African Crop Science Journal, Vol. 20, No. 1, pp. 67 - 75 ISSN 1021-9730/2012 $4.00

APPLYING SNP MARKER TECHNOLOGY IN THE CACAO BREEDING PROGRAMME IN GHANA

J. TAKRAMA, A.M. DADZIE, S.Y. OPOKU, F.K. PADI, B. ADOMAKO, Y. ADU-AMPOMAH, D.S. LIVINGSTONE III[1], J.C. MOTAMAYOR[2], R.J. SCHNELL[1] and D.N. KUHN[1]
Cocoa Research Institute of Ghana (CRIG), P. O. Box 8 Tafo-Akim, Ghana
[1]USDA-ARS, SHRS, 13601 Old Cutler Road, Miami, FL 33158
[2]Mars, Inc., Hackettstown, NJ 07840
Corresponding author's email address: David.Kuhn@ARS.USDA.GOV

(Received 10 July, 2010; accepted 5 January, 2012)

ABSTRACT

In this investigation 45 parental cacao plants and five progeny derived from the parental stock studied were genotyped using six SNP markers to determine off-types or mislabeled clones and to authenticate crosses made in the Cocoa Research Institute of Ghana (CRIG) breeding programme. Investigation was based on the 5' nuclease SNP assay using Illustra Hot Start mix Ready-To-Go PCR strips and BioTek FLx800TBP Fluorescence Microplate Reader. In a group of six cacao plants labeled as PA150 clones and another five labeled as Pound7, one clone in each group was unambiguously determined as off-type or mislabeled. Similarly, in a cohort of 23 PA7 "clones", four genotypes were differentiated. Cross-checking the fidelity of five progeny from the parental stock under study, it was established that no errors were made in the crossing. The most significant outcome of this study, however, was that out of the four categories of 23 PA7 candidate parental trees only one category can be comparable to the reference clone in the International Cacao Germplasm collection, Trinidad (ICG,T); thus informing the need for further work to find the correct clone among these for the breeding programme. It was thus concluded that this simple yet cutting-edge genotyping procedure can be used in applied cocoa breeding programmes in a cocoa producing country. This work represents a first step in the genotypic characterisation of the CRIG germplasm collection and Seed Gardens.

Key Words: Clones, fluorescence microplate reader, genotyping

RÉSUMÉ

Au cours de cette recherche, 45 plants de cacao parentaux et 5 descendants dérivant du stock parental ont été génotypé en utilisant 6 marqueurs SNP, afin de déterminer les clones mal étiquetés et d'authentifier les croisements effectués dans le programme d'amélioration de l'Institut de Recherche sur le Cacao au Ghana (CRIG). Cette étude a été basée sur les 5' nucléases SNP en utilisant des bandes PCR "Hot Start mix Ready-To-Go PCR strips" et un Lecteur Microplat à Fluorescence "BioTek FLx800TBP". Au sein d'un groupe de six plants de cacao étiqueté PA150 et d'un autre groupe de cinq étiqueté Pound 7, il a été déterminé sans ambiguïté qu'un clone par groupe était mal étiqueté. De façon similaire, quatre génotypes différents ont été identifiés dans une même cohorte de clones 23PA7. En vérifiant la fidélité de cinq descendants issus du stock parental étudié, il a été établi qu'aucune erreur n'avait été faite lors du croisement. Le résultat le plus significatif de cette étude a été que, sur quatre catégories de 23 candidats PA7 de souches parentales, une seule pouvait être comparable au clone de référence dans la collection Internationale du Germoplasme de Cacao, Trinidad (ICG,T), démontrant ainsi la nécessité de travaux supplémentaires pour déterminer le clone exact parmi ceux évoqués précédemment. Il a ainsi été conclu que cette méthode avant-gardiste de génotypage, pourtantsimple, peut être utilisée dans les programmes appliqués d'amélioration du cacao dans un pays producteur. Ce travail représente une première étape dans la caractérisation génétique de la collection du germoplasme CRIG et jardins semenciers.

Mots Clés: Clones, lecteur à microplat fluorescent, genotyping

INTRODUCTION

Mislabeling and contamination of controlled crosses is a serious problem in cacao (Takrama *et al.*, 2005). Estimates for misidentification of cacao trees vary from 15 to 44% throughout world cacao germplasm collections (Motilal and Butler, 2003). Pollen contamination has also been identified as another source of off-type trees in progeny from controlled crosses (Schnell *et al.*, 2004). Genotyping in developing countries has not advanced due to countless problems often associated with accessibility to appropriate and affordable molecular genetics equipment and supplies. Inconsistency in the performance of known clones when used as parents in field trials is, in part, due to incorrect genetic identity. The improvement of *T. cacao* using modern breeding techniques can be greatly facilitated by correctly identifying accessions within existing cocoa collections (Young, 1994).

