Genetic Resources and Crop Evolution **51:** 871–882, 2004.

Molecular marker-mediated validation of morphologically defined landraces of Pejibaye (*Bactris gasipaes*) and their phylogenetic relationships

Doriane P. Rodrigues[1,*], Spartaco Astolfi Filho[1] and Charles R. Clement[2]
[1]*Universidade Federal do Amazonas, Instituto de Ciências Biológicas, Av. Gal. Rodrigo Otávio Jordão Ramos 3000, 69077-000 Manaus, AM, Brazil;* [2]*Instituto Nacional de Pesquisas da Amazônia, Cx. Postal 478, 69011-970 Manaus, AM, Brazil; *Author for correspondence*

Received 9 December 2002; accepted in revised form 17 May 2003

Key words: Bactris gasipaes, Crop origin, Landraces, Phylogenetic analysis, RAPD

Abstract

RAPD markers were used to evaluate the genetic variability and structure of seven morphologically defined landraces of pejibaye (*Bactris gasipaes* Kunth, Palmae) to determine their validity and phylogenetic relationships. Two hundred and twenty plants of four Amazonian and three Central American landraces of var. *gasipaes* (the domesticate) and 30 plants of var. *chichagui* (H. Karsten) Henderson (the crop ancestor) maintained at the National Research Institute for Amazonia, Manaus, Amazonas, Brazil, were utilized. Eight RAPD primers yielded 113 markers, with good reproducibility, of which 97 were polymorphic. The four Amazonian landraces had an average heterozygosity of 0.30, with 86% polymorphism, greater than the Central American landraces (0.25; 74.3%) and var. *chichagui* (0.27; 80%). Among landrace genetic diversity (G_{ST}) was 15%, while within (Hs) was 85%, essentially equivalent to the AMOVA within (82.2%) and among (17.8%) variances. The Jaccard similarities, PCA, gene flow coefficients and Exact tests suggested that only one landrace exists in Central America, called Utilis after the first taxon described there, and that the Solimões landrace is part of the Putumayo landrace, rather than a separate entity. The Pará and Pampa Hermosa landraces were validated in accordance with their morphometric interpretations. The dendrogram of Nei's genetic distances among valid landraces and var. *chichagui* supported the hypothesis of a single origin for pejibaye in southwestern Amazonia, with two migration routes: one to the northeast, becoming the Pará landrace, and another to the northwest along the Andes, spreading into western Amazonia (Pampa Hermosa and Putumayo landraces) and across the Andes, reaching Central America (Utilis landrace).

Introduction

The pejibaye or peach palm (*Bactris gasipaes* Kunth, Palmae) is widely distributed in the lowland Neotropics (Mora Urpí et al. 1997), and contains ample genetic diversity in its wild and cultivated populations, due to their different degrees of domestication in different environments. The fruit is a starchy drupe selected for size, color, starch content and form, and is the most important organ for determining degree of domestication. A complex hierarchy of landraces was proposed on morphometric grounds (Mora Urpí 1984; Mora Urpí and Clement 1988; Mora Urpí 1992; Clement 1995). It is important to know if the landraces proposed morphometrically exist genetically, since this will enhance the effectiveness of improvement efforts.

The cultivated populations of pejibaye are grouped into landraces within *B. gasipaes* var. *gasipaes*, while the wild populations, with no characteristics suggesting even incipient domestication, are all in *B. gasipaes* var. *chichagui* (Henderson 2001). *In situ* and *ex situ* morphometric characterization

allowed the classification of the Amazonian landraces (Clement 1986; Mora Urpí and Clement 1988), while Mora Urpí (1992) proposed the Central American landraces. The landrace hierarchy has a primary geographic division defined by the Andes, with an Oriental group in lowland northern South America and an Ocidental group in lowland northwestern South America northwards into Central America. Within these groups, fruit size is the primary factor to distinguish landraces into microcarpa, mesocarpa and macrocarpa, since this trait was most modified by human selection. In the Oriental group, 2 microcarpa, 5 mesocarpa and 2 macrocarpa landraces have been mapped to date, while in the Ociental group, 1 microcarpa and 4 mesocarpa landraces have been mapped (Mora Urpí et al. 1997). These landraces are composed of a variable number of closely related domesticated populations, defined by a specific combination of morphological characteristics, a restricted geographic distribution and probably a distinct ethnic history. This definition of landrace is somewhat different from that of Zeven (1998) and closer to that of Louette (2000).

Doubts have existed about the validity of one Amazonian landrace, the Solimões, since the original morphometric analysis and a recent genetic analysis, using RAPD markers, showed it to be mostly a part of the Putumayo landrace (Sousa et al. 2001). Clement (1986) suggested that the three Central America populations were very similar, although Mora Urpí (1992) later classified them as landraces, so their validity as landraces must be tested.

The origin of pejibaye has been debated extensively and inconclusively for decades (Mora Urpí et al. 1997). The consolidation of the closely related wild taxa into var. *chichagui* simplified the debate by providing a single ancestral taxon. However, most domesticates are derived from single domestication events from specific ancestral populations (Blumler 1992), so the new *Bactris* classification has not resolved the origin debate. Essentially, two schools of thought exist: a single domestication event, most probably in southwestern Amazonia (argued most recently by Clement 1995), although Morcote-Rios and Bernal (2001) have recently argued for the northern Andes; multiple domestication events throughout the distribution of the wild taxa (var. *chichagui*) (argued most recently by Mora Urpí 1999). Molecular techniques can often assist in the phylogenetic analysis that can resolve this type of debate (Matioli 2001).

The modern economic potential of pejibaye is centered on its heart-of-palm, but also includes fruit for direct human consumption as a cooked fruit (its traditional and major modern use), flour for baking (a traditional use with modern potential), animal feed and vegetable oil (Clement and Mora Urpí 1987; Mora Urpí et al. 1997). This repertoire of potential uses can best be transformed into market demand if pejibaye is improved for each specific use. Hence the demand for genetic analysis of landraces, as specific landraces can serve as the basis for improvement for specific uses (Clement and Mora Urpí 1987). This study uses RAPD markers to evaluate the genetic variability and structure of seven morphologically defined landraces of pejibaye to determine their validity and phylogenetic relationships.

