Crop Breeding and Applied Biotechnology 7:52-58, 2007
Brazilian Society of Plant Breeding. Printed in Brazil

DNA marker-assisted sex conversion in elite papaya genotype (*Carica papaya* L.)

Francisco Filho da Silva[1], Messias Gonzaga Pereira[1*], Wellington Ferreira Campos[1], Pedro Corrêa Damasceno Junior[1], Telma Nair Santana Pereira[1], Gonçalo Apolinário de Souza Filho[1], Helaine Christine Cancela Ramos[1], Alexandre Pio Viana[1], and Geraldo Antônio Ferreguetti[2]

Received 29 March 2006

Accepted 24 June 2006

ABSTRACT – *Papaya is a genetically narrow-based crop, with few genetically distinct cultivars available for planting. This study aimed at the sex conversion of genotype 'Cariflora', from the dioecious to the gynoecious - andromonoecious stage, by means of RAPD marker-assisted introgression of the M^2 allele. Hierarchic clustering in the BC_2 generation showed three plants close to the recurrent parent (15, 18 and 89) with desirable phenotypic traits; the genetic divergence matrix indicated that this similarity was 84, 81 and 76 %, respectively. The genetic dissimilarity matrix presented a mean genetic distance between the recurrent parent and the BC_2 plants of 75.7 %, which was less than the expected 87.5 %. The phenotypic selection, applied in the BC_1 and BC_2 populations, partly explains this deviation in favor of the donor parent.*

Key words: Papaya, RAPD markers, assisted backcross, sex expression, Jaccard index.

INTRODUCTION

The cultivated papaya tree (*Carica papaya* L.) has basically three flower forms: female, male and hermaphrodite. According to Hofmeyr (1938) and Storey (1938), the heritability of the gender in this species is monogenic, through three alleles denominated *m*, M^1 and M^2. Storey (1953) claimed that the dominant combinations, M^1M^1, M^2M^2 and M^1M^2 are probably zygotic lethal. According to Storey (1941), the genotypes *mm*, M^1m and M^2m, are denominated gynoecious (female), androecious (male) and andromonoecious (hermaphrodite), respectively. Populations derived from these different tree types are distinguishable in: a) dioecious populations – plants with female flowers (gynoecious) and plants with male flowers (androecious) only; b) gynoecious-andromonoecious populations – plants with female flowers and plants with hermaphrodite flowers (andromonoecious); c) andromonoecious-trioecious populations - plants with female, plants with hermaphrodite and plants with male flowers.

In Brazil, the pear-shaped fruits of the hermaphrodite plants are preferred over round fruits of female plants and are produced to supply the domestic and international markets. However, papaya is sustained by a narrow genetic base, i.e., there are few cultivars and/or commercial hybrids available for planting that would meet the national as much as international market standards. Moreover, the high price of hybrid seed of the papaya 'Formosa' group, generally imported from Taiwan at 4,200 US dollars per kilogram (Silva et al. 2004),

[1] Laboratório de Melhoramento Genético Vegetal, Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual do Norte Fluminense Darcy Ribeiro. Av. Alberto Lamego, 2000. 28.013-600. Campos dos Goytacazes, RJ, Brasil. *E-mail: messias@uenf.br
[2] Caliman Agrícola S.A., C.P. 52, 29.900-970, Linhares, ES, Brasil

has caused many producers to successively plant the F_2, F_3 and F_4 generations of hybrid 'Tainung 01' This poses numerous problems, above all a loss of vigor and segregation of fruit shape.

It is evident that more efforts must be put into genetic improvement programs that broaden the accessible genetic base, create varieties with tolerance or resistance to the main diseases like the Papaya ringspot virus and "meleira", and present desirable agronomic traits as well, to satisfy the market demands. The hybridization of distinct genetic material along with a backcross program represents an important tool to make the introgression of new genes of interest into cultivated varieties possible and broaden the genetic variability in the subsequent generations (Siqueira et al. 1988).

The cross of the dioecious papaya genotype 'Cariflora' with genotypes of the 'Solo' group has resulted in hybrids with excellent yield performance (Marin et al. 2006). However, as a dioecious genetic material, the number of loci in heterozygosis is high and the resulting hybrids are quite heterogeneous. Furthermore, due to dioecy, selfing or the generation of lines are not possible. An alternative for overcoming these limitations and, consequently, amplify the genetic base of the cultivated papaya tree is to resort to the sex conversion of the genotype 'Cariflora' and of other dioecious papaya genotypes, by means of introgression of the M^2 allele (responsible for hermaphroditism). This study contributes to the understanding of the inheritance of morpho-agronomic attributes and the performance of traits associated to the sex expression of these genotypes. It is known that in the dioecious stage, in spite of the high number of loci in heterozygosis, the female are phenotypically more stable than hermaphrodite plants.

