Screening Natural Oil Palm (*Elaeis guineensis* Jacq.) Populations Using SSR Markers

Claude Bakoumé[1], Ratnam Wickneswari[1], Nookiah Rajanaidu[2], Ahmad Kushairi[2], Norbert Billotte[3]

ABSTRACT

The assessment of genetic diversity and genetic structure of natural oil palm populations is carried out mostly for crop improvement purpose. A set of 16 microsatellite markers was used to genotype 494 palms from 49 populations belonging to ten African countries, three breeding and one semi-wild materials. The number of alleles per locus varied from 8 to 22, making a total of 209 alleles over the 16 microsatellite markers. The effective number varied from 1.1 in all populations from Madagascar to 4.7 in population 40 from Nigeria (mean = 3.30. Mean expected heterozygosity was high (0.644) ranging from 0.033 in population 3 from Madagascar to 0.803 in population 40 from Nigeria. The average genetic distance among oil palm populations was 0.684. An UPGMA cluster analysis showed three main clusters. One cluster grouped accessions from Madagascar while populations from Senegal, Guinea and Sierra Leone formed the second cluster. Populations from Ghana, Nigeria, Cameroon, Zaire, Angola, Tanzania and the breeding materials formed the third cluster. The grouping of Deli MPOB and Deli Dabou breeding materials into the same sub-cluster confirmed their common origin. The results acquired lead to redefining new strategies to make the most of the new genetic resources in oil palm improvement.

INTRODUCTION

The species *Elaeis guineensis* Jacq. is an allogamous monocot of the *Arecaceae* family (Hartley 1988). It is diploid (2n = 32 chromosomes), and cultivated for its fruits that contain oils in the mesocarp (palm oil) and in the kernel (palm kernel oil). It exists in wild, semi-wild states in equatorial tropics of Africa (Hartley, 1988).

The oil palm industry had developed from few palms from restricted provenances. The current planting materials are derived from Deli x La Mé and Deli x Eala crosses (Cochard et al. 2009). Deli *dura* provenance corresponds to 4 seedlings (suspected from the same palm) introduced from Africa. Its partner La Mé comes from Bingerville (15 km from Abidjan, Cote d'Ivoire) and Eala from Mbandaka (North-west of Democratic Republic of Congo).

[1] School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia
[2] MPOB (Malaysian Palm Oil Board), P.O. Box 10620, Kuala Lumpur, Malaysia
[3] Cirad (Centre de Coopération Internationale en Recherche Agronomique pour le Développement), UMR 1096, TA 80/03 Avenue Agropolis, 34398 Montpellier, Cedex 5, *France*

Despite the restricted provenances, the discovery of the single gene inheritance of shell presence (Beirnaert and Vanderweyen 1941) and that of the superiority of progenies from inter-origin crosses over those from the intra-origin crosses (Gascon & de Berchoux 1964) contributed to considerable oil yield improvement over the last 50 years. This genetic progress over several generations of selection was coupled with a restriction of the genetic base of the initial set of populations.

There is a general agreement on the need for broadening the narrow genetic base of current oil palm breeding materials for further improvement of yield, and also for ensuring conservation of a wide range of oil palm genetic resources. In the framework of acquisition of natural oil palm genetic resources, the Malaysian Palm Oil Board (MPOB), formerly known as the Palm Oil Research Institute of Malaysia (PORIM), carried out ten plant explorations from 1973 to 1996 in eleven countries in west and central Africa, the centre of distribution of *E. guineensis.* Large spaces, long time, and high observation and upkeep costs are constraints inherent to assessing the performances and the variation in collected natural palms in the field prior to their conservation or introduction to a breeding programme. The assessment of genetic variation using molecular markers allows conserving minimum number of natural oil palm collections that secure maximum amount of genetic diversity (Bakoume et al. 2007). In many places, considerable interest has been shown in rethinking oil palm breeding schemes once molecular markers have given a more understandable genetic structure of both breeding and natural collection materials using protein-based and DNA-based markers (Ghesquiere 1985, Barcelos 1998, Purba et al. 2000, Cochard et al. 2009). The amount of genetic diversity and the understanding of the genetic structure of natural oil palm collections in relation to breeding populations are guides to the introduction of new materials in the base populations of breeding programmes (Mayes et al. 2008, Cochard et al. 2009). Different molecular marker techniques have been used separately or in combination. At MPOB, different molecular marker techniques have been used to characterise the genetic diversity of natural *E. guineensis* Jacq. germplasm collection including RAPD (Shah et al. 1994), AFLP (Kularatne 2002), Isozymes (Hayati et al. 2004), RLFP (Maizura et al. 2006).

