Nuclearmicrosatellite variation inMalagasy baobabs (Adansonia, Bombacoideae, Malvaceae) reveals past hybridization and introgression

Jean-Michel Leong Pock Tsy, Roselyne Lumaret, Elodie Flaven-Noguier, Mathieu Sauve, Marie-Pierre Dubois, Danthu Pascal

HAL Id: cirad-00903908
http://hal.cirad.fr/cirad-00903908

Submitted on 13 Nov 2013

ANNALS OF
BOTANY
Founded 1887

Nuclear microsatellite variation in Malagasy baobabs (*Adansonia*, Bombacoideae, Malvaceae) reveals past hybridization and introgression

Jean-Michel Leong Pock Tsy[1,*], Roselyne Lumaret[2], Elodie Flaven-Noguier[2], Mathieu Sauve[2], Marie-Pierre Dubois[2] and Pascal Danthu[1,3]

[1]*DP Forêts et Biodiversité Madagascar (CIRAD, Université d'Antananarivo, FOFIFA), BP 853, Antananarivo 101, Madagascar,*
[2]*CNRS-UMR 5175, Centre d'Ecologie Fonctionnelle et Evolutive, 1919 route de Mende, F-34293 Montpellier Cedex 5, France and*
[3]*CIRAD, UPR 105, Biens et Services des Ecosystèmes Forestiers Tropicaux, Campus de Baillarguet, F-34398 Montpellier Cedex 5, France*
* For correspondence: E-mail: leong@cirad.mg

• *Background and Aims Adansonia* comprises nine species, six of which are endemic to Madagascar. Genetic relationships between the Malagasy species remain unresolved due to conflicting results between nuclear and plastid DNA variation. Morphologically intermediate individuals between distinct species have been identified, indicative of interspecific hybridization. In this paper, microsatellite data are used to identify potential cases of hybridization and to provide insights into the evolutionary history of the genus on Madagascar.

• *Methods* Eleven microsatellites amplified with new primers developed for *Adansonia rubrostipa* were used to analyse 672 individuals collected at 27 sites for the six Malagasy species and morphologically intermediate individuals. Rates of individual admixture were examined using three Bayesian clustering programs, STRUCTURE, BAPS and NewHybrids, with no *a priori* species assignment.

• *Key Results* Population differentiation was coherent, with recognized species boundaries. In the four Malagasy species of section *Longitubae*, 8·0, 9·0 and 9·5 % of individuals with mixed genotypes were identified by BAPS, NewHybrids and STRUCTURE, respectively. At sites with sympatric populations of *A. rubrostipa* and *A. za*, NewHybrids indicated these individuals to be F_2 and, predominantly, backcrosses with both parental species. In northern Madagascar, two populations of trees combining *A. za* and *A. perrieri* morphology and microsatellite alleles were identified in the current absence of the parental species.

• *Conclusions* The clear genetic differentiation observed between the six species may reflect their adaptation to different assortments of climate regimes and habitats during the colonization of the island. Microsatellite variation reveals that hybridization probably occurred in secondary contact between species of section *Longitubae*. This type of hybridization may also have been involved in the differentiation of a local new stabilized entity showing specific microsatellite alleles and morphological characters, suggesting a potential role of hybridization in the recent history of diversification on Madagascar.

Key words: Nuclear microsatellite (nSSR) variation, genetic structure, hybridization, introgression, Madagascar, *Adansonia*, baobab.

INTRODUCTION

Hybridization has played a significant role in plant evolution (Stebbins, 1950; Arnold, 1997). One of the major evolutionary processes to which hybridization may have contributed is speciation in relation to shifts in climate and distribution (Seehausen, 2004). Specifically, hybridization may facilitate the adaptive radiation process during the colonization of new niches, creating novel adaptive genetic combinations due to the emergence of hybrid offspring that show extreme phenotypes not found in either of their parental species, a phenomenon called transgressive segregation (Grant, 1975). Homoploid hybrid speciation may follow sympatric speciation, as in *Coprosma* (Papadopulos *et al.*, 2013), or may happen after secondary contact between two species, as in *Helianthus* (Rieseberg, 2006). Hybridization has also played a major evolutionary role in the radiation of plant polyploid complexes by creating new allopolyploid species (Thompson and Lumaret, 1992). In species with a long life span and thus a high

genetic inertia in the face of potential rapid environmental change, hybridization is a critical process that allows effective genetic responses to different environmental conditions (Petit and Hampe, 2006). In addition, in long-lived species, hybrids have been shown to persist for long periods in the absence of parental species (Lepais *et al.*, 2009) and citations therein; Buerkle, 2009), which may also increase the opportunity for hybrid speciation (Mallet, 2007).

Madagascar is the world's fourth largest island, and is renowned for its species diversity and endemism (Goodman and Benstead, 2006). Due to the wide diversity of climatic and ecological conditions on the island, the native biota provides a fascinating context for the study of speciation and plant radiation (e.g. Janssen *et al.*, 2008). On Madagascar, the trees of the genus *Adansonia* (Bombacoideae, Malvaceae), the baobabs, are prominent in the dry deciduous forests and thickets of the western half of the island (Wickens and Lowe, 2008). Baobab trees may live for more than 1000 years (Pătrut *et al.*, 2007)

and are characterized by outcrossing breeding systems with self-incompatibility (Baum, 1995*a*). Nine species of this genus occur worldwide (Perrier de la Bâthie and Hochreutiner, 1955; Baum, 1995*b*; Pettigrew *et al.*, 2012). One species, *A. gregorii*, is endemic to north-western Australia, whereas *A. kilima*, identified and described by Pettigrew *et al.* (2012), and *A. digitata* are closely related and are native to mainland Africa. *Adansonia digitata* migrated subsequently to numerous surrounding islands, including Madagascar, and throughout the tropics, by natural and human-mediated terrestrial and overseas dispersal (Leong Pock Tsy *et al.*, 2009). It is tetraploid ($2n = 160$), whereas the other species of the genus are diploid ($2n = 88$) (Baum and Oginuma, 1994; Pettigrew *et al.*, 2012). The six remaining species are endemic to Madagascar and constitute a monophyletic radiation (Baum *et al.*, 1998). *Adansonia grandidieri* and *A. suarezensis* have small white flowers which open during the dry season. They are primarily pollinated by nocturnal mammals (fruit bats and lemurs), and constitute section *Brevitubae* (Baum, 1995*a*; Ryckewaert *et al.*, 2011). The other Malagasy species, *A. perrieri*, *A. za*, *A. rubrostipa* and *A. madagascariensis*, have been classified in section *Longitubae*, which also includes the Australian *A. gregorii* (Hochreutiner, 1908; Perrier de la Bâthie and Hochreutiner, 1955; Baum, 1995*b*). The Malagasy species of section *Longitubae* are distinguished by bright-coloured flowers (yellow, orange or red) which open during the wet season and are pollinated exclusively by hawk moths (Baum, 1995*a*; Ryckewaert *et al.*, 2011).

By using molecular clock inferences and fossil pollen records for closely related neotropical taxa, Baum *et al.* (1998) and Baum (2003) estimated that species differentiation in the baobabs endemic to Madagascar started approximately 4 Ma ago. However, conflicting results, specifically the non-monophyly of several species of section *Longitubae*, were obtained based on plastid DNA markers and were attributed by Baum *et al.* (1998) to incomplete lineage sorting rather than to interspecific hybridization. These authors did not observe morphologically intermediate individuals between distinct species and reported that sympatric species of the Malagasy section *Longitubae* do not show overlapping flowering seasons. Over the past decade, we have made an extensive survey of Malagasy baobab species, which has revealed the existence of morphologically intermediate individuals and populations between distinct species of section *Longitubae*, suggesting the occurrence of hybridization.

In the present study, the pattern of genetic structure in Malagasy baobabs was analysed with polymorphic nuclear microsatellites (nSSR loci) from sympatric and allopatric populations of all six *Adansonia* species endemic to Madagascar and from two small populations possessing intermediate morphology between *A. za* and *A. perrieri*. Our objectives were (1) to assess genetic variation within and between the six Malagasy taxa and determine whether the species are highly differentiated for nSSR markers and whether or not species-specific diagnostic alleles can be used to identify hybrids; and (2) to determine whether genetic patterns reveal hybridization between species, and, if so, to evaluate the frequency and geographical distribution of interspecific exchanges and categorize hybrids (e.g. F_1, F_2 or backcrosses). Such information is important for new insights into the conservation biology of Malagasy baobabs and the role hybridization may have played in their diversification.

MATERIALS AND METHODS

Sampling design

Populations were sampled across much of the range of each Malagasy species of *Adansonia* (Fig. 1; Table 1). Sampling occurred in areas, mostly forests, where samples of 20 or more individuals could be obtained. We collected from four to seven populations of each species except for *A. suarezensis* (three populations) and *A. perrieri* (two populations). These two species are confined to restricted areas in the far north of the island. However, the recent discovery of a small isolated relict population of *A. perrieri*, near Bealanana (the hatched red spot in Fig. 1), indicates that this species has a broader range than previously known. *Adansonia rubrostipa* is generally restricted to coastal areas, on well-drained soils, although a population (site 19 in Fig. 1) was observed 150 km inland and just a few kilometres from large populations of *A. za* and *A. grandidieri* growing at lower elevation, mostly in gallery forests. *Adansonia madagascariensis* is mostly found in the north-west along watercourses (Baum, 1995*b*). *Adansonia za* has the largest range and is mostly found in the south (Fig. 1). Species identification was based on morphological characters (Baum 1995*b*).

Thirteen individuals from four populations (numbered 17, 19, 23 and 25 in Table 1) were identified as morphologically intermediate between *A. rubrostipa* and *A. za* (Fig. 2). In addition, small groups of morphologically homogeneous individuals combining characters of *A. za* and *A. perrieri* were observed (Fig. 2) at several sites in the north. Two of these sites are indicated by black stars in Fig. 1. These individuals have white flowers with intermediate width petals which turn pale yellow with age, as in *A. perrieri*, and an androecium with a short tube and long free filaments, as in *A. za*. Moreover, the trunks of these trees were either completely white or showed a much lighter grey colour than found in *A. za* and *A. perrieri*. A specimen collected at site 14 was deposited at the TEF Herbarium of DRFP-FOFIFA (Antananarivo, Madagascar) with reference number MIGI 13.

