Crop Breeding and Applied Biotechnology 4:384-390, 2004
Brazilian Society of Plant Breeding. Printed in Brazil

ARTICLE

Sample size for number of RAPD markers to estimate genetic diversity in *Eucalyptus*

Edgard Augusto de Toledo Picoli[1], Acelino Couto Alfenas[1*], Cosme Damião Cruz[2], Diogo Faustini Moura[1], and Luiz Antônio dos Santos Dias[2]

Received 23 March 2004

Accepted 6 September 2004

ABSTRACT - *A total of 501 RAPD markers were generated for a population consisting of 84* Eucalyptus *genotypes, 89.8% of which were polymorphic and 10.2% monomorphic. Coherent genotype clustering based on the estimated genetic distances was observed for the species. Bootstrap analyses based on the number of RAPD markers revealed that 393 or more markers achieved values below 5% for stress, lower than 5 for sum of squares, and higher than 95% for correlation statistics. The estimated diversity based on the Shannon index amounted to about 95-96% for* E. urophylla *and 85-88% for* E. grandis. *In an 84-genotype population, each species presented significant differences for the estimated genetic distances. On the other hand, when the sample was reduced to 10 genotypes, the difference was no longer detected by the t-test at 1% probability.*

Key words: Tocher algorithm, coefficient of variation, stress, RAPD, genetic diversity.

INTRODUCTION

Eucalyptus is an important Myrtaceae genus used for charcoal, wood products and essential oil extraction. Pulp for paper earned imports of over 17 billion dollars worldwide. On this market, *Eucalyptus* is one of the most important genera for the cellulose and paper industries, besides the good adaptability and standing at the beginning of its domestication. These aspects awaken the interest for a better yield, product quality, and disease and stress resistance perspectives.

In this context, knowledge on the genetic diversity in a species is important for breeding programs and genetic conservation studies. Breeders have proposed several parameters for estimating genetic relationships among species lineages in which coefficients are based on kinship, pedigree, multivariate quantitative trait, and molecular analysis (Moser and Lee 1994). Molecular markers such as isozymes, RAPD, RFLP, AFLP, and SSR have greatly contributed to germplasm characterization, taxonomic relationships, marker-assisted selection, hybrid performance prediction, genetic distance estimation, mapping and diversity management (Kumar 1999, Joshi et al. 1999, Kwon et al. 2002, Barbosa et al. 2003, Dias et al. 2004).

Since they were first reported, RAPDs (Williams et al. 1990) have been used in studies involving several species (Dias et al. 2004). They produce considerable polymorphism besides being faster and less expensive than other molecular markers

[1]Departamento de Fitopatologia, BIOAGRO, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brasil. *E-mail: aalfenas@ufv.br
[2]Departamento de Biologia Geral, BIOAGRO, UFV

(Skabo et al. 1998). Further advantages are the uncomplicated protocols, lower amount of required DNA, and a wide genome cover. For *Eucalyptus*, they have been successfully used for several purposes (Keil and Griffin 1994, Nesbitt et al. 1995, Gaiotto et al. 1997, Skabo et al. 1998, Junghans et al. 2003). However, there is a wide variation in the number of markers used, from as few as 23 (Gaiotto et al. 1997) to as many as 415 (Baril et al. 1997).

In this context, the present work aimed to evaluate the importance of the number of RAPD markers on the genetic distance estimates and genetic diversity in an *Eucalyptus* population.

