Functional study of oil assembly pathway in oil palm (Elaeis guineensis Jacq.) fruits

Yijun Yuan

université de BORDEAUX

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

Ecole Doctorale des Sciences de la vie et de la santé

Spécialité : Biologie végétale

Par Yijun YUAN

Etude de l'assemblage des acides gras en huile chez le palmier à huile (*Elaeis guineensis* Jacq.)

Sous la direction de : Vincent ARONDEL

Soutenue le 21 décembre 2016

Membres du jury :

M. HOOKS, Mark	Professeur à l'Université de Bordeaux	Président
M. SCHALLER, Hubert	Directeur de recherche CNRS à l'Université de Strasbourg	Rapporteur
Mme LI-BEISSON, Yonghua	Chercheur CEA au CEA de Cadarache	Rapporteur
Mme LAGER, Ida	Chercheur à la Swedish University of Agricultural Sciences	Examinateur

Abstract

Oil palm is the highest oil-yielding crop-plant, accounting for approximately 40% of the total world vegetable oil production. The fruit accumulates oil, made of triacylglycerol (TAG) molecules, in both mesocarp and kernel with totally different fatty acid profiles. Fatty acids are assembled into oil through Kennedy pathway in the endoplasmic reticulum, which is complicated by editing processes involving phosphatidylcholine metabolism. To investigate oil assembly in oil palm, we use lipidomics as a tool to analyze different populations of palm to search for TAG structural diversity, and to further characterize changes in lipid content and composition in mesocarp and kernel during fruit ripening. We used yeast two-hybrid system (split ubiquitin) to test protein-protein interactions for almost all the enzymes (32) involved in oil assembly pathway, and we demonstrated 241 interactions, including 132 strong interactions, 73 medium interactions and 36 weak interactions. Our results suggest that all enzymes might assemble into one or several complexes that may form metabolons. In addition, different isoforms of enzymes showed distinct interaction profiles, providing hints for future studies. Moreover, we also characterized the *in vivo* function of a putative acyltransferase (designated EgWSD1-like) possibly involved in oil assembly and the three major diacylglycerol acyltransferase (DGAT) isoforms of palm mesocarp in the mutant yeast H1246, which is devoid of neutral lipid synthesis. EgWSD1-like only shows wax ester synthase activity in yeast, while three EgDGATs all can restore TAG biosynthesis in yeast with different substrate specificities.

Keywords: *Elaeis guineensis* Jacq., oil assembly, lipidomic analysis, protein/protein interaction, diacylglycerol acyltransferase, wax ester synthase, *Saccharomyces cerevisiae*

Résumé

Le palmier à huile est la première culture oléagineuse, avec environ 40% de la production mondiale, et son fruit accumule deux huiles de composition très différente dans le mesocarpe et l'amande. Chez les plantes, les acides gras sont assemblés en huile dans le réticulum endoplasmique, ceci par la voie dite de Kennedy à laquelle s'ajoutent des mécanismes d'édition impliquant le métabolisme de la phosphatidylcholine. Nous avons utilisé les outils de la lipidomique pour analyser la variabilité au sein de différentes populations de palmier ainsi que pour caractériser l'accumulation d'huile durant le développement du mésocarpe et de l'amande. Puis, nous avons entrepris de tester, dans le système du double hybride de levure, les interactions entre toutes les enzymes de la voie de Kennedy et celles responsables des mécanismes d'édition, et mis en évidence 241 interactions, dont 132 sont fortes, 73 moyennes et 36 faibles. Ces résultats suggèrent que ces enzymes pourraient s'assembler en complexes supra-moléculaires susceptibles de former des métabolons. Certaines isoformes d'une même enzyme ont des profils d'interaction distincts, ce qui ouvre des perspectives pour de futures recherches. De plus, nous avons caractérisé, par expression fonctionnelle dans un mutant de levure dépourvu de TAG, une acyltransférase présumée (EgWSD1-like) ainsi que les trois formes majeures de diacylglycérol acyltransférases du mésocarpe. EgWSD1-like ne restaure que l'activité de synthèse d'esters de cire dans le mutant, tandis que les trois DGAT complémentent toutes la déficience en TAG du mutant, avec d'apparentes spécificités distinctes vis-à-vis des acides gras.

Mots-clés : *Elaeis guineensis* Jacq., assemblage des acides gras en huile, analyses lipidomiques, interactions proteine/proteine, diacylglycerol acyltransférase, synthase d'esters de cires, *Saccharomyces cerevisiae*

Résumé de la thèse en Français :

Introduction :

