



**Cite this article:** Bell KL, Rangan H, Kull CA, Murphy DJ. 2015 The history of introduction of the African baobab (*Adansonia digitata*, Malvaceae: Bombacoideae) in the Indian subcontinent. *R. Soc. open sci.* **2**: 150370. <http://dx.doi.org/10.1098/rsos.150370>

Received: 26 July 2015

Accepted: 14 August 2015

**Subject Category:**

Biology (whole organism)

**Subject Areas:**

plant science

**Keywords:**

*Adansonia digitata*, genetic diversity, Indian subcontinent, population structure, introduction pathways, population genetics

**Author for correspondence:**

Karen L. Bell

e-mail: [karen.bell@emory.edu](mailto:karen.bell@emory.edu)

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsos.150370> or via <http://rsos.royalsocietypublishing.org>.


# The history of introduction of the African baobab (*Adansonia digitata*, Malvaceae: Bombacoideae) in the Indian subcontinent

Karen L. Bell<sup>1</sup>, Haripriya Rangan<sup>2</sup>, Christian A. Kull<sup>3</sup> and Daniel J. Murphy<sup>1</sup>

<sup>1</sup>Royal Botanic Gardens Victoria, Melbourne, Victoria 3004, Australia

<sup>2</sup>School of Geography, University of Melbourne, Victoria 3053, Australia

<sup>3</sup>Institut de Géographie et Durabilité, Université de Lausanne, Lausanne 1015, Switzerland

 KLB, 0000-0002-4445-4026; HR, 0000-0003-4238-2353; CAK, 0000-0002-7516-7898

To investigate the pathways of introduction of the African baobab, *Adansonia digitata*, to the Indian subcontinent, we examined 10 microsatellite loci in individuals from Africa, India, the Mascarenes and Malaysia, and matched this with historical evidence of human interactions between source and destination regions. Genetic analysis showed broad congruence of African clusters with biogeographic regions except along the Zambezi (Mozambique) and Kilwa (Tanzania), where populations included a mixture of individuals assigned to at least two different clusters. Individuals from West Africa, the Mascarenes, southeast India and Malaysia shared a cluster. Baobabs from western and central India clustered separately from Africa. Genetic diversity was lower in populations from the Indian subcontinent than in African populations, but the former contained private alleles. Phylogenetic analysis showed Indian populations were closest to those from the Mombasa-Dar es Salaam coast. The genetic results provide evidence of multiple introductions of African baobabs to the Indian subcontinent over a longer time period than previously assumed. Individuals belonging to different genetic clusters in Zambezi and Kilwa may reflect the history of trafficking captives from inland areas to supply the slave trade between the fifteenth and nineteenth centuries. Baobabs in the Mascarenes, southeast India and Malaysia indicate introduction from West Africa through eighteenth and nineteenth century European colonial networks.

# 1. Background

Innumerable plant species have been transported by humans around the world for a variety of practical and cultural uses. Many of these have geographically widespread or disjunct distributions, with no historical records for when and how they may have been dispersed from their places of origin to new sites [1]. Several studies of cultivated crops (e.g. [2–5]) and weeds (e.g. [6–8]) have sought to investigate the origins of taxa that lack historical documentation by using population genetics to match genotypes in the introduced and native ranges in order to determine source populations. However, genetic evidence of source populations is often insufficient for inferring agency, modes and timing of plant dispersals from their places of origin to new locations. Other evidence, such as archaeobotanical remains, can provide an indication of when the plants may have arrived and indicate possible agents and pathways of introduction [5]. Additional cultural evidence can come from comparing plant terms and uses between places of origin and introduction [9,10]. Combining all these forms of evidence can provide a more comprehensive historical understanding of the global biogeography of plant dispersals and distributions.

The charismatic tree genus *Adansonia* (Malvaceae: Bombacoideae) occurs in Africa, Madagascar and Australia, with a different lineage present in each land mass [11–13]. Molecular divergence dating suggests that the three main lineages shared a common ancestor 5–15 Ma (during the Miocene), with the current distribution hypothesized to be the result of long-distance hydrochory [12,14]. *Adansonia digitata* L., or the African baobab, is widely distributed across continental Africa. *Adansonia digitata* s.l. (i.e. in the broad sense) encompasses two described species, the tetraploid *A. digitata* s.s. and the diploid *A. kilima* [12]. The distribution of *A. digitata* s.s. extends across much of Sub-Saharan Africa, from the northern Sahel to South Africa's Limpopo province, but does not occur in the rainforest areas of central Africa, or higher altitude localities (above 800 m elevation) of eastern Africa [13]. Phylogeographic analyses have demonstrated populations of *A. digitata* s.l. in eastern and western Africa to be genetically distinct, with the eastern populations monophyletic within a paraphyletic grade of western populations [15]. *Adansonia kilima* is recorded in higher altitude localities (650–1500 m elevation), partially overlapping with the range of *A. digitata* s.s., from the eastern slopes of Kilimanjaro to northern South Africa, and west to the Caprivi Strip, Namibia [12]. Phylogenetic analysis has demonstrated that *A. digitata* and *A. kilima* are genetically similar, suggesting that tetraploidy evolved relatively recently [12]. In this study, we refer to this species-complex as *A. digitata* and hereafter use its common name, baobab.

Biogeographic studies of baobabs in the African continent show that there are close associations between past human settlements and the presence of the species [13,16]. The baobab is highly valued for a variety of food and artisanal uses and for its cultural symbolism extending over millennia [13,17–22]. The fruit has a powdery pulp surrounding numerous seeds with a hard endocarp. The sweet–sour pulp is particularly favoured and widely consumed, but the seeds are difficult to digest without additional processing. In some parts of Africa, the seeds are eaten after roasting or grinding into flour [23]. Most commonly though, the seed is discarded or passed intact through the digestive tract after eating the pulp. This may be the means by which human-mediated dispersal of the baobab occurred. Genetic studies also suggest that humans have been the prime agents responsible for distributing the tree species across the African continent [15].

The role of human agency in baobab dispersal is evident in its presence in the Caribbean and parts of tropical South America, where enslaved people were transported from West Africa between the sixteenth and nineteenth centuries to work on sugarcane (*Saccharum* spp.) plantations [24]. African baobab populations are also found in the Indian subcontinent and Sri Lanka [25–27], with sporadic distribution in various places around the Indian Ocean including Yemen, southern Iran [28], Comoros, northwest Madagascar and the Mascarene Islands [25,29], Malaysia and Indonesia [30]. The history of dispersal to these places is less well known. Although it has not been previously hypothesized, long-distance hydrochory is a possibility. Baum *et al.* [14] provided evidence that *Adansonia* in northwest Australia arrived from Africa via long-distance oceanic dispersal in the Miocene. The fruit of the African baobab has been demonstrated to remain viable after immersion in seawater for six months [15]. This could explain the coastal presence of the African baobab on the Indian subcontinent and other locations around the Indian Ocean. However, the lack of species-level divergence, according to morphological taxonomy, between baobabs in Africa and those found in the Indian subcontinent [11,13] would suggest a recent dispersal. This makes human-assisted dispersal a more plausible alternative than transoceanic drift.

The few published surveys of baobabs in the Indian subcontinent attribute the introduction of baobabs to Arab traders [26,27] or medieval Muslim rulers in the subcontinent that maintained African slave

armies [13,25], but without any supporting evidence from genetic, archaeobotanical or ethno-historical analyses [31]. These hypotheses imply a recent history of the baobab in the Indian subcontinent which stand in contrast to significant evidence of a deep prehistory of biotic exchange between Africa and the Indian subcontinent [32–37]. These exchanges include the introduction of cereals such as sorghum (*Sorghum bicolor*), pearl millet (*Pennisetum glaucum*) and finger millet (*Eleusine coracana*), and legumes, such as cowpea (*Vigna unguiculata*), and hyacinth bean (*Lablab purpureus*) from Africa into the Indian subcontinent during the second and first millennia BCE [38–42], and the movement of zebu cattle (*Bos indicus*) in the reverse direction [43]. There was also possibly an earlier introduction of tamarind (*Tamarindus indica*) from the Sudano-Sahelian region to the Indian subcontinent [44–46]. We hypothesize that the baobab, with similarly important nutritious and other use values, may have also been introduced from Africa to the Indian subcontinent in prehistoric times.

In this study, we investigate the history and pathways of introduction of baobabs from Africa to the Indian subcontinent and other parts of the Indian Ocean region. We combined genetic analysis of inferred ancestry and phylogeography with historical evidence of oceanic trade and migrations across the Indian Ocean to: (i) determine source populations of baobabs by examining genetic variation and relationships within and between African, Indian subcontinent and other Indian Ocean populations, (ii) associate these source populations to specific dispersal pathways based on historical trade routes and networks, and (iii) infer possible time periods of introductions based on genetic and historical evidence.

## 2. Material and methods

### 2.1. Sampling

Our sampling in Africa focused on the eastern lineage, based on the assumption that this was the most likely source for dispersal of baobabs across the Indian Ocean. Thorough surveys of baobabs were conducted in Mozambique and Tanzania, with further sampling in Kenya and South Africa. A single sample was included from Senegal as a reference sample for baobabs in the West African genetic group [15]. Baobabs were sampled from multiple introduced populations in India, along with three individuals from the islands of Réunion and Mauritius, and one individual from Penang, in Malaysia. For details, see the electronic supplementary material, appendices S1 and S2.

