Candidate genes linked to QTL regions associated with fatty acid composition in oil palm

Abstract

The present study searched for candidate genes in five linkage groups (LGs) - T2, T3, OT4, OT6 and T9 hosting the QTLs associated with iodine value (IV) and fatty acid composition (FAC) in an oil palm interspecific hybrid population. Each of the five LGs was successfully anchored to its corresponding chromosomal segment where, a wider repertoire of candidate genes was identified. This study further revealed a total of 19 candidate genes and four transcription factors involved in biosynthesis of fatty acids, lipids (including triacylglycerol) and acetyl-CoA, glycosylation and degradation of fatty acids. Their possible involvement in regulating the levels of saturation are discussed. In addition, 22 candidate genes located outside the QTL intervals were also identified across the interspecific hybrid genome. A total of 92 SSR markers were developed to tag the presence of these candidate genes and 50 were successfully mapped onto their respective positions on the genome. The data obtained here complements the previous studies, and collectively, these QTL-linked candidate gene markers could help breeders in more precisely selecting palms with the desired FAC.

Keywords

Elaeis guineensis, interspecific hybrid, fatty acid biosynthesis, triacylglycerol biosynthesis, transcription factors, genetic linkage map

Abbreviations

4CLL1	4-coumarate--CoA ligase 1
AACT	acetoacetyl-CoA thiolase
acbd4	acyl-CoA-binding domain-containing protein 4
ACX4	acyl-CoA oxidase 4
C14:0	myristic acid
C16:0	palmitic acid
C16:1	palmitoleic acid
C18:0	stearic acid
C18:1	oleic acid
C18:2	linoleic acid
CHR	pseudo-chromosome
cM	centiMorgan
CP	cross pollinator
CPO	crude palm oil
CTAB	cetyltriammonium bromide
DNA	deoxyribonucleic acid
EG5	*E. guineensis* genome build
ER	endoplasmic reticulum
FA	fatty acid
FAC	fatty acid composition
FAD3/7/8	omega-3 fatty acid desaturase

FBA	fructose-bisphosphate aldolase
Fwd	forward primer
GLABRA	homeobox protein GLABRA
GPAT3	glycerol-3-phosphate acyltransferase 3
GPDH	glycerol-3-phosphate dehydrogenase
HAGH	hydroxyacyl glutathione hydrolase
HD-Zip	TF homeobox-leucine zipper protein ATHB-13
IV	iodine value
KAR	beta-ketoacyl-ACP reductase
KASII, III	beta-ketoacyl-ACP synthases II, III
LG	linkage group
LIPT2	triacylglycerol lipase 2
LOD	logarithm of odds
LPAAT1	lysophosphatidic acid acyltransferase 1
M13	primer 5'CACGACGTTGTAAAACGAC3'
MACP/MAT	malonyl-CoA:ACP transacylase
MAS	marker-assisted selection
MDH	malate dehydrogenase
MYB	TF myb family PHL8
NCBI	National center for biotechnology information
OxG	the *E. oleifera* x *E. guineensis* interspecific hybrid population
OEP163	outer envelope pore protein 16-3
OTE/FATA	oleoyl-ACP thioesterase
PATE/FATB	palmitoyl-ACP thioesterase
PVE	phenotypic variation explained
QTLs	quantitative trait loci
rf	recombination frequency
Rvs	reverse primer
SAD	stearoyl-ACP desaturase
sn	stereospecific number
SNP	single nucleotide polymorphism
SSR	simple sequence repeats
T128	Nigerian E. guineensis paternal palm
TAG	triacylglycerol
TCP15	TF TCP15
TF	transcription factor
TPE	tris-phosphate buffer
UGT	UDP-glycosyltransferase
Uniprot	Universal protein resource database
UP1026	Colombian E. oleifera maternal palm

Introduction

The oil palm industry is a major contributor to the global vegetable oils and fats market. The production of palm and palm kernel oil is at about 75.2 million tonnes accounting for almost one-third of the world's vegetable oil production (Kushairi et al. 2018). Interestingly, over 84.0 % of production is in South-east Asia (USDA 2017). The high demand for palm oil reflects its position as the most consumed vegetable oil, with India, China and the European Union among the main importing countries (Index Mundi 2016; Kushairi et al. 2018). Despite concerns about sustainable practices, especially in the European Union which the industry is addressing aggressively, the supply of palm oil needs to rise in order to meet increasing demand from the growing population worldwide.

Palm oil is produced and accumulates in the fruit mesocarp tissue and is referred to as crude palm oil (CPO). The physical and chemical characteristics (e.g. melting, crystallization and morphology) of CPO are mainly attributed to its fatty acid composition (FAC). In the commercial *Elaeis guineensis* CPO, FAC comprises a balanced combination of saturated and unsaturated fatty acids (FAs). The saturated FAs consist primarily of palmitic (C16:0; ~ 44.0 %) and stearic (C18:0; ~ 4.5 %) acids whereas the unsaturated FAs consist mainly of oleic (C18:1; ~ 39.0 %) and linoleic (C18:2; ~ 10.0 %) acids. The iodine value (IV) which measures the total level of unsaturation is on average about 53.0 in commercial materials. In comparison, CPO from *E. oleifera* has much higher levels of unsaturated FAs (IV: 70.0 – 93.0) due to high levels of C18:1 (ranges from 47.0 – 69.0 %) and C18:2 (ranges from 2.0 – 22.0 %). In contrast, saturated FAs in the *E. oleifera* CPO only range from 15.0 – 30.0 % for C16:0 and 0.2 – 2.0 % for C18:0 (Mohd Din et al. 2000; Montoya et al. 2014; Corley and Tinker 2016). Increasing the unsaturated FA levels in commercial *E. guineensis* CPO has advantages, especially for marketing palm oil in temperate countries. As such, conventional breeding programmes have been directed at selecting high IV oil palm. A sure way to achieve this is via interspecific hybrid breeding, where there is a desire for the unsaturated characteristic from *E. oleifera* to be introgressed into the commercial planting materials. Interspecific breeding crosses have been created using selected high IV *E. oleifera* palms (> 70.0) and Nigerian *E. guineensis* palms (~ 64.0) which appears close to the upper limit that can be achieved in pure *E. guineensis* materials. Early results showed an additive effect in the interspecific hybrids, which possess an intermediate level of unsaturated FAs compared to both the *E. oleifera* and *E. guineensis* parental palms (Rajanaidu et al. 1990; Rajanaidu et al. 2000; Corley and Tinker 2016).

The application of marker-assisted selection (MAS), especially in interspecific hybrid breeding programmes, can expedite the introgression of unsaturated FAs into elite *E. guineensis* line. In this respect, a number of quantitative trait loci (QTLs) affecting FAC have been previously identified by Singh et al. (2009), Montoya et al. (2013, 2014) and Ting et al. (2016). Markers associated with these QTLs can help breeders in selecting palms with desired FAC, at least in specific genetic backgrounds. In addition to the identification of QTL-linked markers, efforts are also focussed on identifying the genes responsible for the variation in FAC as well as other important economic traits (e.g. yield) in oil palm (Kalyana Babu et al. 2020; Xia et al. 2019; Ting et al. 2018; 2016). The availability of the oil palm genome sequence (Singh et al. 2013) can facilitate the identification of the genes. In the present study, the *E. guineensis* genome build 5 (EG5) successfully revealed a number of important genes and transcription factors (TFs) involved in biosynthesis of FAs and triacylglycerols (TAGs) within the QTL intervals, associated with FAC in an *E. oleifera* x *E. guineensis* (OxG) interspecific hybrid population (Ting et al. 2016). The authors had identified a total of 12 QTLs distributed across six linkage groups (LGs) - OT1, T2, T3, OT4, OT6 and T9 that were linked to IV, myristic acid (C14:0), C16:0, palmitoleic acid (C16:1), C18:0, C18:1 and C18:2. However, previous search for candidate genes was restricted to the QTL intervals in LGOT1. The method was efficient at revealing potential regulatory genes and as such, a similar approach was extended in the present study to mine for candidate genes from QTL intervals on the five other identified LGs. Interestingly, a number of the genes identified in the QTL intervals were similar to those described in other independent studies as

regulating the synthesis of FAs and TAGs in the oil palm mesocarp (Sambanthamurthi et al. 1999; Tranbarger et al. 2011; Bourgis et al. 2011; Dussert et al. 2013; Guerin et al. 2016).

Materials and Methods

The OxG mapping population

The OxG mapping population is an oil palm interspecific hybrid cross between the maternal Colombian *E. oleifera* (coded as UP1026) and a paternal Nigerian *E. guineensis tenera* (coded as T128). The OxG cross consists of 108 F_1 hybrid progenies. F_1 mapping populations have been routinely utilized in genetic linkage and QTL analysis of important economic traits in oil palm as reported by Ong et al. (2019), Bai et al. (2017, 2018), Seng et al. (2016), Lee et al. (2015) and Jeennor and Volkaert (2014).

The spear leaves were sampled and stored at -80 ^{o}C. The frozen leaves were ground into powder in liquid nitrogen and DNA extraction was carried out using the modified cetyltriammonium bromide (CTAB) method (Doyle and Doyle, 1990). DNA purity was assessed using a NanoDrop spectrophotometer (NanoDrop Technologies Inc. DE) and an A260/A280 OD ratio of at least 1.8 was considered acceptable. Quality of the extracted DNA was further verified by comparing DNA digested with *EcoR*I and *Hae*III with undigested DNA on a 0.9 % agarose gel in 1X TPE buffer (90mM tris-phosphate buffer and 2mM EDTA pH 8.0) after electrophoresed at 80 – 100V for 3 hours. The DNA was diluted to 50 ng/uL for genotyping with simple sequence repeats (SSR) markers.

Mining candidate genes and development of candidate markers

Sequence information of the single nucleotide polymorphism (SNP) markers linked to QTLs for IV and FAC (reported by Ting et al. 2016) was downloaded from the publicly accessible Genomsawit database at http://genomsawit.mpob.gov.my. The QTL linked SNP markers were then mapped to the published oil palm reference genome (EG5) (Singh et al. 2013) using BLASTN (Altschul et al. 1997) based on sequence similarity < 1e-5. Markers linked to QTLs and candidate genes reported by Bourgis et al. (2011), Montoya et al. (2013) and Jeennor and Volkaert (2014) were also mapped to EG5. Subsequently, the genomic sequences of the entire chromosomal fragment corresponding to each QTL interval were extracted from EG5 and searched for significant homology (BLASTN and BLASTX) to known genes of interest in the National Center for Biotechnology Information (NCBI) databases (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Relevant information associated with the biological functions of the genes and TFs was obtained from published literature and the Universal Protein Resource (Uniprot) database (http://www.uniprot.org/uniprot/). Genes and TFs involved in regulation of biosynthesis of FAs and TAGs, glycolysis and other possible influential factors were shortlisted as candidates.

The identified candidate genes and TFs were mined for SSRs of various repeats (e.g. mono- di-, tri-, tetra-, penta- and hexa-nucleotides) using default parameters in the MIcroSAtellite identification tool (Thiel et al. 2003). One to five SSRs were selected for each candidate gene and forward (Fwd) and reverse (Rvs) primers (18 – 24 nucleotides) were designed using the default parameters in PRIMER 3 (http://www-genome.wi.mit.edu/genome_software/other/primer3.html). A common M13-tail (5'CACGACGTTGTAAAACGAC3') was attached to the forward primer (Fwd 5'-M13) whereas, another M13 primer was attached to each of the fluorescent dyes (FAM-, NED-, PET- and VIC-M13). Nomenclatures sPSc and _oSSR were used for these candidate SSR markers.

Candidate SSR marker analysis

The genotyping of the SSR markers on the 108 F_1 hybrid progenies and two parental palms was carried out as described previously (Ting et al. 2013, 2014 and 2016). PCR amplification of each SSR marker was carried out in a 10.0 uL mixture containing genomic DNA (50ng/uL), Fwd 5'-M13 primer (2.5 uM), Rvs primer (2.5 uM), one fluorescent dye-M13 (2.5 uM), 1X PCR buffer (NEB, USA), 2 mM of each dNTP (NEB, USA) and 0.5 U *Taq* DNA polymerase (NEB, USA). The PCR conditions were as follows: 95 ºC for denaturation (3 min); 35 cycles consisting of 95 ºC (30 s), 52 – 56 ºC (depending on primers, 30 s) and 72 ºC (30 s) and a final extension at 72 ºC (5 min). Prior to fragment analysis, multiplexing of four to eight PCR products was carried out, depending on the sizes of the expected amplicons. PCR fragments were analysed using capillary electrophoresis and subsequently detected using an ABI3730XL DNA analyser (Applied Biosystems, USA). Sizing and scoring of the SSR alleles were executed using the GeneMapper® 4.1 software (Applied Biosystems, USA).

The genotype profile of the markers was determined as originally described by Billotte et al. (2005). The four segregation profiles that were observed in the OxG mapping population previously (Ting et al. 2016) are illustrated in Online Resource 1: Fig. 1. For profile 1, a polymorphic locus is inherited from one of the parental palms and was scored as *ab* and *aa* for the heterozygous and homozygous genotypes, respectively, with an expected ratio of 1:1. For profile 2, polymorphism involved two common segregating alleles (observed as *ab* in both parents) and was scored as *aa*, *ab* and *bb* in the progenies with the expected ratio of 1:2:1. For polymorphisms that involved three co-segregating alleles (profiles 3), the parental genotypes were scored as *ab* and *ac* whereas, the progenies were scored as *aa*, *ab*, *ac* and *bc*, which are expected to fit a 1:1:1:1 ratio. Finally, for profile 4, the four co-segregating alleles were scored as *ab* and *cd* in the two parents and were expected to segregate as *ac*, *bc*, *ad* and *bd* in the progenies, also in a 1:1:1:1 ratio.

Mapping candidate SSR markers to the OxG linkage map and QTL analysis

The F_1 interspecific hybrid population was analyzed essentially as a pseudo-testcross (Grattapaglia and Sederoff 1994). The SSR loci coded as 'cross pollinator' (CP) were incorporated into the existing data set (Ting et al. 2016) and linkage phases between the SSR alleles were determined using JoinMap® 4.1 (van Ooijen 2006). Segregation of the SSR marker alleles according to expected Mendelian ratios was evaluated using a built-in chi-square analysis and severely distorted markers ($p < 0.0001$) were excluded from linkage analysis. The existing OxG linkage map was used as the backbone in the *Start Order* tabsheet. The new SSR markers were integrated using the maximum likelihood (ML) method and markers were grouped at a recombination frequency (rf) threshold of ≤ 0.2. The rf between markers was transformed into map distance in centiMorgans (cM) using the Haldane mapping function. Markers with a nearest neighbour stress value > 4.0 cM were discarded from each LG and the contribution of each marker to the nearest neighbour fit was also inspected in order to get the best possible map order.

The QTL analysis was performed using three separate methods, namely Interval Mapping (IM), the Multiple-QTL Model (MQM) and the Kruskal-Wallis non-parametric ranking tests where all three methods were implemented via MapQTL 6 (van Ooijen 2009). The logarithm of odds (LOD) thresholds for declaring a significant QTL at genome-wide (GW) and chromosome-wide (CW) in the IM and MQM methods were determined by permutation tests (1,000 times) on the phenotypic data, also implemented via MapQTL 6. Only QTLs that were consistently observed in all three methods were considered significant in this study.

Results

Candidate genes underlying QTL intervals

Eleven QTLs on LGs T2, T3, OT4, OT6 and T9 were previously associated with C14:0 (≥ LOD3.7, GW), C16:1 (≥ LOD4.3, GW), C18:0 (≥ LOD3.8, GW) and C18:2 (≥ LOD1.7 on LGT3 and LOD2.7 on LGOT6, CW) in the OxG mapping population (Ting et al. 2016). Each of the LGs was successfully anchored to the oil palm EG5 genome build on its corresponding pseudo-chromosome (CHR) – 8, 14, 2, 7 and 13, respectively. Markers associated with QTLs, including a number of candidate genes reported previously (Bourgis et al. 2011; Montoya et al. 2013 and Jeennor and Volkaert 2014) were also aligned to CHR 2, 5, 6, 7, 8, 10, 13, 15 and 16. In total 45 candidate genes including four TFs, were identified, most of which are directly involved in biosynthesis of FAs and TAGs (Fig. 1, Table 1 and Online Resource 2: Table 1).