The present paper aims at presenting an assessment of genetic diversity and determining the genetic structure of MPOB's natural oil palm collections using 16 independent SSR markers. Simple-sequence repeats (SSRs), also called microsatellites, are tandem arrays of simple nucleotide motifs that are ubiquitous components of eukaryotic genomes (Delseny et al. 1983). They have co-dominant inheritance, present high degree of polymorphism, and show high reproducibility (Tautz 1993). Furthermore, multiplexing PCR reactions and/or multiloading PCR products on single gels reduce the work for studies requiring a large number of samples (Saghai Maroof et al. 1994).

MATERIALS AND METHODS

Plant materials and DNA preparation

Leaf samples were collected from unopened spear of 464 oil palms representing 45 natural populations originating from Senegal, Guinea and Sierra Leone, Ghana, Nigeria, Cameroon, Zaire, Angola, Tanzania, and Madagascar which were maintained in MPOB's field genebank at Kluang, Johor, Malaysia. Deli *dura* MPOB (Malaysia), Deli *dura* Dabou (Côte d'Ivoire), La Mé (Côte d'Ivoire) breeding materials and Bahia (Brazil) semi-wild material were included in order to compare their genetic diversity to that of oil palm trees from natural African groves. The genomic DNA was extracted and purified according to Doyle and Doyle (1990) CTAB DNA extraction procedure for small quantities of fresh leaf tissue. The DNA concentration of each sample was determined using agarose gel electrophoresis.

PCR amplification and scoring of banding patterns

The genomic DNA samples were amplified using a set of 16 SSR markers developed for oïl palm by the Centre de Coopération Internationale de Recherche Agronomique le pour le Développement (CIRAD) (France. PCR amplification of genomic DNA by the SSR primers was performed according to Billotte et al. (2001) in a Mastercycler® or Mastercycler® gradient thermocycler (Eppendorf AG, Hamburg, Germany). Amplicons were separated on a 7% denaturing polyacrylamide gel run at 250 V for 4 h in a Protean II xi cell (BIO-RAD, USA). PCR products of two to four SSR loci were mixed, and run per gel (multiple loading) for loci showing a reference allele size difference of 30 bp or more. The banding patterns were visualised by silver staining technique (Creste et al. 2001). Two genomic DNAs from Zaire's oil palm collections (ZA21, ZA29) were used as controls for each gel along with 2 lanes for 10 bp size ladder. ZA21, ZA29 served to link reads of the same locus across different gels.

Bands were scored based on their sizes. A counter scoring was carried out in two different laboratories (Forest Genetics laboratory of the Universiti Kebangsaan Malaysia and Biotrop laboratory of CIRAD-France), which validated the alleles detected and the subsequent genotype data. The alleles detected were designated alphabetically from the largest to the smallest size.

Data analysis

The programme POPGENE version 1.32 (Yeh and Boyle 1999) was used to estimate a) allele frequencies, b) effective number of alleles (A_e) (Crow and Kimura 1970), c)corrected expected (H_e) heterozygosities for small and unequal sample size (Nei 1978) and their variances. The approach of Marshall and Brown (1975) were adopted to compare the distribution of l common and rare allele frequency classes for populations studied. The generalised linear model (GLM) procedure in SAS Computer software (SAS Institute 1996) was used to perform One-way ANOVA for expected heterozygosity, to estimate differences between populations, and Duncan's Multiple Range Test (DMRT) to

compare all the natural oil palm populations and breeding materials based on their mean values of A_e and H_e (Glover and Mitchell 2002). All populations studied were grouped using the unweighted pair group method with arithmetic (UPGMA) clustering method described by Sneath and Sokal (1973). A dendogram was generated using the Sequential Agglomerative Hierarchical Nested cluster analysis (SAHN) in NTSYS package, version 2.0 (Rohlf 1990).

RESULTS

Genetic diversity

A total of 209 alleles were detected. The number of alleles per SSR locus ranged from 8 to 22 with an average of 13.1. From the classification of allele frequencies in two classes two rare alleles or alleles at frequency $p < 0.05$ were found at locus mEgCIR0173, one at locus mEgCIR0353 and one at locus mEgCIR0802 in all *E. guineensis* populations from Guinea Conakry, Senegal, and Sierra Leone. Rare alleles found in Deli MPOB breeding material were common in natural oil palm populations.

