

Application of microsatellite markers for hybrid verification and genetic analysis of oil palm (*Elaeis guineensis* Jacq.)

Okoye M. N.^{1,2,3*}, Singh, R.,² Uguru, M. I.,³ and Bakoumé, C.⁴

¹Plant Breeding Division, Nigerian Institute for Oil Palm Research (NIFOR), P.M.B 1030 300001, Benin City, Nigeria

²Advanced Biotechnology and Breeding Centre, Malaysian Palm Oil Board (MPOB), P.O. Box 10620, 50720 Kuala Lumpur, Malaysia

³Department of Crop Science, Faculty of Agriculture, University of Nigeria, Nsukka, Enugu State, Nigeria

⁴Maxi Productivity Sarl, P. O. Box 2137 (or 24240), Douala, Cameroon

Abstract

The legitimacy of parents and progenies used in crop improvement programmes is vital for any meaningful progress in selection. While acknowledging the shortcomings of controlled pollination in oil palm breeding and commercial seed production, the legitimacy of 20 oil palm progenies from the Nigerian Institute for Oil Palm Research (NIFOR) breeding programme was determined using 16 fluorescently-labeled microsatellite markers. The genotyping of parents and progenies was conducted by capillary electrophoresis using the ABI 3730 DNA Genetic Analyzer (Applied Biosystems, USA). Results revealed a complementary expression of the parents' alleles in 18 out of the 20 individual progenies screened, confirming their hybridity and genetic identity. The two illegitimate progenies detected could be attributed to pollination and planting errors, respectively. A subset of three sufficiently informative loci (sMg00016, sMg00179 and sMo00102) was identified for routine quality control genotyping. The detection of illegitimate progenies provided ample evidence to substantiate the importance of assessing hybrid fidelity in breeding programmes. Furthermore, the usefulness of microsatellite markers as a reliable technique for routine assessment and unambiguous identification of oil palm crosses was established. The implications of microsatellite-based hybrid identification in oil palm varietal improvement programmes have been adequately discussed.

Key words: Contamination, Genetic identity, Illegitimacy, NIFOR, Parent-Progeny relationships, seed production

Corresponding Authors Email: maxwellokoye@gmail.com

Introduction

Over the past eight decades at the Nigerian Institute for Oil Palm Research (NIFOR), there has been a steady progress in oil palm breeding with the highest average oil yield of about 4-5

t/ha/yr (Ataga et al., 2018). In accordance with the economic diversification efforts of the federal government to boost national palm oil production, a policy to invest \$500 million in oil palm plantation development was enunciated. This policy aims at increasing the annual

local production of palm oil by 700% over the next eight years (2019 to 2027) to attain self-sufficiency in the commodity (USDA 2019). To meet this projected target, future breeding progress will depend on the legitimacy of individual parents as well as their progeny for breeding crosses and commercial hybrids. While acknowledging the susceptibility of oil palm controlled pollination to various sources of errors (during pollination, seed collection, germination, and field planting), the need for accurate hybrid verification and identification (genetic fingerprints) of oil palm crosses cannot be overlooked (Corley and Tinker, 2016). Therefore, the success of hybrid oil palm production beside other factors depends on the production and timely supply of genetically pure planting materials to the oil palm growers. This ensures that the gains of heterosis can be harnessed through enhanced yield by growing pure F₁ hybrid crops.

A major concern to the sustainability of Nigerian oil palm industry is the fidelity of planting materials. The out-crossing behaviour of the crop coupled with the challenges associated with controlled pollination, inadvertently deviate the expected or theoretical Mendelian segregation ratios, leading to contamination or illegitimacy in fruit forms. Hence the need to objectively confirm the genetic identity of hybrids in a crossing programme for breeding and seed production. This situation is further exacerbated by adulteration of oil palm planting materials by illegal seed or seedling hawkers/producers, who handpick seeds from plantations and raise them as seedlings for sale to farmers (Okwuagwu, 2001). These traffickers market their illegitimate materials to unsuspecting farmers on the pretense of being NIFOR agents to support their fraudulent actions. More worrisome is the inability of the farmers to differentiate between NIFOR *tenera* planting material and the material procured from the seed/seedling vendors. This problem is equally shared by oil palm breeders who find it difficult in identifying the hybrids of crossed progenies before planting and production of fruit bunches. As a step to safeguard the very vulnerable farmers and the oil palm industry from illegal seed traffickers, several measures including microsatellite marker, genetic fingerprinting scheme were initiated to characterize breeding crosses and commercial

planting materials for early identification at or before planting stage (Okoye 2016).

The shell thickness gene (*Sh*) in oil palm fruit forms (*dura*, *pisifera* and *tenera*) plays a major role in identification of fruit type and also influences palm oil yield (Singh et al., 2013). Illegitimacy and contamination (some illegitimate palms in a family) in oil palm is conventionally assayed by using the traditional method of shell thickness assessment and segregation pattern of the fruit forms which are often ambiguous and vulnerable to long term field evaluation. Essentially, oil palm must be grown $\pm 3-4$ years before production of fruit bunches for fruit form determination and subsequent verification of hybrid legitimacy. A reliable method for hybrid identification of oil palm at early seedling stage is therefore crucial both for the integrity of a durable breeding programme and for assuring good quality planting material to the farmers.

