EVALUATION OF TEN SINGLE NUCLEOTIDE POLYMORPHISM MARKERS FOR VERIFICATION OF MALAYSIAN COMMERCIAL COCOA CLONES

Lea Johnsiul[1*], Norasekin Tamchek[1], Nuraziawati Mat Yazik[2]

[1]Cocoa Biotechnology Research Centre, Biotechnology Division, Malaysian Cocoa Board, Kota Kinabalu Industrial Park (KKIP), 88460 Kota Kinabalu, Sabah.

[2]CRDC Bagan Datuk, Upstream Technology Division, Malaysian Cocoa Board, Jalan Sg. Dulang, 36307 Perak.

**Corresponding author: lea@koko.gov.my*

Malaysian Cocoa J. 14: 49-57 (2022)

ABSTRACT - *Evaluation of the previously developed single nucleotide polymorphism panel on cocoa individuals from the Malaysian commercial cocoa clone selection was done to test the discrimination power of the panel. The panel was assessed to determine its ability to distinguish among the training set of 568 cocoa individuals collected from the Malaysian commercial cocoa selection. Analyses demonstrated that the single nucleotide polymorphism panel is effective to fingerprint and discriminate between the individual cocoa in the collection which can serve as a very useful tool to verify potential parent clones in cocoa breeding programmes. Furthermore, the availability of this single nucleotide polymorphism panel will also be particularly useful in tissue culture verification and cocoa seedlings / planting materials supply chain quality control.*

Key words: *Theobroma cocoa*, SNP marker, Malaysian cocoa commercial clone, clone verification, planting material

INTRODUCTION

Malaysian commercial cocoa clones are a set of cocoa clones recommended for planting by various research organisations since 1980s until present. It comprises of fifty-three clones which are categorized into 4 groups, from Borneo Abaca Limited (BAL), Balong River Plantation (BR), Koperasi Pembangunan Desa (DESA), Klon-klon Koko MARDI (KKM), Prang Besar Clone (PBC, SAFIMA Plantation -Red Pod (RP) and Malaysian Cocoa Board Clones (MCBC) (Aizat *et al*., 2020). Groups categorization of the clones are based on their suitability with various agro-climatic conditions in Malaysia, their potential yields based on research and published data in the research stations, pod and beans characteristics, level of tolerance to major pest and diseases, butter fat content and flavour of the beans (Aizat *et al*., 2020).

Malaysian Cocoa Board has published identification based on morphological characters or descriptors to help with the clone's identification (Aizat *et al*., 2020) but morphological characteristics are sometimes not precise and less informative. Furthermore, morphological characteristics are often influenced by many genes, not expressed at all growth stages, and easily influenced by environment and field inputs thus making it difficult to assess quickly, objectively and observations need to be done repeatedly to get correct assessment. Hence, for large collections, this traditional or morphological characteristics approach may not be so effective (Korir *et al.*, 2013). The issue of mislabeling errors is common in many breeding programs for plants as well as animals (Banos *et al*., 2001; Visscher *et al*., 2002; Muñoz *et al*., 2014). Variety of mislabeling errors have been observed in many cocoa collections and estimated at 15 to 44% in global cocoa collection (Motilal and Butler, 2003, Sounigo *et al*., 2006, Takrama *et al*., 2005).

In clonal collections, reported rates include 6.9% (Romero *et al.* 2017), 15-44% (Motilal and Butler 2003), 20-100% (Padi *et al*., 2015), 46.4% (Aikpokpodion *et al*., 2010), 57.4%, and 78% (Olasupo *et al*., 2018). The distribution and use of these individuals in breeding programs alter the expected genetic gains resulting from bi-clonal crosses and will affect all subsequent generations when mislabeled germplasm is used in recurrent selection schemes (Adomako, 2006; Dadzie *et al*., 2013). In hybrid seed gardens of West Africa, the frequent use of off-type parents would be a major contributing factor to failures in meeting predicted productivity (Cervantes-Martinez *et al*., 2006; Padi *et al.*, 2015).

Misidentifications can be due to multiple introductions, frequent transfers of plant from point of collection to planting in the field and the recollection of budwood for planting material propagation. The potential human errors in propagation and planting also contribute to this problem (Takrama *et al*., 2014). The distribution and use of these mislabeled individuals in breeding programmes can alter the expected genetic gains resulting from bi-clonal crosses and will affect all subsequent generations when used in recurrent

selection schemes (Adomako, 2006; Dadzie *et al*., 2013). In many cases, the frequent use of off-type parents could be a major contributing factor to failures in meeting predicted productivity (Cervantes-Martinez *et al.,* 2006; Padi *et al.,* 2015).

