

# PHENOTYPIC VARIABILITY OF *Elaeis oleifera* GERMPLASM REVEALED BY PRINCIPAL COMPONENT AND CLUSTER ANALYSES

ZULKIFLI YAAKUB<sup>1\*</sup>; MARHALIL MARJUNI<sup>1</sup>; SUZANA MUSTAFFA<sup>1</sup>; WAN NOR SALMIAH TUN MOHD SALIM<sup>1</sup>; FATIN MOHD NASIR<sup>1</sup>; NORZIHA ABDULLAH<sup>1</sup>; MOHD DIN AMIRUDDIN<sup>1</sup>; RAJANAIDU NOOKIAH<sup>2</sup>; AHMAD KUSHAIRI DIN<sup>3</sup> and MEILINA ONG-ABDULLAH<sup>1</sup>

## ABSTRACT

Despite the fact that large *Elaeis oleifera* germplasm collections have been established and evaluated, little is known about the genetic diversity of its population. For this purpose, 572 *E. oleifera* palms from eight germplasm collections together with one introgressed population (*Oleifera* x *Oleifera*) were evaluated from 17 sets of field data. The data were then subjected to principal component and cluster analyses. Correlation analysis among traits of *E. oleifera* showed a similar trend as *Elaeis guineensis*, and most of the characters displayed a wide range of variation. In the principal component analysis, the first four principal components (PC1-PC4) with eigenvalue of >1.0, accounted for 96.32% of the total variability, suggesting that *E. oleifera* accessions can be clearly differentiated using the traits studied. Cluster analysis showed that all accessions were clustered into two major groups (consisting of four sub-clusters), generally attributed to their geographical locations in Central and South America. The information obtained from this study aims to provide solutions for a more efficient and manageable approach towards conservation of the *E. oleifera* germplasm.

**Keywords:** Central-South America, conservation, correlation analysis, genetic diversity, oil palm.

**Received:** 30 January 2022; **Accepted:** 28 July 2022; **Published online:** 6 September 2022.

## INTRODUCTION

The genetic diversity of the oil palm species is being threatened by population growth, climate change, habitat loss and selection procedures in oil palm breeding programmes. These actions might impair the oil palm from genetic gains and

withstand environmental threats. For that reason, a large accession of *Elaeis guineensis* from Africa and *Elaeis oleifera* from Latin America has been accumulated. Proper characterisation of these germplasm accessions has led to the identification of valuable traits for future oil palm improvements and optimisation of conservation strategies for its continued accessibility, given oil palm's long breeding cycle. As *E. oleifera* and *E. guineensis* may have diverged about 51 million years ago (Singh *et al.*, 2013), the former is clearly rare compared to its African relative that is cultivated as the preferred commercial variety. *Elaeis oleifera* is found in Honduras, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Suriname, Ecuador, Brazil (Escobar, 1981; Meunier, 1975; Ooi *et al.*, 1981; Rajanaidu, 1983) and Peru (de Blank, 1952; Escobar, 1981; Hardon and Turner, 1967; Meunier, 1975), where its distribution is patchy. They appear

<sup>1</sup> Malaysian Palm Oil Board,  
6 Persiaran Institusi, Bandar Baru Bangi,  
43000 Kajang, Selangor, Malaysia.

<sup>2</sup> 10-14-3 Danau Permai Condo,  
Jalan 3/109F, Taman Danau Desa,  
58100, Kuala Lumpur, Wilayah, Malaysia

<sup>3</sup> No 2, Jalan TH 1,  
Tiara Heights, Salak Tinggi, Sepang,  
43900, Selangor, Malaysia.

\* Corresponding author e-mail: [zulkify@mpob.gov.my](mailto:zulkify@mpob.gov.my)

naturally in groves in open grasslands and riverine areas and their distribution is associated with the migratory movement of indigenous Indians (Santos *et al.*, 1986).

The potential genetic variability provided by the *E. oleifera* species is considered a promising resource for genetic improvement of the commercial oil palm *E. guineensis* species. Although not much commercially planted because of its very low yields, the oil palm breeders' interests in the *E. oleifera* species are its high content of unsaturated fatty acids (>60% oleic acid and >18% linoleic acid), low height increment (5-10 cm yr<sup>-1</sup>) and tolerance to diseases such as bud rot and vascular wilt (Corley and Tinker, 2003; Rao *et al.*, 1989). The ability to transmit their traits has led to the production of interspecific hybrids with *E. guineensis* (Barcelos *et al.*, 2015). In addition, high carotene content and drought tolerance are other interesting characteristics of *E. oleifera*. The main drawback is that it is less productive than *E. guineensis* in terms of oil yield, with some degree of sterility precluding its extensive commercial exploitation (Rajanaidu *et al.*, 2017). As suggested by Choo and Yusof (1996), the possibility for *E. oleifera* to be commercialised is through the pharmaceutical industry and this would promote the development of planting materials with niche traits regardless of their yield performance. In general, *E. oleifera* palms within a population were found to be more uniform in comparison with the *E. guineensis* from West Africa. This observation was attributed to *E. oleifera* palms in South America being found in scattered areas which encouraged inbreeding that led to high levels of homozygosity (Rajanaidu, 1985). This was further supported by molecular marker information indicating *E. guineensis* as more heterozygous than *E. oleifera* (Zulkifli *et al.*, 2012).

Several expeditions to Central and South America to accumulate a series of *E. oleifera* collections for research purposes were initiated by breeders starting in 1982 (Kushairi *et al.*, 2017). A collection of 184 *E. oleifera* accessions are developed and maintained in the form of a field genebank as described by Rajanaidu *et al.* (2017). The collection aimed to capture the genetic variability of *E. oleifera* for utilisation in breeding programmes and as a long-term conservation. At present, *E. oleifera* genetic materials are maintained in the form of a field genebank which requires a large planting area and high maintenance cost. Despite these hurdles, genetic diversity studies are deemed essential for crop improvement and conservation in oil palm breeding programmes. Data collected from these field genebanks on a large number of palm traits provide breeders with the necessary information for maximum efficient selection for the said purpose.

Multivariate analysis has been developed to effectively handle a large amount of accumulated

data for breeders to make more precise decisions (Rencher, 2002). Of all multivariate tools, principal component and clustering analyses are extensively used to classify germplasm materials, expose the variability of accessions and determine the genetic relationships among traits (Iqbal *et al.*, 2008). Principal component analysis (PCA) is a technique from applied linear algebra, a descriptive procedure for analysing relationships that may exist in a set of multivariate datasets. It is designed to identify the patterns of data by reducing the number of dimensions or variables (Ekezie, 2013). On the other hand, clustering is an exploratory data analysis tool for grouping accessions into several homogenous groups (Johnson and Wichern, 2007). Both methods have been extensively used to study genetic variability in crop germplasms including oil palm (Ahmad *et al.*, 2014; Li-Hammed *et al.*, 2015). The objectives of this study are to characterise the Malaysian Palm Oil Board (MPOB) *E. oleifera* germplasm collections using multivariate analysis and to identify the contributing characteristics to the overall variation in yield, bunch quality and vegetative traits. This has an important implication for the sampling strategy for field conservation while assuring a good range of diversity is maintained rather than random sampling.

## MATERIALS AND METHODS

### Plant Materials

*Elaeis oleifera* germplasms from eight countries in Latin America were obtained through bioprospection conducted in 1982, amounting to a total of 182 accessions. The palms were planted at the MPOB Research Station Kluang, Johor, Malaysia in an equilateral triangle design with a 9 m planting distance between palms. This study was carried out on 572 *E. oleifera* palms including one introgressed population [*Oleifera* (Suriname) × *Oleifera* (Brazil), (O × O)] from the germplasm collection (Table 1). The palms were planted either in a randomised complete block design (RCBD) or incomplete randomised design (ICRD).

