Annals of Botany **111**: 47–60, 2013
doi:10.1093/aob/mcs235, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

Divergent pattern of nuclear genetic diversity across the range of the Afromontane *Prunus africana* mirrors variable climate of African highlands

Caroline A. C. Kadu[1,2,†], Heino Konrad[1,†], Silvio Schueler[1], Geoffrey M. Muluvi[2], Oscar Eyog-Matig[3], Alice Muchugi[2], Vivienne L. Williams[4], Lolona Ramamonjisoa[5], Consolatha Kapinga[6], Bernard Foahom[7], Cuthbert Katsvanga[8], David Hafashimana[9], Crisantos Obama[10] and Thomas Geburek[1,*]

[1]*Federal Research and Training Centre for Forests, Natural Hazards and Landscape (BFW), Department of Forest Genetics, Hauptstraße 7, A-1140 Vienna, Austria,* [2]*Kenyatta University, PO Box 43844, Nairobi, Kenya,* [3]*Bioversity International SSA, c/o CIFOR Regional Office, PO Box 2008, Messa, Yaounde, Cameroon,* [4]*School of Animal, Plant & Environmental Sciences, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa,* [5]*Silo National des Graines Forestieres (SNGF), PO Box 5091, Antananarivo-101, Madagascar,* [6]*Tanzania Forestry Research Institute (TAFORI), PO Box 1854, Morogoro, Tanzania,* [7]*Institute of Agricultural Research for Development (IRAD), PO Box 2123 or 2067, Yaounde, Cameroon,* [8]*Faculty of Agriculture and Environmental Science (Forestry Unit), Bindura University of Science Education, P. Bag 1020, Bindura, Zimbabwe,* [9]*National Forestry Resources Research Institute (NaFORRI), PO Box 1752, Kampala, Uganda and* [10]*Coordinador Nacional de la COMIFAC Ministerio de Agricultura y Bosques BP 207, Bata, Equatorial Guinea*
[†]*These authors contributed equally to this work.*
* *For correspondence. E-mail thomas.geburek@bfw.gv.at*

Received: 25 April 2012 Returned for revision: 23 July 2012 Accepted: 1 October 2012

• *Background and Aims* Afromontane forest ecosystems share a high similarity of plant and animal biodiversity, although they occur mainly on isolated mountain massifs throughout the continent. This resemblance has long provoked questions on former wider distribution of Afromontane forests. In this study *Prunus africana* (one of the character trees of Afromontane forests) is used as a model for understanding the biogeography of this vegetation zone.
• *Methods* Thirty natural populations from nine African countries covering a large part of Afromontane regions were analysed using six nuclear microsatellites. Standard population genetic analysis as well as Bayesian and maximum likelihood models were used to infer genetic diversity, population differentiation, barriers to gene flow, and recent and all migration among populations.
• *Key Results Prunus africana* exhibits strong divergence among five main Afromontane regions: West Africa, East Africa west of the Eastern Rift Valley (ERV), East Africa east of the ERV, southern Africa and Madagascar. The strongest divergence was evident between Madagascar and continental Africa. Populations from West Africa showed high similarity with East African populations west of the ERV, whereas populations east of the ERV are closely related to populations of southern Africa, respectively.
• *Conclusions* The observed patterns indicate divergent population history across the continent most likely associated to Pleistocene changes in climatic conditions. The high genetic similarity between populations of West Africa with population of East Africa west of the ERV is in agreement with faunistic and floristic patterns and provides further evidence for a historical migration route. Contrasting estimates of recent and historical gene flow indicate a shift of the main barrier to gene flow from the Lake Victoria basin to the ERV, highlighting the dynamic environmental and evolutionary history of the region.

Key words: Afromontane vegetation, gene flow, microsatellites, population differentiation, population history, *Prunus africana*, East African Rift system.

INTRODUCTION

African geography is characterized by several highlands which are distributed non-continuously in East and West Africa in an archipelago-like way, with the Congo basin posing a serious ecological barrier to distributions of associated species. Further areas of higher elevation exist in southern Africa and Madagascar. In many cases, the occurrence of plant and animal species is similar between East and West African highlands, although the two areas are at present separated by a 2000-km gap in suitable habitat (White, 1983). This pattern is most likely associated with climatic conditions during the Pleistocene (2·5 million to 11 000 years ago), oscillating between cold-arid and warm-moist (Moreau, 1966; Livingstone, 1975; White, 1983; Hamilton and Taylor, 1991; Maley, 1996). Accumulated evidence suggests that during the last glacial maximum (around 20 000 years ago) a drier and colder climate caused a severe recession of tropical rainforest in West Africa, allowing montane species to colonize areas at lower altitudes, at least along rivers as gallery forests, and thus may have favoured migration between eastern and western highlands (White, 1981, 1993; Maley, 1987; Elenga *et al.*, 2000). African biogeography is further complicated by the occurrence of strong ecological barriers, and the East

African Rift system is of outstanding importance in this respect. It is the result of a tectonic uplift of parts of East and Central Africa accompanied by rifting, faulting and volcanism starting in the early Miocene (approx. 25 myr ago). It is basically structured into the Eastern Rift Valley (ERV) and the western Albertine Rift Valley (ARV) with associated mountain areas that are separated by extensive lowlands (savannah). The importance of the East African Rift system as a barrier for distribution and gene flow has been demonstrated by numerous studies of plants, including *Arabis alpina* (Assefa *et al.*, 2007), *Lobelia giberroa* (Kebede *et al.*, 2007) and *Hagenia abyssinica* (Ayele *et al.*, 2009), and other groups of organisms (e.g. Ruiz Guajardo *et al.*, 2010; and references therein).

Environmental changes and geological barriers affect the occurrence of species and ecosystems and evolutionary processes and genetic diversity in the species involved (Hewitt, 2004). For the northern hemisphere, many studies on various kinds of plant species have demonstrated the effect of ice ages and the location of mountain massifs on the genetic diversity and demography of populations (e.g. Petit *et al.*, 2003; Soltis *et al.*, 2006). Although Africa is a main centre of global biodiversity (species and genetic diversity; Linder, 2001), we are just beginning to understand how the climatic and geological history has created this wide range of evolutionary patterns. For instance, regional patterns of genetic diversity have been recently studied to evaluate the impact of climatic discontinuities on the population structure of rainforest species in western Africa (Dainou *et al.*, 2010; Duminil *et al.*, 2010; Born *et al.*, 2011; Koffi *et al.*, 2011) and savannah tree species (Allal *et al.*, 2011). The African highlands appear understudied, but are particularly interesting because their topology has favoured dynamic changes in ecological conditions and associated species distributions (Kebede *et al.*, 2007).

Only a few studies have investigated the genetic diversity of plant species with a wide range of occurrence throughout Africa [e.g. *Adansonia digitata*, Pock Tsy *et al.*, 2009; *Senegalia mellifera* (synonym: *Acacia mellifera*), Ruiz-Guajardo *et al.*, 2010; *Acacia senegal*, Omondi *et al.*, 2010; *Prunus africana*, Kadu *et al.*, 2011]. Although many species of the Afromontane belt are endemics or have a very narrow distribution, some occur on most of the Afromontane locations and have very disjunct distributions (Hedberg, 1969). *Prunus africana* is an important component of Afromontane forests occurring throughout mainland Africa, surrounding islands and Madagascar. It is acknowledged for its medicinal properties for treating benign prostatic hyperplasia and has come under severe pressure due to overexploitation of wild populations (Hall *et al.*, 2000); currently it is listed as 'vulnerable' in the Red List of IUCN (http://www.iucnredlist.org). *Prunus africana* has been described as a 'nomad' species (White, 1983) which is distributed mainly in Afromontane forests but can also colonize specific habitats in lower regions, e.g. riverine forests (Hall *et al.*, 2000). It can also occur as a pioneer species in secondary forests. The seeds are relatively small and birds can disperse them over distances of at least 300 km (White, 1981). This ability for long-distance dispersal is evidenced by the occurrence of this species on volcanic islands in the Gulf of Guinea and the Comoros Islands.

In a previous study, we investigated the plastid phylogeography of the species across Africa (Kadu *et al.*, 2011). Important insights were obtained, and evidence suggested that populations east of the ARV are more closely related to those from western Africa compared with geographically proximate populations in East Africa, strengthening the hypothesis of a Pleistocene migration corridor between eastern and western montane forests. However, plastid markers provide estimates of population processes from only a single gene genealogy (Nordborg, 2001) and are not suitable for gaining deeper insights into gene flow and population history. For a more detailed understanding of biogeographical processes in Africa, knowledge of the more recent population history and current patterns of gene flow among populations is needed. Here, we used six nuclear microsatellite loci to provide further insight on the magnitude and partitioning of genetic variation in *P. africana* over the main part of its range. Extensive sampling of *P. africana* allowed addition of evidence to the existing nuclear data provided by Dawson and Powell (1999) and Muchugi *et al.* (2006). We test the hypotheses that (*a*) the ARV is the main barrier to gene flow on the African continent (Kadu *et al.*, 2011), (*b*) there are differences in population differentiation and divergent rates of historical gene flow among regions and (*c*) gene flow has become disrupted recently among regions. We use a genetic–geographical approach to detect significant barriers to gene flow and apply individual-based Bayesian assignment methods to identify population groups. Estimates of the extent and direction of current and historic gene flow among regions are derived.