Unlike the morphological means (fruit-form determination) of identifying contamination in oil palm, molecular markers are good alternatives because they are not subject to environmental influences and can be readily detected in all plant tissues, notwithstanding the growth or developmental stage (Mondini et al., 2009). The use of microsatellites or simple sequence repeats (SSR) marker technique is well established and accepted for genotype identity and hybrid verification because of their abundance in the genome, co-dominant inheritance, high polymorphism and reproducibility (Amos et al., 1996; Smith et al., 1997). Several studies have successfully employed microsatellite markers as a routine quality control approach to address the issue of illegitimacy and contamination in oil palm breeding programmes and commercial seed production. Especially promising is the current application of microsatellite markers for identity checking in selection and seed production processes by Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) and PalmElit (Pomiès et al., 2019), and Sime Darby Plantation R&D Centre, Malaysia (Teh et al., 2019).

The legitimacy of oil palm materials currently used in the NIFOR oil palm main breeding programme has not been assessed and documented using any molecular marker technique. As part of a quality control approach in the breeding and seed production programme, we present a preliminary fingerprinting report of

some oil palm breeding crosses with molecular markers. Specifically, this study 1) sought out to employ fluorescently-labeled microsatellite markers for confirmation of parentage and hybrid purity of 20 oil palm progenies, and 2) to identify a subset of highly informative SSR markers for routine and low-cost quality control genotyping.

Materials and Methods

Twenty five (25) oil palm genotypes comprising two parental genotypes with contrasting yield components and twenty of their resulting progenies, and three advanced experimental selections from Malaysian Palm Oil Board (MPOB) were evaluated in this study (Table 1). The parental genotypes consist of the thick-shelled *dura* (female) and thin-shelled *tenera* (male) genotypes. The *dura* parent genotype was selected for high bunch weight and low bunch number while the *tenera* genotype was selected for high bunch number and low bunch weight. These two sequentially developed traits are highly heritable and determine the yield of fresh fruit bunch in the oil palm (Okoye et al., 2001). The three advanced experimental selections were used as outgroup in the multivariate analysis to

assess the efficiency of the employed molecular markers in genetic differentiation and discrimination among different oil palm genotypes.

Young leaf samples were collected from unopened spears of 22 individual palms planted at the NIFOR Main Station, Benin City, Nigeria. The samples were stored at -80°C at the Bioscience Centre, International Institute of Tropical Agriculture (IITA) Ibadan, Nigeria prior to DNA isolation. In addition, DNA samples of the three advanced experimental selections were obtained from MPOB Malaysia.

DNA was extracted from the individual palms following the Cetyl-Trimethyl Ammonium Bromide (CTAB) procedure (Doyle and Doyle 1990) with an additional chloroform extraction step. The extracted DNA quality of each sample was assessed using 1% (w/v) agarose gel and the quantity of DNA was confirmed using NANODROP® (ND-1000) Spectrophotometer (Thermo Fisher Scientific Inc., Denver). DNA concentrations were normalized at 25 ng/μl in sterile water and stored at 4°C until polymerase chain reaction (PCR) amplification at the Advanced Biotechnology and Breeding Centre (ABBC), MPOB Selangor, Malaysia.

Table 1: List of parental genotypes along with their putative hybrids and three elite oil palm populations used as an outgroup in this study

S/N	Genotype	Origin	Code
1	P1♀	Serdang Avenue Deli	DD1
2	P2♂	Umuabi OP	T6
3	DT_47/593	Putative hybrid P1 x P2	DT1
4	DT_47/649	Putative hybrid P1 x P2	DT2
5	DT_47/793	Putative hybrid P1 x P2	DT3
6	DT_47/820	Putative hybrid P1 x P2	DT4
7	DT_47/916	Putative hybrid P1 x P2	DT5
8	DT_47/945	Putative hybrid P1 x P2	DT6
9	DT_47/1040	Putative hybrid P1 x P2	DT7
10	DT_47/1041	Putative hybrid P1 x P2	DT8
11	DT_47/1096	Putative hybrid P1 x P2	DT9
12	DT_47/1150	Putative hybrid P1 x P2	DT10
13	DT_47/1151	Putative hybrid P1 x P2	DT11
14	DT_47/1207	Putative hybrid P1 x P2	DT12
15	DT_47/1468	Putative hybrid P1 x P2	DT13
16	DT_47/1679	Putative hybrid P1 x P2	DT14

17	DT_47/1803	Putative hybrid P1 x P2	DT15
18	DT_47/1846	Putative hybrid P1 x P2	DT16
19	DT_47/1847	Putative hybrid P1 x P2	DT17
20	DT_47/1874	Putative hybrid P1 x P2	DT18
21	DT_47/2042	Putative hybrid P1 x P2	DT19
22	DT_47/2043	Putative hybrid P1 x P2	DT20
23	013_0338/0391*	MPOB Kluang <i>Dura</i>	KI-D1
24	014_0174/0655*	MPOB Kluang <i>Pisifera</i>	KI-P1
25	016_911101/1203*	MPOB KULIM <i>Pisifera</i>	Ku-P2

