Genetic differentiation and diversity of *Adansonia digitata* L (Baobab) in Malawi using microsatellite markers

C.R.Y Munthali[1,5] P.W Chirwa[2] W.J Changadeya[3] F.K. Akinnifesi[4]

[1]Mzuzu University, Private Bag 1, Luwinga, Mzuzu, Malawi.

[2]Pretoria University, Lynwood Road, Hatfield, Pretoria, 0002, South Africa

[3]University of Malawi, Chancellor College, P.O. Box 280, Zomba, Malawi

[4]World Agroforestry Centre (ICRAF), Southern Africa Regional Programme, Chitedze Agricultural Research Station, P.O. Box 30798, Lilongwe 3, Malawi

[5]Stellenbosch University, Department of Forest & Wood Science, Stellenbosch, South Africa

Abstract

Baobab (*Adansonia digitata* L) belonging to Bombacaceae family, is one of the most widely used indigenous priority tree species in sub-Saharan Africa, valued in the cosmetic industry for its seed oil, and powdery fruit pulp for juice making. Baobab has high potential for domestication in southern Africa, therefore understanding its genetic diversity and population structuring is warranted. The study investigated the level of genetic diversity and differentiation of five populations of *Adansonia digitata* L. sampled from four diverse silvicultural zones in Malawi. Variation at nine microsatellite loci were examined in 150 individual trees. Low mean genetic diversity was expressed through genetic diversity indices: Nei's genetic diversity (h, 0.18±0.03), Shannon Information Index (I, 0.21±0.07), observed number of alleles (na, 1.47±0.10), effective number of alleles (ne, 1.23±0.04) and percentage polymorphic loci (pp, 48%). The low genetic variation found is attributed to the population growing in marginal areas of genetic centre of diversity of the species, anthropogenic factors and founder effects. Moderate genetic differentiation was observed among populations (Gst = 0.13) indicating the presence of a large number of common alleles resulting in a homogenisation effect. Clustering of individual trees by genetic similarity coefficients indicated that mainland trees were genetically closer than the trees on Likoma Island. Mantel's test showed a weak positive insignificant correlation (Z=0.12; P=0.64) between genetic distance among populations and actual distance on the ground implying that geneflow was not directly influenced by isolation by distance. The results suggest that seed distribution and tree improvement should recognise the presence of ecotypes and conservation measures should protect all the populations due to existence of private alleles which are of adaptive importance.

Key word: Baobab, Genetic diversity, race; polymorphism; provenance

Introduction

Studies in genetic diversity within a species are of paramount importance for understanding how a species would respond to environmental change. Current patterns of genetic diversity can provide important clues to the history of the species and its current population structure (Heywood and Watson 1995). In addition, knowledge about population genetics is fundamental for comprehending micro- environmental processes in plant populations that should be utilised in designing management, breeding and conservation strategies (Kyndt et al. 2009). It has been shown that spatial genetic structuring in tree species is influenced by many biological forces such as gene flow through seed and pollen dispersal, tree density, fragmentation, colonization history, isolation into small numbers, differential mortality, and micro-environmental selection (Kyndt et al. 2009). Genetic variation is the starting point for breeding and offers insurances against genetic erosion. Wild trees are genetically structured through natural processes such as mutation, genetic drift, selection, reproductive isolation, and migration (Buiteveld et al. 2007). Plant molecular studies have shown that fragmentation of habitats and small population sizes negatively affect population genetic diversity (Dawson et al. 2009). For instance, the genetically less diverse populations have reduced chance to buffer the effects of poor environmental conditions or competitions (Pluess and Stöcklin 2004). Buiteveld et al. (2007) have reported that forest ecosystem will only persist if genetic diversity of forest trees is dynamically maintained in view of environmental changes. This demands that genetic knowledge should be included in the forest management in order to keep an appropriate level of genetic diversity to guarantee short-term viability and long- term evolutionary potential. In order to manage germplasm resources effectively in fruit tree domestication, one requires knowledge of the amount and distribution of genetic diversity present in natural populations (Mwase et al. 2006). Zobel and Talbert (1984) have reported that geographic races or ecotypes occur most often in species that have a wider natural geographical range and also encompass a large range of environment as is the case with baobab. These variations could be caused by differences in latitude, altitude, rainfall pattern and other environmental conditions that expose trees to large variation in temperature, soil

type, day length and rainfall. Most forest species have evolved into distinct races (ecotypes, provenances) which should be recognised for tree selection in tree breeding as well as seed distribution for forest planting. Furthermore, it is known that although individuals within a race are similar from past heritage or selection pressures, they may also not necessarily be genetically identical (Zobel and Talbert 1984). Thus it is important to understand the pattern of variation existing in populations for use in domestication, conservation, management and tree breeding. *Adansonia digitata* L (Baobab) has a wide geographical range spanning the drier parts of west Africa, east and south Sudan, drier parts of Angola across to Mozambique and up to northern Transvaal (Guy 1971). Such a distribution should result in formation of distinct geographical races (Zobel and Talbert 1984) that are adapted to various ecological conditions. Hence, the need to explore patterns of genetic diversity in relation to baobab distribution (Sidibe and Williams 2002).

Previously, molecular studies have been done in order to assess genetic diversity in baobabs (Assogbadjo et al. 2009; Kyndt et al. 2009; Pocktsy et al. 2009; Larsen et al. 2009). Assogbadjo et al. (2009) showed that there was genetic structuring and low to high genetic diversity between baobab populations in different climatic regions of Benin (West Africa). Kyndt et al. (2009) found high levels of genetic structuring present in baobabs at regional scale (Benin, Ghana, Burkina Faso and Senegal) and within-population level which was unexpected considering its dispersal by bats and human exchange of seed. However, Assogbadjo et al. (2009) using AFLP markers could not distinguish traditionally classified baobab morphotypes. Pocktsy et al. (2009) established that the tetraploid *A. digitata*, or its diploid progenitor originated in West Africa and migrated subsequently throughout the continent, and beyond, through natural and human–mediated terrestrial and overseas dispersal. Larsen et al. (2009) developed and tested eighteen microsatellite primers (SSR-primers) for tetraploid *Adansonia digitata* and its relatives showing different alleles per locus and different allele sizes. Most of the published results on baobab are from West Africa. There is, however, scanty published information on molecular studies for baobabs existing in southern Africa. In spite of the paucity of genetic diversity information on baobab, domestication of some priority indigenous fruit species has been advanced in southern Africa (Akinnifesi et al. 2008). According to Larsen et al. (2009), it is pertinent to carry out gene flow studies in baobabs to provide insight into dispersal processes that shape the genetic structure. In addition, they indicated that estimates of seed dispersal and differentiation between populations is vital for monitoring impacts from human influence and for forecasting

consequences of climate change. Over time, baobab demography has been influenced substantially by anthropogenic factors (land -use pattern, trampling by domesticated livestock), climate (prolonged drought), elephant damage (Edkins et al. 2007; Wilson 1988), fire, clearing during cultivation, browsing (Chirwa et al. 2006) which have had adverse impact on genetic diversity. It is known that positive correlation exists among the levels of genetic diversity and fitness in plants (A´vila-di´az and Oyama 2007). For baobab domestication to succeed, it requires understanding of the genetic diversity since it is the fabric of evolution, the base material on which adaptation depends with high levels of genetic diversity considered as acceptable for conferring the ability to respond to threats such as diseases, parasites, predators and environmental change (Amos and Harwood 1998). The current study was undertaken to assess genetic diversity and differentiation in subpopulations of baobab sampled from 4 silvicultural zones in Malawi. The aim was to examine whether the delineation of silvicultural zones have subsequently structured the genetic composition of the trees substantially and also to estimate the gene flow among the populations using microsatellite markers.

