PCR-BASED GENOTYPING OF SNP MARKERS IN *THEOBROMA CACAO*

Roslina M.S[1]., Nuraziawati M.Y[2], Aizat J.[3], Ahmad K.M.J.[3], Navies M.[4], Alias A[2], Rosmin K.[1], Neoh J.[5]

[1]Malaysian Cocoa Board, Commercial Zone 1, Norowot Road, Kota Kinabalu Industrial Estate, 88460 Kota Kinabalu, Sabah
[2]Malaysian Cocoa Board, P.O. Box 30, Sg. Dulang Road, 36307 Sg. Sumun, Perak
[3]Malaysian Cocoa Board, Lot 248, Block 14, Daerah Muara Tuang, Daerah Muara Tuang, Bahagian Samarahan, Locked Bag 3131, 93450, Kuching, Sarawak.
[4]Malaysian Cocoa Board, Centre of Cocoa Research and Development, Mile 10, Apas Road, P.O.Box 60237, 91012 Tawau, Sabah
[5]NEOSCIENCE SDN BHD (946222-U) P91, Block C, Level P, Kelana Square,Jalan SS 7/26, Kelana Jaya ,47301 Petaling Jaya
Selangor Darul Ehsan
Corresponding author: roslina@koko.gov.my

Malaysian Cocoa J. (2021) 13(1): 80-89

ABSTRACT - *The availability of reference sequence and sophisticated software does not always guarantee that the discovered SNP can be converted into a valid marker. Validation needs to be performed to ensure that the discovered SNP is in the Mendelian locus. The validation of a marker is the process of designing an assay based on the discovered polymorphism and then genotyping a panel of diverse germplasm and segregating population. Apart from that, the goal of the study was to develop a reliable, rapid, and inexpensive polymerase chain reaction (PCR)-based method to genotype for SNPs previously associated with desirable phenotypes in cocoa. Tetra-primer amplification refractory mutation system PCR (ARMS-PCR) is a simple and sufficient method for detecting different alleles in SNP locus. The allele-specific gene-tagged markers for the target genes are more effective than the genomic random markers as they will not show polymorphism in some recipient backgrounds and sometimes give up results as a false-positive marker. Allele-specific primer designs for the selected SNP from each trait. The primer is designed by using the software Primer1. DNA extracted from cocoa leaves was submitted to PCR amplification followed by agarose gel electrophoresis and determination of banding pattern. Tetra-primer ARMS-PCR was successfully optimized after changes in annealing temperature; annealing and extension times; concentration of MgCl2 and DNA; ratios of inner, outer, forward and reverse primer; and addition of adjuvants. There will be two types of product size expected if using an allele-specific primer set; one indicates that there is no SNP and another size will indicate there will be SNP/SNP is confirmed present in that specific sample. If there is any mismatch (SNP), a specific band size will be produced as compared to no mismatch (no nucleotide changes).*

Keyword: Tetra ARM, PCR, SNPs, Cocoa, Traits

INTRODUCTION

Modern plant breeding in this century makes use of phenotypic data from breeders, genetic variations from plants and development of molecular marker technology to improve plant varieties which contribute to improvement in different plants traits (Mochida *et. al.,* 2010). Most importantly, the technology developed will increase genotypic identities while reducing time and cost for farming. Conventional breeding not only takes a huge effort from the farmers, the uncertainty of whether their plants are able to maintain the good agronomic traits throughout the farming cycles, labours and high operational costs is very common in modern days farming (Pérez *et. al.,* 2012). Therefore, with the development of molecular markers as a tool for targeting on the presence of important qualitative and quantitative traits based on a few loci, especially when the cost of phenotyping greatly exceeds the cost of genotyping (VLK D *et. al.,* 2016). Different genotyping approaches and platforms will contribute to different kinds of results, hence these trade-offs must be

understood by plant breeders to make the best decisions for better crops in the future (Pandey S *et al.*, 2019).

Single-nucleotide polymorphisms (SNPs) are currently the most widely used molecular markers due to their ubiquitous distribution throughout a given genome, as well as their low cost compared to other marker technologies (Bali S *et al.*, 2018; Drenkard *et al.*, 2000). These markers are applicable across the full breadth of living organisms, providing universal interest in SNP technology development (Jiang G.L *et al.*, 2013; Batieno *et al.*, 2018).

Most method for Single nucleotide polymorphisms (SNPs) genotyping has been relying on expensive equipment such as a mass spectrophotometer, capillary electrophoresis, Pyrosequencer or real-time thermocycler (Broccanello *et al.*, 2018) With the use of this equipments, the assays develop will be becoming expensive too due to fluorescent labelling, di-deoxy terminators, a complicated and custom primers design or other reagents used in the assays (Mochida et.al, 2010; Pandey S *et.al*, 2019).

Recently, the single-nucleotide amplified polymorphisms (SNAP) assay, a simple, allele-specific method, was developed for SNP analysis in mapping populations of Arabidopsis (Drenkard *et al.*, 2000). SNAP markers are primers containing a single base mismatch with the 3' end of one allele (the specific allele) (Bundock *et.al*, 2005; Chiapparino *et.al.*, 2004). The tetra-primer amplification refractory mutation system PCR (T-ARMS-PCR) is a fast and economical means of assaying SNP's, requiring only PCR amplification and subsequent electrophoresis for the determination of genotypes.