Germplasm characterisation of plant accessions in gene banks has mainly focused on identification of sources of genes for resistance to plant pathogens or pests, and their transfer to cultivated materials. Characterisation of the CRIG germplasm has been based mainly on agronomic traits but available data have always been limiting. Cacao germplasm characterisation using microsatellite markers has been successful and has allowed a new classification system for cacao to be developed (Motamayor *et al.*, 2008). Microsatellites are well known for their high information content and versatility as molecular tools; however, they are difficult to use in applied breeding programmes in cocoa producing countries.

Single nucleotide polymorphisms (SNPs) have captured the attention of researchers because they constitute the most common type of sequence differences between alleles. They can be used in the detection of associations between phenotype and allelic forms of a gene, especially for common diseases that have multifactorial genetics (Lima *et al*., 2009). Due to their potential for greater number of markers and simpler laboratory requirements, SNPs are being developed rapidly in several crop plants including cocoa. Several studies have successfully used SNPs for genotypic studies (Ranade *et al*., 2001; Tranah *et al*., 2003; Johnson *et al.*, 2004; Livingstone *et al*., 2010a).

Recently, Livingstone *et al*. (2010b) developed a SNP assay for genotyping *Theobroma cacao* suitable for use under field conditions that was tested in Ghana at the CRIG. The assay was used for the retrospective tracking of labeling errors, determination of off-types in progeny trials as well as for the unambiguous identification of parental clones. In this paper we present work that was done at the CRIG to assess the feasibility of a new fluorescence assay for routine use for cacao germplasm characterisation and quality control in commercial cacao seed gardens.

MATERIALS AND METHODS

Materials. Leaf samples of cocoa genotypes used in this study were harvested from CRIG germplasm collections at the Cocoa Research Institute of Ghana. FLx800TBP Fluorescence Microplate Reader with on-board BioTek Gen 5 Data Analysis Software and a Tungsten-Halogen lamp and excitation/emission filters set at 485nm/528nm for FAM dye and 540nm/575nm for the VIC dye, was purchased from Jencons Africa, Nairobi, Kenya. Illustra Hot Start Mix RTG strips for PCR (0.2 ml tubes, 12 x 8 strips x 5) was a product of GE Healthcare, USA. All other reagents were molecular biology grade from various suppliers. Six SNP markers, representing six different loci were used in this study; they were w3s465, w7s104, w8s204, w17s189, Cir160s384, Cir211s1036. The SNP markers were named with a locus and position identifier. For example, in the SNP marker w7s104, the w7 represent the locus *WRKY7*, while s104 represents the SNP position at nucleotide 104 in reference to the GeneBank entry *WRKY7* (AY331163) (Livingstone, 2010a). Table 1 presents the primer and probe sequences used in this study.

DNA extraction and quantification. Leaf samples from 45 parental clones and five progeny produced from crosses from this select group of parents were harvested from the CRIG germplasm collection plots for DNA extraction. Samples which could not be extracted same day as harvested were stored in the cold room until

TABLE 1. Primers and Probes used for SNP assay and their corresponding linkage group used for SNP assays

Marker	LG	Locus Name(Accession #)	5' Primer	Probes	3' Primer
w17s189	1	WRKY 17(EF173893)	tgattacactgttacaccaacttagacg	FAM_ctcttgcg(c)gagatatVIC_tcttgc(t)gagatatc	acgtgtaaagaaaggagggaaaactt
w7s104	7	WRKY 7(AY331163)	ccaaggttttgccttaaataaa	FAM_agctgttt(a)ctactattgVIC_agctgttt(t)ctactctg	gcattaagcagtagcaaatgatgtt
w3s463	2	WRKY 3(AY331157)	ttcaaccttagacaatggagacatttc	FAM_aaacctct(c)aactgttgVIC_aaacctct(g)aactgttg	tccaaaataatcaatagatcacatacataaaag
w8s204	9	WRKY 8(AY331166)	cacttagaaaatggaaagcaacagt	FAM_ttcc(c)gagacttgtacttVIC_ttcc(t)gagacttgtacttga	acctagagccagatgatgaattgtatt
Cir160s384	9	mTcCIR160(AJ566490)	atgatggtgacaacagcaagaaa	FAM_caaggatc(a)ttttgctVIC_aaggatc(g)ttttgct	atgcctattaatcacctaggtgagact
Cir211s1036	8	mTcCir211(AJ566534)	accttaatttatgggaaacgaggt	FAM_caatc(a)gtgctgactgVIC_aatc(t)gtgctgactgat	ccaaacaaaatcttaatgcactgtg
Cir37s112	10	mTcCir37(AJ271942)	aaagtgcgtgtgaaggagttcctatc	FAM_ataatggaaga(c)aacttgVIC_ataatggaaga(g)aacttggt	gcatggaacgatccaagttagtc

