Development of gene based markers for mapping of oil yield traits in African oil palm (*Elaeis guineensis* Jacq.).

Sukanya Jeennor[1,2] and Hugo Volkaert[1,2,3]*

[1]*Center for Agricultural Biotechnology, Kasetsart University, Kampangsaen Campus, Nakhon Pathom 73140, Thailand*

[2]*Center of Excellence on Agricultural Biotechnology: (AG-BIO/PERDO-CHE), Bangkok, 10900, Thailand*

[3]*Agriculture and Food Biotechnology Unit, National Center for Genetic Engineering and Biotechnology, Thailand Science Park, Pathumthani, 12120, Thailand*

**Corresponding author: hugo.v@ku.ac.th*

ABSTRACT

Gene based SNP markers for oil palm were developed from cloned candidate genes that are involved in fatty acid (FA), triacylglycerol (TAG) and pyruvate-acetyl CoA synthesis pathways of oil palm. Based on exon/introns splice junctions, a total of 22 specific primer pairs corresponding to 22 gene loci were designed and tested for polymorphism between 23 parents (13 *dura* and 10 *pisifera*) used in an oil palm breeding program. Single strand conformation polymorphism polyacrylamide gel electrophoresis (SSCP) was used to detect the polymorphism. More than 80% of the gene loci clearly showed polymorphic patterns in target parents which progenies of at least 3, up to all 14 crosses would display segregation for the markers. The gene based SNP markers developed in this study do not only provide potential marker tools for further QTL mapping analysis in multiple crosses but can also be used as anchor points for further comparing and integrating oil palm linkage maps developed from individual progenies.

Keywords: Oil palm; SNP; molecular maker; oil synthesis; SSCP

INTRODUCTION

The advent of molecular marker technology has provided an opportunity for accelerating breeding research in crops to improve productivity through marker-assisted selection (MAS) (Thomas *et al.*,1969). Knowledge about the inheritance and the linkage relationships of molecular markers and traits of economic interest provide breeders with powerful molecular tools that can be used to improve the efficiency of crop improvement

(Shappley *et al.*, 1996). MAS is well suited to a perennial crop like oil palm, in which the economic products are not produced until several years after planting. The use of DNA markers for selection in such crops can greatly reduce the length of breeding cycles needed. With the use of DNA markers, informed decisions can be made at the nursery stage, allowing only the selected palms to be field-planted and could be of great potential value. Improving crop yield traits in practical breeding programs are difficult because they are quantitative traits controlled by many genes with small or large effect. Identification of the chromosomal regions underlying quantitative traits or quantitative trait locus (QTL) could provide the valuable knowledge for developing the molecular marker tool. QTL analysis requires a large number of polymorphic markers that show segregation in the mapping population. Different type of molecular markers both dominant and co-dominant markers have been extensively developed and used for QTL mapping analysis in various economic plants. In highly heterozygous breeding population, dominant marker such as randomly amplified polymorphic DNA (RAPDs) and amplified fragment length polymorphism (AFLPs) markers give only a very partial insight into the inheritance (Grattapaglia *et al.*, 1996; Lerceteau *et al.*, 2000 and Scalfi *et al.*, 2004). Presently, co-dominant marker such as microsatellites (simple sequence repeats; SSRs) and single nucleotide polymorphism (SNP) markers derived from expressed genes or candidate genes are becoming an accepted strategy to identify QTLs that are tightly linked with economically important traits in both animal and plant system (Collard and Mackill, 2008). Molecular markers based on candidate gene that may control the interesting trait have been developed and applied for QTL analysis in several plant species such as *Pinus taeda* (Neale *et al.*, 2002), *Eucalyptus globulus* (Thamarus *et al.*, 2004), black cottonwood (Wegrzyn *et al.*, 2010), maize (Yang *et al.*, 2010) and canola (Smooker *et al.*, 2011). Candidate gene based markers have not yet been report in African oil palm.

