African Journal of Agricultural Research Vol. 6(4), pp. 790-797, 18 February, 2011
Available online at http://www.academicjournals.org/AJAR
DOI: 10.5897/AJAR10.313
ISSN 1991-637X ©2011 Academic Journals

Full Length Research Paper

An assessment of genetic diversity among marula populations using the amplified fragment length polymorphism (AFLP) technique

Kgabo Moganedi[1]*, Mbudzeni Sibara[2], Paul Grobler[3,4] and Elisabeth Goyvaerts[1,5]

[1] Department of Biochemistry, Microbiology and Biotechnology, School of Molecular and Life Sciences, University of Limpopo, Private Bag X 1106 Sovenga, 0727 - South Africa.
[2]Office of Deputy Vice Chancellor, University of Limpopo, Private Bag X 1106 Sovenga, 0727, South Africa.
[3]Department of Biodiversity, School of Molecular and Life Sciences, University of Limpopo, Private Bag X 1106 Sovenga, 0727 - South Africa.
[4]Department of Genetics, University of the Free State, P. O. Box 339, Bloemfontein, 9300 - South Africa.
[5]Kitso Biotech Pty Ltd, 23 Fernvilla P1, Pietermaritzburg 3201, South Africa.

Accepted 26 October, 2010

Marula is a wild-growing dioecious leafy tree species indigenous to Africa. The species exhibits a high phenotypic variability, especially in fruit size and quality and these traits are exploited by indigenous communities for commercial gain. As a first step towards assessing the relationship between the phenotypic and genotypic properties of the marula species, the genetic structure of natural marula populations was assessed using the Amplified Fragment Length Polymorphism (AFLP™) technique. Seven primer combinations were used to assess the degree of genetic diversity within and among natural marula sampled from the Bochum, Tzaneen and Nelspruit areas of South Africa. A total 141 unambiguous bands were amplified, of which 83 (59%) displayed polymorphism in one or more populations. Representative levels of genetic diversity were observed within all local populations. Coefficients of genetic differentiation showed evidence of drift among populations from the Limpopo- and Mpumalanga Provinces. A Bayesian assignment technique supported this trend, and suggested a true genetic structure of two populations, containing the Limpopo and Mpumalanga trees respectively. This diversity provides a plausible foundation for the marked phenotypic variation observed in marula, and suggest genetic diversity to provide for artificial selection during future artificial propagation programmes is present.

Key words: *Sclerocarya birrea*, marula, genetic diversity, genetic structure, amplified fragment length polymorphism.

INTRODUCTION

Marula species belongs to the genus *Sclerocarya* which falls under the mango family, Anacardiaceae. This family is widespread in the warmer regions of the world. It is a large group with over 60 genera and more than 500 species (Palmer and Pitman, 1972). Over 50 species of this family grow in Southern and Southwest Africa as trees, a few of which produce edible fruits and nuts such as mango (*Mangifera indica*), cashew (*Anarcadium occidentale*) and pistachio (*Pistacia vera*). There are two members of the genus *Sclerocarya*, namely, *Sclerocarya birrea* (marula) and *Sclerocarya gillettii* that are native to Africa and Madagascar. These species differ mainly in the number of leaflets, the length of the stalks and the sizes of the fruits. The species *S. birrea* has three subspecies: *caffra, birrea* and *multifoliata*, with *S. birrea* subsp. *caffra* being the most widespread and researched in Southern and Southwest Africa (Palmers and Pitman, 1972; Agufa et al., 2000; Emanuel et al., 2005).

Indigenous communities in South Africa, Botswana,

*Corresponding author. E-mail: kgabom@ul.ac.za. Tel: +27152682996. Fax: +27152683012.

Kenya, Tanzania, Swaziland, Mozambique, Namibia, Zimbabwe, Zambia and Malawi (Palgrave, 1984; Agufa et al., 2000; Eloff, 2000; Wynberg et al., 2002) utilize marula products for various purposes. The fruit is rich in minerals and carbohydrates and is either eaten fresh, processed or brewed. An interest in fruit and seed contents has also emerged (Mdluli, 2005; Hillman et al., 2008; Kleiman et al., 2008; Viljoen et al., 2008). The seeds are rich in oil (50 - 60%) and protein (28%) and are eaten either raw or roasted as nuts (ANU forestry, 2001). The bark and roots of the marula tree are also used in treating a variety of ailments such as ulcers, fever, sore eyes, schistosomiasis and stomach ailments (Palgrave, 1984; Eloff, 2000; ANU forestry, 2001). Marula has been declared a national tree in the Republic of South Africa because of its potential to be developed into a viable commercial crop (Marula natural products, 2003) and has attracted attention for commercial application.