♀- female parent, ♂- male parent, OP- open pollinated, *elite breeding populations forming the outgroup

A total of sixteen fluorescently-labeled microsatellite markers were used to fingerprint progenies and their parents. Nine of these markers (sMg00156, sEg00154, sMo00102, sMg00228, sMg00016, sMg00120, sEg00151, sMg00179 and sMg00087) were developed at MPOB by Singh et al., (2008). A further set of seven markers (mEgCIR3813, mEgCIR0793, mEgCIR0425, mEgCIR3828, mEgCIR3519, mEgCIR0790 and mEgCIR3745) were obtained from CIRAD (Billotte et al., 2005; <http://tropgenedb.cirad.fr/>). Preliminary screening of these markers on the parental genotypes for polymorphism allowed the selection of eight informative loci used to confirm the identity of the 20 oil palm hybrids (Figure 1).

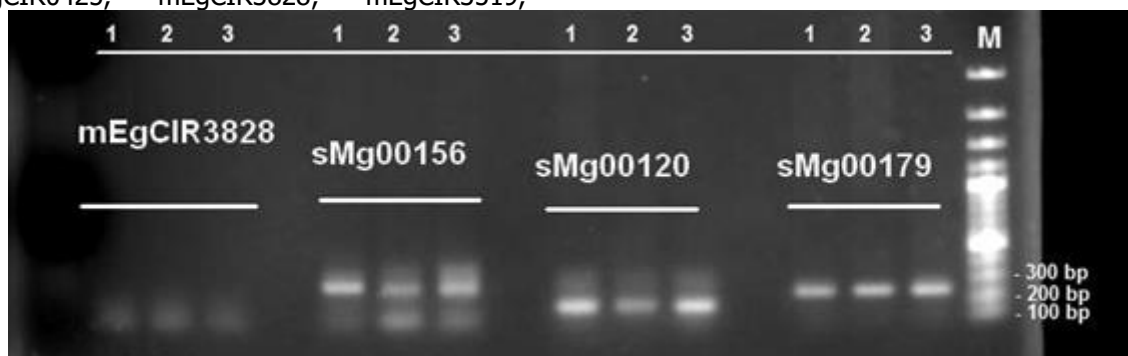


Figure 1: Screening of parents and a hybrid with four SSR markers prior to capillary electrophoresis analyses.

Polymerase chain reaction was conducted in a Perkin Elmer 9700 thermocycler (Life Technologies, Thermo-Fisher Scientific, USA) using the composition and conditions described in Ting et al., (2010). The total reaction volume was 10 μ l containing (prepared in order listed): 2 μ l of 25 ng genomic DNA, 6.625 μ l MilliQ water, 1 \times PCR standard buffer (NEB, USA), 0.2 μ l of 10mM deoxynucleotide triphosphates (dNTPs) (NEB, USA), 0.025 μ l of each primer (M13 tailed forward primers and untailed reverse primers), 0.025 μ l dye, and 0.1 μ l of Taq DNA polymerase (5 U/ μ l) (NEB, USA). The amplification cycle consisted of an initial 3 min denaturation at 95 $^{\circ}$ C, followed by 35 cycles of denaturation at 95 $^{\circ}$ C for 30 sec, primer annealing for 30 sec at 50 - 58 $^{\circ}$ C depending on the primer annealing temperature and an

extension temperature of 72 $^{\circ}$ C for 30 sec, followed by an additional extension at 72 $^{\circ}$ C for 2 min. Amplified fragments were estimated by capillary electrophoresis on a DNA Genetic Analyzer-ABI PRISM 3730 (Applied Biosystems, USA). The alleles were sized with reference to GS 500 LIZ, a formamide containing red DNA size standard. Fragment size in base pairs was determined using GeneMapper[®] software version 4.1 (Applied Biosystems, USA). Sample plots were generated and genotype data for all the markers scored in an Excel matrix. The scoring of the genotype data was performed manually with reference to allele and peak size, except for stutters. Null alleles were assigned to sample individual genotypes that were confirmed to have no amplification products under standard conditions.

The peaks present in the parents and their respective progenies were scored for each of the SSR markers used. These were termed alleles and according to the GeneMapper® software were referred to by their size in base pairs rounded to the appropriate unit. Comparison of the parents' alleles to those of their progenies allowed for the determination of legitimate and illegitimate hybrids. A progeny was considered legitimate if one of its two alleles was maternal and the other one paternal showing that they were inherited from its two parents. On the contrary, a progeny was considered illegitimate or contaminant if at least one of its alleles was not inherited from its parents at a minimum of two microsatellite markers. If only maternal alleles were present, the possibility of the progeny being a product of self-pollination and not that of a cross between putative parents was not discounted. Correspondingly, the presence of an unexpected allele with a maternal allele infers a non-hybrid progeny from contamination by foreign pollen. The number of alleles per locus (N_a), unbiased gene diversity (H_e ; Nei 1978) and the observed heterozygosity (H_o) along with their standard error of means were calculated with Genetic Analysis in Excel computer package 6.5 (Peakall and Smouse, 2012). The software Cervus 3.0.7 was used to assess the probability of identity (PID); the probability that two individuals drawn at random from a population will have the same genotype at multiple loci (Marshall et al. 1998). PID was calculated for the entire data set and on a per locus basis as described in Paetkau et al., (1995) to measure the power of each marker set for individual identification. The polymorphic information content (PIC), according to Anderson