Material and methods

Two hundred and twenty plants of seven morphologically defined landraces of pejibaye [Amazonia – Putumayo (33 plants), Solimões (30), Pará (40), Pampa Hermosa (30); Central America – Tuíra (30), Guatuso (27) and Utilis (30)], 30 plants of two wild populations [Rio Branco, Acre (15); Benjamin Constant, Amazonas (15); used as outgroups], and a single plant of the Juruá landrace (used as control in the gels), all maintained in the Pejibaye Active Bank Germplasm at the National Research Institute for Amazonia (INPA), BR 174, km 38, Manaus, AM, Brazil, were used. Accession and plant numbers available on request. An accession contains nine progeny obtained from a single plant *in situ*, so that these samples represent the landraces at the time of sampling. The accessions were selected to provide good geographic coverage of each landrace, while the plants were randomly sampled from within the accessions.

DNA was extracted with the DNAsy Plant Minikit (Quiagen) from 100 mg of apical meristem of a lateral sucker (Clement et al. 1997), and quantified by comparison with ethidium bromide-stained standard concentrations in 0.9% agarose gels. When suckers were not present, 200 mg of juvenile leaf tissue was used, with extraction following maceration in liquid nitrogen (Weising et al. 1995).

The RAPD fragments were generated following Williams et al. (1990), with minor modifications. Each amplification reaction contained 10 ng of genomic DNA (5 ng/μL), 250 μM of dNTP (2.5 mM), 3 mM of $MgCl_2$ (25 mM), 50 ng of primer (10 ng/μL), 1.5 U of Taq polymerase (CENBIOT/RS), 3 μL of buffer 1$\times$ (200 mM Tris-HCl pH = 8.6; KCl 500 mM) and was completed to a final volume of 30 μL with distilled water.

Three DNA samples from different landraces were used to select among 30 primers [Operon Technologies (Opa); Biosystems (F)]. Eleven primers amplified numerous markers; of these, eight were selected for band resolution and presence of polymorphisms. RAPD reactions were performed with two programs in a Perkin Elmer 9600 thermocycler programmed in accordance with primer resolution after testing aneling temperatures of 40, 45 and 50 °C. Program 1 (primer F-919-3): 1 cycle of 2 min at 94 °C, 40 cycles of 1 min at 92 °C, 1 min at 36 °C, 2 min at 72 °C, plus a final cycle of 3 min at 72 °C. Program 2 (primers Opa-4, Opa-5, Opa-8, Opa-9, Opa-18, Opa-20, FC13): 2 cycles of 1 min at 94 °C, 1 min at 36 °C, 2 min at 72 °C, and 33 cycles of 10 s at 94 °C, 20 s at 40 °C, 2 min at 72 °C, plus a final cycle of 5 min at 72 °C.

After amplification, the products were separated by electrophoresis in 1.5% agarose gels. Permanent records were obtained by photographing the ethidium bromide stained gels under UV light. Band sizes were determined by comparison with a 1 Kb ladder in each gel.

Bands in the gels were classified as intense (1), moderate (2), weak (3) or absent (0), based on visual evaluation of resolution (Grattapaglia 1997). To evaluate the reliability of these interpretations, the Jaccard similarities of the control plant repetitions in the gels were estimated for the possible combinations of intensity: 1 *versus* 2 = 3 = 0; 1 = 2 *versus* 3 = 0; 1 = 2 = 3 *versus* 0. The combination with the greatest similarity among the control repetitions was used (Grattapaglia 1997).

The binary matrix was used to estimate each landrace's heterozygosity [assuming absence as recessive (Weir 1996)] using Nei's (1972) criterion and the percentage of polymorphism, using the TFPGA program (Miller 1997). The within and among landrace variances were estimated by AMOVA, using the WINAMOVA v. 1.55 program (Excoffier et al. 1992). The within and among genetic diversities (H_T, H_S, G_{ST}) was estimated using the Hartl and Clark (1989) criterion, and among landrace gene flow was estimated as $N_m = 0.5(1 - G_{ST})/G_{ST}$ (Slatkin and Barton 1989), using the POPGEN v. 1.31 program (Yeh et al. 1999). An among plant Jaccard similarity matrix was generated and a dendrogram using UPGMA was created, using the NTSys-PC program (Rohlf 1990); also used for the principal components analysis (PCA). The genetic distances among landraces were estimated following Nei (1972) and a dendrogram was generated with UPGMA, using the TFPGA program (Miller 1997). The Exact Test was used (Raymond and Rousset 1995) to determine the existence of significant differences in allelic frequencies between landraces, using the TFPGA program (Miller 1997). The correlation between the gene flow and geographic distance matrices (with and without var. *chichagui*) was estimated according to Mantel (1967), using the TFPGA program (Miller 1997) with 999 permutations. The latter analyses were repeated after preliminary validation of the landraces.

Results and discussion

Genetic analysis of the morphologically defined landraces

The 8 primers used amplified a total of 113 useful markers (bands), of which 97 were polymorphic, with an average of 14.1 bands per primer. These were reasonably trustworthy for this set of plants, since the Jaccard similarities of the control plant averaged 0.953 ($\pm$0.023; maximum = 1.0; minimum = 0.889), when using bands of all intensities together (1 = 2 = 3 *versus* 0).

The overall mean estimated heterozygosity was 0.31, with 89.4% polymorphism (Table 1). The Amazonian landraces had a higher mean (0.30) than the Central American landraces (0.25) and var. *chichagui* (0.27), with 86%, 74% and 80.5% polymorphism, respectively. Estimated heterozygosities based on allozymes (Clement et al. 1997) were much lower in those populations and landraces that can be compared: Benjamin Constant, Amazonas, population (Putumayo landrace) – heterozygosity = 0.066 with allozymes *versus* 0.27

Table 1. Heterozygosity estimates (95% and 99%) and percentage of polymorphism of pejibaye (*Bactris gasipaes*) landraces in Amazonia and Central America, and of populations of var. *chichagui* (Benjamin Constant and Acre) in Amazonia obtained from 113 RAPD markers (97 of which polymorphic).

			% Polymorphism	
Variety, Region, Landrace	Sample (*n*)	Heterozygosity	95%	99%
var. *chichagui*	30	0.27	74.3	80.5
– Acre[1]	15	0.22	60.2	67.2
– Benjamin Constant[2]	15	0.22	59.3	68.1
var. *gasipaes* – Amazonia	133	0.30	83.0	86.0
– Parú	40	0.24	66.4	75.2
– Solimões	30	0.30	76.1	82.3
– Putumayo	33	0.27	73.4	77.0
– Pampa Hermosa	30	0.26	72.6	75.2
var. *gasipaes* – Central America	87	0.25	66.4	74.3
– Tuíra	30	0.22	62.0	64.6
– Utilis	30	0.24	62.8	64.6
– Guatuso	27	0.23	63.8	67.2
Overall	250	0.31	84.9	89.4

[1]See Clement et al. (1989) for details.
[2]See Clement et al. (1999) for details.

with RAPDs; San Carlos, Costa Rica, population (Guatuso landrace) – 0.051 *versus* 0.23; and Yurimáguas, Peru, population (Pampa Hermosa landrace) – 0.141 *versus* 0.26. These differences are partially due to the lower allozyme polymorphism (56, 44 and 69%, respectively), greater intensity of recent selection in the populations studied with allozymes (at least 2 cycles (Clement et al. 1997) *versus* none here), and the greater number of markers that do not possess important physiological functions (Buso et al. 1998). The high RAPD polymorphism better reflects the great genetic variability observed in pejibaye's morphological characteristics (Mora Urpí 1991) than does the allozyme polymorphism. Other palms, such as *Elaeis oleifera* Cortés and *E. guineensis* Jacq., also present high molecular polymorphism and ample morphological diversity (Jack and Mayes 1993; Shah et al. 1994).