MATERIALS AND METHODS

Population sampling

Leaf samples were obtained from 30 populations of *Prunus africana* covering the major block of the geographical range of the species. This included the West African region [Cameroon (three sites), Equatorial Guinea (Bioko) and Nigeria (one site each)]; the East African region [Uganda (four sites), Kenya (eight sites) and Tanzania (six sites)]; the southern region [South Africa (two sites) and Zimbabwe (three sites)]; and Madagascar (two sites) (Fig. 1 and Table 1). We were not able to obtain samples from Ethiopia, Eritrea or the Comoros Islands, which would have made sampling from the main range complete. Population sampling was carried out in 2007–2008 in natural unfragmented mixed stands, or degraded sites in cases where natural forests were absent. From each population, leaf samples from three to 31 trees (459 trees in total; Table 1) were collected and dried in silica gel for storage. In cases of recently fragmented forests with low population densities (mainly populations from South Africa), trees from a wider region spanning up to 100 km were sampled and considered as a single population. Import, export and phytosanitary certificates were obtained in line with CITES and national regulations.

DNA extraction and microsatellite genotyping

Total genomic DNA was extracted from 40–60 mg dry leaf material using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). The supplier's protocol was slightly modified by using 800 μL buffer AP1 and adjusting RNase and buffer

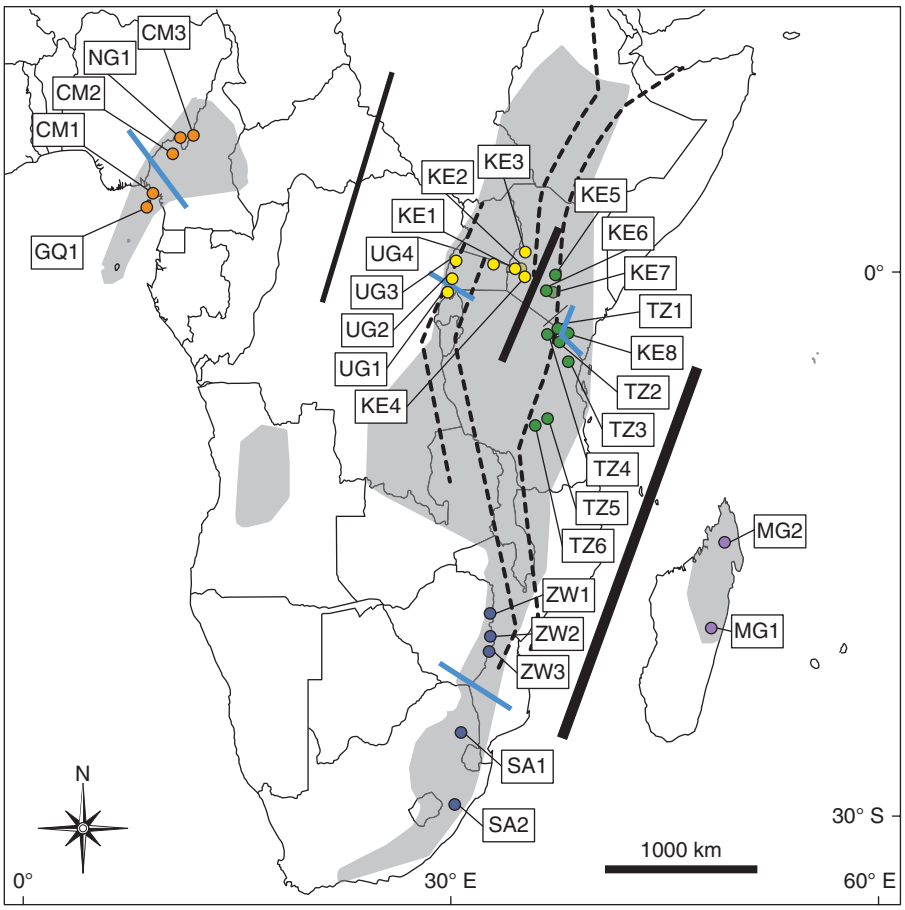

FIG. 1. Sampling locations, group membership and barriers to gene flow for 30 natural populations of *Prunus africana* (the natural distribution range is indicated by the grey shaded area). Broken lines indicate the Rift Valley System. Colours indicate membership to genetic population groups (see Table 1). The main gene-flow barriers among regions as recovered in BARRIER are indicated by black bars; barriers within regions are indicated by blue bars. The confidence of the barrier is indicated in line thickness with strong lines indicating the best-supported barriers.

AP2 volumes accordingly (Farwig *et al.*, 2008). We determined the concentration and purity of the DNA samples using a ND-1000 spectrophotometer (NanoDrop, USA). DNA was stored at 4 °C. Nuclear variation was analysed with six microsatellite loci originally characterized in wild cherry (*P. avium*, primers EMPaS01, EMPaS06 and EMPaS10; Vaughan and Russell, 2004) and peach [*P. persica*, primer pairs U3 (UDP97-403) and U5 (UDP96-018); Cipriani *et al.*, 1999; and P2 (PS12A02); Sosinski *et al.*, 2000]. These were previously optimized for *P. africana* (Cavers *et al.*, 2009). Fragment size was determined by capillary gel electrophoresis using a CEQ8000 sequencer (Beckman-Coulter, USA).

Data analyses

Five sampling locations with <10 individuals (Table 1) were excluded from population diversity analyses, but were retained for the Bayesian assignment and historical gene-flow analysis described below. The microsatellite dataset was checked for the presence of genotyping errors and null alleles using the program MICRO-CHECKER (van Oosterhout *et al.*, 2004). For each locus the observed allele frequencies and the observed and expected levels of heterozygosity (H_o and H_e, respectively) per population were calculated using GENALEX version 6·4 (Peakall and Smouse, 2006). Deviations from Hardy–Weinberg equilibrium (HWE) were assessed using GENEPOP (Raymond and Rousset, 1995) as significant heterozygote excess or deficiency using Markov chain Monte Carlo simulations with 10 000 dememorizations, 100 batches and 10 000 iterations. Genotypic disequilibrium between loci was assessed in FSTAT 2·9·3·2 (Goudet, 1995) using a Bonferroni correction. Allelic richness with rarefaction to 10 individuals was also calculated using FSTAT. The presence of recent changes in population size, as evidenced by heterozygosity excess based on estimates of multi-locus genotypes, was tested using the software BOTTLENECK v1·2·02 (Piry *et al.*, 1999), implementing a Wilcoxon signed rank test and assuming a two-phased model of mutation (95 % stepwise and 5 % multi-step mutations, 15 % variance among multiple steps).

Global F_{ST} and R_{ST} were calculated using ARLEQUIN 3·5 (Excoffier and Lischer, 2010). We also produced an unweighted pair group method arithmetic average (UPGMA) dendrogram based on Cavalli-Sforza and Edwards (1967) chord distances after creating 1000 bootstrapped matrices in Microsatellite Analyser (MSA) (Dieringer and Schlötterer,

TABLE 1. *List of the 30 natural populations of* Prunus africana *analysed in this study*

Code	Population name	Country	Population group	Longitude (°)	Latitude (°)	Elevation (m)	*n*
GQ1	Moka	Equ. Guinea	WG	8·6559	3·3625	1700	13
CM1	Lower Mann's Spring, Mt Cameroon	Cameroon	WG	9·1192	4·1346	1600	20
CM2	Ngashie-Mt Oku	Cameroon	WG	10·5092	6·2048	2450	16
CM3	Mt Danoua	Cameroon	WG	11·9413	7·1871	1400	19
NG1	Ngel Nyaki Forest Reserve	Nigeria	WG	11·0271	7·066	1100	7
UG1	Bwindi Forest	Uganda	EG1	29·7754	−1·0476	1900	20
UG2	Kalinzu Forest Reserve	Uganda	EG1	30·1101	−0·3742	1500	20
UG3	Kibale Forest Natural Park	Uganda	EG1	30·3570	0·5644	1350	20
UG4	Mabira Forest	Uganda	EG1	33·0150	0·3806	1250	20
KE1	Kibiri Forest, Western Province	Kenya	EG1	34·8719	0·1502	1400	12
KE2	Kakamega Forest, Western Province	Kenya	EG1	34·5190	0·1415	1400	17
KE3	Kapcherop, Cherangani Forest, Rift Valley	Kenya	EG1	35·2161	1·0315	1900	18
KE4	Londiani, Rift Valley	Kenya	EG1	35·1836	−0·2737	1750	20
KE5	Chuka, Central province	Kenya	EG2	37·3645	−0·1798	3550	19
KE6	Lari, Central Province	Kenya	EG2	36·687	−1·0311	2250	5
KE7	Ol Danyo Sambuk, Central Province	Kenya	EG2	37·1501	−1·0744	1450	15
KE8	Taita Hills, Eastern Arc	Kenya	EG2	38·2088	−3·2462	900	18
TZ1	Kilimanjaro Catchment Forest Reserve	Tanzania	EG2	37·5237	−3·0036	2000	11
TZ2	Kindoroko Catchment Reserve, Eastern Arc	Tanzania	EG2	37·6267	−3·7389	1750	6
TZ3	Shume Magamba Forest Reserve, Eastern Arc	Tanzania	EG2	38·2521	−4·7544	1700	3
TZ4	Meru Catchment Forest	Tanzania	EG2	36·8074	−3·2925	2000	16
TZ5	Udzungwa, Eastern Arc	Tanzania	EG2	36·7791	−7·7668	1150	10
TZ6	Kidabaga, Eastern Arc	Tanzania	EG2	35·9296	−8·1101	2000	15
ZW1	Nyanga National Park	Zimbabwe	SG	32·7401	−18·2873	1800	19
ZW2	Cashel Valley Chimanimani	Zimbabwe	SG	32·8004	−19·5781	1350	13
ZW3	Chirinda forest Reserve Chipinge	Zimbabwe	SG	32·6950	−20·4090	1200	20
SA1	Mpumalanga	South Africa	SG	30·7306	−25·0225	1300	18
SA2	KwaZulu-Natal	South Africa	SG	30·2725	−29·2872	1500	12
MG1	Marovoay	Madagascar	MG	48·2825	−19·1183	1100	31
MG2	Antsahabiraoka	Madagascar	MG	49·2164	−14·4015	1350	6

