Full Length Research Paper

Genetic diversity of *Uapaca kirkiana* Muel. Årg. populations as revealed by amplified fragment length polymorphisms (AFLPs)

Weston F. Mwase¹,²*, Å. Bjørnstad², B. Stedje³, J.M. Bokosi¹ and M.B. Kwapata¹

¹University of Malawi, Bunda College of Agriculture, Forestry and Horticulture Department, P.O. Box 219, Lilongwe, Malawi.
²Norwegian University of Life Sciences, Plant and Environmental Sciences Department, P.O. Box 5003, Ås, Norway.
³Natural History Museum Botanical Garden, University of Oslo, P.O. Box 1172 Blindern NO-0318 Oslo, Norway.

Accepted 18 May, 2006

*Corresponding author. E-mail: westmwase@yahoo.co.uk.

*Uapaca kirkiana* is a priority fruit tree species for domestication in miombo woodlands of Southern Africa. Natural populations of *U. kirkiana* are declining throughout the woodlands due to deforestation, forest fragmentation and wildfires. Knowledge of population structure and genetic diversity is prerequisite for development of conservation strategies. Amplified fragment length polymorphisms (AFLP) were used to assess the genetic diversity of eight populations from three geographical regions of Malawi. AFLP markers revealed moderate differentiation (GST=0.079) among the populations collected from the three regions, however, there were no significant genetic variations among the regional collections. Analyses of molecular variance (AMOVA) found very high variation (92%) among individuals within populations and 6.8% among populations. The variations between populations indicate that populations can not be considered a single panmictic unit. Analyses of genetic similarity based on unweighted pair group method of arithmetic averages (UPGMA) suggested that the 8 populations fall into three clusters with 5 populations in one cluster, two in another and the lake shore population of Chesamu in its own cluster. Based on results presented it would be cost effective to sample a small number of populations represented by a large number of individuals for germplasm conservation purposes. In view of the wide distribution of *U. kirkiana* in the miombo woodlands in Southern Africa there is need for a more intensive genetic study to include populations growing in different countries to produce a wider picture of levels of distribution of genetic diversity of the species.

**Key words:** AFLP, conservation, dioecious, domestication, genetic diversity, miombo, *Uapaca kirkiana*, UPGMA.

INTRODUCTION

*Uapaca kirkiana* Muel. Årg is a dioecious plant that belongs to family Euphorbiaceae. The genus *Uapaca* consists of around 60 species of which 49 are restricted to tropical Africa whilst the rest occur only in Madagascar (Radcliffe-Smith, 1988). The genus is stable, apparently devoid of polyploidy, with a chromosome number of 2n=26. The range of *U. kirkiana* includes Angola, Burundi, southern Democratic Republic of Congo, Malawi, Mozambique, Tanzania, Zambia and Zimbabwe (Ngulube et al., 1995). Ethnobotanical studies within the miombo ecological zone in Malawi (Maghembe and Seyani, 1992; Maghembe et al., 1994), Tanzania (Karachi et al., 1991) and Zambia (Kwesiga and Chisumpa, 1992) identified *U. kirkiana* as a priority indigenous fruit tree species for conservation and domestication among 50 species of indigenous trees bearing edible fruits. Fruits of *U. kirkiana* have high nutritional value and they can be eaten raw, made into jam and sweetmeats, or used to produce a refreshing
Molecular genetic markers provide a relatively unbiased method of quantifying genetic diversity. Molecular genetics provide fast and effective tools to biologists and geneticists to reveal diversity and variability among plants, animals and microorganisms at DNA level. Information on molecular diversity of *U. kirkiana* is limited and data on genetic variation at the DNA level of Malawian populations is lacking. Several molecular markers are available for studying genetic diversity in plants. Amplified Fragment Length Polymorphisms (AFLPs) (Vos et al., 1995) based on the polymerase chain reaction (PCR) are spread all over the genomes and are hypervariable. Use of AFLPs is advantageous in generation of large number of markers spanning the whole genome without prior knowledge of sequence of the genome. AFLPs, however have disadvantage of being less informative because of their dominance and bi-allelic nature (Heckenbecker et al., 2003). AFLPs allow relatively quick marker development, which is often important in the assessment and conservation of genetic diversity (Marghali et al., 2005). Due to their higher multiplex ratio, discriminatory power and reproducibility among different laboratories, AFLPs have been used successfully in plant population genetic studies (Mulvih et al., 1999; Schmidt and Jenssen, 2000). The primary objectives of the study were to detect and measure genetic variation in natural populations of *U. kirkiana* in Malawi and examine the usefulness of AFLP marker technology in this species.

