Molecular Ecology (2001) **10**, 1895–1907

Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system

M. ELIAS,*† L. PENET,† P. VINDRY,† D. McKEY,* O. PANAUD† and T. ROBERT†
CEFE-CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France, †Evolution et Systématique, bâtiment 360, Université Paris-Sud, 91405 Orsay cedex, France

Abstract

Occurrence of intervarietal or interspecific natural crosses has been reported for many crop plants in traditional farming systems, underlining the potential importance of this source of genetic exchange for the dynamics of genetic diversity of crop plants. In this study, we use microsatellite loci to investigate the role of volunteer seedlings (plants originating from unmanaged sexual reproduction) in the dynamics of genetic diversity of cassava (*Manihot esculenta* Crantz), a vegetatively propagated crop, in a traditional farming system in Guyana. A previous field study showed that farmers incorporate such plants into the germplasm for vegetative propagation, and that many of them are likely to be assigned by farmers to recognized varieties. Under strict vegetative propagation clonality of varieties is expected. The high proportion of polyclonal varieties observed suggests that incorporation of seedlings into the germplasm for propagation is a frequent event. The molecular variability assessed with microsatellite markers shows that there is high differentiation among heterozygous varieties, whereas populations of seedlings do not depart from the proportions expected under Hardy–Weinberg assumptions. Assignment of seedlings to a recognized variety on the basis of morphological similarity greatly increases genetic diversity within the variety. We argue that recombination and gene flow play a major role in the dynamics of genetic diversity of cassava in traditional farming systems. Documenting unmanaged sexual reproduction and its genetic consequences is a prerequisite for defining strategies of *in situ* conservation of crop plant genetic resources.

Keywords: cassava, genetic diversity, *Manihot esculenta*, microsatellite, sexual reproduction, vegetative propagation

Received 12 January 2001; revision received 4 April 2001; accepted 4 April 2001

Introduction

In traditional farming systems, the coexistence of different varieties of a given crop plant in the same field or in neighbouring fields is a common situation (Boster 1983; Clawson 1985; Louette *et al*. 1997). The term 'variety' refers to the group of phenotypes identified by farmers under a single name. Sympatry of cultivated and related wild forms is also often observed in centres of origin of domesticated plants (Jarvis & Hodgkin 1999). Such a situation is very favourable to genetic exchange between individuals displaying different phenotypes, and may therefore lead to the production of new genotypes and phenotypes on which human and natural selection can operate. This should result in a very dynamic evolution of traditional landraces, especially in outcrossing species, as pointed out by several authors (Ellstrand *et al*. 1999; Jarvis & Hodgkin 1999). However, at least two aspects of this question have been poorly investigated: (i) for domesticated plants that are mostly vegetatively propagated by man, the impact of sexual reproduction and recombination on genetic diversity is largely unknown, although examples of management by farmers of individuals resulting from sexual reproduction

Correspondence: Thierry Robert. Fax: + 33 1 69 15 46 97; E-mail: thierry.robert@esv.u-psud.fr

have been reported for potato (*Solanum* spp., Johns & Keen 1986; Quiros *et al*. 1992), cassava (*Manihot esculenta*, Boster 1985; Chernela 1987; McKey & Beckerman 1993; Salick *et al*. 1997; Emperaire *et al*. 1998; Elias *et al*. 2001a,b) and a few other vegetatively propagated crops (e.g. Shigeta 1996). And (ii) the way farming practices encourage or eliminate these sexually produced seedlings found in fields, and thereby influence the occurrence of further introgression between varieties or between domesticated and wild forms, has been very poorly documented, especially for vegetatively propagated plants, and needs to be investigated (Jarvis & Hodgkin 1999).

Cassava is an interesting model for tackling these questions. Cassava is a root crop, vegetatively propagated by means of stem cuttings. The crop originates from the neotropics, and is now grown in all tropical regions. Cassava was probably cultivated as early as 7000 BP (Piperno & Holst 1998), and recent analyses of genetic diversity in the genus *Manihot*, as well as phylogeographic studies, appear to support a single event of domestication in a restricted area in western Amazonia, from the wild form *M. esculenta* ssp. *flabellifolia*, synonym *M. esculenta* ssp. *peruviana* (Roa *et al*. 1997; Olsen & Schaal 1999, 2001). Inter-specific crosses involving cassava can be highly fertile (Jennings 1976; Nassar *et al*. 1986), and hybrids of cassava with wild *Manihot* species have been reported (Second *et al*. 1997). In Amazonian cassava farming systems, different varieties are intercropped in the fields (Boster 1983; Emperaire *et al*. 1998; Elias *et al*. 2001a).

We have undertaken a pluridisciplinary study aimed at documenting the dynamics of cassava diversity in the traditional farming system of the Makushi Amerindians from Guyana. Makushi farming practices have been well studied (Elias *et al*. 2001a). Makushi cassava cultivation is based on a slash-and-burn farming system, involving periods of cultivation of 2 or 3 years (i.e. two or three crops), and periods of fallow lasting 3 to more than 20 years. Previous field studies showed that most cassava varieties grown by the Makushi are fertile and produce seeds before plants are harvested. Seeds are dispersed first by dehiscence of the capsule (autochory), then by ants (myrmecochory), which with unknown frequency bury seeds in their nests (Elias & McKey 2000). Seeds remain dormant in a seed bank during the fallow, and germinate at the same time as the new crop is planted with cuttings, after slashing and burning of fallow vegetation. Volunteer seedlings are not weeded by the farmers, but rather are looked after, cared for, and harvested when they reach maturity, just like plants originating from cuttings. If they are found satisfactory by the farmers, plants originating from seedlings may be multiplied as a new variety, or eventually assigned to a recognized variety with which they share morphological features (Elias *et al*. 2001a,b).

In this paper, we present the first assessment for cassava of the genetic variability generated through sexual reproduction at the field level. Two populations of seedlings, as well as 29 recognized varieties, all collected from the same village, were characterized with microsatellite markers. The objectives of the present study were: (i) to assess the genetic structure of populations of sexually produced seedlings and compare it to that of the varieties; and (ii) to investigate the potential consequences at the genetic level of incorporating seedlings into recognized varieties.

Materials and methods

Plant material

Plant material (Table 1) was collected at Rewa, a 30-household village in Guyana, where at least 76 varieties are cultivated (Elias *et al*. 2001a). Twenty-nine varieties cultivated in Rewa were included in this study. Among them, 26 were used in a study of agro-morphological diversity of the crop at the village level (Elias *et al*. 2001b). On the basis of interviews of farmers and agronomic characterization of most of the varieties, we classified each of them by defining classes for four agronomic traits: the colour of the root (4 classes: 0 = white, 1 = cream, 2 = yellow, 3 = very yellow), the starch content of the root (5 classes: 0 = very watery, 1 = watery, 2 = half watery, 3 = dry, 4 = very dry), the cyanide content of the root or bitterness (4 classes: 0 = sweet, 1 = half bitter, 2 = bitter, 3 = very bitter) and the length of the cycle of cultivation (3 classes: 0 = short, 1 = medium, 2 = long). For each variety, 10 plants were collected, from one or several farmers depending on the variety. We also collected all individuals of two populations of plants originating from sexually produced volunteer seedlings, found in two currently cultivated cassava fields. These plants were characterized for morphological traits (Elias *et al*. 2001b). Hereafter, these plants will be referred to as seedlings. Positions of the seedlings in each field, whose maximum dimension did not exceed 40 m, were mapped. All information concerning plant material is summarized in Table 1.