needed. DNA was extracted by a modification of the CTAB method by Keb-Llanes *et al*. (2002). DNA was quantified using a Sybr Green I assay according to Livingstone *et al*. (2009). According to this protocol, 24 ul of water (molecular biology grade), 1ul of sample DNA, and 25ul of 60 x Sybr Green I (diluted with 10 mM Tris buffer pH 7.5, 1 mM EDTA) were mixed in a 96-well plate. The plate was incubated at room temperature in the dark for 5 min for efficient binding of Sybr Green to DNA and fluorescence measured on a BioTek Fluorescence Microplate reader with excitation and emission wavelength filters set at 485 and 528 nm, respectively. Two-phase standard curves, broad range curve (Fig.1a) and narrow range curve (Fig. 1b), constructed with salmon sperm DNA, were used for the quantification of DNA. After the quantification, all samples were normalised to 4ng μl^{-1}. The normalisation programme computed the respective volumes of DNA and water required to give a final concentration of 4ng μl^{-1} for each sample. DNA once quantified was stored in the freezer and dilutions made accordingly when needed.

The SNP assay. SNP assays were performed by the 5' nuclease (Taqman) assay (Holland *et al*., 1991; Livak, 1999) as modified by Livingstone *et al*. (2010b). The SNPs assays were set up using Illustra Hot Start mix Ready-To-Go PCR strips (RTG strips) each containing all PCR reagents and Taq polymerase as a dried pellet. Each reaction was made up in a total volume of 25ul by combining 2.25µl each of forward and reverse primers (10mM), 0.5µl of FAM-labeled probe (10mM), 0.5µl of VIC-labeled probe (10mM), 18.5µl water and 1µl of template DNA (4ng μl^{-1}) with Ready-To-Go PCR strips. A No-Template-Control (1µl water instead of DNA) was included in the assay. The samples were centrifuged, placed into the FLx 800TBP Fluorescence Microplate Reader and a pre-fluorescence measurement (pre-read, background fluorescence) made and exported into Excel. This was immediately followed by PCR amplification using Eppendorf Master Cycler with the following cycling parameters: one cycle at 95°C for 10 min, 40 cycles at 95°C for 15 sec then 60°C for 1 min. After amplification, a second fluorescence measurement was taken on the FLx 800TBP