Le palmier à huile (*Elaeis guineensis* Jacq.) est la première culture oléagineuse, avec environ 40% de la production mondiale et est également l'oléagineux le plus productif à l'hectare (>4 tonnes/an). Le fruit accumule deux huiles de composition très différente : une huile avec des triacylglycérols (TAG) contenant des acides gras à longue chaine (essentiellement palmitique et oléique) dans le mésocarpe et une autre contenant des acides gras à chaine moyenne (essentiellement laurique) dans l'amande. Le mésocarpe est le tissu d'origine végétale le plus riche en huile connu (environ 90% du poids sec). Toutes ces caractéristiques font du palmier à huile un modèle de choix pour l'étude de la synthèse de l'huile. Celle-ci se déroule en trois étapes, dans les organes hétérotrophes des plantes supérieures : (1) le saccharose qu'exportent les tissus photosynthétiques est converti en acétyl-CoA, essentiellement par la glycolyse et l'action de la pyruvate décarboxylase plastidiale (2) les acides gras sont synthétisés dans le plaste à partir d'acétyl-CoA (3) les acides gras sont exportés vers le réticulum endoplasmique où ils sont assemblés en phospholipides et/ou en huile. Cette dernière étape, qui est l'objet de la présente étude, se fait par la voie dite de Kennedy et consiste en deux acylations successives de glycerol-3-P, donnant l'acide phosphatidique qui est déphosphorylé en diacylglycérol (DAG). Le DAG est acylé une troisième fois pour donner le TAG. Acide phosphatidique et DAG sont également les précurseurs directs des phospholipides membranaires. De plus, il existe des mécanismes d'édition permettant d'enrichir l'huile en acides gras polyinsaturés ; ces mécanismes impliquent le métabolisme de la phosphatidylcholine, molécule qui peut notamment être un intermédiaire direct de la synthèse d'huile sous l'action d'une enzyme spécifique des plantes appelée phosphatidylcholine :DAG choline-P transférase. Malgré la forte interpénétration des voies de synthèse de l'huile et de la phosphatidylcholine, ces deux familles de molécules parviennent à conserver une composition en acides gras distincte et les mécanismes permettant de contrôler cette composition restent encore largement inconnus, malgré l'importance qu'a le degré

d'insaturation des lipides dans les propriétés physicochimiques et fonctionnelles des membranes biologiques.

J'ai utilisé les outils de la lipidomique (chromatographie en phase liquide couplée à la spectrométrie de masse) pour analyser la variabilité de la composition en huile de différentes populations de palmier, ainsi que pour caractériser l'accumulation d'huile durant le développement du mésocarpe et de l'amande, et les espèces moléculaires de phosphatidylcholine, DAG et TAG ont été déterminées (Chapitre 1). Des études de transcriptomique des enzymes de la voie de Kennedy ont montré de faibles changements au cours du développement, suggérant plutôt que la régulation de l'assemblage des acides gras en huile pourrait se faire au niveau post-transcriptionnel. Nous avons émis l'hypothèse que des complexes multi-enzymatiques pourraient diriger le flux d'acides gras spécifiquement vers les TAGs et d'autres vers la phosphatidylcholine, ce qui pourrait expliquer en partie la différence de composition en acides gras. J'ai entrepris de tester cette hypothèse en analysant, dans le système du double hybride de levure (« split-ubiquitine »), les interactions entre les enzymes de la voie (Chapitre 2). De plus, j'ai caractérisé, par expression fonctionnelle dans un mutant de levure dépourvu de TAG, une acyltransférase présumée (EgWSD1-like) ainsi que les trois formes majeures de diacylglycérol acyltransférases (DGAT) du mésocarpe (Chapitre 3).

Chapitre 1 : Analyse de l'huile produite dans le mésocarpe et l'amande au cours du développement et dans différentes populations de palmiers.

J'ai profité d'une étude en cours sur la variabilité de la composition en acides gras de l'huile de palme pour étudier des génotypes représentatifs de cette variabilité afin d'en analyser l'huile par chromatographie liquide couplée à la spectrométrie de masse. Onze arbres, appartenant aux populations Deli Dabou, la Mé et résultant de croisements La Mé x Deli Dabou, ont été sélectionnés. L'analyse des acides gras de l'huile par chromatographie en phase gazeuse montre que les acides palmitique et oléique varient entre 28 et 48% et 32 et 57%, respectivement. L'analyse des espèces moléculaires des TAG par chromatographie liquide couplée à la masse montre que la

plupart des espèces majeures varient légèrement ; une analyse en composante principale montre que les teneurs en acide palmitique et oléique varient de façon opposée sur l'axe principal de variation et sont corrélées avec les espèces de TAG contenant, soit exclusivement l'acide gras considéré, soit deux de cet acide gras sur trois. L'analyse de l'acide gras en position 2 montre un biais en faveur des acides gras insaturés avec 18 carbones, et notamment de l'acide oléique, ceci quel que soit le génotype considéré et sa composition globale en acides gras. Donc, la variation de la composition en acides gras de l'huile impacte peu la composition et la structure des TAG et n'est donc probablement pas liée à l'assemblage des acides gras en huile.

L'analyse de la composition en espèce moléculaires de TAG du mésocarpe au cours du développement du fruit ne montre pas, non plus, de différences importantes. Par contre, les TAG de l'amande peuvent être regroupés en plusieurs catégories selon la présence ou l'absence d'acides gras en chaine moyennes. Cela suggère l'existence de deux systèmes conduisant, l'un à la synthèse d'huile avec seulement des acides gras à chaines moyennes et l'autre avec seulement des acides gras à longue chaine. Ces deux systèmes sont exprimés de façon variable selon le stade de développement. L'amande pourrait donc constituer un modèle d'étude intéressant pour comprendre les mécanismes permettant, au sein du même organe, la synthèse d'huile de composition très différente.

Chapitre 2 : Analyse des interactions protéines/protéines par le système double hybride de levure pour les enzymes impliquées dans l'assemblage des acides gras en huile du palmier.

