Outcrossing rate between 'Haden' and 'Tommy Atkins' mangoes estimated using microsatellite and AFLP markers

Carlos Antonio Fernandes Santos[1] and Francisco Pinheiro Lima Neto[1]

[1]Embrapa Semiárido, BR 428, km 152, Caixa Postal 23, CEP 56302-970 Petrolina, PE, Brazil. E-mail: casantos@cpatsa.embrapa.br, pinheiro.neto@cpatsa.embrapa.br

Abstract – The objective of this work was to estimate outcrossing rates between Haden and Tommy Atkins mango cultivars, using AFLP and microsatellite markers. Progenies of an isolated 'Haden' plant, identified in a 'Tommy Atkins' commercial orchard, in Petrolina, PE, Brazil, were analyzed. Total DNA was isolated from the progeny leaves and used for AFLP and microsatellite reactions. Multilocus outcrossing rates (t_m) were estimated by direct count of AFLP or microsatellite markers and by the mLTR software. Outcrossing rates ranged from 0.85 to 0.87 with the analysis based on seven AFLP markers, and from 0.83 to 0.91 based on three microsatellite primers. No unexpected band patterns were observed for 'Haden' and 'Tommy Atkins'. The estimates obtained with the mLTR software were close to those obtained by direct AFLP and microsatellite allele counting, which indicates that the multilocus model was appropriate for this kind of study. The microsatellites mMiCIR005, mMiCIR030, and mMiCIR036 can be used to elucidate the origin of 'Haden' and 'Tommy Atkins' seedlings.

Index terms: *Mangifera indica*, mLTR, multilocus estimation, SSR.

Taxa de cruzamento entre manga 'Haden' e 'Tommy Atkins' estimada com marcadores microssatélites e AFLP

Resumo – O objetivo deste trabalho foi estimar a taxa de polinização cruzada entre as cultivares de manga Haden e Tommy Atkins, por meio do uso de marcadores AFLP e microssatélites. Progênies de uma planta isolada de 'Haden', identificada em plantio comercial de 'Tommy Atkins', em Petrolina, PE, foram analisadas. O DNA total foi isolado das folhas das progênies e utilizado em reações com marcadores AFLP e microssatélites. Estimativas multilocos (t_m) foram obtidas pela contagem direta de marcas AFLP ou de microssatélites e pelo programa mLTR. A taxa de polinização cruzada variou de 0,85 a 0,87, com análise baseada em sete marcas AFLP, e de 0,83 a 0,91, com base em três microssatélites. Não foram observadas bandas diferentes das esperadas para 'Haden' e 'Tommy Atkins'. As estimativas obtidas com o programa mLTR foram próximas às obtidas com a contagem direta de marcas AFLP e de microssatélites, o que indica que o modelo multilocos foi apropriado para este tipo de trabalho. Os microssatélites mMiCIR005, mMiCIR030 e mMiCIR036 são recomendados para elucidar a origem de mudas de 'Haden' e 'Tommy Atkins'.

Termos para indexação: *Mangifera indica*, mLTR, estimativa multilocos, SSR.

Introduction

Mango (*Mangifera indica* L.) is one of the most important fruit crops in the Brazilian semiarid region, in which commercial orchards are based on Floridian varieties, and 'Tommy Atkins', is cultivated in 85% of an estimated area of 25,000 ha. However, this situation could have a negative impact on the local mango agribusiness, considering changes in overseas consumer preferences or a severe infestation of a specific pest or disease, highlighting the importance of the development of new mango cultivars.

Mango tree inflorescences have male and hermaphrodite flowers (2:1), with a predominance of the first ones on the basis of the panicles. The anthesis is diurnal, asynchronic, and releases a strong and sweet odor. The flowers show dichogamy, characterized by the anthers dehiscence 24 hours after the anthesis. According to Jison & Hedstron (1985), the mango flower structure is totally adapted to cross pollination, and *Apis mellifera*, *Belvosia bicincta*, and *Musca domestica* are the most frequent insect visitors to 'Tommy Atkins' in the São Francisco valley (Siqueira et al., 2008).

Understanding the mating system of the species is fundamental for genetic improvement and conservation programs, since it permits the outlining of strategies that optimize the sampling of genetic variability and

the adoption of genetic-statistical models appropriate for the estimation of genetic parameters. Therefore, it is important to establish strategies for the effective conservation of any species. The mating system, together with the mechanism of pollen and seed dispersion, determines part of the genetic structure of populations (Freitas et al., 2004).

