Use of microsatellite markers in molecular analysis of segregating populations of papaya (*Carica papaya* L.) derived from backcrossing

F.O. Pinto, M.G. Pereira, L.N. Luz, D.L. Cardozo, H.C.C. Ramos and C.M.P. Macedo

Laboratório de Melhoramento Genético Vegetal,
Centro de Ciências e Tecnologias Agropecuárias,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil

Corresponding author: L.N. Luz
E-mail: lucasluzbreeder@gmail.com

ABSTRACT. Brazil is the world leader in papaya production. However, only a small number of cultivars are registered for commercial planting, mainly owing to delays in obtaining cultivars and the high costs of the field phase of breeding programs. These costs can be reduced when molecular tools are combined with conventional breeding methods. In the present study, we conducted a molecular analysis of a self-fertilized population of a first backcrossing generation of BC_1S_1 papaya plants via microsatellite markers both to monitor the level of homozygosity and the gene/allele transfer that confers the Golden trait (fruit color) and to assess the parental genomic proportion in the genotypes studied. Based on the analysis of 20 polymorphic microsatellite loci, 19 genotypes with the Golden trait belonging to BC_1S_1 were evaluated in addition to the parental genotypes. Genetic distance was estimated through weighted index. The genotypes were then grouped using the hierarchical nearest neighbor method, and the analysis of principal coordinates was used to measure the proportion of parental genomes in the segregating genotypes. The mean value of the inbreeding coefficient was 0.36. The

analysis of the principal coordinates revealed that on average, 64% of the recurrent parent genome was present in the population. Together, the analyses allowed the selection of 3 individuals for the next backcross cycle ($33BC_1S_1$-18, $34BC_1S_1$-16, and $37BC_1S_1$-10). These individuals had a higher proportion of the recurrent parent and were grouped close to the recurrent parent in the cluster analysis.

Key words: Microsatellite marker; Assisted selection; Inbreeding

INTRODUCTION

Papaya (*Carica papaya* L.) is a prominent tropical fruit and grows in more than 50 countries around the world. It is required mainly for fresh fruit market, but other sectors related to the production of papaya are gaining importance, such as the extraction of proteolytic enzyme papain, which is widely used in pharmaceutical, food, as a component of softeners meat, as well as in processed fruit market.

C. papaya L. is the only commercial species of the genus. It is diploid, has 9 pairs of chromosomes (Damasceno Junior et al., 2009), a relatively small genome (372 Mb), and 1 pair of sex chromosomes (Liu et al., 2004; Yu et al., 2008). Papaya plants present 3 sexual forms: female, male, and hermaphrodite (Storey, 1953). However, hermaphrodite plants are tracks in commercial crops; therefore, they produce fruit with lower ovarian cavity, a characteristic that facilitates packaging of fruits and transport. Currently, the main hindrances to the continuous expansion of papaya crops are the small number of cultivars available in the market and low genetic variability (Kim et al., 2002; Ma et al., 2004). Ming et al. (2008) cited cultural preferences and geographic isolation as triggers of genetic erosion and consequent low diversity. Breeding programs tend to intensify the use of germplasm from various sources, broadening the genetic base and increasing genetic gains. These improvements will require the use of tools that assist and improve the efficiency of obtaining superior strains, such as the use of assisted selection markers in locating QTL.

The development of papaya strains is carried out through crosses that advance generations for evaluation, selection, and self-fertilization or through backcrossing involving 1 recurrent elite parent and another donor of single inheritance. The latter process has become more effective with the use of DNA markers for monitoring the proportion of the recurrent parent in the lines under development, reducing the number of generations from backcrosses, and hence the time needed to obtain pure lines. The efficiency of marker-assisted selection in plant breeding is well documented and especially practical in crops such as papaya, which has a sequenced genome. In addition, a wide availability of markers are available for the crop (Hospital and Charcosset, 1997; Oliveira et al., 2010).

Another important step in obtaining cultivars is the search for resistance or tolerance to disease. Diseases that impair the visual quality of fruits for fresh consumption deserve closer attention. Skin freckles (SF) cause huge losses in papaya culture, because it affects the appearance of fruit, decreasing its final value. SF is an abiotic disorder that is more intensively observed in the side of the fruit exposed to direct sunlight (Liberato and Zambolim, 2002; Ueno et al., 2002). The cause of this disease is not yet fully understood, although it has been reported in commercial plantings since the 1960s (Ishii and Holtzmann, 1963). Possible

causes have been discussed in the literature (Campostrini et al., 2005; Reis et al., 2008; Oliveira and Vitória, 2011).

