Journal of Oil Palm Research Vol. 24 August 2012 p. 1368-1377

EVALUATION OF MPOB OIL PALM GERMPLASM (*Elaeis guineensis*) POPULATIONS USING EST- SSR

ZULKIFLI, Y*; MAIZURA, I* and RAJINDER, S*

ABSTRACT

MPOB has initiated an oil palm genetic resources programme which includes the collection, evaluation, utilisation and conservation of oil palm germplasm. The objectives of this study were to determine the potential use of expressed sequence tag-simple sequence repeats (EST-SSR) markers to study the genetic variability of the germplasm collection. A total of 330 palms originating from 11 countries in Africa were screened using 10 EST-SSR primers. The data were analysed using the Biosys-1 software to calculate the genetic variability parameters. It was found that the germplasm exhibited a high level of genetic diversity. Most of the loci tested were 100% polymorphic at 0.95 criterion. A total of 46 alleles were detected across all the germplasm populations. Of these, three were considered as rare alleles. The Nigerian germplasm showed the highest number of alleles per locus and the highest number of rare alleles, a high percentage of polymorphic loci and high heterozygosity, suggesting that Nigeria could be the centre of diversity of the wild oil palm. This study also revealed that the Madagascar germplasm is unique and different compared with the oil palm populations from the African mainland. Based on the dendrogram constructed, the germplasm populations could be divided into three major clusters; Cluster 1 consisting of Angola, Tanzania, Cameroon, Nigeria, Democratic Republic of Congo (formerly known as Zaire), Sierra Leone, Guinea and Ghana germplasm, Cluster 2 consisting of Gambia and Senegal germplasm, while the Madagascar germplasm was placed in Cluster 3. The mean genetic distance across the MPOB germplasm populations was 0.251.

Keywords: oil palm, germplasm, EST-SSR, genetic variability.

Date received: 24 December 2010; **Sent for revision:** 10 February 2011; **Received in final form:** 17 February 2012; **Accepted:** 15 May 2012.

INTRODUCTION

Malaysia currently has the largest collection of oil palm germplasm in the world. Since 1973, Malaysia has collected *Elaeis guineensis* germplasm from 11 countries and *E. oleifera* germplasm from seven countries (Rajanaidu, 1994a; Rajanaidu and Jalani, 1994).

The objectives of this collection are to broaden the genetic base of the current oil palm breeding materials, and to ensure the conservation of a wide range of oil palm genetic resources for posterity. At present, the *E. guineensis* and *E. oleifera* genetic materials and other palms from Africa and Central-South America are maintained in the form of a field genebank to safeguard the long-term interest of the Malaysian oil palm industry. For conservation purposes, the major issues with respect to the oil palm field genebank are the large space needed and high maintenance cost.

Studies, using molecular markers, on the genetic diversity in the MPOB oil palm germplasm are important for the establishment of a core collection (*i.e.* germplasm with minimal repetitive genetic and maximal genetic diversity). The estimated genetic distance determined using molecular marker

* Malaysian Palm Oil Board,
P. O. Box 10620, 50720 Kuala Lumpur,
Malaysia.
E-mail: zulkifly@mpob.gov.my

data may prove useful in developing appropriate sampling strategies for conservation and also strategies for future collection expeditions.

Simple sequence repeats (SSR, also known as microsatellites) are regions of DNA consisting of short, tandemly repeated units (1-6 bp in length) found within the coding or non-coding regions of all eukaryotic organisms. The polymorphic level of these markers depends on the variation in the number of repeat units. SSR are co-dominantly inherited, multi-allelic, reproducible, have good genome coverage and are relatively abundant. SSR markers are widely used for the purposes of DNA fingerprinting, paternity testing, linkage map construction and population genetic studies. The first reports of microsatellites in plants were made by Condit and Hubbel (1991) and Gupta *et al.* (1999) suggesting their abundance in plant systems.

Genomic SSR markers, developed from genomic DNA libraries, can correspond to either the transcribed region or the non-transcribed region of the genome, and thus are not very useful in identifying candidate genes. In addition, genomic SSR are frequently not transferable to closely related species (Roder *et al*., 1998; Sourdille *et al*., 2001). In contrast, expressed sequence tag (EST) – SSR markers, which are sequenced portions of complementary DNA copies of mRNA, represent part of the transcribed region of the genome. The conserved nature of the transcribed region may limit their polymorphism; thus, EST-SSR have been reported to be less polymorphic compared with genomic SSR in crop plants (Rajeev *et al*., 2005). However, they have some advantages over genomic SSR, especially in genetic diversity studies of germplasm collections because of their higher level of transferability to related species. They can often be used as anchor markers for comparative mapping and evolutionary studies (Rajeev *et al*., 2005). These markers enable the variation in expressed sequences to be assayed and often have putative functions. These gene-targeted markers have the potential of representing functional markers, which can enhance the role of genetic markers in germplasm evaluation. In several plant species including oil palm (Singh *et al*., 2008), EST-SSR have been identified, developed and used in a variety of studies. They have proven useful in plant systems such as grape (Scott *et al*., 2000), sugar-cane (Cordeiro *et al*., 2001), rye (Hackauf and Wehling, 2002), barley (Thiel *et al*., 2003), wheat (Nicot *et al*., 2004) and coffee (Poncet *et al*., 2006).