et al., (1993), was calculated using PowerMarker v3.25 software (Liu and Muse 2005). The percentage of hybrid genetic purity was calculated using scored data according to the purity index of Bohra et al., (2011): Hybrid purity (%) was determined by dividing the number of true hybrids (comprising alleles from both parents) by the total number of hybrids screened, then multiplied by 100. Pairwise comparisons of the proportion of shared alleles between individual genotypes (plants) were determined by simple matching dissimilarity index. The resulting genetic dissimilarity coefficient was then transformed into a distance matrix averaging over 1000 bootstraps. Cluster analysis was generated from the distance matrix by the unweighted pair group method using the arithmetic averages (UPGMA) algorithm for a better visualization of the genetic relationships among the parent-offspring and MPOB oil palm genotypes used as outgroup. These calculations were performed by the computer program DARwin v5 (Perier and Jacquemoud-Collet 2006).

Results and Discussion

Screening the parents and progenies using SSR markers

The choice of 16 microsatellite markers used for screening the oil palm parent genotypes and their putative hybrids was based on their ability to generate polymorphic bands under optimized PCR conditions. Based on our genotyping results, eight microsatellite markers which were polymorphic among the parents were employed for further analysis of their hybrids (Table 2).

Table 2: Genetic diversity measures for eight polymorphic microsatellite loci tested in 22 oil palm genotypes

Locus	Linkage Group	Repeat Motif	N_a	uH_e	H_o	PID	PIC
mEgCIR0793	2	(GA)15	3.000	0.623	0.955	0.230	0.5318
sEg00154	6	(CAG)5	5.000	0.672	0.955	0.177	0.5979
sMo00102	7	(AG)11	4.000	0.722	1.000	0.141	0.6494
sMg00016	9	(GA)13	6.000	0.717	1.000	0.141	0.6496
mEgCIR3519	10	(GA)15(GT)8	5.000	0.672	0.955	0.177	0.5979
mEgCIR0790	12	(GA)19	4.000	0.644	0.955	0.204	0.5628
sEg00151	13	(CAG)8	2.000	0.512	0.909	0.375	0.3750
sMg00179	14	(AAAG)6	5.000	0.772	1.000	0.104	0.7103

Mean			4.250 ±0.453	0.667 ±0.028	0.966 ±0.011	1.14 x 10 ⁻⁷ #	0.5843
-------------	--	--	-----------------	-----------------	-----------------	---------------------------	--------

Na: observed number of alleles; uHe: expected heterozygosity; Ho: observed heterozygosity; PID: probability of identity; PIC: polymorphism information content, #: combined probability of identity and exclusion

The eight polymorphic SSR loci exhibited a total of 34 alleles among the parents and putative hybrids with a mean value of 4.25 alleles per locus (Table 2). The highest number of alleles (6 alleles per locus) was detected for sMg00016, whereas loci sEg00154, mEgCIR3519 and sMg00179 presented 5 alleles per locus. Similar studies involving the analysis and characterization of microsatellites in oil palm have also detected loci with highly variable number of alleles, as well as the polymorphic information content (2 to 6 alleles per locus, Bakoumé et al., 2011; 3 to 8 alleles per locus, Hama-Ali et al., 2015; 5 to 11 alleles per locus, Budiman et al., 2019). The highest polymorphism information content (PIC) value was detected for the locus sMg00179 (0.7103), while the lowest PIC value of 0.3750 was recorded for the locus sEg00151 with a mean value of 0.5843. This was comparable to PIC values found in previous studies (0.488 to 0.777, Hama-Ali et al., 2014; 0.41 to 0.84, Budiman et al., 2019). The high level of polymorphism observed for the described microsatellite markers support their application in genetic studies of oil palm. Number of alleles detected and PIC value based on the frequencies of different alleles by a particular marker indicates the quality (discriminatory power) of the marker (Powell et al., 1996). The range of gene diversity varied between 0.512 (sEg00151) and 0.772 (sMg00179) with a mean value of 0.667. The observed heterozygosity differed among the locus from 0.909 (sEg00151) to 1.000 (sMo00102, sMg00016 and sMg00179). Lower but comparable results were obtained in a study that involved the analysis of 16 SSR loci in six oil

palm populations from Cameroon where He ranged from 0.47 to 0.62, and Ho from 0.627 to 0.840 (Budiman et al., 2019). This discrepancy may be due to populations evaluated and the fact that the sample size employed in the present study was smaller compared to the aforementioned report. Also, the probabilities of identity values (PID) between two genotypes randomly selected were minimum (0.104) for sMg00179 and maximum (0.375) for sEg00151. When evaluated for all eight microsatellites, the cumulative probability to obtain identical genotypes among the two parents and the 20 putative-hybrids were 1.14 x 10⁻⁷. This result reflects a relatively high genetic polymorphism despite the two generations of selective breeding on the investigated genotypes.