According to Kaiser et al. (1996), stains become patterns of change in fruit skin and form visible eruptions on unripe and ripe fruits. Producers and breeders have reached consensus that the cultivar 'Golden Sunrise', of the Solo group, which produces light green fruits, is tolerant to SF. Therefore, this cultivar is the most common throughout Brazil. Although current understanding of the mechanisms of SF is deepening, the immediate interest of producers is to achieve lines such as 'Golden Sunrise' with light green fruits, a trait of the Golden cultivar. The Golden trait behaves as a recessive inheritance in controlled crosses.

Therefore, we made attempts to transfer the genes responsible for the Golden trait, monitored via the light green color of fruits, to the parental genotypes of the hybrid 'UENF/ Caliman 01'. A backcrossing program has been developed for this purpose in which the parental donor was the cultivar 'Golden Sunrise' and the recurrent genomes were the parents of the hybrid 'UENF/Caliman 01'. The goal is the recovery of the recurrent parent genome (RP) for hybrid recomposition in the "Golden type" version but with light green fruits. For such, microsatellite markers were used to assist backcrosses and characterize the segregant population at the molecular level for a parent of the hybrid.

MATERIAL AND METHODS

Plant material

We evaluated 19 segregating genotypes derived from an initial cross between the genotype JS12 (Formosa group), a parent of the hybrid 'UENF/Caliman 01' as the recurrent parent and the genotype 'Golden Sunrise' (Solo group) as the donor parent. After the F_1 was obtained, the plants were backcrossed with the recurrent parent (JS12) to obtain the BC_1 generation. The BC_1 plants were self-fertilized to obtain the BC_1S_1 generation and the expression of the trait. As expected, a smaller proportion of the plants expressed the Golden trait when self-fertilized given the recessive inheritance of the trait. Only plants of the Golden type were considered for the evaluation and advancement of backcrosses.

Extraction of genomic DNA

The extraction of total genomic DNA was carried out from young leaves using a Plant Genomics DNA Extraction Kit 100 YGP RBC (BioAmerica Biotech, USA) according to manufacturer instructions. The leaves were collected from individual plants that exhibited the Golden trait and from the parents (JS12 and Golden Sunrise), and they were macerated in liquid nitrogen at -196°C.

After extraction, the DNA was quantified via 0.8% agarose gel analysis using High DNA Mass Ladder (Invitrogen, USA). The DNA samples were stained using a mixture of Blue Juice and GelRed™ (Invitrogen) at a 1:1 ratio. The image capture was then carried out using the MiniBis Pro photo-documentation system (Bio-Imaging Systems). The images were subjected to quantification analysis using the Image J software system and diluted to a final concentration of 10 ng/μL.

Simple sequence repeat (SSR) marker analysis

The DNA of the parental genotypes (JS12 and Golden) was used for screening with microsatellite primers to identify polymorphic SSR loci capable of differentiating the 2 genotypes. A total of 114 primers were analyzed, of which only 20 provided a well-defined distinction between the parents. Such SSR markers have been developed by Eustice et al. (2008) from the genetic mapping performed by Chen et al. (2007).

The amplification reactions of the parental and segregating genotypes were then performed according to the descriptions of Ramos et al. (2011), with some modifications. The annealing temperature of the primers ranged from 60° to 63°C, according to each primer. The amplification products were separated via horizontal electrophoresis on 4% agarose gel (Metaphor) stained with GelRed™ and visualized with the MiniBis Pro photo-documentation system (Bio-Imaging Systems).

Data analysis

The data obtained from the amplification of the 20 SSR markers were converted into a numeric code for each allele per locus, and a numeric matrix was developed by assigning values from 1 to the maximum number of alleles per locus - in this case, up to 2 alleles per locus. A locus with 2 alleles represented 11 and 22 for the homozygous forms (A_1A_1 and A_2A_2) and 12 for the heterozygous forms (A_1A_2). The genetic distance among the evaluated genotypes was calculated based on the numeric matrix with the help of the GENES software system (Cruz, 2008). Cluster analysis was carried out with the nearest neighbor hierarchical method using Molecular Evolutionary Genetics Analysis version 5, and graphic dispersion of the genotypes was based on the principal coordinate analysis using Genalex version 6.3 (Kumar et al., 2008; Peakall and Smouse, 2009).

The rates of polymorphism [polymorphic information content (PIC)], the Shannon index, expected heterozygosity (H_E), and observed heterozygosity (H_O) were estimated using PowerMarker version 3.25 (Liu and Muse, 2005) and PopGene version 1.31 (Yeh et al., 1999). The genotypes were also evaluated for their genetic structure using the Bayesian method with Structure 2.3.1 (Pritchard et al., 2000). The admixture model was applied, as well as the correlated allele frequencies, using a burn-in period of 5000 followed by an extension of 50,000 replications during the analysis. The optimal K was 2. The genetic structure of the populations was also assessed using a Bayesian approach that evaluated the SSR genotypes with the Structure software system (Pritchard and Stephens, 2007).

