Genetic characterization of selected Macadamia germplasm and the implications on breeding and conservation in Kenya

Conference Paper · December 2011

DOI: 10.13140/2.1.1566.8164

7 authors, including:

Lucy Gitonga

University of the Witwatersrand

9 PUBLICATIONS **21** CITATIONS

SEE PROFILE

Bramwel Waswa Wanjala

Kenya Agricultural Research Institute

20 PUBLICATIONS **58** CITATIONS

SEE PROFILE

Kamau Ngamau

Jomo Kenyatta University of Agriculture an…

14 PUBLICATIONS **27** CITATIONS

SEE PROFILE

Simon Gichuki

Kenya Agricultural Research Institute

32 PUBLICATIONS **330** CITATIONS

SEE PROFILE

Genetic Characterization of Selected Macadamia Germplasm and the Implications on Breeding and Conservation in Kenya

L.N. Gitonga[1], B.W. Wanjala[1], A.W.T. Muigai[2], K. Ngamau[2], E. Kahangi[2], S.T Gichuki[1] and M.N. Kairichi[1]
[1] Kenya Agricultural Research Institute, PO Box 57811, Nairobi-00200, Kenya
[2] Jomo Kenyatta University of Agriculture and Technology, PO Box 62000, Nairobi-00200, Kenya

Keywords: macadamia, heterozygosity, genetic diversity, genetic differentiation, conservation

Abstract

Macadamia is the most important nut crop in Kenya with an estimated annual production of about 13,510 metric tones with an estimated value of $ 4.8M. The cultivated macadamia is derived from two species; the rough-shelled *M. tetraphylla* and the smooth shelled *M. integrifolia*. The Kenyan macadamia industry relies on selection of superior trees from germplasm derived from open-pollinated seeds of the two species planted between 1946 and 1968. Twenty six macadamia accessions sampled from five populations; Bob Harries, Thika, Kirinyaga, Embu and Meru in central and eastern Kenya were analyzed using 110 polymorphic markers of six AFLP primer combinations. The highest percentage of polymorphic loci of 80.0% was observed in the Bob Harries population and the least, 67.3%, from Thika population. Expected heterozygosity was also highest in the Bob Harries population. Phylogenetic analysis based on UPGMA of genetic distances of the five populations revealed the Bob Harries population, the source of macadamia germplasm in Kenya to be closely related only to the Kirinyaga population and distantly related to the Thika population, a preservation block of selected superior germplasm. Cluster analysis based on the 26 accessions did not distinctly separate the species indicating a high level of heterozygosity and genetic mixture of the original gene pool used in Kenya. Analysis of molecular variance further indicated higher genetic variation within the populations, 96%, than among the populations, 4%, and high genetic differentiation among the accessions (theta=0.1909). The results of this study indicate high genetic diversity within the Kenyan macadamia germplasm. This information can immediately be used for enhanced breeding and more effective sampling of populations for conservation.

INTRODUCTION

Macadamia (family *Proteaceae*), is indigenous to the subtropical coastal region of Australia (Peace et al., 2008). The genus consists of nine species (Costello et al., 2008) but only two; *Macadamia integrifolia* and *M. tetraphylla* are cultivated for their edible nuts (Peace et al., 2003). Macadamia was first introduced in Kenya in 1946 by an Australian farmer, Bob Harries, as six seeds of *M. tetraphylla* who planted them as ornamentals in Thika in Central Kenya (Harries, 2004). Over the next 18 years, he used the open-pollinated seeds from the original six *M. tetraphylla* trees to propagate seedlings for his farm (Harries, 2004). In 1964, seeds of *M. integrifolia*, *M. tetraphylla* and hybrids of the two were imported from Australia, Hawaii and California. Scion material from superior *M. integrifolia* cultivars were also imported from Hawaii including 'HAES 246', 'HAES 328', 'HAES 333', 'HAES 508', 'HAES 660' and 'HAES 669' and grafted seedlings were produced and planted at different agro-ecological (Harries, 2004). These three sources comprise the macadamia germplasm grown in Kenya as a cash crop mainly for export (Wasilwa et al., 2003).