The effective number of alleles per locus (A_e) and the expected heterozygosity (H_e) are given in Table 1. A_e ranged from 1.1 ± 0.2 in populations from Madagascar to 4.7 ± 1.7 in population 40 from Nigeria (mean = 3.3 ± 1.3). H_e ranged from 0.033 ± 0.133 in population 3 from Madagascar to 0.803 ± 0.100 in population 40 from Nigeria (mean = 0.644 ± 0.163).

The analysis of variance for A_e and H_e over the 16 SSR loci revealed significant differences (Pr < 0.0001) between the 49 natural oil palm populations or breeding materials (Table 2). The Duncan's Multiple Range Test (DMRT) of mean values of H_e showed a clear separation between populations from Madagascar which were grouped together and the rest. The fact that different groups of means overlapped rendered the separation of the other populations not clear.

Genetic relatedness

The average genetic distance among the wild and semi-wild oil palm populations was 0.684. Total similitude was found between populations 1 and 3 from Madagascar and total divergence between populations from Madagascar and the rest of the populations from the other countries and the breeding materials. La Mé breeding material showed total divergence from both Deli breeding materials. Bahia semi-wild material was close to population 40 from Nigeria (D = 0.385). The dendogram showed three main clusters (Fig. 2). Populations from Madagascar formed one cluster. Populations from Senegal, Guinea and Sierra Leone formed the second cluster. Populations from Ghana, Nigeria, Cameroon, Zaire, Angola, Tanzania and Deli MPOB, Deli Dabou, La Mé breeding materials and Bahia semi-wild material formed the third cluster. There were situations in the second and third clusters where populations from the different countries grouped together into same sub-clusters.

Table 1 Estimated genetic diversity parameters of the natural accessions and breeding materials.

Country	Pop.	N	A_e	H_e	Country	Pop.	N	A_e	H_e
Senegal	5	14	3.6 (1.7)	0.685 (0.157)	Cameroon	16	10	2.9 (1.0)	0.650 (0.141)
	8	11	2.7 (1.3)	0.566 (0.225)		22	10	2.5 (1.1)	0.566 (0.216)
	13	11	2.4 (1.2)	0.526 (0.235)		28	10	2.8 (1.1)	0.627 (0.151)
Guinea-Conakry	1	10	3.6 (1.6)	0.701 (0.161)		29	10	3.9 (1.6)	0.727 (0.155)
	4	12	3.7 (1.3)	0.722 (0.137)		30	10	3.3 (1.2)	0.678 (0.180)
	12	11	3.5 (1.6)	0.688 (0.167)		31	10	3.3 (1.5)	0.644 (0.243)
Sierra Leone	1	11	3.2 (1.1)	0.672 (0.176)	Zaire	7	10	3.6 (1.5)	0.704 (0.161)
	5	11	3.3 (1.4)	0.665 (0.184)		22	10	3.5 (1.7)	0.698 (0.131)
	13	12	3.7 (1.2)	0.731 (0.112)		29	10	3.5 (1.5)	0.688 (0.209)
Ghana	3	11	3.3 (1.0)	0.686 (0.145)		30	10	3.6 (1.3)	0.724 (0.112)
	9	10	3.8 (1.5)	0.741 (0.118)		36	10	3.0 (1.0)	0.668 (0.130)
	13	10	3.9 (1.6)	0.729 (0.142)	Angola	1	10	3.7 (2.0)	0.690 (0.177)
Nigeria	12	11	3.7 (1.4)	0.736 (0.100)		3	10	3.3 (1.8)	0.649 (0.200)
	35	9	3.5 (1.2)	0.712 (0.151)		5	10	3.6 (1.5)	0.700 (0.180)
	40	9	4.7 (1.7)	0.803 (0.100)		6	10	2.9 (0.9)	0.636 (0.179)
	44	12	3.9 (1.2)	0.751 (0.089)		8	11	3.6 (1.2)	0.714 (0.146)
	45	12	4.2 (2.1)	0.739 (0.136)	Tanzania	1	10	3.9 (1.4)	0.734 (0.146)
Cameroon	1	10	2.8 (0.9)	0.637 (0.127)		3	10	4.4 (1.7)	0.770 (0.125)
	2	10	3.7 (1.6)	0.689 (0.213)		5	10	3.8 (1.4)	0.730 (0.139)
	9	11	3.9 (1.4)	0.739 (0.128)		7	11	3.9 (1.4)	0.745 (0.131)
	13	11	3.2 (1.4)	0.651 (0.203)		9	10	3.5 (1.4)	0.700 (0.155)
Madagascar	1	11	1.1 (0.2)	0.041 (0.126)	La Me	-	6	2.7 (1.0)	0.618 (0.219)
	3	5	1.1 (0.2)	0.033 (0.133)	Deli Dabou	-	9	2.5 (1.0)	0.549 (0.263)
	4	7	1.1 (0.2)	0.055 (0.150)	Bahia	-	4	2.8 (1.2)	0.668 (0.236)
Deli MPOB	1	11	2.7 (1.2)	0.559 (0.266)	Mean			3.3 (1.3)	0.644 (0.163)

