



RAPD marker assessment of self-pollinated inbreeding methods for common bean segregant populations

Ana Luiza Monteiro Castanheira¹ and João Bosco dos Santos^{1*}

Received 29 October 2003

Accepted 2 March 2004

ABSTRACT - RAPD markers were used to assess the potential five inbreeding methods have to release genetic variability. These methods were employed until family selection in the F₅ generation, plus the parents Carioca and Flor de Mayo, and the test cultivar Pérola. DNA was extracted from 16 plants per family for RAPD reaction. Twenty-two primers amplified DNA fragments linked to a number of grain yield loci, weight of 100 seeds, number of days to flowering, and reaction to the oidium and angular leaf spot. To estimate genetic similarities between every genotype pair in each inbreeding method 42 polymorphic bands amplified among the families and parents were used. Genetic similarities were UPGMA-clustered and grouped by multidimensional scaling. The bulk method was most efficient to release genetic variability, followed by bulk in F₂ families, SSD, bulk in F₃ families, and pedigree. The bulk method created families with higher similarities close to the Carioca, which is one of the regionally most grown cultivars.

Key words: *Phaseolus vulgaris*, self-pollinated inbreeding methods, RAPD, genetic similarity, common bean.

INTRODUCTION

Success in common bean breeding by hybridization depends on the trait mean and variability magnitude generated by segregant populations. The magnitude of this variability is determined by the parental genetic diversity as well as the self-pollinated inbreeding methods of segregant populations. Therefore, an adequate choice of parents and the conduction method of segregant populations ensure superior lines with selection.

The Genealogical (Pedigree), Population (Bulk), Single Seed Descent (SSD) and the Bulk within F₂ and F₃ families are some of the methods usually used to conduct the segregant population. The most efficient conduction for segregant population can be selected by assessing the line variability after the advance of the population over several generations. Although this procedure does assess the real variability, released by different conduction procedures and for each trait of interest, it has the disadvantage of requiring a lot of

experimental work. An alternative is to assess the variability released by conduction procedures, which make use of genetic diversity among lines.

Genetic diversity can be determined without experimental assessment by molecular markers such as RAPD (Random Amplified Polymorphic DNA) markers (Williams et al. 1990). However, it should be checked whether the population variability identified by the marker corresponds to the real variability, assessed for one or more traits of interest. Some studies which correlate genetic similarity determined by molecular markers with phenotypic diversity have shown that markers are efficient at predicting genetic diversity when several agronomic traits are taken into account, but not when only grain yield is considered (Duarte 1998, Machado 1999). However, there are results which indicate that in maize and soybean marker-assessed variability is correlated with yield (Ajmone Marsan et al. 1997, Barroso 2000).

¹ Departamento de Biologia, Universidade Federal de Lavras (UFLA), C.P. 37, 37.200-000, Lavras, MG, Brasil. *E-mail: jbsantos@ufla.br

On this background, our study verified whether the RAPD marker-assessed genetic similarity among families originating from different conduction methods of segregant populations indicates the grain yield variability of families.

MATERIAL AND METHODS

Two hundred and ninety-two families were used along with their two parents, Carioca and Flor de Mayo, and a control, the Pérola cultivar, evaluated in a trial carried out by Raposo (1999). Originally, this author conducted the segregant population by different methods to the F₄ generation when 100 to 121 plants were sampled randomly, the families assessed in F₅ for grain yield, and the 64 most productive of each inbreeding method selected for comparison. Fifty-six families were obtained by the genealogical method, 55 by the population method, 62 by the SSD method, 61 by the bulk within F₂ families method and 58 by the bulk within F₃ families method.

DNA extraction and amplification

About two grams of young leaves from 16 plants of each family were used for DNA extraction by a procedure similar to that used by Nienhuis et al. (1995). DNA amplification, electrophoresis, and data recording were also performed following a similar proceeding used by Nienhuis et al. (1995). Twenty-two 10-mers primers were used. Most of the polymorphic bands generated by these primers are linked to one or more QTLs responsible for grain yield, days to flowering, weight of 100 seeds, and reactions to the causal fungi of oidium and angular leaf spot (Melo 2000).