<u>Tetra-primer ARMS-PCR method</u>

In this paper, it is shown that an allele-specific gene-tagged markers for target gene are more effective than the genomic random markers surrounding the target gene (from several kb to a few Mb distance) because some markers will not show polymorphism in some recipient background and sometimes a false positive allele can be selected by recombination between the target gene and genomic random marker (Liu *et al.*, 2012; Kim *et al.*, 2016). However, no information is known for any allele-specific markers for cocoa ever reported.

Customised design of PCR primers for allele-specific PCR needs to be improved by using data and analysis of multiplex PCR performed last year (Broccanello *et al.*, 2018). From this validation/ pilot run, the majority of the expected PCR product size is not tally with the reported phenotype traits. It might be due to the SNPs that present in specific traits does not correlate with gene expression results and thus for cocoa, gene expression profile for each trait might not be a good candidate to be selected as markers.

Single polymerase chain reaction (PCR) is used as a method to test every SNP-gene marker for each trait (with as minimum as; 5 SNPs x 2 primer sets x 5 traits).

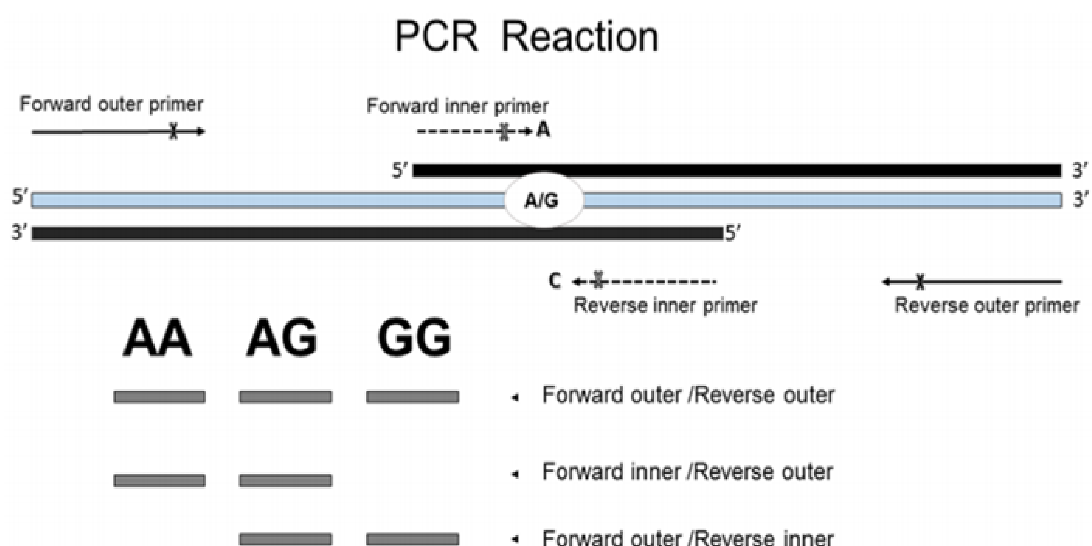


Figure 1: Diagram illustrating the position and orientation of PCR primers for the three primer allele-specific PCR relative to the matching template DNA.

MATERIALS AND METHODS

SNP Selection and Primer Design

Data and sequences from previous cocoa genomes and SNPs mining projects were used as a base for SNP selection and primer design for this study. Two groups of SNPs were categorized using two different approaches;

The first design for **VSD and CPB** was chosen based on;
SNPs in up/downstream regions (which might cause phenotypic variations)

100% occurrences in the control samples

Differentially expressed from the microarray data

While the second design is made for **BLACKPOD**, **HASIL** and **LEMAK** were chosen based on;
Identify SNPs that occur in all the samples from the same trait (Common SNPs)

Remove common SNPs that found in other traits (Unique SNPs for one trait)

Filter based on the sequencing read depth

Flanking primers were designed using the program Primer 1, http://primer1.soton.ac.uk/primer1.html which was also used to calculate the Tm for allele-specific primers which were designed to be as close to the Tm of the opposing flanking primer as possible.

Plant Material, PCR amplification and validation by Sanger Sequencing

Cocoa leaves from known phenotypes were obtained from Lembaga Koko Malaysia (LKM) Bagan Datuk, Perak and used for DNA extraction by following the manufacturer procedures. For verifying SNP and allele specific-PCR reaction, amplifications were carried out in a gradient thermocycler, which enabled the temperature range for allele-specific amplification to be determined. The annealing temperature ranged from 59-65 degrees. The cycling parameters were as follows: 95°C – two minutes, [94°C for 30 sec, 59°-70°C for 30 sec, 72°C for 30 sec] for 35 cycles, 72°C - five minutes, ambient hold.

Optimum quantities of 1:2 ratio of each primer were added to the reaction mix at a final concentration of 0.2uM. 'Hot start' Taq DNA polymerase was used for amplification. The PCR products then were subjected to Sanger sequencing to further confirm genes related to the trait tested.

