BIODIVERSITAS
Volume 20, Number 5, May 2019
Pages: 1233-1240

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d200501

Genetic structure and diversity between and within African and American oil palm species based on microsatellite markers

AZIS NATAWIJAYA[1,2], SINTHO W. ARDIE[1], MUHAMAD SYUKUR[1], ISMAIL MASKROMO[3], ALEX HARTANA[4], SUDARSONO SUDARSONO[1,♥]

[1]PMB Lab., Departement of Agronomy and Horticulture, Faculty of Agriculture, Institut Pertanian Bogor. Jl. Meranti, IPB Dramaga Campus, Bogor 16680, West Java, Indonesia. Tel./Fax: +62-251-8629353, ♥ email: s_sudarsono@ymail.com, sudarsono_agh@apps.ipb.ac.id

[2]Mekarsari Research Station, Mekarsari Fruit Garden (Taman Buah Mekarsari). Jl. Raya Cileungsi – Jonggol Km. 3, Mekarsari, Cileungsi, Bogor 16820, West Java, Indonesia

[3]Indonesian Palms Research Institute, Agency for Agricultural Research and Development. Jl. Raya Mapanget, Mapanget, Talawaan, Manado 95001, North Sulawesi, Indonesia

[4]Plant Genetic Division, Department of Biology, Faculty of Mathematics and Natural Sciences, Institut Pertanian Bogor. Jl. Meranti, IPB Dramaga Campus, Bogor 16680, West Java, Indonesia.

Manuscript received: 16 September 2018. Revision accepted: 1 April 2019.

Abstract. *Natawijaya A, Ardie SW, Syukur M, Maskromo I, Hartana A, Sudarsono S. 2019. Genetic structure and diversity between and within African and American oil palm species based on microsatellite markers. Biodiversitas 20: 1233-1240.* The genus *Elaeis* consists of only two species, *Elaeis guineensis* Jacq. (the African oil palm species) and *E. oleifera* (HBK) Cortes (the American oil palm species). *E. guineensis* (*E.g*) is widely cultivated in southeast Asia and Africa, whereas *E. oleifera* (*E.o*) is naturally existed and cultivated in Central and South America. The objectives of this research were to analyze genetic diversity of eight groups of *E.g* and two groups of *E.o* using co-dominant genetic markers (SSRs) and evaluate their genetic structures. A total of 27 SSR loci was used to genotype a total of 128 accessions of African oil palm species (*E.g*) belonging to three different types (Dura, Pisifera and Tenera) and eight genetic backgrounds (Dumpy Dura and Deli Dura; Avros, Dumpy Avros, Binga, and Angola Pisifera; and Angola and Dumpy Avros Tenera) and 64 accessions of *E.o* collected from two different regions (Tefe and Manaus). The genotype data were used to calculate the population genetic diversity and structures for each oil palm species using the appropriate software. Results of the analysis indicated although they belonged to two different species, *E.g* and *E.o* shared many of the same SSR alleles in their genome and only contain few species-specific SSR alleles. Most of the evaluated genetic parameters were similar between *E.g* and *E.o* oil palm species but *E.o* has higher average number of effective allele than that of *E.g.* The calculated genetic variance is mostly belonged to the within-species variance source while the between species is relatively small. The phylogenetic tree and structure analysis reveal the high genetic variability among the evaluated oil palm groups which would be beneficial for future breeding program at Mekarsari Research Station. The tested *E.o* specific alleles were effective for identifying introgression lines between *Eo* × *E.g* // *E.g* carrying the *E.o* chromosome fragments. Therefore, these *E.o* specific alleles could be used in oil palm backcrossing program to monitor the introgression process.

Keywords: Oil palm species, *Elaeis guineensis*, *Elaeis oleifera*, introgression line, species-specific markers, population structure

INTRODUCTION

Oil palm (*Elaeis* sp.) is one of the most important commercial crops in southeast Asia. The *Elaeis* genus consists of only two species, *E. guineensis* Jacq. (the African oil palm species) and *E. oleifera* (Kunth) Cortes (the American oil palm species) (Singh et al. 2013). *Elaeis guineensis* (*E.g*). is the most productive plant species producing vegetable oils and it is widely cultivated in southeast Asia and Africa (Barcelos et al. 2015). The *E. guineensis* grows well in the tropical lowland with an average of annual rainfall ranges from 1,780-2,280 mm and temperatures from 24-30 ºC. Whereas the *E. oleifera* (*E.o*) populations grow along riverbanks and some grow in submergence in Central and South American forests such as in Colombia, Suriname, and Brazil (Corley and Tinker 2003). The *E.g* species was introduced to Indonesia in 1848 and the *E.o* species in 1950 (Barcelos et al. 2015; Pamin 1998).

The two species of oil palm separated a long time ago, probably when the continental separation occurred. Adaptation to diverse growing environments for a long time leads to differences and modifications to the genetic structure of the population. Schaal et al. (2003) suggest that the rate of gene flow within and between species is one of the factors affecting genetic differentiation. Although it was separated a long time ago, they still retain the same chromosome numbers (2n = 2x = 32), and they are still cross-hybridized (Hardon and Tan 1969). The *E.o* germplasm is widely known as the important donor for oil quality improvement, very low height increment, and for improving resistance to some important diseases in oil palm (Cadena et al. 2013; Montoya et al. 2014; Moretzsohn et al. 2002; Sunilkumar et al. 2015).

The two oil palm species are cross-pollinated species, although the natural pollinators in these species are not the same. Variation within and between population is the source of genetic diversity in cross-pollinated species.