The objectives of this project was to study the ability of oil palm-derived EST-SSR in discriminating between the various oil palm germplasm populations, and to estimate the genetic variability parameters obtained using this marker system.

MATERIALS AND METHODS

Samples

EST-SSR analysis was performed on germplasm from 11 countries, as listed below:

a) Cameroon
b) Ghana
c) Democratic Republic of Congo (formerly known as Zaire)
d) Tanzania
e) Madagascar
f) Gambia
g) Guinea
h) Nigeria
i) Senegal
j) Sierra Leone
k) Angola

Thirty palms were randomly selected from the different populations ranging from 3-11 per country. Selection of the samples covered the entire collected region (north, south, *etc*.) of the countries above. Eight to 12 samples had been selected from the different populations for each region. Spear leaves were harvested from each sampled palm for DNA extraction.

DNA Extraction

DNA extraction was carried out using the modified CTAB method (Dellaporta *et al*., 1983). Two grammes of fresh leaf tissue was ground to a fine powder in the presence of liquid nitrogen using a mortar and pestle. The powder was then transferred into a 30-ml Falcon tube, mixed with 20 ml CTAB buffer and incubated at 60°C for 30 min. An equal volume of chloroform/isomyl alcohol (24:1 v/v) was added, and the mixture centrifuged at 10 000 rpm for 15 min. The supernatant was pipetted into sterile tubes, and two volumes of isopropanol were added to precipitate DNA. DNA was pelleted at 12 000 rpm for 15 min, and the pellet was washed twice with 70% ethanol containing 10 mM ammonium acetate, before dissolving it in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). DNA was then treated with RNAse (50 μg ml^{-1}) at 37°C for 20 min. One-tenth volume of 7.5 M ammonium acetate and two volumes of absolute ethanol were added to precipitate DNA. DNA was spun down at 12 000 rpm for 15 min. The resultant DNA pellet was washed twice with 70% ethanol, and dissolved in 1 ml of TE buffer.

DNA Quantification and Digestibility Testing

The DNA concentration was determined using a spectrophotometer by measuring absorbance at λ = 260.0, 280.0 and 350.0 nm. The optical density

(OD) ratio at 260/280 was calculated to determine the DNA purity, and OD values at λ=260.0 and 350.0 nm were used to calculate the DNA concentration. A digestibility test was carried out using *Eco*RI (six base pair cutter) and *Hae*III (four base pair cutter).

SSR Analysis

The EST-SSR markers used in this study were developed at MPOB. They were CNH 00938, CNH 0887, CNH 1617, CNI 01937, EAP 03160, EO 02978, MF233-033, CNIP 00421, MF233-056 and CNI 01733. In primer labelling, γ-^{33}p was used to detect amplified fragments. The forward primer was 5′ end-labelled at 37°C for 1 hr and 30 min using T4 polynucleotide kinase. The labelling reactions contained 15 pmoles of primer, 0.1 μl of γ-^{33}p and 1 U of T4 polynucleotide kinase in a total volume of 1.0 μl. The PCR reaction consisted of 1 U of *Taq* polymerase, 1.5 mM $MgCl_2$, 0.2 μM of each primer, 0.2 mM dNTPs and 10X PCR buffer. An aliquot of DNA was diluted to a concentration of 50 ng μl^{-1} and used in the SSR analysis.

PCR Conditions

The following PCR programme was used: pre-denaturation at 95°C for 1 min, denaturation at 95°C for 30 s, annealing at 52°C (except for CNH 1617 primer, which used 54°C) for 30 s, and extension at 72°C for 30 s. This programme was repeated for 35 cycles, followed by a final extension at 72°C for 5 min.

Analysis on Acrylamide Gel

The PCR product was mixed with an equal volume (10 μl) of formamide dye (0.3% bromophenol blue), 0.3% xylene cyanol; 10 mM diaminoethanetetra-acteic acid (EDTA); pH 8.0; 97.5% deionised formamide. Prior to loading and electrophoresis, the samples were heated for 3 min at 95°C and then rapidly cooled on ice. The gel was prepared by mixing 100 ml of 6% polyacrylamide (20:1 acrylamide:bisacrylamide) containing 7.5 M urea and 1 M tris-borate-EDTA (TBE). Then 100 μl of 10% ammonium persulphate and 20 μl of tetramethylethylenediamine (TEMED) was also added. The 5 μl of the heat-treated samples was electrophoresed at a constant power of 1600 V for 3 hr. The gels were then vacuum-dried for 1 hr, and exposed against X-ray film for three to four days at -80°C.

Data Analysis

Amplified DNA fragments were scored, and each band within each locus was identified in alphabetical order. The most anodal migrating band was designated as allele A, and the next band was allele B. The data were scored and analysed using the Biosys1 software to calculate the genetic variability parameters such as mean sample size per locus, mean number of alleles per locus, percentage of polymorphic loci, observed and expected heterozygosities (Nei, 1978). The genetic distance values were calculated as described by Nei's (1972), and subsequently used to construct the dendrogram.