There are few known studies on the mango mating system in spite of its importance for genetic studies and of the readiness provided by the use of DNA molecular markers. Degani et al. (1997), in a pioneer and isolated study, estimated the outcrossing rate between 'Maya' and 'Tommy Atkins' mangoes, using isozymes. For Pinto et al. (2002), mango pollinization is critical for monoembryonic or Indian cultivars, since they are the basis of commercial orchards in Brazil, unlike polyembryonic or Filipino cultivars, which produce non-zygotic or maternal embryos.

Dominant DNA markers, such as RAPD and AFLP, have been used to estimate outcrossing rates in *Moringa oleifera* (Muluvi et al., 2004) and *Eucalyptus urophylla* (Gaiotto et al., 1997), despite limitations on the correct genotypic identification of individuals. To overcome this limitation, Ritland (2002) developed the multilocus model to estimate outcrossing rates in plant species, so that dominant markers could be applied in mating system studies. Co-dominant markers, such as microsatellites, do not have this limitation and, therefore, have also been applied in mating system studies, as demonstrated by Borrone et al. (2008) in avocado (*Persea Americana* Mill.).

The objective of this work was to estimate outcrossing rates between Haden and Tommy Atkins mango cultivars, using AFLP and microsatellite markers.

Materials and Methods

A 'Haden' tree isolate was identified in a commercial orchard of 'Tommy Atkins' at a mango-exporting farm, in Petrolina, PE, Brazil. Fruits were harvested from the 'Haden' tree, that was at a minimum distance of 200 m from cultivars other than Tommy Atkins, to avoid crosses different from 'Haden' x 'Tommy Atkins'. 'Haden' progenies were established in field when six months old, with 4x4 m spacing, at the Embrapa experimental station of Mandacaru, BA, Brazil. A dripping irrigation system was used, and fertilizer was applied according to soil analysis.

Genomic DNA was extracted from leaves based on the Doyle & Doyle (1990) protocol, with the following adaptations: 7,500 and 16,500 *g* in the first and second centrifugations, respectively, beta-mercaptoethanol at 2%, and incubation at 60°C for 30 min for all samples, with occasional mixing by gentle tube inversion. After adding Tris-EDTA, the DNA solution was subjected to RNAse treatment to remove co-isolated RNAs. The DNA was quantified and the integrity analyzed in agarose 0.8%, followed by DNA dilution to 40 ng µL^{-1}.

Approximately 200 ng of genomic DNA of each progeny was doubled digested with endonucleases enzymes *EcoRI* and *MseI*, for 2.5 hours. Digested DNA was ligated to specific adaptors with T4 DNA ligase to a final volume of 7.25 µL. Pre-amplification reactions were carried out to a final volume of 15 µL (1.5 µmol L^{-1} of each *EcoRI* and *MseI* primers, 0.2 mmol L^{-1} of each dNTPs, 1x PCR buffer, 2.5 mmol L^{-1} of MgCl$_2$, 1.5 unit of *Taq* DNA Polimerase, 2.0 µL diluted 5-fold ligated DNA solution), with thermocycling to 20 cycles: 94°C for 30 s, 56°C for 1 min, and 72°C for 1 min. After pre-amplification, samples were diluted 20-fold in TE buffer. Selective amplifications were accomplished to a final volume of 10 µL (0.2 µmol L^{-1} of *EcoRI* primer, 0.3 µmol L^{-1} of *MseI* primer, 0.2 mmol L^{-1} dNTPs, 1x PCR buffer (100 mmol L^{-1} Tris-HCl (pH 8.3), 500 mmol L^{-1} of KCl), 2.5 mmol L^{-1} of MgCl$_2$, 0.75 unit of *Taq* DNA polymerase, and 2 µL pre-amplified DNA. Touchdown thermocycling conditions were: one cycle at 94°C for 30 s, 65°C for 30 s, and 72°C for 60 s; 13 cycles starting with an annealing temperature of 65°C, lowered by 0.7°C every subsequent cycle; 23 cycles at 94°C for 30 s, 56°C for 30 s, and 72°C for 60 s. Then, a reaction of 2.0 µL formamide dye (98% formamide, 10 mmol L^{-1} EDTA, 10 mg of each bromophenol blue and xylene cyanol) was added. Reactions were heated for 3 min at 90°C and immediately placed on ice for further polyacrylamide denaturing gels load. The gels were stained with silver nitrate according to Creste et al. (2001).