Integration of candidate gene-linked SSR markers into existing genetic map

A total of 45 FAC and TAG related genes were utilized for development of SSR markers. One to five SSR markers were selected for each candidate gene, resulting in 92 SSR primer-pairs being designed (Table 1). Genotyping of these SSR markers in the OxG mapping population resulted in 50 polymorphic SSR markers, of which 47 were scored according to profile 1 (45 inherited from T128 and two inherited from UP1026) and three were scored as having profile 4. The allelic segregation ratios for the 50 SSR markers met the expected Mendelian ratios at $p \geq 0.0005$. These 50 markers were then included into the existing marker data set for constructing the genetic map and all were successfully mapped into the existing OxG genetic map (Fig. 2 and Online Resource 3: Fig. 2).

Mapping of candidate gene markers to the respective QTLs

The candidate gene markers identified in this study were successfully mapped back to the respective QTL intervals in LGs T2, OT3, OT4, OT6 and T9. In LGT2, the 1.0 cM interval (61.2 – 62.1 cM) related to QTL for C14:0 was mapped to CHR08, but clear candidate genes related to biosynthesis of FAs or TAGs were not detected in the QTL interval. Therefore, the search was extended towards the left and right of the QTL interval and an *oleoyl-ACP thioesterase* (*OTE*/*FATA*), *stearoyl-ACP desaturase* (*SAD*) and *hydroxyacyl glutathione hydrolase 2* (*HAGH*) gene were detected flanking both sides of the interval (Fig. 2). A similar chromosomal region corresponding to the QTL for C14:0 was also reported in an interspecific BC_1 mapping population (Montoya et al. 2013), as determined by the common markers (mEgCIR3649, mEgCIR3282 and mEgCIR0800) mapped on both the studies (Fig. 2).

In LGT9, the QTLs for C14:0, C16:1 and C18:0 were located at regions spanning 17.2 – 32.6, 2.9 – 32.6 and 13.4 – 24.8 cM, respectively (Ting et al. 2016). The QTLs for C14:0 and C18:0 were also found to be located very close to that reported previously in a *tenera* x *dura* mapping population (Montoya et al. 2014). This was revealed by two common SSR markers, namely mEgCIR3592 and mEgCIR3787 that were located within/near the similar QTLs reported by Montoya et al. (2014) (Fig. 2). Taking the regions containing all three QTLs, an interval ranging from 2.9 – 32.6 cM was examined, which identified four potential genes and three TFs. The four candidate genes were *beta-ketoacyl-ACP synthases II and III* (*KASII*, *KASIII*), *malate dehydrogenase* (*MDH*) *and acetoacetyl-CoA thiolase* (*AACT*) whereas, the TFs were *myb family PHL8* (*MYB*), *TCP15* (*TCP15*) and *homeobox-leucine zipper protein ATHB-13* (*HD-Zip*) (Table 1). Three SSR markers, namely sPSc00554, sPSc00571 and sPSc00574, associated with the candidate genes *AACT*, *TCP15* and *HD-Zip* respectively, were successfully mapped within the QTL interval. The *TCP15* linked sPSc00571 was mapped

closest to the QTL peak (LOD5.2 – 12.7) and explained 24.5 – 49.8 % the variation for C14:0, C16:1 and C18:0. The LOD score and phenotypic variation explained (PVE) after mapping of the candidate gene markers were higher than that observed before fine-mapping (LOD4.5 – 10.7 and 21.9 – 44.2 % PVE) with the candidate markers (Ting et al. 2016). Unfortunately, the SSR markers developed for *KASII*, *KASIII*, *MDH* and *MYB*, were not polymorphic (*), and thus could not be mapped onto LGT9. They were placed on the LG based on their relative order compared to other markers (and genes) in CHR08, but their exact map positions could not be determined (Fig. 2).

On LGOT4 (CHR02), *UDP-glycosyltransferase* (*UGT*) was found located underlying the QTL peak for C18:0, defined by SNPM00121 (LOD5.0). *UGTs* are not involved in FA or TAG biosynthesis. They however, catalyse the covalent addition of sugars to a wide range of lipophilic molecules by transferring the glycosyl group from nucleoside diphosphate-activated sugars (e.g. UDP-sugars), and control the levels of many signalling molecules. The molecules include a broad array of hormones (including phytohormones), secondary metabolites and xenobiotics for maintaining good growth and development in plants (Ross et al. 2001; Barvkar et al. 2012; Ostrowski and Jakubowska 2014). Other genes from the QTL interval were *fructose-bisphosphate aldolase* (*FBA*), *outer envelope pore protein* 16-3 (*OEP163*), *4-coumarate--CoA ligase 1* (*4CLL1*) and *beta-ketoacyl-ACP reductase* (*KAR*). Among these, *KAR* is involved in the *de novo* FA chain elongation cycle and the SSR marker associated with this gene, sPSc00584A (LOD4.5) mapped closest to *UGT*. The PVE explained by sPSc00584A at 22.5 % was similar to that observed for SNPM00121 (23.8 %). In fact, another important FA gene namely *stearoyl-ACP desaturase* (*SAD*) which converts C18:0- to C18:1-ACP was identified on LGOT4 but, at a distance of about 36.0 cM from the QTL interval (Fig. 2).

Two putative QTLs for C18:2 were reported on LGs OT3 and OT6, at intervals 46.9 – 65.2 and 38.9 – 54.5 cM, respectively. Three candidate markers - sPSc00664, sPSc00665 and sPSc00666 associated with *omega-3 fatty acid desaturase* (*FAD3/7/8*), were developed within the QTL interval at LGOT3 (corresponded to CHR14). All three SSR markers were successfully mapped back to the QTL peak. The second gene within close proximity was *acyl-CoA oxidase 4* (*ACX4*) which is involved in the peroxisomal degradation of short-chain FAs (C4:0 – C8:0) during beta-oxidation. This process also recycles acetyl-CoA as a carbon and energy source for FA synthesis and plant growth (Poirier et al. 2006; Goepfert and Poirier 2007). Other genes involved in FA and TAG synthesis activities were also detected at the QTL interval in LGOT3. These include *lysophosphatidic acid acyltransferase 1* (*LPAAT1*), *acyl-CoA-binding domain-containing protein 4* (*acbd4*) and *glycerol-3-phosphate acyltransferase 3* (*GPAT3*). A candidate SSR marker, sPSc00694 was developed and mapped close to *LPAAT1* whereas, SSR markers for *acbd4* and *GPAT3* were not polymorphic (*) and could not be mapped (Table 1 and Fig. 2).

For QTL-C18:2 on LGOT6 which corresponded to CHR07, the QTL interval hosted a *palmitoyl-ACP thioesterase* (*PATE/FATB*) gene. None of the markers developed from the gene were polymorphic (*). However, two important FA genes – *SAD* and *OTE*/*FATA* were identified at a distance of 4.4 cM from the QTL interval.

In addition to mining candidate genes from the QTLs identified by Ting et al. (2016), a number of candidate SSR markers were also developed for other published QTLs/genes for FAC and oil yield in oil palm. Other than the five LGs mentioned above, these markers were also mapped onto LGs T5, OT10, OT12, T14, OT15 and T16 (Table 1, Fig. 2 and Online Resource 3: Fig. 2). As an example, the *homeobox protein GLABRA gene* (*GLABRA*), *glycerol-3-phosphate dehydrogenase* (*GPDH*), *plastid-linoleate desaturase* (*FAD7*), *malonyl-CoA:ACP transacylase* (*MACP/MAT*) and *triacylglycerol lipase 2* (*LIPT2*) genes associated with QTLs for oil-to-fruit and oil-to-bunch traits (Jeennor and Volkaert 2014) were successfully mapped to the present LGOT12. The candidate SSR markers developed from these genes mapped close to each of the respective genes as follows: sPSc00442-*GLABRA*, sPSc00443-*GPDH*, sPSc00445-*FAD7*-sPSc00446, sPSc00447-*MACP*-sPSc00448 and sPSc00449-*LIPT2*.

Discussion

The present study builds on previous efforts in searching for candidate genes in the QTL intervals on LGs T2, T9, OT3, OT4 and OT6 (Ting et al. 2016). The use of SSR markers common to those utilized in other studies revealed that several of these QTLs were located near or within the genomic regions linked to FAC in previous studies using a BC_1 and a *tenera* x *dura* mapping populations (Montoya et al. 2013; 2014). This provided confidence to search for candidate genes within the designated QTL intervals. In this study, there were 92 SSR markers developed from FA and TAG related genes, of which 50 (53.0 %) were informative. The 50 SSR candidate markers followed two of the four segregation profiles observed in the OxG mapping population previously. All 50 were successfully mapped to the expected LG, corresponding to the genomic region from which they were designed, confirming the appropriateness of the mapping methodology applied in this study.

Each of the QTL intervals of interest in this study was successfully anchored to the corresponding pseudo-chromosomes and revealed a number of FA and lipid related genes. From the QTL regions associated with C14:0, C16:1 and C18:0 in LGT9 (CHR13), *MDH and KASIII* were identified. The *MDH* encoding enzyme can be found in a range of subcellular locations (e.g. cytosol and mitochondria) and it catalyses the interconversion of malate to oxaloacetate which subsequently can be converted to form phosphoenolpyruvate or can be oxidized to form pyruvate (Wedding 1989; Minárik et al. 2002). This provides the pyruvate source to initiate the synthesis of FAs. *In vitro* experiments in castor bean demonstrated high FA synthesis rate when malate was provided as a precursor (Smith et al. 1992). The enzyme KASIII forms the acetoacetyl-ACP complex from acetyl-CoA and malonyl-ACP in preparation for FA-chain elongation. Both the MDH and KASIII-catalysed reactions take place at a very early stage even before the FA-chain elongation process starts. This suggests that MDH and KASIII activities are important prior to formation of various FAs and could explain the co-localization of the two genes within the same QTL interval associated with C14:0, C16:1 and C18:0 in LGT9.

Another gene, *KASII*, that plays a critical step in elongating C16:0-ACP to form C18:0-ACP was also detected in LGT9. This is one of the most important enzymatic activities for generating and supplying C18:0 for subsequent desaturation into unsaturated FAs by SAD and FADs. In oil palm, KASII activity was found to be positively correlated with unsaturated FA content. The observed relationship was particularly strong with C18:1 and C18:2, suggesting that increased levels of C18:0-ACP are efficiently converted to C18:1-ACP which subsequently is hydrolysed (Sambanthamurthi et al. 1999). The C18:1 released is activated to C18:1-CoA and channelled to endoplasmic reticulum (ER) for TAG assembly or further desaturated to C18:2 prior to TAG assembly. In contrast, KASII activity was found to be negatively correlated with the saturated FAs (Sambanthamurthi et al. 1999). This was supported by the recent transcriptomic co-expression analysis in oil palm, where lower levels of C16:0 were the result of increased *KASII* expression (Guerin et al. 2016). It has been suggested that lower rates of KASII activity increase accumulation of shorter FA chains such as C14:0- and C16:0-ACPs. The increased accumulation of C16:0-ACP allows C16:0-ACP to be desaturated to form C16:1-ACP. Increased accumulation of C16:0-ACP also results in increased hydrolysis by PATE/FATB, activation into C16:0-CoA and assembly of higher levels of C16:0 into TAG in the ER. In *A. thaliana* and cotton seed, silencing or down-regulating the *KASII* gene has led to two- to six-fold increase in C16:0 (Pidkowich et al. 2007; Liu et al. 2017). In Camelina, suppression of the *KASII* gene also led to higher accumulation of palmitate and further reduction of unsaturated FAs (Hu et al. 2017). This provides support for the involvement of *KASII* in the QTL interval linked to C14:0, C16:1 (produced from C16:0) and C18:0. However, the SSR markers designed to the *KASII* gene did not segregate in the mapping family. It will be interesting to extend the analysis in future to search for polymorphic

SNPs, within or flanking the *KASII* gene. The identification of candidate genes that are required for the initiation of FA synthesis (*MDH and KASIII*) and in the accumulation of unsaturated FAs (*KASII*) within the QTL interval in LGT9, suggests that it is an important genomic region influencing FAC in interspecific hybrids.

A number of genes encoding enzymes that show substrate specificity have also been identified in the confidence intervals of QTLs, in accordance to their respective FA preferences. These include *FAD3/7/8*, *acbd4* and *LPAAT1* that were associated with QTL for C18:2 in LGOT3 (CHR14). *FAD3/7/8* encodes desaturase activity to convert C18:2 into C18:3 either in the plastids (by FAD7/8) or in the ER (by FAD3) (Song et al. 2004, Yurchenko et al. 2014). The acbd4 binds oleoyl (C18:1)-CoAs with high affinity and transports them from cytosol to ER for further modification of FAs or synthesis of TAGs (Leung et al. 2004; Xiao et al. 2008). Located next to *acbd4* is *GPAT* which encodes the first step of enzymatic acylation to form TAGs in ER. Generally, *GPAT* is known to have preference for saturated FAs, especially towards C16:0-CoA (Griffiths et al. 1988; Griffiths and Harwood 1991; Xu et al. 2009). However, Sambanthamurthi et al. (2000), Manaf and Harwood (2000) and Dussert et al. (2013) suggested that oil palm GPAT can use both saturated and unsaturated acyl-CoAs (including C18:1-CoA) as substrates. Interestingly in *Brassica napus*, GPAT has a wider range of specificity, allowing addition of variety of fatty acyl-CoAs to the stereospecific number 1 (*sn-1*) position of glycerol-3-phosphate (Gly3P) (Larson et al. 2002). The subsequent acylation is catalysed by lysophosphatidic acid acyltransferase, an enzyme encoded by the *LPAAT* gene. In the current QTL interval, *LPAAT1* was identified and interestingly it has been reported to show high specificity towards unsaturated fatty acyl-CoAs such as C18:1-CoA in humans and most plants (Shindou et al. 2009). In oil palm, LPAAT has also been reported to accept C16:0-CoA as the alternative substrate at the *sn-2* position for producing phosphatidate (Sambanthamurthi et al. 2000). The QTL intervals essentially contain genes that regulate both FA synthesis in the plastid and TAG assembly in the ER. As such, the candidate genes and the SSR markers linked to these genes are ideal candidates to further investigate both FA and lipid biosynthesis in independent oil palm populations.

In this study, there were interesting candidate genes identified outside the QTL confidence intervals such as those for C18:0 in LGOT4 (CHR02) and C18:2 in LGOT6 (CHR07). *SAD* was located at a distance of about 36.0 cM from the QTL for C18:0 whereas on LGOT6, *SAD* and *OTE/FATA* were located at a position about 4.4 cM from the QTL for C18:2. *SAD* and *OTE/FATA* encode stearoyl-ACP desaturase and oleoyl-ACP thioesterases A, respectively and these two enzymes have high specificity towards C18 FA-ACPs. In the oil palm mesocarp, SAD modifies C18:0-ACP to C18:1-ACP while, OTE/FATA hydrolyses and releases C18:1 from C18:1-ACP. Detection of candidate genes outside the confidence interval has also been reported for QTLs associated with C18:1 and C18:2 in watermelon seeds (Meru and McGregor 2014). A point to consider is the observation by Raghaven and Collard (2012) that for a small mapping population (< 194 samples), there is a possibility that the QTL detected may actually be several cM away from its actual position. As such, even though the candidate genes were located outside the QTL confidence interval, they remain as good candidates for further evaluation.

Conclusion

The increasing availability of information on gene function and genome sequence data of plant species (including oil palm) that are accessible in public databases facilitated the present study to uncover potential candidate genes associated with fatty acid composition. This further facilitated development of markers closely linked to these candidate genes within the QTL confidence intervals. In this study, the candidate gene approach once again proved very efficient and was

successfully applied to identify candidate genes and transcription factors from the QTL intervals. More importantly, biological functions of these candidate genes provided potential explanations for their possible involvement in the fatty acid and Kennedy pathways for lipid assembly. Both pathways play an important role in determining the levels of saturation and unsaturation in palm oil. The levels of saturation and unsaturation could possibly be regulated by the expression of these genes. More in-depth evaluation e.g. expression and functional studies will be required to confirm the regulatory effects of these candidate genes. This paper presents an atlas of candidate genes which may be involved in the oil saturation differences between the high IV *E. oleifera* and lower IV *E. guineensis*. Introgression of the high IV character into the African oil palm could lead to new markets and applications for palm oil. The current work represents an important step towards realising these objectives.

Declarations

Funding: This study was funded by the Malaysian Palm Oil Board (MPOB).

Conflict of interest: The authors declare that they have no conflict of interest.

Ethics approval: Not applicable.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data: The sequence information for the candidate SSR markers is available at http://genomsawit.mpob.gov.my.

Code availability: Not applicable.