n, Number of samples.

2003). The computer programs NEIGHBOUR and CONSENSE in the PHYLIP v3·63 package (Felsenstein, 1989) were used for tree construction. Additionally, principal co-ordinate analysis was used to explore multivariate relationships among inter-individual genetic distances (Nei's standard genetic distance; Nei, 1972) and to identify a set of reduced dimension traits (e.g. PC eigenvectors). These analyses were conducted using GENALEX. Consequently analysis of molecular variance (AMOVA) was used to investigate population differentiation at various levels of subdivision defined by the groups detected by principal component analysis (PCA), UPGMA clustering and individual-based population assignment (see below), using both F_{ST} and R_{ST} in separate analyses (10 000 permutations, significance level 0·05). These computations were performed in ARLEQUIN 3·5. To examine genetic barriers between populations at continental and regional scales, we used the software BARRIER 2·2 (Manni *et al.*, 2004). In this approach computational geometry and a Monmonier's maximum-difference algorithm are combined to identify edges associated with high rates of change in a genetic distance matrix using Delaunay triangulation Voronoi polygons. One thousand bootstrapped pairwise relative genetic distance matrices (chord-distance) created in MSA were used for the computations.

Individual-based population assignment methods were used to further analyse population structure. In these computations all available samples were included. The model-based Bayesian clustering method implemented in STRUCTURE v2·3 (Pritchard *et al.*, 2000) was used to infer clusters by assigning individual multi-locus genotypes probabilistically to a user-defined number of *K* clusters. In STRUCTURE assignment of individuals an inference of allele frequencies is performed so that departures from Hardy–Weinberg and gametic-phase disequilibrium within the clusters are minimized. Assuming correlated allele frequencies, we used the admixture model without incorporation of population information. Using *K* values ranging from 2 to 30, run lengths of 500 000 iterations with a burn-in period of 200 000 were chosen based on preliminary runs to determine run parameters. Ten runs per *K* were performed on the total dataset. The average negative logarithmic probability of the data from multiple replicates of each *K* was plotted to evaluate the rate of probability changes with increasing values of *K*. The most likely number of clusters was inferred using the ΔK statistic of Evanno *et al.* (2005).

To quantify average recent migration rates (last few generations) among populations, a Bayesian approach implemented in the software BAYESASS 1·3 (Wilson and Rannala, 2003) was applied. This method makes no assumptions of HWE within the sample and is based on transient multi-locus disequilibrium in multi-locus genotypes of migrants relative to the host population. It calculates inbreeding coefficients for each population separately and the joint probabilities are used to estimate recent migration rates (Wilson and Rannala, 2003). Since the accuracy of estimating population assignment or migration rates with this approach depends on large population sample sizes and/or large numbers of loci, we again chose to omit populations with <10 individuals. A run length of

5000 000 generations, with 999 999 generations burn-in, a sampling frequency of 2000, with δp (maximum change of allele frequency between iterations) at 0·30, a δm of 0·15 (maximum change of migration rate) and a δF of 0·15 (maximum change of inbreeding coefficient). The dataset was run five times with different starting seeds to ensure consistency of results and an average of these runs is reported.

To estimate the direction and magnitude of historical gene flow between population groups determined with STRUCTURE, the maximum likelihood (ML) approach of Beerli and Felsenstein (2001) implemented in MIGRATE version 3·2·7 (Beerli, 2010) was applied. The ML model assumes that populations are at drift–migration equilibrium, had constant size and migration occurred over the coalescent period (approx. 4Ne generations). However, unlike F-statistics, this method allows non-symmetrical migration and differences in population sizes among populations, both biologically realistic scenarios. MIGRATE jointly estimates the mutation-scaled effective population sizes ($\theta = 4N_e\mu$, where μ is the mutation rate for diploid data) and the mutation-scaled effective immigration rate ($M = m/\mu$, where m is the immigration rate) between geographical groups by estimating allele genealogies and then approximates the sum of probabilities across possible genealogies using Metropolis–Hastings Markov chain Monte Carlo sampling (Beerli and Felsenstein, 2001). The Brownian motion model for microsatellites as an approximation of the ladder model (Ohta and Kimura, 1973) was used as mutation model. As a search strategy ten short chains with 5000 recorded trees, followed by three long chains with 50 000 recorded trees with a burn-in of 20 000 and a sampling increment of 100 was used. The first run estimated θ and M from F_{ST} values, whereas subsequent runs incorporated the ML estimates of θ and M from the previous run as starting parameters. MIGRATE was run five times until estimated parameter values converged. The results of the fifth run were used for interpretation. To avoid problems due to different sample size among groups a reduced random subset (37 individuals corresponding to the lowest group size) of each population was used in a parallel computation.

RESULTS

Genetic diversity

The six microsatellite markers exhibited a high level of polymorphism in *P. africana* with 10–45 alleles per locus (142 alleles in total). Nevertheless a high level of divergence among populations was evident as some populations were fixed for single alleles: populations SA1, SA2 and MG1 did not show variation at locus U3, likewise SA1 and CM2 were fixed at locus U5 and population ZW3 was fixed at locus EMPaS01. The number of alleles, number of effective alleles, observed and expected heterozygosity and fixation index are shown in Table 2. Observed heterozygosity ranged from 0·352 in SA1 to 0·808 in UG3 with an average of 0·622. Lower values were mostly found in the southern

TABLE 2. *Genetic diversity indices, private alleles and bottleneck probability for 25 natural* Prunus africana *populations based on six nuclear microsatellite markers*

Pop. code	N_a	N_e	H_o	H_e	F	Ap	R_s	Wilcoxon signed rank test (TPM)[†]
GQ1	5·833	4·027	0·744	0·735	−0·074	2	5·388	0·2188
CM1	7·500	4·866	0·683	0·719	0·158	0	6·090	0·0156
CM2	6·167	3·730	0·531	0·647	0·174*	0	5·287	0·0391
CM3	5·667	3·066	0·658	0·645	−0·051	2	4·681	0·0156
UG1	8·667	5·334	0·733	0·771	0·015*	1	6·789	0·0156
UG2	7·833	4·530	0·692	0·726	0·042	0	6·094	0·0234
UG3	8·167	4·807	0·808	0·722	−0·146**	0	6·588	0·2188
UG4	7·833	5·521	0·658	0·775	0·158***	0	6·606	0·0078[‡]
KE1	9·167	6·936	0·736	0·827	0·104*	1	8·436	0·0156
KE2	8·500	5·533	0·725	0·761	0·036	1	7·081	0·0234
KE3	8·167	5·160	0·620	0·737	0·119***	1	6·696	0·0156
KE4	8·000	5·441	0·692	0·763	0·073**	1	6·596	0·2813
KE5	6·667	4·199	0·781	0·716	−0·130	0	5·662	0·0078[‡]
KE7	4·667	3·200	0·711	0·616	−0·188**	1	4·247	0·0078[‡]
KE8	4·667	2·794	0·491	0·547	0·098*	1	4·079	0·0156
TZ1	6·000	3·770	0·530	0·657	0·133*	0	5·797	0·0156
TZ4	5·500	3·423	0·604	0·629	0·019	1	4·928	0·0156
TZ5	7·333	5·339	0·517	0·753	0·364***	1	7·333	0·0156
TZ6	6·167	3·880	0·578	0·613	−0·011	0	5·499	0·0156
ZW1	5·167	3·284	0·632	0·612	−0·072	1	4·495	0·0156
ZW2	6·000	3·355	0·590	0·574	−0·067	3	5·389	0·0781
ZW3	4·000	2·688	0·575	0·540	−0·088	0	3·548	0·0078[‡]
SA1	4·500	2·937	0·352	0·430	0·182**	0	3·836	0·0078[‡]
SA2	3·500	2·597	0·417	0·467	0·048*	0	3·398	0·0078[‡]
MG1	6·500	3·577	0·500	0·566	0·212	9	4·734	0·0078[‡]

N_a, Number of observed alleles; N_e, number of effective alleles; H_o, observed heterozygosity; H_e, unbiased expected heterozygosity; F, fixation index; *, **, *** deviation from HWE significant at the 0·05, 0·01 and 0·001 level, respectively; Ap, number of private alleles; Rs, average allelic richness.