Leaves were collected into plastic bags and dried with silica gel. If trees were not in leaf, then bark samples were used. Although bark samples are more likely than leaf samples to contain DNA from other organisms, such as endophytes and insects, it is highly unlikely that the *A. digitata*-specific primers would amplify them. To confirm that this had not occurred, we checked results from these samples for the presence of more than four bands per locus, to confirm that no extraneous DNA had amplified.

Since baobab trees tend to occur as isolated individuals throughout their range rather than forming distinct populations, in many locations one tree was sampled. However, for some populations, samples were taken from up to 19 individuals. Voucher specimens were retained for a representative subset of individuals and deposited in the National Herbarium of Victoria (MEL).

### 2.2. DNA isolation and microsatellite genotyping

Plant DNA was isolated from samples using DNeasy Plant Mini Kits (QIAGEN), with modifications to the manufacturer's protocol as described by Pettigrew *et al.* [12]. The microsatellite loci Ad01, Ad02, Ad06, Ad08, Ad09, Ad12, Ad13, Ad14, Ad15 and Ad18 [47] were amplified following the method of James *et al.* [48], modified for multiplexing polymerase chain reactions (PCRs) using the Type-It Microsatellite PCR Kit (QIAGEN). Amplification reactions contained a final concentration of 1x PCR Master Mix (QIAGEN), 0.075  $\mu$ M each multiplexed forward primer appended to the 454A sequencing tag (Applied Biosystems, Foster City, CA, USA), 0.25  $\mu$ M each reverse primer, 0.1  $\mu$ M per multiplexed locus of 454A sequencing tag labelled with either 6-FAM, NED, HEX or PET (Applied Biosystems). Thermal cycling followed the instructions provided with the Type-It Kit. Specifically, initial heat activation of 5 min at 95°C was followed by 28 cycles of denaturation for 30 s at 95°C, annealing for 90 s at 60°C, and extension for 30 s at 72°C, with a final extension of 30 min at 60°C. Following PCR, amplifications using compatible dye types were diluted to equal concentrations and combined, then separated on an ABI 3730XL sequencer with a GS500LIZ size standard (Applied Biosystems) at Macrogen Inc. (Seoul, Korea). Allele sizes were scored using the Geneious microsatellite plugin v. 1.0.0 (Biomatters Ltd).

### 2.3. Genetic diversity

Allele frequencies and measures of genetic diversity were examined for three sets of data: the entire range of the baobab, African populations only and Indian populations only. This allowed us to assess whether any changes could be detected in response to founder effects or admixture of lineages, which may be associated with recent introduction to India, and to test for *in situ* evolution, which may have occurred under the scenario of an ancient introduction. Specifically, we used GenoDive to test whether populations and loci were in Hardy–Weinberg equilibrium (HWE), and we used SPAGeDi v. 1.3a [49] to calculate the number of alleles per locus, expected heterozygosity under HWE, inbreeding coefficient ( $F_{is}$ ) and number of private alleles.

To examine whether there was a recent reduction in effective population size of baobabs, as would be expected from a founder effect, deviations from mutation-drift equilibrium were tested using BOTTLENECK v. 1.2.02 [50]. Expected heterozygosities were simulated with 1000 replications under the Two Phased Model of Mutation with default settings (variance 30, 70% of mutations fitting the stepwise mutation model). Three statistical tests are available in the BOTTLENECK software to compare observed versus expected heterozygosities. The ‘standardized differences test’ is unsuitable for less than 20 loci [50]. We therefore used the ‘sign test’ and ‘Wilcoxon sign-rank test’, which are less sensitive to the low number of loci, and the allele frequency distribution was compared to that expected under mutation-drift equilibrium.

### 2.4. Population structure

A Bayesian classification scheme was used to identify genetic clusters and evaluate their spatial distribution, in order to determine potential connections between source and introduced populations and to assess overall genetic structure. This analysis was implemented in STRUCTURE v. 2.3.3 [51], without *a priori* grouping of individuals into populations. In tetraploid species such as the baobab, the genotype of partial heterozygotes (with two to three different alleles at a locus) is ambiguous. To avoid making assumptions about the genotypes of these individuals, we created a binary matrix of presence or absence of allele variants and treated these as dominant markers [52]. An admixture model, with loci assumed to be unlinked, was used to calculate the likelihood of different values for the number of clusters ( $K$ ). Simulations consisted of a burn-in of 50 000 iterations followed by 100 000 Markov chain Monte Carlo iterations. Ten runs for each  $K = 1–15$  were carried out on an SGI Altix XE Cluster through the Victorian Life Sciences Computing Initiative. The value(s) of  $K$  that best explained the structure in the data was/were determined using STRUCTURE HARVESTER WEB v. 0.6.93 [53], following the  $\Delta K$  method of Evanno *et al.* [54]. To assign individuals to genetic clusters, multiple runs at the selected value of  $K$  were combined using CLUMPP v. 1.1.2 [55], and plotted graphically using DISTRUCT v. 1.1 [56]. Cluster assignment of individuals and populations was visualized spatially using PHYLOGEOVIZ [57].

The statistical significance of spatial genetic clustering identified in the STRUCTURE analysis, for selected values of  $K$ , was tested using analysis of molecular variance (AMOVA) [58] performed in ARLECORE v. 3.5.1.3 [59]. Populations were allocated to regions, based on the predominant STRUCTURE cluster to which the individuals of that population belonged. Samples from West Africa, the Mascarene Islands and Malaysia were excluded due to insufficient numbers. Analyses were based on the  $F_{ST}$  measure of genetic diversity, with 9999 permutations.

Isolation by distance (IBD) was examined among individuals through a Mantel test (linearized genetic distance versus  $\ln(1 + \text{geographical distance})$ , to ensure that this number is non-zero); 999 permutations) conducted in GENALEX v. 6.41 [60]. Genetic distance was calculated as the binary genetic distance in GENALEX. Since geographical structure in the form of genetic clustering and in the form of IBD can be confounded, further Mantel tests and partial Mantel tests were carried out to limit this effect using the VEGAN package [61] in R [62]. A Mantel test of genetic distance versus geographical distance, with 999 permutations, was carried out within the geographical regions (defined above) to remove the confounding effect of genetic clustering from the test of IBD within a region. Conversely, a partial Mantel test of genetic distance versus region (0 for samples from the same region, 1 for samples from different regions), with geographical distance as a covariate, was used to test whether there is genetic structure that cannot be solely explained as IBD [63].

A principal coordinates analysis (PCoA), conducted using GENALEX v. 6.41 [60], was also used to examine genetic clustering defined by STRUCTURE at  $K = 7$ . Clusters with disjunct distribution were subdivided into multiple populations and samples with missing data were excluded from the analysis. This left a total of seven groups under consideration. Principal coordinates were calculated based on

the simple genetic distance (i.e. the total number of allelic differences) between each pair of samples, via a covariance matrix with data standardization. The first two principal coordinates were retained and plotted.

Intraspecific phylogenetic analysis was used to investigate the relationships between regions and to provide further evidence for inferring source populations and dispersal pathways. The biogeographic regions as defined by STRUCTURE at  $K=7$  were used, with the inclusion of West Africa and the Mascarenes, and individuals were assigned to these regions. Loci were bootstrapped to generate 1000 pseudoreplicates and genetic distances between regions were calculated with the  $D_C$  method [64] for each bootstrapped data matrix in PHYLIP v. 3.69 [65]. These distances were then used to construct unrooted neighbour-joining trees, and a 50% majority-rule bootstrap consensus tree was generated in PHYLIP.

## 3. Results

### 3.1. Genetic diversity

In general, genetic diversity parameters were found to be lower in baobabs from India compared to Africa. The number of alleles per locus and heterozygosity were both lower in India, although  $F_{IS}$  values were similar and a small number of private alleles were present in the Indian samples (table 1). Measures of  $G_{IS}$  within populations and loci frequently demonstrated deviation from HWE (electronic supplementary material, appendix S3). Tests for deviations from mutation-drift equilibrium, based on the 'sign test' and 'Wilcoxon test' in BOTTLENECK, were non-significant under all statistical tests, and allele frequencies showed a normal L-shaped distribution. This implies that there have been no recent bottlenecks or reductions in effective population size [50]. However, since tests for genetic bottlenecks based on deviations from mutation-drift equilibrium have low sensitivity [50,66], we cannot rule out the possibility of much older bottlenecks for baobabs in the Indian subcontinent.