Molecular fingerprinting techniques are useful tools for the correction of labeling errors in germplasm collections (Takrama *et al*., 2014; Olasupo *et al*., 2018; de Wever *et al*., 2019) and applied in cocoa since the 1980s (Guiltinan *et al*, 2008). Mislabeled cocoa collections had been identified using various kind of molecular marker techniques such as dominant (Figueira *et al*, 1994, Sounigo *et al*., 1997) and co-dominant markers (N'Goran *et al.*, 2000, Efombagan *et al*., 2008; Motilal *et al*., 2010). Although it has been recommended that all parental stock be genotyped before use in breeding programs, the actual magnitude of impact that off-types have on breeding progress in cacao has never been assessed (Takrama *et al.,* 2005; Padi *et al.,* 2015). The impacts of mislabeled trees in breeding trials for cacao have only recently been considered, and in some studies, poor performance and differences in girth were attributed to off-types (Ofori *et al*., 2012; Padi *et al*., 2015). According to DuVal *et al*. (2017), the impact of off-types to breeding progress, even in the present of <5% of the total population, can alter selections by 48%, and affected heritability estimations of desirable traits. Their results showed that even at a low level of off-types in the collection, there is a 41% difference in estimated heritability for yield, which indicates that it can significantly alter estimations of genetic parameters and selections in a breeding programme.

Recent progress in cocoa genomic has led to the use of single nucleotide polymorphisms (SNPs) in cocoa DNA fingerprinting. Previously, we reported on the establishment of two Malaysian Cocoa Board (MCB) cocoa SNP panel comprises of a fifteen-SNPs set and a ten-SNPs set for verification of the Malaysian cocoa commercial clones (Johnsiul and Asim, 2020 and Johnsiul *et al*., In press). The objective of this present study was to evaluate the discrimination power of the ten SNP markers in distinguishing a training set of 602 cocoa individuals collected from the Malaysian commercial cocoa selection.

MATERIALS AND METHODS

Sample collection and SNP Genotyping

Six hundred and two (602) trees from the Malaysian cocoa clones collection, representing fifty-three Malaysian cocoa commercial clones (each clone represented by five to twenty-one trees) were sampled in this study. Samples were collected from various plots in Cocoa Research and Development Centre (CRDC) Bagan Datuk, CRDC Tawau/Madai and Cocoa Biotechnology Research Centre (CBRC) Kota Kinabalu (Table 1). Five leaf discs were collected from each individual cocoa tree. DNA extraction was performed using the LGC DNA extraction service (https://www.biosearchtech.com/services/dna-rna-extraction-services) and SNP genotyping was performed using KASP™ assays from LGC Genomics (http://www.lgcgroup.com/kasp). KASP™ genotyping assays are based on competitive allele-specific PCR and enable bi-allelic scoring of single nucleotide polymorphisms (SNPs) and insertions and deletions (Indels) at specific loci. The raw data were analyzed using LGC's proprietary Kraken™ software and scored on a Cartesian plot, also known as a cluster plot using a SNPViewer software in order to assign a genotype to each DNA sample.

Data Analysis

Raw data was imported and organized in Microsoft Excel for each of the SNP locus and sample call. The approach used to identify mislabeling (off-types) in the collection was to directly compare the reference clones fingerprint profiles with the genotyped samples. Samples with non-matching SNP patterns with the reference were considered off-types.

Table 1: List of the 53 cocoa clones represented by 602 trees, their plot (when available) and tree stand from CRDC Bagan Datuk, CRDC Tawau and CBRC KKIP Malaysian cocoa collection.