### Data Collection

A total of 17 field data was collected for each palm and classified into three groups; yield, bunch quality components and vegetative measurement. Bunch quality components and vegetative measurements were evaluated using the methods established by Blaak *et al.* (1963) and Breure and Powell (1988), respectively. For yield, three components were evaluated: mean fresh fruit bunch (MFFB), mean bunch number (MBNO) and mean average bunch weight (MABW). Meanwhile, nine parameters for

bunch components were measured: bunch weight (BWT), mean fruit weight (MFW), mesocarp to fruit (MTF), kernel to fruit (KTF), shell to fruit (STF), oil to dry mesocarp (OTDM), fruit to bunch (FTB), oil to bunch (OTB) and kernel to bunch (KTB). Five vegetative traits were also included: frond production (FP), petiole cross-section (PCS), rachis length (RL), leaf area index (LAI) and frond dry weight (FDW).

### Statistical Analysis

Data were systematically extracted from the MPOB-Breeding Information System (MPOB-BIS) (Mohd Din *et al.*, 2012). Based on descriptive statistical analysis, the population mean, standard deviation and coefficient of variation (CV) for each trait were generated using Microsoft Excel version 14.0. PCA and cluster analysis were used to assess the level of diversity and to rank the contributions of the variables. Mean data were analysed for correlation and PCA using SAS version 9.4, after standardisation (mean=0, standard deviation=1). Based on the respective eigenvalues obtained, only a significant proportion of variance was explained by the first few principal components. The linear combinations with variance less than 1 (eigenvalue < 1) were eliminated. In cluster analysis, Ward's hierarchical algorithm based on R-Square was used to construct the phylogenetic relationship.

TABLE 1. NUMBER OF SAMPLES REPRESENTING EACH *Elaeis oleifera* GERMPLASM

Region	Germplasm	No. of samples
Central America	Honduras (HND)	9
	Panama (PAN)	95
	Costa Rica (CRI)	94
South America	Suriname (SUR)	31
	Colombia (COL)	137
	Brazil (BRA)	21
	Ecuador (ECU)	158
	Peru (PER)	10
	Suriname x Brazil (OxO)	17
Total		572

## RESULTS AND DISCUSSION

### Performance of *Elaeis oleifera* Germplasm

Evaluation of germplasm collections is the *sine qua non* of every breeding programme for crop improvement. There is a need for systematic evaluation to describe the collections and in the context of genetic resources, to unravel its various morphological, physiological and developmental characteristics including special features, such as dwarfness, compactness and disease resistance. These activities are required to aggregate useful

genes and gene combinations into beneficial phenotypes and develop a foundation for selection for various breeding programmes. The distribution of variation for yield, bunch quality components and vegetative measurement for *E. oleifera* is given in Tables 2 and 3. The ANOVA showed a highly significant ( $p < 0.01$ ) variation for all studied traits among the germplasms. A high range of variability was recorded for most traits with CV ranging from 5.95% to 78.88%. These results indicate sufficient variability exists for the traits under study, thus providing ample scope for characterisation and selection for future breeding programmes. MFFB yield and MBNO of *E. oleifera* from Central America were higher than the South American germplasms, ranging from 116.82-157.44 kg palm<sup>-1</sup> yr<sup>-1</sup> and 10.23-12.24 bunches palm<sup>-1</sup> yr<sup>-1</sup>, respectively. The OxO introgression progenies displayed the highest bunch weight, *i.e.* MABW of 14.61 kg palm<sup>-1</sup> yr<sup>-1</sup>, followed by the Colombian and Central American *E. oleifera*.

For bunch quality components, all *E. oleifera* accessions recorded a high range of CV for BWT, OTB and KTB, implying that these traits have higher amounts of exploitable genetic variability among the attributes and therefore present a greater potential for these traits to be favoured for selection compared to others (Ndukauba *et al.*, 2015). Besides that, MTF was higher in the South American germplasm, which is in agreement with the mesocarp content of *E. oleifera* in the same region as reported by Rajanaidu *et al.* (2017). In contrast, KTF and STF were lower in the South American collections, except for Colombia. However, despite having higher MTF, OTDM in the South American *E. oleifera* was much lower compared to the ones in Central America. This suggests that increased fruit sets do not affect the OTDM ratio, concurring with previous observations by Haniff and Roslan (2002).

In terms of vegetative measurements, FP was high in *E. oleifera* accessions from Costa Rica, Honduras, Panama and Brazil with CV ranging from 7.82%-13.28%. Other traits such as PCS, RL and LAI were lowest in *E. oleifera* from Suriname, implying that this germplasm is unique in vegetative growth. The lower PCS and RL values contribute to the compactness of this palm (Zulkifli *et al.*, 2019). Thus, the Suriname *E. oleifera* can be considered an important genetic source for the development of compact palms for commercial planting, which allows more palms to be planted per hectare of land with a prolonged economic lifespan. As for FDW, the highest at approximately 45 kg was found in *E. oleifera* from Central America and Colombia with a CV above 20%.

In general, the Colombian germplasm showed similar characteristics as the Central American *E. oleifera*, which is not surprising as both locations are geographically bordering one another. With a few

exceptions, most of the traits in the Oxo population also showed the same trend, which is expected to be the effect of heterosis. Traits, namely the FP, RL, MTF and STF recorded the lowest CV, suggesting that the Colombian germplasm possesses low exploitable genetic variability.

### Correlation analysis

The correlation coefficient,  $r$ , measures the degree (intensity/magnitude) and nature (direction) of association between characters. The correlation analysis of *E. oleifera*'s yield characters is shown in Table 4. MFFB was positively and highly correlated to MBNO ( $r=0.93$ ) and MABW ( $r=0.83$ ). This result is similar to the correlation trends reported by Lopes *et al.* (2012), whereby high positive correlation coefficient values for MFFB-MBNO ( $r=0.84$ ) and MFBB-MABW ( $r=0.52$ ) were observed in *E. oleifera* interspecific hybrids (*E. oleifera*  $\times$  *E. guineensis*). Similarly, high correlation values between MFFB and MABW ( $r=0.96$ ) as well as MBNO ( $r=0.95$ ) were observed on an *E. oleifera* interspecific hybrid (Gomes Junior *et al.*, 2014). The magnitude of the correlation values for MFFB and MBNO is greater than that of MFFB and MABW, implying that MBNO has a greater influence on the total variation in MFFB as compared to MABW, a similar trend observed in *E. guineensis*, with moderate and positive correlation values for MFFB-MBNO ( $r=0.71$ ) and MFFB-MABW ( $r=0.44$ ) (Marhalil *et al.*, 2013). Therefore, selection for a high bunch number could suffice for the improvement of FFB yields in oil palm breeding programmes involving *E. oleifera*. In *E. guineensis* populations, high and moderate positive correlations between MFFB and MBNO were also observed by Okwuagwu *et al.* (2008) ( $r=0.86$ ) and Okoye *et al.* (2009) ( $r=0.65$ ). Correlation between MFFB and MABW in both studies showed positive but low correlation values of only 0.24 and 0.21, respectively.

Meanwhile, MBNO and MABW showed a positive correlation ( $r=0.59$ ), in contrast to *E. guineensis* which normally is negatively correlated as reported in several studies (Gomes Junior *et al.*, 2014; Marhalil *et al.*, 2013; Okoye *et al.*, 2009; Okwuagwu *et al.*, 2008). This suggests that an increase in MABW is associated with a reduction in MBNO and *vice versa*. However, in *E. oleifera*, Lopes *et al.* (2012) reported a moderate positive correlation between these two components ( $r=0.42$ ). This interesting result suggests that *E. oleifera* differs from *E. guineensis* in these traits. This may be due to the unique characteristics of *E. oleifera* bunches, which not only have moderate weight but also possess a high bunch number.