Total genetic diversity (H_T) present in this set of landraces was 0.30, of which 0.25 was within the landraces (H_S) and 0.16 was among the landraces (G_{ST}). Hence, approximately 85% of the genetic diversity was within the landraces, while 15% of the diversity was between them, as also demonstrated for Clement et al. (1997) with allozymes. As expected, these estimates correspond closely to the variance components estimated by AMOVA, which attributed 82.2% ($p < 0.001$) of the variation to the within component and 17.8% ($p < 0.001$) to the among landraces component. These values are similar to those for other allogamous perennial species, both with DNA markers (Kageyama 1990; Bawa 1992; Gillies et al. 1999) and with allozymes (Hall et al. 1994; Chase et al. 1995).

The dendrogram generated from the Jaccard similarities contained eight reasonably well resolved sub-groups (with the predominance of a morphologically defined landrace or population of var. *chichagui*) and five poorly resolved groups (without predominance of a landrace) (Table 2). The fact that 10% of the plants analyzed were attributed to poorly resolved groups may be explained by problems of identification in the plantation, collection and manipulation of the samples in the laboratory, or interpretation of the amplification products, as also observed for Sousa et al. (2001), or by modern distribution of germplasm from its region of pre-Colombian origin to modern urban centers (Mora Urpí and Clement 1988).

The principal components analysis detected significant molecular variation (Figure 1), with the first three components explaining only 19.4% of the variation and the other 110 explaining from 3% to almost zero each. Four main groups which make geographic sense were observed in the three dimensional representation of the three principal components, and which correspond roughly to

Table 2. Number (and %) of individuals attributed to groups in the dendrogram of individual plants of pejibaye (*Bactris gasipaes* var. *gasipaes* and var. *chichagui*) based on Jaccard similarities and their relations to morphometrically defined landraces (based on Mora Urpí et al. 1997).

Group	*n*	*Chichagui*	Pará	Solimões	Putumayo	Pampa H.	Tuíra	Utilis	Guatuso
Resolved groups									
1	8	8 (100)	–	–	–	–	–	–	–
2	56	–	3 (5)	24 (43)	21 (38)	5 (9)	–	1 (2)	2 (4)
3	10	–	8 (80)	–	1 (10)	1 (10)	–	–	–
4	81	–	–	–	–	–	28 (35)	29 (36)	23 (28)
5	18	–	1 (6)	–	–	17 (94)	–	–	–
6	7	7 (100)	–	–	–	–	–	–	–
7	36	2 (6)	27 (75)	–	7 (19)	–	–	–	–
8	12	12 (100)	–	–	–	–	–	–	–
Non-resolved groups									
A	1	–	–	1 (100)	–	–	–	–	–
B	3	1 (33)	–	–	–	2 (67)	–	–	–
C	12	–	–	3 (25)	2 (17)	3 (25)	2 (17)	–	2 (17)
D	5	–	1 (25)	2 (35)	2 (35)	–	–	–	–
E	2	–	–	–	–	2 (100)	–	–	–
Total	250	30	40	30	33	30	30	30	27

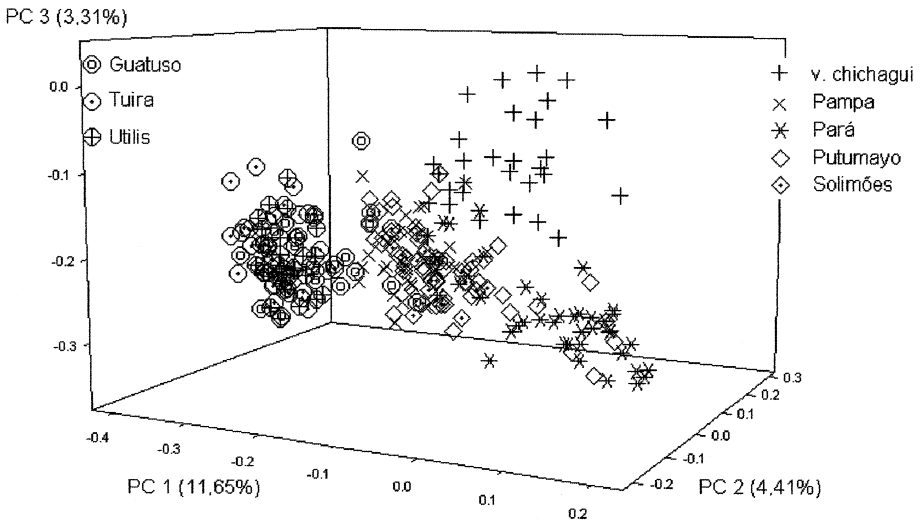

Figure 1. Three-dimensional representation of the first three principal components of the individual plants of pejibaye (*Bactris gasipaes*, var. *gasipaes* = landraces, var. *chichagui* = wild), based on 113 RAPD markers.

the larger groupings in the dendrogram of Jaccard similarities. Within the Central American group, to the extreme west of Figure 1, there is no evident formation of sub-groups, which would be expected if the three Central American populations were indeed landraces; the Jaccard similarity matrix (Table 2) did not separate these either. Within Amazonia, the center and eastern parts of Figure 1, three groups are evident. On the eastern extreme is the Pará landrace, with a few individuals of the Putumayo landrace, Amaturã and São Paulo de Olivença populations, which were also observed by Sousa et al. (2001) and which require further study. In the center and tending towards the west are the Putumayo and Solimões landraces, which show a mixture of plants, as in the Jaccard similarity matrix (Table 2). Further west still is the Pampa Hermosa landrace, with a few individuals