TABLE 1. NUMBER OF ALLELES PRODUCED BY SIMPLE SEQUENCE REPEATS (SSR) PRIMERS USED IN THE STUDY

No.	Primer	No. of alleles produced	Name of alleles
1	EAP03160	2	A, B
2	EO02978	2	A, B
3	CNH00887	4	A, B, C, D
4	CNH01617	5	A, B, C, D, E
5	MF233033	6	A, B, C, D, E, F
6	MF233056	7	A, B, C, D, E, F, G
7	CNI01733	3	A, B, C
8	CNI01937	3	A, B, C
9	CNIP00421	7	A, B, C, D, E, F, G
10	CNH00938	7	A, B, C, D, E, F, G
	Total	46	

RESULTS AND DISCUSSION

Common and Rare Alleles

Forty-six bands were generated from the 10 primers used in the study. The highest number was seven alleles and the lowest two alleles. The results are presented in *Table 1*.

According to Brown (1978), alleles can be classified as either common (allele frequency more than 0.100 for at least one sample) or rare (never exceeding a frequency of more than 0.100). Common alleles are divided into three classes, namely, widespread (occurrence in more than two regions), sporadic (occurrence in two regions), or localised (occurrence in only one region). Rare alleles can be classified as either widespread (occurrence in more than one region) or localised (occurrence in only one region). Results of this screening show that of the 46 alleles, 43 were common, and of these 40 (87.0%) were common-widespread, one (2.2%) was common-sporadic, and two (4.3%) were common-localised. Of the three rare alleles detected, two were rare-widespread and one was rare-localised. *Table 2* summarises the occurrence of rare alleles in the oil palm germplasm populations.

TABLE 2. RARE ALLELES DETECTED IN THE OIL PALM GERMPLASM POPULATIONS

Germplasm population	Rare allele/locus		
	Allele A/CNH00887	Allele B/CNH00938	Allele G/CNH00938
Nigeria	√	√	√
Ghana	√	√	-
Guinea	√	√	-
Sierra Leone	√	√	-
Tanzania	√	√	-
Congo	√	√	-
Angola	√	-	-
Cameroon	√	-	-
Gambia	-	√	-
Senegal	-	√	-

Note: √ indicates presence of allele.

The three rare alleles detected were loci CNH00887 (allele A) and CNH00938 (alleles B and G). These rare alleles occurred in all the germplasm populations except for the population from Madagascar. Allele A/CNH00887 was detected in eight germplasm populations but was absent in the populations originating from the Gambia, Senegal and Madagascar germplasm. Allele B/CNH00938 also occurred in eight germplasm populations with the exceptions of the Angola, Cameroon and Madagascar populations. The G/CNH00938 was found only in the Nigeria germplasm. The Nigeria germplasm possessed all the rare alleles detected in the study.

Genetic Variability

Table 3 summarises the estimates of genetic variability parameters for the germplasm analysed in this study. All the loci were 100% polymorphic (0.95 criterion) except for the Madagascar and Gambia samples. The mean percentage of polymorphism was high (96.4%). The mean number of alleles per locus (A) was 3.7, ranging from 2.7 (Gambia) to 4.2 (Nigeria). The mean expected heterozygosity (H_e) was 0.503, ranging from 0.363 (Gambia) to 0.582 (Tanzania). The mean observed heterozygosity (H_0) was 0.398, with values ranging from 0.162 (Madagascar) to 0.513 (Angola). The highest heterozygosity (H_e) was observed in the Tanzania germplasm (H_e = 0.582), followed by the Congo (H_e = 0.574) and Nigeria (H_e = 0.569). The lowest heterozygosity was observed in populations from Gambia (H_e = 0.363), followed by Madagascar (H_e = 0.372) and Ghana (H_e = 0.480). The heterozygosity revealed by the microsatellite technique in this study was higher compared with that reported in previous studies using RFLP (H_e = 0.135, Barcelos, 1998; H_e = 0.195, Maizura *et al*., 2000), RAPD (H_0 = 0.38, Shah *et al*., 1994; H_0 = 0.38, Rajanaidu *et al*., 2000), isoenzymes (H_e = 0.41, Ghesquiere *et al*., 1987; H_e = 0.380, Purba, 2000; H_e = 0.184, Hayati *et al*., 2004) and AFLPs (H_e = 0.117, Barcelos, 1998; H_e = 0.493, Kularatne, 2000).

The level of polymorphism (P) revealed by the EST-SSR was high (100%) for most of the germplasm except for those from Madagascar and Gambia. This result is in agreement with that

TABLE 3. ESTIMATES OF GENETIC VARIABILITY PARAMETERS FOR 11 OIL PALM GERMPLASM POPULATIONS

Germplasm	Mean sample size per locus*	Mean No. of allele per locus	Percentage of loci polymorphic**	Mean heterozygosity		Chi-square test, X^2 (H_o *vs* H_e)***
				Direct-count (H_o)	Hardy-Weinberg expected (H_e)	
Nigeria	29.5	4.2	100.0	0.490	0.569	0.011
Cameroon	30.0	3.8	100.0	0.503	0.556	0.005
Congo	30.0	3.8	100.0	0.477	0.574	0.016
Tanzania	30.0	3.9	100.0	0.460	0.582	0.026
Madagascar	29.6	3.7	90.0	0.162	0.372	0.119
Angola	30.0	3.5	100.0	0.513	0.537	0.001
Senegal	30.0	3.6	100.0	0.387	0.495	0.024
Gambia	30.0	2.7	70.0	0.327	0.363	0.004
Sierra Leone	30.0	4.0	100.0	0.343	0.490	0.044
Guinea	29.8	4.0	100.0	0.341	0.515	0.059
Ghana	29.5	3.8	100.0	0.376	0.480	0.023
Mean	29.85	3.7	96.4	0.398	0.503	