Authors' contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ting Ngoot-Chin. Bioinformatics analysis was supported by Chan Kuang-Lim and the plant materials used in this study was provided by Kandha Sritharan. The first draft of the manuscript was written by Ting Ngoot-Chin and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

List of tables

Table 1. Candidate SSR markers developed for fatty acid and lipid related genes and transcription factors (TFs) identified within or close to the QTL intervals. The putative biological functions for the listed genes and TFs are shown in Online Resource 2: Table 1.

List of figures

Fig. 1. Postulated relationships among the identified candidate genes in relation to fatty acids (FA), lipids and other related biosynthesis pathways.

Fig. 2. Mapping candidate markers for various fatty acid and lipid related genes onto the respective QTL hosting linkage groups (LGs) in OxG. The updated LGs (After fine-mapping, left panel) are aligned to the previous published map (Before fine-mapping, right panel). Candidate SSR markers are indicated in blue; candidate genes and transcription factors are

indicated in red italics; markers located closest to the QTL are in purple and AFLP markers and common SSR markers are in green. The coloured bars indicate the published QTL regions (by Ting et al. 2016). The QTL regions overlapping with and close to that reported previously by Montoya et al. (2103; 2014), are also indicated for LGT2 (CHR08) and LGT9 (CHR13). The QTL chart including the phenotypic variation explained (PVE) at the QTL peak, is also indicated below each LG.

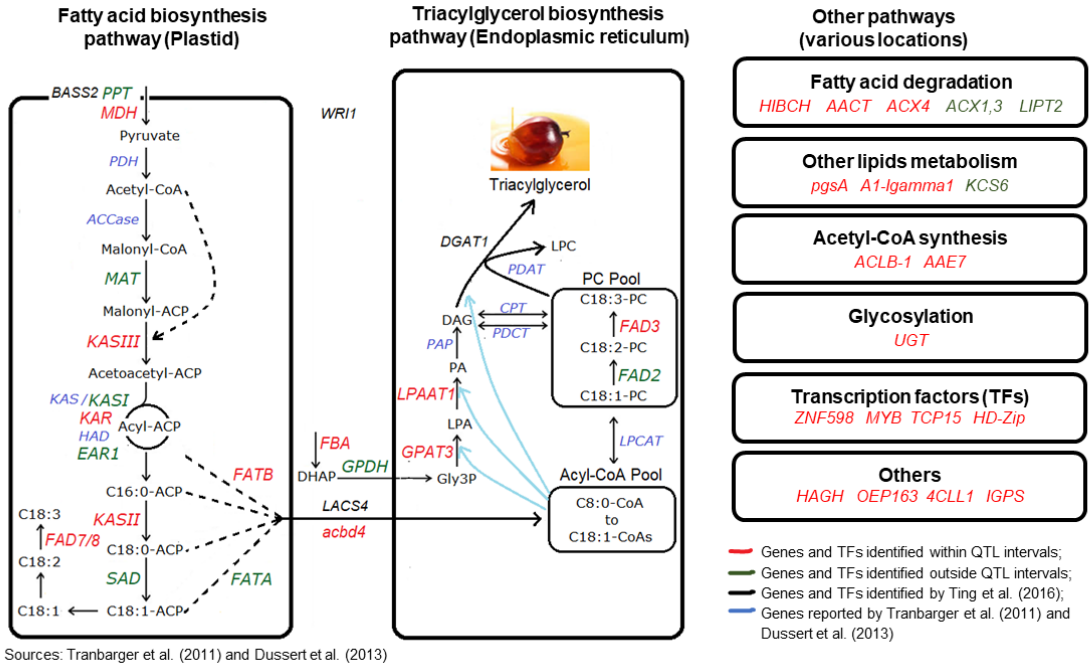
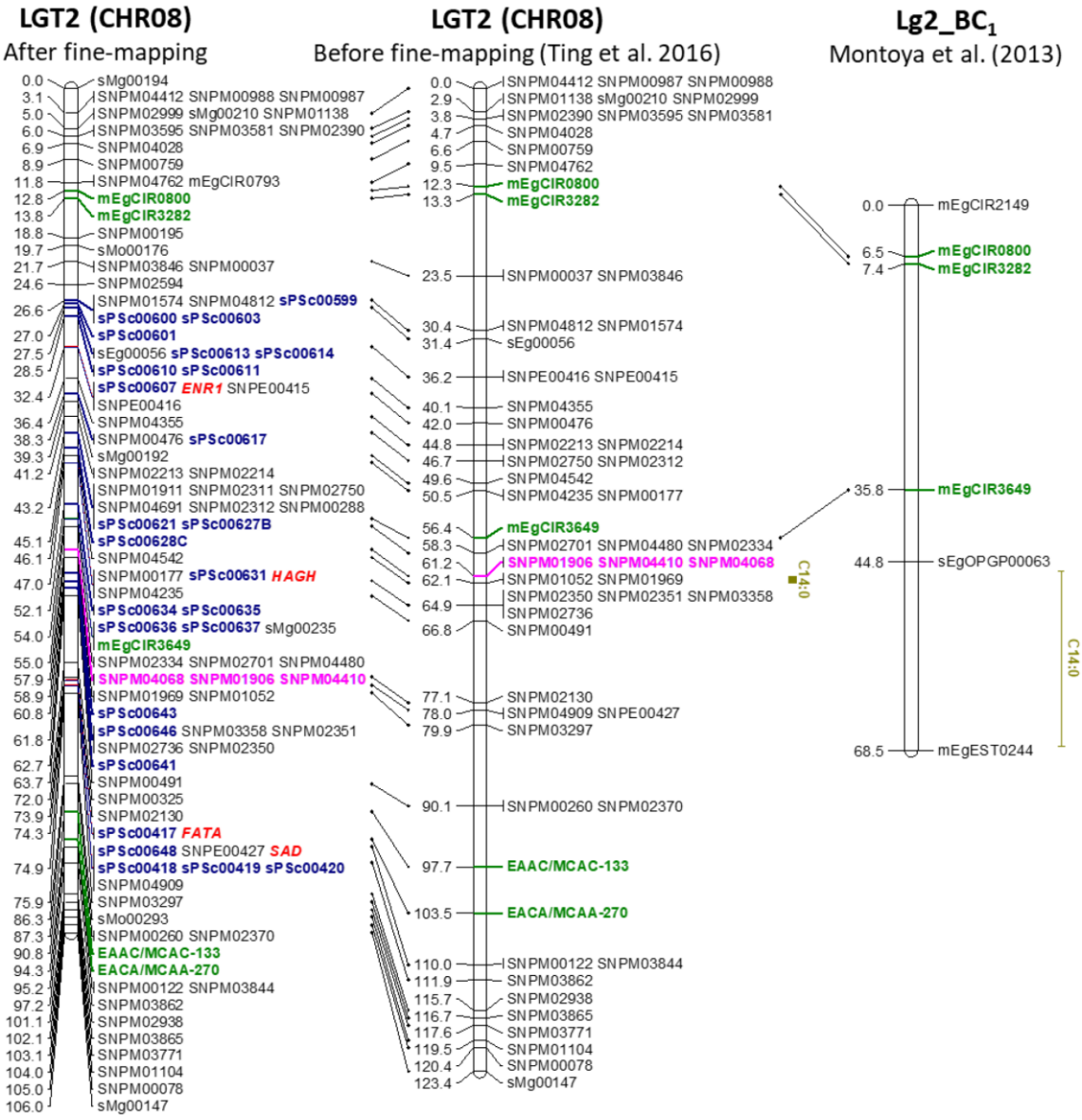
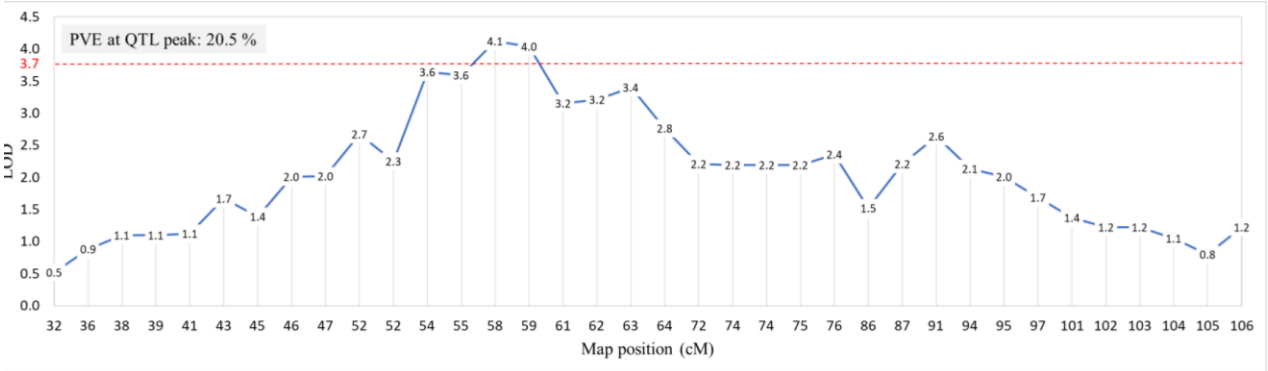


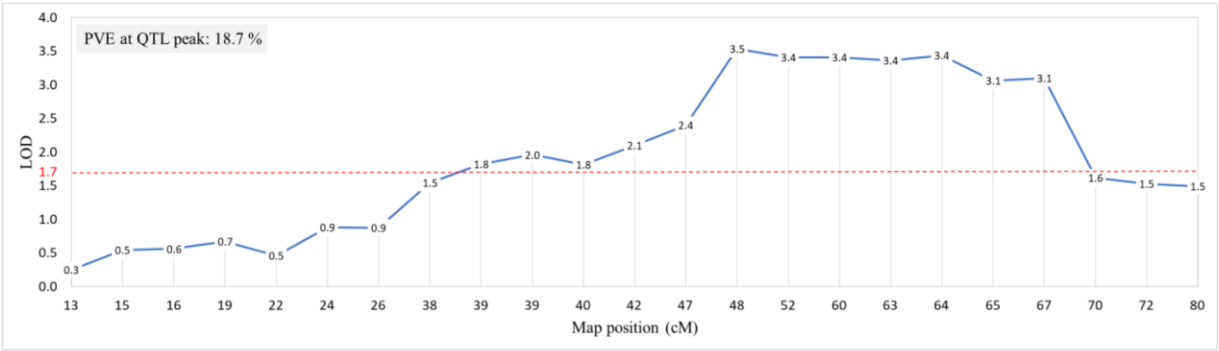
Fig1


LGT2 (CHR08)
After fine-mapping
LGT2 (CHR08)
Before fine-mapping (Ting et al. 2016)
Lg2_BC1
Montoya et al. (2013)
C14:0
C14:0

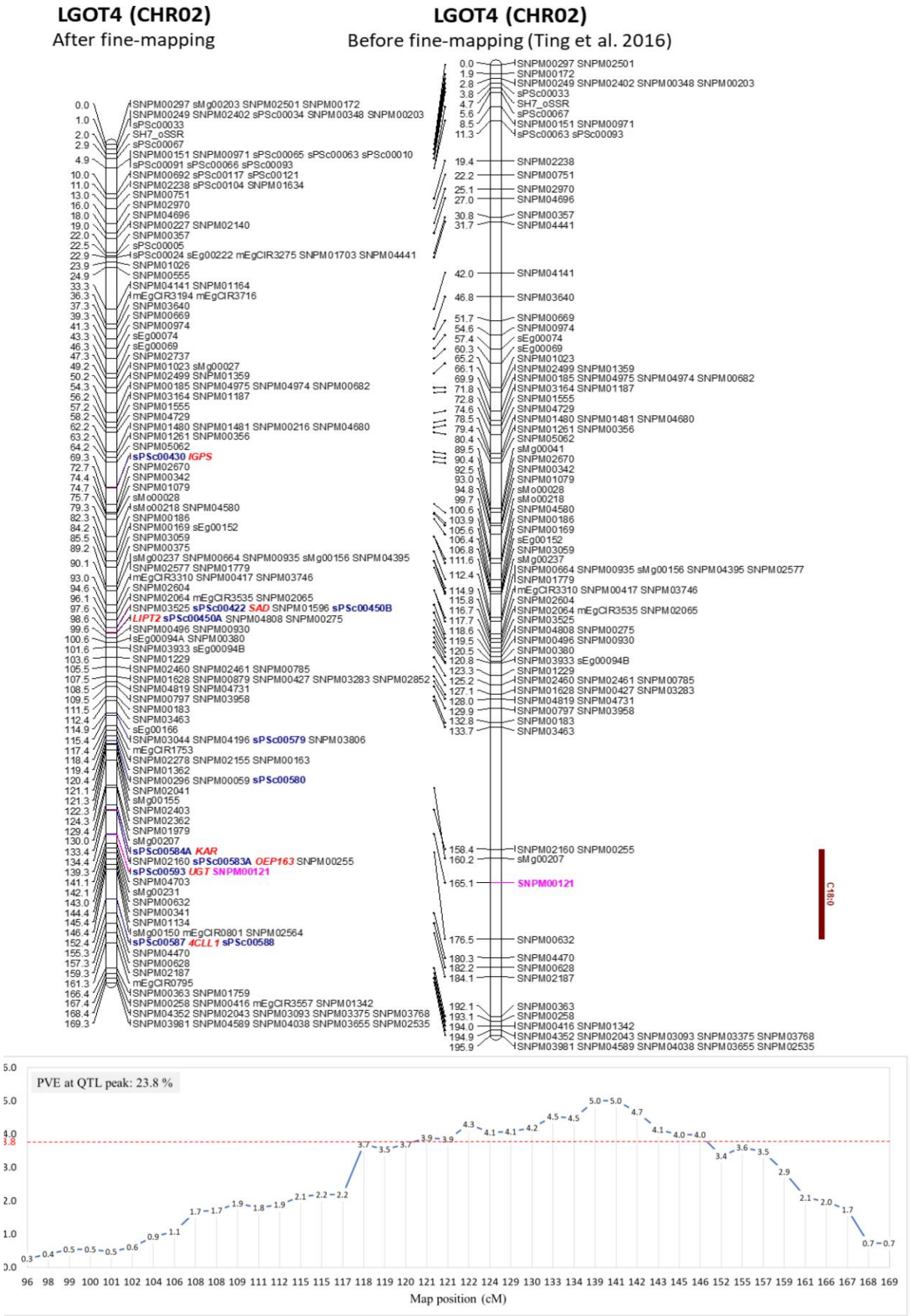

PVE at QTL peak: 20.5 %
LOD
Map position (cM)