When a given species was observed >50 km from another baobab species, the respective populations were considered allopatric. In sympatric areas, species were classified as either mixed or close, depending on whether they were observed in the same site or not, respectively. Information on non-overlapping flowering periods between sympatric baobab species is presented in Table 1. As a comparison, 30, 13 and 31 individuals from three *A. gregorii* populations located in north-western Australia, at Broome, Willare and Telegraph Hill, respectively, were scored for the same genetic markers as those used to analyse the Malagasy taxa.

Nuclear microsatellite locus isolation and genotyping in the Malagasy baobab species

Nuclear microsatellites or simple sequence repeats (nSSRs) are codominant genetic markers that show high rates of evolution and are useful for revealing recent interspecies hybridization and introgression. For the majority of nuclear microsatellites developed by Larsen *et al.* (2009) for *A. digitata*, cross-species transfer was not successful for the Malagasy baobabs and, when it was, locus homology was not always consistent. A new microsatellite
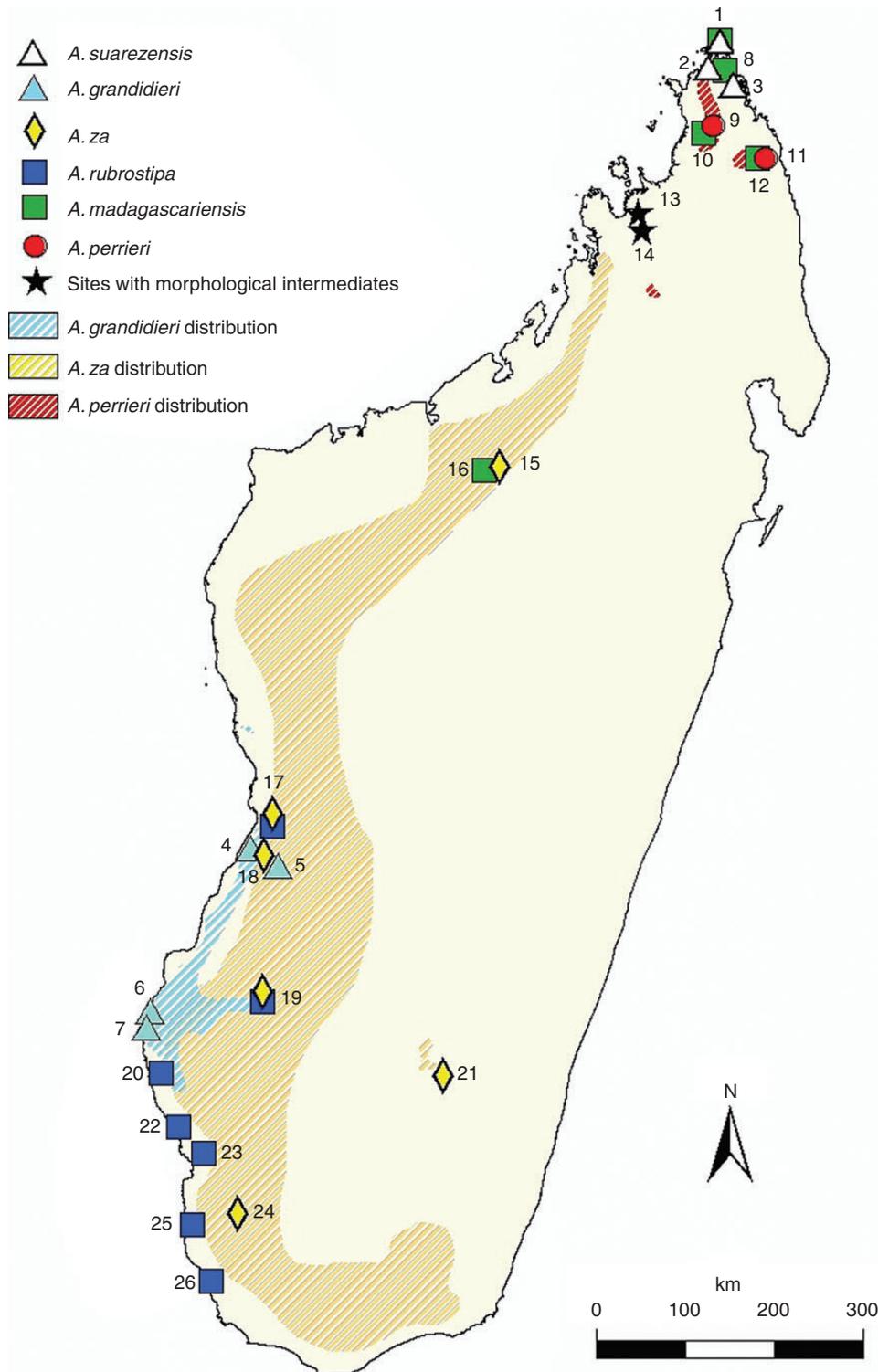
FIG. 1. Geographical distribution of the 26 sites sampled in the study. Species identity is indicated with coloured symbols. Black stars indicate individuals combining morphological characters of *Adansonia za* and *A. perrieri* and sampled at sites 13 and 14. For details of the study sites see Table 1.

library was therefore constructed from the genomic DNA of a single *A. rubrostipa* individual from population 22 (Table 1).

Total plant DNA was extracted from 100 mg of dried leaf tissue preserved in silica gel. The sample was crushed in liquid nitrogen and DNA was extracted using a DNeasy Plant Mini Kit (Qiagen). Isolation of microsatellite loci from an enriched (TG_{10}, TC_{10} and TA_{10}) library was performed following Dubois *et al.* (2005), using biotin-labelled microsatellite

TABLE 1. *Site number, locality, geographical coordinates, sample size and population status of studied populations of* A. suarezensis (Sua), A. grandidieri (Gra) A. madagascariensis (Mad), A. perrieri (Per), A. za (Za) *and* A. rubrostipa (Rub); *information is also presented for morphologically intermediate individuals and the occurrence of multiple baobab species*

			Sample size								
Site no.	Locality	Coordinates	*Sua*	*Gra*	*Mad*	*Per*	*Za*	*Rub*	Int	Status	Additional species
1	Cap Ambre	12 °10′30·66″ S, 49 °14′8·74″ E	22		33					SM*	
2	Beantely	12 °16′8·67″ S, 49 °10′19·32″ E	19							AL	
3	Mahavanona	12 °26′ 44·46″ S, 49 °25′45·37″ E	24							AL	
4	Bekonazy	20 °14′59·22″ S, 44 °25′12·23″ E		31						SC*	*Rub, Za*
5	Ampanihy	20 °25′22·20″ S, 44 °43′15·30″ E		23						SC*	*Rub, Za*
6	Morombe	21 °41′57·62″ S, 43 °26′34·69″ E		17						SC*	*Za*
7	Andavadoaka	22 °3′12·08″ S, 43 °17′16·78″ E		30						AL	
8	Ankorihakely	12 °18′46·98″ S, 49 °20′ 6·23″ E			28					SC*	*Sua*
9	Ambondromifehy	12 °52′9·05″ S, 49 °13′37·31″ E				27				SC*	*Mad*
10	Ankarana	12 °57′7·84″ S, 49 °7′41·83″ E			29					AL	
11	Daraina	13 °17′3·30″ S, 49 °41′9·64″ E				17				SC*	*Mad*
12	Analafiana	13 °27′9·00″ S, 49 °50′27·71″ E			24					SC*	*Per*
13	Ambolobozo	13 °45′23·05″ S, 48 °27′53·89″ E							15 *Za/per*	SC	*Za*
14	Migioko	13 °56′24·94″ S, 48 °31′15·76″ E							10 *Za/per*	SC	*Za*
15	Tsaramandroso	16 °22′4·46″ S, 47 ° 2′56·94″ E					19			SC*	*Mad*
16	Ambahikely	16 °24′12·68″ S, 47 °5′49·30″ E			30					SC*	*Za*
17	Kirindy	20 °4′20·96″ S, 44 °40′29·89″ E					25	32	2 *Za/Rub*	SM	*Gra**
18	Mahabo	20 °19′31·41″ S, 44 °33′54·86″ E					23			AL	
19	Tsitanandro	21 °48′47·73″ S, 44 °32′2·99″ E					21	9	5 *Za/Rub*	SC	*Gra**
20	Ampasikibo	22 °31′51·45″ S, 43 °25′22·70″ E						29		SC	*Za*
21	Ivohibe	22 °32′43·62″ S, 46 °28′39·48″ E					25			AL	
22	Ifaty	23 °4′39·98″ S, 43 °37′2·23″ E						36		AL	
23	Ankiliberengy	23 °20′30·94″ S, 43 °53′39·09″ E						29	3 *Za/Rub*	SC	*Za*
24	Iambory	23 °57′53·99″ S, 44 °15′16·62″ E					27			SC	*Rub*
25	Tsimanampetsotsa	24 °4′36·00″ S, 43 °45′20·87″ E						32	3 *Za/Rub*	SC	*Za*
26	Itampolo	24 °39′5·70″ S, 43 °57′38·61″ E						24		AL	

*Sympatric cases involving baobab species with no overlapping flowering periods.