The metabolic processes that control the biosynthesis of storage oil in plants are very complex and there is no single major step controlling the oil synthesis (Ohlrogge and Browse, 1995). Recent studies in *Arabidopsis thaliana* (Baud *et al.*, 2007) and *Brassica napus* (Li *et al.*, 2006) indicated that pyruvate kinase (PK) that converts phosphoenolpyruvate (PEP) to pyruvate is a critical enzyme for the production of fatty acid biosynthesis precursors. Disruption of this gene lead to a reduction in seed oil content up to 60% in mature seed of *A. thaliana* (Baud *et al.*, 2007). Several key enzymes involved in fatty acid (FA) and triacylglycerol (TAG) synthesis have been reported to be associated with changes in oil accumulation such as acetyl CoA carboxylase (ACCase), glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAT), and diacylglycerol acyltransferase (DGAT). High expression of ACCase in *B.*

napus resulted in relative increase in 5% of seed oil content (Roesler *et al*., 1997). While, expression of safflower GPAT and *Escherichia coli* GPAT in transgenic *A. thaliana* found that both GPATs increased the seed oil content of ranging from 10-22% (Jain *et al*., 2000). The acyltransferase LPAT that produce phosphatidic acid was also reported to be associated with oil content in seed oil. Transgenic lines of *A. thaliana* and *B. napus* containing a mutated yeast LPAT (*SLC1-1*) under a constitutive promoter displayed increases in oil content in mature seed ranging from 8-48% (Taylor *et al*., 2001). Among acyltransferase enzymes in the TAG synthesis pathway, DGAT which catalyzes the final step of the pathway has been most extensively studied in terms of modified seed oil content. Over-expression *DGAT* genes under a constitutive promoter could dramatically increase the oil content of many seed oil plants ranging from 18-50% (Zheng *et al*., 2008 and Taylor *et al*., 2009). Although the activity of this enzyme was found to affect oil accumulation in various oil crops, no such effect was found on the oil accumulation in oil palm (Ramli *et al*., 2005). Thus, this study aimed to develop the candidate gene based marker applied for oil palm mapping analysis. Based on previously studies, candidate genes involved in FA, TAG and pyruvate-acetyl CoA synthesis pathways of oil palm were isolated and characterized. Locus specific primers were designed based on their sequences and intron/exon junction information. Then the PCR-single strand conformation polymorphism (PCR-SSCP) analysis was conducted and tested for polymorphism in 23 parents selected from 14 controlled crosses.

MATERIALS AND METHODS

Plant materials

To screen for polymorphic gene based markers, 23 oil palms consisting of 13 Topi Deli *dura* and 10 Yangambi *pisifera* were selected. These plants had been used as parents in 14 crosses by Univanich Palm Oil Public Company Ltd., Krabi, Thailand.

DNA/RNA extraction

Total genomic DNA was extracted from young leaf tissue using a commercially available kit (Plant genomic DNA isolation kit, Geneaid, Taiwan). Total RNA was extracted from developing fruits with a guanidine-based protocol (Dellaporta *et al*., 1983). The quantity and quality of the extracted DNA/RNA was checked on 1% agarose gel electrophoresis and Nanodrop 8000 spectrophotometer (Thermo scientific, USA).

Cloning of candidate genes

Based on previous reports on genes involved in oil synthesis pathway in model plants or other oil producing crops, more than 20 genes involved in FA, TAG and pyruvate-acetyl CoA synthesis pathways were targeted in this study. DNA sequences of these candidate genes from various plant species were retrieved by BLAST searching (the National Center for Biotechnology Information (NCB)I; http://www.ncbi.nlm.nih.gov/) and aligned with ClustalW (Thomson *et al*., 1994). Based on the sequence alignment, oligonucleotide primers corresponding to each candidate gene were designed from the conserved regions in homologous genes (Table 1). The PCR amplification was performed using genomic DNA or cDNA of oil palm as template. To clone gene by RT-PCR, approximately 3 µg of total RNA was reverse transcribed using the oligo-dT primer (17 bp) and 5 U SuperScript II first strand synthesis system (Invitrogen, CA).

