MOLECULAR ECOLOGY

Molecular Ecology (2009) **18**, 2897–2907 doi: 10.1111/j.1365-294X.2009.04231.x

Population genetics of *Manihot esculenta* ssp. *flabellifolia* gives insight into past distribution of xeric vegetation in a postulated forest refugium area in northern Amazonia

ANNE DUPUTIÉ,* MARC DELÊTRE,† JEAN-JACQUES DE GRANVILLE‡ and DOYLE MCKEY*

*CEFE UMR5175, 1919 Route de Mende, 34293 Montpellier Cedex 5, France, †Botany Department, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland, ‡Unité 84 Biodival, Institut de Recherche pour le Développement, Herbier de Guyane, Route de Montabo, BP165, 97323 Cayenne Cedex, France

Abstract

The Guianas have often been proposed as a forest refugium; however, this view has received little testing. Studies of population genetics of forest taxa suggest that the central part of French Guiana remained forested, while the southern part (currently forested) may have harboured more open vegetation. Insights into the population structure of species restricted to non-forested habitats can help test this hypothesis. Using six microsatellite loci, we investigated the population genetics of French Guianan accessions of *Manihot esculenta* ssp. *flabellifolia*, a taxon restricted to coastal savannas and to rocky outcrops in the densely forested inland. Coastal populations were highly differentiated from one another, and our data suggest a recent colonization of these savannas by *M. esculenta* ssp. *flabellifolia* in a west-to-east process. Coastal populations were strongly differentiated from inselberg populations, consistent with an ancient separation of these two groups, with no or low subsequent gene flow. This supports the hypothesis that the central part of the region may have remained forested since the Last Glacial Maximum, impeding the establishment of *Manihot*. Contrary to coastal populations, inselberg *Manihot* populations were strikingly homogeneous at a broad spatial scale. This suggests they were connected until recently, either by a large continuous savanna area or by smaller, temporary disturbed areas shifting in space.

Keywords: Amazonia, dispersal, French Guiana, hybridization, *Manihot*, refugia

Received 13 January 2009; revision received 26 March 2009; accepted 31 March 2009

Introduction

Climate variation due to Milankovitch oscillations has triggered repeated shifts in species distributions (Hewitt 2000). During glacial periods, climate in the tropics was cooler and drier than now, whereas it was warmer and wetter during interglacials (Hewitt 2000). Several theories explaining patterns of species distribution in Amazonia have been proposed. On the one hand, the long-debated refugia theory proposed by Haffer (1969) postulates that forests have expanded during interglacials, and receded during glacial times, becoming restricted to small islands (refugia) in a matrix of savanna vegetation or dry forest (Pennington *et al*. 2000). Allopatric speciation of forest species would have been favoured by such stable forest refugia, and the regions of high forest species endemism, such as the Guianas (Haffer 1969; de Granville 1982, 1993; Bush 1994), would have been forest refugia. On the other hand, Bush's (1994) disturbance–vicariance hypothesis proposes that high endemism in the Guianas is due not to long-term stability of this region, but to numerous climate changes in this region. High endemism would be due to high rates of speciation, favoured by selection pressures due to environmental change. According to this view, vegetation adapted to drier climates covered much of the Guianas during glacial times, with rainforest expanding during interglacials.

Data supporting one or the other hypothesis in the Guianas are scarce. We are aware of only one site in inner French Guiana where palynological studies have been conducted, in the central part of French Guiana (Charles-Dominique *et al*. 1998; Ledru 2001). This site remained forested for the last 10 000 years (Charles-Dominique *et al*. 1998), although

Correspondence: Anne Duputié, Fax: +33467612138;
E-mail: anne.duputie@ens-lyon.org

primary forest was temporarily replaced by secondary forest at least twice, around 1500 years before present (BP) and around 1000 BP (Ledru 2001). Results of charcoal analyses conducted on soil cores (Tardy 1998), together with those arising from population genetic studies of a rainforest tree species, *Vouacapoua americana* (Dutech *et al*. 2003, 2004), were consistent with this central region being a forest refugium since the Last Glacial Maximum (LGM). However, no studies so far have considered the remote southern part of French Guiana, which is characterized by a drier forest type than in the remainder of the country (Gond *et al*. in press). Vegetation models for the LGM suggest that the central part of French Guiana, as well as the western coastal region, were covered by rainforest vegetation during the LGM, whereas southern and eastern regions harboured somewhat drier vegetation [either dry forest or savannas; Bush (1994); van der Hammen & Hooghiemstra (2000); Pennington *et al*. (2006)]. However, climate simulations by Mayle *et al*. (2004) infer the continuous presence of evergreen broad-leaved forest in the whole region ever since the LGM.

Phylogeographic studies may provide insights into past vegetation. To cast light on the history of the vegetation of the Guianas, we examined the population genetics of a plant restricted to dry environments, at the scale of French Guiana. The species we considered, *Manihot esculenta* ssp. *flabellifolia* (Pohl) Ciferri, is the wild progenitor of cassava (Olsen & Schaal 1999), with which it has been shown to occasionally hybridize in nature, at least in the coastal savannas of French Guiana (Duputié *et al*. 2007). *M. esculenta* ssp. *flabellifolia* is distributed on an arc partly encircling the Amazon basin, from eastern Bolivia and Peru eastward to northeastern Brazil, northward to the Guianas and then westward to Venezuela. In the southwestern part of its distribution range, the plant grows in transition zones between forest and *cerrado* and can persist as a climbing shrub growing in openings in the forest (Allem 1994; Olsen & Schaal 1999), but it also grows as a bushy shrub in open areas in this region (A. Duputié and D. McKey, personal observation). In the northeastern part of its distribution range, the taxon is found in open environments such as savannas and granitic outcrops. Its distribution range in the Guianas is very patchy, as it is found in a narrow stretch along the coast, and forms small populations on rocky outcrops (inselbergs) far away from the coast, a distribution pattern it shares with numerous other species restricted to dry environments (de Granville 1982). In inner French Guiana where forest openings and river banks could offer suitable habitats, *M. esculenta* ssp. *flabellifolia* has never been found. It is also absent from inselbergs located in northeastern French Guiana (Fig. 1).