RESULTS AND DISCUSSIONS

PCR Primers

Primers were designed using Primer1 software and were tested each by Sanger sequencing for validation. Figure 2 showing a summary of PCR designed from Primer 1 programme and their Sanger primer pair. More accurate visualization of outer and inner primers was shown in Figure 3.

For each trait tested, five possible SNP locations with their respective allele-specific PCR primers were listed in Figure 4.

Optimization of T-ARMS primer control sample

This step was done to ensure that the primer is working. Adjustment of the PCR reaction conditions such as annealing temperature, primer ratio and sample concentration was performed. Both data from PCR fragment and NGS data from Sanger sequencing must be analysed and matched.

Primer-BLAST from

SNP allele:

Primer pair 1

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CATGTTAGAATGGAGGAAGGGACAAG	26	61.36	46.15	4.00	0.00
Reverse primer	GCAAAGCATCAATTTGTAGGTTTCTCAC	28	62.10	39.29	7.00	2.00

Products on target templates

>LT594789.1 Theobroma cacao genome assembly, chromosome: II

```
product length = 165
Forward primer  1         CATGTTAGAATGGAGGAAGGGACAAG  26
Template        33423285  ..........................  33423260

Reverse primer  1         GCAAAGCATCAATTTGTAGGTTTCTCAC  28
Template        33423121  .........................A..  33423148
```

Sanger:

Primer pair 1

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CATGTTAGAATGGAGGAAGGGACAAG	26	61.36	46.15	4.00	0.00
Reverse primer	AATCAGCACCTTTTTTTCTTTTTTTCTGAAAT	30	60.35	26.67	7.00	5.00

Products on target templates

>LT594789.1 Theobroma cacao genome assembly, chromosome: II

```
product length = 220
Forward primer  1         CATGTTAGAATGGAGGAAGGGACAAG  26
Template        33423285  ..........................  33423260

Reverse primer  1         AATCAGCACCTTTTTTTCTTTTTTTCTGAAAT  30
Template        33423066  ..............................  33423095
```

Figure 2: Summary of PCR designed by Primer 1 and their Sanger primer pair

```
   3. g21690 pos: 2659

lcl|ctg_len_3827_22914:2359-2959
600 → 300 A→G
GTATCTCCTTGCCAAATTGAGCTGCTACCTAATCAGTATCAGCTGCCTTATTGGCATTTCTCTTAATCAGTTTTA
ATGTAGAATAATGAATAAGGGAACAGGAAGTATGGAAAGTATCAAGTCTTCCACAAGCCGCTATGCAGGCCATTA
AAGCAATCCCAGCATGTTAGAATGGAGGAAGGGACAAGATGATTTGGCCGCAGAATGCAACGCGTGTGAATGCAC
AATTTTAATATGAGAAGGTGGAAATAGAAATGGATTCCCAAATTGTCCATTCTGAAATAACAAGGAGCATGCAAA
TTAGAAACCTACAAATTGATGCTTTGCTCATGTATGTGTTTTGCTGAATCTTatttcagaaaaaagaaaaaaggt
gctgattttatgaatattttgatcctaattttatataaatatatatatatagatttttttaaaaattataaatat
agtttaaatataaaaaacaatatttttacaAAGATTGATATCAACATTTTATAAGTTAACTGAAATGATAGAAAC
TCAAAAAATTTCGACAAattatactataaaattaagatattaatcatgcgatttaataaaaaattagtactTCTA

Forward inner primer (A allele):                Melting temperature
274 ATTCTGAAATAACAAGGAGCATGCCAA 300                       66

Reverse inner primer (G allele):
327 GCAAAGCATCAATTTGTAGGTTTCTCAC 300                      65

Forward outer primer (5' - 3'):
163 CATGTTAGAATGGAGGAAGGGACAAG 188                        65

Reverse outer primer (5' - 3'):
382 AATCAGCACCTTTTTTTCTTTTTTTCTGAAAT 353                  65

Product size for A allele: 110
Product size for G allele: 165 (FO & RI)
Product size of two outer primers: 220 (for sanger)
```

Figure 3: Location of Forward and Reverse for Outer primers and location for Forward and Reverse primers for Inner primers

Trait: VSD

	SNP	Position	Forward Inner	Reverse Inner	Forward Outer	Reverse Outer
1	VSD_g29198	lcl|ctg_len_4444_26276:2157 (T>C)	TCTATTTTGCTGCATTCATGAATATGAACC	CTTGTTTCGGTGTGCTTTCCTTTGCGA	TGTGAATATGTTAGTGCTGAAGGCGAAGG	CCTTCTTTCGGTAAAGTTATAATGGGGCG
2	VSD_g248	lcl|ctg_len_9252303_163:1759980 (G>A)	CTTACGTGCAACGAAAGTTACTATTTCAAA	CAACTGAACTTTGAGTCAATTTCCGC	GGCTATTGTCTGATCAACCCATTGAATA	ACAACATTGCTAGCAGATAATGGCTCTC
3	VSD_g306	lcl|ctg_len_9252303_163:2329341 (C>T)	CTGCCGATTCCTATTTCATCTTTGGT	AATCTCGAGGCAAGAATTAGTATATCTG	TAACAGAGAGCCATTAAAGTAGAAAGCG	TAAAGGCCATGTGTTCATAACTTTTGTT
4	VSD_g2239	lcl|ctg_len_10940617_1267:5988856 (T>C)	ACGAGTAGATACAAGTTCATCCTAAGAT	AGTCTATGATTTTCTGAGTCCAAAAG	TTTAGAGATATATTTGGTTGGAAATGAT	TCATCAAACAAAAAAGTGATTAAGTTTA
5	VSD_g41941	lcl|ctg_len_2810_43862:2231 (C>T)	CATGTATACCAGAGTTGGAATCGGGC	AGGTTTTCTTTGGGCTTCCTGGTCTA	TGCTCTCTGGTGTGTGTTATGTACCCA	ACCCGACAATATGTGGTGGAATATTACG