Bakoume et al. (2007) reported that analysis results of the natural genetic diversity suggested that wild African populations (*E.g*) could be separated into three groups, the extreme west of Africa, equatorial Africa, and the Madagascar Island. The highest allelic diversity was found among the Nigerian oil palm populations, indicating Nigeria was probably the possible African oil palm center of origin. Using MPOB collection, Malaysia, Din et al. (2000) reported the morphological differences and diversity of American oil palm population (*E.o*) from Panama, Costa Rica, Colombia, and Honduras. Using 13 qualitative characters, Caicedo et al. (2017) reported the morphological diversity of 59 accessions of *E.o* in Colombia. Molecular analysis of the *E.o* species would certainly complement the previously collected morphological diversity data.

Utilization of some DNA-based molecular markers to study the population genetics of oil palm has been reported by some oil palm researchers (Bakoumé et al. 2014; Barcelos et al. 2002; Budiman et al. 2019; Hayati et al. 2004; Ithnin et al. 2017; Natawijaya et al. 2018; Okoye et al. 2016). Among molecular markers, microsatellite or simple sequence repeats (SSRs) as codominant and multiallelic markers have been widely used to estimate the oil palm genetic diversity. Therefore, the use of SSR markers to compare genetic diversity of *E.g* and *E.o* would add beneficial information for the oil palm breeding program.

Mekarsari Research Station, West Java, Indonesia has some elite accessions of the African oil palm species (*E.g*) and some accessions of the American oil palm species (*E.o*). The objectives of this study were to compare the genetic diversity of individuals within either the oil palm types and the groups or among the two oil palm species (*E.g* and *E.o*), to identify the existence of the *E.o* species specific SSR alleles and evaluate effectiveness of the species specific SSR alleles to identify accessions of the second backcross generation of *E.o* × *E.g* Dura E206 // *E.g* Dura E206 introgression lines (composite Dura) carrying the the *E.o* chromosome fragments.

MATERIALS AND METHODS

Plant materials and DNA extraction

The genetic materials (Table 1) consisted of 128 accessions of *E.g* belonging to three different types (Dura, Pisifera, and Tenera, with various genetic backgrounds) and 64 accessions of *E.o*, originated from two different regions in Latin America (Manaus and Tefe). All the oil palm materials were introduced from Malaysia and grown in the Mekarsari Research Station, West Java, Indonesia. Fresh leaf samples were taken from the research station and immediately used for DNA isolation.

The research was conducted at Plant Molecular Biology Laboratory, Department of Agronomy and Horticulture, Faculty of Agriculture, Bogor Agricultural University, Bogor, Indonesia. Total DNA samples from fresh leaves were extracted using Plant Genomic DNA Mini Kit for DNA isolation (Geneaid). The quantity and quality of the isolated DNA were measured using a spectrophotometer at 260 nm and 280 nm wavelengths. Subsequently, the DNA quality and quantity were checked by agarose gel electrophoresis using 0.8% agarose in 0.5 M TBE buffer, and they were stained with GelRed™ for visualization.

Primers and SSR analysis

Detection of SSR fragment polymorphisms was performed using 27 SSR markers. The markers were developed by Billote et al. (2001) from *E.g* and are available publicly at http://tropgenedb.cirad.fr/tropgene/JSP/interface.jsp?module =OILPALM.

Polymerase Chain Reaction (PCR) consisted of 4.75 µl ddH2O, 6.25 KAPA 2G FAST master (Taq DNA polymerase, dNTPs, MgCl2 and reaction buffer), 0.5 µl forward primer, 0.5 µl reverse primer and 1 µl of 50 ng / µl DNA template. The PCR was initiated with the first denaturation stage at 95 ºC. for 3 minutes, the second denaturation stage was at 95 ºC. for 15 seconds, the annealing stage for 30 seconds, the extension stage was at 72 ºC for 50 seconds and the final extension stage was at 72 ºC for 8 minutes. The process is run with 35 cycles before the temperature is lowered to 4 ºC. The results of PCR reactions were stored at 4 ºC. Amplification of SSR fragments using BIO-RAD T100 DNA Thermal Cycler. The PCR products were confirmed with 0.8% agarose gel in 1x Sodium Boric (SB) buffer.

Table 1. List of oil palm genetic materials used in this study

Species	Type[1]	Group[2]	Number of samples	Special phenotypes associated with the group
Elaeis guineensis (*E.g*)	Dura	Dumpy Dura	50	Slow growth
		Deli Dura	20	Big bunch size
	Pisifera	Avros	10	Big bunch size
		Dumpy Avros	10	Slow growth
		Binga	8	Thick mesocarp
		Angola	10	Long stalk
	Tenera	Angola	10	Long stalk, virescence
		Dumpy Avros	10	Slow growth
Elaeis oleifera (*E.o*)	-	Manaus	14	Slow growth, high oil quality
	-	Tefe	50	Slow growth, high oil quality

Note: [1] Identity based on the shell thickness of oil palm fruits. The oil palm fruits of Dura type have a thick shell, Pisifera type-shelless or thin shell, and Tenera type-a medium shell. [2] Identity was based on their genetic background or places of origin of the oil palm materials

The PCR products were separated by 6% polyacrylamide gel electrophoresis (PAGE) using 1x SB Buffer (Brody and Kern 2004). Vertical electrophoresis used the Cole-Parmer® Dedicated Height Sequencers tool. Each PCR product was mixed with loading dye and denatured for 10 minutes then placed in crushed ice. Pre-run was done at 100 watts for 30 minutes. Electrophoresis was carried out at 60 watts for 90 min and 50 bp DNA ladder was used as the fragment size control.