427

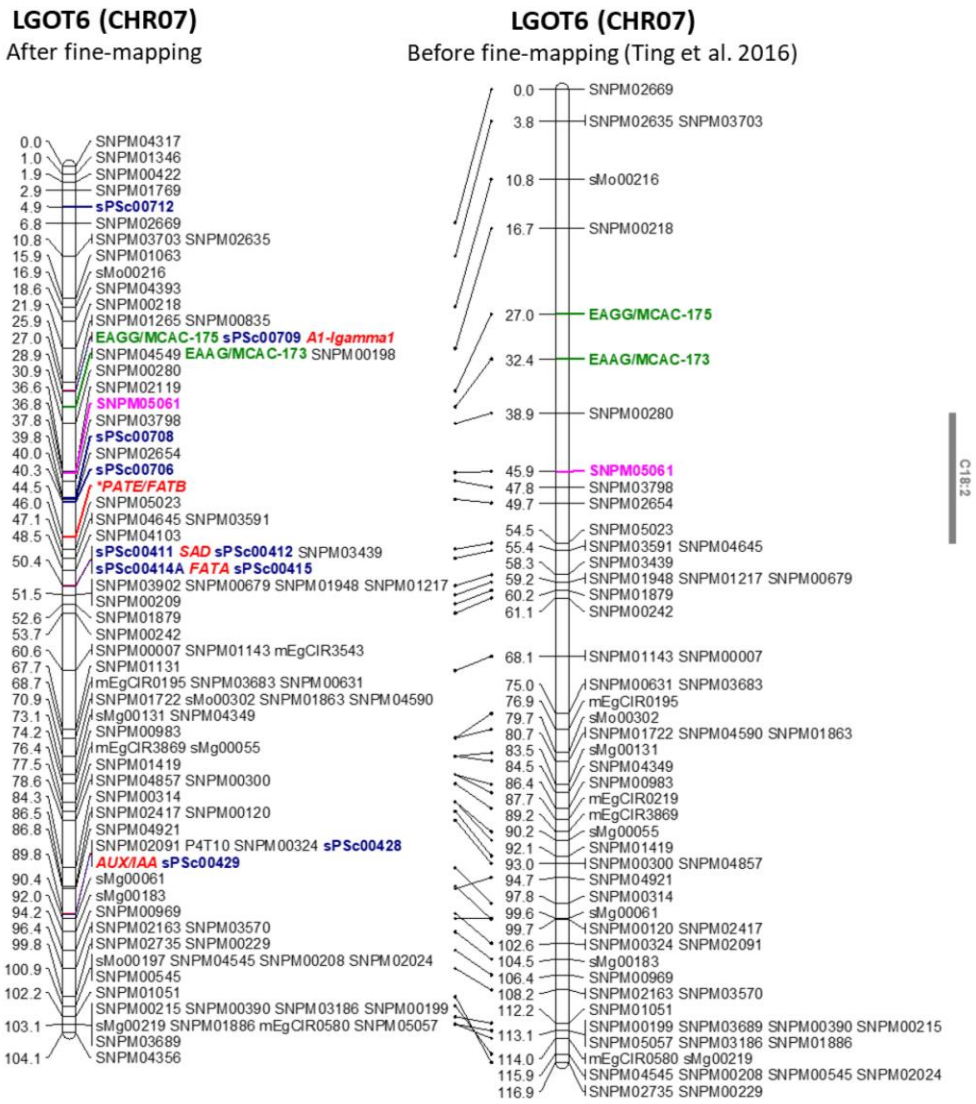

LGOT3 (CHR14)
After fine-mapping
LGT3 (CHR14)
Before fine-mapping (Ting et al. 2016)
0.0 sMg00011
2.0 SNPM04971 SNPM01461
3.8 SNPM02639
4.8 SNPM00748 SNPM03619
7.6 SNPM02349
9.5 SNPM03909
10.5 SNPM00128
11.4 SNPM01679
12.3 SNPM00550
13.3 sMo00204
15.1 SNPM02983
16.1 SNPM00588
18.9 SNPM01211
21.8 SNPM04197 mEgCIR0369 mEgCIR1772 SNPM01899 SNPM03053 SNPM00362
23.7 SNPM02899
25.6 SNPM01030
38.1 SNPM01170
39.1 SNPM02590
39.3 sPSc00691
40.0 sPSc00690 SNPM03409
41.8 sMo00252
46.6 *LPAAT
46.8 sPSc00694 SNPM00561 SNPM04100
47.8 sPSc00695 AAE7 SNPM02538
48.5 *ACLB-1
49.0 *acbd4
50.0 *GPAT3
51.0 *ZNF598
51.6 SNPM00261 SNPM01446
52.5 sPSc00678
52.6 SNPM04312
59.5 sPSc00188A
60.6 sPSc00672 sPSc00671 sMg00217B SNPM03122
63.4 sPSc00664 sPSc00665 FAD3/7/8 sPSc00666 SNPM00462 sPSc00667 ACX4
64.4 sPSc00669B
65.3 sPSc00655 pgsA
67.2 SNPM04474 sPSc00193 SNPM04198
70.0 SNPM01767 sPsc00164
71.9 sPSc00183
79.8 sPSc00652
0.0 SNPM01461
0.9 SNPM02639
1.9 SNPM00748 SNPM03619
4.7 SNPM02349
6.6 SNPM03909
7.6 SNPM00128
8.5 SNPM01679
9.4 SNPM00550
12.3 SNPM02983
13.2 SNPM00588
16.1 SNPM01211
18.9 SNPM01899 SNPM04197 SNPM03053 SNPM00362
20.8 SNPM02899
22.7 SNPM01030
35.3 SNPM01170
36.2 SNPM02590
37.1 SNPM03409
38.1 SNPM02215
39.0 mEgCIR0425
46.0 SNPM00561 SNPM04100
46.9 SNPM02538
50.7 SNPM01446 SNPM00261
63.3 SNPM00462
65.2 SNPM04474 sPSc00193 SNPM04198
68.0 sPsc00164 SNPM01767
C18:2

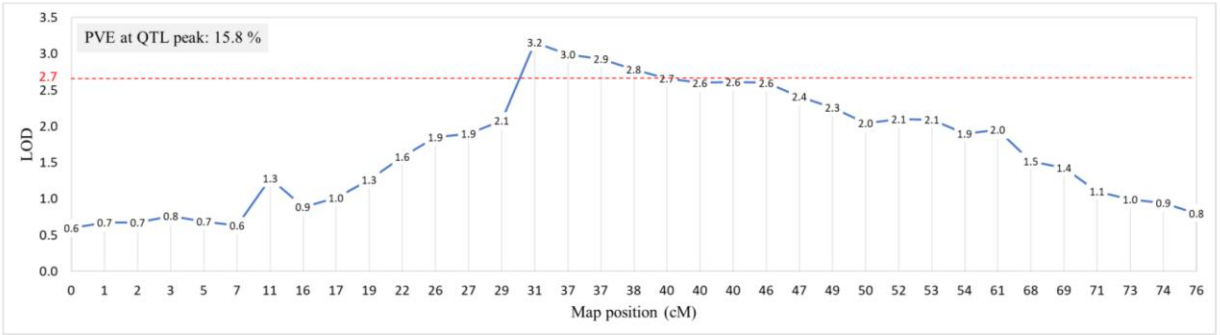

PVE at QTL peak: 18.7 %
LOD
Map position (cM)
0.3 0.5 0.6 0.7 0.5 0.9 0.9 1.5 1.8 2.0 1.8 2.1 2.4 3.5 3.4 3.4 3.4 3.4 3.1 3.1 1.6 1.5 1.5
13 15 16 19 22 24 26 38 39 39 40 42 47 48 52 60 63 64 65 67 70 72 80

428

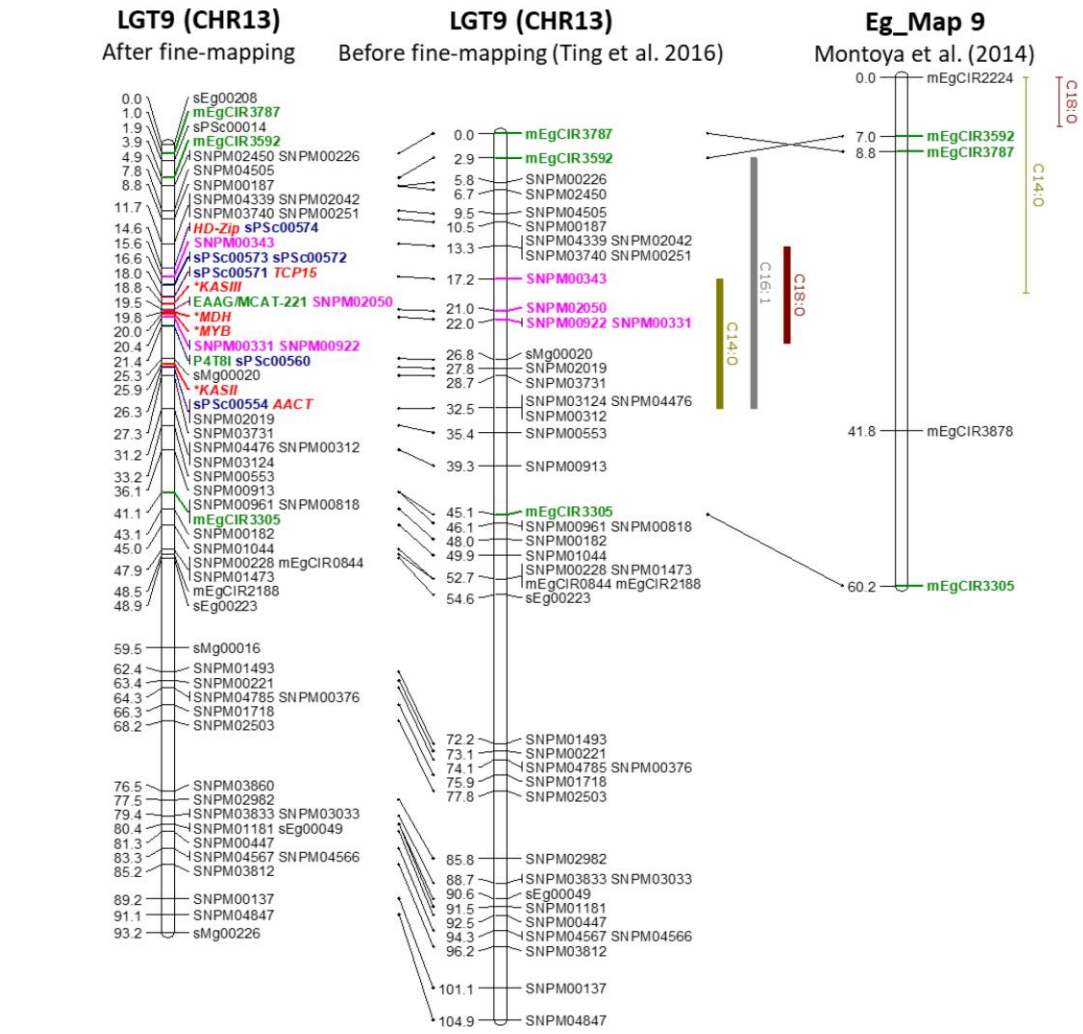

LGOT4 (CHR02)
After fine-mapping
LGOT4 (CHR02)
Before fine-mapping (Ting et al. 2016)
C18:0
PVE at QTL peak: 23.8 %
Map position (cM)

429


LGOT6 (CHR07)
After fine-mapping
LGOT6 (CHR07)
Before fine-mapping (Ting et al. 2016)
C18.2


PVE at QTL peak: 15.8 %
LOD
Map position (cM)

430



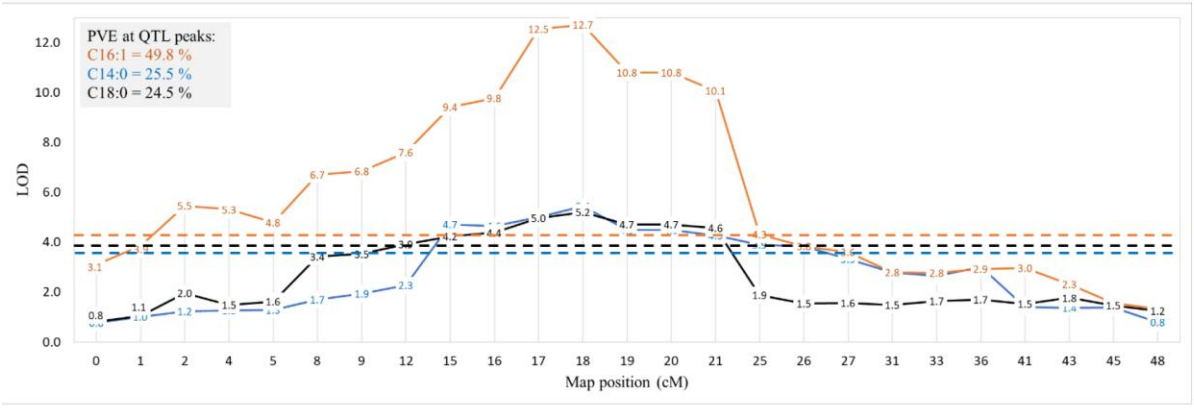


Table 1. Candidate SSR markers developed for fatty acid and lipid related genes and transcription factors (TFs) identified within or close to the QTL intervals. The putative biological functions for the listed genes and TFs are shown in Supplementary Table 1.

Linkage group (LG)	EG5 chromosome (CHR)	QTL[#]	Candidate SSR marker	Gene/TF	Source
T2	8	-	sPSc00417; sPSc00647*; sPSc00648	*Oleoyl-ACP thioesterase (OTE/FATA)*	Bourgis et al. (2011); Montoya et al. (2013); Jeennor and Volkaert (2014); NCBI
		-	sPSc00418; sPSc00419; sPSc00420	*Stearoyl-ACP desaturase (SAD)*	Bourgis et al. (2011); Montoya et al. (2013)
		-	sPSc00607	*Enoyl-ACP reductase I (ENR1/EAR1)*	Jeennor and Volkaert (2014); NCBI
		-	sPSc00631	*Hydroxyacyl glutathione hydrolase 2 (HAGH)*	Singh et al. (2013); NCBI
OT3	14	C18:2	sPSc00654*; sPSc00655	*CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase 2 (pgsA)*	Singh et al. (2013); NCBI
		C18:2	sPSc00664; sPSc00665; sPSc00666	*Omega-3 fatty acid desaturase (FAD3/7/8)*	Singh et al. (2013); NCBI
		C18:2	sPSc00667; sPSc00668*; sPSc00669B	*Acyl-coenzyme A oxidase 4 (ACX4)*	Singh et al. (2013); NCBI
		C18:2	sPSc00679*	*Zinc finger protein 598 (ZNF598)*	Singh et al. (2013); NCBI
		C18:2	sPSc00681*; sPSc00682*	*Glycerol-3-phosphate acyltransferase 3 (GPAT3)*	Singh et al. (2013); NCBI
		C18:2	sPSc00685*; sPSc00686*	*Acyl-CoA-binding domain-containing protein 4 (acbd4)*	Singh et al. (2013); NCBI
		C18:2	sPSc00683*	*ATP-citrate synthase beta chain protein 1 (ACLB-1)*	Singh et al. (2013); NCBI
		C18:2	sPSc00694	*Lysophosphatidic acid acyltransferase 1 (LPAAT1)*	Singh et al. (2013); NCBI
		C18:2	sPSc00695	*Acetate/butyrate--CoA ligase (AAE7)*	Singh et al. (2013); NCBI
OT4	2	C18:0	sPSc00583A	*Outer envelope pore protein 16-3 (OEP163)*	Singh et al. (2013); NCBI
		C18:0	sPSc00584A; sPSc00585*	*beta-ketoacyl-ACP reductase (KAR)*	Singh et al. (2013); NCBI
		C18:0	sPSc00587; sPSc00588	*4-coumarate--CoA ligase 1 (4CLL1)*	Singh et al. (2013); NCBI
		C18:0	sPSc00591*	*Fructose-bisphosphate aldolase (FBA)*	Singh et al. (2013)
		C18:0	sPSc00593; sPSc00594*	*UDP-glycosyltransferase (UGT)*	Singh et al. (2013); NCBI
		-	sPSc00422; sPSc00423*	*Stearoyl-ACP desaturase (SAD)*	Bourgis et al. (2011); Montoya et al. (2013)
		-	sPSc00430; sPSc00431*	*Indole-3-glycerol phosphate synthase (IGPS)*	Jeennor and Volkaert (2014); NCBI

		-	sPSc00450A; sPSc00450B	*Triacylglycerol lipase 2 (LIPT2)*	Jeennor and Volkaert (2014); NCBI
T5	16	-	sPSc00432; sPSc00433	*Acyl-coenzyme A oxidase 1 (ACX1)*	Jeennor and Volkaert (2014); NCBI
OT6	7	C18:2	sPSc00709	*Phospholipase A1-Igamma1 (A1-Igamma1)*	Singh et al. (2013); NCBI
		C18:2	sPSc00696*; sPSc00697*; sPSc00698*	*Palmitoyl-acyl carrier protein thioesterase (PATE/FATB)*	Singh et al. (2013); NCBI
		-	sPSc00411; sPSc00412; sPSc00413*	*Stearoyl-ACP desaturase (SAD)*	Bourgis et al. (2011); Montoya et al. (2013)
		-	sPSc00414A; sPSc00415	*Oleoyl-ACP thioesterase OTE/(FATA)*	Bourgis et al. (2011); Montoya et al. (2013)
T9	13	C14:0; C16:1; C18:0	sPSc00554	*Acetoacetyl-CoA thiolase (AACT)*	Singh et al. (2013); NCBI
		C14:0; C16:1; C18:0	sPSc00556*; sPSc00557*; sPSc00558*	*beta-ketoacyl-ACP synthase II (KASII)*	Singh et al. (2013); NCBI; Jeennor and Volkaert (2014)
		C14:0; C16:1; C18:0	sPSc00564*; sPSc00565*	*TF myb family PHL8 (MYB)*	Singh et al. (2013); NCBI
		C14:0; C16:1; C18:0	sPSc00566*; sPSc00567*	*Malate dehydrogenase (MDH)*	Singh et al. (2013); NCBI
		C14:0; C16:1; C18:0	sPSc00568*; sPSc00569*	*beta-ketoacyl-ACP synthase III (KASIII)*	Singh et al. (2013); NCBI; Jeennor and Volkaert (2014)
		C14:0; C16:1; C18:0	sPSc00570*; sPSc00571	*TF TCP15 (TCP15)*	Singh et al. (2013); NCBI
		C14:0; C16:1; C18:0	sPSc00574; sPSc00575*	*TF Homeobox-leucine zipper protein ATHB-13-like (HD-Zip)*	Singh et al. (2013); NCBI
OT10	6	-	sPSc00439; sPSc00440*	*Lysophospholipase 2 (LYPLA2)*	Jeennor and Volkaert (2014)
		-	sPSc00436; sPSc00437*; sPSc00438	*Phosphoenolpyruvate/phosphate translocator 1 (PPT)*	Jeennor and Volkaert (2014); NCBI
OT12	5	-	sPSc00399*; sPSc00400*; sPSc00401B	*beta-ketoacyl-ACP synthase II (KASII)*	Bourgis et al. (2011); NCBI
		-	sPSc00441*; sPSc00442	*Homeobox protein GLABRA (GLABRA)*	Jeennor and Volkaert (2014)
		-	sPSc00443; sPSc00444*	*Glycerol-3-phosphate dehydrogenase (GPDH)*	Jeennor and Volkaert (2014); NCBI