The PCR contained approximately 25 ng of template, 1XPCR buffer with $(NH_4)_2SO_4$, 2 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 µM of each primer and 0.5 U of Taq DNA polymerase (Fermentas UAB, Lithuania) in a total volume of 20 µl. Reactions were incubated for 3 min at 94°C, followed by 35 cycles of 94°C for 35 s, primer-specific optimised annealing temperature for 40 s, 72°C for 1 –2 min and a final extension of 72°C for 5 min. The amplified products were purified using the PCR-Min-elute kit (Qiagen) following the manufacturer's protocol and then cloned into pGEM-T vector (Promega). To identify the clones that contained the PCR product, colonies were used directly for PCR using the universal T7 and SP6 primers or the degenerate primers. Plasmid DNA from transformant clones was extracted using a high speed plasmid mini kit (Geneaid, Taiwan) and sent for sequencing (1st Base, Malaysia).

Sequence analysis and primer design

To identify the homologous sequences corresponding to each candidate gene, the obtained sequences were compared with homologs in GenBank database using the BLASTn program. All oil palm sequences were also aligned to the rice *(Oryza sativa*) reference genome sequence, to identify the exon/intron splice junctions. A total of 28 locus specific primer pairs were designed using the Primer3 software program (http://frodo.wi.mit.edu/primer3/) targeting individual genomic DNA fragments containing one or more introns of each candidate gene. All sequence data of candidate genes used for primer design were deposited in GenBank (Table 2).

Polymorphisms and heterozygosity analysis

To detect polymorphism between the "*dura*" and "*pisifera*" crossing parents, genomic DNA fragments were amplified by PCR as mentioned above. The annealing temperature for the PCR was set at 55°C except for the EgFAD3 marker for which amplification was done at 50°C annealing temperature. Polymorphism analysis was done on non-denaturing polyacrylamide gels using single strand conformation polymorphism (SSCP) conditions (Orita *et al.*, 1989). 2 µl of PCR products was mixed with 4 µl of SSCP-loading dye (95% (v/v) formamide, 0.025% bromphenol blue, 0.025% xylene cyanol and 10 mM NaOH),

Table 1 List of consensus primer corresponding to candidate genes involved in oil synthesis. The consensus primer, optimum amplified PCR condition and derived homologous sequences with intron analyzed were presented.

Name of candidate gene	**Abbreviation**	**Consensus primers (5'-3')[*]**		**PCR Template[**]**	**PCR annealing temp (°C)**	**PCR product size (bp)**	**No. of intron**
		Forward	**Reverse**				
Fatty acid synthesis							
Beta-ketoacyl ACP synthase I	*KASI*	GATGGTTTTGTNATGGGWGARGG	CCACCAAATCCAAANGARTTNGA	gDNA	55	--	--
Beta-ketoacyl ACP synthase II	*KASII*	GGTTGGATGGGTCCWAAYTAYTC	CCTTCTCCCATGACAAANCCRTC	gDNA	55	658	2
3-Ketoacyl-ACP-reductase	*KRACP*	ACTGGTGCTTCTAGAGGNATHGG	CTGTGTACAAAGGAATACNCCNGT	cDNA	52	339	4
3-Hydroxyacyl-ACP dehydrase	*DHACP*	CCTATAATGCCTGGTGTYCTNATG	TAAGCTTTTCCTTCCATYTTNGC	cDNA	52	224	1
Enoyl-ACP reductase	*EACP*	GCTGGATCTTCAAAYTGGACNGT	GCTGCTTTTGCTGAACTCATNCC	cDNA	50	298	4
3-Ketoacyl-CoA synthase	*KCS*	GGTGGAATGGGTTGYAGYGCNGG	CCAAACTGCACTRTTRCAYTTRAA	gDNA	52	738	0
Steroyl-ACP desaturase	*SACPD*	TGGACGAGGGCWTGGACTGCNGARGA	GGCATTGAGATYTTYTTCYTCATCAT	gDNA	55	897	1
Plastid-oleate desaturase	*FAD6*	ATAGGTCATGATTGTGCNCAYAARTC	GGAGCTGTGTGATGTACCATNGTRAA	cDNA	48	479	3
Triacylglycerol synthesis							
Acyl-CoA: lysophosphatidate acyl transferase 1	*LPAT1*	AGCTTCTTGGATATNTAYACMCT	GGAAAGAAAAATWCRGANGCTCC	gDNA	55	722	1

Table 1 Continued.