In contrast to cassava, which farmers propagate clonally using stem cuttings, vegetative reproduction appears to play no role in the biology of *M. esculenta* ssp. *flabellifolia*. Observations suggest that the taxon is pollinated mostly by meliponine bees. Seeds are dispersed first by ballistic autochory, through explosive dehiscence of the fruit, and then are secondarily dispersed by ants [notably *Ectatomma brunneum* and *Dorymyrmex pyramicus* ssp. *guyanensis* in the Guianas (Elias & McKey 2000)]. Seeds germinate in response to disturbance and/or opening of the environment (Pujol *et al*. 2002). As meliponine bees have a flight range of only 0.1–1 km (Roubik 1989), and as seed dispersal does not exceed a few metres (Elias 2000), most dispersal events occur over a very short range. Because the germination of *Manihot esculenta* ssp. *flabellifolia* is conditioned to canopy opening, and as its closest relatives are restricted to open environments (A. Duputié *et al*., unpublished), one can hypothesize that this taxon has been restricted to open environments over the past few thousand years.

Genetic characteristics of *Manihot* populations from coastal and inner French Guiana may help cast light on the history of vegetation in this part of Amazonia. If the northern part of the region (excluding the coastal savanna corridor) has remained forested during the past few thousand years, as postulated by Haffer (1969), de Granville (1982) and Dutech *et al*. (2004), one should expect this zone to have formed a barrier to the dispersal of *M. esculenta* ssp. *flabellifolia*. Populations from the coastal savannas should therefore be strongly differentiated from populations of inner savannas. If the southern part of the region also remained forested, as Haffer (1969) postulated, species restricted to dry environments should have been historically restricted to small xeric islands (inselbergs). Under this scenario, populations of *M. esculenta* ssp. *flabellifolia* in the southern part of the region would have been founded through occasional long-distance colonization of a few migrants. Hence, a strong isolation by distance should be expected for the inselberg populations of southern French Guiana. These populations would be expected to be strongly differentiated from one another and to have low intrapopulation diversity, due to genetic drift and to founder effects. On the other hand, if the southern part of the Guianas was formerly covered by drier vegetation types, a hypothesis defended by Bush (1994), inselberg populations of *Manihot* would be the remnants of formerly more extensive populations. Interpopulation differentiation would be due to drift and mutation acting independently in each population, eventually obscuring underlying signals of isolation by distance. Intrapopulation diversity could have remained high.

Here, we (i) assess whether genetic structure of *Manihot* populations provides any evidence for past continuous dry vegetation areas in French Guiana (particularly in the southern region), and (ii) determine the relationships between inselberg and coastal populations of *Manihot* in this region. To that aim, it is first necessary to assess whether, and to what extent, gene flow from domesticated cassava has affected the genome of its wild relative in French Guiana.

Materials and methods

Sampling

Leaf samples were picked from 373 individuals of *Manihot esculenta* ssp. *flabellifolia* (*sensu* Allem 1994) from all over French Guiana, between 2002 and 2007. Ten populations were sampled along the coast and four on inselbergs (filled squares on Fig. 1; sampling locations are presented in greater detail in Table S1, Supporting information). One of these, SM, is the population where hybrids with cassava had been found by Duputié *et al*. (2007). Only purely wild individuals from this population were included. On the coast, we found no *Manihot* populations west of Savane Grand Macoua or east of Cayenne. As inland inselbergs are only easily accessible by helicopter, we were only able to sample a limited number. The taxon probably can be found on other inselbergs we were not able to visit (dots on Fig. 1). On the other hand, it was not present on all inselbergs we visited in the southern region, and has never been found on inselbergs in the northeastern region (open squares on Fig. 1). The map showing the occurrence of *Manihot* on inland inselbergs was constructed using all samples present at the herbarium of Cayenne (CAY) and based on the compilation of unpublished notes and field observations of researchers who have visited these inselbergs.

In each population, five to 60 individuals were sampled randomly. Our sample size for each population is not representative of the actual population size: coastal populations are very large, whereas inselberg populations are of several tens to a few hundred individuals. Four very close inselbergs were sampled for population MA, three for population WA, and two for population RT (Fig. 1). Although our sampling design is not balanced, removal of the four populations with < 19 individuals did not modify our conclusions (see Supporting information).

Genotyping

DNA was extracted using DNeasy 96 Plant kit (QIAGEN GmbH). On each extraction plate, six wells were dedicated to data control (blank and replicates). We scored seven microsatellite loci: GA12, GA21, GA57, GA126 (Chavarriaga-