### 3.2. Population structure

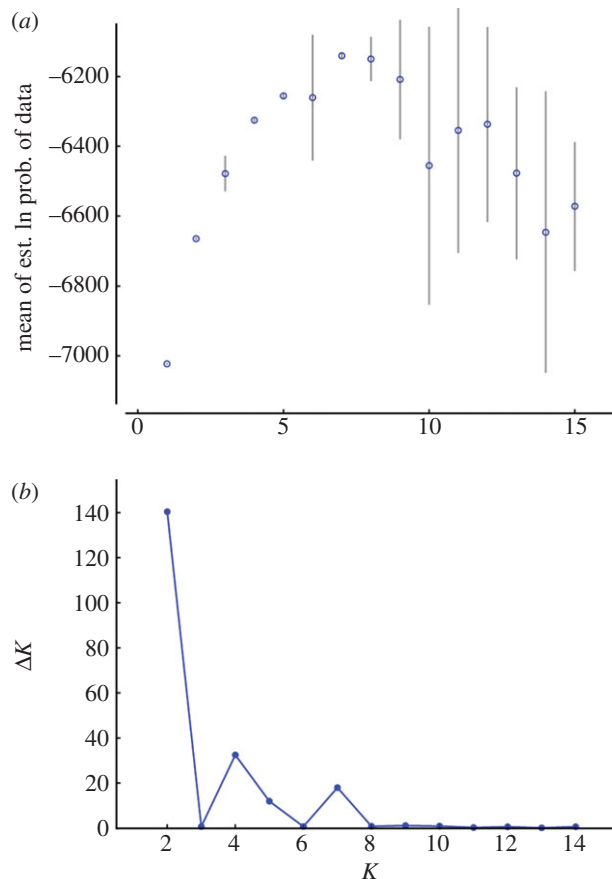
Likelihood values for STRUCTURE runs at different values of  $K$  increased rapidly with increasing  $K$ , up to  $K=4$ , after which the likelihood increased more slowly or decreased, and the standard deviations were greater (figure 1*a*). The highest value for  $\Delta K$ , indicating the number of clusters that best explained the data [54], was recorded at  $K=2$  (figure 1*b*). A peak at  $K=2$  in STRUCTURE analyses can be an artefact (e.g. [42,67]), so we also examined clustering based on the secondary peaks at  $K=4$  and 7. The inferred clusters at higher  $K$ -values nested within the clusters inferred at lower  $K$ -values (figures 2–4). Most individuals were assigned to populations with high certainty at each of these values of  $K$ , although a small number of individuals were ambiguously assigned. Across all  $K$ -values,  $\alpha$  was low, implying that these ambiguous assignments are due to uncertainty of origin, rather than admixture.

At  $K=2$ , the samples were divided into an Inland and a Coastal cluster (figure 2*a,b*). The Inland cluster included individuals from inland East Africa, southeast India (Chennai), Malaysia and Senegal. The Coastal cluster included individuals from coastal and lowland East Africa and India (excluding Chennai). Individuals from Tete, in Mozambique's Zambezi River valley, and from Réunion and Mauritius included a mixture of both clusters.

At  $K=4$ , the clusters were: (1) southern Africa (SA) and Malaysia, (2) southern coast of East Africa between Mozambique and southern Tanzania (SCEA), (3) inland Tanzania (IT), and (4) coastal region between Mombasa–Dar es Salaam and India (MDES). Baobabs from Tete included individuals with similarity to the SA cluster and others similar to MDES, and those from Senegal, the Mascarenes and Chennai equal likelihood of belonging to all four clusters (figure 3*a,b*).

At  $K=7$ , the clusters were: (1) SA, (2) SCEA, (3) inland Tanzania excluding Kilimanjaro (ITXK), (4) Kilimanjaro (KM), (5) Senegal, Chennai, Mascarenes and Malaysia (WA), (6) western India (WI), and (7) central India (CI). Individuals from the coastal region between Mombasa and Dar es Salaam were unable to be assigned to clusters, but showed similarities to clusters 4, 5, 6 and 7 (figure 4*a,b*).

The AMOVA supported the clustering patterns identified in the STRUCTURE analysis at  $K=2$ , 4 and 7 (table 2). For all tested population structures (from different values of  $K$ ), genetic clustering was significantly higher than expected by chance. The partitioning of genetic variance among clusters, among populations within clusters, and within populations was similar across the different population structures that we tested. For subsequent analyses, we defined regions based on the structuring at  $K=7$ , as a more fine-scaled explanation of the population structure.



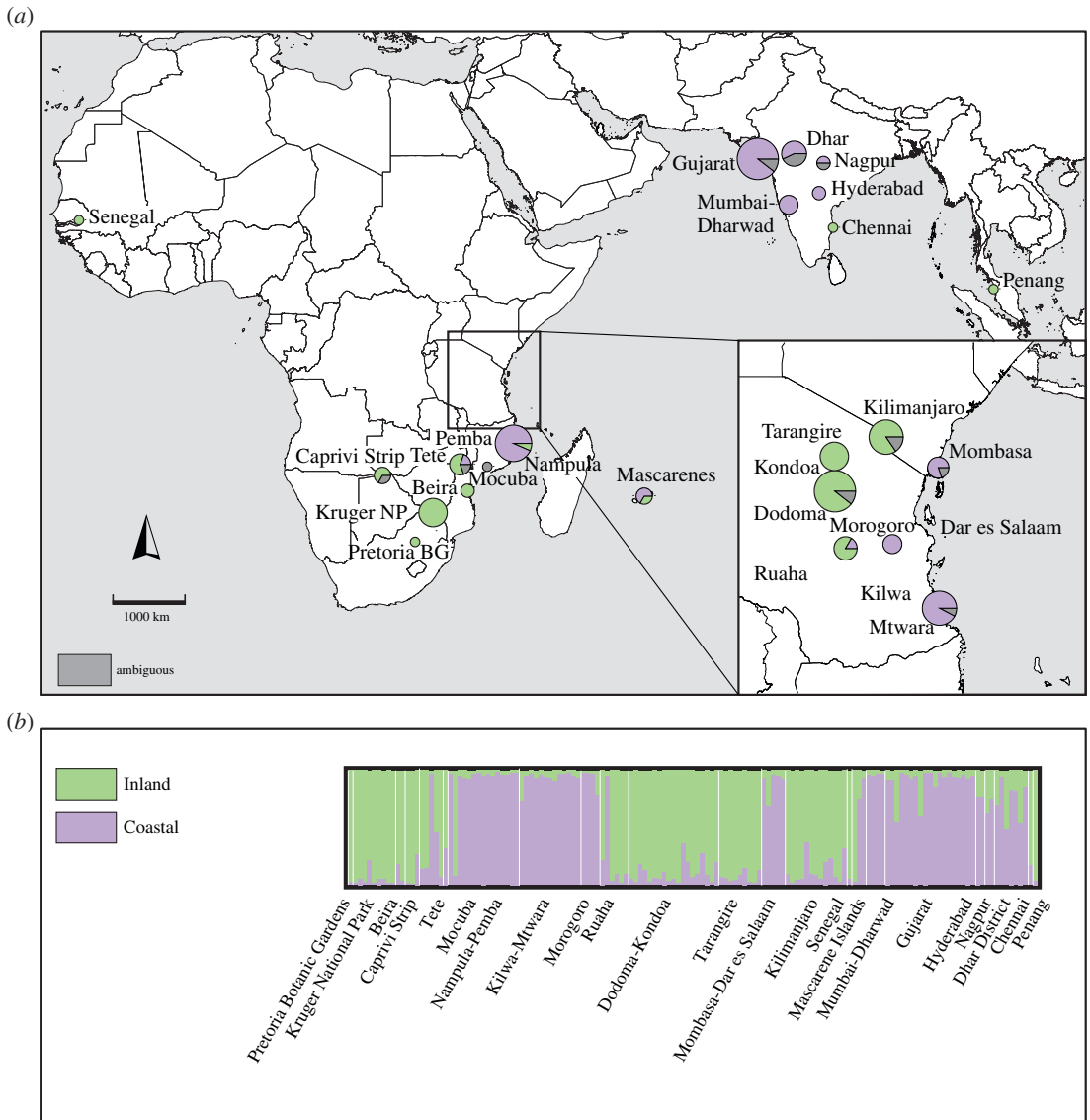
**Figure 1.** Comparison of the results of STRUCTURE analyses for genetic clustering of *Adansonia digitata* for  $K$  (number of clusters) between 1 and 15. (a) Mean and standard deviation of  $\ln(\text{likelihood})$  for each value of  $K$ . (b)  $\Delta K$  statistic of Evanno *et al.* [54] for each value of  $K$ .

**Table 1.** Population statistics (mean  $\pm$  s.d. across loci) for all *Adansonia digitata* populations, for African populations only and for Indian populations only.  $N_A$ , number of alleles per locus;  $H_e$ , expected heterozygosity;  $F_{IS}$ , inbreeding coefficient.

region	$n$	$N_A$	private alleles	$H_e$	$F_{IS}$
Africa	106	$15.5 \pm 3.6$	$5.8 \pm 1.9$	$0.823 \pm 0.113$	$-0.011 \pm 0.108$
India	35	$10.2 \pm 3.5$	$0.9 \pm 1.0$	$0.773 \pm 0.153$	$-0.006 \pm 0.098$
India (excluding Chennai)	34	$10.2 \pm 3.5$	$0.6 \pm 0.8$	$0.766 \pm 0.160$	$-0.008 \pm 0.101$
all	144	$16.2 \pm 3.7$	n.a.	$0.819 \pm 0.119$	$-0.007 \pm 0.095$

Mantel tests and partial Mantel tests provided additional support for the clustering by the STRUCTURE analysis. The Mantel test for IBD across the entire range was significant ( $R_{XY} = 0.129$ ;  $p = 0.001$ ), demonstrating that individuals from distant populations were more genetically divergent than individuals from nearby populations. When permutations were carried out only within  $K = 7$  genetic clusters, the Mantel test was no longer significant ( $p = 0.142$ ). The partial Mantel test of the  $K = 7$  genetic clusters, with geographical distance as a covariate, was significant ( $R_{XY} = 0.070$ ;  $p = 0.001$ ). The latter tests showed that the pattern identified in the first Mantel test could be explained better as genetically cohesive clusters of populations, rather than a more gradual pattern of isolation by distance. Statistical support from AMOVA and Mantel tests aligned well with the STRUCTURE analysis, justifying use of the  $K = 7$  clusters for further analyses.

