INHERITANCE OF SSR AND SNP LOCI IN AN OIL PALM INTERSPECIFIC HYBRID BACKCROSS (BC2) POPULATION

ZULKIFLI, Y*; RAJINDER, S*; MOHD DIN, A*; TING, N C*; RAJANAIDU, N*; KUSHAIRI, A*; MUSA, B; MOHAMAD, O‡ and ISMANIZAN, I‡‡**

ABSTRACT

The oil palm species, Elaeis oleifera *although it produces lower oil yield than its African counterpart* Elaeis guineensis*, does possess certain unique traits that are of interest to breeders.* E. oleifera *has slower annual height increment, higher level of carotenoids and unsaturated fatty acids. A number of interspecific hybrid crossing programmes have been initiated to introgress the unique characteristics of* E. oleifera *into high yielding planting materials. As such, an appropriate backcross interspecific hybrid population (BC2) could be used to develop genetic maps and identify quantitative trait loci (QTL) linked to traits of interest, which will be potentially helpful in reducing the number of breeding cycles. The objective of this study was to determine the heritability of SSR and SNP loci in oil palm. Seventy-five palms from a BC2 population were selected as the mapping population. A total of 4966 (4451 SNP and 515 SSR) markers were screened and 2135 (43%) were polymorphic and mostly contributed by the female parent. The level of polymorphism exhibited by the BC2 population was similar for both marker systems. Genotyping results revealed a sufficient number of co-dominant polymorphic markers that are important in genetic mapping studies. Chi-square test showed that 87.6% of the markers were in accordance to the expected Mendelian segregation ratios. The results indicated that they are stably inherited and suitable for genetic mapping studies and also have proven to be useful as fingerprinting tools to assess the genetic purity in controlled crosses.*

Keywords: Mendelian segregation, *Elaeis guineensis, Elaeis oleifera*, polymorphism, fingerprinting.

Date received: 25 July 2013; **Sent for revision:** 2 February 2014; **Received in final form:** 25 July 2014; **Accepted:** 4 August 2014.

* Malaysian Palm Oil Board, 6 Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia. E-mail: zulkifly@mpob.gov.my

** United Plantation Berhad, 36009 Teluk Intan, Perak, Malaysia.

‡ Faculty of Plantation and Agrotechnology, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia.

‡‡ Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

INTRODUCTION

Molecular markers are fragments of DNA linked to a certain location within the genome. These fragments can be used as molecular markers when they differ in their DNA sequence in multiple organisms or organism lines. These variations in DNA sequence, called polymorphisms, can be associated or linked with different forms (alleles) of nearby genes which may be associated with specific traits (Srivastava and Mishra, 2009). The development of cloning techniques and subsequently the polymerase chain

reaction (PCR) has increased and improved the utility of molecular markers. They are now extremely useful for identification and authentication of plant and animal species. The main advantage is that the technology is less affected by age, physiological condition of samples and environmental factors (Shaw *et al.*, 2002). Molecular markers are also not tissue-specific and only a small amount of sample is needed for detection at any phase of an organism's development. Any differences in DNA sequences near the gene can be used as a marker to locate the gene and track the desired results in a breeding programme. Molecular markers have been applied successfully in many diversity studies, fingerprinting analysis and also to localise genes associated with traits of interest such as maturity, plant height, and disease resistance (Lee *et al*., 1996; Chelkowski *et al*., 2003; Liu *et al*., 2005; Alwala *et al.*, 2006).

DNA markers can be broadly divided into three classes based on the method of their detection: hybridisation-based [restriction fragment length polymorphism (RFLP)], [single nucleotide polymorphism (SNP)]; PCR-based [random amplified polymorphic DNA (RAPD)], amplified fragment length polymorphism (AFLP), simple sequence repeat (SSR) and sequencing-based (Gupta *et al.*, 1999; Collard *et al.*, 2005). RFLP and RAPD were the first molecular markers to be widely used (Grodzicker *et al.*, 1975; Botstein *et al.*, 1980; Helentjaris *et al*., 1986; Kiss *et al.*, 1993; Adam-Blondon *et al*., 1994). However, the RFLP technique is not amenable for automation, is time-consuming, laborious, while RAPD markers are unreliable and show low reproducibility.

In recent years, SSR and SNP have gained popularity as DNA markers. SSR have been the genetic markers of choice for DNA fingerprinting, linkage map construction and population genetic studies because of their high allelic diversity and are usually selectively neutral (Smith *et al.*, 1997). SSR are tandem repeats of mono-, di-, tri-, tetra- and penta-nucleotides DNA sequences that are widely present in the genomes of plants and animals. They display high level of genetic variation due to the differences in the number of repeat units. Successful construction of an oil palm SSR enriched genomic library was first reported by Billotte *et al*. (2001) who demonstrated the potential of the SSR markers in revealing the genetic relationship of several oil palm populations.

The major constraint of using SSR markers from genomic libraries is the high development cost, technical difficulty and it can be a time-consuming process. In addition, genomic SSR markers may not be very useful in identifying genes of interest because the markers can correspond to either the transcribed region or the non-transcribed region of the genome (Zulkifli *et al.*, 2012). As an alternative, sequences that code for genes that are expressed in the genome have been exploited for developing molecular markers. These sequences, called expressed sequence tags (EST), have been instrumental in developing markers in plants and many other organisms (Kota *et al*., 2001; Kantety *et al*., 2002; Michalek *et al.*, 2002). They are partial or complete sequences of complementary DNA (cDNA) obtained from mRNA isolated from different tissues and therefore represent expressed genes in these tissues with putative function. Among the molecular markers that can be developed from EST are SNP (Rafalski, 2002) and SSR (Varshney *et al*., 2002). The availability of large numbers of EST for many important plants species has allowed an inexpensive computational approach to the development of SSR and SNP markers (Cho *et al*., 2000; Eujayl *et al*., 2001; Gupta *et al*., 2003).

