



Strategies in identifying individuals in a segregant population of common bean and implications of genotype x environment interaction in the success of selection

M.P. Mendes¹, M.A.P. Ramalho² and A.F.B. Abreu³

¹Departamento de Genética e Melhoramento de Plantas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidades de São Paulo, Piracicaba, SP, Brasil

²Universidade Federal de Lavras, Lavras, MG, Brasil

³EMBRAPA Arroz e Feijão, Goiânia, GO, Brasil

Corresponding author: M.P. Mendes

E-mail: celapmendes@yahoo.com.br

Genet. Mol. Res. 11 (2): 872-880 (2012)

Received April 13, 2011

Accepted October 18, 2011

Published April 10, 2012

DOI <http://dx.doi.org/10.4238/2012.April.10.3>

ABSTRACT. The objective of this study was to compare the BLUP selection method with different selection strategies in $F_{2,4}$ and assess the efficiency of this method on the early choice of the best common bean (*Phaseolus vulgaris*) lines. Fifty-one $F_{2,4}$ progenies were produced from a cross between the CVIII8511 x RP-26 lines. A randomized block design was used with 20 replications and one-plant field plots. Character data on plant architecture and grain yield were obtained and then the sum of the standardized variables was estimated for simultaneous selection of both traits. Analysis was carried out by mixed models (BLUP) and the least squares method to compare different selection strategies, like mass selection, stratified mass selection and between and within progeny selection. The progenies selected by BLUP were assessed in advanced generations, always selecting the greatest and smallest sum of the

standardized variables. Analyses by the least squares method and BLUP procedure ranked the progenies in the same way. The coincidence of the individuals identified by BLUP and between and within progeny selection was high and of the greatest magnitude when BLUP was compared with mass selection. Although BLUP is the best estimator of genotypic value, its efficiency in the response to long term selection is not different from any of the other methods, because it is also unable to predict the future effect of the progenies x environments interaction. It was inferred that selection success will always depend on the most accurate possible progeny assessment and using alternatives to reduce the progenies x environments interaction effect.

Key words: Plant breeding; Genetic improvement; Selection efficiency

INTRODUCTION

Several characteristics should be considered in the selective process to obtain common bean cultivars that meet producer and consumer requirements. Erect plant architecture associated with high grain yield has received much attention. The most erect plants have several advantages, including less lodging, that makes crop treatment easier and permits mechanical harvesting, reduction in harvest losses because there is less pod contact with the soil, better grain quality and lower incidence of some diseases.

When the objective is to select more erect and productive plants, first segregant populations must be obtained for these traits for later assessment and selection of the most promising populations. After identifying the populations, the next step is conducting the segregant population. The alternatives available include the BULK and BULK within F_2 progenies. By these methods, starting at the F_2 generation, the populations are advanced for some generations to F_4 or F_5 when progenies are selected for more intense assessment in replicated experiments.

There are several alternatives that can be used to identify the individuals/progenies to continue the selective process. They include mass selection and between and within selection, the latter when progenies derived from F_2 plants are used. Another alternative that is little used in common bean breeding is the BLUP procedure (best linear unbiased predictor) that is focused on mixed models analysis. According to Bernardo (2002), the BLUP procedure is useful because it permits analysis of unbalanced data, supplying, in these conditions, more reliable predictions than those obtained by the least squares (LS) method. It also allows the estimate of genetic values closer to the true genotypic value (Resende, 2002; Nunes et al., 2008). However, these selection alternatives have not yet been compared under field conditions nor has the efficiency of these methods been assessed in the identification of individuals that will originate the best lines. Thus, this study aimed to compare alternatives in the identification of the best individuals in the conduct of populations by the bulk method within F_2 progenies and assess the selection efficiency after identification of the best lines.

MATERIAL AND METHODS

The experiments were carried in Lavras, MG, in the experimental area of the Universidade Federal de Lavras, situated at 21°14'S latitude, 44°59'W longitude and 919 m altitude; in Lambari, MG, in the experimental area belonging to the Empresa de Pesquisa Agropecuária

de Minas Gerais (EPAMIG), situated at 21°58'S latitude, 45°21'W longitude and 887 m altitude, and in Patos de Minas, MG, also in the experimental area of EPAMIG, situated at 18°34'S latitude, 4°31'W longitude and 832 m altitude.