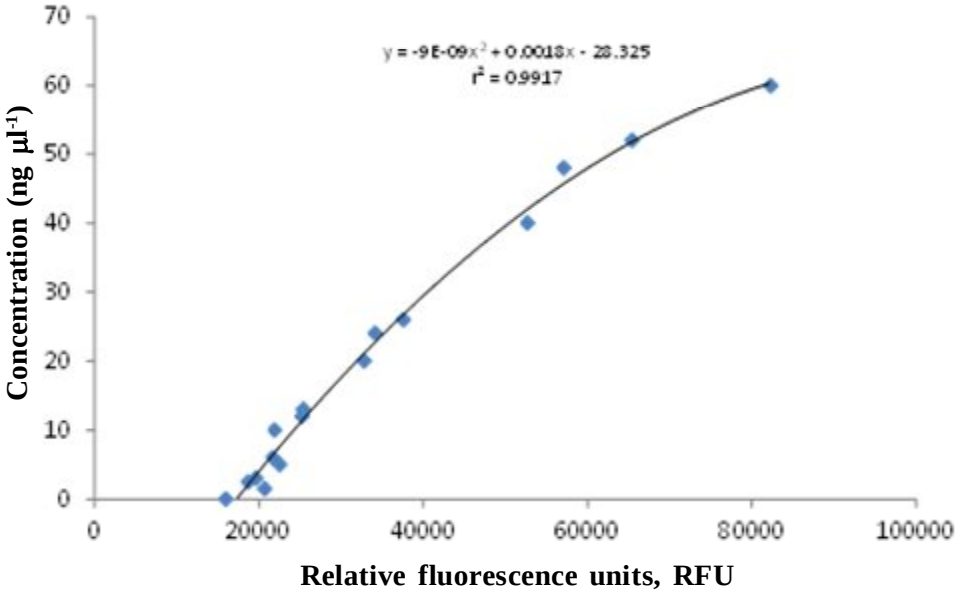

Figure 1a. DNA Quantification standard curve (broad range curve). RFU = relative fluorescence unit.

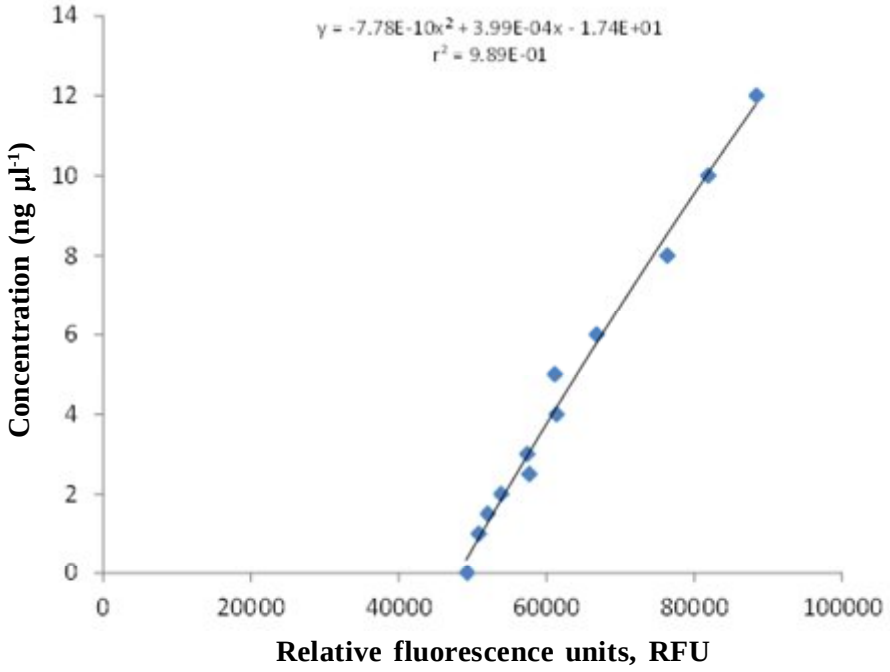

Figure 1b. DNA quantification standard curve (narrow range curve). RFU = relative fluorescence units.

Microplate Reader. The fluorescence values (Relative Fluorescence Units, RFU) for each plate were exported using the on-board Gen 5 control software to an Excel spreadsheet which subtracts background fluorescence, and plots FAM vs. VIC fluorescence for each sample. Samples homozygous for the FAM-labeled allele cluster along the y-axis whilst those homozygous for

the VIC-labeled allele cluster along the x-axis. Heterozygous samples cluster between the two. Allele data for each SNP marker with respect to each genotype was compiled in the Excel spreadsheet (Livingstone *et al.,* 2010b).

Genotypic fingerprints. The sequence of allele calls for all markers constitutes a fingerprint for a given genotype. These fingerprints for all members of a named clone were compared and 'clones' having identical fingerprints were considered identical for the named genotype. All others were considered as 'off-types' or mislabeled clones. Generally, only individuals that differed at two or more loci (markers) were considered different from each other, although samples that differed at one marker point were re-analysed and when confirmed were viewed with less confidence awaiting future confirmation with more markers (Livingstone *et al*., 2010b). Any SNP markers which repeatedly produced poor clustering on VIC-FAM graphs were not used in determining fingerprints of the clones.

RESULTS AND DISCUSSION

Figure 1a and 1b show dual standard curves used for the quantification of fluorescence. The narrow range curve standard (0-12 ng μl^{-1}) and the broad range curve standard (0-48 ng μl^{-1}) were included such that a gain sensitivity on the plate reader was set to autoadjust to 48 ng μl^{-1} for the broad range curve standard and to 12 ng μl^{-1} for the narrow range curve standard. This sets the fluorescence value of each sample near the maximum allowed by the plate reader. By always autoadjusting (i.e. calibrating) the gain of the standards to the same maximum fluorescence, variation in plate reader's measurements were eliminated (Livingstone *et al.,* 2009).