EST can be used to identify candidate genes for QTL analysis of adaptive traits. EST-SSR have been identified, developed and used for many research purposes in a variety of plant species such as grape (Scott *et al*., 2000), sugar-cane (Cordeiro *et al*., 2001), rye (Hackauf and Wehling, 2002), barley (Thiel *et al*., 2003), wheat (Nicot *et al*., 2004), coffee (Valerie *et al*., 2006) and oil palm (Singh *et al*., 2008). However, the EST-SSR markers have been reported to be less polymorphic compared to the SSR markers derived from genomic libraries (Eujayl *et al*., 2002; Rajeev *et al.*, 2005; Chabane *et al.*, 2005). On the other hand, SSR themselves have some drawbacks which are: a lengthy and costly development phase, and have a relatively low throughput. Furthermore, there are difficulties for automation and data management, especially when compared to SNP.

SNP are the smallest observable unit of DNA polymorphism which can differentiate individual variations detected at the level of a single nucleotide base in the genome. SNP markers are also highly conserved. Since they occur throughout the genome, they represent the most frequent type of polymorphism and thus are available in large abundance for almost any organism including plants. For example, there are around 50 million SNP that have been identified in the human genome (Sherry *et al.*, 2001) and found approximately every kilobase in the genome. SNP markers are biallelic, have lower information content than SSR, but are more abundant in the genome, are amenable to high-throughput methods with lower genotyping error rates (Hamblin *et al.*, 2007). SNP as such, offer the potential for generating high density genetic maps and are also extremely useful for determining haplotypes for genes or quantitative trait loci (QTL). For plants, SNP are becoming popular for construction of genetic linkage maps such as for tomato (Shirasawa *et al*., 2010), cucurbita (Esteras *et al.*, 2012) and apple (Laima *et al*., 2012).

Oil palm can be distinguished into two species, *Elaeis guineensis* from Africa and *Elaeis oleifera* native to South America. *E. guineensis* is well-known as a high oil yield planting material while for *E.oleifera* the oil yield is much lower with oil to bunch ratio of 5% as compared to 25% for *E. guineensis*. However, *E. oleifera* has potential as a source to improve the competitiveness and sustainability of oil palm. *E. oleifera* exhibits a lower annual height increment, more resistant to disease, higher carotene content and higher level of unsaturated fatty acids compared to *E. guineensis* (Rajanaidu *et al*., 2000). Direct exploitation of *E. oleifera* would not be economically viable due to the low bunch and oil yields. Introgressions between *E. oleifera* and *E. guineensis* have been made to form interspecific hybrids with intermediate values for oil yield, iodine value and carotene content (Hardon, 1969; Rajanaidu *et al*., 1995; Mohd Din, 2000). The next step after forming the interspecific hybrids is to fix the genes of interest by development of backcrosses. Almost five to six generations of backcrosses to the recurrent parent may be required to fix the genes. Being a perennial crop (at 10-12 years per breeding cycle) it may take quite a long time to achieve the final product. However, molecular DNA technology such as marker-assisted selection (MAS) in the oil palm breeding programmes will help breeders to accelerate the selection process for traits of interest. The successful breeding of oil palm largely depends on the availability of genetic variability for traits of interest. This study utilised SNP from oil palm as well as SSR markers (both EST and genomic) to screen a backcross two (BC2) interspecific hybrid mapping population. The level of polymorphism and mode of inheritance revealed by the markers in the mapping family were examined.

MATERIALS AND METHODS

Palm Materials and DNA Extraction

A total of 75 backcross palms (BC2) derived from the cross 1084/TP51/22.32 x 335/5.2-5/23.96 were selected. The female parent, 1084/TP51/22.32, was a *E. guineensis tenera* derived from a cross involving the Nigerian *tenera* palm, T128 and a Serdang *pisifera* palm. The male parent, 335/5.2-5/23.96, was an interspecific backcross (BC1) generated from a cross involving the Colombian-based interspecific hybrid and the Nigerian *tenera* palm, T128. Spear leaves were harvested from each palm including the parental palms for DNA extraction. DNA extraction was carried out using the modified CTAB method (Dellaporta *et al*., 1983).

SSR Analysis

A total of 515 SSR markers were screened for polymorphism on the BC2 population then followed by genotyping analysis. All these markers were obtained from the MPOB in-house oil palm sequence database (*GeneThresher*™, *PalmGenes* and *PalmDNABase*) and derived from both *E. guineensis* and *E. oleifera* (Singh *et al*., 2008; Ting *et al*., 2010). The forward primer was 5′ end-labelled with γ-^{33}P at 37°C for 1 hr and 30 min using T4 polynucleotide kinase. The labelling reactions contained 15 pmoles of primer, 0.1 μl of γ-^{33}P and 1 U of T4 polynucleotide kinase in a total volume of 1.0 μl. The PCR reaction consisted of 1 U of *Taq* polymerase, 1.5 mM $MgCl_2$, 0.2 μM of each primer, 0.2 mM dNTPs and 10X PCR buffer. An aliquot of DNA was diluted to a concentration of 50 ng μl^{-1} and used in the SSR analysis. The following PCR programme was used: pre-denaturation at 95°C for 1 min, denaturation at 95°C for 30 s, annealing (temperature depends on primer) for 30 s, and extension at 72°C for 30 s. This programme was repeated for 35 cycles, followed by a final extension at 72°C for 5 min.