MATERIAL AND METHODS

Plant material, DNA extraction, and RAPD reactions

Leaf samples of *E. grandis* and *E. urophylla* genotypes from 84 different genotypes from a CENIBRA S/A breeding program were collected and stored at -80 ºC. DNA was extracted from these samples according to Ferreira and Grattapaglia (1995), diluted to 3 ng μL^{-1}, and stored at -20 ºC until being used for the RAPD reactions. RAPD reactions were conducted as described by Ferreira and Grattapaglia (1995) with Operon primers (Operon Technologies Inc), namely OPA 2, OPA 3, OPA 4, OPA 5, OPB 1, OPB 2, OPB 5, OPB 6, OPB 7, OPB 8, OPB 9, OPB 10, OPB 11, OPB 12, OPB 13, OPB 14, OPB 15, OPB 16, OPB 17, OPB 18, OPC 1, OPC 2, OPC 5, OPC 9, OPD 6, OPD 7, OPD 8, OPD 9, OPE 1, OPE 2, OPE 3, OPE 4, OPF 1, OPF 2, OPF 3, OPF 5, OPF 9, OPG 3, OPG 4, OPG 5, OPG 7, OPH 1, OPH 2, OPH 4, OPH 6, OPJ 12, OPJ 13, OPJ 14, OPJ 16, OPJ 18, OPJ 19, OPX 1, OPX 2, OPX 3, OPX 4, OPX 5, OPX 6, OPX 7, OPX 9, OPX 15, OPAC 8, OPAE 9, and OPAN 19. Each reaction contained 13 μL with 0.4 mM primer, 20 ng DNA, 100 μM of dGTP, dATP, dCTP, dTTP (Promega), and 1 Taq DNA polymerase unit. The reactions were poured into 96 wells in polycarbonate plates and cycled in a PTC-100 Thermocycler (MJ Research Inc.). Electrophoresis was carried out in agarose gel 1.5% in half concentrated TBE buffer containing 0.3 μ g L^{-1} ethidium bromide. The images were captured by the Eagle Eye Stratagene system and saved in JPG format for a visual analysis.

Genetic distance estimation

After electrophoresis, the RAPD profiles were scored visually for the presence (1) or absence (0) of bands. Based on these profiles, the distance matrix using the estimates of the arithmetic complement for the Nei and Li index was obtained using the software Genes (Cruz 2001). The population considered in all analyses was defined by one sub-population of 44 *E. urophylla* genotypes and one of 40 *E. grandis* genotypes, which came from 14 and the latter from 10 provenances, respectively.

Genetic distances were estimated for all genotypes and for one sample of ten genotypes, the former five from each species, respectively. In this analysis, the total number of markers or only the polymorphic markers for both sub-populations were used to estimate the genetic distances and further tested by the *t*-test at 1% probability using Gqmol software (Cruz 2003).

Effect of the number of markers on the estimated genetic distance

For the 84 genotypes, two hundred sets of 10 to 500 markers were randomly selected from the total of 501 RAPD markers. In this bootstrap analysis, the selected markers were always returned to the pool of markers, in a way that they could be picked up in the next sample. The genetic distances for all 84 genotypes were calculated per set of markers. Subsequently, the correlation (C), stress (S) and sum of squares (SS) of these estimated distances, in relation to the real genetic distances, were calculated for each of the 200 samples in each set of number of markers using software Gqmol. There was generated a total of 98000 distance matrixes for each C, S, and SS. In all cases, the real genetic distance was considered to be the distance based on the estimative using the 501 RAPD markers for each genotype.

Clustering and genetic diversity estimation

A dendrogram based on the unweighted pair-group method using an arithmetic average (UPGMA) was obtained with software Treecon (Van de Peer 1997). Tocher clustering algorithm was performed with Gqmol software. The graphical dispersion analyses of the coordinates plotting was performed based on the generated distance matrix in order to search for dispersion and grouping patterns considering species, as proposed by Cruz and Viana (1994).

The genetic diversity was calculated considering the total number of markers, the polymorphic markers for the entire population, and only the markers that were polymorphic for both sub-populations (Table 1). Five indexes of variation (IV) were employed. Index IV_1 corresponds to the total of polymorphic bands in the sub-population divided by the total of bands in the *E. urophylla* and *E. grandis* sub-populations, while IV_2 was calculated dividing the total of polymorphic bands in the sub-populations by the total of bands in the whole population. Index IV_3 was obtained with the ratio of the total of polymorphic bands in the sub-populations and the total of polymorphic bands in the whole population, and index IV_4 was calculated dividing the number of markers with frequency differences superior to 50% or absent in one of the populations by the total of bands in the population. Finally, the Shannon index (Lewontin 1972) and the percentage of the total diversity estimated based on this index were also considered for the sub- and the entire population.