Name of candidate gene	Abbreviation	Consensus primers (5'-3')*		PCR Template**	PCR annealing temp (°C)	PCR product size (bp)	No. of intron
		Forward	Reverse				
Acyl-CoA diacylglycerol acyltransferase	*DGAT*	TTCATGATGGCTCCTACAYTNTGYTA	GCATTCCACCAATCYTTRTARAAYTC	cDNA	52	334	3
Pyruvate acyl-CoA synthesis							
Plastid NADPH-dependent malate dehydrogenase	*PMDH*	GAAGACTTTGCAAACCAYAAYGC	CTTTGAGTCCACAAGCCARATYTTYTT	cDNA	48	282	2
Phosphoenolpyruvate carboxylase	*PEPC*	GAGGTCATGATTGGCTAYTCNGAYTC	GTTTGTGTCCAAGCAAADATCCANGG	gDNA	55	964	1
Biotin carboxyl carrier protein	*BCCP*	CCTATGGCTGGRACNTTYTA	TCATTCATCAGTTTCATNGCYTC	gDNA	55	391	1
Alpha-carboxyl transferase	*CT*	AACTTTGGGATGCCWACTCCNCAYGG	TTCCACAAGATTGCWGCACANGCYTC	gDNA	52	1,390	3
Biotin carboxylase	*BC*	ATGTTGCATCCTGGNTATGGTTTYC	CCATCATTTCCAAAWGCAGCNGC	cDNA	50	359	4

* Degenerated base: Y = C / T, R = A/G , M = A/C, W = A/T, S = G/C, K = G/T, N = C/G/A/T

** Two types of PCR template; gDNA (genomic DNA) and cDNA (complementary DNA) then denatured at 95°C for 8 min and immediately placed in ice-cold water to stabilize single strands. Then 2 µl aliquots were loaded on top of the gels (0.33-0.4x Sequagel MD, National Diagnostics in 0.6 x TBE) with 1xTBE as electrophoresis buffer at constant 10-12 Watt for 14-22 hr in a 4°C refrigerator. DNA bands in the polyacrylamide gel were visualized with silver staining. The level of polymorphism and heterozygosity for each marker among the 23 breeding parents were calculated by the Excel Microsatellite Toolkit (Park, 2001).

RESULTS AND DISCUSSION

Identification and isolation of candidate genes in oil palm

Candidate genes implicated in FA, TAG and pyruvate-acetyl coA synthesis pathways have been extensively studied in plant oil crops as well as oleaginous microorganism. In oil palm, homologous sequences of a few genes were already available in public resources (GenBank). However, many other related genes had not yet been characterized. A total of 16 candidate genes involved in FA, TAG, pyruvate-acetyl CoA synthesis pathways were cloned in this study using degenerate primers deduced from alignments of homologous sequences (Table 1). Under optimized PCR amplification conditions, PCR fragments ranging between 200-1,400 bp could be amplified from genomic DNA or cDNA of oil palm (Table 1). Sequencing and BLASTn analysis of 69 clonal inserts revealed 16 unique loci coding for 16 of the targeted genes. Two isoforms were recovered for the 3-ketoacyl-CoA synthase (KCS) gene while no homologous sequence corresponding to beta-ketoacyl ACP synthase I (KASI) gene was obtained. All sequences have strong homology with genes from monocotyledonous plants especially *O. sativa*. Comparing the cloned genomic or cDNA sequences with genomic sequences of the model plants *Arabidopsis* and *Oryza*, it was found that the homologous oil palm gene fragments contained one to four introns in their sequences.