Fifty-one $F_{2,4}$ progenies derived from the cross between two lines of carioca-type grain were used: RP-26 with erect plant architecture and CVIII8511 that does not have erect plant architecture but has good productivity and cooking qualities. A randomized block design was used with 20 replications and plots of one plant. The greatest care was taken to obtain the same distance among the plants. Three seeds were used per hill plot. Fifteen days after emergence, the plants were thinned leaving one plant per hill plot. The traits assessed were: 1) grain yield, in grams per plot; 2) plant architecture, using a scale of scores modified by Collicchio et al. (1997) ranging from 1 to 9 where score 9 refers to the II-type plant, erect, with a stem with high insertion of the first pods and score 1 for the III-type plant, with long internodes and complete lodging.

In order to obtain information on the two traits simultaneously, the data were standardized and the index obtained by the sum of the standardized variables ($\sum Z$; Mendes et al., 2009) was used. As the Z variable can assume both positive and negative values, a constant of value three was added to make all the values positive. In this case, the population mean per trait, instead of zero, assumed the value three. The use of this index in the simultaneous selection of the two traits was based on the assumption that greater Z values are better.

The analysis of variance was performed by the mixed models method, BLUP, using the data per plant in the $F_{2,4}$ generation. For this analysis, the Computerized Genetic Selection software - SELEGEN (Resende, 2007), model 59, was used to obtain individual genotypic values. The $F_{2,4}$ generation analysis was also performed by the LS method considering the randomized block design, with 20 replications and plots of one plant. The phenotypic value data were submitted to different selection strategies for later comparison with the data supplied by BLUP. In the mass selection, the 100 best and 100 worst individuals were selected in the $F_{2,4}$ generation in function of the $\sum Z$ regardless of the progeny or replication to which they belonged. In the stratified mass selection, the plants were divided in strata and each stratum was a replication. Thus, each stratum contained one plant from each progeny and there was a total of 51 plants per stratum. The five best and worst individuals were selected from each stratum in function of the $\sum Z$. As there were 20 replications, there was a total of 100 progenies in each group. To perform the between and within progeny selection, analysis of variance was carried out first using the LS method to obtain the $\sum Z$ means of the 51 progenies. The six best progenies were selected from these means (11.7% between progeny selection intensity) and within these, the 16 best of the 20 existing plants. The same was done with the group of worst progenies, totaling 96 plants in each group.

Using the mixed models methodology (BLUP) the 100 best and 100 worst plants were identified in function of the $\sum Z$ originating the $F_{4,5}$ progenies. These progenies were sown in November 2009 to increase the quantities of seed and obtain the $F_{4,6}$ progenies. Eighty-one $F_{4,6}$ progenies from the group of the best progenies and the same number from the group of the worst progenies were assessed for plant architecture and grain yield scores. These progenies were assessed in different experiments in the 2009 dry season, in Lavras. A single 9 x 9-lattice design was used.

In the 2009 winter season, the $F_{4,7}$ progenies were assessed in a single experiment, containing the 71 best and 71 worst progenies selected in function of the plant architecture scores, plus the two controls. The experiment was also carried out in Lavras using a triple 12 x 12-lattice design.

After analysis of variance, the 24 best and 23 worst $F_{4,8}$ progenies were selected. They were assessed with the two controls, in the wet season, sown in November 2009, in Lavras and

Lambari. A triple 7 x 7-lattice design was used.

The same F_{4,9} progenies were assessed in the dry season, sown in February 2010 in three locations, Lavras, Lambari and Patos de Minas. The experiment was carried out similarly to the previous one, using a triple 7 x 7-lattice design. Similarly, the 12 best and 11 worst F_{4,10} progenies were assessed in the 2010 winter season, in a triple 5 x 5-lattice design, but only in Lavras.