Table 1. Number of polymorphic and monomorphic RAPD bands and diversity indexes (IV_1 to IV_4) and Shannon index, considering the entire *Eucalyptus* population, and both *Eucalyptus* sub-populations

Total of RAPD markers (polymorphic and monomorphic)											
	Number of Genotypes	Monomorphic bands	Polymorphic bands	Allele frequency differing loci*	Total	IV_1	IV_2	IV_3	IV_4	Shannon Index	Diversity percentage of Shannon index
E. urophylla	44	70	430	-	500	86.0	85.8	95.6	-	100.12	96
E. grandis	40	92	400	-	492	81.3	79.8	88.9	-	88.73	85
Total pop.	84	51	450	163	501	-	89.8	-	32.5	104.52	
Polymorphic markers considering the entire *Eucalyptus* population											
	Number of Genotypes	Monomorphic bands	Polymorphic bands	Allele frequency differing loci	Total	IV_1	IV_2	IV_3	IV_4	Shannon Index	Diversity percentage of Shannon index
E. urophylla	44	19	430	-	449	95.8	95.6	95.6	-	100.12	96
E. grandis	40	41	400	-	441	90.7	88.9	88.9	-	88.73	85
Total pop.	84	0	450	163	450	-	100.0	-	36.2	104.52	
Polymorphic markers considering both *Eucalyptus* sub-populations											
	Number of Genotypes	Monomorphic bands	Polymorphic bands	Allele frequency differing loci	Total	IV_1	IV_2	IV_3	IV_4	Shannon Index	Diversity percentage of Shannon index
E. urophylla	44	0	380	-	380	100	100	100	-	91.97	95
E. grandis	40	0	380	-	380	100	100	100	-	85.54	88
Total pop.	84	0	380	143	380	-	100	-	37.6	96.85	

* loci that presented over 50% variation considering in the relationship of the allele frequencies of *Eucalyptus* populations ($F_{E.\ urophylla}/F_{E.\ grandis}$), or absent in one of the populations.

RESULTS AND DISCUSSION

RAPD profiles resulted in 501 bands (DNA amplified fragments), 450 of which were polymorphic and 51 monomorphic considering both *Eucalyptus* sub-populations (Table 1). One of these bands was exclusive to *E. grandis* and nine others were restricted to *E. urophylla*. When sub-populations were considered individually, 70 and 92 monomorphic bands were observed for *E. urophylla* and *E. grandis*, respectively. Additionally, 163 markers displayed a difference of at least 50% for the allele frequency between species or were absent in one of them.

Successful eucalypt genotype distinction based on RAPD data was observed in the UPGMA dendrogram and the graph dispersion (Figures 1 and 2) in agreement with Keil and Griffin (1994) who had already reported successful eucalypt genotype characterization using RAPDs. It is worth noting that, besides being a simple and rapid technique and providing good polymorphism levels and genome cover, RAPDs are dominant markers and present difficulties of repeatability, consequently providing a lower information content compared to other markers. For example, co-dominant markers such as microsatellites provide better individual discrimination, reliability and a high polymorphism level (Ferreira and Grattapaglia 1995, Glaubitz et al. 2001, Jones et al. 2002), although their development has the disadvantage of being more time and cost-demanding.

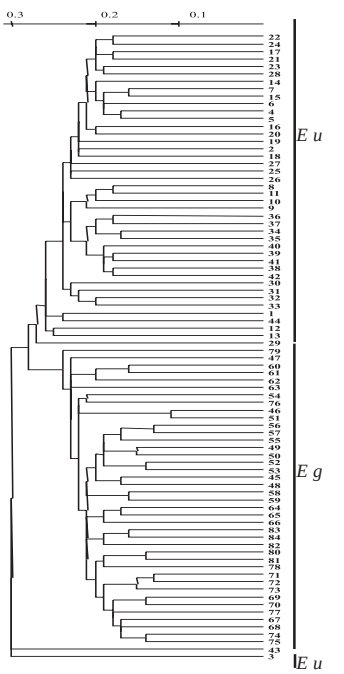

Figure 1. UPGMA dendrogram of 84 *E. grandis* (*E g*) and *E. urophylla* (*E u*) genotypes from genetic distances estimated by the arithmetic complement of Nei and Li index