Applying SSR markers for hybrid identification

The eight microsatellite markers used to screen the parent genotypes were employed to verify that the 20 progenies were indeed genetically descended from their putative parents. The alleles recorded for each hybrid with the microsatellite loci tested are shown in Table 3. The number of different allele combinations observed with each primer pair ranged from 2 (sEg00151) to 6 (sMg00016) and the number of hybrids distinguished ranged from 1 (sEg00151) to 5 (sMg00016). These allelic data could be used by other oil palm breeders to check the identity of oil palm with similar pedigree as microsatellites are transferable between laboratories (Peakall et al., 1998). In fact, the parents evaluated in this study are widely used for commercial seed production.

Table 3: Allele combinations (in base pairs) at eight microsatellite loci in the oil palm parents and hybrids

	mEgCIR0793	sEg00154	sMo00102	sMg00016	mEgCIR3519	mEgCIR0790	sEg00151	sMg00179
P1♀	170/170	260/266	254/258	283/289	241/241	231/231	239/239	237/242
P2♂	163/171	261/261	252/260	284/297	260/267	233/243	236/236	232/238
DT1	163/170	260/261	258/260	284/289	241/260	231/243	236/239	232/242
DT2	163/170	260/261	254/260	284/289	241/260	231/233	236/239	232/242
DT3	163/170	261/266	254/260	284/287	241/267	231/243	236/239	232/242

DT4^z	<u>170/175</u>	260/261	258/260	284/289	241/267	<u>231/235</u>	236/239	<u>237/239</u>
DT5	163/170	261/261	252/254	284/289	241/267	231/243	236/238	237/238
DT6	163/170	261/266	252/254	284/289	241/267	231/243	236/239	237/238
DT7	170/171	261/266	254/260	284/289	241/260	231/233	236/239	232/242
DT8	163/170	261/266	252/254	284/289	241/267	231/233	236/239	237/238
DT9	163/170	260/261	254/260	289/297	241/267	231/233	236/239	232/242
DT10	163/170	260/261	254/260	284/289	241/267	231/233	236/239	237/238
DT11	170/171	261/266	258/260	289/297	241/260	231/233	236/239	237/238
DT12^z	170/171	<u>262/265</u>	252/254	283/284	<u>255/258</u>	231/243	236/239	237/238
DT13	170/171	260/261	254/260	289/297	241/260	231/233	236/239	232/242
DT14	163/170	261/266	252/254	284/289	241/260	231/243	236/239	232/242
DT15	163/170	260/261	254/260	284/289	241/267	231/233	236/239	237/238
DT16	170/171	260/261	252/254	289/297	241/260	231/243	236/239	232/242
DT17	163/170	260/261	254/260	289/297	241/260	231/233	236/239	237/238
DT18	163/170	260/261	254/260	284/289	241/260	231/233	236/239	237/238
DT19	163/170	261/266	258/260	284/289	241/267	231/233	236/239	237/238
DT20	170/171	261/266	252/254	283/284	241/267	231/243	236/239	237/238
Na	3	5	4	6	5	4	2	5
K	2	3	3	5	3	3	1	3

Na: number of different allele combinations; K: number of hybrids distinguished by each primer pair; ^z illegitimate progeny with mismatching alleles underlined

One allele from each of the two intercrossing parents was amplified at each SSR locus of the hybrid. A hybrid was also considered illegitimate if one or both of its two alleles were not inherited from its two parents. The hybrid purity index for the entire tested locus was 90%. Out of 20 individual palms screened, two illegitimate hybrids were observed (Table 3). A closer examination of the illegitimate/contaminated progenies revealed that one of the progenies (DT4) presented single allele genotype mismatch at three loci, with all matching alleles derived from the female parent. This suggests that pollen from different male parents may have been used to pollinate the maternal tree rather than the intended pollen, or the pollen used was contaminated with pollen from a different source. Overall however, a pollination error or a planting error could possibly explain this event. The other illegitimate sample (DT12) did not have any allele in common with either parent (Table 3). An error during the seed and seedling handling stages (incorrect labeling or mixing of families in nurseries) is the most likely explanation here. All the legitimate progenies inherited both parents' alleles and were heterozygous at all the SSR loci tested.

Notwithstanding the long history of crossing programmes in oil palm breeding, controlled pollination is difficult and susceptible to various sources of error resulting in illegitimacy or contamination in the controlled crosses. The reasons for illegitimacy or contamination in oil palm breeding programmes are several and have been elaborately discussed by Corley (2005) and Corley and Tinker (2016). More recent utilization of SSR for characterizing genetic diversity of oil palm breeding populations by Budiman et al., (2019), revealed four illegitimate individuals among two *dura* self progenies of PT Astra Agro Lestari (AAL) breeding materials in Indonesia. In a study that used 30 microsatellite markers for illegitimacy and sibship assignments in oil palm, Hama-Ali et al., (2015) found three illegitimate palms among 200 progenies of four half sib families of FELDA's breeding materials in Malaysia. Thongthawee et al., (2012) used eight SSRs for parentage analysis in six full sib families of the breeding plantation of Univanich Palm Oil Public Company Ltd., Krabi, Thailand and detected four non-hybrids. They speculated that illegitimate palms probably resulted from pollen contamination during the control cross and errors during the nursery or planting stage. Taken

together, these observations underscore obvious risk of errors from pollen collection to field planting of controlled crosses, and the need for careful supervision of the process coupled with strict adherence to quality control procedures. The two non-essential genotypes found in this study can be considered excellent when compared to previous reports (Budiman et al., 2019; Hama-Ali et al., 2015; Thongthawee et al., 2012) as well as other studies on different plant species (Subashini et al., 2014).