Marula is functionally dioecious and out crossing. The species exhibits a high degree of intraspecies phenotypic differences for traits such as tree size, number of leaves in an inflorescence, fruit yield, fruit size, fruit juice and sugar content (Holtzhausen, 2001; Leakey et al., 2002; Muok et al., 2007). There is no published study that has investigated and attributed the relative contributions of genetic and environmental factors to these phenotypic variations. It is noteworthy that the phenotypic characteristics of a tree are often stable throughout its lifespan when growing in an environment of constant climatic conditions. Based on interviews during sample collection, it appears that local communities use these characteristics to tag the trees of high fruit quality. However, phenotypic properties may not offer a good measure of the actual genetic diversity that exists in a population because of the effect of environmental factors, especially when trees are grown at different locations of different climatic or physical conditions because morphological characteristics could be the result of environmental influences and their interactions with genetic properties. This implies that some genes of an organism may never be expressed unless these genes are presented with the right environment.

In this paper, we report on the use of the Amplified Fragment Length Polymorphism (AFLP) technique to assess genetic differentiation among conspecific populations and the genetic diversity within marula populations. In particular, the aims of this study were to assess the extent of genetic divergence between three marula populations in the Limpopo and Mpumalanga Provinces of South Africa and to determine the level of genetic diversity within marula populations.

MATERIALS AND METHODS

Study sites and sample collection

Leaf samples from 20 marula trees were collected from each of the three sampling areas Bochum (Bo) and Tzaneen (Tz) areas in the Limpopo Province, and from fields around Nelspruit (Ne) in the Mpumalanga Province (Figure 1). Equal numbers of both sexes were collected. Samples were labelled Bo 1-20, Tz 1-20 and Ne 1-20 indicating the source of the respective samples, with odd numbers representing male samples and even numbers female samples.

Genetic analysis

Leaf samples were ground to a fine powder under liquid nitrogen. The DNA was isolated following the method of Kleinhofs et al. (1993). The DNA from each genotype was double digested with *Tru*9I (M- Roche-Biochemicals) and *Eco*RI (E- Roche-Biochemicals) at appropriate conditions prescribed by the manufacturers and *Mse*I (M) and *Eco*RI (E) adaptors were ligated to the fragments. *Tru*9I and *Mse*I are isochizomers. The generic *Mse*I and *Eco*RI adaptors and primers were prescribed in Vos et al. (1995). Preselective amplification was performed in a 20 µl reaction mixture containing 1x *Ex*Taq buffer, 0.75U *Ex*Taq polymerase (Takara biochemicals), 0.5 µM of each primer (M-A/E-G and M-A/E-T), 0.25 mM dNTP mix and 5 µl restriction digested template. The cycling conditions consisted of 30 cycles, each with 94°C for 1 min, 56°C for 90 and 90 s at 72°C, followed by 7 min at 72°C. Polymerase Chain Reaction (PCR) reaction mixtures for selective amplification were the same as for preselective amplification except that primers with more selective nucleotides were used, that is, M-ACA/E-GTT, M-AGT/E-GCA, M-AGC/E-GAC, M-ACA/E-GAC, M-ACA/E-TAT, M-AGA/E-TGC and M-AGT/E-TCC. A touchdown cycling program was performed for selective amplification of the template using the following cycling conditions: 13 cycles of 94°C for 1 minute; 1 minute at 65→ 56°C (Δt = 0.7°C per cycle) and 1 minute at 72°C and a further 17 cycles at 94°C for 1 min; 1 min at 56°C and 1 min at 72°C, with an extended 7 min at 72°C. The amplification products were analysed on a 5% denaturing polyacrylamide gel [PAA gel solution: 5% acrylamide/bisacrylamide (19:1), 7.5 M urea, 89 mM Tris (pH 8.3), 89 mM borate and 2 mM EDTA]. The fingerprints were visualised with silver staining following the procedure in the Promega© protocols and applications guide (1996). AFLP fingerprints were manually scored for the presence (1) and the absence (0) of bands.

Statistical analysis

Distinct monomorphic and polymorphic bands were analysed using POPGENE (Yeh et al., 1997), ARLEQUIN ver 3.0 (Excoffier et al., 2005) and NTSYS-pc (version 2.02i; Rohlf, 1997) softwares. POPGENE was used to determine observed number of alleles (Ao); expected heterozygosity (*He*) within populations and genetic distance (D; Nei, 1972) among populations**.** ARLEQUIN software was used to determine the spatial distribution of genetic diversity in marula, by implementing an Analysis of Molecular Variance (AMOVA) and to calculate conventional Fst values. Associate gene flow (Nm) values were calculated using Nm= [0.25*(1-Fst))/Fst]. The NTSYS-pc program was used to calculate genetic similarity as a measure of genetic differentiation using the dice coefficient; and a Principal Co-ordinates Analysis (PCO) based on the genetic similarity matrix was then determined.