Molecular data analysis

Each polymorphic band was treated as a single trait. A matrix of 0 and 1 values was elaborated from the codes: presence (1) and absence (0) of 42 polymorphic bands in the families resulting from each inbreeding method. The Dice coefficient (Rohlf 1992) was used to estimate the genetic similarity (gs_{ij}) between each line pair by the expression $gs_{ij} = 2a/(2a + b + c)$, where a corresponds to the presence of the band in a pair of families, b corresponds to band presence only in the first family and c corresponds to band presence only in the second family.

The error (S_{gs}) associated with each similarity index was estimated according to Skroch et al. (1992), by the expression

$$S_{gs} = \sqrt{gs_{ij} \times \frac{1 - gs_{ij}}{(n - 1)}}$$

where n is the total number of combinations a , b , and c for each pair of families.

Similarities were cluster-analyzed by the UPGMA method (Rohlf 1992). Genetically different families were identified in the dendrograms from the estimate of the maximum significant similarity value (gs_m). The gs_m was estimated by the t test at the 1% probability level using the equation

$$gs_m = 1 - (t \times \bar{s}_{gs})$$

where t is the tabulated value of t with $n-2$ degrees of freedom and \bar{s}_{gs} is the mean error of all gs_{ij} . A biplot was worked out by the Multidimensional Scale Method (Rohlf 1992).

Besides the cluster analysis, an another analysis of genetic similarity (gs_{ij}) among the families of each inbreeding method was carried out by the IV method of the Griffing (1956) diallel procedure, according to the model

$$gs_{ij} = \bar{gs} + ggs_i + ggs_j + sgs_{ij} + e_{ij}$$

where \bar{gs} is the mean genetic similarity value for the families of each method; ggs_i the general genetic similarity of family i , ggs_j the general genetic similarity of family j ; sgs_{ij} the specific genetic similarity for the families i and j , and e_{ij} is the error associated with gs_{ij} (Barroso 2000). The ggs and grain yield of the families (Raposo 1999) of each method were correlated.

RESULTS

The extreme values of genetic similarity estimated by each inbreeding method with their respective amplitude and mean error are showed in Table 1. The amplitude of these values was greater in the population method, due to the greater diversity of the families, while the genealogical method showed the least amplitude.

Each inbreeding procedure brought forth: 23 genetically different families by the genealogical method, 43 by the population method, 28 by the bulk with F₂ families, 23 by the

Table 1. Minimum, maximum, and genetic similarity amplitude values of the families generated by each inbreeding method, with the respective mean errors

| Inbreeding method | Genetic similarity | | | Mean error |
|---------------------|--------------------|---------|--------|------------|
| | Minimum | Maximum | Range | |
| Pedigree | 0.3333 | 0.8800 | 0.5467 | 0.0791 |
| Bulk | 0.1142 | 0.8615 | 0.7473 | 0.0810 |
| Bulk/F ₂ | 0.2000 | 0.9122 | 0.7122 | 0.0760 |
| Bulk/F ₃ | 0.3243 | 0.9180 | 0.5937 | 0.0766 |
| SSD | 0.2777 | 0.8929 | 0.6152 | 0.0771 |

bulk method within F₃ families, and 36 by SSD. These numbers were obtained from the dendrograms constructed for each inbreeding method, and an example is shown in Figure 1 for the population method.

