Amiruddin, N., Chan, P-K., Azizi, N., Morris, P. E., Cham, K-L., Ong, P. W., Rosli, R., Masura, S. S., Murphy, D. J., Sambanthamurthi, R., Haslam, R. P., Chye, M-L., Harwood, J. L. and Low, E-T. L. 2019. Characterisation of Oil Palm Acyl-CoA-Binding Proteins and Correlation of their Gene Expression with Oil Synthesis. *Plant and cell physiology.* pcz237.

Title:

Characterisation of Oil Palm Acyl-CoA-Binding Proteins and Correlation of their Gene Expression with Oil Synthesis

Running Head:

Co-expression of oil palm ACBPs with oil synthesis

Corresponding authors:

E-T. L. Low

Advanced Biotechnology and Breeding Centre, Malaysian Palm Oil Board, No. 6, Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia

Email: lowengti@mpob.gov.my, Tel: +603 8769 4504, Fax: +603 8926 1995

J. L. Harwood

School of Biosciences, University of Cardiff, Cardiff, CF10 3AX, United Kingdom

Email: harwood@cardiff.ac.uk, Tel: +44 2920 874108

Subject Area:

(3) regulation of gene expression

Number of black and white figures: 0

Number of colour figures: 8

Number of tables: 1

Number of supplementary materials:

Supplementary File = 1

Supplementary Figures = 1

Title:

Characterisation of Oil Palm Acyl-CoA-Binding Proteins and Correlation of their Gene Expression with Oil Synthesis

Running Head:

Co-expression of oil palm ACBPs with oil synthesis

Nadzirah Amiruddin[1], Pek-Lan Chan[1], Norazah Azizi[1], Priscilla Elizabeth Morris[1], Kuang-Lim Chan[1], Pei Wen Ong[1], Rozana Rosli[1], Subhi Siti Masura[1], Denis J Murphy[2], Ravigadevi Sambanthamurthi[1], Richard P Haslam[3], Mee-Len Chye[4], John L Harwood[5,*], Eng-Ti Leslie Low[1,*]

[1] Advanced Biotechnology and Breeding Centre, Malaysian Palm Oil Board, No. 6, Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia

[2] Genomics and Computational Biology Research Group, University of South Wales, Pontypridd, CF37 1DL, United Kingdom

[3] Department of Plant Sciences, Rothamsted Research, Harpenden, AL5 2JQ, United Kingdom.

[4] School of Biological Sciences, The University of Hong Kong, Pokfulam, Hong Kong

[5] School of Biosciences, University of Cardiff, Cardiff, CF10 3AX, United Kingdom

***Corresponding authors:**

E-T. L. Low, Email: lowengti@mpob.gov.my, Tel: +603 8769 4504, Fax: +603 8926 1995

J. L. Harwood, Email: harwood@cardiff.ac.uk, Tel: +44 2920 874108

Abstract

Acyl-CoA-binding proteins (ACBPs) are involved in binding and trafficking acyl-CoA esters in eukaryotic cells. ACBPs contain a well-conserved acyl-CoA-binding domain (ACBD). Their various functions have been characterized in the model plant Arabidopsis and, to a lesser extent, in rice. In this study, genome-wide detection and expression analysis of *ACBP*s were performed on *Elaeis guineensis* (oil palm), the most important oil crop in the world. Seven *E. guineensis ACBPs* were identified and classified into four groups according to their deduced amino acid domain organization. Phylogenetic analysis showed conservation of this family with other higher plants. All seven *EgACBPs* were expressed in most tissues while their differential expression suggests various functions in specific tissues. For example, *EgACBP3* had high expression in inflorescences and stalks while *ACBP1* showed strong expression in leaves. Because of the importance of *E. guineensis* as an oil crop, expression of *EgACBPs* was specifically examined during fruit development. *EgACBP3* showed high expression throughout mesocarp development, while *EgACBP1* had enhanced expression during rapid oil synthesis. In endosperm, both *EgACBP1* and *EgACBP3* exhibited increased expression during seed development. These results provide important information for further investigations on the biological functions of EgACBPs in various tissues and, in particular, their roles in oil synthesis.

Key words

acyl-CoA-binding proteins; endosperm; lipid metabolism; mesocarp; oil accumulation; oil palm (*Elaeis guineensis*)

Introduction

Oil palm (*Elaeis guineensis*) is the major global oil crop in terms of output, accounting for some 35% of the total vegetable oil production (Weselake et al*.* 2017; Kushairi et al. 2018). Its basic chemistry and commercial uses are summarized in Gunstone et al. (2007) and Tang and Pantzaris (2017) while background information, including early biochemistry, is reviewed by Sambanthamurthi et al. (2000) and Salas et al. (2000). More recent coverage of oil palm as a major global crop is reviewed by Murphy (2014). One of the reasons for the commercial success of oil palm is its uniquely high oil yield per hectare, which is 8-10 times that of major oil seeds (Harwood et al. 2017). Ironically, despite some of the adverse publicity that oil palm has received (see Basiron 2005), its exceptionally high yield has significant environmental advantages in terms of land use impact compared to other oil crops (Murphy, 2014). Moreover, in view of the recent severe regulatory restrictions on the use of *trans*-fatty acids in foods (Akoh 2017), the nutritional qualities of palm oil and its fractions which are *trans*-fatty acid-free have been emphasized (Stanley, 2008).

Oil palm fruits produce two distinct types of vegetable oil, namely palm oil and palm kernel oil (Sambanthamurthi et al. 2000). The ratio of palm oil to palm kernel oil is about 8.5 to 1. Palm oil, from the fruit mesocarp, contains about 44% palmitate, 4% stearate, 40% oleate and 11% linoleate. It can be conveniently fractionated into palm olein, palm stearin, palm mid-fraction and other products. In contrast, the oil from the kernel (seed), which comprises mainly endosperm tissue, contains around 49% laurate, 16% myristate, 8% palmitate and 14% oleate with smaller amounts of octanoate and decanoate (Ibrahim 2013). Together with coconut oil, palm kernel oil is widely used in the production of high-lauric compounds that have some uses in the food sector but are mainly used to generate surface-active products such as soaps and detergents for non-food purposes (Gunstone et al. 2007). Given the limitation of available agricultural land, the desire to preserve primary tropical forests and the significance of oil palm in world vegetable oil production, research into enhancing oil palm productivity has increased recently (Murphy 2014; Kushairi et al. 2018), leading to greater interest in the regulation of triacylglycerol (TAG) biosynthesis in *E. guineensis*.

The basic pathways of plant lipid metabolism have been described in detail in chapters within Murphy (2005) and aspects of oil accumulation in crops are described further in Weselake et al. (2009), Bates et al. (2013) and Chen et al. (2015). For oil palm, the basic biochemistry was reviewed by Sambanthamurthi et al. (2000) while the genes involved in fatty acid biosynthesis were identified and characterized by Siti Nor Akmar et al. (1999), Wan Omar et al. (2008), Ramli et al. (2012) and Chan et al. (2017). The oil palm diacylglycerol acyltransferase (DGAT) genes that catalyze an important step in TAG biosynthesis have also been identified (Rosli et al. 2018). More importantly, two major developments have shed some light into the regulation of lipid synthesis in oil palm. First, Ramli et al. (2002, 2009) used callus cultures to examine the regulation of lipid synthesis using flux control analysis. They were able to document quantitative information about the greater control exerted by fatty acid synthesis compared to lipid assembly. Second, studies of gene expression during fruit ripening on mesocarp and kernel found that the expression of *WRINKLED1* (*WRI1*) coordinated well with genes encoding several enzymes in fatty acid biosynthesis genes and high rates of lipid accumulation (Tranbarger et al. 2011).

The importance of *WRI1* was reinforced when Bourgis et al. (2011) showed that *WRI1* was expressed 57-fold higher in oil palm during fruit ripening compared to date palm while most genes for TAG assembly had similar levels. These results confirmed the conclusions of Ramli *et al.* (2002, 2009) from flux control experiments.

During lipid synthesis in oil crops (Weselake et al. 2009, Bates et al. 2013), different subcellular compartments are involved. Fatty acids generated in the plastids join the acyl-CoA pool in the cytosol and are used for lipid assembly on the endoplasmic reticulum. As in other organisms, plants contain acyl-CoA-binding proteins (ACBPs) to prevent the harmful effects of elevated acyl-CoA concentrations, facilitate movement of acyl-CoA esters within the cell and aid enzyme-substrate binding (Xiao and Chye 2011, Du et al. 2016).

All ACBPs contain an acyl-CoA-binding domain (ACBD) for the intracellular transport of acyl-CoA esters. ACBPs function in the protection and formation of acyl-CoA ester pools which act as intermediates and regulators in lipid metabolism and cellular signalling (Knudsen et al. 1994, Faergeman et al. 1997). The highly conserved 90-amino acid ACBD is capable of binding medium- to long-chain acyl-CoA esters and transporting them from the plastid to the endoplasmic reticulum, prior to the formation of TAG (Harwood 1996). ACBPs have been reported from a number of different plants and have been characterized in *Brassica napus*, *Arabidopsis thaliana* (reviewed in Du et al. 2016) and *Oryza sativa* (Meng et al. 2011). Many species encode multiple ACBP genes (six in both rice and *A. thaliana*) which are categorized into Class I-IV, based on their size and the presence of Kelch motifs or ankyrin-repeats (Du et al. 2016). Members of Class III are larger than those of Class I. Class II contains one ACBD plus ankyrin-repeats that drive interactions with other proteins and Class IV contains one ACBD plus Kelch motifs that potentially facilitate protein-protein interactions (Xiao and Chye 2011, Du et al. 2016).

In comparison to the information on the role of ACBPs in stress responses and development (Du et al. 2016), less is known of their involvement in lipid biosynthesis especially in oil crops. In view of the pre-eminence of oil palm in the world oil market and that ACBPs have not been reported in *E. guineensis*, experiments were initiated to address this important deficiency.

Results and Discussion

The E. guineensis ACBP family

All ACBPs are expected to function as acyl-CoA transporters by virtue of possessing the core ACBD. A total of seven *ACBPs* were identified in the *E. guineensis* genome, and were designated as *EgACBP1* to *EgACBP7*. Using SMART (Letunic and Bork, 2018), the EgACBPs were categorized into four classes, aided by the domain architecture of AtACBPs as a reference (Fig. 1). Similar to the AtACBPs, there is only one Class I member, namely EgACBP1. Motif analysis on corresponding orthologues using Motif Multiple En for Motif Elicitation (MEME) (Bailey et al. 2009) shows all ACBDs are at the same location (See Supplemental Fig S1A), in agreement with multiple sequence alignment analysis (Fig. 2A). Multiple alignments of the Class I ACBPs had previously identified two important motifs (YKQA and KWDAW) that are related to acyl-CoA-binding and the coenzyme-A head group-binding, which are proposed to be conserved in all four classes (Kragelund et al. 1993, Xiao and Chye 2011). Xiao and Chye (2009) demonstrated that small AtACBPs could bind linoleoyl-CoA esters, whereas Class I BnACBPs could bind oleoyl-CoA and palmitoyl-CoA esters as reported by Raboanatahiry et al. (2015). Thus, it appears that the acyl-CoA-binding preferences differ among Class I ACBP members, despite them having highly conserved binding motifs. Based on phylogenetic analysis, the acyl-CoA preference of EgACBP1 is likely to be more similar to the OsACBPs, being clustered in the same monocot clade (Fig. 3).