AL allopatric; SC, sympatric and close; SM, sympatric and mixed.

oligoprobes and streptavidin-coated magnetic beads. Recombinant clones were screened with TG_{10}, TC_{10} and TA_{10} and AGE1 (AAACAGCTATGACCATGATTAC) or AGE2 (TTGTAAAACGACGGCCAGTG) oligonucleotides using a modified PCR method (Waldbieser, 1995). Three hundred and eighty-four clones were screened, 174 of which gave a positive signal; 163 were sequenced using an ABI PRISM 3100 sequencer (Applied Biosystems). Fifty-one sequences including both repeated and flanking regions, allowing the determination of PCR primers, were designed using PRIMER3 (v. 0·4·0).
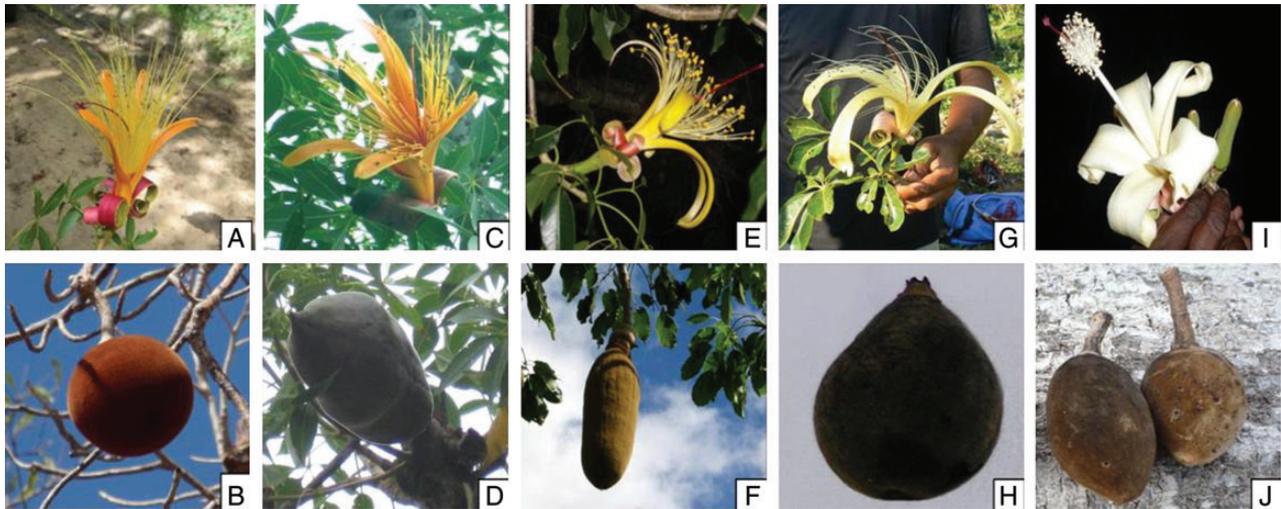
Fɪɢ. 2. Flower and fruit characteristics of *Adansonia rubrostipa* (A, B), *A. za* (E, F) and *A. perrieri* (I, J) and flowers and fruits morphologically intermediate between *A. rubrostipa* and *A. za* (C, D) at site 25 and between *A. za* and *A. perrieri* (G, H) at site 14.

Twelve of the 18 loci with the largest numbers of either perfect or interrupted repeats were finally retained for further amplification based on amplification success and the level of polymorphism. GenBank accession numbers are listed in Supplementary Data Table S2. PCRs were carried out separately for each locus and then multiplexed for four subsets of loci (*Ar01–Ar04*, Ar05–Ar06, Ar07–Ar09 and *Ar10–Ar12*) in 10 µl final volumes including 0·2 µM each primer and 1 µl genomic DNA and using the Qiagen multiplex PCR kit. In each multiplex, the primers were directly labelled using different fluorescent dyes (Supplementary Data Table S2). PCRs were conducted using Mastercycler (Eppendorf) under the following conditions: 15 min activation of HotStart *Taq* DNA polymerase (Qiagen) at 95 °C, 30 cycles including 30 s of initial denaturation at 94 °C, 90 s of annealing at 58 °C for all loci, and 60 s of extension at 72 °C, and then 30 min of final extension at 60 °C. Three microlitres of diluted PCR products (1/100) were pooled in 15 µl of HI-DI™ formamide (Applied Biosystems) and 0·2 µl of GeneScan-500 LIZ size standard, and analysed on an ABI PRISM 3130xL DNA Analyzer (Applied Biosystems) at the IFR 119 facility, University of Montpellier 2, France. Fragment analysis and scoring were carried out using GeneMapper v. 3·7 (Soft Genetics). The DNA fragment sizes measured from the peaks were converted into discrete alleles by comparison with reference lists of allele sizes. For each species and at each locus, the lists were compiled from cumulative frequency plots of size distribution and verified at intervals by sequencing.

Variability was first screened in 36 and 24 *A. rubrostipa* individuals from populations 22 and 26, respectively, and allele segregation was checked by analysing the open pollination progeny of three mother trees collected in population 22. Seeds instead of leaves were analysed for this purpose. The 12 *A. rubrostipa* microsatellite DNA primers were amplified in four, two, one, two, one and one populations of *A. madagascariensis*, *A. za*, *A. perrieri*, *A. grandidieri*, *A. suarezensis* and *A. gregorii*, respectively, under the conditions described above. Finally, the analysis of microsatellite variability was extended to the complete sampling

design presented in Table 1 after being checked for locus homology between the seven species by sequencing and comparing sequence alignments using BioEdit v. 7·0·9 (Hall, 1999). Cycle sequencing was carried out using the same protocol and PCR conditions as described above. Sequences were visualized on an ABI 3730xL automatic sequencer at LGC Genomics. Sequence alignments between species were obtained using the ClustalW module implemented in BioEdit v. 7·0·9. Alignments were subsequently revised manually.

Statistical analyses of microsatellite data

Patterns of genetic variation. At each of the 12 loci, alleles were characterized by the length of the DNA fragments generated during amplification. Failed PCRs were repeated once and unsuccessful amplifications were treated as missing data. Any individual that did not amplify at three or more loci were excluded from the data set. For each Malagasy species, total and private (species-specific) allele numbers were calculated per microsatellite locus and over all loci in all the sampled populations. The corresponding data from the three small populations of *A. gregorii* were added as a point of comparison. In addition, for the species of section *Longitubae*, except *A. perrieri*, allele numbers were averaged separately over allopatric and sympatric populations and differences were tested using Mann–Whitney *U*-tests. Bonferroni correction was used for multiple comparisons.

In the Malagasy species, departure of the inbreeding coefficient (F_{IS}) from zero was tested at each locus for each population using the permutation option of GENETIX 4·05 (Belkhir *et al.*, 2001). Each test involved 10 000 permutations. The program FreeNA (Chapuis and Estoup, 2007) was applied to estimate the frequency of null alleles (non-amplifying alleles, usually due to mutations in the priming sites) in the microsatellite data set. Linkage disequilibrium (LD) between pairs of loci was estimated for each population. LD significance was tested with 1000 permutations using GENETIX 4·05.

Moreover, in the six endemic Malagasy species, allele number per locus and per population (A_m), rarefied allelic richness (A_r),

observed heterozygosity (H_o), the inbreeding coefficient (F_{IS}) and two other standard F statistics, F_{IT} and F_{ST} (Wright, 1965), were estimated per locus and over all loci and were averaged across populations of each species. We used HP-Rare (Kalinowski, 2005) to estimate allelic richness with rarefaction. GENETIX 4·05 was used to estimate the other parameters. However, to infer levels of genetic differentiation in each species, overall unbiased estimates for small populations of H_S, H_T and the standardized measure of genetic differentiation, G'_{ST} (Hedrick, 2005), were calculated per locus using SMOGD (Crawford, 2010) and were averaged over all loci. In addition, F_{ST} values were estimated within and between the Malagasy species in a pairwise comparison. At each locus and overall, the global and pairwise F_{ST} values were computed with and without the excluding null alleles (ENA) correction method of Chapuis and Estoup (2007). The significance of genetic differentiation between species or between populations within a species was tested by comparison of observed F_{ST} with the distribution of F_{ST} under the hypothesis of no genetic structure, using 10 000 random permutations of individuals between species or between populations. The F_{ST} analogue θ of Weir and Cockerham was calculated for each locus and over loci using GENETIX 4·05 (Weir and Cockerham, 1984).

Finally, a neighbour-joining network was reconstructed using PHYLIP v. 3·67 (Felsenstein, 2007) from a pairwise genetic distance matrix for the Malagasy and Australian populations. Bootstrap analysis with 1000 replicates was performed. The stepwise weighted genetic distance (D_{sw}), an extension of Nei's minimum genetic distance (Nei, 1973) that is appropriate for the analysis of highly polymorphic DNA loci (Shriver *et al.*, 1995), was used. In addition, in cases when populations of a species exceeded 3, the correlation between population subdivision estimated by $F_{ST}/(1 - F_{ST})$, and the logarithm of geographical distance (Rousset, 1997) was quantified (Z and R2 values) by a Mantel test using GENETIX 4·05. Relationships between genetic structure and geographical distance were considered as significant when the true Z value was greater than 95 % of the Z values obtained with 10 000 permutations ($P < 5$ %).

Population structure and admixture analyses using multilocus genotype data. In Malagasy *Adansonia* samples and in 74 individuals of *A. gregorii* scored at the same loci, genetic structure was analysed from multilocus genotypes using three different Bayesian approaches. Two clustering analyses were performed using STRUCTURE 2·0 (Pritchard *et al.*, 2000) and BAPS 5·3 (Corander *et al.*, 2003; Corander and Marttinen, 2006) to assign individuals to groups and identify putative hybrids. STRUCTURE uses Markov chain Monte Carlo algorithms (MCMC), whereas BAPS is based on stochastic optimization to infer the posterior mode of genetic structure and does not require multiple trials for each possible group number value (K). In all the analyses, data were used to calculate posterior probabilities of membership to assumed K groups of individuals, without prior information on their taxonomic identity and population of origin. Using BAPS, the maximum number of clusters was initially set to 30 (the assumed maximum number of populations present in the sample). $K = 2–15$ was selected continuously and then at values of 20, 25 and 30. The K value exhibiting the minimum log-likelihood was selected in order to obtain the optimal partition of individuals. Five repetitions were carried out for each K value. Admixture analysis was subsequently performed using the

clusters obtained in the optimal partition with 100 iterations to estimate the admixture coefficients for the individuals under study, plus 100 reference individuals from each cluster, and ten iterations to estimate the admixture coefficients for the reference individuals. Admixture of individuals was based on either mixture clustering or pre-defined clustering by using the K value of the optimal partition and adding the multilocus genotypes identified in populations 13 and 14, which included exclusively morphological intermediate individuals that were not pre-assigned to clusters, as an extra group. With STRUCTURE, the admixture setting with correlated allele frequencies was used and 10^5 burn-in periods and 10^6 MCMC simulations were performed. Data were run with K values ranging from 1 to 10 in order to find the most likely number of clusters. Ten runs were performed for each K value and interpreted with STRUCTURE HARVESTER (Earl and von Holdt, 2012). This program produces a plot of the mean likelihood value per K value and calculates the highest value of the second-order rate of change (ΔK) according to the method of Evanno *et al.* (2005) to detect the number of K groups that best fitted the data set. The results of 30 replicates at the best fit K were post-processed using CLUMPP 1·1·2 (Jakobsson and Rosenberg, 2007) to align the multiple outcomes and determine the optimal clustering. In addition, to visualize the patterns of genetic variation in the baobabs, a factorial correspondence analysis, which portrays the relationships between individuals based on the detection of the best linear combination of allele frequencies, was performed across populations based on the multilocus genotypes using microsatellite loci and GENETIX 4·05. The individuals were mapped onto a plane determined by the two axes accounting for the maximum percentage of total variation.