Materials and Methods

Sample collection and DNA extraction

The study populations were selected based on the silvicultural zones (A, Ba, L and J) delineated by Hardcastle (1978) (Table 1, Fig. 1) based on climate and geomorphology. In addition, Likoma Island population, apart from occurring in silviculture zone L, was included due to its geographical isolation. According to Hardcastle (1978), silviculture zone L has mean annual rainfall (MAR) of >1600 mm with predominantly weathered ferralitic soils. Silviculture zone Ba has MAR ranging from 710 to 850 mm and characterized by calcimorphic soils overlaying vertisols. Silviculture zone J has MAR ranging from 1200 to 1600 mm and characterized by ferrallitic soils whilst Silviculture zone A has MAR ranging from 710 to 840 mm with vertisol soils.

A total of 150 individuals, representing five wild populations (Karonga, Likoma Island, Salima, Mwanza and Chikwawa) were sampled (Fig.1). Two populations were collected in silvicultural zone L as stated above because one was an isolated population from an Island. Thirty randomly selected trees at a minimum distance of 100 m represented a population. Four young leaflets were collected, dried in the field and preserved in silica gel in sealed 300

ml plastic bottles. Total genomic DNA was isolated from leaf tissue following a modified CTAB method (Gawal and Jarrent 1991).

INSERT **Table 1** Physical description of populations (site/provenance)

INSERT Fig. 1 Map of Malawi showing location of Chikwawa, Mwanza, Salima, Likoma Island and Karonga sampled *Adansonia digitata* L populations.

Microsatellite analysis

Primer sequences specific for nine microsatellite loci described by Larsen et al. (2009) were used in this study (Table 2). The Polymerase Chain Reaction (PCR) conditions were optimised for the nine microsatellites to produce scorable amplification products. PCR cocktail included the following reaction reagents: 5.7 µl PCR grade water (double distilled water), 1 µl of 10mM DNTP mix, 1.25 µl of 10 x PCR buffer, 1.6 µl of 25 mM Magnesium Chloride ($MgCI_2$), 0.75 µl of both forward and reverse *A. digitata* microsatellite primers, 0.06 µl of 5 u/µl Taq DNA polymerase in storage buffer and 2 µl of 25ng/ µl template DNA. Each reaction tube had final volume of 12.5 µl PCR master mix. PCR programme comprised ten cycles of amplification initialised with a denaturation step at 94 ^{o}C for 30 sec, annealing step at primer specific temperature for 15 sec (Table 2) and extension step at 72 ^{o}C for 30 sec. Another 30 cycles of amplification followed consisting of denaturing step at 89 ^{o}C for 30 sec, annealing step at primer specific temperature for 15 sec, extension step at 72 ^{o}C for 30 sec and final extension at 65 ^{o}C for 20 m. The soaking temperature was 4^{o}C. The PCR products were separated on 6% polyacrylamide gels stained with silver nitrate as described in Promega Silver Sequence™ DNA Sequencing System Technical Manual. The microsatellites bands were assigned base pair scores using pGem DNA marker and X174 DNA/*Hinf* 1 as band size standard markers (Promega, 2000). Sizing of bands was kept consistent though two themocyclers were used by persistently assigning the same band size in base pairs to a particular band right through all gels.

According to Giang et al. (2003) microsatellites as a co-dominant genetic marker enables detection of both homozygotes and heterozygotes that improves the insufficiency of AFLP

and could be used for examining mating system. As few as five or six microsatellite loci can answer many conservation genetic questions (Glaubitz and Moran 2000). Estimating the exact number of copies of individual alleles is difficult among polyploidy species. Therefore data is often analysed as a binary data matrix and SSR markers are treated as dominant markers (Changadeya 2009). Hence, the presence or absence of each PCR amplification product was scored as "1" or "0", respectively and data matrix was generated. Using the data matrix, POPGENE Version 1.31 freeware (Yeh et al. 1999) was used to calculate measures of genetic variation within and between baobab populations. The following variables were calculated: observed number of alleles (na), effective number of alleles (ne) (Kimura and Crow 1964), number of polymorphic loci (p), percentage of polymorphic loci (pp), Nei's genetic diversity (h) (Nei 1973) and Shannon's information index (I) (Lewontin 1974). Significance of the various genetic indices was determined by 95% Confidence Interval. The presence of zero in the interval meant no significant difference. Genetic differentiation (G_{st}), which measures among-population component of genetic variation, was calculated to determine the proportion of total variation that was due to differences between population allele frequencies. Total heterozygosity (H_t), gene diversity of individuals relative to their population (H_s) and gene flow (N_m) were also determined. Dendrogram based on Nei's 1972 using the Unweighted Pair-group Method based on Arithmetic averages (UPGMA) modified from NEIGHBOR procedure of PHYLIP Version 3.5 was constructed. Further, Nei's unbiased measure of pairwise genetic identity and genetic distance (Nei 1978) were calculated. Individual tree similarity was analysed using NTSYSpc version 2.11c (Rolhf 2001). Pairwise similarity matrices were used to construct dendrograms from the Sequential Agglomerative Hierarchical and Nested (SAHN) clustering method using the Unweighted Pair-group Method with Arithmetical averages (UPGMA) (Sneath and Sokal 1973). Mantel's test was done to measure the relationship between genetic distance (Nei 1972 genetic distance) and actual geographical distance.

INSERT **Table 2** Microsatellites used in this study with loci name, forward and reverse sequences and annealing temperature

RESULTS

Allelic diversity

A total of 193 alleles were scored among the five populations. The highest allele number was scored at locus Ad18 (34) and the least number of alleles was scored at Locus Ad05 (9). Variation in allele size ranged from 140 to 318 base pairs (bp). Total number of alleles and allele size (bp) varied among the populations across all the nine loci (Table 3).

INSERT **Table 3** Total number of alleles (A) and allele size range (SR) in base pairs (bp) in five *Adansonia digitata* L at nine loci

Genetic diversity as estimated by; observed mean number of alleles (na), expected mean number of alleles (ne), Nei's (1973) gene diversity (h), Shannon's Information Index (I) was not significantly different among populations (Table 4). The average number of alleles (na) ranged from 1.39±0.10 (in Salima) to 1.59±0.10 (in Karonga) with an average of 1.49±0.10. The effective number of alleles (ne) ranged from 1.20±0.06 (in Salima) to 1.29±0.7 (in Karonga) with an average of 1.23±0.04. The Nei's (1973) gene diversity (h) ranged from 0.12±0.03 (in Salima) to 0.18±0.04 (in Karonga) with an average of 0.18±0.13. The Shannon's Information index (Lewontin 1972) (I) ranged from 0.18 ± 0. 04 (in Salima) to 0.27± 0.05 (in Karonga) with an average of 0.21± 0.07. The number of polymorphic loci (p) ranged from 25 (in Salima) to 38 (in Karonga) with an average of 31 and the percentage polymorphic loci (pp) ranged from 39% (in Salima) to 59% (in Karonga) with an average of 48%.