Silver staining followed the method of Creste et al. (2001) with some modifications as explained in the following steps. The staining process had five stages. The first stage was gel fixation for 10 minutes, followed by rinsing the plate with aquadest for 1 minute. The second stage was nitric acid washing for 3 minutes, then rinsed with aquadest for 1 minute. The third stage was silver nitrate staining stage for 20 minutes. The fourth was the developing stage in the developing solution (1.5 ml of formaldehyde and 200 µl of sodium thiosulfate) for 5-7 minutes until the DNA bands appear on the glass plate. The plate was then washed with aquadest quickly, about 5-10 seconds. The fifth stage was the stopping reaction for 5 minutes in stop solution (50 ml acetic acid glacial), then washing for 5 minutes in 1 L aquadest. The plate was then dried at room temperature with an upright position overnight until it completely dry. Finally, the visualization and scoring were done on the light table.

Data analysis

The molecular marker data were analyzed using DARwin 5 (Perrier and Jacquemod-Collet 2006), GenAlEx 6.5 (Peakall and Smouse 2006) and STRUCTURE (Evanno et al. 2005). The DARwin 5 was used to calculate Neighbour-Joining trees based on genetic distance using 1000 bootstrap. GenAlEx 6.5 was used to calculate the number of alleles per locus (Na), the number of effective alleles (Ne), polymorphic information content (PIC), and gene diversity index (H; Shannon 1948). The identification of species-specific alleles is based on the allele frequency information at each locus in both *E.o* and *E.g*, analyzed using GenAlEx software version 6.5 (Peakall and Smouse 2006).

Detection of *E.o* specific alleles in the *E.g* × *E.o* introgression lines

Effectiveness of the *E.o* specific SSR alleles identified in this study was tested by evaluating 23 accessions of the second backcross generation of *E.o* × *E.g* Dura E206 // *E.g* Dura E206 introgression lines (composite Dura). The ability to detect the *E.o* chromosome fragment was evaluated by identifying *E.o* specific SSR allele in the *E.g* genetic background.

Detection of *E.o* specific SSR allele was performed using mEgCIR3376 primer pairs. The SSR alleles were amplified by PCR using previously described conditions. The PCR products were separated by 6% polyacrylamide gel electrophoresis (PAGE) using 1x SB Buffer (Brody and Kern 2004). Vertical electrophoresis used the Cole-Parmer® Dedicated Height Sequencers tool. Electrophoresis was carried out at 60 watts for 90 min. The PAGE was stained using silver staining as described previously. The final visualization and scoring were done on the light table.

RESULTS AND DISCUSSION

Allelic diversity among oil palm species

The 27 SSR loci used in this study were distributed across sixteen linkage group based on Billote et al. (2005). Across the two oil palm species evaluated, the SSR loci yielded number of total alleles ranged from 3 to 10 alleles and the average across loci was 6.85 alleles per locus (Table 2). The number of shared alleles between the two-species ranged from 2 to 8 alleles per locus and the average across loci was 5.26 allele per locus (Table 2). Out of the 27 SSR loci, six loci yielded both *E.g* and *E.o* specific alleles, nine loci yielded *E.g* specific alleles, five yielded *E.o* specific alleles and seven yielded no species-specific allele (Table 2). The number of *E.g* specific alleles generated from each of the evaluated SSR locus ranged from 0 to 3 *E.g* specific alleles per locus and the average across loci was 0.96 (Table 2). On the other hand, the number of *E.o* specific alleles generated from each of the evaluated SSR locus also ranged from 0 to 3 *E.o* specific alleles per locus but the average across loci was 0.63 (Table 2). The value of the polymorphic information content (PIC) ranged from 0.10 for mEgCIR2813 to 0.79 for the mEgCIR0588 SSR primer pairs (Table 2). The average PIC across the two oil palm species and the 27 SSR marker loci was 0.56 (Table 2).

The observed number of alleles (Na), the number of effective allele (Ne), the Shannon's information index (I), the observed (Ho) and the expected (He) heterozygosity, and the fixation index (F) for each SSR marker locus across either the two oil palm species (*E.g* and *E.o*) were presented in Table 3. Most of the genetic parameters were similar between *E.g* and *E.o* oil palm species. However, *E.o* oil palm species have higher average number of effective allele (Ne) than that of *E.g* (Table 3). Moreover, the average number of alleles (Na) and the number of effective alleles (Ne) obtained from these results was higher than those obtained by Tinche et al. (2014) (Cameroon population = 2.76), Abdullah et al. (2011) (Nigeria population = 3.3), Ajambang et al. (2012) (Cameroon's natural population = 4.71) and Billotte et al. (2001) (La Me × Dura Deli = 5.25).

Genetic differences among oil palm species

Using analysis of molecular variance, the total genetic variance can be partitioned into two variance sources (among individuals within species and between species). Variation among individual within the oil palm species contributed 89%, while variation between species contributed only 11% of the total genetic variance (Figure 1). The values of variation between species indicated the existence of genetic dissimilarity.

As for the population parameters across individuals within the species, the *E.g* population has a higher number

of alleles (Na) than the *E.o* population but the number of effective allele (Ne), number of private alleles, and the Shannon diversity index (I) in *E.g* was lower than in *E.o* (Table 4). Furthermore, the genetic heterozygosity in the *E.o* population was wider than in the *E.g* population (Table 4). The high number of alleles reflects the high genetic diversity of the population and the existence of wide genetic variability determines the genetic progress in population improvement program. The genetic distance value between *E.g* and *E.o* oil palm species accessions based on the evaluated 27 SSR marker loci was 0.4 (Table 4).