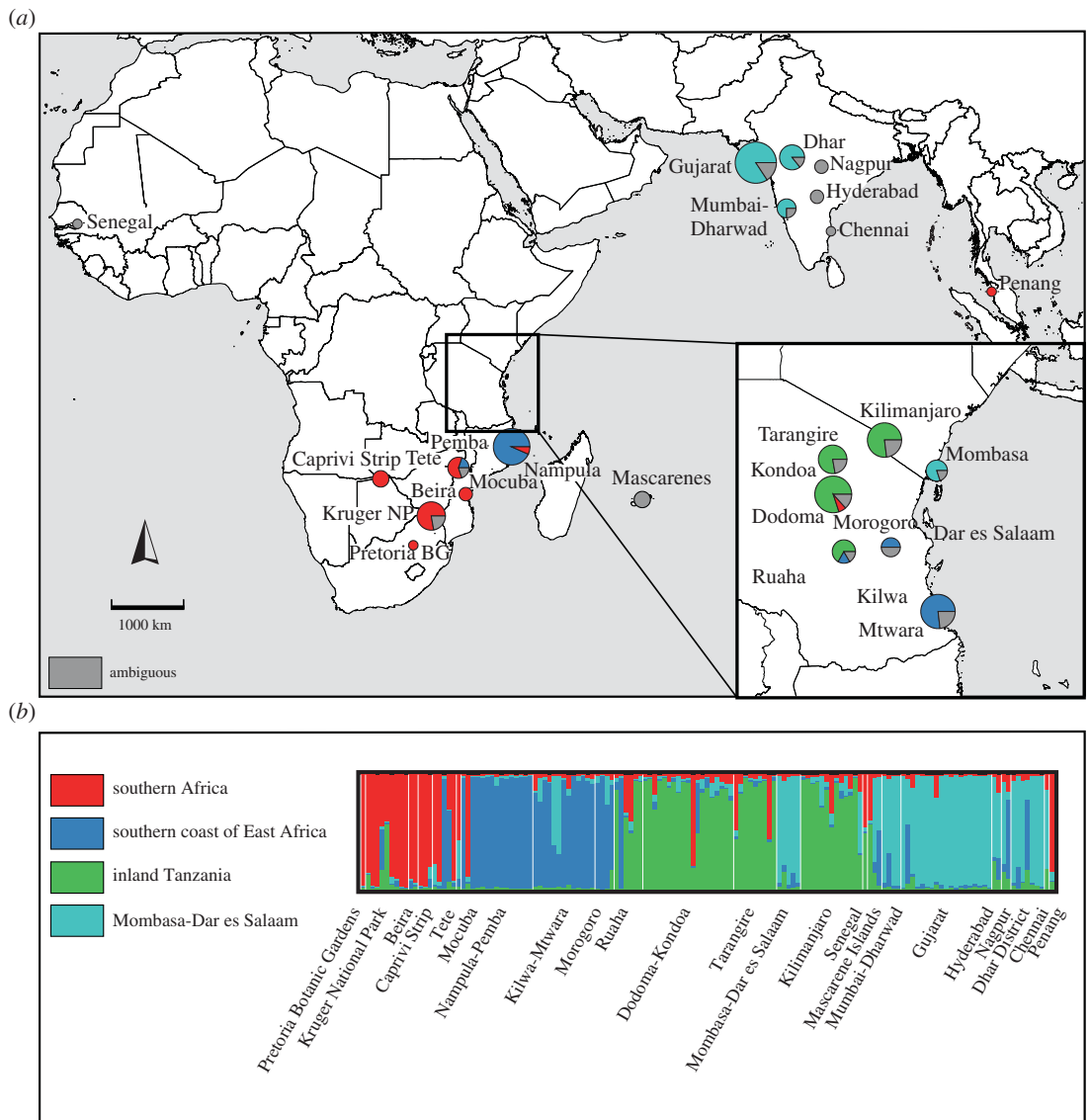
The PCoA demonstrated weak differentiation between the clusters identified in the STRUCTURE analysis, with substantial overlap between clusters in coordinate space (figure 5). For the first principal coordinate ( $x$ -axis; explaining 23.51% of variation), individuals from the MDES,



**Figure 2.** STRUCTURE analysis at  $K = 2$ . (a) Geographical localities of *Adansonia digitata* populations, and proportion of individuals within populations belonging to each genetic cluster. Assignment is considered to be ambiguous if the  $Q$ -value is less than 0.75. The size of the pie chart is proportional to the number of sampled individuals. (b) Proportional assignment of *Adansonia digitata* individuals to genetic clusters.

SCEA, KM, CI and WI clusters tended to have positive values; those from the SA cluster tended to have negative values, and those from ITXK were closer to zero. The second principal coordinate ( $y$ -axis; explaining 20.07% of variation), differentiated CI from WI individuals, and KM from SCEA, but with some overlap between adjacent clusters. Considering both coordinates, the WI individuals appear closer to those from SCEA and the CI individuals closer to MDES.

The unrooted neighbour-joining tree of baobab individuals from eastern Africa, India, West Africa and Mascarene Islands supports inferences from STRUCTURE analysis (figure 6). The Indian and eastern African individuals clustered together, and the West African and Mascarene individuals clustered with relatively long branches leading to the latter two populations. Within the southern and eastern African regional populations, KM and ITXK formed a clade, as did SA and SCEA. The MDES cluster formed a clade with WI, and the CI cluster was sister to MDES and WI. The branches leading to the Indian populations were long relative to internal branches, consistent with the private alleles detected.



**Figure 3.** STRUCTURE analysis at  $K = 4$ . (a) Geographical localities of *Adansonia digitata* populations, and proportion of individuals within populations belonging to each genetic cluster. Assignment is considered to be ambiguous if the  $Q$ -value is less than 0.75. The size of the pie chart is proportional to the number of sampled individuals. (b) Proportional assignment of *Adansonia digitata* individuals to genetic clusters.

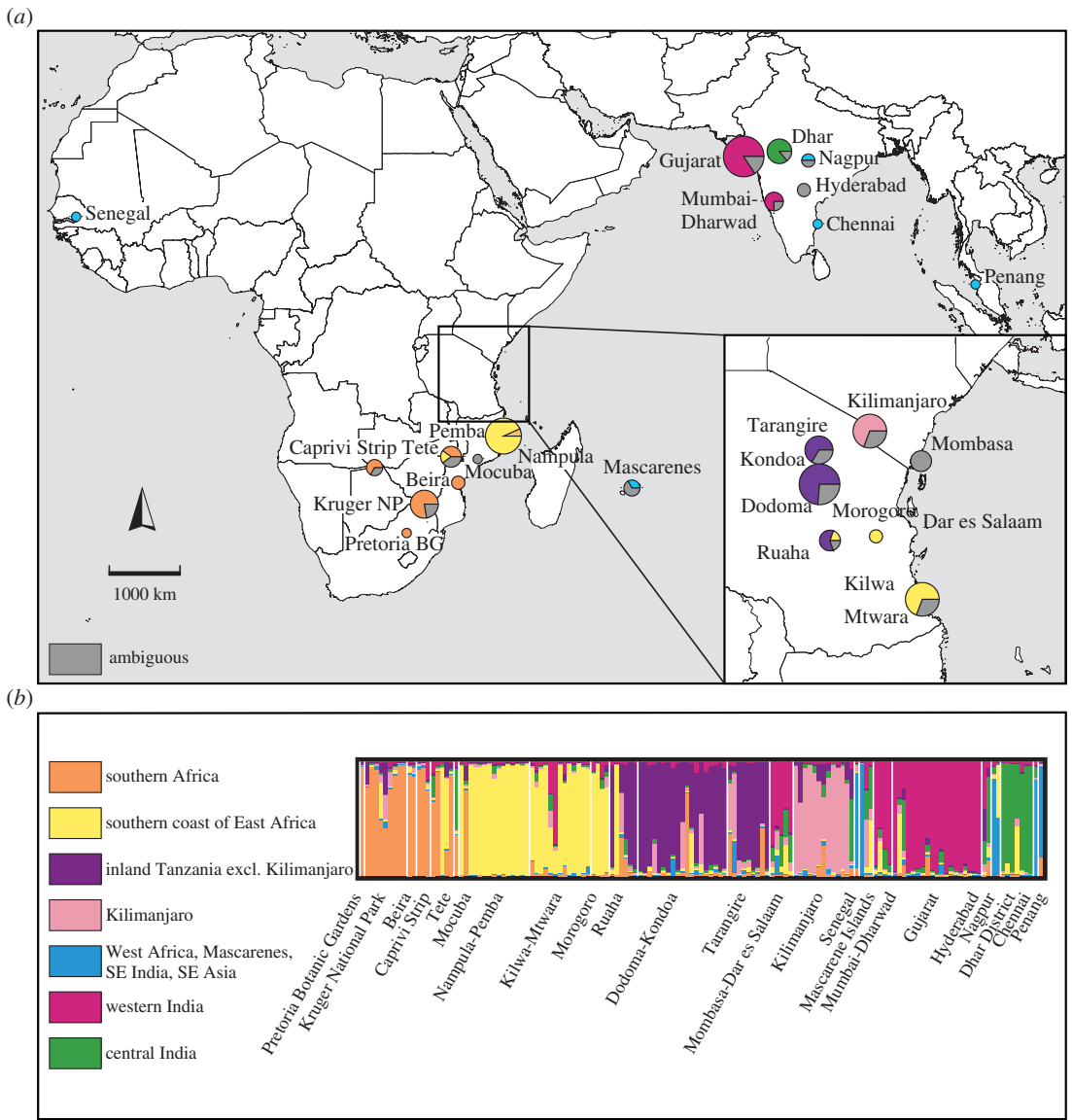
## 4. Discussion

### 4.1. Population structure and introductory relationships

The low genetic diversity in the Indian baobab populations compared with those from Africa is consistent with introduction to the Indian subcontinent. The STRUCTURE analysis showed clusters associating introduced populations in India with source populations in Africa at  $K=2$  and 4. It also provides evidence for at least two introductions to India. The Chennai population appears most closely related to West African populations and might thus represent a modern introduction. The two major clusters from the Indian subcontinent (CI and WI) are closely related to (and likely derived from) the Mombasa-Dar es Salaam populations. However, they also differed from African clusters and contained private alleles.