PCR analyses with intron-flanking primer set

To take advantage of the elevated DNA polymorphism in introns compared to exons, several studies have developed gene based markers using intron-targeted primer pairs by aligning sequences from expressed gene with corresponding genomic regions of the same or related species in which the intron locations were well characterized (Wei *et al.*, 2005 and Tamura *et al.*, 2009). We adopted a similar strategy to detect polymorphism in target loci in the oil palm genome. When sequence information was obtained, PCR primers were designed on either side of potential intron/exon junctions as determined by alignment with rice genomic and cDNA sequences. PCR amplification revealed that 22 primer pairs corresponding to individual gene successfully amplified a single fragment from oil palm genomic DNA. The PCR fragment sizes amplified from each marker are shown in Table 2. Based on PCR-SSCP analysis of the amplicons, 18 primer pairs (81% of amplicons) revealed a clearly polymorphic pattern between the two target parental lines indicating the presence of one or more polymorphisms in their sequences. To evaluate the reliability of the band shifts detected in the PCR-SSCP assay, randomly 20 fragments that showed polymorphism and 18 that did not were PCR amplified with 6 gene based markers (*EgKRACP*, *EgEACP*, *EgPKCP*, *EgPMDH, EgPEPC*, and *EgCoAs*) and sequenced from the *dura* x *pisifera* parent pair used in linkage mapping. All PCR fragment sequence polymorphisms were observed corresponding to the DNA conformation variation in the SSCP gels due to the presence of one or more SNPs in the sequence of the PCR fragments. When no variation in the banding pattern on the SSCP gels was observed no sequence polymorphism could be detected in the fragment sequences. The PCR-SSCP proved to be a powerful assay to screen for SNPs in candidate genes even when sequence information is not available. The SSCP assay was found to be reliable, effective, and simple without the need for specialized equipment or reagents.

Polymorphism analysis in breeding parents

Polymorphism analysis of 23 elite oil palm parent trees with the 22 developed locus specific primers reproduced clearly 2-6 different banding patterns for individual loci (Fig. 1). Unidentified banding patterns were found in 4 loci (*EgKCS2*, *EgSACPD*, *EgFAD3*, and *EgCT*). The number of alleles, expected heterozygosity (H_e), observed heterozygosity (H_o) and polymorphism information content (PIC) for each marker are shown in Table 2 The level of heterozygosity was generally moderate with PIC ranging from 0.15 to 0.63 (average = 0.41). Assessment of the particular alleles found among the parent pairs used in the 14 crosses indicated that segregation among the progeny could be expected within an average of 8 crosses with a maximum of all 14 and minimum of 3 crosses (Table 2). This indicates that it would be possible to map most of these markers in multiple crosses, which would ultimately allow the integration of marker and phenotype data across several progenies. The co-dominant feature and anchor points of gene based markers could be a valuable tool for comparing and integrating linkage maps in multiple crosses, in order to get high confidence genetic maps for oil palm breeding.

Table 2 List of gene based markers for oil palm, their GenBank accession no., PCR product size number of polymorphic allele, number of progenies for which the marker would be segregating, expected and observed heterozygosity (He and Ho) and polymorphism information content (PIC) values based on 23 oil palm breeding parents.