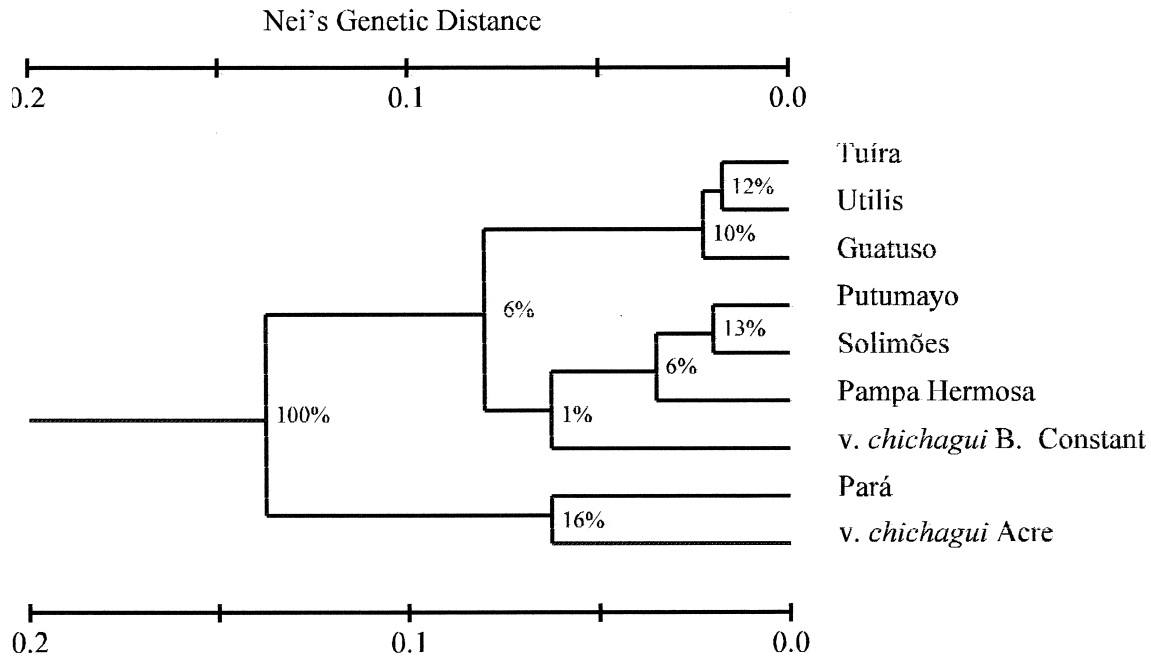

Figure 2. Dendrogram of Nei's (1972) genetic distances based on 113 RAPD markers among two populations of wild pejibaye (*Bactris gasipaes* var. *chichagui*) and seven cultivated pejibaye (var. *gasipaes*) landraces defined on morphometric criteria in Amazonia and Central America. The consistency values refer to the percentage of markers that supports each junction.

from Central America. Finally, the two populations of var. *chichagui* were distributed over the three Amazonian groups, suggesting either current gene flow (see below) or that they are ancestral to the Amazonian landraces. Within this group, a set of plants tended towards the west (Benjamin Constant population) and another towards the east (Acre population).

The dendrogram of the landraces based on Nei's (1972) genetic distances was structurally similar to the dendrogram of the Jaccard similarities at the macro level, presenting three groups (Figure 2): the Central American group was composed of the Tuíra and Utilis landraces, and then the Guatuso landrace; the western Amazonian group was composed of the Solimões and Putumayo landraces, then the Pampa Hermosa, and finally the var. *chichagui*, Benjamin Constant population; and the southeastern Amazonian group was composed of the Pará landrace and var. *chichagui*, Acre population. The general structure of this dendrogram is similar to that of Rojas-Vargas et al. (1999), based on allozymes in five populations within the geographic space analyzed here, suggesting consistency between the two studies.

The Exact Test (Raymond and Rousset 1995) showed that the Tuíra and Utilis landraces were not different ($p = 0.97$), as observed in the dendrogram of Jaccard similarities. The Tuíra and Guatuso landraces also were not different ($p = 0.09$), although they were much more different than Tuíra and Utilis. It is possible that the Guatuso landrace is somewhat different due to selection against spines, since this is the principal spineless landrace (population) in Central America (Clement et al. 1988; Clement and Manshardt 2000). The Exact Test also showed that the Putumayo and Solimões landraces were not different ($p = 0.11$), which agrees with the dendrogram of Jaccard similarities and the observations of Sousa et al. (2001). The var. *chichagui* populations (Acre and Benjamin Constant) were different, although the small number of individuals does not allow the precision of the other comparisons. The Pará landrace was different from all other landraces and from the var. *chichagui* Acre population, although they group together in the dendrogram of Nei's genetic distances. The Pampa Hermosa landrace was also different from all others, which may be partially explained by

Table 3. Gene flow among pejibaye (*Bactris gasipaes*) populations of var. *chichagui* and landraces (var. *gasipaes*) estimated with 113 RAPD markers and a mean of 30 plants per population.

	v.c. Acre	v.c. BC	Pará	Solimões	Putumayo	Pampa H.	Tuíra	Utilis
v.c. BC	2.7	–						
Pará	4.7	3.3	–					
Solimões	2.5	5.5	4.1	–				
Putumayo	4	6.3	7.9	16.3	–			
Pampa H.	2.4	4.4	3.5	11	11	–		
Tuíra	1.4	2.3	1.9	4.3	4.3	4.8	–	
Utilis	1.6	2.7	2.1	5.1	5.1	5.7	15.7	–
Guatuso	1.7	3.2	2.6	5.6	5.9	5.7	10.3	12.4

selection against spines in that region also (Clement et al. 1988; Clement and Manshardt 2000).

The mean gene flow among these landraces and wild populations was 2.7. Gene flow among the Central American landraces was high (Table 3), which agrees with the Nei genetic distances and again suggests only one landrace in the region. Gene flow between the Putumayo and Solimões landraces was higher still (16.3), strongly suggesting only one landrace there also, which agrees with the Nei genetic distances and the hypothesis of Sousa et al. (2001). Gene flow between the Central American and western Amazonian landraces was low (4.0), which contradicts, *a priori*, Clement's (1986) hypothesis of their relatedness.

The correlation between the gene flow and geographic distance matrices, with var. *chichagui* included, was -0.62 ($Z = 974.8$; Z min $p = 0.003$; $R^2 = 0.38$). When var. *chichagui* was excluded, this correlation increased to -0.83 ($Z = 624.2$; Z min $p = 0.009$; $R^2 = 0.69$). The difference between these correlations suggests that gene flow between the landraces and the populations of var. *chichagui* is minor even when in geographic proximity. This agrees with Clement et al.'s (1999) hypothesis of little introgression between var. *gasipaes* and var. *chichagui* in Benjamin Constant, even though the populations are sympatric. Two factors may be involved: limited flowering synchrony in Benjamin Constant, as the var. *chichagui* plants tended to fruit at and beyond the end of the var. *gasipaes* harvest; the plants sampled were from different sub-populations in the Benjamin Constant municipality, which might not have representatives of var. *chichagui* immediately adjacent to the var. *gasipaes* plants. These two hypotheses will require further study.