An initial screening with 'Haden' and 'Tommy Atkins' parents was done with 30 *M. indica* microsatellites published by Duval et al. (2005). PCR reactions were carried out at a final volume of 20 µL (30 ng genomic DNA, 1x PCR *Taq* DNA polymerase buffer, 2.0 mmol L^{-1} of MgCl$_2$, 0.2 mmol L^{-1} dNTPs, 0.2 µmol L^{-1} of each primer, 0.5 units of *Taq* DNA

polymerase). The PCR amplification program was used for initial denaturation at 94°C for 4 min, followed by 32 cycles of 94°C for 45 s, 55°C for 60 s, 72°C for 60 s, and a final step of 72°C for 4 min. Microsatellites with one allele in 'Tommy Atkins' and absent in the 'Haden' parent were selected to genotype 'Haden' progenies. After PCR amplification, AFLP and microsatellite fragments were visualized in 6% polyacrylamide gels and stained with silver nitrate, according to Creste et al. (2001).

AFLP and microsatellite analysis were carried out on 108 and 94 individuals, respectively. AFLP polymorphic bands were scored as 1 for presence and as 0 for absence, and the chi-square test for 1:1 and 3:1 segregation was carried out. Two outcrossing estimates were obtained by direct allele counting, in which 'Haden' progenies with at least two AFLP alleles or one microsatellite allele from 'Tommy Atkins' were assumed to be a crossing between the two analyzed parents; and by multi-locus estimation, in which single-locus (t_s) and multilocus (t_m) outcrossing rates were estimated according to the mixed-mating model (mixed outcrossing and selfing), using the mLTR software (Ritland, 2002). This model specifies that both selfing and outcrossing occur in the population. One hundred bootstrap resampling families were used in mLTR analysis. Maximum-likelihood estimates were obtained for all AFLP marker loci used in the estimation of outcrossing rates. The allelic genotype for the maternal plant 'Haden' was specified in microsatellite analysis.

Results and Discussion

Seven easily scored AFLP bands present in the pollinator 'Tommy Atkins' and absent in the maternal tree 'Haden' were identified (Figure 1). In addition, seven polymorphic bands were observed per AFLP primer combination.

Five AFLP loci present in 'Tommy Atkins' and absent in 'Haden' had a ratio of 1:1, while the other two had a segregation ratio of 3:1 in the 'Haden' progenies (Table 1). A 1:1 segregation was expected, since 'Haden' was the maternal progenitor of 'Tommy Atkins' (Schnell et al., 2006), establishing a backcross configuration. The observed segregation is an indication of the high heterozygosity generally found in fruit trees, as reported by Weeden et al. (1994).

The estimated direct outcrossing rate was 0.87 in the 108 'Haden' progenies evaluated with seven AFLP loci, considering that 94 progenies had at least two bands from the 'Tommy Atkins' pollinator, while ten progenies had only one band, and six others did not have any band from 'Tommy Atkins'. A greater number of AFLP loci would be necessary to declare the progenies that had only one band from 'Tommy Atkins' as true hybrids between the two analyzed mango cultivars.

The multilocus outcrossing rate (t_m) estimated by the mLTR software was 0.85, and the single-locus (t_s) estimate was 0.95, indicating the predominance of the outcrossing system in the species, with selfing of 0.15. The 0.85 multilocus estimate was very close to the estimate obtained with the presence of at least two bands from the putative pollinator.
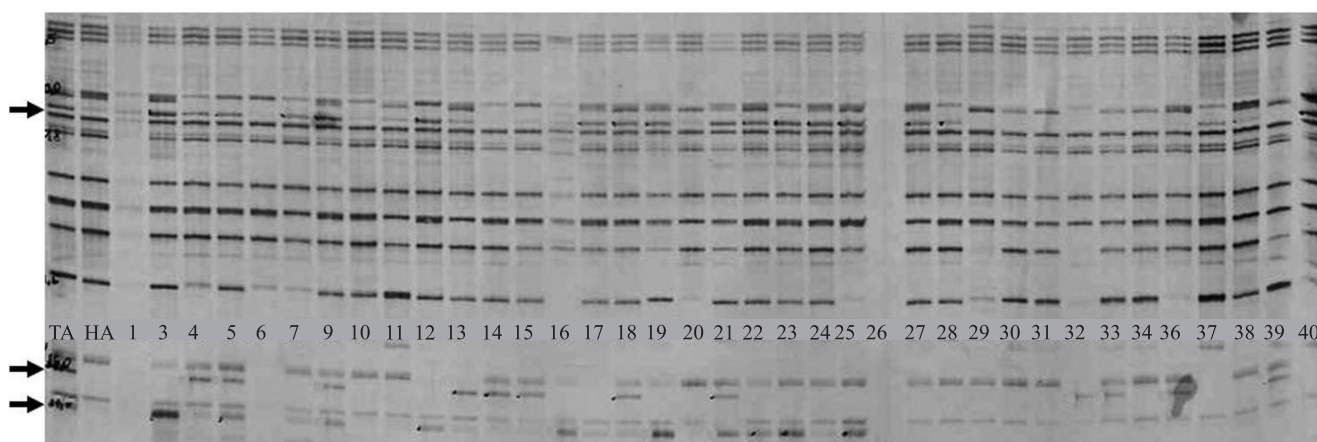
Figure 1. AFLP band pattern of the AFLP primer combination E-ACC/M-CTA for 37 progenies of 'Haden'. TA, 'Tommy Atkins'; HA, 'Haden'. Lanes 1 to 40 were 'Haden' progenies, with 'Tommy Atkins' as putative pollinator.