There could be two reasons for the presence of private alleles. First, mutations may have occurred *in situ*, suggesting that these populations have been present for more than one generation. Hence the presence of private alleles could imply that the existing trees are not founding individuals. Around 20% of the trees sampled in India for the current study are above 1400 cm girth and, using Swart’s estimate



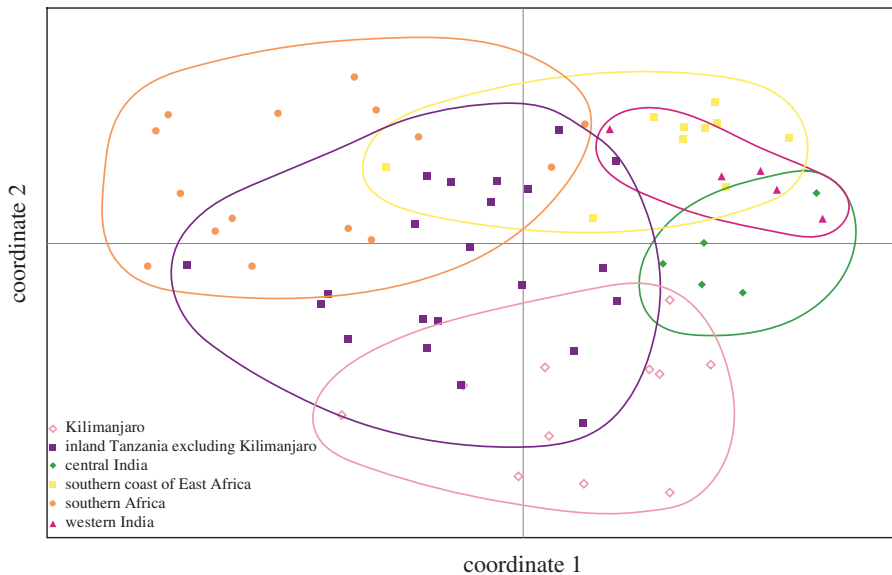


**Figure 4.** STRUCTURE analysis at  $K = 7$ . (a) Geographical localities of *Adansonia digitata* populations, and proportion of individuals within populations belonging to each genetic cluster. Assignment is considered to be ambiguous if the  $Q$ -value is less than 0.75. The size of the pie chart is proportional to the number of sampled individuals. (b) Proportional assignment of *Adansonia digitata* individuals to genetic clusters.

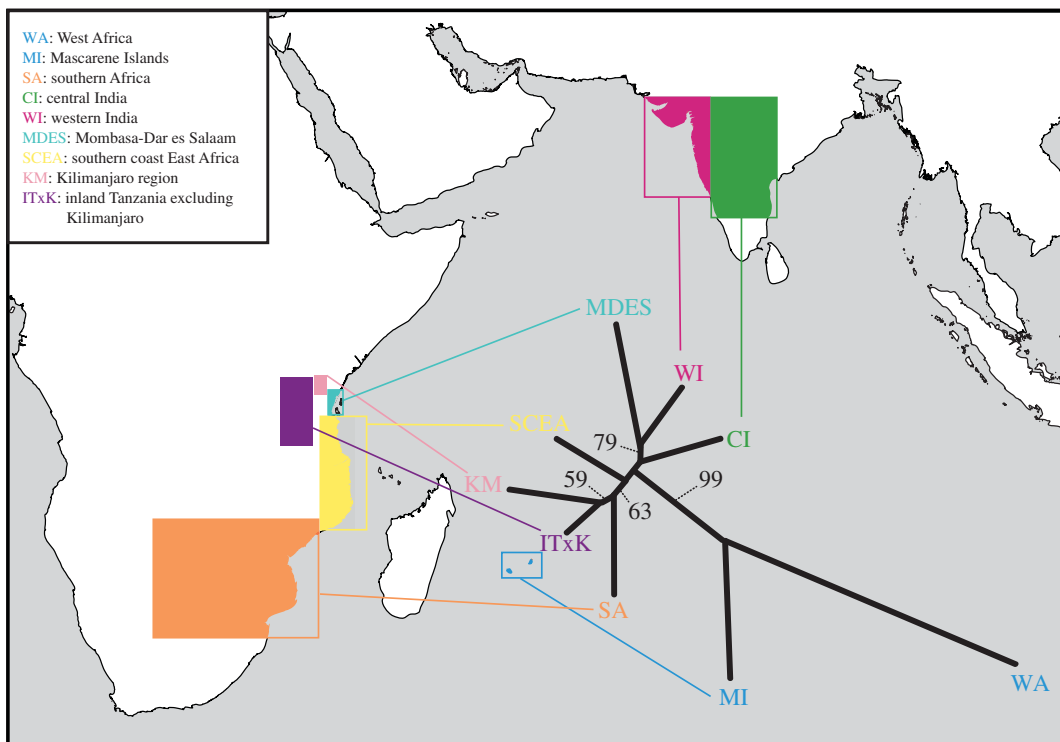
**Table 2.**  $F$  statistics,  $p$ -values, and proportions of variance at the level of individuals, populations and regions for AMOVA with populations clustered into regions based on STRUCTURE analysis at  $K = 2, 4$  and  $7$ .

	$K = 2$	$K = 4$	$K = 7$
variance among regions (%)	5.86	10.34	13.79
variance among populations (%)	11.42	6.49	2.52
variance within populations (%)	82.71	83.17	83.69
$F_{CT}$	0.059 ( $p = 0.00000$ )	0.103 ( $p = 0.00000$ )	0.138 ( $p = 0.00000$ )
$F_{SC}$	0.121 ( $p = 0.00000$ )	0.072 ( $p = 0.00000$ )	0.029 ( $p = 0.02931$ )
$F_{ST}$	0.173 ( $p = 0.00000$ )	0.168 ( $p = 0.00000$ )	0.163 ( $p = 0.00000$ )

of baobab age based on girth measurement and radiocarbon dating [68], could be more than a thousand years old (electronic supplementary material, appendix S4). Second, the absence of evidence of a recent bottleneck could suggest introductions from multiple source populations, or earlier bottlenecks that



**Figure 5.** Graphical representation of the first and second axes (explaining 23.51% and 20.07% of variation, respectively) of a PCoA, based on genetic distances between all samples of *Adansonia digitata* for which there were no missing data.



**Figure 6.** Unrooted neighbour-joining phylogeny of *Adansonia digitata* populations, based on genetic distances ( $D_c$ ) at the regional level. Geographical regions are identified by coloured boxes on the surrounding map. Bootstrap values above 50% are represented on branches.

could not be detected as deviations from mutation-drift equilibrium. These earlier bottlenecks could be the result of introduction by ancient humans.

Within the African clusters, the STRUCTURE analysis revealed individuals with ancestry from multiple clusters at  $K=4$  for SCEA and ambiguously assigned individuals at  $K=7$  for MDES, which either indicates dispersal from different regional clusters or the presence of unique individuals not matching other sampled populations. The phylogenetic analysis shows that MDES populations were the most likely African source for introductions to the Indian subcontinent, but the long branches leading to CI and

WI may indicate relationships to other non-sampled sources, perhaps to those from further north near the Horn of Africa. However, the evidence of individuals assigned to both MDES and SCEA clusters suggests that both may have been source populations for the Indian introductions. The PCoA shows some clustering of populations, but with substantial overlap. This is consistent with substantial and repeated dispersals between African sources and Indian populations, preventing complete differentiation. The PCoA supports the close genetic similarity of WI and CI baobabs with the MDES cluster.

## 4.2. Correlations with historical trade routes and networks

The long history of trade networks across the Indian Ocean can provide evidence for reconstructing the introduction of baobabs from Africa to the Indian subcontinent and other regions of the Indian Ocean [37–39,69,70]. The PCoA shows close similarity between individuals from MDES and WI and CI regions. However, the STRUCTURE analysis provides evidence for SCEA as an additional source for introductions to India. The history of trade interactions between East African coastal centres and those in the Indian subcontinent suggests that the MDES populations were possibly the main source for introductions to CI, and the SCEA populations a significant source for WI introductions [46,52,71,72].