Pour analyser l'organisation supra-moléculaire des enzymes d'assemblage de l'huile, j'ai tout d'abord cloné les isoformes des gènes liés à l'assemblage des acides gras en huile, et représentées dans le mésocarpe, par RT-PCR. J'ai ensuite testé deux-à-deux les interactions protéines-protéines par le système « split-ubiquitine » de double hybride chez la levure *Saccharomyces cerevisiae*. Sur plus de 600 paires testées, j'ai mis en évidence 241 interactions, dont 132 sont fortes, 73 moyennes et 36 faibles. Ces résultats indiquent que l'ensemble des enzymes de la voie (enzymes de la

voie de Kennedy vers la synthèse de TAG et de phosphatidylcholine) sont susceptibles de s'assembler en un complexe supra-moléculaire. Les enzymes spécifiques de la synthèse des TAGs (DGAT1, DGAT2, phospholipid :diacylglycerol acyltransferase) montrent toutes des interactions avec les enzymes d'édition, ce qui suggère que des métabolons, incluant enzymes de synthèse de TAG et enzymes d'édition des acides gras destinée spécifiquement à la synthèse des TAG, sont susceptibles d'exister. Les interactions s'étendent à des enzymes liées à la désaturation des acides gras, comme FAD2, ainsi qu'à des enzymes du métabolisme des acyl-CoA (acyl-CoA synthetase, protéines fixatrices d'acyl-CoA). Certaines isoformes d'une même enzyme impliquée dans la synthèse de phosphatidylcholine interagissent avec des DGAT et d'autres non, ce qui suggère que ces isoformes pourraient être dédiées aux mécanismes d'éditions impliqués spécifiquement dans la synthèse des TAG.

J'ai ensuite testé des protocoles d'isolement de membranes à partir de mésocarpe frais ou congelé, prélevé à différents stades de développement, et j'ai obtenu des fractions dont l'analyse protéomique a montré qu'elles étaient enrichies en enzymes de synthèse des TAGs. Ces membranes seront utilisées par la suite pour tenter de mettre en évidence les complexes multi-enzymatiques hypothétiques, ceci après solubilisation et analyse par électrophorèse native en présence de bleu de Coomassie.

Chapitre 3 : Expression hétérologue d'acyltransférases dans la levure *S. cerevisiae*

J'ai exprimé trois DGAT (une DGAT1 et deux DGAT2) dans un quadruple mutant de levure qui ne synthétise pas de TAG. L'analyse des lipides montre que l'expression de chaque enzyme permet de rétablir un niveau de TAG comparable à celui de la levure sauvage. L'analyse de la composition en acides gras des TAG suggère que les 3 DGAT ont une spécificité différente vis-à-vis de la composition en acides gras de leurs substrats, tandis que les compositions en diacylglycérol et phosphatidylcholine sont peu affectées. Des travaux seront réalisés pour étudier la spécificité de substrats de ces enzymes *in vitro,* à partir de membranes isolées de ces

souches de levure. EgWSD1-like ne restaure que l'activité de synthèse d'esters de cire dans le mutant.

Conclusions :

Les travaux de lipidomique indiquent peu de variations dans la composition et la structure des TAGs selon le stade de développement du mésocarpe ou selon la nature du génotype et suggèrent l'existence de deux systèmes distincts pour la synthèse de l'huile dans l'amande. Les études d'interactions indiquent que toutes les enzymes de la voie d'assemblage des TAG pourraient s'assembler en complexes supra-moléculaires qui pourraient former des métabolons. Enfin, les trois DGATs exprimées majoritairement dans le mésocarpe présentent des spécificités de substrat différentes en ce qui concerne leur composition en acides gras.

ACKNOWLEDGEMENTS

I would like to thank all members of my thesis committee, Mark HOOKS, Hubert SCHALLER, Yonghua LI-BEISSON and Ida LAGER, for accepting to judge my work.

I am very grateful to Jean-Jacques BESSOULE for permitting me to work in his laboratory, and to China Scholarship Council (CSC) for the thesis scholarship it awarded to me.

To my supervisor Vincent ARONDEL, thank you a lot for taking good care of me all the time whether it concerned work or life in France. Otherwise, it would have been much harder for me to survive here without speaking French.

Many thanks to Laetitia FOUILLEN for her very important involvement in the lipidomic part of my thesis and to Frederic DOMERGUE, thank you for your guidance and patience. I am also very grateful to Fred BEISSON and Bertrand LÉGERET for solving an important lipid contamination problem and to Marina for statistics analyses.

I thank all the people in the laboratory who advised me and helped me with my experiments, especially Francois, Didier, Jerome, Valerie.

I would also like to thank all laboratory members for the good “ambiance “ in the laboratory, especially the students Camille, Claireline, Houda, Julien, William, Charlotte, Paul, Nicolas, Aurélie, and also Pathy, Stephanie, Agnès, Magali, Florence, Denis, Amélia, Guillaume.

I am extremely thankful to my previous professors: Dongdong LI, Yusheng ZHENG and Yinhua CHEN, who recommended me to study aboard and accepted to guarantee my scholarship. You made another choice for my life which I never thought before.

I would like to specially thank Yonghua LI-BEISSON. It is very lucky for me that I met you at the beginning of my study here. Even we met very few times, you are more like a relative for me in this foreign country.

I am very grateful to my boyfriend and my family for constantly supporting me.

These three years will be one of the most impressive memories in my life. Thank all the people I know here, if I miss someone, please forgive me.