SNP assays for the six markers were performed with 45 parents and five progeny. For each assay, i.e., marker against the 50 clones, a cluster plot was generated in Excel. The resulting plots were used to identify allele groupings (clusters) homozygous for FAM or VIC or heterozygous for both FAM and VIC and the identified genotypes are stored in an excel table. Figure 2a and 2b show two representative FAM vs VIC scatter plots for two markers. The Illustra Hot Start mix RTG assay was used for most of the assays; Figure 2a is representative of this assay type. Regular Taq polymerase (Sigma-Aldrich, St. Louis, Mo) assays were also used when the Illustra Hot Start mix RTG ran out and Figure 2b illustrates this use.

While the use of the Illustra Hot Start RTG PCR 96 well plates was the most convenient assay, their high cost limited their routine usage. Several attempts were made to use regular Taq polymerase to constitute the SNP assays but we were only successful with Taq from Sigma-Aldrich (Fig. 2b). Taq polymerases from other sources including Accupower PCR PreMix (Bioneer Co., Korea) used regularly in our laboratory for other work did not generate any fluorescence in the SNP assay. Taq from other sources tended to quench the fluorescence probably due to some unknown compounds in those preparations. Other sources of enzymes are being explored.

Table 2 is a tabulation of all the genotyping results as determined from cluster plots (e.g., Fig. 2a and 2b). Cacao trees bearing same name tags were most similar in their fingerprints differing only at one to three SNP loci. The data unambiguously projected PA150_2 as different from the other five PA150 clones and presumed an off-type or mislabeled tree. Similarly Pound7_2 is clearly different from the remaining four in its category. Furthermore, PA150_2 is identical to PA121; and T85/799 is the same as T63/971 suggesting some mislabeling occurred. Of the 23 trees labeled PA7, four different genotypes were distinguished suggesting the propagation of offtypes. Additional genotyping experiments will be needed to confirm these results and to distinguish the real PA7 from the remainder.

That only individuals that differed at two or more loci (markers) were considered not identical, it was concluded that all five crosses (Table 3) made were presumed correct, as parental alleles were found in all progeny except at one locus each for PA150 x Pound7 and PA7 x IMC67. In the case of PA150 x Pound7, among the six likely parents genotyped only PA150_5 provided all but one progeny allele. For PA7 x IMC67, more stands of IMC67 need to be genotyped to make an unambiguous determination and furthermore the correct PA7 should be unequivocally

