BIODIVERSITAS
Volume 18, Number 2, April 2017
Pages: 465-475

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d180204

Development of SNAP markers based on nucleotide variability of WRKY genes in coconut and their validation using multiplex PCR

ANNEKE PESIK[1,2], DARDA EFENDI[1], HENGKY NOVARIANTO[3], DINY DINARTI[1], SUDARSONO SUDARSONO[1,♥]

[1]PMB Lab, Department of Agronomy and Horticulture, Faculty of Agriculture, Institut Pertanian Bogor. Jl. Meranti, Darmaga Campus, Bogor 16680. West Java, Indonesia. Tel./Fax.: +62-2518629353, ♥email: s_sudarsono@ymail.com

[2]Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Pattimura, Jl. Ir. M. Putuhena, Kampus Poka, Ambon 97233, Maluku, Indonesia

[3]Indonesian Palm Research Institute, P.O. Box 1004, Manado 95001, North Sulawesi, Indonesia

Manuscript received: 30 August 2016. Revision accepted: 9 February 2017.

Abstract. *Pesik A, Efendi D, Novarianto H, Dinarti D, Sudarsono S. 2017. Development of SNAP markers based on nucleotide variability of WRKY genes in coconut and their validation using multiplex PCR. Biodiversitas 18: 465-475.* Development of molecular markers benefits coconut breeding program. The availability of DNA sequences for coconuts opens the new possibility of developing molecular markers such as single nucleotide amplified polymorphism (SNAP). The study aimed to evaluate nucleotide sequence diversities of WRKY gene in the GenBank database, design gene specific primers for generating WRKY gene specific SNAP markers, optimize multiplex PCR technique and validate SNAP marker effectiveness for evaluating Kopyor coconut germplasm. Based on 35 sequences data of coconut WRKY genes that are available in the GenBank database, we have identified eight informative SNPs and used them to generate SNP-specific primers. We have designed sixteen primer pairs and validated them using singleplex PCR. Subsequently, we optimized the Ta and primer concentrations either for duplex or triplex PCR. Duplex PCR using two sets of primer pairs was more reliable for genotyping Kopyor coconut germplasm than triplex PCR. We have successfully demonstrated the duplex PCR for genotyping Banten Tall, Jember Tall, Kalianda Tall, Pati Dwarf and Sumenep Tall Kopyor coconuts. Therefore, one may use the developed SNAP markers as a simple alternative of co-dominant markers for genetic analysis of coconuts. One may use the developed SNAP markers to investigate the possible association between markers and Kopyor endosperm and other relevant phenotypes in coconuts.

Keywords: Kopyor coconut, multiplex PCR, SNP

INTRODUCTION

Advance in plant molecular biology and biotechnology has opened up many opportunities in plant science research. Researchers have isolated and characterized a large number of different genes from many plant species. The NCBI GenBank DNA database (http://www.ncbi.nlm.nih.gov/) curates most of those DNA sequence data. Those gene sequence data are freely accessible, and plant scientists can use the data to support and accelerate their plant research activities. Many research areas have benefitted from the accumulated DNA sequences in the NCBI GenBank DNA database, especially in the development of molecular markers. The existence of more information about the gene sequences in the NCBI GenBank DNA database opens up the new possibilities for developing and evaluating markers based on the single nucleotide polymorphisms (SNPs).

Single nucleotide polymorphism (SNP) is a single nucleotide DNA variation at specific location throughout the plant genome. The SNPs has been used as plant genetic markers (Simko et al. 2012) in sugar beet and apple genotyping studies (Bakooie et al. 2015; Chagné et al. 2008). Tea and tomato variety identification also used SNP markers (Fang et al. 2014; Olmos et al. 2015; Wang et al. 2015). Similarly, cotton and other plants high-resolution genetic linkage map and gene tagging have used SNP in their evaluation (Cho et al. 2011; Mammadov et al. 2012; Li et al. 2014). Massive availabilities of SNP markers open the possibility of tagging genes controlling important agronomic characters and supporting breeding programs (Ren et al. 2013). The frequency of SNP occurrences in the plant genome is as many as one SNP for every 31 base pair (bp) up to every 124 bp of plant genome (Eulgem et al. 2000; Ching et al. 2002; Zhang and Wang 2005; Fang et al. 2014).

The simplest marker that can be developed utilizing SNP is the single nucleotide amplified polymorphism (SNAP) marker. The SNAP markers can be developed and applied to analyze genetic diversity, kinship and pollen dispersal of target plants (Singh et al. 2013, Larekeng et al. 2015a,b). Kopyor coconut is an exotic coconut mutant originated from Indonesia. The endosperm of Kopyor coconut is detached from its shell and formed endosperm crumbs in the coconut shell, and its development benefits Indonesian coconut farmers (Maskromo et al. 2015). The Kopyor endosperm in coconut is probably caused by a mutation in a single regulatory gene causing pleiotropic effects to some coconut plant developments. Therefore, identification and characterization of the mutant gene

controlling Kopyor endosperm in coconut would be beneficial (Sukendah et al. 2009; Sudarsono et al. 2015). However, there has been no conclusive evidence that the isolated genes are ones responsible for Kopyor characters in coconut.

The WRKY is a group of genes encoding superfamily of transcription factor proteins having one or two conserved domain consisting of WRKYGQK residues in the primary amino acid sequences (Eulgem et al. 2000). The WRKY plays important roles in some plant growth and developments, responses to biotic and abiotic stresses, senescence, and the zygotic embryo formation (Eulgem et al. 2000; Zhang and Wang 2005). There have been some endosperm-specific WRKY gene fragments from coconuts deposited in the GenBank database which can be used to develop SNAP markers. Since two PCR (polymerase chain reaction) reactions are needed to genotype a single locus for each SNP site, multiplex PCR is proposed to generate more efficient procedures for and speed up the genotyping (Liu and Wu 2012). Multiplex PCR techniques have previously been developed to amplify either SSR or SNP markers (Hayden et al. 2008; Kim et al. 2012; Liu and Wu 2012; Wen and Zang 2012; Mousset et al. 2015). Researchers also used multiplex PCR for plant genotyping, linkage and parental analysis of many plants (Masi et al. 2003; Chaerani et al. 2009; Migliaro et al. 2012; Larekeng et al. 2015a,b).

However, there has been a discussion about some limitations in plant multiplex PCR technique (Hayden et al. 2008; Sheng et al. 2003). The objectives of the study were to evaluate coconut WRKY gene accessions available in the NCBI GenBank DNA database and their nucleotide sequence diversities, design gene specific primers for SNAP markers based on identified SNP in the WRKY genes, optimize multiplex PCR technique and validate SNAP marker effectiveness for evaluating Kopyor coconut germplasm. Availability of such multiplex PCR technique for generating WRKY gene specific SNAP markers may be used to investigate the possible association between variability of the WRKY gene and the Kopyor endosperm in coconut. If such associations exist, the developed SNAP markers may facilitate and accelerate the selection of coconut-carrying the Kopyor phenotypes.

MATERIALS AND METHODS

Coconut WRKY genes and their nucleotide sequence diversity

Searches for accessions of the WRKY genes from NCBI GenBank DNA database were conducted using the appropriate keywords (WRKY, coconut, and *Cocos nucifera*)and positive hits identified. The identified WRKY sequences were used to search for other highly similar genes from coconut using basic local alignment search tools (BLAST, http://blast.ncbi.nlm.nih.gov/). Each sequence in a group of similar WRKY genes was arranged in fasta format and imported to sequence analysis software. Sequences of each group of WRKY genes were evaluated offline using Genious® Basic 6.1.8. (Biomatters). Multiple sequence alignment (MSA) for accessions of each WRKY group was conducted to identify locations of the SNP sites. We used the identified non-synonymous SNPs in primer design steps.