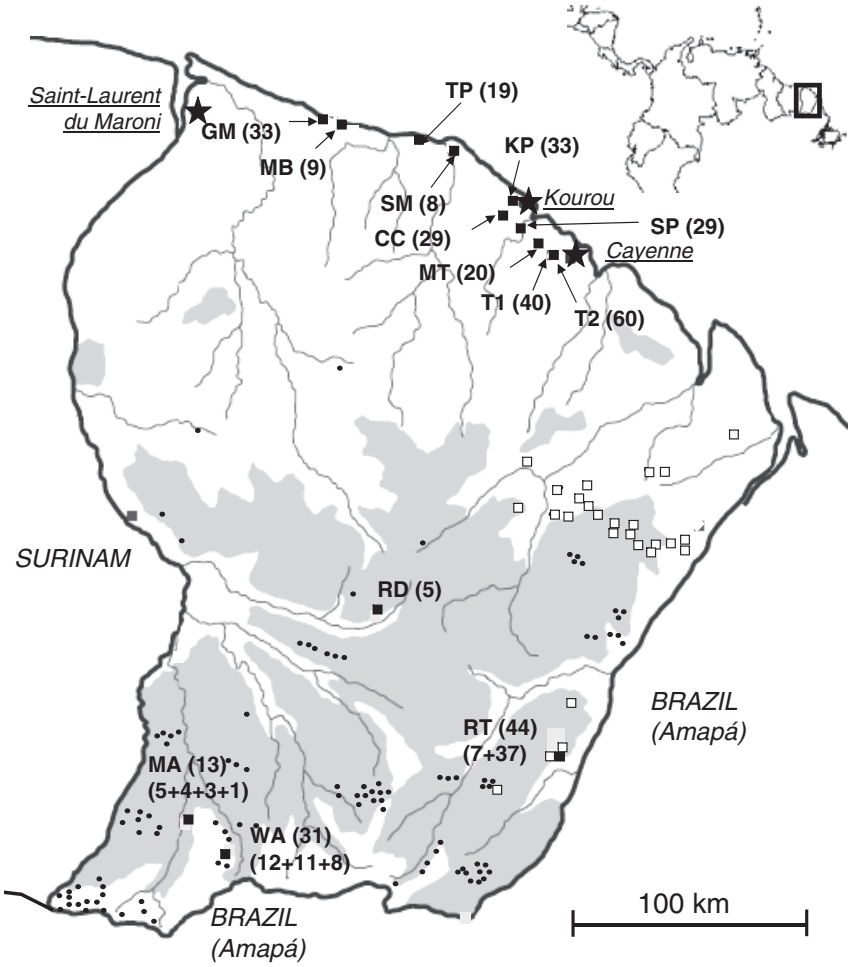

Fig. 1 Map of the sampling locations: filled squares indicate locations where *Manihot* was sampled; open squares indicate inselbergs where there was no *Manihot*. Unsampled inselbergs appear as dots. Numbers of individuals sampled are indicated in brackets. For the inland inselberg populations, when several close inselbergs were sampled, the numbers of individuals sampled on each is indicated after the total number for the population. GM, Grand Macoua; MB, Savane Mammaribo; TP, Trou Poissons; SM, Savane Manuel; CC, Carapa Hill; KP, Kourou Piste; SP, Savane des Pères; MT, Savane Matiti; T1, Tonate, population 1; T2, Tonate, population 2; RD, Roche Dachine; RT, Roche Touatou; WA, Wanapi; MA, Haut-Marouini. Grey areas correspond to altitudes > 200 m above sea level.

Aguirre *et al*. 1998), SSR55, SSR68 and SSR169 (Mba *et al*. 2001). All loci were amplified jointly, using the multiplex polymerase chain reaction *Taq* from QIAGEN, following the manufacturer's recommendations for a final volume of 10 μL. Amplifications were conducted on PTC-100 thermocyclers (MJ Research) and genotyping was performed on an ABI 3130 sequencer. Genotypes were then checked by eye under GeneMapper 3.0 software (Applied Biosystems). Control blank wells never showed peaks, and no mismatch between the replicates was detected. Typing error rate was therefore low.

Data analysis

Before performing any analysis, we first looked for individuals introgressed by domesticated cassava. Duputié *et al*. (2007) showed that, although the microsatellites we used were designed for domesticated cassava, they amplify readily in the wild taxon of French Guianan savannas, with alleles different from those found in domesticated cassava. Therefore, we sought individuals bearing alleles characteristic of domesticated individuals. The individuals used as a reference for domesticated alleles were a combination of the cultivated cassava clones used in Duputié *et al*. (2007), and of over 400 cassava plants collected in different locations all over French Guiana (unpublished data).

Allelic richness A_R [computed using the rarefaction method as suggested by El Mousadik & Petit (1996), and averaged over all loci], Weir and Cockerham's *F* statistics (*f*, estimator of F_{IS}, and θ, estimator of F_{ST}; Weir & Cockerham 1984), and Nei's gene diversity *H* (expected heterozygosity, Nei 1987) were computed with FSTAT (Goudet 1995). *P* values for *f* were assessed using the exact test implemented in GenePop (Raymond & Rousset 1995) with default settings. Comparisons of allelic richness and gene diversities were conducted in FSTAT using 10 000 permutations of genotypes among populations. Population differentiation tests were conducted in FSTAT, by permuting genotypes among populations 5000 times. As multiple tests were performed, significance of the *P* values was assessed using Benjamini & Hochberg's (2000) sharpened test, as implemented in Verhoeven *et al*. (2005). Isolation by distance was assessed according to Rousset (1997) in GenePop.

Genetic structure of the populations was further explored using the Bayesian clustering algorithm implemented in Structure version 2.2 (Pritchard *et al*. 2000). The optimal number of clusters was assessed using the method of Evanno *et al*. (2005), running the program 10 times for each value of the allowed number of clusters *K*, with *K* varying from 1 to 10. The program was given no prior information on populations, and run under the admixture model with correlated allelic frequencies, with 100 000 Markov chain Monte Carlo iterations and a burn-in of 10 000 iterations.

Selfing rate was estimated using the maximum-likelihood method of David *et al*. (2007), implemented in the software RMES. This method is not biased by the presence of null alleles. Tests for bottlenecks could not be performed because Hardy–Weinberg equilibrium was not met in most populations and for most loci (Cornuet & Luikart 1996). Interpopulation genetic distances D_{CE} (Cavalli-Sforza & Edwards 1967) were computed using the software Populations (Langella 1999), and a neighbour-joining tree of populations was drawn after 100 bootstraps on individuals.

Results

Identification of hybrids and introgressed individuals

The alleles encountered in wild and domesticated populations were strongly differentiated [as already found by Duputié *et al*. (2007)]. This excludes the possibility that historical gene flow between the two taxa could have blurred the relationships between populations of the wild taxon.

Individuals bearing alleles characteristic of domesticated cassava were encountered in three coastal populations: three individuals in MB bore one allele characteristic of domesticated populations, and 13 individuals in each of populations T1 and T2 exhibited one to three typical domesticated alleles. These 29 individuals had genotypes consistent with their being backcrosses of hybrids (F_1 or later generations) with wild *Manihot*. They were removed from subsequent analyses (based on 344 individuals), together with locus GA57, which was fixed for a single allele in all remaining individuals.