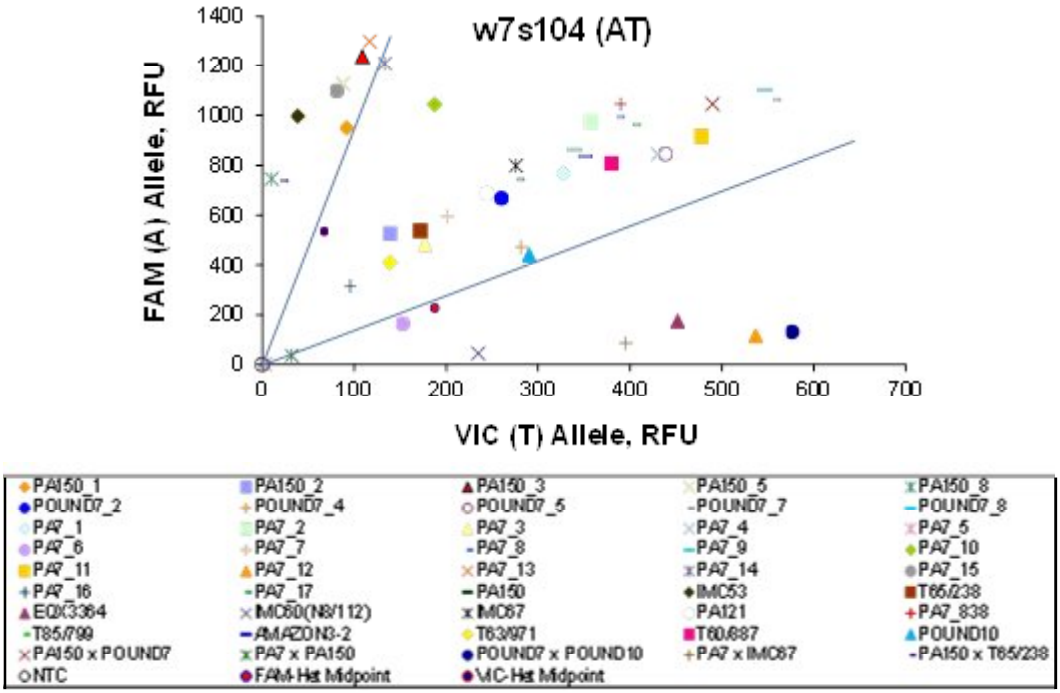

Figure 2a. Cluster graph for SNP marker w7s104 plotted for all genotypes (Illustra Hot start RTG PCR plates used for assay). Arrows are virtual demarcation lines separating genotypes.

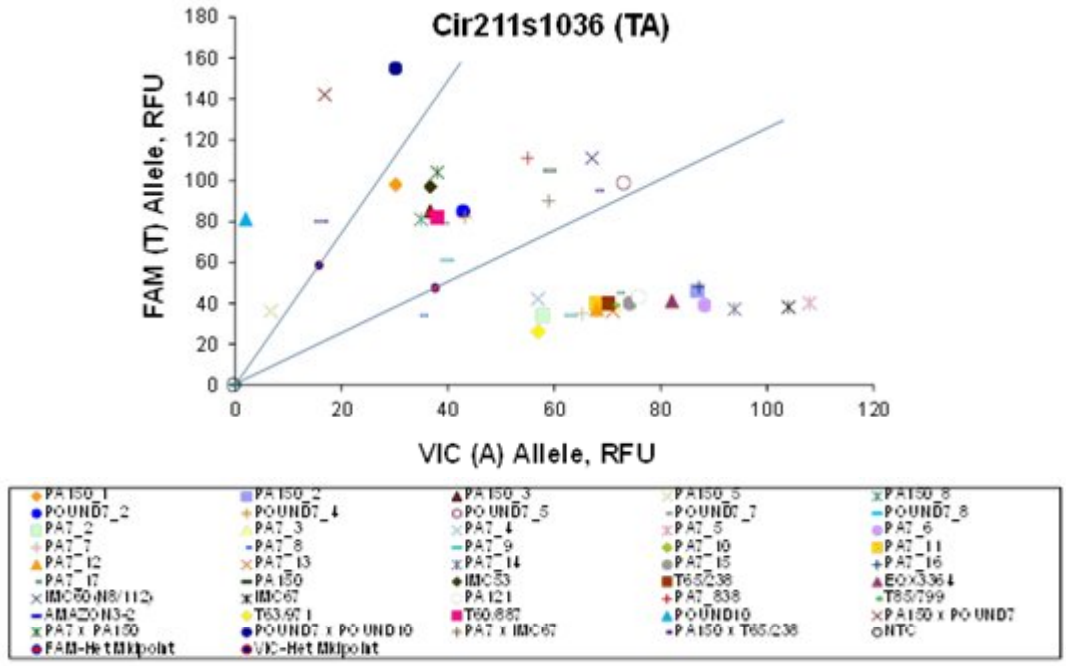

Figure 2b. Cluster graph for SNP marker Cir211s1036 for all genotypes using regular Taq pol from Sigma-Aldrich. Arrows are virtual demarcation lines separating genotypes.