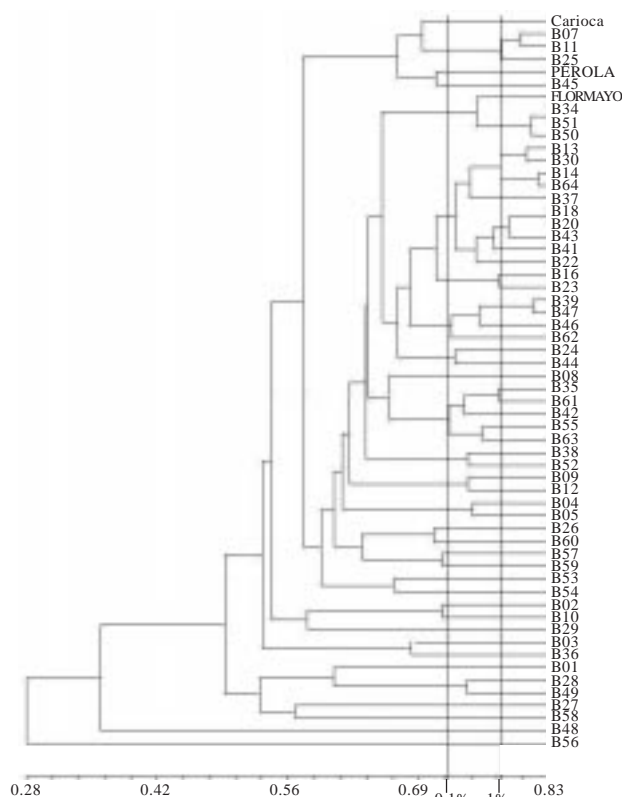


Figure 1. Genetic similarities of families obtained by the population method.

Families considered closest to the Carioca cultivar were also selected in the biplot of each method by the distance between the Carioca and Pérola cultivars as separation criteria. The population method presented most (10) families close to the Carioca cultivars: B 03, B 04, B 05, B 07, B 09, B 11, B 25, B 29, B 36, and B 45, as shown in Figure 2. With the same procedure, the SSD and bulk within the F₂ families both presented nine families close to the Carioca cultivar. The genealogical method had only one family close to the Carioca cultivar, and the bulk method within F₃ families did not show any family close to the Carioca cultivar. These results are in line with those shown in Table 1 and with the number of genetically different families produced by each inbreeding method. The biplot reduced the number of similarity estimates in two dimensions, and did therefore not include the totality of similarity variation. The precision of the biplot of similarities is given by the stress estimate (Rohlf 1992) with a value of 0.1906 for the population method, expressing good precision of the graphic representation.

When the families closest to the Carioca cultivar were analyzed based on similarity values, the result differed little from that shown in the biplots (Table 2). As already expected, these disagreements occurred in consequence of the grouping method and, mainly, because the similarities considered in the analysis were estimated only for the Carioca cultivar. In the plots, the two dimensions involved all similarity estimates.

In the case of the genealogical method, the family closest to the Carioca cultivar according to the plot was not included among the ten families with the highest similarities. By the population method, five of the ten families selected by the