Locus name	GenBank accession no.	Specific primer sequence (5'-3')	PCR product size (bp)	# allele	Number of progenies segregating	H_e	H_o	PIC
Fatty acid synthesis								
EgKASII	JN203213	F-CATCATATTTTAAGAGGTGAAACCGT R-CGTCGCGACCCTGGCATATGATT	688	2	5/14	0.36	0.25	0.29
EgKASIII	DQ459441*	F-CAGAATGTCCTAGTAATTGGTGC R-TTGCTTGAGATGAGCTCAGTTTC	730	4	3/14	0.60	0.64	0.53
EgKRACP	JN003497	F-CTGGGTTCATTGCATCTCCT R-GTGGAGGCTTAGCACACAAT	1,000	3	8/14	0.54	0.36	0.42
EgDHACP	JN003470	F-CCTGGTGTTCTTATGGTTGAGGT R-TCTCTAACATGCATGGACGAAG	560	3	8/14	0.67	0.43	0.59
EgEACP	JN203234	F-GGAGGAAGCACAAGATTAGAGT R-AAACCAGGAATGCCGCGGTAT	600	3	10/14	0.62	0.43	0.53
EgKCS1	JN203215	F-CATCCCAGTACCTATGCCTTG R-GCCTAATCTTCATCTTGAGGAC	405	2	4/14	0.49	0.14	0.36
EgKCS2	JN203216	F-CACCCCAATACCTATGCTGTC R-GCTTAATCTTCATCTTGAAAAG	405	≥2	-	-	-	-
EgSACPD	JN003489	F-TATCTATACCTTTCTGGTAGAGT R-TCATCTGAGGCAATTATACCACA	732	≥5	-	-	-	-
EgFAD3	EL687400*	F-TCGTTCTTGGCCACGATTGG R- GACTCGTCATTCTCGACGTG	400	≥4	-	-	-	-
EgFAD6	JN003493	F-TGGAGAAAGGAAATTGACTCCTC R-CCTAGGACCCCTGTCTTATAAAC	1,200	2	5/14	0.17	0.18	0.15
Triacylglycerol synthesis								
EgGPAT1	AF251795*	F-GAGATCCTAGAGCAAACAAG R-GCTTCTCTTCCATGTCAGAGA	1,500	4	12/14	0.70	0.57	0.63

Table 2 Continued.

Locus name	**GenBank accession no.**	**Specific primer sequence (5'-3')**	**PCR product size (bp)**	**# allele**	**Number of progenies segregating**	**H_e**	**H_o**	**PIC**
EgLPAT1	JN203217	F-ATGCTCCTTATGGGTGTCGTTCC R-TTCCTCACCAAATCCATGCAACG	601	3	12/14	0.51	0.54	0.39
EgPAP28	JN003504	F-GACGGAAAGGAGGTTTGATTC R-AAACCAAGACCTGCAAAGGACC	717	2	9/14	0.30	0.36	0.25
EgDGAT	JN003563	F-CGCCAGCTTGTTAAGTTGGTGAT R-GATAGCTTCAAGACTCTCTCTAT	300	4	14/14	0.56	0.75	0.50
Pyruvate-acetyl-CoA synthesis								
EgPKCP	EL687112*	F-AGCCTCAGGATTGAAAGGTG R-GTCTCCTTACAGACGTTGAAG	1,200	2	10/14	0.40	0.54	0.32
EgPMDH	JN003502	F-CTGCTTGCAAAATATGGCACTAC R-TTCCGAGTCTCTTCCAATGGAGC	1,100	3	7/14	0.55	0.32	0.44
EgPEPC	JN003566	F-GCATATATGGAGGAGCACTTGTG R-CTGCTGCCAATATTCATCCTACC	643	3	12/14	0.62	0.61	0.54
EgCoAs	EY396234*	F-GTTGTTGGAGACTCACGATGT R-CCATGACTTCTTGATGCAAAG	600	3	12/14	0.53	0.57	0.42
EgPDHE1a	EL682342*	F-AACATGATCTGGCCGGTGAAT R-ATCTGTCTTAGGAAGATCGGA	400	2	3/14	0.14	0.14	0.12
EgBCCP	JN203212	F-TTTTACCGATGCCCAGCACC R-AACAACCTGGCCTTTTTGCAC	340	6	4/14	0.44	0.25	0.39
EgCT	JN003495	F-GGGATGCCAACTCCTCATGGGT R-CCTGGAGTATCAATGAATGTGAT	575	≥6	-	-	-	-
EgBC	JN003494	F-GCTGTCAAGCTTGCACATGAA R-ACCTGGAGTAACTTTACAAACTC	1,000	2	4/14	0.57	0.14	0.46