Analyses of genetic variability with microsatellite markers

DNA extraction was performed on leaves dried for 48 h at 35 °C, using the protocol described by Colombo *et al*. (1998). We used eight microsatellite primers defined by Chavarriaga-Aguirre *et al*. (1998), named GAGG5, GA12, GA131, GA140, GA21, GA126, GA134 and GA136. The same loci and two additional ones were used by Roa *et al*. (2000) to characterize 38 accessions of cultivated cassava forming a subsample of a core collection maintained at the International Center for Tropical Agriculture, Cali, Colombia (the subsample is referred to as the 'core of the

Table 1 Varieties from the Makushi village that were used for microsatellite analysis. Sample size, number of farmers from whose fields samples were obtained, and classification for the four agronomic traits studied are given for each variety

Local name	English name	Code	Sample size	No. of farmers	Cyanide content (4 classes)	Colour of the root (4 classes)	Starch content (5 classes)	Cultivation cycle (3 classes)
Ainis pîye	Inez stick	2	10	2	2	1	2	2
Amo'ko pîye	Grand father stick	4	10	2	2	2	3	1
Santra pîye	Sandra stick	5	10	1	3	3	4	2
Anra pîye	Crane stick	6	10	5	2	1	2	1
Fatpoi pîye	Fat boy stick	7	10	4	3	2	3	1
Kraiwa pîye	Brazilian stick	8	10	4	4	2	2	1
Eti pîye	Eddie stick	12	10	5	2	3	3	2
Kaima pîye	Pumpkin stick	14	10	5	3	4	1	2
Kasiri pîye	Kashiri stick	19	10	5	2	1	4	2
Kini' pîye	Dry stick	21	10	2	3	3	4	3
Kuraatuma pîye (short)	Caiman stick (short)	22	10	5	4	1	3	1
Kurarî pîye	Corral stick	23	10	5	3	3	3	3
Kuraswa pîye	Crash Water stick	24	10	4	2	4	2	2
Itakon ye	White man stick cousin	26	10	2	3	3	2	3
Papîro pîye	Pablo stick	35	10	2	3	2	3	2
Paakaima ye	Buffalo stick	36	10	1	3	3	3	2
Pîrîkwa pîye	Bird stick	39	10	5	2	1	3	1
Paranakîrî pîye	White man stick	40	10	5	2	3	4	3
Sapîrî pîye	Fine fish stick	43	10	4	3	1	3	2
Siya pîye	Shea stick	45	10	2	3	2	3	3
Siment pîye	Cement stick	47	10	4	2	1	4	1
Supra pîye	Cutlass stick	48	10	3	2	1	3	3
Tarekaya pîmoi pîye	Water turtle egg stick	49	10	5	3	4	1	1
Wo'ye	Drink stick	54	10	2	2	3	3	2
Siwal pîye	Sea wall stick	57	10	1	2	4	3	2
Kana	Sweet cassava	58	10	2	1	1	3	1
Kuraatuma pîye (tall)	Caiman stick (tall)	59	10	1	4	1	3	1
Reni pîye	Renie stick	60	10	3	4	2	4	2
Zacari pîye	Zaccharie stick	112	10	2	3	2	3	2
Population of seedlings in the first field		A	39		1			
Population of seedlings in the second field		N	42		1			

core'), as well as accessions of wild cassava and wild *Manihot* species.

Polymerase chain reactions (PCRs), detection and sizing of alleles were performed at two different locations, with different methods of detection (silver staining and radioactivity). Control samples were used to ensure that the same alleles were obtained whatever the method used. Nonradioactive and radioactive PCRs were performed following the methods developed by Chavarriaga-Aguirre *et al*. (1998). Nonradioactive PCR products were detected by silver staining as described by Le Thierry d'Ennequin *et al*. (2000). Radioactive PCR products were run on 5% denaturant-page polyacrylamide gels for 2 h.

Statistical analyses

Genetic variability in varieties and populations of seedlings. Allelic frequencies, observed heterozygosity (H_O), gene diversity (H_E) calculated as the Hardy–Weinberg expected heterozygosity (Nei 1978) and *F*-statistics (Wright 1978; Weir & Cockerham 1984) were calculated for the cultivated varieties and the seedlings using BIOSYS-2 software (Swofford *et al*. 1997). BIOSYS-2 was also used for testing departure from the genotypic proportions expected under Hardy–Weinberg assumptions for the population of varieties and the two populations of seedlings at each locus. GENETIX software (Belkhir *et al*. 2000) was used to calculate the coefficient of correlation associated with unbiased linkage disequilibrium values between loci for the two populations of seedlings (Cockerham & Weir 1977), as defined by Weir (1979). Multidimensional analyses of the diversity among varieties and seedlings were performed using Factorial Correspondence Analysis (FCA) with GENETIX. The FCA was performed on the centroids of the varieties; individuals belonging to recognized varieties, or to the populations of seedlings, were plotted as supplementary individuals.

Relationship between spatial distribution and genetic variability in the populations of seedlings. For each population of seedlings, the correlation between matrices of genetic and morphological distances was assessed with a Mantel test (Mantel 1967), using GENETIX. Using the AUTOCORG software (Hardy *et al*. 2000), spatial distances between seedlings in each population were distributed among 10 classes. Classes were defined in a way that in each field all classes included approximately the same number of pairs of individuals (approximately 74 for population A, and 40 for population N). Moran's *I* statistic (Hardy & Vekemans 1999) was computed for each distance class.

Relationships between molecular variability and morphological diversity in varieties and seedlings. Morphological characterization was available for 26 of the 29 varieties used in this study, as well as for most of the seedlings (14 characters for the varieties, nine for the seedlings, Elias *et al*. 2001b). We used this information to assess the correlation between morphological and genetic distances among the varieties, and among the seedlings of each population. Genetic distances between varieties were estimated using the chord distance of Cavalli-Sforza (Cavalli-Sforza & Edwards 1967) computed with BIOSYS-2. For morphological distances between varieties, we used the Mahalanobis distance (Dagnélie 1975) between centroids of varieties, which was previously computed (Elias *et al*. 2001b). For inter-individual pairwise distances in each population of seedlings, we used the Nei & Li (1979) distance on microsatellite data, and euclidian distance on standardized morphological data, computed with STATISTICA software (STATISTICA 1997).

Fertility of the varieties. To estimate the potential contribution of the varieties to sexual reproduction, fertility was assessed by recording the total number of inflorescences (at all stages of development) on 17–38 individuals per variety, planted in the experimental design described by Elias *et al*. (2001b). However, fertility was measured only in nine individuals for variety 35, while, because of tardily detected synonymy between varieties, 79 individuals were available for variety 49. Differences between varieties were investigated by an analysis of variance, followed by a Tukey–Kramer test (Sokal & Rohlf 1985), performed with SAS software (SAS 1996).

Evaluation of genetic consequences of incorporation of seedlings into recognized varieties. Observations in the field showed that seedlings are often assigned to recognized varieties by farmers, on the basis of morphological similarities (Elias *et al*. 2001a). In a previous study, in order to simulate the incorporation of seedlings into the varieties, we performed a discriminant analysis on seedlings of populations A and N and on individuals of 26 of the 29 varieties of the present study, characterized by morphological descriptors involved in identification of varieties by farmers (Elias *et al*. 2001b). In our simulation, a total of 43 seedlings (26 from population A and 17 from population N) were assigned to a recognized variety. Genetic consequences of the incorporation of the seedlings into varieties were evaluated in terms of number of genotypes per variety, and of number of alleles per locus, before and after the incorporation of seedlings.