INSERT **Table 4** Mean genetic diversity values of five *Adansonia digitata* L populations based on nine microsatellite loci

Population differentiation and gene flow

Among population differentiation (G_{st}) and gene flow (N_m) values are presented in Table 5. G_{st} values varied depending on the combination of populations (Table 5). Taking all the populations together showed the highest differentiation of 13%. Including Salima in the northern region populations of Karonga and Likoma Island increased the population differentiation from 8 to 13 %, representing a leap of 6%. The lowest differentiation of 4% was between Chikwawa and Mwanza both occurring in southern region of Malawi. Gene flow (N_m) among the populations ranged from 3.4 to 13.0. While the overall migration rate

among all the populations was low (3.5), Chikwawa and Mwanza had the highest number (13.0) of migrants per generation (Table 5).

INSERT **Table 5** Mean genetic diversity, differentiation and gene flow of all five and subdivided populations

The genetic identity and genetic distance between pairs of populations are shown in Table 6. The genetic distance was smallest (0.0099) between Salima and Mwanza populations where as the largest distance (0.0561) was between Likoma Island and Chikwawa populations. Likewise, the highest genetic identity (0.9902) was between Mwanza and Salima populations whilst the lowest genetic identity (0.9455) was between Chikwawa and Likoma Island populations.

Mantel's test which correlated genetic and geographical distance among populations showed a weak positive insignificant correlation (Z=0.12; *P*=0.64) between genetic distance among populations and actual distances on the ground.

INSERT **Table 6** Nei's genetic identity (above diagonal) and genetic distance (below diagonal)

Population structuring

The UPGMA tree constructed based on the genetic distance (Nei, 1972) is given in Fig. 2. The dendrogram divided the populations into two major clusters namely, Likoma Island cluster (LA) and Karonga (KA), Mwanza (MN), Salima (SA) and Chikwawa (CK) cluster.

INSERT **Fig. 2** Dendrogram for five populations (Karonga (KA), Chikwawa (CK), Salima (SA), Mwanza (MN) and Likoma Island (LA) based on Nei's (1972) Genetic distance method= UPGA—Modified from NEIGHBOR procedure of PHYLIP version 3.5

Genetic relationship in individual trees

The dendrogram showing similarity among the individual trees from the five populations is shown in Fig. 3. The genetic similarity coefficients ranged from 0.66 (for Likoma tree 6) to 1.00 (for Likoma 13 and 14 trees and Chikwawa 14 and 15 trees). The clustering was not clearly based on silvicultural zonation. Broadly, the genetic similarity coefficients predominantly grouped trees from the mainland trees separate from the Likoma Island trees.

INSERT Fig. 3 Genetic relationships among 150 trees in five populations (Karonga, Chikwawa, Mwanza, Salima and Likoma Island) analysed with microsatellites.

Discussion

Allelic diversity

Assessment of genetic variation is important for executing plant domestication, conservation and breeding programmes (Sreekumar and Renuka 2006). The common genetic diversity parameters in assessing genetic richness in baobabs has been percentage of polymorphic loci (pp) and Nei's genetic diversity index (h) (Assogbadjo et al. 2006; Assogbadjo et al. 2009; Kyndt et al. 2009). The proportion of polymorphic loci amplified in this study ranged from 39% to 59% with an average of 48% whereas the Nei's genetic diversity (h) ranged from 0.12±0.03 to 0.18±0.04 with an average of 0.18±0.03. In contrast to these results, relatively higher values of percentage polymorphic loci (pp) and Nei's genetic diversity (h) have been reported in baobabs in West Africa. For instance, Assogbadjo et al. (2006) reported (pp) ranging from 91.2 to 94.9% and (h) ranging from 0.28 to 0.37; Assogbadjo et al. (2009) reported (pp) ranging from 94.1% to 100%, (h) ranging from 0.29 to 0.37 whereas Kyndt et al. (2009) reported (pp) ranging from 41.7% to 96.1%, (h) ranging from 0.22 to 0.35. The high levels of polymorphism suggest high levels of genetic variation in a species (Sreekunar and Renuka 2009). Changadeya (2009) reported that percentage polymorphic loci (pp)

values greater than 50% depict high genetic variation. The average percentage polymorphic loci (pp) and Nei's genetic diversity (h) values in the present study show that the variation in baobab is generally low in Malawi. The low genetic diversity in Malawian populations could be explained by the fact that the populations are growing in the fringes of the centre of diversity for *Adansonia digitata* L. which is in West Africa (Pocktsy et al. 2009). Baobabs' low genetic diversity in Malawi could also be considered to have resulted from founder effect of the original founding population as well as subsequent young populations influenced by human dispersal (Wickens 1982; Pocktsy et al. 2009). Genetic drift which lowers genetic diversity is common in small tree populations (Yea 2000) which is suspected in founding baobab populations in Malawi.

The results on genetic variation within the population (Table 4) though not significantly different showed consistently that Karonga had the highest diversity in all genetic indices followed by Likoma Island and Chikwawa populations whilst Salima followed closely by Mwanza had the least diversity. The difference in genetic diversity could be attributed to differences in mutation rates within populations occurring in varying geographical areas (Changadeya et al. 2012). In addition, different anthropogenic factors and natural phenomenon such as death due to drought, fungal diseases and destruction by elephants have shaped genetic diversity within the populations (Guy (1971; Wilson 1988; Edkins et al. 2007). Furthermore, differences in time of tree flowering within and between populations happening in Malawi may reduce the effective number of trees in baobab populations differently; setting in varying degrees of random drift impacting on genetic diversity (Dawson et al. 2009). Mwase et al. (2006) indicated that distribution range and population size have strong relationship within population genetic variation in tropical tree species with restricted populations displaying significantly less variation than those with wide distribution. In this study, distribution range and population size seem not to correspond to genetic diversity levels. For instance, baobab densities and distribution are higher in Salima and Mwanza (see Chirwa et al. 2006) and yet these populations have shown low genetic variation. The low polymorphism such as the one found in Salima and Mwanza might be a sign of genetic erosion and populations could be considered endangered since large pool of genetic diversity is prerequisite for a population to survive environmental pressures (Yea, 2000).

When populations were pooled, mean heterozygosity values (H_t and H_s) (Table 5) showed that genotypes for Karonga and Likoma Island populations were the most diverse.

However, the heterozygosity decreased when Salima from central region was combined with Karonga and Likoma Island populations The results further showed least heterozygosity in Chikwawa and Mwanza as a group both occurring in southern Malawi. From the findings, it may be inferred that the hotspot for genetic variation in Malawi reside in the populations in the north in silviculture zone L. These populations should be given special prominence in *insitu* conservation. However, presence of rare alleles in all populations implies that conservation should essentially cover all the populations.