		-	sPSc00445; sPSc00446	*Plastid-linoleate desaturase (FAD7)*	Jeennor and Volkaert (2014); NCBI
		-	sPSc00447; sPSc00448	*Malonyl-CoA:ACP transacylase (MACP/MAT)*	Jeennor and Volkaert (2014); NCBI
		-	sPSc00449	*Triacylglycerol lipase 2 (LIPT2)*	Jeennor and Volkaert (2014)
T14	11	-	SA1_oSSR	*Stearoyl-ACP desaturase (SAD)*	Singh et al. (2013); NCBI
OT15	10	-	sPSc00390*; sPSc00391; sPSc00392*; sPSc00393*	*beta-ketoacyl-ACP synthase II (KASII)*	Bourgis et al. (2011); Montoya et al. (2013); NCBI
		-	sPSc00394; sPSc00395; sPSc00396*	*beta-ketoacyl-ACP synthase I (KASI)*	Bourgis et al. (2011); Montoya et al. (2013); NCBI
T16	15	-	sPSc00402*; sPSc00403*; sPSc00404*; sPSc00405	*beta-ketoacyl-ACP synthase II (KASII)*	Bourgis et al. (2011); Montoya et al. (2013); NCBI

*Non-polymorphic markers; #QTLs reported by Ting et al. (2016)

Supplementary material

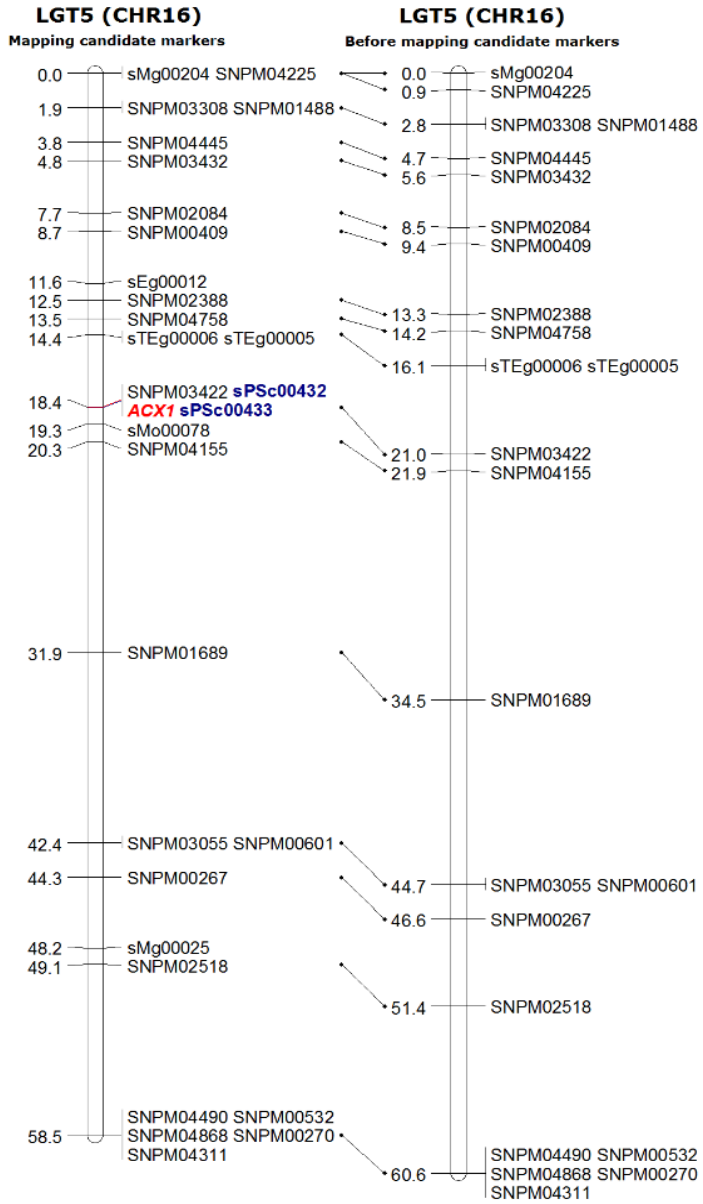
Fig. 1. The polymorphism profiles observed for the segregating SSR alleles in the OxG mapping population.

Polymorphism pattern and scoring of SSR alleles												Expected Segregation ratio
Profile	Maternal	Paternal	Progenies									
			1	2	3	4	5	6	7	8	9	10	
1	*a* – *b* –	*a* –	– –	–	– –	– –	–	–	– –	–	– –	–	1:1 (*ab*:*aa*)
2	*a* – *b* –	*a* – *b* –	–	– –	– –	–	–	– –	–	– –	–	– –	1:2:1 (*aa*:*ab*:*bb*)
3	*a* – *b* –	*a* – *c* –	–	– –	– –	– –	–	– –	– –	– –	–	– –	1:1:1:1 (*aa*:*ab*:*ac*:*bc*)

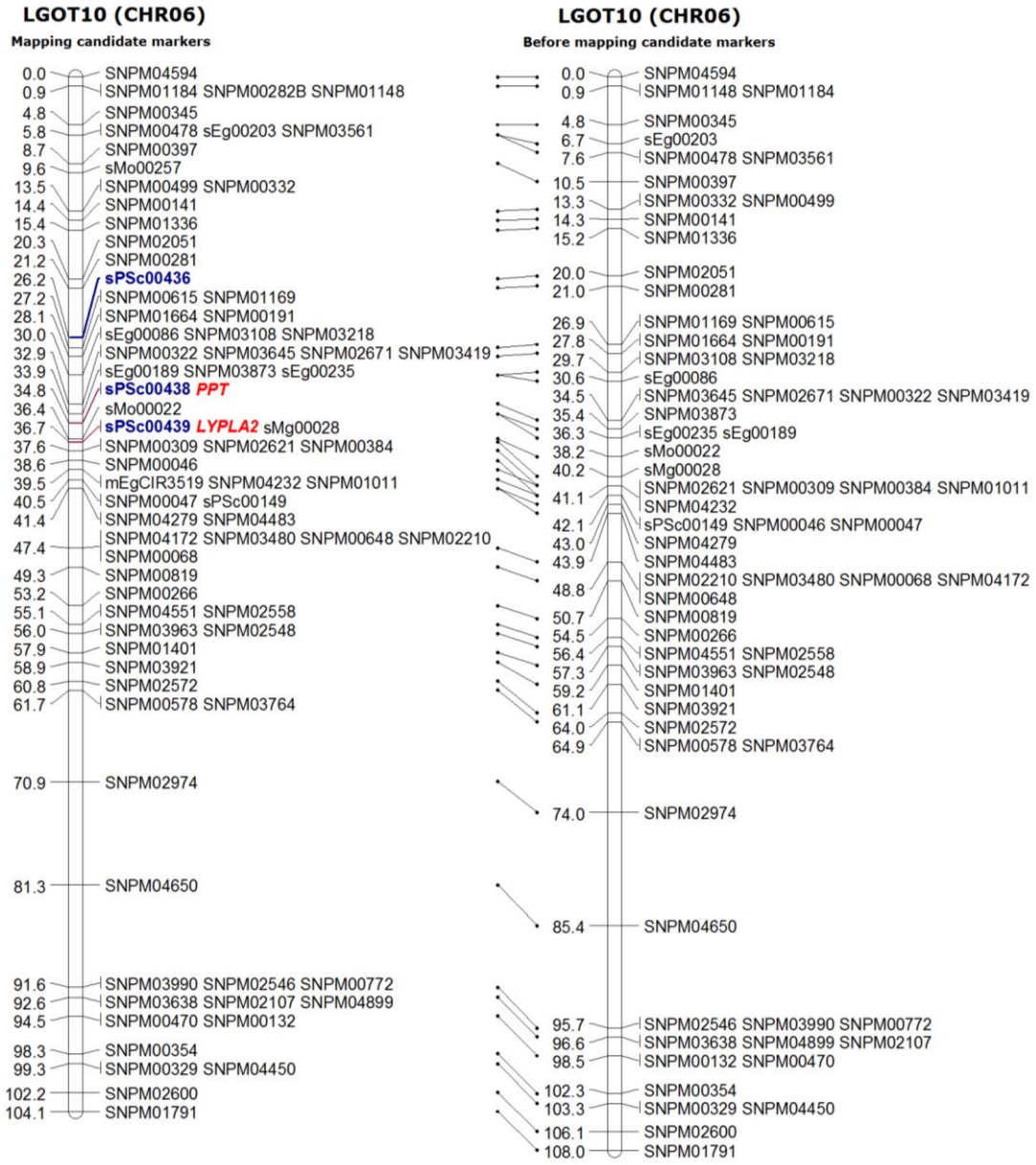
Online Resource 1: Fig. 1. The polymorphism profiles observed for the segregating SSR alleles in the OxG mapping population.

443 Online Resource 2: Table 1. Putative biological function for the candidate genes and transcription factors (TFs) identified
444 from various QTL regions associated with palm oil iodine value (IV) and fatty acid composition (FAC) on the OxG
445 genetic linkage map (Ting et al. 2016).

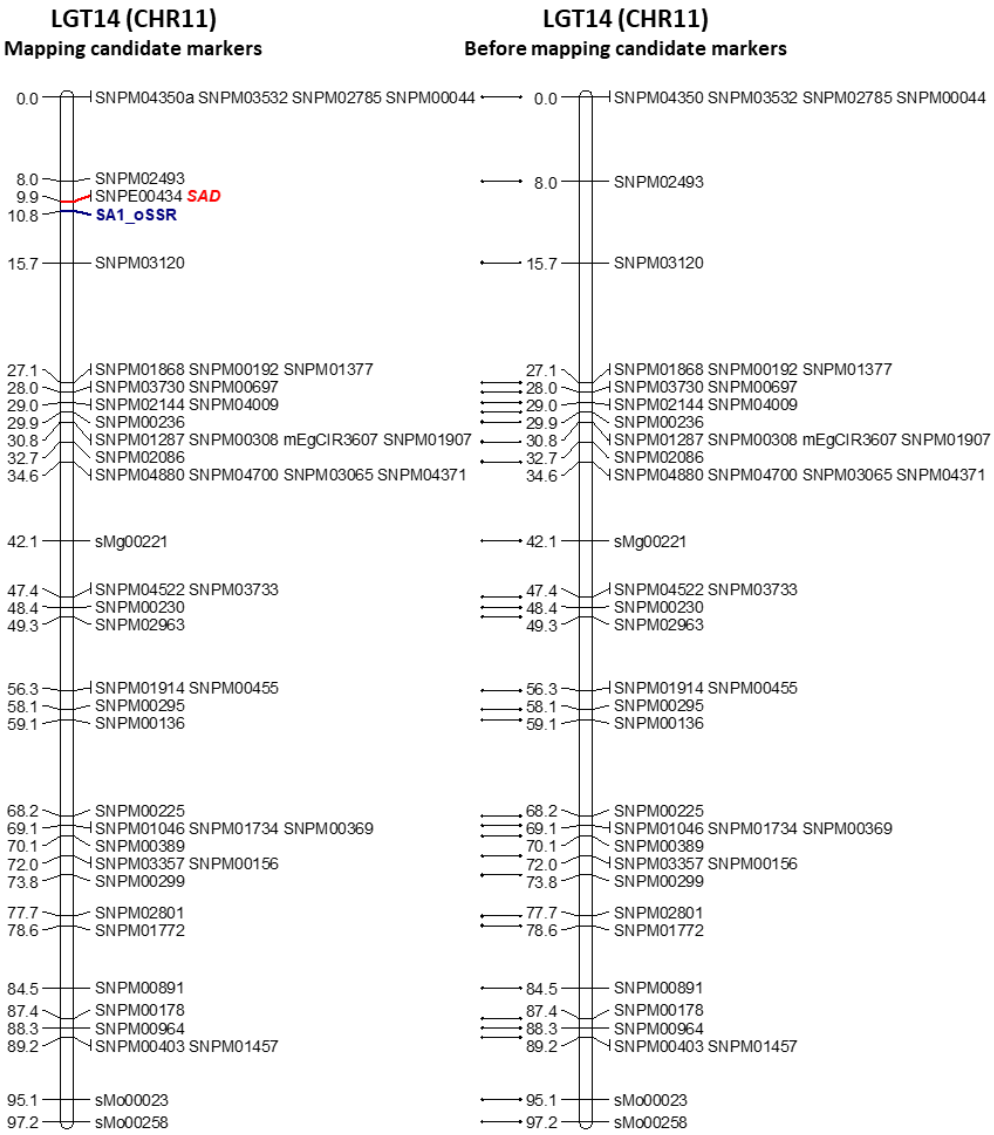
Fig. 2. Mapping candidate SSR markers (in blue font) onto the OxG linkage map in linkage groups (LGs) T5, OT10, OT12, T14, OT15 and T16. These candidate SSR markers were developed from QTLs associated with FAC and oil yield published previously by Bourgis et al. (2011), Montoya et al. (2013) and Jeennor and Volkaert (2014). The updated OxG LGs (Mapping candidate markers) are aligned to the previous map (Before mapping candidate markers) published by Ting et al. (2016). Candidate genes and transcription factors are indicated in italic red.