The microsatellites mMiCIR005, mMiCIR030, and mMiCIR036 were selected to genotype the 94 'Haden' x 'Tommy Atkins' progenies, since the presence and absence of an allele was observed in 'Tommy Atkins' and 'Haden', respectively (Figure 2). One allele of 'Haden' from the microsatellite loci was observed in 'Tommy Atkins', as 'Haden' was the maternal tree of 'Tommy Atkins' (Schnell et al., 2006). The microsatellite mMiCIR036 showed typical backcross situation, in which the maternal 'Haden' was homozygote and its descendent, 'Tommy Atkins', was heterozygote. The other two microsatellites were heterozygote in both parents, validating the general high heterozygosity expected in fruit trees.

Seventy-eight out of ninety-four 'Haden' progenies analyzed with the three microsatellites were true

hybrids of 'Haden' x 'Tommy Atkins', which was confirmed by the presence of one microsatellite of 'Tommy Atkins' in the 'Haden' progeny (Figure 2). This number corresponds to a direct outcrossing rate of 0.83. No unexpected profile was observed in the analyzed 'Haden' progenies, indicating that only alleles of 'Tommy Atkins' were present or absent in the progenies.

The three microsatellites used in the present study were able to identify 91% of true hybrids, since individuals, such as 1, 10, 12, 17, and 28 (Figure 2), were not declared as selfed 'Haden' progenies, which indicates that a greater number of microsatellites should be used to confirm this hypothesis. However, three individuals were identified as selfed 'Haden' progenies, because they only had alleles of the maternal 'Haden', as observed for individual 40.

The estimate for t_m obtained with the mLTR software was 0.91, while t_s estimate was 0.89 ($\pm$0.001), indicating a predominance of the outcrossing rate in *M. indica*, with 0.09 of selfing. This estimate was superior to the direct estimate, considering the presence of at least one allele from 'Tommy Atkins' in the 'Haden' progenies, probably due to the inclusion of individuals considered doubtful in the previous allelic analysis.

Multilocus outcrossing rates estimated in the present study, ranging from 85 to 87% with seven AFLP markers and from 83 to 91% with three microsatellite primers, were close to those estimated by Degani et al. (1997), of 0.79 between 'Maya' and 'Tommy Atkins'. A segregation rate of 3:1 was observed with AFLP markers and two microsatellite markers as heterozygotes in both parents, corroborating the

Table 1. Number of bands present (1) and absent (0), PCR failed reaction, and chi-square test for segregation 1:1 and 3:1 to seven AFLP bands scored in 108 individuals, with 'Haden' as maternal progenitor and 'Tommy Atkins' as putative pollinator.