The plots in all the experiments consisted of two 2-m long rows. They were fertilized with 400 kg/ha of the 8-28-16 formula of N, P₂O₅ and K₂O, respectively, and 20 days after emergence, 200 kg/ha ammonia sulfate was applied in cover. The other crop treatments were those recommended for the common bean crop in the region.

The data for the yield, plant architecture score and $\sum Z$ of the F_{4,6}, F_{4,7}, F_{4,8}, F_{4,9} and F_{4,10} generations were submitted to analysis of variance using the LS method. Analyses were made per generation, considering as random the effect of progenies, replication and error. The other effects were considered to be fixed. Joint analysis was then performed on the common progenies in the different environments.

The BLUP efficiency was assessed by three alternatives: the mean of the progenies identified with the greatest and least $\sum Z$ estimate in the F_{2,4} compared to the other generations; percentage of coincidence in the ranking among the progenies considered as best or worst in F_{2,4} in the different generations and the number of common progenies among the 15 with the greatest or least $\sum Z$ in the F_{4,6} and F_{4,7} generations.

RESULTS AND DISCUSSION

As there were two traits to be considered simultaneously in the selection, yield and stand score, the best option was to use a selection index. There are many index options in the literature (Bernardo, 2002; Cruz and Carneiro, 2006). In the present study, the index obtained by the sum of the standardized variables ($\sum Z$) was used. This index is easy to estimate and interpret (Mendes et al., 2009). In the discussion, the emphasis is directed to this index in detriment to the isolated variables.

First, using the F_{2,4} generation data, the analyses were compared using the BLUP procedure and the LS method. Table 1 shows that when the 10 best progenies identified by the two analysis procedures were considered, the coincidence was 100%. Bernardo (2002) commented that when the design is completely balanced, BLUP and LS supply the same information. Because in the F_{2,4} generation each plot consisted of one plant, and 88 plots were lost, that is, 8.6% of the total, it was inferred that with this loss level there was no advantage in BLUP compared to LS. Similar

Table 1. Best F_{2,4} progenies identified by index Z using the least squares (LS) method and best linear unbiased prediction (BLUP).

| Classification | Index Z | |
|----------------|---------|----|
| | BLUP | LS |
| 1 | 47 | 47 |
| 2 | 15 | 15 |
| 3 | 14 | 14 |
| 4 | 43 | 43 |
| 5 | 23 | 23 |
| 6 | 7 | 32 |
| 7 | 32 | 7 |
| 8 | 5 | 5 |
| 9 | 24 | 24 |
| 10 | 4 | 4 |

results have been reported in this comparison, as in the case of the eucalyptus crop (Reis, 2010).

The individual BLUP gives an estimate that involves all the model variables, for example, the merit of the progeny, the individual in the progeny and even the replication where it is located (Resende, 2007). Thus, between and within progeny selection, especially because LS and BLUP identified the same progenies, should present great coincidence in the individuals to be selected. This indeed occurred. With the ΣZ as reference, the coincidence was greater than 80% (Table 2).

Table 2. Coincidence (in %) of the plants selected in the $F_{2.4}$ generation by different selection strategies compared to BLUP.

| Strategy | ΣZ | |
|--------------------------------------|------------|-----------|
| | 100 Best | 100 Worst |
| Mass selection | 44 | 36 |
| Stratified mass selection | 42 | 39 |
| Between and within progeny selection | 83 | 84 |

BLUP = best linear unbiased prediction.

Mass selection is the procedure that common bean breeders adopt to choose individuals that will originate progenies for assessment in advanced generations. This is normally carried out visually. However, the efficiency of visual selection is constantly questioned (Silva et al., 1994; Cutrim et al., 1997). In the present study, visual selection was not considered because the yield data and plant architecture scores were obtained from the $F_{2.4}$ plants. The proportion of coinciding individuals selected by mass and BLUP selection was less than 40%. Considering that there were 20 replications in the experiment, stratified bulk selection was also carried out, and there were 51 plants in each stratum, one from each progeny. This strategy also did not increase coincidence compared to mass selection (Table 2).