A third Bayesian model-based clustering approach for identifying hybrids employed NewHybrids v. 1·1beta (Anderson and Thompson 2002). This method identifies hybrid individuals on the basis of the posterior probability (q) of belonging to different pure parental or hybrid categories generated during two generations of potential interbreeding. NewHybrids considers only two hybridizing species. Consequently, analyses with NewHybrids were restricted to three pairs of sympatric or parapatric species of section *Longitubae*. Six distinct genotype classes were defined: pure species I, pure species II, F_1 hybrids, F_2 hybrids and backcrosses with pure species I ($B \times I$) and pure species II ($B \times II$). NewHybrids was run without prior information on the genotype of an individual, using the uniform prior for both θ and the mixing proportion π, and the program was left to run for 100 000 sweeps after burn-in.

The HYBRIDLAB 1·1 program (Nielsen *et al.*, 2006) was used to assess the power of Bayesian approaches and to evaluate whether these methods could accurately identify admixture and classify hybrids. For each species pair, 100 genotypes of each parental species (except *A. perrieri*, for which 40 genotypes were available) showing a q value above 0·98 in STRUCTURE and BAPS were selected to generate ten data sets of 100 genotypes of each parental and hybrid class. Using each data set, the simulated genotypes were then used in STRUCTURE and in BAPS under the same settings as our real data. Ten independent runs with $K = 2$ were performed for each data set and the results were summed up with CLUMPP. The simulated genotypes were also used in NewHybrids to assess the efficiency of the analysis and estimate the threshold value for individual allocation to one of the six genotype classes.

The analyses of simulated data showed minimum q values of simulated parental genotypes, which varied according to species pair and ranged from 0·836 to 0·901 when performed using STRUCTURE and from 0·845 to 0·914 when performed using BAPS. Setting the threshold q value (Tq) to these respective levels, the efficiency (i.e. the proportion of individuals in a group that were correctly identified (Vähä and Primmer, 2006) ranged from 89 and 91 % for *A. za/A. perrieri* to 99 and 98 % for *A. za/A. rubrostipa* using STRUCTURE and BAPS, respectively. The complement (from 1 to 11 %) corresponded to simulated genotypes of backcross hybrids, which showed q values higher than the minimum expected for parental genotypes. To reduce percentage error and increase the power of our analyses, we set a more conservative threshold q value of 0·95, yielding an efficiency of admixture identification ranging between 99·5 and 100 % according to species pair and program. Using NewHybrids with a threshold posterior probability of 0·95 provided a proportion of correctly identified parental individuals ranging from 97·1 % for *A. za/A. perrieri* to 99·2 % for *A. za/A. rubrostipa*. The proportions were slightly smaller for F_1 hybrids (from 87·3 to 92·8 %), and considerably smaller for F_2 hybrids (from 40·2 to 74·1 %) and first-generation backcrosses (from 52·8 to 79·6 %). Therefore, with NewHybrids, following Burgarella *et al.* (2009), threshold values were used in two ways: (1) for each of the three analysed species pairs, a threshold of 0·95 was applied only to the purebred category, assuming that individuals with $q \geq Tq$ are purebreds and that all others are hybrids; and (2) based on simulation results, threshold values of 0·80 and of 0·50 were set for assignment to all the categories of hybrids and first-generation backcrosses, when *A. za/A. rubrostipa* and the other two species pairs were considered, respectively.

RESULTS

Microsatellite polymorphism in Malagasy baobabs

In the present study, transfer of the 12 microsatellites developed for *A. rubrostipa* was successful for the five other Malagasy species and for most individuals of *A. gregorii*. The results for the 12 selected SSR loci in two populations of *A. rubrostipa* are shown in Supplementary Data Table S1. Between four and 15 alleles were identified per locus per population and significant heterozygote deficiency ($P < 0.05$) was observed exclusively at *Ar07* and *Ar10* in the populations of sites 22 and 26, respectively.

At the 12 loci, allele size range in each baobab species is indicated in Supplementary Data Table 2S. Three of the 12 loci (*Ar01*, *Ar07* and *Ar10*) exhibited 1 bp size differences between alleles instead of the 2 bp differences expected from loci with di-nucleotide repeats, and at seven loci (*Ar03*, *Ar05*, *Ar06*, *Ar07*, *Ar08*, *Ar10* and *Ar12*) a 1 bp allele size difference was observed between various species groups. Our sequence data indicate that all these 1 bp differences were due to indels either in the repeat region or in the flanking regions of the microsatellites.

Substantial polymorphism was observed at the 12 loci for each species, with the exception of *Ar07* in *A. suarezensis* and *A. grandidieri*, and at *Ar09*, which was monomorphic in *A. perrieri* and *A. madagascariensis*, whereas two alleles were identified in *A. za* (Supplementary Data Table 2S). Allele sequencing at *Ar09* in the last three species indicated that they share the same specific base change from C to A at position 105, located in the microsatellite repeat region, suggesting that the change may be responsible for stopping additional motif repeats.

Adansonia rubrostipa had the highest total number of alleles (203) and the highest number (57) of private (species-specific) alleles, and the lowest total (83) and private (six) allele numbers were found in *A. suarezensis* and *A. madagascariensis*, respectively. Many private alleles were observed at low frequency but several were predominant, e.g. one allele at *Ar02* and at *Ar07* was observed at 59–82 and 97–99 %, respectively, exclusively in the species of section *Brevitubae*. At *Ar04*, one allele occurring at 97 % was specific to *A. madagascariensis*. Compared with allopatric populations, a higher number of alleles were shared by the species of section *Longitubae* in areas of sympatry ($P < 0.05$).

Genetic diversity and differentiation

In the Malagasy species, 84 of the 340 population/locus combinations showed significant departures from zero for the inbreeding coefficient (F_{IS}), involving all the loci except *Ar08*. Nine and 75 departures were due to heterozygote excesses and deficiencies, respectively. Most of the significant deficiencies were observed at three loci, *Ar02*, *Ar12* and, more specifically, *Ar10*, in several species, and at several loci in two populations, *A. suarezensis* (site 2) and *A. rubrostipa* (site 17). The data obtained at *Ar10* and those from the two populations mentioned above were considered as possibly reflecting the occurrence of null alleles and were removed from the data set to avoid bias in genetic diversity and population differentiation estimates.

TABLE 2. *Mean allele number per locus and per population* (A)$_m$, *rarefied allelic richness* (A)$_r$, *observed heterozygosity* (H)$_o$, *diversity within populations* (H)$_s$, *total diversity* (H)$_T$, *standardized measure of genetic differentiation* (G′)$_{ST}$ *and multi-locus estimation of* F *statistics* F$_{IS}$, F$_{it,}$ *and* F$_{ST}$ *by the jack-knife method over 11 nuclear microsatellite loci in the six baobab species endemic to Madagascar*

Species	A_m	A_r	H_o	F_{IS}	F_{it}	F_{ST}	H_s	H_T	G'_{ST}
A. suarezensis	4·598 (0·590)	4·205 (0·534)	0·607 (0·008)	−0·087 (0·047)	0·032 (0·037)	0·111 (0·031)	0·551 (0·047)	0·596	0·242
A. grandidieri	5·522 (0·586)	5·032 (0·470)	0·565 (0·050)	0·049 (0·012)	0·159 (0·017)	0·116 (0·015)	0·581 (0·072)	0·646	0·334
A. rubrostipa	9·078 (1·655)	7·502 (0·829)	0·769 (0·063)	0·012 (0·020)	0·072 (0·018)	0·061 (0·005)	0·776 (0·055)	0·817	0·421
A. za	8·846 (0·976)	7·975 (0·626)	0·669 (0·099)	0·044 (0·019)	0·103 (0·024)	0·052 (0·014)	0·691 (0·023)	0·737	0·291
A. perrieri	8·222 (1·420)	7·474 (0·650)	0·648 (0·029)	0·068 (0·029)	0·207 (0·041)	0·144 (0·029)	0·679 (0·042)	0·746	0·573
A. madagascariensis	5·820 (0·472)	5·262 (0·251)	0·534 (0·020)	0·040 (0·030)	0·150 (0·045)	0·113 (0·022)	0·549 (0·069)	0·617	0·362

Standard deviations are indicated in parentheses.

When the other 11 microsatellite loci were considered, of the nine population/locus combinations that showed a significant excess of heterozygotes, eight were identified at five loci in the two populations of *A. suarezensis* (sites 1 and 3), which also showed a significant excess of heterozygotes across all loci (Supplementary Data Table 3S). Null allele frequency was estimated to range from 0·00 to 0·11 (mean 0·015), varying according to species and loci, and was assumed not to be associated with significant biases in genetic diversity estimates. In addition, when LD was evaluated in the Malagasy species, and after Bonferroni correction, pairwise loci showed a significant LD value that did not exceed 5·1 %. These pairs were observed in all the species and ranged from four in *A. suarezensis* and *A. perrieri* to 18 in *A. za*. In most cases, LD was not strong: the correlation coefficients were below 0·4 in 95 % of the significant tests.