SNAP primer design based on identified SNP in WRKY gene

The gene specific SNAP primers were designed based on previously identified SNP sites. The identified SNPs having bi-allelic alternative alleles were selected, and fragment sequences adjusted to the required submission format for online SNAP primer design using WebSnapper (http://ausubellab.mgh.harvard.edu). Subsequently, all identified SNPs in the evaluated WRKY genes were individually used to design gene-specific SNAP primers. The outputs of every primer set have been assessed to select at least three primers (two forward primers [reference and alternate forward primers] and one reverse primer) to be used for genotyping a single SNP locus. The genotyping of a single locus was carried out using two PCR amplification reactions, the first PCR reaction using a pair of reference (Ref) and reverse (Rev) primers and the second using a pair of alternate (Alt) and Rev primers, respectively.

.

Validation of SNAP primers in a singleplex PCR

The validation of the effectiveness of the designed gene-specific SNAP primers was done using a singleplex PCR. The PCR amplification was done using genomic DNA of Pati Dwarf Kopyor coconut as DNA template and KAPA2G PCR kit (Kapa Biosystems Inc., USA). The mixtures for each PCR reaction consisted of 6.0 µL 5x PCR buffer; 0.6 µL, 25 mM $MgCl_2$, 0.6 µL, 10 mM dNTPs, 0.75 µL, 10 µM of each of forward and reverse primer; 0.1 µL, 5 unit/µL Taq DNA Polymerase and sterile ultrapure water (ddH_2O)-added to the final volume to 25 µL. For each primer pairs, the optimum annealing temperature was optimized using thermocycling gradient PCR. The evaluated primer annealing temperature (Ta) in the gradient PCR ranged from 48.0 to 60.0°C and annealing temperature yielding the best amplicons was selected. PCR amplification was done using Biorad T100™ DNA Thermal Cycler (Biorad). The amplification steps were as follow: one cycle pre-denaturation stage at 95°C for 3 min; 35 cycles consisting of denaturation at 95°C for 15 s, primer annealing at 49.5°C-56.4°C (according to the selected Ta based on the gradient PCR) for 15 s, and primer extension at 72°C for 1 s; and one cycle of final extension at 72°C for 10 minutes.

Optimization of SNAP primers in a multiplex PCR

Duplex and triplex PCR were proposed to minimize the number PCR reactions in genotyping using SNAP markers. This activity tested the ability of two (duplex PCR), and three sets of primer pairs in a single tube of PCR reaction (triplex PCR) to produce different sizes of the target amplicons. For each combination of primers, the optimum annealing temperature was optimized using gradient PCR. To evaluate the effect of tested primer annealing temperature (Ta) in the gradient PCR, the Ta ranged from

48.0 to 59.1°C were tested and one yielding clear amplified products for all target amplicons was selected. We tested two concentrations (0.375 µL, 10 µM and 0.75 µL, 10 µM) of each set of primer pairs to evaluate the effect of primer concentrations in PCR reaction mix on amplicon quality and quantity. The PCR amplification was done using genomic DNA of either Pati Dwarf or Kalianda Tall Kopyor coconut and KAPA2G PCR kit (Kapa Biosystems Inc., USA). The mixtures for each PCR reaction consisted of 6.0 µL 5x PCR buffer; 0.6 µL, 25 mM $MgCl_2$, 0.6 µL, 10 mM dNTPs, the sets of forward and reverse primer pairs, 0.1 µL, 5 unit/µL Taq DNA Polymerase and sterile ultrapure water (ddH_2O)-added to the final volume of 25 µL. PCR amplification was done using Biorad T100™ DNA Thermal Cycler (Biorad), and the amplification steps were as described in the singleplex evaluation.

Validation of SNAP primers in a duplex PCR

The optimized multiplex PCR conditions evaluated in previous activities were selected and used to validate their ability to generate amplicons using diverse coconut genomic DNA. We used an array of genomic DNAs from 15 accessions each of Banten Tall (B), Jember Tall (J), Kalianda Tall (K), Pati Dwarf (P) and Sumenep Tall (S) Kopyor coconuts as a template to validate the effectiveness of the selected PCR conditions. The PCR amplification was done using optimized measures as described in the previous activities and KAPA2G PCR kit (Kapa Biosystems Inc., USA). The mixtures for each PCR reaction and the optimum annealing temperature were as described previously. PCR amplification was done using Biorad T100™ DNA Thermal Cycler (Biorad), and the amplification steps were as outlined in the singleplex evaluation.

Fractionation and visualization of amplicons

Fractionation of amplicons of all PCR reactions according to their fragment sizes was using 1% agarose (Vivantis Inc., USA) gel electrophoresis in 1x sodium borate electrophoresis buffer. The electrophoresis was done using a constant current of 80 volts for 30 min. The amplicon visualization was done using 33% (v/v) GelRed™ (Biotium Inc., USA) DNA staining solution and viewed under UV lamp (Vilber Lourmat Super Bright TFX-20 Sigma-Aldrich, MX). The electropherograms were documented using a digital camera. The amplicon sizes were estimated using standard 100 bp DNA ladder (Vivantis Inc., USA).

Construction of phylogenetic tree

We evaluated a total of 75 individuals of Banten Tall (B), Jember Tall (J), Kalianda Tall (K), Pati Dwarf (P) and Sumenep Tall (S) Kopyor coconuts. We calculated dissimilarity matrix based on genotypic data for two ploidy levels and used simple matching dissimilarity index. We set at 10,000 iterations for the number of bootstrap analysis. We did the tree construction by weighted Neighbour Joining approach and used the calculated dissimilarity matrix. Calculation of the dissimilarity matrix, bootstrap, and tree construction for the Kopyor coconut accessions was done using Dissimilarity Analysis and Representation for WINDOWS (DARwin) software version 6.05 (Perrier and Jacquemoud-Collet 2010).

RESULTS AND DISCUSSION

Coconut WRKY gene and their nucleotide sequence diversities

Searching and BLAST analysis for WRKY gene from coconut in the NCBI GenBank DNA database having at least 200 base pair (bp) in fragment size resulted in a total of at least 35 WRKY gene accessions (Table 1). Based on their sequence identity, we grouped the accessions into WRKY2, WRKY6, WRKY7, WRKY19, and WRKY21. The fragment sizes of WRKY accessions ranged from 236-972 bp. Subsequently, we conducted multiple sequence alignment (MSA) for all accessions in each WRKY group. Based on MSA results, two bi-allelic SNP sites were identified from each of the WRKY2, WRKY6 and WRKY21 groups, while one SNP was from each of the WRKY7 and WRKY19 groups, respectively (Table 1). Table 2 listed summaries of the identified SNP sites for all WRKY groups while Figure 1 presented an example of identified bi-allelic SNP site after MSA of 7 accessions belonging to the WRKY6 group.