Standard deviations are in parentheses. N: number of samples per population. A: mean number of alleles per locus. A_e: effective number of alleles per locus. H_o: observed heterozygosity. H_e: expected heterozygosity.

Table 2 Analysis of variance for effective number of alleles per locus (A_e) and expected heterozygosity (H_e) in the 49 natural oil palm populations and breeding materials

		A_e			H_e	
Source	**DF**	**MS**	**F value**	**DF**	**MS**	**F value**
Model (Populations over all loci)	63	33.43	17.89**	63	0.43	20.40**
Residual	720	1.87		720	0.02	

DF: degree of freedom. SS: sum of squares. MS: mean square. F: calculated parameter of Fisher. ** significant difference at 1%.

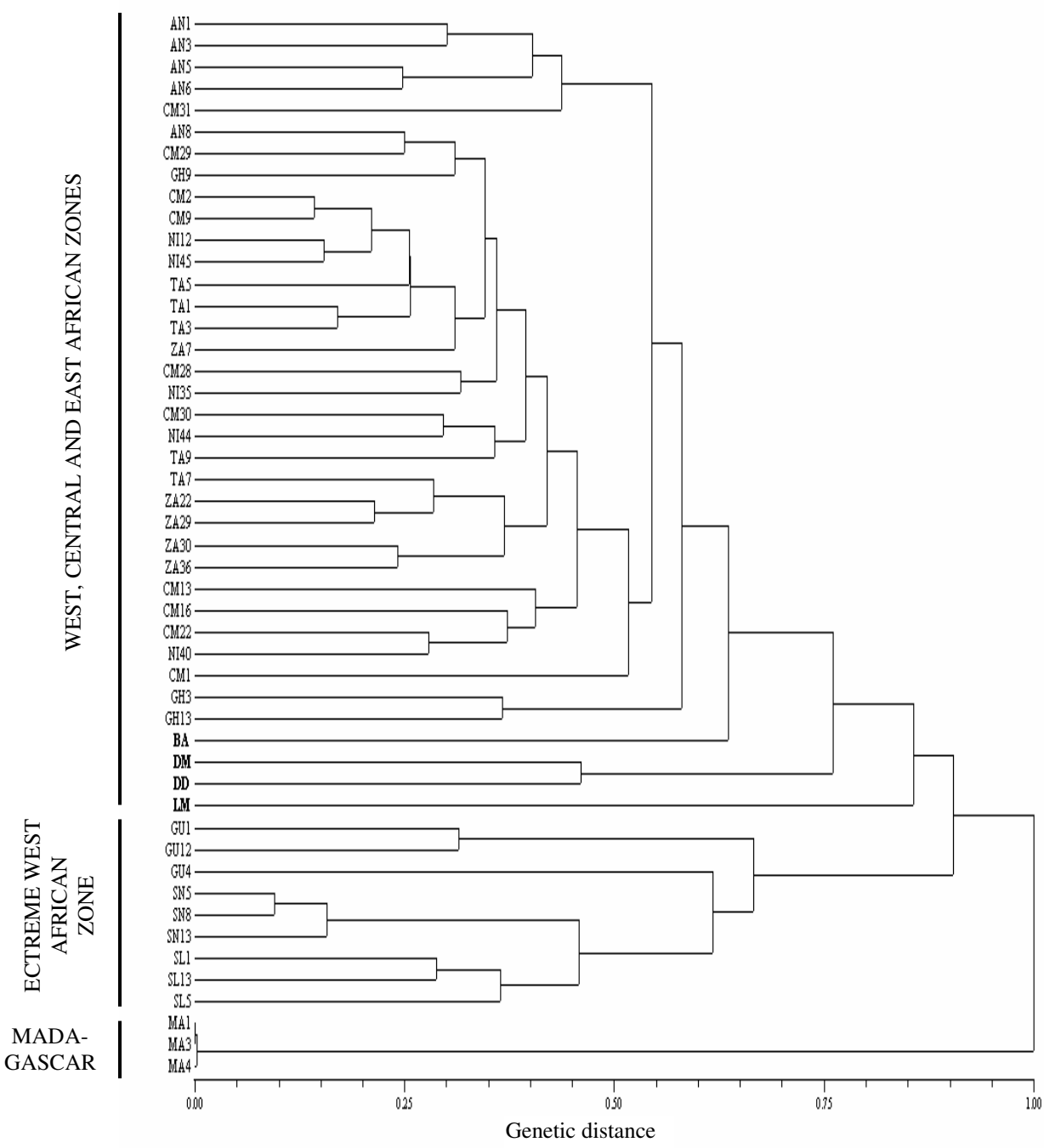

Figure 1: UPGMA clustering of 49 oil palm populations based on the genetic distance of Nei (1978)

DISCUSSION

Genetic diversity

Allelic frequencies at each locus varied from population to population and some alleles occurred only in one or some populations. In an out-crossing plant like oil palm where random mating is expected, genetic drift and reproductive isolation are the most common factors that affect allele frequencies.

The presence of some alleles only in populations from Senegal, Guinea and Sierra Leone, countries with prevailing low rainfall and dry weather may be related to adaptive traits. The presence of the same rare alleles ($p < 0.05$) across populations from different countries with similar dry weather conditions indicates that their existence maybe also due to selection. Zeng et al. (2004) considered unique alleles (or alleles not detected in commercial cultivars) in breeding lines and land races of rice were related to their adaptation to saline soils. Rare alleles in Deli MPOB, which are common in natural oil palm populations, denoted also a reduction in allelic diversity over several generations of selection.

The expected heterozygosity of 0.644 was similar to 0.68 reported as high by Billotte et al. (2001) for *E. guineensis*. The high genetic diversity observed might have resulted from the out-crossing behaviour of oil palm as earlier reported on *Quercus petraea* (Cottrell et al. 2003).It can explain the oil palm's plasticity as regard to adaptation to various environments and to its actual large distribution area. H_e value in Deli materials represented only 82% of that of La Mé because the latter had undergone fewer generations of selection compared to Deli materials. The lowest genetic diversity found in populations from Madagascar are congruent with previous results reported by Rajanaidu et al. (2000). Random genetic drift and inbreeding have acted over a sufficient number of generations in these small populations.