Allelic richness (A_r), based on a rarefaction sample of 32 genes, was significantly lower in *A. suarezensis* and *A. madagascariensis* than in *A. rubrostipa*, *A. za* and *A. perrieri* ($P < 0.05$), and was significantly higher in the two last species than in *A. grandidieri* ($P < 0.05$) (Table 2). *Adansonia rubrostipa* showed the highest observed heterozygosity, which was significantly higher than in the two *Brevitubae* species ($P < 0.05$) and in *A. madagascariensis* ($P < 0.01$), which had the lowest proportion of heterozygotes. Diversity within populations was significantly higher ($P < 0.05$) in *A. rubrostipa* than in *A. suarezensis*, which shared the lowest diversity value with *A. madagascariensis* (Table 2). Genetic differentiation between populations (G'_{ST}) was high in *A. perrieri* and *A. rubrostipa* and very low in *A. suarezensis* and *A. za*. As shown in Supplementary Data Table S4, which provides detailed values for each locus, significantly higher G'_{ST} and D values were observed in *A. perrieri* at five loci, namely *Ar01*, *Ar04*, *Ar05*, *Ar06* and *Ar08*, than in the other baobab species ($P < 0.001$). In *A. perrieri*, the two populations were genetically very distinct. At the 11 microsatellite loci, in the eastern population (site 11) 87 alleles were observed. Of these alleles, 93 %, including the seven private ones, were shared with the western population (site 9), which possessed 55 additional alleles, most of them observed at high frequency, or even exclusively in the populations of *A. za*.

Pairwise F_{ST} values between populations with and without correction for null alleles were close (Supplementary Data Table S5). All comparisons between populations of different species were significantly different from zero ($P < 0.001$). Mean values grouped per species ranged from 0·048 (in *A. za*) to 0·111 (in *A. perrieri*) within species and from 0·163 to 0·416 between species. Pairwise F_{ST} values between the seven *Adansonia* species ranged from 0·118 for *A. za* and *A. perrieri* to 0·353 and 0·359 for *A. madagascariensis* and the two species of section *Brevitubae*, respectively (Table 3).

Pairwise D_{sw} genetic distances between populations ranged within species from 0·025 (*A. za*) to 0·162 (*A. perrieri*), and between species from 0·117 to 0·479 (Table 3). The unrooted neighbour-joining tree obtained from the pairwise distance matrix is presented in Fig. 3. The bootstrap values of the consensus tree ranged from 37 to 100 %, all but two being >50 %. The 30 populations were grouped into seven distinct clusters, most positioned at the ends of long branches and corresponding to the seven described species. Moreover, populations from sites 13 and 14 were distinct and associated with the *A. za* cluster. The populations from sites 1, 3, 4, 5, 6 and 7 were grouped into two diverging and closely related clusters corresponding to species of section *Brevitubae*. The three *A. gregorii* populations formed a separate clade from the Malagasy species.

A positive correlation between genetic and geographical distances was obtained using the Mantel test over the four populations of *A. grandidieri* ($Z = 7.65$, $P = 0.046$, R2 = 0·634), the five populations of *A. rubrostipa* ($Z = 6.27$, $P = 0.026$, R2 = 0·720), and the five populations of *A. madagascariensis* ($Z = 14.52$, $P = 0.043$, R2 = 0·852). The test was not significant for the six populations of *A. za* ($Z = 9.49$, $P = 0.371$, R2 = 0·011).

Admixture analyses

The model-based clustering analyses, STRUCTURE and BAPS, were first conducted based on the genotype data (11 loci) of the 746 baobab individuals (672 from the six Malagasy species and 74 from *A. gregorii*). Disregarding species assignment of individual trees based on morphology, an optimal number of seven clusters was obtained using STRUCTURE. The approach of Evanno *et al.* (2005) also indicated seven groups, and ΔK (7) was much larger than all other ΔK values (Supplementary Data Fig. S1). The analysis from CLUMPP indicated sufficient agreement between the 30 STRUCTURE runs for $K = 7$ ($H' = 0.965$), and the seven genetically distinct clusters corresponded to our taxonomic assignment of individuals to the seven *Adansonia* species (Fig. 4A). Results from BAPS also revealed an optimal number of $K = 7$ [probability of ($K = 7$) = 1] and the same groups were obtained from the two Bayesian methods. However, 49 and 43 mixed genotypes, i.e. individuals having <95 % probability of being assigned to their own species and corresponding to putative hybrid trees, were identified using STRUCTURE and BAPS (with both

TABLE 3. *Matrices of pairwise* F_{ST} *values between seven* Adansonia *species* (*upper right in italics*) *and range in pairwise distance values between the 30 populations of baobabs analysed at 11 microsatellite loci and grouped according to their respective species* (*lower left*)

| Species	suarezensis	grandidieri	rubrostipa	za	perrieri	madagascariensis	gregorii	Inter
suarezensis	0·081	*0·271*	*0·220*	*0·303*	*0·277*	*0·359*	*0·306*	*0·408*
grandidieri	0·264–0·338	0·033–0·140	*0·229*	*0·284*	*0·279*	*0·353*	*0·298*	*0·372*
rubrostipa	0·230–0·316	0·237–0·374	0·036–0·085	*0·188*	*0·172*	*0·262*	*0·206*	*0·264*
za	0·336–0·344	0·294–0·497	0·192–0·291	0·025–0·085	*0·118*	*0·178*	*0·234*	*0·130*
perrieri	0·315–0·323	0·313–0·386	0·195–0·279	0·117–0·205	0·128	*0·185*	*0·238*	*0·171*
madagascariensis	0·391–0·436	0·380–0·479	0·281–0·377	0·184–0·237	0·186–0·274	0·031–0·162	*0·292*	*0·251*
gregorii	0·312–0·419	0·310–0·458	0·227–0·350	0·259–0·362	0·269–0·356	0·287–0·427	0·086–0·108	*0·313*

Continued
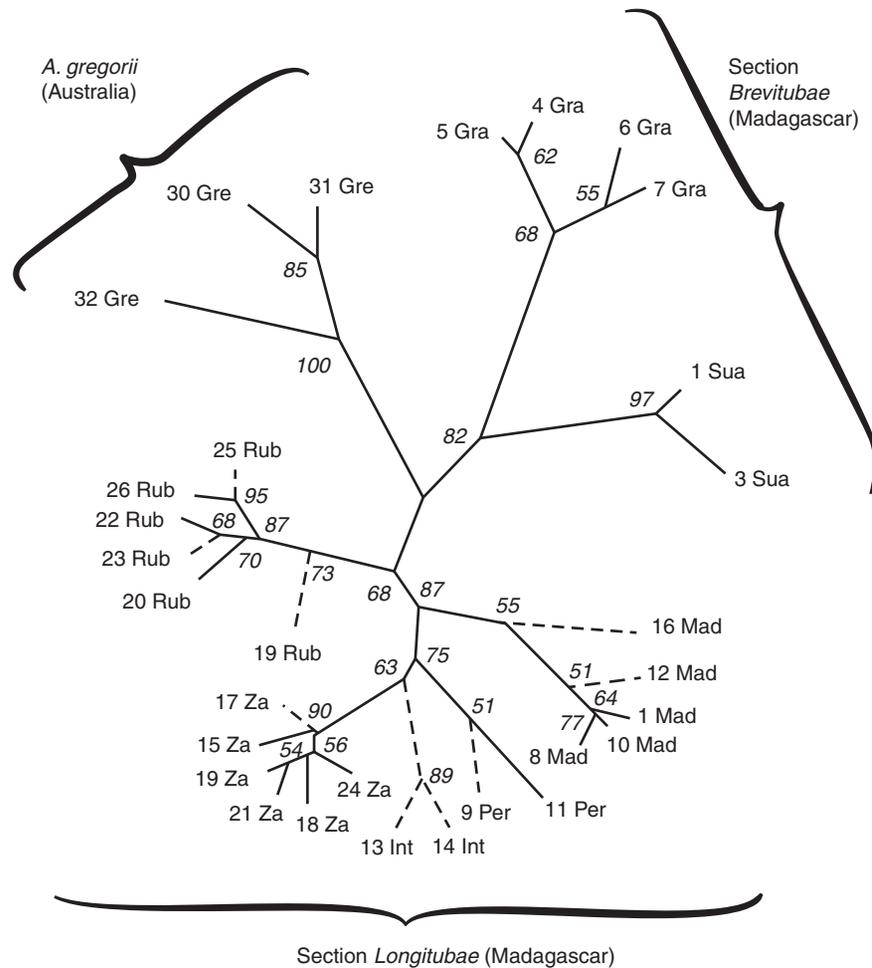
FIG. 3. Neighbour-joining tree using stepwise weighted genetic distances (D_{sw}) between all pairs of *Adansonia* populations included in this study. Site numbers and species identification are presented in Table 1. Populations including individuals with admixed genotypes are indicated with a dashed line. Bootstrap values higher than 50 % are in italics.

admixture methods), respectively. These individuals were restricted to four groups of Malagasy members of section *Longitubae*. The same genotypes were generally identified as mixed by both Bayesian approaches and, in most cases, with approximately the same probability of membership (Table 4).

With NewHybrids, 46 individuals were classified as hybrids (Supplementary Data Tables S4 and S6), while the others were assigned to their respective purebred species with an average posterior probability of 0·993 (range 0·979–0·999) in *A. za/ A. madagascariensis*, 0·980 (0·962–0·999) in *A. za/A. perrieri* and 0·997 (0·992–0·999) in *A. za/A. rubrostipa*.

At four sites (17, 19, 23 and 25), individual trees showed characters morphologically intermediate between *A. za* and *A. rubrostipa*, of which 13 and 12 individuals were identified as putative hybrids by STRUCTURE and BAPS, respectively (Table 4). Moreover, at site 19 a few individuals classified morphologically as *A. za* and *A. rubrostipa* were identified as genetically mixed between the two species by both Bayesian approaches (Table 4). Two, two and six of all these mixed individuals were assigned by NewHybrids to the F_2, backcrosses with *A. za* and backcrosses with *A. rubrostipa* genotype classes, respectively, with posterior probabilities >0·80. An additional

individual was assigned by NewHybrids to the pure *A. za* class with posterior probability >0·95, and the nine remaining individuals, although confirmed as hybrids by their posterior probabilities <0·95 of belonging to a parental form, were not classified with confidence to a specific hybrid class (Supplementary Data Table S6). No individual was assigned to the F_1 hybrid class.

