Genetic characterization of papaya plants (*Carica papaya* L.) derived from the first backcross generation

H.C.C. Ramos[1], M.G. Pereira[1], F.F. Silva[2], L.S.A. Gonçalves[1], F.O. Pinto[1], G.A. de Souza Filho[1] and T.S.N. Pereira[1]

[1]Laboratório de Melhoramento Genético Vegetal,
Centro de Ciências e Tecnologias Agropecuárias,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil
[2]Instituto Federal de Educação, Ciência e Tecnologia do Tocantins,
Campus Araguatins, Araguatins, TO, Brasil

Corresponding author: H.C.C. Ramos
E-mail: helainecr@uenf.br

ABSTRACT. The limited number of papaya varieties available reflects the narrow genetic base of this species. The use of backcrossing as a breeding strategy can promote increases in variability, besides allowing targeted improvements. Procedures that combine the use of molecular markers and backcrossing permit a reduction of the time required for introgression of genes of interest and appropriate recovery of the recurrent genome. We used microsatellite markers to characterize the effect of first-generation backcrosses of three papaya progeny, by monitoring the level of homozygosity and the parental genomic ratio. The homozygosity level in the population ranged from 74 to 94%, with a mean of 85% for the three progenies (52-08, 52-29 and 52-34). The high level of inbreeding found among these genotypes increases the expectation of finding more than 95% fixed loci in the next generation of self-fertilization of superior genotypes. The mean proportion of the recurrent parent genome found in first-generation backcross progeny was 50.1%; 52-34 had a larger

genomic region in common with the recurrent genitor and the lowest level of homozygosity. The progeny 52-08 was genetically closest to the donor genitor, and it also had the highest level of homozygosity. We found that linking conventional procedures and molecular markers contributed to an increase in the efficiency of the breeding program.

Key words: Backcross; Microsatellite markers; Parental genome; Inbreeding; Genetic variability

INTRODUCTION

The papaya plant (*Carica papaya* L.), a typically tropical crop, has been cultivated in over 50 countries around the world (FAO, 2009), for fresh fruit consumption and also for exploiting its latex, with the extraction of the proteolytic enzyme papain, which is widely used in pharmaceutical and food industries. As with other fruit families, such as Myrtaceae, Malpighiaceae, Musaceae, and Passifloraceae, the genus *Carica* is the only one from the family Caricaceae with economic importance and commercial exploitation. This species is diploid with nine chromosome pairs (Damasceno Junior et al., 2009) and a relatively small genome of 372 Mb (Arumuganathan and Earle, 1991), formed by a pair of incipient and homomorphic sexual chromosomes, with differences only at the molecular level (Liu et al., 2004; Yu et al., 2008). The plants have three basic sexual forms (female, male and hermaphrodite), characterizing this species as trioecious (Storey, 1953).

Despite the increasing importance of the papaya crop in Brazil, because of the commercialized volume or the area where it is established, papaya is supported on a narrow genetic base, with few available cultivars for commercial planting. It is evident that there is a need to direct efforts to breeding programs that promote the expansion of the genetic base by developing varieties that have desirable agronomic characteristics and, mainly, to meet market demand (Silva et al., 2007a). In this context, genetic improvement mediated by backcrossing programs allows the hybridization of distinct genetic material, establishing an important strategy of breeding for the introgression of genes of interest in crop varieties, allowing the expansion of genetic diversity in subsequent generations (Siqueira et al., 1988), and also making targeted and predictable improvements.

The backcross has gained extensive applications in recent years due to the increasing efforts to transfer transgenes to commercial elite genotypes (Lewis and Kernodle, 2009). Although it is considered to be a well-known method for introgression or substitution of a target allele, the large number of generations required to recover the recurrent parent genome and the presence of portions of the donor parent genome linked to the favorable allele transferred (dragged), there are two problems inherent to the backcross (Benchimol et al., 2005). On the one hand, a well-established alternative in plant breeding is the procedure of molecular marker-assisted backcross (Tanksley, 1983; Young and Tanksley, 1989; Visscher et al., 1996; Schmierer et al., 2004), which helps in classical procedures, accelerating the recovery of recurrent parent genome and reducing the number of generations required for introgression of the gene of interest (Hospital and Charcosset, 1997). This strategy can also increase genetic gain and economic efficiency in relation to classical procedures (Kuchel et al., 2005).

