Wilhelm Eigo Hagiwara; João Bosco dos Santos\* and Sidney L. M. do Carmo

Departamento de Biologia, Universidade Federal de Lavras (UFLA), Caixa Postal 37, CEP 37200-000, Lavras, MG, Brazil. (\*Corresponding Author. E-mail: jbsantos@ufla.br)

# ABSTRACT

Aiming at using RAPD markers to accelerate the recovery of desirable phenotypes in backcross programs, 167 common bean plants belonging to  $BC_1$ ,  $BC_2$  and  $BC_3$  resistant to the 2047 race of the fungus *Colletotrichum lindemuthianum* were selected after inoculation. The DNA of those plants and of the parents G2333 (donor), ESAL696 and CI140 (recurrents) was extracted and 70 polymorphic bands were obtained using 34 *primers*, through the RAPD procedure. Genetic similarities between BC plants and parents for each group were estimated. It was verified that RAPD was efficient in estimating genetic diversity among genotypes, as well as parentage for the  $BC_1F_4$  and  $BC_2F_1$  populations in relation to their parents, according to their genealogy. The efficiency of the genetic similarity for determining the parentage of the backcross populations in relation to their parents confirms its usefulness in selecting plants more similar to the recurrent parent, even in  $BC_1$ .

KEY WORDS: Phaseolus vulgaris L., genetic similarity, Colletotrichum lindemuthianum.

## INTRODUCTION

In the genetic improvement of common bean (*Phaseolus vulgaris* L.) in Brazil, emphasis has been given to the selection of lines with resistance to pathogens (Ramalho et al., 1993; Vieira et al., 1999). The backcross method has been thoroughly used, mainly because the donor parent of the resistance allele is little adapted and possesses several undesirable phenotypes. Depending on the level of the donor's adaptation up to 6 backcrosses are recommended to recover 99,2% of the alleles of the recurrent parent, which demands time for the obtaining of the improved line.

Considering that the group of alleles responsible for the expression of desirable agronomic and grain quality phenotypes are segregating since the first backcrossed generation, there is the possibility of recovering some superior plants even in the first backcross. That is possible with the use of molecular markers, like RAPD, which identify the genotypes with a larger proportion of the alleles of the recurrent parent already in the  $F_1$  generation of the first backcross and without need to wait for the conclusion of the life cycle of the plant, making possible a time reduction in the breeding program. The identification of the genotypes through RAPD can be done through the estimates of genetic similarity, that corresponds to the parentage of the segregating plants of each backcross and the recurrent parent. There is evidence that the genetic diversity identified by RAPD markers is correlated with morpho-agronomic diversity (Duarte et al., 1999; Machado et al., 2000). Several simulation results using RAPD markers to aid in the recovery of alleles of the recurrent parent indicated reduction varying from one to four backcross generations (Hospital et al., 1992; Openshaw et al., 1994; Visscher et al., 1996; Frisch et al., 1999). Experimental results with several crops have also already been obtained, showing the usefulness of RAPD in aiding the reduction of the number of backcrosses (Utumi, 1996; Carvalho, 1997; Arruda, 1998; Alzate-Marin et al., 1999, 2001).

Thus, the objective of this research was to estimate the genetic similarity, based on RAPD markers, among plants from different backcrosses, derived from crosses of the donor parent G2333 with the recurrent ESAL696 and CI140, aiming to identify those more similar to the recurrent parents.

# MATERIAL AND METHODS

#### Genotypes

The plants were obtained from the cross of the non-adapted G2333 with the ESAL696 line.