Clone Name	Number of trees	Location	Plot Number	Tree Position
BAL 209	5	Bagan Datuk	28C	1, 2, 4, 6, 3
BAL 209	3	KKIP		1, 2, 3
BAL 244	4	Bagan Datuk	19B	4, 5, 4, 5
BAL 244	5	Bagan Datuk	30D	1, 2, 3, 4, 5,
BAL 244	3	KKIP		1, 2, 3
BR 25	12	Bagan Datuk	1A	2, 6, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
BR 25	3	Bagan Datuk	19B	3, 1, 6
BR 25	3	KKIP		1, 2, 3
DESA 1	2	Bagan Datuk	19C	6, 7,
DESA 1	3	Bagan Datuk	19B	10, 1, 3
KKM 15	5	Bagan Datuk	26B	2, 5, 6, 8
KKM 15	3	KKIP		1, 2, 3
KKM 17	5	Bagan Datuk	26B	2, 5, 1, 2, 3,
KKM 17	6	Bagan Datuk	19B	4, 6, 8, 4, 5, 6
KKM 17	3	KKIP		6, 1, 2
KKM 19	6	Bagan Datuk	26B	10, 8, 5, 8, 9, 10,
KKM 19	6	Bagan Datuk	20B	1, 2, 4, 5, 6, 7
KKM 19	3	KKIP		4, 2, 6
KKM 1	2	Bagan Datuk	23D	2, 6
KKM 1	3	Bagan Datuk	19B	12, 1, 4
KKM 1	3	KKIP		1, 2, 3
KKM 22	2	Bagan Datuk	26B	1, 4
KKM 22	3	Bagan Datuk	19B	10, 6, 8
KKM 22	3	KKIP		3, 4, 5
KKM 25	5	Bagan Datuk	26B	2, 4, 2, 3, 6
KKM 25	5	Bagan Datuk	5A	1, 2, 3, 4, 5
KKM 25	3	KKIP		2, 1, 3
KKM 26	8	Bagan Datuk	26B	10, 4, 5, 10, 12, 1, 2, 3
KKM 26	3	KKIP		1, 2, 3
KKM 27	5	Bagan Datuk	26B	2, 5, 2, 4, 5
KKM 27	3	KKIP		1, 4, 6
KKM 28	5	Bagan Datuk	26B	10, 4, 1, 2, 3
KKM 28	6	Bagan Datuk	19B	12, 1, 6, 4, 5, 6
KKM 2	5	Bagan Datuk	26B	2, 4, 1, 2, 3
KKM 2	3	KKIP		1, 2, 3
KKM 3	4	Bagan Datuk	26B	2, 4, 5, 6
KKM 3	4	Bagan Datuk	19B	1, 2, 3, 4
KKM 3	1	Bagan Datuk	20B	5
KKM 3	3	KKIP		1, 2, 7
KKM 4	4	Bagan Datuk	26B	1, 3, 2, 3
KKM 4	3	KKIP		1, 3, 4
KKM 5	2	Bagan Datuk	26B	1, 8

Table 1. Continued

KKM 5	6	Bagan Datuk	22B	1, 2, 3, 4, 5, 6
KKM 5	3	KKIP		1, 2, 3
KKM 6	10	Bagan Datuk	26B	1, 4, 1, 6, 8, 1, 2, 3, 4, 5
KKM 6	3	KKIP		1, 2, 3
MCB C10	2	Bagan Datuk	12A	1, 6
MCB C10	2	Bagan Datuk	31D	1, 2
MCB C10	2	Tawau		2, 3
MCB C11	7	Bagan Datuk	31D	12, 16, 1, 2, 3, 4, 5
MCB C11	3	Bagan Datuk		10, 2, 4
MCB C11	3	Tawau		2, 3, 1
MCB C12	2	Bagan Datuk	31D	12, 16
MCB C12	3	Bagan Datuk		1, 6, 9
MCB C12	3	Tawau		3, 2, 1
MCB C13	2	Bagan Datuk	31D	12, 14
MCB C13	3	Bagan Datuk		10, 11, 4
MCB C13	3	Tawau		1, 2, 3
MCB C14	2	Bagan Datuk	3B	12, 7
MCB C14	3	Bagan Datuk		1, 3, 5
MCB C14	3	Tawau		3, 2, 1
MCB C1	1	Bagan Datuk	1B	10
MCB C1	3	Bagan Datuk	17C	8, 10, 3
MCB C1	3	Bagan Datuk	6B	1, 2, 3
MCB C1	3	KKIP		2, 3, 7
MCB C1	3	Tawau		1, 2, 3
MCB C2	7	Bagan Datuk	1B	1, 5, 1, 2, 3, 4, 5
MCB C2	7	Bagan Datuk	17C	6, 3, 6, 7, 8, 9, 10
MCB C2	3	KKIP		11, 5, 4
MCB C2	3	Tawau		1, 2, 3
MCB C3	7	Bagan Datuk	1B	1, 2, 1, 2, 3, 4, 5