The PCR product was mixed with an equal volume (10 µl) of formamide dye (0.3% bromophenol blue, 0.3% xylene cyanol; 10 mM diaminoethanetetra-acetic acid (EDTA); pH 8.0; 97.5% deionised formamide) and heated for 3 min at 95°C and then rapidly chilled on ice. Five µl of the heated sample was run on an acrylamide gel at a constant power of 1600V for 3 hr. The gel was prepared by mixing 100 ml of 6% polyacrylamide (20:1 acrylamide:bisacrylamide) containing 7.5 M urea and 1 M tris-borate-EDTA (TBE). The 100 µl of 10% ammonium persulphate and 20 µl of tetramethylethylenediamine (TEMED) were also added to the gel prior to pouring the mixture to help solidify the gel. The gels were vacuum-dried after the run for 1 hr, and exposed against X-ray film for 7-14 days at -80°C.

SNP Analysis

A total of 4451 SNP markers were obtained from the *GeneThresher*™ oil palm genome sequences for genotyping (Leslie *et al*., 2012). DNA samples were diluted to 200 ng μl^{-1} and then genotyped with Infinium HD Assay Ultra by following the manufacturer's protocol (Illumina Inc., USA). The DNA samples were at first denatured and neutralised to prepare them for amplification. The denatured DNA was isothermally amplified in an overnight step. The whole-genome amplification uniformly increases the amount of the DNA sample by several

thousand-fold without introducing large amounts of amplification bias. Then, the amplified product was fragmented by a controlled enzymatic process. After isopropanol precipitation, the fragmented DNA was collected by centrifugation at 4°C. This was followed by resuspending the precipitated DNA in hybridisation buffer. The BeadChip was appropriately prepared for hybridisation. Following hybridisation, the Illumina iScan was used to scan the BeadChip. The data was analysed using the Genome Studio software (Illumina Inc., USA).

Data Analysis

The SSR data was scored and coded according to Billotte *et al*. (2005) and Singh *et al*., (2009) for a cross involving two heterozygous parents. The SNP data as indicated above was scored directly using Genome Studio software (Illumina Inc., USA). Chi-square test was performed to determine if markers followed the expected Mendelian ratios (1:1, 1:2:1 or 1:1:1:1). The chi-square analysis was performed using JoinMap version 4.1 at p = 0.05. Skewed markers were identified for exclusion from subsequent studies.

RESULTS AND DISCUSSIONS

Genotyping Analysis

Currently, molecular markers especially SSR and SNP are widely used to identify genomic loci associated with traits of interest in an appropriate mapping population. These markers are very informative and occur abundantly in the genome. Appropriate statistical methods have been developed to determine the association between markers and trait of interest. It is important that prior to making use of SSR and SNP markers, the inheritance of both markers from parents to progenies should be evaluated. The mapping population selected for this purpose was a BC2 hybrid between *E. oleifera* and *E. guineensis*. This population showed good variation in vegetative traits and fatty acid composition among the progenies, which is important for linking markers to these quantitative traits. In the SSR analysis, a small subset of 12 palms was initially used to identify polymorphic markers. Of the 515 SSR markers screened, 216 (41.9%) were found to be polymorphic in the BC2 population. The percentage of polymorphic markers in this population was higher than that reported for other studies in oil palm using other marker systems including RFLP and SSR (Singh *et al.,* 2008; 2009). The 216 polymorphic markers then were used to genotype the mapping population. As the SNP analysis was done using the Infinium platform, all palms were genotyped simultaneously on the 4451 SNP and the results were encouraging. The data analysed was of high quality as the call rate observed was above 97%. Of the 4451 SNP markers, 1919 (43%) were polymorphic. The level of polymorphism exhibited by both the SSR and SNP markers in the BC2 population was similar. All the polymorphic markers were scored either as dominant or codominant (*Table 1*). For both marker systems, majority of the polymorphic markers were found in the female parent (*E. guineensis tenera*) as shown in *Table 1*.

Of the 4966 markers tested, 2135 markers (43%) were polymorphic in the BC2 population. The analysis showed that 25.9% of the polymorphic markers were co-dominant and 74.1% markers were of dominant nature. Most of the segregating dominant markers were inherited from the female parent (*E. guineensis tenera*). The ratio of informative markers in the female parent to the male parent is 2:1. Nevertheless, considering the relatively high number of co-dominant markers detected, there are sufficient markers for the development of genetic linkage map for both parental palms. The results indicated that the female *E. guineensis* parental palm was more heterozygous than the backcross one (BC1) interspecific hybrid. This was not surprising as the BC1 parental palm was derived from the Colombian *E. oleifera.* In fact, Rajinder and Cheah (1999) reported that the *E. oleifera* palm UP1026 was very homozygous. This was attributed to the

TABLE 1. RESULTS OBTAINED FROM GENOTYPING OF SIMPLE SEQUENCE REPEAT (SSR) AND SINGLE NUCLEOTIDE POLYMORPHISM (SNP) IN A BC2 MAPPING POPULATION

Marker	No. of markers screened	No. of polymorphic markers	Parent genotypes of segregating markers		
			Dominant female parent	Dominant male parent	Codominant in both parent
SSR	515	216	88	67	61
SNP	4 451	1 919	931	497	491
Total	4 966	2 135	1 019	564	552
Percentage	-	43%	47.7%	26.4%	25.9%

fact that oleifera palms in South America are found in scattered areas which encouraged inbreeding resulting in low levels of heterozygosity (Rajanaidu, 1985). However, the heterozygosity observed in the BC1 parental palm was an improvement compared to the very low level reported previously (Rajinder and Cheah, 1999).