Among the different classes of markers available for molecular analysis, microsatellites (single sequence repeat - SSR) have been widely used for marker-assisted selection in backcross programs (Benchimol et al., 2005; Xi et al., 2008). In the last years, SSR has

become one of the most popular molecular markers due to the massive amount of sequences available in databases, reflecting the progress of genome research (Leal et al., 2010). In plant breeding, as well as genetic analysis, this marker has had a variety of applications due to its multi-allelic nature, reproducibility, high information content, codominant inheritance, high abundance, and extensive coverage of the genome (Gupta and Varshney, 2000) but distributed in a non-random way (Wang et al., 2008). In the papaya genome, microsatellites are the more abundant type of tandem repetition, with a density of one every 0.7 kb; however, it represents only 0.19% of the entire genome of this species (Moore and Ming, 2008; Wang et al., 2008).

In our study, microsatellite markers were used to characterize the effect of the first-generation backcross in three families of papaya, aiming to monitor the level of homozygosity, as well as the genomic proportion of parental genotypes in the offspring, in addition to identifying favorable segregating genotypes for generation advance.

MATERIAL AND METHODS

Plant material

We evaluated 43 genotypes of the S_2 generation, where 17, 14 and 12 genotypes were respectively derived from the 52-34, 52-29 and 52-08 families, besides the genitors Cariflora (recurrent parent) and SS783 (donor parent of hermaphroditism) of a backcrossing program (Silva et al., 2007b). The evaluated segregating families came from a single plant (52) selected in the first backcross generation (BC_1) between an F_1 plant and the recurrent genitor Cariflora (Figure 1). The same was submitted to two generations of self-fertilization and selection, resulting in the S_2 progeny analyzed in this study. These progeny were selected for favorable phenotypic attributes, having the red pulp in common, a feature much appreciated by the fresh fruit market and, therefore, intended to be kept in selected genotypes for generation advance in order to obtain genotypes with such attributes, and other ones equally important for the papaya breeding program.
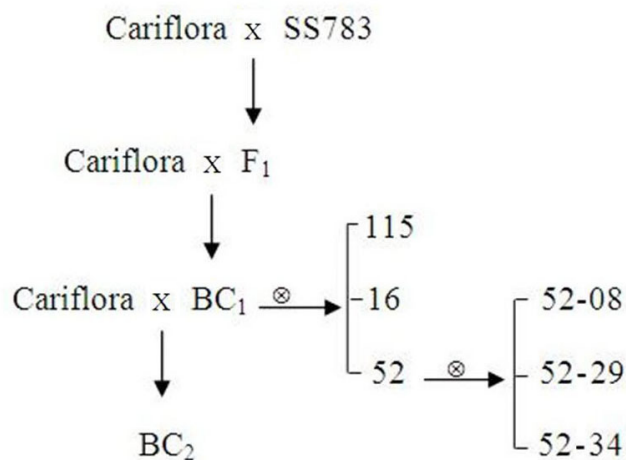

Figure 1. Breeding procedures of the improvement used for obtaining the three progeny evaluated in this study.

The recurrent genitor (Cariflora) is a dioecious selection and thus has a segregating behavior. It has yellow flesh, moderate fruit firmness and good flavor. The fruits weigh around 1.67 kg with a rounded shape and large internal cavity. On the other hand, the donor genitor (SS783) is an elite homozygous cultivar, belonging to the 'Solo' group, showing fruits with a pear shape, red flesh, good quality, and average weight of 0.52 kg. The cross between these genotypes results in vigorous and productive hybrids (Marin et al., 2006).

Extraction of genomic DNA

The genetic materials analyzed were obtained from an experiment established at the Caliman Agrícola Company S.A., located in Linhares, ES, Brazil.

Total cellular DNA was extracted from young leaves of the strains using the CTAB method (Doyle and Doyle, 1990) with some modifications suggested by Daher et al. (2002). The genetic material from the donor genitor (SS783) and the recurrent genitor (Cariflora) was collected in bulk in attempt to sample the largest possible number of allelic forms present in each of the genitors for each locus analyzed. After DNA extraction, DNA was quantified by 0.8% agarose gel electrophoresis and diluted to a concentration of 10 ng/μL, using the High DNA Mass Ladder marker (Invitrogen, USA). The gel was stained in an ethidium bromide solution and the image captured by the Eagle-Eye II photo-documentation system.

