departamento de Ciências Exatas, Universidade Federal de Lavras, Lavras, MG, Brazil.*

Abstract

We used a mixed model approach and computer simulation to evaluate the inclusion of parentage information as determined by the genealogy established in the pedigree method. The simulations were based on a purely additive genetic model for one quantitative trait of 20 unlinked segregating loci with equal effects and an allelic frequency of 0.5 for heritability values of 10%, 25%, 50% and 75% for selection based on an $F_{4:5}$ progeny mean. We simulated 1000 experiments for each heritability value, corresponding to the evaluation of 256 $F_{4:5}$ progenies. The phenotypic values of the progenies were analyzed according to two models, one ignoring and one considering the additive genetic parentage among the progenies. The additive relationship coefficients among $F_{4:5}$ progenies ranged from 0.0 to 1.75. The evaluated selection procedures were the phenotypic progeny mean (*M*) and the best linear unbiased predictor including parentage ($BLUP_A$). The inclusion of parentage among progenies using the $BLUP_A$ procedure resulted in higher selection gains than when the relationship information was ignored, which possibly recompenses the additional work invested to obtain these records, above all in the case of low - heritability traits.

Key words: autogamous crops, *BLUP*, computer simulation, plant breeding.

Received: February 2, 2007; Accepted: June 11, 2007.

Introduction

The pedigree method, proposed towards the end of the 19th century, is widely applied to improvement programs of self-fertilized plants and is mainly based on recording the genealogies among progenies over the selfing generations (Ramalho *et al.*, 2001). However, not only is does this procedure require time and dedication from the breeder but the usefulness of this method for the selection process is somewhat restricted. One possibility of using this parentage information in support of the selection process would be in progeny evaluations in experiments with replications. Such an approach could be useful since breeders of autogamous species are primarily interested in selecting progenies that, during homozygosis, accumulate a higher quantity of favorable alleles that associate the best additive genetic values (*AGV*), bearing in mind that the ultimate aim is the establishment of lines (Fehr, 1987). For quantitative traits, however, the phenotype does not always reflect the associated *AGV*. In this case, it would be important to use methodologies that optimize the use of the available information, in order to classify the progenies as closely as possible to the ranking given by the true *AGV* (White and Hodge, 1989). Several fixed model and mixed model procedures have been proposed to predict the *AGV* of progenies, including the best linear unbiased estimator (*BLUE*) method, the best linear predictor (*BLP*) technique and the best linear unbiased predictor (*BLUP*) approach (White and Hodge, 1989; Mrode, 1996; Lynch and Walsh, 1998; Resende, 2002).

The *BLUP* procedure has been the most widely used in the prediction of the genetic merit in animals (Mrode, 1996) and, more recently, it has been widely applied in plant improvement (Bernardo, 2002; Resende, 2002). Under unbalanced conditions this procedure not only has the advantage of making predictions more reliable compared to those obtained by the ordinary least square method but also incorporates information on related plants and thus optimizes the use of the available data in progeny comparisons (Bernardo, 2002).

Since we found no reports on the use of genealogy established by the pedigree method in progeny selection in

Send correspondence to José Airton Rodrigues Nunes. Departamento de Planejamento e Política Agrícola, Centro de Ciências Agrárias, Universidade Federal do Piauí, Campus Socopo, Bairro Ininga, 64049-550 Teresina, PI, Brazil. E-mail: jarnunes@ufpi.br.

self-pollinated crops and field experiments produce unreliable information (Wang *et al.*, 2003) we evaluated the efficiency of selection incorporating this genetic relationship using a mixed model computer simulation.

Methodology

The program was implemented in the Delphi 6.0 environment (Cantú, 2002). A simplified genetic model was assumed for any quantitative trait considering 20 loci of independent segregation, with equal and additive effects and an allelic frequency of 0.5 without dominance. The simulations considered heritability values of 10%, 25%, 50% and 75% for selection based on an $F_{4:5}$ progeny mean (h_p^2). For each h_p^2 heritability we simulated 1000 F_2 populations with 20 segregating loci consisting of 64 plants each. The plant multiplication rates were assumed to be equal, with each plant generated 40 offspring.

