

Development and Characterization of *Elaeis oleifera* Microsatellite Markers (Pembangunan dan Pencirian Penanda Mikrosatelit *Elaeis oleifera*)

NOORHARIZA MOHD ZAKI*, ISMANIZAN ISMAIL, ROZANA ROSLI,
TING NGOOT CHIN & RAJINDER SINGH

ABSTRACT

Ten *Elaeis oleifera* microsatellite markers were developed and characterised from 1500 sequences of the *E. oleifera* genomic library. The markers were utilised to assess the genetic diversity of *E. oleifera* germplasm collections from four South American countries (Colombia, Costa Rica, Panama and Honduras). The number of alleles per-locus varied from 2 to 11 and the observed and expected heterozygosity ranged from 0.0685 to 0.9853 and 0.1393 to 0.8216 respectively. Majority of the markers showed transferability to *Elaeis guineensis* while two markers showed transferability across *Arecaceae* taxa. These *E. oleifera* microsatellite markers are expected to become useful tools to determine the population structure and conservation of *E. oleifera* populations.

Keywords: *E. oleifera*; genomic; microsatellite; transferability

ABSTRAK

Sebanyak sepuluh penanda mikrosatelit genomik *Elaeis oleifera* telah dibangunkan dan dicirikan daripada sejumlah 1500 klon di dalam perpustakaan genomik *E. oleifera*. Keupayaan penanda yang dibangunkan ini digunakan bagi menganalisis kepelbagaian genetik di dalam koleksi germplasma dari empat negara di Amerika Selatan (Colombia, Costa Rica, Panama dan Honduras). Bilangan alel per-locus yang dicerap adalah di dalam lingkungan 2 hingga 11 manakala heterozigositi yang dicerap dan dijangka adalah masing-masing 0.0685 hingga 0.9853 dan 0.1393 hingga 0.8216. Majoriti penanda mikrosatelit menunjukkan kebolehpindahan kepada *Elaeis guineensis* manakala dua penanda berjaya menunjukkan kebolehpindahan merentasi takson famili *Arecaceae*. Penanda-penanda mikrosatelit yang berjaya dibangunkan ini di jangka dapat menjadi penanda dalam menentukan struktur populasi dan pemuliharaan populasi *E. oleifera*.

Kata kunci: *E. oleifera*; genomik; kebolehpindahan; mikrosatelit

INTRODUCTIONS

The oil palm belongs to the genus *Elaeis* and comprises of two species; the economically important *E. guineensis*, native to Africa and *E. oleifera*, found in Central and South America. The *E. oleifera* species is seen as a promising genetic resource and is currently being used in oil palm breeding programs for the development of inter-specific hybrids. The agronomic traits in *E. oleifera* that have attracted the attention of oil palm breeder includes slow trunk growth, improved oil quality (Moretzsohn et al. 2002) and resistance to disease such as *Fusarium* wilt and lethal yellowing (Hardon & Tan 1969)

Microsatellite markers are being widely applied for studies such as genome mapping, marker assisted selection of crop plants and molecular ecology studies. The development of microsatellite markers will be very valuable for genetic studies of oil palm. This is especially so, since microsatellite markers show co-dominant inheritance, multiallelic in nature, abundant,

allow extensive genome coverage and can be detected using PCR. Microsatellites also appear to be the most promising molecular marker system for understanding the population genetic structure and gene flow (Singh et al. 2008).

Molecular marker studies on oil palm have mainly focused on the commercial *E. guineensis* although some studies have been reported for *E. oleifera* (Moretzsohn et al. 2002; Singh et al. 2008). The limited studies on *E. oleifera* have either used neutral markers such as random amplified polymorphic DNA (RAPD) or microsatellite markers developed from *E. guineensis* (Bilotte et al. 2001; Singh et al. 2008). Thus, in this study, we report the development of genomic microsatellite markers from *E. oleifera* and their application in characterizing four *E. oleifera* germplasm collections. We also investigated the transferability of the microsatellite loci to the other species of the same genus and across palm taxa, respectively.

TABLE 1. Characterisation of the 10 *E.oleifera* genomic microsatellite loci based on 75 individuals from 4 countries in South America

