Allelic Diversity and Relatedness of Wild *Elaeis oleifera* Populations Using Microsatellite Markers

Teh, C. K.[1], Maizura, I.[2], Bakoume, C.[3] and Wickneswari, R.[1]

ABSTRACT

Genetic diversity and genetic structure of plant collections are useful for conservation and breeding purposes. A total of 540 trees of Elaeis oleifera *(American oil palm) belonging to 19 populations from four countries in Central and South America maintained by Malaysian Palm Oil Board (MPOB) at Kluang Oil Palm Research Station were screened using 14 microsatellite loci. On average, 74.4% of loci were polymorphic at 0.99 criterion and 109 alleles were detected across the populations. The average number of alleles per locus (A_o) ranged from 1.8 in population 1 from Colombia to 4.6 in population 14 from Costa Rica (mean = 3.0 ± 0.9) and the mean effective number of alleles per locus was low (A_e = 1.4 ± 0.1). Unweighted pair group method with arithmetic mean (UPGMA) cluster analysis revealed three main clusters i.e. Panama populations formed one cluster, Colombian populations a second cluster and the third cluster comprised of Honduras populations and Costa Rican populations.*

INTRODUCTION

The *Elaeis* genus consists of two species i.e. *Elaeis guineensis* and *Elaeis oleifera*. The fruits of both species contain oil in their kernel (palm kernel oil) and mecocarp (palm oil) which are suitable for human consumption as well as oleo-chemical industries. According to O'Holohan (1997) and Hartley (1988), *E. guineensis* originates from tropical Guinea's rainforest in Central and West Africa, while *E. oleifera* is naturally distributed in Central and South America. Palm oil is now the largest contributor to world total fats and oils production after surpassing soybean oil since year 2003 (Oil World Annual report 2008). In year 2008, palm oil production represented 116% of that of soybean oil. However, soybean oil is by far preferred compared to palm oil because of the relative high quality of the former.

To defend Malaysian palm oil, the oil quality and the economic life of oil palm plantation must be improved in order to strengthen its global competitive power instead of solely increase the oil production. To improve the crop, new genetic resource is essential to broaden the genetic base of the present commercial plantations. High oil quality and low vertical growth rate are peculiarities of the American oil palm (*E. oleifera*). According to Choo & Yusof (1996), the *E. oleifera* oil is highly unsaturated (i.e. high iodine value or IV) with high composition of linoleic acid and oleic acid, low content of palmitic acid and high carotene content. The carotene profile

[1] School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.
[2] Malaysian Palm Oil Board (MPOB), No. 6, Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia.
[3] Sime Darby Technology Centre Sdn. Bhd., No. 2, Jalan Tandang, 46050 Petaling Jaya, Selangor, Malaysia.

of *E. guineensis dura*, *pisifera* and *tenera* are 428 ppm, 997 ppm and 673 ppm, respectively (Sundram et al. 2003). Furthermore, the annual height increment of *E. oleifera* is extremely low, about 5 – 10 cm which is less than one-fifth of that of *E. guineensis*. Reduced interest of planters to this species is in relation with its low oil yield (0.5 t ha^{-1} yr^{-1}). Hence, breeders have initiated interspecific hybrid breeding in order to transfer the desirable agronomic traits in *E. oleifera* to *E. guineensis* (Hardon 1969; Amblard et al. 1995). An extensive *E. oleifera* germplasm collection was also carried out in its natural distribution areas in Latin America (Rajanaidu 1986).

The oil palm germplasm are field evaluated based on their phenotypes including, oil quality, vegetative and physiological characteristics. But, phenotypes are influenced by environmental factors which can lead to questionable conclusions. Hence, molecular approaches which address genetic variation of natural populations without prior reference to phenotypic characteristics are being adopted extensively. In this study, microsatellite markers were used to study the genetic diversity of *E. oleifera* populations from Central and South America. Microsatellite markers are tandem arrays of simple nucleotide motifs that are ubiquitous component of eukaryotic genomes (Delseny et al. 1983). These codominant markers have a high degree of polymorphism which confers high discriminative power. Besides, microsatellite markers were found to be reliable, reproducible and cost effective (Smith et al. 1997). The information obtained would be useful for oil palm breeding and conservation of *E. oleifera* genetic resources.