Removing all populations where introgressed individuals have ever been found [T1, T2 and MB (this study), SM (Duputié *et al*. 2007) and MT (A. Duputié, unpublished)] did not change the conclusions (see Supporting information).

Intrapopulation genetic diversity

A total of 36 alleles (4–9 per locus) were encountered overall. Allelic richness A_R, population fixation index *f* (estimator of F_{IS}) and Nei's gene diversity *H* for all 14 populations are presented in Table 1. On average, allelic richness was higher in inselberg and western coastal populations (GM, MB, TP, SM) than in eastern coastal populations (Table 1; Mann–Whitney tests, inselberg vs. eastern populations, $P = 0.010$; western vs. eastern populations, $P = 0.028$, respectively). Over all populations, allelic richness ranged from 1.95 (for locus GA21) to 3.88 (for locus SSR68). Gene diversity followed the same pattern as allelic richness.

There was a strong overall heterozygote deficit, with $f = 0.183$ (95% confidence interval, 0.106–0.295). Eight

Table 1 Number of individuals N, allelic richness A_R, gene diversity H and intrapopulation fixation index f for the 14 sampled populations and the associated P value [exact test as implemented in GenePop (Raymond & Rousset 1995); P values < 0.05 are indicated in bold]. Population abbreviations are the same as in Fig. 1

	Inselbergs				Coast (west to east)									
	MA	RD	RT	WA	GM	MB	TP	SM	CC	KP	SP	MT	T1	T2
N	13	5	44	31	33	5	19	8	29	33	29	20	27	47
A_R	2.538	1.767	2.246	2.503	2.378	2.24	2.462	1.897	1.304	1.343	1.75	1.811	1.855	2.011
H	0.571	0.313	0.424	0.489	0.514	0.531	0.514	0.354	0.101	0.112	0.339	0.335	0.317	0.331
f	0.387	0.147	0.205	0.231	0.149	0.183	0.077	0.294	0.081	–0.049	0.365	0.347	0.007	0.101
P value	**< 0.001**	0.392	**0.001**	**< 0.001**	**0.016**	0.144	0.25	**0.024**	0.399	0.743	**< 0.001**	**< 0.001**	0.407	**0.002**

Table 2 Pairwise F_{ST} values between all populations (lower-left matrix) and their significance (upper-right matrix, Benjamini and Hochberg's sharpened test, $\alpha = 0.05$; NS, nonsignificant population differentiation)

		Inselbergs				West of Kourou				Kourou		Near Kourou		Tonate (east)	
		MA	RD	RT	WA	GM	MB	TP	SM	CC	KP	SP	MT	T1	T2
Inselbergs	MA	—	NS	*	*	*	NS	*	*	*	*	*	*	*	*
	RD	0.117	—	*	NS	*	NS	*	*	*	*	*	*	*	*
	RT	0.068	0.131	—	NS	*	*	*	*	*	*	*	*	*	*
	WA	0.045	0.109	–0.002	—	*	*	*	*	*	*	*	*	*	*
West of Kourou	GM	0.142	0.365	0.289	0.249	—	NS	*	*	*	*	*	*	*	*
	MB	0.103	0.409	0.334	0.281	0.075	—	*	*	*	*	*	*	*	*
	TP	0.177	0.396	0.345	0.3	0.039	0.105	—	*	*	*	*	*	*	*
	SM	0.244	0.554	0.388	0.341	0.077	0.194	0.076	—	*	*	*	*	*	*
Kourou	CC	0.5	0.809	0.545	0.524	0.422	0.471	0.483	0.516	—	*	*	*	*	*
	KP	0.5	0.8	0.549	0.531	0.421	0.465	0.48	0.489	0.041	—	*	*	*	*
Near Kourou	SP	0.316	0.574	0.455	0.418	0.253	0.136	0.194	0.245	0.301	0.299	—	*	*	*
	MT	0.277	0.578	0.442	0.411	0.143	0.126	0.122	0.114	0.41	0.374	0.179	—	*	*
Tonate (east)	T1	0.377	0.63	0.502	0.473	0.287	0.276	0.249	0.342	0.582	0.566	0.253	0.208	—	*
	T2	0.373	0.612	0.493	0.467	0.282	0.265	0.216	0.272	0.484	0.464	0.175	0.164	0.03	—

of the 14 populations exhibited a significant deficit of heterozygotes. Both inland and coastal populations were affected. This deviation from Hardy–Weinberg equilibrium can be due to high selfing rates, to intrapopulation genetic structuring (Wahlund effect, notably in the inselberg populations, three of which are aggregates of populations found on different rocks), or to the presence of null alleles. High selfing rates are excluded by an analysis using RMES software, as values (results not shown) in no case differed significantly from zero. Therefore, *Manihot esculenta* ssp. *flabellifolia* is mainly an outcrosser. Both latter explanations for heterozygote deficit are plausible. Indeed, population substructure at a very short spatial scale (less than 75 m) has already been observed in one population of this taxon (population SM, Duputié *et al*. 2007). Null alleles also are likely to be present, because the primers we used had initially been designed for domesticated cassava. Removing the locus exhibiting the highest frequency of null alleles did not modify the results presented here (see Supporting information).

Population structuring

Population differentiation was high overall: $\theta = 0.363$ (95% confidence interval, 0.258–0.427). Pairwise F_{ST} values are given in Table 2, together with their significance. Most tests for population differentiation were significant. The only two large populations which were not differentiated were two inselberg populations (RT and WA).