G2333 possesses the Co-42 allele for resistance against all races of the fungus Colletotrichum *lindemuthianum* that occur in Brazil. The F generation was crossed with ESAL696 and generation  $BC_1F_1$  was obtained. This generation was inoculated with the race 2047 of the pathogen and the resistant plants were selected. The generation  $BC_1F_4$  was obtained by self-fertilization. Plants  $BC_1F_3$  derived from resistant plants  $BC_1F_1$ were crossed with line CL40 that possesses excellent agronomic phenotypes, as recurrent parent, and a population similar to  $BC_2F_1$  was obtained, since both recurrent parents are adapted. The generation  $BC_2F_2$  was also obtained by selffertilization. The population  $BC_3F_1$  was obtained using CL40 as recurrent parent. The populations  $BC_1F_4$ ,  $BC_2F_1$ ,  $BC_2F_2$  and  $BC_3F_1$  were inoculated again with the 2047 race of the fungus *C*. *lindemuthianum*, the resistant plants were selected and divided in two groups; 73 BC<sub>1</sub>F<sub>4</sub> plants formed group 1 plus the parent ESAL696; and 88  $BC_{2}F_{1}$ plants,  $2 BC_2F_2$  and  $4 BC_3F_1$ , formed group 2 plusthe parents ESAL696 and CI 40.

# **DNA Extraction and RAPD Analysis**

DNA was extracted from each of the 167 resistant plants of the  $BC_1F_4$ ,  $BC_2F_1$ ,  $BC_2F_2$  and  $BC_3F_1$  generations and also from the parents, according to the procedure described by Nienhuis et al. (1995).

The DNA obtained was amplified by the RAPD procedure with 34 primers from "Operon Technologies" (California, USA), that identified polymorphisms in the parents. Each RAPD reaction was prepared in volume of 10ml, according to the procedure used by Nienhuis et al. (1995). The reactions were done in glass capillary tubes, in an air thermocycler (Idaho Technology, Idaho Falls, Idaho). The thermocycler was programmed for 40 cycles, in which the first two cycles were 60 seconds for the denaturing of DNA at 91°C, seven seconds for the annealing of primer at 42°C and 70 seconds for the elongation at 72°C. Thirty eight subsequent cycles differed in comparison to the first two cycles in the denaturing time, which was reduced by one second. Finally a four-minute stage was programmed at 72°C for the final elongation.

After the DNA amplification the fragments were separated by electrophoresis in a 1% agarose gel

in TBE buffer at 65 volts for 4 hours. The DNA fragments were stained with a ethydium bromide solution with 0.5mg/ml concentration, visualized in an ultraviolet light transluminator and photographed with 667 polaroid film.

The bands were visually classified as intense, medium and faint, based on the degree of resolution and amplification. Only the intense and medium bands were used for the analysis. In the gel, each band was considered as a unique character. From the bands obtained by those primers a matrix of 0 and 1 was built, in which 1 indicates presence of the band and 0 its absence. That matrix was used to obtain the estimates of genetic similarity between every pair of plants.

# Analysis of genetic similarity

The genetic similarity (gsij) was estimated using the Nei & Li coefficient (Rohlf, 1992), by the expression (gsij) = 2a/(2a + b + c), *a* corresponding to the presence of a certain band in the individuals *i* and *j*; *b* the presence of the band in *i* and absence in *j*; and *c* the absence of the band in *i* and presence in *j*. The errors associated with each similarity (sgs) were estimated using the expression (Skroch et al., 1992b): sgs =  $[gsij (1 - gsij)/(n - 1)]\frac{1}{2}$ , where *n* is the total number of *a*, *b* and *c* band patterns between each pair of plants.

The grouping of the similarities was accomplished through a dendrogram, using the NTSYS-PC 2,0 program (Rohlf, 1992). The hierarchical aglomerative method of the unweighted pair-group mean arithmetics (UPGMA) was used.

The genetically different plants were identified in the dendrogram by considering the estimate of the maximum significant value of similarity (*gsm*). The *gsm* was estimated through the *t* test using the expression:  $gsm = 1 - (t \ \overline{s}_{gs})$ , *t* being the value of *t* with n-2 degrees of freedom and  $\overline{s}_{gs}$  the mean error of the *gsij*.