The valid landraces and their relationships

The set of genetic analyses presented above confirms the validity of the Pará and Pampa Hermosa landraces. Although a few plants were mixed with other landraces in the dendrogram of Jaccard similarities and the PCA, the number of these was small compared to other landraces. Part of this mixture may be due to the introduction of germplasm from the upper Solimões River (Putumayo landrace) into Belém, Manaus and Yurimáguas, as mentioned by Mora Urpí and Clement (1988). Nonetheless, in general the genetic and morphometric analyses of these landraces are in agreement.

The hypothesis of three landraces in Central America was not supported. Rather, the dendrogram of Jaccard similarities and the PCA showed the absence of a structure that might support the hypothesis. The Central American heterozygosities and percentages of polymorphism were low in comparison to the other landraces and similar among themselves, which suggests a strong affinity among these populations. The dendrogram of Nei's genetic distances and the Exact Tests support these affinities, and the gene flows were very high among these three populations. This set of evidence suggests the existence of only one landrace in Central America, contrary to the proposal of Mora Urpí (1992). The Utilis name should be conserved, in homage to the botanist S. Oersted, who described cultivated pejibaye in Central America as *Guilielma utilis* in 1858.

The case of the Solimões and Putumayo landraces is similar to that of Central America: the set of evidence suggests that the Putumayo landrace is not different from the Solimões landrace, corroborating the hypothesis of Sousa et al. (2001). The

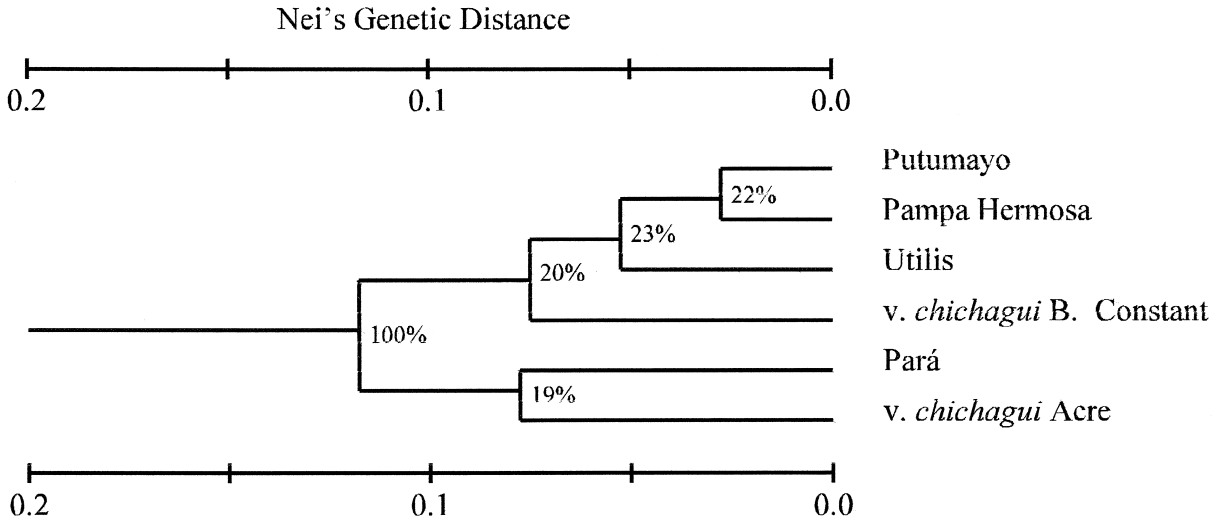

Figure 3. Dendrogram of Nei's (1972) genetic distances based on 113 RAPD markers among two populations wild pejibaye (*Bactris gasipaes* var. *chichagui*) and four cultivated pejibaye (var. *gasipaes*) landraces validated with genetic criteria in Amazonia and Central America. The consistency values refer to the percentage of markers that supports each junction.

only contrary evidence is the high heterozygosity of the Solimões landrace, but this can be explained by introgression between the Putumayo and Pará landraces, as these landraces come from different branches of the dendrogram of Nei's genetic distances (Figure 2) and would be expected to generate enhanced heterozygosity. If this hypothesis is correct, the Solimões landrace is not a landrace but a contact zone between landraces, the limits of which remain to be determined. The Putumayo name should be conserved.

Re-analysis of the matrix considering four valid landraces generated a dendrogram of Nei's genetic distances (Figure 3), with the Utilis landrace more closely related to the western Amazonian landraces, as suggested by Clement (1986) on morphometric grounds. The western branch of Figure 3 is remarkably similar to relationships presented by Rojas-Vargas et al. (1999) and Clement et al. (1997) using allozymes. Rojas-Vargas et al. found strong similarities between the populations of Guapiles (Utilis landrace) and Darién (Tuíra landrace), and that these were more closely related to Yurimáguas (Pampa Hermosa landrace) than to Belém (Pará landrace) and Chaparé (Bolivia; not ncluded here) on our eastern branch. The Nei genetic identities estimated from 17 allozyme loci (Clement et al. 1997) were nearly identical to those presented in Figure 3 for the same three populations of the same three landraces. The Exact Test demonstrated that all the valid landraces are different and gene flow among them is lower than in the original matrix. The correlation between the gene flow and geographic distance matrices was -0.86 ($Z = 202.9$; Z min $p = 0.001$; $R^2 = 0.74$), without var. *chichagui*. The distribution of the landraces of *B. gasipaes* var. *gasipaes* confirmed by the molecular characterization is demonstrated in Figure 4, together with the landraces not yet studied with molecular techniques.