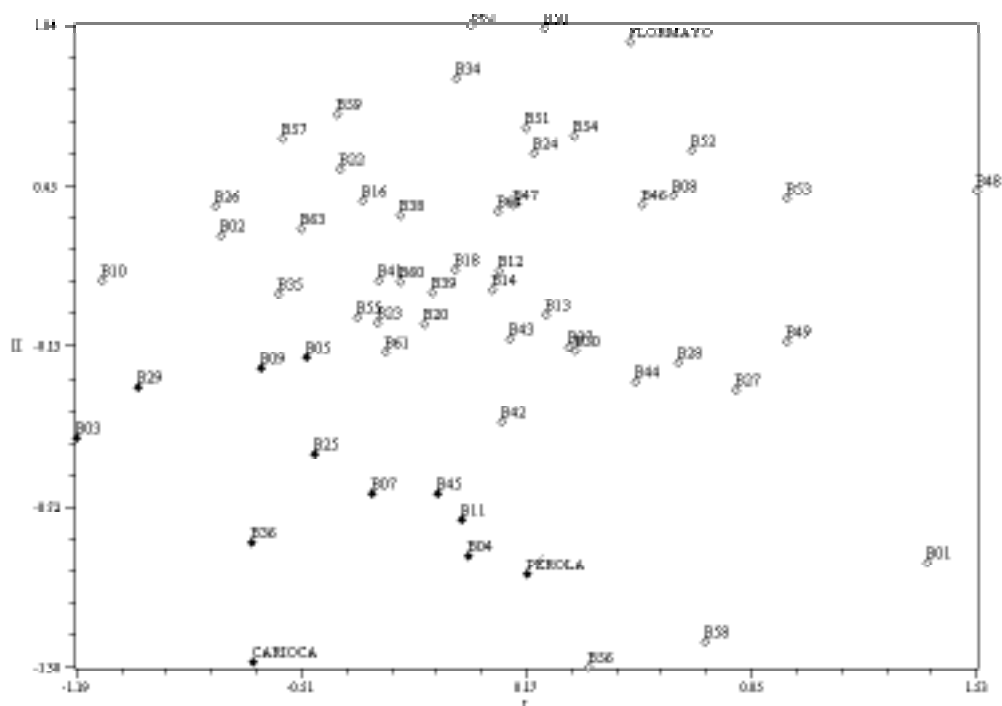


Figure 2. Biplot of the population method families, based on genetic similarity, with emphasis on the families which diverged from the Carioca cultivar, in the maximum equal to that of the Pérola cultivar (full losanges).

Table 2. The ten most similar families and the ten most divergent families, in relation to the Carioca cultivar, obtained by the bulk method, with their respective grain yields

| Most similar families | S _{ij} | Grain yield | Most divergent families | S _{ij} | Grain yield |
|-----------------------|-----------------|-------------|-------------------------|-----------------|-------------|
| B 11 | 0.727 | 499.080 | Flor de Mayo | 0 | 629.914 |
| B 07 | 0.682 | 597.916 | B 48 | 0.138 | 604.626 |
| B 25 | 0.682 | 605.376 | B 56 | 0.231 | 686.860 |
| B 09 | 0.667 | 565.578 | B 50 | 0.233 | 634.319 |
| B 39 | 0.653 | 682.750 | B 62 | 0.333 | 601.156 |
| Pérola | 0.651 | 628.708 | B 54 | 0.341 | 562.756 |
| B 20 | 0.627 | 599.266 | B 08 | 0.350 | 656.681 |
| B 14 | 0.600 | 551.569 | B 34 | 0.356 | 610.413 |
| B 41 | 0.600 | 561.049 | B 28 | 0.359 | 554.914 |
| B 03 | 0.579 | 639.949 | B 59 | 0.359 | 560.488 |
| Mean | 0.647 | 593.124 | | 0.270 | 616.513 |

biplot were found among the ten with greatest similarities to the Carioca, but the B 14, B 20, B 39, and B 41 families did not appear selected in the plot. By the bulk within F₂ families method, only two of the families selected by the plot were found among the ten with highest similarities for the Carioca. By the SSD method, only two families selected by the plot were found among the ten families with greater similarities compared to the Carioca.

Taking the mean of the ten most similar families to the Carioca cultivar, the five inbreeding methods were similar. The population method created the most divergent families for the mean of the ten families least similar to the Carioca. Furthermore, Table 2 shows that the population method, besides having created the most divergent families, also produced those most similar to the Carioca cultivar, which is one of the most grown in the region.

To identify the contribution of each family to the genetic similarity of the others, the similarity analysis by the diallel method showed that the general genetic similarity (ggs) was predominant. It was 51 times greater when the magnitudes of the mean square (MS) of the ggs were compared with the MS of the specific genetic similarity (ggs) in the population method. The MS of the ggs was 28 times greater in the genealogical method, 87 times greater in the bulk within F₂ families, 40 times greater in the bulk within F₃ families and 36 times greater in the SSD method.

Only the population method presented association with the correlation ($r = 0.49^{**}$) of the ggs estimates per family from each method with the respective grain yields.

DISCUSSION

The effect of each inbreeding method on the segregant population occurred up to the F₄ generation, when individual plants were sampled. Each plant generated a family, which

was assessed in F₅, and the 64 families with highest grain yield were selected (Raposo 1999). The markers identify the variability among families resulting from this selection. Theoretically, it would be expected that the magnitudes of genetic variability among the families of each inbreeding method were those presented in Table 3 (Ramalho and Venkovsky 1978).

Considering the components of expected genetic variance among families in the F₅ generation, the additive genetic variance (σ_A^2) is predominant. Furthermore, σ_A^2 is usually superior to the σ_D^2 of most traits in common bean (Ramalho et al. 1993). Specifically regarding genetic similarity, the part of genetic variance caused by the additive effects, corresponding to general genetic similarity (ggs), represents almost the totality. Consequently, it is expected that the population, SSD, and genealogical methods would release greater variability among families, followed by the bulk within F₃ families and bulk within F₂ families with the least potential.