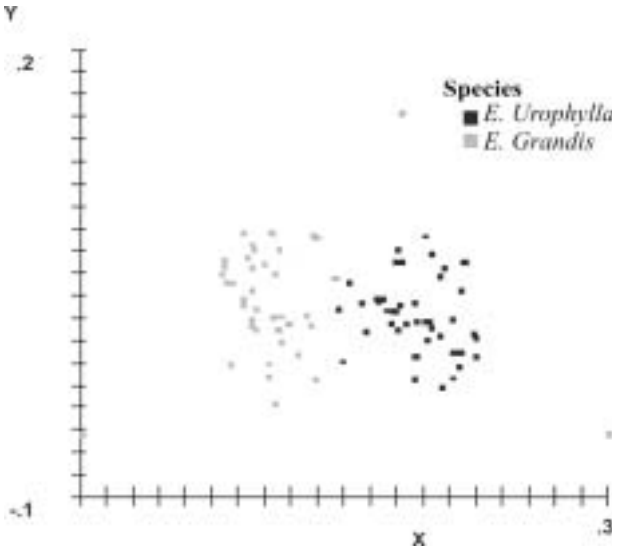

Figure 2. Coordinate plotting associated with the estimated RAPD genetic distances according to *E. urophylla* and *E. grandis* genotypes. X and Y correspond to the coordinates. Correlation between original and estimated distances equal to 0.5984, and distortion equal to 64.7%

Despite the comparison of results from different studies may be difficult because of the type of markers (Powell et al. 1996, Crouch et al. 1999), the technical protocol (Breyne et al. 1999) and the calculated indexes used (Bussell et al. 1999), moderate (Sun et al. 2001, Ipek et al. 2003) and high correlations (Patzak 2001, Raina et al. 2001, Renganayaki et al. 2001, Simioniuc et al. 2002) were observed for the divergence/similarity matrices based on RAPD and other molecular markers. Additionally, RAPD is still more frequently used compared to other markers in studies involving several species (Dias et al. 2004).

In this context, it was hypothesized that the number of sampled genotypes could influence genetic distance estimates. Table 2 shows that, regardless of the RAPD markers' polymorphism, the *E. urophylla* sub-population had a significantly different mean genetic distance from *E. grandis* in an 84-genotype population. However, when the sample size was reduced to 10 genotypes (5 *E. grandis* and 5 *E. urophylla*) the genetic distance was no longer significant between the sub-populations by the t-test at 1% probability. It must be considered that, due to sampling, this result may be valid only for these genotypes in particular. Nevertheless, a lack of correlation was observed between the population size and genetic variability in certain eucalypt species (Rosseto et al. 1999).

The effect of the number of markers was also investigated based on the bootstrap analysis of RAPD band random samples. The analysis of the estimated genetic distances showed that less than 10, 5, and 1% for stress statistics were obtained, respectively, with 71, 338, and 472 bands, considering all 200 samples for each number of markers (Figure 3A). Similarly, correlations of 0.90 and 0.95 were detected with 317 and 393 markers (Figure 3B), respectively, while sum of squares values below 10 were observed for 183 and below 5 for 368 markers (Figure 3C). Interestingly, an analogous number of markers was established in other approaches using a variation coefficient of 5% (Tivang et al. 1994, Fanizza et al. 1999).

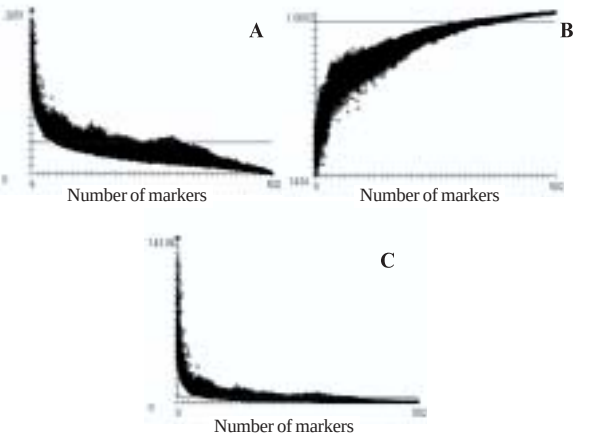

Figure 3. Bootstrap analysis. Plot of all of the 200 bootstrap analyses evaluating the effect of the number of markers on the genetic distance matrices for **A**: stress statistical; **B**: correlation and **C**: sum of squares