SNAP primer design based on selected SNPs in WRKY gene

The outputs of each primer design in the WebSNAPPER entry are a list of alternative SNAP primers for each SNP site. For each of identified SNP site, we only need one set of SNAP primers, consisting of at least three primers per set, two forwards (Ref and Alt) and one common reverse (Rev) primers. We select those primers from the output of the WebSNAPPER. The selected Ref and Alt of the SNAP forward primers were chosen from a list of WebSNAPPER output based on the mismatch position of the last four nucleotides at their 3' end to the reference of target nucleotide sequences (Sutanto et al. 2013). The further away from the mismatch position from 3'end of the primer the more desirable (Bru et al. 2008). Table 3 present list of the selected SNAP primer sets based on the identified bi-allelic SNPs of coconut WRKY genes.

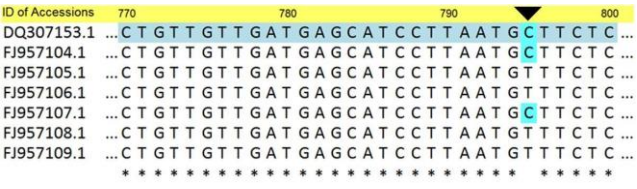

Figure 1. An example of multiple sequence alignment results over seven accessions of WRKY6 DNA fragments and the identified bi-allelic single nucleotide polymorphism (SNP) site #3 at the base position of 795. The bi-allelic nucleotide variants are either T (57.1%) or C (42.9%).

Table 1. List of existed accessions of coconut (*Cocos nucifera*) WRKY gene fragments in NCBI GenBank DNA database and numbers of identified informative non-synonymous single nucleotide polymorphisms (SNPs) after multiple sequence alignment (MSA) over accessions for each WRKY group

Gene group	Coconut WRKY DNA in NCBI GenBank database			Number of informative SNPs [a)]
	Total sequences	Fragment sizes	NCBI accessions no.	
WRKY2	8	236-959	FJ956959.1, FJ956960.1, FJ956961.1, FJ956962.1, FJ956963.1, FJ956964.1, DQ307150.1, DQ861409.1	2
WRKY6	7	724-927	FJ957104.1, FJ957105.1, FJ957106.1, FJ957107.1, FJ957108.1, FJ957109.1, DQ307153.1	2
WRKY7	6	607-639	FJ957178.1, FJ957179.1, FJ957180.1, FJ957181.1, FJ957182.1, FJ957183.1	1
WRKY19	7	354-525	FJ957389.1, FJ957390.1, FJ957391.1, FJ957392.1, FJ957393.1, FJ957394.1, DQ307157.1	1
WRKY21	7	764-972	FJ957030.1, FJ957031.1, FJ957032.1, FJ957033.1, FJ957034.1, FJ957035.1, DQ307158.1	2
Total	35	-	-	8

Note [a)] The informative SNP site was the SNP having bi-allelic alternative nucleotides

Table 2. List of single nucleotide polymorphisms (SNPs) identified after multiple sequence alignment (MSA) over accessions for each group of WRKY genes available in NCBI GenBank DNA database and their characteristics

WRKY gene fragment identity	SNP locus number and positions [a)]		Alternatives bi-allelic SNP alleles [b)]	SNP allele ratio (Percentage) [c)]
WRKY2	#1	697	A/G	6/8 (75.0)
	#2	721	G/C	7/8 (87.5)
WRKY6	#1	156	C/T	6/7 (85.7)
	#3	795	T/C	4/7 (57.1)
WRKY7	#1	561	T/A	5/6 (83.3)
WRKY19	#1	74	G/A	6/7 (85.7)
WRKY21	#1	447	G/C	5/7 (71.4)
	#3	553	A/G	6/7 (85.7)

Note: [a)] indicate the SNP locus numbers and the nucleotide position of the SNPs in the evaluated WRKY DNA fragments; [b)] indicate the nucleotide identities of major and minor alleles of the SNPs; [c)] indicate the ratios of the major alleles by total samples and the percentage of major allele.

Table 3. List of selected single nucleotide amplified polymorphism (SNAP) primer sets developed based on the identified bi-allelic single nucleotide polymorphism (SNP) of coconut WRKY genes

Primer ID (SNP Locus) [a)]	Primer sequences [b)]	Primer length	Ta [c)] (°C)	Product size (bp) [d)]
WRKY2#1 Ref	TCAAAAGTCGTACGTGAACCACGG*G*	25	52.8	319
WRKY2#1 Alt	CATCAAAAGTCGTACGTGAACCAGGG*A*	27	52.8	321
WRKY2#1 Rev	GCCCTCTCAACGTGCTTCCGG	21	52.8	-
WRKY2#2 Ref	TGATGATGGTTACCGCTGGGG*G*	22	52.8	260
WRKY2#2 Alt	TGATGATGGTTACCGCTGGGG*C*	22	52.8	260
WRKY2#2 Rev	GCCCTCTCAACGTGCTTCCGG	21	52.8	-
WRKY6#1 Ref	AAAATATAATTCAGGGAATTAAGCA*G*	26	49.5	107
WRKY6#1 Alt	AAAATATAATTCAGGGAATTAAGCA*A*	26	49.5	107
WRKY6#1 Rev	AAAGTAGTGAAAGGGAATCCAAA	23	49.5	-
WRKY6#3 Ref	ATAAATATCACATATCCCTGAGCA*A*	25	52.8	352
WRKY6#3 Alt	TATAAATATCACATATCCCTGAGCA*G*	26	52.8	353
WRKY6#3 Rev	CTACAGGTATTGTTAAATGTGCCA	24	52.8	-
WRKY7#1 Ref	CATCCGATCATTCGTGC*G*	18	56.4	354
WRKY7#1 Alt	CATCCGATCATTCGTGC*C*	18	56.4	354
WRKY7#1 Rev	ATCGATATCACTTGTGGTCTGAA	23	56.4	-
WRKY19#1 Ref	CGTCTTCTGCAAACTAAGCTC*A*	22	52.8	211
WRKY19#1 Alt	GTCTTCTGCAAACTAAGCTC*G*	21	52.8	210
WRKY19#1 Rev	ATGATATATATACAAGACTACGCCGAT	27	52.8	-
WRKY21#1 Ref	ATTGCCTATTAAAAATAACAATTTAC*G*	27	52.8	272
WRKY21#1 Alt	ATTGCCTATTAAAAATAACAATTTAC*C*	27	52.8	272
WRKY21#1 Rev	TACCATATATTAATGCGGATGAGAT	25	52.8	-
WRKY21#3 Ref	AGCATCTTGATCAATTACGATAC*C*	24	54.6	237
WRKY21#3 Alt	AAGCATCTTGATCAATTACGATAC*A*	25	54.6	238
WRKY21#3 Rev	TACCATATATTAATGCGGATGAGAT	22	54.6	-

Note [a)] Primer ID labeled with Ref, Alt, or Rev indicates the reference, alternate or reverses primers, respectively. [b)] Underlined and italic letters at the 3' end of the primers indicate the nucleotide variants of the SNP site. [c)] Ta, the optimized primer annealing temperature based on singleplex gradient PCR. Alternative primer pairs for the PCR reactions are either (Ref+Rev) or (Alt+Rev)

Validation of SNAP primers in a singleplex PCR

The primer sets were tested in singleplex PCR system using genomic DNA of Pati Dwarf Kopyor coconut t to validate the effectiveness of the designed SNAP primers. Fig 2 presents validation results of the 16 primer pairs for the genotypes of eight SNP loci of the coconut WRKY genes. The results showed all primer pairs were able to produce amplicons of the expected sizes. The PCR amplification using genomic DNA of Pati Dwarf Kopyor coconut and primer pairs of the WRKY2#1 SNP locus produced 319 bp amplicon for Ref and 321 bp for Alt; WRKY2#2-260 bp; WRKY6#1-107 bp; WRKY6#3-353 and 352 bp; WRKY7#1-354 bp; WRKY19#1-211 and 210 bp; WRKY21#1-272 bp; and WRKY21#3-237 and 238 bp, respectively (Figure 2). Some of the amplicons are very close in their sizes such as: between those of WRKY2#2 and WRKY21#1 (260 bp and 272 bp) and between WRKY6#3 and WRKY7#1 (352 and 354 bp); therefore, it may be difficult to fractionate the amplicons in a 1% agarose gel electrophoresis.