MCB C3	8	Bagan Datuk	17C	1, 7, 4, 6, 7, 8, 9, 10
MCB C3	3	Tawau		1, 2, 3
MCB C3	3	KKIP		6, 1, 5
MCB C4	2	Bagan Datuk	3D	11, 3
MCB C4	6	Bagan Datuk	17C	12, 14, 16, 2, 4, 7
MCB C4	3	KKIP		7, 8, 9
MCB C4	3	Tawau		1, 2, 3
MCB C5	5	Bagan Datuk	1B	3, 2, 3, 4, 5
MCB C5	7	Bagan Datuk	17C	2, 14, 15, 7, 8, 9, 10
MCB C5	3	KKIP		1, 2, 3
MCB C5	3	Tawau		1, 2, 3
MCB C6	2	Bagan Datuk	3E	12, 3
MCB C6	3	Bagan Datuk	7C	7, 9, 1
MCB C6	3	KKIP		1, 2, 3
MCB C6	2	Tawau		2, 3
MCB C7	5	Bagan Datuk	7C	3, 4, 4, 2, 12
MCB C7	3	KKIP		1, 2, 3
MCB C7	2	Tawau		1, 3
MCB C8	5	Bagan Datuk	7C	10, 2, 5, 10, 2
MCB C8	2	KKIP		5, 1
MCB C8	3	Tawau		1, 2, 3
MCB C9	4	Bagan Datuk	7C	11, 4, 4, 5
MCB C9	3	Tawau		1, 2, 3
MCB C9	3	KKIP		2, 3, 4
PBC 112	9	Bagan Datuk	28D	2, 4, 1, 3, 5, 6, 7,8, 9
PBC 112	2		19B	7, 9
PBC 112	2	KKIP		2, 3
PBC 123	12	Bagan Datuk	1A	2, 7, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
PBC 123	3	Bagan Datuk	19B	6, 8, 10
PBC 123	3	KKIP		1, 2, 3
PBC 130	15	Bagan Datuk	27D	1, 2, 1, 2, 4, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
PBC 130	2	KKIP		6, 1
PBC 131	5	Bagan Datuk	27D	1, 4, 2, 3, 5
PBC 131	3	KKIP		1, 2, 3
PBC 137	7	Bagan Datuk	28D	2, 6, 2, 3, 4, 5, 6
PBC 137	3	Bagan Datuk	19B	1, 4, 3
PBC 137	3	KKIP		1, 2, 3
PBC 139	5	Bagan Datuk	19B	12, 8, 11, 14, 9
PBC 139	3	KKIP		1, 2, 4
PBC 140	9	Bagan Datuk	30D	2, 4, 5, 7, 3, 1, 2, 3, 4

Table 1 Continued

PBC 140	3	KKIP		2, 3, 4
PBC 159	2	Bagan Datuk	30D	3, 6
PBC 159	2	Bagan Datuk	19B	1, 5
PBC 159	8	Bagan Datuk	32D	1, 2, 3, 4, 5, 6, 7, 8
PBC 159	3	KKIP		3, 2, 6
PBC 179	9	Bagan Datuk	19B	1, 2, 11, 12, 9, 1, 2, 3, 4,
PBC 179	3	KKIP		1, 2, 3
PBC 221	4	Bagan Datuk	3C	4, 8, 9, 10
PBC 221	10	Bagan Datuk	28D	2, 3, 1, 2, 3, 4, 5, 6, 7, 8
PBC 221	3	KKIP		7, 9, 3
QH 1003	1	Bagan Datuk	29D	3
QH 1003	3	b	19B	11, 1, 3
QH 1003	3	KKIP		1, 3, 5
QH 1176	5	Bagan Datuk	29D	1, 2, 1, 5, 3
QH 1176	1	KKIP		3
QH 1287	2	Tawau		1, 2
QH 1287	3	KKIP		4, 6, 8
QH 186	2	Tawau		1, 2
QH 186	3	KKIP		1, 2, 4
QH 22	15	Bagan Datuk	29D	1, 4, 11, 5, 3, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
QH 22	3	Bagan Datuk	28D	1, 3, 4,
QH 22	2	KKIP		1, 4
QH 240	1	Tawau		2
QH 240	6	KKIP		1, 3, 9, 1, 3, 9
QH 326	5	Bagan Datuk	29D	1, 2, 5, 7, 3
QH 326	3	Bagan Datuk	19B	1, 2, 3
QH 326	6	Bagan Datuk	28D	1, 2, 3, 4, 5, 7
QH 326	2	KKIP		1, 2
QH 37	11	Bagan Datuk	19B	1, 2, 4, 7, 9, 1, 2, 3, 4, 5, 6
QH 441	2	Bagan Datuk		1, 4
QH 441	7	Bagan Datuk	19B	2, 5, 3, 1, 2, 3, 4
QH 441	2	KKIP		3, 4
QH 968	14	Bagan Datuk	29D	1, 2, 4, 5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
RP 1	5	Bagan Datuk	3C	2, 6, 2, 5, 7

RESULTS

SNP Genotyping

Six hundred and two (602) samples genotyped with the ten SNP panels generated high call rates (>90%) across all the tested cocoa samples. The failures of fifty-one loci to be called were probably due to DNA quality or problem with PCR amplification. Thirty-seven samples with one or more loci failed to be called were excluded from this study and five hundred and thirty-one were used in this study.