The average of 160 RAPD markers used in genetic distance estimation studies (Dias et al. 2004) together with the present results suggest an adjustment in the number of markers used in order to obtain more accurate and affordable estimates. It is also observed that a number between 350 and 400 markers would provide sound genetic distance estimates, even when considering different species and markers. Nevertheless, other features may also affect this aspect, such as the level of primer polymorphism. Fanizza et al. (1999) found a correlation of 0.89 and 0.90 with 462 and 470 bands with primers with more than 6 and less than 5 bands per primer, respectively. Otherwise, the correlation declined with a higher or lower level of polymorphism, although

Table 2. Average genetic distance of *E. grandis* and *E. urophylla* sub-populations

Types of bands	84-genotype population		10-genotype population	
	E. urophylla	*E. grandis*	*E. urophylla*	*E. grandis*
Polymorphic and monomorphic markers[1]	22.57 a	19.18 b	20.37 a	19.57 a
Only polymorphic markers[2]	30.96 a	27.96 b	24.29 a	23.76 a

Considering the total 501 polymorphic and monomorphic RAPD markers[1] and the markers that were polymorphic for the sub-populations only[2]. Numbers followed by the same letter in the same line and population do not differ statistically by the t-test at the 1% probability level

this result might be associated with the number of markers used as well.

Fanizza et al. (1999) also argued that identical profiles did not add further information to the variety discrimination and that their inclusion might affect the computation of the genetic distances. Although reasoning that only polymorphic bands will not alter dendrograms and grouping (data not shown), the use of monomorphic bands would provide more accurate estimates. The importance of this observation is confirmed as estimated genetic distances are used for hybrid performance prediction, supported by reports of possible non-linear relation between divergence/ heterosis (Fabrizius et al. 1998, Sant et al. 1999).

Considering that 393 RAPD markers lead to stress statistics of less than 5% and good values of sum of squares and correlation, this amount was used for Tocher grouping, which displayed consistent clustering with the dendrogram and where *E. urophylla* and *E. grandis* species were separated (Figure 1, Table 3). According to this analysis, 12 to 21 groups were detected, where two major groups containing *E. urophylla* or *E. grandis* genotypes were formed in agreement with the original grouping based on the genetic distances estimated with 501 RADP markers. The other groups comprised a smaller number of individuals, mostly composed of one genotype, but always respecting the species assortment. Though rearrangements were observed even in the Tocher grouping consisting of 12 sets, they were coherent with the species grouping of 12 clusters obtained from the original distances (Table 3). This had been expected since similarity assessment does not obligatorily correspond to morphological resemblance, given that genetic distance based on RAPD is a relative measure and an approach

Table 3. Tocher grouping based on genetic distances estimated from randomly chosen RAPD markers. **A**: Original Tocher grouping based on 501 RAPD markers, **B** through **H**: Tocher groupings based on 393 randomly chosen RAPD markers and presenting equal number of groups

group	genotypes A	group	genotypes B	group	genotypes C	group	genotypes D
I	46 51 76 75 74 55 50 57 56 53 68 67 69 70 77 48 45 71 49 84 60 59 58 60 78 64 66 65 72 73 83 61 81 63 52 82 54 47 79	I	46 51 76 75 74 77 67 68 69 55 56 53 57 70 50 49 45 48 52 71 84 59 60 80 78 61 58 72 64 81 65 66 73 83 54 82 62 63 47 79	I	46 51 76 75 77 74 67 53 55 50 49 57 56 68 69 59 84 70 60 80 71 48 81 52 61 72 64 65 78 73 83 66 58 45 63 54 82 62 47 79	I	46 51 76 75 74 55 67 77 53 56 57 68 69 70 49 50 71 48 52 45 60 61 64 80 59 84 73 72 78 65 81 66 83 58 82 63 62 47 54 79
II	36 37 35 42 39 38 41 40 34 21 33 8 9 11 15 24 7 17 23 22 5 14 4 20 10 28 6 19	II	36 37 35 42 39 38 40 41 9 21 17 8 34 11 24 15 22 23 5 7 20 33 31 6 14 16 28	II	36 37 35 42 39 38 40 41 9 21 8 11 17 15 24 22 7 5 33 23 6 10 28 20 14 3 19 16 4 2 44	II	36 37 35 42 39 17 41 21 38 9 40 8 11 24 15 7 23 22 34 33 31 5 14 16 20 10
III	1 44	III	2 4 1 3	III	26 32 30	III	2 4 1 6 44 28
IV	2 3	IV	12 19 25	IV	31 43	IV	12 19
V	25 26 27 16	V	26 32 30	V	34	V	26 32
VI	31 32	VI	27	VI	18	VI	27 29
VII	30	VII	29	VII	1	VII	30
VIII	18	VIII	18	VIII	27	VIII	25
IX	43	IX	13	IX	12	IX	3
X	29	X	10	X	25	X	18
XI	13	XI	44	XI	13	XI	13
XII	12	XII	43	XII	29	XIII	43