To supplement the frequency-based analyses, an assignment test based on a Bayesian approach was used to identify the true number of populations (clusters) and assign individual marula trees probabilistically to each cluster. Structure software Pritchard et al. (2000) and Falush et al. {2003} was used for this analysis. A model with assumption of admixture ancestry and correlated allele frequencies was selected. The parameter ln Pr (X|K) was calculated

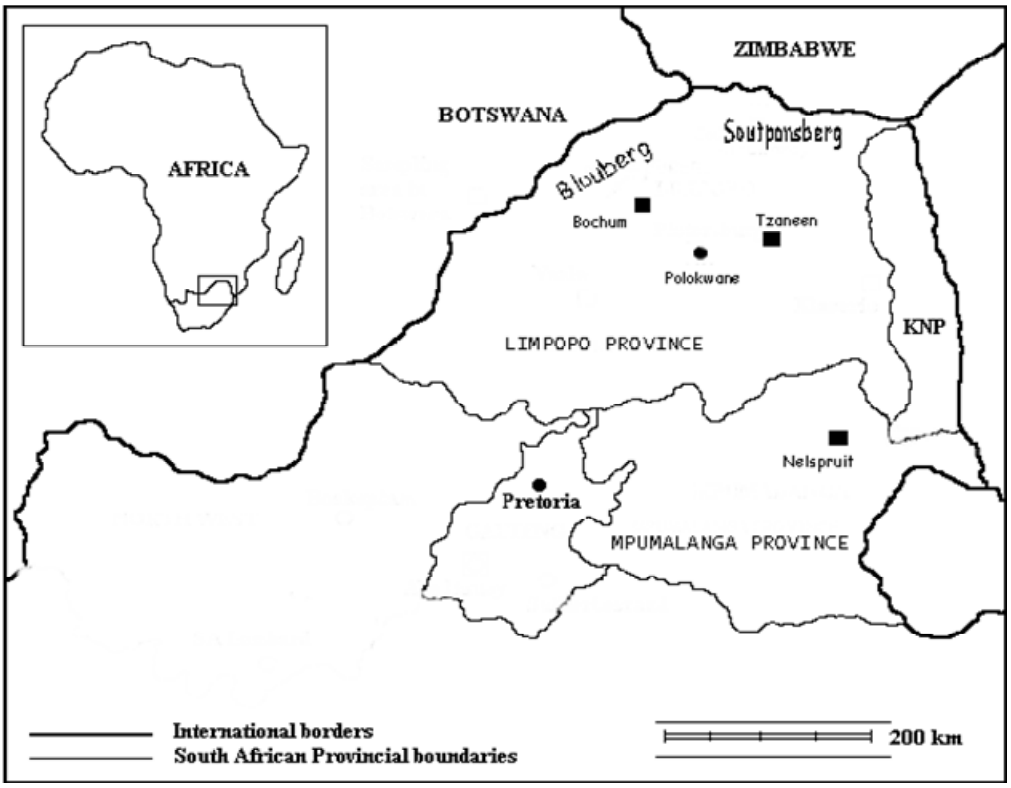

Figure 1. A geographical map of Limpopo and Mpumalanga provinces showing the three areas (■) where natural marula trees were sampled.

Table 1. Genetic distance (D; below the diagonal); genetic variability measures (diagonal; with standard deviation in brackets) and Fst and gene flow estimates (above the diagonal) among marula populations.

	Bochum	Nelspruit	Tzaneen
Bochum	Ao=1.411(0.494) He=0.168 (0.201)	Fst=0.219 (p=0.001) Nm=0.892	Fst=0.134 (p=0.001) Nm=1.616
Nelspruit	0.058	Ao=1.461(0.500) He=0.184 (0.212)	Fst=0.191 (p=0.001) Nm=1.059
Tzaneen	0.053	0.055	Ao=1.404 (0.493) He=0.171 (0.217)

for K values (number of populations) of 1-3, with 10 independent runs for each K, to estimate the true number of populations. All runs consisted of a burn-in period of 100,000 steps, followed by 200,000 iterations.

RESULTS

A total of 141 bands were amplified using the seven combinations of selective primers. Of these bands, 83 (59%) were polymorphic in one or more populations. Intraspecific genetic variability was very similar in all the three populations (Table 1), with the expected heterozygosity values within populations ranging from 0.168 to 0.184.