The levels of polymorphism revealed by both marker systems was generally high although lower than that reported in a previous study using AFLP, 61.4% (Rajinder and Cheah, 1999). In this study, 40 SSR markers revealed more than two alleles showing their multi-allelic character, which makes SSR attractive as DNA markers. SNP, on the other hand, are biallelic, where an individual reveals two alleles for a particular marker. In this respect, the information content per SNP marker is lower compared to SSR. It has been estimated that it will take approximately five SNP markers to equal the information revealed by one SSR marker, meaning that ~2000 SNP will be required to equal a 10 cM SSR map (Asgar , 2011). In addition, rare alleles exhibited by some SSR markers have been associated with traits of interest (Chattopadhyay *et al.*, 2008). In SNP, rare alleles are not informative in most cases. SNP however, are abundant in all genomes and hold more promise for generating very high density genetic maps. Furthermore, SNP markers are amenable to high throughput analysis and have a lower error rate compared to SSR markers. In fact, SNP themselves may represent the polymorphisms associated with the gene of interest under study, thereby enabling new applications in molecular breeding (Collard *et al.*, 2005; Vroh Bi *et al*., 2006; Jones *et al*., 2002). Data is also easily compared across laboratories and different assay technologies whereas standardisation of SSR data across laboratories can be challenging due to variation in allele sizing.

In this study, all the polymorphic SSR primers were subsequently genotyped on all 75 individuals of the BC2 and the results are presented in *Table 2*. From the 216 polymorphic primers tested, 395 easy to read amplification products ranging from 104 to 330 bp in size were generated. In the inheritance analysis, 61 primers tested showed co-dominant segregation profile while the rest were dominant. *Figures 1a* to *1c* show an example of profiles obtained in the SSR analysis. Genotyping of all the polymorphic SNP markers generated about 2410 segregating bands, which were further analysed using Genome Studio® software to determine the segregation profile. The segregation profiles observed and scored are summarised in *Table 2*.

Data from SSR and SNP were divided into five files (two for dominant and three for co-dominant). The first file contained data for bands segregating in the ratio 1:1 present in the female parent but absent in the male parent (BC1); the second file also segregating in the 1:1 ratio, but present in the male parent and absent in the female parent; the third file contained data involving bands present in both parents and segregating with a 1:2:1 ratio in the progeny. The fourth and fifth files also contained data present in both parents but with 1:1:1:1 segregation profiles.

In general, co-dominant markers are more useful than dominant markers due to higher information content. A co-dominant molecular marker can distinguish homozygous and heterozygous genotypes and help determine allele frequencies. The co-dominant markers will also allow easy comparison and integration of genetic maps generated in different studies. In contrast, dominant markers are scored on the basis of the presence or absence of the band, in which case homozygous and heterozygous individuals cannot be distinguished.

The segregating bands obtained were evaluated using the chi-square tests for goodness-of-fit to the expected segregation ratios based on 5% significance level ($p<0.05$) as shown in *Table 3*. Of the 2135 informative markers from both marker systems, 1871 (87.6%) markers were inherited in the expected Mendelian manner. A total of 264 markers (12.4%) showed distortion from the expected segregation ratios. The small progeny size and viability selection that occur on the distorted markers or loci linked to the markers are a possible explanation. The later phenomenon has been attributed to distorted inheritance of alleles due to competition among gametes or from abortion of gamete or zygotes (Lyttle, 1991). Departures from Mendelian ratios can also result from the presence of null alleles (Recee *et al.*, 2004). Nevertheless, in hybrids of species or subspecies, skewed segregation ratios are common (Whitkus, 1998). It is heartening to note that the level of distortion was lower than that reported previously for oil palm (Rajinder and Cheah, 1999; Singh *et al.*, 2009). In genetic mapping for quantitative traits, Mendelian segregation is the basic assumption of genetic linkage maps. Therefore, severely distorted markers are usually discarded prior to QTL mapping to avoid the unexpected consequences of distorted markers on the results.

Evaluating Fidelity of Controlled Crosses

Large-scale quality control testing also benefits from molecular genotyping by using these markers to develop genetic fingerprint to trace samples and estimate the purity of a population. Contaminants can be detected in cases where individuals show bands not found in either parents. In this study, one such individual was detected as a possible contaminant. The palm (No. 261) revealed an additional band not found in the parent in 173 (8.1%) of the 2135 polymorphic markers examined (*Figures*