Genetic improvement of macadamia in Kenya started in 1977 with selection of trees with superior characteristics based morphological characteristics of tree form, leaf and nuts and biochemical analysis of nuts to determine quality (Ondabu et al., 1996; Wasilwa et al., 2003). Kenya is the fourth largest producer of macadamia nuts after

Australia, Hawaii and South Africa (Wilkie, 2008). In the year 2007, 11,100 t in-shell nuts were produced while 2008 a total of 13,510 metric tones were produced with an estimated value of US$ 4,887,194 (Ksh 351,878,000 (1 Ksh=0.72 US$)) (Muthoka et al., 2008; Wilkie, 2008). However, the breeding process has been slow and to date only four cultivars have been recommended (Ondabu et al., 2007). In breeding of new cultivars it is important to know the extent of genetic variation within and between populations (Nei and Kumar, 2000). Among several genetic markers for molecular diversity studies, amplified fragment length polymorphisms (AFLP) offer several advantages in that only small amounts of DNA are needed for analysis, use of only two primers gives reproducible results and many restriction fragments can be amplified by changing the nucleotide extensions on the adaptor sequence, high resolution is obtained because of the stringent PCR conditions and no prior knowledge of the genome sequence is needed (Vos et al., 1995). The objective of this study was to determine genetic diversity in selected macadamia germplasm in Kenya using AFLP markers for enhanced breeding and conservation.

MATERIALS AND METHODS

Twenty six (26) macadamia accessions were sampled from five populations (Table 1) and maintained in greenhouse as grafted seedlings to provide young shoots for DNA extraction. DNA was extracted from 0.4 g of young leaf tissues following cetyl trimethyl ammonium bromide, $CH_3(CH_2)_{15}N^+(CH_3)_3Br$ (CTAB) method (Saghai-Maroof et al., 1984). AFLP genotyping followed Vos et al. (1995). A total of eight AFLP primer combinations (*Mse*I and *Eco*RI) were screened in order to select pairs that were polymorphic and easy to score. These were *M*-CTC × *E*-AGG, *M*-CAT × *E*-ACC, *M*- CTA × *E*-ACC, *M*-CAA × *E*-AAC, *M*-CAG × *E*-ACT, *M*-CAG × *E*-ACC, *M*-CAG × *E*-AAC, *M*-CAG × *E*-AGG. Out of these, the first six pairs gave consistent results and they were selected for subsequent PCR as they resulted to good distinct bands on the gel.

Samples were then analyzed by ABI 373 Genetic Analyzer and results of AFLP reactions evaluated using GeneScan software (version 3.1.2., Applied Biosystems) and the allele size was determined with Genotyper® software (version 2.0, Applied Biosystems). Only polymorphic AFLP markers were scored with a number 1 for presence and 0 for absence of band.

The percentage of polymorphic loci and the expected heterozygosities were calculated using GenALEx version 6.2 (Peakall and Smouse, 2005) and calculation of expected heterozygosity assumed random mating. Genetic distance matrices for the five populations were calculated following the method of Huff et al. (1993) based on Nei's genetic distance (Nei, 1972). The distance matrices were subjected to UPGMA (unweighted paired group method using arithmetic averages) (Sneath and Sokal, 1973) cluster analysis to assess the grouping of populations using the tools for population genetic analyses (TFPGA) version 1.3. (Miller, 2000). Principle component analysis (PCA) was calculated using the covariance matrix with data standardization using GenALEX version 6.2 (Peakall and Smouse, 2005). Analysis of molecular variance (AMOVA) following Excoffier et al. (1992) was calculated using GenALEx version 6.3 (Peakall and Smouse, 2005) while calculation of population differentiation, P (theta) was done using TFPGA and assuming Hardy-Weinburg equilibrium.

RESULTS AND DISCUSSION

A total of 110 alleles were observed in the five macadamia populations. Allele frequencies across all loci ranged from 0.0 to 1.0. The Bob Harries population was the most genetically diverse with highest percentage of polymorphic loci of 80.0% and highest heterozygosity of 0.30 significantly higher than all other populations (Fig. 1). The Thika population had the least percentage of polymorphic loci of 67.3% and the least heterozygosity of 0.22. The other populations had polymorphisms intermediate between these two populations. These results support the theory that the Bob Harries population formed the base gene pool from which macadamia germplasm in Kenya was distributed