Results

Structure of the genetic variability of varieties and populations of seedlings

Distribution of allelic diversity among loci. Table 2 shows for each locus the number of alleles, the mean and standard deviation of allelic sizes, the mean observed heterozygosity, and gene diversity for the overall sample of cassava varieties. Table 2 also presents summary statistics averaged over all loci for the varieties and for each population of seedlings.

For each locus, sizes of alleles and their respective frequencies in the sample of the 290 individuals are shown in Table 3.

All alleles detected for loci GA12 and GA126 had been detected by Roa *et al*. (2000) in the subsample of the core collection of cultivated cassava. For all other loci, a total of nine alleles that were not detected in the 'core of the core' were found, including four frequent alleles (allelic frequency greater than 0.1). Six of them had already been detected in another study involving Amazonian cassava landraces (data not shown). Among these alleles newly detected in accessions of cultivated cassava, four were also found in accessions of wild forms of *Manihot esculenta* or other wild *Manihot* species (Roa *et al*. 2000).

Five alleles that were detected in the 'core of the core' (Roa *et al*. 2000) were not present in our sample.

Structure of the genetic variability of cultivated varieties. In the overall sample of the varieties, 75 different multilocus genotypes were found. Varieties 2, 14, 23, 26, 36, 39, 40 and 59 were monomorphic on the basis of the eight loci used, i.e. for each of them all 10 individuals had identical multilocus genotypes, although all of them except varieties 36 and 59 were obtained from at least two farmers. Intravarietal polymorphism was detected for the 21 remaining varieties. Except for variety 24, intravarietal polymorphism was even detected among individuals sampled from the same farmer. Number of multilocus genotypes for polymorphic varieties varied from two for varieties 5, 19, 24, 41, 45, 49 and 57 to six for varieties 8 and 43, and even seven for variety 58, the sweet cassava. Within each polymorphic variety, one multilocus genotype was

Table 2 Number of alleles, mean and standard deviation of allelic sizes, observed heterozygosity (H_O) and gene diversity (H_E) for each microsatellite locus in the overall sample of 29 varieties; and averaged values of the same parameters averaged over all loci for the varieties, and for the two populations of seedlings (A and N)

Loci	Population	Sample size	No. of alleles per locus	Mean of allelic sizes ± standard deviation (bp)	H_O	H_E
GAGG5		290	3	120.48 ± 4.48	0.560	0.605
GA12		290	3	144.00 ± 3.35	0.542	0.522
GA131		290	9	113.20 ± 4.85	0.691	0.663
GA140	varieties	290	5	165.30 ± 5.98	0.830	0.720
GA21		290	4	113.00 ± 1.19	0.380	0.406
GA126		290	7	185.00 ± 8.40	0.790	0.696
GA134		290	4	317.20 ± 6.10	0.377	0.494
GA136		290	5	154.50 ± 4.86	0.559	0.633
Averaged	varieties	290	5.0 ± 0.7	—	0.591 ± 0.060	0.592 ± 0.039
over all	A	39	3.5 ± 0.2	—	0.510 ± 0.041	0.551 ± 0.048
loci	N	42	3.6 ± 0.3	—	0.492 ± 0.074	0.512 ± 0.059

Table 3 Sizes of alleles and their respective frequencies at each locus. Boldface type indicates alleles that were not detected in the subsample of the core collection of cultivated cassava (Roa *et al*., personal communication). Among them, those that are underlined were detected in wild *Manihot* species (Roa *et al*. 2000). Allele 104 of locus GA21, detected in one seedling of population A (see text) was not represented in the population of varieties

Locus	GA12			GA21					GA136					
size	137	145	147	104	109	112	114	**120***	**144***	150	154	156	158	
freq.	0.185	0.644	0.171	0	0.019	0.241	0.732	0.007	0.133	0.101	0.205	0.014	0.546	
Locus	GAGG5			GA126							GA134			
size	117	**119***	127	178	182	184	188	190	212	219	309	317	326	**332***
freq.	0.519	0.165	0.315	0.203	0.440	0.011	0.260	0.032	0.048	0.016	0.192	0.678	0.030	0.100
Locus	GA131									GA140				
size	96	102	104	**108**	**112***	**113**	114	**115**	116	158	162	**166***	168	172
freq.	0.028	0.039	0.074	0.005	0.021	0.018	0.325	0.018	0.474	0.323	0.113	0.146	0.042	0.376

*Alleles already detected by Elias *et al*. (submitted).

usually predominant over the others. Moreover, while some genotypes included in polymorphic varieties were close to each other, others were distant (e.g. those included in variety 43, Fig. 1a). Identical multilocus genotypes were detected for individuals that belonged to different varieties. Pairs of varieties involved are 7 and 47, 7 and 54, 47 and 54, 54 and 57, 45 and 54, and 22 and 59. Varieties within each of these pairs are morphologically similar (Elias *et al.* 2001b), and the fact that identical multilocus genotypes were detected is probably due to confusion of the farmers between these varieties.

Genetic differentiation among varieties was significant [averaged over the eight loci, $F_{ST} = 0.363 \pm 0.021$ ($P < 10^{-4}$), ranging from 0.300 for locus GA140 to 0.467 for locus GA136]. A strong excess of heterozygotes was detected within variety [averaged over the eight loci, $F_{IS} = -0.551 \pm 0.041$ ($P < 10^{-4}$), ranging from −0.293 for locus GA134 to −0.667 for locus GA126]. Differentiation among varieties and excess of heterozygotes within varieties resulted at the whole population level in either a slight excess or a slight deficit of heterozygotes, depending on the locus [averaged over the eight loci, $F_{IT} = 0.013 \pm 0.049$ ($P = 0.18$), ranging from −0.124 for locus GA126 to 0.246 for locus GA134].

Structure of genetic variability in the populations of seedlings. The number of alleles per locus, observed heterozygosity and gene diversity for the two populations of seedlings are presented in Table 2. Both populations of seedlings had comparable levels of diversity, but neither of them had all the allelic diversity that was found in the set of all varieties. However, allele 104 of locus GA21 found in one seedling of population A was not detected among the 29 varieties studied. The observed mean heterozygosity across all loci