Coefficients of genetic differentiation suggested significant geographic genetic structuring within marula. Values from AMOVA showed that 19.831% of total variation is found among the three regional populations, with 80.169% within populations. Pair-wise Fst values showed significant differentiation ($P < 0.001$) between all populations pairs. The Fst values between the populations from Bochum and Tzaneen (Fst=0.134) were however, slightly lower compared to values between these two populations and the third, from the Nelspruit area (Fst=0.191-0.219). Gene flow values supported this trend (Table 1), with Nm values indicative of less exchange of genetic material between Nelspruit population and either of the populations from Bochum or

Table 2. Proportion of membership of each pre-defined marula population in each of two clusters, from a fully Bayesian assignment test following Pritchard et al. (2000).

	Cluster 1	Cluster 2
Bochum	0.732	0.268
Nelspruit	0.069	0.931
Tzaneen	0.758	0.242

Tzaneen. Genetic distance between population pairs were closely comparable, but provided some support for the observation of closer identity between the populations from the Limpopo Province (D=0.53) compared to distances between these populations and the Nelspruit population (D=0.55-0.58; Table 1).

The pattern from Fst and D values is also evident from the PCO plot, which is based on the relative similarities between individuals and shows a cluster – though separate from the Nelspruit individuals when compared to the interspersed individuals from Bochum and Tzaneen areas. It is noteworthy that none of the three clusters on the PCO plot showed any tendency of the individuals of the same sex to cluster together (Figure 2).

Bayesian analysis suggested a genetic structure consisting of two marula populations, based on the posterior probabilities of K=1-3 (Figure 3). The -Ln probabilities showed a plateau at this K value and the standard deviation between different runs of K=2 was very low (0.745 compared to 1.866-1.944). The proportion of membership of each pre-defined population in each of two clusters is presented in Table 3 and Figure 4. A total of 73.2% of marula trees from the Bochum area and 75.8% from Tzaneen were assigned to one cluster, whereas 93.1% of individuals from the Nelspruit area were assigned to the second cluster.

A general phenotypic variation among natural marula was observed among populations from the Bochum, Tzaneen and Nelspruit areas. It was noted during this study that marula fruits from Tzaneen and Nelspruit are bigger and juicier than marula fruits from Bochum.

DISCUSSION

The AFLP analysis proved to be useful in revealing genetic differences between and within marula populations. The He analyses indicated a high intraspecific genetic diversity in all populations. The He values observed in this study (0.168-0.184) are slightly lower though comparable to values reported by Portis et al. (2005) in *Cynara cardunculus* (He=0.167-0.218) and Han et al. (2006) in *Gardenia jasminoides* (0.140-0.207) and were described as indicative of high variation by these authors. The substantive similarity between our results and these AFLP-based studies indicate representative level of diversity in wild populations of tree species.

A gradual increase in intraspecific genetic diversity from Northwest (Bochum) to southeast (Nelspruit) for He and Ao values was apparent. This trend is probably an indication of isolation by distance and is explained by the geographic closeness between the Bochum and Tzaneen populations, with the Nelspruit population more distant. The results from AMOVA provide further support for a hypothesis of geographic structure in Marula. AMOVA indicated that 19.831% of total genetic diversity was found between regional populations. This among-region value is substantially higher than the 9.10% calculated among regional populations of *G. jasminoides* by Han et al. (2006), and slightly exceeds the value of 17.02% calculated between different species of the genus *Prunus* by Aradhya et al. (2004), suggesting significant differentiation. Further, comparisons between population pairs using Fst (Table 2) confirm the hypothesis of significant differentiation among populations, with significant (P=0.001) divergence between all population pairs (Table 2).

Han et al. (2006) noted that Nm<1.0 suggests signifi-cant genetic drift (following Wright 1940; 1943), whereas Nm values above 4.0 indicate sufficient gene flow to overcome genetic drift. The Nm values of 0.892 and 1.059 calculated between marula from the Bochum and Nelspruit areas, and the Tzaneen and Nelspruit areas, are thus on or below the threshold leading to potentially significant biological differentiation, whereas Nm between the Bochum and Tzaneen populations (1.616) indicates less interruption of gene flow which is explained by a continuum of marula stands that extends between the two regions when compared to the distance between either Bochum or Tzaneen region and Nelspruit region. A certain degree of genetic uniformity is expected within a confined area but the outcrossing nature of marula allows a good exchange of genetic material between adjacent areas which is what has been observed between Bochum and Tzaneen populations. The genetic relatedness among populations is further confirmed by D values of 0.053-0.058 (Table 1) that are relatively similar for all pairwise comparisons between populations, but with least distance among the populations from the Bochum and Tzaneen areas. Further strong evidence for genetic structure in marula is found in results from the assignment test, with the Bochum and Tzaneen marula stands grouped into one cluster and marula stands from Nelspruit region forming a distinct second, well defined cluster containing 93.1% of trees in that population (Table 2, Figs. 3 and 4). The pattern of some divergence in the Nelspruit population compared to the other two populations also supports the trend observed from coefficients of genetic diversity, which showed a gradient from northwest to southeast. The Nelspruit marula population thus shows some evidence of genetic drift relative to Bochum and Tzaneen populations.