EgACBP2 which consists of ACBD and ankyrin-repeat domains, is the sole member in Class II. It resembles AtACBP1 and AtACBP2. Two C-terminal ankyrin-repeat domains were located in EgACBP2 (Fig. 1), while only one ankyrin-repeat domain was identified in each of AtACBP1 and AtACBP2 (Fig. 1). Two conserved binding motifs YKIA and KWNAW are identical in EgACBP2 and OsACBP4. However, the conserved motif KWNAW differs from both AtACBPs (KWQAW), which was observed to be identical with BnACBPs instead (Fig. 2B). This agrees with phylogenetic analysis of corresponding amino acid sequences which classifies EgACBP2 and OsACBP4 in one clade, and each AtACBP with two BnACBPs in a separate clade (Fig. 3). Being positioned near the C-terminal, the ankyrin-repeat domains would allow Class II ACBPs some additional functions. These widespread structural motifs could mediate protein-protein interactions and ensure diverse functions including transcription initiation, cytoskeletal integrity, ion transport, cell signalling and cell cycle regulation (Mosavi et al. 2002). Additionally, they are responsible for targeting proteins to specialized compartments such as the endoplasmic reticulum or the plasma membrane (Bennet and Chen 2001). Transmembrane domains were detected in ankyrin-repeat AtACBPs, without which the proteins would not be targeted into the plasma membrane (Li and Chye, 2003). This domain is located at the N-terminal in EgACBP2 (Fig. 1). Ankyrin-repeat-containing AtACBPs have been reported to be involved in many important biological functions such as interacting with ethylene-responsive element binding protein (AtEBP) or farnesylated protein 6 (AtFP6), and in heavy-metal accumulation responses, ABA signalling, stem cuticle formation, lead [Pb (II)] accumulation in roots and drought tolerance (Xiao and Chye 2011, Du et al. 2016).

Unlike the AtACBPs which have only a single Class III member (AtACBP3), three EgACBPs (EgACBP3, EgACBP4 and EgACBP5) were identified to be Class III. The ACBD was identified at different positions in each of these EgACBPs (Fig. 1). Interestingly, the binding domain in EgACBP5 is located nearer to the C-terminal in comparison to EgACBP3 and EgACBP4 (Fig. 1). Similarly, a sole member Class III AtACBP, AtACBP3 also features a C-terminally located ACBD, in contrast to the N-terminal occurrence in other classes of AtACBPs (Fig. 1) (Meng et al. 2011, Xiao and Chye 2011). EgACBP5 is the largest of the three Class III EgACBP members (Fig. 1). In the large Class III EgACBPs, the transmembrane domains were found in residues 7-29 in both EgACBP3 and EgACBP4. Two were identified at position 10 to 29 and 42 to 61 in EgACBP5 (Fig. 1). Signal peptides were predicted between residues 1 to 26 in Class III ACBPs, EgACBP3 and EgACBP4 (Fig. 1). A signal peptide is a sequence that contains 16-30 amino acids at the N-terminal of newly synthesised proteins destined for the secretory pathway, including those that are inserted into cellular membranes such as the endoplasmic reticulum (Lodish et al. 2000). Deduced transmembrane domains detected on the large AtACBP (Leung et al. 2006) and OsACBP5 (Meng and Chye 2014) also overlap with signal peptides. EgACBP3, EgACBP4 and OsACBP5 (Meng and Chye 2014) of Class III ACBPs are localized in the endoplasmic reticulum, as predicted by PSORT (Supplemental File S1), while AtACBP3 is targeted to apoplasts (Leung et al. 2006). Phylogenetic analysis shows also that EgACBP5 diverged from the other two large EgACBPs (Fig. 3), which suggests that it could have a non-redundant function when compared to other class members. The large EgACBPs were clustered in a clade differently from other plants, which were observed to cluster together in another clade. Such divergence could indicate differences in domain sequence and structure, as well as functions of individual genes as a result of a changing environment and domain combination (Yang and Bourne 2009). In 2011, Meng et al. reported different acyl-CoA-binding specificities between the large AtACBPs and OsACBPs. This shows that differences exist in ACBP orthologues, and their functions may not always be identical to *A. thaliana*. Class III AtACBPs were reported to be involved in multiple biological functions such as in plant defence signalling during fungal infection, circadian regulation and response to hypoxia (see Xiao and Chye 2011, Du et al. 2016). Given the different phylogenetic positions (Fig. 3) of AtACBPs and EgACBPs, EgACBPs are likely to be involved in other functions.

EgACBP6 and EgACBP7 belonging to Class IV EgACBPs, contain an ACBD accompanied by Kelch motifs (Fig. 1). The ACBDs in Kelch motif-containing EgACBPs are located nearest to the N-terminal in comparison to other ACBPs, which are located on positions ranging from 31 to 101 in Kelch motif AtACBPs (Du et al. 2016). In this study, each of the two largest ACBPs, EgACBP6 and EgACBP7, contain four Kelch-motif domains (Fig. 1). The deduced amino acid sequences of two conserved motifs in these proteins show perfect identities with OsACBP6, at YQQA and KWTSW (Fig. 2). Phylogenetic analysis further confirmed this, by positioning both orthologues in the same clade (Fig. 3). Kelch motif AtACBPs were diverged and positioned separately, showing AtACBP5 is more orthologous to the *B. napus* genes, Ais76197 and Ais76198 (Raboanatahiry et al. 2015). Similar to the ankyrin-repeat motifs, the additional structure provided by the Kelch motifs would be likely to enhance biological functions, such as protein-protein interactions (Du et al. 2016). In AtACBPs, both ankyrin-repeat and Kelch motif ACBPs interact with other proteins in response to biotic and abiotic stress factors. For instance, Kelch motif-containing AtACBP4 was proposed to be involved in AtEBP (element binding protein)-mediated defence, while the ankyrin-repeat-containing AtACBP2 may be involved in ethylene or jasmonate signalling besides facilitating Pb (II) accumulation in roots (Xiao and Chye 2011, Du et al. 2016). These observations suggest that Class IV EgACBPs may potentially possess diverse functions. However, given that oil palm and rice show closer relationships in the phylogenetic tree, it seems probable that the function of the Kelch motif EgACBPs would be more similar to that of rice.

Expression patterns of EgACBPs in oil palm tissues

As a preliminary to more in-depth analysis of the expression of EgACBPs in different tissues, we examined a variety of samples using Roche-454 methodology. Transcriptome data (Fig. 4A, 4B) showed the expression of all seven *EgACBPs* in most of the ten tissues (leaf, root, pollen, mesocarp, endosperm, pith, sepal, immature fruit, spikelet and stalk) examined. *EgACBP1* showed high expression in leaves and white roots of seedlings (Fig. 4A). In contrast, *AtACBP6* (an orthologue of *EgACBP1*) was observed to exhibit stronger expression in siliques and developing seeds than roots and leaves (Engeseth et al. 1996). This result indicates that *EgACBP1* is more involved in leaf and root development compared to *AtACBP6*. Also, β-glucuronidase (GUS) quantitative and GUS-histochemical assays in Arabidopsis plants have previously shown that *AtACBP6* is actively regulated during germination and seed development as its expression was detected in pollen grains, microspores and tapetal cells, as well as in cotyledons, hypocotyl and cotyledonary-staged embryos (Du et al. 2016). Thus, *EgACBP1* may also be involved in seed development and germination if it is to resemble *AtACBP6*, but further investigations need to address this aspect.

It was noteworthy that *EgACBP1* also showed good expression in both mesocarp and endosperm tissues (Fig.4B) with higher levels at stages where oil deposition is greater. This is consistent with data from Arabidopsis (Du et al. 2016), rapeseed (Yurchenko and Weselake 2011) and rice (Meng et al. 2011) where the Class I ACBPs all play a role in seed development and lipid biosynthesis, including oil accumulation. In contrast, *EgACBP3* showed high expression in the stalk. For the endosperm, *EgACBP4* gave moderate expression (but not in the mesocarp) and it was noticeable that *EgACBP1* showed good expression while *EgACBP3* showed barely detectable levels at 10 weeks after anthesis (WAA) in contrast to 15WAA.

For more detailed analyses, the expression patterns of the *EgACBPs* were determined in 15 oil palm tissue samples using reverse transcription quantitative polymerase chain reaction (RT-qPCR) (Fig. 5A, 5B, 5C). Expression of all seven *EgACBPs* was detected in spear leaf, mature leaf, white root, primary root and lateral root (Fig. 5A), with mature leaves usually being the highest. *EgACBP3* and *EgACBP5*, members of Class III ACBPs, showed somewhat higher expression values in the lateral roots, which may indicate that both could play an important role in their development. In rice, all ACBPs, including the Class III protein OsACBP5, were thought to play a role in both leaf and root development (Du et al. 2016, Meng et al. 2011). In Arabidopsis, the large Class III protein AtACBP3, showed increased expression during development and light regulation of leaves (Xiao and Chye 2011).

Interestingly, most *EgACBPs* were observed to have higher expression in mature compared to spear leaves, suggesting their overlapping role in the development of such tissues. This parallels the situation in rice, where all six ACBPs appeared to be involved in leaf development (Meng et al. 2011). In *A. thaliana* also, *AtACBP1*-*AtACBP6* were all reported to be expressed in mature plants. For instance, *AtACBP1* was found to be expressed in silique, root, stem, leaf and flower (Xiao and Chye 2011). On the other hand, *AtACBP2* was demonstrated to participate in stem cuticle formation and was expressed in all plant organs. Both Class II AtACBPs play overlapping roles by promoting ABA signalling during germination and seedling development. *AtACBP3* was reported to be involved in leaf senescence and was detected in floral organs and vegetative tissues. *AtCBP4* and *AtACBP5* showed expression in leaves and roots of mature plants (Xiao and Chye 2011, Du et al. 2016). In addition, various organs (leaf, stalk, silique and flower) were used to examine spatial expression of *AtACBP6,* as demonstrated by Chen et al. (2008).

Another five vegetative tissues were also tested including non-embryogenic and embryogenic calli, polyembryoids, and both male and female inflorescences (Fig. 5B). Two encoding Kelch motif ACBPs, *EgACBP6* and *EgACBP7* displayed higher expression in embryogenic calli and polyembryoids than in non-embryogenic calli. In *A. thaliana*, involvement of the Kelch motif containing ACBPs in such tissues has not yet been reported although their expression has been documented in flowers, siliques, mature plants and seedlings (Engeseth et al. 1996, Xiao and Chye 2011). RT-qPCR expression profiles showed their high expression in inflorescences, while GUS assays revealed expression of *AtACBP4* in pollen grains, and *AtACBP5* in microspores and tapetal cells (Du et al. 2016). *AtACBP4* and *AtACBP5* showed inversed expression in anther development, with *AtACBP5* being expressed at early stages in microspores, and AtACBP4 at later stages in pollen and endothecium (Ye et al. 2017). Intriguingly, *EgACBP3* also showed high expression in both male and female inflorescence (Fig. 5B), suggesting its involvement during flowering. This correlates well with *AtACBP3* which was observed to be expressed in stigma, indicating a potential role related to reproduction (Zheng et al. 2012).