The analysis of the proportion of the genome of the recurrent parent in self-fertilized individuals of the first backcross generation (BC_1S_1) was performed by recording the alleles of the recurrent parent in each individual of the segregating generations for the 20 SSR loci analyzed. Thus, the percentage of the recurrent genome could be calculated, which allowed the generation of graphics from the data obtained in individuals with the Golden trait.

RESULTS AND DISCUSSION

The segregating genotypes (BC_1S_1) were analyzed with the 20 polymorphic microsatellite markers that were effective in the differentiation of parental genotypes among the 114 markers evaluated. The SSR loci generated a total of 40 alleles, 2 alleles per locus (Table 1).

Table 1. Sequence of 20 microsatellite primer pairs used in the analysis of 19 Golden genotypes derived from BC_1S_1 and their parents.

Locus	Primer sequence (5'→3')	Ta (°C)	N_A	H_E	H_O	PIC	I
SSR 4	GAATCACGCGATGAGTCACA	60	2	0.51	0.57	0.37	0.69
SSR 8	GAAGCCTCAGTGAATCCAAA	60	2	0.45	0.28	0.34	0.63
SSR 10	TTTCTCCCACATGCACCACA	62	2	0.39	0.23	0.31	0.57
SSR 20	CTCCAAGAAAACTTTGACATGGG	60	2	0.45	0.09	0.34	0.63
SSR 40	CGGTAGCGACTCATCGGACT	60	2	0.50	0.90	0.37	0.68
SSR 41	TGGTGGATGTTGATGCATGTT	60	2	0.49	0.23	0.36	0.67
SSR 49	GCAGTCTCTTCTCTCCAAGGTCA	63	2	0.45	0.28	0.34	0.68
SSR 50	GCATGATGGTTCGGTTCAT	60	2	0.51	0.19	0.37	0.69
SSR 51	AAAGATGACCGGAGCCGT	60	2	0.50	0.23	0.37	0.68
SSR 70	TTCATCGTCTCGCTGAAATTGA	60	2	0.17	0.19	0.15	0.31
SSR 76	AACCAGAACTCCAAAAAGAAATCC	60	2	0.47	0.71	0.35	0.65
SSR 82	GACGTCAGCATAAAACTGGCAA	60	2	0.51	0.70	0.37	0.69
SSR 85	AGCACTGAGAAACGCCAAGC	60	2	0.28	0.33	0.23	0.45
SSR 86	CCTTTGCATTGTCATGGCCT	60	2	0.43	0.33	0.33	0.61
SSR 88	TGTTCTTCAATTGGTGAGAGTTTG	60	2	0.17	0.09	0.15	0.31
SSR 89	TCAAATGGCTGAAGCAAAGG	60	2	0.34	0.14	0.28	0.51
SSR 93	ATTTGACGTGGCAGCACCTT	60	2	0.37	0.28	0.29	0.54
SSR 101	GGTTGGATTGCTTTGCCTTAAA	60	2	0.48	0.19	0.36	0.66
SSR 107	GCAGCCCGTACAGAAGAGGA	60	2	0.39	0.23	0.31	0.57
SSR 114	CACACACTCGCAATGGTAAAGAA	60	2	0.51	0.09	0.37	0.69
Mean				0.42	0.31	0.32	0.60

Ta = annealing temperature; N_A = number of alleles; H_E = expected heterozygosity; H_O = observed heterozygosity; PIC = polymorphism information content; I = Shannon's index.

The greater the number of alleles, the greater the probability of describing the studied population. The number of polymorphic markers is occasionally associated with the number of alleles found in the population, as demonstrated by Trindade and Dantas (2008), who reported the identification of 86 alleles of 27 polymorphic markers in papaya. However, this relationship is greatly affected by the type of marker used and the nature of the population studied. Ramos et al. (2011) reported finding 39 alleles, using only 16 polymorphic markers in papaya. In the present study, the number of alleles per locus was consistent with the nature of biparental segregating population used.

The analysis of SSR loci allowed an estimation of H_E, which ranged from 0.17 to 0.51 with an average of 0.42, and H_O, which ranged from 0.09 to 0.90 and averaged 0.31. Notably, considering the parents used, H_E was anticipated to be 0.50 in agreement with the H_E obtained with the SSR loci. Furthermore, for the BC_1S_1 generation, the H_O would be 0.25. In other words, for the number of individuals evaluated (19) as well as the number of SSR loci (20), good agreement occurred between the mean values of heterozygosity (H_E: 0.42 *vs* 0.50; H_O: 0.31 *vs* 0.25). This relative discrepancy is certainly due to the size of the population studied and the number of loci considered.