Population structuring and genetic differentiation

The UPGMA tree constructed based on the genetic distance (Nei 1972) grouped populations into two major clusters with an Island population clustered separately from the mainland populations (Fig.2). On the mainland, there were also minor clusters with Karonga and Chikwawa belonging to separate groups whilst Mwanza and Salima were in the same group. The clustering did not correspond to silvicultural zonation (Table 1) of Malawi (Hardcastle 1978). For instance, Karonga and Likoma Island belonging to the same silvicultural zone L are clustered separately. Furthermore, Salima and Mwanza belonging to different silvicultural zones of Ba and J respectively are clustered together. This is supported by, Mantel's test which showed no direct relationship between genetic distance among populations and actual geographical distance. This implies that close populations in the country are not necessarily genetically similar. Similarly, populations occurring in the same or different silvicultural zone are not necessarily genetically similar or divergent. This mirrors the findings in Fig 2 where Chikwawa and Mwanza which are geographically the closest (56 km) were clustered separately whilst Salima and Mwanza 204 km apart were clustered together. Similar trends in population structure are depicted in the values of genetic identity and distance (Table 5). The clustering may be considered as ecotypes or races of baobabs existing in Malawi. Ecotypes are generally taken as populations adaptable to specific habitats which may have an implication in domestication, seed distribution, conservation and tree breeding strategies. The Gst values showed that the sampled populations were differentiated by 13% (which comes from alleles private to each population) meaning that 87% were common among the populations. High similarity of alleles may deduce sharing of common ancestral alleles and the low frequency of alleles generated subsequent to evolution (Esselman et al. 2000). The preponderance of common alleles could be due to water dispersal (Wilkens 1982) in the Great African Rift Valley where all the sampled populations exist.

The genetic similarity coefficients (Fig.3) showed that individual trees from the five populations were 66% to 100 % genetically similar. Generally, the results depicted that mainland trees are genetically closer than the Island trees. The contributing factor may be the differences in origin of the founding populations in different silvicultural zones. Mixture of individuals from different populations in the same cluster may mean that those individuals came from the same origin (refugia) (A´vila-di´az and Oyama 2007) or as a result of multiple introductions from several refugia. The other possible explanation may be that even though populations occur in different silvicultural zones, they may not have undergone significant differentiation (Maghuly et al. 2006). Human influence of seed dispersal may have also likely mixed the genotypes in baobab populations in the country due to long history of slave trade along the water systems of the Great African Rift Valley. Arabs might have moved baobab seed from one locality to another (Wickens 1982). The results mirror those reported by Assogbadjo et al. (2006) and Pocktsy et al. (2009). Assogbadjo et al. (2006) found that some genotypes of *A. digiata* L. populations growing in different climatic zones of Benin belonged to more than one gene pool. Pocktsy et al. (2009) also found that some *A. dig*itata L. haplotypes were found belonging to more than one region which suggested natural colonization and/or human introductions to the areas involved.

The G-statistics (G_{st}) values ranging from 0.050 to 0.150 indicate moderate genetic differentiation; 0.151 to 0.250 is representative of large gene differentiation; and above 0.250 represents very large gene differentiation (Yeh 2000). In this study, genetic differentiation of 13% amongst the five populations and that of 6% for three populations of Karonga, Likoma Island and Salima could be considered moderate; whereas genetic differentiation of 4% for populations from Chikwawa and Mwanza was considered low. When compared to baobab results in West Africa: Kyndt et. al. (2009) found differentiation among populations ranging from 2% to 28% whilst Assogbadjo et al. (2006) found 17.63%. The genetic differentiation found in baobab in West Africa was largely attributed to isolation –by- distance pattern and restricted geneflow with some human influence (Kyndt et al. 2009). Kelly et al. (2004) reported that extensive gene flow leads to low population differentiation. This may be the case with baobab populations in Malawi considering that high gene flow revealed Nm>1 (Nm = 3.4 to 13). This could be as a result of bat pollination and water dispersal (Wilkens, 1982) along Lake Malawi and Shire River (Fig.1). However, physical isolation may as well account for the differentiation despite the species having long geneflow (Assogbadjo et al.

2006). For instance, Karonga is genetically separated (Table 6, Fig.2) from other mainland populations due to mountain barriers, Likoma Island due to the massive water body and Chikwawa due to Shire Highlands and Escarpments. However, Mantel's test result suggests that apart from long geneflow there may be other underlying factors differentiating populations thriving in Malawi. Changadeya (2009) interpreted the low population differentiation in bananas as being a result of the existence of many common alleles shared among populations. Sharing of 87% common alleles among baobab populations may be the cause of low differentiation observed. In addition, gene flow events in tetraploid species like baobab involve the movement of twice the number of genes transported than in a diploid species leading to less differentiation among the populations (Nassar et al. 2003). This may have further homogenised the populations in Malawi. The presence of 13% private alleles means that the populations might not be taken as panmictic. Thus, domestication, conservation and tree breeding strategies should recognise this in future.

Conclusion

At national level, the results have shown that Karonga, Likoma Island and Chikwawa are the genetic diversity hotspots in Malawi. Conservation should therefore prioritise these populations. Compared to populations in West Africa, all Malawian populations have shown indications of genetic erosion. For management purposes, it is important that the genetic diversity in all populations do not decrease further. Impoverished level of genetic diversity in Salima and Mwanza should be a matter of great concern since baobab is an important Agroforestry species in the areas. Low genetic diversity is not healthy for outbreeding species facing environmental and climatic change. It is hypothesized that human factors have a major role on the genetic erosion being experienced in baobab populations. This is an area for further study especially to establish how human demographics have impacted on baobab diversity. Examining genetic diversity in adult and young trees would reveal impact of anthropogenic factors. The impact of time of flowering phenology on genetic diversity should also be investigated since mating system is critical in tree genetic structure. The results have shown moderate genetic differentiation among the populations. Moderate genetic structuring among populations, implies that many alleles are common among the populations with few rare alleles present due to environmental adaptation. Thus the populations should not be taken as panmictic (single interbreeding unit). Presently, it is unknown how the genotypes would respond once seed is distributed to a unique habitat.

Provenance and family trials are required. The long distance gene flow found in this study may imply that spatial genetic differentiation might be low within the population. Therefore, it will be rational to collect seed for domestication and tree breeding from as many mother trees as possible to achieve broad genetic base. The practice in seed collection in wild plants of taking trees spaced at 50-100 m as genetically unrelated needs review with the extensive gene flow happening in baobab. The UPGMA tree constructed based on the genetic distance (Nei 1972) grouped populations into two major clusters which may be inferred as ecotypes of baobab. The present grouping contrasts the Hardcastle's (1978) silvicultural zonation implying the criterion for tree seed distribution for forestry operations should be reviewed in Malawi. The clustering of individual trees through genetic similarity coefficients has generally shown that mainland trees are genetically closer than trees from the Island. Specific clustering has shown that genotypes might belong to several gene pools due to natural distribution, anthropogenic influence and water dispersal. Present results are important in tree domestication, management and improvement strategies at national level. However, baobab is widely distributed in southern Africa; it is therefore recommended to assess the genetic diversity at regional scale. The information gathered will guide in domestication, conservation, breeding strategies and general management (seed collection, seed distribution and afforestation activities) at national and regional level.