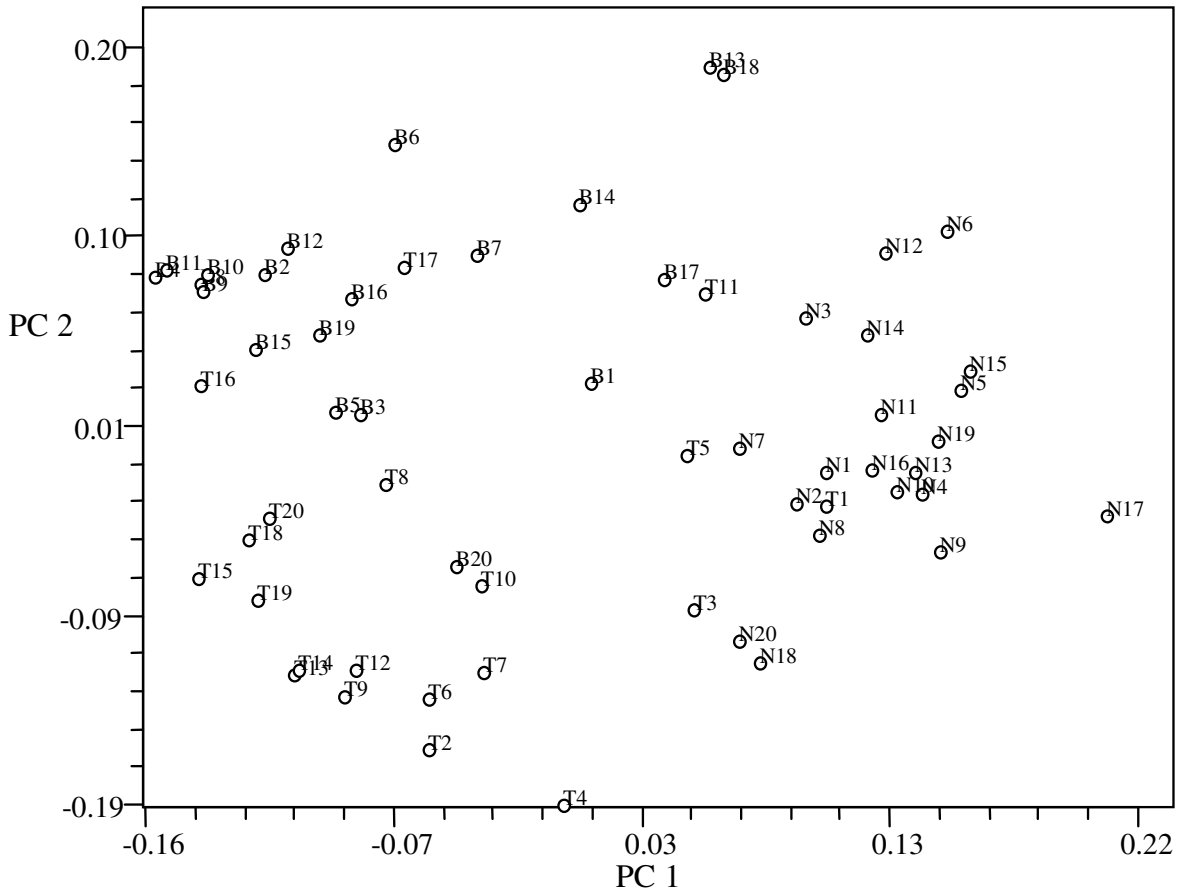

Figure 2. PCO plot showing the dispersal of the natural genotypes in three South African populations, in relation to each other.