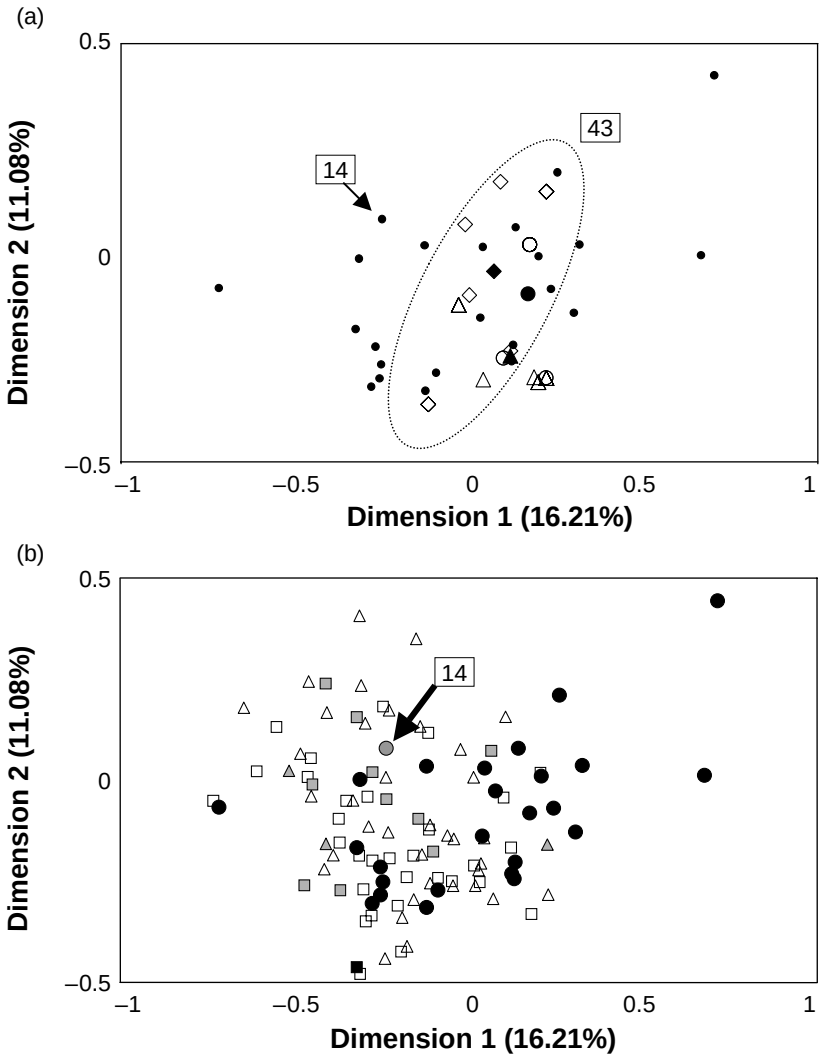

Fig. 1 Factorial Correspondence Analysis (FCA) performed on the centroids of the varieties (black symbols). To improve the clarity of the representation, centroids of varieties 40, 58 and 112, which have a marginal position, are not represented here. (a) As illustrative examples, individuals of varieties 7, 43 and 47 are plotted as supplementary points, and represented by open symbols (triangles, diamonds, and large circles, respectively). Corresponding centroids are represented by the same symbols, but symbols are filled. Some varieties are monoclonal, such as variety 14 (indicated by the arrow). Other varieties embrace several genotypes, such as variety 43 (diversity of included genotypes shown by the ellipse), but these genotypes may be either close to each other, or very different [Nei & Li (1979) genetic distance averaged over the 15 pairwise distances between the six different genotypes of this variety: 0.309 ± 0.162, ranging from 0.040 to 0.546]. Varieties 7 and 47 embrace several multilocus genotypes, and are genetically indistinguishable, even sharing one multilocus genotype. (b) Seedlings of populations A and N are plotted as supplementary individuals, and are represented with triangles and squares, respectively. To exemplify genetic consequences of incorporation of seedlings into a recognized variety, grey triangles and squares represent individuals from each population that are likely to be assigned to variety 14 (grey circle, indicated by the arrow) on the basis of morphological similarities (Elias *et al*. 2001b). Some seedlings assigned to variety 14 are genetically close to this variety, whereas others are genetically distant from the individuals of the variety [average Nei & Li (1979) genetic distance between the genotypes of seedlings and the genotype of the clonal variety 14: 0.393 ± 0.147, ranging from 0.083 to 0.583, but genetically close to individuals belonging to different varieties].

was very close to the Hardy–Weinberg expected heterozygosity. Moreover, at each locus, after a Bonferroni correction for multiple independent tests, both populations of seedlings globally did not depart from the genotypic proportions expected under Hardy–Weinberg assumptions.

There was significant, although weak, linkage disequilibrium for only six of 28 pairs of loci (GA131, GA140), (GAGG5, GA136) (GA140, GA136) (GA126, GA134) (GA126, GA136) and (GA134, GA136) for population A, and only for the pair (GA126, GA12) for population N (Table 4).

There was slight genetic differentiation between the two populations of seedlings [averaged over the eight loci, $F_{ST} = 0.023 \pm 0.008$ ($P < 10^{-4}$), ranging from –0.004 for locus GA126 to 0.039 for locus GA12], as well as a slight

Table 4 Values of Weir's (1979) coefficient of correlation corresponding to the unbiased estimate of linkage disequilibrium (Cockerham & Weir 1977) for each pair of loci in the two populations of seedlings (top right half matrix: population A; lower left half matrix: population N). Significance of the coefficient of correlation: NS = non significant, *$P < 0.01$, **$P < 0.001$, ***$P < 0.0001$

	GAGG5	GA12	GA131	GA140	GA21	GA126	GA134	GA136
GAGG5	—	0.063 NS	0.152 NS	0.143 NS	0.127 NS	0.165 NS	0.224 NS	0.237***
GA12	0.178 NS	—	0.190 NS	0.137 NS	0.059 NS	0.103 NS	0.150 NS	0.094 NS
GA131	0.140 NS	0.116 NS	—	0.164**	0.100 NS	0.116 NS	0.083 NS	0.144 NS
GA140	0.164 NS	0.138 NS	0.145 NS	—	0.186 NS	0.115 NS	0.116 NS	0.208*
GA21	0.190 NS	0.057 NS	0.148 NS	0.175 NS	—	0.188 NS	0.177 NS	0.169 NS
GA126	0.183 NS	0.150*	0.121 NS	0.186 NS	0.132 NS	—	0.237*	0.304***
GA134	0.124 NS	0.136 NS	0.108 NS	0.112 NS	0.080 NS	0.146 NS	—	0.246***
GA136	0.199 NS	0.167 NS	0.123 NS	0.142 NS	0.090 NS	0.175 NS	0.213 NS	—

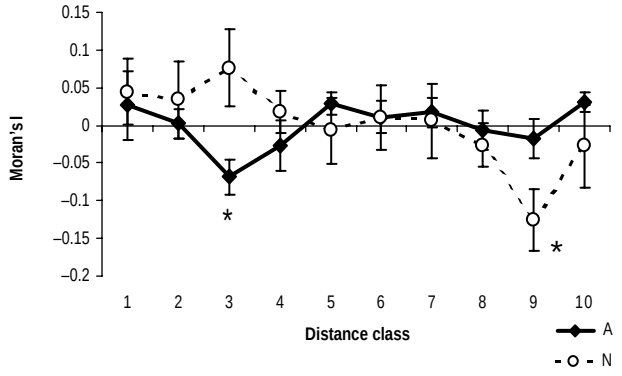

Fig. 2 Correlogram of Moran's *I* per distance class for each of the populations of seedlings. Means and standard errors of Moran's *I* over the eight loci per distance class. * indicates Moran's *I* significantly different from 0 at $P < 0.05$.

differentiation between each of them and the population constituted by the varieties [averaged over the eight loci, $F_{ST} = 0.026 \pm 0.012$ ($P < 10^{-4}$) between population A and the varieties, and $F_{ST} = 0.033 \pm 0.014$ ($P < 10^{-4}$) between population N and the varieties]. Figure 1(b) shows that the diversity of the two populations of seedlings overlaps to a large extent the diversity of the set of varieties.