(Harries, 2004). The theory states that the two species *M. integrilfolia* and *M. tetraphylla* were imported by Bob Harries family at different occasions and from Australia, Hawaii and California from which open-pollinated seeds were used to propagated seedlings for farmers. The low mean number of different alleles, mean number of effective alleles and expected heterozygosity in the Thika population are indicative of a relatively uniform germplasm. This can be explained by the fact that these accessions were selected by the Kenya Agricultural Research Institute (KARI) based on mostly high yields and nut and kernel characteristic (Ondabu et al., 1996; Kiuru et al., 2004) and they were mostly those related to *M. integrifolia* and (*M. integrifolia* × *M. tetraphylla*) hybrids. Hence, they are expected to be closely related in their genetic make-up encoding for these traits. This fact is also supported by the presence of highest number of locally common alleles.

One private allele was observed in the Embu population and since this population also had high number of locally common alleles, high percentage of polymorphic loci and high heterozygosity it makes it a good target for sampling for maximum possible genetic variation for conservation (Maguire et al., 2002; Lanteri and Barcaccia, 2006).

Pair-wise comparisons between the five macadamia populations indicated the largest genetic distance occurred between the Embu and the Kirinyaga populations according to Nei genetic distance and unbiased genetic distance of 0.139 and 0.100 respectively. However, the least Nei genetic distance of 0.050 occurred between Meru and Thika populations and between Bob Harries and Kirinyaga populations using Nei genetic distance of 0.055 (Table 2).

Results of cluster analysis based on UPGMA (Nei, 1972) genetic distance revealed the Bob Harries population to be closely related to the Kirinyaga population and clustered together in a separate sub-cluster. Meru and Thika populations were also closely related and distantly related to both Bob Harries and Kirinyaga populations. The Embu population clustered as a separate sub-cluster (Fig. 2). The numbers at the juncture of two branches are the percentage of 1000 bootstrap trees with the same branch structure.

Cluster analysis based on the 26 accessions did not separate them into distinctly *M. integrifolia*, *M. tetraphylla* or hybrids. This implies high level of heterozygosity and genetic admixture of the original gene pool used in Kenya confirming the need to preserve this germplasm.

Principle component analysis of the five macadamia populations generated four principle components with the first three principle components accounting for 90.4% of the total variation; PC1 (45.98%), PC2 (31.86) and PC3 (12.53). The most closely clustered populations were Thika and Meru. The Bob Harries and Kirinyaga populations also clustered relatively close together and distant from Thika, Meru and Embu populations and were separated by the first principle component. The Embu population clustered alone, distant from the other four populations (Fig. 3).

Results of genetic relationships among the macadamia population based on genetic distances, phylogenetic analysis and principle component analysis consistently showed that the most closely related populations were Thika and Meru as they had the least genetic distance and grouped in one sub-cluster. From the results of principle component analysis accessions of Thika population clustered closely with those of the Meru population. The close genetic relationship of the Meru to the Thika population could be due to the accessions selected for the study. These accessions clustered as either *M. integrilfolia* or *M.* hybrids and none as *M. tetraphylla* similar to those of Thika (Ondabu et al., 1996). The Embu population consistently clustered distantly from the other populations, confirming genetic uniqueness.

Results of analysis of molecular variance indicated that there were genetic differences among the populations ($\Phi PT=0.043$) but that these differences were not significant ($P=0.056$; probability based on 9999 permutations) and the variation among the populations was only 4% compared to variation within the populations of 96%. The *F* statistical analysis generated θ (theta P) values ranging from -0.0927 to 0.6680 with an average θ (theta P) value of 0.1909 (standard deviation of 0.0194) across all loci. The non significance in genetic variation among the population supports the information that most

of the germplasm of the five populations originated from a common gene pool from Bob Harries population (Harries, 2004). However, the high theta value of 0.1909 indicates great genetic differentiation in to distinct genotypes. This differentiation is most likely a result of previous gene flow, and genetic drift and natural and pressure of selection by farmers (Hardner et al., 2005).

CONCLUSIONS AND RECOMMENDATIONS

The results of the current study indicated that AFLP markers were effective in elucidating the genetic diversity that exists within the selected macadamia populations. AFLP markers can help in identifying priority populations for sampling for maximum possible genetic variation based on allele richness or number of alleles per locus, evaluated as locally common alleles, that is, those that are common in one to several populations but not in the whole population and presence of unique alleles.