Associated to the classic improvement procedures, the use of DNA markers has opened up completely new perspectives for the papaya crop by means of assisted selection.

On this background, the present study had the objective of converting the gender of the 'Cariflora' genotype from the dioecious to the gynoecious-andromonoecious stage, by means of the introgression of the M^2 allele partly assisted by RAPD markers (Random Amplified Polymorphic DNA), so this material can be used to create endogamic lines and superior hybrids.

MATERIAL AND METHODS

Plant material

Hermaphrodite papaya plants of the BC_1 and BC_2 generations were used: the BC_1 generation was derived from an initial cross between the dioecious genotype 'Cariflora' (recurrent parent) and the elite variety 'Sunrise Soil 783' (SS 783) (M^2 allele donor parent). Later, the backcrossing procedure was conducted generating the BC_1 and BC_2 segregating populations for the gender (hermaphrodite and female) and other phenotypic traits. For the step of molecular genotyping, performed in the BC_2 generation only, young leaves of hermaphrodite plants as well as leaves in bulk were collected from both parents to extract genomic DNA.

The 'Cariflora' genotype is a dioecious selection with fruits of yellow pulp, moderate firmness, a mean weight of around 1.67 kg, besides pleasant taste and flavor (Conover et al. 1986). According to these authors, the tolerance level of this genotype to papaya ringspot virus (PRSV) is high in the edaphoclimatic conditions of southern Florida – USA. PRSV, also known as mosaic virus or papaya ringspot virus, represents one of the gravest constraints to the maintenance of production centers of this crop, imposing a migratory character instead, in view of the damage done and the absence of resistant varieties (Oliveira et al. 1994). The cross of 'Cariflora' with genotypes of the 'Solo' group results in very vigorous and productive hybrids, though fairly heterogeneous, with a high degree of loci in heterozygosis. On the other hand, genotype 'SS 783' is an elite variety of pear-shaped fruit with a mean weight of 0.52 kg, with red pulp of good quality.

Installation and evaluation of the experiments

The backcross generations were planted on a commercial stand, in March of 2003 and January of 2004, respectively, BC_1 and BC_2. In the first trial, the final spacing was 1.5 m between plants by 3.6 m between single rows and, in the second trial, 1.5 m x 2.0 m x 3.6 m in a double-row system. Fertilization, management, pest and disease control and the crop treatments were the customary, as used in commercial plantations.

Ninety-four and 87 hermaphrodite BC_1 and BC_2 plants were evaluated 180 days after transplanting (DAT) and 140 DAT, respectively, for morpho-

agrononomic attributes related to plant vigor and flower and fructification traits.

The phenotypic evaluations of the BC_1 population were used to select the plant for backcrossing with the recurrent parent 'Cariflora', in other words, this selection was based on the phenotype only; the plant with the best phenotypic attributes and most similar to the recurrent parent was chosen.

For the third backcross, aside from the phenotypic evaluations, RAPD markers were used to help select the BC_2 plant most similar to the recurrent parent. Young leaves were collected from hermaphrodite plants, just after the beginning of fructification. The leaf samples were taken to the laboratory where they were ground in liquid nitrogen, filled in 15 mL tubes (Falcon) and deep-frozen (-86 ºC).

Extraction of genomic DNA from hermaphrodite BC_2 plants

Of the ground leaf samples of the BC_2 generation and the two bulks of the parents, approximately 200 mg tissue was transferred to 1.5 mL Eppendorf tubes that were cooled in liquid nitrogen. The genomic DNA was extracted by the CTAB method (Doyle and Doyle 1990), with some modifications described by Daher et al. (2002). The DNA concentrations of the samples were estimated by electrophoresis in agarose gel (0.8 %) and stained with ethidium bromide (0.005 %), using the standard of a previously spectrophotometrically (Spekol, Zeiss, Germany) determined concentration of papaya DNA. The samples were standardized at a concentration of 10 $ng.\mu L^{-1}$.