We used those singleplex results as the basis for selecting appropriate sets of primer pairs to be used in a multiplex PCR. It was suggested that singleplex PCR be done first before attempting to set up a multiplex PCR (Sheng et al. 2003). Singleplex PCR before multiplexing could also be used to estimate optimal annealing temperatures to produce the correct amplicon sizes (Sint et al. 2012). Results of singleplex PCR could be used as references to indicate which primer pairs produced specific and which produced non-specific amplicons.

These singleplex PCR results may also be used to select the primer combinations for multiplex PCR based on their amplicon sizes. Based on singleplex PCR results, duplex PCR might not be possible using the combination of WRKY19#1 and WRKY21#3 primer sets; WRKY21#3 and WRKY2#2; WRKY2#2 and WRKY21#1; WRKY21#1 and WRKY2#1, or WRKY2#1 and WRKY6#3 or WRKY7#1. Other than those primer sets, we used the primer combinations in the duplex PCR since size differences of the amplicons were long enough for fragment separation in 1% agarose gel electrophoresis. However, the choices of primer combinations for triplex PCR was limited, and it could be WRKY6#1, plus either WRKY19#1, WRKY21#3, or WRKY2#2, and either WRKY6#3 or WRKY7#1.

The genotype of each SNP locus is estimated using two singleplex PCR reactions. The first PCR reaction using a pair of the Ref-Rev primers was to determine the reference allele while the second PCR using a pair of the Alt-Rev primer pairs was to determine the alternate variant allele. If the first PCR reaction yielded an amplicon of the expected size while the second PCR was negative, the genotype for this particular SNP locus was a homozygous Ref. On the other hand, if the first PCR reaction yielded no amplicon and the second PCR yielded the expected size of amplicon, the genotype for this particular SNP locus was a homozygous Alt. Meanwhile, if both the first PCR and the second PCR reactions positively yielded the expected amplicon sizes, the genotype for this particular SNP locus is a heterozygous Ref-Alt. The Figure 3 present illustration for the genotyping scores and the predicted genotypes for a single SNP locus using a SNAP marker. Therefore, genotyping with SNAP markers using singleplex PCR cost more and probably requires more resources.

Optimization of SNAP primers in a multiplex PCR

Multiplex PCR is PCR reaction using two or more sets of primer pairs to simultaneously produce two or more amplicons in a single PCR reaction. Therefore, multiplex PCR requires optimization of various PCR conditions to obtain desirable amplicons. Multiplex PCR requires carefully designed primers to prevent potential annealing among primers or primer dimers. It also requires a balance annealing temperature for all the primer combinations so that desirable amplicons are all amplified. Moreover, it also requires balance primer to primer and primer to template ratios. All of those factors need optimization for multiplex PCR. In this evaluation, duplex PCR (two sets of primer pairs in a single PCR reaction) was tested using the developed gene specific SNAP primers.

The annealing temperature and concentration of primers strongly influence the efficiency of amplicon amplification, especially in a multiplex PCR. Figure 4 presented an example of the annealing temperature (Ta) optimization in a gradient duplex PCR with primer annealing temperature (Ta) from Ta=48.0 to Ta=59.1°C. If the Ta was too low (Ta=48.0 or 48.8°C), it results in non-specific binding of the primer and the template. Primer annealing in low-temperature may result in amplification of non-target amplicons. On the other hand, if it was too high (Ta=59.1°C, Figure 4.A; Ta>50.3°C, Figure 4.B), it missed one of the target amplicons from the PCR products. In the duplex PCR, either WRKY2#1 or WRKY2#2 primer pairs always yielded amplicon regardless of the Ta while the WRKY6#1 primer pairs in combination with WRKY2#2 yield clear amplicon at Ta=50.3°C and in combination with WRKY2#1°C yielded a clear amplicon at Ta=48.8°C. In this evaluation, we used Ta=50.3 for the duplex PCR using two pairs of WRKY2#2 and WRKY6#1 primers and Ta=48.8-50.3°C for those using WRKY2#1 and WRKY6#1. Further optimization showed that the optimum Ta for WRKY6#1 in duplex PCR was 50.3°C (The data were not presented). These results indicated optimization of annealing temperature for each primer combination in multiplex PCR were necessary.

Primer to primer and primer to template ratios are other aspects of PCR conditions that need optimization to increase the rate of success in multiplex PCR (Sheng et al. 2003). In this evaluation, only duplex PCR (two sets of primer pairs in a single PCR reaction) was tested using the developed gene specific SNAP primers. We evaluated two levels of primer concentrations, either 0.375 µL of 10 µM or 0.75 µL of 10 µM for each primer in a duplex PCR reaction and used the same amount of DNA template of Kalianda Tall Kopyor coconut genome in all PCR reactions. Figure 5 presented results of the primer concentration optimization in a duplex PCR.

More primer concentration in the PCR reaction mix produced more amplicons than less; however, less primer was also able to generate comparable PCR product. Using primers at 0.375 µL of 10 µM in duplex PCR produced sharper amplicons in agarose gel electrophoresis than that of 0.75 µL of 10 µM (Figure 5). Therefore, it was necessary to optimize the concentration of added primers into the duplex PCR mix. It was also important to adjust primer concentration according to the sizes of amplicons (Liu and Wu 2012) since there was competition among primers in multiplex PCR (Sint et al. 2012). In multiplex PCR, some primers and their concentrations affect the efficiency of PCR reactions (Xu et al. 2012).

The effectiveness of duplex and triplex PCR was tested using combination primers producing two and three different sizes of amplicons. For duplex PCR, a mixture of WRKY6#1 (Ref or Alt) and WRKY2#2 (Ref or Alt) primer pairs always yielded the 107 bp and 260 bp amplicons (Figure 6.A). In duplex PCR using a combination of WRKY6#1 and WRKY2#1, WRKY6#1 (Ref or Alt) primer pairs always yield a clear 107 bp amplicon. On the other hand, the WRKY2#1 Ref primer pairs yielded 319 bp amplicon while those of WRKY2#1 Alt failed to produce the desirable amplicon (Figure 6).

For triplex PCR, amplifications using a combination of either WRKY6#1, WRKY2#2 and WRKY2#1 or WRKY6#1, WRKY21#1 and WRKY7#1 primers yielded only two clearly separable amplicons while the third amplicon was missing (Figure 6). The triplex PCR conditions were probably not optimized for amplification of the largest amplicons (either the 319 bp or 354 bp amplicon). However, we conducted no further triplex PCR optimization in this evaluation and only used duplex PCR for genotyping of Banten Tall, Jember Tall, Kalianda Tall, Pati Dwarf and Sumenep Tall Kopyor coconuts.