When mesocarp and endosperm samples were examined (Fig. 5C), *EgACBP3* showed good expression at 10WAA in developing mesocarp, while *EgACBP1* gave gradually increased expression during the 10-20WAA period. For endosperm tissue, expression of *EgACBP1* and *EgACBP3* increased markedly between 10 and 15WAA, corresponding to the active period of oil accumulation. This suggests a role for these two EgACBPs during lipid metabolism and oil accumulation in both mesocarp and endosperm. The Class I small ACBPs have also been implicated in the synthesis and storage of TAG in Arabidopsis (Du et al. 2016), rice (Meng et al. 2011), sunflower (Aznar-Moreno et al. 2016) and oilseed rape (Yurchenko and Weselake 2011). As reported recently by Guo et al. (2019), overexpression of OsACBP2 (which is also a small ACBP) in transgenic rice increases its grain size as well as its oil content. For the Class III large ACBPs, while OsACBP5 has been noted for seed development in rice (Du et al. 2016, Meng et al. 2011), the Arabidopsis AtACBP3 was not thought to be involved (Du et al. 2016).

The significant expression of *EgACBP3* in mesocarp at 10WAA also suggests a potential role before high oil synthesis, as oil accumulation becomes rapid at later times (Sambanthamurthi et al. 2000). For the other *EgACBPs*, most had low expression in mesocarp and/or showed little increase during the whole oil accumulation period.

In the endosperm, *EgACBP1* and *EgACBP3* showed a notable increase in expression from 10 to 15WAA in contrast to the expression of most other *EgACBPs* which decreased (Fig. 5C). Since 10WAA is at the very start of the endosperm oil accumulation period and 15WAA is during rapid TAG biosynthesis, these data are consistent with a role for *EgACBP1* and *EgACBP3* in kernel oil production.

Expression profiles of EgACBPs during oil accumulation

In oil palm, the most important economic trait is the production of oil in the fruit mesocarp and endosperm. In the mesocarp, high rates of oil accumulation start around 15WAA and are complete by about 20WAA (Sambanthamurthi et al. 2000). In the endosperm, significant oil accumulation starts at about 12WAA and is essentially complete by 16WAA (Oo et al. 1985). In the mesocarp, *EgACBP1* and *EgACBP3* were well expressed as judged by transcriptome data (Fig. 6A). *EgACBP3* showed good expression throughout the whole developmental period, suggesting its involvement in both fruit ripening and senescence. For *EgACBP1*, increasing expression was observed up to the mature stage (22WAA). Its orthologue, *AtACBP6* was reported to have better binding with 16:0-CoA compared to 18:1-CoA (Xiao and Chye 2011) and it would be interesting to investigate if EgACBP1 had a substrate preference for 16:0-CoA as this would have important implications in efforts to increase the oleic acid content of palm oil at the expense of palmitic acid. *EgACBP2* and *EgACBP4* to *EgACBP7* showed relatively low expression, which did not alter significantly during mesocarp development. For the endosperm, *EgACBP4* showed high expression throughout seed development. *EgACBP1* and *EgACBP3* both demonstrated increased expression toward the end of oil accumulation (Fig. 6B). *EgACBP2* and *EgACBP5* to *EgACBP7* showed lower expression, which did not alter significantly during development.

When the mesocarp developmental time points were examined using RT-qPCR (Fig. 7), *EgACBP1* again showed increased expression during the active period of oil accumulation and elevated towards the later stages of development (15-22WAA), while *EgACBP3* demonstrated higher expression than most of the other *EgACBPs* throughout oil production (Fig. 7A). *EgACBP2*, *EgACBP4*, *EgACBP6* and *EgACBP7* showed constant and rather low expression throughout the mesocarp development as they had for the Illumina transcriptome data (Fig. 6A). Similar expression profiles were observed for *EgACBP5* in both methods although with higher expression in RT-qPCR (Fig. 7A). RT-qPCR data for endosperm development (Fig. 7B) agreed with the Illumina results (Fig. 6B) for *EgACBP1*, which revealed high expression at 15 and 18WAA, and *EgACBP3* had notable expression in mature kernel (18WAA). In addition, there was good agreement between the two methods for *EgACBP2*, *EgACBP6* and *EgACBP7,* which did not show any obvious correlation of expression with endosperm development (Fig. 7B). In contrast, both *EgACBP4* and *EgACBP5* demonstrated elevated expression only at 18WAA, as measured by RT-qPCR.

Although there was general agreement between the transcriptomic (Illumina) and the RT-qPCR results, some analyses differed, especially for the endosperm. RNA extraction at later development times is usually difficult and, indeed, from work with other mature seeds, we know that reproducibility of RT-qPCR data can be a problem. Thus, the elevated responses at 18WAA for *EgACBP4* and *EgACBP5* could be due to contaminants that were not completely purified from the RNA samples. For example, natural compounds such as carotenoids that emit fluorescence at the same wavelength (Roshchina 2008) as SYBR green are found in abundance in oil palm fruits. In addition, the expression of *EgACBP4* in the RT-qPCR data was relatively low compared to the transcriptome data (Fig. 6). More detailed examination of the raw reads mapped to *EgACBP4* showed that exon 1 had high levels of reads mapped to it compared to the other 4 exons. This increase appears to be due, at least in part, to raised levels of non-specific mapping of reads to exon 1 rather than an alteration in splicing.

Acyl-CoA pools during fruit development

In view of the predominant role of ACBPs in binding acyl-CoAs, we examined pools of the latter in both endosperm and fruit mesocarp during development. The results are shown in Fig. 8 and Supplemental File S1. Since there is a marked difference in the fatty acyl composition of the oils produced by palm fruit mesocarp (Fig. 8A) and endosperm (Fig. 8B) (Sambanthamurthi et al. 2000; Gunstone et al. 2007), it was expected that the acyl-CoA pools would reflect this. Fig. 8 shows the striking difference in the two pools. For the mesocarp (Fig. 8A), the acyl-CoA pool is dominated by the major acyl groups found in palm oil, namely palmitate, oleate and linoleate. There was a significant increase in oleoyl-CoA during development until the acyl-CoA composition near maturity (20WAA) reflected well the typical palm oil composition (44% palmitate, 40% oleate, 11% linoleate: Sambanthamurthi et al. 2000; Gunstone et al. 2007). During development, the mesocarp acyl-CoA pool showed significant reductions in the linolenoyl and very long chain acyl components. Presumably, this reflects the reduced need for thylakoid and surface wax components, respectively.

The endosperm (Fig. 8B) acyl-CoA pool is consistent with the need for synthesis of the short/medium chain constituents of palm kernel oil. Towards the end of the oil accumulation period (15WAA), the composition of the acyl-CoA pool (35% laurate, 10% myristate, 10% palmitate, 16% oleate) was nearing the final composition of palm kernel oil and was quite distinct from that of the mesocarp. Thus, the alterations in fruit tissue acyl-CoA pools during development were in excellent agreement with the supply of acyl groups needed for TAG accumulation in the mesocarp and endosperm tissues, respectively.

Conclusions

Seven actively expressed members of the *E. guineensis* ACBP family were identified and characterized. The results strengthen the hypothesis that plant ACBPs are non-redundant and that each ACBP appears to play a role, as reflected by its expression in various tissues. These initial findings with *E. guineensis* agree well with other observations made for the AtACBPs and OsACBPs (as reviewed in Du et al. 2016). In general, the *EgACBPs* were expressed in all tissues studied in a manner consistent with them having roles as constitutive genes. However, two *EgACBPs* were very highly expressed in specific tissues indicating their possible roles in tissue development. For example, *EgACBP3* showed high expression in inflorescences (Fig. 5B), while two Kelch motif-containing EgACBPs, *EgACBP6* and *EgACBP7,* showed significant expression in tissue culture samples. The significantly higher expression of *EgACBP6* and *EgACBP7* in embryogenic calli compared to non-embryogenic calli is interesting and suggests a role in embryogenesis. The role of *AtACBP1* and *AtACBP2* in development, specifically in callus-induction and embryogenesis, has been established (Chen et al. 2010). The *acbp1acbp2* double mutant not only failed in callus induction but was embryo lethal. However, it should be noted that both *AtACBP1* and *AtACBP2* belong to the ankyrin-repeat-containing Class II ACBPs while *EgACBP6* and *EgACBP7* are Class IV ACBPs. Hence, the data reported here indicates a possible new function for Class IV ACBPs in tissue culture.

In mesocarp and endosperm tissues, *EgACBP1* and *EgACBP3* showed high expression during oil synthesis, indicating their potential roles in transporting acyl-CoA esters during TAG accumulation. Because *EgACBP3* has a transmembrane domain, the two *EgACBPs* may have complementary roles with EgACBP3 also involved in facilitating Kennedy pathway enzyme activity. In comparison, *AtACBP6* the orthologue of *EgACBP1*, is also known to function in lipid exchange and lipid accumulation (Engeseth et al. 1996). Consistently, acyl-CoA pools of fatty acids as well as TAG synthesis during seed maturation in transgenic *A. thaliana* were reported to be affected following transformation with a transformed Class I *BnACBP* (Yurchenko et al. 2009). Yurchenko and Weselake (2011) also demonstrated the housekeeping role of this 10-kDa soluble plant ACBP in maintaining the intracellular acyl-CoA pool, based on *in vitro* studies and transgenic plant analyses. Yurchenko et al. (2014) provide further evidence that the expression of the 10 kDa *B. napus* ACBP (BnACBP) in the developing seeds of *A. thaliana* effectively altered the fatty acid composition of their seed oil and the acyl-CoA pool composition. Their data, as well as the results reported here, suggest specific experiments to delineate the role of EgACBP1 during oil accumulation in oil palm.

Here we have shown the characterisation of seven ACBPs in oil palm, as well as their differential expression in various tissues. Such data provide important background information for further examination of their function in oil palm. In particular, the correlation of EgACBP1 and EgACBP3 expression with oil accumulation in both mesocarp and endosperm tissues is significant for future investigations on the role of ACBPs during TAG biosynthesis in the pre-eminent commercial oil crop, *E. guineensis*.