Abbreviations

AAPT	Aminoalcoholphosphotransferase
ABA	Abscisic acid
ABI3	ABSCISIC ACID INSENSITIVE3
ACBP	Acyl-CoA binding protein
ACCase	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
BSTFA	*N,O*-bis(trimethylsilyl) trifluoroacetamide with trimethylchlorosilane
CCT	Choline-phosphate cytidylyltransferase
CDP	Cytidinediphosphate
CDP-DAG	CDP-diacylglycerol
CDP-DAGS	CDP-diacylglycerol synthase
CoA	Coenzyme A
CTP	Cytidine 5'-triphosphate
Cub	C-terminal half of ubiquitin
DAG	Diacylglycerol
DEPC	Diethylpyrocarbonate
DGAT	Diacylglycerol acyltransferase
DGD	Digalactosyl diacylglycerol
DGDG	Digalactosyldiacylglycerol
DGPP	Diacylglycerol pyrophosphate
DNA	Deoxyribonucleic acid
DTT	1,4-Dithiothreitol
ECT	Ethanolamine cytidylyltransferase
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
EST	Expressed sequence tag
FA	Fatty acid
FAD2	Oleate desaturase
FAD3	Linoleate desaturase
FAEE	Fatty acid ethyl ester
FAIE	Fatty acid isoamyl ester
FAME	Fatty acid methyl ester
FAR	Fatty acyl reductase
FAS	Fatty acid synthase
Fat	Acyl-ACP thioesterase
FAX1	FATTY ACID EXPORT 1
FFA	Free fatty acid
FID	Flame ionization

FUS3	FUSCA3
GC	Gas chromatography
G3P	Glycerol-3-phosphate
GPAT	Glycerol-3-phosphate acyltransferase
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
IE	Isoamyl ester
KAS	Ketoacyl-ACP synthase
KO	Knockout
LACS	Long chain acyl-CoA synthetase
LB	Luria-Bertani
LC	Long-chain
LCFA	Long-chain fatty acid
(HP)LC	(High-performance) Liquid chromatography
LDAP	Lipid droplet associated protein
LEC	LEAFY COTYLEDON
LPA	Lyso-phosphatidic acid
LPAT	Lyso-phosphatidic acid acyltransferase
LPC	Lyso-phosphatidylcholine
LPCAT	Lyso-phosphatidylcholine acyltransferase
LPEAT	Lyso-phosphatidylethanolamine acyltransferase
LPGAT	Lyso-PG acyltransferase
LPP	Lipid phosphate phosphatase
LPSAT	Lyso-PS acyltransferase
MAG	Monoacylglycerol
MAT	Acyl carrier protein malonyltransferase
MBOAT	Membrane-bound O-acyltransferase
MCFA	Medium-chain fatty acid
MGAT	Monoacylglycerol acyltransferase
MGD	Monogalactosyl diacylglycerol
MRM	Multiple reaction monitoring
MS	Mass spectrometry
NubG	The mutated N-terminal half of ubiquitin
NubI	N-terminal half of ubiquitin
OD	Optical Density
OPPP	Oxidation pentose phosphate pathway
PA	Phosphatidic acid
PAH	Phosphatidic acid phosphohydrolase
PAP	PA phosphatase
PC	Phosphatidylcholine
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDAT	Phospholipid:DAG acyltransferase
PDCT	PC:DAG cholinephosphotransferase
PDH	Pyruvate dehydrogenase

PE	Phosphatidylethanolamine
PEG	Polyethylene glycol
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PPP	Pentose phosphate pathway
PS	Phosphatidylserine
PUFA	Polyunsaturated fatty acid
PVPP	Polyvinylpolypyrolidone
RNA	Ribonucleic acid
RNAi	RNA interference
RPKM	Reads per kilobase per million ESTs
RT-PCR	Reverse transcription polymerase chain reaction
RuBP	Ribulose 1,5-bisphosphate
SAD	Stearoyl-ACP desaturase
SC	Synthetic complete medium
-TL	SC lacking tryptophan and leucine
-HTL	SC lacking histidine, tryptophan and leucine
-AHTL	SC lacking adenine, histidine, tryptophan and leucine
SC-U	SC lacking uracil
SDS	Sodium dodecyl sulfate
SE	Sterol ester
SOC	Super Optimal Broth
TAE	Tris-acetate-EDTA
TAG	Triacylglycerol
TCA	Tricarboxylic acid
TLC	Thin layer chromatography
UBPs	Ubiquitin specific proteases
VLC	Very long chain
VLCFA	Very long chain fatty acid
WAP	Weeks after pollination
WE	Wax ester
WRI	Wrinkled
WSD1	Wax ester synthase
WS	Wax ester synthase
Y2H	Yeast two-hybrid system

Fatty acids (number of carbon atoms : number of double bonds)

8:0	caprylic acid	**18:0**	stearic acid
10:0	capric acid	**18:1**	oleic acid
12:0	lauric acid	**18:2**	linoleic acid
14:0	myristic acid	**18:3**	linolenic acid
16:0	palmitic acid	**20:0**	Arachidic acid

TABLE OF CONTENTS

Chapter 1. Introduction

1.1 Brief introduction on lipids

While other components of living organisms, such as proteins, nucleic acids and carbohydrates are defined based on their chemical structures, lipids are defined based on their physical properties: lipids are those molecules that are insoluble in water and soluble in organic solvents. Therefore, lipids comprise a few families of compounds that are chemically unrelated. This definition makes sense in view of the function of most lipids, which are main components of biological membranes, as membranes are hydrophobic/amphiphilic structures that separates /delimitates aqueous compartments. Here, we will focus exclusively on lipids that contain fatty acids (FAs).