Results of the Neighbor-Joining analysis of two oil palm species based on 27 SSR loci were presented in Figure 2 for *E.g* and Fig.3 for *E.o* oil palm species. Using six *E.o* accessions as outgroups, groupings of the evaluated *E.g* accessions were presented in Figure 2. The Deli Dura (DeD) group was further divided into 3 sub-groups (DeD-1, 2, and 3) while the Dumpy Dura (DyD) group was divided into 4 sub-groups (DyD1, 2, 3, and 4). For each of the Pisifera type, Avros Pisivera (AVP), Binga Pisifera (BIP), and Dumpy Avros Pisifera groups were divided into two sub-groups (AVP1 and 2; BIP 1 and 2; and DyAP 1 and 2), respectively. All accessions belonging to the Angola Pisifera (ANP) belonged in one group. As for the Tenera type, the Angola Tenera (ANT) belonged in one group while Dumpy Avros Tenera (DyAT) group was further sub-divided into two sub-groups (DyAT 1 and 2).

Using six *E.g* accessions as outgroups, groupings of the evaluated *E.o* accessions were presented in Figure 3. The *E.o* accessions originated from Manaus were further divided into two sub-groups (*E.o* Manaus-1 and 2) while those from Tepe were divided into six sub-groups (*E.o* Tepe-1, 2, 3, 4, 5, and 6), respectively. The *E.o* Manaus-1 was genetically closely related to the *E.o* Tepe-1 while the *E.o* Manaus-2 was closely related to the *E.o* Tepe-6 (Figure 3).

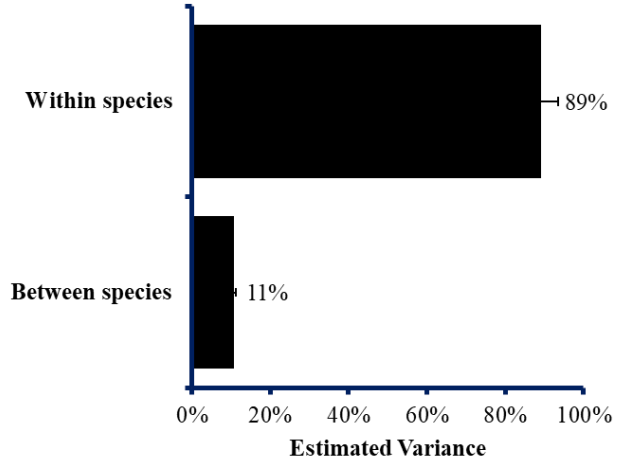

Figure 1. Partitioning of variation between and within the two species of oil palm (*Elaeis guineensis* and *E. oleifera*)

Table 2. The number of total alleles, shared alleles, and species-specific alleles and the polymorphic information content (PIC) for each SSR locus in two oil palm species (*Elaeis guineensis* and *E. oleifera*).

Locus name	Number of total alleles	Number of shared alleles	Number of species specific alleles: *E. guineensis*	Number of species specific alleles: *E. oleifera*	PIC
mEgCIR3399	9	5	2	2	0.55
mEgCIR3788	8	5	3	0	0.69
mEgCIR0878	9	8	1	0	0.67
mEgCIR3847	7	5	2	0	0.65
mEgCIR0555	10	8	1	1	0.73
mEgCIR0783	8	6	2	0	0.66
mEgCIR0801	6	3	2	1	0.34
mEgCIR1729	6	6	0	0	0.56
mEgCIR3358	6	6	0	0	0.50
mEgCIR3691	6	4	0	2	0.61
mEgCIR3543	5	5	0	0	0.60
mEgCIR0803	6	3	0	3	0.27
mEgCIR3902	8	7	0	1	0.74
mEgCIR2347	7	5	2	0	0.67
mEgCIR3350	7	7	0	0	0.44
mEgCIR1730	6	6	0	0	0.60
mEgCIR3745	8	4	3	1	0.53
mEgCIR3300	7	5	1	1	0.61
mEgCIR3428	7	4	2	1	0.59
mEgCIR3546	3	3	0	0	0.33
mEgCIR3376	7	7	0	0	0.54
Eg12M411	7	6	1	0	0.73
mEgCIR0588	10	8	2	0	0.79
mEgCIR2813	4	2	0	2	0.10
Eg9M4035	7	6	1	0	0.73
mEgCIR0243	6	4	0	2	0.51
mEgCIR3534	5	4	1	0	0.51
Average	6.85	5.26	0.96	0.63	0.56

Table 4. Population genetic differences between the evaluated *Elaeis guineensis* and *E. oleifera* accessions

Parameters	*Elaeis guineensis*	*Elaeis oleifera*
Na (Number of alleles)	5.78 ± 0.33	5.22 ± 0.25
Ne (Number of effective alleles)	2.47 ± 0.17	3.16 ± 0.20
No private alleles	3.00 ±0.20	3.56 ±0.26
I (Shannon diversity index)	1.06 ± 0.07	1.25 ± 0.07
Genetic heterozygosity	0.54 ± 0.04	0.64 ± 0.03
Genetic distance between *E.g* and *E.o*	0.4	

Table 3. Genetic parameters for each locus in two oil palm species *Elaeis guineensis* (*E.g*) and *E. oleifera* (*E.o*) for the 27 SSR marker loci