MATERIALS AND METHODS

Plant materials and DNA extraction

The *E. oleifera* germplasm collections are currently planted in MPOB experimental plots located at Johore, Malaysia. A total of 540 oil palms were sampled from populations from Honduras (2), Costa Rica (6), Colombia (5) and Panama (6) (Table 1). Nicaragua and Suriname populations were not included in the study due to limited number of palms available in the field. About 30 progenies were sampled from each population. The genomic DNA was isolated from 3 g of each leave sample and purified using CTAB extraction and purification protocol by Doyle & Doyle (1990) with minor mortifications. The purified DNA samples were quantified by optical density (OD) reading using spectrophotometer (UV/VIS Spectrometer Lambda Bio, Perkin Elmer, USA).

Table 1: Number of populations, families and progenies included in the study

Country	No. of populations	No. of families	No. of progenies
Honduras	2	5	52
Costa Rica	6	20	168
Panama	6	11	172
Colombia	5	14	148
Total	19	50	540

Transferability of *E. guineensis* microsatellite markers to *E. oleifera*

A set of 18 microsatellite markers (at least one per linkage group for good coverage of the whole genome) developed by Billote et al. (2005) from *E. guineensis* was used in this study. They are characterized by a very high polymorphism in *E. guineensis* (Bakoume 2006). Three DNA samples from oil palms each one belonging to different population was used to check for transferability of the 18 microsatellite markers to *E. oleifera*.

Microsatellite amplification

The total volume of each PCR reaction was 12.5 µl, comprising 1x PCR buffer (50 mM KCl, 10mM Tris-HCl (pH 9.1 at 21˚C), 0.1% Triton^{TM}X-100), 1.5 mM Mg_2Cl, 200 µM dNTPs mix, 1.0 U Taq DNA polymerase (Vivantis, Malaysia), 0.25 µM of each primer, 0.38 µM fluorescent dye (Applied Biosystems, USA) and 50 ng of DNA template. The PCR conditions included initial denaturation at 94 ˚C for 5 minutes, followed by 35 cycles of denaturation at 94 ˚C for 30 seconds, annealing at 52 ˚C for 1 minute and elongation at 72 ˚C for 2 minutes, then final elongation step of 72 ˚C for 10 minutes. The M13-tailed forward primer was labeled with 4-colour fluorescent dyes (Blair et al. 2002). The fluorescent dyes used in the study were FAM (blue), PET (red), VIC (green) and NED (yellow). The sample set was subdivided into four panels and each of them consisted of four microsatellite primers labeled with four fluorescent dyes, respectively.

Fragment analysis and genotyping

To multiload the PCR products, 2µl of PCR product for each primer in a panel was transferred into a 0.2 ml PCR tube and mixed gently by pipetting. Then, 2µl of the mixed PCR products were transferred into 96-well PCR microplate (Axygen, USA). The PCR product was pre-treated by adding 7.80 µl of Hi-Di^{TM} Formamide (Applied Biosystems, USA) and 0.20 µl of GeneScan^{TM}-500 LIZ^{TM} Size Standard (Applied Biosystems, UK) to make up to 10.0 µl of the final volume, the 96 well PCR microplate was sealed by 96 well plate septa (Applied Biosystems, USA). Then, the mixture was vortexed thoroughly and then denatured at 95 ˚C for 5 minutes in a thermocycler. The microplate was placed on ice immediately after the denaturation for at least 5 minutes. Subsequently, the fragment analysis via capillary electrophoresis was carried out using ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, USA). All the raw data were imported from the database of ABI PRISM® 3100 data collection software to GeneMapper V3.1 software. The genotypic data were finally generated in size (bp) and reviewed in GeneMapper software for edition purposes.

Data analysis

The generated genotypic data of *E. oleifera* were analyzed using PowerMarker V3.25 (Liu & Muse 2005) based on Weir's (1996) concepts to estimate the genetic diversity parameters. It comprises allele frequencies, average number of allele per locus (*A*), effective number of allele per locus (A_e) and percentage of polymorphic loci. An unrooted dendrogram was drawn according to the genetic distance of Nei et al. (1983) using the Unweighted Pair Group Methods

with Arithmetic Mean (UPGMA) clustering methods (Sneath & Sokal 1973). One-way ANOVA test was done in order to check whether there were differences among all populations studied for A_e and H_e using Minitab 14 (2003).