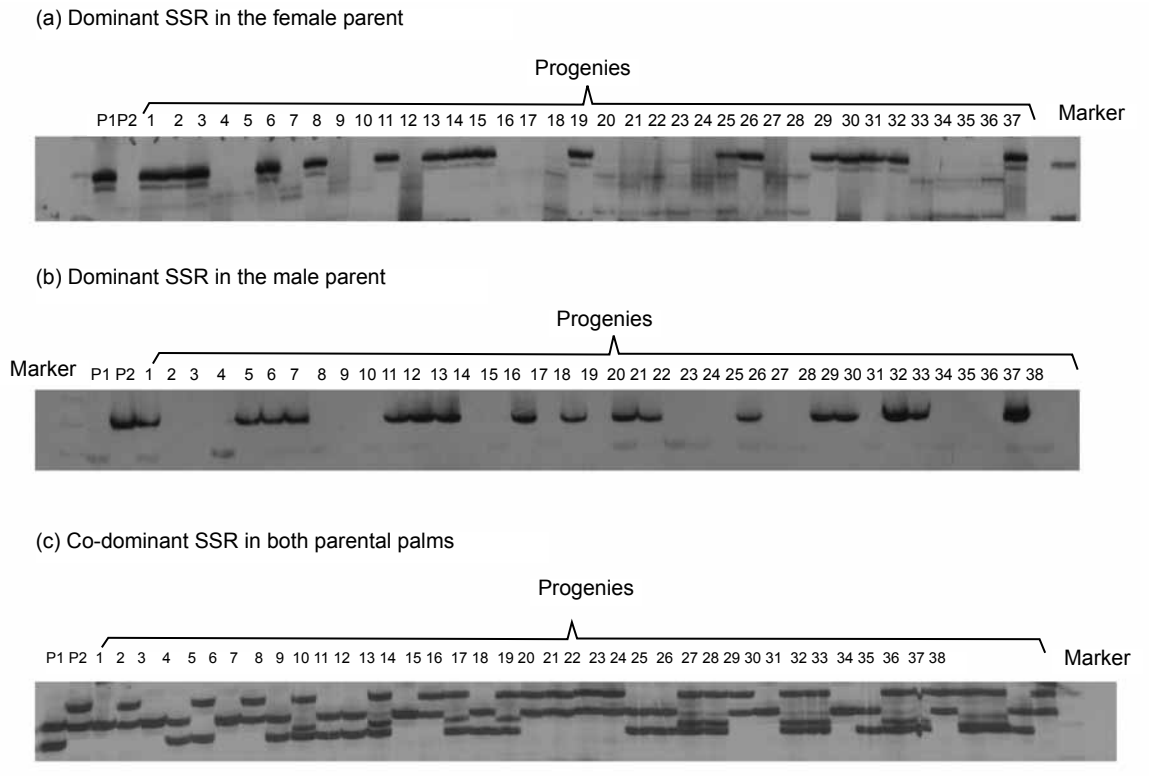

Figure 1 (a - c). Autoradiogram of different inheritance profiles obtained in simple sequence repeat (SSR) analysis. P1 and P2 refer to parents E. guineensis tenera *and BC1, respectively.*

TABLE 2. SEGREGATION OF SIMPLE SEQUENCE REPEAT (SSR) AND SINGLE NUCLEOTIDE POLYMORPHISM (SNP) MARKERS OBSERVED IN THE BC2 MAPPING POPULATION

	No. of alleles	Parental genotypes		Progeny genotypes				Expected segregation ratio	No. of segregating markers	
		Female	Male	1	2	3	4		SSR	SNP
1	Single allele (dominant)	▬		▬		▬		1:1	88	931
2	Single allele (dominant)		▬	▬			▬	1:1	67	497
3	Two alleles (codominant)	▬ ▬	▬ ▬	▬	▬ ▬	▬	▬ ▬	1:2:1	21	491
4	Three alleles (codominant)	▬ ▬	▬ ▬	▬	▬ ▬	▬ ▬	 ▬ ▬	1:1:1:1	34	-
5	Four alleles (codominant)	▬ ▬	 ▬ ▬	▬ ▬	 ▬ ▬	▬ ▬	 ▬ ▬	1:1:1:1	6	-

TABLE 3. CHI-SQUARE ANALYSIS FOR GOODNESS-OF-FIT TO THE EXPECTED MENDELIAN SEGREGATION RATIOS

Marker type	No. of segregating markers	Chi-square analysis on Mendelian segregation ratio (p≥0.05)	
		Fit markers	Distorted markers
SSR	216	179	37
SNP	1 919	1 692	227
Total	2 135	1 871 (87.6%)	264 (12.4%)

Note: SSR - simple sequence repeat. SNP - single nucleotide polymorphism.