Acknowledgement

The study was co-financed by Mzuzu University and Stellenbosch University. Special thanks go to Mr Y Kazembe and Miss E Aquaron of Molecular Biology and Ecology Research Unit (MBERU), Department of Biological Sciences, Chancellor College, University of Malawi for helping during the extraction of DNA. Mr S Kamowa of Mzuzu University is thanked for active participation in field work. Lastly, we thank Pretoria University for hosting the first author at the time of writing the article.

Reference

A´vila-di´az I, Oyama K (2007) Conservation Genetics of an endemic and endangered epiphytic *Laelia speciosa* (Orchidaceae). American Journal of Botany 94 (2): 184-193

Akinnifesi FK, Silesh G, Ajayi OC, Chirwa PW, Kwesiga FR, Harawa R (2008) Contribution of Agroforestry Research and development to Livelihood of Smallholder Farmers in Southern Africa: 2 Fruit, Medicinal, Fuelwood and Fodder Tree Systems. Agricultural Journal 3 (1): 76-88

Amos W, Harwood J (1998) Factors affecting levels of genetic diversity in natural populations. Philosophical. Transactions: Biological Sciences. Vol., 353 No. 1366:177-186

Assogbadjo AE, Kyndt T, Chadare FJ, Sinsin B, Gheysen G, Eyog-Matig O, Van Damme P (2009) Genetic fingerprinting using AFLP cannot distinguish traditionally classified baobab morphotypes. Agroforestry Systems 75: 157-165

Assogbadjo AE, Kyndt T, Sinsin B, Gheysen G, Van Damme P (2006) Patterns of Genetic and Morphometric Diversity in Baobab (*Adansonia digitata*) Populations Across Different Climatic Zones of Benin (West Africa). Annals of Botany 97: 819-830

Buiteveld J, Vendramin GG, Leonardi S, Kamer K, Geburek T (2007) Genetic diversity and differentiation in European beech (Fagus sylvatica L.) stands varying in management history. Forest Ecology and Management 247: 98-106

Changadeya W, Kaunda E, Ambali AJD (2012). Molecular characterisation of Musa L. cultivars cultivated in Malawi using microsatellite markers. African Journal of Biotechnology 11(18): 4140-4157

Changadeya W (2009) Genetic Morphological and Socio-Economic Characteristion of Musa (Banana) Cultivars Growing in Malawi. Dissertation. University of Malawi

Chirwa M, Chithila V, Kayambazinthu D, Dohse C (2006). Distribution and Population Structures of *Adansonia digitata* in Malawi. FRIM, Zomba, Malawi

Chung MI, Nason J, Chung MG, Kim K, Park C, Sun B, Pak J (2002) Landscape-level spatial genetic structure in *Quercus acutissima* (Fagaceae). American Journal of Botany 89(8): 1229-1236

Dawson IK, Lengkeek A, Weber JC, Jamnaclass R (2009) Managing genetic variation in tropical trees: Linking knowledge with action in agroforestry ecosystems for improved conservation and enhanced livelihoods. Biodiversity Conservation 18: 969-986

Edkins MT, Kruger LM, Harris K, Midgley JJ (2007) Baobabs and elephants in Kruger National Park: nowhere to hide. Blackwell Publishing Ltd. African. Journal of. Ecology. Doi:10.1111/j.1365-2028.2007.00798.x

Esselman E J, Crawford DJ, Soren B, Stuessy TF, Anderson GJ, Silva MO (2000) RAPD Marker Diversity within and divergence among species of *Dendroseris* (Asteraceae: Lactuceae). American Journal of Botany 87(4): 591-596

Gawal NG, Jarret RL (1991) Cytoplasmic genetic diversity in banana and plaintain. Euphytica, 52:19-23

Giang LH, Hong PN, Tuan MS, Harada K (2003) Genetic variation of *Avicennia marina* (Forsk) Vierh. (Avicenniaceae) in Vietnam revealed by microsatellite and AFLP markers. Genes Genet. Syst. 78: 399-407

Glaubitz JF, and Moran GF (2000) Genetic Tools: The Use of Biochemical and Molecular Markers. In: Young A, Boshier, D and Boyle T (eds.), Forest Conservation Genetics. Principles and Practice. CSIRO Publishing, United Kingdom

Guy GL (1971) The Baobabs: Adansonia Spp. (Bombacaceae). Journal of the Botanical Society of Southern. Africa 57: 31-37

Hardcastle PD (1978) A preliminary silvicultural classification of Malawi. Forestry Research Institute of Malawi. Forestry Research Record No. 57

Heywood VH, Watson RT (1995) Global Biodiversity Assessment. Cambridge University Press, United Kingdom

Kelly BA, Hardy OJ, Bouvet J (2004) Temporal and Spatial genetic structure in *Vitellaria paradoxa* (shea tree) in an agroforestry system in southern Mali. Molecular Ecology 13:1231-1240

Kyndt T, Assogbadjo AE, Hardy, OJ, Kakaïa RG, Sinsin, B, Van Damme P, Gheysen G (2009) Spatial genetic structuring of baobab (*Adansonia digitata*, Malvaceae) in the traditional Agroforestry systems of West Africa. American Journal of Botany 96(5):950-957

Larsen AS, Vaillant A, Verhaegen D, Kjaer ED (2009) Eighteen SSR-primers for tetraploid *Adansonia digitata* and its relatives. Conservation Genet Resour. DOI 10.1007/s12686-009-9075-y

Maghuly F, Pinsker W, Praznik W, Fluch S (2006) Genetic diversity in managed subpopulations of Norway spruce [*Picea abies* (L) Karst]. Forest Ecology and Management 222: 266-271

Mwase WF, Bjornstad A, Stedje B, Bokosi JM, Kwapata MB (2006) Genetic diversity of *Uapaka kirkiana* Muel. **A**rg. Populations as revealed by amplified fragment length polymorphism (AFLPs). African Journal of Biotechnology Vol. 5 (13): 1205-1213

Nassar JM, Hamrick JL, Fleming TH (2003) Population Genetic Structure of Venezuelan Chripterophilous Columnar Cacti (Cactaceae). American Journal of Botany 90 (11): 1628-1637

Pluess AR and Stöcklin J (2004) Population genetic diversity of the clonal plant *Geum reptans* (Rosaceae) in the Swiss Alps. American Journal of Botany 91(12): 2013-2021

Pocktsy JL, Lumaret R, Mayne D, Mohamed VA, Abutaba YIM, Sagna M, Raoseta S, Danthu P (2009) Chloroplast DNA phylogeography suggest a West African centre of origin for the baobab, *Adansonia digitata* L (Bombacoideae, Malvaceae). Molecular Ecology 18: 1707-1715

Rholf JR (2001) NTSYpc Version 2.11c Numerical Taxonomy and Multivariate Analysis System, Exeter Software, New York

Sidibe M, and Williams JT (2002) Baobab. *Adansonia digitata*. International Centre for Underutilised Crops, Southampton, UK

Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco

Sreekumar VB, Renuka C (2006) Assessment of genetic diversity in *Calamus thwaitesii* BECC. (Arecaceae) using RAPD markers. Biochemical Systematics and Ecology 34, 397-405

Wickens GE (1982) The Baobab: Africa's Upside-down Tree. Kew Bulletin, Vol. 37, No. 2: 173-209

Wilson RT (1988) Vital Statistics of the baobab (*Adansonia digitata*). African. Journal of. Ecology. Volume 26: 197-206