Observed and expected similarity between each parent and the plants derived from the backcrosses of the  $BC_1F_4$  and  $BC_2F_1$  generations were also obtained. The observed similarity corresponds to the mean of the similarities observed between the plants from each backcross with each parent. The expected

similarity among plants of each backcross with the parents (*egs*) was calculated according to Skroch et al. (1992a). The expected similarity among BC, plants with the parents ESAL696 (egsBC, g,) and G2333 ( $egsBC_1g_2$ ) were calculated by the following expressions:  $egsBC_1g_1 = \%g_1 + \%g_2$  $gsg_1g_2$ ;  $egsBC_1g_2 = \%g_1$  '  $gsg_1g_2 + \%g_2$ . The expected similarity among BC<sub>2</sub> plants with the parents ESAL696 ( $egsBC_2g_1$ ), G2333 ( $egsBC_2g_2$ ) and CI140 ( $egsBC_2g_3$ ) were calculated by the expressions:  $egsBC_2g_1 = \%g_1 + \%g_2 \ gsg_1g_2 +$  $\%g_3 \circ gsg_1g_3; egsBC_2g_2 = \%g_1 \circ gsg_1g_2 + \%g_2 +$  $\%g_3 ' gsg_2g_3; egsBC_2g_3 = \%g_1 ' gsg_1g_3 + \%g_2$  $gsg_2g_3 + \%g_3$ , in which:  $\%g_1$ ,  $\%g_2$  and  $\%g_3$  is the expected proportion of the alleles of the parents ESAL696, G2333 and CI140, respectively, in the backcrossed population;  $gsg_1g_2$ ,  $gsg_1g_3$  and  $gsg_2g_3$  are the genetic similarities estimated between the parents, and  $g_1, g_2$  and  $g_3$  are the parents ESAL696, G2333 and CI140, respectively.

### **RESULTS AND DISCUSSION**

### **RAPD** analysis

The 34 primers used generated a total of 70 polymorphic bands in the backcrossed plants and parents. In a simulation study for marker assisted selection, Openshaw et al. (1994) recommended the use of four markers per chromosome, which would give 44 markers in the case of the common bean (n = x = 11). It should still be considered that the chromosomes of common bean are extremely short, compared with other species (Vieira et al., 1999). Hospital et al. (1992) showed that in the initial generations an increase of the number of markers to more than three per chromosome is not efficient. In his simulation research, Visscher et al. (1996) found gains of one to two selection generations in relation to the phenotypic selection using spaced markers from 10 to 20 centimorgans (cM), which in common bean would give approximately 60 to 120 markers considering the total length of common bean linkage map of 1226 cM, obtained by Freyre et al. (1998). Additionally, Johns et al. (1997) verified by resampling that 50 bands produced the same grouping obtained with 106 bands in a study with 69 lines of Chile common beans. In a similar study, Nienhuis et al. (1995) verified that for a number above 100 bands there is practically no improvement of the efficiency in the estimate of genetic distances. Thus, the number of polymorphic bands used in the present research can be considered appropriate.

### **Evaluation of genetic similarity**

For a better visualization of the genetic divergence among the plants, a dendrogram was obtained for each group (Figures 1 and 2). In those figures the line representing the maximum value of similarity (gsm) at 1% level of probability indicates that on the right side of it the plants are considered similar. Note that G2333 was the most distant in the two dendrograms, due to the smaller proportion of their alleles in the backcrossed plants. The line CI140 was included in the dendrogram with the plants  $BC_1F_4$ , in spite of not having participated as a parent for obtaining that generation. Evidently, it was distant from the backcrossed plants, however, not as much as G2333, due to its larger similarity with ESAL696. Only four plants  $BC_1F_4$  were similar to ESAL696 (Figure 1), showing that a wide variation exists among the plants. In Figure 2, ESAL696 was more distant from the plants  $BC_{2}$ and BC<sub>2</sub> in relation to CI140, however, there was a small group of plants more similar to ESAL696 than CI140. No plant was equal to any of the recurrent parents, which was expected due to the lower proportion of recurrent parents alleles in the BC<sub>2</sub> and BC<sub>3</sub> plants in comparison to the BC<sub>1</sub>F<sub>4</sub> plants with ESAL696. About 23% of the plants from the  $BC_1$  population are different from each other, while in the BC<sub>2</sub> and BC<sub>3</sub> population, 27% are different (Figure 1 and 2). That result was not expected because in BC<sub>1</sub> there are, an average 75% of the alleles of the recurrent parent and 25% of the donor, while in BC<sub>2</sub>, an average 87.5% of alleles of the recurrent and 12.5% of the donor was expected. Therefore, more variation was expected in BC<sub>1</sub>. However, as already pointed out, BC<sub>2</sub> was obtained using a second recurrent parent, CI140, that is considerably different from ESAL696 (Table 3, Figure 2). The procedure of changing the recurrent parent after BC<sub>1</sub> contributed to the increase of the genetic variability among the