Multilocus Matching

Due to the very small number of SNP used, stringent scoring was applied where all ten SNP loci to match with reference before considered as true to type. An example of the multilocus matching was presented in Table 2. Majority of the clones in this study were found mislabeled. These trees are called off-types or homonymous mislabeling because they shared the same clone name with the reference clone but had different multilocus SNP profiles. The off-types percentage ranged from 0% to 75%. (Figure 1). Off the fifty-three clones, thirteen clones which were DESA 1, KKM 22, KKM 4, MCB C10, MCB C12, MCB C14, MCB C6, MCB c9, PBC 131, QH 1176, QH 186, QH 240 and RP 1 were all matched among the samples.

Synonymous mislabeling was found in two tree samples of clone MCB C3 collected from CRDC Tawau (Table 2) where the trees were labelled MCB C3 but based on multilocus matching, the SNP profile matched the profile of MCB C2 reference.

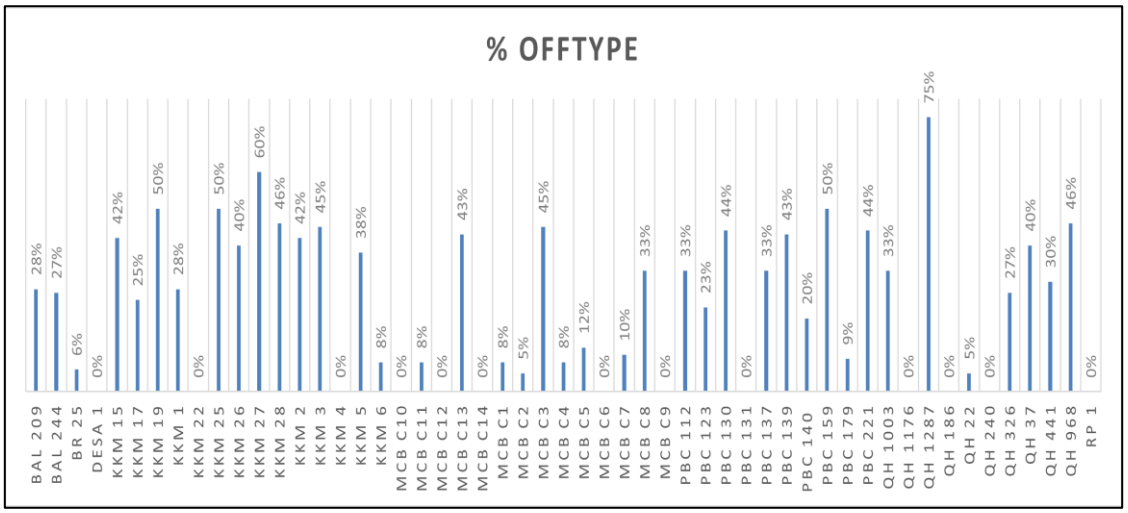

Figure 1. Percentage of off-types in each clone in this study which ranges from 0% (all trees in true to type) to the highest off-type percentage of 75

DISCUSSIONS

The Malaysian cocoa clones are specially selected cocoa clones recommended for planting in various Malaysian agro-climatic environment conditions with desirable traits such as high yield, pest and disease tolerant, good flavour and beans sizes (Aizat *et al*., 2020). These cocoa clones are the planting materials propagated and distributed to farmers all over Malaysia. Frequent multiple introductions, transfers of plants from nurseries to planting in the field, recollection of budwoods for propagation materials and production of grafted seedlings in the nurseries increase the potential of mislabeling errors throughout these processes. The impact of mislabeling errors unfortunately rarely noticed during the early stage of cocoa planting, nevertheless the distribution and use of mislabeled clones in breeding programmes and mislabeled planting materials in farmers' fields can affect the predicted productivity, expected yield and other desirable traits of the clones (Dadzie *et al*., 2013, Padi *et al*., 2015). eventually when the trees start bearing fruits and infested with pests and diseases.

Previously, microsatellite markers were employed for clone identification and mislabeling issues (Johnsiul and Awang, 2019) however genotyping results were not straightforward, often had to be repeated multiple times to ensure elimination of genotyping errors. Another drawback of microsatellite markers is difficulty in comparing generated data with data generated by other organisations due to being binned differently that

may lead to wrong assessment (Takrama *et al*., 2014).