Trait: CPB

	SNP	Position	Forward Inner	Reverse Inner	Forward Outer	Reverse Outer
1	CBP_g659	lcl|ctg_len_9252303_163:7655727 (G>A)	AAGAAGAAGAGAATTTCGAACTTTAGCCA	TTTTATCCCTCTTCAAGAGGGTTTTTTC	CCTTTGAACTCATACTCGCTTAGATCAA	TTTGAATGTATGTATAAATGACCTTGGCA
2	CBP_g515	lcl|ctg_len_9252303_163:5117784 (T>A)	TTGGGATAAGAAAAGTGTAATTTATAGGTT	ATATTGGATTATACCGTGAGAGACGT	GACCATTTTTAACTTACTGACTTGTTCC	ATTTAAGGGTACTTTTATGGTCAAAATG
3	CBP_g21690	lcl|ctg_len_3827_22914:2659 (A>G)	ATTCTGAAATAACAAGGAGCATGCCAA	GCAAAGCATCAATTTGTAGGTTTCTCAC	CATGTTAGAATGGAGGAAGGGACAAG	AATCAGCACCTTTTTTCTTTTTTCTGAAAT
4	CBP_g4230	lcl|ctg_len_4505_5195:368 (T>G)	ATCCTATGCAAATTTAAATTACAGCCG	GAAAAATAGCTCACTATAACTGAAGCTCCA	ATGTAATTAATCTGAAGGACACAAGGGA	TCTCTTTCCAATTACAAGTGTATCAGCA
5	CBP_g6449	lcl|ctg_len_22025072_7341:200267 (C>T)	TACGCATGAGTGAAATGATCTTTTGCT	TTCTTTCTTCCTTTTTGGTTCCAATTG	CATGAGGAATCATTTTCGAGAAGAAAGA	TTTGCTCATTGCTATGGTCTCAAGTTAA

Trait: BP

	SNP	Position	Forward Inner	Reverse Inner	Forward Outer	Reverse Outer
1	BP_SNP1	lcl|ctg_len_15929230_21830:5860447 (A>T)	TTTGACACAATGTAAAGGAGTAAGGTTGTT	ACACGTGAGCTTTGCCTTAACCAAGTT	CATAGGCACTAACACCAAAGAAAAAGGTC	TACACGAAAACACACACATTCACTGAAC
2	BP_SNP2	lcl|ctg_len_13028994_24846:12211586 (T>C)	CCCAACCCGAGGCGATATCTGTCATGCC	GGAAGAATGCGGGGAGAGGGGATGCA	ACCCCTCTCTGCCAGGTTGCTTCCTCCAA	CGGAAAGGATCGCTAGCCATGTCTTGGTCG
3	BP_SNP3	lcl|ctg_len_994145_42265:218234 (G>A)	GTGCTGTTTCGGTTGACTTGCGAGGTA	GTGGAGCCCTTCAAACCTTAGTATCGGC	TCTTCTCCACCGTAAAGATATCGCTGCC	TTTGAAGCCAAAACCATTAACAAAGCCG
4	BP_SNP4	lcl|ctg_len_13028994_24846:8645527 (C>T)	AGGGGTTTCATTATGTGACTCAACGGT	CTGTTATTGAGCCTTCCACATCGATTG	CAGATTCAGGTTCATGGTCATGAGTGTTA	ATCAAACACCTATCCAGCATGGCAATTA
5	BP_SNP5	lcl|ctg_len_2699747_834:2125458 (C>T)	CTAACATCAAGTGGCGGTTGTGGAATC	CACCATTGTACCTCTCTTGTTGCCCA	CGCAACTGGAGACAAGGCATAATCATTT	CCGAGGTATTGGAATGGGAATTTTTGAG