Isolation by distance was significant when all populations were included [regression of $F_{ST}/(1 - F_{ST})$ with ln(distance), Mantel test after 10 000 permutations, 14 populations, $P = 0.041$; Fig. S1, Supporting information]. However, as populations from Kourou (KP, CC) were very strongly differentiated from all other populations, irrespective of distance (note the high values

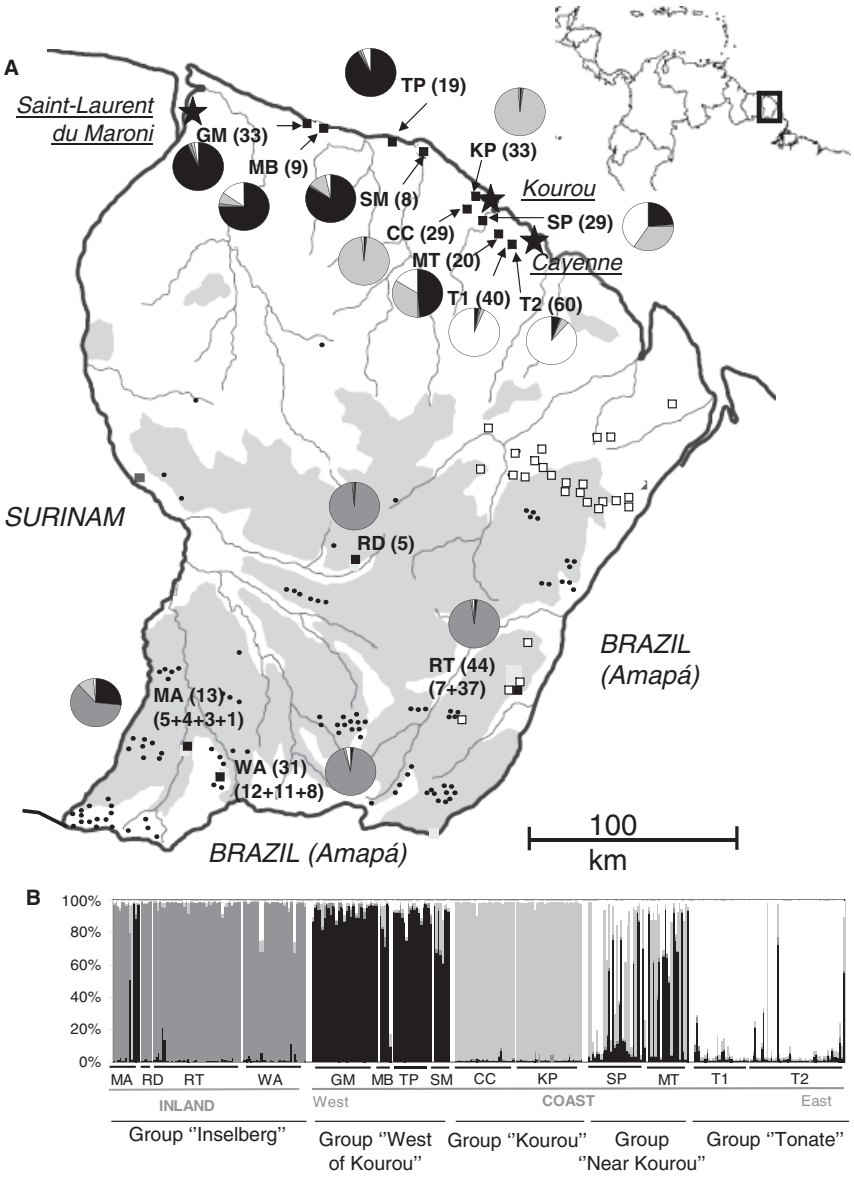

Fig. 2 (A) Map of the populations showing the average proportion of the genome inferred by Structure to be drawn from each of the four clusters. (B) Proportion of the genome of each individual assigned to each of the four clusters. Each individual is represented by a vertical bar. Grey shades are as in panel A.

of F_{ST} in Table 2), the isolation-by-distance pattern was much stronger when they were removed from the analysis (12 populations, $P = 0.002$; Fig. S1). Among inselberg populations, there was no evidence for isolation by distance ($P = 0.795$ but $N = 4$ populations; Fig. S1). Among coastal populations, isolation by distance was not significant when the populations from Kourou were included (regression of $F_{ST}/(1 - F_{ST})$ with geographic distance, 10 populations, Mantel test, $P = 0.138$), but was when these populations were removed (8 populations, $P = 0.011$; Fig. S1).

The optimal number of clusters formed by Bayesian clustering was four. Most individuals had a clear ancestry in one of the clusters (with Structure inferring that more than 90% of their genome was drawn from this cluster). For each population, the average fraction of genome assigned to each of the four clusters is presented in Fig. 2A. All inland individuals were assigned to cluster 1, except four individuals from Marouini. Coastal individuals from populations located west of Kourou fell into cluster 2, populations from Kourou were gathered in cluster 3, and populations from Tonate were classified in cluster 4. Individuals of populations SP and MT shared their ancestry between clusters 2, 3 and 4. When Structure was allowed to form a fifth cluster, these individuals remained of mixed ancestry. Individual assignments are presented in Fig. 2B.

From this analysis, we identified five genetic groups, four of which correspond to the clusters identified by Structure, the fifth one including the populations of unclear ancestry (SP and MT) (see map in Fig. 3). Allelic richness, number of private alleles, f and Nei's