**MATERIALS AND METHODS**

**Collection of plant material**

A total of 96 samples were collected from 8 localities with 12 individuals representing a population (Table 1) along the geographical range of natural distribution of *U. kirkiana* in Malawi (Figure 1). Selected populations were heterogeneous for forest land tenure, habitat type and elevation. Collections were made from three land tenure categories namely public land in forest reserves, customary and leasehold land. About 5 g of young leaves were harvested from young leaves in the field and preserved in silica gel.
in 50 ml vials. Young leaves were sampled with minimum distance between sampled trees set as 2 m to reduce the chance of sampling from sprouts from similar individuals.

**DNA extraction**

Total genomic DNA extraction was from 200 mg leaf sample following the protocol of Patterson et al. (1993) with minor modifications. Following precipitation of impurities and RNase digestion isolated DNA was resuspended in 200 µl Tris-EDTA buffer and stored at -20°C until further analysis. DNA quality was checked by ethidium bromide staining on a 1% agarose gel and concentration was estimated by visual assessment relative to 1 Kb DNA (Gibco Life Technologies) ladder of different known concentrations (Sambrook et al., 1989).

**AFLP analysis**

AFLP reactions were performed on individual DNA samples from the different natural populations. AFLP analysis was carried out according to the protocol of Vos et al. (1995) using restriction enzymes EcoRI and MseI. Two replicates of PCR reactions from independent DNA extractions were performed to determine the accuracy of the analysis. Initially, 64 primer pairs originated by the combination of 8 EcoRI and 8 MseI primers were tested in ten *U. kirkiana* from 10 populations. From the primer screening exercise, six primer pairs (Table 2) were selected based on clarity and reproducibility of electrophoretic patterns and applied to all the 8 populations. Five microlitres of the 1:10 diluted pre-amplification was used as template for selective amplification using primers with three and four selective nucleotides at the 3’end. Two replicates of the PCR reactions from independent DNA extractions were performed to determine the accuracy of the analysis. AFLPs were
Table 2. Level of polymorphism and fingerprinting of AFLP markers in *Uapaca kirkiana*.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer combination</th>
<th>Total bands</th>
<th>Polymorphic bands</th>
<th>Polymorphism rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E33/M59</td>
<td>EcoR1-AAG/ Mse1-CTA</td>
<td>92</td>
<td>10</td>
<td>10.9</td>
</tr>
<tr>
<td>E33+C/ M47</td>
<td>EcoR1-AAGC/ Mse1-CAA</td>
<td>95</td>
<td>13</td>
<td>13.7</td>
</tr>
<tr>
<td>E35/M47</td>
<td>EcoR1-ACA/Mse1-CAA</td>
<td>107</td>
<td>14</td>
<td>7.6</td>
</tr>
<tr>
<td>E36/M62</td>
<td>EcoR1-ACC/ Mse1-CTT</td>
<td>106</td>
<td>16</td>
<td>15.1</td>
</tr>
<tr>
<td>E40+A/M60</td>
<td>EcoR1-AGCA/Mse1-CTC</td>
<td>104</td>
<td>25</td>
<td>24.0</td>
</tr>
<tr>
<td>E41+A/M59</td>
<td>EcoR1-AGGA/ Mse1-CTA</td>
<td>108</td>
<td>32</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>612</strong></td>
<td><strong>110</strong></td>
<td><strong>18.0</strong></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td><strong>102</strong></td>
<td><strong>18.3</strong></td>
<td><strong>5.6</strong></td>
</tr>
</tbody>
</table>

Table 3. AMOVA for 96 individuals of *Uapaca kirkiana* from eight populations in Malawi using 110 AFLP markers. Statistics include sums of squares, mean squares, variance component estimates, percentage of total variance and the probability (p) of obtaining a more extreme component estimate by chance alone (estimated from 1000 sampling realizations).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>SS</th>
<th>MS</th>
<th>Variance component</th>
<th>Total (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among regions</td>
<td>2</td>
<td>68.872</td>
<td>34.436</td>
<td>0.19623</td>
<td>1.21</td>
<td>&lt;0.094</td>
</tr>
<tr>
<td>Among three altitude groups</td>
<td>2</td>
<td>58.114</td>
<td>29.06</td>
<td>-0.04179</td>
<td>0.26</td>
<td>&lt;0.535</td>
</tr>
<tr>
<td>Among populations</td>
<td>5</td>
<td>141.008</td>
<td>28.2016</td>
<td>1.10478</td>
<td>6.79</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>Within population</td>
<td>88</td>
<td>1316.714</td>
<td>14.963</td>
<td>14.962</td>
<td>92.00</td>
<td>&lt;0.000</td>
</tr>
</tbody>
</table>