The Mantel tests did not show any correlation between spatial and genetic distances (population A: Z = 9107, $P = 0.455$, $r = 0.069$, $P = 0.436$; population N: Z = 4560, $P = 0.128$; $r = 0.077$, $P = 0.190$). Similarly, for each distance class of each of the populations of seedlings, values of Moran's *I* statistic were never significantly higher than 0 (Fig. 2).

Relationships between genetic variability and morphological diversity

Varieties. There was a significant positive association of genetic and morphological distances between the varieties (parameters of the Mantel test: Z = 19924, $P = 0.038$; $r = 0.204$, $P = 0.054$), but this association was weak (Fig. 3). *F*-statistics associated with the classification of varieties into agronomic classes defined for bitterness, root colour, starch content and cycle of cultivation showed no molecular differentiation between classes (Table 5).

Seedlings. Population A of seedlings showed a weak association between the structure of morphological variability and the structure of molecular variability assessed by microsatellite markers (Z = 2003, $P = 0.060$; $r = 0.132$, $P = 0.050$), while no significant association was found for population N (Z = 1636, $P = 0.359$; $r = 0.026$, $P = 0.371$).

Potential contribution of the varieties to sexual reproduction

The number of inflorescences varied greatly within variety (coefficient of variation on average higher than 100%, data not shown). Individual contribution to sexual reproduction is therefore expected to vary substantially. This trait also varied significantly among varieties ($F_{29,816} = 12.39$, $P < 10^{-3}$, see also Elias *et al*. 2001b), from almost or totally sterile varieties (e.g. varieties 5, 26, or 112) to varieties bearing on average more than 15 inflorescences (e.g. varieties 40, 47 or 7). Analysis of paternity between seedlings and varieties was attempted, but the number and polymorphism of the microsatellite loci used in this study were not sufficient to allow a powerful discrimination between the potential parental pairs. It was, therefore, not possible to determine whether some varieties contributed significantly more to the seedlings than others.

Consequences at the genetic level of the incorporation of seedlings into cultivated varieties

Seedlings were usually not genetically close to the individuals of the variety into which they were likely to be incorporated on the basis of morphological characters, although in some rare cases they were, as for variety 14 and some seedlings incorporated into it in our simulation (Fig. 1b). In most cases, they were closer to individuals of other varieties. As a result, incorporation of seedlings into

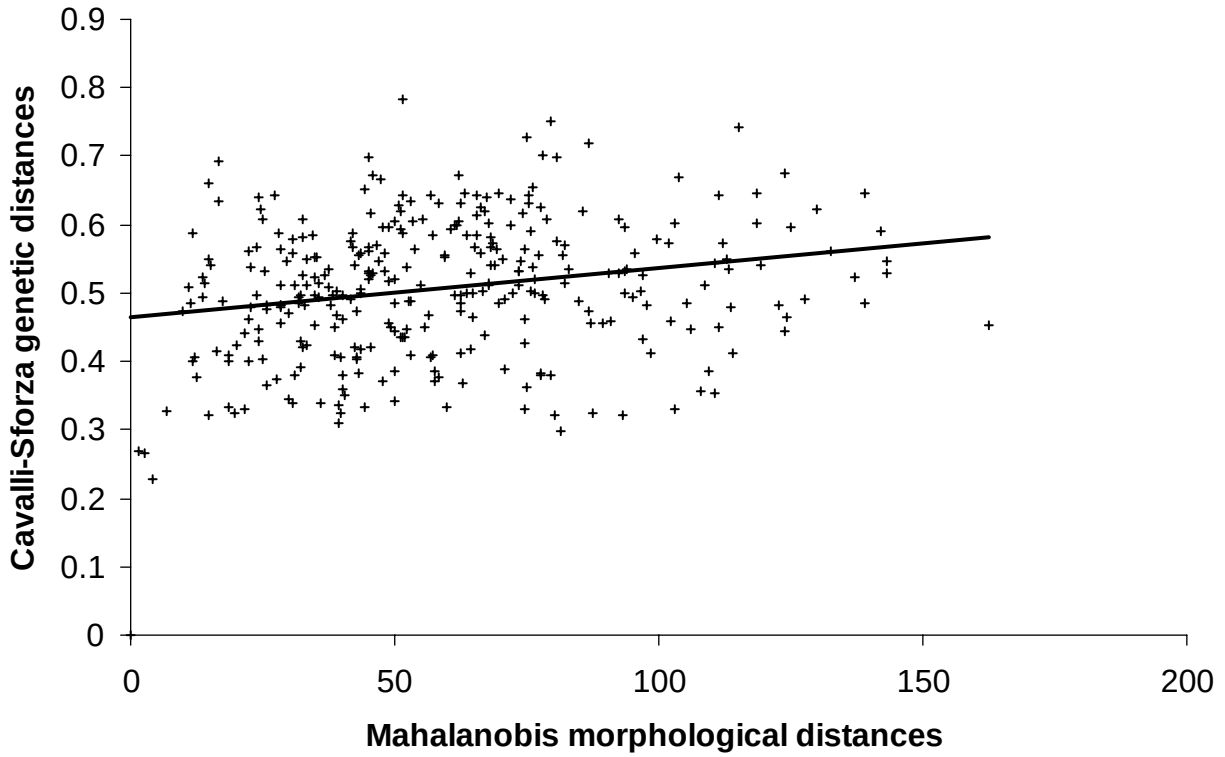

Fig. 3 Relationship between Cavalli-Sforza genetic distances and Mahalanobis morphological distances between varieties. Each point on the graph represents a pair of varieties.

Table 5 Distribution of molecular variability among varieties and classes of four agronomic traits: values of hierarchical multilocus *F*-statistics (Wright 1978) and molecular variance component

	Bitterness		Root colour		Starch content		Cycle of cultivation	
	F	Variance component	*F*	Variance component	*F*	Variance component	*F*	Variance component
Varieties — classes	0.342	1.637	0.333	1.571	0.330	1.548	0.340	1.624
Varieties — total	0.334	1.576	0.334	1.577	0.334	1.577	0.334	1.577
Classes — total	−0.013	−0.060	0.001	0.001	0.006	0.029	−0.010	−0.046

Table 6 Number of seedlings assigned to each variety (see Elias *et al*. 2001b), and number multilocus genotypes and mean number of alleles per locus of each variety before and after incorporation of the seedlings

Variety	No. of seedlings assigned	No. of multilocus genotypes		Mean no. of alleles	
		Before	After	Before	After
5	4	2	6	1.8 ± 0.2	3.0 ± 0.3
6	6	3	9	2.0 ± 0.3	3.0 ± 0.3
7	1	4	4	2.6 ± 0.2	2.6 ± 0.2
12	4	4	8	2.9 ± 0.3	3.3 ± 0.3
14	14	1	15	1.5 ± 0.2	3.4 ± 0.3
23	1	1	2	1.4 ± 0.2	1.8 ± 0.3
39	6	1	7	1.6 ± 0.2	3.0 ± 0.3
47	1	4	5	2.5 ± 0.2	2.9 ± 0.2
49	2	2	6	1.8 ± 0.3	2.3 ± 0.3
54	1	5	6	2.3 ± 0.2	2.6 ± 0.2
57	1	2	3	1.8 ± 0.2	2.5 ± 0.3
58	2	6	8	2.8 ± 0.4	3.3 ± 0.3

varieties decreases the genetic differentiation between varieties. Incorporation of seedlings into a variety increased the number of genotypes represented in the variety and introduced new alleles in the variety (Table 6), which increases the gene diversity of the variety. For variety 7, however, the sole seedling that was incorporated in the simulation had exactly the same multilocus genotype as one individual of this variety. All other seedlings, whether or not they were incorporated into any variety in the simulation, had a multilocus genotype unique among all individuals (seedlings or recognized varieties) sampled.