1.1.1 Definitions, structures and functions of lipids

Fatty acids are usually abbreviated using two numbers separated by a colon: first refers to number of carbon atoms in the chain and second to number of double bonds carried by the chain. For example, oleic acid, which contains 18 carbon atoms and one double bond is abbreviated 18:1. Fatty acids are very rarely abundant as such in living organisms. They are usually esterified to other molecules to form complex lipids. Most common are glycerolipids: two fatty acid chains are esterified to position *sn*-1 and position *sn*-2 of glycerol, while position *sn*-3 maybe occupied by a polar head that contains (phospholipids), or not, a phosphate group, or by a third fatty acid (Triacylglycerol or TAG), as illustrated in Fig.1-1 (Buchanan *et al.*, 2000). There are other types of lipids that contain fatty acids, such as steryl esters (fatty acid esterified to a sterol molecule), wax esters (fatty acid esterified to fatty alcohol) or sphingolipids (fatty acid amidified to a serine backbone) for example. Major fatty acids in plants glycerolipids have a chain length of 16 or 18 carbons with 1 to 3 double bonds, including saturated 16:0 and 18:0, monounsaturated 18:1, and polyunsaturated 18:2 and 18:3 (Ohlrogge & Browse, 1995; Bates & Browse, 2012). Other fatty acids with longer chains and carrying other functional groups, such as hydroxyl groups, can be integrated in sphingolipids, or surface lipids (waxes, cutin) or suberin (Pollard *et al.*, 2008; Bernard & Joubès, 2013). In a limited number of plant families can be found

unusual fatty acids (Dyer *et al*., 2008); some of them contains hydroxyl (*Ricinus communis*; Lin *et al*., 2003) or epoxy groups (*Vernonia galamensis*; Baye *et al*., 2005), others are shorter (*Cocos nucifera*; Dyer *et al*., 2008), etc. These fatty acids are usually found in TAGs exclusively, not in membranes glycerolipids. Aside from variability of fatty acids with regards to chain length and saturation level, complex lipids differ by nature of head group, nature of fatty acid at position *sn*-1 and *sn*-2, so that there are thousands of possible different species of these glycerolipids. Species of glycerolipids are usually designated by numbers summing up both total carbon atoms and double bonds of their fatty acid components. For example, a phosphatidylcholine (PC) molecule that contains one stearic (18:0) and one linoleic (18:2) fatty acid will be designated PC 36:2. One needs to note that a PC molecule with two oleic acids (18:1) will also be designated PC 36:2.

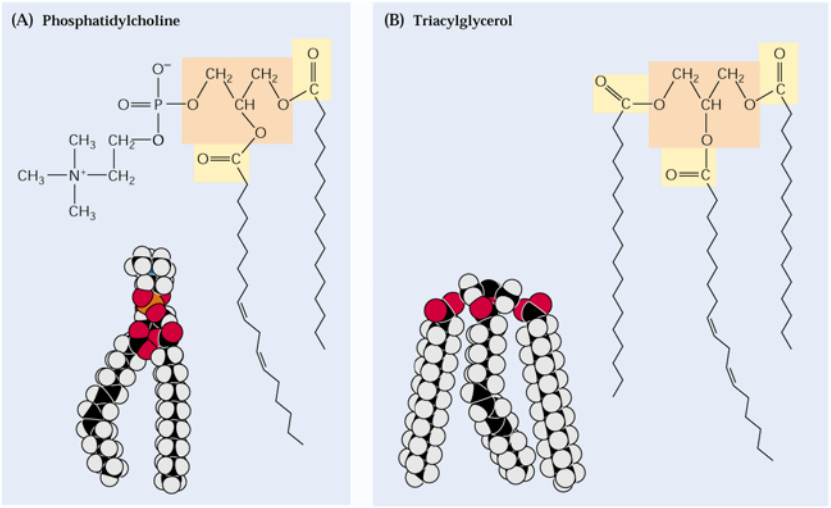

Fig.1-1 Space-filling and conformational models of (A) the phospholipid phosphatidylcholine and (B) triacylglycerol (Fig.10.5 (Page 463) from Buchanan *et al*., 2000). The ester linkages are highlighted in yellow, and the glycerol backbone is in orange.

Saturated fatty acids have much higher melting points than unsaturated fatty acids of corresponding size. For example, the melting point of stearic acid (18:0) is 69 ℃ while that of oleic, linoleic and linolenic acids are 13.4, -5 and -11 ℃, respectively. This property has major impacts on membrane properties, as membranes needs to be fluid to properly carry out their functions. Fatty acid composition is key to plant sensitivity to external temperature and *Arabidopsis* mutants with less unsaturated fatty

acids become very sensitive to low temperatures when compared to wild type (Ohlrogge & Browse, 1995). More generally, polyunsaturated fatty acids (PUFAs) confer fluidity and flexibility to cellular membranes and affect many cellular and physiological processes including cold adaptation and survival, modulation of ion channels, endocytosis/exocytosis, pollen formation, pathogen defense, chloroplast development in plants, and activities of membrane-associated enzymes that are sensitive to the biophysical properties of lipid membranes (Wallis *et al*., 2002).

Together with nature of polar head group, unsaturation levels of glycerolipid can impact the shape of these lipids, which will affect the supramolecular structure into which these lipids may assemble. One can distinguish bilayer forming lipids, such as PC and digalactosyl diacylglycerol (DGD) and non-bilayer forming lipids, such as phosphatidylethanolamine (PE) or monogalactosyl diacylglycerol (MGD). These self-assembly properties are important to maintain membrane integrity and to induce membrane budding/fusion events for example (Garab *et al*., 2016).