separated on 5% denaturing polyacrylamide sequencing gels (19:1 acrylamide: bis-acrylamide, 7 M urea) in 1X TBE buffer (1 M Tris, 1 M boric acid, 20 mM EDTA). The gels were pre-run for 40 min before 3 µl of the mix was loaded and run at 85 W for 3 h. Bands were visualised in sodium carbonate developing solution for 5 min and silver-stained gels were rinsed in ultra pure water and dried at room temperature in fumehood.

Data analysis

AFLP bands between 100 and 550 bp were scored twice manually and independently from the glass plates and counter checked. Polymorphic AFLP fragments were scored as 1 and 0 for presence and absence of bands, respectively, and data were entered into a binary matrix for each marker as discrete variables. The binary matrices were analysed by the software Arlequin 2.0 (Schneider et al., 2000) for calculating average difference between all the genotypes in the population (Tajima, 1993) and average gene diversity over loci (Nei, 1987). For each primer pair, the numbers of polymorphic and monomorphic bands were determined; however, bands that were monomorphic across all the 96 individuals were excluded from analyses. A matrix of corrected average pairwise differences between all pairs of populations generated by Arlequin was used for a constructing dendrograms through cluster analysis using the unweighted pair group method based on arithmetic averages (UPGMA) in the programme NTSYS-pc version 2.1 (Rohlf, 2000). Simple Matching and Jaccard’s coefficients of similarity (Sneath and Sokal, 1973) were calculated for all pairwise comparisons among populations and the coefficients presented similar results (only used the analyses based on Jaccard’s coefficients were used). Unscrambler 9.5 from CAMO in Oslo was used to investigate further relationships among the populations using principal component analysis (PCA) to show clusters of relatedness among *U. kirkiana* populations.

RESULTS

Level of polymorphism

The six primer combinations generated a total of 612 AFLP bands out of which 110 were polymorphic representing 18% polymorphism. Total number of polymorphic loci scored per primer pair ranged from 10 (E-AAG/M-CTA) to 32 (E41+A/MCTA) with an average of 18 loci per primer combination (Table 2) and these numbers suggest a low degree of variability.

Genetic diversity and differentiation

Analysis of molecular variance (AMOVA) (Excoffier et al., 1992) revealed low but significant genetic variation among populations of *U. kirkiana* but there were no significant genetic variations among *U. kirkiana* in the three regions and between the altitudes (low, medium and high) (Table 3). The mean Nei’s (1973) genetic diversity for all the 110 loci in the 8 populations was 0.27 ± 0.141 and the mean Nei’s genetic diversity for populations ranged from H=0.223 to H=0.322. Tsamba population had the lowest average gene diversity.
**Table 4.** Genetic diversity indices in 8 Malawian *Uapaca kirkiana* populations. The indices calculated are frequency of polymorphic loci, mean number of pairwise differences between individuals and average gene diversity over loci.

<table>
<thead>
<tr>
<th>Population</th>
<th>Frequency of polymorphic loci</th>
<th>Average difference</th>
<th>Average gene diversity (H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likhubula</td>
<td>0.82</td>
<td>34.74</td>
<td>0.322</td>
</tr>
<tr>
<td>Chimpeni</td>
<td>0.76</td>
<td>29.02</td>
<td>0.279</td>
</tr>
<tr>
<td>Tsamba</td>
<td>0.61</td>
<td>23.89</td>
<td>0.223</td>
</tr>
<tr>
<td>Chongoni</td>
<td>0.76</td>
<td>28.41</td>
<td>0.268</td>
</tr>
<tr>
<td>Dzalanyama</td>
<td>0.69</td>
<td>22.82</td>
<td>0.240</td>
</tr>
<tr>
<td>Chimaliro</td>
<td>0.78</td>
<td>29.35</td>
<td>0.282</td>
</tr>
<tr>
<td>Perekezi</td>
<td>0.76</td>
<td>27.23</td>
<td>0.279</td>
</tr>
<tr>
<td>Chesamu</td>
<td>0.78</td>
<td>27.23</td>
<td>0.269</td>
</tr>
<tr>
<td>Average</td>
<td>0.75 ± 0.08</td>
<td>27.84 ± 12.01</td>
<td>0.270 ± 0.14</td>
</tr>
</tbody>
</table>

whereas Likhubula had the highest average gene diversity (Table 4).