The results obtained, in the first moment, allowed inference that mass selection was not efficient compared to BLUP. However, it has already been argued that selection by different strategies does not necessarily identify the same progeny/individual, but rather individuals similar in terms of performance. Under this condition, the efficiency of the mass selection strategy would be underestimated. To demonstrate this fact, the selection differential of the different strategies was estimated (Table 3). It was observed that the selection differential was greatest in the mass selection strategy. It could be argued from this angle that mass selection was more efficient, but it should be emphasized that there is a shrinkage effect in BLUP that reduces the extreme values and thus the selection differential estimate is smaller. Furthermore, in the selection gain (SG) expression the selection differential should be multiplied by the heritability (h^2). In the case of BLUP, $h^2 = 1.0$. Thus, the estimate of the gain with selection would be the selection differential estimate. No reports were found of the use of BLUP to estimate the individual genotypic value to be used in common bean selection. According to Resende (2002), BLUP estimates the true genotypic value. This factor, however, is only valid if the heritability of the trait is very high (Nunes et al., 2008). If the h^2 is low, the estimate can be considered as the best genotypic predictor, but not necessarily the true genotypic value.

In the between and within progeny selection the h^2 between and h^2 within progeny selection should be used as weights, especially the latter should be of smaller magnitude. In the case of mass selection, it would be $SG = ds \times h_1^2$, where ds = differential of selection and h_1^2 is the heritability for selection at individual level, a value that is not normally high (Moreto et al, 2007). It can be inferred that the expected gain from mass selection and BLUP would be similar if h_1^2

was 74.3%, that is, $2.6/3.5 = 0.7428$. Although this h_1^2 estimate was not obtained, it would be difficult to be of this magnitude. However, the efficiency of mass selection compared to BLUP should not be considered of only 40% in function of the coincidence of the selected individuals.

Table 3. Estimates of the selection differential considering the $\sum Z$ in the $F_{2,4}$ generation for the different selection strategies.

| Strategy | Overall mean of $F_{2,4}$ (m_o) | Average of selected individuals from $F_{2,4}$ (m_s) | Selection differential = $m_s - m_o$ |
|--------------------------------------|-------------------------------------|--|--------------------------------------|
| BLUP | 8.47 | 11.07 | 2.6 |
| Mass selection | 8.47 | 12.1 | 3.5 |
| Stratified mass selection | 8.47 | 11.97 | 3.5 |
| Between and within progeny selection | 8.47 | 11.61 | 3.14 |

BLUP = best linear unbiased prediction.

Another focus of the present study was to verify the individual selection efficiency in the $F_{2,4}$ generation in more advanced generations. In this context, some alternatives were used. As divergent selection was applied in the $F_{2,4}$ generation, that is, the individuals with the best and worst performance, greatest and least $\sum Z$, the efficiency of this selection was observed in $F_{4,6}$. In the mean of the 79 progenies of each group, the progenies with greatest $\sum Z$ in $F_{4,6}$ presented $\sum Z$ 26% greater to the least $\sum Z$ in the $F_{4,6}$ generation (Table 4), indicating that the selection was efficient. This same finding was observed in the other generations. It should be pointed out, however, that it did not reflect only the selection efficiency in $F_{2,4}$ because selection was made again in a divergent manner in each generation. It was also expected that proportionally, the percentage of superiority of the progenies with greater $\sum Z$ would increase, but this did not occur.

Table 4. Mean of the $\sum Z$ in the different generations considering the group of progenies identified in $F_{2,4}$ as best and worst and the percentage by which the best plants surpassed the worst.

| Generation | $\sum Z$ | |
|------------|------------|-------|
| | Best | Worst |
| $F_{4,6}$ | 6.64 (26%) | 5.29 |
| $F_{4,7}$ | 6.68 (26%) | 5.31 |
| $F_{4,8}$ | 6.71 (27%) | 5.30 |
| $F_{4,9}$ | 6.60 (24%) | 5.32 |

Another way to assess selection efficiency in $F_{2,4}$ would be by ranking the progenies in $F_{4,6}$, that is, verify the proportion of plants identified with greater $\sum Z$ in the group of the best in $F_{4,6}$. Of the 79 plants classified with greatest $\sum Z$ in $F_{2,4}$, 58 retained the ranking. The same result was observed among the worst (Table 5). In the last generation, the coincidence increased slightly. It should be emphasized that the reverse selection effect was accumulative, as already mentioned. Again, it was evident that the early selection was efficient.