Duplex PCR for genotyping Kopyor coconut accessions

We have optimized duplex PCR using a combination of two sets of WRKY specific SNAP primer pairs for genotyping of coconuts. Four selected primer sets were used to genotype 15 accessions each of Banten Tall (B), Jember Tall (J), Kalianda Tall (K), Pati Dwarf (P), and Sumenep Tall (S) Kopyor coconuts and validate the potential use of the developed SNAP markers. In this evaluations, the selected primer sets used in duplex PCR were WRKY6#3 (amplicon: 352 bp)-WRKY19#1 (amplicon: 211 bp) and WRKY6#1 (amplicon: 107 bp)-WRKY2#2 (amplicon: 260 bp). Figure 7 presented the representative results of the duplex PCR evaluation for WRKY6#1, and WRKY2#2 primer sets. Under the optimized duplex PCR conditions, the four studied primer sets were able to generate the targeted amplicons efficiently.

Table 4 summarized genotyping results of 75 samples of five Kopyor coconut populations (B, J, K, P and S Kopyor coconuts) using the four SNAP marker loci. The genotypes of majority individuals in all five populations (62 out of 75 individuals) were heterozygous in four loci (Table 4). Six samples (1 K, 3 P, 1 J, and 1 B) were identified as homozygous in either one or two out four evaluated SNP loci. Moreover, we identified seven individuals (3 B, 1 K, 2 P and 1 S) as having 0 (no PCR amplified product for both Ref and Alt alleles) in one or two out of four evaluated SNP loci (Table 4).

Figure 8 presents tree construction based on dissimilarity matrix calculated using the genotype data from 75 individuals of B, J, K, P and S Kopyor populations. Most of the individuals belonged to the group 1, consisted of 62 mix accessions of B, J, K, P and S Kopyor coconut populations having the same genotypes plus the J1 and J6 individuals from Jember. The group 2 consisted of a mix of total eight of individuals (2 B, 1 K, 3 P, and 2 S) while the group 3 consisted of four samples (2 B, 1 K and 1 S), respectively (Figure 8).

In multiplex PCR, selection of the amplicon sizes for the target loci needs serious consideration especially if the size fractionation were carried out using agarose gel electrophoresis. Such selection is necessary to avoid overlapping among PCR amplicons (Culley et al. 2013) which result in ambiguous scoring. Agarose gel electrophoresis can only differentiate 30 bp differences in fragments of less than 300 bp while larger DNA fragment requires larger size differences to be able to score the amplicons unambiguously (Le et al. 2012; Liu and Wu 2012; Culley et al. 2013). Although multiplex PCR has advantages, the potential disadvantages should also be considered before applying the approach such as selection among different sets of primers and low amplification efficiency because of primer competition (Xu et al. 2012).

We have analyzed 35 accessions of coconut WRKY genes available in the NCBI GenBank DNA database, identified some bi-allelic SNP sites and used the selected eight SNPs to develop SNAP markers. Each of the generated primer sets was able to successfully generate amplicon using singleplex PCR, indicating the effectiveness of the primer sets to generate SNAP markers. Since genotyping using the SNAP marker require two PCR reactions for each evaluated SNP locus, optimized duplex PCR using two sets of the generated primers in one PCR reaction could be used to save resources. Using the generated WRKY specific SNAP markers, we have demonstrated the effectiveness and informativeness of duplex PCR to evaluate genotypes of five populations of Kopyor coconut samples originated from various places in Indonesia.

In summary, based on 35 sequence data of coconut WRKY genes available in the NCBI GenBank DNA database, we identified eight informative SNPs and used to generate WRKY gene specific SNAP primers. Sixteen primer pairs were designed and selected based on the eight SNP sites. Validation of the designed primers using singleplex PCR revealed they were effective for generating SNAP markers from Kopyor coconut accessions. Subsequently, we have evaluated optimization of the Ta, primer concentrations, and either duplex or triplex PCR. Under the current conditions, the duplex PCR using two sets of primer pairs was more reliable for genotyping coconut germplasm using SNAP markers than triplex PCR. We have successfully demonstrated duplex PCR using four WRKY specific SNAP markers for genotyping samples of

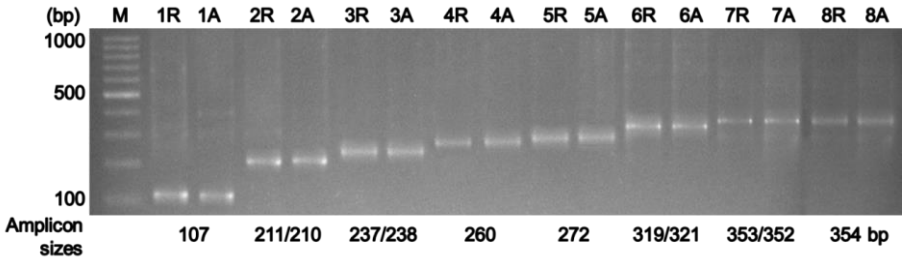

Figure 2. Singleplex PCR reactions to validate the effectiveness of the selected SNAP specific primers and amplify SNAP markers from coconut. The PCR amplifications were conducted using the following primers: (1) WRKY6#1 (Ref and Alt, 107 bp); (2) WRKY19#1 (Ref and Alt, 211/210 bp); (3) WRKY21#3 (Ref and Alt, 237/238 bp); (4) WRKY2#2 (Ref and Alt, 260 bp); (5) WRKY21#1 (Ref and Alt, 272 bp); (6) WRKY2#1-319/321 bp, (7) WRKY6#3-353/352 bp and (8) WRKY7#1-354 bp. The R-or A-labels represent either the Reference (Ref) or Alternate (Alt) primers. The primer annealing temperature settings were as listed in Table 3. All primers were used with their appropriate reverse primers while the PCR reactions were done using the same Pati Dwarf Kopyor coconut DNA as a template

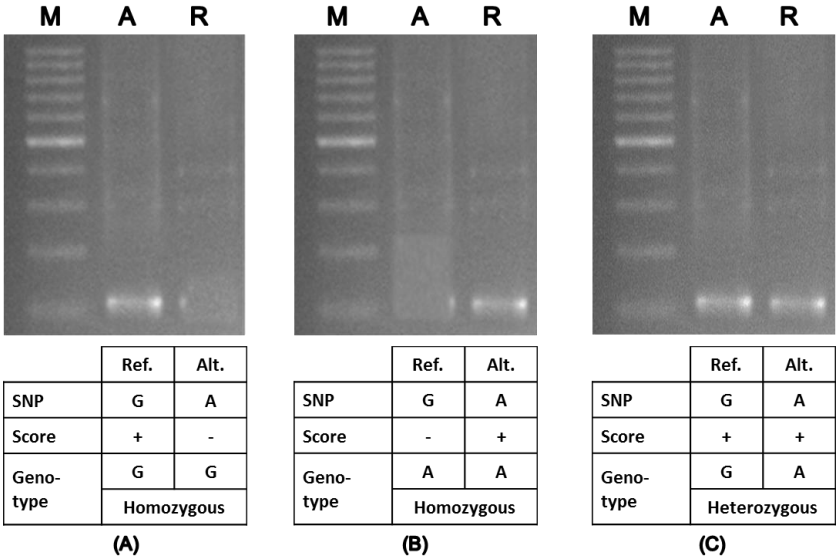

Figure 3. Genotyping of coconut with SNAP markers using singleplex PCR results in three possible coconut genotypes. (A) Homozygous for Ref, it scores (+) for Ref and (-) for Alt. The genotypes for the plant is GG, (B) homozygous for Alt, its scores (-) for Ref and (+) for Alt. The genotypes for the plant is AA and (C) heterozygous for Ref and Alt, its score (+) for Ref and (+) for Alt. The genotype of the plant is AG