TABLE 2. Diversity panel of genotypes at the six SNP loci

Genotype name	w17s189	w7s104	w3s463	w8s204	Cir160s384	CIR211s1036
PA150_1	C	A	G	C/T	A/G	T/A
PA150_3	C	A	G	C/T	A/G	T/A
PA150_5	C	A	G	C/T	A/G	A
PA150_8	C	A	G	C/T	A/G	T/A
PA150_9	C	A	G	C/T	A/G	T/A
PA150_2	C/T	A/T	G	C	A	T
PA121	C/T	A/T	G	C	A	T
POUND7_2	C/T	A/T	G/C	T	A	T/A
POUND7_4	T	A/T	G/C	C	A	T/A
POUND7_5	T	A/T	G/C	C	A	T/A
POUND7_7	T	A/T	G/C		A	T/A
POUND7_8	T	A/T	G/C	C	A	T/A
POUND10	T	A/T	G	C	A	A
PA7_1	C/T	A/T	G			T
PA7_2	C/T	A/T	G	C		T
PA7_3	C/T	A/T	G	C	G	T
PA7_7	C/T	A/T	G	C	G	T
PA7_8	C/T	A/T	G	C	G	T
PA7_9	C/T	A/T	G	C	G	T
PA7_10	C/T	A/T	G	C	G	T
PA7_16	C/T		G	C	G	T
PA7_17	C/T	A/T	G	C	G	T
PA7_838	C/T	A/T	G	C	G	T
PA7_4	T	A/T	G	C	G	T
PA7_5	T		G	C	G	T
PA7_6	T		G	C	G	T
PA7_11	T	A/T	G	C	G	T
PA7_12	T	T	G	C	G	T
PA7_13	C	A	G	C	G	T
PA7_14	C	A	G		G	T
PA7_15	C	A	G	C	G	T
PA7_18	C/T	A	G	C	G	T
PA7_19	C/T	A	G	C	G	T
PA7_20	C/T	A	G	C	G	T
PA7_21	C/T	A	G	C	G	T
PA7_22	C/T	A	G	C	G	T
IMC53	C	A	G/C	C/T	A	T/A
T65/238	C/T	A/T	G/C	C	A/G	T
EQX3364	T	T	C	C/T	A	T
IMC60(N8/112)	T	T	G/C	C	A	T/A
IMC67	T	A/T	G/C	C/T	A	T
T85/799	C/T	A/T	G/C	C	A	T
T63/971	C/T	A/T	G/C	C	A	T
AMAZON3-2	C/T	A/T	C	C	A	A
T60/887	C/T	A/T	G/C	C/T	A/G	T/A
PA7 x PA150	C/T		G	C/T	A/G	T/A
PA150 x POUND7	T	A/T	G	T	A	A
POUND7 x POUND10	T	T	G/C	C/T	A	A
PA7 x IMC67	T	T	G	C/T	A/G	T/A
PA150 x T65/238	C	A	G	C/T	G	T/A

TABLE 3. Diversity Panel for parental and progeny genotypes at the six SNP loci

Genotype name	w17s189	w7s104	w3s463	w8s204	Cir160s384	CIR211s1036
PA150_5	C	A	G	C/T	A/G	A
PA150xT65/238	C	A	G	C/T	G	T/A
T65/238	C/T	A/T	G/C	C	A/G	T
PA150_5	C	A	G	C/T	A/G	A
PA150 x POUND7	T[1]	A/T	G	T	A	A
POUND7_2	C/T	A/T	G/C	T	A	T/A
POUND7_2	C/T	A/T	G/C	T	A	T/A
POUND7 x POUND10	T	T	G/C	C/T	A	A
POUND10	T	A/T	G	C	A	A
PA7_4	T	A/T	G	C	G	T
PA7 x PA150	C/T		G	C/T	A/G	T/A
PA150_5	C	A	G	C/T	A/G	A
PA7_4	T	A/T	G	C	G	T
PA7 x IMC67	T	T	G	C/T	A/G	T/A[1]
IMC67	T	A/T	G/C	C/T	A	T

[1]genotypes that could not be inherited from the parents shown

identified. Data from this work indicate the need to authenticate clones in CRIG's breeding programme by genotyping reference clones obtainable from the Cacao Quarantine Facility in Reading (UK). Working with tree crops in a cacao producing country where resources are not unlimited, the less number of highly polymorphic markers used the more useful the method. The strategy adopted, therefore, is not to increase the markers indefinitely but to genotype all stands perceived as parents. Further work is in progress genotyping most of the crosses in Seed Gardens together with several candidate parents.

In large breeding programmes, such as that at CRIG, parental clones may become mislabeled or rootstock may be mistaken for the cloned parents. Selected parents are placed in Seed Gardens and manual pollinations used to produce hybrid pods or/and seedlings; these are distributed to farmers for commercial planting. The inclusion of misidentified parental trees in the pollination programme in the past has led to mistrust and disappointment among some farmers who were supplied with the wrong "hybrid" seedlings. The identification of four categories of PA7 poses a serious mislabeling problem in the CRIG breeding programme and steps are being taken to address it. The validation of parental stocks is of uttermost importance as the Ghana cocoa industry relies on seeds supplied by government sponsored Seed Gardens. This work clearly demonstrates a quick and easy way to use the SNP assay to identify off-types or mislabeled clones as well as verify crosses in the applied cacao breeding programme at CRIG.

ACKNOWLEDGEMENT

This work was supported by the USDA under ARS SCA 58-6631-6-086F programme. We thank Eric Brenyah and Clement Aryee of the Physiology/Biochemistry Division of CRIG for technical assistance. This work is published with permission of the Executive Director of CRIG.