Yeh FC (2000) Population genetics. In: Young A, Boshier D and Boyle T (eds.), Forest Conservation Genetics. Principles and Practice. CSIRO Publishing, United Kingdom

Zobel B, Talbert J (1984) Applied Forest Tree Improvement. John Wily and Sons, Inc. New York

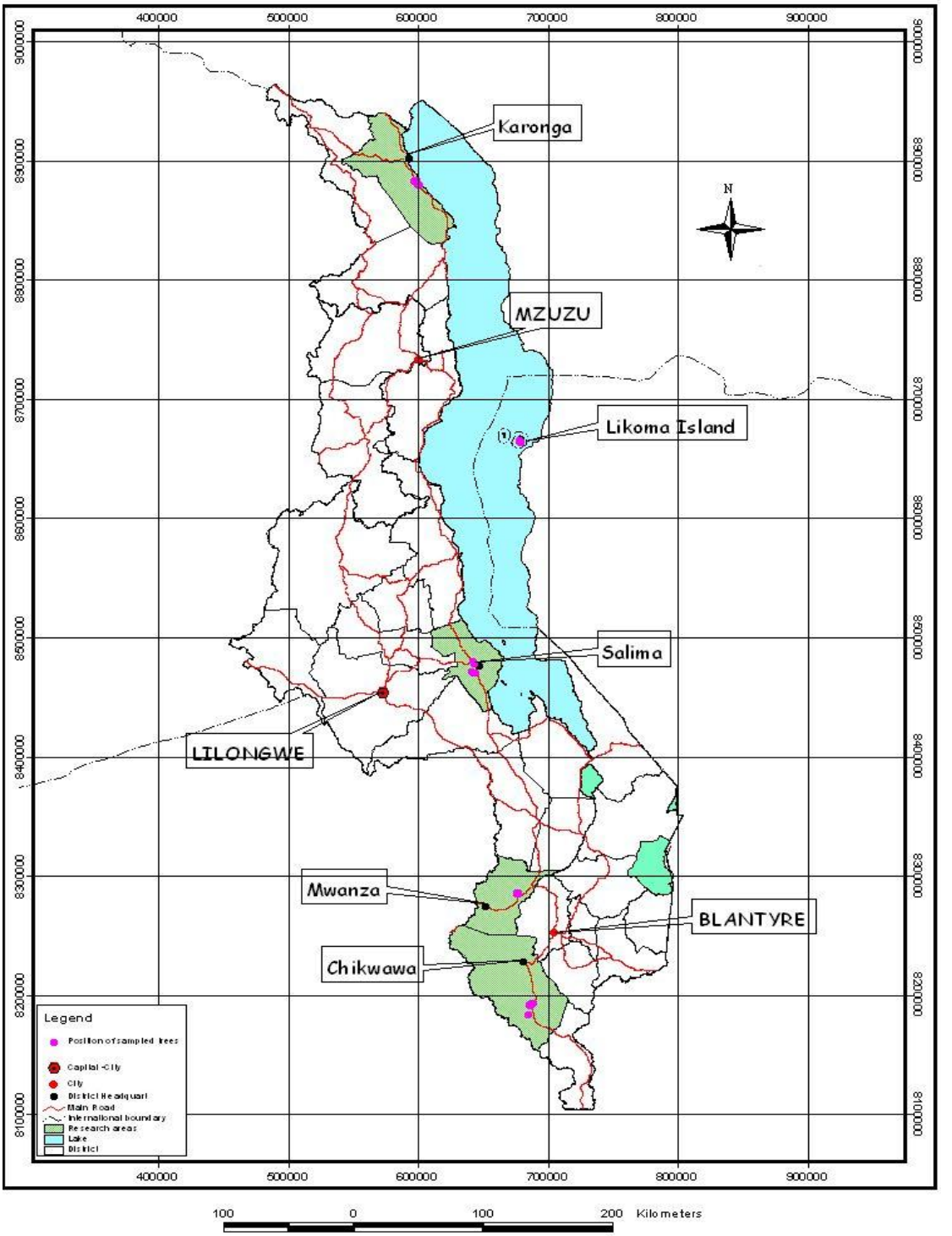


Fig. 1 Map of Malawi showing location of Chikwawa, Mwanza, Salima, Likoma Island and Karonga sampled *Adansonia digitata* L populations.

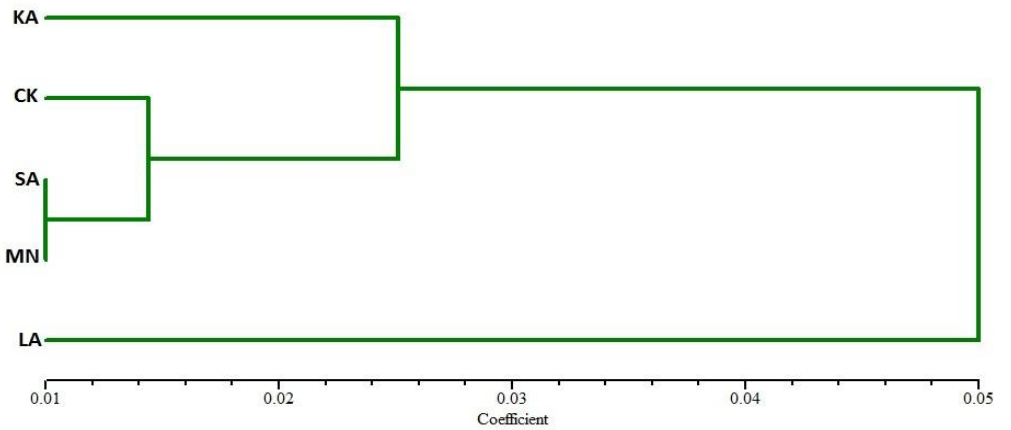
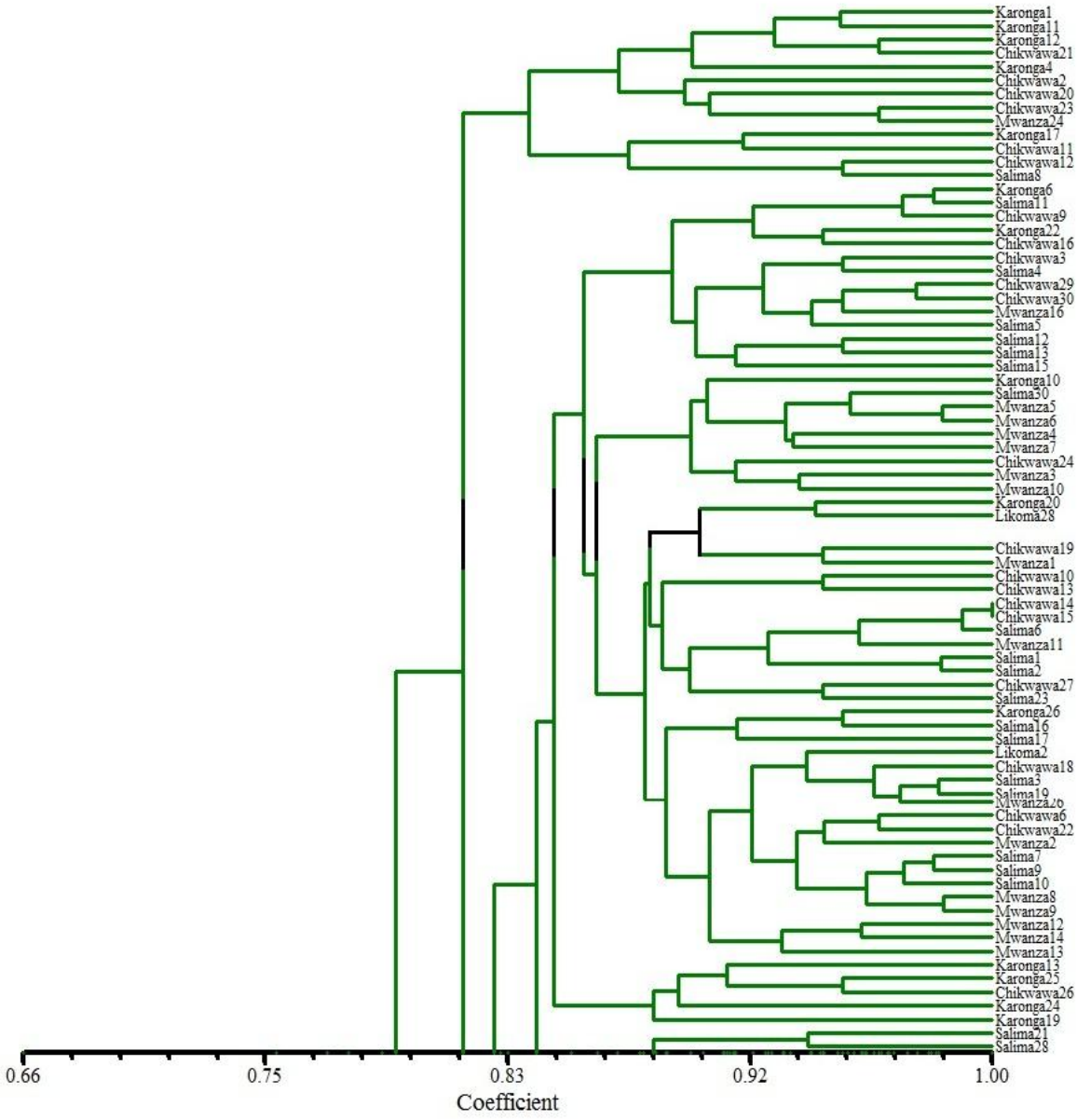