Initially, the generations were advanced by the pedigree method with no visual selection. A segregating F_2 population of 64 simulated plants gave rise to the 64 $F_{2:3}$ progenies with 40 plants each. Two plants were randomly selected from each $F_{2:3}$ progeny, resulting in 128 $F_{3:4}$ progenies and the process repeated in the following generation to finally obtain 256 $F_{4:5}$ progenies with 40 plants each (Figure 1).

The phenotypic values for the plants of each $F_{4:5}$ progeny (y_i) were simulated by adding normally distributed random errors to the genotypic values (*GV*), by the following model:

$$y_i = \mu + g_i + w_i,$$

where μ is a constant (100 in the present case), g_i is the genotypic effect of plant i ($i = 1, 2, ..., 40$) and w_i is the environmental deviation associated to y_i.

The g_i effect result from the cumulative effect of the 20 loci as already described in the genetic model. The additive effect (a_l) of the l^{th} locus was assumed equal to 1.0, where $l = 1, 2, ..., 20$. The value of g_i taking locus B with two alleles (B^1 and B^2) as reference is given by:

$$g_i = \sum_{l=1}^{20} a_l \quad \text{where } a_l = \begin{cases} 1 & \text{if } B^1B^1 \\ 0 & \text{if } B^1B^2. \\ -1 & \text{if } B^2B^2 \end{cases}$$

The w_i effects were randomly attributed based on a normal distribution with constant variance, *i.e.*, $N(0, \sigma_w^2)$. The variance component σ_w^2 is the environmental variance among plants, which can be obtained by:

$$\sigma_w^2 = \left[\frac{1 - h_{F_2}^2}{h_{F_2}^2}\right]\sigma_G^2,$$

where σ_G^2 is the genetic variance among the F_2 plants (*i.e.*, $\sigma_G^2 = \sigma_A^2 + \sigma_D^2$), σ_A^2 is the F_2 additive variance with, in this case, an allelic frequency of 0.5 $\sigma_A^2 = \sum_{l=1}^{20} a_l^2 / 2$, and since a_l was assumed equal to 1.0 for all loci, then $\sigma_A^2 = L/2$, σ_D^2 is the F_2 variance dominance and since dominance was assumed to be absent, then $\sigma_D^2 = 0$, and $h_{F_2}^2$ is the F_2 generation individual heritability.

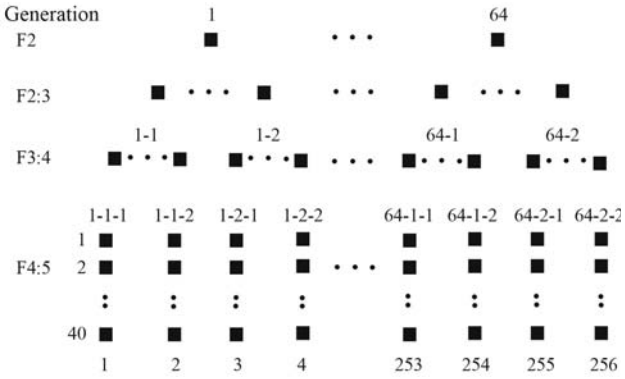

Figure 1 - Scheme of conduction by pedigree method.

The 40 simulated genotypes or plants per $F_{4:5}$ progeny were divided into two virtual plots of 20 plants ($n = 20$) to produce two replications ($r = 2$) for each progeny. In the following equations, random errors were considered to be normally distributed among plots, with $e \sim N(0, \sigma_e^2)$ in relation to the mean phenotypic values of the plots. The σ_e^2 variance component is the environmental variance among plots.