At last, lipids can also be signaling molecule and, aside from hormones such as jasmonate and oxilipins, derived from PUFAs and brassinosteroids, derived from sterols, phosphatidic acid is a well-studied second messenger involved in several transduction pathways (Munnik, 2001; Testerink & Munnik, 2005; Wang *et al*., 2006).

Therefore, given the physiological importance of lipids, it is clearly essential that cells control tightly lipid composition of its membranes.

1.1.2 Lipids and sub-cellular compartments

While subcellular membranes can widely differ with regards to lipid composition (Fig.1-2; Buchanan *et al*., 2000), their lipid content is largely conserved throughout the plant kingdom. Endoplasmic reticulum (ER) and closely derived membranes contain mostly phospholipids, while plastid photosynthetic membranes contain galactolipids such as MGD and DGD (polar headgroup is made of one or two

galactose residue), and also sulfolipids (polar headgroup is a sulfated sugar) (Fig.1-2; Buchanan *et al*., 2000). Interestingly, upon phosphate deprivation, DGD replaces in part PC molecules outside of plastids, and is transported from the chloroplast to several other organelles and membranes, such as plasma membrane, tonoplast, and mitochondria (Jouhet *et al*., 2007; Li-Beisson *et al*., 2013). Phosphatidylglycerol (PG) is found mostly in plastid, while a derivative, cardiolipid, is exclusively found in mitochondria and sphingolipids and sterols are mostly found in the plasma membrane. It is still poorly known how the cell can regulate and control the lipid composition of its subcellular membranes.

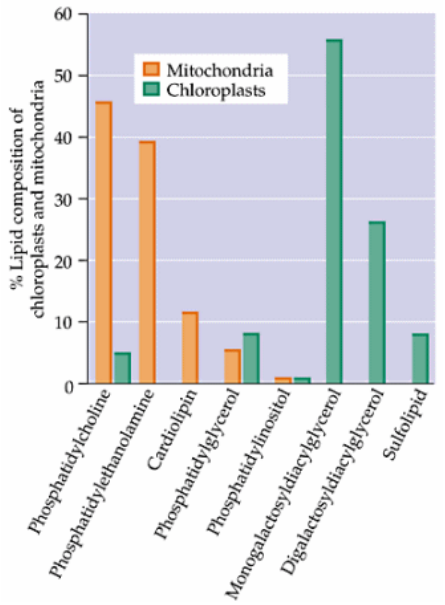

Fig.1-2 Comparison of the lipid compositions of chloroplasts and mitochondria (Fig.10.42 (Page 500) from Buchanan *et al*., 2000).

In higher plants, fatty acid synthesis occurs in plastid only and yields saturated (16:0 and 18:0) and mono-unsaturated (18:1) species. Then, fatty acids can be used by the plastid to make its own lipids such as PG, MGD, DGD and sulfolipids. The fatty acids esterified to these lipids can be further desaturated to di-(16:2 and 18:2) and tri-(16:3 and 18:3) enoic acids by plastid desaturases. Fatty acids can also be exported to the ER where they are incorporated into PC and 18:1 can also be, while esterified to PC, further desaturated to 18:2 and 18:3 by ER desaturases. Chloroplast pathway to glycerolipid synthesis and desaturation is qualified as “prokaryotic pathway” and ER

pathway to phospholipid desaturation as “eukaryotic pathway”. Frequently, the diacylglycerol backbone (DAG, two fatty acids esterified to a glycerol molecule) of PC synthesized and desaturated by the eukaryotic pathway is used by the chloroplast to synthesize chloroplast lipids (Benning, 2009). The two types of lipids can be distinguished because lipids made in the plastid contain exclusively 16-carbon fatty acids at *sn*-2 position while ER lipids contain 18-carbon fatty acids at same position.

1.1.3 Triacylglycerol

TAGs, which are the focus of our study, are made in the ER (very small amounts can be found in plastids) and derive from the eukaryotic pathway. These lipids do not assemble into bilayers and are stored as droplets into organelles called oil- or lipid bodies. Seed oil bodies are best studied. They bud from a region of the ER where both TAGs and oleosins, a class of structural proteins specific to seed oil bodies, are synthesized. They are made of a TAG core (about 95%) and a monolayer of phospholipids in which are embedded amphipathic proteins including oleosins, steroleosins and caleosins (Huang, 1996). Oleosins are considered to be structural proteins that prevent coalescence of lipid bodies. Contrarily to what was found in yeast, no enzyme of TAG biosynthesis has ever been demonstrated in plant lipid bodies. Therefore, it is still considered in plants that TAG synthesis occurs in the ER only (except for very small synthesis in plastid).

During post-germinative growth, TAGs in oil bodies will be hydrolyzed to fatty acids by SDP1 lipase (Eastmond, 2006). Fatty acids will be then transferred to glyoxysomes where they will undergo beta-oxidation, yielding acetyl-CoA molecules that will enter glyoxylic cycle and serve ultimately to synthesize sugars through neoglucogenesis, fueling post-germinative growth before photosynthesis fully operates (Graham, 2008).