[†] Test for population bottlenecks based on the Wilcoxon signed rank test under the two-phased model of mutation (95 % stepwise, 5 % multi-step mutations); [‡] significant at the $P < 0·01$ level.

populations (Zimbabwe and South Africa) and in Madagascar. The lower levels of genetic diversity are also reflected in measures of average allelic richness, which ranged from 3·398 in SA2 to 8·436 in KE1. The highest average values of allelic richness were found in East African populations west of the ERV (Table 1 and Supplementary Data Table S1 available online). In many populations, private alleles were detected, highlighting the strong regional divergence among the populations of *P. africana* in Africa (frequency of private alleles per population is given in Table 2). When counting private alleles at the regional level, Madagascar and East Africa (east of the ERV) harboured the highest number (nine each). Also, East African populations west of the ERV had many private alleles (eight), whereas other regions harboured fewer (southern African populations, five; western African populations, four). The frequencies of these private alleles were generally $<0{\cdot}05$; however, in Madagascar three alleles at locus U5 had frequencies of 0·14, 0·11 and 0·23, respectively; similarly one allele in locus P2 occurred at a frequency of 0·31 in southern African populations.

Although several populations showed heterozygote deficiency or excess, only few highly significant deviations from HWE ($P < 0{\cdot}001$) were detected (populations UG4, KE3 and TZ5; Table 2). Deviations were probably caused by the occurrence of null alleles, which affected all loci except locus P2 to varying degrees (Supplementary Data Table S2 available online) but probably did not affect further analysis. No evidence for linkage disequilibrium between loci within populations was found after sequential Bonferroni correction. Assuming a two-phased mutation model, the Wilcoxon signed rank test implemented in BOTTLENECK ($P < 0{\cdot}01$) detected a recent decrease in population size in southern Africa (SA1, SA2, ZW3), Kenya (KE5, KE7), Uganda (UG4) and Madagascar (MG1) (Table 2).

Genetic differentiation and barriers to gene flow

Across all populations a high global F_{ST} value of 0·27 was obtained. Population differentiation, taking into account allele size and a stepwise mutation model, was significantly higher with $R_{ST} = 0{\cdot}70$, indicating a strong phylogeographic pattern in the sample. Relationships among populations based on a bootstrapped Cavalli-Sforza and Edwards (1967) chord distance matrix are shown in the UPGMA tree (Fig. 2): the population from Madagascar appeared central in the unrooted tree and southern African and populations east of the ERV appeared close to each other, as did populations from West Africa and west of the ERV (with population CM3 on the same branch with these east African populations). Despite geographical proximity, populations east and west of the ERV were the most diverged on the tree. A similar result was obtained in the PCA (Fig. 3); in this analysis population CM2 from Cameroon clustered with populations from East Africa west of the ERV, and these two groups were not as clearly differentiated as the other groups.

Main barriers to gene flow as detected with BARRIER 2·2 are indicated in Fig. 1. The first barrier (91·8 % support) was

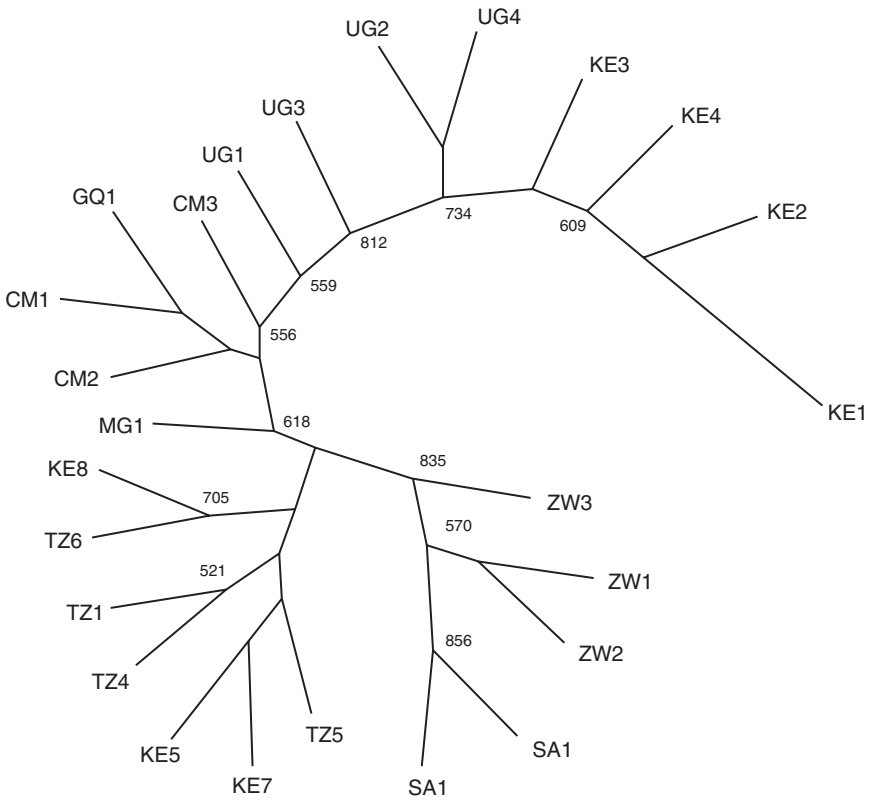

FIG. 2. Majority-rule consensus tree showing relationships between 25 *Prunus africana* populations based on 1000 UPGMA trees from Cavalli-Sforza and Edwards (1967) chord distances matrices. Numbers indicate bootstrap support for nodes.

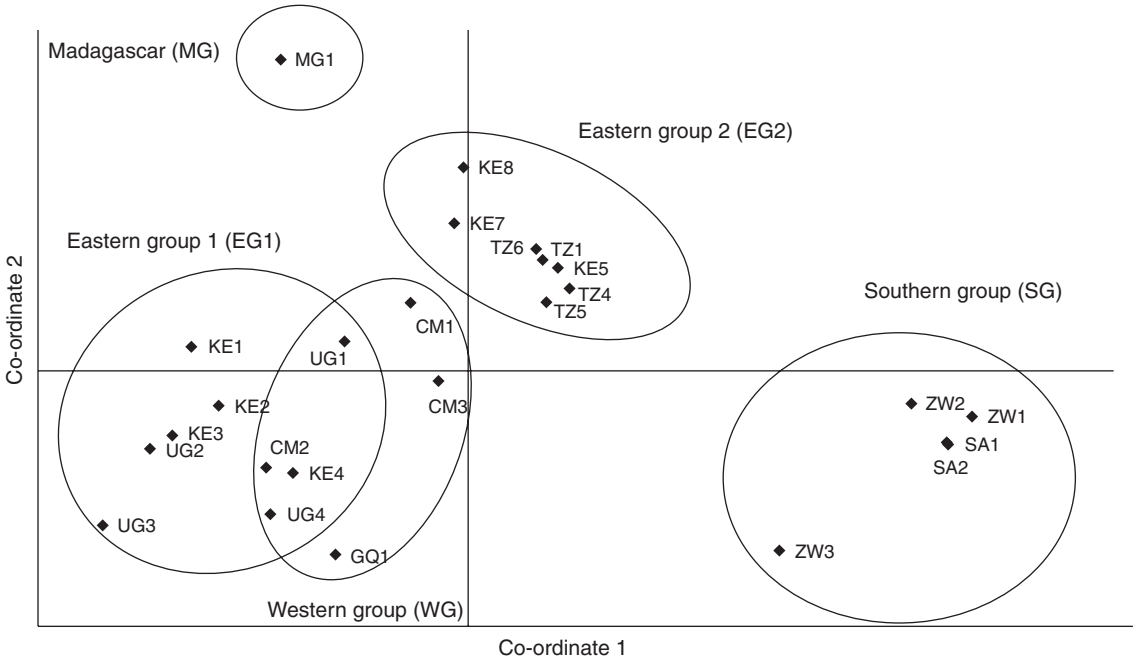

FIG. 3. Principal co-ordinate analysis showing the multivariate relationships of 25 *Prunus africana* populations. The first two axes explain a total of 67 % variation. The first and second axes explain 50 % and 17 % of the variance, respectively. Circles indicate regional population groups.

found between Madagascar and mainland Africa. The second barrier (54·8 % support) was found between populations east and west of the ERV. The third barrier between populations in West and East Africa was only weakly supported (34 %), similar to all further barriers. Because of the strong regional structuring of the dataset, a separate analysis of gene-flow barriers was conducted within the four groups with more than two populations (Fig. 1). In West Africa the most important gene-flow barrier among populations (61 %) was detected between CM1 and CM2; in populations west of the ERV the main barrier (55 %) separated population UG1 from neighbouring populations; in populations east of the ERV the sharpest break in allele frequencies was detected between population KE8 and all other populations (76 %); in the southern African group the strongest barrier (72 %) was found between populations of South Africa and Zimbabwe.