In the simulation the relation σ_w^2 / σ_e^2 was considered fixed at eight ($c = 8$). The error terms varied according to values assumed for h_p^2 heritability:

$$h_p^2 = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_d^2}{nr}},$$

where σ_p^2 is the genetic variance among $F_{4:5}$ progenies ($\sigma_p^2 = 7/4\sigma_A^2$), σ_d^2 is the phenotypic variance within a plot ($\sigma_D^2 = \sigma_{Gd}^2 + \sigma_w^2$, where σ_{Gd}^2 is the genetic variance within $F_{4:5}$ progenies given by $\sigma_{Gd}^2 = 1/8\sigma_A^2$). Thus the individual F_2 heritability ($h_{F_2}^2$) was determined as a function of the pre-fixed h_p^2 heritability values as:

$$h_{F_2}^2 = \frac{\sigma_{Gd}^2\left(1 + \frac{n}{c}\right)}{\frac{nr\sigma_p^2(1 - h_p^2)}{h_p^2} - \sigma_{Gd}^2 + \sigma_{Gd}^2\left(1 + \frac{n}{c}\right)}.$$

We analyzed 1000 experiments corresponding to the evaluation of 256 simulated $F_{4:5}$ progenies, derived from the pedigree method. The analysis was based on the mean phenotypic data of the plots, using a completely randomized experimental design with two replications.

According to the description of the conduction by pedigree method, each F_2 plant generated four $F_{4:5}$ progenies (Figure 1). Based on this detailed pedigree, the matrix

of the additive genetic parentages among the related progenies was determined, considering the F_2 population as non-inbred. The phenotypic progeny data were then analyzed according to two models:

Model G_I

The genetic relationship among progenies was ignored. The mean phenotypic data of the plots of the 256 $F_{4:5}$ progenies were analyzed using a linear mixed model (Henderson *et al.*, 1959) $\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{a} + \mathbf{e}$, where **y** is a 512 x 1 vector of the mean phenotypic data of the plots, **X** is a 512 x 1 fixed effect design matrix, β is a scalar fixed effect of the constant, **Z** is a 512 x 256 random effects of progenies design matrix, **a** is a 256 x 1 progeny random effects vector with $a \sim N(0, G)$ and $G = A\sigma_a^2$, while **e** is a 512 x 1 vector of errors with $e \sim N(0, R)$ and $R = I\sigma_e^2$. The *G* matrix was designated by $I\sigma_p^2$ (*i.e.*, **A** = **I**), indicating that the progenies were assumed to be unrelated. In this case, the σ_a^2 component is equal to the genetic variance among $F_{4:5}$ progenies (σ_p^2).

Model G_A

In this model the genetic relationship among progenies was considered by the inclusion of parentage among progenies. The mixed model for analysis was identical to model G_I, except that the *G* matrix was designated by $\mathbf{A}\sigma_A^2$, with **A** containing the additive relationship coefficients among $F_{4:5}$ progenies, corresponding to twice the Malecot's coancestry coefficient (Bernardo, 2002: section 2.3.5.2), and σ_a^2 refers directly to the F_2 additive variance among plants (σ_A^2). In animal breeding the *A* matrix is referred to as the *numerator relationship matrix* (Mrode, 1996) and, in this case, it was given by:

$$\mathbf{A} = \mathbf{I}_{64} \otimes \begin{bmatrix} 1{,}75 & 1{,}50 & 1{,}00 & 1{,}00 \\ 1{,}50 & 1{,}75 & 1{,}00 & 1{,}00 \\ 1{,}00 & 1{,}00 & 1{,}75 & 1{,}50 \\ 1{,}00 & 1{,}00 & 1{,}50 & 1{,}75 \end{bmatrix},$$

where $\otimes$ is the Kronecker product.

The solutions for the random (**â**) and fixed effects ($\hat{\beta}$) for both models were obtained by solving the following equation (Henderson *et al.*, 1959):

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

To obtain the previous solutions, the components of genetic and non-genetic variances were assumed to be unknown. These variance components were estimated using the restricted maximum likelihood (REML) method (Patterson and Thompson, 1971). Since the REML method employs an iterative process, the expectation-maximization (EM) numeric algorithm was applied (Dempster *et al.*, 1977).