Selection of informative markers for quality control

The selection of informative microsatellite markers for hybrid confirmation is vital in the reciprocal recurrent selection (RRS) method of oil palm breeding. This is cogent, in view of the possibility of contamination at the different stages of oil palm breeding activities. Another concern, mostly from a practical standpoint, is the initial high cost, research equipment, and tedious procedures involved in microsatellite marker development (Bakoumé et al., 2011). Corley (2005) in his review, suggested that the number of markers necessary for hybrid confirmation might be fewer than five due to the high polymorphism of microsatellites at the locus level. Out of the eight detected polymorphic SSR

markers, a suite of three markers (sMg00016, sMg00179 and sMo00102) presented sufficient information content among the 22 genotypes representing two parental palms and their 20 hybrids. This marker suite had the highest expected heterozygosity (uHe), PIC values and the lowest probabilities of identity (PID; Table 2). They are located in three different oil palm linkage groups (or chromosomes) which explains their significant discriminatory power to distinguish closely related oil palm genotypes.

The efficiency of the cited markers in genetic differentiation and discrimination among different oil palm genotypes was further validated using multivariate analysis. The results of the UPGMA dendrogram constructed based on shared allele frequency showed grouping of the 25 genotypes (22 NIFOR parent-progeny palms and 3 MPOB advanced lines used as outgroup) into two major clusters (Figure 2). Cluster I consisted of the two parents and their hybrids while the MPOB experimental selections were grouped in cluster II. This outcome indicated the usefulness of the markers to differentiate among closely related genotypes as well as the 25 oil palm genotypes under investigation. In a related study, Singh et al., (2007) obtained satisfactory separation of six oil palm ortet-ramet sets using five SSR loci.

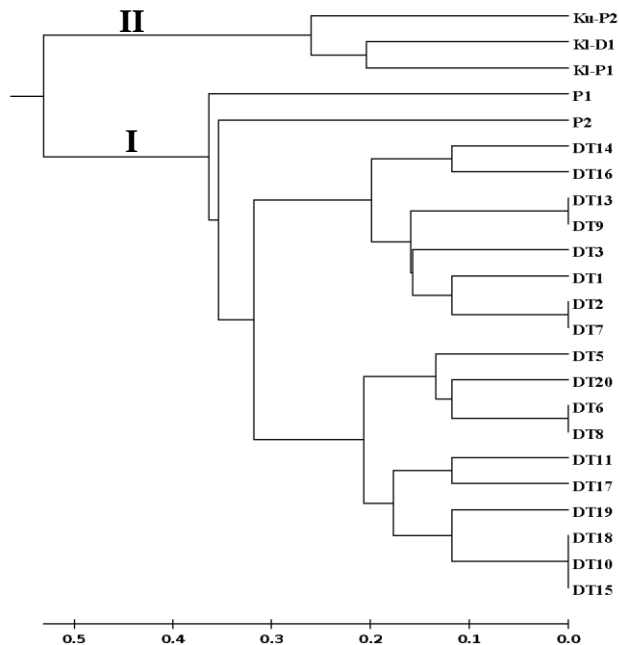


Figure 2. UPGMA dendrogram showing the ability of the three sets of SSR loci to discriminate among the parental genotypes, their hybrids and the three elite populations used as outgroup. The dendrogram was

specifically used to assess the discriminating ability of the SSR for fingerprinting. All the genotypes were derived from a single cross with parents of different origin so their relationship is not in doubt. The two parents and their resulting progeny are in one cluster while the MPOB advanced materials used as outgroup/control were assigned to a different cluster.

Implications to oil palm varietal improvement

Reports show that molecular markers such as microsatellites or simple sequence repeats (SSR) are more faithful in the acceleration of conventional oil palm breeding with respect to verification and identification of hybridity (Singh et al., 2007; Thawaro and Te-chato, 2010; Thongthawee et al., 2012; Bakoumé et al., 2011; Hama-Ali et al., 2015; Budiman et al., 2019). Results from this study indicate that it is possible to accurately and rapidly determine true hybrids in oil palm using PCR-based techniques compared to the cumbersome morphological observation that has been adopted in the NIFOR oil palm breeding programme. In contrast to the morphological method of hybrid identification which usually takes 3-4 years, SSR analysis takes only 1-3 months depending on the number of populations or sample size. Nevertheless, given the occurrence of inter-type contamination in oil palm, it is not easy for SSR markers to distinguish *dura* fruit form from *tenera* or *pisifera* fruit forms. Therefore, the traditional method of hybrid identification will continue to complement molecular marker technology. Early detection of true hybrids indicates precise selection of plants in the field. In an effort to develop marker-assisted selection (MAS) tools for economically important agronomic traits as well as disease resistance in the oil palm breeding programme, these legitimate hybrids could be used as mapping population. The use of MAS will increase selection efficiency, reduce the breeding cycle and enhance variety development. Introduction of an additional phase involving SSR markers in the modified RRS programme of NIFOR will be desirable for the identification of contaminants in addition to selection of recombinant genotypes that will maximize heterosis among populations. According to Babu et al., (2017), significant cost savings could be made by eliminating illegitimate crosses before field planting.