Materials and Methods

Sample collection and expression analysis using RT-qPCR

All the samples (spear leaf, mature leaf, white root, primary root, lateral root, endosperm (8, 10, 12, 15 and 18WAA), mesocarp (5, 8, 10, 12, 15, 18, 20 and 22WAA), inflorescences (male and female) and tissue culture material (non-embryogenic and embryogenic calli, and polyembryoids)) (Supplemental File S1) were immediately frozen in liquid nitrogen (N_2) and stored at -80°C. Leaves (spear and mature), endosperm, mesocarp and inflorescences (female and male) were sampled from the Malaysian Palm Oil Board (MPOB) Research Station, Kluang, Johor, Malaysia. Roots (white, primary and lateral) were sampled from the oil palm nursery in MPOB, Bangi, Selangor, Malaysia. Tissue culture samples (non-embryogenic and embryogenic calli, and polyembryoids) were provided by the tissue culture laboratory at the Advanced Biotechnology and Breeding Centre, MPOB, Bangi, Selangor, Malaysia. For mesocarp and endosperm samples, two independent bunches were harvested for each developmental stage. All mesocarp and endosperm samples were collected from Angola *dura* palms while the remaining tissues samples were from *tenera* palms. Samples were ground to a fine powder in liquid N_2 using mortar and pestle, and total RNA extracted according to Ong et al. (2019). The samples were subjected to on-column RNase free DNase I treatment using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, USA, Valencia, CA). The quality and quantity of total RNA were assessed with a Nanodrop ND-1000TM spectrophotometer (Thermo Fisher Scientific, USA). The integrity of the RNA was checked using electrophoretic fractionation on an Agilent 2100 Bioanalyzer using RNA 6000 Nano LabChip (Agilent Technologies, CA).

RT-qPCR was performed according to Chan et al. (2014). The first-strand cDNAs were synthesized using the High-capacity cDNA Reverse-Transcription Kit according to the manufacturer's instructions (Applied Biosystems). The SYBR Green based RT-qPCR experiment was carried out using the Eppendorf Mastercycler® ep *realplex* (Eppendorf, Germany). The PCR program was as follows: 95 ^{o}C, 3 min for 1 cycle; followed by 95 ^{o}C, 3 sec, annealing temperature of primer pairs at 60 or 63 ^{o}C (Table 1), 20 sec for 40 cycles and a melting curve analysis at, 60 ^{o}C to 95 ^{o}C, for 15 sec with 0.4 ^{o}C temperature increment at each step. Each sample was analyzed in triplicate wells comprising 10 ng cDNA template, 0.2 µM of reverse primer, 0.2 µM of forward primer, 1X KAPA SYBR FAST Universal 2X qPCR Master Mix (KAPA Biosystems) in a final volume of 20 µl. Several potential reference genes were surveyed for use in RT-qPCR analysis. Two genes (TCONS-20183, TCONS-59914) (Supplemental File S1) (Morris et al. 2018) were found to be the best combination by GeNorm (Vandesompele et al. 2002) when compared to other potential reference genes including classical housekeeping genes (actin, GAPDH, tubulin, ubiquitin). Hence, these were used as reference genes for normalization of expression in the RT-qPCR analysis using Pfaffl (Pfaffl 2001).

Identification of EgACBPs and ACBD using bioinformatic tools

Identification of EgACBPs was carried out using an ACBD sequence (PF00887), acquired from PFAM (http://pfam.xfam.org/) (Finn et al. 2016). This domain was used as a query to identify a set of ACBD-containing proteins against oil palm Gene Models (Chan et al. 2017) using HMMSEARCH (https://www.ebi.ac.uk/Tools/hmmer/search/hmmsearch) (Finn et al. 2015). ACBD-containing proteins were classified based on the domain architecture of the conserved domains. The conserved domains were analyzed using SMART (http://smart.embl-heidelberg.de/) (Letunic and Bork 2018). Predictions of subcellular localizations of EgACBPs were performed using WoLF PSORT (https://wolfpsort.hgc.pc) (Horton et al. 2006). EgACBP sequences can be downloaded from PalmXplore (Sanusi et al. 2018) and Genomsawit (Rosli et al. 2014) portals. Corresponding protein sequences of EgACBPs and those of other organisms were aligned using ClustalW (Thompson et al. 1994) program to identify two conserved binding motifs in the ACBD. Analysis of motifs in ACBPs from *E. guineensis* and higher plants, namely *A. thaliana*, *O. sativa* and *B. napus*, was performed using MEME (Bailey et al. 2009).

Phylogenetic analysis

ACBPs of *E. guineensis* and higher plants (*A. thaliana*, *O. sativa* and *B. napus*) were used in phylogenetic analysis. Three non-plant reference organisms were added to the tree: *Caenorhabditis elegans* (roundworm), *Homo sapiens* (human) and *Mus musculus* (mouse). Protein sequences were acquired from the NCBI database (http://ncbi.nlm.nih.gov). Alignment was performed with ClustalW software (Thompson et al. 1994). The phylogenetic tree was generated using Neighbourhood Joining (NJ) methods (Saitou and Nei 1987) in MEGA7 software (Tamura et al. 2013).

Expression analysis using transcriptome data

Expression of *EgACBPs* across oil palm tissues from BioProject PRJNA201497 (leaf, root, seedling white root, pollen, mesocarp, endosperm) and PRJNA345530 (pith (1 day after anthesis, DAA), sepal (1DAA), fruit (1DAA), spikelet (1DAA), stalk (1DAA)) were determined by mapping Roche-454 sequencing transcriptome data of each tissue to the *E. guineensis* P5 genome (Singh et al. 2013) using the Tuxedo suite pipeline (Bowtie2.1.0, TopHat2.0.9 (Trapnell et al. 2009), Cufflinks 2.2.1 (Trapnell et al. 2011), Cuffmerge 2.2.1, CuffDiff 2.2.1). Each of two spear leaf and two root samples were collected from Deli *dura* palm (thick-shelled), MPOB 0.122/70 and *pisifera* palm (shell-less), MPOB 0.182/77. F17 leaf, pith, sepal, fruit, spikelet and stalk were also collected from *pisifera* palm, MPOB 0.182/77. Seedling white roots, mesocarp and endosperm were collected from *tenera* (Dura x Pisifera) palms (thin-shelled fruit). Two pollen (female sterile and female fertile) samples, were collected from *pisifera* palms.

To estimate the expression of *EgACBPs* in oil producing tissues, two independent Illumina RNA-Seq libraries each of fruit mesocarp and endosperm samples (Angola *dura* palm) were read mapped to the P5 genome by using the same Tuxedo suite pipeline. Fragments Per Kilobase of Transcript per Million mapped fragments (FPKM) for each library were calculated, and the mean FPKM from two biological replicates (each with three technical repeats) was measured. The mesocarp and endosperm sample collection was described in the first section of Materials and Methods.

Acyl-CoA profiling

Samples (mesocarp and endosperm) were frozen in liquid nitrogen, ground to powder and then extracted after Larson et al, (2001) for reverse-phase LC with electrospray ionisation tandem mass spectrometry (multi reaction monitoring: using a SCIEX 4000QTRAP instrument) in positive ion mode. LS-MS/MS MRM analysis followed the methods described in Haynes et al. (2008). Acyl-CoAs were separated using an Agilent 1200 LC system; Gemini C18 column, 2mm inner diameter, 150 mm with 5 µm particles. For the purpose of identification and quantification, standard acyl-CoA esters with chain lengths from C14 to C20 were synthesised from free acids or lithium salts (Sigma-Aldrich, St. Louis, Missouri, USA). Heptadecanoyl-CoA (ammonium salt) was used as an internal standard in each analytical run. Retention times were also confirmed using acyl-CoA standards from Avanti. Each tissue and development stage was measured using quintuplicate samples.

Supplemental data

Supplemental File S1. Location of EgACBPs in genome, protein sizes and ACBD location, list of tissues samples used for RT-qPCR, primer sequences of reference genes and acyl-CoA profiling data.

Supplemental Figure S1. Motif and conserved domain analysis on EgACBPs and its orthologues using Motif Multiple En for Motif Elicitation (MEME).

Funding

This work was supported by the Malaysian Palm Oil Board. J.L. Harwood and M-L Chye thank The Royal Society, UK for financial support (International Exchanges Grant: IE60011). R.P. Haslam is funded by the BBSRC Tailoring Plant Metabolism Institute Strategic Programme grant BBS/E/C/00010420 (Rothamsted Research).

Disclosures

The authors declare no conflict of interest.

Acknowledgments

We wish to thank the Director General of MPOB, Dr. Ahmad Parveez Ghulam Kadir, for permission to publish this paper. We would also like to extend our appreciation to the staff of Genomics, Tissue Culture and Transgenic Technology Laboratories of the Advanced Biotechnology and Breeding Centre for the supply of tissue samples and their assistance in RNA extraction for RT-qPCR. Special thanks also go to Dr. Benjamin Lau Yii Chung, Dr. Meilina Ong-Abdullah, Dr. Rajinder Singh, Dr. Nathan M. Springer and Dr. Nick Kent for their valuable advice.

Abbreviations

ACBD	acyl-CoA-binding domain
ACBP	acyl-CoA-binding protein
DAA	days after anthesis
DGAT	diacylglycerol acyltransferase
FPKM	Fragments Per Kilobase of transcript per Million mapped fragments
GUS	β-glucuronidase
LPAAT	lysophosphatidic acid acyltransferase
TAG	triacylglycerol
WAA	weeks after anthesis
WRI1	WRINKLED1

References

Akoh, C.C. (ed.) (2017) Food Lipids: chemistry, nutrition and biotechnology. CRC Press, Boca Raton.

Aznar-Moreno, J.A., Venegas-Calerón, M, Du, Z.Y., Garcés, R., Tanner, J.A., Chye, M.L., et al. (2016) Characterization of a small acyl-CoA-binding protein (ACBP) from *Helianthus annuus* L. and its binding affinities. *Plant Physiol. Biochem.* 102: 141-150.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., et al. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 37: W202-208.

Basiron, Y. (2005) Palm oil. In *Bailey's Industrial Oil and Fat Products*. Edited by Shahida, F. pp. 333-430. Wiley-Interscience, New York.

Bates, P.D., Stymne, S, and Ohlrogge, J. (2013) Biochemical pathways in seed oil synthesis. *Current Opin. Plant Biol.* 16: 358-364.

Bennet, V. and Chen, L. (2001) Ankyrins and cellular targeting of diverse membrane proteins to physiological sites. *Curr. Opin. Cell Biol*. 13: 61-67.

Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G.F., Drira, N., Ohlrogge, J.B., et al. (2011) Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc. Natl. Acad. Sci. USA* 108: 12527-12532.

Chan, P.L., Rose, R.J., Abdul Murad, A.M., Zainal, Z., Low, E.T.L., Ooi, L.C., et al. (2014) Evaluation of reference genes for quantitative real-time PCR in oil palm elite planting materials propagated by tissue culture. *PLoS One* 9: e99774.

Chan, K.L., Tatarinova, T.V., Rosli, R., Amiruddin, N., Azizi, N., Halim, M.A.A., et al. (2017) Evidence-based gene models for structural and functional annotations of the oil palm genome. *Biol. Direct* 12: 21.