plants, which in fact is favorable for the selection, increasing the chance of getting combinations of desirable phenotypes. The BC<sub>3</sub> plants, represented by the treatments 37, 60, 75 and 76 were not classified among the more similar to CI140, probably because those four plants do not represent the BC<sub>3</sub> population, and all of them came from the same BC<sub>2</sub> plant. Although larger variability was observed in BC<sub>2</sub> rather than in BC<sub>1</sub>, it is worth noting that the population of BC<sub>2</sub> plants is genetically more distinct from the donor parent, with a medium similarity value of 36%, in comparison with the BC<sub>1</sub> plants, whose medium similarity with the donor is 41% (Table 3).

Among the genetic similarities estimated between every pair of plants of each group, the 5 most similar and the 5 most distant plants of groups 1 and 2 compared to the parents are shown in Tables 1 and 2, respectively. The estimated similarities between the donor parent G2333 and the recurrents ESAL696 and CI140 were low, 0.20 and 0.25, respectively, while between the recurrent parents this value was larger (0.44), reflecting the degree of divergence among them. Even in BC<sub>2</sub> several plants presented similarities above 0.50 with G2333 (Table 2), indicating that many plants still present high proportions of its alleles. However, many plants recovered the alleles of the recurrent parent in proportions above the expected mean



**Figure 1** - Dendrogram of the genetic similarities among the  $BC_1F_4$  plants and the parents. The numbers refer to the resistant plants selected in the population  $BC_1F_4$ .

2001, Sociedade Brasileira de Melhoramento de Plantas



**Figure 2** - Dendrogram of the genetic similarities among the  $BC_2$ ,  $BC_3$  plants and the parents. The numbers refer to the resistant plants selected in the populations  $BC_2$  and  $BC_3$ .

(Table 3), reaching 0.91 of similarity with ESAL696 in plants from the first backcross (Table 1). In the plants of the second backcross (Table 2) there were also many plants that recovered the genotype of the recurrent parents above the expected mean (Table 3), reaching 0.78 of similarity with CI140 and 0.79 with ESAL696. However, those estimates are only valid for the sampled loci for RAPD and they do not necessarily reflect the recovering of the desired phenotypes of the recurrent parents.

Considering the  $BC_1F_4$  and  $BC_2F_1$  populations and the expected similarities based on the genealogy and based on the parents, the obseved similarities between parent and backrossed plants were quite high (Table 3), showing the efficiency of RAPD in the prediction of the degree of parentage among the backcross plants for the sampled loci by the markers. Analyzing 17 lines coming from two backcrosses in common bean, Skroch et al. (1992a) also found a good agreement in the value of expected similarity based on genealogy and that observed by RAPD markers.

	More distant				More similar					
BC <sub>1</sub> plant	84	17	83	167	3	145	70	82	85	16
ESAL 696	0.48	0.51	0.51	0.58	0.59	0.86	0.87	0.87	0.89	0.91
	More distant				More similar					
BC <sub>1</sub> plant	33	34	108	164	104	83	82	167	17	84
G2333	0.24	0.25	0.27	0.27	0.27	0.56	0.57	0.57	0.59	0.63

**Table 1** - Estimates of genetic similarities between the five plants more distant in relation to the parents and of the five more similar, of the BC<sub>1</sub>F<sub>4</sub> generation.

**Table 2** - Estimates of genetic similarities between the five backcrossed plants (BC) more distant in relation to the parents and the five more similar 1/.