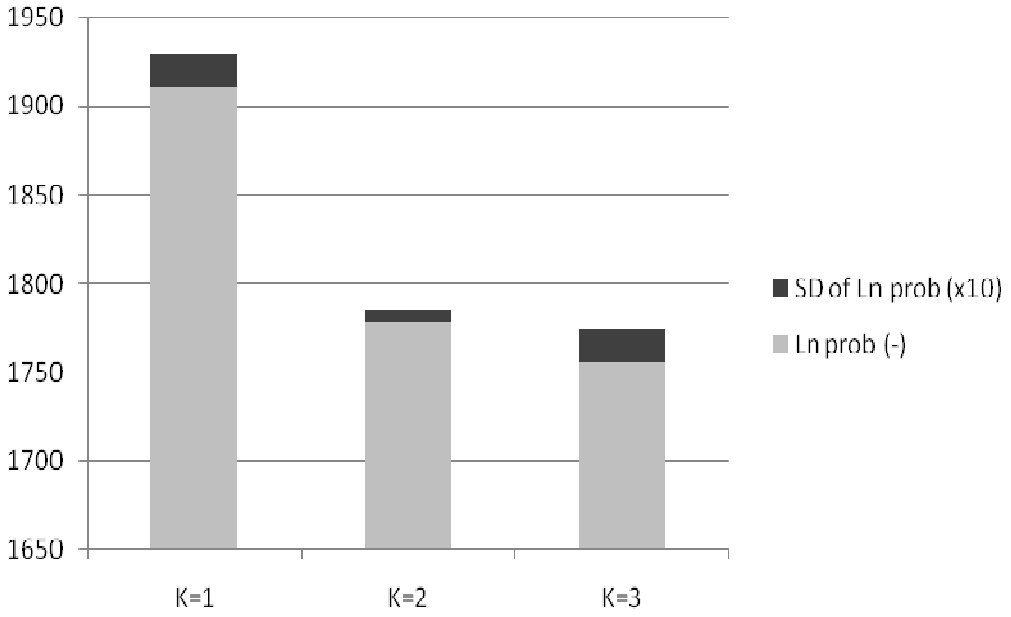

Figure 3. Probability of K=1-3 in three marula populations. Light grey bars show the average –Ln probability for 10 simulations of each value of K; darker areas above bars indicate the standard deviation for 10 simulations.

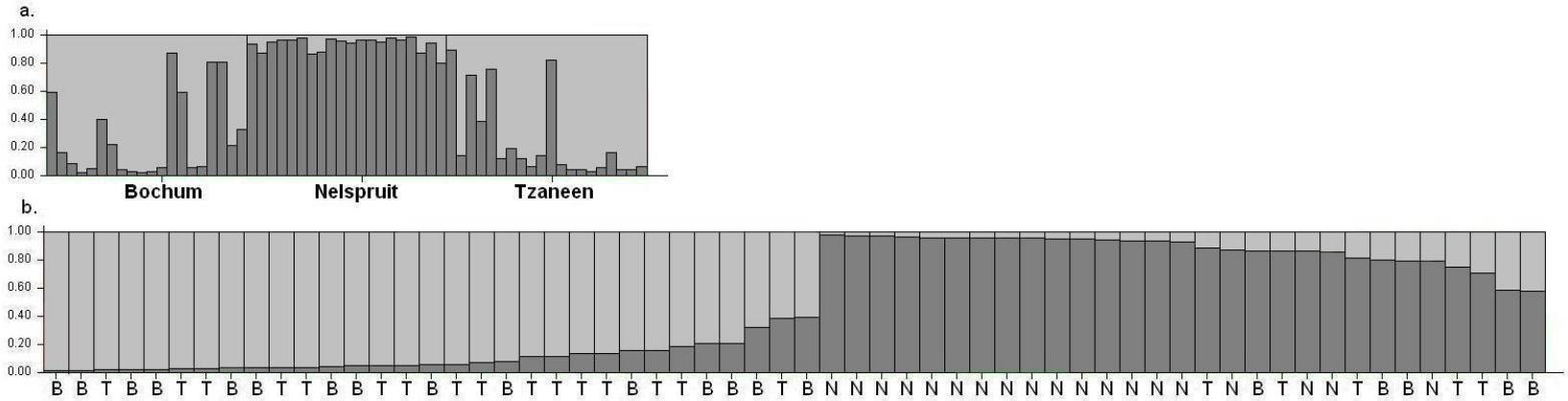

Figure 4. Bar plots (from structure). Each tree is represented by a single vertical line broken into two differently shaded segments, with lengths proportional to membership of each of the two inferred clusters. (a) Three geographic populations in original order. (b) Membership of individuals to two true clusters of marula. Abbreviations used: B=Bochum; N=Nelspruit; T=Tzaneen.

Naturally occurring marula trees are widely dispersed throughout the Limpopo and Mpumalanga Provinces, with a high density in non-residential areas. The present study shows that this exchange of genetic material induces gene flow within local localities, as demonstrated by representative within-population levels of genetic diversity. Over greater geographic distances, some genetic structure in marula is apparent as shown by various comparisons of marula trees from the Nelspruit area compared to the remaining populations. This may be pure drift caused by isolation through distance or it may represent true adaptive differentiation due to a gradient of environmental factors or a combination of both.