Individual-based population assignment

Individual-based assignment using admixture analysis in STRUCTURE revealed generally high assignment coefficients for the specific clusters (Fig. 4), illustrating strong population differentiation. Searching for the most likely group structure (Evanno *et al.*, 2005) identified grouping into five clusters, because Δln P(D) reached a maximum at five groups. These five main clusters corresponded to the main areas of occurrence: West Africa (WG), East Africa west of the ERV (EG1), East Africa east of the ERV (EG2), southern Africa (SG) and Madagascar (MG). When K was set to 6 and 7, further clusters differentiating populations from West Africa and from East Africa were detected (Fig. 4), but these groupings showed less-clear ancestry coefficients. At $K = 7$ population, UG1 showed heavy admixture from the West African cluster. Above a K value of 10 no biologically meaningful splitting in the dataset occurred. These five groups were also supported by the results of the distance tree and PCA (Figs 2 and 3) and thus were taken as the basis for further analyses. AMOVA analysis using F_{ST} as the differentiation measure partitioned the molecular variance to 14·4 % among groups, 12·6 % within populations among groups and 73·0 % within populations. However, when taking allele size and stepwise mutation into account using R_{ST}, 62·5 % of variance was found among groups, 7·5 % within populations among groups and only 30·0 % within populations (Table 3), suggesting a significant effect of mutation on regional differentiation.

Recent gene flow

Estimates of recent migration between populations obtained with BAYESASS were relatively low, but differed (significantly) among the five regions (Table 4). In general, recent migration occurred mainly within population groups. Only a few estimates >1 % suggested inter-group migration, and this only applied to migration between WG and EG1 and EG2 and SG, respectively. No evidence for recent migration between spatially close populations east and west of the ERV was found. Populations from West Africa, southern Africa and Madagascar (one population) showed evidence of little current gene flow among them, with average self-recruitment of 0·90, 0·83 and 0·99, respectively, but values were higher in populations from East Africa. Within this region, populations belonging to the EG1 on average showed the highest level of recent migration among populations

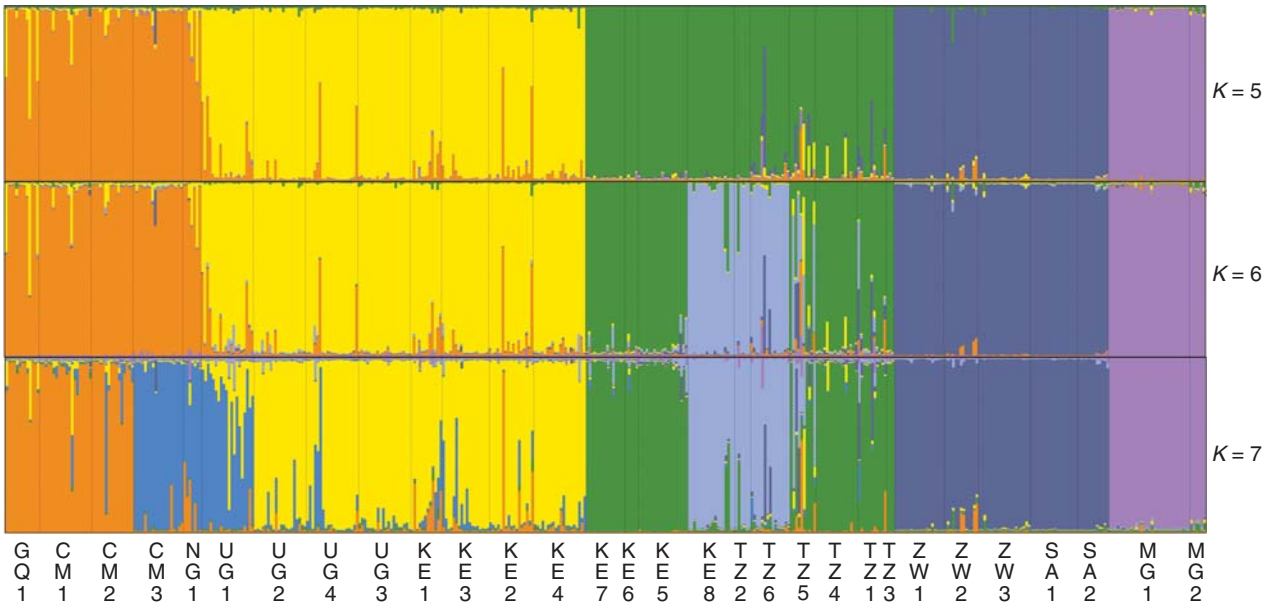

FIG. 4. Results of individual population assignment (admixture mode) performed with STRUCTURE for 30 populations of *Prunus africana* for five, six and seven assumed ancestral clusters (*K*). The run with lowest ln P(D) out of a total of ten runs is depicted.

TABLE 3. *Results of analysis of molecular variance (AMOVA) based on* F_{ST} *and* R_{ST}*, respectively, from data of six nuclear microsatellite loci of 25* Prunus africana *populations from seven countries*

	F_{ST}			R_{ST}		
Source of variation	Sum of squares	Variance components	Percentage variation	Sum of squares	Variance components	Percentage variation
Among groups	320·96	0·39	14·44	50 8857·65	750·14	62·49
Among populations within groups	266·36	0·34	12·57	66 982·60	89·95	7·49
Within populations	1662·02	1·98	72·99	30 2357·47	360·38	30·02
Total	2249·33	2·71		87 8197·73	1200·47	

Groups correspond to the five biogeographic regions detected.

(self-recruitment average: 0·79), whereas populations bordering the eastern side of the Rift (KE5 and KE7) were mostly self-recruiting (0·88). Higher migration estimates (self-recruitment 0·78) were found in the populations from Tanzania (populations around Mount Kilimanjaro), which also appeared to have a minor migration rate to populations of the southern range (Table 4). The population from Madagascar had the lowest migration rate of all populations; sharing of migrants within WG was inferred only between CM1 and the island population GQ1.

Historical gene flow

Historical rates of gene flow among groups derived from the coalescent computations implemented in MIGRATE were similarly low and divergent among regions (Table 5). Estimates of migration rates (given as scaled migration rate *M*) were similar between computations involving the whole dataset and a random subset (not shown). Estimates of *M* between Madagascar and the mainland populations were low. In the SG population, historical gene flow was only indicated to and from EG2. There was, opposed to the results obtained for recent migration rates, evidence for moderate bidirectional historical gene flow between populations east (EG2) and west of the ERV (EG1). The highest historical gene-flow estimates were derived for populations west of the ERV (EG1) and West Africa (WG). The derived direction of gene flow was primarily from EG1 to WG. Estimates of scaled population size (θ) were similar in EG1 and EG2, were highest for MG, but were comparatively low for WG and SG. Estimates of θ were similar for the different computations, although in the random subset (smaller sample size) θ for EG1 was about 20 times larger than estimated with the full dataset. However, the θ estimates need to be interpreted cautiously, as subgroup structure and observed deviations from model assumptions can influence estimates (cf. Beerli and Felsenstein, 2001).

DISCUSSION

Genetic diversity and population differentiation

The within-population genetic diversity (H_e) in *P. africana* was comparable to other hardwood species, e.g. *Fraxinus excelsior* in Europe (0·736–0·814; Heuertz *et al.*, 2004) or tropical species, e.g. *Vitellaria paradoxa* (0·570–0·710; Logossa *et al.*, 2011). Genetic diversity differed among regions. Allelic richness was highest in populations from Kenya and Uganda west of the ERV, suggesting that *P. africana* has occurred in large populations for a long time in these habitats. Levels of genetic diversity were lowest in populations from southern Africa and Madagascar. Population bottlenecks were detected in populations from southern Africa, Kenya, Uganda and Madagascar; these were probably anthropogenic as severe deforestation and changes in land use starting approx. 1000–2000 years BP have had a dramatic impact on Afromontane ecosystems (Gade, 1996; Finch *et al.*, 2009).

Population differentiation with a global F_{ST} of 0·27 and R_{ST} of 0·70 was high, reflecting the large and disjunct distribution area of the species and the strong divergence among populations even within regions. This was highlighted by the findings of the AMOVA (based on F_{ST}), as 14·4 % of the variation was partitioned among regions and 12·6 % among populations within regions. Stepwise mutations had a strong effect on population differentiation as shown by $R_{ST} > > F_{ST}$, indicating long-term isolation of groups. Also the frequent occurrence of private alleles and null alleles within regional samples underlines the strong divergence of populations. Null alleles might have affected the consecutive analysis; however, we believe that the inference we make is based on strong evidence, and is corroborated by previous findings based on plastid DNA markers (Kadu *et al.*, 2011).

Five main groups of populations were resolved in the analysis corresponding to main biogeographic regions: Afromontane forests of western Africa (WG), eastern Africa west of the ERV (EG1), eastern Africa east of the ERV (EG2), southern Africa (SG) and Madagascar (MG). This grouping was supported by the UPGMA distance tree, principal co-ordinate analysis and the results of individual population assignment. A similar pattern had been obtained for *P. africana* from RAPD data by Dawson and Powell (1999), but the relationship between WG and EG1 derived in that study was even closer; we show that gene flow between WG and EG1 has been disrupted for a relatively short period, but this has allowed populations to diverge.