Using STRUCTURE and BAPS, a few genotypes mixed between *A. za* and *A. madagascariensis* and between *A. za* and *A. perrieri* were identified in a population of *A. madagascariensis* (site 16) and in one of *A. perrieri* (site 9), respectively (Table 4). Using NewHybrids, one of the four mixed genotypes identified by STRUCTURE was assigned to pure *A. madagascariensis* and the second to pure *A. perrieri*, with posterior probability >0·95, whereas the other two individuals were confirmed as putative hybrids, but could not be assigned to a single hybrid class (Table 4 and Supplementary Data Table S6).

In addition, using the three Bayesian approaches, the 25 genotypes analysed from sites 13 and 14 were identified as mixed between *A. perrieri* and *A. za*, but having greater affinity to the latter species. Using NewHybrids, one of these individuals was assigned to the F_2 genotype class with posterior probability >0·50, and the other 24 individuals, although confirmed to

FIG. 4. (A) Results obtained using STRUCTURE for seven clusters (*K*) and from the analysis of 746 individual baobabs scored at 11 microsatellite loci and graphically displayed by DISTRUCT (Rosenberg, 2004). Each individual is represented by a thin vertical line partitioned into *K* coloured segments proportional to its membership in the corresponding genetic cluster. Site numbers are indicated above the diagram. (B) Relative positions obtained from the correspondence analysis (CA) of multi-locus genotypes in *Adansonia grandidieri* (*Gra*), *A. suarezensis* (*Sua*), *A. rubrostipa* (*Rub*), *A. perrieri* (*Per*), *A. za* (*Za*), *A. madagascariensis* (*Mad*) and *A. gregorii* (*Gre*) from 21 baobab populations free of admixed genotypes. Axis 1 and axis 2 account for 15·59 and 13·06 % of the total variation, respectively. (C) Relative positions obtained from the CA of genotypes in 21 populations analysed from the four Malagasy species of section *Longitubae*. Axes 1 and 2 account for 20·22 and 12·80 % of the total variation, respectively. Species are indicated using the same symbols as in (B). The admixed genotypes identified from the Bayesian analysis (see part 3A) are shown as yellow squares (Int, intermediate morphology) with an indication of their population number and morphotype.

possess hybrid genotypes by their posterior probabilities <0·95 of belonging to parental classes, were not assigned to a single hybrid class. In both populations, at all loci analysed except *Ar07* the genotypes combined alleles observed in either species, including private ones, or alleles shared by both. Low polymorphism, with the same allele occurring at high frequency (>85 %), was observed at six loci in the two populations. At *Ar08*, the genotypes were characterized by 98 % of alleles never found in *A. za* and predominant in *A. perrieri*, whereas at *Ar02*, *Ar04*, *Ar06* and *Ar11* the alleles were those usually

observed in *A. za*. At *Ar07*, three alleles ranging from 231 to 235 bp were identified exclusively in these two populations and were present in all but three individuals.

Seven well-defined groups of multilocus genotypes were obtained from factorial correspondence analysis, corresponding to the seven baobab species (Fig. 4B). The genotypes of *A. perrieri*, *A. madagascariensis* and *A. za* formed closely related genotype groups. Among the populations of Malagasy *Longitubae* (Fig. 4C), most of the admixed genotypes identified using STRUCTURE (see above) occurred in intermediate

TABLE 4. *Admixed genotype distribution, range of individual assignment* (Q_i) *and inferred ancestry of the 49 individuals identified with admixed origin in the four clusters corresponding to* Adansonia madagascariensis *(Mad),* A. perrieri *(Per),* A. za *(Za) and* A. rubrostipa *(Rub).*

Taxonomic status according to morphology	Site no.	No. of admixed genotypes	Probability of membership (%)				Inferred ancestry
			Mad	*Per*	*Za*	*Rub*	
A. madagascariensis	16	2 (1)	85–94 (72)		06–15 (28)		1 *Mad* [1]
A. perrieri	09	2 (1)	(05)	91–94 (70)	06–09 (25)		1 *Per* [1]
A. za	19	1 (1)			89 (81)	11 (19)	[1]
A. rubrostipa	19	6 (3)	02		10–37 (29–33)	61–90 (67–71)	4 *B* × *Rub* [2]
Intermediate *A. perrieri/A. za*	13	15 (15)	02–04	11–58 (10–52)	40–90 (67–92)		1 F_2 [14]
	14	10 (10)	01–02 (02)	10–49 (10–45)	51–90 (55–90)		[10]
Intermediate *A. za/A. rubrostipa*	17	2 (1)			90–92 (86)	08–10 (14)	1 *Za* [1]
	19	5 (5)	(01)		52–67 (54–72)	33–48 (28–46)	2 *B* × *Za* [3]
	23	3 (3)			20–27 (17–27)	73–80 (73–83)	1 *B* × *Rub* [2]
	25	3 (3)			41–65 (40–68)	35–59 (32–60)	2 F_2, 1 *B* × *Rub*

The Q_i values were calculated in STRUCTURE and BAPS (in parentheses) from individuals analysed at 11 microsatellite loci. Ancestry was calculated in NEWHYBRIDS using six genotype classes [species *I*, species *II*, F_1, F_2, *B* × species (backcross with each of the parental species)]. Numbers of unassigned individuals are indicated in square brackets.

positions between the *A. rubrostipa* and *A. za* genotype groups or between those of *A. perrieri* and *A. za*.

DISCUSSION

Patterns of genetic structure in Malagasy baobabs

With the exception of *A. perrieri*, moderate to low F_{ST} values were obtained between populations of the baobab species. This was particularly the case in *A. za*, which nevertheless has the largest geographical range. Compared with herbaceous plants and shrubs, trees are characterized by high levels of within-population genetic diversity and by high homogeneity between populations, as a result of their outcrossing mating systems, their capacity for extensive gene flow and their large population sizes (Petit and Hampe, 2006). In *A. perrieri*, the western population was genetically divergent from the eastern one (pairwise $F_{ST} = 0.111$). Morphological differences were also observed between the two populations. These included differences in the inside calyx lobe colour, which was entirely white in the eastern population and light or dark pink, as for the other Malagasy species of section *Longitubae*, in the western one. These morphological and selectively neutral microsatellite differences may reflect a long period of isolation between the geographically distant populations, although complete isolation followed by rapid adaptation to new environmental conditions cannot be ruled out.

Compared with section *Longitubae* (except *A. perrieri*), the lower genetic diversity observed in section *Brevitubae*, specifically in *A. suarezensis*, may reflect their restricted geographical range. A significant excess of heterozygotes was identified at many loci in the two populations of *A. suarezensis* for which we also noticed the predominance of older adult trees without fruits and the complete lack of regeneration. In tree species, heterozygosity has been shown to increase with population age, and an excess of heterozygotes has often been found in the oldest age-classes (e.g. Nijensohn *et al.*, 2005). Thus, the excess of heterozygotes in *A. suarezensis* may reflect the general ageing of the

populations. In *A. grandidieri*, genetic variation showed geographic structure: a clear discrimination occurred between the northern populations (sites 4 and 5), which possessed high microsatellite diversity and consisted of tall individuals growing in deep soils, and the two southern populations (sites 6 and 7). The latter populations were characterized by alleles reduced to those most frequently observed in the species, and consisted of small or even dwarf stocky trees growing on limestone in notably dry conditions. Such a genetic difference between the two population groups may reflect founder effects or may indicate that the southern populations have undergone drastic selection.

Microsatellite variation was also structured geographically in all species of section *Longitubae*, except *A. za*. The lack of geographical genetic structure in that species is striking, particularly given its wide distribution (Fig. 1) and its occurrence in different habitats. Baum (1996) assumed that this broad observed distribution was probably related to the high ecological plasticity of the species. In the present study, *A. za* was revealed to possess high microsatellite diversity with low genetic differentiation between populations (Table 2), suggesting a recent expansion of the species.

Significant F_{ST} values were observed between baobab species, associated with a large number of specific alleles, indicating substantial isolation from one another. The Malagasy baobab species were also shown to differ in physiological requirements and habitat preference (Wickens and Lowe, 2008, and references therein), suggesting that each species evolved in its own niche and that ecological differentiation may have contributed substantially to species isolation. This scenario is consistent with most radiation patterns described for different Malagasy endemic groups. After arrival, at early stages of colonization, the founder populations were subjected to a wide variety of climatic regimes and habitats across this mini-continental island, which has been isolated from other landmasses deep in geological time (Leigh *et al.*, 2007; Buerki *et al.*, 2013).

Overall, the phenogram shown in Fig. 3 is consistent with the maximum likelihood phylograms in *Adansonia* of Baum *et al.*

(1998), more specifically the tree based on ITS variation. However, the considerable divergence between *A. suarezensis* and *A. grandidieri* was not evident in the analysis of Baum *et al.* (1998). In addition, for section *Longitubae*, nuclear microsatellite variation provides decisive evidence for close relationships between *A. madagascariensis*, *A. perrieri* and *A. za*, which share numerous alleles and the same nucleotide change at *Ar09*. Consistently with previous studies based on plastid DNA and ITS sequences (Baum *et al.*, 1998; Pettigrew *et al.*, 2012), our results clearly indicate that *A. gregorii* is distantly related to the Malagasy species, suggesting that its classification in section *Longitubae* should be re-examined with additional data.

Species boundaries and evaluation of interspecific genetic exchange

The three Bayesian approaches performed well and consistently identified admixed genotypes from the polymorphism observed at the 11 microsatellites used in this study. However, the uncertainties in the definition of hybrid classes for the species pairs *A. za*/*A. madagascariensis* and *A. za*/*A. perrieri* highlight the inherent difficulties of dealing with closely related species, specifically with a limited number of loci. Overall, our study illustrates that hybridization has occurred in contact zones between baobab species. Approximately 7 % of the sampled individuals possessed admixed genotypes that were restricted to the species of section *Longitubae*. In these species, admixed genotypes were identified in 8·0, 9·0 and 9·5 % of the individuals according to BAPS, NewHybrids and STRUCTURE, respectively. These genotypes were concentrated in sympatric areas where the populations showed higher allelic diversity and lower interspecies variation compared with allopatric populations. This suggests that they are predominantly the result of hybridization, which may be favoured in contact areas due to physical proximity between trees for pollen transfer.