Because in the breeding program the interest is to identify a small group of superior progenies and thus reduce the assessment work, the agreement in the ranking of the 15 best or worst progenies in the different generations was considered to assess the efficiency of early selection. In this case, the coincidence was very small: between the $F_{2,4}$ and $F_{4,6}$ generations only three among the best and one among the 15 worst plants (Table 6). By this criterion, early selection was not efficient.

Table 5. Ranking of the best (B) and worst (W) progenies and their origins in the different assessment generations.

| Generation | Number of progenies and their origin in F _{2,4} | Ranking in the generation of evaluation for $\sum Z$ | Origin in F _{2,4} |
|-------------------|--|--|----------------------------|
| F _{4,6} | 79 B and 79 W | 79 Worst | 21 B 58 W |
| | | 79 Best | 21 W 58 B |
| F _{4,7} | 71 B and 71 W | 71 Worst | 14 B 57 W |
| | | 71 Best | 14 W 57 B |
| F _{4,8} | 23 W and 24 B | 23 Worst | 5 B 18 W |
| | | 24 Best | 5 W 19 B |
| F _{4,9} | 23 W and 24 B | 23 Worst | 6 B 17 W |
| | | 24 Best | 6 W 18 B |
| F _{4,10} | 11 W and 12 B | 11 Worst | 1 B 10 W |
| | | 11 Best | 1 W 11 B |

Table 6. Progenies that coincided between the 15 best and the 15 worst plants in the F_{2,4}, F_{4,6} and F_{4,7} generations and the strategy used in their selection.

| Generation | Selected progenies | Selection strategy |
|-------------------------------------|--------------------|--|
| F _{2,4} / F _{4,6} | B21 | All |
| | B34 | All |
| | B9 | All |
| | W55 | BLUP and between/within |
| F _{2,4} / F _{4,7} | B18 | All |
| | B36 | All |
| | W46 | BLUP |
| | W47 | BLUP, stratified mass and between/within |
| F _{4,6} / F _{4,7} | B8 | All |
| | B75 | BLUP and between/within |
| | B67 | BLUP |
| | W35 | All |
| | W27 | BLUP, mass and between/within |

BLUP = best linear unbiased prediction.

The efficiency of early selection, in other contexts, has already been assessed in the common bean crop (Patiño and Singh, 1989; Silva et al., 1994; Rosal et al., 2000). The results showed that the efficiency of early selection varied with the generation, but it was usually low. Bernardo (1991) showed that the correlation between the phenotypic performance of the plant or progeny in the F_i generation and the respective genotypic performance in the G_j generation ($r_{F_i G_j}$) was a function of $r_{F_i G_j} = r_{G_i G_j} \sqrt{h_x^2}$. As $r_{G_i G_j}$ is only a function of the coefficient of the inbreeding and this correlation is usually high (Bernardo, 1991) the $r_{F_i G_j}$ is a function predominantly of the heritability of the trait at the time of selection.

The h^2 for selection on the mean of the progenies for the $\sum Z$ in F_{2,4} was estimated by Mendes et al. (2009) as being 67%, that is, a relatively high and favorable value for early selection. However, it was pointed out that selection was made at individual level and therefore it was expected that the value would be lower than this.

The realized heritability (h^2_R) was estimated for the ΣZ by selecting 15% of the $F_{2:4}$ plants and response in the $F_{4:6}$ plants. The values obtained were 15.8% among the greatest ΣZ and 33% among those with least ΣZ , that is, lower values than reported by Mendes et al. (2009). The ratio of the difference among the h^2 estimates can, in principle, be a function predominantly of the progenies x environments interaction. This interaction was not considered in the expression presented by Bernardo (1991). The efficiency of early selection was reduced by the occurrence of the interaction. The h^2_R between $F_{2:4}$ and $F_{4:7}$ considering the 15% with greatest ΣZ was still less, only 8.7%. Again, the progenies x environments/generations interaction contributed to reducing the h^2_R . The existence of the progenies x environments interaction has been frequently reported in the literature for the common bean crop in the region (Pirola et al., 2002; Nunes et al., 2005; Bruzi et al., 2007), that reduces the efficiency of early selection.