Recent patterns of migration

Estimates of recent migration rates between populations showed that gene flow between populations has been largely restricted in the western and Madagascan groups (Table 4). Given that Madagascar is an island, the result is not surprising there, whereas in West Africa gene flow was restricted to the island of Bioko (GQ1) and CM1, the nearest population from the Cameroonian mainland, and appeared disrupted among the other mainland populations. A different pattern was observed in East Africa, in particular in EG1 – here gene flow was estimated to contribute to genetic diversity in most populations to a higher degree. On the other hand, Kenyan populations east and close to the ERV appeared to be less connected (average self-recruitment 0·88). Actually the ERV was identified as the second most important barrier for gene flow throughout the range of *P. africana*, though some populations especially from the Eastern Arc appeared to be better connected. Eastern Arc populations also appear to share a small amount of migrants with southern African populations, and are thus most probably the original sources for these populations, which were also similar in self-recruitment. These findings are in general agreement with the estimates of population divergence obtained from plastid DNA (Kadu *et al.*, 2011) – western and eastern lineages diverged early, and the population migration from EG1 to West Africa appears to be a relatively recent event. East of the Rift, effects of fragmentation are more evident and appear to be older, as evidenced by the fixation of unrelated plastid haplotypes in spatially close populations (Kadu *et al.*, 2011). The better connectivity among Eastern Arc populations is probably associated with the more stable climate in that region (Hamilton and Taylor, 1991; Finch *et al.*, 2009).

The Rift Valley system as major gene-flow barrier: major shift of a gene-flow barrier in East Africa

The situation in East Africa is particularly remarkable. While only one group was supported in the other main geographical regions, East Africa holds two major groups; also the most evident genetic barrier on the mainland was detected in this region. This result is in line with previous findings in other species by Dawson and Powell (1999) and Muchugi *et al.* (2006). Our MIGRATE analysis (Beerli and Felsenstein, 2001) of historical gene-flow patterns over the coalescent period ($4N_e$ generations) corroborated earlier studies and provided important new insight into population history. In contrast to estimates of recent gene flow, historical estimates indicate low levels of gene flow between EG1 and EG2, providing evidence that, during a putatively colder earlier period of the Pleistocene, both groups were connected via gene flow. At the same time, all analyses (chord distance tree, assignment, migration) support the strong divergence between populations east and west of the ERV. Ancestral population size appears to have also been large in populations of East Africa (on both sides of the ERV), providing strong evidence that this region was central for population differentiation and spread.

The divergent climatic conditions between East and West Africa probably best explain the observed pattern: during the Pleistocene in West Africa the montane vegetation spread to lowland areas several times, in particular as riverine gallery forests, and thus allowed ample gene flow between populations (Hamilton and Taylor, 1991). Accumulated evidence also suggests that montane forests in the ARV had more constant climatic conditions (e.g. Jolly *et al.*, 1997; Marchant *et al.*, 1997), and Marchant and Taylor (1998) report that Afromontane forests were more or less stable in south-west Uganda from at least 42 000 years until approx. 1000 years BP (start of human influence). In contrast, east of the ERV the climate was in general dryer and more variable at least during the last 10 000 years, and montane vegetation became sparse several times as exemplified by pollen records (reviewed in White, 1983; Street-Perrot and Perrot,

TABLE 4. *Recent gene-flow estimates (as rate of migration from population* i *into population* j*, average from five computations) among 25 populations of* Prunus africana *(s.d. in brackets)*

	Population *j*													
	WG				EG1								EG2	
Population *i*	GQ1	CM1	CM2	CM3	UG1	UG2	UG3	UG4	KE1	KE2	KE3	KE4	KE5	KE7
GQ1	**0·8598 (0·1403)**	0·0892 (0·1093)		0·0141 (0·0283)										
CM1	0·0652 (0·0799)	**0·8627 (0·1443)**												
CM2			**0·9799 (0·0023)**											
CM3				**0·8933 (0·1198)**										
UG1					**0·8281 (0·0952)**									
UG2					0·0440 (0·0836)	**0·8626 (0·1474)**	0·0115 (0·0230)	0·0861 (0·1055)	0·0395 (0·0455)	0·0540 (0·0702)	0·0579 (0·0826)	0·0454 (0·0907)		
UG3							**0·8306 (0·1260)**							
UG4							0·0123 (0·0162)	**0·8543 (0·1408)**						
KE1					0·0173 (0·0283)	0·0327 (0·0654)	0·0310 (0·0621)		**0·7450 (0·1089)**	0·0398 (0·0795)	0·0425 (0·0850)	0·0452 (0·0903)		
KE2				0·0242 (0·0485)	0·0215 (0·0293)				0·0516 (0·0636)	**0·8008 (0·1427)**	0·0838 (0·1026)	0·0872 (0·1069)		
KE3					0·0160 (0·0240)	0·0137 (0·0273)	0·0494 (0·0724)				**0·6836 (0·0013)**			
KE4				0·0212 (0·0423)		0·0404 (0·0808)			0·0138 (0·0276)	0·0194 (0·0389)	0·0252 (0·0504)	**0·7423 (0·1188)**		
KE5													**0·8592 (0·1449)**	0·0478 (0·0726)
KE7													0·0897 (0·1098)	**0·9024 (0·1144)**

Rate of self-recruitment is shown in bold. Values below 0·01 are not depicted.

1993; Elenga *et al.*, 2000). Thus, east of the ERV only smaller and more isolated populations of *P. africana* may have survived the last glacial maximum, favouring higher population differentiation due to drift effects (UPGMA tree, Fig. 2; plastid DNA, Kadu *et al.*, 2011). In the Eastern Arc Mountains of Tanzania and Kenya (populations KE8, TZ2, TZ5 and TZ6) the situation is different again, as that region is buffered against extreme drought by orographic rainfall from the Indian Ocean (Hamilton, 1982; Lovett, 1993), which may also explain the better connectivity among these populations inferred from recent gene-flow estimates and the domination of few plastid haplotypes in the region (Kadu *et al.*, 2011).

In our previous study using plastid DNA (Kadu *et al.*, 2011), populations on the west side of the ERV shared haplotypes with eastern populations, not with western populations close to the ARV (and other western clade populations). In this study, we show that these populations (UG4, KE1, KE2, KE3, KE4) clearly cluster with populations west of the ERV with nuclear microsatellites. With the calculated low levels of historical gene flow and the disrupted recent gene flow between EG1 and EG2, our results suggest a shift in gene-flow barriers probably during the early to mid-Pleistocene: apparently the ERV has replaced the Lake Victoria basin as the main barrier to gene flow in eastern Africa for the species since at least the mid Holocene (cf. patterns of recent gene flow). Shifts in gene-flow barriers have only been rarely reported, and this finding highlights the dynamic environmental history of the continent and its impact on the evolutionary history of the species.

The importance of the East African Rift system as a gene-flow and dispersal barrier in the species deserves special attention – apparently relatively recent gene flow was possible across distances of >2000 km, but not across the 100 km of savannah separating the mountainous areas east and west of the ERV. The ecological reasons for this are intriguing, but are still mostly obscure. The ERV was also found to be a major gene-flow barrier in several other Afromontane plant species, e.g. *Arabis alpina* (Assefa *et al.*, 2007), *Lobelia gibberoa* (Kebede *et al.*, 2007), *Hagenia abyssinica* (Ayele *et al.*, 2009) and *Acacia senegal* (Omondi *et al.*, 2010) and in other organisms (e.g. Nicolas *et al.*, 2008, and references therein). Strong population divergence between populations associated with the ERV was also observed in a recent study on *Senegalia mellifera*, a savannah species (Ruiz Guajardo *et al.*, 2010). There, the striking breaks in allelic patterns were attributed to selection due to site (soil) conditions. The most likely explanation, however, is that adaptations to climate (precipitation) may be more important in the moisture-demanding *P. africana* (Mbatudde *et al.*, 2012). The observed

TABLE 4. *Continued*

	Population *j*										
	EG2					SG					MG
Population *i*	KE8	TZ1	TZ4	TZ5	TZ6	ZW1	ZW2	ZW3	SA1	SA2	MG1
KE5	0·0429 (0·0859)	0·0267 (0·0343)	0·0490 (0·0773)	0·0484 (0·0422)	0·0171 (0·0342)						
KE7		0·0260 0·0343		0·0357 0·0445							
KE8	**0·8629** **(0·1466)**				0·0037 0·0073						
TZ1		**0·6931** **(0·0004)**									
TZ4	0·0434 (0·0868)	0·0734 (0·0458)	**(0·8988)** **(0·1126)**	0·0277 (0·0340)	0·0735 (0·0901)						
TZ5			0·0126 (0·0252)	**0·7134** **(0·0382)**	0·0371 (0·0742)						
TZ6					**0·7598** **0·1066**						
ZW1						**0·8633** **(0·1468)**	0·0949 (0·0775)		0·0438 (0·0876)	0·0304 (0·0609)	
ZW2							**0·6887** **(0·0005)**				
ZW3		0·0182 (0·0063)		0·0235 (0·0040)		0·0445 (0·0889)	0·0330 (0·0660)	**0·9845** **(0·0008)**			
SA1						0·0443 (0·0887)	0·0351 (0·0633)		**0·9233** **(0·1202)**	0·1218 (0·0609)	
SA2										**0·6907** **(0·0005)**	
MG1											**0·9900** **(0·0003)**

discordance between plastid and nuclear variation further strengthens the indication of a selective component determining group distribution, as nuclear variation may also reflect site adaptation (cf. Petit *et al.*, 2001; Tsitrone *et al.*, 2003), and obviously was influenced by gene flow to a lesser extent than cytoplasmic variation.