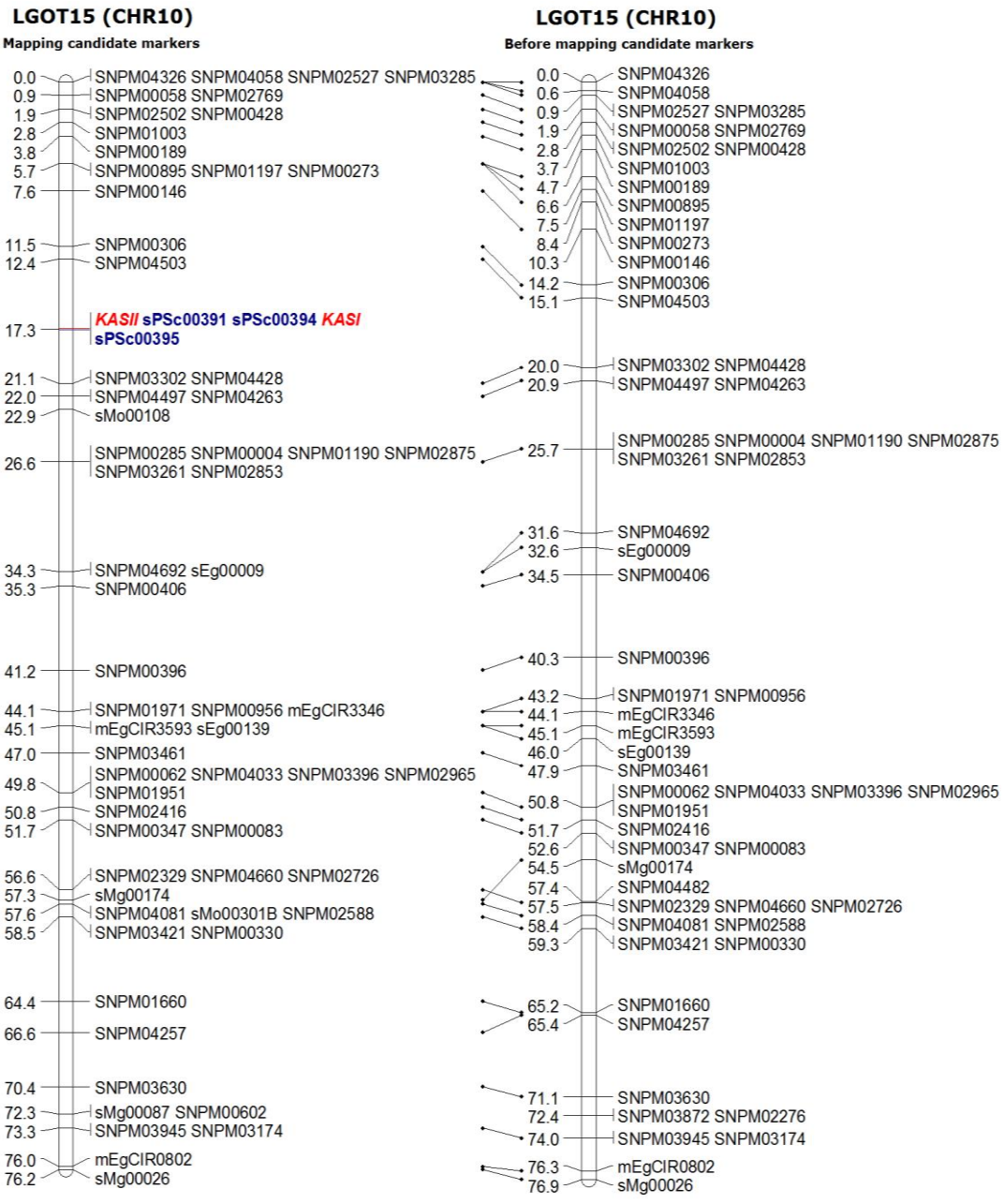
446


LGOT10 (CHR06)
Mapping candidate markers
0.0 SNPM04594
0.9 SNPM01184 SNPM00282B SNPM01148
4.8 SNPM00345
5.8 SNPM00478 sEg00203 SNPM03561
8.7 SNPM00397
9.6 sMo00257
13.5 SNPM00499 SNPM00332
14.4 SNPM00141
15.4 SNPM01336
20.3 SNPM02051
21.2 SNPM00281
26.2 sPSc00436
27.2 SNPM00615 SNPM01169
28.1 SNPM01664 SNPM00191
30.0 sEg00086 SNPM03108 SNPM03218
32.9 SNPM00322 SNPM03645 SNPM02671 SNPM03419
33.9 sEg00189 SNPM03873 sEg00235
34.8 sPSc00438 PPT
36.4 sMo00022
36.7 sPSc00439 LYPLA2 sMg00028
37.6 SNPM00309 SNPM02621 SNPM00384
38.6 SNPM00046
39.5 mEgCIR3519 SNPM04232 SNPM01011
40.5 SNPM00047 sPSc00149
41.4 SNPM04279 SNPM04483
47.4 SNPM04172 SNPM03480 SNPM00648 SNPM02210
SNPM00068
49.3 SNPM00819
53.2 SNPM00266
55.1 SNPM04551 SNPM02558
56.0 SNPM03963 SNPM02548
57.9 SNPM01401
58.9 SNPM03921
60.8 SNPM02572
61.7 SNPM00578 SNPM03764
70.9 SNPM02974
81.3 SNPM04650
91.6 SNPM03990 SNPM02546 SNPM00772
92.6 SNPM03638 SNPM02107 SNPM04899
94.5 SNPM00470 SNPM00132
98.3 SNPM00354
99.3 SNPM00329 SNPM04450
102.2 SNPM02600
104.1 SNPM01791
LGOT10 (CHR06)
Before mapping candidate markers
0.0 SNPM04594
0.9 SNPM01148 SNPM01184
4.8 SNPM00345
6.7 sEg00203
7.6 SNPM00478 SNPM03561
10.5 SNPM00397
13.3 SNPM00332 SNPM00499
14.3 SNPM00141
15.2 SNPM01336
20.0 SNPM02051
21.0 SNPM00281
26.9 SNPM01169 SNPM00615
27.8 SNPM01664 SNPM00191
29.7 SNPM03108 SNPM03218
30.6 sEg00086
34.5 SNPM03645 SNPM02671 SNPM00322 SNPM03419
35.4 SNPM03873
36.3 sEg00235 sEg00189
38.2 sMo00022
40.2 sMg00028
41.1 SNPM02621 SNPM00309 SNPM00384 SNPM01011
SNPM04232
42.1 sPSc00149 SNPM00046 SNPM00047
43.0 SNPM04279
43.9 SNPM04483
48.8 SNPM02210 SNPM03480 SNPM00068 SNPM04172
SNPM00648
50.7 SNPM00819
54.5 SNPM00266
56.4 SNPM04551 SNPM02558
57.3 SNPM03963 SNPM02548
59.2 SNPM01401
61.1 SNPM03921
64.0 SNPM02572
64.9 SNPM00578 SNPM03764
74.0 SNPM02974
85.4 SNPM04650
95.7 SNPM02546 SNPM03990 SNPM00772
96.6 SNPM03638 SNPM04899 SNPM02107
98.5 SNPM00132 SNPM00470
102.3 SNPM00354
103.3 SNPM00329 SNPM04450
106.1 SNPM02600
108.0 SNPM01791

447


LGT14 (CHR11)
Mapping candidate markers
0.0 SNPM04350a SNPM03532 SNPM02785 SNPM00044
8.0 SNPM02493
9.9 SNPE00434 SAD
10.8 SA1_oSSR
15.7 SNPM03120
27.1 SNPM01868 SNPM00192 SNPM01377
28.0 SNPM03730 SNPM00697
29.0 SNPM02144 SNPM04009
29.9 SNPM00236
30.8 SNPM01287 SNPM00308 mEgCIR3607 SNPM01907
32.7 SNPM02086
34.6 SNPM04880 SNPM04700 SNPM03065 SNPM04371
42.1 sMg00221
47.4 SNPM04522 SNPM03733
48.4 SNPM00230
49.3 SNPM02963
56.3 SNPM01914 SNPM00455
58.1 SNPM00295
59.1 SNPM00136
68.2 SNPM00225
69.1 SNPM01046 SNPM01734 SNPM00369
70.1 SNPM00389
72.0 SNPM03357 SNPM00156
73.8 SNPM00299
77.7 SNPM02801
78.6 SNPM01772
84.5 SNPM00891
87.4 SNPM00178
88.3 SNPM00964
89.2 SNPM00403 SNPM01457
95.1 sMo00023
97.2 sMo00258
LGT14 (CHR11)
Before mapping candidate markers
0.0 SNPM04350 SNPM03532 SNPM02785 SNPM00044
8.0 SNPM02493
15.7 SNPM03120
27.1 SNPM01868 SNPM00192 SNPM01377
28.0 SNPM03730 SNPM00697
29.0 SNPM02144 SNPM04009
29.9 SNPM00236
30.8 SNPM01287 SNPM00308 mEgCIR3607 SNPM01907
32.7 SNPM02086
34.6 SNPM04880 SNPM04700 SNPM03065 SNPM04371
42.1 sMg00221
47.4 SNPM04522 SNPM03733
48.4 SNPM00230
49.3 SNPM02963
56.3 SNPM01914 SNPM00455
58.1 SNPM00295
59.1 SNPM00136
68.2 SNPM00225
69.1 SNPM01046 SNPM01734 SNPM00369
70.1 SNPM00389
72.0 SNPM03357 SNPM00156
73.8 SNPM00299
77.7 SNPM02801
78.6 SNPM01772
84.5 SNPM00891
87.4 SNPM00178
88.3 SNPM00964
89.2 SNPM00403 SNPM01457
95.1 sMo00023
97.2 sMo00258

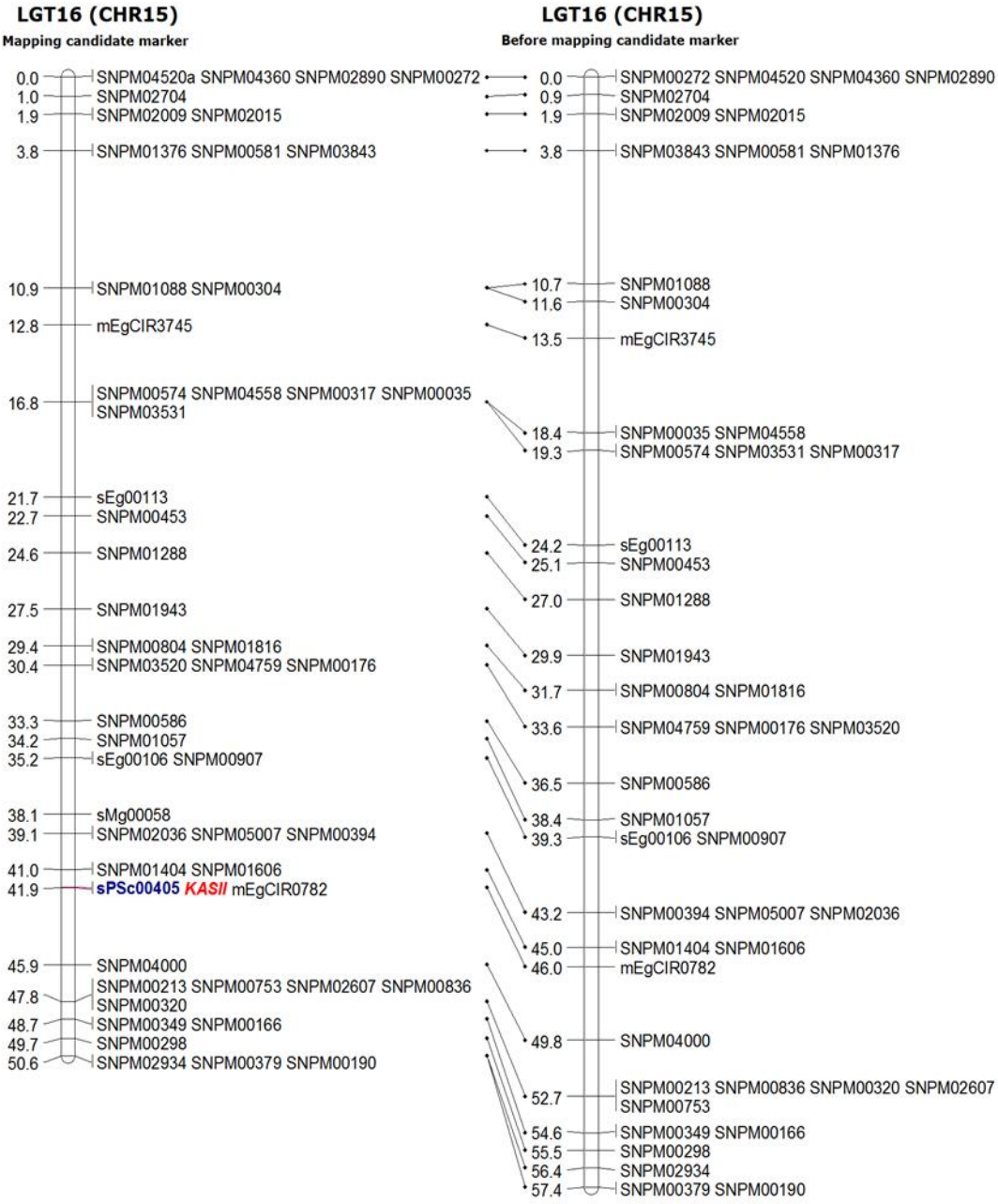
448


LGOT15 (CHR10)
Mapping candidate markers
0.0 SNPM04326 SNPM04058 SNPM02527 SNPM03285
0.9 SNPM00058 SNPM02769
1.9 SNPM02502 SNPM00428
2.8 SNPM01003
3.8 SNPM00189
5.7 SNPM00895 SNPM01197 SNPM00273
7.6 SNPM00146
11.5 SNPM00306
12.4 SNPM04503
17.3 KASII sPSc00391 sPSc00394 KASI sPSc00395
21.1 SNPM03302 SNPM04428
22.0 SNPM04497 SNPM04263
22.9 sMo00108
26.6 SNPM00285 SNPM00004 SNPM01190 SNPM02875 SNPM03261 SNPM02853
34.3 SNPM04692 sEg00009
35.3 SNPM00406
41.2 SNPM00396
44.1 SNPM01971 SNPM00956 mEgCIR3346
45.1 mEgCIR3593 sEg00139
47.0 SNPM03461
49.8 SNPM00062 SNPM04033 SNPM03396 SNPM02965 SNPM01951
50.8 SNPM02416
51.7 SNPM00347 SNPM00083
56.6 SNPM02329 SNPM04660 SNPM02726
57.3 sMg00174
57.6 SNPM04081 sMo00301B SNPM02588
58.5 SNPM03421 SNPM00330
64.4 SNPM01660
66.6 SNPM04257
70.4 SNPM03630
72.3 sMg00087 SNPM00602
73.3 SNPM03945 SNPM03174
76.0 mEgCIR0802
76.2 sMg00026
LGOT15 (CHR10)
Before mapping candidate markers
0.0 SNPM04326
0.6 SNPM04058
0.9 SNPM02527 SNPM03285
1.9 SNPM00058 SNPM02769
2.8 SNPM02502 SNPM00428
3.7 SNPM01003
4.7 SNPM00189
6.6 SNPM00895
7.5 SNPM01197
8.4 SNPM00273
10.3 SNPM00146
14.2 SNPM00306
15.1 SNPM04503
20.0 SNPM03302 SNPM04428
20.9 SNPM04497 SNPM04263
25.7 SNPM00285 SNPM00004 SNPM01190 SNPM02875 SNPM03261 SNPM02853
31.6 SNPM04692
32.6 sEg00009
34.5 SNPM00406
40.3 SNPM00396
43.2 SNPM01971 SNPM00956
44.1 mEgCIR3346
45.1 mEgCIR3593
46.0 sEg00139
47.9 SNPM03461
50.8 SNPM00062 SNPM04033 SNPM03396 SNPM02965 SNPM01951
51.7 SNPM02416
52.6 SNPM00347 SNPM00083
54.5 sMg00174
57.4 SNPM04482
57.5 SNPM02329 SNPM04660 SNPM02726
58.4 SNPM04081 SNPM02588
59.3 SNPM03421 SNPM00330
65.2 SNPM01660
65.4 SNPM04257
71.1 SNPM03630
72.4 SNPM03872 SNPM02276
74.0 SNPM03945 SNPM03174
76.3 mEgCIR0802
76.9 sMg00026

449

No.	Gene/TF	NCBI accession	Putative function for the encoded enzymes/protein/TF	Reference
1	*Malate dehydrogenase* (*MDH*)	XM_010938857.1	The encoded enzyme catalyzes the interconversion of malate to oxaloacetate and *vice versa* which, subsequently can be converted to form phosphoenolpyruvate (PEP) or can be oxidized to form pyruvate as the source to initiate the synthesis of fatty acids (FAs).	Wedding (1989); Minárik et al. (2002)
2	*Phosphoenolpyruvate/phosphate translocator 1* (*PPT*)	XM_010925820.2	The encoded translocator transports phosphoenolpyruvate (PEP) produced from glycolysis into the plastids.	Flügge et al. (2011); Bourgis et al. (2011)
3	*ATP-citrate synthase beta chain protein 1* (*ACLB-1*)	XM_010939930.2	The encoded enzyme synthesizes cytosolic acetyl-CoA which can be used for the elongation of FAs. May also supply substrate to the cytosolic acetyl-CoA carboxylase, which generates malonyl-CoA for the synthesis of very long chain FAs.	Uniprot (http://www.uniprot.org/uniprot/Q93VT8)
4	*Malonyl-CoA:ACP transacylase* (*MACP/FabD/MAT*)	XR_831945.2; XM_010922720.2	The encoded enzyme catalyses malonylation *via* transferring malonyl group from malonyl-CoA to the acyl carrier protein (ACP). The resulted malonyl-ACP as the substrate for the subsequent condensation reaction catalysed by KASIII	Heath and Rock (1995); Zhang et al. (2007); Arthur et al. (2009); Hong et al. (2010)
5	*beta-ketoacyl-ACP synthase III* (*KASIII*)	XM_010938804.1; JN003561.1	The encoded KASIII enzyme forms the acetoacetyl-ACP complex from malonyl-ACP and acetyl-ACP, in preparation for FA-chain elongation.	Clough et al. (1992); Yuan et al. (2012)
6	*beta-ketoacyl-ACP reductase (KAR) / 3-oxoacyl-ACP reductase* (*FabG*)	XM_010915609.1	The encoded enzyme catalyses 3-ketoacyl-ACP to form 3-hydroxyacyl-ACP, the first reduction reaction during the C4 – C14 FAs-chain elongation.	Hoang et al. (2002); Feng et al. (2015)

450



Online Resource 3: Fig. 2. Mapping candidate SSR markers (in blue font) onto the OxG linkage map in linkage groups (LGs) T5, OT10, OT12, T14, OT15 and T16. These candidate SSR markers were developed from QTLs associated with FAC and oil yield published previously by Bourgis et al. (2011), Montoya et al. (2013) and Jeennor and Volkaert (2014). The updated OxG LGs (Mapping candidate markers) are aligned to the previous map (Before mapping candidate markers) published by Ting et al. (2016). Candidate genes and transcription factors are indicated in italic red.