Chen, G., Woodfield, H.K., Pan, X., Harwood, J.L. and Weselake, R.J. (2015) Acyl-trafficking during plant oil accumulation. *Lipids* 50: 1057-1068.

Chen, Q.F., Xiao, S. and Chye, M.L. (2008) Overexpression of the Arabidopsis 10-kilodalton acyl-CoA-binding protein ACBP6 enhances freezing tolerance. *Plant Physiol.* 148: 304-315.

Chen, Q.F., Xiao, S., Qi, W., Mishra, G., Ma, J., Wang, M., et al. (2010) The Arabidopsis acbp1 acbp2 double mutant lacking acyl-CoA-binding proteins ACBP1 and ACBP2 is embryo lethal. *New Phytol.* 186: 843-855.

Du, Z.Y., Arias, T., Meng, W. and Chye, M.L. (2016) Plant acyl-CoA-binding proteins: an emerging family involved in plant development and stress responses. *Prog. Lipid Res.* 63: 165-181.

Engeseth, N.J., Pacovsky, R.S., Newman, T. and Ohlrogge, J.B. (1996) Characterization of an acyl-CoA-binding protein from *Arabidopsis thaliana*. *Arch. Biochem. Biophys.* 331: 55-62.

Faergeman, N.J. and Knudsen J. (1997) Role of long-chain fatty acyl-CoA esters in the regulation of metabolism and cell signalling. *Biochem J.* 323: 1-12.

Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., et al. (2015) HMMER web server: 2015 update. *Nucleic Acids Res.* 43: W30-W38.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., et al. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44: D279-D285.

Gunstone, F.D., Harwood, J.L. and Dijkstra, A.J. (eds.) (2007) *The Lipid Handbook, third edition.* CRC Press, Boca Raton.

Guo, Z.H., Haslam, R.P., Michaelson, L.V., Yeung, E.C., Lung, S.C., Napier, J.A. and Chye, M.L. (2019) The overexpression of rice ACYL-CoA-BINDING PROTEIN2 increases grain size and bran oil content in transgenic rice. *Plant J*. doi: 10.1111/tpj.14503.

Harwood, J.L. (1996) Recent advances in the biosynthesis of plant fatty acids. *Biochim. Biophys. Acta* 1301: 7-56.

Harwood, J.L., Woodfield, H.K., Chen, G. and Weselake, R.J. (2017) Modification of oil crops to produce fatty acids for industrial applications. In *Fatty Acids: Chemistry, Synthesis and Applications*. Edited by Ahmad, M.U. pp. 187-236. Elsevier, London. doi.org/10.1016/B978-0-12-809521-8.00005-2.

Haynes, C.A., Allegood, J.C., Sims, K., Wang, E.W., Sullards, M.C. and Merril, A.H. (2008) Quantitation of fatty acyl-coenzyme as in mammalian cells by liquid chromatography-electrospray ionisation tandem mass spectrometry. *J. Lipid Res.* 49: 1113-1125.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J. and Nakai, K. (2017) WoLF PSORT: protein localization predictor. *Nucleic Acid Res*. 35: W585-W587.

Ibrahim, N.A. (2013) Characteristics of Malaysian palm kernel and its products. *J. Oil Palm Res.* 25: 245-252.

Knudsen, J., Faergeman, N.J., Skøtt, H., Hummer, R. and Børsting, C. (1994) Yeast acyl-CoA-binding protein: acyl-CoA binding affinity and effect on intracellular acyl-CoA pool size. *Biochem J.* 302: 479-485.

Kragelund, B.B., Andersen, K.V., Madsen, J.C., Knudsen, J., and Poulsen, F.M. (1993) Three dimensional structure of the complex between acyl-coenzyme A binding protein and palmitoyl-coenzyme A. *J. Mol. Biol.* 230: 1260-1277.

Kushairi, A., Loh, S.K., Azman, I., Hishamuddin, E., Ong-Abdullah, M., Noor Izuddin, Z.B.M., et al. (2018) Oil palm economic performance in Malaysia and R&D progress in 2017. *J. Oil Palm Res.* 30: 163-195.

Larson, T.R. and Graham, I.A. (2001) A novel technique for the sensitive quantification of acyl CoA esters from plant tissues. *Plant J*. 25: 115-125.

Letunic, I. and Bork, P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46: D493-496.

Leung, K.C., Li, H.Y., Xiao, S., Tse, M.H. and Chye, M.L. (2006) Arabidopsis ACBP3 is an extracellularly targeted acyl-CoA-binding protein. *Planta* 223: 871-881.

Li, H.Y. and Chye, M.L. (2003) Membrane localization of Arabidopsis acyl-CoA-binding protein ACBP2. *Plant Mol. Biol.* 51: 483-492.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J.E. (2000) Protein sorting: organelle biogenesis and protein secretion. *Mol. Cell Biol.* 4: 675-750.

Meng, W., Su, Y.C.F., Saunders, R.M.K. and Chye, M.L. (2011) The rice acyl-CoA-binding protein gene family: phylogeny, expression and functional analysis. *New Phytol.* 4: 1170-1184.

Meng, W. and Chye, M.L. (2014) Rice acyl-CoA-binding proteins OsACBP4 and OsACBP5 are differentially localized in the endoplasmic reticulum of transgenic Arabidopsis. *Plant Signal Behav*. 9(8): e29544.

Morris, P.E., Ooi, L.C.L, Ong, P.W., Kamaruddin, K., Marjuni, M., Amiruddin, M.D., et al. (2018) Transcription Quantitative Real-Time PCR (RT-qPCR) across mesocarp tissues of MPOB-Angola *Dura*. International Conference and Workshop on Comparative Genomics and Interactomics for Agriculture, 1-2 November 2018, Selangor, Malaysia.

Mosavi, L.K., Minor, D.L. Jr. and Peng, Z.Y. (2002) Consensus-derived structural determinants of the ankyrin-repeat motif. *Proc. Natl. Acad. Sci. USA* 99: 16029-16034.

Murphy, D.J. (ed.) (2005) Plant Lipids: Biology, Utilisation and Manipulation. Blackwell Publishing, Oxford.

Murphy, D.J. (2014) The future of oil palm as a major global crop: opportunities and challenges. *J. Oil Palm Res.* 26: 1-24.

Oo, K.C., Tah, S.K., Khor, H.T. and Ong, A.S.H. (1985) Fatty acid synthesis in the oil palm (*Elaeis guineensis*): incorporation of acetate by tissue slices of the developing fruit. *Lipids* 20: 205-210.

Ong, P.W., Chan, P-L. and Singh, R. (2019) Isolation of high quality total RNA from various tissues of oil palm (*Elaeis guineensis*) for reverse transcription quantitative real-time PCR (RT-qPCR). *J. Oil Palm Res.* In press.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acid Res*. 29: e45.

Raboanatahiry, N.H., Lu, G. and Li, M. (2015) Computational prediction of acyl-CoA-binding proteins structure in *Brassica napus*. *PLoS One* 10: e129650.

Ramli, U.S., Baker, D.S., Quant, P.A. and Harwood, J.L. (2002) Control analysis of lipid biosynthesis in tissue cultures from oil crops shows that flux control is shared between fatty acid synthesis and lipid assembly. *Biochem. J.* 364: 393-401.

Ramli, U.S., Salas, J.J., Quant, P.A. and Harwood, J.L. (2009) Use of metabolic control analysis to give quantitative information on control of lipid biosynthesis in the important oil crop, *Elaeis guineensis* (oil palm). *New Phytol.* 184: 330-339.

Ramli, U.S., Sambanthamurthi, R., Omar, A.R., Parveez, G.K.A., Manaf, M.A.A., Abrizah, O., et al. (2012) The isolation and characterisation of oil palm (*Elaeis Guineensis* Jacq.) β-ketoacy-acyl carrier protein (ACP) synthase (KAS) II cDNA. *J Oil Palm Res*. 24: 1480–91.

Roshchina V.V. (2008) Flourescing world of plant secreting cells (1st Edition). Science Publishers, New Hampshire.

Rosli, R., Chan, P.L., Chan, K.L., Amiruddin, N., Low, E.T.L., Singh, R., et al. (2018) *In silico* characterization and expression profiling of the diacylglycerol acyltransferase gene family (DGAT1, DGAT2, DGAT3 and WS/DGAT) from oil palm, *Elaeis guineensis*. *Plant Sci.* 275: 84-96.

Rosli, R., Halim, M.A.A., Chan, K.L., Azizi, N., Jayanthi, N., Nik Shazana, N.M.S., et al. (2014) Genomsawit website. MPOB Information Series No. 134.

Saitou, N and Nei M. (1987) The neighbor-joining method: A new method for reconstructing phylogenic trees. *Mol. Biol. Evol.* 4: 406-425

Salas, J.J., Sanchez, J., Ramli, U.S., Manaf, A.M., Williams, M. and Harwood, J.L. (2000) Biochemistry of lipid metabolism in olive and other oil fruits. *Prog. Lipid Res*. 39: 151-180.

Sambanthamurthi, R., Sundram, K. and Tan, Y. (2000) Chemistry and biochemistry of palm oil. *Prog. Lipid Res.* 39: 507-558.

Sanusi, N.M.S., Rosli, R., Halim, M.A.A., Chan, K.L., Nagappan, J., Azizi, N., et al. (2018) PalmXplore: oil palm gene database. *Database* 2018: 1-9.

Singh, R., Ong-Abdullah, M., Low, E.T.L., Manaf, M.A., Rosli, R., Nookiah, R., et al. (2013) Oil palm genome sequence reveals divergence of infertile species in old and new worlds. *Nature* 500: 355-359.

Siti Nor Akmar, A., Cheah, S.C., Ooi, L.C.L., Sambanthamurthi, R. and Murphy, D.J. (1999) Characterization and regulation of the oil palm (*Elaeis guineensis*) stearoyl-ACP desaturase genes. *J. Oil Palm Res. Special Issues* 1-17.

Stanley, J.C. (2008) The nutritional reputation of palm oil. *Lipid Technol.* 20: 112-114.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.

Tang, T.S. and Pantzaris, T.P. (2017) *Pocketbook of oil palm uses.* Malaysian Palm Oil Board, Malaysia.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence. *Nucleic Acids Res.* 22: 4673–4680.

Tranbarger, T.J., Dussert, S., Joët, T., Argout, X., Summo, M., Champion, A., et al. (2011) Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening and functional specialization in lipid and carotenoid metabolism. *Plant Physiol.* 156: 564-584.

Trapnell, C., Pachter, L. and Salzberg, S.L. (2009) Discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., et al. (2011) Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat. Biotechnol.* 28: 511-515.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 3: research0034.1-research0034.11.

Wan Omar, W.S., Willis, L.B., Rha, C.K., Sinskey, A.J., Ramli, U.S., Abdul Masani, M.Y., et al. (2008) Isolation and utilization of acetyl-CoA carboxylase from oil palm (*Elaeis guineensis*) mesocarp. *J. Oil Palm Res.* 2: 97-107.