For bunch quality characters, the highest positive correlation was detected between KTF

and STF ( $r=0.90$ ), while the highest negative correlation was observed between MTF and STF ( $r=-0.99$ ). KTF was highly and positively correlated to a few traits such as BWT, STF, OTDM and KTB but negatively correlated with MFW, MTF, FTB and OTB. As expected, MTF was negatively correlated to all bunch quality components except MFW, FTB and OTB. The positive correlation between FTB and OTB ( $r=0.44$ ) suggests that more fruits are required to obtain a higher oil content. In general, the results indicate that an increase in OTB was associated with increases in MTF, MFW and FTB, but was also affected by the negative associations with KTF and STF. Selection of genotypes for low KTF would implicitly lead to palms with relatively low STF but high MTF.

The vegetative character PCS was positively and highly correlated with other vegetative characters such as RL ( $r=0.98$ ), LAI ( $r=0.97$ ) and FDW ( $r=0.97$ ). In addition, the strong association of these traits with yield characters in this study was similar to previous observations by Rafii (1996) in *E. guineensis* and Mohd Din *et al.* (2000) in *E. oleifera*. Palms with higher LAI values will allow more light interception as LAI is related to photosynthesis and dry matter accumulation. Higher LAI is associated with increased yields due to a higher photosynthesis rate, which leads to better dry matter production (Rajanaidu *et al.*, 2011). Therefore, in selecting for yield, higher vegetative growth should be targeted as well as a greater proportion of the photosynthesised dry matter could be diverted to fruit bunches (Corley and Tinker, 2003).

### Principal Component Analysis (PCA)

The partitioning of total variance into its components facilitates the use of oil palm germplasm as a genetic resource for crop improvement programmes. A well-defined relationship between traits through PCA helps know the traits contributing most to variation. The first four principal axes with eigenvalues of  $>1.0$  accounted for 96.32% of the total variation suggesting that the *E. oleifera* accessions can be clearly differentiated using these traits. The eigenvalue is often used to determine the number of major principal components and an eigenvalue of  $>1.0$  has been the rule of thumb to select useful PCs with practical significance and therefore should be retained (Iezonni and Pritts, 1991). The contribution of each variable to the extracted principal components (PCs) is presented in Table 5. Eigenvalues associated with the first two principal components decreased drastically from 11.17 to 2.40. PC1 contributed 65.71% of the total variation in the data set whereas PC2 to PC4 explained the additional 14.12%, 9.99% and 6.50% variations, respectively. These results showed a higher

TABLE 2. THE MEAN AND DISTRIBUTION OF VARIATION FOR YIELD AND VEGETATIVE MEASUREMENT TRAITS IN *Elaeis oleifera* GERMLASMS

Country	N	MFFB		MBNO		MABW		FP		PCS		RL		LAI		FDW	
		Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)
SURINAME	31	8.69	68.52	4.46	37.91	1.94	61.61	19.16	8.95	1.96	21.54	1.25	8.88	0.43	20.6	7.81	16.03
COLOMBIA	137	111.75	44.88	8.61	43.44	13.01	15.53	21.99	14.38	17.68	20.96	4.28	11.24	4.99	27.83	44.18	22.72
COSTA RICA	94	116.82	44.58	10.23	42.20	11.62	19.29	23.82	13.28	16.79	17.81	4.16	14.06	5.10	24.94	45.97	21.89
HONDURAS	9	157.44	30.32	12.24	30.04	12.86	8.36	23.00	11.30	17.72	18.09	3.98	8.48	5.76	20.98	46.93	26.05
PANAMA	95	123.42	35.30	10.86	32.90	11.43	16.88	23.64	11.74	16.43	20.40	4.20	11.55	5.12	26.32	44.73	22.69
BRAZIL	21	38.55	65.30	6.64	61.86	6.10	24.77	23.05	7.82	5.07	20.68	1.99	10.60	1.19	34.39	16.71	17.28
O x O	17	81.45	40.14	5.65	41.54	14.61	14.67	22.53	9.82	15.64	15.22	3.72	12.78	4.10	19.44	40.68	16.19
ECUADOR	158	24.26	45.68	4.66	37.89	5.12	21.60	18.83	10.19	10.54	20.41	3.35	10.23	2.68	23.62	24.21	20.44
PERU	10	12.53	76.10	1.54	69.62	7.58	26.85	17.10	9.33	13.42	15.68	3.79	9.86	3.03	16.36	27.03	16.99

Note: MFFB (kg palm<sup>-1</sup> yr<sup>-1</sup>) - mean fresh fruit bunch; MBNO (bunches palm<sup>-1</sup> yr<sup>-1</sup>) - mean bunch number; MABW (kg palm<sup>-1</sup> yr<sup>-1</sup>) - mean average bunch weight; FP (no.) - frond production; PCS (cm<sup>2</sup>) - petiole cross section; RL (m) - rachis length; LAI - leaf area index; FDW (kg) - frond dry weight.

TABLE 3. THE MEAN AND DISTRIBUTION OF VARIATION FOR BUNCH QUALITY COMPONENT TRAITS IN *Elaeis oleifera* GERMLASMS

Country	N	BWT		MFW		MTF		KTF		STF		OTDM		FTB		OTB		KTB	
		Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)	Mean	CV (%)
SURINAME	31	3.18	78.88	3.41	18.56	48.95	8.04	7.17	27.38	43.88	6.90	27.77	9.72	51.91	13.19	3.28	24.82	3.40	40.09
COLOMBIA	137	12.38	33.24	3.28	16.38	36.99	11.05	13.04	16.51	49.97	6.43	38.37	15.80	46.32	14.45	3.85	39.91	4.00	35.42
COSTA RICA	94	10.57	41.03	3.81	18.83	36.72	10.52	12.53	17.32	50.74	6.07	40.30	13.93	51.33	13.98	4.84	40.57	3.73	43.65
HONDURAS	9	10.27	27.65	3.96	15.36	34.79	14.64	13.18	16.87	52.03	6.45	38.37	12.32	44.98	23.48	3.14	29.41	3.15	47.75
PANAMA	95	10.36	39.64	3.45	21.68	35.79	11.19	12.68	18.85	51.53	5.95	38.16	17.39	47.58	12.91	4.15	39.56	3.82	41.94
BRAZIL	21	11.03	41.86	10.72	29.53	47.32	9.68	9.94	16.49	42.74	8.72	40.43	12.45	50.40	21.28	5.49	32.64	3.79	28.65
OxO	17	14.47	26.74	3.87	14.67	36.77	11.56	12.09	13.75	51.14	6.46	36.04	22.55	53.72	13.13	3.56	58.61	4.55	24.45
ECUADOR	158	6.47	31.16	11.78	17.64	62.13	6.01	6.69	20.86	31.18	9.38	23.85	13.41	55.14	13.56	4.24	35.58	2.81	36.44
PERU	10	6.82	26.21	11.73	26.64	55.86	10.15	6.16	30.99	37.98	11.84	30.12	19.86	59.29	19.80	5.16	39.13	2.97	33.22

Note: BWT (kg) - bunch weight; MFW (g) - mean fruit weight; MTF (%) - mesocarp to fruit; KTF (%) - kernel to fruit; STF (%) - shell to fruit; OTDM (%) - oil to dry mesocarp; FTB (%) - fruit to bunch; OTB (%) - oil to bunch; KTB (%) - kernel to bunch.