SSR locus name	Na		Ne		I		Ho		He		F	
	E.g	*E.o*	*E.g*	*E.o*	*E.g*	*E.o*	*E.g*	*E.o*	*E.g*	*E.o*	*E.g*	*E.o*
mEgCIR3399	6	7	1.97	2.99	1.07	1.37	0.46	0.8	0.49	0.67	0.06	-0.2
mEgCIR3788	8	5	3.1	4.02	1.39	1.47	0.55	0.34	0.68	0.75	0.18	0.54
mEgCIR0878	8	7	2.49	4.83	1.14	1.69	0.56	0.73	0.6	0.79	0.07	0.08
mEgCIR3847	6	5	2.78	3.23	1.14	1.3	0.7	0.73	0.64	0.69	-0.1	-0.06
mEgCIR0555	8	8	3.45	5.57	1.48	1.84	0.69	0.27	0.71	0.82	0.02	0.67
mEgCIR0783	7	5	2.66	3.06	1.26	1.27	0.58	0.82	0.62	0.67	0.08	-0.22
mEgCIR0801	5	3	1.17	2.02	0.38	0.73	0.09	0.67	0.15	0.5	0.37	-0.32
mEgCIR1729	6	5	2	3.36	1.04	1.36	0.38	0.55	0.5	0.7	0.25	0.22
mEgCIR3358	5	5	2.09	2.75	0.95	1.14	0.56	0.73	0.52	0.64	-0.07	-0.14
mEgCIR3691	4	5	2.67	3.03	1.12	1.25	0.54	0.44	0.63	0.67	0.14	0.34
mEgCIR3543	4	4	2.6	2.88	1.08	1.13	0.56	0.55	0.62	0.65	0.09	0.16
mEgCIR0803	3	5	1.24	1.86	0.39	0.88	0.13	0.11	0.19	0.46	0.31	0.75
mEgCIR3902	7	7	3.49	3.91	1.47	1.61	0.45	0.73	0.71	0.74	0.36	0.03
mEgCIR2347	6	5	3.17	3.52	1.35	1.35	0.61	0.59	0.68	0.72	0.1	0.17
mEgCIR3350	6	6	1.74	2.12	0.93	1.03	0.48	0.54	0.43	0.53	-0.12	-0.02
mEgCIR1730	5	5	2.33	3.72	0.99	1.4	0.84	0.77	0.57	0.73	-0.47	-0.06
mEgCIR3745	7	5	2.12	2.75	1.1	1.24	0.46	0.41	0.53	0.64	0.13	0.36
mEgCIR3300	6	5	2.56	3.52	1.11	1.38	0.45	0.3	0.61	0.72	0.26	0.59
mEgCIR3428	6	5	1.69	3.71	0.74	1.45	0.45	0.77	0.41	0.73	-0.09	-0.05
mEgCIR3546	3	3	1.76	1.29	0.69	0.46	0.22	0.06	0.43	0.23	0.49	0.72
mEgCIR3376	6	6	1.9	2.83	0.93	1.26	0.25	0.69	0.47	0.65	0.48	-0.06
Eg12M411	7	6	3.16	3.69	1.38	1.47	0.5	0.86	0.68	0.73	0.27	-0.18
mEgCIR0588	9	7	4.73	5.06	1.75	1.68	0.61	0.13	0.79	0.8	0.23	0.84
mEgCIR2813	2	3	1.02	1.35	0.05	0.46	0	0.02	0.02	0.26	1	0.94
Eg9M4035	7	6	4.06	3.6	1.58	1.47	0.44	0.45	0.75	0.72	0.42	0.37
mEgCIR0243	4	4	2.49	2.13	1.02	0.9	0.34	0.3	0.6	0.53	0.43	0.43
mEgCIR3534	5	4	2.18	2.6	0.99	1.13	0.3	0.34	0.54	0.61	0.45	0.44
Average	5.78	5.22	2.47	3.16	1.06	1.25	0.45	0.51	0.54	0.64	0.20	0.23

Note: Na = number of alleles, Ne = number of efective alleles, I = Shannon's information index, Ho = observed heterozigosity, He = expected heterozigosity, F = fixation index. *E.g-Elaeis guineensis* and *E.o-E. oleifera*

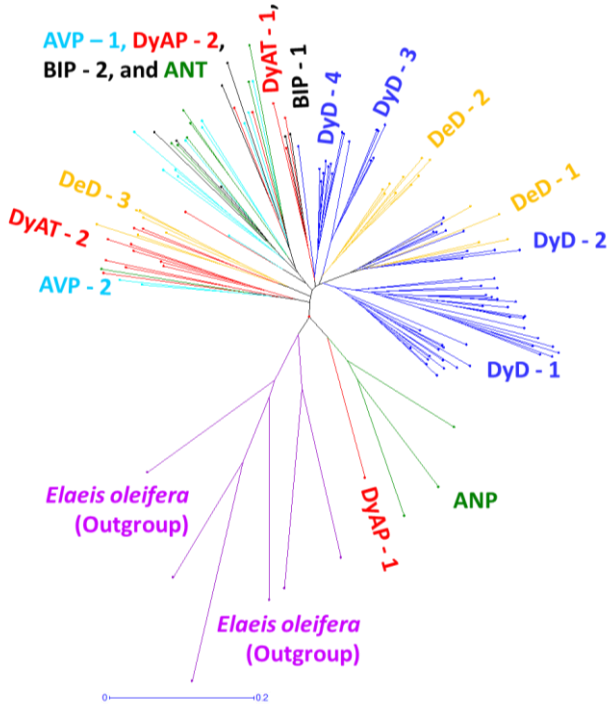

Figure 2. Neighbor-Joining Tree of various groups of the *Elaeis guineensis* species of oil palm. The tree was constructed based on 27 SSRs and using six accessions of *E. oleifera* species as outgroups. ANP-Angola Pisifera, ANT-Angola Tenera, AVP-Avros Pisifera, BIP-Binga Pisifera, DeD-Deli Dura, DyAP-Dumpy Avros Pisifera, and DyAT-Dumpy Avros Tenera, and DyD-Dumpy Dura