Note: * Decimal values mean that there were samples that did not amplify for a particular locus.
** A locus is considered polymorphic if the frequency of the most common allele does not exceed 0.95.
*** Values are considered to be not significantly different if $X^2 < 3.841$ at α=0.05.

observed using RFLP analysis (Maizura *et al*., 2006), where the lowest level of polymorphism was also reported in populations from Madagascar and Gambia. As expected, the EST-SSR markers showed a higher level of polymorphism compared with that determined using RAPD (P = 88.6%) (Rajanaidu *et al*., 2000), AFLP with P = 93.4 (Kularatne, 2000) and 61% (Purba *et al*., 2000), isozymes (P = 54.5%) (Hayati, 2002) and RFLP (P = 56.3%) (Maizura *et al*., 2006). Screening of oil palm germplasm using genomic SSR markers also indicated a high level of polymorphism among the germplasm populations, with P = 93.1% (Bakoume, 2006). In general, the oil palm germplasm populations showed high values of A and P, indicating that they have high genetic variability and can potentially be used in crop improvement programmes.

The populations from Tanzania, Congo, Nigeria, Cameroon and Angola exhibited high levels of heterozygosity. Heterozygosity refers to the presence of different alleles at one or more loci on homologous chromosomes. The measure of the level of heterozygosity across loci can be used as an indicator of the amount of genetic variability. These populations may possess unique genes associated with traits which may not be found in the current commercial oil palm planting materials. The current germplasm populations have indeed proven to be useful in developing new and improved planting materials. For example, the planting materials known as PS1 (dwarf with high yield) (Rajanaidu, 1994b), PS2 (high iodine value with high yield) (Rajanaidu, 1994c) and PS3 (high kernel content) (Rajanaidu, 1996) were developed from selected palms from Nigeria. Selected palms from Tanzania were introduced as thin-shelled *tenera* (PS5) (Kushairi *et al*., 2003a) and in high bunch index breeding populations (PS7) (Junaidah *et al*., 2004). The Angolan palms with large fruit *dura* were selected as PS6 (Kushairi *et al*., 2003b), while individuals from the Tanzania and Angola populations are being used in the long stalk breeding programme, namely as PS10 (Noh *et al*., 2005). Besides that, materials from Tanzania also show high carotene content (Mohd Din *et al*., 2006). Progenies created from Cameroon x Congo crosses show tolerance to *Ganoderma* (Idris *et al*., 2004), and high vitamin E content is observed in materials from Congo, Tanzania, Angola, Cameroon and Nigeria (Kushairi *et al*., 2004). Generally, the high heterozygosity obtained in this study as well as that reported by Bakoume (2006) support the adaptability of oil palm germplasm which was analysed in various environments, and its large distribution area.

The chi-square test was employed to determine the difference between the values of the observed heterozygosity and the Hardy-Weinberg heterozygosity as shown in *Table 3*. The observed heterozygosity values for all the germplasm showed no significant difference from the Hardy-Weinberg expectation (in H-W equilibrium), which means that the genetic variation in the oil palm germplasm had remained constant.

The oil palm populations from Gambia, Madagascar and Ghana exhibited low heterozygosity. Generally, the populations near the centre of a species range are contiguous and genetically diverse, whereas marginal populations are isolated, smaller and less genetically variable. In case of Madagascar, which is isolated from mainland Africa, the limited gene flow among the populations may have contributed to their low heterozygosity. Besides that, only a small number of populations were collected during the expeditions, and these populations probably only carried a part of the total alleles present in Madagascar, which would explain the low heterozygosity observed. Maizura *et al.* (2000), Rajanaidu *et al*. (2000) and Hayati *et al*. (2004) also reported low genetic heterozygosity for the Madagascar germplasm. However, Kularatne (2000) reported high heterozygosity in the populations from Madagascar. The dissimilarity observed and reported could be due to the different marker techniques employed. However, there are two possible mechanisms operating in small and isolated populations that could influence the level of heterozygosity in the Gambia and Madagascar germplasm. In small populations, genetic drift may occur whereby allele frequency changes more rapidly than in large populations. In small populations, there are a limited number of individuals that represent the alleles in the gene pool. The absence of some of these individuals will therefore directly affect the allele frequencies. The limited number of individuals in a small population may also increase the tendency for inbreeding (mating among related individuals), resulting in low heterozygosity. In the current study, the population from Ghana also showed low heterozygosity, although the country is situated at central and western parts of Africa. Similar results were reported by Kularatne (2000) and Hayati *et al.* (2004). As elaborated by Hartley (1988) and mentioned by Maizura *et al.* (2006), the human activities of collecting alcoholic beverage from the oil palm had caused a tremendous amount of felling of the palms in the wild. In addition, the establishment of cocoa plantations had also caused massive forest clearing and the wiping out wild oil palms.