Another point to take into consideration is that some factors may affect the genetic variance released among families derived from different inbreeding methods. These include the inherent sampling problems of each method. Such problems affect the

Table 3. Additive genetic variance (σ_A^2) and the dominance (σ_D^2) expected among the F₅ families resulting from different segregant population conduction methods

| Method | σ_A^2 | σ_D^2 |
|---------------------|-------------------|--------------------|
| Pedigree | 1.75 σ_A^2 | 0.109 σ_D^2 |
| Bulk | 1.75 σ_A^2 | 0.109 σ_D^2 |
| Bulk/F ₂ | 1.00 σ_A^2 | 0.016 σ_D^2 |
| Bulk/F ₃ | 1.50 σ_A^2 | 0.047 σ_D^2 |
| SSD | 1.75 σ_A^2 | 0.109 σ_D^2 |

effective size (N_e) of the population and, as Raposo (1999) points out, the sampling effect impinges most on the population method, while SSD is the only method not influenced by this effect. This author mentioned that, in the F_5 generation, the N_e of the population method was 0.55, while the N_e of the SSD method was 16. The author also emphasized that it is perfectly possible to use a population with many individuals in the population method, which decreases the chance of loss by sampling. For example, it is common to use 2000 or more plants for the population method (this would be equivalent to the effective size of 1040 plants in F_5 , which would be the equivalent to the use of 65 plants in the SSD method).

Results for genetic similarity among the families derived from the different methods are not in line with the theoretically expected results. Among them, the population method stands out for releasing greatest variability. The bulk within the F_2 generation, expected to be the potentially least efficient, was ranked second, while SSD was ranked third. On the other hand, the genealogical method, expected to be one of the most efficient, was ranked last.

However, the most important aspect to consider is that the results for genetic similarity were similar to those obtained by Raposo (1999) who used grain yield to assess the potential of the different inbreeding methods at releasing variation (Table 4). This shows that the employed markers supplied similar data to those obtained from grain yield. Thus the results of genetic similarity and grain yield show that the population method was the most efficient to release variability, followed,

Table 4. Estimate of correlations (r) among ggs and grain yield in $g\ plot^{-1}$, for families derived from each inbreeding method with the respective genetic variances (σ_G^2), and heritability (h^2) estimated by Raposo (1999)

| Method | r | σ_G^2 | h^2 |
|-------------|----------|--------------|-------|
| Bulk | 0.4860** | 4314.85 | 57.15 |
| Bulk/ F_2 | 0.1465 | 3701.78 | 53.36 |
| Bulk/ F_3 | -0.1647 | 2214.88 | 40.64 |
| Pedigree | -0.2488* | 1725.08 | 34.70 |
| SSD | -0.1333 | 3642.90 | 52.96 |

* $P < 0.05$; ** $P < 0.01$.

in decreasing order, by the bulk within F_2 families, SSD, bulk within F_3 families, and the genealogical method.

A specific feature of the population method is the easiness of handling the segregant population as the generations advance, thus allowing the use of more individuals, which should represent the released genetic variability better. Furthermore, natural selection acts with maximum intensity in this method and more adapted individuals are expected to be selected. Results showed, as expected, that this method was most efficient at releasing the population variability, and mainly, at favoring the achievement of some families, which are more similar to the Carioca cultivar, and therefore probably more adapted.

The positive correlation observed in families generated by the population method suggests that there is an association of the band patterns with grain yield. This means that families with similar values of ggs also have similar grain yields. Thus, again, the population method did not only favor the achievement of more divergent families, but must also have aided the selection of the most adapted, since the method exploits effects of natural selection. Favorable effects of natural selection can be exploited much more intensely in a tropical country such as Brazil, where common bean is cultivated year-round, as observed by Corte (1999). This author obtained gains of 2.4% per generation based on grain yield, after conducting the population in bulk for 18 generations. These gains are greater than those obtained by artificial selection with other inbreeding methods to advance the segregant populations (genealogical, bulk within families, SSD methods). These results indicate the superiority of the population method under tropical conditions compared to other methods, mainly to the genealogical method, which is still being used in some programs and the SSD method, which does not use natural selection.