Weselake, R.J., Taylor, D.C., Rahman, M.H., Shah, S., Laroche, A., McVetty, P.B.E., et al. (2009) Increasing the flow of carbon into seed oil. *Biotechnol. Adv.* 27: 866-878.

Weselake, R.J., Woodfield, H.K., Field, C.J. and Harwood, J.L. (2017) Production of edible oils through metabolic engineering. In *Food Lipids*. Edited by Akoh, C.C. pp. 973-995. CRC Press, Boca Raton.

Xiao, S. and Chye, M.L. (2009) An Arabidopsis family of six acyl-CoA-binding proteins has three cytosolic members. *Plant Physiol. Biochem.* 47: 479-484.

Xiao, S. and Chye, M.L. (2011) New roles for acyl-CoA-binding proteins (ACBPs) in plant development, stress responses and lipid metabolism. *Prog. Lipid Res.* 50: 141-151.

Yang, S. and Bourne, P.E. (2009) The evolutionary history of protein domains viewed by species phylogeny. *PLoS One* 4: e8378.

Ye, Z.W., Xu, J., Shi, J., Zhang, D. and Chye, M.L. (2017) Kelch-motif containing acyl-CoA-binding proteins AtACBP4 and AtACBP5 are differentially expressed and function in floral lipid metabolism. *Plant Mol. Biol.* 93: 209-225.

Yurchenko, O.P., Nykiforuk, C.L., Moloney, M.M., Ståhl, U., Banaś, A., Stymne, S., et al. (2009) A 10-kDa acyl-CoA-binding protein (ACBP) from *Brassica napus* enhances acyl exchange between acyl-CoA and phosphatidylcholine. *Plant Biotechnol. J.* 7: 602-610.

Yurchenko, P., Singer, S.D., Nykiforuk, C.L., Gidda, S., Mullen, R.T., Moloney, M.M., et al. (2014) Production of a *Brassica napus* low-molecular mass acyl-coenzyme A-binding protein in Arabidopsis alters the acyl-coenzyme A pool acyl composition of oil in seeds. *Plant Physiol.* 165: 550-560.

Yurchenko, O.P. and Weselake, R.J. (2011) Involvement of low molecular mass soluble acyl-CoA-binding protein in seed oil biosynthesis. *New Biotechnol.* 28: 97-109.

Zheng, S.X., Xiao, S. and Chye, M.L. (2012) The gene encoding Arabidopsis acyl-CoA-binding protein 3 is pathogen inducible and subject to circadian regulation. *J. Exp. Bot.* 63: 2985-3000.

Table 1. *EgACBP* specific primers for RT-qPCR

ACBP genes	Primer sequences (5'-3')	Amplicon length (bp)	Annealing temperature (^oC)	Amplification efficiency (%)	Regression coefficient (R^2)
EgACBP1	F-CCAACCATGAATGATGCTTCCTA R-TAAACCCAAGCAGAACCAGTGAG	143	60	90	0.9987
EgACBP2	F-CCACCTTGTGAAACCCGATTCCG R-ACCCTAGAAAGTCGCGGAGGTTC	101	60	92	0.9984
EgACBP3	F-TGAGGAGAGGAGGGAAGGATTGG R-CTTCTTCTTCCACCTCGCTAGCC	125	60/63	74	0.9932
EgACBP4	F-TGGTGGTGGGATCTGGAAAAGGA R-ACCCACATGATCACCTTGAGTGC	119	60	92	0.9966
EgACBP5	F-GTGGCTGAGGTGGAGATTGGAAG R-GTTCATTCTCGTTCTCCGGGGTC	106	60	91	0.9932
EgACBP6	F-CCAGAATCACCAATTGTGCTGCC R-AGAAGGATCTTTGGTGTGCCCAC	105	60	100	0.9909
EgACBP7	F-CAGCAAATGGACCAGTTGGCATG R-ACAGTAGGCTCCACAGTGACACT	126	60	95	0.9976

Fig. 1. Schematic domain structures of *A. thaliana* and *E. guineensis* ACBPs. *E. guineensis* ACBPs are designated as EgACBP1-EgACBP7. The acyl-CoA-binding domain (ACBD) (brown), ankyrin-repeat motif (red), Kelch motif (green), transmembrane domain (yellow), coiled-coil region (purple) and signal peptide (orange) based on SMART analysis are labelled accordingly.

Fig. 2. Alignment of the conserved acyl-CoA-binding domain (ACBD) in four classes of ACBPs. The ACBD is indicated in a box with two conserved motifs; YKQA and KWDAW (each boxed "----") with their corresponding PFAM HMM Logos shown below the alignments. These motifs correspond to the acyl-CoA-binding site shown in Fig. 1 that are essential in binding acyl-CoA esters, and were proposed to be conserved in all four species. Their respective positions are indicated. (A) Class I ACBP orthologues, EgACBP1; AtACBP6, **AEE31396**; OsACBP1, **BAG86980**; OsACBP2, **BAG86809**; OsACBP3, **ABF97253**; BnACBP6, **BnaAnng25690D**, **BnaA05g36060D**, **BnaA08g07670D**, **BnACnng15340D;** (B) Class II ACBP orthologues, EgACBP2; AtACBP1, **AED96361**; AtACBP2, **AEE85391**; OsACBP4, **BAF16206**; BnACBP1, **BnaA02g10270D**, **BnaC02g44810D**; BnACBP2, **BnaA01g16660D**, **BnaC01g20440D;** (C) Class III ACBP orthologues, EgACBP3; EgACBP4; EgACBP5; AtACBP3, **AEE84874**; OsACBP5, **BAG93201**; BnACBP3, **BnaA01g13710D**, **BnaC01g16110D**, **BnaA03g46540D**, **BnaC07g38820D;** and (D) Class IV ACBP orthologues, EgACBP6; EgACBP7; AtACBP4, **AEE74237**; AtACBP5, **AED93708**; OsACBP6, **ABF99748**; BnACBP4, **Ais76194.1**, **Ais76199.1**, **Ais76201.1**, **Ais76196.1**, **Ais76195.1**, **Ais76200.1**; BnACBP5, **Ais76197.1**, **Ais76198.1**. Species abbreviations: Eg: *Elaeis guineensis*, At: *Arabidopsis thaliana*, Os: *Oryza sativa*, Bn: *Brassica napus*.

Fig. 3. Phylogenetic tree of *E. guineensis* ACBPs plus three higher plants and three other non-plant reference organisms. The amino acid sequences of all the selected proteins were aligned using the ClustalW program and subjected to phylogenetic analysis by the Neighbour-Joining method using MEGA7 software.

Fig. 4. Expression of *E. guineensis ACBPs* in different tissues of oil palm based on transcriptome data (Roche-454). Tissues tested were (A) spear leaf and root from *dura* 0.212/70; spear leaf, F17 leaf and root from *pisifera* 0.182/77; seedling white root from *tenera* (Dura x Pisifera); pollen from female fertile and sterile palms (*pisifera*); (B) 10 and 15 weeks after anthesis (WAA) mesocarp (*tenera*); developing endosperm at 10 and 15WAA (*tenera*); and pith, sepal, fruit, spikelet and stalk at 1 day after anthesis (DAA) (all from *pisifera* 0.182/77). The expression was normalized using Fragments Per Kilobase Million mapped fragments (FPKM). No expression was detected (ND) for *EgACBP1* in root (*pisifera*); for *EgACBP3* in endosperm at 10WAA; for *EgACBP4* in F17 leaf (*pisifera*), root (*pisifera*) and mesocarp at 10WAA; for EgACBP5 in root (dura). *EgACBP*s expression was not detected in pollen (female fertile), except *EgACBP5*. In pollen (female sterile), no expression of *EgACBPs* was detected except for *EgACBP5* and *EgACBP6*. Note that *dura* palms have thick-shelled kernels while *pisifera* have shell-less kernels (see Materials and Methods). F17 refers to leaves of frond 17.

Fig. 5. Expression of *E. guineensis ACBPs* in different oil palm tissues as monitored by RT-qPCR. Tissues tested were (A) spear and mature leaves; white, primary and lateral roots; (B) non-embryogenic and embryogenic calli; polyembryoids; female and male inflorescences; (C) developing mesocarp at 10, 15 and 20 weeks after anthesis (WAA); and developing endosperm at 10 and 15WAA. All samples (leaves, roots, tissue culture materials, inflorescences) were collected from *tenera* palm. Mesocarp and endoperm samples were from Angola *dura* palm. Spear leaf was used as a reference for comparison of expression. Y-axis shows average relative expression of *EgACBPs* calculated using Pfaffl algorithm against the reference genes (TCONS-20183 and TCONS-59914). Standard deviation (SD) is represented by error bars. Three technical replicates were used for each tissue sample.

Fig. 6. Expression of *E. guineensis ACBPs* in oil-producing tissues (A) developing mesocarp and (B) developing endosperm based on transcriptome data (Illumina). The expression was normalized using Fragments Per Kilobase Million mapped fragments (FPKM). Mean FPKM and standard deviation (SD) from two biological replicates, each with three technical replicates were calculated. The results shown as means $\pm$ SD. Single biological samples were used for 8 weeks after anthesis (WAA) and 15WAA (EgACBP1 only) endosperm (marked with *). All mesocarp and endosperm samples were collected from Angola *dura* palm.

Fig. 7. Expression of *E. guineensis ACBPs* in oil-producing tissues (A) developing mesocarp and (B) developing endosperm using RT-qPCR. Y-axis shows average relative expression of *EgACBPs* calculated using Pfaffl algorithm against the reference genes (TCONS-20183 and TCONS-59914). Standard deviation (SD) is represented by error bars. Three technical replicates were used for each of two biological samples. Single biological samples were used for 5 and 8 weeks after anthesis (WAA) mesocarp (marked with *), while two biological samples for 10, 12, 15, 18, 20 and 22WAA. For endosperm, a single biological sample was used for 8WAA (marked with *), while two biological samples for 10, 12, 15 and 18WAA. Three technical replicates were analysed for each. All mesocarp and endosperm samples were collected from Angola *dura* palm.

Fig. 8. Acyl-CoA pools in (A) mesocarp and (B) endosperm tissues from developing oil palm fruits (Angola *dura*). Data shown as means +/- standard deviations (SD) from five replicates.