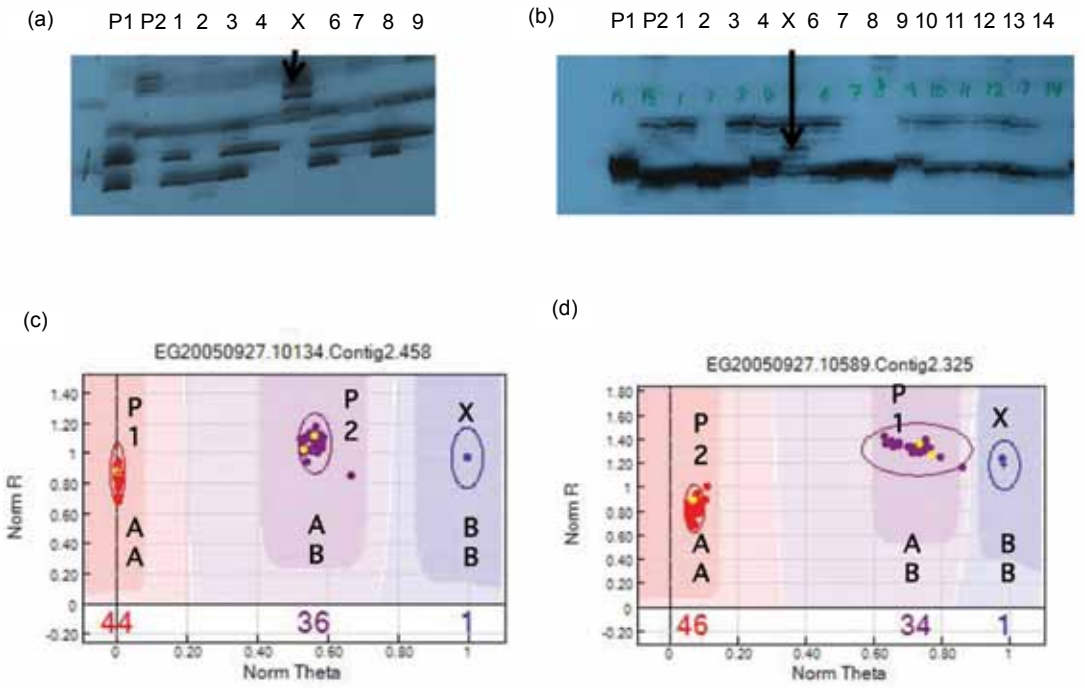

Figure 2 (a-d). Detection of contaminant (marked X) in simple sequence repeat (SSR) (a,b) and single nucleotide polymorphism (SNP) (c,d) genotyping analysis. Examples demonstrate that the contaminant had unexpected genotype compared to the parents. P1 and P2 refer to parents E. guineensis tenera *and BC1, respectively.*

2a to *2d*). It is important in mapping and QTL analysis to remove the contaminant from analysis to avoid distortion of data. The contaminant could have arisen from pollen contamination of controlled crosses, mix-ups or mislabelling at laboratory and nursery stages. Pollen contamination has been reported to occur in oil palm (Chin, 1995; Singh *et al.*, 1999), and other controlled crosses, such as in olive (Aurora and Antonio, 2007) and pine (Elliot *et al.*, 2006).

Mapping Population

Apart from efficient DNA marker systems, an appropriate segregating populations is also important to develop a genetic map for QTL analysis. In this study, two different species (*E. guineensis* and *E. oleifera*) were used to create a hybrid population. Subsequently, a selected OG hybrid palm was backcrossed to the *E. guineensis* parental palm, giving rise to backcross one (BC1). Palm from BC1 was selected to backcross to the *E. guineensis* palm again, giving rise to BC2 progeny used in this study. The repeated backcrossing was carried out to introgress the desirable *E. oleifera* traits into *E. guineensis*. At the same time, it creates a population segregating for the traits of interest, namely vegetative characteristics (*e.g.* dwarfness) and fatty acid composition (FAC) (Hardon, 1969; Rajanaidu *et al.*, 1995). In fact, the phenotypic data showed that the BC2 population actually showed a good variation for both characters which is desirable for QTL analysis (data not shown). In interspecific backcross programme, several backcross cycles are necessary to get a pair of strains that are essentially identical at all loci except for the region surrounding

the gene under selection. Theoretically, by the sixth generation of backcrossing, the entire genotype of the *E. guineensis* parent would be recaptured, together with the desired traits from *E. oleifera*. However, with markers linked to traits and application of marker assisted selection (MAS), the number of cycles could be reduced to achieve the same objective.

CONCLUSION

This is one of the first reports revealing the inheritance of SNP in combination with SSR markers in oil palm. In fact, this is the first attempt to evaluate a large number (4966) of markers in an interspecific BC2 mapping population segregating for vegetative traits and FAC. The polymorphism revealed (43%) is considerable and there are sufficient numbers of informative markers to generate genetic maps for both parental palms. In fact, there are sufficient co-dominant markers identified to integrate both parental palms. The fact that 87.6% of the polymorphic markers, both SNP and SSR, are inherited in expected Mendelian ratio from parents to offsprings also shows their potential for use in developing genetic maps. Furthermore, with the ability to detect contaminants in the current mapping materials, the SSR and SNP markers are also useful in conducting DNA fingerprinting on other controlled crosses to ensure the genetic identity and purity of crosses in the breeding programmes.

ACKNOWLEDGEMENT

We wish to thank the Director-General of MPOB for permission to publish these research findings. The BC2 population used in this study was obtained from United Plantations Berhad, Teluk Intan, Perak, Malaysia.

REFERENCES

ADAM-BLONDON, A F; SÉVIGNAC, M; DRON, M and BANNEROT, H (1994). A genetic map of common bean to localize specific resistance genes against anthracnose. *Genome, 37*: 915-924.

ALWALA, S; SUMAN, A J; ARRO, A; VEREMIS, J C and KIMBENG, C A (2006). Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Science, 46*: 448-455.

ASGAR, S (2011). *DNA Markers for Marker-assisted Breeding*. Eurofins STA Laboratories eNewsletter.

AURORA, D and ANTONIO, M (2007). Cross-compatibility of the parents as the main factor for successful olive breeding crosses. *American Society for Horticultural Science, 6*: 830-835.

BILLOTTE, N; RISTERUCCI, A M; BARCELOS, E; NOYER, J L; AMBLARD, P and BAURENS, F C (2001). Development, characterization and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome, 44*: 413-425.

BILLOTTE, N; MARSEILLAC, N; RISTERUCCI, A M; ADON, B and BROTTIER, P (2005). Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics, 110*: 754-765.

BOTSTEIN, D; WHITE, R L; SKOLNICK, M and DAVIS, R W (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics, 32*: 314-331.

CHABANE, K; ABLETT, G A; CORDEIRO, G M; VALKOUN, J and HENRY, R J (2005). EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. *Genetic Resources & Crop Evolution, 52*: 903-909.

CHATTOPADHYAY, T; CHATTERJEE, M; MANDAL, N and BHATTACHARYYA, S (2008). Biochemical and SSR marker based characterization of some Bengal landraces of rice suffixed with 'sail' in their name. *The Indian Journal of Genetics and Plant Breeding, 68*: 15-20.