group	genotypes E	group	genotypes F	group	genotypes G	group	genotypes H
I	46 51 76 75 74 77 55 56 57 68 53 52 50 49 69 70 67 48 71 84 60 61 64 59 80 78 81 66 58 73 72 65 45 82 83 63 62 47 54 79	I	46 51 76 75 74 55 50 33 69 56 57 49 68 67 70 84 59 48 45 77 52 80 71 60 61 58 81 72 73 64 66 83 82 78 65 63 47 62 54 79	I	46 51 76 75 74 77 55 69 53 50 49 56 57 70 68 67 84 71 52 48 45 60 61 80 59 66 64 78 72 73 65 81 83 58 54 82 62 63 47	I	46 51 76 75 77 74 55 56 68 69 57 70 53 67 50 49 84 48 71 52 60 59 45 80 58 61 78 72 64 66 73 83 65 62 81 63 54 47 82 79
II	36 37 35 42 39 38 34 40 41 21 8 9 11 17 7 15 24 22 33 23 5 20 31 19 6 4 14 16 28	II	36 37 35 42 39 40 38 34 41 8 11 9 21 17 24 15 33 7 5 22 23 31 16 6 14 10 20 28 19 2	II	36 37 35 42 39 38 40 34 41 21 9 8 11 17 24 15 33 22 7 23 20 5 6 16 14 19 28 10	II	36 37 35 42 39 40 38 41 17 21 9 15 24 33 8 11 7 22 5 23 20 4 10 28 6 14 16 19 1 2 3
III	2 3	III	1 4 44	III	2 4 1 44	III	26 32
IV	1 44 27	IV	26 32 30	IV	26 32	IV	27 44
V	26 32	V	25	V	31 43	V	31 43
VI	12 25	VI	27	VI	12 25	VI	34
VII	10	VII	18	VII	27 29	VII	18
VIII	13	VIII	29	VIII	30	VIII	30
IX	18	IX	3	IX	3	IX	25
X	29	X	13	X	18	X	13
XI	30	XI	43	XI	13	XI	12
XII	43	XII	12	XII	79	XII	29

where geographically distant individuals may display similar estimates (Sale et al. 1996, Skabo et al. 1998, Rossetto et al. 1999).

Assuming that marker frequency and polymorphism in the sub-populations and in the entire population enclose information on genetic population diversity, some indexes were used to evaluate *E. grandis* and *E. urophylla* diversity. It is observed that although similar to *E. grandis*, *E. urophylla* had higher diversity values (Table 1). Diversity indexes IV_1, IV_2, IV_3, and IV_4 increased as the monomorphic bands were discarded for the whole population and subsequently the sub-populations. This may represent some inconvenience, as monomorphic bands are not always considered in this approach. This may be circumvented with the Shannon index, which in spite of decreasing with the same approach remained almost unaltered in its relative proportion of the estimated diversity within the sub-populations. In addition to the number of markers, the Shannon index considered the allele frequency, and when the entire population was used as reference, the sub-population proportion of diversity was less affected compared to the other indexes.

On the other hand, the IV_4 diversity index (Table 1) indicated that there are markers either restricted to one sub-population or with a great frequency difference between each other. Similar information was found for the subspecies *E. globulus* where 63.8% of the scored RAPD bands differed significantly in frequency between sub-species cores and sites (Nesbitt et al. 1995).