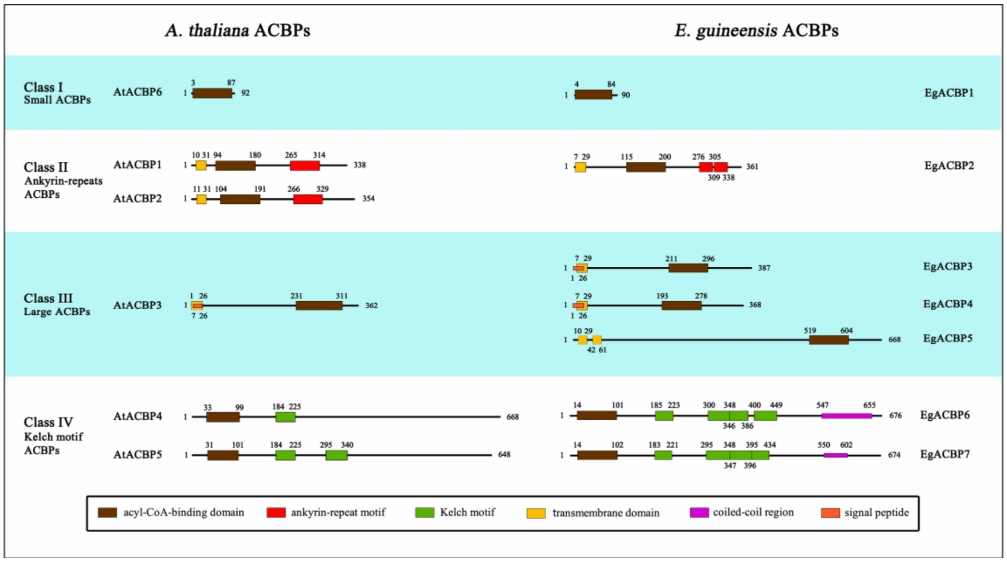

Fig. 1. Schematic domain structures of A. thaliana and E. guineensis ACBPs. E. guineensis ACBPs are designated as EgACBP1-EgACBP7. The acyl-CoA-binding domain (ACBD) (brown), ankyrin-repeat motif (red), Kelch motif (green), transmembrane domain (yellow), coiled-coil region (purple) and signal peptide (orange) based on SMART analysis are labelled accordingly.

233x129mm (300 x 300 DPI)

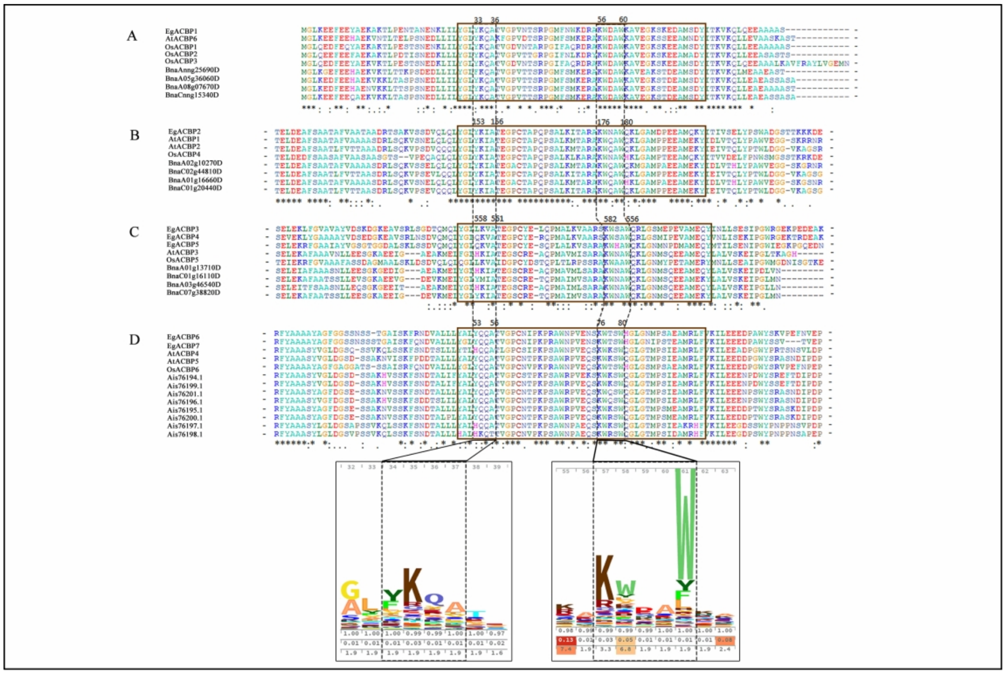

Fig. 2. Alignment of the conserved acyl-CoA-binding domain (ACBD) in four classes of ACBPs. The ACBD is indicated in a box with two conserved motifs; YKQA and KWDAW (each boxed "----") with their corresponding PFAM HMM Logos shown below the alignments. These motifs correspond to the acyl-CoA-binding site shown in Fig. 1 that are essential in binding acyl-CoA esters, and were proposed to be conserved in all four species. Their respective positions are indicated. (A) Class I ACBP orthologues, EgACBP1; AtACBP6, AEE31396; OsACBP1, BAG86980; OsACBP2, BAG86809; OsACBP3, ABF97253; BnACBP6, BnaAnng25690D, BnaA05g36060D, BnaA08g07670D, BnACnng15340D; (B) Class II ACBP orthologues, EgACBP2; AtACBP1, AED96361; AtACBP2, AEE85391; OsACBP4, BAF16206; BnACBP1, BnaA02g10270D, BnaC02g44810D; BnACBP2, BnaA01g16660D, BnaC01g20440D; (C) Class III ACBP orthologues, EgACBP3; EgACBP4; EgACBP5; AtACBP3, AEE84874; OsACBP5, BAG93201; BnACBP3, BnaA01g13710D, BnaC01g16110D, BnaA03g46540D, BnaC07g38820D; and (D) Class IV ACBP orthologues, EgACBP6; EgACBP7; AtACBP4, AEE74237; AtACBP5, AED93708; OsACBP6, ABF99748; BnACBP4, Ais76194.1, Ais76199.1, Ais76201.1, Ais76196.1, Ais76195.1, Ais76200.1; BnACBP5, Ais76197.1, Ais76198.1. Species abbreviations: Eg: Elaeis guineensis, At: Arabidopsis thaliana, Os: Oryza sativa, Bn: Brassica napus.

233x155mm (300 x 300 DPI)

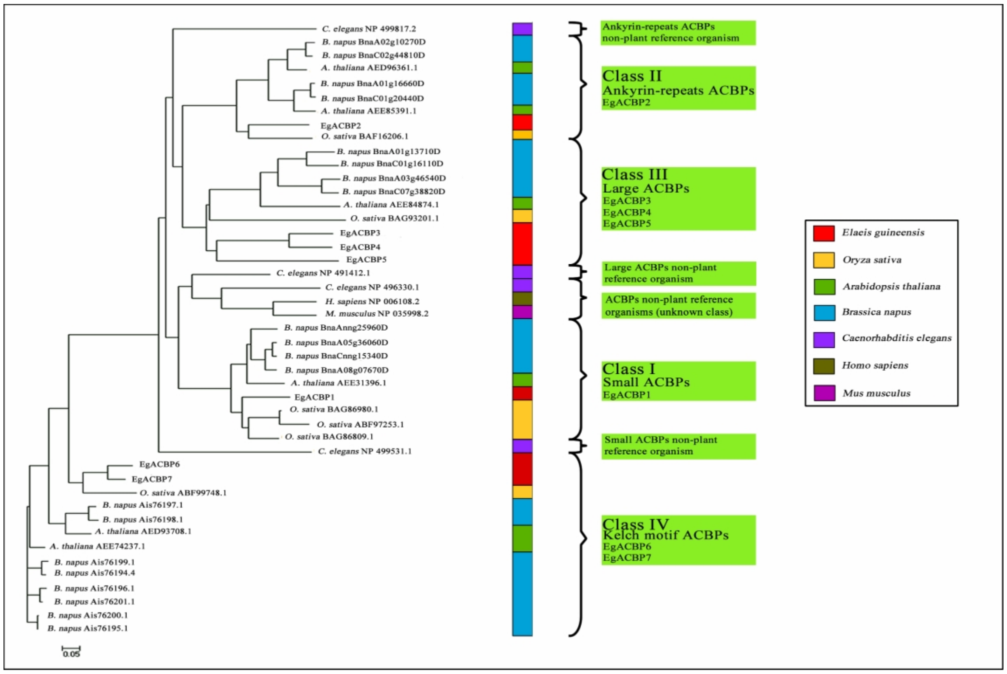

Fig. 3. Phylogenetic tree of E. guineensis ACBPs plus three higher plants and three other non-plant reference organisms. The amino acid sequences of all the selected proteins were aligned using the ClustalW program and subjected to phylogenetic analysis by the Neighbour-Joining method using MEGA7 software.

233x155mm (300 x 300 DPI)

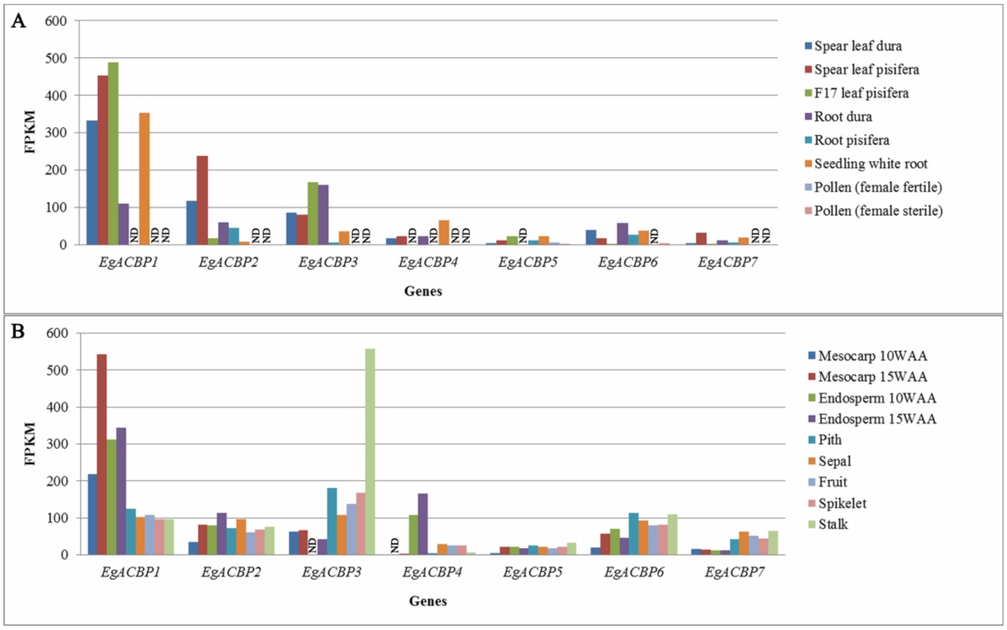

Fig. 4. Expression of E. guineensis ACBPs in different tissues of oil palm based on transcriptome data (Roche-454). Tissues tested were (A) spear leaf and root from dura 0.212/70; spear leaf, F17 leaf and root from pisifera 0.182/77; seedling white root from tenera (Dura x Pisifera); pollen from female fertile and sterile palms (pisifera); (B) 10 and 15 weeks after anthesis (WAA) mesocarp (tenera); developing endosperm at 10 and 15WAA (tenera); and pith, sepal, fruit, spikelet and stalk at 1 day after anthesis (DAA) (all from pisifera 0.182/77). The expression was normalized using Fragments Per Kilobase Million mapped fragments (FPKM). No expression was detected (ND) for EgACBP1 in root (pisifera); for EgACBP3 in endosperm at 10WAA; for EgACBP4 in F17 leaf (pisifera), root (pisifera) and mesocarp at 10WAA; for EgACBP5 in root (dura). EgACBPs expression was not detected in pollen (female fertile), except EgACBP5. In pollen (female sterile), no expression of EgACBPs was detected except for EgACBP5 and EgACBP6. Note that dura palms have thick-shelled kernels while pisifera have shell-less kernels (see Materials and Methods). F17 refers to leaves of frond 17.

