

## SSR mining in oil palm EST database: application in oil palm germplasm diversity studies.

### ABSTRACT

This study reports on the detection of additional expressed sequence tags (EST) derived simple sequence repeat (SSR) markers for the oil palm. A large collection of 19243 *Elaeis guineensis* ESTs were assembled to give 10258 unique sequences, of which 629 ESTs were found to contain 722 SSRs with a variety of motifs. Dinucleotide repeats formed the largest group (45.6%) consisting of 66.9% AG/CT, 21.9% AT/AT, 10.9% AC/GT and 0.3% CG/CG motifs. This was followed by trinucleotide repeats, which is the second most abundant repeat types (34.5%) consisting of AAG/CTT (23.3%), AGG/CCT (13.7%), CCG/CGG (11.2%), AAT/ATT (10.8%), AGC/GCT (10.0%), ACT/AGT (8.8%), ACG/CGT (7.6%), ACC/GGT (7.2%), AAC/GTT (3.6%) and AGT/ACT (3.6%) motifs. Primer pairs were designed for 405 unique EST-SSRs and 15 of these were used to genotype 105 *E. guineensis* and 30 *E. oleifera* accessions. Fourteen SSRs were polymorphic in at least one germplasm revealing a total of 101 alleles. The high percentage (78.0%) of alleles found to be specific for either *E. guineensis* or *E. oleifera* has increased the power for discriminating the two species. The estimates of genetic differentiation detected by EST-SSRs were compared to those reported previously. The transferability across palm taxa to two *Cocos nucifera* and six exotic palms is also presented. The polymerase chain reaction (PCR) products of three primer-pairs detected in *E. guineensis*, *E. oleifera*, *C. nucifera* and *Jessinia bataua* were cloned and sequenced. Sequence alignments showed mutations within the SSR site and the flanking regions. Phenetic analysis based on the sequence data revealed that *C. nucifera* is closer to oil palm compared to *J. bataua*; consistent with the taxonomic classification.

**Keyword:** Genetic variation; Transferability; Phenetic analysis; *Elaeis guineensis*.