The present results highlight the importance of the number of markers and its polymorphism for saving time and resources in studies involving genetic distance estimates. This is especially true since this approach is based on experimental data of highly polymorphic species such as eucalypt, besides reinforcing the usefulness of RAPD markers for eucalypt breeding programs.

ACKNOWLEDGMENTS

The authors wish to thank Cenibra S/A for supporting this research conducted at the Laboratório de Patologia Florestal e Genética da Interação Planta-Patógeno/BIOAGRO/UFV and for granting it financial support.

Tamanho amostral para número de marcadores RAPD para estimar diversidade genética em *Eucalyptus*

RESUMO - *Um total de 501 marcadores RAPD foi gerado para uma população de 84 genótipos de* Eucalyptus*, dos quais 89,8% foram polimórficos e 10,2% monomórficos. Agrupamentos coerentes baseados nas estimativas de distância genética foram obtidos para as espécies. Análises de re-amostragem baseadas no número de marcadores mostraram que 393 ou mais marcadores resultaram em valores de estresse menores que 5%, menores que 5 para soma de quadrados e maiores que 95% para correlação. A diversidade estimada com base no índice de Shannon foi de 95-96% para* E. urophylla *e 85-88% para* E. grandis. *Considerando uma população com 84 genótipos, cada espécie apresentou diferenças significativas para as distâncias genéticas estimadas. Por outro lado, quando a população foi reduzida para 10 genótipos, a diferença não foi detectada pelo teste t a 1% de probabilidade.*

Palavras-chave: Algoritmo de Tocher, coeficiente de variação, estresse, RAPD, diversidade genética.

REFERENCES

Barbosa AMM, Geraldi IO, Benchimol LL, Garcia AAF, Souza Jr. CL and Souza AP (2003) Relationship of intra- and interpopulation tropical maize single cross hybrid performance and genetic distances computed from AFLP and SSR markers. **Euphytica 130**: 87-99.

Baril CP, Verhaegen D, Vigneron P, Bouvet JM and Kremer A (1997) Structure of the specific combining ability between two species of *Eucalyptus*. I. RAPD data. **Theoretical and Applied Genetics 94**: 796-803.

Breyne P, Roumbaut D, Van Gysel A, Van Montagu M and Gerats T (1999) AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. **Molecular and General Genetics 261**: 627-634.

Bussell JD (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petreae* (Lobeliaceae). **Molecular Ecology 8**: 775-789.

Crouch JH, Crouch HK and Constandt H (1999) Comparison of PCR-based molecular analyses of *Musa* breeding populations. **Molecular Breeding 5**: 233-244.

Cruz CD (2001) **Programa Genes: aplicativo computacional em genética e estatística**. Editora UFV, Viçosa, 648p.

Cruz CD (2003) **Programa Gqmol**. http://www.ufv.br.dbg.gqmol/gqmol.htm

Cruz CD and Viana JMS (1994) A methodology of genetic divergence analysis based on sample unit projection on two-dimensional space. **Brazilian Journal of Genetics 17**: 69-73.

Dias LAS, Picoli EAT, Rocha BR and Alfenas AC (2004) *A priori* choice of hybrid parents in plants. **Genetics and Molecular Research 3**: 356-368.

Fabrizius MA, Busch RH, Khan K and Huckle L (1998) Genetic diversity and heterosis of spring wheat crosses. **Crop Science 38**:1108-1112.

Fanizza G, Colonna G, Resta P and Ferrara G (1999) The effect of the RAPD markers on the evaluation of genotypic distances in *Vitis vinifera*. **Euphytica 107**: 45-50.

Ferreira ME and Grattapaglia D (1995) **Introdução ao uso de marcadores moleculares em análise genética**. 2nd ed., EMBRAPA-CENARGEN, Brasília, 220p.

Gaiotto FA, Bramucci M and Grattapaglia D (1997) Estimation of outcrossing rates in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. **Theoretical and Applied Genetics 95**: 842-849.

Glaubitz JC, Emebiri LC and Moran GF (2001) Dinucleotide microsatellite from *Eucalyptus sieberi*: inheritance, diversity, and improved scoring of single-base differences. **Genome 44**: 1041-1045.

Ipek M, Ipek A and Simon PW (2003) Comparison of AFLPs, RAPD markers, and isozymes for diversity assessment of garlic and detection of putative duplicates in germplasm collections. **Journal of the American Society for Horticultural Sciences 128**: 246-252.