TABLE 4. THE CORRELATION COEFFICIENT ANALYSIS, R, OF *Elaeis oleifera* GERMPLASMS BETWEEN 17 TRAITS

Correlation matrix																	
	MFFB	MBNO	MABW	BWT	MFW	MTF	KTF	STF	OTDM	FTB	OTB	KTB	FP	PCS	RL	LAI	FDW
MFFB	1.0000																
MBNO	0.9302	1.0000															
MABW	0.8306	0.5937	1.0000														
BWT	0.6405	0.4768	0.8689	1.0000													
MFW	-0.664	-0.6329	-0.513	-0.3146	1.0000												
MTF	-0.8651	-0.7957	-0.7817	-0.6826	0.8499	1.0000											
KTF	0.9346	0.8737	0.8481	0.7989	-0.7198	-0.9453	1.0000										
STF	0.8121	0.7407	0.7334	0.6174	-0.8749	-0.9917	0.8956	1.0000									
OTDM	0.7244	0.7104	0.6709	0.7651	-0.4521	-0.8276	0.8519	0.7939	1.0000								
FTB	-0.7827	-0.8824	-0.43	-0.3913	0.6573	0.7418	-0.7925	-0.6987	-0.633	1.0000							
OTB	-0.3489	-0.2929	-0.2185	0.0321	0.6734	0.3728	-0.2832	-0.3958	0.158	0.439	1.0000						
KTB	0.3287	0.2327	0.5674	0.7611	-0.5508	-0.6694	0.6255	0.6668	0.6117	-0.2824	-0.1114	1.0000					
FP	0.7925	0.8549	0.6337	0.7308	-0.5792	-0.8404	0.9079	0.7885	0.8749	-0.7401	-0.0755	0.6211	1.0000				
PCS	0.7885	0.545	0.9027	0.6369	-0.3582	-0.5722	0.6748	0.516	0.4537	-0.3073	-0.152	0.2111	0.4084	1.0000			
RL	0.6847	0.4481	0.8193	0.5646	-0.2129	-0.4243	0.5549	0.3614	0.3549	-0.1888	-0.0209	0.1125	0.3038	0.9779	1.0000		
LAI	0.8977	0.7132	0.8913	0.6133	-0.4779	-0.6756	0.7719	0.6184	0.5284	-0.4805	-0.239	0.2011	0.5371	0.974	0.9295	1.0000	
FDW	0.8997	0.7089	0.9367	0.7053	-0.5032	-0.7251	0.8195	0.6672	0.5955	-0.4695	-0.204	0.3271	0.6065	0.9703	0.9189	0.9895	1.0000

Note: MFFB - mean fresh fruit bunch; MBNO - mean bunch number; MABW - mean average bunch weight; BWT - bunch weight; MFW - mean fruit weight; MTF - mesocarp to fruit; KTF - kernel to fruit; STF - shell to fruit; OTDM - oil to dry mesocarp; FTB - fruit to bunch; OTB - oil to bunch; KTB - kernel to bunch; FP - frond production; PCS - petiole cross section; RL - rachis length; LAI - leaf area index; FDW - frond dry weight.

TABLE 5. CONTRIBUTION OF EACH VARIABLE TO THE EXTRACTED FOUR PCs WITH RESPECTIVE EIGENVALUES AND PROPORTION OF TOTAL VARIABILITY

Field evaluation	Traits	PC1	PC2	PC3	PC4
Bunch yield	MFFB	0.287	0.039	-0.146	0.181
	MBNO	0.255	-0.105	-0.151	0.417
	MABW	0.271	0.201	0.073	-0.226
Bunch quality	BWT	0.235	0.081	0.380	-0.175
	MFW	-0.215	0.294	0.268	0.270
	MTF	-0.281	0.182	-0.006	0.082
	KTF	0.295	-0.084	0.050	0.050
	STF	0.268	-0.215	-0.010	-0.132
	OTDM	0.242	-0.117	0.349	0.224
	FTB	-0.218	0.275	0.187	-0.346
	OTB	-0.090	0.212	0.582	0.375
	KTB	0.171	-0.232	0.387	-0.488
	Vegetative measurement	FP	0.253	-0.198	0.206
PCS		0.233	0.392	-0.092	-0.083
RL		0.200	0.469	-0.065	-0.040
LAI		0.257	0.300	-0.159	0.034
FDW		0.268	0.274	-0.073	-0.023
Eigenvalue		11.171	2.400	1.698	1.105
Variance (%)		65.71	14.12	9.99	6.50
Cumulative (%)		65.71	79.83	89.82	96.32

Note: MFFB - mean fresh fruit bunch; MBNO - mean bunch number; MABW - mean average bunch weight; BWT - bunch weight; MFW - mean fruit weight; MTF - mesocarp to fruit; KTF - kernel to fruit; STF - shell to fruit; OTDM - oil to dry mesocarp; FTB - fruit to bunch; OTB - oil to bunch; KTB - kernel to bunch; FP - frond production; PCS - petiole cross section; RL - rachis length; LAI - leaf area index; FDW - frond dry weight.

percentage of variation contributed by the first two components compared to Camillo *et al.* (2014), who reported a total variability of 68.80% for the first two components of Brazil's *E. oleifera* germplasm. High phenotypic variability was observed on traits such as rachis length, height and fruit weight in *E. oleifera* (Rey *et al.*, 2004; Rios *et al.*, 2011). In *E. guineensis* however, a lower genetic variability (55.1%) was quantified for the first two components of yield and fatty acid traits (Li-Hammed *et al.*, 2016).

Variables with significant positive and negative impacts on each PC, especially PC1, are attributed mostly to diversity and are the primary source of variation beneficial for the selection of diverse genetic backgrounds (Denton and Nwangburuka, 2011). PC1 was mostly related to KTF (0.295), MFFB (0.287) and MABW (0.271). This PC exhibited positive effect for all variables except for MFW (-0.215), MTF (-0.281), FTB (-0.218) and OTB (-0.090). These characteristics play a vital role in differentiating the *E. oleifera* accessions and should be considered during selection in breeding

programmes. For PC2, the largest positively significant component loadings were from RL (0.469), followed by PCS (0.392) and LAI (0.300) while the negatives were from KTB (-0.232), STF (-0.215) and FP (-0.198). The largest positive and highly significant coefficient in PC3 was observed in bunch quality components, *i.e.* OTB (0.582), KTB (0.387) and BWT (0.380), and the least negative contributions were from LAI (-0.159), MBNO (-0.151) and MFFB (-0.146). PC4 was positively contributed by MBNO (0.417), OTB (0.375) and MFW (0.270), meanwhile negatively contributed by KTB (-0.488), FTB (-0.346) and MABW (-0.226). In general, vegetative characteristics and bunch quality components had a positive significant influence on PC2 and PC3, respectively. Meanwhile, most of the traits with high positive loading for PC1 and PC4 were mainly associated to yield and bunch quality components traits.

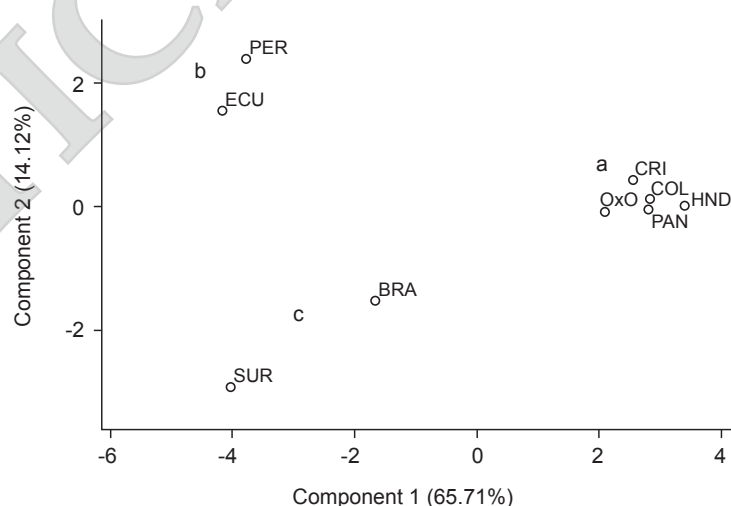
The relationship among the eight locations associated with MPOB *E. oleifera* and Oxo populations was analysed from the score plot

between PC1 and PC2 (Figure 1). This provides a map of how the germplasms relate to each other. The germplasms from Central America, Colombia and OxO were closely connected and positively inclined towards PC1 as they performed well in the variables exhibiting the highest positive loading in PC1 such as MFFB (81.45-157.44 kg palm<sup>-1</sup> yr<sup>-1</sup>), MABW (11.43-14.61 kg palm<sup>-1</sup> yr<sup>-1</sup>) and FDW (40.68-46.93 kg). These populations were also positioned at the centre of PC2 due to being highest in vegetative traits such as RL (3.72-4.28 m), PCS (15.64-17.17 cm<sup>2</sup>) and LAI (4.10-5.76).