CHELKOWSKI, J; GOLKA, L and STEPIEN, L (2003). Application of STS markers for leaf rust resistance genes in near-isogenic lines of spring wheat cv. Thatcher. *Applied Genetics, 44*: 323-338.

CHIN, C W (1995). Oil palm planting materials and quality control. *Proc. of the PORIM National Oil Palm Conference*. PORIM, Bangi. p. 38-47.

CHO, Y G; ISHII, T; TEMNYKH, S; CHEN, X; LIPOVICH, L; MCCOUCH, S R; PARL,W D; AYERS, N and CARTINHOUR, S (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theoretical & Applied Genetics, 100*: 713-722.

COLLARD, B C Y; JAHUFER, M Z Z; BROUWER, J B and PANG, E C K (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica, 142*: 169-196.

CORDEIRO, G M; CASU, R; MCINTYRE, C L; MANNERS, J M and HENRY, R J (2001). Microsatellite markers from sugarcane *(Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Science*, *160*: 115–1123.

DELLAPORTA, S L; WOOD, J and HICKS, J B (1983). A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter, 1:* 19-21.

ELLIOTT, F G; SHEPHERD, M J and HENRY, R J (2006).Verification of interspecific pine hybrids using paternally inherited chloroplast microsatellites. *Forest Genetics*, *12*: 81-87.

ESTERAS, C; GOMEZ, P; MONFORTE, A J; BLANCA, J; VICENTE-DOLERA, N; ROIG, C, NUEZ, F and PICO, B (2012). High-throughput SNP genotyping in *Cucurbita pepo* for map construction and quantitative trait loci mapping. *Genomics*, *13*: 80.

EUJAYL, I; SORRELLS, M; BAUM, M; WOLTERS, P and POWELL, W (2001). Assessment of genotypic variation among cultivated durum wheat based on EST SSRs and genomic SSRs. *Euphytica. 119*: 39-43.

EUJAYL, I; SORRELLS, M; BAUM, M; WOLTERS, P and POWELL, W (2002). Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theoretical & Applied Genetics, 104*: 399-407.

GRODZICKER, T; WILLIAMS, J; SHARP, P and SAMBROOK, J (1975). Physical mapping of temperature sensitive mutants of adenovirus. *Cold Spring Harbor Symposia Quantitative Biology, 39*: 439-446.

GUPTA, P K; VARSHNEY, R K; SHARMA, P C and RAMESH, B (1999). Molecular markers and their applications in wheat breeding. *Plant Breeding, 118*: 369-390.

GUPTA, P K; RUSTGI, S; SHARMA, S; SING, R; KUMAR, N and BALYAN, H S (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genomics & Genetics, 270*: 315-323.

HACKAUF, B and WEHLING, P (2002). Identification of microsatellite polymorphisms in an expressed portion of the rye genome. *Plant Breeding, 121*: 17-25.

HAMBLIN, M T; WARBURTON, M L and BUCKLER, E S (2007). Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS One, 2(12)*: 1367-1371.

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.

HELENTJARIS, T; SLOCUM, M; WRIGHT, S; SCHAEFER, A and NIENHUIS, J (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical & Applied Genetics, 61*: 650-658.

KANTETY, R V; ROTA, M L; MATTHEWS, D E and SORRELLS, M E (2002). Data mining for simple-sequence repeats in expressed sequence tags from barley, maize, rice, sorghum, and wheat. *Plant Molecular Biology, 48*: 501-510.

JONES, E S; BREESE, W A; LIU, C J; SINGH, S D; SHAW, D S and WITCOMBE, J R (2002). Mapping quantitative trait loci for resistance to downy mildew in pearl millet: field and glasshouse screens detect the same QTL. *Crop Science, 42*: 1316-1323.

KISS, G; CSANADI, G; KALMAN, K; KALO, P and OKRESZ, L (1993). Construction of a basis genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Molecular Genomics & Genetics, 238*: 129-137.

KOTA, R; VARSHNEY, R K; THIEL, T; DEHMER, K J and GRANER, A (2001). Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas (Lund), 135*: 145-151.

LAIMA, A; FELICIDAD, F; JOHANNES, J; ELISA, B; KATHERINE, E; ROBERTO, V; RICCARDO, V; JIM, D; MICHELA, T and DANIEL, S (2012). Development of a dense SNP-based linkage map of an apple rootstock progeny using the *Malus Infinium* whole genome genotyping array. *Genomics*, *13*: 203.

LEE, S H; BAILEYM, M A; MIAN, A R; CARTERJRD, T E and SHELY, A (1996). Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Science, 36*: 728-734.

LESLIE, C O; RAHIMAH, A R; LESLIE, L E T and RAJINDER, S (2012). A 1536-plex oil palm SNP set for genetic mapping and identification of markers associated with oil palm fruit colour. *Proc. of the International Plant & Animal Genome Conference*. San Diego, California. P0201.

LIU, X M; SMITH, C M; FRIEBE, B R and GILL, B S (2005). Molecular mapping and allelic relationships of Russian wheat aphid resistance genes. *Crop Science, 45*: 2273-2280.

LYTTLE, T W (1991). Segregation distorters. *Annual Review of Genetics, 25*: 511-557.

MICHALEK, V; WESCHKE, W; PLEISNER, K P and GRANER, A (2002). EST analysis in barley defines a unigene set comprising 4000 genes. *Theoretical & Applied Genetics, 104*: 97-103.