Trait: Hasil

	SNP	Position	Forward Inner	Reverse Inner	Forward Outer	Reverse Outer
1	HASIL_SNP1	lcl|ctg_len_16346715_50720:7766731 (T>A)	CACCTACCAATGGCTTTCGCAATGATAA	TTTGACAATTGAAGGGGGTATCGAATGA	GAGGTAAGTCTGCAGCCATCAGCTATTCAG	CCCTTCTCTCTCCAGTGGCTTTAGCTGT
2	HASIL_SNP2	lcl|ctg_len_12080_11245:9,236 (C>T)	GGATTGATGCTGTTCATGGACACAACCAC	GTGAGGGAAAATGGTTGCCTTGTAGCCA	AGCTCCAAAAGCTTCAGCAAAGACTTGGC	TAGCAAAGCAGATGAACTTGCCTGGTAGCA
3	HASIL_SNP3	lcl|ctg_len_12696376_5496:12,378,978 (A>G)	TCGAGGCTTTAACACATCAGATAGAGATTG	GATACGAAGGAGCCAGAGGATGGGAT	TTAGTGCTGTTGATGCAGAAGGGATATG	AAAAATTAGCCCTGAAGCGGAAAAATGT
4	HASIL_SNP4	lcl|ctg_len_15052430_36031:8,978,480 (C>T)	GAGGTCTGGGGTGGGGGTTTGGTGAT	CGGTTCCCTCTCTGTTTTACCCTTTAAATG	GTTTCACTGGGGATGGCTTGGACTTTTT	ACCGTTCTTTCCCCTCCCCTTTCTTTTC
5	HASIL_SNP5	lcl|ctg_len_15052430_36031:8978422 (T>A)	GCTTTAGTGTTTCTTTTTTCCCTTTGACA	CAACCAGAGATACAACACACAAAGCATTA	CTTGGACTTTTTGTTGCACTTACATCCT	CCTCCCCTTTCTTTTCTTTTCTTTTCTCT
6	HASIL_SNP6	lcl|ctg_len_15052430_36031:8,978,480 (C>T)	GAGGTCTGGGGTGGGGGTTTGGTGAT	CGGTTCCCTCTCTGTTTTACCCTTTAAATG	GTTTCACTGGGGATGGCTTGGACTTTTT	ACCGTTCTTTCCCCTCCCCTTTCTTTTC

Trait: Lemak

	SNP	Position	Forward Inner	Reverse Inner	Forward Outer	Reverse Outer
1	LEMAK_SNP1	lcl|ctg_len_11964671_28462:10599914 (G>A)	CTGAGGGACAGCCCTTCCTGTCACTTA	TCTATAAATTGCTTGAGTATTCTCACCCAC	GTTCTAGCAGAGCAAGAGGCAGGTCTA	CAGCGCATCACTTTTAGATCCATAGTTTC
2	LEMAK_SNP2	lcl|ctg_len_7557474_17708:6,037,395 (C>T)	CCCAAGAACACTAAGGCACAACCAAC	TATCAAATGGATCAAAAGGAGGCGAA	TGTATTGACATTGTTAATGCAATTTGGGG	GGAATGTCATTGTTGGTACTTGGATGGT
3	LEMAK_SNP3	lcl|ctg_len_1443295_10887:400,163 (G>A)	CTAAAGTGCCTACATCCACTCCCGTA	TTTAGTGGATTGAAAATTCCTTAGTGAACC	AGCTAAATGCAATAAATAGTAAGAAGGGGG	AATCCCGTTAAAAAGAAAGTGTAAGGGT
4	LEMAK_SNP4	lcl|ctg_len_15052430_36031:5922085 (G>A)	ATGCAGGGATTGTGGCTTTATGTTGA	TCAGATAATCAATTAAGTGGCAAGATTACC	TTGCAACTTATCATGTGAATAACAATGGAT	ACAATTTTGCTTTCAAGGAACAAGTTGA
5	LEMAK_SNP5	lcl|ctg_len_22025072_7341:19128876 (T>A)	GTGCAGTCCTCTATATCGATCTCCTACTT	TCCCAGGCTTAGCTTAGAAAACAAGT	CTGCTTACTACCTTAAGAAACATTGGGA	TATAGGCAAGTATTTGGTTTCATATCGC
6	LEMAK_SNP6	lcl|ctg_len_11964671_28462:1594917 (C>G)	GTTCCATGATGCAGCAAGAAAGTGACC	GCAACCTGATAGATATGGCCTAGCCTAGTC	TGCTTTTGTAAAGACAGGGAAATACGGG	GGATGATTACGTTTTCATTGTCAATCCTGA

Figure 4 : Listing of each traits with SNP location and possible allele-specific PCR primers

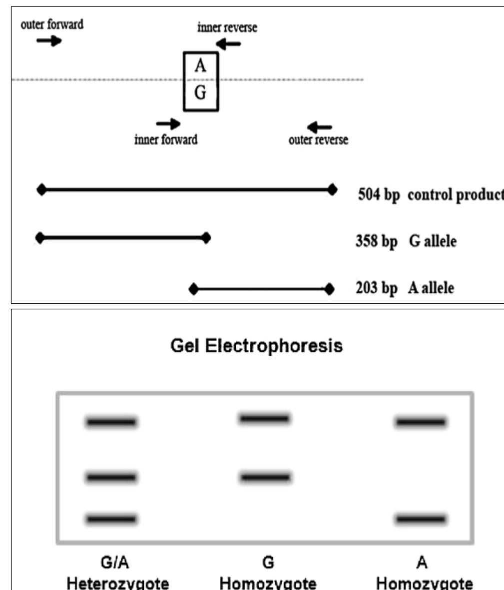


Figure 5: Showing data representation from gel electrophoresis of allele-specific PCR

Validation of T-ARMS PCR result

Trait VSD: VSD_g41941
□lcl|ctg_len_2810_43862:2231 (C>T) (up regulated)