The predictions of the progeny random effects (**â**) based on the overall adjusted mixed model are *BLUP* predictions (Henderson, 1975). After an adjustment of the G_I model the predictions were denoted as $BLUP_I$, while for the G_A model the predictions were designated $BLUP_A$. Additionally, the phenotypic progeny means (*M*) for each simulated experiment were also obtained.

It should be noted that due to the balancing conditions under which the simulation were conducted and the use of an orthogonal experimental design with no missing data the $BLUP_I$ predictions do not have selective advantage in relation to the phenotypic means of the progenies *(M)* (Kennedy and Sorensen, 1988). Thus, only the results using the mean *M* will be shown, and these should be understood as being equal to $BLUP_I$.

For each pre-fixed h_p^2 heritability, corresponding to 1000 simulated experiments, we obtained the mean estimates of the genetic variance among the $F_{4:5}$ progenies ($\hat{\sigma}_p^2$) and heritability on an $F_{4:5}$ progeny mean basis ($\hat{h}_p^2$) for both models (G_I and G_A). The selection procedures of the $F_{4:5}$ progenies (mean *M* and $BLUP_A$) were evaluated and compared based on the true genotypic values (*GV*) so for both procedures we estimated the Spearman correlations (r_S), proportions of coincidence in the 5%, 10% and 25% selection fractions for lower and upper extremes on the ranking of progenies, and the mean *GV* for different percentages (0.4% (best progeny), 5%, 10% and 25%) of the superior selected progenies.

The relative efficiency (*RE*) of $BLUP_A$ in relation to the mean *M* was determined by $RE = \{[r_{S(BLUPA,\ GV)} - r_{S(M,\ GV)}] / r_{S(M,\ GV)}\}$ x 100, where $r_{S(BLUPA,\ GV)}$ is the Spearman correlation between $BLUP_A$ and *GV* of the selected progenies, and $r_{S(M,\ GV)}$ is the Spearman correlation between the mean *M* and *GV* of the selected progenies. The relative efficiency was obtained also using proportions of coincidence. We also calculated the relative gain (*RG*) of $BLUP_A$ in relation to the mean *M* using $RG = \{[MGV_{BLUPA} - MGV_M] / MGV_M\}$ x 100, where MGV_{BLUPA} is the mean genotypic values of the selected progenies calculated by $BLUP_A$ while MGV_M is the mean genotypic values of the selected progenies calculated by the mean *M* method.

Results

For both models, the mean estimates of the genetic parameters associated with the $F_{4:5}$ progenies were close to the pre-fixed parametric values for all the h_p^2 heritabilities studied (Table 1). Nevertheless in all the evaluations the genetic parameter estimates by the G_A model, which includes parentage among progenies, were more accurate than those produced by the G_I model. For instance, for 25% h_p^2 heritability the standard error associated with the $\hat{\sigma}_p^2$ estimate in the G_A model was 33.5% but was 44.4% for the G_I model. However, when 50% h_p^2 heritability was considered the same percentages were very similar at 21.3% for the G_A

Table 1 - Mean estimates of the genetic variance among $F_{4:5}$ progenies ($\hat{\sigma}^2_p$) and heritability on a $F_{4:5}$ progeny mean basis ($\hat{h}^2_p$) and standard errors (values in brackets) according to model G_I (ignoring the pedigree information) and model G_A (considering the pedigree information) for different values of heritability h^2_p.

h^2_p (%)	Model G_I		Model G_A	
	$\hat{\sigma}^2_p$	$\hat{h}^2_p$ (%)	$\hat{\sigma}^2_p$	$\hat{h}^2_p$ (%)
10	20.48(17.90)	11.12(9.22)	18.04(12.55)	10.02(6.58)
25	17.31(7.68)	24.23(9.43)	17.59(5.90)	24.78(6.85)
50	17.35(3.88)	49.25(6.96)	17.45(3.71)	49.46(6.30)
75	17.48(2.80)	74.65(3.75)	17.45(2.63)	74.72(3.60)

model and 22.4% for the G_I model (Table 1). This demonstrates that it is advantageous to take into account genealogy (as normally occurs when using the pedigree method), although this advantage decreases as the character heritability increases ($h^2_p \geq 50\%$).