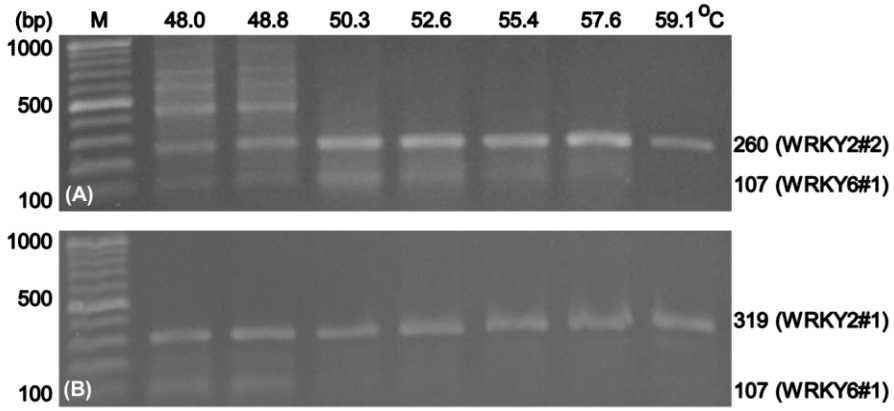

Figure 4. Annealing temperature affects SNAP marker amplification of coconut in duplex PCR. We evaluated effects of annealing temperatures in duplex PCR and each PCR reaction. The two primer pairs consisted of (A) WRKY6#1 Ref and WRKY2#2 Alt or (B) WRKY6#1 Alt and WRKY2#1 Ref-were used to amplify two different sizes of amplicons. Amplification of DNA template using WRKY6#1, both Ref and Alt yielded 107 bp, WRKY2#1-319 bp and WRKY2#2-260 bp amplicons, respectively. All tested primers were used with their appropriate reverse primers while the PCR reactions were done using the same Kalianda Tall Kopyor coconut DNA template

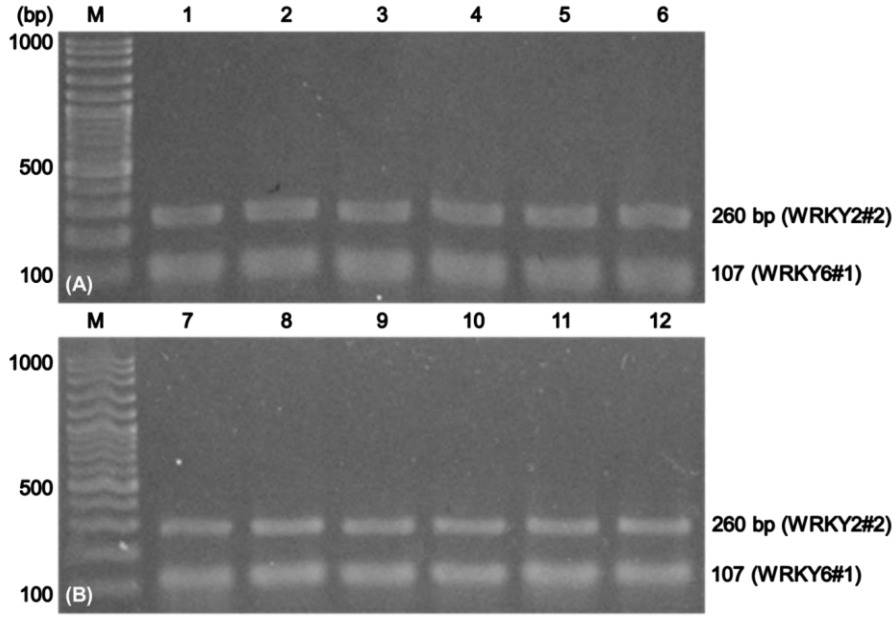

Figure 5. Primer concentration affects SNAP marker amplification of coconut in duplex PCR. Amplicons yielded from duplex PCR using primer concentration of either (A) 0.75 µL of 10 µM or (B) 0.375 µL of 10 µM for each primer. For each duplex PCR reaction, two primer pairs consisting of WRKY6#1 Ref and WRKY2#2 Alt were used to amplify two different amplicon sizes. Duplex PCR amplification of DNA template using both WRKY6#1 and WRKY2#2 primer pairs was done using Ta=55.4ºC, and the expected amplicon size was either 107 bp or 260 bp, respectively. All tested primers were used with their appropriate reverse primers while the PCR reactions were done using the same Kalianda Tall Kopyor coconut DNA template

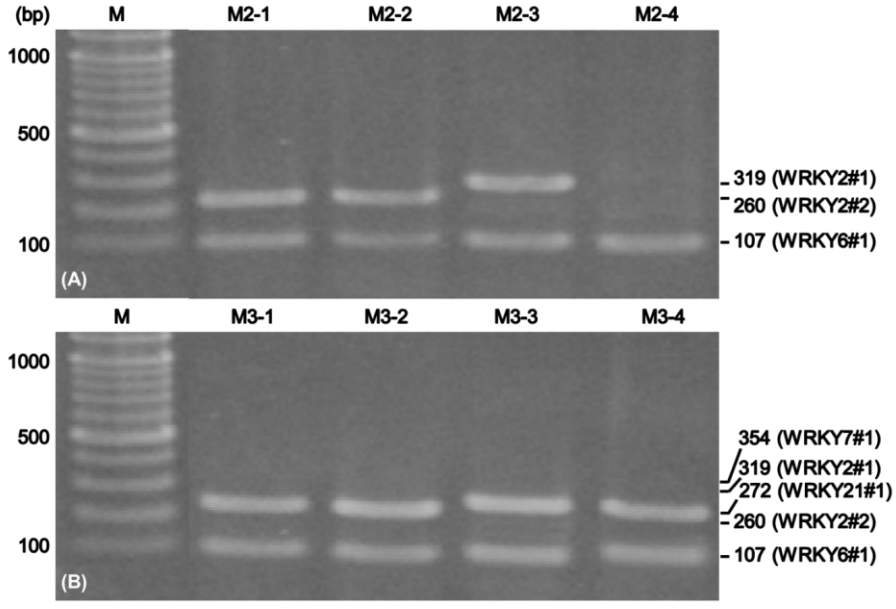

Figure 6. Multiplex PCR to amplify SNAP marker from coconut using (A) duplex PCR and (B) triplex PCR. For each duplex PCR reaction, the primer pairs used for M2-1 were WRKY6#1 Ref and WRKY2#2 Ref while for M2-2-WRKY6#1 Alt and WRKY2#2 Alt; for M2-3 were WRKY6#1 Ref and WRKY2#1 Ref while for M2-4-WRKY6#1 Alt and WRKY2#1 Alt. For each triplex PCR, the primer pairs used for M3-1 were WRKY6#1 Ref, WRKY2#2 Ref, and WRKY21#1 Ref while M3-2-WRKY6#1 Alt, WRKY2#2 Alt and WRKY21#1 Alt; for M3-3 were WRKY6#1 Ref, WRKY21#1 Ref, and WRKY7#1 Ref while M3-4-WRKY6#1 Alt, WRKY21#1 Alt and WRKY7#1 Alt. The Ta was set at 55.4ºC (M2-1, M2-2, M3.1, and M3.2); 48.8ºC (M2-3 and M2-4) or 52.6ºC (M3.3 and M3.4). All primers were used with their appropriate reverse primer while the PCR reactions were done using the same Pati Dwarf Kopyor coconut DNA template