The Ecuador and Peru accessions were positioned together based on their high values in MFW, MTF, FTB while moderate in RL, PCS and LAI. Oil palm germplasm collected from Suriname and Brazil were negatively correlated with both PC1 and PC2. This is likely due to the poor performance of both germplasms in most variables such as FDW, MFFB, PCS and RL, in comparison to the other germplasms. Suriname oil palms have been observed to be vegetatively smaller and displayed low bunch production, resulting in low bunch yields compared to *E. oleifera* from other parts of the South American continent (Rao *et al.*, 1989). Field observations of Suriname germplasm exhibited unique physical characteristics among the *E. oleifera* accessions. Notably, OxO performance improved due to hybrid vigour contributed by Brazil and Suriname introgression. The superiority of OxO over both parents is indicated by the positive correlation of OxO to both PC1 and PC2 (Figure 1) which helps detect deviation from normality. Most of the traits exhibited by OxO improved, especially

in yield and bunch quality components such as MFFB (81.45 kg palm<sup>-1</sup> yr<sup>-1</sup>), MABW (14.61 kg palm<sup>-1</sup> yr<sup>-1</sup>), BWT (14.47 kg), KTF (12.09%) and FTB (53.72%).

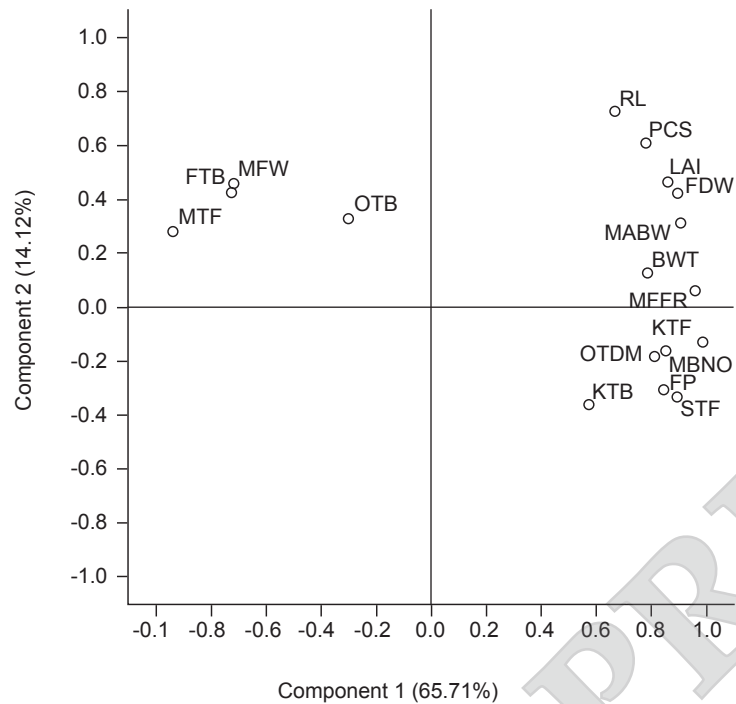
Biplot diagrams provide information to select genotypes that might have favourable combinations of traits that can be strategised in a breeding programme. Most of the variables grouped at quadrants 1 and 4 indicated that those variables were positively correlated to PC1, whilst FTB, OTB, MFW and MTF positioned at quadrant 2 were negatively correlated to PC1 but positively to PC2 (Figure 2). A PCA biplot was constructed to study the relationship between variables and to identify the best variables contributed by each germplasm (Figure 3). Variables with a small angle between them imply a positive and large correlation between them, indicating that an increase in one will increase in the other (Li-Hammed *et al.*, 2016). On the other hand, variables opposite to each other with large angles in between are negatively correlated. For instance, KTB is highly positively correlated with STF due to the small angle between them and is located in the same quadrant. However, they are negatively correlated to MFW and FTB. Therefore, oil palms from Ecuador and Peru have high FTB values but low KTB values. The Suriname oil palm is distinct as there is no interesting or unique variable associated with it. For other populations, most variables in yield parameters and bunch quality components such as MFFB, MBNO, BWT, FP, OTDM and KTF are good characters for selection in the breeding programme as they contribute positively to the primary component (PC1).



Note: (a) Costa Rica (CRI), Colombia (COL), Honduras (HND), Panama (PNM) and Suriname x Brazil (OxO) were positioned together and positively correlated toward both PC1 and PC2. (b) Peru (PER) and Ecuador (ECU) were negatively correlated to PC1 and positive toward PC2. (c) Brazil (BRA) and Suriname (SUR) were negatively correlated to PC1 and PC2.

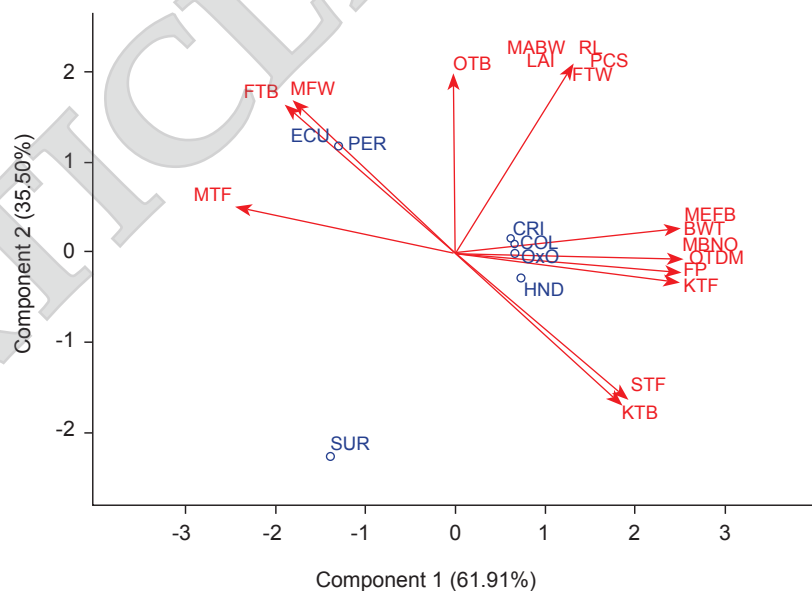
Figure 1. PCA score plot of PC1 and PC2 among the eight locations associated to MPOB *Elaeis oleifera* and OxO populations. Germplasms close to each other have similar field data performance profiles, whereas those far from each other are dissimilar.





Note: Variables in quadrant 1; RL - rachis length; PCS - petiole cross-section; LAI - leaf area index; FDW- frond dry weight; MABW - mean average bunch weight; BWT - bunch weight and MFFB - mean fresh fruit bunch were positively correlated to both PC1 and PC2. Variables in quadrant 2; FTB - fruit to bunch; MFW - mean fruit weight; MTF - mesocarp to fruit and OTB - oil to bunch negatively correlated to PC1 but positively correlated towards PC2. KTF - kernel to fruit; OTDM - oil to dry mesocarp; MBNO - mean bunch number; FP - frond production; KTB - kernel to bunch and STF - shell to fruit together in quadrant 4 were positive to PC1 but negative to PC2.

Figure 2. Loadings plot showing the relationship of 17 variables on PC1 and PC2. Variables contributing similar information are correlated and grouped in the same quadrant.



Note: Germplasm from Peru (PER) and Ecuador (ECU) have a good character in fruit to bunch (FTB) and mean fruit weight (MFW). Brazil (BRA), Costa Rica (CRI), Colombia (COL), Honduras (HND), Panama (PAN) and Suriname x Brazil (OXO) were the best populations for mean fresh fruit bunch (MFFB), mean bunch number (MBNO), bunch weight (BWT), frond production (FP), oil to dry mesocarp (OTDM) and kernel to fruit (KTF). No interesting variable is associated with Suriname (SUR).

Figure 3. PCA Biplot of MPOB *Elaeis oleifera* germplasms shows how strongly each variable influences the principal components.