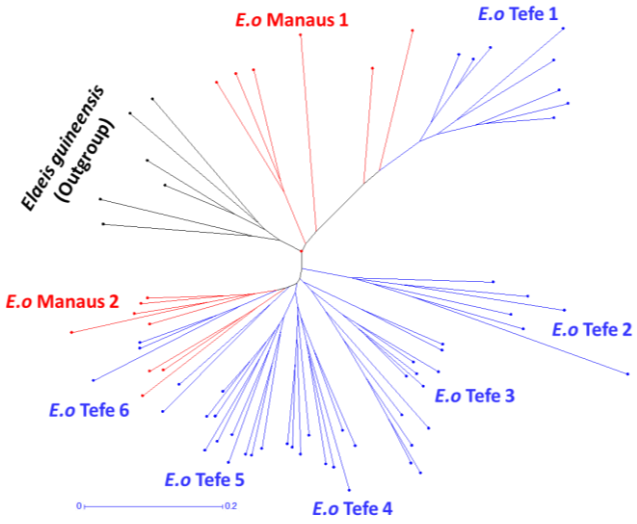

Figure 3. Neighbor-Joining Tree of various groups of the *Elaeis oleifera* species of oil palm. The tree was constructed based on 27 SSRs and using six accessions of *E. guineensis* species as outgroups. *E.o* Manaus-accessions of *E. oleifera* from Manaus region and *E.o* Tefe-accessions of *E. oleifera* from Tefe region, South America

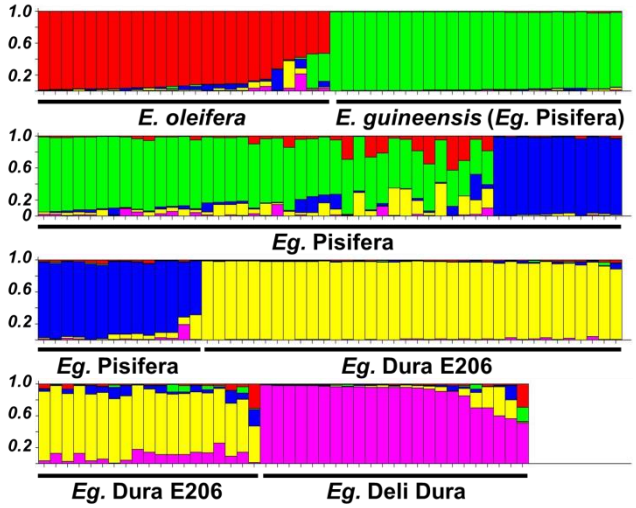

Figure 4. Population structure of two species oil palm based on 27 SSR loci

The Structure analysis is used to visualize the proportion of the genome in each oil palm individual and the result was presented in Figure 4. Some accessions in the *E.o* species still shared the small proportion of the *E.g* genome as represented by few shared SSR alleles. Similarly, some accessions of the *E.g* also shared the small proportions of the *E.o* genome as represented by few shared SSR alleles. The presence of shared alleles between the two oil palm species in this study supported the Barcelos et al. (2002) hypothesis which stated that the two oil palm species were separated due to microevolution.

Identification of *E.o* specific markers for backcrossing program

The *E.o* species-specific SSR alleles could potentially be used to assist and monitor introgression of the *E.o* genomic fragments in the marker-assisted backcross breeding. The success of introgressing *E.o* genomic fragments into *E.g* genetic background was monitored by evaluating the presence of *E.o* specific SSR alleles among the *E.o* × *E.g* // *E.*g backcross progenies. Figure 5 presented the results of the *E.o* specific SSR alleles testing to identify backcross progenies carrying the *E.o* genomic fragments within the *E.g* genetic background. Results presented in Figure 5 clearly indicated the effectiveness of using the identified *E.o* specific alleles to point out the correct backcross progenies (Figure 5). Sample accessions no. 1-4, 9-10, 20 and 23 carried the *E.o* specific alleles (either *E.o*-1 or *E.o*-2 allele) in addition to the *E.g* specific allele (*E.g*-1 allele).

Discussion

The *E.g* accessions belonging to Mekarsari Research Station were divided into three major types based on the shell thickness (Dura, Pisifera or Tenera). Recent results showed that the three types of African oil palm (*E.g*), i.e. dura (thick-shelled), pisifera (shelless) and tenera (thin-shelled, which is a dura × pisifera hybrid) could be attributed to two independent mutations in the regulatory gene (SEEDSTICK, STK gene) controlling ovule identity and seed development (Singh et al. (2013). Using the homozygosity mapping by sequencing method, Singh et al. (2013) found two mutations occurring independently which determine the oil palm phenotype as the dura, pisifera, or tenera type. The results of Singh et al. (2013) provided a clue about the origin of the dura, pisifera, and tenera in *E. guineensis* of cultivated and wild palms from sub-Saharan Africa. The differences of shell thickness characteristic are considered to have an important adaptation value for the reproductive success in *E.g* species but not in the *E.o* species (Sing et al. 2013).