Discussion

Genetic structure of populations of seedlings

Seeds germinating in a newly burnt old fallow are certainly overwhelmingly those from the bank of dormant seeds produced by the individuals that were cultivated in the same parcel a few years previously (Elias & McKey 2000). In the case of the two populations of seedlings studied here, both fields from which they were collected had been cultivated in a previous cycle for two consecutive years, then left fallow for five years, before again being cultivated in the year of our study.

Neither population of seedlings globally departed from Hardy–Weinberg proportions, and linkage disequilibria between loci were weak. This suggests that the mating system was predominantly outcrossed. Sexual traits of cassava, such as strong protogyny (individuals are monoecious), insect pollination, and the existence of inbreeding depression (Kawano *et al*. 1978) are also consistent with an outcrossed mating system. Assessment of cassava breeding system using morphological markers (Pereira 1978) also supports the view that cassava is allogamous. However, in light of large differences among varieties in fertility and in frequency in the fields, it is highly improbable that all varieties equally participated in sexual reproduction.

The seedlings whose genetic composition we analysed in this study were morphologically differentiated from the set of varieties studied here, especially the seedlings of population N (Elias *et al*. 2001b). However, allelic frequencies in both populations of seedlings were little differentiated from those of the overall sample of varieties, suggesting that the seedlings originated from crosses within a pool of varieties that were not genetically different from the overall sample of varieties, in terms of allelic frequencies at the microsatellite loci. Because of the weak correlation between structure of morphological variability and of genetic variability of microsatellite markers, higher contribution of some varieties sharing particular morphological traits (e.g. degree of branching, which, because of architectural constraints, is strongly associated with the fertility of the varieties, Elias *et al*. 2001b) does not necessarily imply differences in allelic frequencies at the microsatellite loci studied between these varieties and the whole set of varieties.

In neither of the fields sampled did molecular variability of the seedlings display a spatially structured pattern. Two reasons may account for this. First, seeds can be scattered several metres away from the plants that produced them by a double dispersal mechanism, involving explosive dehiscence of the fruit, followed by dispersal of seeds by ants (Elias & McKey 2000). Second, because of the polyclonality of the varieties, and their spatial patterns in the fields (the same variety can be planted at several locations in the fields), effective sibs, i.e. seeds produced by genetically identical parents, can be produced at distant locations. However, slight but significant genetic differentiation between the two populations of seedlings was observed. This could be explained by genetic differences between progenitors of each population.

Sexual reproduction and dynamics of genetic variability of cassava

Cassava is a vegetatively propagated crop, and only a limited number of individuals per variety contribute stem cuttings to the next generation of planting (Elias *et al*. 2001a), leading to a bottleneck at the variety level. A strong genetic differentiation between varieties is therefore expected, which was indeed observed. However, this study using microsatellite markers confirmed that most varieties (72%) were not monoclonal, even when individuals were collected from a single farmer, and in spite of genetic drift due to the bottleneck combined with vegetative propagation. In a previous study including 20 of the varieties used in the present study, we had already shown, with amplified fragment length polymorphism (AFLP) markers, that a large proportion of the varieties (66%) were not monoclonal (Elias *et al*. 2000). Polyclonality has also been detected among traditional varieties of cassava cultivated in Brazil (Colombo 1997; Second *et al*. 1997; Sambatti *et al*. 2001) and in Malawi (Mkumbira *et al*., in preparation).

Several reasons can account for existence of polyclonal varieties. First, 11 multilocus genotypes differed only by one allele from another multilocus genotype of the same variety, a difference which could be reasonably attributed to mutation. In other cases, intravarietal polyclonality can be attributed to the confusion of genetically distinct individuals that share similar morphological features (Elias *et al*. 2001b). This was indeed the case for five pairs of varieties, which are also frequently confused by farmers (M. Elias, personal observation). Confusion between morphologically similar varieties cannot account for all cases of

polyclonality detected, because most varieties involved are distinctive enough to be identified without risk of confusion (Elias *et al*. 2001b). Previous results have led us to propose that incorporation of individuals originating from sexually produced seeds into the germplasm used for vegetative propagation, including the assignment of some of these individuals to recognized varieties, plays a major role in the dynamics of the genetic diversity of cultivated varieties of cassava in traditional farming systems (Elias *et al*. 2000, 2001b; Elias & McKey 2000). Assignment of seedlings to recognized varieties increases the genetic diversity within varieties, in terms of number of genotypes and allelic diversity, all the more as there is little or no association between morphological and genetic distances among seedlings.

We showed a strong excess of heterozygous genotypes within varieties of cassava included in this study. This excess of heterozygotes may simply result from the varieties being monoclonal, or polyclonal with a predominant genotype. For a monoclonal variety, the homozygous state at a given locus leads to computation of no F_{IS} value (F_{IS} cannot be estimated for a fixed homozygous genotype), while the heterozygous state leads to F_{IS} values equal to –1. While heterozygote excess cannot alone provide evidence for heterosis, reports of high heterozygosity of cassava varieties compared to wild forms or wild species (Olsen & Schaal 1999; Roa *et al*. 2000), and the existence of inbreeding depression (Kawano *et al*. 1978), are both consistent with heterosis. In contrast, each population of seedlings conformed to the genotypic proportions expected under Hardy–Weinberg assumptions. In the field, it was not possible to have consistent information about the decision of farmers as to whether particular seedlings would be kept or discarded. While some farmers told us they systematically multiplied every seedling they found in their farms, they admitted they often discarded the least productive ones after a few generations of vegetative propagation. Other farmers stated that they immediately select the seedlings to be multiplied (i.e. when they harvested them), discarding the rest. Under heterosis, i.e. if heterozygous genotypes at selected loci are on average more vigorous than homozygous ones, one should expect the subset of seedlings actually multiplied (on the basis of evaluation of their productivity, among other traits) to be highly heterozygous also at neutral loci. Preliminary data tend to support this hypothesis (M. Elias, unpublished data), but this point needs to be confirmed.

The high diversity found in seedlings both at the molecular (this paper) and morphological (Elias *et al*. 2001b) levels suggests that seedlings are produced by frequent and broad intervarietal recombination, favoured by intercropping. Existence of gene flow between varieties via incorporation of seedlings, together with the maintenance of heterozygosity through vegetative propagation, may account for the relatively moderate genetic differentiation among varieties, with values of F_{ST} much lower than one, despite the bottleneck affecting the varieties each planting generation.

In our study, we detected in one seedling an allele that was not detected in any of the 29 varieties. This allele is probably a rare allele present in one or a few varieties that were not included in the study sample, but we cannot discard the possibility that this seedling was produced by a variety bearing this allele that is no longer cultivated in the village. Because seeds usually germinate several years after they are produced (Elias & McKey 2000), they may include alleles that have become rare or have even been lost since the seeds were produced.