SSR analysis

For the amplification reactions with the SSR marker, we selected 32 primers used by Santos et al. (2003), and 45 clones containing microsatellite sequences developed by Pérez et al. (2006), available at GenBank (http://www.ncbi.nlm.nih.gov). The clones containing the DNA sequence that flanks the microsatellite region were used to design the primer pairs (forward and reverse), using the programs Expression version 1.0.0.0 and Oligo version 6.68 (Genamics). To obtain greater specificity of the primers in the amplification, we established criteria in the design of the primers: minimum size of the primers (>14 bp), and melting temperatures between 35 and 45°C. Besides these criteria, we also avoided the nucleotide sequences that had G and C at their ends (mainly the 3'OH) because of the strength of its link with the corresponding bases to avoid nonspecific amplification.

The DNA of the genitors (Cariflora and SS783) was initially used to perform the optimization of the reaction and the screening of 77 synthesized primers. The amplification reactions were performed in a final volume of 20 μL, containing 10 ng DNA, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl$_2$, 100 μM dNTPs, 2 μM of each primer, and 1 U Taq DNA polymerase. The amplifications for the optimization of the annealing temperature were performed in an Eppendorf thermal cycler, according to the following program: 4 min of denaturation at 94°C, 32 amplification cycles (94°C at 30 s, 53°C at 1 min, 72°C at 1 min), followed by a final extension of 7 min at 72°C. The amplification products were separated on an 8% polyacrylamide gel without denaturing agent, stained with ethidium bromide and photo-documented using the Eagle Eye II system. After the optimization of the reaction conditions, 16 primers were selected, because they had a greater complementarity, reproducibility and presence of polymorphism among the genitors (Table 1).

Data analysis

The data obtained from the amplification of microsatellite primers were converted into code numbers for each allele per locus. Thus, if the locus has three alleles, it has the representation 11, 22 and 33 to homozygous forms (A_1A_1, A_2A_2 and A_3A_3) and 12, 13 and 23 for the heterozygous (A_1A_2, A_1A_3 and A_2A_3). Subsequently, the genetic distance was calculated between genotypes using the weighted index (Cruz, 2008). The cluster analysis between genotypes was performed by the hierarchical unweighted pair-group method with arithmetic means (UPGMA) analysis, using the GENES (Cruz, 2008) and R programs (www.r-project.org).

The following parameters were estimated: observed number of alleles, effective number of alleles and Shannon index. The analysis of the level of homozygosity of the genotypes evaluated proceeded from the calculation of the ratio between the number of homozygous loci and the total number of loci analyzed per genotype.

RESULTS AND DISCUSSION

Of the 77 microsatellite loci analyzed in this study, only 16 showed polymorphism between the parental genotypes and were therefore selected for the amplification reactions in the population. A total of 39 alleles were found among the parental genotypes (Cariflora and SS783) with a minimum of two and a maximum of three alleles, giving an average of 2.4 alleles/locus. Considering only the S_2 population, the occurrence of one to three alleles was observed, with an average of 1.7 alleles/locus (Table 1). This proportion of alleles is in line with what is expected for this backcross progeny, considering the nature of the segregating recurrent progenitor used in this study. The average value for the observed allele numbers and the effective number of alleles was 1.87 and 1.6, respectively. The Shannon information index ranged from 0 to 0.99, with an average of 0.46. This result suggests that there is low variability in the population, quite expected due to the selection process performed in the population, since the three progeny were from a single plant from the BC_1 generation.

The level of homozygosity observed in the 43 genotypes derived from backcrossing ranged from 74 to 94% with an average of the three progeny of 85% (Figure 2). Considering the three families separately, it can be seen that a higher level of homozygosity was found among the genotypes belonging to the 52-08 family, where the values ranged from 82 to 94% with an average of 88.5%. The other treatments showed an average of 83% homozygosity, with an individual variation between 75 and 94% in the 52-34 family, and 74 to 94% among the genotypes belonging to the 52-29 family.

For plants, which are in the first backcross generation and second generation of self-fertilization, it is expected that an average of 87.5% of their loci are homozygous, according to Silva et al. (2007a) when monitoring genetic variability in the progenitor of 'Formosa' papaya from the 'UENF/CALIMAN 01' hybrid using random amplified polymorphic DNA (RAPD) markers. Considering that the progeny in this study was derived from a single plant of the first generation of backcrossing, this sample is not representative of the entire BC_1 population. Even so, the level of homozygosity found is quite satisfactory, since it considerably approaches the expected value for the population.