244x152mm (300 x 300 DPI)

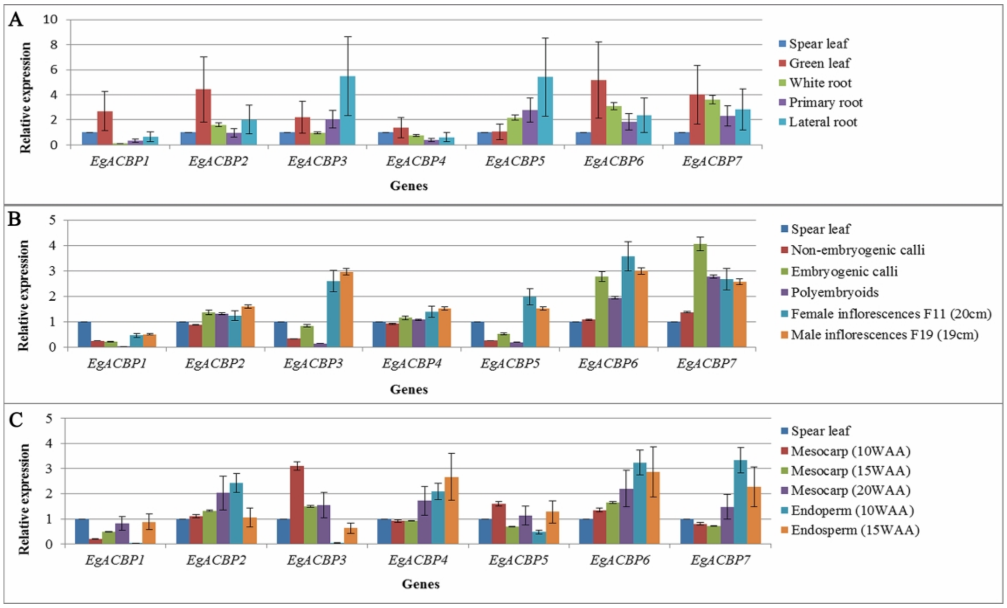

Fig. 5. Expression of E. guineensis ACBPs in different oil palm tissues as monitored by RT-qPCR. Tissues tested were (A) spear and mature leaves; white, primary and lateral roots; (B) non-embryogenic and embryogenic calli; polyembryoids; female and male inflorescences; (C) developing mesocarp at 10, 15 and 20 weeks after anthesis (WAA); and developing endosperm at 10 and 15WAA. All samples (leaves, roots, tissue culture materials, inflorescences) were collected from tenera palm. Mesocarp and endoperm samples were from Angola dura palm. Spear leaf was used as a reference for comparison of expression. Y-axis shows average relative expression of EgACBPs calculated using Pfaffl algorithm against the reference genes (TCONS-20183 and TCONS-59914). Standard deviation (SD) is represented by error bars. Three technical replicates were used for each tissue sample.

244x146mm (300 x 300 DPI)

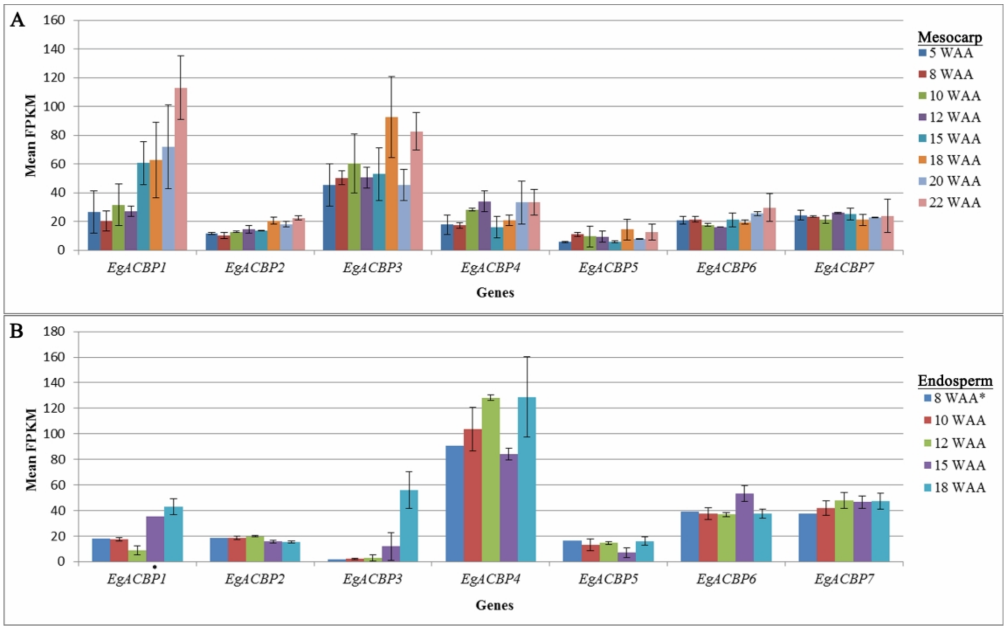

Fig. 6. Expression of E. guineensis ACBPs in oil-producing tissues (A) developing mesocarp and (B) developing endosperm based on transcriptome data (Illumina). The expression was normalized using Fragments Per Kilobase Million mapped fragments (FPKM). Mean FPKM and standard deviation (SD) from two biological replicates, each with three technical replicates were calculated. The results shown as means + SD. Single biological samples were used for 8 weeks after anthesis (WAA) and 15WAA (EgACBP1 only) endosperm (marked with *). All mesocarp and endosperm samples were collected from Angola dura palm.

244x152mm (300 x 300 DPI)

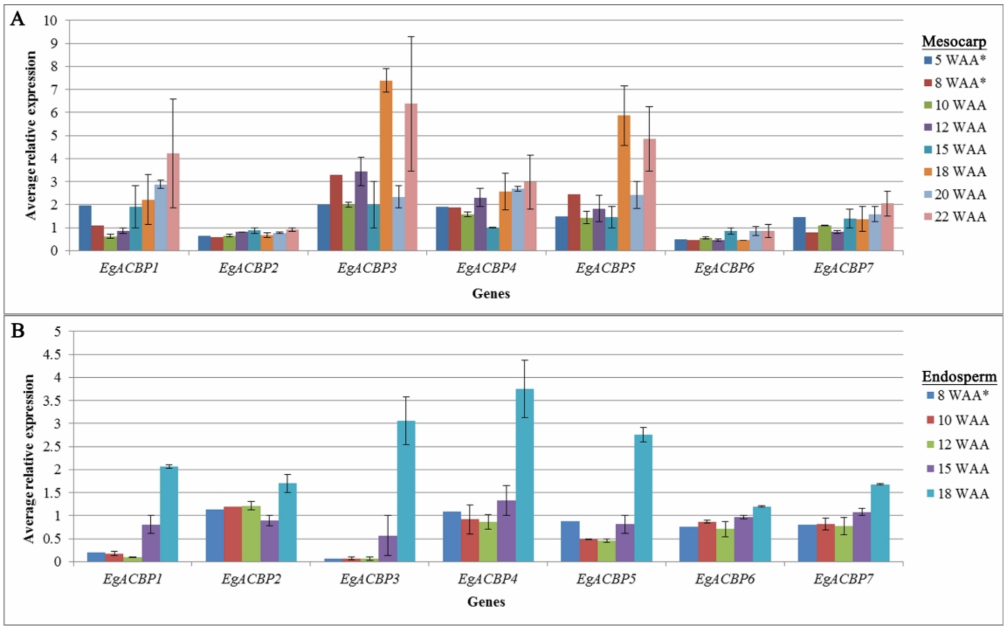

Fig. 7. Expression of E. guineensis ACBPs in oil-producing tissues (A) developing mesocarp and (B) developing endosperm using RT-qPCR. Y-axis shows average relative expression of EgACBPs calculated using Pfaffl algorithm against the reference genes (TCONS-20183 and TCONS-59914). Standard deviation (SD) is represented by error bars. Three technical replicates were used for each of two biological samples. Single biological samples were used for 5 and 8 weeks after anthesis (WAA) mesocarp (marked with *), while two biological samples for 10, 12, 15, 18, 20 and 22WAA. For endosperm, a single biological sample was used for 8WAA (marked with *), while two biological samples for 10, 12, 15 and 18WAA. Three technical replicates were analysed for each. All mesocarp and endosperm samples were collected from Angola dura palm.

244x152mm (300 x 300 DPI)

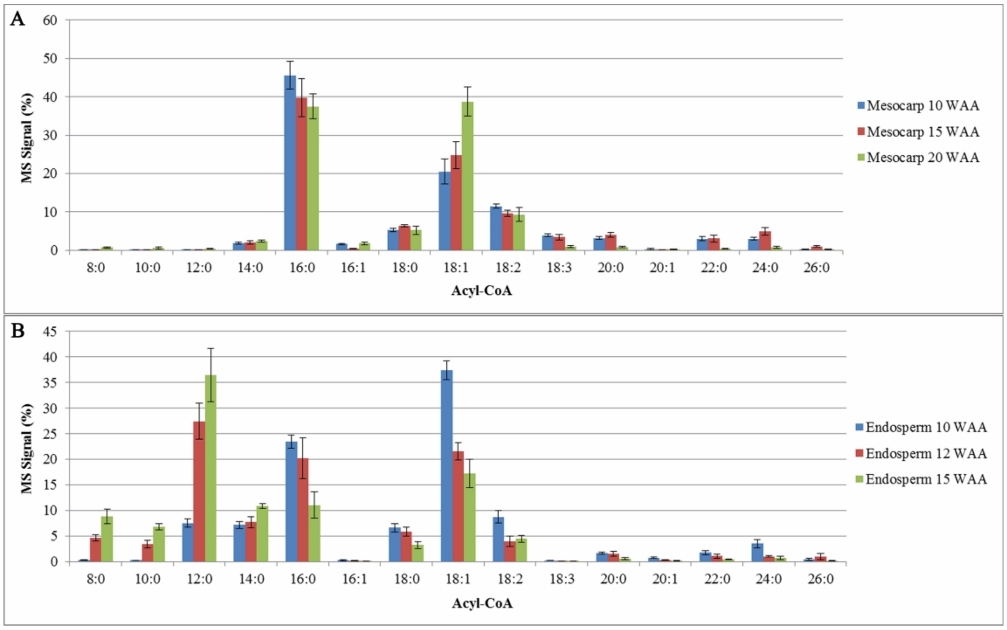

Fig. 8. Acyl-CoA pools in (A) mesocarp and (B) endosperm tissues from developing oil palm fruits (Angola dura). Data shown as means +/- standard deviations (SD) from five replicates.

244x152mm (300 x 300 DPI)