It should be emphasized that all the three $F_{4:6}$ progenies that were among the 15 best $F_{2:4}$ plants would have been selected by all the selection strategies. Generally, the same fact was observed in the other cases where there was coincidence. It was observed that although BLUP is the best estimator of genotypic value, its efficiency in the response to long term selection is not different from any of the other methods, because it is also unable to predict the future effect of the progenies x environments interaction. It was inferred that selection success will always depend on the most accurate possible progeny assessment and using alternatives to reduce the progenies x environments interaction effect.

ACKNOWLEDGMENTS

The authors are grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for supporting this study.

REFERENCES

- Bernardo R (1991). Correlation between testcross performance of lines at early and late selfing generations. *Theor. Appl. Genet.* 82: 17-21.
- Bernardo R (2002). Breeding for Quantitative Traits in Plants. Stemma Press, Woodbury.
- Bruzi AT, Ramalho MAP and Abreu AFB (2007). Desempenho de famílias de cruzamento entre linhagens de feijões andinos e mesoamericanos em produtividade e resistência a *Phaeoisariopsis guseola*. *Cienc. Agrotec.* 31: 650-655.
- Collicchio E, Ramalho MAP and Abreu AFB (1997). Associação entre o porte da planta do feijoeiro e o tamanho dos grãos. *Pesq. Agropec. Bras.* 32: 297-304.
- Cruz CD and Carneiro PCS (2006). Modelos Biométricos Aplicados ao Melhoramento Genético. Editora UFV, Viçosa.
- Cutrim VA, Ramalho MAP and Carvalho AM (1997). Eficiência da seleção visual na produtividade de grãos de arroz (*Oryza sativa* L.) irrigado. *Pesq. Agropec. Bras.* 601-606.
- Mendes FF, Ramalho MAP and Abreu AFB (2009). Índice de seleção para escolha de populações segregantes de feijoeiro-comum. *Pesq. Agropec. Bras.* 44: 1312-1318.
- Moreto AL, Ramalho MAP, Nunes JAR and Abreu AFB (2007). Estimação dos componentes da variância fenotípica em feijoeiro utilizando o método genealógico. *Cienc. Agrotec.* 31: 1035-1042.
- Nunes JAR, Ramalho MAP and Abreu AFB (2005). Graphical method in studies adaptability and stability of cultivars. *Annu. Rep. Bean Improv. Coop.* 48: 182-183.
- Nunes JAR, Ramalho MAP and Ferreira DF (2008). Inclusion of genetic relationship information in the pedigree selection method using mixed models. *Genet. Mol. Biol.* 31: 73-78.
- Patiño H and Singh SP (1989). Respuesta a la selección visual para rendimiento en generations F_2 e F_3 en frijol, *Phaseolus vulgaris* L. (Temas actuales en el mejoramiento genético del frijol común, Documento 47). CIAT, Colombia, 417.
- Pirola LH, Ramalho MAP, Carneiro JES and Abreu AFB (2002). Natural selection and family x location interaction in the common (dry) bean plant. *Genet. Mol. Biol.* 25: 343-347.

- Reis CA (2010). Proposta de Índice de Seleção no Melhoramento Genético do Eucalipto. Doctoral thesis, Departamento de Genética e Melhoramento de Plantas, Universidade Federal de Lavras, Lavras.
- Resende MDV (2002). Genética Biométrica e Estatística no Melhoramento de Plantas Perenes. EMBRAPA Informação Tecnológica, Brasília.
- Resende MDV (2007). Matemática e Estatística na Análise de Experimentos e no Melhoramento Genético. Embrapa Florestas, Colombo.
- Rosal CJS, Ramalho MAP, Gonçalves FMA and Abreu AFB (2000). Seleção precoce para a produtividade de grãos no feijoeiro. *Bragantia* 59: 189-195.
- 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. Prática* 18: 181-185.