In contact zones between *A. rubrostipa* and *A. za*, 19 admixed genotypes were identified, most of them in the populations of the former taxon. In most cases, these corresponded to individuals showing morphologically intermediate characters. On the basis of the NewHybrids program, the majority were identified as backcrosses, predominantly with *A. rubrostipa*, indicating bidirectional introgressive hybridization. Two geographically close F_2 hybrids were also identified in a large population of *A. rubrostipa* (site 25), whereas F_1 hybrids were not detected in the same area or other analysed populations. F_1 hybrids between Malagasy baobab species are apparently uncommon, suggesting that the formation of first-generation hybrids is less likely to occur than the interbreeding of hybrids with purebreds or with other hybrids. The occurrence of F_2 hybrids and first-generation backcross individuals in three of the five *A. rubrostipa*/*A. za* sympatric populations probably reflects recent hybridization between the two species which overlap in flowering period (Ryckewaert *et al.*, 2011). The absence of admixed individuals outside the contact zones may be due to the occurrence of post-zygotic barriers with substantial selection against the hybrids between the two species, which have distinct habitat preferences (see above). As neutral markers, microsatellites may not constitute the direct targets of selection, but they may be linked to genes under selection.

STRUCTURE detected two admixed genotypes between *A. madagascariensis* and *A. za* in northern Madagascar. One of these was confirmed as admixed by BAPS and NewHybrids but was not assigned to a specific hybrid class or first-generation backcross by the latter. The individual was collected in the only *A. madagascariensis* study population occurring sympatrically with *A. za* (site 16). In this portion of the island, these two species have non-overlapping flowering periods (Baum, 1995*a*; Ryckewaert *et al.*, 2011), making hybridization unlikely. The admixed genotype identified by the Bayesian programs could therefore reflect the occurrence of shared ancestral polymorphism in the two species (Twyford and Ennos, 2012). However, compared with the other four populations analysed in *A. madagascariensis*, the population at site 16 possessed many additional alleles, all predominant in *A. za*, at several loci. In addition, the lower genetic divergence between sympatric/parapatric populations compared with that between allopatric populations is consistent with hybridization. In Malagasy baobabs, the beginning and length of the flowering period primarily depend on the length of the dry season (Baum, 1995*b*). With different climatic conditions in the past, overlapping flowering periods may have occurred between *A. madagascariensis* and *A. za*. Hence, ancient hybridization between the two species cannot be completely ruled out.

Genetic status of the morphologically intermediate individuals observed at sites 13 and 14

The 25 genotypes analysed from the two sites were admixed between *A. perrieri* and *A. za*, with a genomic predominance of the latter in all but one genotype assigned to the F_2 class. These individuals, combining morphological characters of *A. perrieri* and of *A. za*, may have originated from hybridization between the two species, which show substantial overlap in flowering periods and are pollinated by the same long-tongued hawk moth species (Ryckewaert *et al.*, 2011). The most likely explanation for the predominance of the *A. za* nuclear genome is introgressive hybridization. Indeed, backcrossing may have been favoured by the synchronous flowering phenology of early hybrids and *A. za*, by a local demographic superiority of *A. za* or by better gamete compatibility, as previously reported (e.g. Arnold 1997; Lepais *et al.*, 2009). Genetic drift and bottlenecks, more specifically if acting in the first generations after hybridization or in small populations, or if selection is acting on genes closely linked to neutral loci, could also explain the observed predominance of the *A. za* alleles. In addition, neither significant excess of heterozygotes for alleles derived from each species (heterospecific) nor strong parental linkage disequilibrium was detected, which indicates the probable rupture of parental associations and suggests that hybridization of populations at sites 13 and 14 is not recent. This result is also consistent with the findings, in the same region, of five additional small baobab populations showing the same morphological characteristics as those observed at sites 13 and 14. In addition, in all these populations the PCR–RFLP analysis of plastid DNA, maternally inherited in Bombacoideae (Dick *et al.*, 2007; Leong Pock Tsy *et al.*, 2009), revealed the occurrence of a single haplotype. This was identical to the predominant haplotype identified in *A. perrieri* and was distinct from all those observed in *A. za* and in the other *Adansonia* species (J.-M. Leong Pock Tsy *et al.*, unpubl. res.). This result reveals incongruence between

uniparentally and biparentally inherited markers, and constitutes additional support for the hybrid origin of the site 13 and 14 populations. However, *A. za* and *A. perrieri* individuals are not currently identified in the region and are rare in its periphery (Fig. 1), indicative of a contemporary predominance of mating between the individuals originating from the initial hybridization over their backcrossing with the parental species. *Adansonia perrieri* is a rare species threatened by human activities, mainly forest destruction and possibly climate change (Baum, 1995*b*), and is considered an endangered Malagasy baobab species (IUCN, 2012). Our finding of a southern group of old *A. perrieri* trees growing near Bealanana, at high elevation (1170 m) and under notably more moist conditions, suggests that this hydrophilous species probably occurred further south than its current distribution in the recent geological past and may have been in contact with *A. za*. In long-lived trees such as baobabs, hybrids have been shown to persist for long periods in the absence of parental species (Buerkle, 2009; Lepais *et al.*, 2009; Pillon *et al.*, 2009). However, the maintenance of populations derived by hybridization may also reflect their better adaptation to local conditions compared with the parental species. The baobabs of sites 13 and 14 grow in a region currently characterized by compact and periodically flooded soils that are suitable for *A. perrieri* but not for *A. za*, as the latter tends to occur on more well-drained soils and areas with a dry season of at least 7 months (Cornet, 1974), which is too long for *A. perrieri* to survive.

In the baobab populations of sites 13 and 14, little genetic variation was observed at many loci and at *Ar07* most individuals shared the same private alleles, although the possibility that these were not identified in the Malagasy species due to limited sample size cannot be ruled out. The mechanism giving rise to the genetic structure observed in these populations, specifically the predominance of *A. za* alleles at several specific loci and of *A. perrieri* alleles at another locus, is unknown and further studies are needed with an increased number of loci. In addition, no evidence of polyploidy was obtained from genotype examination at the 11 microsatellite loci or from the DNA content and/or chromosome number analysed in three individuals from each population by flow cytometry and chromosome counts, respectively (J. Razanameharizaka, pers. comm.).

The situation observed at sites 13 and 14 contrasts with the substantial morphological variability we observed in transient hybrid or introgressed baobabs, such as at site 19, and may reflect a process of stabilization leading to speciation. Several examples of homoploid hybrid speciation have been reported in plants using population genetic methods (Arnold, 1997; Gross and Rieseberg, 2005), including forest-tree species (Mir *et al.*, 2006), but few on islands (Fjellheim *et al.*, 2009; Papadopulos *et al.*, 2013). In the past, several botanists reported the morphological particularities of the baobabs they encountered in the region where sites 13 and 14 are currently located, and they described them as constituting a distinct species, *A. alba* Jum. & H.Perr., or as *A. za* var. *bozy* (Jum. & H.Perr.) H.Perr. (Jumelle and Perrier de la Bâthie, 1909; Perrier de la Bâthie and Hochreutiner, 1955). Subsequently, Baum (1995*b*) visited the south of the same region and considered that the local baobabs could be assigned to *A. za*. As indicated above, the results of the present study support a stabilized hybrid constitution for the baobabs of that region. These also show morphological particularities and grow in specific habitat conditions,

which suggests they may constitute a genetic entity distinct from *A. za* and *A. perrieri*.

In the present study, analysis of microsatellite variation at 11 loci has provided new insights into the structure of genetic variation in the six endemic Malagasy baobabs. Notable differences between the two recognized sections were found. The species of section *Brevitubae* were isolated geographically from each other and showed low genetic diversity, especially *A. suarezensis*. This species is characterized by a restricted range with ageing populations and, on the basis of simulation analyses integrating climate changes, it has been classified recently as critically endangered with a high probability of extinction in the short term (Vieilledent *et al.*, 2013). Conversely, in species of section *Longitubae*, substantial genetic diversity was found, with interspecific gene flow occurring in contact zones and the development of a new homoploid hybrid entity. Further ecophysiological investigation is required to determine its niche precisely. However, the exclusive location of the hybrid populations in an area where the other seven baobab species found in Madagascar do not occur already constitutes a novel argument for including the area in the Protected Area Network (PAN) of Madagascar, as proposed by Kremen *et al.* (2008). Madagascar is unique on a global scale for its high levels of microendemism (Wilmé *et al.*, 2006; Vences *et al.*, 2009). In plants, few studies have dealt with diversification processes on this island (e.g. Janssen *et al.*, 2008; Strijk *et al.*, 2012). In our study, nuclear molecular markers were used for the first time to document hybridization in Malagasy plant species. Our finding of a new differentiated entity originating from the hybridization of two baobab species indicates that diversification is probably still ongoing in Madagascar, suggesting that the island should not be considered solely as a biodiversity refugium.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxford-journals.org and consist of the following. Fig. S1: plot of mean likelihood L(K) values per K value obtained from the Bayesian analysis of 746 individual baobabs that were genotyped at 11 microsatellite loci using STRUCTURE, and plot for detecting the number of K groups that best fit the data. Table S1: Sample size, allele numbers, observed heterozygosity and inbreeding coefficients in the two populations of *Adansonia rubrostipa* used to screen genetic variability at 12 microsatellite loci. Table S2: Dye colours, GenBank Accession numbers, forward and reverse primer sequences and repeat motifs for 12 microsatellite loci developed for *Adansonia rubrostipa*. Table S3: Genetic parameters in each of the 24 populations analysed at 11 nSSR loci in the six Malagasy baobab species and in the two populations characterized by individuals combining morphological traits of *A. perrieri* and *A. za*. Table S4: Genetic parameters at each of the 11 nSSR loci in the six Malagasy baobab species and in the group of populations characterized by individuals combining morphological traits of *A. perrieri* and *A. za*. Table S5: Mean pairwise F_{ST} values obtained using and not using the ENA correction for 24 Malagasy baobab population analysed at 11 nSSR loci and grouped by described species of *Adansonia*. Table S6: Inferred ancestry calculated in NewHybrids for the 49 individual baobabs identified by STRUCTURE with an admixed origin between baobab species of the *Longitubae* section.