* Accession codes starting with JN were characterized in this study, other homologous sequences were obtained from public resources

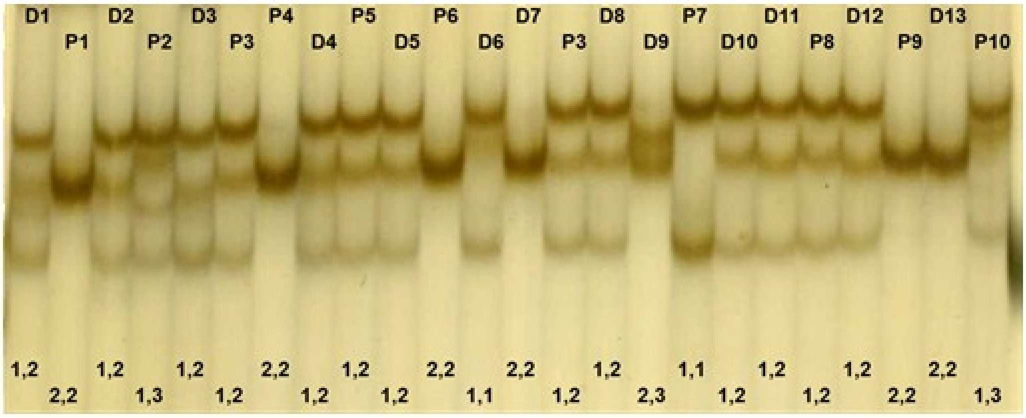

Figure 1 SSCP banding patterns for the *EgCoA* locus in a set of 23 oil palm trees, one of which was repeated as an internal control. Three different alleles were observed. The deduced genotypes for each individual are given below. D = *dura* parent, P = *pisifera* parent.

CONCLUSION

Several important genes involved in oil synthesis of oil palm were isolated and developed into gene based markers. The PCR-SSCP assay among a small set of breeding parents of *dura* and *pisifera* types found that a total 18 markers corresponding to 7 genes involved in FA synthesis, 4 genes in the TAG synthesis and 7 genes in pyruvate-acetyl CoA synthesis showed clear polymorphism patterns. The co-dominant markers developed in this study can be conveniently applied for a variety of purposes including linkage map analysis, association mapping using genome scanning, detection of QTLs especially for oil yield traits, and comparative mapping in oil palm.

ACKNOWLEDGEMENTS

SJ was supported by the Thailand Graduate Institute of Science and Technology (TGIST) from National Science and Technology Development Agency (NSTDA). This research was supported by grants from the National Center for Genetic Engineering and Biotechnology (BIOTEC), NSTDA and the Center of Excellence on Agricultural Biotechnology, Science and Technology Postgraduate Education and Research Development Office, Office of Higher Education Commission, Ministry of Education. (AG-BIO/PERDO-CHE), Thailand. We thank Julapark Chunwongse, Kasetsart University, for his critical comments. We also thank the Univanich Palm Oil Public Company Limited, Krabi, Thailand for kindly providing plant samples.

REFERENCES

Baud S, Wuillime S, Dubreucq B, de Almeida A, Vuagnat C, Lepiniec L, Miquel M and

Rochat C (2007) Function of plastidial pyruvate kinases in seeds of *Arabidopsis thaliana*. Plant J 52: 405–419.

Collard BC and Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos Trans R Soc Lond B Biol Sci 363: 557–572.

Dellaporta SL, Wood J and Hicks JB (1983) A plant DNA miniprepration: version II. Plant Mol Biol Rep 1.

Grattapaglia D, Bertolucci FL, Penchel R and Sederoff RR (1996) Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics 144: 1205–1214.

Jain RK, Coffey M, Lai K, Kumar A and MacKenzie SL (2000) Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. Biochem Soc Trans 28: 958–961.

Lerceteau E, Plomion C and Andersson B (2000) AFLP mapping and detectionofquantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. Mol Breed 6: 451–458.