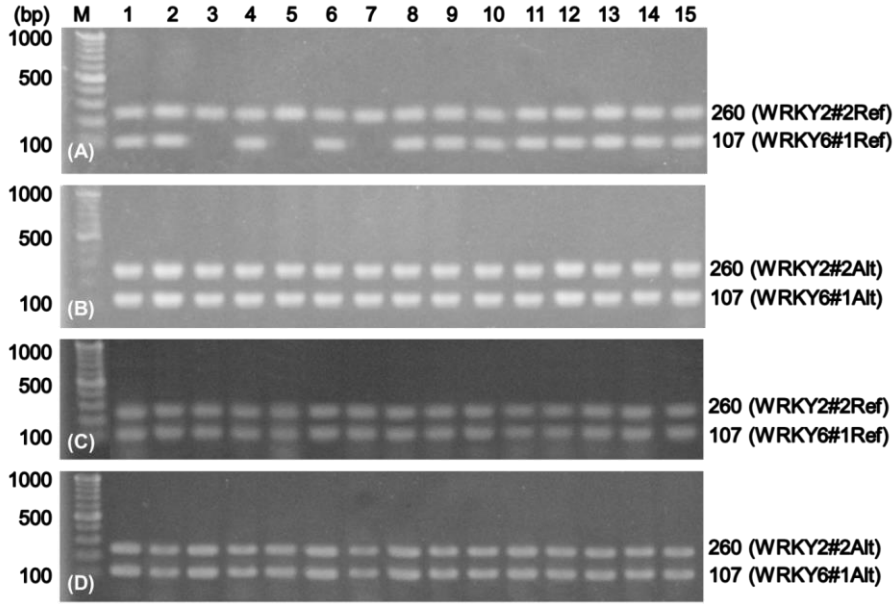

Figure 7. Representative results of SNAP marker application to genotype Pati Dwarf (A and B) and Jember Tall (C and D) Kopyor coconut using duplex PCR. The genotyping was conducted using two pairs of WRKY specific SNAP markers in a duplex PCR reaction. The two primer pairs are WRKY6#1Ref and WRKY2#2Ref (A and C) and WRKY6#1Alt and WRKY2#2Alt (B and D). We used all of those primers in duplex PCR with their appropriate reverse primers. 1-15, 15 samples of either Pati Dwarf or Jember Tall Kopyor coconuts. M, a size control of 100 bp DNA ladder

Table 4. Predicted genotypes of four SNP loci and their frequency in Banten Tall (B), Jember Tall (J), Kalianda Tall (K), Pati Dwarf (P), and Sumenep Tall (S) Kopyor coconuts based on SNAP markers. The four loci of SNAP marker data were generated using duplex PCR

Allele variants at WRKY SNP loci:								The frequency of individuals in a population:					
WRKY6#3		WRKY19#1		WRKY6#1		WRKY2#2		B	J	K	P	S	Total
T	C	G	A	C	T	G	C	11	14	13	12	12	62
T	C	G	A	C	T	G	G	-	-	1	-	-	1
T	C	G	A	T	T	G	C	-	-	-	3	-	3
C	C	G	A	C	T	G	C	-	1	-	-	-	1
T	C	G	A	T	T	C	C	1	-	-	-	-	1
T	C	G	A	C	T	0	0	-	-	-	-	1	1
T	C	G	A	0	0	G	C	-	-	1	-	-	1
0	0	G	A	C	T	G	G	1	-	-	-	-	1
T	C	0	0	C	C	G	C	1	-	-	-	-	1
T	C	0	0	0	0	G	C	-	-	-	2	-	2
0	0	0	0	T	T	G	G	1	-	-	-	-	1

Note: The bi-allelic nucleotides for WRKY6#3 SNP locus are T/C, WRKY19#1-G/A, WRKY6#1-C/T and WRKY2#2-G/C. The "0" indicates no PCR amplified product. SNP loci shaded with green were heterozygous for Reference (Ref) and Alternate (Alt); yellow-homozygous for either Ref or Alt allele and red-no PCR amplified product using either Ref or Alt primer pair.

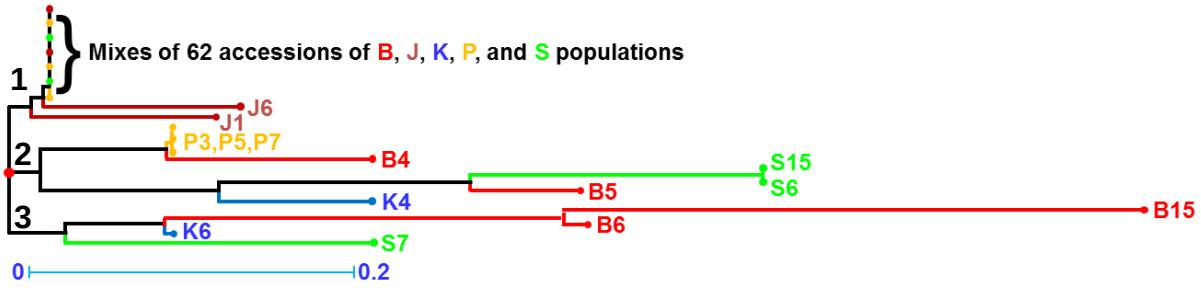

Figure 8. Unrooted Neighbour-Joining tree (weighted) realized based on dissimilarity matrix for Banten Tall (B), Jember Tall (J), Kalianda Tall (K), Pati Dwarf (P), and Sumenep Tall (S) Kopyor coconuts based on four loci of SNAP marker genotype data. The SNAP markers were generated using duplex PCR. The tree shows a total of 75 individuals (15 individuals per population) of B, J, K, P and S Kopyor coconuts

Banten Tall, Jember Tall, Kalianda Tall, Pati Dwarf and Sumenep Tall Kopyor coconuts. Therefore, we can use the developed SNAP markers as a simple alternative and co-dominant markers for genetic analysis of coconuts. To be more useful, more gene-specific primers to generate SNAP markers are needed. Such approach should be feasible with the availability of more gene sequences deposited in the NCBI GenBank DNA database, including ones from coconut genomes.

ACKNOWLEDGEMENTS

Part of this research was supported by KKP3N-Kopyor Coconut Research Project, under the Project Number: 708/LB.620/1.1/2/2013, Agency for Agricultural Research and Development (AARD), Ministry of Agriculture, Republic of Indonesia (2013-2015). Research grants and Post Graduate Scholarship from the Ministry of Research, Technology and Higher Education, the Directorate General of Higher Education, Republic of Indonesia supported Anneke Pesik. The authors are thankful to Ismail Maskromo for his help in sample collection and to Zikril Illahi for his help in lab activities.

AUTHOR CONTRIBUTIONS

Anneke Pesik has participated in the data collections and validation-executed the day to day activities in the laboratory and contributed to the writing of the manuscript. Darda Efendy, Hengky Novarianto, and Diny Dinarti have supervised Anneke Pesik's laboratory activities-assisted and reviewed the manuscript preparation. Sudarsono Sudarsono has been the principal investigator for the Indonesian Kopyor Coconut Project and the corresponding author-designed and organized the research, identified SNP and designed SNAP primers, develop SNAP markers, conducted data analysis, prepared supporting data for publication and wrote the manuscript.