Table 1. Microsatellite primers used for molecular analysis of 49 genotypes and the number of alleles found for each primer.

Locus	Motif	Primer sequence	T_A (°C)	Allele number
mCpCIR01	$CT_{(18)} GA_{(3)}$	*F: ATCGTCTCCTTTTTCTGGTT*	57	2
		R: TCTGCCTCCCAATACACTAAT		
mCpCIR02	$TC_{(24)}$	*F: AGCCACAACCTACGGGAAAT*	57	3
		R: AGTAACGGAGGAAAATGAGT		
mCpCIR05	$TC_{(18)}$	*F: ATCGTCTCCTTTTTCTGGTT*	57	2
		R: TTCTGCCTCCCAATACACTA		
mCpCIR08	$CT_{(20)} AC_{(5)}$	*F: ACCCACCAGCAATCTCCAT*	59	3
		R: AGCAAACCACTCACTCTCATA		
mCpCIR09	$CT_{(9)}$	*F: TGACGATAAAACCCTAACGA*	59	2
		R: TAAGAAACAGCGAAACCCTA		
mCpCIR16	$CT_{(9)}$	*F: TACACTGCCTAACACCCATT*	59	3
		R: AACCAACCATAACTGCCTTT		
mCpCIR17	$GA_{(14)}$	*F: ACAAACAAGTCCCCAAATCT*	59	2
		R: TACACTGCCTAACACCCATT		
mCpCIR28	$TC_{(8)}$	*F: ATCAAGGAAGTGCAAATTT*	56	2
		R: ATGAGCCAATGAGAAGAGGGA		
mCpCIR35	$TC_{(20)}$	*F: ACATACAAAACACTTACCACCA*	56	3
		R: TCAGACATACTGCATCTCAA		
mCpCIR39	$CT_{(10)}$	*F: ATAGCAAACAGAAAAACCCA*	59	2
		R: ATAGAAAGAGAAAGCGA		
mCpCIR40	$TC_{(13)} TC_{(21)}$	*F: TCGGTTCTCAGGTTTCTTCTAA*	57	2
		R: ACAATCACAGGCACACAT		
mCpCIR45	$GA_{(14)}$	*F: AAAAGGACGAAAAGGAGACT*	56	3
		R: TTTGAACTACCTACACGAACT		
S285	$GAT_{(3)}$	*F: AATGTGTGAGAATAGGTT*	59	3
		R: AATCTATCCTCCTCATGTA		
S414	$AC_{(7)}$	*F: ATTCTTAGCCAGATGATGT*	59	2
		R: ATTGCATGTACACATACCGT		
S422	$GAT_{(8)}$	*F: ACGCATCACACGTATATCTA*	56	3
		R: ATAACCTCGCTACATCCTCT		
S552	$GAT_{(4)}$	*F: AACAAGTGGAACTCCTATA*	52	2
		R: CAATGGAACTTCTGCTACTA		

T_A = annealing temperature.
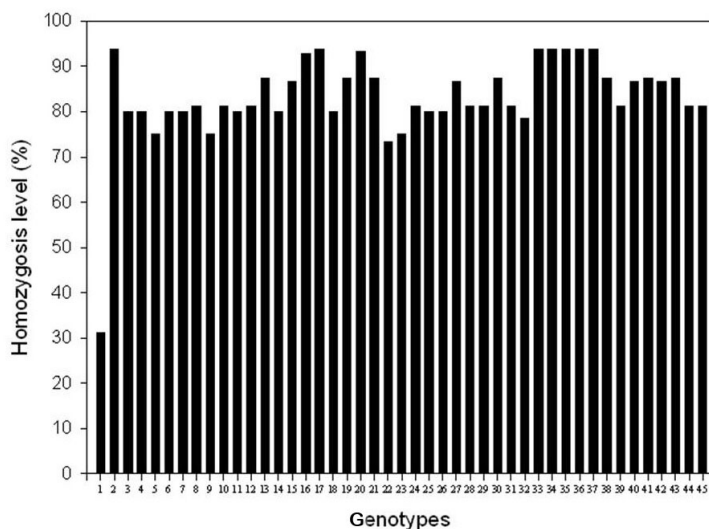
Figure 2. Level of parental homozygosis of Cariflora (1) and SS783 (2), and of the 52-34 (3 to 19), 52-29 (20 to 33), and 52-08 (34 to 45) populations.