REFERENCES

Holland, P.M., Abramson, R.D., Watson, R. and Gelfand, D.H. 1991. Detection of specific polymerase chain reaction product by

utilising the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences of the United States* 88:7276-7280.

Johnson, V. J., Yucesoy, B. and Luster, M. I. 2004. Genotyping of single nucleotide polymorphisms in cytokine genes using real-time PCR allelic discrimination technology. *Cytokine* 27:135-141.

Keb-Llanes, M., Gonzalez, G., Chi-Manzanero, B. and Infante, D. 2002. A Rapid and simple method for small-scale DNA extraction in Agavaceae and other tropical plants. *Plant Molecular Biology Reporter* 20: 299a – 299e

Lima, L.S., Gramacho, K.P., Carels, N., Novais, R., Gaiotto, F.A., Lopes, U.V., Gesteira, A.S., Zaidan, H A., Cascardo, JC M., Pires, JL. and Micheli, F. 2009. Single nucleotide polymorphisms from *Theobroma cacao* expressed sequence tags associated with witches' broom disease in cacao. *Genetics and Molecular Research* 8:799-808.

Livak, K.J. 1999. Allelic determination using fluorogenic probes and the 5' nuclease assay. *Genetic Anal Biomol Eng.* 14:143-149.

Livingstone, D.S., Freeman, B., Motamayor, J.C., Dadzie, A.M., Takrama, J., Schnell, R. J. and Kuhn, D.N. 2010a. A SNP assay for genotyping Theobroma cacao suitable for use under field conditions. *Plant and Animal Genome Conference*, San Diego, CA. Town & Country Convention Center, 9-13 January, 2010.

Livingstone, D.S., Motamayor, J.C., Schnell, R. J., Cariaga, K., Freeman, B., Meerow, A.W., Brown, J.S. and Kuhn, D.N. 2010b. Development of single nucleotide polymorphism markers in Theobroma cacao and comparison to simple sequence repeat markers for genotyping of Cameroon clones. *Molecular Breeding* (DOI 10.1007/s11032-010-9416-2).

Livingstone, D., Freeman, B., Tondo, C.L., Cariaga, K.A., Oleas, N.H., Meerow, A.W., Schnell, R.J. and Kuhn, D. 2009. Improvement of high-throughput genotype after implementation of a dual – curve Sybr Green I – based quantification and normalization procedure. *Horticultural Science* 44:1228 - 1232.

Motamayor, J.C., Lachenaud, P., da Silva, e Mota, J.W., Loor, R., Kuhn, D.N., Steven Brown, J. and Schnell, R.J. 2008. Geographic and genetic population differentiation of the Amazonian Chocolate Tree (*Theobroma cacao* L.). *PLoS ONE* 3(10): e3311. doi:10.1371/journal.pone. 0003311.

Motilal, L. and Butler, D. 2003. Verification in global cacao germplasm collections. *Genetic Resources and Crop Evolution* 50:799-807.

Ranade, K., Chang, M-S., Ting, C-T., Pei, D., Hsiao, C-F., Olivier, M., Pesich, R., Hebert, J., Chen, Y-D. I., Dzau, V. J., Curb, D., Olshen, R., Risch, N., Cox, D.R. and Botstein, D. 2001. High-throughput genotyping with single nucleotide polymorphisms. *Genome Research* 11:1262-1268.

Schnell, R.J., Brown, J.S., Olano, C.T., Power, E. J., Krol, C.A., Kuhn, D.N. and Motamayor, J.C. 2004. Frequency of off-type progeny among original ICS 1 x SCA 6 reciprocal families and parental clones used for disease resistance selection in Trinidad. *INGENIC Newsletter* 9:34-39.

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 cacao breeding programme using microsatellites. *INGENIC Newsletter* 10:2-8.

Tranah, G.J., Lescault, P.J., Hunter, D.J. and De Vivo, I. 2003. Multiple displacement amplification prior to single nucleotide polymorphism genotyping in epidemiologic studies. *Biotechnology Letters* 25:1031-1036.

Young, A. and Severson, D. 1994. Comparative analysis of steam distilled floral oils of cacao cultivars (*Theobroma cacao* L., Sterculiaceae) and attraction of flying insects: Implications for a Theobroma pollination syndrome. *J. Chemical Ecol.* 20:2687-2703.