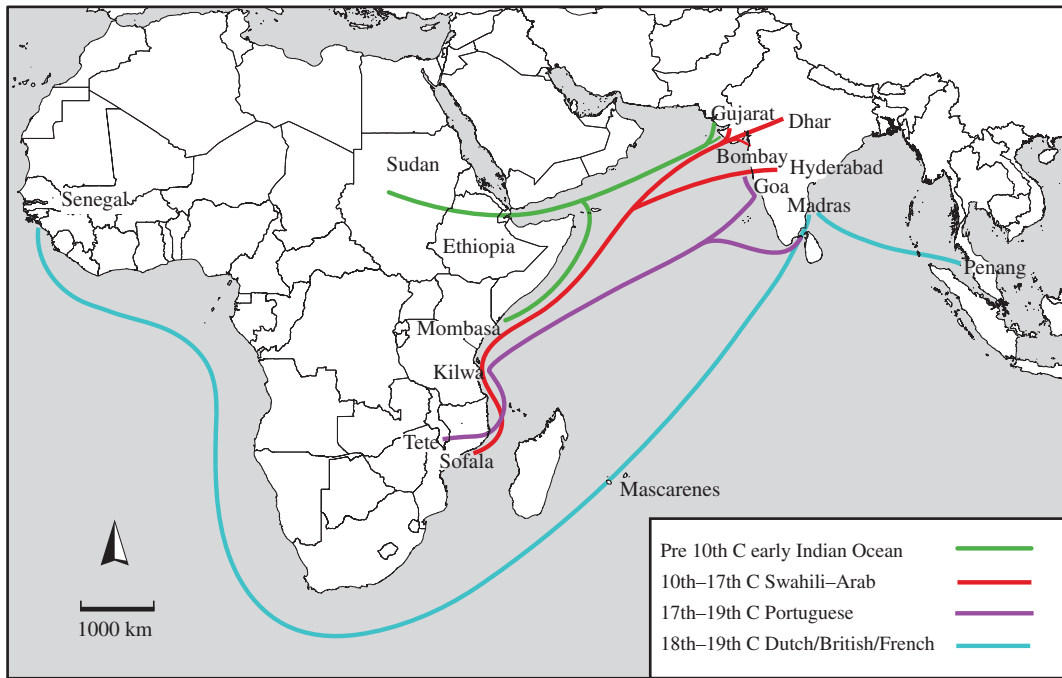
As noted earlier, the presence of private alleles in the Indian populations may indicate introductions from biogeographic regions of Africa not sampled in this study. These could be from the Sudanian Region or the Sahelian Transition Zone as per White's [69] chorological and vegetation mapping, or the Somalian and Ethiopian biogeographic regions as identified by Linder *et al.* [73]. The Somalian Region defined by Linder *et al.* extends across most of northeast Africa including the Horn, most of Kenya and the eastern half of Ethiopia; their Ethiopian Region includes significant elements of White's Sudanian and Afromontane (e.g. KM) regions. Hence the private alleles found in the Indian clusters could indicate introductions from populations in present-day Sudan, Ethiopia and Somalia, all of which have had histories of interactions with the Indian subcontinent extending over more than four millennia [74].

The presence of individuals in Tete-Mocuba with ancestry in the SA and SCEA clusters suggests that baobab dispersal occurred from these neighbouring populations into this area (figures 3 and 4;  $K = 4, 7$ ). These source cluster areas match White's [69] Zambezi Region, Kalahari-Highveld Transition Zone and Zanzibar-Inhambane Mosaic. Following Linder *et al.*'s [73] biogeographic classification, these source clusters occur in the Zambezi Region incorporating Mozambique; the Southern African Region incorporating the Kalahari-Namib and the South African Zones; and the Kilimanjaro Zone. The presence of individuals of divergent ancestries in this population corresponds with the history of long-term trade networks that connected southern and eastern inland regions with the East African coast [46,74]. A striking corroboration of genetic and historic evidence of the inland and coastal interactions is the presence of different cluster ancestries among the individuals from Tete-Mocuba in inland Mozambique. Tete, which is located on the banks of the Zambezi River, was the farthest inland trading centre that could be reached by boat from the coast; beyond this point, the Zambezi rapids made it difficult for boat navigation. Tete was an important trading centre for African people that were captured and sold as slaves to Portuguese and Swahili traders [75]. Captives were brought to Tete from places around Lake Malawi, present-day eastern Zimbabwe and the southern highlands of Tanzania, and transported by boat to trading ports such as Quelimane, Cabo Delgado, Ilha do Moçambique, Kilwa and Zanzibar [70].

The assignment of individuals from West Africa, Mascarene Islands, southeast India and Malaysia to the same cluster in the STRUCTURE analysis at  $K = 7$  (figure 4), and to the same clade in the intraspecific phylogenetic tree (figure 6), suggests that West Africa may be the source of introductions to these places. However, this inference needs to be treated with caution since it is based on limited samples. The history of trade between West Africa and other parts of the Indian Ocean also extends over several millennia through a combination of land and oceanic routes and networks across the Sahel to ports on the Mediterranean Sea and Red Sea [76]. From the sixteenth century onwards, European colonial networks linked the Atlantic and Indian Oceans via the Cape of Good Hope [72,76,77].

## 4.3. Inferences on timing of baobab introductions

The presence of private alleles in the Indian clusters may represent introductions further back in time or indicate introductions from baobab populations in northeast Africa. Transoceanic drift could have been a factor. Baobab fruit could have floated on ocean currents to western Indian shores and spread inland



**Figure 7.** Inferred pathways of introduction of baobabs from continental Africa to India and beyond in relation to oceanic trade during different historic periods.

without human assistance. This would have been a slow process resulting in strong genetic divergence, or even speciation. However, the lack of species-level divergence between African and Indian baobabs makes this scenario less probable. It is more likely that the ancient introductions were from northeast Africa. Baobabs occur in Sudanian–Sahelian transition zone which includes present-day Sudan, Ethiopia and Somalia. These are the areas where food crops such as sorghum, pearl and finger millets, cowpea, hyacinth bean and useful trees such as the tamarind originated [32,35,37,46], and arrived in the Indian subcontinent at least 4000 years ago [39–41,44,72]. It is likely that ancient Africans belonging to farming, nomadic and fishing communities of this zone were familiar with the numerous uses of baobabs for food, medicinal, artisanal and symbolic purposes [20,21,23,78]. Given the similar time scales, a more parsimonious explanation for the introduction of baobabs to India is that these people carried the lightweight, nutritious and hardy baobab fruit with them as they travelled overland or on sea journeys across the Indian Ocean.

Figure 7 shows the pathways of introduction of baobabs to India and beyond in relation to oceanic trade during different historical periods. Ancient and early historical introductions are consistent with trade connections between Sudan and Ethiopia in northeast Africa and western India. The expansion of Swahili–Arab trade and cultural networks around the western Indian Ocean between the tenth and seventeenth centuries matches with evidence of MDES and SCEA sources in the Indian clusters [77]. The close position of MDES and CI baobab clusters in the population-level phylogeny may reflect ancestry of the CI populations in the Sudanian–Sahelian transition zone. This corresponds with historical evidence of regular recruitment of soldiers from Ethiopia and Sudan during this period for the armies of Muslim kingdoms in central India [31].

The activities of the Portuguese in the Indian Ocean region between the sixteenth and nineteenth centuries provide insights for the short genetic distance between SCEA and WI clusters. During this period, Portuguese colonists and traders in eastern Africa obtained soldiers and slaves from present-day Malawi, northern Mozambique and southern Tanzania for their territorial enclaves in western India [71,79]. As the enslaved people were marched from these areas to the ports, they would have picked up the fallen baobab fruit and eaten them along the way. Being familiar with the tree, they would have known the nutritive qualities of its mature fruit, its lightness and durability of the pod, and the long-term viability and edibility of the pulp and seeds, and thus carried the fruit with them on their journey to the seaports and beyond [80]. The presence of individuals from multiple clusters in the southeast African coastal populations is most strongly reflective of the Portuguese slave trade in East Africa during this period, when men and women captured from places around Lake Malawi

and inland Mozambique were brought to Tete and then moved to ports such as Quelimane and Ilha do Moçambique, to be shipped to colonial sugarcane plantations in the Mascarene Islands and Brazil [79]. The ambiguous cluster assignments in the MDES cluster may be the result of baobab introductions from various sampled and unsampled populations. This would be consistent with heavy traffic related to slave trading by Swahili–Arab traders during this period. These traders obtained captives from highland and inland regions of southern Tanzania and shipped them from Kilwa, Pemba and Zanzibar to sugarcane plantations in Réunion and Mauritius, and date palm (*Phoenix sylvestris*) plantations in Oman and around the Persian Gulf [70].

The grouping of baobab individuals from West Africa, Mascarene Islands, southeast India and Malaysia in the same genetic clusters is consistent with the history of Dutch, English and French colonial networks between Africa and Asia. During the eighteenth and nineteenth centuries, both the English and Dutch East India Companies recruited ship crews, soldiers and labourers from West Africa for their factories and forts in southeast India, Ceylon and the East Indies [31,81]. In addition, English, Dutch and French botanists travelled between their colonies in West Africa, the Mascarenes, South and Southeast Asia, often transferring plants for commercial production or ornamental purposes [40]. A combination of both kinds of human agency may have introduced West African baobabs in the Mascarenes (Mauritius), eastern India (Madras) and Southeast Asia (Penang).

## 5. Conclusion

By combining genetic analysis of baobabs with historical evidence of oceanic trade, this study provides three important findings that improve current understanding of the history of their dispersal across the Indian Ocean.