Madagascar

Madagascar holds the most diverged populations and as suggested by Kadu *et al.* (2011) the colonization of that island was probably the result of a single dispersal event that dates back to the oldest splitting events within the species. This finding is in agreement with previous studies showing very low rates of propagule exchange between Madagascar and mainland Africa (Meve and Liede, 2002). The high estimated ancestral population size also indicates long-term isolation and an independent evolution of this island population. The long-term climatic stability of this large island has allowed undisturbed evolution and accumulation of a rich flora (Vences *et al.*, 2009). It is also evident that the colonization of Madagascar was independent from the colonization of southern Africa, because the lowest rate of historical gene flow was found among these groups (Table 5). Sampling from the Comoros Islands would be important for evaluating possible gene-flow routes between mainland Africa and Madagascar. The evidence for recent population decline due to recent fragmentation (as exemplified by the detected bottleneck) and also morphological divergence (C. A. C. Kadu *et al.*, unpubl. res.) strengthens the demand for the protection of these populations.

Derived populations in southern and western Africa

The results clearly show that the populations in southern African are descendants from EG2 and founding population sizes in this region were small as compared with other areas. The split appears to be relatively recent and is probably mainly driven by drift (monomorphic loci, private alleles). Samples in southern Africa clearly clustered in the STRUCTURE computations, although they were collected over a relatively large geographical area. This suggests that populations in this region have only recently become fragmented and are probably derived from a single population expansion event. A significant barrier to gene flow within the region was only detected between populations from South Africa and Zimbabwe.

The MIGRATE analysis also sheds some light on the relationship between the East and West African populations of the species. Historical migration estimates between WG and EG1 were high, especially in the direction from EG1 to WG; considering both these results and patterns of diversity and previous results from plastid DNA (Kadu *et al.*, 2011), colonization from east to west appears to be the most plausible hypothesis. In fact, the only other study (to our knowledge)

TABLE 5. *Results of historic gene-flow analysis among population groups derived from MIGRATE*

		M (m/μ)				
	θ ($4N_e\mu$)	WG →	EG1 →	EG2 →	SG →	MG →
WG	2·4562 (2·2561–2·6808)		**2·6182** (2·3228–2·9376)	0·5895 (0·4542–0·7487)	0·1465 (0·0645–0·2315)	0·2266 (01480–0·3290)
EG1	4·3005 (4·0447–4·5784)	**1·388** (1·2220–1·5681)		**1·0171** (0·8767–1·1718)	0·1007 (0·0610–0·1547)	0·3796 (0·2966–0·4769)
EG2	3·803 (3·5489–4·0820)	0·0898 (0·0491–0·1486)	**1·3086** (1·1290–1·5063)		**1·5507** (1·3543–1·7653)	0·3543 (0·2647–0·4622)
SG	2·7108 (2·4981–2·9483)	0·5901 (0·4586–0·7447)	0·2525 (0·1701–0·3579)	**1·4215** (1·2102–1·6556)		0·0001 (0–0·0171)
MG	7·1183 (6·3558–8·0076)	0·2157 (0·1635–0·2778)	0·0655 (0·0389–0·1021)	0·1321 (0·0923–0·1819)	0·0023 (0–0·0139)	

θ = the estimated mutation-scaled population size, *M* = the mean estimated mutation-scaled migration rate (receiving populations in rows).
The upper and lower 99·5 % profile likelihood percentiles are given in brackets.
Values for $M > 1$ are in bold.

scrutinizing gene-flow patterns between West and East Africa is the study of Ackermann and Bishop (2010) on gorillas, in which an opposite direction of migration was inferred. Estimates of recent migration still indicate low levels of gene flow between these regions in *P. africana*. The most similar populations in WG were Bioko Island (GQ1) and CM1, indicating that the island population is a direct descendent of CM1; other populations from that region recently appear to have had much less contact. Given the long-term absence of the species north of the Congolian plateau and the existence of extant satellite populations in the Congo, Zambia and Angola, the southern migration corridor (*sensu* White/Moreau) south of the central African rain-forest region may have served as the migration corridor for this movement. In addition, aridification of Africa during the Pleistocene has driven forests south (Hamilton and Taylor, 1991), making this route of migration even more plausible. Sampling from these areas (western Tanzania, Zambia and Angola) would be needed to solve this question finally. Also future collections of *P. africana* from riverine populations in Southern Africa and other regions appear promising.

Conclusions

Based on the results of this study, we conclude that the biogeography of *P. africana* is multifaceted and has been determined by rare long-distance dispersal events coupled with constant migration at intermediate geographical ranges and strong gene-flow barriers. Pleistocene climatic variation was probably a decisive factor in the biogeography of the species. The observed patterns corroborate the main result of our previous plastid DNA study (Kadu *et al.*, 2011). Comparing nuclear data with the plastid DNA study, we found evidence for the dislocation of a historical immigration barrier to a more recent barrier to gene flow over several hundreds of kilometres, exemplifying the highly dynamic environmental history of African highlands. More generally the dynamic history of the complete continent, similar to South America (Colinvaux *et al.*, 2000), needs to be investigated in future studies on species from Afromontane regions (east, west and south) throughout the range to resolve biogeographical patterns in order to understand causes and factors impacting evolution of the species. Investigations into this and other species with disjunct distributions in West and East Africa (and the Rift Valley in particular) and other parts of the continent are strongly encouraged to elucidate further the underlying processes. Promising candidates are suggested by White *et al.* (1981), and include *Myrica salicifolia*, *Nuxia congesta* and *Podocarpus latifolius*. Other species that could be considered are *Podocarpus falcatus*, distributed from Ethiopia to South Africa, and *Maytenus gracilipes* occurring from Ethiopia to Zaire and to Cameroon.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Table S1: details of allelic richness for the 25 *P. africana* populations studied. Table S2: estimates of null allele frequency for the 25 *P. africana* populations studied.

ACKNOWLEDGEMENTS

We thank all persons involved in carrying out the sample collections. This research was financed by the Austrian Development Agency through collaboration between BFW and Bioversity International. We kindly acknowledge colleagues in Cameroon, as well as Alistair Jump, Stephen Cavers and Hazel Chapman for providing additional samples of *P. africana*. We thank Thomas Thalmayr and Johanna Kohl for help with computers and creating the figures. We gratefully acknowledge Tony Cunningham for invaluable initial advice on sampling. The authors also thank Dr Michael Fay and two anonymous reviewers for their valuable suggestions and comments on the manuscript.

LITERATURE CITED

Ackermann RR, Bishop JM. 2010. Morphological and molecular evidence reveals recent hybridization between gorilla taxa. *Evolution* **64**: 271–290.
Allal F, Sanou H, Millet L, *et al.* 2011. Past climate changes explain the phylogeography of *Vitellaria paradoxa* over Africa. *Heredity* **107**: 174–186.

Assefa A, Ehrich D, Taberlet P, Nemomissa S, Brochmann C. 2007. Pleistocene colonization of afro-alpine 'sky islands' by the artic-alpine *Arabis alpina*. *Heredity* **99**: 133–142.

Ayele TB, Gailing O, Umer M, Finkeldey R. 2009. Chloroplast DNA haplotype diversity and postglacial recolonization of *Hagenia abyssinica* (Bruce) J.F. Gmel. in Ethiopia. *Plant Systematics and Evolution* **280**: 175–185.

Beerli P. 2010. Migrate-n*, version 3·2 documentation*. Available from http://popgen.sc.fsu.edu/Migrate/Migrate-n.html.

Beerli P, Felsenstein J. 2001. Maximum-likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the USA* **98**: 4563–4568.

Born C, Alvarez N, McKey D, *et al*. 2011. Insights into the biogeographical history of the Lower Guinea Forest Domain: evidence for the role of refugia in the intraspecific differentiation of *Aucoumea klaineana*. *Molecular Ecology* **20**: 131–142.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics* **19**: 233–257.

Cavers S, Munro RC, Kadu CAC, Konrad H. 2009. Transfer of microsatellite loci for the tropical tree *Prunus africana* (Hook.f.) Kalkman. *Silvae Genetica* **58**: 276–279.

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterization and cross-species amplification in *Prunus*. *Theoretical and Applied Genetics* **99**: 65–72.

Colinvaux PA, De Oliveira PE, Bush MB. 2000. Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews* **19**: 141–169.