	More distant			More similar						
BC plant	18	19	148	372	52	66	138	163	137	31
ESAL 696	0.55	0.56	0.56	0.57	0.57	0.76	0.76	0.76	0.79	0.79
		N	Iore dista	int		More similar			lar	
BC plant	106	60 <sup>2</sup>	91	75 <sup>2</sup>	24	19	121	63	89	61
G2333	0.27	0.28	0.29	0.30	0.30	0.52	0.52	0.52	0.53	0.53
	More distant					More similar				
BC plant	100 <sup>1</sup>	58	51	152 <sup>3</sup>	99	52	56	114	24	106
CI140	0.32	0.41	0.41	0.49	0.52	0.77	0.77	0.77	0.78	0.78

<sup>1/</sup> BC plant numbers without superscript are from  $BC_2F_1$  generation; <sup>2</sup> from  $BC_3F_1$ ; <sup>3</sup> from  $BC_2F_2$ .

The small deviations between the observed and expected similarities in this study are probably the result of chance, due to the sampling of the plants in selfing generations. In the case of BC<sub>1</sub>, the plants coming from the segregating  $F_4$  generation may have had a deviation in allele frequencies, when compared with BC<sub>1</sub>F<sub>1</sub>. The same may have happened with BC<sub>2</sub> population due to the use of F<sub>3</sub> generation. However, those results suggest that the genetic similarities between the backcrossed plants and the recurrent parents are an efficient indication of the proportion of parent alleles in each plant and, consequently, it should be useful to guide in the

#### 2001, Sociedade Brasileira de Melhoramento de Plantas

selection of those with a larger proportion of alleles of the recurrent parent. It is important to point out that in common bean the plant cycle is very short, and as many as three generations can be accomplished in one year. So, the reduction of the number of backcrosses will not contribute to a significant gain of time for obtaining an improved line. It is evident that a result like this would be much more useful for a perennial crop. However, it is also important to stress that the RAPD information may help to recover larger proportion of alleles of the recurrent parent already in the first backcross, which is important even for common bean.

	BC <sub>1</sub>	$F_4$	$BC_2F_1$			
	Observed	Expected	Observed	Expected		
ESAL696	0.72	0.80	0.68	0.62		
G2333	0.41	0.40	0.36	0.33		
CI140	-	-	0.63	0.70		

**Table 3** – Mean observed and expected genetic similarities of the populations  $BC_1F_4$  and  $BC_2F_1$  with the parents.

#### **RESUMO**

# Uso de RAPD para acelerar a seleção em programas de melhoramento por retrocruzamentos em feijão-comum

Uso de RAPD para acelerar a seleção em programas de melhoramento por retrocruzamentos em feijão comum, visando utilizar marcadores RAPD para acelerar a recuperação dos fenótipos desejáveis em programa de retrocruzamento, foram selecionadas 167 plantas  $RC_1$ ,  $RC_2$  e  $RC_3$  resistentes ao fungo Colletotrichum lindemuthianum, após a inoculação com a raça 2047. Foram extraídos o DNA dessas plantas e dos genitores G2333 (doador), ESAL696 e CI140 (recorrentes) e obtidos 70 bandas polimórficas utilizando-se 34 primers, por meio do RAPD. Foram estimadas as similaridades genéticas entre as plantas e genitores da cada grupo. Constatouse que o RAPD foi eficiente para estimar a diversidade genética entre os genótipos, bem como o parentesco das populações  $RC_1F_4 e RC_2F_1$  com os genitores de acordo com a genealogia delas. A eficiência do RAPD na predição do parentesco confirma sua utilidade para auxiliar na seleção de plantas nas populações segregantes mais semelhantes ao genitor recorrente, mesmo em RC<sub>1</sub>.

## ACKNOWLEDGMENTS

This research was supported by a grant from FAPEMIG. W.E. Hagiwara and S.L.M. do Carmo were supported by a scholarship from CAPES.