From this result, it can be inferred that the number of SSR loci analyzed was sufficient to effectively determine the degree of homozygosity of the progeny evaluated. Furthermore, we observed in the population genotypes with up to 94% of the fixed loci, creating an expectation of obtaining in the next generation of self-fertilizing plants with a high degree of stability and a level of homozygosity greater than or equal to 95%, which is quite satisfactory in a breeding program.

In a BC_1 population from a crossing between two contrasting genotypes, it is expected that an average of 75% of the progeny genome is similar to the recurrent genitor (Collard et al., 2005). However, when analyzing this ratio considering all amplified alleles (Figure 3A), we found an average ratio of 50.1% of the recurrent genitor genome (Cariflora) in the progeny evaluated, a value lower than expected for the BC_1 population.
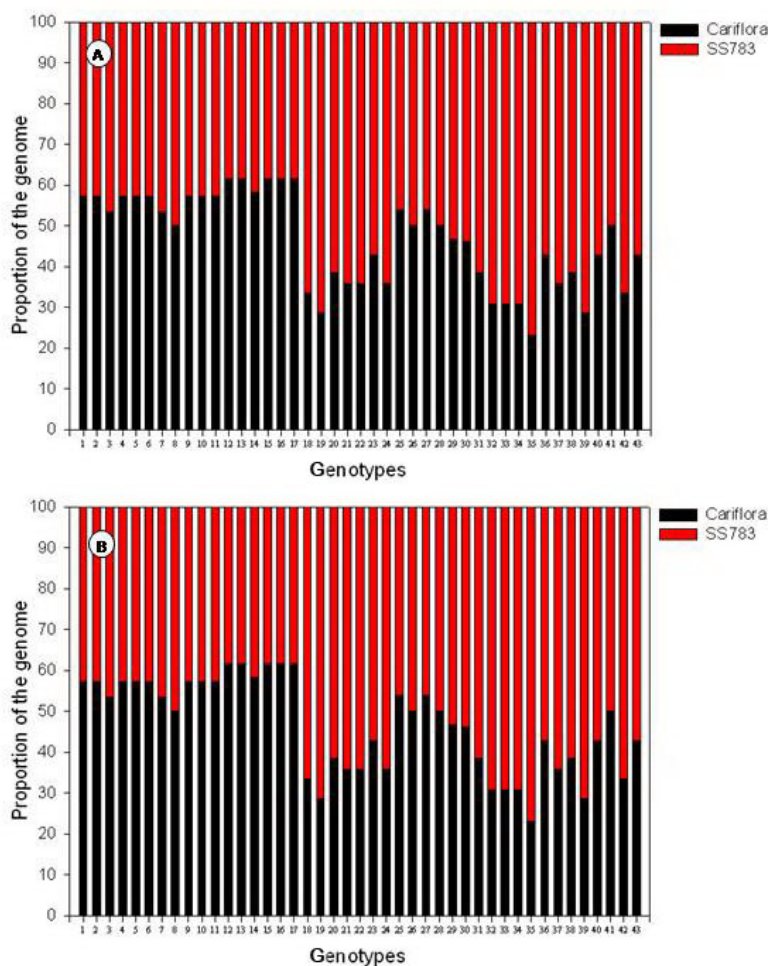

Figure 3. Genomic proportion of 52-34 (1 to 17), 52-29 (18 to 31) and 52-08 (32 to 43) populations. **A.** Based on all amplified alleles. **B.** Based on the homozygous loci.

Among the three families evaluated, the one that had the greatest genome proportion of the recurrent genitor was the 52-34 family with an average of 58%, with individual values ranging from 50 to 62%. The second highest frequency was found in the 52-29 family, with an average of 50% and a ratio by genotype ranging from 41 to 60%. In the 52-08 family, the average was 41.4%, where 12 genotypes showed only a higher proportion of donor genome. The parental genome proportion in the progeny was also analyzed based on the homozygous alleles (Figure 3B), considering only the alleles able to characterize each genitor. Note that the hierarchical order in relation to the average proportion of the Cariflora genome remains the same as the one found for the analysis of the total alleles; however, on average, only the 52-34 family showed a higher genome proportion of the recurrent genitor in relation to the donor genome.