Optimization steps of T-ARMS primer using a proper control sample

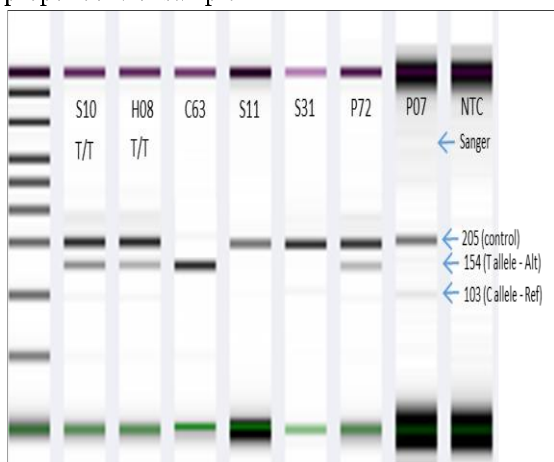


Figure 6: PCR results for VSD before confirmation with Sanger sequencing

Validation of T-ARMS PCR result:

Sequencing results of S10 and H08, both showed homozygous genotype, T/T. The data shown by Sanger sequencing was in concordance with the genotype shown by the T-ARMS PCR.

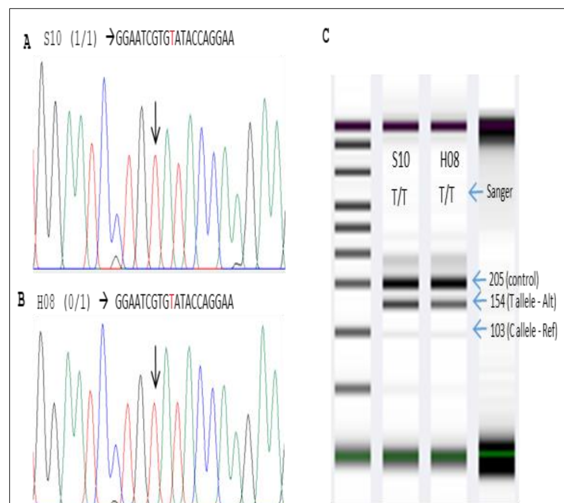
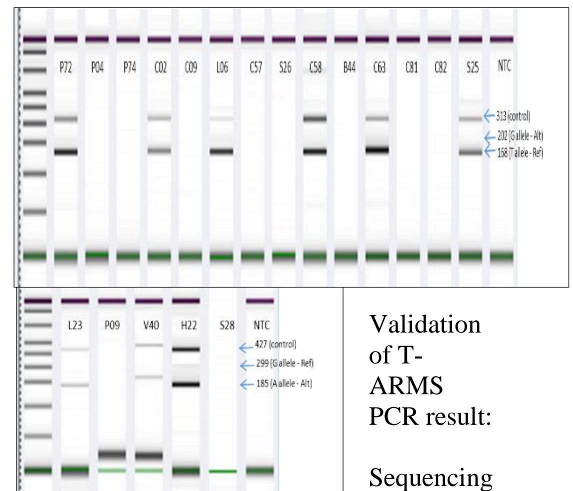


Figure 7: Validation pf PCR results by comparing with Sanger sequencing



Validation of T-ARMS PCR result:

Sequencing results of fifteen samples showed that twelve of the data shown by Sanger sequencing was in concordance with the genotype showed by the T-ARMS PCR

Figure 9: PCR results for CPB with different samples

Trait CPB: CPB_g659 lcl|ctg_len_9252303_163:7655727 (G>A) (up regulated)

Optimization steps of T-ARMS primer using a proper control sample

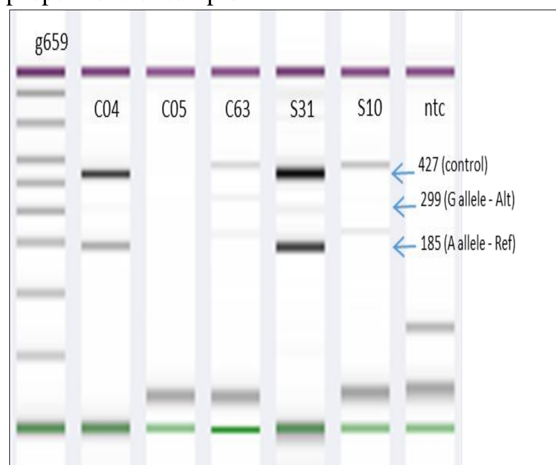


Figure 8: PCR results for CPB before confirmation with Sanger sequencing

Trait BP: BP_SNP2 -> lcl|ctg_len_13028994_24846:12211586 (T>C)

Optimization steps of T-ARMS primer using a proper control sample

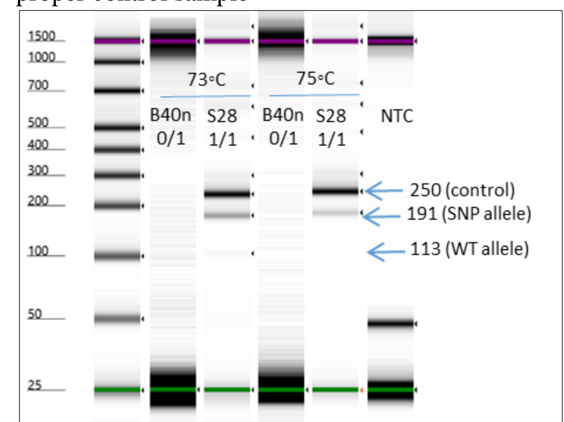


Figure 10: PCR results for BP before confirmation with Sanger sequencing

Validation of T-ARMS PCR result:
Sequencing result of eighteen samples showed that fifteen of the data showed by Sanger sequencing was in concordance with genotype showed by the T-ARMS PCR.