The estimated PIC can also be used to quantify the genetic polymorphism of each locus in a segregating population. PIC values ranged from 0.15 to 0.37 with a mean value of 0.32. The lowest value (0.15) was found for the loci SSR 70 and SSR 88; and the highest value (0.37) was found for 6 loci: SSR 4, SSR 40, SSR 50, SSR 51, SSR 82, and SSR 114. The PIC values indicated the discriminatory power of the loci analyzed. According to Cruz et al. (2011), the greater the number of alleles, the higher the tendency of the PIC value to approach H_E. In this case, on average, the agreement was very high, because H_O was 0.31 and the average PIC value was 0.32.

The analysis of genetic diversity within populations can be complemented by the Shannon index, which is a measure of the genotypic richness of a population. This index has been proposed by Alfenas (1991) to measure the diversity or richness of species in ecology and can also be used as a measure of phenotypic or genotypic diversity within a population (Picoli, 2004). The Shannon index dictates that the closer to unity the estimated values are, the greater the diversity. Shannon's index values for the genotypes evaluated ranged from 0.31 to 0.69. The mean value was 0.60, which revealed the existence of a moderate variation in the population studied, although it is sufficient for the continuity of the breeding program because it refers to self-fertilized plants of the first backcross generation (BC_1S_1). These values are expected to increase in subsequent backcrosses (BC_2) and self-fertilization (BC_1S_2) generations.

The analysis of diversity among genotypes was based on estimated H_E, which ranged from 0.05 to 0.60 with a mean of 0.32 (Table 2) and H_O, with values between 0.04 and 0.49 and an average value of 0.42. According to Cruz et al. (2011), for backcross populations, the values of PIC, Shannon's index, and H_O tend to decrease over generations. The same effect occurs with the advancement of self-fertilization generations, which leads to a progressive variability loss. As mentioned, the H_O average anticipated for the generation studied was 0.25, a value close to the average observed (0.31).

Table 2. Parents and segregating genotypes (BC_1S_1) and respective values of observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (f).

Individual	Genotypes	H_O	H_E	f
1	JS12	0.04	0.05	0.90
2	Golden	0.25	0.20	0.60
3	$30BC_1S_1$-1	0.45	0.40	0.22
4	$30BC_1S_1$-5	0.46	0.35	0.30
5	$30BC_1S_1$-6	0.49	0.50	-0.02
6	$30BC_1S_1$-10	0.46	0.45	0.06
7	$31BC_1S_1$-21	0.42	0.35	0.32
8	$33BC_1S_1$-1	0.45	0.25	0.51
9	$33BC_1S_1$-8	0.45	0.40	0.21
10	$33BC_1S_1$-14	0.46	0.15	0.70
11	$33BC_1S_1$-18	0.42	0.30	0.36
12	$33BC_1S_1$-19	0.46	0.35	0.32
13	$34BC_1S_1$-4	0.42	0.20	0.54
14	$34BC_1S_1$-16	0.34	0.05	0.89
15	$34BC_1S_1$-21	0.43	0.25	0.50
16	$37BC_1S_1$-7	0.48	0.25	0.50
17	$37BC_1S_1$-10	0.36	0.15	0.69
18	$37BC_1S_1$-14	0.49	0.35	0.32
19	$38BC_1S_1$-7	0.49	0.60	-0.22
20	$38BC_1S_1$-9	0.48	0.52	-0.02
21	$38BC_1S_1$-12	0.48	0.60	-0.17
Mean		0.42	0.32	0.36

The inbreeding coefficient or fixation index (f) is an extraordinarily important parameter for breeding programs aimed at developing lines because it allows the determination of the level of homozygosity and heterozygosity in a population. The analysis of the f values of the self-fertilized individuals from the first backcrossing generation of this study yielded values ranging from -0.22 to 0.90 and average values of 0.36 (see Table 2). For the anticipated H_E and H_O values of 0.50 and 0.25, respectively, in the present study, the panmictic index (H_O/H_E) was 0.5; hence, the value of f was 0.50. In other words, the comparison between the observed

(0.36) and expected (0.50) average f values leads to the conclusion that deficiency exists in inbreeding, with a consequent excess of loci in heterozygosis. Again, the best explanation is associated with aspects of sample size. It must be pointed out that knowing the values of each individual is important. Although average values are expected, each individual has its own coefficients. The relevance of using SSR markers is evident considering the importance of the knowledge of such coefficients in decision making related to selection for the next generation. In the present case, the identification of segregating individuals is possible: those with the Golden trait and f value ($34BC_1S_1$-16, $33BC_1S_1$-14, and $37BC_1S_1$-10). Their f values were, respectively, 0.89, 0.70, and 0.69. In other words, interesting candidates for generation advancement can already be identified in the BC_1S_1 generation.