Polymerase chain reaction (PCR)

The RAPD amplification reactions were conducted according to the modified protocol of Williams et al. (1990), in a final volume of 20 µL, containing Tris-HCl 10 mM (pH 8.3), KCl 50 mM, $MgCl_2$ 2.4 mM, 100 µM of each of the deoxyribonucleotides, 0.3 µM of each primers, 20 ng GENOMIC DNA and one unit of the enzyme *Taq* DNA polymerase (Pharmacia Biotech, EUA). The amplifications were performed in a thermocycler (Perkin Elmer GeneAmp PCR System 9700) programmed for 95 ºC for 1 minute followed by 45 cycles of 1 min at 94 ºC, 1 min at 36 ºC and 2 min at 72 ºC and a final extension of 7 min at 72 ºC. After the amplifications, the temperature of the samples was reduced to 4 ºC. The amplification products were analyzed by electrophoresis in 1.0 % agarose gel and visualized after staining with ethidium bromide under UV light and photo documented with the Eagle Eye II image system.

Molecular analysis

The readings of the RAPD loci of each genotype considered band presence (1) or absence (0). The arithmetic complement of the Jaccard index (Alfenas et al. 1991) that corresponds to the dissimilarity values was calculated as $1 - I_{AB}$ which is equal to $1 - a/(a+b+c)$, where **a** refers to the number of bands common to both genotypes, **b** to the number of bands found in genotype A only and **c** to the number of bands found in genotype B only.

Based on the genetic distance matrix the clustering analysis was performed using the "Hierarchic method of the most distant neighbor" (Cruz and Regazzi 1997). The statistical analyses were performed using the software packages Genes (Cruz 2001) and statistica version 95 (StatSoft Inc, Tulsa, Oklahoma, EUA).

RESULTS AND DISCUSSION

The BC_1 as well as the BC_2 populations presented broad variability for most traits, especially for variations of the elongate hermaphrodite flower (normal) to sterile and deformed female forms (carpelloid and pentandric), as well as to carpelloid and pentandric fruits. Then, in the BC_1 as much as the BC_2 population selection was performed against these flower variations, reinforced by the occasion of the leaf collection for the genomic DNA extraction. It is furthermore worth highlighting that the donor parent of allele M^2 (SS 783) presented a reduced expression of flower deformations and, consequently, of carpelloid and pentandric fruits in earlier evaluations (data not shown). The strong expression of these traits in the BC_1 and BC_2 populations (data not shown) is therefore probably inherited from parent 'Cariflora', since these traits would never find expression in the dioecious condition in this genotype.

Thirty - three primers were used to genotype 89 genotypes (87 trees of the BC_2 population plus the two parents) and a total of 98 bands was obtained (Table 1). Of the 98 bands, 54 were polymorphic, i.e., in mean, each primers generated 1.64 polymorphic bands. The

Table 1. Primers and number of monomorphic and polymorphic markers generated in the BC_2 papaya generation

Initiator	Markers	
	Monomorphic	Polymorphic
OPA 14	02	04
OPB 05	02	01
OPC 15	01	01
OPD 05	02	02
OPD 18	02	01
OPD 20	01	02
OPF 12	02	01
OPG 15	02	01
OPK 10	01	02
OPM 10	02	01
OPM 15	01	03
OPM 16	02	02
OPM 20	01	01
OPN 06	01	02
OPO 19	01	02
OPP 03	01	01
OPR 04	01	02
OPR 08	02	01
OPT 01	01	02
OPV 14	01	01
OPZ 10	01	01
OPAA 15	02	01
OPAC 11	02	01
OPAC 19	01	01
OPAF 15	01	03
OPAM 09	01	02
OPAN 10	02	02
OPAP 01	02	01
OPAP 07	01	02
OPAS 12	01	02
OPAU 08	01	01
OPAV 06	01	01
OPAV 19	0	03
Total	45	54

total number of bands per primers varied from 2 to 6 and 55.10 % of the bands showed polymorphisms.

In an analysis of the genetic distance matrix, the values of genetic dissimilarity in the BC_2 generation varied from 0.0145 to 0.3461, representing an amplitude of 0.3316, which was 0.4706 between the parents. The mean genetic dissimilarity coefficient of the BC_2 generation was 0.1802 and the standard deviation from the mean equal to 0.0426; i.e., the trees of generation BC_2 are genetically fairly close.