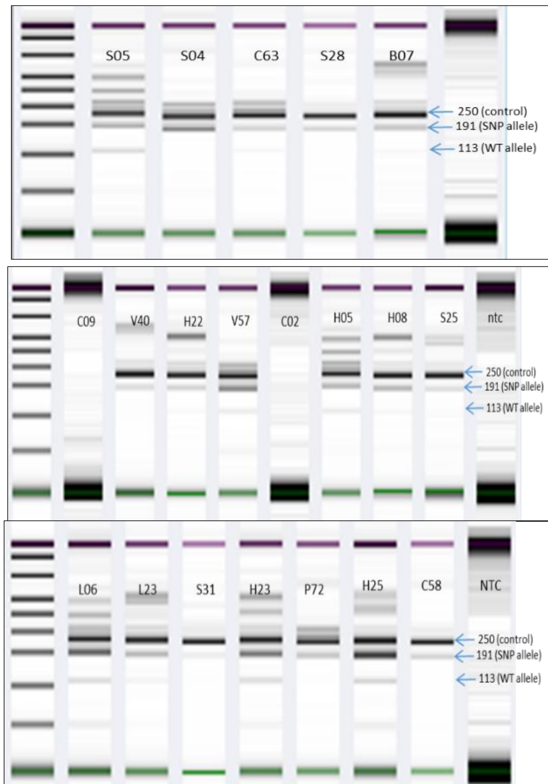


Figure 11: PCR results for BP with different samples

<u>Trait HASIL: HASIL_SNP1 -> lcl|ctg_len_16346715_50720:7766731 (T>A)</u>
Optimization steps of T-ARMS primer using a proper control sample

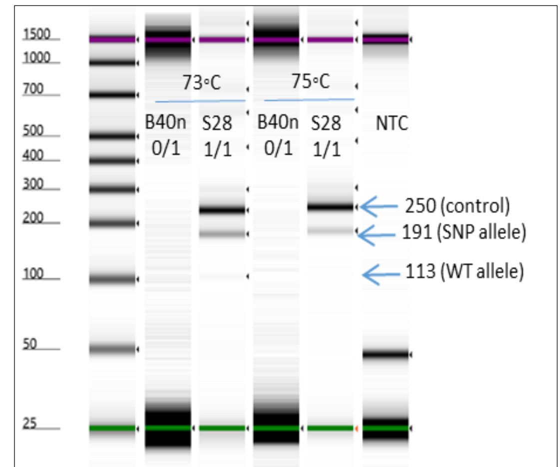


Figure 12: PCR results for HASIL before confirmation with Sanger sequencing

Validation of T-ARMS PCR result:
Sequencing result of eighteen samples showed that twelve of the data shown by Sanger sequencing was in concordance with genotype showed by the T-ARMS PCR.

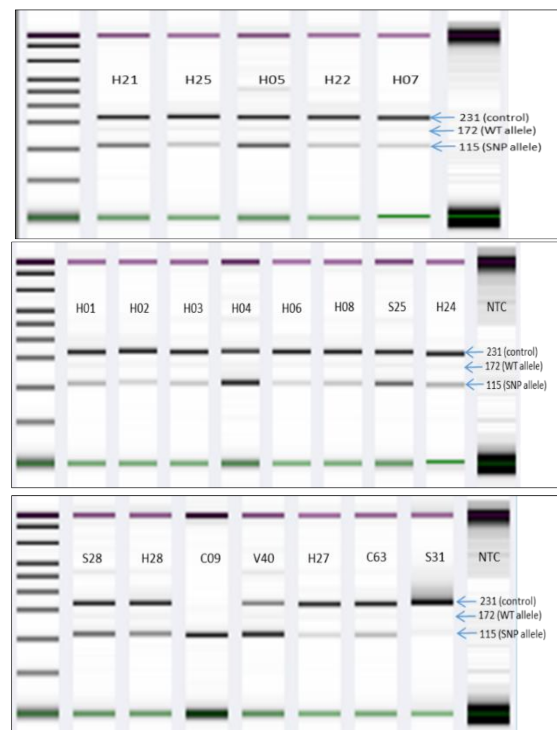


Figure 12: PCR results for HASIL with different samples

Trait LEMAK: LEMAK_SNP4 -> lcl|ctg_len_15052430_36031:5922085 (G>A)Optimization steps of T-ARMS primer using a proper control sample

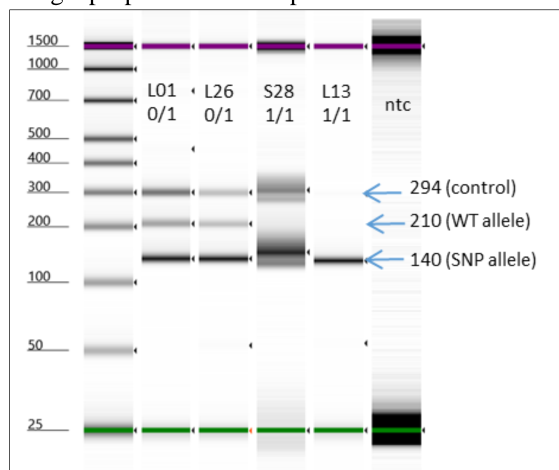


Figure 13: PCR results for LEMAK before confirmation with Sanger sequencing

Validation of T-ARMS PCR result:
Sequencing result of twenty samples showed that fifteen of the data shown by Sanger sequencing was in concordance with genotype showed by the T-ARMS PCR.