The observed phenotypic differences that have been reported in marula from different regions suggest that some adaptive differences, stemming from genetic differentiation, do occur. Thiong'o et al. (2000) also observed a wide phenotypic variability within marula in the wild which can be exploited to produce superior individuals. Phenotypic variations in fruit size, fruit yield per tree, flesh, juice and sugar contents between trees within regions have been reported by Holtzhausen (2001) and Leakey et al. (2002).

Such differences can be exploited during future artificial breeding programs to suit the needs of specific enterprises. Similarly, phenotypic differences among natural marula were observed among populations from the Bochum, Tzaneen and Nelspruit areas. The differences in juice content observed among marula trees in the three regions could be the result of differential amount of free water that is available from rainfall. Bochum is relatively dry when compared to Tzaneen and Nelspruit which both receive a higher amount of rainfall during spring (South African weather service, 2004). Nevertheless, genetic differentiation and adapta-tion to specific environmental conditions could equally explain the phenotypic differences.

Notwithstanding the slight geographic genetic structuring observed in marula, the genetic data obtained in the current study do not suggest reduced exchangeability or reproductive isolation between regional populations, and therefore support the presence of only one subspecies in the two provinces studied (Limpopo and Mpumalanga). Muok et al. (2007) studied genetic and morphological relationship between marula populations in Kenya and Tanzania and resulted in restructuring of the marula populations in Kenya from two to three subspecies.

The observation from this study suggested less exchange of genetic material between Nelspruit and the other two populations Bochum and Tzaneen. However, a high intraspecific variation within Nelspruit population was revealed. Noting the geographical location in relation to neighbouring areas and countries, some level of gene flow may be probable between the Nelspruit marula population and Swaziland and/or Mozambique populations based on (1) the genetic separation of Nelspruit population from the

Tzaneen and Bochum populations, (2) Swaziland and Mozambique are both hosts to the same marula species (Palgrave, 1984; Agufa et al., 2000) and (3) gene flow is enhanced in insect- based pollinated species because of good transfer of pollen over large distance (Fatemi and Gross, 2009). In conclusion, the current study provides the first data on AFLP-based genetic variation within and among South African marula populations. Most studies on marula have focused on the fruit content and its utilization in producing commercial products, not on its natural genetic diversity. An understanding of the genetic structure is important for the protection of the biodiversity of this species especially noting that marula has been declared a national tree in the Republic of South Africa because of its economic value. Some genetic structure exists based on the observations in this study; however, intraspecific diversity outranks any differentiation resultant from genetic drift. There is an underlying supposition that genetic, environmental and epigenetic factors interact in determining the marula phenotype based on the diverse phenotypic characteristics alongside the narrow genetic variation. The study of the relationship between the environment, genetic properties and specific phenotypic characteristics will provide additional data on the foundation of phenotypic variation, so that important attributes could be preserved and utilized for commercial selection in marula.

ACKNOWLEDGEMENTS

The authors would like to thank the Flanders Interuniversity Institute for Biotechnology (VIB) at the University of Gent, Belgium for providing skills, funds and space to carry out the laboratory work. Our gratitude goes to National research Foundation for financial assistance; and we further thank Ms.'s N. Colpaert and M Legodi; Messrs. H. Labbeau, D. B. Moganedi and S. Mtileni for their inputs and technical assistance during the execution of this study.

REFERENCES

Agufa CAC, Simons AJ, Maghembe J, Dawson IK (2000). Molecular genetic variation within and between populations of *Sclerocarya birrea* measured by RAPD and chloroplast RFLP-PCR analysis: Implications for genetic management of the species. Http://www.bangor.ac.uk/afforum/research/monographs/output.htm. Accessed 1 March 2004.

ANU forestry (2001). Marula fruit. <http://www.anu.edu.au/Forestry/wood/nwfp/toesu/Toesu2.html Accessed 22 May 2001>.

Aradhya MK, Weeks C, Simon CJ (2004). Molecular characterization of variability and relationships among seven cultivated and selected wild species of *Prunus* L. using amplified fragment length polymorphism. Sci. Hortic., 103: 131-144.