Fig. 2 Dendrogram for five populations (Karonga (KA), Chikwawa (CK), Salima (SA) , Mwanza (MN) and Likoma (LA) based on Nei's (1972) Genetic distance method= UPGA—Modified from NEIGHBOR procedure of PHYLIP version 3.5


Karonga1
Karonga11
Karonga12
Chikwawa21
Karonga4
Chikwawa2
Chikwawa20
Chikwawa23
Mwanza24
Karonga17
Chikwawa11
Chikwawa12
Salima8
Karonga6
Salima11
Chikwawa9
Karonga22
Chikwawa16
Chikwawa3
Salima4
Chikwawa29
Chikwawa30
Mwanza16
Salima5
Salima12
Salima13
Salima15
Karonga10
Salima30
Mwanza5
Mwanza6
Mwanza4
Mwanza7
Chikwawa24
Mwanza3
Mwanza10
Karonga20
Likoma28
Chikwawa19
Mwanza1
Chikwawa10
Chikwawa13
Chikwawa14
Chikwawa15
Salima6
Mwanza11
Salima1
Salima2
Chikwawa27
Salima23
Karonga26
Salima16
Salima17
Likoma2
Chikwawa18
Salima3
Salima19
Mwanza26
Chikwawa6
Chikwawa22
Mwanza2
Salima7
Salima9
Salima10
Mwanza8
Mwanza9
Mwanza12
Mwanza14
Mwanza13
Karonga13
Karonga25
Chikwawa26
Karonga24
Karonga19
Salima21
Salima28
0.66
0.75
0.83
0.92
1.00
Coefficient

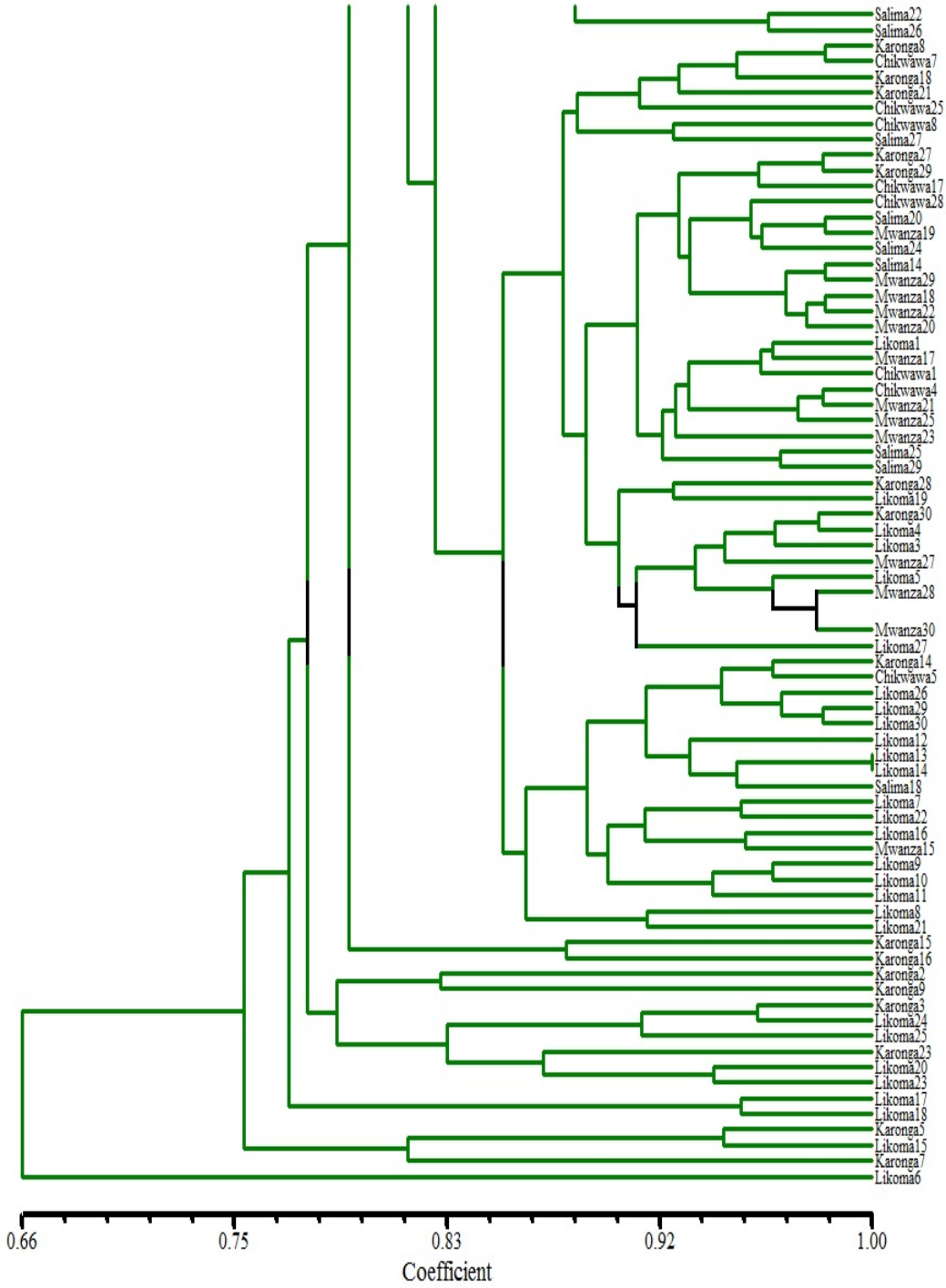


Fig. 3 Genetic relationships among 150 trees in five populations (Karonga, Chikwawa, Mwanza, Salima and Likoma Island) analysed with microsatellites.