## Cluster Analysis

PCA is usually complemented with other analyses such as cluster analysis to provide a clear representation of the genetic diversity of the populations under study. The dendrogram (Figure 4) represents the hierarchical cluster analysis based on Ward's method, an approach commonly suggested for clustering quantitative variables (Siracli *et al.*, 2013). This method classifies the materials depending on the environment or correlation of the individuals as claimed by Granato *et al.* (2018). Two main clusters were observed, consisting of four sub-clusters. To envision the clustering pattern of *E. oleifera* germplasm, the mean value for each sub-cluster is presented in Table 6. Cluster 1 was divided into two sub-clusters in which Peru and Ecuador were coupled together in sub-cluster 1, while Brazil and Suriname formed another sub-cluster. MFW, MTF and OTB values for populations in Cluster 1, consisting of Peru, Ecuador, Suriname and Brazil, were greater than the populations in Cluster 2. However, the average MFW, MTF, FTB, OTB and RL values were slightly lower in the Brazil and Suriname populations, therefore comprising a different sub-cluster than Peru and Ecuador.

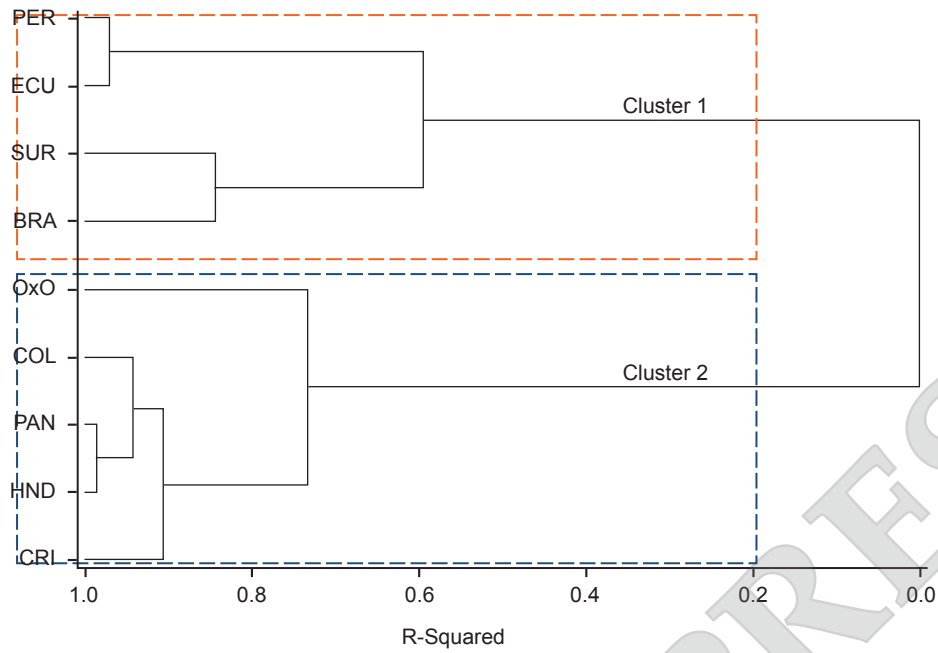
The second cluster (Cluster 2) comprised two sub-clusters, with OxO as a singleton and the second sub-cluster consisting of Colombia, Panama, Honduras and Costa Rica. These populations had higher mean values for MFFB, MBNO, MABW, BWT, KTF, STF, OTDM, PCS, LAI and FDW. In addition, the highest mean values of MABW, BWT and KTB but with lower values for other traits, possibly separated OxO from the other populations in Cluster 2. As mentioned earlier, the traits improvement in OxO resulted in a separate cluster from their original parents, Brazil and Suriname. This may be due to differences in gene combinations compared to both parents, resulting in hybrid vigour (Bingham, 1998). The results of the phylogenetic relationship also reflected the geographical origin of the populations. Populations in Costa Rica, Panama and Honduras tend to cluster together as they form part of Central America. Through sharing a border with Panama, it is not surprising that Colombia in South America is included in Cluster 2. High gene flow may have occurred between populations within countries where similar trends were also observed in *E. guineensis* (Zulkifli *et al.*, 2012) and Napirira bean (Fan and Beta, 2017).

It is a known fact that utilising and conserving germplasms involve proper consideration of developing genetic resources for varietal breeding programmes while avoiding loss of genetic diversity, respectively. Combining the results from correlation studies, principal component and cluster analyses provided knowledge on the patterns of genetic diversity, which are vital for developing

proper breeding strategies for improvement and conservation. Crossing between morphologically distant populations, in this case between populations in Clusters 1 and 2, would theoretically result in a maximum degree of heterosis which could lead to improvement in certain traits of interest (Rahim *et al.*, 2010; Singh *et al.*, 2008). For conservation purposes, the focus should be directed towards germplasm with high variability to develop optimal sampling strategies and minimise any redundancies in the collection. Selection and conservation should also take into account unique populations such as Suriname, as this population may possess interesting genes for further exploitation.

The wealth of diversity in the *E. oleifera* populations in this study can be observed from the wide variability in traits such as MFFB, MBNO, LAI, BWT, FTB, OTB and KTB. However, their trait values are still incomparable to that of the commercial *E. guineensis* (Noh *et al.*, 2012). Although *E. oleifera* is not widely planted because of its very low yield, it has several interesting characteristics, such as compactness, which may be introgressed to improve *E. guineensis*. Compact palms generally demonstrate lower height increment, have shorter fronds (RL) and small petioles (PCS) (Barcelos *et al.*, 2015; Zulkifli *et al.*, 2020). All *E. oleifera* populations in this study showed lower RL and PCS mean values compared to the commercial *E. guineensis* (5.6 m and 29.0 cm<sup>2</sup>, respectively) (Noh *et al.*, 2012). It would then be reasonable to cross populations with extreme mean values for both of these traits, *i.e.* Suriname and Brazil, with commercial *E. guineensis* to derive compact offsprings. The correlation of both traits with other characters should be taken into account for the selection of parental palms in the breeding programme. Therefore, the high variability in *E. oleifera* accessions revealed through multivariate analyses would help breeders explore under-utilised genetic resources and guide future strategies for oil palm genetic improvement.

Although the Malaysian oil palm industry had significant success in generating new planting materials to gain a strong footing, the genetic variation of the planting materials was still considered narrow (Arasu and Rajanaidu, 1976; Rajanaidu *et al.*, 2000). The oil palm industry in Malaysia realised the risk of relying on a limited genetic base and recognised the necessity for an adequate genetic base for effective selection in breeding programmes. This realisation provided the impetus to systematically search for new oil palm genetic materials including *E. oleifera*. The exploitation of the *E. oleifera* germplasm materials in this study will provide potential new genes for incorporation into the existing breeding stock, as well as the flexibility of further developing planting materials including for niche markets. Information obtained by the characterisation of



Note: R-squared value measures the extent to which groups or clusters are different from each other. Germplasm was grouped in two major clusters. Cluster 1 (orange box) comprised germplasm from Peru (PER), Ecuador (ECU), Suriname (SUR) and Brazil (BRA). Cluster 2 (blue box) contained germplasm from Costa Rica (CRI), Honduras (HND), Panama (PAN), Colombia (COL) and Suriname x Brazil (OxO).

Figure 4. Cluster analysis of MPOB *Elaeis oleifera* germplasms based on 17 variables.