The selection units (mean *M* and $BLUP_A$) were evaluated regarding the correct ranking of $F_{4:5}$ progenies using the true associated genotypic values (*GV*) as reference. As expected, the mean correlation estimates r_S of the evaluated procedures were directly proportional to the h^2_p heritability values (Table 2). The h^2_p heritability represents a determination coefficient between the *M* and *GV* means, so that the mean values of the correlation estimates ($r_{S(M,\ GV)}$) can be used to verify the quality of the simulations, since they are approximate estimators of $\sqrt{h^2_p}$ (Falconer and Mackay, 1996). The $r_{S(M,\ GV)}$ correlation values were near the expected ($\sqrt{h^2_p}$) values for all the h^2_p heritabilities studied (Table 2), *e.g.* for 25% h^2_p heritability the mean $r_{S(M,\ GV)}$ correlation estimate was 0.48 and therefore close to the population value of 0.5.

The $r_{S(BLUPA,\ GV)}$ mean correlations between $BLUP_A$ and *GV* were superior to the $r_{S(M,GV)}$ mean correlation values for all the h^2_p heritability values studied (Table 2), demonstrating that the incorporation of genetic relationships results in greater efficiency regarding the correct classification of progenies, particularly in situations where h^2_p heritability was less than 50%. For example, for 10% h^2_p heritability the relative efficiency (*RE*) of $BLUP_A$ to mean *M* was 43.33% while for 50% h^2_p heritability the *RE* dropped to only 14.5%, this being confirmed by the high $r_{S(M,\ BLUPA)}$ correlation (0.87) between $BLUP_A$ and mean *M* (Table 2).

The identification of the progenies in the extremes on their ranking is of greater relevance for breeders than the classification of all the progenies evaluated. For this we estimated the coincidence proportions ($C_{(BLUPA,\ GV)}$) of selected progenies using the $BLUP_A$ and mean *M* methods and compared the results with selected progenies based on the real *GV* (Table 3) and found that for a fixed selection fraction (*s*) value the corresponding proportions of estimated coincidences in the lower and upper selected extremes were identical.

Table 2 - Mean estimates of the Spearman correlation (r_S) and standard errors (in parentheses) between genotypic values (*GV*), phenotypic means (*M*) and *BLUP* predictions considering the additive parentage ($BLUP_A$) among $F_{4:5}$ progenies, conducted by the pedigree method for different values of heritability on a $F_{4:5}$ progeny mean basis (h^2_p).

h^2_p (%)	$r_{S(M\ ,\ GV)}$	$r_{S(BLUPA\ ,\ GV)}$	$r_{S(BLUPA\ ,\ M)}$
10	0.30(0.06)	0.43(0.09)	0.69(0.04)
25	0.48(0.05)	0.62(0.06)	0.76(0.04)
50	0.69(0.04)	0.79(0.04)	0.87(0.03)
75	0.85(0.02)	0.89(0.02)	0.95(0.01)

For all h^2_p heritability and selection fractions *s* values the $C_{(BLUPA,\ GV)}$ between $BLUP_A$ and *GV* were higher than the $C_{(M,\ GV)}$ between the mean *M* and *GV*, (Table 3), supporting our r_S estimates (Table 2). As mentioned above, the *RE* of the $BLUP_A$ in relation to mean *M* in the coincidences with *GV* was proportionally greater for lower h^2_p heritability values and selected fractions (*s*). For example, for 10% h^2_p heritability and *s* = 5% $C_{(BLUPA,\ GV)}$ was 0.21 and $C_{(M,\ GV)}$ 0.15 (an *RE* of 40%), while at the same h^2_p heritability but with *s* = 25% *RE* and was only 15.4%. When h^2_p heritability was 50% *RE* dropped to 26.3% for *s* = 5% and 13.3% for *s* = 25% (Table 3). This indicates that the efficiency of $BLUP_A$ could possibly be higher when breeders work with a trait of low heritability and apply high selection intensity.