Although an average of 75% of the recurrent genome is expected in the progeny of the first backcross, it must be considered that within the BC_1 population this proportion may vary from 50 to 100%. Thus, it can be inferred that the $52BC_1$ plant that originated the progeny studied is among the genotypes that have a lower proportion of recurrent genome. Moreover, it is important to note that only one plant of the first backcross was sampled in this study ($52BC_1$), and therefore, this becomes a given point.

Silva et al. (2007b), working in the same breeding program, when evaluating the BC_2 generation using RAPD markers in order to select superior genotypes for generation advance, did not observe the expected proportion (75.7% instead of 87.5%). A possible explanation for the low genomic proportion of the recurrent genitor is the effect of the phenotypic selection made, favoring other attributes of the donor genitor, besides the transferred character (hermaphroditism). This somewhat complicates the recovery of the recurrent genome in backcross cycles, tending to divert this recovery in favor of the donor genitor (Davies et al., 2006; Oliveira et al., 2008).

Through cluster analysis, considering a genetic distance value of 0.5, the formation of a large group can be noted, which includes all the progeny evaluated and the parental donor (SS783), and a single group formed only by the recurrent genotype (Cariflora). Subdividing the larger group (considering a genetic distance of 0.28) resulted in the formation of four smaller groups: groups I and III were formed by the 52-34 and 52-08 families, respectively; the parental donor clustered among the genotypes of the last family; and groups II and IV were formed by genotypes of the 52-29 family (Figure 4).

Although the four subgroups are almost equidistant, the 52-34 family was closer to the parental donor. This profile grouping confirms the results found in the analysis of the genomic proportion of parents in the progeny evaluated.

Based on these results, it can be inferred that among the families evaluated, 52-08 is genetically the closest to the donor genitor, and it also has the highest level of homozygosity. On the other hand, despite having a lower level of homozygosity, the 52-34 family combines a higher proportion of genomic regions common for the recurrent genitor, where this family shows results that are more in line with that expected in this backcross program. Considering that the population studied experienced only one cycle of backcrossing, the occurrence of genotypes with genomic proportion of Cariflora above 60% can be considered to be a satisfactory result. Based on this principle, about 100% of the genotypes that fit this profile would be selected within the 52-34 family, emphasizing that it would be the most promising one for the purpose of recovering recurrent genitor genome and to obtain genotypes genetically closer to Cariflora genitor.
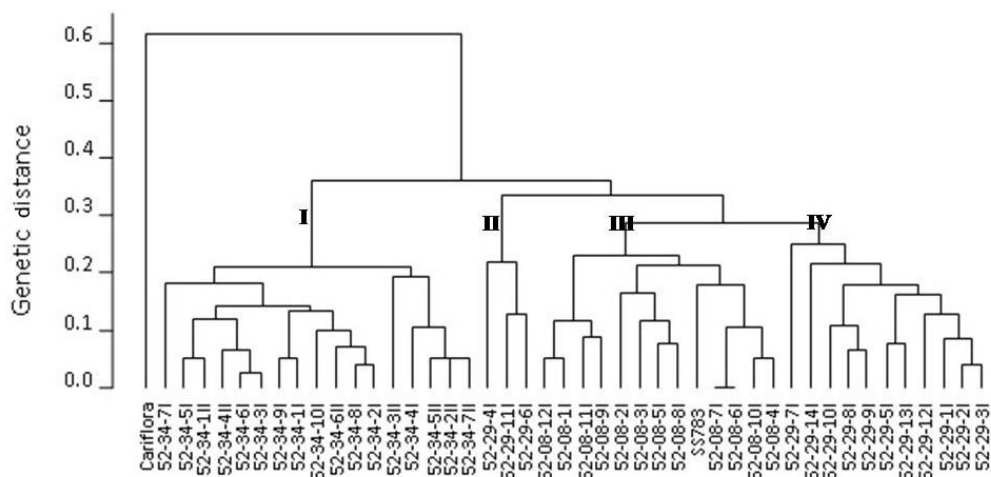
Figure 4. Dendrogram of genetic dissimilarity among 43 genotypes belonging to three S_2 families and parental genotypes of Cariflora and SS783, obtained by UPGMA (Cophenetic correlation = 0.86).