References

460 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST:
461 a new generation of protein database search programs. Nucleic Acids Research 25:3389–3402.
462 https://doi.org/10.1093/nar/25.17.3389

463 Arthur CJ, Williams C, Pottage K, Płoskoń E, Findlow SC, Burston SG, Simpson TJ, Crump MP, Crosby J (2009)
464 Structure and malonyl CoA-ACP transacylase binding of *streptomyces coelicolor* fatty acid synthase acyl carrier
465 protein. ACS Chemical Biology 4:625–636. https://doi.org/10.1021/cb900099e

466 Bai B, Wang L, Lee M, Zhang YJ, Rahmadsyah R, Alfiko Y, Ye BQ, Wan ZY, Lim CH, Suwanto A, Chua N-H, Yue
467 GH (2017) Genome-wide identification of markers for selecting higher oil content in oil palm. BMC Plant Biology
468 17:93. https://doi.org/10.1186/s12870-017-1045-z

469 Bai B, Wang L, Zhang YJ, Lee M, Rahmadsyah R, Alfiko Y, Ye BQ, Purwantomo S, Suwanto A, Chua N-H, Yue GH
470 (2018) Developing genome-wide SNPs and constructing an ultrahigh-density linkage map in oil palm. Scientific
471 Reports 8:691. https://doi.org/10.1038/s41598-017-18613-2

472 Barvkar VT, Pardeshi VC, Kale SM, Kadoo NY, Gupta VS (2012) Phylogenomic analysis of UDP glycosyltransferase 1
473 multigene family in *Linum usitatissimum* identified genes with varied expression patterns. BMC Genomics 13:175.
474 https://doi.org/10.1186/1471-2164-13-175

475 Billotte N, Marseillac N, Risterucci AM, Adon B, Brotteir P, Baurens FC, Singh R, Herran A, Asmady H, Billot C,
476 Amblard P, Durrand-Gasselin T, Courtois B, Asmono D, Cheah SC, Rohde W, Charrier A (2005) Microsatellite-based
477 high density linkage map in oil palm (*Elaeis guineensis* Jacq.). Theoretical and Applied Genetics 110:754-765.
478 https://doi.org/10.1007/s00122-004-1901-8

479 Bourgis F, Kilaru A, Cao X, Ngando-Ebongue GF, Drira N, Ohlrogge JB, Arondel V (2011) Comparative transcriptome
480 and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning.
481 Proceedings of the National Academy of Sciences of the United States of America 108:12527–12532.
482 https://doi.org/10.1073/pnas.1106502108

483 Cahoon EB, Ripp KG, Hall SE (2001) Formation of conjugated Δ^8, Δ^{10}-double bonds by Δ^{12}-oleic-acid desaturase-related
484 enzymes. Journal of Biological Chemistry 276:2637–2643. https://doi.org/10.1074/jbc.M009188200

485 Cao J, Li JL, Li D, Tobin JF, Gimeno RE (2006) Molecular identification of microsomal acyl-CoA:glycerol-3-phosphate
486 acyltransferase, a key enzyme in de novo triacylglycerol synthesis. Proceedings of the National Academy of Sciences
487 of the United States of America 103:19695–19700. https://doi.org/10.1073/pnas.0609140103

488 Cao S, Zhou X-R, Wood CC, Green AG, Singh SP, Liu L, Liu Q (2013) A large and functionally diverse family
489 of *Fad2* genes in safflower (*Carthamus tinctorius* L.). BMC Plant Biology 13:5. https://doi.org/10.1186/1471-2229-
490 13-5

491 Chapman KD, Ohlrogge JB (2012) Compartmentation of triacylglycerol accumulation in plants. Journal of Biological
492 Chemistry 287:2288–2294. https://doi.org/10.1074/jbc.R111.290072

493 Clough RC, Matthis AL, Barnum SR, Jaworski JG (1992) Purification and characterization of 3-ketoacyl-acyl carrier
494 protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis.
495 Journal of Biological Chemistry 267:20992–20998.

496 Corley RHV, Tinker PB (2016) The oil palm, 5th edition. John Wiley & Sons, New Jersey.

497 Deng W, Yan F, Zhang X, Tang Y, Yuan Y (2015) Transcriptional profiling of canola developing embryo and
498 identification of the important roles of BnDof5.6 in embryo development and fatty acids synthesis. Plant & Cell
499 Physiology 56:1624–1640. https://doi.org/10.1093/pcp/pcv074

500 Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. FOCUS 12:13-15.

501 Dussert S, Guerin C, Andersson M, Joët T, Tranbarger TJ, Pizot M, Sarah G, Omore A, Durand-Gasselin T, Morcillo F
502 (2013) Comparative transcriptome analysis of three oil palm fruit and seed tissues that differ in oil content and fatty
503 acid composition. Plant Physiology 162:1337–1358. https://doi.org/10.1104/pp.113.220525

504 El-Kouhen K, Blangy S, Ortiz E, Gardies AM, Ferté N, Arondel V (2005) Identification and characterization of a
505 triacylglycerol lipase in Arabidopsis homologous to mammalian acid lipases. FEBS Letters 579:6067–6073.
506 https://doi.org/10.1016/j.febslet.2005.09.072

507 Feng SX, Ma JC, Yang J, Hu Z, Zhu L, Bi HK, Sun YR, Wang HH (2015) *Ralstonia solanacearum* fatty acid composition
508 is determined by interaction of two 3-ketoacyl-acyl carrier protein reductases encoded on separate replicons. BMC
509 Microbiology 15:223. https://doi.org/10.1186/s12866-015-0554-x

510 Flügge U-I, Häusler RE, Ludewig F, Gierth M (2011) The role of transporters in supplying energy to plant plastids.
511 Journal of Experimental Botany 62:2381–2392. https://doi.org/10.1093/jxb/erq361

512 Frentzen M (1998) Acyltransferases from basic science to modified seed oils. European Journal of Lipid Science and
513 Technology 100:161–166. https://doi.org/10.1002/(SICI)1521-4133(19985)100:4/5<161::AID-LIPI161>3.0.CO;2-P

514 Froman BE, Edwards PC, Bursch AG, Dehesh K (2000) ACX3, a novel medium-chain acyl-coenzyme A oxidase from
515 Arabidopsis. Plant Physiology 123:733–742. https://doi.org/10.1104/pp.123.2.733

516 Ghosh A, Pareek A, Sopory SK, Singla-Pareek SL (2014) A glutathione responsive rice glyoxalase II, OsGLYII-2,
517 functions in salinity adaptation by maintaining better photosynthesis efficiency and anti-oxidant pool. The Plant
518 Journal 80:93–105. https://doi.org/10.1111/tpj.12621

519 Goepfert S, Poirier Y (2007) Beta-oxidation in fatty acid degradation and beyond. Current Opinion in Plant Biology
520 10:245–251. https://doi.org/10.1016/j.pbi.2007.04.007

521 Graham IA, Eastmond PJ (2002) Pathways of straight and branched chain fatty acid catabolism in higher plants. Progress
522 in Lipid Research 41:156–181. https://doi.org/10.1016/s0163-7827(01)00022-4

523 Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-
524 testcross: mapping strategy and RAPD markers. Genetics 137:1121–37.

525 Griffiths G, Harwood JL (1991) The regulation of triacylglycerol biosynthesis in cocoa (*Theobroma cacao*) L. Planta
526 184:279–284. https://doi.org/10.1007/BF00197958

527 Griffiths G, Stymne S, Stobart AK (1988) The utilisation of fatty-acid substrates in triacylglycerol biosynthesis by tissue-
528 slices of developing safflower (*Carthamus tinctorius* L.) and sunflower (*Helianthus annuus* L.) cotyledons. Planta
529 173:309–316. https://doi.org/10.1007/BF00401017

530 Guerin C, Joët T, Serret J, Lashermes P, Vaissayre V, Agbessi MD, Beulé T, Severac D, Amblard P, Tregear J, Durand-
531 Gasselin T, Morcillo F, Dussert S (2016) Gene coexpression network analysis of oil biosynthesis in an interspecific
532 backcross of oil palm. The Plant Journal 87:423–441. https://doi.org/10.1111/tpj.13208

533 Hanson J, Johannesson H, Engstroem P (2001) Sugar-dependent alterations in cotyledon and leaf development in
534 transgenic plants expressing the HDZhdip gene ATHB13. Plant Molecular Biology 45:247–262.
535 https://doi.org/10.1023/A:1006464907710

536 Hamberger B, Hahlbrock K (2004) The 4-coumarate: CoA ligase gene family in *Arabidopsis thaliana* comprises one rare,
537 sinapate–activating and three commonly occurring isoenzymes. Proceedings of the National Academy of Sciences of
538 the United States of America 101:2209–2214. https://doi.org/10.1073/pnas.0307307101

539 Heath RJ, Rock CO (1995) Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of
540 fatty acid elongation in *Escherichia coli.* Journal of Biological Chemistry 270:26538–26542.
541 https://doi.org/10.1074/jbc.270.44.26538

542 Hoang TT, Sullivan SA, Cusick JK, Schweizer HP (2002) Beta-ketoacyl acyl carrier protein reductase (FabG) activity of
543 the fatty acid biosynthetic pathway is a determining factor of 3-oxo-homoserine lactone acyl chain lengths.
544 Microbiology 148:3849–3856. https://doi.org/10.1099/00221287-148-12-3849

545 Hong SK, Kim KH, Park JK, Jeong KW, Kim Y, Kim EE (2010) New design platform for malonyl-CoA-acyl carrier
546 protein transacylase. FEBS Letters 584:1240–1244. https://doi.org/10.1016/j.febslet.2010.02.038

547 Hu Z, Wu Q, Dalal J, Vasani N, Lopez HO, Sederoff HW, Qu R (2017) Accumulation of medium-chain, saturated fatty
548 acyl moieties in seed oils of transgenic *Camelina sativa*. PLoS One 12:e0172296.
549 https://doi.org/10.1371/journal.pone.0172296

550 Index Mundi (2016) Palm oil imports by country in 1000 mt.
551 https://www.indexmundi.com/agriculture/?commodity=palm–oil&graph=imports.

552 Jeennor S, Volkaert H (2014) Mapping of quantitative trait loci (QTLs) for oil yield using SSRs and gene-based markers
553 in African oil palm (*Elaeis guineensis* Jacq.). Tree Genetics & Genomes 10:1–14. https://doi.org/10.1007/s11295-
554 013-0655-3

555 Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin
556 of plant acyl-ACP thioesterases. The Plant Cell 7:359–371. https://doi.org/10.110/tpc.7.3.359

Kalyana Babu B, Mathur RK, Ravichandran G, Anita P, Venu MVB (2020) Genome wide association study (GWAS) and identification of candidate genes for yield and oil yield related traits in oil palm (*Eleaeis guineensis*) using SNPs by genotyping-based sequencing. Genomics 112:1011–1020. https://doi.org/10.1016/j.ygeno.2019.06.018

Kim HJ, Silva JE, Iskandarov U, Andersson M, Cahoon RE, Mockaitis K, Cahoon EB (2015) Structurally divergent lysophosphatidic acid acyltransferases with high selectivity for saturated medium chain fatty acids from *Cuphea* seeds. The Plant Journal 84:1021–1033. https://doi.org/10.1111/tpj.13063

Kriechbaumer V, Weigang L, Fiesselmann A, Letzel T, Frey M, Gierl A, Glawischnig E (2008) Characterisation of the tryptophan synthase alpha subunit in maize. BMC Plant Biology 8:44. https://doi.org/10.1186/1471-2229-8-44

Kurihara T, Ueda M, Tanaka A (1988) Occurrence and possible roles of acetoacetyl-CoA thiolase and 3-ketoacyl-CoA thiolase in peroxisomes of an n-alkane-grown yeast, *Candida tropicalis.* FEBS Letters 229:215–218. https://doi.org/10.1016/0014-5793(88)80830-5

Kushairi A, Loh SK, Azman I, Elina H, Ong-Abdullah M, Zanal Bidin MNI, Razmah G, Shamala S, Parveez GKA (2018) Oil palm economic performance in Malaysia and R&D progress in 2017. Journal of Oil Palm Research 30:163–195. https://doi.org/10.21894/jopr.2018.0030

Larson TR, Edgell T, Byrne J, Dehesh K, Graham IA (2002) Acyl CoA profiles of transgenic plants that accumulate medium–chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. The Plant Journal 32: 519–527. https://doi.org/10.1046/j.1365-313x.2002.01440.x

Lee M, Xia JH, Zou Z1 Ye J, Rahmadsyah, Alfiko Y, Jin J, Lieando JV, Purnamasari MI, Lim CH, Suwanto A, Wong L, Chua NH and Yue GH (2015) A consensus linkage map of oil palm and a major QTL for stem height. Scientific Reports 5:8232. https://doi.org/10.1038/srep08232

Leung K-C, Li H-Y, Mishra G, Chye M-L (2004) ACBP4 and ACBP5, novel *Arabidopsis* acyl-CoA-binding proteins with kelch motifs that bind oleoyl-CoA. Plant Molecular Biology 55:297–309. https://doi.org/10.1007/s11103-004-0642-z

Li QT, Lu X, Song QX, Chen HW, Wei W, Tao JJ, Bian XH, Shen M, Ma B, Zhang WK, Bi YD, Li W, Lai YC, Lam SM, Shui GH, Chen SY, Zhang JS (2017) Selection for a zinc-finger protein contributes to seed oil increase during soybean domestication. Plant Physiology 173:2208–2224. https://doi.org/10.1104/pp.16.01610

Liu YF, Li QT, Lu X, Song QX, Lam SM, Zhang WK, Ma B, Lin Q, Man WQ, Du WG, Shui GH, Chen SY, Zhang JS (2014) Soybean GmMYB73 promotes lipid accumulation in transgenic plants. BMC Plant Biology 14:73. https://doi.org/10.1186/1471-2229-14-73

Liu Q, Wu M, Zhang B, Shrestha P, Petrie J, Green AG, Singh SP (2017) Genetic enhancement of palmitic acid accumulation in cotton seed oil through RNAi down-regulation of ghKAS2 encoding β-ketoacyl-ACP synthase II (KASII). Plant Biotechnology Journal 15:132–143. https://doi.org/10.1111/pbi.12598

Lokesh U, Kiranmai K, Pandurangaiah M, Sudhakarbabu O, Nareshkumar A, Sudhakar C (2013) Role of plant fatty acid elongase (3 keto acyl-CoA synthase) gene in cuticular wax biosynthesis. Research and Reviews: Journal of Agriculture and Allied Sciences 2:35–42.

592 Ma J, Liu F, Wang Q, Wang K, Jones DC, Zhang B (2016) Comprehensive analysis of TCP transcription factors and their
593 expression during cotton (*Gossypium arboreum*) fiber early development. Scientific Reports 6:21535.
594 https://doi.org/10.1038/srep21535

595 Manaf AM, Harwood JL (2000) Purification and characterisation of acyl-CoA, glycerol 3-phosphate acyltransferase from
596 oil palm (*Elaeis guineensis*) tissues. Planta 210:318–328. https://doi.org/10.1007/PL00008140

597 Massengo-Tiassé RP, Cronan JE (2009) Diversity in enoyl-acyl carrier protein reductases. Cellular and Molecular Life
598 Sciences 66:1507–1517. https://doi.org/10.1007/s00018-009-8704-7

599 Meru G, McGregor C (2014) Quantitative trait loci and candidate genes associated with fatty acid content of watermelon
600 seed. Journal of the American Society for Horticultural Science 139:433–441.
601 https://doi.org/10.21273/jashs.139.4.433

602 Minárik P, Tomásková N, Kollárová M, Antalík M (2002) Malate dehydrogenases--structure and function. General
603 Physiology and Biophysics 21:257–265.

604 Mohd Din A, Rajanaidu N, Jalani BS (2000) Performance of *Elaeis oleifera* from Panama, Costa Rica, Colombia and
605 Honduras in Malaysia. Journal of Oil Palm Research 12:71–80.