High genetic diversity was detected within the macadamia populations. This diversity was highest in the Bob Harries population that was initially used for propagation of planting materials. This population constitutes valuable germplasm that requires preservation. Genetic diversity of percentage of polymorphic alleles and heterozygosity was also high in the Embu population that also contained a private allele implying it also to be a unique germplasm block worth of conservation. These two populations contained accessions of both *M. integrilfolia* and *M. tetraphylla*, and the hybrids and hence suitable for conservation of the two species of macadamia in Kenya.

ACKNOWLEDGEMENTS

The Kenya Agricultural Research Institute provided financial and logistical support. Thanks also to the macadamia farmers for providing samples for analysis.

Literature Cited

Costello, G., Gregory, M. and Donaitu, P. 2008. Southern macadamia species recovery plan 2008-2012. Report to Department of the environment, water, heritage and the arts, Canberra by Horticulture Australia Limited, Sydney.

Excoffier, L., Smouse, P.E. and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA halotypes: application to human mitochondrial DNA restriction sites. Genet. 131:479-491.

Hardner, C.M., Scott, J., Williams, E., Mathews, K. and Levitt, C. 2005. Sub-sampling methods for nut and kernel traits. p.80-100. In: C.M. Hardner (ed.), Assessment methods for selection and management in macadamia. Horticulture Australia, Ltd., Sydney.

Harries, M. 2004. Involvement of the Harries family in the development of macadamia in Kenya 1946-1974. Proceedings of the macadamia stakeholders meeting. 15 June 2004. KARI, Nairobi, Kenya.

Huff, D.R., Peakall, R. and Smouse, P.E. 1993. RAPD variation within and among natural populations of outcrossing buffalograss *Buchloe dactyloides* (Nutt) Engelm. Theoret. & Applied Genet. 86:927-934.

Kiuru, P., Nyaga, A.N. and Wasilwa, L. 2004. A review of macadamia research in Kenya. Proceedings of the macadamia stakeholders meeting. 15 June 2004. KARI, Nairobi, Kenya.

Lanteri, S. and Barcaccia, G. 2006. Molecular marker based analysis for crop germplasm preservation. p.105-120. In: J. Ruane and A. Sonnino (eds.), The Role of Biotechnology in exploring and protecting agricultural genetic resources. FAO, Rome, Italy.

Maguire, T.L., Peakall, R. and Saenger, P. 2002. Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (*Avicenniaceae*) detected by AFLPs and SSRs. Theoretical and Applied Genetics 104:388-398.

Muthoka, N.M., Kiuru, P.D.N., Mbaka, J., Nyaga, A.N., Muriuki, S.J.N. and Waturu, C.N. 2008. Macadamia Nut Production and Research in Kenya. Afr. J. of Plant Sci. &

Biotechnol. 2:(1):46-48.
Nei, M. 1972. Genetic distance between populations. American Naturalist 106:283-392.
Nei, M. and Kumar, S. 2000. Mol. Evol. & Phylogenetics. Oxford University Press, New York.
Ondabu, N., Nyaga, A.N. and Tominaga, K. 1996. Macadamia clonal selection in Kenya. 5th Biennial KARI Scientific Conf., Nairobi, Kenya.
Ondabu, N.O., Lusike, A.W. and Watani, W.W. 2007. Adaptability of Kenyan Macadamia selections and Hawaiian varieties in four agro-ecological zones. Afri. Crop Sci. Conf. Proc. 8:391-394.
Peace, C., Ming, R., Schmidt, A., Manners, J. and Vithanage, V. 2008. Genomics of Macadamia, a recently domesticated tree nut crop. p.313-332. In: P.H. Moore, R. Ming and D.P. Delmer (eds.), Genomics of Tropical Crop Plants. Springer Verlag, New York.
Peace, C.P., Vithanage, V., Turnbull, C.G.N. and Carroll, B.J. 2003. A genetic map of macadamia based on randomly amplified DNA fingerprinting (RAF) markers. Euphytica 134(1):17-26.
Peakall, R. and Smouse, P. 2005. GenAlEx 6. Methods and Statistics. Appendix 1.
Saghai-Maroof, M.A., Soliman, K.M., Jorgenson, R. and Allard, R.W. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proceedings of Natl. Acad. of Sci. USA 81:8014-8081.
Sneath, I.H.A. and Sokal, R.R. 1973. Numerical taxonomy. W.H. Freeman and Company, San Francisco.
Vos, P., Hogers, R., Bleeker, M., Reijans, T., van de Lee, T., Hornes, M., Frifters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nuc. Acids Res. 23:4407-4414.
Wasilwa, L.A., Ondabu, N. and Watani, G. 2003. Adaptability of Kenyan macadamia selections and Hawaiian varieties in four agro ecological zones. A paper presented at the 1st KARI Adaptive Research Conference, 16-19 June 2003, Nairobi, Kenya.
Wilkie, J.H. 2008. Macadamia industry update. Paper presented at the 5th China International Foodstuff Exposition, 5-7 July 2008, Liuhua, China.