The sex conversion of genotype 'Cariflora' was obtained in only three backcross generations, because in the second generation, RAPD markers assisted the identification of candidate plants for recombination with female 'Cariflora' plants, establishing conditions that would allow a recovery of over 95 % of the recurrent genome in the third backcross.

The dendrogram of the genetic distance of the 87 trees of the RC_2 generation and the recurrent (1) and donor (2) parents is presented in Figure 1. The formation of two large groups was observed, which contained the parents, 'Cariflora' (1) and 'SS 783' (2). The vast majority of the genotypes grouped at a genetic distance of less than 20 %, though the two large groups were linked at a genetic distance of approximately 47 %, giving the two big groups consistency.

The first big group, which contains the recurrent parent (1), comprises 48.3 % of the genotypes, while 51.7 % were grouped in the second, which holds the donor parent of the M^2 allele (2). Two considerations must be looked into: the plant selection in the first backcross generation in order to establish the BC_2 population was based only on phenotypic observations; before using the DNA markers in molecular genotyping, in order to identify plants with higher genomic proportion of the recurrent parent, a phenotypic selection was performed, preserving the trait of the donor to be transferred to the recurrent parent. Consequently, the recovery of the recurrent genome in backcross cycles tends to be biased in favor of the donor parent. In the present study, the selection for hermaphroditism was performed in each one of the backcross generations in which the plants segregated for gender at a mean proportion of 1:1 (female plant: hermaphrodite plant). Apart from the selection for hermaphroditism, other phenotypic attributes were selected, which most likely favored the greater participation of the donor genome.

The genetic dissimilarity matrix also allowed an estimation of the distance between the recurrent parent and the BC_2 plants selected for hermaphroditism. A mean genetic similarity of the recurrent parent and the BC_2 generation of 75.7 % was verified, versus an expected 87.5 %.

Guimarães et al. (2006) report that in a study with soybean that targeted an increased protein content, 16 BC_1F_2 superior plants for this trait were selected, in which the mean genetic similarity with the recurrent