Conclusion

This is the first study confirming the ascendance of some progenies in NIFOR oil palm main breeding programme using microsatellite markers. Eight polymorphic microsatellite

markers were employed for the verification of some oil palm progenies from the breeding programme. In principle, results from the progeny screening highlight the efficiency of the microsatellite markers for analysis of genetic variation and parentage verification. Two individual palms were off-types probably due to genotype mix-ups from pollen contamination during controlled crosses and planting errors, respectively. This also draws attention to the need for constant care, control and organization at all the stages of seed production; from raising of seedlings in the nursery, to planting in the field. It is therefore recommended as a standard practice to test the legitimacy of all crosses in the breeding programme. The three most informative SSR markers (sMg00016, sMg00179 and sMo00102) could be confidently utilized for routine fingerprinting, hybrid purity tests and certification of controlled crosses in oil palm breeding programmes. Although more numbers of plant samples will be required for large scale application in commercial seed production, minor modifications in terms of measurement precision and probabilities will be worked out to circumvent technical drawbacks. The legitimate progenies identified will be useful for reliable inheritance studies and comparison of genetic diversity determined by microsatellite markers to that revealed by agronomic markers, which have rarely been reported in literature.

Acknowledgments

The support and assistance received from the Genomics Unit of the Advanced Biotechnology & Breeding Centre, MPOB in the form of laboratory attachment for microsatellite genetic analyses is gratefully acknowledged. We wish to thank Dr. Ravigadevi Sambanthamurthi for her logistic support, Ms. Rahimah Abd Rahman and Dr. Ting Ngoot-Chin for their technical assistance in the microsatellite assays and capillary electrophoresis. The assistance of Dr. Maizura Ithnin in statistical analysis, and Messrs. Innocent Ani and Hugh Okoye in field work is also acknowledged. The valuable suggestions and comments of Dr. Mehmood Hassan of ICRAF on

an early version of this manuscript is highly appreciated.

References

Amos, W., Sawcer, S. J., Feakes, R. W. and Rubinsztein, D. C. (1996). Microsatellites Show Mutational Bias and Heterozygote Instability. *Nature Genetics* 13: 390-391. doi:10.1038/ng0896

Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D. and Sorrells, M.E. (1993). Optimizing parental selection for genetic linkage maps. *Genome* 36:181-86.

Ataga, C.D., Okoye, M.N., Enaberue, L.O. and Ikuenobe, C.E. (2018). Oil Palm Breeding Programme in Nigeria. In: International Oil Palm Conference (IOPC)-Smoothing the market disequilibria, Medan, North Sumatra, Indonesia, 17 – 19th July 2018.

Babu, B.K., Mathur, R.K., Kumar, P.N., Ramajayam, D., Ravichandran, G., Venu, M.V.B., *et al.* (2017). Development, identification and validation of CAPS marker for SHELL trait which governs *dura*, *pisifera* and *tenera* fruit forms in oil palm (*Elaeis guineensis* Jacq.). *PLoS ONE* 12(2): e0171933. doi:10.1371/journal.pone.0171933

Bakoumé, C., Aziah, M.Y., Praveena, T., Chee, K.T., Suzaini, Y., Hamidah, M., Jangi, M.S., Basiran, M.N., Khairudin, H. and Harikrishna, K. (2011). DNA sequence-based markers for verification of ramet-to-ortet relationship in oil palm (*Elaeis guineensis* Jacq.). *Am. J. Plant Sci.*, 2: 539-48. doi:10.4236/ajps.2011.24064

Billotte, N., Marseillac, N., Risterucci, A.M., Adon, B., Brottier, P., Baurens, F.C., Singh, R., Herran, A., Asmady, H., Billot, C., Amblard, P., DurandGasselien, T., Courtois, B., Asmono, D., Cheah, S.C., Rohde, W., Ritter, E. and Charrier, A. (2005). Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theor. Appl. Genet.*, 110: 754-65. doi: 10.1007/s00122-004-1901-8 PMID: 15723275

Bohra, A., Dubey, A., Saxena, R.K., Penmetsa, R.V., Poornima, K.N., Kumar, N. and Farmer,

A.D. (2011). Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus* spp.). *BMC Plant Biol* 11:56

Budiman, L. F., Apriyanto, A., Pancoro, A. and 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): 937-49.

Corley, R.H.V. (2005). Illegitimacy in oil palm breeding-a review. *J Oil Palm Res* 17(1):64–69.

Corley, R. H. V. and Tinker, P. B. (2016). *The Oil Palm*, 5th Ed. United Kingdom. WILEY Blackwell, p.156.