The Nigeria oil palm germplasm exhibited a high level of polymorphism, high H_e, highest mean number of alleles per locus (A) as well as number of rare alleles, suggesting that this area may likely be the centre of diversity for oil palm. This study may not be enough to make such a conclusion, but similar results have been reported by Hayati (2002)

and Maizura *et al*. (2006). The finding of fossils similar to that of oil palm pollen from the Miocene period in the Niger delta (Zeven, 1964; Rees, 1965) further supports the hypothesis. In addition, Nigeria is located in central Africa and may have experienced massive gene flow from palms in the neighbouring countries, which would explain the high level of genetic diversity observed.

Genetic Relatedness among the Germplasm Populations

The genetic distance matrix among the populations is shown in *Table 4*. Clustering of the oil palm germplasm populations is presented in *Figure 1*. The mean genetic distance among the MPOB germplasm was 0.251. The lowest genetic distance was observed between the Tanzania and Congo germplasm (0.026). A similar observation had been reported by Hayati (2002). This could be due to the fact that Tanzania and Congo are neighbouring countries. Madagascar and Guinea which are separated by geographical distance showed the highest genetic distance (0.711).

The dendrogram constructed using the Biosys1 software revealed three genetic groups (*Figure 1*). The largest cluster, Cluster 1, consists of the germplasm from Angola, Tanzania, Congo, Cameroon, Nigeria, Sierra Leone, Guinea and Ghana. In Cluster 1, there are four sub-clusters as follows:

- sub-cluster 1 – Angola, Tanzania and Congo germplasm;
- sub-cluster 2 – Nigeria and Cameroon germplasm;

TABLE 4. MATRIX OF GENETIC DISTANCE (Nei, 1972) AMONG MPOB GERMPLASM POPULATIONS

Population	1	2	3	4	5	6	7	8	9	10	11
1	Angola	***									
2	Cameroon	0.119	***								
3	Gambia	0.290	0.357	***							
4	Ghana	0.220	0.221	0.299	***						
5	Guinea	0.196	0.192	0.381	0.171	***					
6	Madagascar	0.399	0.503	0.470	0.544	0.711	***				
7	Nigeria	0.164	0.065	0.308	0.163	0.161	0.425	***			
8	Senegal	0.252	0.321	0.171	0.199	0.256	0.452	0.254	***		
9	S.Leone	0.195	0.134	0.277	0.253	0.112	0.597	0.112	0.304	***	
10	Tanzania	0.075	0.156	0.240	0.218	0.189	0.381	0.170	0.223	0.165	***
11	Congo	0.067	0.117	0.196	0.183	0.184	0.421	0.146	0.202	0.175	0.026

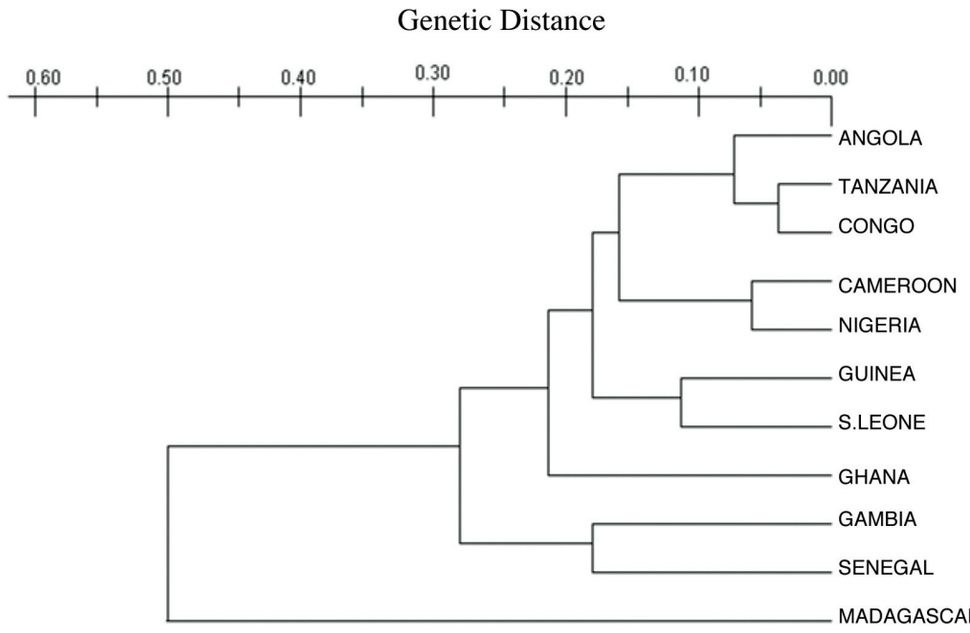

Figure 1. Dendrogram constructed based on Nei's distance (Nei, 1972) for the 11 MPOB oil palm germplasm populations included in the study.

- sub-cluster 3 – Sierra Leone and Guinea germplasm; and
- sub-cluster 4 – Ghana germplasm.