Tables

Table 1. Accession code, population and locations of 26 macadamia accessions used for molecular characterization.

Accession	Population	Longitude	Latitude
H1	Bob Harries	S01°02.589’	E036°59.455’
HWI	Bob Harries	S01°02.620’	E036°59.521’
KLM	Bob Harries	S01°02.624’	E036°59.539’
MYK	Bob Harries	S01°02.612’	E036°59.498’
EB1	Embu	S00°27.747’	E037°27.766’
EBH	Embu	S00°27.764’	E037°27.758’
EBT1	Embu	S00°27.725’	E037°27.777’
EBT2	Embu	S00°27.750’	E037°27.769’
EBT3	Embu	S00°27.761’	E037°27.767’
EBT4	Embu	S00°27.750’	E037°27.748’
K-15	Kirinyaga	S00°31.085’	E037°17.441’
K-3	Kirinyaga	S00°31.124’	E037°17.581’
K-5	Kirinyaga	S00°31.125’	E037°17.563’
KRT2	Kirinyaga	S00°33.019’	E037°14.219’
KRT3	Kirinyaga	S00°33.014’	E037°14.230’
MU-23	Meru	S00°04.274’	E037°36.446’
MU-24	Meru	S00°04.252’	E037°36.329’
MU-25	Meru	S00°04.461’	E037°36.321’
MU-27	Meru	S00°03.569’	E037°34.526’
MU-SM	Meru	S00°04.093’	E037°38.736’
H508	Thika	S01°00.002’	E037°03.650’
H660	Thika	S00°59.928’	E037°03.608’
KB-3	Thika	S00°59.787’	E037°03.619’
M-20	Thika	S00°59.983’	E037°03.640’
M-25	Thika	S00°59.792’	E037°03.506’
MT	Thika	S00°59.782’	E037°03.447’

Table 2. Pairwise population matrices for five macadamia populations based on Nei genetic distance and Nei Unbiased genetic distance after genotyping with AFLP markers.

	Nei Genetic Distance		Pairwise Population Matrix of Nei Genetic Distance		
	Bob Harries	Embu	Kirinyaga	Meru	Thika
Bob Harries	0.000				
Embu	0.107	0.000			
Kirinyaga	0.055	0.139	0.000		
Meru	0.117	0.099	0.090	0.000	
Thika	0.127	0.084	0.101	0.050	0.000