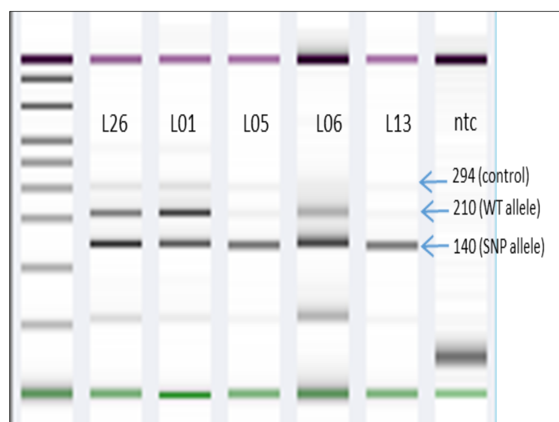


Figure 14: PCR results for LEMAK with different samples

CONCLUSIONS

A.SNP Selection
More than one SNP is required to represent each trait. Selection is based on the result of targeted sequencing performed from the last project. One problem is the list of SNPs from targeted sequencing is a lot. From one region, there is more than one SNP present.

If the region is selected, we must ensure that there is no more than one SNP present.

B.Control sample/Positive control
A proper positive control sample needs to be present in SNP marker development. It is advisable to utilize the control or reference sample which is used as a reference in WGS data mapping.

C.T-ARMS PCR validation
More than 20 samples are recommended for the primer validation to ensure that the selected SNP accurately represents the trait. More is needed for this stage.

REFERENCES

Bali S., Robinson B.R., Sathuvalli V., Bamberg J., Goyer A. (2018). *Single Nucleotide Polymorphism (SNP) markers associated with high folate content in wild potato species*. PLoS ONE 13 (2): 1-17.

Batieno T.B.J, Souleymane O., Tignegre J.B., Huynh B.L., Kusi F., Poda S.L., Close T.J., Roberts P., Danquah E., Ofori K. and Ouedraogo T.J. (2018). *Single nucleotide polymorphism (SNP)-based genetic diversity in a set of Burkina Faso cowpea germplasm*, African Journal of Agricultural Research; Vol. 13(19), pp. 978-987

Broccanello C., Chiodi C., Funk A., McGrath J. M., Panella L., & Stevanato P. (2018). *Comparison of three PCR-based assays for SNP genotyping in plants*, Plant Methods 14:28

Bundock P., Cross M., Shapter F. and Henry R, (2005), *Allele-specific PCR markers for single nucleotide polymorphisms in barley*, Molecular Plant Breeding CRC, Australia.

Chiapparino E., Lee D., and Donini P. (2004), *Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS–PCR*, Genome 47: 414–420 (2004)

Drenkard, E., Richter, B. G., Rozen, S., Stutius, L. M., Angell, N. A., Mindrinos, M., & Ausubel, F. M. (2000). *A Simple Procedure for the Analysis of Single Nucleotide Polymorphisms Facilitates Map-Based Cloning in Arabidopsis*, Plant Physiol. Vol. 124

Jiang G.L, (2013), *Molecular Markers and Marker-Assisted Breeding in Plants, Plant Breeding from Laboratories to Fields*: Chap 3

Kim S.G., Lee J.S., Shin S., Bae H.H., Kim J.T., Son BY, Baek SB, (2016), *Developing PCR-Based SNP Markers for Distinguishing Korean Waxy Corn F1 hybrids,* Plant Breed. Biotech., (August) 4(3):315-323

Liu R, Maia A.T, Russell R., Caldas C., Ponder B.A. and Ritchie M.E., (2012), *Allele-specific expression analysis methods for high-density SNP microarray data, Bioinformatics*; Vol. 28, pages 1102–1108

Mochida K., Shinozaki K., (2010), *Genomics and Bioinformatics Resources for Crop Improvement*, Plant Cell Physiol. 51(4): 497–523

Pandey S., Mishra S., Chandra K. (2019), *Tools and Resources for SNP Mining in Crop Plants*, International Journal of Current Microbiology and Applied Sciences, Vol. 8 Num 01

Pérez-de-Castro A.M., Vilanova S., Cañizares J., Pascual L, Blanca J.M, Díez M.J, Prohens J, and Picó B., (2012), *Application of Genomic Tools in Plant Breeding*, Current Genomics: 13, 179-195

Vlk D., Řepkova J., (2016) *Application of Next-Generation Sequencing in Plant Breeding*, Czech J. Genet. Plant Breed