The oil palm germplasms in Mekarsari Research Station generally has high genetic diversity (1.06 ± 0.07 for *E. guineensis* and 1.25 ± 0.07 for *E. oleifera*). The differences in genetic diversity between these oil palm species may be related to the number evaluated samples in this study. The *E.o* accessions used in this study were less than those of the *E.g* accessions. We also found that oil palm accessions existed in Mekarsari Research Station (either *E.g* or *E.o* species) showed higher diversity than those of the previous studies as reported by Barcelos et al. (2002).

All of the SSR loci used in this study were developed from *E.g* genome (Billote et al. 2005). However, the evaluated SSR markers could generate amplicons from almost all of the *E.o* accessions. Previous studies also reported that there are some genome similarities between *E.g* and *E.o* (Singh et al. 2009; Singh et al. 2013; Montoya et al. 2014; Ithnin et al. 2017). The genetic similarity between oil palm species is still high despite the geographical distance between the natural distribution of these species. The high degree of similarity is an indicator which justifies the high crossability between *E.g* and *E.o.* However, analysis of the genetic structure of *E.g* and *E.o* accessions based on 27 SSR loci revealed that the two are different oil palm species.

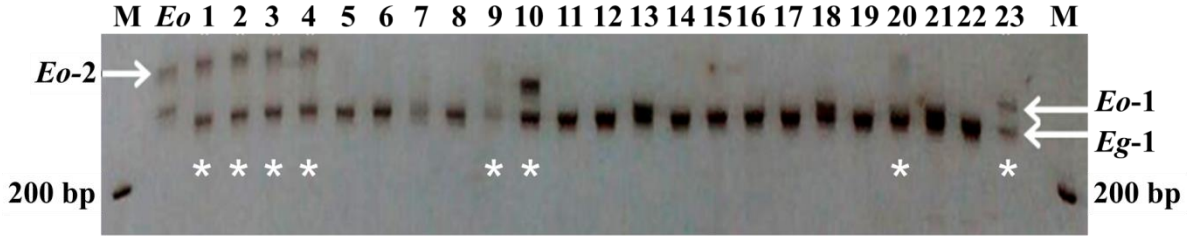

Figure 5. Visualization of the *E. guineensis* (*E.g*) and *E. oleifera* (*E.o*)-specific SSR alleles. * Indicates the Dura Composite introgression line (*E.o* × *E.g* // *E.g*) individuals carrying the *E. oleifera* specific alleles in the *E.g* genetic background; The SSR marker locus used to generate the *Elaeis oleifera*-specific SSR alleles was mEgCIR3376

In contrast, analysis of the genome proportion using STRUCTURE further showed that the two oil palm has few similarities in the genetic background. Barcelos et al. (2002) also found few common alleles between these species based on RFLP and AFLP marker analysis. They suggested that these two species present a relatively small genetic divergence, suggesting that the gene flow interruption between them is relatively recent. The micro-evolutionary mechanism which differentiated between the two oil palm species is the common mechanism which resulted in the wide genetic diversity in palm kingdom.

For oil palm genetic improvement program, *E. oleifera* germplasm offers some desirable characters that are absent in the *E. guineensis* gene pool. Those characters include very low height increment and high oil quality. Introgression and backcrossing programs between *E.o* to *E.g* using a large breeding population should be conducted for developing a new plant type of commercial oil palm. For backcross breeding purposes, utilization of the species-specific SSR markers associated with particularly desirable characters is essentials. In this study, we have successfully demonstrated the use of species-specific SSR markers to identify the legitimate progenies derived from *E.o* × *E.g* // *E.g* back cross. Therefore, the *E.o* specific SSR marker identified in this study could be used for developing marker-assisted backcrossing (MAB) to accelerate and monitor the *E.o* genomic fragment introgression process into *E.g* genetic background.

REFERENCES

Ajambang W, Sudarsono, Asmono D, Toruan-Mathius N. 2012. Microsatellite markers reveal Cameroon's wild oil palm population as a possible solution to broaden the genetic base in the Indonesia-Malaysia oil palm breeding programs. African J Biotechnol 11: 13244-13249. DOI: 10.5897/AJB11.3897.

Bakoume C, Wickneswari R, Rajanaidu N, Kushairi A, Amblard P, Billotte N. 2007. Allelic diversity of natural oil palm (*Elaeis guineensis* Jacq.) populations detected by microsatellite markers: implications for conservation. Plant Genet Res: Characterization and Utilization 5 (2): 104-107.

Bakoume C, Wickneswari R, Siju S, Rajanaidu N, Kushairi A, Billotte N. 2014. Genetic diversity of the world's largest oil palm (*Elaeis guineensis* Jacq.) field gene bank accessions using microsatellite markers. Genet Res Crop Evol J 62: 349-360. DOI: 10.1007/s10722-014-0156-8.

Barcelos E, Amblard P, Berthaud J, Seguin M. 2002. Genetic diversity and relationship in American and African oil palm as revealed by RFLP and AFLP molecular markers. Pesqui Agropecu Brasileira 37: 1105-1114. DOI: 10.1590/S0100-204X2002000800008.

Barcelos E, Rios SA, Cunha RNV, Lopes R, Motoike SY, Babiychuk E, Skirycz A and Kushnir S. 2015. Oil palm natural diversity and the potential for yield improvement. Front Plant Sci 6: 190. DOI: 10.3389/fpls.2015.00190

Billotte N, Marseillac N, Risterucci AM, Adon B, Brottier P, Baurens FC, Singh R, Herran A, Asmady H, Billotte C, et al. 2005. Microsatellite-based high-density linkage map in oil palm (*Elais guineensis* Jacq.). Theor Appl Genet 110: 754-765.

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

Brody JR, Kern SE. 2004. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. BioTechniques 36: 214-216.