Genetic relatedness

The average genetic distance (Nei 1978) among the wild and semi-wild oil palm populations was rather higher than that revealed by isozymes (D = 0.113 Hayati et al. 2004). The low genetic distance between Deli materials as a whole, La Mé and Bahia were derived from different and distant natural populations. The proximity of Bahia semi-wild material to population 40 from Nigeria is congruent with its south-eastern Nigeria origin which dates to the period of slave trade (Hartley 1988). The dendogram showed one cluster formed by populations from West Africa (Ghana, Nigeria), Central Africa (Cameroon, Zaire, Angola), and East Africa (Tanzania). This may indicate that the oil palms from these three geographical zones might have derived from one or similar genetic background(s). Deli MPOB and Deli Dabou breeding materials clustered together supporting their common origin. Materials from Madagascar have already shown morphological different notably the fused filaments of the staminal tube and anthers

which were erected in the male flower, a smaller and rounded fruits surrounded by larger bracts, and a larger trunk. .

CONCLUSION

The present study has provided information on the amount of genetic diversity of populations and genetic relatedness among natural oil palm populations from different African countries High genetic diversity found in the MPOB's oil palm germplasm collections implies a high amount of additive genetic variance useful for progress in oil palm breeding. The structure of the genetic diversity allows thinking new strategies for introgression of breeding populations with genetic resources of germplasm collections.

ACKNOWLEDGEMENTS

The present study was carried out in the Forest Genetics Laboratory of Universiti Kebangsaan Malaysia (UKM) with the joint support from the Malaysian Palm Oil Board (MPOB) in Malaysia, the Centre de Coopération Internationale de Recherche Agronomique pour le Développement (Cirad) in France and the Institute of Agricultural Research for Development (IRAD) in Cameroon.

REFERENCES

Barcelos E, Amblard P, Berthaud J, Seguin M (2000) Genetic diversity and relationship in American and African oil palm revealed by RFLP and AFLP molecular markers. Pesquisa Agropecuaria Brazilia 37(8):1105-1114.

