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-lokus yang dicerap adalah di dalam lingkungan 2 hingga 11 manakala heterozigositi yang dicerap 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 a *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>
<thead>
<tr>
<th>Locus</th>
<th>Primer pair sequence (5’-3’)</th>
<th>Probe Database UID</th>
<th>Ta (°C)</th>
<th>Repeat motif</th>
<th>Size (bp)</th>
<th>A_o</th>
<th>H_o</th>
<th>H_e</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>sMo00018</td>
<td>F: TTAATGAGAGAGAGAGGAGCAGGAC R: TGGAGCCATGAGAAGAGTA</td>
<td>9947963</td>
<td>54</td>
<td>(CT)_14</td>
<td>242-252</td>
<td>3</td>
<td>0.2778</td>
<td>0.4507</td>
<td>0.0146</td>
</tr>
<tr>
<td>sMo00020</td>
<td>F: CTTTTCTCTCTTCCTTCCTTTTG R: CCTCCCCCTCCTCACCATA</td>
<td>9947964</td>
<td>58</td>
<td>(AG)_15</td>
<td>182-208</td>
<td>9</td>
<td>0.6338</td>
<td>0.8216</td>
<td>0.0092</td>
</tr>
<tr>
<td>sMo00027</td>
<td>F: TTACAGTGAGGCGATATGTAATGCTACAT R: CTGTATGTCAAAACCTTCTGCA</td>
<td>9947965</td>
<td>50</td>
<td>(TC)_15</td>
<td>184-213</td>
<td>6</td>
<td>0.4533</td>
<td>0.6674</td>
<td>0.0003</td>
</tr>
<tr>
<td>sMo00128</td>
<td>F: TAGCTCCAACAGCTTGCTTATGATATGA R: GCTCCGGTCTCCTGATTTTATCT</td>
<td>9947966</td>
<td>56</td>
<td>(AAT)_12</td>
<td>182-203</td>
<td>6</td>
<td>0.4167</td>
<td>0.7071</td>
<td>0.0001</td>
</tr>
<tr>
<td>sMo00129</td>
<td>F: TTAGTATGTTGTTGCTAGAATGGGGTGATAG R: GCTCCAGCTTCTGCTTCTTCTC</td>
<td>9947967</td>
<td>56</td>
<td>(TTT)_3</td>
<td>212-232</td>
<td>7</td>
<td>0.3378</td>
<td>0.7032</td>
<td>0.0000</td>
</tr>
<tr>
<td>sMo00130</td>
<td>F: TAGAGAAAAGATCGGAGGGCAGCTACGCTACG CTGCCTGAGGAGAGATAGGTTTATCAGGCA</td>
<td>9947968</td>
<td>56</td>
<td>(AAG)_11</td>
<td>173-214</td>
<td>11</td>
<td>0.9853</td>
<td>0.8068</td>
<td>1.0000</td>
</tr>
<tr>
<td>sMo00132</td>
<td>F: ATAGCCAGAGGGCAAAGACTGTG R: GCAAACACAGGGACTCAGACTA</td>
<td>9947969</td>
<td>56</td>
<td>(TTA)_13</td>
<td>159-180</td>
<td>2</td>
<td>0.0685</td>
<td>0.1393</td>
<td>0.1433</td>
</tr>
<tr>
<td>sMo00134</td>
<td>F: TCCCAATAGTGTTCATTACAAAGCAG R: GATTAGCAAAAAGGCAAAAGGGGCAATTA</td>
<td>9947970</td>
<td>56</td>
<td>(ATTA)_6</td>
<td>252-256</td>
<td>2</td>
<td>0.1944</td>
<td>0.4753</td>
<td>0.0000</td>
</tr>
<tr>
<td>sMo00138</td>
<td>F: AGGTTTGTGCTGCTCAATTATAT R: GGCAATCTTTTTGACCTGTAAGGAGGGGGT</td>
<td>9947971</td>
<td>56</td>
<td>(TTTTC)_6</td>
<td>180-194</td>
<td>3</td>
<td>0.1200</td>
<td>0.3588</td>
<td>0.0000</td>
</tr>
<tr>
<td>sMo00152</td>
<td>F: GGAACAGAGGACGAAAGAGAGAAGAA R: TGATCATGAGCCCTAAGGATCAGGGGAGG</td>
<td>9947972</td>
<td>56</td>
<td>(AC)_6(AG)_1</td>
<td>245-256</td>
<td>2</td>
<td>0.2319</td>
<td>0.4999</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

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
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-32p 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 heterozygosities (HWE) and exact Hardy–Weinberg (heterozygosity (Nei 1978)). Exact Hardy-Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)).

Among the genetic variability measures calculated were number of alleles per locus, expected and observed heterozygosities (HWE) and exact Hardy–Weinberg (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)). Exact Hardy–Weinberg equilibrium (heterozygosity (Nei 1978)).

Six newly developed 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 apricot and peach (Jung et al. 2005) and coffee (Aggarwal et al. 2007).

The number of alleles per locus (A) 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>
<thead>
<tr>
<th>Family</th>
<th>Aracaceae</th>
<th>Genus</th>
<th>E. oleifera</th>
<th>E. guineensis</th>
<th>Cocos nucifera</th>
<th>Jessinia bataua</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSR locus/Species</td>
<td></td>
<td></td>
<td>184-213</td>
<td>184-213</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>sMo00027</td>
<td>184-213</td>
<td>184-213</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sMo00129</td>
<td>218-232</td>
<td>212</td>
<td>176</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sMo00130</td>
<td>173-214</td>
<td>173-214</td>
<td>200-220</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sMo00134</td>
<td>252-256</td>
<td>252-256</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sMo00138</td>
<td>192-194</td>
<td>186-190</td>
<td>186-190</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sMo00152</td>
<td>256-258</td>
<td>254</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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**


Noorhariza Mohd Zaki*, Rozana Rosli, Ting Ngoot Chin & Rajinder Singh
Malaysian Palm Oil Board
6, Persiaran Institutio
Bandar Baru Bangi
43000 Kajang, Selangor, Malaysia

Ismanizar 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