RESULTS AND DISCUSSION

Only 14 out of 18 microsatellite loci chosen among those developed for *E. guineensis* were transferable to *E. oleifera*. The amplicons' sizes were comparable to those obtained from *E. guineensis*. The unsatisfactory loci i.e. mEgCIR1753, mEgCIR3300, mEgCIR3574 and mEgCIR3785 showed inconsistent and poor amplifications. They were not specific to any classes of microsatellite. The mutations within the DNA sequences complementary to PCR primers (flanking regions of microsatellite) probably suppress their binding causing a complete loss of allele amplification (Dumas et al. 1998). The mutation rate between species is relatively higher than that intraspecies, so null alleles tend to occur in cross-taxa amplification. Finally, 14 microsatellite loci were assayed across all populations. A total of 109 alleles were detected among 540 individuals (Table 2). The number of alleles per locus ranged from 3 to 13 with an average of 7.8.

Table 2: Microsatellite loci assayed, number of alleles and detected size range of alleles

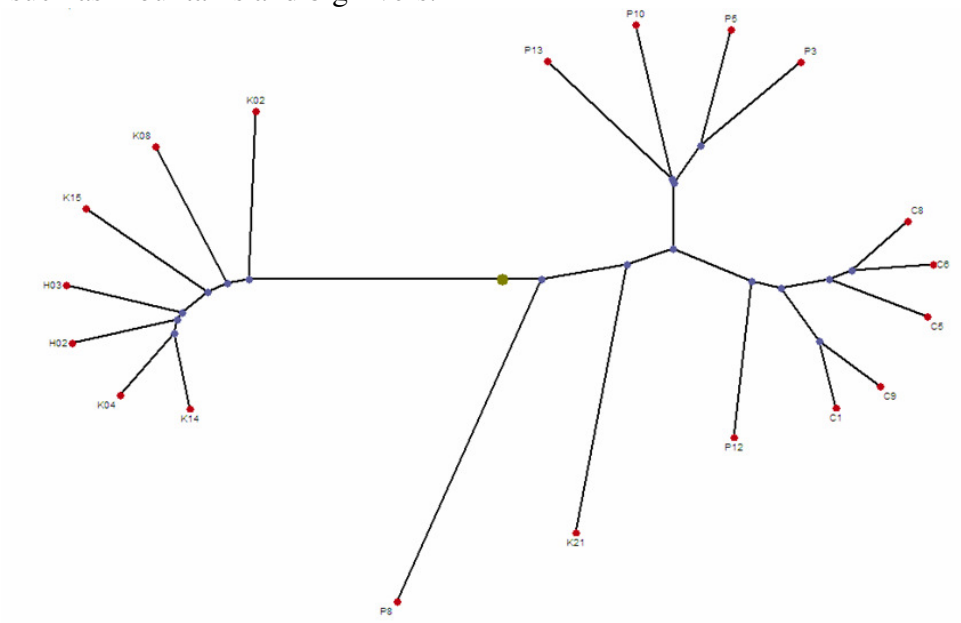
Linkage group	Microsatellite loci	Motif	No. of allele	*Size (bp)	Size range (bp)
1	mEgCIR0802	$(GA)_{12}$	8	250	230-272
2	mEgCIR3282	$(GA)_{20}$	12	245	232-272
3	mEgCIR0173	$(GA)_{18}$	5	132	122-154
5	mEgCIR3691	$(GA)_{14}$	6	181	160-260
6	mEgCIR3543	$(GA)_{17}$	6	232	200-260
7	mEgCIR2387	$(GA)_{11}$	13	243	250-314
8	mEgCIR3363	$(GA)_{17}$	6	195	175-228
9	mEgCIR3886	$(GA)_5GT(GA)_{20}$	7	187	184-218
11	mEgCIR3362	$(GA)_{19}$	7	151	130-180
12	mEgCIR1730	$(GT)_{17}(GT)_5$	11	269	260-290
13	mEgCIR0832	$(GA)_{19}$	3	240	145-240
14	mEgCIR3546	$(GA)_{15}$	11	286	286-336
15	mEgCIR3292	$(GA)_{20}$	6	173	140-212
16	mEgCIR0353	$(GT)_{11}(GA)_{15}$	8	102	50-110
Total			109		
Mean			7.8		

*according to Billotte et al. (2005); TOPGENE database (CIRAD)
-microsatellite loci from linkage groups 4 and 10 were not assayed.