Breeders want the selected progenies to have the highest possible genetic values, which ultimately reflect the gain achieved with selection, disregarding the progeny by environment interaction. In the selected fractions (*s*) comparing the *GV* means of the $BLUP_A$-selected progenies with the mean *M* for the pedigree method it can be seen verify

Table 3 - Mean values of the proportions of coincidences (*C*) and standard errors (values in brackets) in the selection proportions (*s*) of 5%, 10% and 25% of the superior or inferior $F_{4:5}$ progenies, conducted by the pedigree method, ranked by the parametric genotypic values (*GV*), phenotypic means (*M*) and *BLUP* considering the additive parentage ($BLUP_A$) for different values of heritability on a $F_{4:5}$ progeny mean basis (h^2_p).

s (%)	h^2_p (%)	$C_{(M\ ,\ GV)}$	$C_{(BLUPA,\ GV)}$	$C_{(BLUPA,\ M)}$
5	10	0.15(0.09)	0.21(0.13)	0.40(0.11)
	25	0.24(0.11)	0.33(0.14)	0.48(0.11)
	50	0.38(0.12)	0.48(0.14)	0.60(0.10)
	75	0.57(0.12)	0.63(0.13)	0.75(0.09)
10	10	0.22(0.07)	0.29(0.11)	0.47(0.08)
	25	0.33(0.08)	0.42(0.11)	0.54(0.07)
	50	0.46(0.09)	0.56(0.10)	0.66(0.07)
	75	0.63(0.08)	0.69(0.08)	0.79(0.06)
25	10	0.39(0.05)	0.45(0.07)	0.60(0.05)
	25	0.48(0.05)	0.56(0.06)	0.66(0.04)
	50	0.60(0.05)	0.68(0.06)	0.75(0.04)
	75	0.74(0.04)	0.78(0.04)	0.85(0.03)

Table 4 - Mean genotypic values and standard errors (values in brackets) in the selection proportions (s) of 0.4% (best progeny), 5%, 10% and 25% of the superior $F_{4:5}$ progenies, conducted by the pedigree method, ranked by the phenotypic means (*M*) and *BLUP* considering the additive parentage ($BLUP_A$) for different values of heritability on a $F_{4:5}$ progeny mean basis (h_p^2).

s (%)	h_p^2 (%)	Selection procedure	
		Phenotypic mean	$BLUP_A$
0.4	10	103.9(3.8)	104.7(3.7)
	25	105.5(3.4)	106.9(3.3)
	50	108.0(2.9)	108.7(2.7)
	75	109.6(2.4)	109.9(2.3)
5	10	102.6(1.2)	103.7(1.5)
	25	104.2(1.2)	105.4(1.4)
	50	105.9(1.1)	106.7(1.2)
	75	107.3(1.1)	107.6(1.1)
10	10	102.3(0.9)	103.2(1.1)
	25	103.6(0.9)	104.7(1.1)
	50	105.1(0.9)	105.8(1.0)
	75	106.2(0.9)	106.5(0.9)
25	10	101.7(0.6)	102.3(0.8)
	25	102.6(0.7)	103.4(0.7)
	50	103.7(0.7)	104.2(0.7)
	75	104.6(0.6)	104.8(0.6)

that the $BLUP_A$ procedure offers an advantage at all the h_p^2 heritabilities studied, although with lower relative gains (*RG*). The *RG* increased continuously as h_p^2 heritability and *s* decreased (Table 4), *e.g.*, for 10% h_p^2 heritability and $s = 0.4\%$ the *RG* for $BLUP_A$ was 0.77%, while for $s = 25\%$ it was 0.59%. With higher h_p^2 heritabilities *RG* and at $h_p^2 = 50\%$ $RG = 0.65\%$ for $s = 0.4\%$ and 0.48% for $s = 25\%$.