TABLE 6. WARD'S CLUSTER MEANS OF 17 TRAITS IN MPOB *Elaeis oleifera* GERmplasm

Traits	Cluster 1 (Sub 1)	Cluster 1 (Sub 2)	Cluster 2 (Sub 1)	Cluster 2 (Sub 2)
MFFB	18.39	23.62	81.45	127.36
MBNO	3.10	5.55	5.65	10.48
MABW	6.35	4.02	14.61	12.23
BWT	6.64	7.10	14.47	10.90
MFW	11.73	7.06	3.87	3.63
MTF	59.00	47.91	36.77	36.07
KTF	6.43	8.55	12.09	12.86
STF	34.58	43.49	51.14	51.07
OTDM	26.99	34.10	36.04	38.80
FTB	57.22	51.16	53.72	47.55
OTB	4.70	4.38	3.56	3.99
KTB	2.89	3.60	4.55	3.64
FP	17.96	21.10	22.53	23.11
PCS	11.98	3.52	15.64	17.09
RL	3.57	1.62	3.72	4.15
LAI	2.85	0.81	4.10	5.24
FDW	25.62	12.26	40.68	45.20

Note: MFFB (kg palm<sup>-1</sup> yr<sup>-1</sup>) - mean fresh fruit bunch; MBNO (bunches palm<sup>-1</sup> yr<sup>-1</sup>) - mean bunch number; MABW (kg palm<sup>-1</sup> yr<sup>-1</sup>) - mean average bunch weight; BWT (kg) - bunch weight; MFW (g) - mean fruit weight; MTF (%) - mesocarp to fruit; KTF (%) - kernel to fruit; STF (%) - shell to fruit; OTDM (%) - oil to dry mesocarp; FTB (%) - fruit to bunch; OTB (%) - oil to bunch; KTB (%) - kernel to bunch; FP (no.) - frond production; PCS (cm<sup>2</sup>) - petiole cross section; RL (m) - rachis length; LAI - leaf area index; FDW (kg) - frond dry weight.

these germplasms is useful as they open new avenues for crop improvement and have the potential for producing novel crop varieties. The ability to select palms for particular traits using new genetic resources would thus represent a major step forward in expediting oil palm breeding.

## CONCLUSION

The richness of any germplasm collection is quantified in terms of the magnitude of the genetic variability of its accessions. It can be concluded from this study that multivariate analyses classified the *E. oleifera* germplasms generally based on their phenotypic characters. Most of the important traits of *E. oleifera* displayed a wide range of variation. Cluster analysis also grouped the *E. oleifera* germplasms into two distinct clusters, which were highly associated with their geographical locations in Central and South America. The optimum traits for measurement can be identified by focusing on variables that contribute most to the variation, to assist the selection of specific populations for regeneration purposes and selective breeding for oil palm improvement programmes. This phenotypic data will be integrated with molecular-based diversity patterns so that breeding and conservation strategies could be systematically formulated for *E. oleifera* populations.

## ACKNOWLEDGEMENT

The authors would like to thank the Director-General of MPOB for permission to publish this paper. The authors also would like to thank Ooi Siew Eng for editing and providing valuable comments on the manuscript.

## REFERENCES

Ahmad, H F; Awar, S I; Aziz, O and Ali, M A (2014). Multivariate analysis of some metric traits in bread wheat (*Triticum aestivum* L.). *European J. Biotechnol. Biosci.*, 1: 22-26.

Arasu, N T and Rajanaidu, N (1976). Oil palm genetic resources. *International Developments in Oil Palm* (Earp, DA, and Newall, W eds.). International Society of Planters, Kuala Lumpur. p. 16-25.

Barcelos, E; Rios, S A; Cunha, R N V; Lopes, R; Motoike, S Y; Babiychuk, E; Skiryycz, A and Kushnir, S (2015). Oil palm natural diversity and the potential for yield improvement. *Front. Plant Sci.*, 6: 190. DOI: 10.3389/fpls.2015.00190.

Bingham, E T (1998). Role of chromosome blocks in heterosis and estimates of dominance and overdominance: Concept and breeding of heterosis in crop plant. *Crop Sci. Soc. Amer.*, 25: 71-87. DOI: 10.2135/cssaspepub25.c6.

Blaak, G; Sparnaaij, L D and Menendez, T J (1963). Breeding and inheritance in the oil palm (*Elaeis guineensis* Jacq.) II: Method of bunch quality analysis. *J. W. Afr. Inst. Oil Palm Res.*, 4: 146-155.

Breure, C J and Powell, M S (1988). The one-shot method of establishing growth parameters in oil palm. *Proc. of the International Oil Palm Conferences: Agriculture*. PORIM, Bangi. p. 203-209.

Camillo, J; Braga, V C; de Abreu Mattos, J K; Lopes, R; da Cunha, R N V; Padilha, J and Scherwinski Pereira, J E (2014). Seed biometric parameters in oil palm accessions from a Brazilian germplasm. *Pesq. Agropec. Bras.*, 49: 604-612. DOI: 10.1590/S0100-204X2014000800004.

Choo, Y M and Yusof, B (1996). PORIM Information Series. *Elaeis oleifera* palm for the pharmaceutical industry. PORIM, Bangi. TT 42.

Corley, R H V and Tinker, P B (2003). Care and maintenance of oil palms. *The Oil Palm*. 4<sup>th</sup> Edition. Blackwell Science, Oxford. p. 287-326.

de Blank, S (1952). A reconnaissance of the American oil palm. *Trop. Agric. Trin.*, 29: 90-101.

Denton, O A and Nwangburuka, C C (2011). Heritability, genetic advance and character association in six yield related characters of *Solanum anguivi*. *Asian J. Agric. Res.*, 5: 201-207. DOI: 10.3923/ajar.2011.201.207.

Ekezie, D D (2013). Principal component analysis, an aid to interpretation data. A case study of oil palm (*Elaeis guineensis* Jacq.). *J Emerg. Trend Engine Appl. Sci.*, 4: 237-241.

Escobar, R., 1981. Preliminary results of the collection and evaluation of the American oil palm *Elaeis oleifera* (H.B.K.) Cortes in Costa Rica, Panama and Colombia. *The Oil Palm in Agriculture in the Eighties* (Pushparajah, E and Chew, P eds.). Incorporated Society of Planters, Kuala Lumpur. p. 17-20.

Fan, G and Beta, T (2017). Discrimination of geographical origin of Napirira bean (*Phaseolus vulgaris* L.) based on phenolic profiles and antioxidant activity. *J. Food Comp. Analy.*, 62: 217-222. DOI: 10.1016/j.jfca.2017.07.001.