AFLP band	Number of individuals		Failed PCR	Chi-square	
	1	0		1:1	3:1
E-ACC/ M-CTA-1	28	75	5	21.4**	0.2[ns]
E-ACC/ M-CTA-2	62	41	5	4.2[ns]	68.0**
E-ACC/ M-CTA-3	51	52	5	0.1[ns]	33.0**
E-ACA/M-CTA-1	33	75	0	16.3**	1.7[ns]
E-ACG/M-CTG-1	55	47	6	0.6[ns]	45.5**
E-ACT/ E-CAC-1	61	37	10	5.8[ns]	72.5**
E-ACT/ E-CAC-2	43	55	10	1.4[ns]	18.6**

[ns]Nonsignificant. **Significant at 5% probability by the chi-square test.
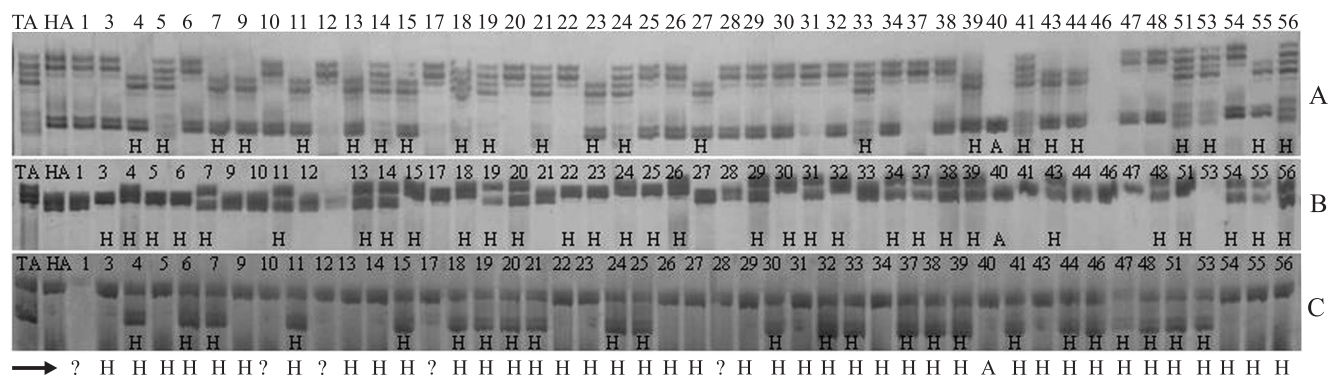
Figure 2. Polyacrylamide 6% gel with 46 individuals of the putative cross between 'Haden' (HA) and 'Tommy Atkins' (TA) genotyped with mMiCIR005 (panel A), mMiCIR030 (panel B), and mMiCIR036 (panel C) microsatellites. H, hybrid; A, selfing; ?, could be hybrid or selfing.

high heterozygosity expected for mango. This high heterozygosity in perennial fruit trees has been used by plant breeders to develop crossing strategies, known as double pseudo-testcross, in order to map genes in these species (Weeden et al., 1994).

According to Allard (1999), multilocus estimation of outcrossing rates, as those obtained in the present study, are more reliable, since each progeny can be classified for many loci, with one gamete coming from the maternal plant and the second gamete being a product of selfing or of another pollinator parent. As reported by Muchugi et al. (2008), the predominance of cross pollinization in tropical trees is a reasonable explanation for the high genetic variability that has been observed in these species. Borrone et al. (2008), for example, estimated outcrossing rates of 0.74 and 0.96 for Simmonds and Tonnage avocado cultivars, respectively, based on eight microsatellite markers.

The estimates obtained with the multilocus model, available in the mLTR software (Ritland, 2002), were very close to those obtained by direct AFLP and microsatellite allele counting in the 108 and 94 'Haden' progenies, indicating that the multilocus model is very reliable for mating studies. This model has been successfully applied by different authors (Gaiotto et al., 1997; Muluvi et al., 2004; Muchugi et al., 2008) to species without available co-dominant markers.

The outcrossing estimates obtained in the present study and by Degani et al. (1997) indicate that mango is a predominantly open pollinated species and that biometrical models applied to mangoes should take into account the deviation from random outcrossing to estimate genetic parameters and the constitution of broad germplasm samples to preserve the genetic variability of the species, as recommended by Freitas et al. (2004), while evaluating *Myracrodruon urundeuva*.

The minimum distance of 200 m in relation to the other trees, except 'Tommy Atkins', in the commercial orchard, prevented crossings other than 'Haden' x 'Tommy Atkins'. Therefore, no strange or foreign alleles were observed by AFLP (Figure 1) and microsatellite profiles (Figure 2). Degani et al. (1997) and Dag et al. (2009), who evaluated pollen dispersion on mango as far as 108 m in Israel, found a substantial reduction in the outcrossing rate between orchards of 'Maya' and 'Tommy Atkins'.

For a typical backcross scheme, in which one individual always has one allele from the mother, dominant markers are more appropriate than co-dominant markers, since a high number of polymorphic bands can be obtained, so that true hybrids can be quickly identified.