The assayed populations shared the same common alleles at eight or 57% of microsatellite loci showed that genetic drift is strongly present in the assayed populations. According to Templeton (2006), the evolutionary changes in allele frequency caused by genetic drift accumulate with time and eventually populations tend to lose or to be fixed alleles. The

genetic drift effect will be terminated once the lost or fixation of alleles happens within population. The means for A and A_e over the populations were 3.0 (± 0.9) and 1.4 (± 0.1), respectively. The mean A_e was almost 50% lower than that reported in *E. guineensis* (A_e = 3.3) from different geographic regions using 16 microsatellite markers (Bakoume 2006). The low values of A_e may be due to the small and isolated *E. oleifera* populations as reported by Rajanaidu (1986). According to Nei et al. (1975), the allelic variation in small and isolated populations might decline obviously because of demographic bottleneck events. The percentage of polymorphic loci at 0.99 criterion ranged from 50% to 100% with an average of 74.4%. This confirms the high polymorphism of *E. guineensis*-derived SSR markers in *E. oleifera*. The one-way ANOVA test did not detect any significant difference among populations ($p < 0.05$).

An unrooted-dendrogram of genetic relatedness among the 19 *E. oleifera* populations showed that the genetic structure of *E. oleifera* corresponded to the geographic distribution (Figure 1). Panama populations and Colombian populations formed two distinct clusters and the third cluster comprised of Honduras populations and Costa Rican populations. All Costa Rican populations are closely related except population 21 which has closer kinship with Colombian populations and Panama populations. The finding maybe associated to the human movement either from southeast Panama or northern Colombia to southern Costa Rica. It might relate to the importance of *E. oleifera* for the indigenous people in that region. For instance, Indians use *E. oleifera* to make fire with dried fruits, to make 'wine', to use it as natural dye, as mosquito repellent and to feed domestic animals (Moretzsohn et al. 2002; Corley & Tinker 2003). Interestingly, cocoa (*Theobroma cacao*) is also a diploid (2n = 20), perennial and allogamous species which is similar to *E. oleifera*. This tropical species in Latin America also revealed the same genetic patterns and probably spread by the Indian's migration from Southern America to Central America and Southern Mexico in the pre-Columbian times (Schultes 1984). Furthermore, population 12 from Panama was also apart from the other Panama populations due to its proximity to Colombia. Population 8 from Panama seems to be isolated by the geographical barriers such as mountains and big rivers.



C – Colombia, H – Honduras, K – Costa Rica and P – Panama

Figure 1: UPGMA unrooted-dendrogram of 19 *E. oleifera* populations based on genetic distance of Nei et al. (1983)

CONCLUSION

The allelic diversity of *E. oleifera* in MPOB is relatively low. The results from the study provide a tool for selection of minimum number of populations to conserve that secure the current diversity. It would be necessary to carry out supplementary plant explorations in larger natural distribution areas of the species for further allelic enrichment of germplasm. Besides, the current study could be confirmed by using biochemical markers such as isozymes.

ACKNOWLEDGEMENTS

The authors thank the Director-General of Malaysian Palm Oil Board (MPOB) for permission to present this paper. The research work was supported by MPOB Graduate Student Assistantship Scheme.

REFERENCES

Amblard, P., Noiret, J. –M., Kouamé, B., Potier, F. & Adon, B. 1995. Performances comparées des hybrides interspécifiques et du matériel commercial *E. guineensis*. *Oléaginueux Crops gras Lipides* 2: 335 – 340.

Bakoume, C. 2006. Genetic diversity of natural oil palm (*Elaeis guineensis* Jacq.) populations using microsatellite markers. Ph.D Thesis. Universiti Kebangsaan Malaysia.

Billotte, N., Marseillac, N., Risterucci, A. -M., Adon, B., Brottier, P., Baurens, F. –C., Sigh, R., Herrán, A., Asmady, H., Billot, C., Amblard, P., Durand-Gasselin, T., Courtois, B., Asmono, D., Cheah, S. C., Rohde, W., Ritter, E. & Charrier, A. 2005. Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.) *Theory Applied Genetics* 110: 754 – 765.

Billote, N., Risterucci, A. M., Barcelos, E., Noyer, J. L., Amblard, P. & Baurens, F. C. 2001. Development, characterisation and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome* 44: 413 – 425.