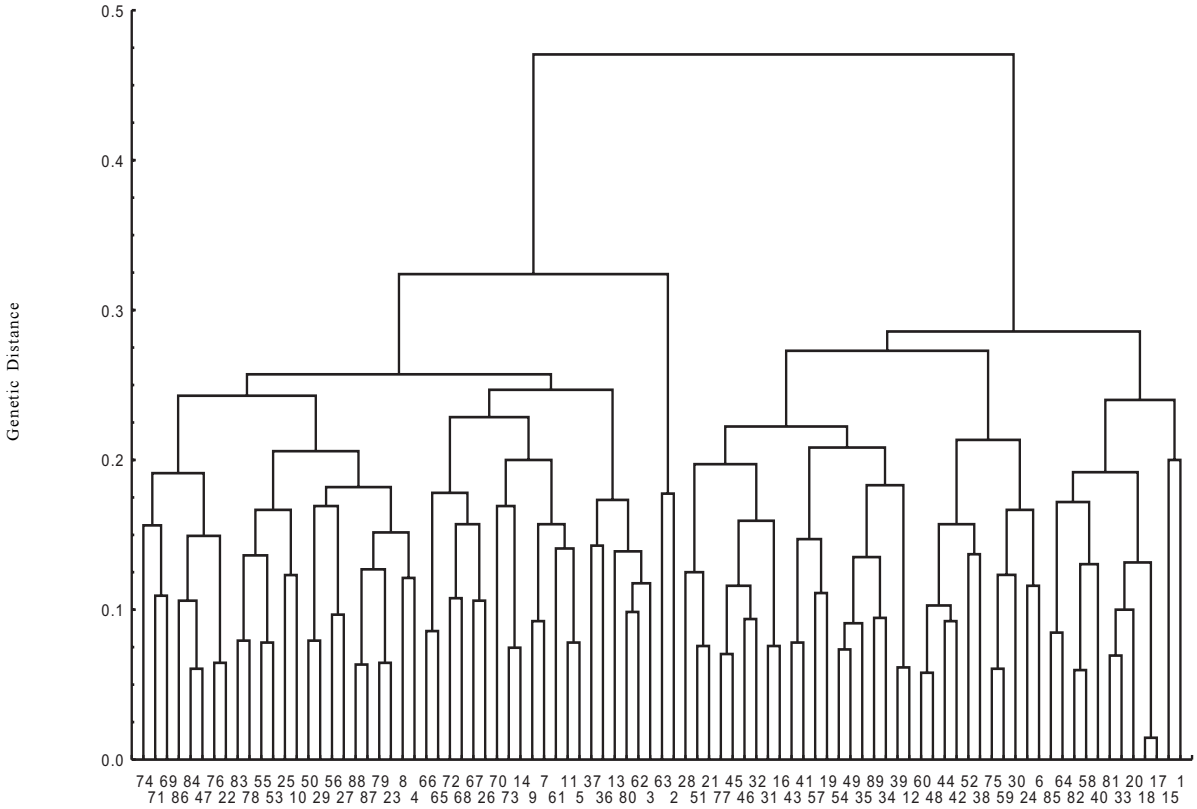

Figure 1. Dendrogram of the genetic dissimilarity of 87 RC_2 trees and of the recurrent (1) and donor parents (2), based on the genetic distance matrix generated by RAPD markers, obtained by the 'most distant neighbour' method

parent was 53.5 %, versus an expected 75 %. This marked difference was ascribed to the fact that the selected trait, as derived from the donor parent, is of quantitative heritability, involving several genes distributed across the genome. On the other hand, Mesquita et al. (2005), using 68 SSR markers in two assisted backcross cycles, obtained corn plants selected for low ear insertion (trait of the donor parent) with 98.2 % of mean genetic similarity with the recurrent parent. In a conventional program, this result could be expected in the fifth backcross cycle.

In the present study, the plant closest to the recurrent parent was the one labeled 15 (Figure 1). This one had however been discarded for phytosanitary reasons. The second and third in the hierarchy were the plants 17 and 18. Yet, within this group, the plants of best phenotypic performance were the numbers 18 and 89, mainly the latter. We therefore decided to use the plants 18 and 89 in recombination with 10 female plants of genotype 'Cariflora' to obtain the BC_3 generation and thus obtain sex-converted 'Cariflora', segregating in the proportion 1:1 (female and hermaphrodite), different from the original 'Cariflora' that segregates 1:1 for female and male plants. The reason for selecting plant 89 was basically to prioritize the transference of the allele to hermaphroditism, while preserving a maximum of the desirable phenotypic traits of a genetic improvement program of papaya, such as vigor, earliness, low insertion of the first fruit, reduced incidence of flower deformations (carpelloidy and pentandry), low incidence of female sterility or sex reversion, and yield potential. According to Vieira et al. (2005), the use of molecular markers in the study of genetic dissimilarity together with phenotypic information is important for the selection of genotypes and genetic mapping.

The genetic divergence matrix, based on the 54 RAPD markers, indicates a similarity between the recurrent parent and the plants represented by the numbers 15, 18 and 89 of 84 %, 81 % and 76 %,

respectively. The final decision on which one of the recombined generations ('Cariflora' x 18 or 'Cariflora' x 89) will be used as reference for future improvement studies depends on the phenotypic performance of both.

Marker-assisted selection has had a great impact on plant improvement, making the identification of superior genotypes in segregating populations possible and, above all, reducing the time for the release of new cultivars. In this context, the use of such markers in programs of gene introgression via backcross is maybe the most concrete application of molecular marker technique in improvement (Ferreira and Grattapaglia 1998).

Summing up, the use of molecular markers together with phenotypic evaluations was effective in the sex conversion of the dioecious genotype 'Cariflora' into gynoecious - andromonoecious, boosting the future establishment of papaya varieties and superior hybrids.

ACKNOWLEDGEMENTS

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for granting one of the authors a doctorate scholarship and the Financiadora de Estudos e Projetos (FINEP) and the Company Caliman Agrícola S/A for support with funds and data processing.

Conversão sexual em genótipo elite de mamoeiro (*Carica papaya* L.) assistida por marcadores de DNA

RESUMO – *A cultura do mamoeiro sustenta-se em uma estreita base genética, resultando em poucos cultivares geneticamente distintos para o plantio. Este trabalho visou à conversão sexual do genótipo 'Cariflora', do seu estado dióico para o ginóico-andromonóico, por meio da introgressão do alelo M^2 assistida pelos marcadores RAPD. O agrupamento hierárquico, na geração RC_2, indicou haver três plantas próximas ao genitor recorrente (15, 18 e 89), que apresentavam características fenotípicas desejáveis, e a matriz de divergência genética indicou que esta semelhança foi de 84, 81 e 76 %, respectivamente. A matriz de dissimilaridade genética indicou que a distância genética média entre o genitor recorrente e as plantas RC_2 foi de 75.7 %, portanto inferior ao esperado de 87.5 %. Contudo, a seleção fenotípica, aplicada nas populações RC_1 e RC_2 explica em parte esse desvio a favor do genitor doador.*

Palavras-chave: Mamoeiro, marcadores RAPD, retrocruzamento assistido, expressão sexual, Índice de Jaccard.