The significant correlation for the population method also shows that the QTLs identified by Melo (2000), although of small effect, were useful to partially explain the grain yield.

ACKNOWLEDGEMENTS

The authors thank the CNPq (National Council of Scientific and Technological Development) for financial support to carry out this study.

Potencial de métodos de condução de populações segregantes de feijão avaliado por marcadores RAPD

RESUMO - Cinco métodos de condução de população segregante de feijão foram avaliados por marcadores RAPD quanto ao potencial de liberação de variabilidade genética. Os métodos foram aplicados até a geração F_4 , quando foram selecionadas famílias que foram avaliadas juntamente com os genitores 'Carioca' e 'Flor de Mayo' e a testemunha 'Pérola'. O DNA foi extraído de 16 plantas de cada família para as análises RAPD. Foram utilizados 22 primers que amplificaram 42 bandas

polimórficas, a maioria ligadas a QTLs da produtividade de grãos, peso de 100 sementes, número de dias para o florescimento e reação ao oídio e mancha angular. Para cada método de condução da população segregante, foram estimadas as similaridades genéticas das famílias, genitores e testemunha, as quais foram agrupadas pelos métodos UPGMA e escala multidimensional. O método da população demonstrou ser o mais eficiente para liberar a variabilidade genética, seguindo do bulk dentro de famílias F_2 , SSD, bulk dentro de famílias F_3 e genealógico. Além disso, o método da população gerou as famílias mais semelhantes ao 'Carioca', que é uma das cultivares mais cultivadas na região.

Palavras chaves: *Phaseolus vulgaris*, métodos de condução da população segregante, similaridade genética, RAPD.

REFERENCES

- Ajmone Marsan P, Castiglione P, Fusari F, Kuiper M and Motto M (1997) Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. **Theoretical and Applied Genetics** 96:219-227.
- Barroso PAV (2000) **Comparações entre métodos quantitativos e métodos baseados em marcadores RAPD para a predição do comportamento de populações de soja**. PhD Thesis, ESALQ, USP, Piracicaba, 162p.
- Corte HR (1999) **Comportamento de populações segregantes de feijão, avançadas pelo método do "bulk", por dezessete gerações**. MSc. Thesis, UFLA, Lavras, 95p.
- Duarte JM (1998) **Estudo da divergência genética em raças de feijão por meio de marcadores RAPD**. MSc. Thesis, UFLA, Lavras, 78p.
- Griffing JB (1956) Concept of general and specific combining ability in relation to diallel systems. **Australian Journal of Genetics and Plant Breeding** 51:448-453.
- Machado CF (1999) **Procedimentos para a escolha de genitores de feijão**. MSc. Thesis, UFLA, Lavras, 118p.
- Melo LC (2000) **Mapeamento de QTLs em feijoeiro, por meio de marcadores RAPD, em diferentes ambientes**. MSc. Thesis, UFLA, Lavras, 148p.
- Nienhuis J, Tivang J, Skroch P and Santos JB (1995) Genetic relationships among cultivars and lines of lima bean (*Phaseolus lunatus* L.) as measured by RAPD markers. **Journal of the American Society for Horticultural Science** 120:300-306.
- Ramalho MAP and Vencovsky R (1978) Estimação dos componentes da variância genética em plantas autógamas. **Ciência e Prática** 2:117-140.
- Ramalho MAP, Santos JB and Zimmermann MJO (1993) **Genética quantitativa em plantas autógamas: aplicações no melhoramento do feijoeiro**. Editora da UFG, Goiânia, 271p.
- Raposo FV (1999) **Comparação de métodos de condução de população segregante na cultura do feijoeiro**. MSc. Thesis, UFLA, Lavras, 73p.
- Rohlf FJ (1992) **Numerical taxonomy and multivariate analysis system**. Exeter Software, New York, 470p.
- Skroch P, Tivang J and Nienhuis J (1992) Analysis of genetic relationships using RAPD marker data. In: **Applications of RAPD technology to plant breeding**. Crop Science Society of America, Minneapolis, p.26-30.
- Williams JG, Kubelik AR, Livak KJ, Rafalski LA and Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research** 18:6531-6535.