Dainou K, Bizoux JP, Doucet JL, Mahy G, Hardy OJ, Heuertz M. 2010. Forest refugia revisited: nSSRs and cpDNA sequences support historical isolation in a wide-spread African tree with high colonization capacity, *Milicia excelsa* (Moraceae). *Molecular Ecology* **19**: 4462–4477.

Dawson IK, Powell W. 1999. Genetic variation in the Afromontane tree *Prunus africana*, an endangered medicinal species. *Molecular Ecology* **8**: 151–156.

Dieringer D, Schlötterer C. 2003. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**: 167–169.

Duminil J, Heuertz M, Doucet JL, *et al*. 2010. CpDNA-based species identification and phylogeography: application to African tropical tree species. *Molecular Ecology* **19**: 5469–5483.

Elenga H, Peyron O, Bonnefille R, *et al*. 2000. Pollen-based biome reconstruction for southern Europe and Africa 18 000 yr BP. *Journal of Biogeography* **27**: 621–634.

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.

Excoffier L, Lischer HEL. 2010. ARLEQUIN suite ver 3·5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**: 564–567.

Farwig N, Braun C, Boöhning-Gaese K. 2008. Human disturbance reduces genetic diversity of an endangered tropical tree, *Prunus africana* (Rosaceae). *Conservation Genetics* **9**: 317–326.

Felsenstein J. 1989. PHYLIP – phylogeny inference package (Version 3·2). *Cladistics* **5**: 164–166.

Finch J, Leng MJ, Marchant R. 2009. Late Quaternary vegetation dynamics in a biodiversity hotspot, the Uluguru Mountains of Tanzania. *Quaternary Research* **72**: 111–122.

Gade DW. 1996. Deforestation and its effects in highland Madagascar. *Mountain Research and Development* **16**: 101–116.

Goudet J. 1995. FSTAT (Version 1·2): a computer program to calculate *F*-statistics. *Journal of Heredity* **86**: 485–486.

Hall JB, O'Brien EM, Sinclair FL. 2000. Prunus africana*: a monograph*. School of Agricultural and Forest Sciences Publication Number 18, University of Wales, Bangor. 104 pp.

Hamilton AC. 1982. *Environmental history of East Africa: a study of the Quaternary*. London: Academic Press.

Hamilton AC, Taylor D. 1991. History of climate and forests in tropical Africa during the last 8 million years. *Climatic Change* **19**: 65–78.

Hedberg O. 1969. Evolution and speciation in a tropical high mountain flora. *Botanical Journal of the Linnean Society* **1**: 135–148.

Heuertz M, Hausman JF, Hardy OJ, Vendramin GG, Frascaria-Lacoste N, Vekemans X. 2004. Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution* **58**: 976–988.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions Royal Society London B* **359**: 183–195.

Jolly D, Taylor DM, Marchant RA, Hamilton AC, Bonnefille R, Riolett G. 1997. Vegetation dynamics in central Africa since 18 000 yr BP: pollen records from the interlacustrine highlands of Burundi, Rwanda and western Uganda. *Journal of Biogeography* **24**: 495–512.

Kadu CAC, Schueler S, Konrad H, *et al*. 2011. Phylogeography of the Afromontane *Prunus africana* reveals a former migration corridor between East and West African highlands. *Molecular Ecology* **20**: 165–178.

Kebede M, Ehrich D, Taberlet P, Nemomissa S, Brochmann C. 2007. Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology* **16**: 1233–1243.

Koffi KG, Hardy OJ, Doumenge C, Cruaud C, Heuertz M. 2011. Diversity gradients and phylogeographic patterns in *Santiria trimera* (Burseraceae), a widespread African tree typical of mature rainforests. *American Journal of Botany* **98**: 254–264.

Linder HP. 2001. Plant diversity and endemism in sub-Saharan tropical Africa. *Journal of Biogeography* **28**: 169–182.

Livingstone DA. 1975. Late quaternary climatic change in Africa. *Annual Review of Ecology and Systematics* **6**: 249–280.

Logossa ZA, Camus-Kulandaivelu L, Allal F, *et al*. 2011. Molecular data reveal isolation by distance and past population expansion for the shea tree (*Vitellaria paradoxa* C.F. Gaertn) in West Africa. *Molecular Ecology* **20**: 4009–4027.

Lovett JC. 1993. Climatic history and forest distribution in eastern Africa. In: Lovett JC, Wasser SK. eds. *Biogeography and ecology of the rain forests of eastern Africa*. Cambridge: Cambridge University Press, 23–29.

Maley J. 1987. Fragmentation de la forêt dense humide africaine et extension des biotopes montagnards au quaternaire recent: nouvelles donnees polliniques et chronologiques: implications paleoclimatiques et biogeographiques. *Palaeoecology of Africa* **18**: 307–334.

Maley J. 1996. The African rain forest: main characteristics of changes in vegetation and climate from the Upper Cretaceaous to the Quarternary. *Proceedings of the Royal Society of Edinburgh Section B* **104**: 31–73.

Manni FR, Guérard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**: 173–190.

Marchant RA, Taylor DM. 1998. A late holocene record of montane forest dynamics from south-western Uganda. *The Holocene* **8**: 375–381.

Marchant R, Taylor D, Hamilton A. 1997. Late Pleistocene and Holocene history at Mubwindi Swamp, southwest Uganda. *Quaternary Research* **47**: 316–328.

Mbatudde M, Mwanjololo M, Kakudidi EK, Dalitz H. 2012. Modelling the potential distribution of endangered *Prunus africana* (Hook.f.) Kalkm. in East Africa. *African Journal of Ecology* **50**: 393–403.

Meve U, Liede S. 2002. Floristic exchange between mainland Africa and Madagascar: case studies in Apocynacae-Asclepiadoideae. *Journal of Biogeography* **29**: 865–873.

Moreau RE. 1966. *The bird faunas of Africa and its islands*. New York, NY: Academic Press.

Muchugi A, Lengkeek AG, Kadu CAC, Muluvi GM, Njagi ENM, Dawson IK. 2006. Genetic variation in the threatened medicinal tree *Prunus africana* in Cameroon and Kenya: implications for current management and evolutionary history. *South African Journal of Botany* **72**: 498–506.

Nei M. 1972. Genetic distance between populations. *American Naturalist* **106**: 283–392.

Nicolas V, Mboumba JF, Verheyen E, *et al*. 2008. Phylogeographic structure and regional history of *Lemniscomys striatus* (Rodenita: Muridae) in tropical Africa. *Journal of Biogeography* **35**: 2074–2089.

Nordborg M. 2001. Coalescent theory. In: Balding D, Bishop M, Cannings C. eds. *Handbook of statistical genetics*. Chichester, UK: Wiley, 179–212.

Ohta T, Kimura M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research* **22**: 201–204.

Omondi SF, Kireger E, Dangasuk OG, *et al*. 2010. Genetic diversity and population structure of *Acacia senegal* (L.) Willd. in Kenya. *Tropical Plant Biology* **3**: 59–70.

van Oosterhout C, Hutchinson WF, Wills PM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.

Petit RJ, Pineau E, Bialozyt R, Brewer S, Cheddadi R, Combs B. 2001. From spatial patterns of genetic diversity to post glacial migration processes in forest trees. In: Silvertown J, Antonovic J. eds. *Integrating ecology and evolution in a spatial context.* Oxford: Blackwell Science, 295–318.

Petit RJ, Aguinagalde I, de Beaulieu J, *et al*. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.

Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**: 502–503.

Pock Tsy JML, 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.

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

Raymond M, Rousset F. 1995. GENEPOP (version 1·2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248–249.

Ruiz Guajardo JC, Schnabel A, Ennos R, Preuss S, Otero-Arnaiz A Stone G. 2010. Landscape genetics of the key African acacia species *Senegalia mellifera* (Vahl): the importance of the Kenyan Rift Valley. *Molecular Ecology* **19**: 5126–5139.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15**: 4261–4293.

Sosinski B, Gannavarapu M, Hager LD, *et al*. 2000. Characterization of microsatellite markers in peach *Prunus persica* (L.) Batsch. *Theoretical and Applied Genetics* **101**: 421–428.

Street-Perrott FA, Perrott RA. 1993. Holocene vegetation, lake levels and climate of Africa. In: Wright HE Jr, Kutzbach JE, Webb T III, Ruddiman WF, Street-Perrott FA, Bartlein PJ. eds. *Global climates since the Last Glacial Maximum*. MN: University of Minnesota Press, 318–356.

Tsitrone A, Kirkpatrick M, Levin DA. 2003. A model for chloroplast capture. *Evolution* **57**: 1776–1782.

Vaughan SP, Russell K. 2004. Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium. Molecular Ecology Notes* **4**: 429–431.

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.

White F. 1981. The history of the Afromontane archipelago and the scientific need for its conservation. *African Journal of Ecology* **19**: 33–54.

White F. 1983. Long-distance dispersal and the origins of the Afromontane flora. *Sonderbände des naturwissenschaftlichen Vereins in Hamburg* **7**: 87–116.

White F. 1993. Refuge theory, ice-age aridity and the history of tropical biotas: an essay in plant geography. *Fragmenta Floristica et Geobotanica* Suppl. 2: 385–409.

Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.