Very little is known on plant non-seed oil bodies. Recently, oil bodies from avocado mesocarp have been characterized and found to contain no oleosin but lipid

droplet associated proteins (LDAP), closely related to proteins previously associated to rubber particles in Hevea tree (Gidda *et al*., 2016). It is thought that LDAP play a structural role similar to that of oleosins. Again, proteomic data do not present evidence for any TAG biosynthetic enzymes. In oil palm, a proteomic study performed upon fruit development showed that LDAP accumulation curve matches almost exactly that of oil (Gidda *et al*., 2013). Lipid droplets that contain TAGs can also be found in tissues that store no oil. For example, they can be found in leaves but amount of TAGs represent about 1% of total lipids, except in physiological situations such as senescence (Troncoso-Ponce *et al*., 2013), when lipids, as well as other cell constituents, are being recycled to be absorbed by other parts of plants before leaf abscission.

Plant oils are an important commodity and world plant oil market represents about 150 billion dollars a year. It is used mostly for human consumption (80%) but also in oleochemistry, to make coatings, lubricants, detergents etc, and also biofuels. Consumption increased by about 5% a year, including 1.5% for biofuels only, the rest being mostly linked to increased demand in emerging countries such as China and India. The economical interest of plant oil influences both fundamental and applied research and one of the most prominent axis of research concerns the transgenic production of unusual fatty acids that might be used by oleochemical industry (Napier, 2007). For example, important work has been carried out on ricinoleic acid production in transgenic *Arabidopsis* and this research has led to interesting findings on the editing mechanisms responsible for oil fatty acid composition (Bates *et al*., 2014).

TAG synthesis is the focus of our study and the rest of introduction part will be devoted to the pathway leading to TAG synthesis.

1.2 Fatty acids for oil biosynthesis in plant

In plant, *de novo* fatty acid (FA) synthesis occurs in the plastid, including chloroplasts in green tissues and proplastids or other types of plastid in non-green

tissues (Chapman & Ohlrogge, 2012; Li-Beisson *et al*., 2013). The acetyl group of acetyl-CoA provides the 2-carbon building block for fatty acid synthesis and a fatty acid chain results from successive condensation of acetyl groups, from which oxygen atoms are subsequently removed to yield an aliphatic chain. The elongating fatty acid chain is esterified to a protein called acyl carrier protein (ACP). One double bond may be introduced at delta 9 position of 18:0-ACP. Acyl-ACPs can serve as substrate to a soluble acyltransferase localized in plastid, leading to synthesis of lysophosphatidic acid that will serve to synthesize plastid lipids by the prokaryotic pathway. Otherwise, acyl chains destined to lipid assembly in the ER (eukaryotic pathway) are cleaved off from ACP. In most cases are released 16:0, 18:0 and 18:1 fatty acids, which are subsequently esterified to coenzyme A (CoA) to yield acyl-CoA. Acyl-CoAs are then exported to ER. Before giving detailed information on this process, summarized in Fig.1-3 (Chapman & Ohlrogge, 2012), we will briefly talk about origin of carbon destined to synthesis of acetyl-CoA. Because externally-provided radioactive acetate is immediately and almost exclusively incorporated into fatty acids, it has been long thought that acetic acid was precursor to acetyl-CoA destined to fatty acid synthesis. However, recent *in vivo* studies clearly indicate that pyruvate, not acetate, is direct precursor to acetyl-CoA, *via* the plastidial enzyme pyruvate dehydrogenase (PDH) which catalyzes the oxidation decarboxylation reaction generating CO_2, acetyl-CoA and NADH (Johnston *et al*., 1997). A theoretical view of possible pathways indicates that pyruvate might originate from glycolysis, malate (through malic enzyme) and pentose phosphate pathway (PPP) linked to photosynthesis or to sugar oxidation (oxidative PPP = OPPP). Because most of the oil-storing tissues are non-photosynthetic, carbon used for FA synthesis is mainly derived from glycolysis of sucrose in the cytosol, which is the major form of sugar photosynthesized by green tissues of the parent plant and transported into developing seeds/storage organ for oil accumulation (Fig.1-3; Baud & Lepiniec, 2010; Chapman & Ohlrogge, 2012). The product of glycolysis - pyruvate - can be generated from hexose phosphate in the cytosol or plastid. Additionally, the OPPP exists in the non-green seeds, while green seeds can generate ribulose 1,5-bisphosphate (RuBP) through the non-oxidative PPP,

and ribulose 1,5-bisphosphate carboxylase (Rubisco) can capture the CO_2 from PDH resulting in 20% higher yields of acetyl-CoA for oil synthesis from sucrose (Allen *et al*., 2009; Hay & Schwender, 2011; Schwender *et al*., 2004). However, it is generally considered that, in all cases, glycolysis remains the main supplier of carbon for fatty acid synthesis in oil-rich organs. Glycolysis of sucrose produces, in addition to carbon units, ATP and reducing power in such amounts that they should cover 80-90% of what is needed to produce TAGs. However, they may not be located in appropriate compartment (mostly plastid). One supposes that plastid transporters and malate dehydrogenase shuttles allows transfer of these to plastids for fatty acid synthesis.