The Gambia and Senegal germplasm were grouped in Cluster 2 whereas the Madagascar germplasm formed a separate cluster (Cluster 3).

There is a strong association between genetic distance and geographical location. With the exceptions of Guinea and Sierra Leone, the populations from central Africa (Nigeria, Cameroon, Angola, Tanzania, Congo and Ghana) formed one cluster. Similar observations were also reported by Shah *et al*. (1994), Maizura *et al*. (2000) and Rajanaidu *et al.* (2000). Despite the countries in west Africa (Guinea, Sierra Leone and Ghana) being closer to each other compared with other countries in Cluster 1, it was surprising that the germplasm of Guinea and Sierra Leone were grouped together with the materials from central Africa. However, Zeven (1967) considered that the main oil palm belt begins in Guinea, spreading to the east and south through Sierra Leone, Ghana and other central African countries, and this may explain the genetic similarity of the germplasm from these countries. The grouping of Gambia and Senegal germplasm in Cluster 2 reflects that Gambia is entirely surrounded by Senegal, and that gene flow may have occurred between the populations of the two countries contributing to a higher degree of genetic similarity. The Madagascar germplasm formed a single region cluster implying the unique characteristics of these palms as compared with those in mainland Africa. Palms from Madagascar are generally short in stature, with oil having a high iodine value and high linoleic acid (C18:2) (Rajanaidu *et al*., 2000).

The results of this study provide an overview of the genetic variability and relationships amongst the MPOB germplasm populations, other than revealing the potential use of EST-SSR markers to distinguish among the germplasm from the various countries. Despite using a small number of markers and that the sample size in the analysis might not have covered the total level of heterozygosity and relationships among the germplasm populations (which may not be enough for selection and conservation purposes), the results of this study will help to determine the design of future studies on genetic diversity for individual populations in each of the germplasm. Focus should be directed more on germplasm with high heterozygosity compared with germplasm with low heterozygosity. Further studies will be continued to estimate the genetic variability and relationships for each individual population in germplasm originating from each country. Such information can help the breeders to identify redundancies in the collection, and

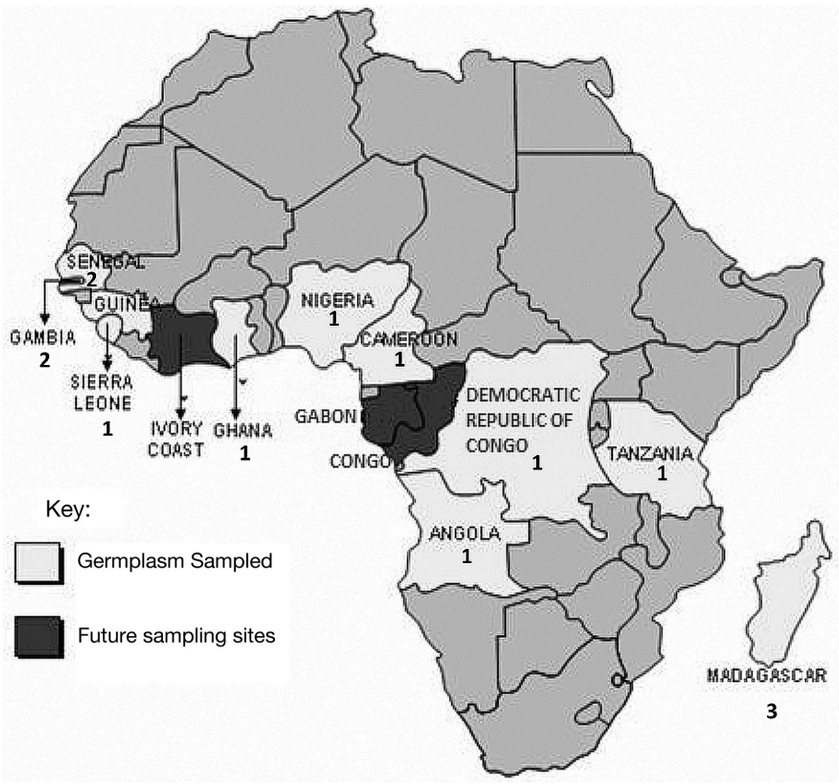

Figure 2. Map of Africa indicating the distribution of oil palm germplasm and the clustering results based on 10 EST-SSR markers.

develop optimal sampling strategies for field conservation. This can secure a greater range of diversity than sampling at random. Selection and conservation should also take into account those oil palm populations which are high in rare alleles and have high H_e as these populations may possess interesting and unique genes which can be exploited further.

CONCLUSION

Studies on genetic variation within the MPOB germplasm collection or natural populations are crucial for effective conservation and exploitation of genetic resources. Introgression studies can be done using genetic materials ranging from germplasm to advanced planting materials for crop improvement programmes. The results of this study indicate the usefulness of EST-SSR markers in revealing genetic variability and relationships among populations in the MPOB germplasm collection. The study will be extended in the future to estimate genetic variability and relationships for individual populations in each of the germplasm from different countries.