Implications of the new landrace hypothesis

The genetic proximity (Figure 3) between the Utilis landrace and the western Amazonian landraces (Putumayo and Pampa Hermosa) suggests historical gene flow along a corridor between these areas, the most likely flow being the migration of domesticated populations from Amazonia to Central America, as suggested by Prance (1984) and Clement (1986). The low heterozygosity and percentage of polymorphism in the Utilis landrace strengthens the migration hypothesis, since migrations tend to involve only a sample of alleles from the source population, and inbreeding and genetic drift can further diminish this sample and change allele frequencies (Hartl 1981); this also explains the location of Utilis on the western branch of the

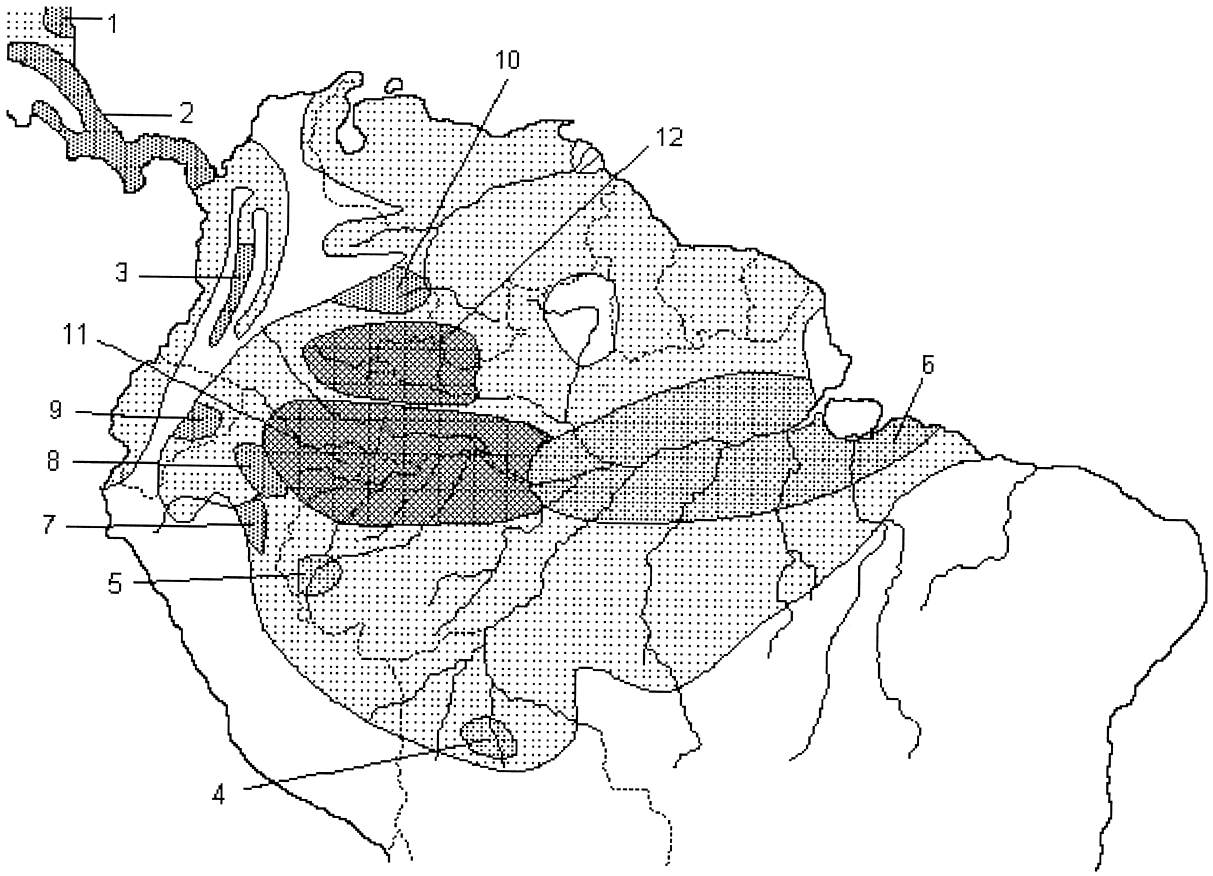

Figure 4. Approximate distribution of *B. gasipaes* var. *gasipaes* (light shading) in the lowland Neotropics, with the approximate distribution of valid (defined by molecular characterization and morphometric data) and still to be validated landraces. Central America and northwestern South America landraces – 1. Rama, 2. Utilis (now including Guatuso and Tuíra), 3. Cauca. Amazonian landraces – 4. Tembé, 5. Juruá, 6. Pará, 7. Pampa Hermosa, 8. Tigre, 9. Pastaza, 10. Inirida, 11. Putumayo (now including Solimões), 12. Vaupés.

dendrogram. Clement et al. (1997) reported that the San Carlos population (ex-Guatuso landrace; now Utilis landrace) contained a subset of the Benjamin Constant population (Putumayo landrace) allozyme alleles, as well as two unique alleles of unknown origin, possibly derived from introgression with the local populations of var. *chichagui* along the migration corridor in Ecuador, Colombia, Panama and Costa Rica (see Mora Urpí 1999 for a listing of these wild populations). In our study, the Central American populations also contained a subset of the Amazonian markers, with a unique marker in Central America (the 640 pb fragment produced by the F-919-3 primer).

If this hypothesis is valid, the Utilis landrace probably resulted from a rather late introduction of already domesticated pejibaye from western Amazonia into Central America, probably shortly before the Christian era. Corrales and Mora Urpí (1990) reported carbonized pejibaye seeds in the Costa Rican lowlands dated to about 2250 BP; human intervention in southern Mesoamerican ecosystems was much earlier, as was the transition to food production (Piperno and Pearsall 1998). Hence, *in loco* domestication is unlikely, contrary to the proposal of Mora Urpí (1992). While this hypothesis does not negate Morcote-Rios and Bernal's (2001) new hypothesis of a Colombian Andean origin, this is unlikely given the position of the var. *chichagui* populations on the final dendrogram (Figure 3) and the fact that northern Andean populations of var. *chichagui* do

not have the domesticate's seed morphology (Ferreira 1999).

Ferreira (1999) reported that all cultivated pejibaye (*B. gasipaes* var. *gasipaes*) have seed shape and germinal pore position similar to the southern populations of var. *chichagui*, represented in this study by the Acre and Benjamin Constant populations. Although gene flow among these wild pejibayes and geographically proximate cultivated pejibayes was low (Table 3), the positions of the two wild populations in the final dendrogram (Figure 3) relative to the cultivated pejibayes are those expected from the Ferreira (1999) report. If Costa Rica or Colombia was the origin of the Utilis landrace (or even all cultivated pejibayes), the Utilis landrace should appear on a separate branch of the dendrogram from the var. *chichagui* populations.

The two branches of the final dendrogram support the hypothesis of the origin of the pejibaye in southwestern Amazonia because of the position of the two populations of var. *chichagui*. Although the dendrogram could also support a double origin of pejibaye in southwestern Amazonia, the use of only two populations suggests that the parsimonious hypothesis is still a single event.