Conclusions

1. The estimates obtained with the mLTR software are very close to those obtained by direct AFLP and microsatellite allele counting.

2. The microsatellites mMiCIR005, mMiCIR030, and mMiCIR036 are able to identify the origin of 'Haden' and 'Tommy Atkins' seedlings.

References

ALLARD, R.W. **Principles of plant breeding**. 2nd ed. New York: Wiley, 1999. 254p.

BORRONE, J.W.; OLANO, C.T.; KUHN, D.N.; BROWN, J.S.; SCHNELL, R.J.; VIOLI, H.A. Outcrossing in Florida avocados as measured using microsatellite markers. **Journal of the American Society of Horticulture Science**, v.133, p.255-261, 2008.

CRESTE, S.; TULMANN NETO, A.; FIGUEIRA, A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. **Plant Molecular Biology Reporter**, v.19, p.299-306, 2001.

DAG, A.; DEGANI, C.; GAZIT, S. Gene flow in mango orchards and its impact on yield. **Acta Horticultarae**, v.820, p.347-350, 2009.

DEGANI, C.; YUTKO, O.; EL-BATSRI, R.; GAZIT, S. Outcrossing rate in adjacent 'Maya' and 'Tommy Atkins' mango blocks. **Scientia Horticulturae**, v.70, p.25-30, 1997.

DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, v.12, p.3-15, 1990.

DUVAL, M.F.; BUNEL, J.; SITBON, C.; RISTERUCCI, A.M. Development of microsatellite markers for mango (*Mangifera indica* L.). **Molecular Ecology Notes**, v.5, p.824-826, 2005.

FREITAS, M.L.M.; SEBBENN, A.M.; MORAES, M.L.T.; LEMOS, E.G.M. Mating system of a population of *Myracrodruon urundeuva* F.F. & M.F. **Genetics and Molecular Biology**, v.27, p.425-431, 2004.

GAIOTTO, F.A.; BRAMUCCI, M.; GRATTAPAGLIA, D. Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. **Theoretical and Applied Genetics**, v.95, p.842-849, 1997.

JISON, L.F.; HEDSTROM, I. Pollination ecology of mango (*Mangifera Indica* L.) (Anacardiaceae) in the Neotropic region. **Turrialba**, v.35, p.269-277, 1985.

MUCHUGI, A.; MULUVI, G.M.; SIMONS, A.J.; WACHIRA, F.N.; JAMNADASS, R.H. Estimation of out-crossing rate in a

natural breeding population of *Warburgia ugandensis* using AFLP marker. **African Journal of Biotechnology**, v.7, p.139-146, 2008.

MULUVI, G.M.; SPRENT, J.I.; ODEE, D.; POWELL, W. Estimates of outcrossing rates in *Moringa oleifera* using Amplified fragment length polymorphism (AFLP). **African Journal of Biotechnology**, v.3, p.146-151, 2004.

PINTO, A.C. de Q.; SOUZA, V.A.B. de; ROSSETTO, C.J.; FERREIRA, F.R.; COSTA, J.G. da. Melhoramento genético. In: GENU, P.J. de C.; PINTO, A.C. de Q. (Ed.). A **cultura da mangueira**. Brasília: Embrapa Informação Tecnológica, 2002. p.53-92.

RITLAND, K. Extensions of models for the estimation of mating systems using n independent loci. **Heredity**, v.88, p.221-228, 2002.

SCHNELL, R.J.; BROWN, J.S.; OLANO, C.T.; MEEROW, A.W.; CAMPBELL, R.J.; KUHN, D.N. Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite markers. **Journal of the American Society for Horticultural Science**, v.131, p.214-224, 2006.

SIQUEIRA, K.M.M. de; KIILL, L.H.P.; MARTINS, C.F.; LEMOS, I.B.; MONTEIRO, S.P.; FEITOZA, E. de A. Estudo comparativo da polinização de *Mangifera indica* L. em cultivo convencional e orgânico na região do Vale do Submédio do São Francisco. **Revista Brasileira de Fruticultura**, v.30, p.303-310, 2008.

WEEDEN, N.F.; HEMMAT, M.; LAWSON, D.M.; LODHI, M.; BELL, R.L.; MANGANARIS, A.G.; BROWN, S.K.; YE, G.N. Development and application of molecular marker linkage maps in woody fruit crops. **Euphytica**, v.77, p.71-75, 1994.