Overall, the high proportion of polyclonal varieties detected with both AFLP markers (Elias *et al*. 2000) and microsatellite markers, despite the strong bottleneck affecting each variety at every planting generation, suggests that incorporation of seedlings into recognized varieties, and more generally into the material of propagation, is much more frequent than usually suspected. In the literature, although presence and use of seedlings have been reported among many groups of traditional cassava farmers (Boster 1984; Chernela 1987; Salick *et al*. 1997; Emperaire *et al*. 1998), their occurrence and incorporation have often been regarded as rare phenomena by the authors. We argue that frequency of seedlings may have been underestimated among other groups of cassava farmers in the literature. Makushi farmers, for instance, do not spontaneously offer information about volunteer seedlings, and interviews would not give the impression that their incorporation is a frequent practice, unless the investigator poses precise questions about this aspect. Even when such questions are posed, farmers tend to underestimate the actual frequency of this practice (based on censuses of plants in farmers' fields, Elias *et al*. 2001a). The Makushi farming system may therefore be representative of other traditional cassava farming systems, and the use of sexually produced seedlings in traditional cassava farming in Amazonia may be the rule rather than the exception.

Occurrence of gene flow between varieties in traditional cassava farming systems also raises the question of whether gene flow might occur between cultivated and wild forms of cassava, or with other wild *Manihot* species. Introgression of wild forms or other wild *Manihot* species into cultivated forms could account for the presence in our study sample of alleles that were detected only in wild forms of *M. esculenta* or other *Manihot* species by Roa *et al*. (2000). Alternatively, these alleles could have been present in the initial domesticated pool, but were not represented in the 'core of the core'. These alleles may also have arisen independently by mutation in cultivated cassava, resulting in convergent allelic size with alleles of wild *Manihot*.

Sexual reproduction and evolution of cassava

Exclusively vegetative propagation is expected to produce linkage disequilibria. Association between neutral molecular markers and loci involved in agronomic and morphological characters is therefore expected. In the present study, there was only little correspondence between the structure of molecular polymorphism revealed by microsatellite markers, and that of agro-morphological variability, confirming previous results obtained using AFLP markers (Elias *et al*. 2000, 2001b). Selection on agronomic characters (which can be directional, disruptive or diversifying, depending on the trait under consideration) as well as diversifying selection for morphological characters (Boster 1985), may have taken place repeatedly and independently. Such selection, combined with sexual reproduction that leads to recombination between neutral loci and genes involved in agronomic and morphological characters, may account for the observed patterns of molecular and agro-morphological diversity. Our results testify to the contemporary importance of sexual reproduction in the dynamics of cassava due to the incorporation by farmers of sexually produced plants into the material for vegetative propagation. The contemporary importance of sexual reproduction in dynamics of cassava diversity may reflect the role of sexual reproduction during the domestication of cassava, and during its evolution. Sexual reproduction may have been particularly crucial during the domestication of cassava, because wild forms have a very low aptitude for vegetative reproduction (Antonio Costa Allem, personal communication). Recombination may also have greatly facilitated the accumulation of domesticated characters (Elias & McKey 2000). Similar processes may have acted during the evolution of other vegetatively propagated crops (e.g. potato, Johns & Keen 1986; Quiros *et al*. 1992; sweet potato, Yen 1968; and ensete, Shigeta 1996), challenging the classical view of evolution of such crops, for which sexual reproduction is commonly regarded as anecdotal or even nonexistent (Zohary 1984).

Conclusion

This study underlines the importance of genetic exchange within populations of domesticated cassava in traditional farming systems, for the maintenance and the dynamics of genetic diversity in this vegetatively propagated crop. Ethnological and ecological observations have shown that such genetic exchange is favoured by traditional farming practices (Elias & McKey 2000; Elias *et al*. 2001b). To fully explore the consequences of the incorporation of sexually produced plants into the germplasm for propagation, more data are needed concerning the behaviour of the farmers towards these plants (criteria for multiplying the seedlings, frequency of multiplication of seedlings, real frequency of assignment of seedlings into a recognized variety), the breeding system of cassava and ecology of its sexual reproduction (phenology, fertility, pollination, dispersal, seed predation, dormancy and germination).

There is an urgent need for defining strategies of dynamic conservation of genetic resources (FAO 1996); our findings argue in favour of *in situ* conservation of genetic resources of crop plants as an important component of such a dynamic strategy (see also McKey *et al*. 2001). Ecological and social determinants that allow a dynamic management of genetic resources of crop plants in traditional farming systems may be relatively fragile, especially when faced with forces such as modernization of farming practices and economic pressures.

Finally, the existence of considerable genetic exchange within populations of crop plants, and the potential for hybridization with wild relatives, even in crops thought to be insulated from this process because they are 'vegetatively propagated', raise the question of the consequences of release of genetically modified varieties, associated with high risks of diffusion of transgenes.

Acknowledgements

The European Commission (DGVIII, Future of Tropical Forest Peoples Program [APFT]) and the government of France (Programme Environnement, Vie et Société, CNRS) provided funds for fieldwork. A grant from the Bureau des Ressources Génétiques (France) supported all genetic analyses. We are grateful to the government of Guyana, especially the Ministry of Amerindian Affairs and the Environmental Protection Agency, for allowing us to carry out our fieldwork in Guyana. We are also grateful to personnel of the Amerindian Research Unit of the University of Guyana and to the National Agronomic Research Institute (Georgetown, Guyana) for their advice and support. Many thanks to Paola Bertolino, for her extensive help with laboratory work, to Nicolas Bierne, for useful comments on an earlier draft of this paper, and for providing the subprogramme of GENETIX performing FCA on centroids of populations, and to John Bowers for help with the attempts to analyse paternity. Special thanks are extended to the Makushi Amerindians of Rewa, for their inestimable contribution to our study, and for their hospitality.

References

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000) *GENETIX 4.02, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UPR 9060, Université de Montpellier II, Montpellier (France).

Boster JS (1983) A comparison of the diversity of Jivaroan gardens with that of the tropical forest. *Human Ecology*, **11**, 47–68.

Boster JS (1984) Classification, cultivation and selection of Aguaruna cultivars of *Manihot esculenta* (Euphorbiaceae). In: *Ethnobotany in the Neotropics. Advances in Economic Botany, 1* (eds Prance GT, Kallunki JA), pp. 34–47. New York Botanical Garden, Bronx, New York.

Boster JS (1985) Selection for perceptual distinctiveness: evidence from Aguaruna cultivars of *Manihot esculenta*. *Economic Botany*, **39**, 310–325.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**, 550–570.

Chavarriaga-Aguirre P, Maya MM, Bonierbale MW *et al.* (1998) Microsatellites in cassava (*Manihot esculenta* Crantz): discovery, inheritance and variability. *Theoretical and Applied Genetics*, **97**, 493–501.

Chernela JM (1987) Os cultivares de mandioca na área do Uaupés (Tukano). In: *SUMA: Etnólogica Brasileira 1* (ed. Ribeiro B), pp. 151–158. Etnobiologia, FINEP, Petrópolis.

Clawson DL (1985) Harvest security and intraspecific diversity in traditional tropical agriculture. *Economic Botany*, **39**, 56–67.

Cockerham CC, Weir BS (1977) Digenetic descent measures for finite populations. *Genetical Research*, **30**, 121–147.

Colombo C (1997) *Etude de la diversité génétique des maniocs américains (*Manihot esculenta *Crantz) par les marqueurs moléculaires (RAPD et AFLP).* PhD Dissertation, ENSAM, Montpellier.