The use of SNP markers has significantly improved the efficiency of clones and off-types identification. The present study used the minimum number of markers, which is only ten SNP markers for identification of mislabeling errors specifically for the fifty-three Malaysian cocoa clones. Based on the results obtained in this study, the selected ten SNP markers were able to distinguish between clones and identify off-types within the fifty-three Malaysian cocoa clones. As shown in Table 2, the ten SNP markers were able to distinguish between clones and identified homonymous and synonymous errors within the Malaysian cocoa clones.

Comparison of multilocus SNP profiles with the reference SNP profiles are straightforward. The use of a small number of SNP markers for the Malaysian cocoa clones will help to reduce cost and suitable for screening large numbers of Malaysian cocoa planting materials.

Table 2: Examples of DNA fingerprints based on multilocus matching of 10 SNPs between the reference and Malaysian cocoa clones (of which only part of three Malaysia cocoa clones were presented).

Genotype	Sample ID	Assessment	Error Type	CSNP1	CSNP2	CSNP3	CSNP4	CSNP5	CSNP6	CSNP7	CSNP8	CSNP9	CSNP10
DESA 1	DESA_1 Ref	Reference	NA	B	D	C	F	G	B	F	H	H	I
DESA 1	DESA_1_7_BD_19C	True to type	NA	B	D	C	F	G	B	F	H	H	I
DESA 1	DESA 1_10_BD_19C	True to type	NA	B	D	C	F	G	B	F	H	H	I
DESA 1	DESA 1_1_BD_19B	True to type	NA	B	D	C	F	G	B	F	H	H	I
DESA 1	DESA 1_3_BD_19B	True to type	NA	B	D	C	F	G	B	F	H	H	I
QH 37	QH_37_Ref	Reference	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH_37_2_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH 37_4_BD_19B	Off-type 1	Homonymous	A	F	G	B	G	F	F	H	H	H
QH 37	QH 37_7_BD_19B	Off-type 1	Homonymous	A	F	G	B	G	F	F	H	H	H
QH 37	QH 37_9_BD_19B	Off-type 1	Homonymous	A	F	G	B	G	F	F	H	H	H
QH 37	QH_37_1_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH_37_2_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH_37_3_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH_37_4_BD_19B	Off-type 2	Homonymous	A	D	G	B	H	B	F	H	G	H
QH 37	QH_37_5_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
QH 37	QH_37_6_BD_19B	True to type	NA	B	F	C	B	H	B	D	H	H	H
MCB C3	MCB C3_Ref	Reference	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB_C3_1_BD_1B	Off-type 1	Homonymous	A	D	C	B	G	B	I	I	H	G
MCB C3	MCB_C3_2_BD_1B	Off-type 1	Homonymous	A	D	C	B	G	B	I	I	H	G
MCB C3	MCB C3_1_BD_17C	Off-type 2	Homonymous	A	F	G	A	G	B	D	I	I	H
MCB C3	MCB C3_7_BD_17C	Off-type 3	Homonymous	B	F	G	A	I	B	F	H	I	I
MCB C3	MCB C3_4_BD_17C	Off-type 4	Homonymous	B	F	G	B	G	B	I	I	I	G
MCB C3	MCB C3_3_TW	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_1_TW	Off-type 5	**Homonymous MCB C3 Synonymous MCB C2**	A	D	G	B	G	B	D	H	H	H
MCB C3	MCB C3_2_TW	Off-type 5	**Homonymous MCB C3 Synonymous MCB C2**	A	D	G	B	G	B	D	H	H	H
MCB C3	MCB C3_6_KK	Off-type 6	Homonymous	A	D	C	A	H	B	F	I	I	H
MCB C3	MCB C3_1_KK	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_5_KK	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_2_BD_1B	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_4_BD_1B	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_5_BD_1B	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C3	MCB C3_6_BD_17C	Off-type 2	Homonymous	B	F	G	A	I	B	F	H	I	I
MCB C3	MCB C3_7_BD_17C	True to type	NA	A	F	G	A	G	B	I	I	I	H
MCB C2	MCB_C2_Ref	Reference	NA	A	D	G	B	G	B	D	H	H	H
MCB C2	MCB_C2_5_BD_1B	Off-type 1	Homonymous	A	D	G	A	G	A	D	H	H	G
MCB C2	MCB C2_11_KK	True to type	NA	A	D	G	B	G	B	D	H	H	H
MCB C2	MCB C2_5_KK	True to type	NA	A	D	G	B	G	B	D	H	H	H
MCB C2	MCB C2_3_TW	True to type	NA	A	D	G	B	G	B	D	H	H	H
MCB C2	MCB C2_4_1B	True to type	NA	A	D	G	B	G	B	D	H	H	H

* NA denotes Not Applicable.