606 Montoya C, Lopes R, Flori A, Cros D, Cuellar T, Summo M, Espeout S, Rivallan R, Risterucci A-M, Bittencourt D,
607 Zambrano JR, Alarcón GWH, Villeneuve P, Pina M, Nouy B, Amblard P, Ritter E, Leroy T, Billotte N (2013)
608 Quantitative trait loci (QTLs) analysis of palm oil fatty acid composition in an interspecific pseudo-backcross from
609 *Elaeis oleifera* (H.B.K.) Cortés and oil palm (*Elaeis guineensis* Jacq.). Tree Genetics & Genomes 9:1207–1225.
610 https://doi.org/10.1007/s11295-013-0629-5

611 Montoya C, Cochard B, Flori A, Cros D, Lopes R, Cuellar T, Espeout S, Syaputra I, Villeneuve P, Pina M, Ritter E, Leroy
612 T, Billotte N (2014) Genetic architecture of palm oil fatty acid composition in cultivated oil palm (*Elaeis guineensis*
613 Jacq.) compared to its wild relative *E. oleifera* (H.B.K) Cortés. PLoS One 9:e95412.
614 https://doi.org/10.1371/journal.pone.0095412

615 Niu YF, Wang X, Hu DX, Balamurugan S, Li DW, Yang WD, Liu JS, Li HY (2016) Molecular characterization of a
616 glycerol-3-phosphate acyltransferase reveals key features essential for triacylglycerol production in *Phaeodactylum*
617 *tricornutum*. Biotechnology for Biofuels 9:60. https://doi.org/10.1186/s13068-016-0478-1

618 Nurniwalis AW, Zubaidah R, Siti Nor Akmar A, Zulkifli H, Manaf AM, Massawe F, Chan KL, Parveez GKA (2015)
619 Genomic structure and characterization of a lipase class 3 gene and promoter from oil palm. Biologia Plantarum 59:
620 227–236. https://doi.org/10.1007/s10535-015-0500-7

621 Ohlrogge JB, Jaworski JG (1997) Regulation of fatty acid synthesis. Annual Review of Plant Physiology and Plant
622 Molecular Biology 48:109–136. https://doi.org/10.1146/annurev.arplant.48.1.109

623 Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J (1994) Arabidopsis *FAD2* gene encodes the enzyme that
624 is essential for polyunsaturated lipid synthesis. The Plant Cell 6:147–158. https://doi.org/10.1105/tpc.6.1.147

Ong A-L, The C-K, Kwong Q-B, Tangaya P, Appleton DR, Massawe F, Mayes S (2019) Linkage-based genome assembly improvement of oil palm (*Elaeis guineensis*). Scientific Reports 9:6619. https://doi.org/10.1038/s41598-019-42989-y

Oura T, Kajiwara S (2014) Substrate specificity and regioselectivity of Δ12 and ω3 fatty acid desaturases from *Saccharomyces kluyveri*. Bioscience, Biotechnology, and Biochemistry 72:3174–3179. https://doi.org/10.1271/bbb.80361

Ouyang J, Shao X, Li J (2000) Indole-3-glycerol phosphate, a branchpoint of indole-3-acetic acid biosynthesis from the tryptophan biosynthetic pathway in *Arabidopsis thaliana*. The Plant Journal 24:327–333. https://doi.org/10.1046/j.1365-313x.2000.00883.x

Ostrowski M, Jakubowska A (2014) UDP–glycosyltransferases of plant hormones. Advances in Cell Biology 4:43–60. https://doi.org/10.2478/acb-2014-0003

Paul S, Gable K, Beaudoin F, Cahoon E, Jaworski J, Napier JA, Dunn TM (2006) Members of the Arabidopsis FAE1-like 3-ketoacyl-CoA synthase gene family substitute for the Elop proteins of *Saccharomyces cerevisiae*. Journal of Biological Chemistry 281:9018–9029. https://doi.org/10.1074/jbc.M507723200

Pedersen L, Henriksen A (2005) Acyl-CoA oxidase 1 from *Arabidopsis thaliana*. Structure of a key enzyme in plant lipid metabolism. Journal of Molecular Biology 345:487–500. https://doi.org/10.1016/j.jmb.2004.10.062

Pidkowich MS, Nguyen HT, Heilmann I, Ischebeck T, Shanklin J (2007) Modulating seed β-ketoacyl-acyl carrier protein synthase II level converts the composition of a temperate seed oil to that of a palm-like tropical oil. Proceedings of the National Academy of Sciences of the United States of America 104:4742–4747. https://doi.org/10.1073/pnas.0611141104

Poirier Y, Antonenkov VD, Glumoff T, Hiltunen JK (2006) Peroxisomal beta-oxidation--a metabolic pathway with multiple functions. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1763:1413–1426. https://doi.org/10.1016/j.bbamcr.2006.08.034

Raghavan C, Collard BCY (2012) Effect of small mapping population sizes on reliability of quantitative trait locus (QTL) mapping. African Journal of Biotechnology 11:10661–10674. https://doi.org/10.5897/AJB11.2032

Rajanaidu N, Rao V, Halim HA, Ong ASH (1989) Genetic resources, new developments in oil palm breeding. Elaeis 1:1–10.

Rajanaidu N, Kushari A, Rafii M, Mohd Din A, Maizura I, Jalani BS (2000) Oil palm breeding and genetic resources. In Basiron Y, Jalani BS and Chan KW (eds), Advances in oil palm research. Malaysian Palm Oil Board, Bangi, pp 171–237.

Reinbothe S, Quigley F, Springer A, Schemenewitz A, Reinbothe C (2004) The outer plastid envelope protein Oep16, role as precursor translocase in import of protochlorophyllide oxidoreductase A. Proceedings of the National Academy of Sciences of the United States of America 101:2203–2208. https://doi.org/10.1073/pnas.0301962101

Ross J, Li Y, Lim E-K, Bowles DJ (2001) Higher plant glycosyltransferases. Genome Biology 2:reviews3004.1 https://doi.org/10.1186/gb-2001-2-2-reviews3004

Sambanthamurthi R, Abrizah O, Umi Salamah R (1999) Biochemical factors that control oil composition in the oil palm. Journal of Oil Palm Research Special Issue:24–33.

Sambanthamurthi R, Oo KC (1990) Thioesterase activity in the oil palm (*Elaeis guineensis*) mesocarp. In Quinn PJ and Harwood JL (eds). Portland Press Ltd., London, pp 166–168.

Sambanthamurthi R, Sundram K, Tan YA (2000) Chemistry and biochemistry of palm oil. Progress in Lipid Research 39:507–558. https://doi.org/10.1016/s0163-7827(00)00015-1

Seng TY, Siti Hawa MS, Chin CW, Ting NC, Singh R, Faridah QZ, Tan SG, Syed Alwee SSR (2011) Genetic linkage map of a high yielding FELDA Deli x Yangambi oil palm cross. PLoS One 6(11):e26593. https://doi.org/10.1371/journal.pone.0026593

Seng T-Y, Ritter E, Mohamed Saad SH, Leao L-J, Singh R, Qamaruz Zaman F, Tan S-G, Syed Alwee SSR, Rao V (2016) QTLs for oil yield components in an elite oil palm (*Elaeis guineensis*) cross. Euphytica 212:399–425. https://doi.org/10.1007/s10681-016-1771-6

Shanklin J, Somerville C (1991) Stearoyl-acyl-carrier-protein desaturase from higher plants is structurally unrelated to the animal and fungal homologs. Proceedings of the National Academy of Sciences of the United States of America 88:2510–2514. https://doi.org/10.1073/pnas.88.6.2510

Shindou H, Hishikawa D, Harayama T, Yuki K, Shimizu T (2009) Recent progress on acyl CoA, lysophospholipid acyltransferase research. Journal of Lipid Research 50 Suppl:S46–51. https://doi.org/10.1194/jlr.R800035-JLR200

Singh R, Ong-Abdullah M, Low ETL, Abdul Manaf MA, Rosli R, Nookiah R, Ooi LC-L, Ooi S-E, Chan K-L, Halim MA, Azizi N, Nagappan J, Bacher B, Lakey N, Smith SW, He D, Hogan M, Budiman MA, Lee EK, DeSalle R, Kudrna D, Goicoechea JL, Wing RA, Wilson RK, Fulton RS, Ordway JM, Martienssen RA, Sambanthamurthi R (2013) Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. Nature 500:335–339. https://doi.org/10.1038/nature12309

Singh R, Tan SG, Panandam JM, Rahimah AR, Leslie Ooi LCL, Low ETL, Sharma M, Jansen J, Cheah SC (2009) Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. BMC Plant Biology 9:114. https://doi.org/10.1186/1471-2229-9-114

Smith RG, Gauthier DA, Dennis DT, Turpin DH (1992) Malate- and pyruvate-dependent fatty acid synthesis in leucoplasts from developing castor endosperm. Plant Physiology 98:1233–1238. https://doi.org/10.1104/pp.98.4.1233

Song SJ, Lee DE, Jung S, Kim H, Han O, Cho BH, Lee IJ, Back K (2004) Characterization of transgenic rice plants expressing an *Arabidopsis FAD7*. Biologia Plantarum 48:361–366. https://doi.org/10.1023/B:BIOP.0000041087.17353.d8

Soto G, Stritzler M, Lisi C, Alleva K, Pagano ME, Ardila F, Mozzicafreddo M, Cuccioloni M, Angeletti M, Ayub ND (2011) Acetoacetyl-CoA thiolase regulates the mevalonate pathway during abiotic stress adaptation. Journal of Experimental Botany 62:5699–5711. https://doi.org/10.1093/jxb/err287

692 Steinkamp T, Hill K, Hinnah SC, Wagner R, Röhl T, Pohlmeyer K, Soll J (2000) Identification of the pore-forming
693 region of the outer chloroplast envelope protein OEP16. Journal of Biological Chemistry 275:11758–11764.
694 https://doi.org/10.1074/jbc.275.16.11758

695 Thiel T, Michalek W, Varshney RK, Graner A (2003) Exploiting EST databases for the development and characterization
696 of gene–derived SSR-markers in barley (*Hordeum vulgare* L.). Theoretical and Applied Genetics 106:411–422.
697 https://doi.org/10.1007/s00122-002-1031-0

698 Ting N-C, Jansen J, Nagappan J, Ishak Z, Chin CW, Tan S-G, Cheah S-C, Singh R (2013) Identification of QTLs
699 associated with callogenesis and embryogenesis in oil palm using genetic linkage maps improved with SSR markers.
700 PLoS One 8:e53076. https://doi.org/10.1371/journal.pone.0053076

701 Ting N-C, Jansen J, Mayes S, Massawe F, Sambanthamurthi R, Ooi LC-L, Chin CW, Arulandoo X, Seng T-Y, Syed
702 Alwee SSR, Ithnin M, Singh R (2014) High density SNP and SSR-based genetic maps of two independent oil palm
703 hybrids. BMC Genomics 15:309. https://doi.org/10.1186/1471-2164-15-309

704 Ting N-C, Yaakub Z, Kamaruddin K, Mayes S, Massawe F, Sambanthamurthi R, Jansen J, Low LET, Ithnin M, Kushairi
705 A, Arulandoo X, Rosli R, Chan K-L, Amiruddin N, Sritharan K, Lim CC, Nookiah R, Mohd Din A, Singh R (2016)
706 Fine-mapping and cross-validation of QTLs linked to fatty acid composition in multiple independent interspecific
707 crosses of oil palm. BMC Genomics 17:289. https://doi.org/10.1186/s12864-016-2607-4

708 Ting N-C, Mayes S, Massawe F, Sambanthamurthi R, Jansen J, Syed Alwee SSR, Seng T-Y, Ithnin M, Singh R (2018)
709 Putative regulatory candidate genes for QTL linked to fruit traits in oil palm (*Elaeis guineensis* Jacq.). Euphytica
710 214:214. https://doi.org/10.1007/s10681-018-2296-y

711 Tranbarger TJ, Dussert S, Joët T, Argout X, Summo M, Champion A, Cros D, Omore A, Nouy B, Morcillo F (2011)
712 Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening and functional specialization in lipid
713 and carotenoid metabolism. Plant Physiology 156:564–584. https://doi.org/10.1104/pp.111.175141

714 USDA, Oil crops yearbook, world vegetable oils supply and distribution, 2012/13–2016/17.
715 https://www.ers.usda.gov/data-products/oil-crops-yearbook/oil-crops-yearbook/#Vegetable oils and animal fats.

716 van Ooijen JW (2006) JoinMap® 4.1, software for calculation of genetic linkage maps in experimental populations.
717 Kyazma B.V., Wageningen, the Netherlands.

718 van Ooijen JW (2009) MapQTL® 6, software for the mapping of quantitative trait loci in experimental populations of
719 diploid species. Kyazma B.V., Wageningen, the Netherlands.

720 Vigeolas H, Waldeck P, Zank T, Geigenberger P (2007) Increasing seed oil content in oil-seed rape (*Brassica napus* L.)
721 by over-expression of a yeast glycerol-3-phosphate dehydrogenase under the control of a seed-specific promoter. Plant
722 Biotechnology Journal 5:431–441. https://doi.org/10.1111/j.1467-7652.2007.00252.x

723 Wedding RT (1989) Malic enzymes of higher plants, characteristics, regulation, and physiological function. Plant
724 Physiology 90:367–371. https://doi.org/10.1104/pp.90.2.367

Wendel AA, Cooper DE, Ilkayeva OR, Muoio DM, Coleman RA (2013) Glycerol-3-phosphate acyltransferase (GPAT)-1, but not GPAT4, incorporates newly synthesized fatty acids into triacylglycerol and diminishes fatty acid oxidation. Journal of Biological Chemistry 288:27299–27306. https://doi.org/10.1074/jbc.M113.485219

Wu G-Z, Xue H-W (2010) *Arabidopsis* β-ketoacyl-[acyl carrier protein] synthase I is crucial for fatty acid synthesis and plays a role in chloroplast division and embryo development. The Plant Cell 22:3726–3744. https://doi.org/10.1105/tpc.110.075564

Xia W, Luo T, Dou Y, Zhang W, Mason AS, Huang D, Huang X, Tang W, Wang J, Zhang C, Xiao Y (2019) Identification and validation of candidate genes involved in fatty acid content in oil palm by genome-wide association analysis. Frontiers in Plant Science 10:1263. https://doi.org/10.3389/fpls.2019.01263

Xiao S, Li HY, Zhang JP, Chan SW, Chye ML (2008) Arabidopsis acyl-CoA-binding proteins ACBP4 and ACBP5 are subcellularly localized to the cytosol and ACBP4 depletion affects membrane lipid composition. Plant Molecular Biology 68:571–583. https://doi.org/10.1007/s11103-008-9392-7

Xu J, Zheng Z, Zou J (2009) A membrane-bound glycerol-3-phosphate acyltransferase from *Thalassiosira pseudonana* regulates acyl composition of glycerolipids. Botany 87:544–551. https://doi.org/10.1139/B08-145

Yang T, Xu R, Chen J, Liu A (2016) β-ketoacyl-acyl carrier protein synthase I (KASI) plays crucial roles in the plant growth and fatty acids synthesis in Tobacco. International Journal of Molecular Sciences 17:E1287. https://doi.org/10.3390/ijms17081287

Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB (2012) A land-plant-pecific glycerol-3-phosphate acyltransferase family in Arabidopsis, substrate specificity, sn-2 preference, and evolution. Plant Physiology 160: 638–652. https://doi.org/10.1104/pp.112.201996

Yuan Y, Sachdeva M, Leeds JA, Meredith TC (2012) Fatty acid biosynthesis in *Pseudomonas aeruginosa* is initiated by the FabY class of β-ketoacyl acyl carrier protein synthases. Journal of Bacteriology 194:5171–5184. https://doi.org/10.1128/JB.00792-12

Yuan Y, Yu S, Yu J, Zhan Z, Li M, Liu G, Wang X, Huang L (2014) Predicting the function of 4-coumarate:CoA ligase *(LJ4CL*1) in *Lonicera japonica*. International Journal of Molecular Sciences 15:2386–2399. https://doi.org/10.3390/ijms15022386

Yurchenko OP, Park S, Ilut DC, Inmon JJ, Millhollon JC, Liechty Z, Page JT, Jenks MA, Chapman KD, Udall JA, Gore MA, Dyer JM (2014) Genome-wide analysis of the omega-3 fatty acid desaturase gene family in *Gossypium*. BMC Plant Biology 14:312. https://doi.org/10.1186/s12870-014-0312-5

Zhang L, Liu W, Xiao J, Hu T, Chen J, Chen K, Jiang H, Shen X (2007) Malonyl-CoA, acyl carrier protein transacylase from *Helicobacter pylori*, crystal structure and its interaction with acyl carrier protein. Protein Science 16:1184–1192. https://doi.org/10.1110/ps.072757307

Zhang X, Li M, Wei D, Xing L (2008) Identification and characterization of a novel yeast omega3-fatty acid desaturase acting on long-chain n-6 fatty acid substrates from Pichia pastoris. Yeast 25:21–27. https://doi.org/10.1002/yea.1546

759 Zhang Y, Maximova SN, Guiltinan MJ (2015) Characterization of a stearoyl-acyl carrier protein desaturase gene family
760 from chocolate tree, *Theobroma cacao* L.. Frontiers in Plant Science 6:239. https://doi.org/10.3389/fpls.2015.00239