- Gomes Junior, R A; de Lima Gurgel, F; de Azevedo Peixoto, L; Bhering, L L; Cunha, R N V; Ricardo, R; Pina, A J A; Veiga, A S (2014). Evaluation of interspecific hybrids of palm oil reveals great genetic variability and potential selection gain. *Ind. Crops Prod.*, 52: 512-518. DOI: 10.1016/j.indcrop.2013.10.036.
- Granato, D; Santos, J S; Escher, G B; Ferreira, B L and Maggio, R M (2018). Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends in Food Sci. Tech.*, 72: 83-90. DOI: 10.1016/j.tifs.2017.12.006.
- Haniff, H M and Roslan, M N M (2002). Fruit set and oil palm bunch components. *J. Oil Palm Res.*, 14: 24-33.
- Hardon, J J and Turner, P D (1967). Observations on natural pollination in commercial plantings of oil palm (*Elaeis guineensis*) in Malaya. *Experiment. Agri.*, 3: 105-116. DOI: 10.1017/S0014479700021864.
- Iezonni, A P and Pritts, M P (1991). Applications of principal component analysis to horticultural research. *Hortic. Sc.*, 26: 334-338. DOI: 10.21273/HORTSCI.26.4.334.
- Iqbal, Z; Arshad, M; Ashraf, M; Mahmood, T and Waheed, A (2008). Evaluation of soybean [*Glycine max* (L.) Merrill] germplasm for some important morphological traits using multivariate analysis. *Pak. J. Bot.*, 40: 2323-2328.
- Johnson, R A and Wichern, D W (2007). *Applied Multivariate Statistical Analysis*. 6<sup>th</sup> Edition. Pearson, New Jersey. 455 pp.
- Kushairi, A; Singh, R and Ong-Abdullah, M (2017). The oil palm industry in Malaysia: Thriving with transformative technologies. *J. Oil Palm Res.*, 29: 431-439. DOI: 10.21894/jopr.2017.00017.
- Li-Hammed, M A; Kushairi, A; Rajanaidu, N; Mohd Sukri, H; Che Wan Zanariah, C W N and Jalani, B S (2015). Multivariate analysis of vegetative and physiological traits in oil palm (*Elaeis guineensis* Jacq) germplasm. *Expert Opin. Environ. Biol.*, 4: 1-5. DOI: 10.4172/2325-9655.1000120.
- Li-Hammed, M A; Kushairi, A; Rajanaidu, N; Mohd Sukri, H; Che Wan Zanariah, C W N and Jalani, B S (2016). Genetic variability for yield, yield components and fatty acid traits in oil palm (*Elaeis guineensis* Jacq) germplasm using multivariate tools. *Intl. J Agri Forest Plant.*, 2: 219-226.
- Lopes, R; Cunha, R N V and Resende, M D V (2012). Bunch yield and genetic parameters of progenies between caiaué and African oil palm. *Pesqui. Agropecu. Bras.*, 47: 1496-1503. DOI: 10.1590/S0100-204X2012001000012.
- Marhalil, M; Rafii, M Y; Afizi, M M A; Arolu, I W; Noh, A; Mohd Din, A; Kushairi, A; Norziha, A; Rajanaidu, N and Latif, M A (2013). Genetic variability in yield and vegetative traits in elite germplasm of MPOB-Nigerian *dura* x AVROS *pisifera* progenies. *J. Food Agric. Environ.*, 11: 515-519.
- Meunier, J (1975). *Le palmier à huile américain* *Elaeis melanococca*. *Oléagineux.*, 30: 51-61.
- Mohd Din, A; Rajanaidu, N and Jalani, B S (2000). Performance of *Elaeis oleifera* from Panama, Costa Rica, Colombia and Honduras in Malaysia. *J Oil Palm Res.*, 12: 71-80.
- Mohd Din, A; Rajanaidu, N; Kushairi, A; Marhalil, M and Zaharah, R (2012). MPOB Breeding Information System (MPOB-BIS). MPOB, Bangi. MPOB Information Series 599.
- Ndukauba, J; Nwofia, G E; Okocha, P I and Ene-Onon, E E (2015). Variability in Egusi-Melon genotypes (*Citrullus lanatus* [Thumb] Matsum and Nakai) in derived Savannah environment in South-Eastern Nigeria. *Int. J. Plant Res.*, 5: 19-26. DOI: 10.5923/j.plant.20150501.04.
- Noh, A; Rafii, M Y; Saleh, G; Kushairi, A and Latif, M A (2012). Genetic performance and general combining ability of oil palm *Deli dura* x AVROS *pisifera* tested on inland soils. *Sci. World J.*, 1-8. DOI: 10.1100/2012/792601.
- Ooi, S; Barcelos, E; Muller, A and Nascimento, J (1981). Oil palm genetic resources- Native *E. oleifera* populations in Brazil offer promising sources. *Pesqui. Agropecu. Bras.*, 16: 385-395.
- Okoye, M; Okwuagwu, C and Uguru, M (2009). Population improvement for fresh fruit bunch yield and yield components in oil palm (*Elaeis guineensis* Jacq.). *Eurasian J. Sci. Res.*, 4: 59-63.
- Okwuagwu, C O; Okoye, M N; Okolo, E C; Ataga, C D and Uguru, M I (2008). Genetic variability of fresh fruit bunch yield in *Deli dura* x *tenera* breeding populations of oil palm (*Elaeis guineensis* Jacq.) in Nigeria. *J. Trop. Agri.*, 46: 40-45.
- Rafii, M Y (1996). *Genotype x environment interaction and stability analyses in 40 oil palm (Elaeis guineensis Jacq.) progenies over six locations*. PhD thesis. Universiti Kebangsaan Malaysia.

- Rahim, M A; Mia, A A F; Mahmud, F; Zeba, N and Afrin, K S (2010). Genetic variability, character association and genetic divergence in mungbean (*Vigna radiata* L. Wilczek). *Plant Omic J.*, 3: 1-6. DOI: 10.3316/INFORMIT.982467342869605.
- Rajanaidu, N (1983). Collection of oil palm (*Elaeis oleifera*) genetic material in Central and South America. PORIM Bulletin, Bangi. p. 1-11.
- 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; 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 Volume 1* ( Basiron, Y; Jalani, and Chan, eds.). Malaysian Palm Oil Board (MPOB) Bangi, Selangor. p. 171-237.
- Rajanaidu, N; Kushairi, A; Mohd Din, A; Marhalil, M; Fadila, A M and Isa, Z A (2011). Selection criteria (15-15-15) to develop sustainable oil palm planting materials. Paper presented at the International Seminar on Breeding for Sustainability in Oil Palm, Kuala Lumpur. 18 November 2011.
- Rajanaidu, N; Kushairi, A and Mohd Din, A (2017). Utilization of germplasm. *Monograph Oil Palm Genetic Resources*. MPOB, Bangi. 289 pp.
- Rao, V; Chin, C W and Rajanaidu, N (1989). Biology and performance of Suriname *Elaeis oleifera* (H.B.K.) Cortes. *J. Oil Palm Res.*, 1: 109-118.
- Rencher, A C (2002). Methods of multivariate analysis. 2<sup>nd</sup> edition. *Wiley series in probability and mathematical statistics*. John Wiley & Sons., New York. 738 pp.
- Rey, B L; Gomez, C P L; Ayala, D I M; Delgado A W and Rocha, S P J (2004). *Colecciones genéticas de palma de aceite Elaeis guineensis (Jacq.) Y Elaeis oleifera (H.B.K.) de Cenipalma: Características de importancia para el sector palmicultor*. *Revista Palmas.*, 25: 39-48.
- Rios, S D A; Cunha, R N V D; Lopes, R; Barcelos, E; Teixeira, P C; Lima, W A A D and Abreu, S C (2011). *Caracterização fenotípica e diversidade genética em subamostras de Caiauê (Elaeis oleifera)*. *Unimontes Científica.*, 13: 49-56.
- Santos, d.M; Barcelos, E and Nascimento, J C (1986). Genetic resources of *E. oleifera* (H.B.K.) Cortes in the Brazilian Amazon. *Proc. of the International Workshop on Oil Palm Germplasm and Utilization*. PORIM, Kuala Lumpur. p. 95-101.
- Singh, M; Kaur, K and Singh, B (2008). Cluster algorithm for genetic diversity. *World Academy of Sci. Eng. Tech.*, 18: 453-457.
- Singh, R; Low, E T L; Ooi, L C L; Meilina, O A; Ting, N C; Jayanthi, N; Rajanaidu, N; Mohd Din, A; Rozana, R; Mohamad Arif, M A; 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: 340-344. DOI: 10.1038/nature12356.
- Siracli, S; Dogan, N and Dogan, I (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *J. Inequal App.*, 203: 1-8. DOI: 10.1186/1029-242X-2013-203.
- Zulkifli, Y; Maizura, I and Rajinder, S (2012). Genetic diversity study of *Elaeis guineensis* germplasm using EST-SSR. *J. Oil Palm Res.*, 24: 1368-1377.
- Zulkifli, Y; Maizura, I; Rajinder, S; Marhalil, M; Mohd Din, A and Ong-Abdullah, M (2019). An RFLP marker, CB75A, predicts for palmitic acid content across selected oil palm germplasm and interspecific-hybrids. *J. Oil Palm Res.*, 31: 315-320. DOI: 10.21894/jopr.2018.0061.
- Zulkifli, Y; Katialisa, K; Rajinder, S; Suzana, M; Marhalil, M; Ting, N C; Mohd Din, A; Leslie, L E T, Leslie, O C L; Kandha, S; Rajanaidu, N; Jansen, J and Meilina, O A (2020). An integrated linkage map of interspecific backcross 2 (BC2) populations reveals QTLs associated with fatty acid composition and vegetative parameters influencing compactness in oil palm. *BMC Plant Biol.*, 20: 356. DOI: 10.1186/s12870-020-02563-5.