Table 1 Physical description of populations (site/provenance)

Population	Silviculture Zone	Average annual Stress Period (weeks)	Altitude (m.a.s.l)	Mean nnual rainfall (mm)	Mean annual temperature (^{0}C)	Soil
Karonga (pop1)	L	7" Field capacity: 15*, 12"Field capacity": 10	475-1000	>1600	23-25	Ferrisols dominant regosols
Likoma Island (pop2)	L	7" Field capacity: 15 12" Field capacity": 10	475-1000	>1600	23-25	Ferrisols, alluvial calcimorphic, regosols, lithosols
Chikwawa (pop3)	A	7" Field capacity:35 12" Field capacity: 35	<200	710-840	>25	Vertisols
Salima (pop4)	Ba	7"Field capacity:28 12" Field capacity: 26	200-1200	710-850	20-25	Alluvial calcimorphic soils above the vertisols
Mwanza (pop5)	J	7" Field capacity:16 12" Field capacity: 11	900-1500	1200-1600	19-21	Sandy ferrallitic

Source: Hardcastle (1978)

* = Information on average annual stress calculations are reported by Hardcastle (1978)

Table 2 Microsatellites used in the study with loci name, forward and reverse sequences and annealing temperature

Locus name	Primer sequences 5'-3'		Annealing* temperature (^{o}C)
	F	R	
Ad01	CATTGCCAGGAATGCTTTTGC	GGATTGCCAGGTCTACTAC	55
Ad03	GGATCAAATTATGGTTAAGGC	CCAATTTTGAGCCAATTCTCA	50
Ad04	GTTGCTTGTGTGCTTACCC	CATCCCTCTCCCCATTCC	55
Ad05	CTCAACAAGGTTCGGATGTCGTATG	GTCTGCCGGGTGTTTTGCATG	58
Ad07	TAGAAAATTAGCAGATAAGTGC	GATTTCGGTGATATGTTGTAG	41
Ad08	TCTAAAGCCTGTAAGGAAAAATGGG	TTCTCCGTTCACTCTGTACTTCC	54.5
Ad09	TACCACTTCTCCAGATGCTAC	ACTGGCTAGAGATGCGTTG	53.7
Ad14	CTTGATTGGAATACGGGAAATGGAG	CCAAACCAATTGGACTTTGACCTTC	56
Ad18	ACCGCTTCCGTTCTCATTCC	ACCACCACTACACCGTCATTG	56.5

*Annealing temperature after optimisation

Source of primers: Larsen et al (2009)

Table 3 Total number of alleles (A) and allele size range (SR) in base pairs (bp) in five *Adansonia digitata* L at nine loci.

*Population	Ad01		Ad03		Ad04		Ad05		Ad07		Ad08		Ad09		Ad14		Ad18	
	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR
CK	4	182-198	14	140-206	17	150-250	7	306-322	12	182-256	12	208-314	13	186-256	9	194-222	23	142-316
KA	9	180-200	19	140-252	12	150-248	2	308-312	7	184-200	5	142-258	10	188-252	10	186-210	19	140-318
LA	7	182-200	13	140-206	13	150-214	4	314-320	17	150-250	14	140-204	12	186-250	7	196-222	10	142-318
MN	4	186-200	12	140-166	14	150-310	4	306-312	10	182-204	12	140-312	8	188-208	7	190-206	9	200-262
SA	5	182-198	12	140-202	11	150-214	3	310-314	13	113-252	8	242-314	11	192-258	9	190-214	6	200-262
All	11	180-200	23	140-252	24	150-310	9	306-322	25	113-256	29	140-318	22	186-258	16	186-222	34	140-318

*The population names have been code shortened as follows: CK= Chikwawa, KA=Karonga, LA=Likoma Island, MN=Mwanza, SA=Salima

Table 4 Mean genetic diversity values of five *Adansonia digitata* L populations based on nine microsatellite loci

Population	N	Na	Ne	H	I	p	pp
Chikwawa	26	1.50±0.10 ns[a]	1.22± 0.07 ns	0.13± 0.04 ns	0.21±0.05 ns	32	50
Karonga	25	1.59±0.10 ns	1.29±0.07ns	0.18±0.04 ns	0.27±0.05 ns	38	59
Likoma	26	1.53±0.10 ns	1.23±0.06 ns	0.14±0.04 ns	0.22±0.05 ns	33	52
Mwanza	27	1.42±0.10 ns	1.21±0.06 ns	0.12±0.03 ns	0.19±0.05 ns	27	42
Salima	26	1.39±0.10 ns	1.20±0.06 ns	0.12±0.03 ns	0.18±0.05 ns	25	39
Mean	26	1.49±0.10 ns	1.23±0.04 ns	0.18± 0.03 ns	0.21±0.07 ns	31	48

Mean values are followed by standard error

n = number of samples, na = observed number of alleles, ne = Effective number of alleles (Kimura and Crow, (1964), h = Nei's (1973) gene diversity, I = Shannon's Information index (Lewontin 1972). P = number of polymorphic loci, pp = percentage of polymorphic loci;a means within the columns followed by ns were not significantly different at P=0.05 based on 95% Confidence Interval.

Table 5 Mean genetic diversity, differentiation and gene flow of all five and subdivided populations

Population grouping	n	H_t	H_s	G_{st}	N_m
All five populations	130	0.158± 0.03	0.138± 0.02	0.13	3.5
Karonga & Likoma	51	0.173± 0.03	0.159± 0.03	0.08	5.7
Karonga, Likoma & Salima	78	0.166 ± 0.03	0.145± 0.02	0.13	3.4
Chikwawa & Mwanza	58	0.133±0.03	0.128± 0.03	0.04	13.0

Mean values are followed by standard error; n= number of individuals; Gst = Genetic differentiation; Nm = gene migration; Ht = gene diversity over all groups; Hs= gene diversity of individual relative to their population.

Table 6 Nei's genetic identity (above diagonal) and genetic distance (below diagonal)

Population	Karonga	Likoma	Chikwawa	Salima	Mwanza
Karonga	-	0.9669	0.9785	0.9739	0.9775
Likoma	0.0337	-	0.9455	0.9482	0.9592
Chikwawa	0.0217	0.0561	-	0.9838	0.9887
Salima	0.0264	0.0532	0.0164	-	0.9902
Mwanza	0.0227	0.0417	0.0114	0.0099	-