Budiman LA, Apriyanto A, Pancoro A, Sudarsono S. 2019. Genetic diversity analysis of Tenera× Tenera and Tenera× Pisifera Crosses and D self of oil palm (*Elaeis guineensis*) parental populations originating from Cameroon. Biodiversitas 20 (4): 937-949.

Cadena T, Prada F, Perea A, Romero HM. 2013. Lipase activity, mesocarp oil content, and iodine value in oil palm fruits of *Elaeis guineensis*, *Elaeis oleifera*, and the interspecific hybrid OxG (*E. oleifera* × *E. guineensis*). J Sci Food Agric 93: 674-680. DOI: 10.1002/jsfa.5940.

Caicedo LPM, Perez SEB. 2017. Morphological characterization of the American oil palm collection *Elaeis oleifera* (Kunth) Cortes. Acta Agron 66 (1): 135-140.

Corley RHV, Tinker PB. 2003. The Oil Palm. 4th ed. Blackwell Science, Oxford, UK.

Corley RHV. 2009. How much palm oil do we need? Environ Sci Pol 12: 134-139.

Creste S, Neto AT, Figueira A. 2001. Detection of single sequence repeats polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Mol Biol Rep 19: 299-306.

Din AM, Rajanaidu N, Jalani BS. 2000. Performance of *Elaeis oleifera* from Panama, Costa Rica, Colombia and Honduras in Malaysia. J Oil Palm Res 12 (1): 71-80.

Dransfield J, Uhl NW, Asmussen CB, Baker WJ, Haarley MM, Lewis CE. 2005. A new phylogenic classification of the palm family *Arecaceae*. Kew Bulletin. 60: 559-569.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.

Hardon J, Tan G. 1969. Interspecific hybrids in the genus *Elaeis* I. crossability, cytogenetics, and fertility of F1 hybrids of *E. guineensis* × *E. oleifera*. Euphytica 18: 372-379. DOI: 10.1007/BF00397784

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 (7): 1274-84.

Ithnin M, Teh CK, Ratnam W. 2017. Genetic diversity of Elaeis oleifera (HBK) Cortes populations using cross-species SSRs: implication's for germplasm utilization and conservation. BMC Genet 18: 37. DOI: 10.1186/s12863-017-0505-7

Montoya C, Cochard B, Flori A, Cros D, Lopes R, Cuellar T, et al. 2014. Genetic architecture of palm oil fatty acid composition in cultivated oil palm (*Elaeis guineensis* Jacq.) compared to its wild relative *E. oleifera* (H.B.K) Cortes. PLoS ONE 9: e95412. DOI: 10.1371/journal.pone.0095412.

Moretzsohn MC, Ferreira MA, Amaral ZPS, Coelho PJA, Grattapaglia D, Ferreira ME. 2002. Genetic diversity of Brazilian oil palm (*Elaeis oleifera* H.B.K.) germplasm collected in the Amazon Forest. Euphytica 124: 35-45.

Natawijaya, A., SW. Ardie, I. Maskromo, M. Syukur, A. Hartana, S. Sudarsono. 2017. Variability of some agronomic traits within and between Dura elite family collection of Mekarsari Fruit Garden. Buletin Palma 18 (1): 23-32. [Indonesian]

Nei M. 1972. Genetic distance between populations. Am Nat 106: 282-292.

Okoye MN, Uguru MI, Bakoumé C, Singh R, and Okwuagwu CO. 2016. Assessment of genetic diversity of NIFOR oil palm main breeding parent genotypes using microsatellite markers. Amer J Plant Sci 7: 218-237.

Pamin K. 1998. A hundred and fifty years of oil palm development in Indonesia: From the Bogor Botanical Garden to the Industry. Proceedings 1998 International Oil Palm Conference: Commodity of the past, today and the future. Pp 3-25.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 28: 2537-2539. DOI: 10.1093/bioinformatics/bts460.

Rajanaidu N, Khusairi A, Rafii M, Din M, Maizura I, Jalani BS. 2000. Oil palm genetic resources. In: Basiron Y, Jalani BS, Chan KW (eds.). Advances in Oil Palm Research. Malaysian Palm Oil Board, Kuala Lumpur.

Schaal BA, Gaskin JF, Caicedo AL. 2003. The Wilhelmine E. Key 2002 Invitational Lecture. Phylogeography, haplotype trees, and invasive plant species. J Hered 94: 197-204.

Singh R, Low ET, Ooi LC, Ong-Abdullah M, Ting NC, Nagappan J, Nookiah R, Amiruddin MD, Rosli R, Manaf MA, Chan KL, Halim MA, Azizi N, Lakey N, Smith SW, Budiman MA, Hogan M, Bacher B, Van Brunt A, Wang C, Ordway JM, Sambanthamurthi R,

Martienssen RA. 2013. The oil palm SHELL gene controls oil yield and encodes a homologue of Seedstick. Nature 500: 340-344. DOI: 10.1038/nature12356

Singh R, Tan SG, Panandam JM, Rahman RA, Ooi LC, Low ET, et al. 2009. Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. BMC Plant Biol 9: 114. DOI: 10.1186/1471-2229-9-114.

Sunilkumar K, Mathur RK, Sparjanbabu DS, Pillai RSN. 2015. Evaluation of interspecific oil palm hybrids for dwarfness. J Plantation Crops 43 (1): 29-34.

Tinche, Asmono D, Dinarty D, Sudarsono. 2014. Genetic diversity of oil palm (*Elaeis guineensis*) populations from Nigeria based on SSR marker analysis Bulletin Palma 15 (1): 14-23. [Indonesian].