Li RJ, Wang HZ, Mao H, Lu YT and Hua W (2006) Identification of differentially expressed genes in seeds of two near-isogenic *Brassica napus* lines with different oil content. Planta 224: 952–962.

Neale DB, Sewell MM and Brown GR (2002) Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. Ann For Sci 59: 595–605.

Ohlrogge J and Browse J (1995) Lipid biosynthesis. Plant Cell 7: 957–970.

Orita M, Iwahana H, Kanazawa H, Hayashi K and Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc Natl Acad Sci U S A 86: 2766–2770.

Park SDE (2001) The Excel Microsatellite Tool Kit. Website: http://animalgenomics.ucd.ie /sdepark/ms-toolkit/ [accessed April 2008] .

Ramli US, Salas JJ, Quant PA and Harwood JL (2005) Metabolic control analysis reveals an important role for diacylglycerol acyltransferase in olive but not in oil palm lipid accumulation. FEBS J 272: 5764–5770.

Roesler K, Shintani D, Savage L, Boddupalli S and Ohlrogge J (1997) Targeting of the *Arabidopsis* homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. Plant Physiol 113: 75–81.

Scalfi M, Troggio M, Piovani P, Leonardi S, Magnaschi G, Vendramin GG and Menozzi P (2004) A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.) Theor Appl Genet 108: 433–441.

Shappley ZW, Jenkins JN, Watson CE, Kahler AL and Meredith WR (1996) Establishment of molecular markers and linkage groups in two F2 populations of Upland Cotton. Theor

Appl Genet 92: 915–919.

Smooker AM, Wells R, Morgan C, Beaudoin F, Cho K, Fraser F and Bancroft I (2011) The identification and mapping of candidate genes and QTL involved in the fatty acid desaturation pathway in *Brassica napus*. Theor Appl Genet 122: 1075–1090.

Tamura K, Yonemaru J, Hisano H, Kanamori H, King J, King IP, Tase K, Sanada Y, Komatsu T and Yamada T (2009) Development of intron-flanking EST markers for the Lolium/Festuca complex using rice genomic information. Theor Appl Genet 118: 1549–1560.

Taylor DC, Katavic V, Zou J, MacKenzie SL, Keller WA, An J, Friesen W, Barton DL, Pedersen KK, Giblin EM, *et al*. (2001) Field-testing of transgenic rapeseed cv.Hero transformed with a yeast *sn-2* acyltransferase results in increased oil content, erucic acid content and seed yield. Mol Breed 8: 317–322.

Taylor DC, Zhang Y, Kumar A, Francis T, Giblin EM, Barton DL, Ferrie JR, Laroche A, Shah S, Zhu W, *et al.* (2009) Molecular modification of triacylglycerol accumulation by over-expression of DGAT1 to produce canola with increased seed oil content under field conditions. Botany 87: 533–543.

Thamarus K, Groom K, Bradley A, Raymond CA, Schimleck LR, Williams ER and Moran GF (2004) Identification of quantitative trait loci for wood and fibre properties in two full-sib properties of *Eucalyptus globulus*. Theor Appl Genet 109: 856–864.

Thomas RI, Watson L and Hardon JJ (1969) Inheritance of some components of yield in the Deli *dura* variety of oil palm. Euphytica 18: 92–100.

Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Wegrzyn JL, Eckert AJ, Choi M, Lee JM, Stanton BJ, Sykes R, Davis MF, Tsai CJ and Neale DB (2010) Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. New Phytol 188: 515–532.

Wei H, Fu Y and Arora R (2005) Intron-flanking EST-PCR markers: from genetic marker development to gene structure analysis in *Rhododendron*. Theor Appl Genet 111: 1347–1356.

Yang X, Guo Y, Yan J, Zhang J, Song T, Rocheford T and Li JS (2010) major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. Theor Appl Genet 120: 665–678.

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, *et al.* (2008) A phenyalanine in DGAT is a key determinant of oil content and composition in maize. Nat Genet 40: 367–372.