These results also indicate that the selection of plants of the progeny evaluated for the next backcross generation would cause a delay in the improvement program, increasing the number of generations needed for an adequate recovery of the recurrent genome.

Although we found a low proportion in the genome of the recurrent genitor in the population studied, it cannot be considered a problem, since the two genotypes used in the initial crossing (Cariflora and SS783) have desirable agronomic characteristics. The parental donor is an elite genotype and it therefore has great agronomic potential, eliminating any concern from the improvement standpoint on their genomic proportion in the progeny. Furthermore, the recurrent genitor has a good general combining ability when crossed with genotypes from the 'Solo' group, for fruit yield (Marin et al., 2006), and possible resistance to the mosaic virus (Conover et al., 1986).

Moreover, as the transfer of hermaphroditism by backcrossing is not the only priority in this breeding program, i.e., it also seeks to select good segregation with a high level of inbreeding to be available in future stages as varieties or hybrid genitors with high degree of stability, the parental genomic proportion in these progeny becomes less important. However, knowledge of the genomic proportions helps when making a decision for the next steps in backcrossing.

Practical studies aiming to compare the strategies of marker-assisted backcrossing and based on conventional procedures have been performed for some crops (Knapp, 1997; Hospital et al., 1997; Davies et al., 2006; Oliveira et al., 2008). The results indicate that there is a statistically significant difference between the two strategies, with a 14.5% increase in efficiency when the microsatellite markers were used to monitor the selection. It has been suggested that the use of markers with a distance of 10 to 20 cM provides an advantage of one to two generations of backcrossing in relation to random or phenotypic selection (Visscher et al., 1996).

Thus, the present results agree with the principle that molecular markers are important

tools to guarantee the adequate monitoring of the recurrent parent genome (Faleiro et al., 2004; Silva et al., 2007b; Xi et al., 2008; Oliveira et al., 2008) and the level of inbreeding (Oliveira et al., 2010), suggesting their effective application in the papaya improvement program. Therein, the combination of classical procedures with molecular markers emerges as an advantageous strategy, which could provide greater efficiency in the process of backcrossing, contributing to the reduction in the number of generations needed to obtain the converted Cariflora, besides helping to obtain truly superior genotypes.

ACKNOWLEDGMENTS

REFERENCES

Arumuganathan K and Earle ED (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9: 208-218.

Benchimol LL, Souza CL Jr and Souza AP (2005). Microsatellite-assisted backcross selection in maize. *Genet. Mol. Biol.* 28: 789-797.

Collard BCY, Jahufer MZZ, Brouwer JB and Pang ECK (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142: 169-196.

Conover RA, Litz RE and Malo SE (1986). Cariflora, a Papaya for South Florida with Tolerance to Papaya Ringspot Virus. Agricultural Experiment Station, University of Florida (Circular, 329), Florida.

Cruz CD (2008). Programa GENES: Diversidade Genética. Universidade Federal de Viçosa, Viçosa.

Daher RF, Pereira MG, Tupinambá EA, Amaral Júnior AT, et al. (2002). Assessment of coconut tree genetic divergence by compound sample RAPD marker analysis. *Crop Breed. Appl. Biotechnol.* 3: 431-438.

Damasceno Junior PC, Costa FR, Pereira TNS, Freitas Neto M, et al. (2009). Karyotype determination in three Caricaceae species emphasizing the cultivated form (*C. papaya* L.). *Caryologia* 62: 10-15.

Davies J, Berzonsky WA and Leach GD (2006). A comparison of marker-assisted and phenotypic selection for high grain protein content in spring wheat. *Euphytica* 152: 117-134.

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Faleiro FG, Ragagnin VA, Moreira MA and de Barros EG (2004). Use of molecular markers to accelerate the breeding of common bean lines resistant to rust and anthracnose. *Euphytica* 138: 213-218.

FAOSTAT (2009). Food and Agriculture Organization of the United Nations Statistical Database. Available at [http://faostat.fao.org/site/567/default.aspx#ancor]. Accessed August 20, 2009.

Gupta PK and Varshney RK (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113: 163-185.

Hospital F and Charcosset A (1997). Marker-assisted introgression of quantitative trait loci. *Genetics* 147: 1469-1485.

Hospital F, Moreau L, Lacoudre H, Charcosset A, et al. (1997). More on the efficiency of marker-assisted selection. *Theor. Appl. Genet.* 95: 1181-1189.