**Cluster analyses**

The dendrogram generated from AFLP data using UPGMA is shown in Figure 2. There is no clear structuring of the variability relative to region, however three clusters in UPGMA can be defined by cutting the dendrogram at a genetic similarity value of 0.078 and the populations separate according to geographic origin with first cluster composed of 3 populations from southern Malawi (Likhubula, Chimpeni and Tsamba) and two from Central Malawi populations (Chongoni and Dzalanyama). The second cluster is composed of Chimaliro from the Central region which clustered with Northern region population of Perekezi leaving Chesamu from Northern region as a third separate cluster.

**Principal component analysis**

Results of principal component analysis (Figure 3) show that only 36% of the total variation can be explained by the three PCA based on first-second and third eigenvectors which account for 22, 8 and 6%, respectively. The eight populations fall in one cluster although some individuals from populations 2, 4, 5 and 8
were separated from the cluster. Excluding the six individuals, the analysis did not show separation of the eight populations into different clusters (data not shown).

**DISCUSSION**

The data exhibit very high genetic diversity within population compared to among population of *U. kirkiana*; this is in agreement with its life history traits and geographical distribution. Very little is known about the mating system of *U. kirkiana*, however, work by Ngulube et al. (1998) has shown that the species is highly outcrossing. High within population variability is expected from such a plant with wide and continuous distribution (Hamrick and Godt, 1989). The highest levels of genetic diversity ($H=0.322$) was found within the broader distribution of Likhubula population in Southern Malawi with the restricted population of Tsamba in the same region showing lowest level of diversity. Distribution range and population size have been identified as major correlates of within population genetic variation in tropical tree species with restricted populations showing significantly less variation than those with a broader distribution (Loveless, 1992). In general, populations collected from protected forest reserves had higher genetic diversity than populations collected from customary and private forest suggesting negative impact of human activities on genetic diversity. This is consistent with field observations that the numbers of *U. kirkiana* trees and saplings in forest reserves were higher than on customary and private forests (Mwase et al., in preparation).

The results based on AMOVA show strong intrapopulation variation with small differentiation among the eight populations. The results agree with the general observation that woody, perennial, outbreeding species maintain most of their variation within populations (Hamrick et al., 1992). The lack of differentiation among regions and different altitude is indicative of high levels of gene flow and this is typical of woody outcrossing species. The low, medium and high altitude areas are randomly distributed in Malawi and they are often adjacent to each other in which case gene flow would be high. The negative variance among regional samples implies that alleles are more related than within populations. The lack of differentiation is also indicative of lack of sufficient time for significant differentiation along geographic lines. *U. kirkiana* is a pioneer plant species growing in open canopies in miombo woodland and low population differentiation can also be attributed to its longevity traits which are not very typical of early successional species. According to Wright’s (1978) guidelines for interpretation of $F_{ST}$ values, the range 0.0-0.005 indicates little, 0.05-0.15 moderate, 0.15-0.25 large and >0.25 very large population differentiation. The value

![Figure 3. Principal component analyses of populations of *Uapaca kirkiana* using 110 AFLP polymorphic markers.](image)
of $F_{ST}$ 0.079 suggests moderate differentiation and is consistent with the life history of *U. kirkiana*. Tropical trees tend to possess the most genetic diversity within populations (Hamrick and Loveless, 1989), and species reproducing both sexually and asexually show less differentiation than species reproducing only sexually (Hamrick and Godt, 1989). Although *U. kirkiana* reproduces both sexually and vegetatively by sprouting of prostate stems; this can be viewed as a strategy for reproductive success which is common in woody plants exposed to environmental conditions. The results agree with Agufa (2000) who reported high genetic diversity within populations of 91% for *U. kirkiana* samples collected from Zambia, Malawi, Zimbabwe and Tanzania with very low diversity among countries. The values are similar to those obtained in other studies of outcrossing tree species; for example, *Swietenia macrophylla* (Gillies et al., 1999) and *Populus tremuloides* (Yeh et al., 1995) where 87.9 and 97.4% of their genetic variation resided within populations, respectively. Lack of genetic variation among the three geographical region suggests considerable gene flow due to exchange of plant material taking place over longer distances of more than 200 km. Geographical distances between populations vary from 120 to 800 km. Natural seed dispersal by rodents and birds are thought to take place for very longer distances of up to 50 km. Taking into account the biological features of the plants, high effective gene flow would be expected. Human transportation of fresh fruits across large distances between forests and towns might have improved genetic exchange through seed dispersal within species natural distribution. Farmers selling fruits by road side markets and at markets in the towns may be responsible for long distance seed dispersal. The extensive gene flow could to a certain extent counteract effects of genetic drift in populations in the different geographical regions and act as a homogenization factor between nearby populations.