The 10 EST-SSR primers used in this study successfully differentiated and clustered the germplasm populations from the various countries. Estimates of the genetic variability parameters for the germplasm populations were obtained and high genetic diversity was observed among the MPOB germplasm populations. EST-SSR have some advantages over genomic SSR and other markers such as RFLP and AFLP because of their potential as markers representing genes linked to traits of agronomic interest. This has important applications in diversity and phylogenetic analyses of germplasm populations. However, the functions of the SSR motif and genes that contain SSR in the plant genome need to be studied further by localising EST with traits of interest on genetic linkage maps. EST-SSR markers also have been useful for integrating phenotypic and genotypic variations, and have provided breeders and geneticists with an efficient tool to link both characteristics (Gupta and Varshney, 2000).

The high genetic diversity shown in this study revealed that the germplasm collection is a good source for new genes related to novel traits which can be introgressed into the current planting materials. Introgressing new genes will increase genetic variation for selection and may contribute many valuable economic traits.

Based on the estimated genetic parameters (P, A and H_e), it can be concluded that generally, the natural African oil palm populations assayed maintained high levels of genetic variation. High percentage of polymorphism loci, highest mean number of alleles per locus, high H_e and a number of rare alleles were obtained in the population from Nigeria. With support from other findings, this suggests that this population may represent the centre of diversity of wild oil palm, as previously proposed.

ACKNOWLEDGEMENT

We wish to thank the Director-General of MPOB for permission to publish these research findings.

REFERENCES

BAKOUME, C (2006). *Genetic Diversity of Natural Oil Palm* (Elaeis guineensis *Jacq.*) *Populations Using Microsatellite Markers*. Ph.D thesis, Universiti Kebangsaan Malaysia.

BARCELOS, E (1998). Etude de la diversite genetique du genre *Elaeis* [*E. oleifera* (Kunth) Cortes et *E. guineensis* Jacq.] par marqueurs moleculaires (RFLP et AFLP). These de Doctorat, Ecole Nationale Superieure Agronomique de Montpellier.

BORNET, B; GORAGUER, F; JOLY, G and BRANCHARD, M (2002). Genetic diversity in European and Argentinian cultivated potatoes (*Solanum tuberosum* subsp. *tuberosum*) detected by inter-simple sequence repeats (ISSRs). *Genome, 45*: 481-484.

BROWN, A H D (1978). Isoenzyme, plant population structure and genetic conservation. *Theor. Appl. Genetics*, *52*: 145-157.

CONDIT, R and HUBBEL, S P (1991). Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome, 34*: 66-71.

CORDEIRO, G M; CASU, R; McINTYRE, C L; MANNERS, J M and HENRY, R J (2001). Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Science*, *160*: 1115-1123.

DELLAPORTA, S L; WOODS, J and HICKS, J B (1983). A plant DNA minipreparation: version II. *Plant Mol. Biol., 1*: 19-21.

GHESQUIERE, M; BARCELOS, E; SANTOS, M D M and AMBLARD, P (1987). Enzymatic polymorphism in *Elaeis oleifera* H.B.K (*E. melanococca*), analysis of population in Amazon Basin. *Oleagineux, 42(4)*: 143-153.

GUPTA, P K and VARSHNEY, R K (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica, 113*: 163-185.

GUPTA, P K; VARSHNEY, R K; SHARMA, P C and RAMESH, B (1999). Molecular markers and their applications in wheat breeding. *Plant Breeding, 118*: 369-390.

HACKAUF, B and WEHLING, P (2002). Identification of microsatellite polymorphism in an expressed portion of the rye genome. *Plant Breeding*, *121*: 17-25.

HARTLEY, C W C (1988). *The Oil Palm* (*Elaeis guineensis* Jacq.). 3rd ed. Longman Group Limited, New York.

HAYATI, A (2002). *Isozyme Variation Study in African Oil Palm* (Elaeis guineensis *Jacq.*) *Germplasm Collections*. M.Sc. thesis, Universiti Kebangsaan Malaysia.

HAYATI, A; WICKNESWARI, R; MAIZURA, I and RAJANAIDU, N (2004). Genetic diversity of oil palm (*Elaeis guineensis* Jacq.) germplasm collections from Africa: implication for improvement and conservation of genetic resources. *Theoretical and Applied Genetics*, *108*: 1274-1284.

JUNAIDAH, J; KUSHAIRI, A; ISA, Z A; MOHD DIN, A; NOH, A and RAJANAIDU, N (2004). Palm Series 7 (PS7) – high bunch index breeding population. *MPOB Information Series No. 221.* MPOB, Bangi.

KULARATNE, R S (2000). *Assessment of Genetic Diversity in Natural Oil Palm* (Elaeis guineensis *Jacq.*) *Using Amplified Fragment Length Polymorphism Markers*. Ph.D thesis, Universiti Kebangsaan Malaysia.

KUSHAIRI, A; RAJANAIDU, N; MOHD DIN, A; ISA, Z A; NOH, A and JUNAIDAH, J (2003a). Palm Series 5 (PS5) – breeding populations selected for thin shell *tenera*. *MPOB Information Series No. 183*.

KUSHAIRI, A; RAJANAIDU, N; MOHD DIN, A; ISA, Z A; NOH, A and JUNAIDAH, J (2003b). Palm Series 6 (PS6) – breeding populations selected for large fruit *dura*s. *MPOB Information Series No. 184*.