First, it shows that there have been multiple instances of dispersal from continental Africa to the Indian subcontinent, possibly even extending back into prehistoric times. Second, we infer that the dispersal of baobabs in the Indian subcontinent and around the Indian Ocean was due to the agency of African migrants. Their long-standing familiarity with the baobab fruit and the tree's ubiquitous presence along inland trade routes would have made it a free and important food for surviving long journeys on land and by sea. The evidence of multiple introductions, and sources from different regions of continental Africa show, without doubt, that the presence of baobabs in the Indian subcontinent, the Mascarenes and Penang signals the forgotten history of African migrants to these places [82].

Finally, our study highlights the need for further genetic analyses of African baobab populations incorporating additional samples from West Africa, Somalia, Ethiopia, Sudan, coastal Yemen, Oman, Iran, Pakistan and Sri Lanka. These analyses can be combined with more detailed historical, cultural and linguistic evidence to establish new hypotheses of baobab dispersal across the Indian Ocean.

**Data accessibility.** The datasets supporting this article have been uploaded as the electronic supplementary material.

**Authors' contributions.** H.R., C.A.K. and D.J.M. initiated the research project and framed the overall research objectives; H.R. and C.A.K. carried out the fieldwork in East Africa; H.R. conducted fieldwork with local assistance in India; K.L.B. carried out laboratory procedures and analysed the data; K.L.B. and H.R. led the write-up of the paper, with contributions from C.A.K. and D.J.M. All authors gave final approval for publication.

**Competing interests.** We have no competing interests.

**Funding.** This research was funded through the Australian Research Council (DP1093100 to H.R., D.J.M. and C.A.K.). Genetic analysis for this study was supported by grant no. VR0165 from the Victorian Life Sciences Computation Initiative on its Peak Computing Facility at the University of Melbourne.

**Acknowledgements.** We thank Jack Pettigrew and Pascal Danthu for sharing samples. We thank Kara Rasmanis for her help in preparing the high-quality diagrams and maps. We are grateful to David Baum, Elizabeth James, Collin Ahrens, Rebecca Jordan and the anonymous reviewers for their comments and suggestions that considerably strengthened the manuscript.

## References

1. Bean AR. 2007 A new system for determining which plant species are indigenous in Australia. *Aust. Syst. Bot.* **20**, 1–43. (doi:10.1071/SB06030)
2. Rashford J. 1987 The search for the baobab tree in Jamaica. *Jamaica J.* **20**, 2–11.
3. Meyer RS, Karol KG, Little DP, Nee MH, Litt A. 2012 Phylogeographic relationships among Asian eggplants and new perspectives on eggplant domestication. *Mol. Phylogenet. Evol.* **63**, 685–701. (doi:10.1016/j.ympev.2012.02.006)
4. Perrier X *et al.* 2011 Multidisciplinary perspectives on banana (*Musa spp.*) domestication. *Proc. Natl Acad. Sci. USA* **108**, 11 311–11 318. (doi:10.1073/pnas.1102001108)
5. Roullier C, Benoît L, McKey DB, Lebot V. 2013 Historical collections reveal patterns of diffusion of

- sweet potato in Oceania obscured by modern plant movements and recombination. *Proc. Natl Acad. Sci. USA* **110**, 2205–2210. (doi:10.1073/pnas.1211049110)
6. Goolsby JA, De Barro PJ, Makinson JR, Pemberton RW, Hartley DM, Frohlich DR. 2006 Matching the origin of an invasive weed for selection of a herbivore haplotype for a biological control programme. *Mol. Ecol.* **15**, 287–297. (doi:10.1111/j.1365-294X.2005.02788.x)
  7. Hawkins JA, Boutaoui N, Cheung KY, van Klippen RD, Hughes CE. 2007 Intercontinental dispersal prior to human translocation revealed in a cryptogenic invasive tree. *New Phytol.* **175**, 575–587. (doi:10.1111/j.1469-8137.2007.02125.x)
  8. Williams DA, Overholt WA, Cuda JP, Hughes CR. 2005 Chloroplast and microsatellite DNA diversities reveal the introduction history of Brazilian peppertree (*Schinus terebinthifolius*) in Florida. *Mol. Ecol.* **14**, 3643–3656. (doi:10.1111/j.1365-294X.2005.02666.x)
  9. Duvall CS. 2006 On the origin of the tree *Spondias mombin* in Africa. *J. Human Geogr.* **32**, 249–266.
  10. Helm R, Crowther A, Shipton C, Tengeza A, Fuller D, Boivin N. 2012 Exploring agriculture, interaction and trade on the eastern African littoral: preliminary results from Kenya. *Azania: Archaeol. Res. Africa* **41**, 39–63. (doi:10.1080/0067270X.2011.647947)
  11. Baum DA. 1995 A systematic revision of *Adansonia* (Bombacaceae). *Ann. Missouri Bot. Garden* **82**, 440–471. (doi:10.2307/2399893)
  12. Pettigrew JD *et al.* 2012 Morphology, ploidy and molecular phylogenetics reveal a new diploid species from Africa in the baobab genus *Adansonia* (Bombacoideae; Malvaceae). *Taxon* **61**, 1240–1250.
  13. Wickens GE, Lowe P. 2008 *The baobabs: pachycauls of Africa, Madagascar, and Australia*. Dordrecht, The Netherlands: Springer.
  14. Baum DA, Small RL, Wendel JF. 1998 Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Syst. Biol.* **47**, 181–207. (doi:10.1080/106351598260879)
  15. Leong Pock Tsy J-M *et al.* 2009 Chloroplast DNA phylogeography suggests a West African centre of origin for the baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). *Mol. Ecol.* **18**, 1707–1715. (doi:10.1111/j.1365-294X.2009.04144.x)
  16. Duvall CS. 2007 Human settlement and baobab distribution in south-western Mali. *J. Biogeogr.* **34**, 1947–1961. (doi:10.1111/j.1365-2699.2007.01751.x)
  17. Blench RM. 2007 The intertwined history of the silk-cotton and baobab. In *Fields of change; progress in African ethnobotany* (ed. RTJ Cappers), pp. 1–20. Gröningen, Germany: Barkhuis and Gröningen University.
  18. De Caluwé E, Halamová K, Van Damme P. 2009 Baobab (*Adansonia digitata* L.) a review of traditional uses, phytochemistry and pharmacology. In *African natural plant products: new discoveries and challenges in chemistry and quality* (eds HR Juliani, JE Simon, C-T Ho), pp. 51–84. Washington, DC: American Chemical Society.
  19. Kamatou GPP, Vermaak I, Viljoen AM. 2011 An updated review of *Adansonia digitata*: a commercially important African tree. *South Afr. J. Bot.* **77**, 908–919. (doi:10.1016/j.sajb.2011.08.010)
  20. Hilu KW, De Wet JM. 1976 Domestication of *Eleusine coracana*. *Econ. Bot.* **30**, 199–208. (doi:10.1007/BF02909728)
  21. De Wet JM, Prasada Rao KE, Brink DE, Mengesha MH. 1984 Systematics and evolution of *Eleusine coracana* (Gramineae). *Amer. J. Bot.* **71**, 550–557. (doi:10.2307/2443330)
  22. Marshall F, Hildebrand E. 2002 Cattle before crops: the beginnings of food production in Africa. *J. World Prehistory* **16**, 99–143. (doi:10.1023/A:1019954903395)
  23. Gebauer J, El-Siddiq K, Ebert G. 2002 Baobab (*Adansonia digitata* L.): a review on a multipurpose tree with promising future in the Sudan. *Gartenbauwissenschaft* **67**, 155–160.
  24. Boivin N, Fuller DQ. 2009 Shell middens, ships and seeds: exploring coastal subsistence, maritime trade and the dispersal of domesticates in and around the Ancient Arabian Peninsula. *J. World Prehistory* **22**, 113–180. (doi:10.1007/s10963-009-9018-2)
  25. Armstrong P. 1979 The history, natural history and distribution of *Adansonia*: a plant genus of the Indian Ocean littoral. In *The Indian Ocean in Focus: Int. Conf. on Indian Ocean Studies, Perth, Australia, Section I Environment and Resources*, pp. 1–21.
  26. Maheshwari JK. 1971 The baobab tree: disjunctive distribution and conservation. *Biol. Conserv.* **4**, 57–60. (doi:10.1016/0006-3207(71)90058-9)
  27. Vandercone R, Sajithran TM, Wijeyamohan S, Santiapillai C. 2004 The status of the baobab (*Adansonia digitata* L.) in Mannar Island, Sri Lanka. *Curr. Sci.* **87**, 1709–1713.
  28. Parsa A. 1959 Medicinal plants and drugs of plant origin in Iran. I. *Plant Food Hum. Nutr.* **5**, 375–394. (doi:10.1007/BF01099755)
  29. Perrier de la Bâthie H. 1932 Les plantes introduites à Madagascar. *Revue de Botanique Appliquée et d'Agriculture Tropicale*. **12**, 296–301. (doi:10.3406/jatba.1932.5126)
  30. Sidibe M, Williams JT. 2002 *Baobab. Adansonia digitata*. Southampton, UK: International Centre for Underutilised Crops.
  31. Burton-Page J. 1969 The problem of the introduction of *Adansonia digitata* into India. In *The domestication and exploitation of plants and animals* (eds PJ Ucko, GW Dimbleby), pp. 331–335. London, UK: Duckworth.
  32. Manning K, Pelling R, Higham T, Schwenniger J-L, Fuller DQ. 2011 4500-year old domesticated pearl millet (*Pennisetum glaucum*) from the Tilemsi Valley, Mali: new insights into an alternative cereal domestication pathway. *J. Archaeol. Sci.* **38**, 312–322. (doi:10.1016/j.jas.2010.09.007)
  33. Weber S. 1998 Out of Africa, the initial impact of millets in South Asia. *Curr. Anthropol.* **39**, 267–274. (doi:10.1086/204725)
  34. Fuller DQ, Harvey E. 2006 The archaeobotany of Indian pulses: identification, processing and evidence for cultivation. *Environ. Archaeol.* **11**, 219–246. (doi:10.1179/174963106x123232)
  35. Hoogervorst T. 2013 If only plants could talk: reconstructing pre-modern biological translocations in the Indian Ocean. In *The sea, identity and history: from the Bay of Bengal to the South China Sea* (eds S Chandra, HP Ray), pp. 67–92. Delhi, India: Society for Indian Ocean Studies, Manohar Press.
  36. Achaya KT. 1994 *Indian food: a historical companion*. Delhi, India: Oxford University Press.
  37. Blench RM. 2003 The movement of cultivated plants between Africa and India in prehistory. In *Food, fuel and fields: progress in African archaeobotany* (eds K Neumann, A Butler, S Kahlheber), pp. 273–292. Cologne, Germany: Heinrich-Barth-Institute.
  38. Fuller D, Boivin N. 2009 Crops, cattle, and commensals across the Indian Ocean. *Etudes Océan Indien* **42–43**, 13–46. (doi:10.4000/oceanindien.698)
  39. Rangan H, Carney J, Denham T. 2012 Environmental history of botanical exchanges in the Indian Ocean world. *Environ. Hist.* **18**, 311–342. (doi:10.3197/096734012X13400389809256)
  40. Burkill HM. 1985–2004 *The useful plants of West Tropical Africa*, vol. 6. London, UK: Royal Botanic Gardens.
  41. Andreakis N, Kooistra WHCF, Procaccini G. 2009 High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Mol. Ecol.* **18**, 212–226. (doi:10.1111/j.1365-294X.2008.04022.x)
  42. Lucek K, Roy D, Bezault E, Sivasundar A, Seehausen O. 2010 Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Mol. Ecol.* **19**, 3995–4011. (doi:10.1111/j.1365-294X.2010.04781.x)
  43. Lesur J, Hildebrand EA, Abawa G, Gurtherx X. 2014 The advent of herding in the Horn of Africa: new data for Ethiopia, Djibouti and Somaliland. *Q. Int.* **343**, 148–158. (doi:10.1016/j.quaint.2013.11.024)
  44. Weiss EA. 1979 Some indigenous plants used domestically by East African fishermen. *Econ. Bot.* **33**, 35–51. (doi:10.1007/BF02858210)
  45. White F. 1983 *The vegetation of Africa*. Paris, France: UNESCO.
  46. Sheriff A. 2005 Slave trade and slave routes of the East African coast. In *Slave routes and oral tradition in southeastern Africa* (eds B Zimba, E Alpers, A Isaacman), pp. 13–38. Maputo, Mozambique: Filson.
  47. Larsen AS, Vaillant A, Verhaegen D, Kjær ED. 2009 Eighteen SSR-primers for tetraploid *Adansonia digitata* and its relatives. *Conserv. Genet. Resour.* **1**, 325–328. (doi:10.1007/s12686-009-9075-y)
  48. James EA, Brown GK, Citroen R, Rossetto M, Porter C. 2011 Development of microsatellite loci in *Triglochin procera* (Juncaginaceae), a polyploid wetland plant. *Conserv. Genet. Resour.* **3**, 103–105. (doi:10.1007/s12686-010-9301-7)
  49. Hardy OJ, Vekemans X. 2002 SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620. (doi:10.1046/j.1471-8286.2002.00305.x)
  50. Cornuet J-M, Luikart G. 1996 Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014.
  51. Pritchard JK, Stephens M, Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