REFERENCES

Adomako, B. (2006). Combining Ability Analysis of Blackpod Disease Incidence in Cocoa Genotypes in Ghana. *Trop Sci* **46**: 201-204.

Aizat, J., Haya, R, Nuraziawati, M. Y., Mohammad Jaafar, H., Zailani, M. J., Jamaludin. R., Suhaida. S., Mohd Qahar Muzzakar. M. A., Mohd Faizal, A., Zamri, K., Sarina, A. S., Ishak, M. Y., Abdul Rahman, R., and Albert, LSC. (2020). Malaysia Cocoa Clones. Second Edition. *Malaysian Cocoa Board.*

Banos, G., Wiggans, G. R., and Powell, R.L. (2001). Impact Of Paternity Errors in Cow Dentification on Genetic Evaluations and International Comparisons. *J Dairy Sci* **84**: 2523-2529.

Cervantes-Martinez, C., Brown, J.S., Schnell, R.J., Phillips-Mora, W., Takrama, J.F., and Motamayor, J.C. (2006). Combining Ability for Disease Resistance, Yield, And Horticultural Traits of Cacao (*Theobroma cacao* L.) clones. *J Am Soc Hortic Sci* **131**: 231-241.

Dadzie, A.M., Livingstone, D.S., Opoku, S.Y., Takrama, J., Padi, F.K., Offei, S. K, Danquah, E.Y., Motamayor, J.C., Schnell, R.J., Kuhn, D.N. (2013). Conversion Of Microsatellite Markers to Single Nucleotide Polymorphism (SNP) Markers for Genetic Fingerprinting of *Theobroma cacao* L. J *Crop Improv* **27**: 215-241.

De Wever, J., Everaert, H., Coppieters, F., Rottiers, H., Dewettinck, K., Lefever, S., and Messens, K. (2019). The Development of a Novel SNP Genotyping Assay to Differentiate Cacao Clones. *Scientific Reports*, **9**: 9512.

DuVal, A., Gezan, S.A., Mustiga, G., Stack, C., Marelli, J.-P., Chaparro, J., and Motamayor, J.C. (2017). Genetic Parameters and the Impact of Off-Types for *Theobroma cacao* L. in a Breeding Program in Brazil. *Front. in Plant Sci*. **8**: 1–12.

Efombagn, I.B.M., Motamayor, J.C., Sounigo, O., Eskes, A., Nyasse, S., Cilas, C., Schnell, R., Manzanares-Dauleux, M. J., Kolesnikova-Allen, M. (2008) Genetic Diversity and Structure of Farm and Genbank Accessions of Cacao (*Theobroma cacao* L.) in Cameroon Revealed by Microsatellite Markers. *Tree Genetics & Genomes* **4:** 821–831.

Figueira A., Janick J., Levey M., Goldsbrough P., (1994) Reexamining the Classification of *Theobroma cacao* L. Using Molecular Markers. *J. Am. Soc. Hortic. Sci*. **119**:1073 - 1083.

Guiltinan M., Verica J., Zhang D., Figuiera A. (2008). Genomics of *Theobroma cacao*, "the Food of the Gods" In: Moore P., Ming R. (eds) *Genomics of Tropical Crop Plants*. Springer. New York. Pp. 145 – 170.

Johnsiul L., and Asim, S. (2020) Development of A Core SNP Panel for Verification of The Malaysian Commercial Cocoa Clones. *International Journal of Agriculture, Forestry and Plantation*, Vol. **10**: 373 – 378.

Johnsiul L., Asim, S., and Mat Yazik, N. (2020). Preliminary Evaluation on The SNP Markers for The Malaysian Commercial Cocoa Clones. *Malaysian Cocoa Journal* (In press).

Johnsiul L., and Awang, A. (2019). Evaluation Of Clonal Uniformity in Class One Malaysian Commercial Cocoa Clones Based on SSR Markers. *International Journal of Agriculture, Forestry and Plantation*, **8**:12 - 17.

Korir, N. K., Han, J., Shangguan, L., Wang, C., Kayesh, E., Zhang, Y., and Fang, J. (2013). Plant Variety and Cultivar Identification: Advances and Prospects. *Critical Reviews in Biotechnology*. **33(2):** 111-25.