MOHD DIN, A (2000). *Genetic Variation for Yield, Bunch Component and Vegetative Traits in Oil Palm* (Elaeis oleifera) *and Interspecific Hybrids*. Ph.D thesis. Universiti Kebangsaan Malaysia, Bangi.

NICOT, N; CHIQUET, V; GANDON, B; AMILHAT, L; LEGEAI, F; LEROY, P; BERNARD, M and SOURDILLE, P (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theoretical & Applied Genetics, 109:* 800-805.

RAFALSKI, A (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology, 5*: 94-100.

RAJANAIDU, N (1985). *Elaeis oleifera* collection in Central and South America. *Proc. of the International Workshop on Oil Palm Germplasm and Utilization*, Selangor, Malaysia. p. 84-94.

RAJANAIDU, N; CHIN, C W and JALANI, B S (1995). Performance of *Elaeis oleifera* (Surinam) x *E. guineensis* hybrids. ISOPB International Seminar. Barranquilla, Colombia.

RAJANAIDU, N; KUSHAIRI, A; RAFII, M; MOHD DIN, A; MAIZURA, I and JALANI, B S (2000). Oil palm breeding and genetic resources. *Advances in Oil Palm Research*. Vol. 1. MPOB, Bangi. p. 171-237.

RAJEEV, K V; ANDREAS, G and MARK, E S (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology, Vol. 23 No. 1.*

RAJINDER, S and CHEAH, S C (1999). Analysis of the inheritance of AFLP markers in an interspecific cross of oil palm using the pseudo-testcross strategy. *J. Oil Palm Res. (Special Issue)*: 64-73.

REECE, K S; RIBEIRO, W L; GAFFNEY, P M; CARNEGIE, R B and ALLEN, K (2004). Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): confirmation of null alleles and non-Mendelian segregation ratio. *Heredity, 95*: 346-352.

SCOTT, K D; EGGLER, P; SEATON, G; ROSSETTO, M; ABLETT, E M; LEE, L S and HENRY, R J (2000). Analysis of SSRs derived from grape ESTs. *Theoretical & Applied Genetics, 100*: 723-726.

SHAW, P C; WANG, J and BUT, P P H (2002). *Authentication of Chinese Medicinal Materials by DNA Technology*. World Scientific, Singapore. p. 1-23.

SHERRY, S T; WARD, M H; KHOLODOV, M; BAKER, J; PHAN, L; SMIGIELSKI, E M and SIROTKIN, K (2001). dbSNP: The NCBI database of genetic variation. *Nucleic Acids Research, 29*: 308-311.

SHIRASAWA, K; ISOBE, S; HIRAKAWA, H; ASAMIZU, E; FUKUOKA, H; JUST, D; ROTHAN, C; SASAMOTO, S; FUJISHIRO, T; KISHIDA, Y; KOHARA, M; TSURUOKA, H; WADA, T; NAKAMURA, Y; SATO, S and TABATA, S (2010). SNP discovery and linkage map construction in cultivated tomato. *DNA Research, 17*: 381-391.

SINGH, R; NOORHARIZA, M Z; TING, N C; ROZANA, R; TAN, S G and LESLIE, L E T (2008). Exploiting an oil palm EST database for the development of gene-derived and their exploitation for assessment of genetic diversity. *Biologia, 63*: 227-235.

SINGH, R; TAN, S G; PANANDAM, J M; RAHIMAH, A R and OOI, L C L (2009). Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. *BMC Plant Biology, 9*: 114.

SMITH, J S C; CHIN, E C L; SHU, H, SMITH, O S and WALL, S 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 & Applied Genetics*, *95*: 163-173.

SRIVASTAVA, S and MISHRA, N (2009). Genetic markers – a cutting-edge technology in herbal drug research. *Chemical & Pharmaceutical Research, 1*: 1-18.

THIEL, T; MICHALEK, R K; VARSHNEY, R K and GRANER, A (2003). Exploiting EST databases for the development and characterization of gene-derived SSR markers in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics, 106*: 411-422.

TING, N C; NOORHARIZA, M Z; ROZANA, R; LOW, E T; MAIZURA, I; CHEAH, S C; TAN, S G and SINGH, R (2010). SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. *J. Genetics, 89*: 135-145.

VALERIE, P; MYRIAM, R; CHRISTINE, T; ANNE, C; SERGE, H; ALEXENDRE, K and PERLA, H (2006). SSR mining in the coffee tree EST database: potential use of ESTSSRs as markers for the *Coffea* genus. *Molecular Genetics Genomics*, *276*: 436-449.

VARSHNEY, R K; THIEL, T; STEIN, N; LANGRIDGE, P and GRANER, A (2002). In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular & Molecular Biology Letters, 7*: 537-546.

VROH- BI, I; McMULLEN, M D; VILLEDA, H S; SCHROEDER, S; GARDINER, J; POLACCO, M; SODERLUND, C; WING, R; FANG, Z and COE, E H (2006). Single nucleotide polymorphisms and insertion deletions for genetic markers and anchoring the maize fingerprint contig physical map. *Crop Science, 46*: 12-21.

WHITKUS, R (1998). Genetics of adaptive radiation in Hawaiian and Cook Island species of *Tetramolopium* (Asteraceae). II. Genetic linkage map and its implications for interspecific breeding barriers. *Genetics, 150*: 1209-1216.

ZULKIFLI, Y; MAIZURA, I and RAJINDER, S (2012). Evaluation of MPOB oil palm germplasm (*Elaeis guineensis*) populations using EST- SSR. *J. Oil Palm Res. Vol. 24*: 1368-1377.