Conversely, values much higher than those found in this study have been described for papaya. Oliveira et al. (2010) reported f values between 0.63 and 1.00 and identified 11 pure lines with 100% inbreeding and 18 lines with values close to the maximum (0.95 and 0.96). However, the authors analyzed 83 pure lines and only 3 segregating populations.

The average ratio of recovery of the RP was estimated using microsatellite markers. Regarding the average proportion of the RP, each backcross generation is expected to be similar to the recurrent parent in $1-(1/2)^{t+1}$ on average (Allard, 1960). Considering all genotypes evaluated to belong to the BC_1S_1 generation, the analysis of the RP proportion in the population showed an average of 64% of the RP, with individual values ranging significantly from 42 to 95% (Figure 1). For the current generation (BC_1), this proportion was expected to be 75%.
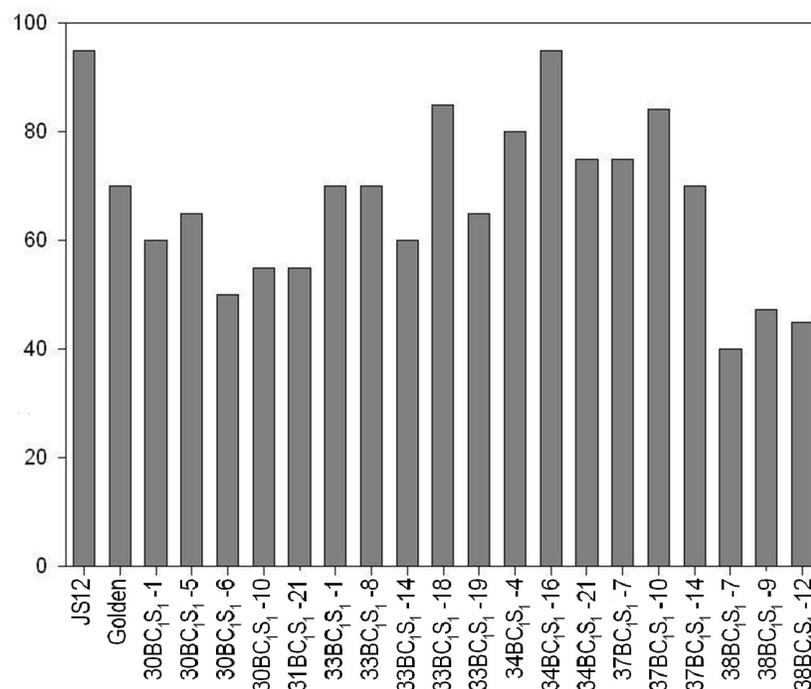

Figure 1. Proportion of recurrent parent genome (PR) in the 19 genotypes with the Golden trait that belongs to the BC_1S_1 generation. JS12 = recurrent parent; Golden = parental donor.

The variation in individual values of the population indicated that the RP genomic proportions of some genotypes were within the expected rate and that the proportions of others were above or below it. It also demonstrated that the genomic similarity between each Golden individual and the recurrent parent depended on both the number of backcross generations and the level of recombination. The results obtained showed 3 genotypes with the Golden trait that had a greater ratio of the recurrent parent: $33BC_1S_1$-18 (90%), $34BC_1S_1$-16 (95%), and $37BC_1S_1$-10 (88%).

Low proportion levels of the recurrent parent were also obtained by Silva et al. (2007). However, although few genotypes with satisfactory genomic proportions are presented here, a considerable proportion of the recurrent parent is already present in the genotypes of the Golden type, and with just 1 additional backcross generation in these progenies, the proportion may be close to 100%.

Ramos et al. (2011) reported that the selection of phenotypic attributes can also cause a deviation in selection favoring the donor parent and delaying the recovery of the recurrent genome. Frisch et al. (1999) suggested that gene blocks related to the genes of interest from the donor parent can also contribute.

The genetic dissimilarity matrix generated was subjected to the clustering technique using the nearest neighbor method (Figure 2). With a cutoff point of approximately 0.15 (average distance), 4 groups of Golden-type individuals and recurrent and donor parents were formed in the extremities. Notably, the segregating individuals were grouped closer to the recurrent parent, as expected. In addition to the recurrent parent (JS12), group I joined the individuals $37BC_1S_1$-10, $31BC_1S_1$-21, $33BC_1S_1$-1, and $37BC_1S_1$-1. These individuals were genetically closest to the recurrent parent and were possible candidates for the next selection to the advance self-fertilization and backcrossing generations.
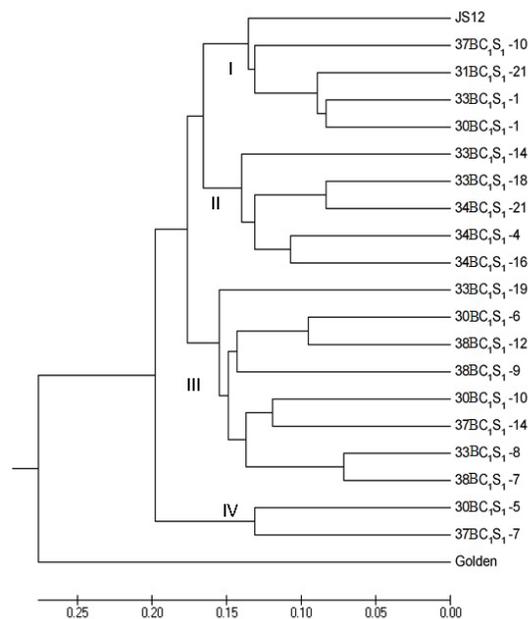