Motilal, L., and Butler, D. (2003). Verification Of Identities in Global Cacao Germplasm Collections. *Genet Resour Crop Evol* **50**: 799-807.

Motilal L., Zhang D., Umaharan P., Mischke S., Mooleedha V., Meinhardt LW (2010). The Relic Criollo Cacao in Belize- Genetic Diversity and Relationship with Trinitario and Other Cacao Clones Held in The International Cocoa Genebank, *Trinidad. Plant Genetic Resources* **8**:106-110.

Munoz, P.R., Resende, M.F.R., Huber, D.A., Quesada, T., Resende, M.D.V., Neale D.B., Wegrzyn, J.L., Kirst, M., and Peter, G.F. (2014). Genomic Relationship Matrix for Correcting Pedigree Errors in Breeding Populations: Impact on Genetic Parameters and Genomic Selection Accuracy. *Crop Sci* **54**: 1115-1123.

N'Goran, J. A. K., Laurent, V., Risterucci, A. M., Lanaud, C. (2000). The Genetic Structure of Cacao Populations (*Theobroma cacao* L.) Revealed by RFLP Analysis. *Euphytica* **115**: 83-90.

Ofori, A., Opoku, S., and Padi, F. (2012). Variation Among a Core Set of Cocoa *Theobroma cacao* L.) Clones for Key Agronomic Traits in Ghana. In: *Proceedings of the 17th International Cocoa Research Conference*, Cameroon, 15-20th October 2012.

Olasupo, F.O., Adewale, D., Aikpokpodion, P.O., Muyiwa, A., Bhattacharjee, R., Gutierrez, O., Motamayor, J.C., Schnell, R., Zhang, D. (2018). Genetic Identity and Diversity of Nigerian Cacao Genebank Collections Verified By Single Nucleotide Polymorphisms (SNPs): A Guide To Field

Genebank Management And Utilization. *Tree Genet Genomes* **14**: 32.

Padi, F.K., Ofori, A., Takrama, J., Djan, E., Opoku, S.Y., Dadzie, A.M., Bhattacharjee R., Motamayor, J.C., and Zhang, D. (2015). The Impact of SNP Fingerprinting and Parentage Analysis on The Effectiveness of Variety Recommendations in Cacao. *Tree Genet Genomes* **11**: 44.

Romero Navarro, J.A., Phillips-Mora, W., Arciniegas-Leal, A., Mata-Quirós, A., Haiminen, N., Mustiga, G., Livingstone, D., Van bake, H., Kuhn, D.N., Parida, L., Kasarskis, A., Motamayor, J.C. (2017). Application of Genome Wide Association and Genomic Prediction for Improvement of Cacao Productivity and Resistance to Black and Frosty Pod Diseases. *Front Plant Sci* Vol.**8**.

Sounigo, O., Risterucci, A. M., Clément, D., Fouet, O., Lanaud, C. (2006). Identification Of Off-Types of Clones Used in The International Clone Trial Using DNA Analyses. In*: Global Approaches to Cocoa Germplasm Utilization and Conservation: Final Report of the CDC/CCO/IPGRI Project On "Cocoa Germplasm Utilization and Conservation: A Global Approach*" (1998-2004). Eskes Albertus (ed.), Efron Yoel (ed.). CFC, ICCO, IPGRI. Amsterdam: CFC, pp. 82-86. (CFC Technical paper, 50) ISBN 978-92-9043-734-5

Sounigo, O., Bekele, F., Bidaisee, G., Christopher, Y., and Umaharan, R. (1997). Comparison Between Genetic Diversity Data Obtained from Morphological, Biochemical and Molecular Studies. *The University of the West Indies, Cocoa Research Unit, Report* for 1997:20-29.

Takrama, J. F., Cervantes-Martinez, C., Phillips-Mora, W., Brown, J. S., Motamayor, J. C., and Schnell., R. J. (2005) Determination of Off-Types in a Cocoa Breeding Programme Using Microsatellites. *Ingenics Newsletter*. Issue **10**: 2-8.

Takrama, J., Kun, J., Meinhardt, L.W., Mischke, S., Opoku, S.Y., Padi, F.K. and Zhang, D. (2014). Verification of Genetic Identity of Introduced Cacao Germplasm in Ghana Using Single Nucleotide Polymorphism Polymorphism (SNP) Markers. *J African Biotech* **13**: 2127-2136.

Visscher, P.M., Woolliams, J.A., Smith, D., and Williams, J.L. (2002). Estimation of Pedigree Errors in The UK Dairy Population Using Microsatellite Markers and the Impact on Selection. *J Dairy Sci* **85**: 2368-2375.