Blair, M. W., Hedetale, V. & McCouch, S. R. 2002. Fluorescent-labeled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 105: 174 – 184.

Choo, Y. M. & Yusof, B. 1996. *Elaeis oleifera* palm for the pharmaceutical industry. *PORIM Information Series No. 42*: 1 – 4.

Corley, R. H. V. & Tinker, P. B. 2003. The classification and morphology of the oil palm. In: *The Oil Palm*. 4th Edit., pg. 27 – 51. Oxford: Blackwell Publishing Company.

Delseny, M., Laroche, M. & Penon, P. 1983. Detection of sequences with Z-DNA forming potential in higher plants. *Biochemical and Biophysical Research Communications* 116: 113 – 120.

Doyle, J. & Doyle, L. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13 – 15.

Dumas, V., Herder, S., Bebba, A., Cadoux-BarnabÃ©, C. Bellec, C. & Cuny, G. 1998. Polymrphic microsatellites in *Simulium damnosum* s.l and their use for differentiating two savannah populations: Implications for epidemiological studies. *Genome* 41: 154 – 161.

Hardon, J. J. 1969. Interspecific hybrids in the genus *Elaeis* II. Vegetative growth and yield of F1 hybrids *E. guineensis* x *E. oleifera*. *Euphytica* 18: 380 – 388.

Hartley, C.W.S. 1988. *The Oil Palm (Elaeis guineensis)*. 3rd Edit. New York: Longman.

Liu, K. & Muse, S.V. 2005. PowerMarker: Intergrated analysis environment for genetic marker data. *Bioinformatics* 21(9): 2128 – 2129.

Minitab 14 Statistical Software. 2003. [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com).

Moretzsohn, M. C., Ferreira, M. A., Amaral, Z. P. S., Coelho, P. J. A., Grattapaglia, D. & Ferreira, M. E. 2002. Genetic diversity of Brazilian oil palm (*Elaeis oleifera* H.B.K.) germplasm collected in the Amazon Forest. *Euphytica* 124: 35 – 45.

Nei, M., Maruyama, T. & Chakraborty, R. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29: 1 – 10.

Nei, M. & Takezaki, N. 1983. Estimation of genetic distances and phylogenetic trees from DNA anlysis. *Proceeding of the 5th World Congress on Genetics Appied to Livestock* 21: 405-412.

O'Holohan, D. R. 1997. *Malaysian Palm Oil: The Story of a Major Edible Vegetable Oil and its Role in Human Nutrition*. Kuala Lumpur: Malaysian Palm Oil Promotion Council.

Oil world. 2008. *Oil World Annual*. ISTA Mielke, Hamburg.

Rajanaidu, N. 1986b. *Elaeis oleifera* collections in Central and South America. *Proceeding of International Workshop on Oil Palm Germplasm and Utilization*, pg. 84 – 94. Bangi: Palm Oil Research Institute Malaysia.

Romaro, C., Pedryc, A., Muñoz, V., Llácer, G. & Badenes, M. L. 2003. Genetic diversity of different apricot geographical groups determined by SSR markers. *Genome* 46: 244 – 252.

Schultes, R. E. 1984. Amazonian cultigens and their northward and westward migrations in pre-Columbian times. In: Stone, D. (ed.). *Pre-Columbian Plant Migration, Papers of the Peabody Museum of Archeology and Ethnology* 76, pg. 69 – 83. Massachusetts: Harvard University Press: Cambridge.

Smith, J. S. C., Chin, E. C. L., Shu, H., Smith, O. S., Wall, S. J., Senior, M. L., Mitchell, S. E., Kresovich, S. & Ziegle, J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.), comparisons with data from RFLPs and pedigree. *Theoretical and Applied Genetics* 95: 163 – 173.

Sneath, P. H. A. & Sokal, R. R. 1973. Numerical taxonomy. San Francisco: Freeman.

Sundram, K., Sambanthamurthi, R. & Tan, Y. A. 2003. Palm fruit chemistry and nutrition. *Asia Pacific Journal of Clinical Nutrition* 12(3): 355 – 362.

Templeton, A. R. 2006. *Population Genetics and Microevolutionary Theory*. New Jersey: John Wiley & Sons, Inc.

Weir, B. S. 1996. Genetic data analysis II. Sunderland, MA: Sinauer Associates, Inc.