Discussion

The fact that the dominance effect is not included in our genetic model does not constitute a severe restriction because the simulation involved $F_{4:5}$ progenies that represent only 7/64 of the dominance variance (Ramalho *et al.*, 2001). Furthermore, most of the characters of self-fertilized plants, including grain yield, usually show a non-expressive dominance effect (Souza and Ramalho, 1995; Novoselovic *et al.*, 2004). Van Oeveren and Stam (1992) have also verified that the dominance has little importance in computer simulations of autogamous crops.

A restriction of the simulation was the lack of visual selection, normally occurring in the pedigree method, during the conduction stages (Fehr, 1987). However, there are many literature reports on the inefficiency of visual selection for characters with low (< 50%) heritability, which is the case for most characters of economic importance (Silva *et al.*, 1994; Cutrim *et al.*, 1997). Thus, taking two random plants to generate subsequent progenies probably causes no expressive effect on the results, especially for h_p^2 heritabilities lower than 50%.

It is worth mentioning that the $BLUP_A$ and mean *M* estimators are phenotypic data functions that both predict additive genetic values (*AGV*) associated with progenies. The best estimator is therefore the one that results in the *AGV* ranked closest to the ranking by the true *AGV* (White and Hodge, 1989). It should be noted that, with the adoption of the G_A model, the predictions of the random effect of progenies ($\hat{\mathbf{a}}$) or $BLUP_A$ correspond to the predictions of the additive genetic value (*AGV*) of the progenies (Lynch and Walsh, 1998), indicating the theoretical superiority of the $BLUP_A$ procedure in relation to mean *M*.

An important aspect must be mentioned concerning the meaning of unbiasedness for *BLUP*, more specifically for $BLUP_A$. As mentioned above, in the present context $BLUP_A$ is a predictor of the *AGV* of progenies (**a**) derived from the same breeding population, whose expectation, by definition, is zero $[E(\mathbf{a}) = 0]$ (Falconer and Mackay, 1996). In this context, $BLUP_A$ is unbiased in the sense that $E(\hat{\mathbf{a}}) = E(\mathbf{a})$ (Robinson, 1991), where $\hat{\mathbf{a}}$ denotes the *AGV* predictors. The conclusion that can be drawn is, differently from the concept of unbiasedness for estimators of fixed effects, that the unbiasedness property for *BLUP* does not refer to predictions of individual random effects $[E(\hat{\mathbf{a}}) = \mathbf{a}]$ but to the expected value of these effects. Summing up, when $h_p^2 \to 100\%$, $\hat{\mathbf{a}} = E(\mathbf{a}/\mathbf{y}) \to \mathbf{a}$, while with $h_p^2 \to 0$ we have $\hat{\mathbf{a}} = E(\mathbf{a}/\mathbf{y}) \to 0$, demonstrating that the shrinkage effect in $BLUP_A$ predictions is more marked when the h_p^2 values are low, resulting in lower $r_{S(M,\ BLUPA)}$ correlation estimates. Thus the results of simulation showed in a concordant way that when h_p^2 heritability diminishes information on parentage becomes more important, so that with higher heritability h_p^2 (> 50%) the genotypic values are already well-determined by the mean phenotypic values (*M*) (Duarte and Vencovsky, 2001).

In general, our simulation showed that the inclusion of parentage among the progenies of the pedigree method using the $BLUP_A$ procedure resulted in slightly higher selections gains and more accurate estimates of genetic parameters than when this relationship information was ignored. This possibly compensates for the additional work invested in obtaining these records, especially when investigating low-heritability traits. Our results are supported by other published research showing that higher selection gains can be reached when using the $G\text{-}_A$ model or $BLUP_A$ procedure (Durel *et al.*, 1998; Bromley *et al.*, 2000). A study by Panter and Allen (1995) comparing two *BLUP* models (with and without the inclusion of information about genetic parentage between lines) for prediction of soybean crossings showed no marked differences between the *BLUP* models, yet the model which takes parentage into consideration performed better.

Acknowledgments

This research was financially supported by the Brazilian Agencies CAPES and CNPq. The authors gratefully acknowledge Dr. Eduardo Bearzoti for his excellent comments and suggestions.