52. Curtin P, Feierman S, Thompson L, Vansina J. 1995 *African history from earliest times to independence*. London, UK: Longmans.
53. Earl DA, vonHoldt BM. 2012 STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361. (doi:10.1007/s12686-011-9548-7)
54. Evanno G, Regnaut S, Goudet J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620. (doi:10.1111/j.1365-294X.2005.02553.x)
55. Jakobsson M, Rosenberg NA. 2007 July 15, 2007 CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806. (doi:10.1093/bioinformatics/btm233)
56. Rosenberg NA. 2004 DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137–138. (doi:10.1046/j.1471-8286.2003.00566.x)
57. Tsai Y-HE. 2011 PhyloGeoViz: a web-based program that visualizes genetic data on maps. *Mol. Ecol. Resour.* **11**, 557–561. (doi:10.1111/j.1755-0998.2010.02964.x)
58. Excoffier L, Smouse PE, Quattro JM. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.
59. Excoffier L, Lischer H. 2009 *Arlequin*. 3.5.1.3 ed. *An integrated software package for population genetics data analysis*. Bern, Switzerland: University of Berne, Swiss Institute of Bioinformatics.
60. Peakall R, Smouse PE. 2006 GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295. (doi:10.1111/j.1471-8286.2005.01155.x)
61. Oksanen J, Kindt R, Legendre P, O'Hara B. 2009 *vegan* Community Ecology Package. See <http://CRAN.R-project.org/package=vegan>.
62. R Development Core Team. 2008 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
63. Miermans PG. 2012 The trouble with isolation by distance. *Mol. Ecol.* **21**, 2839–2846. (doi:10.1111/j.1365-294X.2012.05578.x)
64. Cavalli-Sforza LL, Edwards AWF. 1967 Phylogenetic analysis: models and estimation procedures. *Amer. J. Hum. Genet.* **19**, 233–257. (doi:10.2307/2406616)
65. Felsenstein J. 2009 *PHYLIP: Phylogeny Inference Package*. 3.69 ed. Seattle, WA: Department of Genome Sciences and Department of Biology, University of Washington.
66. Peery MZ *et al.* 2012 Reliability of genetic bottleneck tests for detecting recent population declines. *Mol. Ecol.* **21**, 3403–3418. (doi:10.1111/j.1365-294X.2012.05635.x)
67. Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Sánchez GJ, Doebley J. 2008 Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *Amer. J. Bot.* **95**, 1240–1253. (doi:10.3732/ajb.0800097)
68. Swart ER. 1963 Age of the baobab tree. *Nature* **198**, 708–709. (doi:10.1038/198708b0)
69. Boivin N, Crowther A, Helm R, Fuller DQ. 2013 East Africa and Madagascar in the Indian Ocean world. *J. World Prehistory* **26**, 213–281. (doi:10.1007/s10963-013-9067-4)
70. Chaudhuri KN. 1985 *Trade and civilisation in the Indian Ocean*. Cambridge, UK: Cambridge University Press.
71. Cardoso HC. 2010 The African slave population of Portuguese India. *J. Pidgin Creole Lang.* **25**, 95–119. (doi:10.1075/jpcl.25.1.04car)
72. Alpers E. 2005 Mozambique, 'Mozambiques' slave trade and diaspora on a global scale. In *Slave routes and oral tradition in southeastern Africa* (eds B Zimba, E Alpers, A Isaacman), pp. 39–62. Maputo, Mozambique: Filson.
73. Linder HP, de Klerk HM, Born J, Burgess ND, Fjeldsø J, Rahbek C. 2012 The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *J. Biogeogr.* **39**, 1189–1205. (doi:10.1111/j.1365-2699.2012.02728.x)
74. Beaujard P. 2012 *Les mondes de l'Océan Indien*. Paris, France: Armand Colin.
75. Livingstone D. 1861 *A popular account of missionary travels and researches in South Africa*. London, UK: John Murray.
76. de Silva Jayasuriya S, Pankhurst R. 2003 *The African diaspora in the Indian Ocean*. Trenton, NJ: Africa World Press.
77. van Kessel I. 2007 Belanda Hitam: the Indo-African communities in Java. *Afric. Asian Stud.* **6**, 243–270. (doi:10.1163/156920907X212222)
78. Schumann K, Wittig R, Thiombiano A, Becker U, Hahn K. 2012 Uses, management, and population status of the baobab in eastern Burkina Faso. *Agroforestry Syst.* **85**, 263–278. (doi:10.1007/s10457-012-9499-3)
79. Gustad G, Dhillon SS, Sidibé D. 2004 Local use and cultural and economic value of products from trees in the parklands of the Municipality of Cinzana, Mali. *Econ. Bot.* **58**, 578–587. (doi:10.1663/0013-0001(2004)058[0578:LUACAE]2.0.CO;2)
80. Carney J, Rosomoff RN. 2011 *In the shadow of slavery: Africa's botanical legacy in the Atlantic world*. Berkeley, CA: University of California Press.
81. Harlan JR. 1971 Agricultural origins: centers and noncenters. *Science* **174**, 468–474. (doi:10.1126/science.174.4008.468)
82. Rangan H, Bell KL. 2015 Elusive traces: baobabs and the African diaspora in South Asia. *Environ. Hist.* **21**, 103–133. (doi:10.3197/096734015X14183179969827)