LITERATURE CITED

Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.

Arnold ML. 1997. *Natural hybridisation and evolution.* New York: Oxford University Press.

Baum DA. 1995a. The comparative pollination and floral biology of baobabs (*Adansonia* - Bombacaceae). *Annals of the Missouri Botanical Garden* **82**: 322–348.

Baum DA. 1995b. A systematic revision of *Adansonia* (Bombacaceae). *Annals of the Missouri Botanical Garden* **82**: 440–471.

Baum DA. 1996. The ecology and conservation of the baobabs of Madagascar. *Primate Report*, Special Issue **46**: 311–328.

Baum DA. 2003. Bombacaceae, *Adansonia.* In: Goodman SM, Benstead JP. eds. *The natural history of Madagascar.* Chicago, University of Chicago Press, 339–342.

Baum DA, Oginuma K. 1994. A review of chromosome numbers in Bombacaceae with new counts for *Adansonia* (Bombacaceae). *Taxon* **43**: 11–20.

Baum DA, Small RL, Wendel JF. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae), as inferred from multiple data sets. *Systematic Biology* **47**: 181–187.

Belkhir K, Borsa P, Chikli L, Raufaste N, Bonhomme F. 2001. *GENETIX, logiciel sous Windows TM pour la génétique des populations.* Montpellier, France: Laboratoire Génome, Populations, Interactions CNRS UMR 5000, UMII.

Buerki S, Devey DS, Callmander MW, Phillipson PB, Forest F. 2013. Spatio-temporal history of the endemic genera of Madagascar. *Botanical Journal of the Linnean Society* **171**: 304–329.

Buerkle CA. 2009. Ecological context shapes hybridization dynamics. *Molecular Ecology* **18**: 2077–2079.

Burgarella C, Lorenzo Z, Jabbour-Zahab R, et al. 2009. Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). *Heredity* **102**: 442–452.

Chapuis M-P, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**: 621–631.

Corander J, Marttinen P. 2006. Bayesian identification of admixture events using multi-locus molecular markers. *Molecular Ecology* **15**: 2833–2843.

Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* **163**: 367–374.

Cornet A. 1974. *Essai de cartographie bioclimatique à Madagascar. Notice explicative No. 55.* Paris: ORSTOM.

Crawford NG. 2010. SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources* **10**: 556–557.

Dick CW, Bermingham E, Lemes MR, Gribel R. 2007. Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology* **16**: 3039–3049.

Dubois MP, Jarne P, Jouventin P. 2005. Ten polymorphic microsatellite markers in the wandering albatross *Diomedea exulans. Molecular Ecology Notes* **5**: 905–907.

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

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.

Felsenstein J. 2007. *PHYLIP Phylogeny, inference package, version 3·67.* Seattle, WA: Department of Genome Sciences, University of Washington.

Fjellheims S, Jørgensen MH, Kjos M, Borgen L. 2009. A molecular study of hybridization and homoploid hybrid speciation in *Argyranthemum* (Asteraceae) on Tenerife, the Canary Islands. *Botanical Journal of the Linnean Society* **159**: 19–31.

Goodman SM, Benstead JP. 2006. Updated estimates of biotic diversity and endemism for Madagascar. *Oryx* **39**: 73–77.

Grant V. 1975. *Genetics of flowering plants*, 2nd edn. New York: Columbia University Press.

Gross BL, Rieseberg LH. 2005. The ecological genetics of homoploid hybrid speciation. *Journal of Heredity* **96**: 241–252.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor program for Windows 95/98/NT. *Nuclear Acids Symposium Series* **41**: 95–98.

Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution* **59**: 1633–1638.

Hochreutiner BPG. 1908. Un nouveau baobab et révision du genre *Adansonia. Annales du Conservatoire et du Jardin botaniques de Genève* **1908**: 136–143.

IUCN (International Union for Conservation of Nature). 2012. Red List of Threatened Species *version 2012·1.* Available from http://www.iucnredlist.org. Accessed 1 October 2012.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Janssen T, Bystriakova N, Rakotondrainibe F, Coomes D, Labat JN, Schneider H. 2008. Neoendemism in Madagascan scaly tree ferns results from recent, coincident diversification bursts. *Evolution* **62**: 1876–1889.

Jumelle H, Perrier de la Bâthie H. 1909. Les baobabs du nord-ouest de Madagascar. *Les Matières Grasses* **1909**: 1306–1308.

Kalinowski ST. 2005. HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes* **5**: 187–189.

Kremen C, Cameron A, Moilanen A, et al. 2008. Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science* **320**: 222–226.

Larsen AS, Vaillant A, Verhaegen D, Kjær ED. 2009. Eighteen SSR-primers for tetraploid *Adansonia digitata* and its relatives. *Conservation Genetics Resources* **1**: 325–328.

Leigh EG, Hladik A, Hladik CM, Jolly A. 2007. The biogeography of large islands, or how does the size of the ecological theater affect the evolutionary play? *Revue d'Ecologie (Terre Vie)* **62**: 105–168.

Leong Pock Tsy JM, Lumaret R, Mayne D, et al. 2009. Chloroplast DNA phylogeography suggests a West African centre of origin for the baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). *Molecular Ecology* **18**: 1707–1715.

Lepais O, Petit RJ, Guichoux E, et al. 2009. Species relative abundance and direction of introgression in oaks. Molecular *Ecology* **18**: 2228–2242.

Mallet J. 2007. Hybrid speciation. *Nature* **446**: 279–283.

Mir C, Toumi L, Jarne P, Sarda V, Di Giusto F, Lumaret R. 2006. Endemic North African *Quercus afares* Pomel originates from hybridisation between two genetically very distant oak species (*Q. suber* L. and *Q. canariensis* Willd.): evidence from nuclear and cytoplasmic markers. *Heredity* **96**: 175–184.

Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA* **70**: 3321–3323.

Nielsen EEG, Bach LA, Kotlicki P. 2006. HYBRIDLAB (version 1·0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes* **6**: 971–973.

Nijensohn SE, Schaberg PG, Hawley GJ, Dehayes DH. 2005. Genetic subpopulation structuring and its implications in a mature eastern white pine stand. *Canadian Journal of Forest Research* **35**: 1041–1052.

Pătrut A, Von Reden KF, Lowy DA, et al. 2007. Radiocarbon dating of a very large African baobab. *Tree Physiology* **27**: 1569–1574.

Papadopulos AST, Price Z, Devaux C, et al. 2013. A comparative analysis of the mechanisms underlying speciation on Lord Howe Island. *Journal of Evolutionary Biology* **26**: 733–745.

Perrier de la Bâthie H, Hochreutiner BPG. 1955. Bombacacées (Bombacaceae). In: Humbert H. ed. *Flore de Madagascar et des Comores*. Paris: Société Botanique de France, 1–17.

Petit RJ, Hampe A. 2006. Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics* **37**: 187–214.

Pettigrew JD, Bell KL, Bhagwandin A, *et al*. 2012. Morphology, ploidy and molecular phylogenetics reveal a new diploid species from Africa in the baobab genus *Adansonia* (Malvaceae: Bombacoideae). *Taxon* **61**: 1240–1250.

Pillon Y, Munzinger J, Amir H, Hopkins HCF, Chase MW. 2009. Reticulate evolution on a mosaic of soils: diversification of New Caledonian endemic genus *Codia* (Cunoniaceae). *Molecular Ecology* **18**: 2263–2275.

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

Rieseberg LH. 2006. Hybrid speciation in wild sunflowers. *Annals of the Missouri Botanical Garden* **93**: 34–48.

Rosenberg NA. 2004. DISTRUCT: a program for graphical display of population structure. *Molecular Ecology Notes* **4**: 137–138.

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

Ryckewaert P, Razanamaro O, Rasoamanana E, Rakotoarimihaja T, Ramavovololoma P, Danthu P. 2011. Les Sphingidae, probables pollinisateurs des baobabs malgaches. *Bois et Forêts des Tropiques* **307**: 55–68.

Seehausen O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* **19**: 198–207.

Shriver MD, Jin L, Boerwinkle E, Deka R, Ferrell RE, Chakraborky R. 1995. A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Molecular Biological Evolution* **12**: 914–920.

Soltis DE, Soltis PS, Pires JC, *et al*. 2004. Recent and recurrent polyploidy in Tragopogon (Asteraceae): cytogenetic, genomic, and genetic comparisons. *Biological Journal of the Linnean Society* **82**: 485–501.

Stebbins GL. 1950. *Variation and evolution in plants*. New York, Columbia University Press.

Strijk JS, Noyes RD, Strasberg D, *et al*. 2012. In and out of Madagascar: dispersal to peripheral islands, insular speciation and diversification of Indian Ocean daisy trees (*Psiadia*, Asteraceae). *PLoS One* **7**: e42932. doi:10.1371/journal.pone.0042932.

Thompson JD, Lumaret R. 1992. The evolutionary dynamics of polyploidy plants: origins, establishment and persistence. *Trends in Ecology and Evolution* **7**: 302–307.

Twyford AD, Ennos RA. 2012. Next-generation hybridization and introgression. *Heredity* **108**: 179–189.

Vähä J-P, Primmer CR. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Evolution* **15**: 63–72.

Vences M, Wollenberg KC, Vieites DR, Lees DC. 2009. Madagascar as a model region of species diversification. *Trends in Ecology and Evolution* **24**: 456–465.

Vieilledent G, Cornu C, Cuni Sanchez A, Leong Pock Tsy J-M, Danthu P. 2013. Vulnerability of baobab species to climate change and effectiveness of the protected area network in Madagascar: towards new conservation priorities. *Biological Conservation* **166**: 11–22.

Waldbieser GC. 1995. PCR-based identification of AT-rich tri and tetranucleotides repeat loci in an enriched plasmid library. *Biotechniques* **19**: 742–744.

Weir BS, Cockerham C. 1984. Estimating F-Statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Wickens GE, Lowe P. 2008. *The baobabs: pachycauls of Africa, Madagascar and Australia*. Berlin: Springer.

Wilmé L, Goodman SM, Ganzhorn JU. 2006. Biogeographic evolution of Madagascar's microendemic biota. *Science* **312**: 1063–1065.