Beirnaert A and Vanderweyen R (1941) Contribution à l'étude génétique et biométrique des variétés d'*Elaeis guineensis* Jacq. Pub. de l'INEAC. Série Scientifique 27.

Billotte N, Risterucci AM, Barcelos E, Noyer L, Amblard P, Baurens FC (2001) Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome* 44: 413-425.

Cottrell JE, Munro RC, Tabbener HE, Milner AD, Forrest GI, Lowe AJ (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. *For Ecol Manag* 176: 287-303.

Creste S, Tulmann Neto, Figueira A (2001) Detection of single sequence repeat polymorphism in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol Biol Rep* 19: 200-306.

Crow JF, Kimura M (1970) An introduction to population genetic theory. Harper and Row, New York.

Cochard B, Adon B, Rekima S, Billotte N, Desmier de Chenon R, Koutou A, Nouy B, Omoré A, Purba AR, Glazsmann JC, Noyer JL (2009) Geographic and genetic structure of African oil palm diversity suggests new approaches to breeding. *Tree Gentics & Genomes*. Doi: 10-1007/s11295-009-0203-3, 1 May 2009.

Delseny M, Laroche M, Penon P (1983) Detection of sequences with Z-DNA forming potential in higher plants. Biochem Biophys Res Commun 116: 113–120.

Doyle J, Doyle L (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Gascon JP. De Berchoux C. (1964) Caracterisation de la production d'*elaeis guineensis* Jacq. De diverses origins et de leurs croisements. *Oleagineux* 2: 75.

Ghesquiere M (1985) Enzyme polymorphism in oil palm (*Elaeis guineensis* Jacq.). II. Variability and genetic of seven origins of oil palm. *Oleagineux* 40: 529-540

Glover T, Mitchell K (2002) An introduction to biostatistics. 1st edn. The McGraw-Hill Companies, New-York.

Hartley CWS (1988) The oil palm, 2nd edn. Longman, London.

Hayati A, Wickneswari R, Maizura I, Rajanaidu N (2004) Genetic diversity of oil palm (*Elaeis guineensis* Jacq.) germplasm collections from Africa: implications for improvement and conservation of genetic resources. *Theor Appl Genet* 108: 1274-1284.

Kularatne RS (2000) Assessment of genetic diversity in natural oil palm (*Elaeis guineensis* Jacq.) using amplified fragment length polymorphism markers. PhD thesis, Universiti Kebangsaan Malaysia, Bangi.

Maizura I, Rajanaidu N, Zakri AH, Cheah SC (2006) Assessment of genetic diversity in oil palm (*Elaeis guineensis* Jacq.) using restriction fragment length polymorphism (RFLP). *Genetic Resources and Crop Evolution* 53:187-1295.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Purba AR, Noyer JL, Baudouin L, XPerrier X, Hamon S, Lagoda PJL (2000) A new aspect of genetic diversity of Indonesian oil palm (*Elaeis guineensis* Jacq.) revealed by isoenzyme and AFLP markers and its consequences for breeding. *Theor Appl Genet* 101: 956-961.

Rajanaidu N, Maizura I, Cheah SC (2000) Screening of oil palm natural populations using RAPD and RFLP molecular markers. In: Rajanaidu, N, Ariffin D (eds) Proceedings of International Symposium on Oil Palm Genetic Resources and Utilization, Kuala Lumpur, pp AA1-AA28.

Rohlf FJ (1990) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis, Version 2.0. State University of New York.

Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinary poplymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. *Proc Natl Acad Sci U.S.A*. 91: 5466-5470.

SAS Institute (1996) SAS procedures guide. Release 6.12 Ed SAS Institute Inc., Cary, NC.

Shah, F.H., Rashid, O., Simons, A.J. & Dunsdon, A. 1994. The utility of RAPD markers for the determination of genetic variation in oil palm (*Elaeis guineensis*). *Theor Appl Genet* **89**: 713-718.

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

Tautz D (1993) Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys JA (eds) DNA Fingerprinting: State of the Science, Birkhauser Verlag, Basel, pp 21-28.

Yeh FC, Boyle T (1999) POPGENE Version 1.32. The user-friendly software for population genetic analysis. University of Alberta and CIFOR, Calgari, Alta.

Zeng L, Kwon TR, Liu X et al (2004) Genetic diversity analyzed by microsatellite markers among rice (*Oryza sativa* L.) genotypes with different adaptations to saline soils. *Plant Science* 116: 1275-1285.