Figure 2. Dissimilarity dendrogram obtained by UPGMA of the 19 genotypes with the Golden trait from the BC_1S_1 generation, in addition to the parental genotypes JS12 and Golden. Cophenetic correlation coefficient = 0.91.

The second group brought together the individuals $34BC_1S_1$-14, $33BC_1S_1$-18, $34BC_1S_1$-21, $34BC_1S_1$-4, and $34BC_1S_1$-16. The third group combined the most individuals - namely, $33BC_1S_1$-19, $30BC_1S_1$-6, $38BC_1S_1$-12, $38BC_1S_1$-9, $30BC_1S_1$-7, $37BC_1S_1$-14, $33BC_1S_1$-8, and $38BC_1S_1$-7. Individuals $30BC_1S_1$-5 and $37BC_1S_1$-7 formed group IV and were the most genetically distant from the recurrent parent.

The analysis of genetic relationships among genotypes was also assessed with graphic dispersion via principal coordinate analysis (Figure 3). The first 2 coordinates together explained 55.6% of the total data variation, whereas 36.28% of this variation was explained by coordinate 1 and 19.29% by coordinate 2. Although moderate, these values indicated some reliability of the dispersion shown. A clear separation exists between the genotypes of the recurrent 'JS12' and donor 'Golden Sunrise' parents. Verification that individuals with Golden-type trait were scattered closer to the recurrent parent than to the donor parent was also possible.
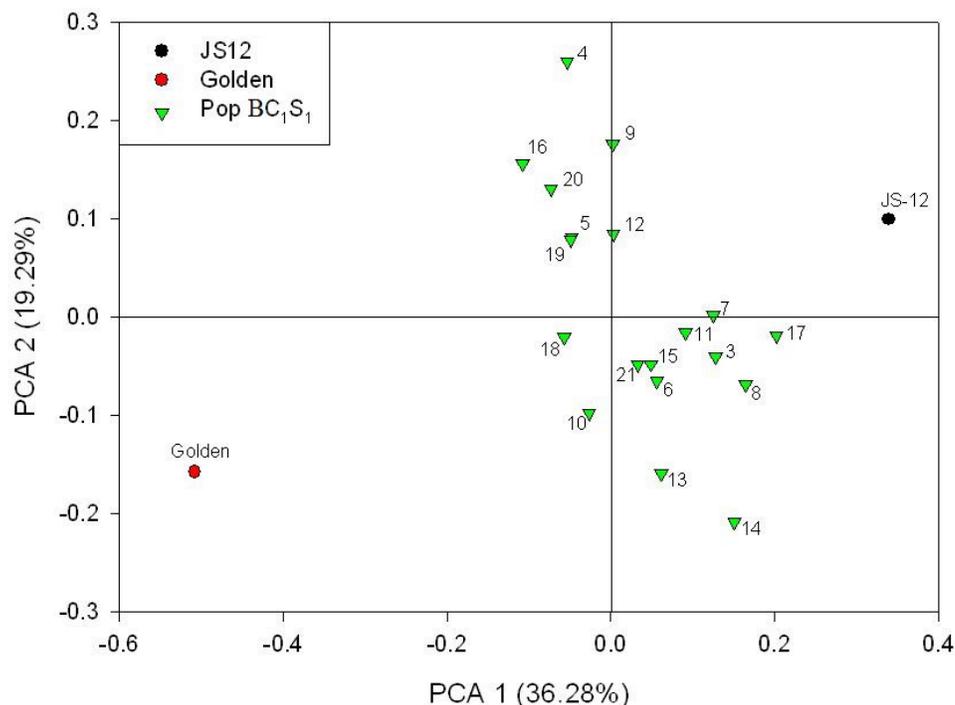

Figure 3. Principal coordinate analysis (PCA), considering 19 genotypes of papaya derived from backcrossing, the recurrent parent (JS12) and parental donor (Sunrise Golden), based on the distance matrix obtained by analysis of microsatellite markers.