REFERENCES

Alfenas AC, Peters I, Brune W and Passador GC (1991) **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. Universidade Federal de Viçosa, Viçosa, 242p.

Conover RA, Litz RE, Malo SE (1986) Cariflora, a papaya for South Florida with tolerance to papaya ringspot virus. **Florida: Agricultural Experiment Station, University of Florida, (Circular, 329),** 3p.

Cruz CD (2001) **Programa Genes: versão windows - aplicativo computacional em genética e estatística.** Universidade Federal de Viçosa**,** Viçosa, 648p.

Cruz CD and Regazzi AJ (1997) **Modelos biométricos aplicados ao melhoramento genético**. 2nd ed, Universidade Federal de Viçosa, Viçosa, 340p.

Daher RF, Pereira MG, Tupinambá EA, Amaral Júnior AT, Aragão WM, Ribeiro FE, Oliveira LO and Sakiyama NS (2002) Assessment of coconut tree genetic divergence by compound sample RAPD marker analysis. **Crop Breeding and Applied Biotechnology 2**: 431-438.

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

Ferreira M E and Grattapaglia D (1998) **Introdução ao uso de marcadores moleculares em análise genética**. Embrapa – Cenargen, Brasília-DF. 220p.

Guimarães CT, Schuster I, Magalhães JV, Souza Júnior CL (2006) Marcadores moleculares no melhoramento de plantas. In: Borém A and Caixeta ET (eds.) **Marcadores moleculares**. Editora UFV, Viçosa, p. 107-144.

Hofmeyr JDJ (1938) Genetical studies of *Carica papaya* L. **African Department Agricultural For Science, Technical Bulletin 187**: 1-46.

Marin SLD, Pereira MG, Amaral Júnior AT, Martelleto LAP and Ide CD (2006) Heterosis in papaya hybrids from partial diallel

of 'Solo' and 'Formosa' parents. **Crop Breeding and Applied Biotechnology 6**: 24-29.

Mesquita AGG, Guimarães CT, Parentoni SN and Paiva E (2005) Recuperação do genitor recorrente em milho utilizando retrocruzamento assistido por marcadores microssatélites. In: Borém A e Caixeta ET (eds.) **Marcadores moleculares.** Universidade Federal de Viçosa, Viçosa, p.107-144.

Oliveira AMG, Farias ARN, Santos Filho HP, Oliveira JRP, Dantas JLL, Santos LB dos, Oliveira M de A, Souza Junior MT, Silva MJ, Almeida AO de, Nickel O, Medina VM and Cordeiro ZJM (1994) **Mamão para a exportação: aspectos técnicos da produção**. Brasília, DF: EMBRAPA-SPI, 52p. (FRUPEX. Série Publicações Técnicas, 9).

Silva MCA, Tarsitano MAN and Corrêa LS (2004) Análise do custo de produção e lucratividade do mamão Formosa Cultivado no município de Santa Fé do Sul (SP). **Revista Brasileira de Fruticultura 26**: 40-43.

Siqueira WJ, Fonseca MIS and Sondhal MR (1988) Regeneração de plantas híbridas entre *Lycopersicon esculentum* e *L. peruvianum* a partir de calos com dois anos de cultura *in vitro*. **Bragantia 47**: 1-8.

Storey WB (1938) The primary flower types of papaya and the fruit types that develop from them. **Proceedings of the American Society for Horticultural Science 35**: 83-85.

Storey WB (1941) The botany and sex relationships of the papaya. In: Papaya production in the Hawaiian Islands. **Hawaii Agricultural Experiment Station, Bulletin 87**: 5-22.

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

Vieira EA, Zimmer PD, Oliveira AC, Carvalho FIF, Malone G and Benin G (2005) Emprego de modelos gráficos na seleção de genitores de milho para hibridização e mapeamento genético. **Ciência Rural 35**: 986-994.

Williams JG, Kubelik AR, Livak KJ, Rafalsky LA and Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research 18**: 6531-6535.