KUSHAIRI, A; RAJANAIDU, N; SUNDRAM, K and MAIZURA, I (2004). PS8: vitamin E breeding population. *MPOB Information Series No. 229*.

MAIZURA, I; RAJANAIDU, N; ZAKRI, A H and CHEAH, S C (2000). Oil palm (*Elaeis guineensis* Jacq.) genetic diversity determined by restriction fragment length polymorphism. *Proc. of the 4th National Congress on Genetics – Genetic Manipulation: Challenges and Advances.* p. 123-129.

MAIZURA, I; RAJANAIDU, N; ZAKRI, A H and CHEAH, S C (2006). Assessment of genetic diversity in oil palm (*Elaeis guineensis* Jacq.) using restriction fragment length polymorphism (RFLP). *Genetic Resources and Crop Evolution*, *53*: 187-195.

MOHD DIN, A; NOH, A; MOHD ISA, Z A; MAIZURA, I; KUSHAIRI, A and RAJANAIDU, N (2006). PS11 – high carotene *Elaeis guineensis* breeding population. *MPOB Information Series No. 312.*

NEI, M (1972). Genetic distance between populations. *American Naturalist*, *106*: 283-292.

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

NICOT, N; CHIQUET, V; GANDON, B; AMILHAT, L; LEGEAI, F; LEROY, P; BERNARD, M and SOURDILLE, P (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theoretical and Applied Genetics*, *109*: 800-805.

NOH, A; KUSHAIRI, A; MOHD DIN, A; MAIZURA, I; ISA, Z A and RAJANAIDU, N (2005). Palm Series 10 (PS10) – breeding populations selected for long stalk. *MPOB Information Series No. 267*.

PONCET, V; RONDEAU, M; TRANCHANT, C; CAYREL, A; HAMON, S; KOCHKO, A and HAMON, P (2006). SSR mining in the coffee tree EST database: potential use of EST-SSRs as markers for the *Coffea* genus. *Molecular Genetics Genomics*, *276*: 436-449.

PURBA, A R; NOYER, J L; BAUDOUIN, L; PERRIER, X; HAMON, S and LAGODA, P J L (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. *Theoretical and Applied Genetics, 101*: 956-961.

RAJANAIDU, N (1994a). *PORIM Oil Palm Genebank*. PORIM, Bangi.

RAJANAIDU, N (1994b). PORIM Series No.1 – PORIM elite oil palm. Planting Materials. *PORIM Information Series No. 15*.

RAJANAIDU, N (1994c). PORIM Series 2 – PORIM elite oil palm. *PORIM Information Series No. 16*.

RAJANAIDU, N (1996). PORIM elite oil palm Series 3 (mother palm) – high kernel. *PORIM Information Series No. 41*.

RAJANAIDU, N and JALANI, B S (1994). Oil palm genetic resources – collection, evaluation. utilization and conservation. Paper presented at the PORIM Colloqium on Oil Palm Genetic Resources. PORIM, Bangi. 13 September 1994.

RAJANAIDU, N; MAIZURA, I and CHEAH, S C (2000). Screening of oil palm natural population using RAPD and RFLP molecular markers. *Proc. of the International Symposium on Oil Palm Genetic Resources and Utilization*. MPOB, Bangi. p. AA1-AA28.

RAJEEV, K V; ANDREAS, G and MARK, E S (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, *23*: 48-55.

REES, A R (1965). Evidence for the African origin of the oil palm. *Principles*, *9*: 30.

RODER, M S; KORSUN, V; WENDEHAKE, K; PLASCHKE, J; TIXIER, M H; LEROY, P and GANAL, M W (1998). A microsatellite map of the wheat genome. *Genetics, 149*: 2007-2023.

SHAH, F H; RASHID, O; SIMONS, A J and DUNSDON, A (1994). The utility of RAPD markers for determination of genetic variation in oil palm (*Elaeis guineensis*). *Theoretical and Applied Genetics*, *89*: 713-718.

SCOTT, K D; EGGLER, P; SEATON, G; ROSETTO, M; ABLETT, E M; LEE, L S and HENRY, R J (2000). Analysis of SSRs derived from grape ESTs. *Theoretical and Applied Genetics*, *100*: 723-726.

SINGH, R; NOORHARIZA, M Z; TING, N C; ROZANA, R; TAN, S G; LOW, E T; MAIZURA, I and CHEAH, S C (2008). Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity. *Biologia*, *63(2)*: 227-235.

SOURDILLE, P; TAVAUD, M; CHARMET, G and BERNARD, M (2001). Transferability of wheat microsatellite to diploid Triticeae species carrying the A, B and D genome. *Theoretical and Applied Genetics*, *103*: 346-352.

THIEL, T; MICHALEK, W; VARSHNEY, R K and GRANER, A (2003). Exploiting EST database for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, *106*: 411-422.

ZEVEN, A C (1964). On origin of the oil palm. *Grana Palynol, 5*: 50.

ZEVEN, A C (1967). *The Semi-wild Oil Palm and its Industry in Africa*. Centre for Agricultural Publishing and Documentation (PUDOC). Wageningen. 178 pp.