REFERENCES

Bakooie M, Pourjam E, Mahmoudi SB, Safaie N, Naderpour M. 2015. Development of an SNP marker for sugar beet resistance/susceptible genotyping to root-knot nematode. J Agr Sci Tech 17: 443-454.

Bru D, Laurent FM, Philippot L. 2008. Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16s rRNA gene as an example. App Env Microbiol 74 (5): 1660-1663.

Chaerani, Hidayatun N, Utami DW. 2009. Development of a set of microsatellite DNA markers present in genetic variation of the rice and soybean [in Indonesia]. J AgroBiogen 5 (2): 57-64.

Chagné, Gasic DK, Crowhurst RN, Han Y, Bassett HC, et al. 2008. Development of a set of SNP markers present in expressed genes of the apple. Genomics 92 (5): 353-358.

Ching A, Caldwell KS, Jung M, Dolan M, Smith OSH, et al. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genetics 19 (3): 1-14.

Cho J, Lu C, Kohel RJ, Yu JZ. 2011. Development and mapping of gene-tagged SNP markers for gland morphogenesis in cotton. J Cotton Sci 15: 22-32.

Culley TM, Stamper TI, Stokes RL, Brzyski JR, Hardiman NA, et al. 2013. An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. Appl Plant Sci 1 (10). doi:10.3732/apps.1300027.

EulgemT, Rushton PJ, Robatzek S, Somssich IE. 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci 5: 199-206.

Fang WP, Meinhardt LW, Tan HW, Zhou L, Mischke S, Zhang D. 2014. Varietal identification of tea (*Camellia sinensis*) using nanofluidic array of single nucleotide polymorphism (SNP) markers. Hortic Res 1: 14035. DOI: 10.1038/hortres.2014.35.

Hayden JM, Nguyen TM, Waterman A, Chalmers KJ. 2008. Multiplex-ready PCR: a new method for multiplexed SSR and SNP genotyping. BMC Genomics 80 (9): 1-12.

Kim JS, Jang HW, Kim JS, Kim HJ, Kim JH. 2012. Molecular identification of *Schisandra chinensis* and its allied species using multiplex PCR based on SNPs. Genes Genomics 2012 (34): 283-290.

Larekeng SH, Maskromo I, Purwito A, Mattjik NA, Sudarsono S. 2015a. Pollen dispersal based on SSR analysis proves Kalianda to Kopyor coconut pollinations. Bulletin Palma 16 (1):77-92. [Indonesian]

Larekeng SH, Maskromo I, Purwito A, Mattjik NA, Sudarsono S. 2015b. Pollen dispersal and pollination pattern studies in Pati Kopyor coconut using molecular markers. Intl J CORD 31 (1):46-60.

Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Le XTK, et al. 2012. Development and Evaluation of a Single-Step Duplex PCR for Simultaneous Detection of *Fasciola hepatica* and *Fasciola gigantica* (Family Fasciolidae, Class Trematoda, Phylum Platyhelminthes). J Clinical Microbiol 50 (8): 2720-2726.

Li X, Gao W, Guo H, Zhang X, Fang DD, Lin Z. 2014. Development of EST-based SNP and InDel markers and their utilization in tetraploid cotton genetic mapping. BMC Genomics 15 (1): 1-11.

Liu L, Wu Y. 2012. Development of a genome-wide multiple duplex-SSR protocol and its applications for the identification of selfed progeny in switchgrass. BMC Genomics 13 (522): 1-13.

Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S. 2012. SNP markers and their impact on plant breeding. Intl J Plant Genomics 2012: 1-11.

Masi P, Spagnoletti ZPLS, Donini P. 2003. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.). Mol Breed 11: 303-313.

Maskromo I, Tenda TE, Tulalo MA, Novarianto H, Sukma D, et al. 2015. Fenotipic and genetic diversity of three Dwarf Kopyor coconut from Pati, Central Java, Indonesia. J Littri 2 (1): 1-8. [Indonesian]

Migliaro D, Morreale G, Gardiman M, Landolfo S, Crespan M. 2012. Direct multiplex PCR for grapevine genotyping and varietal identification. Plant Genet Resour 2012: 1-4.

Mousset M, Flaven E, Justy F, Pouzadoux J, Gode C, et al. 2015. Characterization and multiplexing of 21 microsatellite markers for the herb *Noccaea caerulescens* (Brassicaceae). Appl Plant Sci 3 (12): 1-6.

Olmos CC, Vilanova S, Pascual L, Roselló J, Cornejo JC. 2015. SNP markers applied to the characterization of Spanish tomato (*Solanum lycopersicum* L.) landraces. Sci Hortic 194: 100-110.

Ren J, Sun D, Chen L, You FM, Wang J, et al. 2013. Genetic diversity revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. Intl J Mol Sci 14: 7061-7088.

Sheng ZL, Becquet V, Hua LS, Zhang D. 2003. Optimization of multiplex PCR and multiplex gel electrophoresis in sunflower SSR analysis using infrared fluorescence and tailed primers. Acta Botanica Sinica 45 (11): 1312-1318.

Simko I, Eujayl I, Hintum TJL. 2012. Empirical evaluation of DArT, SNP, and SSR marker systems for genotyping, clustering, and assigning sugar beet hybrid varieties into populations. Plant Sci 184: 54-62.

Singh N, Choudhury DR, Singh AK, Kumar S, Srinivasan K, et al. 2013. Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. PLoS One 8 (12): e84136. doi: 10.1371/journal.pone.0084136.

Sint D, Raso L, Traugott M. 2012. Advances in multiplex PCR: balancing primer efficiencies and improving detection success. Methods Ecol Evol 2012 (3): 898-905.

Sudarsono, Novarianto H, Hossang MLA, Dinarti D, Yuliasti, et al. 2015. Genomic and molecular approach for developing exotic coconut cultivars from Indonesia, propagation of the planting materials and control of the insect pests. Final report, KKP3N Kopyor Coconut Project, Agency for Agricultural Research and Development

(AARD), Ministry of Agriculture, Republic of Indonesia. [Indonesian]

Sukendah, Volkaert H, Sudarsono. 2009. Isolation and analysis of DNA fragment of genes related to Kopyor trait in coconut plants. Indonesian J Biotech 14 (2): 1169-1178.

Sutanto A, Hermanto C, Sukma D, Sudarsono. 2013. Development of SNAP marker based on resistance gene analogue genomic sequences in banana (*Musa* spp.). J Hort 23 (4): 300-309. [Indonesian]

Wang B, Tan HW, Fang W, Meinhardt LW, Mischke S, et al. 2015. Developing single nucleotide polymorphism (SNP) markers from transcriptome sequences for identification of longan (*Dimocarpus longan*) germplasm. Hortic Research 65: 14065. DOI: 10.1038/hortres.2014.65.

Wen D, Zang C. 2012. Universal multiplex PCR: a novel method of simultaneous amplification of multiple DNA fragments. Plant Methods 32 (8): 2-9.

Xu W, Zhai Z, Huang K, Zhang N, Yuan Y, et al. 2012. A novel universal primer-Multiplex-PCR method with sequencing gel electrophoresis analysis. PLoS One 7 (1): e22900. DOI: 10.1371/journal.pone.0022900.

Zhang Y, Wang L. 2005. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol Biol 5 (1): 1-12.