The analysis of the genetic structure of self-fertilized individuals of the first backcross generation using the Structure software system was carried out using the selection of an optimal K value of 2, which abruptly changed the graph (Figure 4). This change allowed the separation of population individuals into 2 genomic regions: red and green (Figure 5).
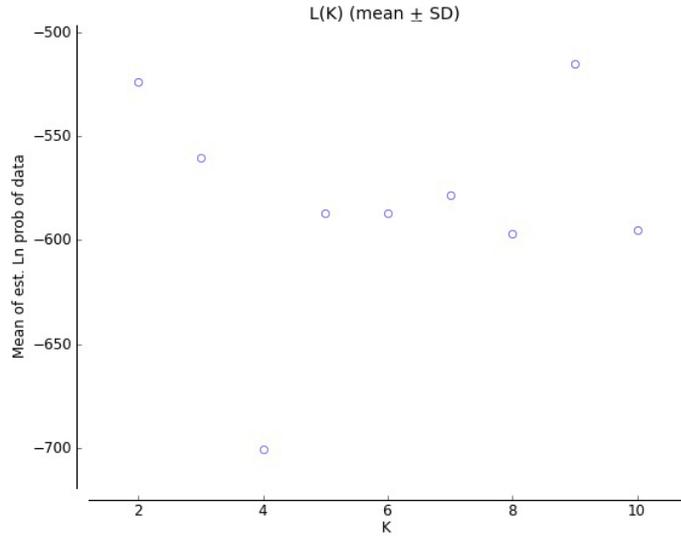
Figure 4. Graphics presenting the achievement of the optimal K for the analysis of the genetic structure of BC_1S_1 individuals.
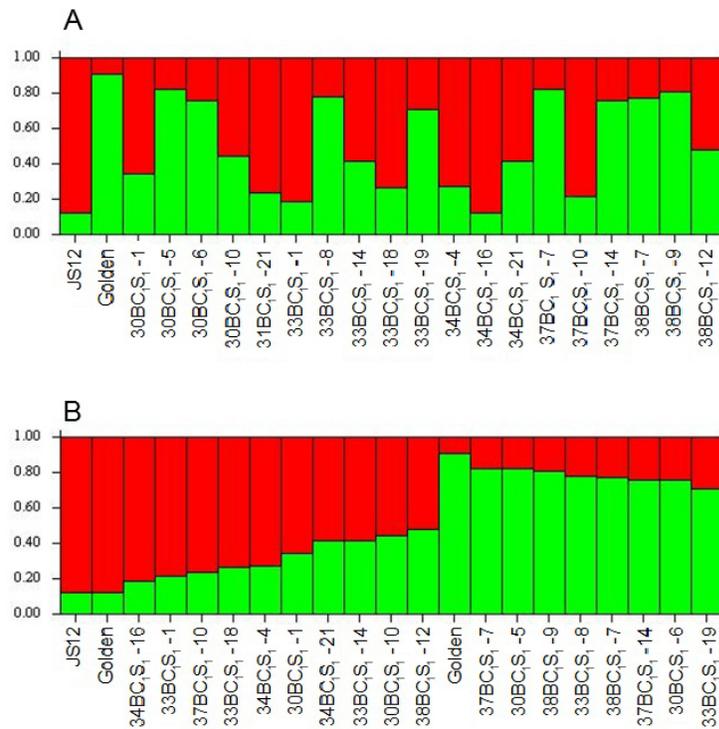

Figure 5. Analysis of genetic structure of the 19 progenies of papaya and 2, JS12 and Golden, parental genotypes BC_1S_1.

The contrast of colors in Figure 5A facilitates the visualization of the recurrent 'JS12' and donor 'Golden Sunrise' parents as well as the progenies. Most genotypes evaluated presented greater similarity on average with recurrent parent JS12, corroborating the graph obtained with principal coordinate analysis. In Figure 5B, the progenies with genomic regions more similar to the recurrent parent according to genetic structure were the genotypes $34BC_1S_1$-16, $33BC_1S_1$-1, $37BC_1S_1$-10, $31BC_1S_1$-21, $33BC_1S_1$-18, $34BC_1S_1$-4, $30BC_1S_1$-1, $34BC_1S_1$-21, $33BC_1S_1$-14, $30BC_1S_1$-10, and $38BC_1S_1$-12, representing 58% of the total individuals of this population that have the Golden-type trait.

In general, the various analyses agreed. Comparing diversity estimates and the recording of the recurrent parent alleles in progeny, 4 individuals were the most similar to the recurrent parent being: 37BC1S1-10, 31BC1S1-21, 33bC1S1-1, and 30BC1S1-1 and may be able to advance the generations.