Eloff JN (2000). Antibacterial activity of Marula (*Sclerocarya birrea* (A. Rich.) Hochst. subsp. *caffra* (Sond.) Kokwaro) (Anacardiaceae) bark and leaves. Fitoterapia, 71: 570- 573

Emanuel PL , Shackleton CM, Baxter JS (2005). Modelling the sustainable harvest of *Sclerocarya birrea* subsp. *caffra* fruits in the South African lowveld. Forest Ecol. Manage., 214: 91- 103

Excoffier L, Laval G, Schneider S (2005). Arlequin, version 3.0: An Integrated Software Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Laboratory (CMPG), Institute of Zoology, University of Berne, Switzerland.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics, 164: 1567–1587.

Fatemi M, Gross CL (2009). Life on the Edge-High levels of genetic diversity in a cliff population of *Bertya ingramii* attributed to *B. rosmarinifolia* (Euphorbiaceae). Biol. Conserv., 142: 1461-1468

Han J, Zhang W, Cao H, Chen S, Wang Y (2006). Genetic diversity and biogeography of the traditional Chinese medicine, *Gardenia jasmoides*, based on AFLP markers. Biochem. Syst. Ecol., 35: 138-145.

Hillman Z, Mizrahi Y, Beit-Yannai E (2008). Evaluation of valuable nutrients in selected genotypes of marula (*Sclerocarya birrea* ssp. *caffra*). Sci. Hortic., 117: 321-328

Holtzhausen C (2001) Ennobling the African Marula. Report and abstracts, Marula workshop, 14-16 February 2001, Magaliespark Timeshare, Hartebeespoortdam

Kleiman R, Ashley DA, Brown JH (2008). Short communication: Comparison of two seed oils used in cosmetics, moringa and marula. Ind. Crops Prod., 28: 361-364

Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Dukrna D, Bollinger J, Knapp SJ, Lui B, Sorrells M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993). A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor. Appl. Genet., 86: 705-712

Leakey R, Shackleton S, du Plessis P, Pate K, Lombard C (2002). Characterization of phenotypic variation in marula (*Sclerocarya birrea*) fruits, nuts and kernels in South Africa and Namibia. . Agroforestry and Novel Crops Unit, James Cook University, Australia

Marula natural products (2003) Marula news. http://www.marula.org.za/news3.htm Accessed 23 July 2003

Mdluli KM (2005). Partial purification and characterisation of polyphenol oxidase and peroxidase from marula fruit (*Sclerocarya birrea* subsp. *caffra*). Food Chem., 92 (2): 311- 323

Muok B, Matsumura A, Ishii T, Odee DW (2007). Genetic diversity within *Sclerocarya birrea* populations in Kenya. J. Arid Environ., 71: 1-11

Nei M (1972). Genetic distance between populations. Am. Nat., 106: 283-292.

Palgrave KC (1984). Trees of Southern Africa, 4th edn. C. Struik publishers, Cape Town.

Palmer E, Pitman N (1972). Trees of Southern Africa. A. A. Balkema, Cape Town.

Portis E, Acquadro A, Comino C, Mauromicale G, Saba E, Lanteri S (2005). Genetic structure of island populations of wild cardoon [*Cynara cardunculus* L. var. sylvestris (Lamk) Fiori] detected by AFLPs and SSRs. Plant Sci., 169: 199-210.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000). Association mapping in structured populations. Am. J. Hum. Genet. 67: 170–181.

Promega (1996). Protocols and Applications guide: the source for discovery, 3rd edn, USA.

Rohlf FJ (1997). NTSYSpc version 2.02i. Numerical taxonomy and Multivariate Analysis System. Exeter Software, Setauket, New York

South African weather services (2004). Climate data. http://www.weathersa.co.za/climate/rainfall Accessed 5 March 2004

Thiong'o M (2000). Selection of superior indigenous fruit trees based on nutritional analysis: the case of Marula. Http://www.bangor.ac.uk/afforum/research/monographs/output.htm Accessed 1 March 2004.

Viljoen AM, Kamatou GPP, Başer KHC (2008). Short communication: Head-space volatiles of marula (*Sclerocarya birrea* subsp. *caffra*). South. Afr. J. Bot., 74: 325-326.

Vos P, Rogers R, Bleeker M, Reijans M, Van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res., 23: 4407-

4414.
Wright S (1940). Breeding structure of populations in relation to speciation. Am. Nat., 74: 232-248.
Wright S (1943). Isolation by distance. Genetics, 28: 114-138.
Wynberg R, Cribbins J, Leakey R, Lombard C, Mander M, Shackleton S, Sullivan C (2002). Knowledge of *Sclerocarya birrea* subsp. *caffra* with emphasis on its importance as a non-timber forest product in South and southern Africa: A Summary. South. Afr. Forest. J., 196: 67-78.
Yeh FC, Yang R, Boyle T (1997). POPGENE version 1.32. Microsoft windows-based freeware for population genetic analysis.