The low percentage (36%) of the variation explained by the first three components in the PCA could be attributed to high extent of genetic variability between individuals of each population than the low variation among the populations. The PCA support the AMOVA analysis. The PCA results analysis shows that the 8 populations did not segregate into distinguishable clusters; this is consistent with AMOVA results which indicated that genetic diversity was higher between individuals within a population than between populations. Even though the genetic differentiation among populations was significant, it accounted for only 6.8% of the variation. However the PCA showed two individuals from populations 2 (Chimpeni) and 4 (Chongoni) clustering together and this agrees with results of the UPGMA but the clustering by PCA did not agree in its entirety with UGMA. This could be attributed to the sensitivity of the PCA since this analysis is expected to be more informative about differentiation among major groups while the cluster analysis of UPGMA provides resolution among closely related populations. Populations tended to cluster together with other populations from the same region. However, Chimaliro from central region clustered with northern population of Perekezi indicating relatively high geographic proximity of the two populations. The two populations, even though belong to different regions, are close to each other and this facilitates gene exchange. Chesamu population is located on a physically isolated lakeshore area and the isolation may have influenced its genetic diversity through genetic drift.

Populations from communal forests of Tsamba had the lowest genetic diversity probably due to a founder effect since this is a population of recent origin. This result agrees with the prediction that marginal populations should be less variable than populations in the primary range (Stewart and Nilsen, 1995). Low genetic diversity in this low altitude region could also be attributed to extensive exploitation of the trees species by local communities. Changes in the landscape caused by deforestation and habitat destruction alter population density, abundance, diversity and abundance of pollinator communities (Lowe et al., 2005) thereby impinging on the mating system. The results signify the importance of focusing on populations which show more variability for conservation purposes. The present data pattern suggests that Likhubula population should be considered more distinct for the purposes of seed collection and management. However, for the purpose of conservation of genetic resources seed collection ought to be done across different populations to ensure a more representative sample of genetic variation. The country has shown low levels of genetic variation among the populations but very high within population. The results have some implications for conservation of *U. kirkiana* in Malawi. When selecting trees for clonal propagation, there is need to maximise selection from individual populations further suggesting that germplasm evaluation and conservation of *U. kirkiana* should focus on extensive sampling of more individuals especially from populations that are genetically diverse.

**CONCLUSION**

The analysis of AFLP exhibited very low genetic variation among *U. kirkiana* populations but very high variation between individuals in a population characteristic of an outbreeding species. These results are of great relevance for domestication and conservation of *U. kirkiana*. Priority on conservation and collection of germplasm should be given to populations with higher genetic diversity and it would be cost effective to sample a small number of populations in a region represented by a large number of individuals from a population. The genetic structure of the eight populations appears within the same range with higher genetic diversity values for population from forest.
reserves than populations from customary land. Overall, the study has demonstrated usefulness of AFLP markers in detecting the genetic diversity over a large number of randomly sampled loci and it has proved its usefulness in discriminating *U. kirkiana* populations and individuals within a population. Finally, our results indicate that a more intense monitoring of genetic diversity using AFLP should be conducted for *U. kirkiana* provenances in East and Southern African countries to identify candidate populations for conservation among the countries.

**ACKNOWLEDGEMENTS**

This research was supported by the Norwegian Council for Higher Education NU FU PROMA 69/2003 coordinated by Dr. M. Appelgren and Prof. M. Kwapata. We wish to express our gratitude to Anne-Guri Marøy at the SKP laboratory for technical assistance in the AFLP genotyping and A. Gedebo for help in scoring the AFLP data. We thank Dr. Siri Fjellheim for her helpful comments on the analysis of molecular data. We also thank G. Dawa L. Kalaundi and A. Kangoya for facilitating the field sample collection exercise.

**REFERENCES**

Agua CAC (2002). Genetic variation in *Sclerocaryea birea* and *Uapaca kirkiana* -indigenous fruit trees of the miombo woodlands. MSc thesis Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya.