Jones RC, Steane DA, Potts BM and Vaillancourt RE (2002) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. **Canadian Journal of Forestry Research 32**: 59-66.

Joshi SP, Ranjekar PK and Gupta VS (1999) Molecular markers in plant genome analysis. **Current Science 77**: 230-240.

Junghans DT, Alfenas AC, Brommonschenkel SH, Valle LAC, Oda S and Mello EJ (2003) Fine genetic mapping of the rust (*Puccinia psidii*) resistance gene *Ppr1* in *Eucalyptus grandis*. **Theoretical and Applied Genetics 108**: 175-180.

Keil M and Griffin AR (1994) Use of random amplified polymorphic DNA (RAPD markers in the discrimination and verification of genotypes in *Eucalyptus*. **Theoretical and Applied Genetics 89**: 442-450.

Kumar LS (1999) DNA markers in plant improvement: An overview. **Biotechnology Advances 17**: 143-182.

Kwon SJ, Ha WG, Hwang HG, Yang SJ, Choi HC, Moon HP and Ahn SN (2002) Relationship between heterosis and genetic divergence in 'Tongil'-type rice. **Plant Breeding 121**: 487-492.

Lewontin RC (1972) The apportionment of human diversity. In: Dobzhansky T, Hecht MK and Steere WC (eds.) **Evolutionary Biology**. Volume 6, Meredith Corporation, New York, p. 381-389.

Moser H and Lee M (1994) RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. **Theoretical and Applied Genetics 87**: 947-956.

Nesbitt KA, Potts BM, Vaillancourt RE, West AK and Reid JB (1995) Partioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). **Heredity 74**: 628-637.

Patzak J (2001) Comparison of RAPD, STS, ISSR e AFLP molecular methods used for assessment of genetic diversity in hop (*Humulus lupulus* L.). **Euphytica 121**: 9-18.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, RAPD and SSR (microsatellite) markers for germplasm analysis. **Molecular Breeding 2**: 225-238.

Raina SN, Rani V, Kojima T, Ogihara Y, Singh KP and Devarumath RM (2001) RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. **Genome 44**: 763-772.

Renganayaki K, Read JC and Fritz AK (2001) Genetic diversity among Texas bluegrass genotypes (*Poa arachnifera* Torr.) revealed by AFLP and RAPD markers. **Theoretical and Applied Genetics 102**: 1037-1045.

Rosseto M, Jeziersky G, Hopper SD and Dixon KW (1999) Conservation genetics and clonality in two critically endangered eucalypts from the highly endemic south-western Australian flora. **Biological Conservation 88**: 321-331.

Sale MM, Potts BM, West AK and Reid JB (1996) Molecular differentiation within and between Eucalyptus risdonii, E. amygdalina and their hybrids using RAPD markers. **Australian Journal of Botany 44**: 559-569.

Sant VJ, Patankar AG, Sarode ND, Mhase LB, Sainani MN, Deshmukh RB, Ranjekar PK and Gupta VS (1999) Potential of DNA markers in detecting divergence and in analyzing heterosis in Indian elite chickpea cultivars. **Theoretical and Applied Genetics 98**: 1217-1225.

Simioniuc D, Uptmoor R, Friedt W and Ordon F (2002) Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. **Plant Breeding 121**: 429-435.

Skabo S, Vaillancourt RE and Potts BM (1998) Fine structure of *Eucalyptus globulus* ssp. g*lobulus* forest revealed by RAPDs. **Australian Journal of Botany 46**: 583-594.

Sun GL, William M, Liu J, Kasha KL and Pauls KP (2001) Microsatellite and RAPD polymorphisms in Ontario corn hybrids are related to the commercial sources and maturity ratings. **Molecular Breeding 7**: 13-24.

Tivang JG, Nienhuis J and Smith OS (1995) Estimation of sampling variance of molecular marker data using the bootstrap procedure. **Theoretical and Applied Genetics 89**: 259-264.

Van de Peer Y (1997) **Treecon for Windows, versión 1.3b**.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphisms, amplified by arbitrary primers are useful as genetic markers. **Nucleic Acid Research 18**: 6531-6535.