Doyle, J.J. and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Hama-Ali, E.O., Alwee, S.S.R.S., Tan, G.S., Panandam, J. M., Ling, H.C., Namasivayam, P. and Peng, H.B. (2015). Illegitimacy and sibship assignments in oil palm (*Elaeis guineensis* Jacq.) half -sib families using single locus DNA microsatellite markers. *Mol Biol Rep* 42:917– 925. DOI 10.1007/s11033014-3829-7

Liu, K. and Muse, S.V. (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128–2129.

Marshall, T.C., Slate, J., Kruuk, L.E.B. and Pemberton, J.M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7: 639-55. <http://dx.doi.org/10.1046/j.1365-294X.1998.00374.x>

Mondini, L., Noorani, A. and Pagnotta, M.A. (2009). Assessing plant genetic diversity by molecular tools. *Diversity* 1:19-35.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583-90.

- Okoye, M. N. (2016). Molecular characterization of some NIFOR and elite oil palm breeding populations using microsatellite markers. Ph.D thesis. University of Nigeria. 148 pp.
- Okoye, M. N., Okwuagwu, C.O. and Ataga, C.D. (2001). The traits of economic importance in oil palm, their inheritance and implication in breeding and selection. Paper presented at the In-house Research and Development Seminar on the 17th of October, 2001, ERLSD Hall, NIFOR Benin City. Nigeria.
- Okwuagwu, C. O. (2001). The development of NIFOR high yielding and stable EWS and the consequences of adulteration of oil palm planting materials. A technical paper presented at the meeting of the Major Oil palm Growers in Nigeria, NIFOR. 24th April 2001, ERLSD Hall. p11-13
- Paetkau, D., Calvert, W., Stirling, I. and Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.*, 4: 347-54.
- Peakall, R., Gilmore, S., Keys, W., Morgante, M., Rafalski, A. (1998). Cross-species amplification of soybean (*Glycine Max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution* 15: 1275-1287.
- Peakall, R. and Smouse P.E. (2012). GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28: 2537-2539. <http://bioinformatics.oxfordjournals.org/content/28/19/2537>
- Perrier, X. and Jacquemoud-Collet, J.P. (2006). DARwin software <http://darwin.cirad.fr/darwin>
- Pomiès, V., Riou, V., Flory, A., Manez, A., Jacob, F. and Cochard, B. (2019). Quality Approach in Oil Palm Selection and Seed Production by Using ID Checking Test with 12 SSR Markers. Proceedings of the PIPOC 2019 International Palm Oil Congress (Agriculture, Biotechnology & Sustainability) AP39 Vol.2:230-234.
- Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2: 225-238.
- Singh, R., Jayanthi, N., Tan, S.G., Jothi Malar, P. and Cheah, S.C. (2007). Development of simple sequence repeat (SSR) markers for oil palm and their application in genetic mapping and fingerprinting of tissue culture clones. *Asia Pac J Mol Biol Biotechnol.* 15(3):121–31.
- Singh, R., Zaki, N.M., Ting, N.C., Rosli, R., Tan, S.G., Low, E.T.L., Ithnin, M. and Cheah, S.C. (2008). Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity. *Biologia*, 63: 227-35.
- Singh, R., Low, E.T.L., Ooi, L.C.L., Ong-Abdullah, M., Ting, N.C., Jayanthi, N., Rajanaidu, N., Mohd Din, A., Rozana, R., Mohamad Arif, A.M., Chan, K.L., Mohd Amin, H., Norazah, A., Lakey, N., Smith, S.W., Budiman, M.A., Hogan, M., Bacher, B., Van Brunt, A., Wang, C., Ordway, J.M., Sambanthamurthi, R. and Martienssen, R.A. (2013). The oil palm *SHELL* gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature*, 500, 340-44.
- Smith, J. S. C., Chin, E. C. L., Shu, H., Smith, O. S., Wall, S. J., Senior, M. L., Mitchel, S. E., Kresovich, S. and Ziegler, 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 and Applied Genetics* 95: 163-173. doi:10.1007/s001220050544
- Subashini, V., Shanmugapriya, A. and Yasodha, R. (2014). Hybrid purity assessment in Eucalyptus F1 hybrids using microsatellite markers. *Biotech* 4: 367–373 DOI 10.1007/s13205 013-0161-1
- Teh, C.K., Lee, H.L., Abidin, H., Ong, A. L., Mayes, S., Chew, F. T. and Appleton, D. (2019). A practical genome-enabled legitimacy assay for oil palm breeding and seed production. *BMC Plant Biology*19:470.

Thawaro, S. and Te-chato, S. (2010). Verification of legitimate *tenera* oil palm hybrids using SSR and propagation of hybrids by somatic embryogenesis. Songklanakarin J. Sci. Technol.32 (1):1-8.

Thongthawee, S., Tittinutchanon, P. and Volkaert, H. (2012). Microsatellites for parentage analysis in an oil palm breeding population. Thai J. Genet 3(2):172–81.

Ting, N.C., Noorhariza, M.Z., Rozana, R., Low, E.T., Ithnin, M. and Cheah, S.C. (2010). SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. J. Genet 89:135-45.

USDA FAS Report (2019) Nigeria: Oilseeds and Products Annual. <https://www.fas.usda.gov/data/nigeria-oilseeds-and-products-annual>