In crops such as papaya, speed in the advancement of breeding populations is especially significant considering the long generation time. When this process is combined with molecular marker characterization, it becomes even more attractive owing to the survey of the proportion of the recurrent parent in the progenies. In specific cases of hybrid recomposition, new lines must have as much genetic and phenotypic similarity as possible, except for the trait intended to be changed, to ensure the suitability of the hybrids generated from the new lines. The approach outlined in this study was effective in segregating not only individuals with the Golden trait but also those with high f value and high RP.

ACKNOWLEDGMENTS

REFERENCES

Alfenas AC (1991). Eletroforese de Proteínas e Isoenzimas de Fungos de Essências Florestais. Universidade Federal de Viçosa, Viçosa.

Allard RW (1960). Principles of Plant Breeding. Wiley, New York.

Campostrini E, Lima HC, Oliveira JG, Monnerart PH, et al. (2005). Teores de Ca e variáveis meteorológicas: relações com a mancha fisiológica do mamão no norte fluminense. *Bragantia* 64: 601-613.

Chen C, Yu Q, Hou S, Li Y, et al. (2007). Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. *Genetics* 177: 2481-2491.

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

Cruz CD, Ferreira FM and Pessoni LA (2011). Biometria Aplicada ao Estudo da Diversidade Genética. Universidade Federal de Viçosa, Viçosa.

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.

Eustice M, Yu Q, Lai CW, Hou S, et al. (2008). Development and application of microsatellite markers for genomic analysis of papaya. *Tree Genet. Genomes* 4: 333-341.

Frisch M, Bohn M and Melchinger AE (1999). Comparison of selection strategies for marker-assisted backcross of a gene. *Crop Sci.* 39: 1295-1301.

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

Ishii Y and Holtzmann OV (1963). Papaya mosaic disease in Hawai. *Plant Desease Rep.* 47: 947-951.

Kaiser C, Allan P, White BJ and Dehrmann FM (1996). Some morphological and physiological aspects of freckle on papaya (*Carica papaya* L.) fruit. *J. South Afr. Soc. Hortic. Sci.* 6: 37-40.

Kim MS, Moore PH, Zee F, Fitch MM, et al. (2002). Genetic diversity of *Carica papaya* as revealed by AFLP markers.

Genome 45: 503-512.

Kumar S, Nei M, Dudley J and Tamura K (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* 9: 299-306.

Liberato JR and Zambolim L (2002). Controle das Doenças Causadas por Fungos, Bactérias e Nematóides em Mamoeiro. In: Controle de Doenças de Plantas: Fruteiras (Zambolim L, Vale FXR and Costa H, eds.). Editora UFV, Viçosa, 1023-1170.

Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129.

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.

Ma H, Moore PH, Liu Z, Kim MS, et al. (2004). Highdensity linkage mapping revealed suppression of recombination loco in papaya. *Genetics* 166: 419-436.

Ming R, Yu Q, Bias A, Chen C, et al. (2008). Genomics of Papaya, A Common Source of Vitamins in the Tropics. In: Genomics of Tropical Crop Plants (Moore PH and Nilno R, eds.). Springer, New York.

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 JG and Vitória AP (2011). Papaya: Nutritional and pharmacological characterization, and quality loss due to physiological disorders: An overview. *Food Res. Int.* 44: 1306-1313.

Peakall R and Smouse P (2009). GenAlEx Tutorials - Part 1: Introduction to Population Genetic Analysis. Australian National University, Canberra.

Picoli EAT (2004). Influence of RAPD number of markers and sample size on Eucalyptus genetic distance and diversity. *Crop Breed. Appl. Biotechnol.* 4: 384-390.

Pritchard JK and Stephens PDP (2007). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multiloco genotype data. *Genetics* 155: 945-959.

Ramos HC, Pereira MG, Silva FF, Goncalves LS, et al. (2011). Genetic characterization of papaya plants (*Carica papaya* L.) derived from the first backcross generation. *Genet. Mol. Res.* 10: 393-403.

Reis LF, Campostirni E and Netto AT (2008). Skin freckles on Formosa papaya: relationships with soil water potential total soluble solids of latex and climate variables. *Ciênc. Agrotec.* 32: 1473-1480.

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

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

Trindade AVT and Dantas JLL (2008). Embrapa Mandioca e Fruticultura Tropical: Programa de Pesquisa e Resultados Obtidos com a Cultura do Mamão. Papaya Brasil, Vitória.

Ueno B, Neves EF, Machado Filho JA, Yamanishi OK, et al. (2002). Mancha Fisiológica em Frutos de Mamoeiro no Oeste da Bahia: Relatório de Trabalho da Parceria Universidade de Brasília com os Produtores de Mamão da Associação dos Irrigantes do Oeste da Bahia (AIBA). Brasília, 109.

Yeh FC, Boyle T, Rongcai Y, Ye Z, et al. (1999). POPGENE. Microsoft Window-Based Freeware for Population Genetic Analysis. Version 1.31, Manual.

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.