Knapp SJ (1997). Marker assisted selection as a strategy for increasing the probability of selection superior genotypes. *Crop Sci.* 38: 1164-1174.

Kuchel H, Ye G, Fox R and Jefferies S (2005). Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy. *Mol. Breed.* 16: 67-78.

Leal AA, Mangolin CA, do Amaral Júnior AT, Gonçalves LS, et al. (2010). Efficiency of RAPD versus SSR markers for determining genetic diversity among popcorn lines. *Genet. Mol. Res.* 9: 9-18.

Lewis RS and Kernodle SP (2009). A method for accelerated trait conversion in plant breeding. *Theor. Appl. Genet.* 118: 1499-1508.

Liu Z, Moore PH, Ma H, Ackerman CM, et al. (2004). A primitive Y-chromosome in papaya marks incipient sex chromosome evolution. *Nature* 427: 348-352.

Marin SLM, Pereira MG, Amaral Júnior AT, Martelleto LAP, et al. (2006). Heterosis in papaya hybrids from partial diallel of 'Solo' and 'Formosa' parents. *Crop Breed. Appl. Biotechnol.* 6: 24-29.

Moore PH and Ming R (2008). Papaya genome: a model for tropical fruit trees and beyond. *Trop. Plant Biol.* 1: 179-180.

Oliveira EJ, Silva AS, Carvalho AM, Santos LF, et al. (2010). Polymorphic microsatellite marker set for *Carica papaya* L. and its use in molecular-assisted selection. *Euphytica* 173: 279-287.

Oliveira LK, Melo LC, Brondani C, Peloso MJ, et al. (2008). Backcross assisted by microsatellite markers in common bean. *Genet. Mol. Res.* 7: 1000-1010.

Pérez OJ, Dambier D, Ollitrault P, Coppens DG, et al. (2006). Microsatellite markers in *Carica papaya* L.: isolation, characterization and transferability to *Vasconcellea* species. *Mol. Ecol. Notes* 6: 212-217.

Santos SC, Ruggiero C, Silva CLSP and Lemos GM (2003). A microsatellite library for *Carica papaya* L. cv Sunrise Solo. *Rev. Bras. Frutic.* 25: 263-267.

Schmierer DA, Kandemir N, Kudma DA, Jone BL, et al. (2004). Molecular marker-assisted selection for enhanced yield in malting barley. *Mol. Breed.* 14: 463-473.

Silva FF, Pereira MG, Campos WF, Damasceno Júnior PC, et al. (2007a). Monitoring of the genetic variability in papaya parent 'Formosa' of 'UENF/CALIMAN 01' hybrid via RAPD. *Crop Breed. Appl. Biotechnol.* 7: 36-42.

Silva FF, Pereira MG, Campos WF, Damasceno Júnior PC, et al. (2007b). DNA marker-assisted sex conversion in elite papaya genotype (*Carica papaya* L.). *Crop Breed. Appl. Biotechnol.* 7: 52-58.

Siqueira WJ, Fonseca MIS and Sondhal MR (1988). Regeneration of *Lycopersicon esculentum* x L. *peruvianum* hybrid plants from two year old callus culture. *Bragantia* 47: 1-8.

Storey WB (1953). Genetics of the papaya. *J. Heredity* 44: 70-78.

Tanksley SD (1983). Molecular markers in plant breeding. *Plant Mol. Biol. Rep.* 1: 3-8.

Visscher PM, Haley CS and Thompson R (1996). Marker-assisted introgression in backcross breeding programs. *Genetics* 144: 1923-1932.

Wang J, Chen C, Na JK, Yu Q, et al. (2008). Genome-wide comparative analyses of microsatellites in papaya. *Trop. Plant Biol.* 1: 278-292.

Xi ZY, He FH, Zeng RZ, Zang ZM, et al. (2008). Characterization of donor genome contents of backcross progenies detected by SSR markers in rice. *Euphytica* 3: 369-377.

Young ND and Tanksley SD (1989). RFLP analysis of the size of chromosomal segments retained around the Tm-2 locus of tomato during backcross breeding. *Theor. Appl. Genet.* 3: 353-359.

Yu Q, Hou S, Feltus FA, Jones MR, et al. (2008). Low X/Y divergence in four pairs of papaya sex-linked genes. *Plant J.* 53: 124-132.