Various other authors have argued for southwestern Amazonia as a center of origin also, as proposed by Ferreira (1999) based on seed morphology and argued most recently by Clement (1995). However, the final dendrogram does not identify which southwestern population(s) gave rise to pejibaye, as the two southwestern populations studied here occur on different branches. The relation between the Pará landrace and the var. *chichagui*, Acre population, is greater than that between the Pará landrace and the var. *chichagui*, Benjamin Constant population, in terms of Nei's genetic distances (Figure 3) and gene flow (Table 3), although the geographic distance between them is similar. The same is true for the Putumayo landrace and the var. *chichagui*, Benjamin Constant population, *versus* the Acre population. Gene flow among the var. *chichagui* populations is also low, as expected given the geographic distance between them.

Huber (1904) identified two populations of what are now var. *chichagui* in southwestern Amazonia, one of which was very similar and geographically close to the Acre population studied here, and proposed that this was the region of origin of pejibaye. Mora Urpí (1984) later argued that the pejibaye of Bolivia, locally called Tembé, was morphologically very similar to the pejibaye of the Pará landrace. Rojas-Vargas et al. (1999) reported similarities between Tembé and Pará based on allozymes. Hence, the eastern branch of the final dendrogram presented here (Figure 3) agrees with previous studies. The western branch is less well studied and will require more sampling sites to confirm.

Assuming that this origin hypothesis is valid, the final dendrogram (Figure 3) suggests two migration routes out of the source area: one to the northeast, in the direction of eastern Amazonia, finally resulting in the Pará landrace; another to the northwest, in the direction of western Amazonia, resulting in the Pampa Hermosa and Putumayo landraces, before crossing the Andes (perhaps in Ecuador, as suggested by Prance 1984) and reaching Central America, resulting in the Utilis landrace. This new hypothesis requires the genetic analysis of more populations of pejibaye, both cultivated and wild, especially in northwestern South America.

Conclusions

The hypothesis of landraces in pejibaye was supported by the genetic evidence presented here and complements the morphometric evidence presented over the last two decades. The Pará and Pampa Hermosa landraces were supported as originally presented. The ex-Solimões landrace was shown to be part of the Putumayo landrace, which now extends into central Brazilian Amazonia. The Central American populations analyzed were shown to be components of a single landrace in that region, called Utilis. These four landraces are closely related, suggesting a single domestication event, as is generally observed in crop species. This new interpretation of these four landraces also supports the hypothesis of the origin of pejibaye in southwestern Amazonia and suggests two migration routes: one to the northeast and another to the northwest. Further genetic analysis is needed to test these hypotheses, which are also amenable to study by archaeologists, linguists and anthropologists, principally to identify the ethnic groups that may have been involved.

Acknowledgements

The authors thank the Banco da Amazônia, S.A. (BASA) for financial support for the project 'Marcadores moleculares (RAPDs) na discriminação das raças primitivas de pupunha (*Bactris gasipaes*) mantidas no Banco Ativo de Germoplasma', the Fundação Djalma Batista for administrative support, Dra. Marilene L.A. Bovi, Instituto Agronômico, Campinas, São Paulo, and Dr. Evandro Ferreira, INPA, Rio Branco, Acre, for numerous critical suggestions to improve the manuscript. This study is part of the first author's Master's Thesis.

References

Bawa K.S. 1992. Mating systems, genetic differentiation and speciation in tropical rain forest plants. Biotropica 24(2b): 250–255.

Blumler M.A. 1992. Independent inventionism and recent genetic evidence on plant domestication. Economic Botany 46(1): 98–111.

Buso G.S.C., Rangel P.H.and Ferreira M.E. 1998. Analysis of genetic variability of South American wild rice populations (*Oryza glumaepatula*) with isozymes and RAPDs markers. Molecular Ecology 7: 107–117.

Chase M.R., Boshier D.H.and Bawa K.S. 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 1. Genetic variation in natural populations. Am. J. Botany 82: 468–475.

Clement C.R. 1986. Descriptores mínimos para el pejibaye (*Bactris gasipaes* H.B.K.) y sus implicaciones filogenéticas. Masters' Thesis, Escuela de Biologia, Universidad de Costa Rica, San José, Costa Rica.

Clement C.R. 1995. Pejibaye (*Bactris gasipaes*). In: Smartt J. and Simmonds N.W. (eds), Evolution of Crop Plants, 2nd edn, pp. 383–388, Longman, London, UK.

Clement C.R. and Manshardt R.M. 2000. A review of the importance of spines for pejibaye heart-of-palm production. Scientia Horticulturae 83: 11–23.

Clement C.R. and Mora Urpí J. 1987. The pejibaye (*Bactris gasipaes* H.B.K., Arecaceae): multi-use potential for the lowland humid tropics. J. Econ. Botany 41(2): 302–311.

Clement C.R., Chávez F.W.B. and Moreira Gomes J.B. 1988. Considerações sobre a pupunha como produtora de palmito. Anais 1° Encontro Nacional de Pesquisadores em Palmito, CNPF/IAPAR/IAC, Curitiba, Paraná, Brasil. pp. 225–247.

Clement C.R., Aguiar J.P.L., Arkcoll D.B., Firmino J.L. and Leandro R.C. 1989. Pupunha brava (*Bactris dahlgreniana* Glassman): progenitora da pupunha (*B. gasipaes* H.B.K.)? Boletin do Museu Paraense Emílio Goeldi, série Botânica 5(1): 39–55.

Clement C.R., Aradhya M.K. and Manshardt R.M. 1997. Allozyme variation in spineless pejibaye (*Bactris gasipaes*, Palmae). Economic Botany 51(2): 149–157.

Clement C.R., Aguiar J.P.L. and Aued-Pimentel S. 1999. A pupunha brava (*Bactris dahlgreniana* Glassman, Palmae) no estado do Amazonas, Brasil. Acta Botanica Venezuelica 22(1): 29–44.

Corrales U.F. and Mora Urpí J. 1990. Sobre el proto-pejibaye en Costa Rica. Pejibaye (*Guilielma*) 2(2): 1–11.

Excoffier L., Smouse P.E. and Quatro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes application to human mitocondrial DNA restriction data. Genetics 131: 479–491.

Ferreira E. 1999. The phylogeny of pupunha (*Bactris gasipaes* Kunth, Palmae) and allied species. In: Henderson A. and Borchsenius F. (eds), Evolution, Variation and Classification of Palms. Memoirs of the New York Botanical Garden, vol. 83, pp. 225–236, New York Botanical Garden Press, Bronx, NY.

Gillies A.C.M., Navarro C., Lowe A.J., Newton A.C., Hernandez M., Wilson J. and Cornelius J.P. 1999. Genetic diversity in Mesoamerican populations of mahogany (*Swietenia macrophylla*), assessed using RAPDs. Heredity 83: 722–732.