### REFERENCES

- Alzate-Marin, A.L.; Costa, M.R. and Dias, I.S.N. et al. 1999. Retrocruzamento assistido por marcadores moleculares visando acelerar a introgressão de genes de resistência à antracnose do feijoeiro. p.62-64. In: Anais da Reunião Nacional de Pesquisa de Feijão, 4th, Salvador, 1999. EMBRAPA-CNPAF-APA, Goiânia.
- Alzate-Marin, A.L.; Costa; M.R.; Sarorato, A.; Rava, C.A.; Barros, E.G.; Moreira, M.A. 2001. Use of markers as a tool to investigate the presence of disease resistance genes in common bean cultivars. Crop Breeding and Applied Biotechnology. 1:125-133.
- Arruda, M.C.C. 1998. Resistência do feijoeiro comum à antracnose: herança, identificação de marcadores moleculares e introgressão do gene Co-4 no cultivar Rudá. M.S. Thesis. Universidade Federal de Viçosa, Viçosa.
- Carvalho, W.L. 1997. Efeito de lipoxigenases sobre níveis de inibidores de proteases de sementes de soja e uso de marcadores moleculares e fenotípicos para acelerar programa de retrocruzamento visando eliminar inibidor de protease. D.S. Thesis. Universidade Federal de Viçosa, Viçosa.
- Duarte, J.M.; Santos, J.B. and Melo, L.C. 1999. Genetic divergence among common bean cultivars from different races based on RAPD markers. Genetics and Molecular Biology. 22:419-426.
- Freyre, R.; Skroch, P.W. and Geffroy, V. et al. 1998. Towards an linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. Theoretical and Applied Genetics. 97:847-856.

- Frisch, M.; Bohn, M. and Melchinger, A.E. 1999. Comparison of selection strategies for markerassisted backcrossing of a gene. Crop Science. 39:1295-1301.
- Hospital, F.; Chevalet, C. and Mulsant, P. 1992. Using markers in gene introgression breeding programs. Genetics. 132:1199-1210.
- Johns, M.A.; Skroch, P.W. and Nienhuis, J. et al. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Science. 37:605-613.
- Machado, C. de F; Santos, J.B.; Nunes, G.H. de S. and Duarte, J.M. 2000. Efficiency of genetic distance based on RAPD markers for choosing parents of common bean. Journal of Genetics and Breeding. 54:251-258.
- Nienhuis, J.; Tivang, J.; Skroch, P. and Santos, J.B. 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.
- Openshaw, S.J.; Jarboe, S.G. and Beavis, W.D. 1994. Marker-assisted selection in backcross breeding. In: Lower, R. (Ed.) Joint Plant Breeding Symposium on Analysis of Molecular Data. 1994. Oregon State University, Corvallis.

- Ramalho, M.A.P.; Santos, J.B. and Zimmermann, M.J. 1993. Genética quantitativa aplicada ao melhoramento de plantas autógamas: aplicações ao melhoramento do feijoeiro. Ed. UFG, Goiânia.
- Rohlf, F.J. 1992. Numerical taxonomy and multivariate analysis system. New York.
- Skroch, P.W.; Santos, J.B. and Nienhuis, J. 1992. Genetic relationships among *Phaseolus vulgaris* genotypes based on RAPD marker data. Annu. Rep. Bean Imp. Coop. 35:23-24.
- Skroch, P.W.; Tivang, J. and Nienhuis, J. 1992. Analysis of genetic relationships using RAPD marker data. p. 26-30. In: Proceedings of the Symposium on Applications of RAPD Technology to Plant Breeding, Minneapolis, 1992. Crop Science Society of America, Madison.
- Utumi, M.M. 1996. Melhoramento genético da proteína da soja: eliminação de lipoxigenasese das subunidades a' e G4, uso de RAPD-PCR na seleção de genótipos e na identificação de marcadores ligados ao gene Cgy1. D.S. Thesis. Universidade Federal de Viçosa, Viçosa.
- Vieira, C.; Borém, A. and Ramalho, M.A.P. 1999. Melhoramento do feijão. p.273-349. In: Borém, A. (Ed.). Melhoramento de espécies cultivadas. Ed. UFV, Viçosa.
- Visscher, P.M.; Haley, C.S. and Thompson, R. 1996. Marker-assisted introgression in backcross breeding programs. Genetics. 144:1923-1932.

Received: May 16, 2001; Accepted: September 05, 2001.