Locus	Primer pair sequence (5'-3')	Probe Database UID	Ta (°C)	Repeat motif	Size (bp)	A _o	H _o	H _e	P-values
sMo00018	F: TTAAATGAGAGAGAGAGAGGAC R: TGGAGCCATGAGAAAGAGTA	9947963	54	(CT) ₁₄	242-252	3	0.2778	0.4507	0.0146
sMo00020	F: CCTTCTCCTCCCTCCTTTTG R: CCTCCCTCCCTCTCACCATA	9947964	58	(AG) ₁₅	182-208	9	0.6338	0.8216	0.0092
sMo00027	F: TTACAGTTGAGGCAGTATGTCAAT R: CTGTAIGTCAAACCTTCTGCAC	9947965	50	(TC) ₁₄	184-213	6	0.4533	0.6674	0.0003
sMo00128	F: TAGCTCCAACAGTTGCCTTAT R: GGTCCCGTCTTAIGATTTAFTCT	9947966	56	(AAT) ₁₂	182-203	6	0.4167	0.7071	0.0001
sMo00129	F: TTAGTATGGGTGTCATAAGTGG R: GCTTCCAGCTCCTCTTCTACC	9947967	56	(TTC) ₁₃	212-232	7	0.3378	0.7032	0.0000
sMo00130	F: TAAGCAAAAATCAGGGCCTC R: GGCTGGTGAAAATAAGTTTACAAAAG	9947968	56	(AAG) ₁₁	173-214	11	0.9853	0.8068	1.0000
sMo00132	F: ATAGCCAGAGGGCAAAACTGT R: GCAACACACGGGACTCAAAACTA	9947969	56	(TTA) ₁₃	159-180	2	0.0685	0.1393	0.1433
sMo00134	F: TCCCAATAGTCGTTACAAACCAG R: GATTAGCAAAAAGGGCAAAAAGG	9947970	56	(ATTA) ₆	252-256	2	0.1944	0.4753	0.0000
sMo00138	F: AGGGTTGTCGCTCCAATTTAT R: GGCATCTTTTGTGACCTGTAGAAG	9947971	56	(TTTTC) ₆	180-194	3	0.1200	0.3588	0.0000
sMo00152	F: GGAACAAGGACAAGAAAGAAA R: TGTATCAAGCCTCAAGTATCTGG	9947972	56	(AC) ₆ (AG) ₁₁	245-256	2	0.2319	0.4999	0.0000

Notes: A_o - number of alleles detected; H_o - observed heterozygosity; H_e - expected heterozygosity

P-values <0.05 indicates significant deviation from Hardy-Weinberg equilibrium

MATERIALS AND METHODS

The *E. oleifera* genomic library was constructed by using the GeneThresher® Technology (Budiman et al. 2005). A total of 1500 *E. oleifera* genomic sequences were screened for microsatellites containing mono-, di-, tri-, tetra-, penta- and hexanucleotide repeats. The identification and localization of the microsatellite markers were performed by using the MISA software as described by Thiel et al. (2003). Primers were designed using the Primer3 program (Rozen & Skaletsky 2000). Ten primers were randomly chosen from all repeat types (except mononucleotide repeats) to be tested against four *E. oleifera* germplasm collections namely Colombia, Costa Rica, Panama and Honduras, each of which consisted of 15-20 palms.

The 5' end forward primer was labelled with γ -³³P dATP (GE Healthcare Biosciences, UK, 3000Ci/mmol) and PCR was performed essentially as described by Singh et al. (2008). The PCR products were separated on 6.0% acrylamide gel containing 7 M urea in 5X TBE buffer at constant power of 1,600 V for 3 hours. Sizing of each allele was done using 30-330 bp AFLP® DNA ladder (Invitrogen, USA).

Only fragments that could be visibly scored were analyzed using POPGENE version 1.32 (Yeh & Boyle 1999). Among the genetic variability measures calculated were number of alleles per locus, expected and observed heterozygosity (Nei 1978). Exact Hardy-Weinberg equilibrium (HWE) tests were carried out by using the software GENEPOP version 4.0 (Raymond & Rousset 1995). Six *E. oleifera* microsatellite primers which produced clear banding profiles in *E. oleifera* were utilised to study the cross species/genera amplification against *E. guineensis*, *Cocos nucifera* and *Jessinia bataua*.

RESULTS AND DISCUSSION

Of the 363 *E. oleifera* clones containing microsatellite motifs, 360 (99.2%) were suitable for primer design. With

the exclusion of mononucleotides, the prevalent repeats were di-nucleotides (61.5%), followed by tri-nucleotides (26.9%), penta-nucleotides (5.4%), tetra-nucleotides (4.6%), and hexa- and hepta-nucleotide (0.8%) repeats. Among the di-nucleotide motif, the AC/GT motif was the most prevalent (47.5%), followed by AT/AT (45%) and AC/GT (7.5%). The abundance of the AC/GT motif was reported previously in the *E. guineensis* oil palm species by Singh et al. (2008) and Low et al. (2008), as well as in other plant species such as apricot and peach (Jung et al. 2005) and coffee (Aggarwal et al. 2007)

The number of alleles per locus (A_o) varied from two to 11 with an average of 5.1 alleles per marker. The detected alleles per locus (5.1) are higher than that reported for *Elaeis* using EST-microsatellites (mean=2.56; Singh et al. 2008) but slightly lower compared to what was reported by Billotte et al. (2001) (mean=5.25) in *E. guineensis* accession using genomic microsatellites. The observed and expected heterozygosities ranged from 0.0685 to 0.9853 and 0.1393 to 0.8216, respectively (Table 1). The observed heterozygosity for the majority of the loci (90%) was lower than expected. Eight loci deviated significantly from the HWE. This is not surprising because the germplasm used in this study comprised small groups of palms scattered across the four South American countries (Rajanaidu 1985). This could have encouraged inbreeding resulting in a relatively high homozygous genome.