Figures

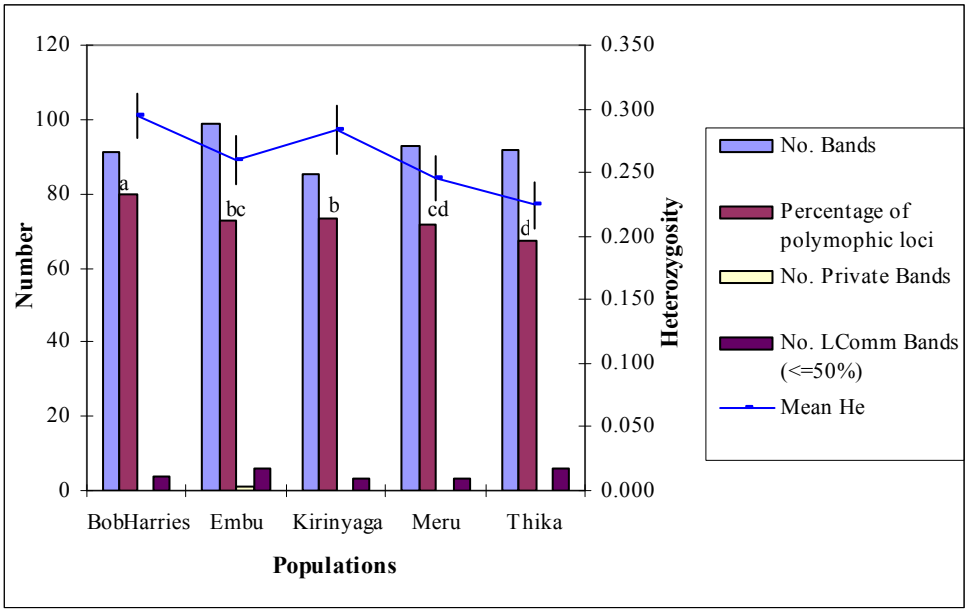

Fig. 1. Banding patterns across five macadamia populations genotyped with six AFLP primer combinations.

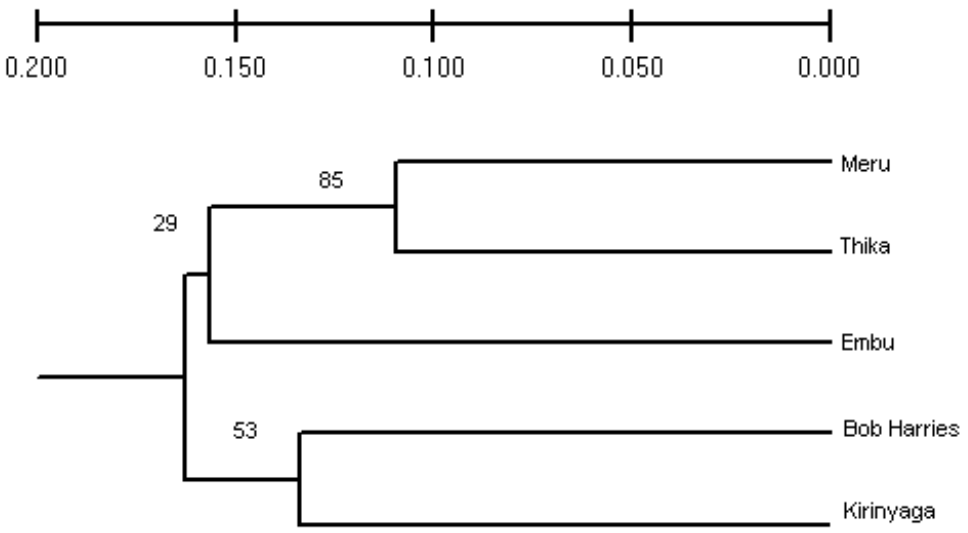

Fig. 2. Dendrogram showing genetic relationships between the five macadamia populations after genotyping with AFLP markers.

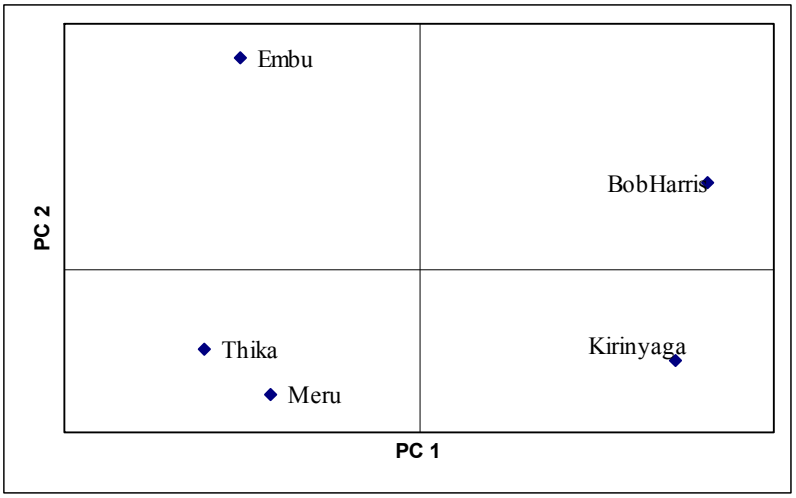

Fig. 3. Two dimensional graph representing the relationship between PC1 and PC2 using Nei's genetic distances of five macadamia populations genotyped with AFLP markers.