References

Bernardo R (2002) Breeding for Quantitative Traits in Plants. Stemma Press, Woodbury, 359 pp.

Bromley CM, Van Vleck LD, Johnson BE and Smith OS (2000) Estimation of genetic variance in corn from F_1 performance with and without pedigree relationship among inbred lines. Crop Sci 40:651-655.

Cantú M (2002) Dominando o Delphi 6: A Bíblia. MAKRON Books, São Paulo, 1104 pp.

Cutrim VA, Ramalho MAP and Carvalho AM (1997) Eficiência da seleção visual na produtividade de grãos de arroz (*Oryza sativa* L.) irrigado. Pesq Agropeq Bras 32:601-606.

Dempster A, Laird N and Rubin D (1977) Maximum likelihood from incomplete data via the EM Algorithm. JR Stat Soc Ser B 39:1-38.

Duarte JB and Vencovsky R (2001) Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. Sci Agric 58:109-117.

Durel CE, Laurens F, Fouillet A and Lespinasse Y (1998) Utilization of pedigree information to estimate genetic parameters from large unbalanced data sets in apple. Theor Appl Genet 96:1077-1085.

Falconer DS and Mackay TFC (1996) Introduction to Quantitative Genetics. 4th ed. Longman, London, 464 pp.

Fehr WR (1987) Principles of Cultivar Development: Theory and Technique. MacMillan Publishing Company, New York, 527 pp.

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. Biometrics 31:423-447.

Henderson CR, Kempthorne O, Searle SR and Von Krosigk CM (1959) The estimation of environmental and genetic trends from records subject to culling. Biometrics 13:192-218.

Kennedy BW and Sorensen DA (1988) Properties of mixed-model methods for prediction of genetic merit under different genetic models in selected and unselected populations. In: Weir B, Goodman MM and Namkoong G (eds) Second International Conference Quantitative Genetics. North Carolina State University, Raleigh, pp 91-103.

Lynch M and Walsh B (1998) Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc., Sunderland, 948 pp.

Mrode RA (1996) Linear Models for the Prediction of Animal Breeding Values. Biddles, Guildford, 184 pp.

Novoselovic D, Baric M, Drezner G, Gunjaca J and Lalic A (2004) Quantitative inheritance of some wheat plant traits. Genet Mol Biol 27:92-98.

Panter DM and Allen FL (1995) 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 Sci 35:405-410.

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

Ramalho MAP, Abreu AFB and Santos JB (2001) Melhoramento de espécies autógamas. In: Nass LL, Valois ACC, Melo IS and Inglis MCV (eds) Recursos Genéticos e Melhoramento de Plantas. Fundação MT, Rondonópolis, pp 201-230.

Resende MDV (2002) Genética Biométrica e Estatística no Melhoramento de Plantas Perenes. Embrapa Informação Tecnológica, Brasília, 975 pp.

Robinson GK (1991) That *BLUP* is a good thing: The estimation of random effects. Stat Sci 6:15-51.

Silva HD, Ramalho MAP, Abreu AFB and Martins LA (1994) Efeito da seleção visual para produtividade de grãos em populações segregantes do feijoeiro. II. Seleção entre famílias. Cienc Prat 18:181-185.

Souza GA and Ramalho MAP (1995) Estimates of genetic and phenotypic variance of some traits of dry bean using a segregating population from the cross Jalo x Small White. Rev Bras Genet 18:87-91.

Van Oeveren AJ and Stam P (1992) Comparative simulation studies on the effects of selection for quantitative traits in autogamous crop: Early selection versus single seed descent. Heredity 69:342-351.

Wang J, van Ginkel M, Podlich D, YE G, Trethowan R, Pfeiffer W, Delacy IH, Cooper M and Rajaram S (2003) Comparison of two breeding strategies by computer simulation. Crop Sci 43:1764-1773.

White TL and Hodge GR (1989) Predicting Breeding Values with Applications in Forest Tree Improvement. Kluwer Academic Publishers, Dordrecht, 363 pp.

Senior Editor: Ernesto Paterniani