Six newly developed *E. oleifera* genomic microsatellite markers which produced prominent banding profiles in *E. oleifera* on a polyacrylamide gel (SSR analysis) were used to study the cross-transferability in Arecaceae taxa (Table 2). Successful amplification (transferability) of either similar or varying sized fragments was obtained with all the markers within the *Elaeis* genus. Of the six primers, two (sMo00130 and sMo00138) were also amplifiable in *Cocos nucifera* and *Jessinia bataua*. The ability of these markers to amplify fragments with either similar or different sizes indicates their efficiency in revealing sequence conservation among the different species in Arecaceae family. Similar trend was also reported by

TABLE 2. Transferability of *E.oleifera* microsatellite loci across different species and genus in the Arecaceae family. (allele size observed (bp) based on a minimum of 2 individuals)

Family	Arecaceae			
	<i>Elaeis</i>		<i>Cocos</i>	<i>Jessinia</i>
	Genus	<i>Elaeis</i>	<i>Cocos</i>	<i>Jessinia</i>
SSR locus/ Species	<i>oleifera</i>	<i>guineensis</i>	<i>nucifera</i>	<i>bataua</i>
sMo00027	184-213	184-213	NA	NA
sMo00129	218-232	212	176	NA
sMo00130	173-214	173-214	200-220	180
sMo00134	252-256	252-256	NA	NA
sMo00138	192-194	180-186	186-190	180
sMo00152	256-258	254	NA	NA

Note: NA- No amplification

Singh et al. (2008) for the microsatellites developed from the *E. guineensis* expressed sequence tags (ESTs) collections.

CONCLUSION

This preliminary study we believe is the first report of SSR markers from *E. oleifera*. The markers will be useful tools for genetic and evolutionary studies in oil palm. The transferability across oil palm species and taxa reveals the suitability of *E. oleifera* genomic SSR markers for comparative genetic studies in the Arecaceae family.

ACKNOWLEDGEMENTS

The authors wish to thank the Director-General of MPOB for permission to publish this manuscript.

REFERENCES

- Aggarwal, R.K., Hendre, P.S., Varshney, R.K., Bhat, P.R., Krishnakumar, V. & Singh L. 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor. Appl. Genet.* 114: 359-372.
- Billotte, 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: 413-25.
- Budiman, M.A., Rajinder, S., Low, E.T.L., Nunberg, A., Citek, R., Rohlfling, T. J., Bedell, A., Lakey, N.D., Martienssen, R.A. & Cheah, S.C. 2005. Sequencing of the oil palm genespace. *Proceeding of Agriculture, Biotechnology and Sustainability Conference*. Petaling Jaya, Malaysia. 628-639.
- Hardon, J.J. & G. Y. Tan, 1969. Interspecific hybrids in the *Elaeis*. I. Cross-ability, cytogenetics and fertility of F₁ hybrids of *E. guineensis* x *E. oleifera*. *Euphytica* 18: 372-379.
- Jung, S. Albert, A., Jesudurai, C., Tomkins, J. & Main, D. 2005. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. *Funct. Integr. Genomics* 5: 136-143.
- Low, E.T.L., Halimah, A., Boon, S.H., Elyana, M.S., Tan, C.Y., Ooi, L.C.L., Cheah, S.C., Abdul Rahim, R., Wan, K.L. & Singh, R. 2008. Oil palm (*Elaeis guineensis* Jacq.) tissue culture ESTs: identifying genes associated with callogenesis and embryogenesis. *BMC Plant Biol.* 8: 62.
- 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. 1978. Estimation of average heterozygosity and genetic distance from a small number of individual. *Genetics* 89: 583-590.
- Rajanaidu, N. 1985. *Elaeis oleifera* collection in Central and South America. *Proceeding of the International Workshop on Oil Palm Germplasm and Utilization*, Selangor, Malaysia. 84-94.
- Raymond, M. & Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86: 248-249.
- Rozen, S. & Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Method Mol. Biol.* 132: 365-386.
- Singh, R., Zaki, N.M., Ting, N.C., Rosli, R., Tan, S.G., Low, E.T.L., Ithnin, M. & Cheah, S.C. 2008. Exploiting an oil palm EST the development of gene-derived and their exploitation for assessment of genetic diversity. *Biologia* 63: 1-9.
- Thiel, T., Michalek, W., Varshney, R.K. & Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106: 411-422.
- Yeh, F.C. & Boyle, T. 1999. POPGENE version 1.32. *The user-friendly software for population genetic analysis*. University of Alberta and CIFOR, Calgary, Alta.
- Noorhariza Mohd Zaki*, Rozana Rosli, Ting Ngoot Chin & Rajinder Singh
Malaysian Palm Oil Board
6, Persiaran Institusi
Bandar Baru Bangi
43000 Kajang, Selangor, Malaysia
- Ismanizan Ismail
Pusat Pengajian Biosains dan Bioteknologi
Fakulti Sains dan Teknologi
Universiti Kebangsaan Malaysia
43600, Bangi, Selangor, Malaysia
- Institut Biologi Sistem (INBIOSIS)
Universiti kebangsaan Malaysia
43600, Bangi, Selangor, Malaysia

*Corresponding author; email: hariza@mpob.gov.my

Received: 7 October 2009

Accepted: 11 January 2010