Colombo C, Second G, Losada Valle T, Charrier A (1998) Genetic diversity characterization of cassava cultivars (*Manihot esculenta* Crantz). (I) RAPD markers. *Genetics and Molecular Biology*, **21**, 105–113.

Dagnélie P (1975) *Analayses statistiques à plusieurs variables*. Les presses agronomiques de Gembloux, Belgique.

Elias M, Mckey D (2000) The unmanaged reproductive ecology of domesticated plants in traditional agroecosystems: an example involving cassava and a call for data. *Acta Oecologica*, **21**, 223–230.

Elias M, Mckey D, Panaud O, Anstett MC, Robert T (2001b) Traditional management of cassava morphological and genetic diversity by the Makushi Amerindians (Guyana, South America): perspectives for on-farm conservation of crop genetic resources. *Euphytica*, in press.

Elias M, Panaud O, Robert T (2000) Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. *Heredity*, **85**, 219–231.

Elias M, Rival L, Mckey D (2001a) Perception and management of cassava (*Manihot esculenta* Crantz) diversity among Makushi Amerindians of Guyana (South America). *Journal of Ethnobiology*, in press.

Ellstrand NC, Prentice HC, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics*, **30**, 539–563.

Emperaire L, Pinton F, Second G (1998) Une gestion dynamique de la diversité variétale du manioc en Amazonie du Nord-Ouest. *Nature, Science et Société*, **6**, 27–42.

Fao (1996) *The State of the World's Plant Genetic Resources for Food and Agriculture*. FAO, Rome.

Hardy OJ, Vanderhoeve S, Meerts M, Vekemans X (2000) Spatial autocorrelation of allozyme and quantitative markers within a natural population of *Centaurea jacea* (*Asteraceae*). *Journal of Evolutionary Biology*, **13**, 656–667.

Hardy O, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83**, 145–154.

Jarvis DI, Hodgkin T (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular Ecology*, **8**, S159–S173.

Jennings DL (1976) Cassava *Manihot esculenta* (Euphorbiaceae). In: *Evolution of Crop Plants* (ed. Simmonds NW), pp. 81–84. Longman, London.

Johns T, Keen SL (1986) Ongoing evolution of the potato on the altiplano of Western Bolivia. *Economic Botany*, **40**, 409–424.

Kawano K, Amaya A, Daza P, Rios M (1978) Factors affecting efficiency of hybridization and selection in cassava. *Crop Science*, **18**, 373–376.

Le Thierry d'ennequin M, Panaud O, Toupance B, Sarr A (2000) Assessment of genetic relationships between *setaria italica* and its wild relative *Setaria viridis* using AFLP markers. *Theoretical and Applied Genetics*, **100**, 1061–1066.

Louette D, Charrier A, Berthaud J (1997) In situ conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. *Economic Botany*, **51**, 20–38.

Mantel N (1967) The detection of disease clustering and generalised regression approach. *Cancer Research*, **27**, 209–220.

McKey D, Beckerman S (1993) Chemical ecology, plant evolution and traditional cassava cultivation systems. In: *Tropical Forest, People and Food* (eds Hladik CM, Hladik A, Pagezy H, Linares OF, Koppert GJA, Froment A), pp. 83–112. UNESCO, Paris.

McKey D, Emperaire L, Elias M *et al.* (2001) Gestions locales et dynamiques régionales de la diversité variétale du manioc en Amazonie. *Génétique, Sélection, Evolution*, in press.

Nassar NMA, da Silva JR, Vieira C (1986) Hibridação interespecifica entre mandioca e espécies silvestres de *Manihot*. *Ciência e Cultura*, **38**, 1050–1055.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **83**, 583–590.

Nei M, Li W (1979) Mathematical model for studying genetic variations in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the USA*, **76**, 5269–5273.

Olsen KM, Schaal BA (1999) Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences of the USA*, **96**, 5586–5591.

Olsen KM, Schaal BA (2001) Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae), and its wild relatives: further evidence for a southern Amazonian origin of domestication. *American Journal of Botany*, **88**, 131–142.

Pereira AS (1978) Taxa de fecundação cruzada no cultivar de mandioca Branca-de-Santa-Catarina. *Bragantia*, **37**, 95–96.

Piperno DR, Holst I (1998) The presence of starch grains on prehistoric stone tools from the humid neotropics: indication of early tuber use and agriculture in Panama. *Journal of Archeological Science*, **25**, 765–776.

Quiros CF, Ortega R, Van Raamsdonk LWD *et al.* (1992) Amplification of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmers. *Genetic Resources and Crop Evolution*, **39**, 107–113.

Roa AC, Maya MM, Duque MC *et al.* (1997) AFLP analysis of relationships among cassava and other *Manihot* species. *Theoretical and Applied Genetics*, **95**, 741–750.

Roa AC, Chavarriaga-Aguirre P, Duque MC *et al.* (2000) Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellite: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647–1655.

Salick J, Cellinese N, Knapp S (1997) Indigenous diversity of cassava: generation, maintenance, use and loss among the Amuesha, Peruvian upper Amazon. *Economic Botany*, **51**, 6–19.

Sambatti JBM, Martins PS, Ando A (2001) Folk taxonomy and evolutionary dynamics of cassava: a case Study in Ubatuba, Brazil. *Economic Botany*, **55**(1), 93–105.

Sas (1996) *SAS/STAT user's guide, release 6.12*. SAS Institute, Cary, N.C.

Second G, Allem AC, Mendes RA *et al.* (1997) Molecular marker (AFLP)-based *Manihot* and cassava numerical taxonomy and

genetic structure analysis in progress: implications for their dynamic conservation and genetic mapping. *African Journal of Root and Tuber Crops*, **2**, 140–147.

Shigeta M (1996) Creating landrace diversity: the case of the Ari people and Ensete (*Ensete ventricosum*) in Ethiopia. In: *Redefining Nature Ecology, Culture and Domestication* (eds Ellen RF, Fukui K), pp. 233–268. Berg, Oxford.

Sokal RR, Rohlf FJ (1985) *Biometry*. W.H. Freeman, New York.

STATISTICA (1997) *STATISTICA 5.1*. StatSoft Inc., Tulsa, OK. http://www.statsoft.com.

Swofford DL, Selander RB, Black WC (1997) *BIOSY-2, a computer program for the analysis of allelic variation in genetics*. Colorado State University, Ft. Collins, USA. ftp://lamar.edu/pub/wcb4.

Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics*, **35**, 235–254.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright S (1978) *Evolution and the Genetics of Populations, Vol 4. Variability Within and Among Natural Populations* University of Chicago Press, Chicago.

Yen DL (1968) Natural and human selection in the Pacific sweet potato. In: *Evolution and Environment: a Symposium presented on the Occasion of the One Hundredth Anniversary of the Foundation of Peabody Museum of Natural History at Yale University* (ed. Drake ET), pp. 387–412. Yale University Press, New Haven and London.

Zohary D (1984) Modes of evolution in plants under domestication. In: *Plant Biosystematics* (ed. Grant WF). Academic Press, New York.

Marianne Elias is interested in evolutionary biology. This study, part of her PhD project, was carried out as part of a 5-year EU program on the future of tropical forest peoples (APFT). Doyle McKey is a tropical biologist at the university of Montpellier who also participated in the APFT programme. Thierry Robert and Olivier Panaud are interested in the evolutionary and conservation biology of domesticated plants. Laurent Penet is a graduate student and Prune Vindry an undergraduate student in their laboratory.