Grattapaglia D. 1997. Pseudo-testcross mapping strategy using RAPD markers. In: Micheli M.R. and Bova R. (eds), Fingerprinting Methods Based on Arbitrarily Primed PCR, Springer Verlag, Berlin, pp. 201–217.

Hall P., Chase M.R. and Bawa K.S. 1994. Low genetic variation but high population differentiation in a common tropical forest tree species. Conservation Biol. 8: 471–482.

Hartl D.L. 1981. A Primer of Population Genetics. Sinauer Associates, Sunderland, MA.

Hartl D.L. and Clark A.G. 1989. Principles of Population Genetics, 2nd edn. Sinauer Associates, Sunderland, MA.

Henderson A. 2001. *Bactris* (Palmae). Flora Neotropica Monograph 79, New York Botanical Garden, New York.

Huber J. 1904. A origem da pupunha. Boletim do Museu Paraense Emilio Goeldi 4(1–4): 474–476.

Jack P.L. and Mayes S. 1993. Use of molecular markers for oil palm breeding. II. Use of DNA markers (RFLPs). Oléagineux 48(1): 1–8.

Kageyama P.Y. 1990. Genetic structure of tropical tree species of Brazil. In: Bawa K.S. and Hadley M. (eds), Reproductive Ecology of Tropical Forest Plants, UNESCO, Paris, pp. 375–387.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Louette D. 2000. Traditional management of seed and genetic diversity: what is a landrace? In: Brush S.B. (ed.), Genes in the Field: On-farm Conservation of Crop Diversity, IDRC, IPGRI, Lewis Publ., Boca Raton, Florida, pp. 109–142.

Matioli S.R. 2001. Biologia e Evolução. Editora Holos, Ribeirão Preto, São Paulo, Brasil.

Miller M.P. 1997. Tools for Population Genetic Analysis (TFPGA), version 1.3. Northern Arizona University, Flagstaff, Arizona.

Mora-Urpí J. 1984. El pejibaye (*Bactris gasipaes* H.B.K): origen, biología floral y manejo agronómico. In: Palmeras Poco Utilizadas de América Tropical, Food Agriculture Organization/Centro Agronómico Tropical de Investigación y Enseñanza – CATIE, Turrialba, Costa Rica, pp. 118–160.

Mora-Urpí J. 1991. Diversidad genética en pejibaye. II. Origen y domesticación. In: Mora Urpí J., Szott L.T., Murillo M. and Patinõ V.M. (eds), IV Congreso Internacional sobre Biología, Agronomía e Industrialización del Pijuayo, Editorial Universidad de Costa Rica, San José, Costa Rica, pp. 21–29.

Mora-Urpí J. 1992. Pejibaye (*Bactris gasipaes*). In: Hernández Bermejo J.E. and León J. (eds), Cultivos Marginados: Otra Perspectiva de 1492, Producción y protección vegetal N° 26, Food and Agriculture Organization – FAO/Jardín Botánico de Córdoba (España), Rome, pp. 209–220.

Mora-Urpí J. 1999. Origen y domesticacíon. In: Mora Urpí J. and Gainza Echeverria J. (eds), Palmito de Pejibaye (*Bactris gasipaes* Kunth): Su Cultivo e Industrializacíon, Editorial Universidad de Costa Rica, San José, Costa Rica, pp. 17–24.

Mora-Urpí J. and Clement C.R. 1988. Races and populations of peach palm found in the Amazon basin. In: Clement C.R. and Coradin L. (eds), Final Report (revised): Peach Palm (*Bactris gasipaes* H.B.K.) Germplasm Bank, U.S. A.I.D. project report, Instituto Nacional de Pesquisas da Amazônia/Centro Nacional de Recursos Genéticos, Manaus, Amazonas, Brasil, pp. 78–94.

Mora-Urpí J., Weber J.C. and Clement C.R. 1997. Peach Palm. Bactris gasipaes Kunth. Promoting the conservation and use of underutilized and neglected crops. 20. Institute of Plant Genetics and Crop Plant Research – IPK, Gatersleben/ International Plant Genetic Resources Institute – IPGRI, Rome.

Morcote-Rios G. and Bernal R. 2001. Remains of palms (Palmae) at archaeological sites in the New World: A review. Botanical Rev. 67(3): 309–350.

Nei M. 1972. Genetic distance between populations. American Naturalist 106: 283–292.

Piperno D.R. and Pearsall D.M. 1998. The Origins of Agriculture in the Lowland Neotropics. Academic Press, San Diego.

Prance G.T. 1984. The pejibaye, *Guilielma gasipaes* (HBK) Bailey, and the papaya, *Carica papaya* L. In: Stone D. (ed.), Pre-Columbian Plant Migration. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 76, Harvard University Press, Cambridge, pp. 85–104.

Raymond M.L. and Rousset R. 1995. An exact test for population differentiation. Evolution 49: 1280–1283.

Rohlf F.J. 1990. NTSys-pc – Numerical taxonomy and multivariate analysis system, version 1.6. Exeter Software, Setauket, NY.

Rojas-Vargas S., Ramírez P. and Mora Urpí J. 1999. Polimorfismo isoenzimático en cuatro razas y un híbrido de *Bactris gasipaes* (Palmae). Revista de Biologia Tropical 47(4): 755–761.

Shah F.H., Rashid O., Simons A.J. and Dunsdon A. 1994. The utility of RAPD markers for the determination of genetic variation in oil palm (*Elaeis guineensis*). Theor. Appl. Genetics 89: 713–718.

Slatkin M. and Barton N.H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43: 1349–1368.

Sousa N.R., Rodrigues D.P., Clement C.R., Astolfi-Filho S. and Nagao E.O. 2001. Discriminação de raças primitivas de pupunha (*Bactris gasipaes*) na Amazônia brasileira com marcadores moleculares RAPDs. Acta Amazonica 31(4): 539–545.

Weir B.S. 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer Associates, Sunderland, MA.

Weising S.A., Nybom H., Wolf K. and Meyer W. 1995. DNA Fingerprinting in Plants and Fungi. CRC Press, London.

Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful genetic markers. Nucl. Acid Res. 18: 6531–6535.

Yeh F.C., Yang R.-C. and Boyle T. 1999. Microsoft Windows-based Freeware for Population Genetic Analysis – POPGEN, versão 1.31. Downloaded in August 2000 from: http://www.ualberta.ca/~fyeh/

Zeven A.C. 1998. Landraces: A review of definitions and classifications. Euphytica 104: 127–139.