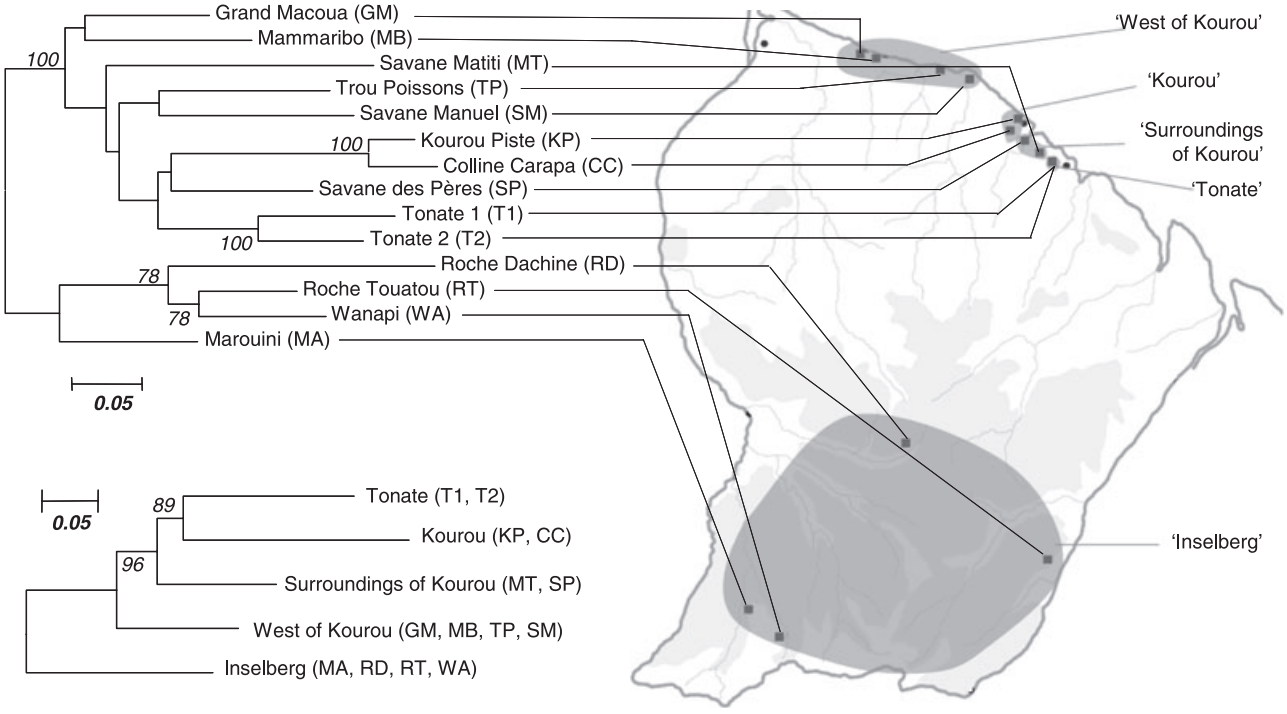

Fig. 3 Neighbour-joining trees of populations (upper left) and groups of populations (lower left) based on D_{CE} distance (Cavalli-Sforza & Edwards 1967). Trees are based on 100 bootstraps on individuals. Only bootstrap values > 50% are shown. Abbreviations correspond to those in Fig. 1. The five genetic groups identified by Structure are mapped (right-hand panel).

	Inselberg (MA, RD, RT, WA)	West of Kourou (GM, MB, TP, SM)	Kourou (CC, KP)	Near Kourou (SP, MT)	East of Kourou (T1, T2)
A_R	3.93	3.90	1.83	2.65	3.30
PA/TA	5/25	5/27	1/11	0/15	1/19
H	0.48	0.52	0.11	0.37	0.33
f	0.27	0.18	0.03	0.42	0.08
P value	0.001	0.001	0.379	0.001	0.034

Table 3 Allelic richness A_R, number of private alleles and total number of alleles *PA/TA*, gene diversity *H*, intrapopulation fixation index *f* and associated *P* value (exact test performed in GenePop) for the five genetic groups of *Manihot esculenta* ssp. *flabellifolia* detected in French Guiana

genetic diversity for these five groups are presented in Table 3. Populations of the groups 'Inselberg' and 'West of Kourou' each bore five private alleles (i.e. alleles that were not encountered in any other group), whereas only one private allele was found in each of the groups 'Kourou' and 'Tonate', and none in the 'Near Kourou' group. All groups except 'Kourou' exhibited significant heterozygote deficits.

As tests for bottlenecks require Hardy–Weinberg equilibrium in the populations (Cornuet & Luikart 1996; Luikart *et al*. 1998), no tests were performed. Nevertheless, very low allelic richness in the populations from Kourou (KP, CC) and the fixation in these populations of a single allele at three loci that are polymorphic in other populations suggest that these two populations are issued from recent bottleneck events.

Coast–inselberg population differentiation

While Structure evidenced that coastal populations could be organized into four distinct genetic groups (one of which was of mixed ancestry), the population tree inferred from D_{CE} distances highlighted another aspect of genetic structure of the taxon in French Guiana (Fig. 3). Inselberg and coastal populations were clearly separated: although several groups were identified among coastal populations, all of them belong to the same genetic unit, and inselberg populations to another one. Among coastal populations, the two westernmost populations were basal, supported by a bootstrap value of 100%.

Out of the 36 alleles we recorded, five were private to inselberg populations and 11 to coastal populations, further indicating a long isolation between inselberg and

coastal populations. As most of these alleles were rare, smaller sample size in inselbergs than on the coast probably prevented the detection of some private alleles in inselberg populations.

Discussion

Hybridization and introgression between wild Manihot *and domesticated cassava*

Hybridization between the two taxa has already been shown to occur in several savannas from coastal French Guiana (Duputié *et al*. 2007). Here, we found evidence for introgression of the wild taxon by domesticated genes in three further populations, all in the coastal savannas: T1, T2 and MB. These populations grow in human-altered environments (roadside, edge of a soccer field) and a few plants of domesticated cassava have been observed in their vicinity. Interestingly, no evidence for wild-crop hybridization was found in other zones of close sympatry between the wild and the domesticated taxa, such as in population KP.

Inselberg populations did not present any evidence of hybridization. On the one hand, this was not surprising, as inselbergs lie in very isolated areas, far away from any village. On the other hand, until recently the Amerindians, who have been the only inhabitants of inner French Guiana for the past centuries, moved their villages periodically. Therefore, some Amerindian villages could have been located at some time next to inselbergs where wild *Manihot* grew, or Amerindian farmers may even have planted wild plants together with cassava, as has occasionally been observed in recent times (J. Chave, personal communication). Gene flow would have been facilitated by such situations; however, if they occurred, neutral alleles issued from the domesticated taxon have disappeared since.

Evidence for past large savanna areas in southern French Guiana

All inselberg populations were classified in a single group by Bayesian clustering. Unexpectedly, differentiation was less marked within the group of inselberg populations than between inselberg and coastal populations, or among coastal populations. There was no isolation by distance between inselberg populations, which were very weakly differentiated.

This pattern could be explained by long-distance pollen or seed flow between inselberg populations, or from large source populations, which would homogenize French Guianan *Manihot* inselberg populations. However, long-distance pollen or seed flow seems inconsistent with the biology of the species and the patchy nature of the suitable habitats: inselbergs are often isolated from one another by several tens of kilometres. Further sampling of inselbergs, as well as analysis of *Manihot* samples from the neighbouring regions of Amapá, Suriname and Guyana, would help gain insight into the mechanisms by which it colonizes inselbergs.

Alternatively, shared ancestral polymorphism could explain the striking genetic homogeneity of inselberg populations: *Manihot* populations observed on inselbergs in the southern region could be remnants of formerly more extended, homogeneous populations. Connections between inselberg populations could have resulted either from the presence of large savanna areas, or from occasional large-scale fires occurring in a mosaic of dry and gallery forest, creating labile and moving open areas where *Manihot* would have been able to grow. Several large-scale fire episodes have been recorded at various locations in French Guiana (Tardy 1998), even in zones continuously covered with rainforest (Charles-Dominique *et al*. 1998). *Manihot* populations from southern French Guiana, now found only on inselbergs, could therefore have been connected, at least periodically, either by a large savanna area, or by periodically disturbed dry forest areas. Demographic stochasticity, combined with genetic drift, may have led to local extinctions of small populations. Low frequency of gene flow — most probably through pollen — between inselbergs is nevertheless required to account for the homogeneity that is still observed between these inselberg populations.

Our hypothesis is consistent with the conclusions of Bush (1994), Dutech *et al*. (2004), and Noonan & Gaucher (2005) that southern French Guiana must not have been a stable forest refugium area. Because only one species is studied here, the patterns that emerge from this study need to be confirmed by the study of other species, either restricted to open or dry environments, or to forest habitats. Further studies of population structure of other taxa restricted to inselbergs could help validate this scenario and yield further insights into past climate in the southern part of French Guiana.

Colonization of the coastal savannas

Population differentiation was much higher between inselberg populations and any coastal population than among inselberg populations. Similar coast–interior isolation has been documented for the forest tree *Vouacapoua americana* (Dutech *et al*. 2004) and for the poison frog genus *Atelopus* (Noonan & Gaucher 2005). As coastal savannas were formed only around 5000 BP (Tissot *et al*. 1988), this raises the question of how they were colonized by *Manihot*. Although it does not include populations from outside French Guiana, the neighbour-joining population tree suggests that western coastal populations have been isolated for a longer time than eastern ones. Together with our find-

ings of higher genetic diversity and higher number of private alleles in western than in eastern populations, this would be consistent with *Manihot* having colonized the coastal savannas from west to east, through a series of bottleneck events. Sampling of the neighbouring regions (Amapá, Suriname, Guyana and Venezuela) is needed to confirm this hypothesis. Significant isolation by distance among coastal populations indicates low gene flow between coastal *Manihot* populations (consistent with the hypothesis that long-distance gene flow is probably limited).

As inselberg and coastal *Manihot* populations are strongly differentiated, with no occurrence of the taxon on inselbergs north of Saül, and as the central part of French Guiana seems to have remained forested for the past 10 000 years (Charles-Dominique *et al*. 1998, 2003), the central region probably remained somewhat wetter than the southern region, at least after the LGM [as has been proposed by de Granville (1982) and Dutech *et al*. (2004)]. Therefore, the colonization of coastal savannas by *Manihot* populations probably did not occur directly from the inselbergs inland, but rather from savannas located westwards. These 'source' savannas could correspond to the Sipaliwini and Rupununi savannas of inland Suriname and Guyana (de Granville 1982).

Conclusion

Our data strongly support a long period of isolation between inselberg and coastal populations of *Manihot esculenta* ssp. *flabellifolia*. While inselberg populations are strikingly homogeneous for neutral genetic markers, coastal populations show a strong pattern of isolation by distance. As the hypothesis of ongoing gene flow from unsampled large source populations seems incompatible with the low gene dispersal abilities of the species, we hypothesize that inselberg populations of southern French Guiana were connected with one another until recent times, either through continuous savanna areas or through periodically disturbed dry forest areas. The taxon appears to have colonized coastal savannas in the past few thousand years, from the western savannas of inner Suriname or Guyana, through a succession of bottlenecks. Although they should be completed by the study of other taxa, and of *Manihot* samples from savannas westwards, our results therefore argue in favour of the occurrence of formerly drier vegetation in the southern part of French Guiana, whereas the central and northern parts have probably remained forested since the LGM.

Acknowledgements

We are grateful to Pierre Charles-Dominique, Jérôme Chave, Olivier Hardy, Finn Kjellberg, Philippe Jarne, Hélène Joly and François Massol for critical reading of previous versions of this manuscript. Three anonymous reviewers helped us improve previous versions of this manuscript. We thank Guillaume Léotard and Benoît Pujol for the loan of *Manihot* samples, and Corinne Sarthou and Philippe Gaucher for collecting *Manihot* from inselberg populations during their expeditions to study other organisms. This study was funded by grants to D. McKey from various agencies of the French Government (Mission pour la Création du Parc de la Guyane, programme 'Écosystèmes Tropicaux' of the Ministry of Ecology and Sustainable Development, programme 'Impact des Biotechnologies dans les Agroécosystèmes' of the Ministry of Research and Contrat Plan État Région Guyane). A. Duputié was supported by a doctoral thesis grant from the French Ministry of Research. Genotyping was done in the 'Service Commun des Marqueurs Moléculaires en Écologie' of the CEFE, with the kind help of Laure Benoît and Hélène Vignes. Data used in this work were partly produced through the molecular genetic analysis technical facilities of the IFR 119 'Montpellier Environnement Biodiversité'.

References

Allem AC (1994) The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution*, **41**, 133–150.

Benjamini Y, Hochberg Y (2000) On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioural Statistics*, **25**, 60–83.

Bush MB (1994) Amazonian speciation: a necessarily complex model. *Journal of Biogeography*, **21**, 5–17.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis — models and estimation procedures. *American Journal of Human Genetics*, **3**, 233–257.

Charles-Dominique P, Blanc P, Larpin D *et al*. (1998) Forest perturbations and biodiversity during the last ten thousand years in French Guiana. *Acta Oecologica*, **19**, 295–302.

Charles-Dominique P, Chave J, Dubois MA, de Granville JJ, Riéra B, Vezzoli C (2003) Colonization front of the understorey palm *Astrocaryum sciophilum* in a pristine rain forest of French Guiana. *Global Ecology and Biogeography*, **12**, 237–248.

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.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

David P, Pujol B, Viard F, Castella V, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474–2487.

Duputié A, David P, Debain C, McKey D (2007) Natural hybridization between a clonally propagated crop, cassava (*Manihot esculenta* Crantz) and a wild relative in French Guiana. *Molecular Ecology*, **16**, 3025–3038.

Dutech C, Maggia L, Tardy C, Joly HI, Jarne P (2003) Tracking a genetic signal of extinction-recolonization events in a Neotropical tree species: *Vouacapoua americana* Aublet in French Guiana. *Evolution*, **57**, 2753–2764.

Dutech C, Joly HI, Jarne P (2004) Gene flow, historical population dynamics and genetic diversity within French Guianan populations of a rainforest tree species, *Vouacapoua americana*. *Heredity*, **92**, 69–77.

Elias M (2000) *Sélection naturelle, sélection humaine, et diversité chez une plante domestiquée propagée de façon végétative: le cas de la culture du manioc par les Indiens Makushi du Guyana*. PhD Thesis, Université Montpellier II, Montpellier, France.

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.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, **92**, 832–839.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Gond V, Bernard JZ, Brognoli C *et al*. (in press) Analyse multi-échelles de la caractérisation des écosystèmes forestiers guyanais et des impacts humains à partir de la télédétection spatiale. In: *Actes du Colloque Ecosystèmes Forestiers des Caraïbes*, Martinique, December 5th–10th, 2005. Karthala Editions, Paris. http://www.cirad.fr/ur/ressources forestieres/content/download/8/51/4924/file/Article%20bouclier%20VG.pdf (accessed May 6th, 2009).

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

de Granville JJ (1982) Rain forest and xeric flora refuges in French Guiana. In: *Biological Diversification in the Tropics* (ed. Prance G), pp. 137–158. Columbia University Press, New York.

de Granville JJ, Cremers G, Hoff M (1993) L'endémisme en Guyane française. In: *Phytogéographie Tropicale — Réalités et Perspectives* (eds Guillaumet J, Belin M, Puig H). ORSTOM, Paris, France.

Haffer J (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131–137.

van der Hammen T, Hooghiemstra H (2000) Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Reviews*, **19**, 725–742.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Langella O (1999) Populations version 1.2.30. http://bioinformatics.org/~tryphon/populations/.

Ledru MP (2001) Late Holocene rainforest disturbance in French Guiana. *Review of Palaeobotany and Palynology*, **115**, 161–176.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.

Mayle FE, Beerling DJ, William D, Bush MB (2004) Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 499–514.

Mba REC, Stephenson P, Edwards K *et al*. (2001) Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theoretical and Applied Genetics*, **102**, 21–31.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Noonan BP, Gaucher P (2005) Phylogeography and demography of Guianan harlequin toads (*Atelopus*): diversification within a refuge. *Molecular Ecology*, **14**, 3017–3031.

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

Pennington RT, Lewis GP, Ratter JA (2006) *Neotropical Savannas and Dry Forests: Plant Diversity, Biogeography and Conservation* (eds Pennington RT, Lewis G, Ratter JA). CRC Press, New York.

Pennington RT, Prado DE, Pendry CA (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, **27**, 261–273.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pujol B, Gigot G, Laurent G *et al*. (2002) Germination ecology of cassava (*Manihot esculenta* Crantz, Euphorbiaceae) in traditional agroecosystems: seed and seedling biology of a vegetatively propagated domesticated plant. *Economic Botany*, **56**, 366–379.

Raymond M, Rousset F (1995) GenePop Version 1.2. Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Roubik DW (1989) *Ecology and Natural History of Tropical Bees*. Cambridge University Press, New York.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Tardy C (1998) *Paléoincendies naturels, feux anthropiques et environnements forestiers de Guyane Française du Tardiglaciaire à l'Holocène récent – Approches chronologique et anthracologique*. PhD Thesis. Université Montpellier II, Montpellier, France.

Tissot C, Djuwansah MR, Marius C (1988) Evolution de la mangrove en Guyane au cours de l'Holocène: étude palynologique. *Symposium de l'Association des Palynologues de Langue Française*, Vol. 25, pp. 125–137. Travaux de la Section Scientifique et Technique de l'Institut Français de Pondichéry, Pondicherry.

Verhoeven KJF, Simonsen KL, McIntyre LM (2005) Implementing false discovery rate control: increasing your power. *Oikos*, **108**, 643–647.

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

This work is part of Anne Duputié's PhD thesis, which aims at understanding some ecological and genetic aspects of the domestication of cassava. Marc Delêtre is a PhD student now working on cassava's genetic diversity and how people perceive and manage this crop in Gabon. Jean-Jacques de Granville is a botanist working as a senior researcher at the IRD centre in Cayenne, and Doyle McKey is a tropical biologist interested in co-evolution, be it between humans and their crops, or between other kinds of organisms.

Supporting information

Additional supporting information may be found in the online version of this article:

Fig. S1 Isolation by distance between French Guianan populations of *Manihot esculenta* ssp. *flabellifolia*.

Analysis of population structure, without the populations where introgressed individuals were found.

Analysis of population structure, without the small populations (*N* < 19).

Null allele quantification.

Table S1 Sampling locations for *Manihot esculenta* ssp. *flabellifolia* in French Guiana

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.