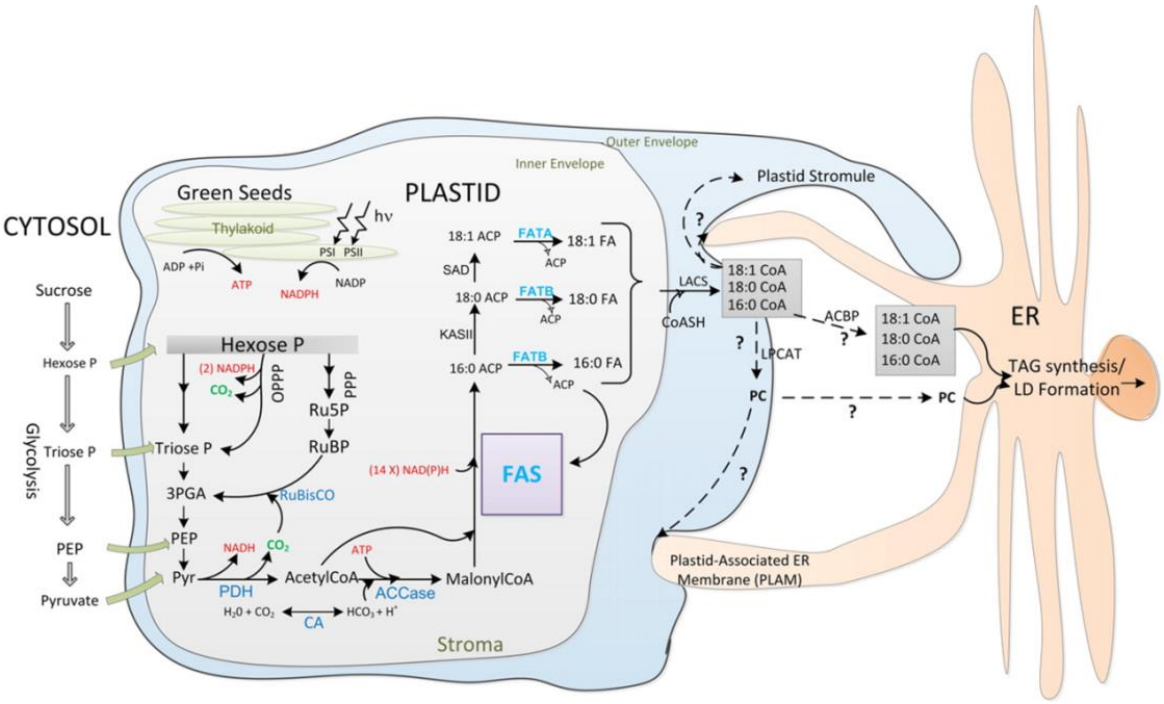

Fig.1-3 Overview of pathways and compartmentation of FA supply for TAG synthesis in plants (Fig.1 from Chapman & Ohlrogge, 2012). Synthesis starts from sucrose which is converted to pyruvate. In plastid, pyruvate serves to synthesize acetyl-CoA, which is the building block used by fatty acid synthase (FAS) to make fatty acids. Those are released from acyl carrier protein in plastid and then transported to the endosplasmic reticulum as acyl-CoAs. Acyl-CoAs are then used for synthesis of TAGs (see text for details). **Abbreviations**: ACCase, acetyl-CoA carboxylase; CA, carbonic anhydrase, PDH, pyruvate dehydrogenase; KAS, ketoacyl-ACP synthase; SAD, stearoyl-ACP desaturase; FAT, acyl-ACP thioesterases; LACS, long chain acyl-CoA synthethases; LPCAT, lyso-PC acyltransferase; ACBP, acyl-CoA binding protein; (O)PPP, (oxidative) pentose phosphate pathway; PS, photosystems; Hexose P, Hexose phosphate; Triose P, triose phosphate; 3PGA, 3-phosphoglycerate; Pyr, pyruvate; PEP, phosphoenolpyruvate; PC, phosphatidylcholine; RuBP, ribulose 1,5-bisphosphate; Ru5P, ribulose 5-phosphate; TAG, triacylglycerol; LD, lipid droplets; ER: endoplasmic reticulum.

1.2.1 Fatty acid synthesis

Aside from acetyl-CoA, the other substrate for FA synthesis is malonyl-CoA, which is generated from the carboxylation of acetyl-CoA catalyzed by acetyl-CoA carboxylase (ACCase) in the plastid. ACCase is the first committed enzyme to FA biosynthesis (Konishi *et al*., 1996; Li-Beisson *et al*., 2013) and has been shown to be the target of several post-translational regulatory mechanisms, which will be described below (see 1.2.2). Before entering fatty acid synthesis pathway, the malonyl group is transferred from CoA to an ACP by malonyl:acyl carrier protein malonyltransferase (MAT) (Baud & Lepiniec, 2010).

FA synthase (FAS), a multi-enzyme (non-covalently bound) complex synthesizes the fatty acids with up to 16 or 18 carbons using acetyl-CoA as a starting unit and malonyl-ACP as the elongator through a series of condensation, dehydration, and reduction reactions (Baud & Lepiniec, 2010; Li-Beisson *et al*., 2013). There are three isoforms of ketoacyl-ACP synthase (KAS), the enzymes responsible for condensation reactions, which are playing different roles in acyl chain elongation. KASIII catalyzes the initial condensation of acyl-CoA and malonyl-ACP to 4:0-ACP. Subsequent condensations (up to 16:0-ACP) are catalyzed by KASI, and KASII provides the final elongation to 18:0-ACP, which may be then desaturated by a stromal stearoyl-ACP desaturase (SAD) to yield 18:1-ACP. Then, these acyl groups are released from ACP by acyl-ACP thioesterases (Fat) yielding free FAs in the plastid stroma. This hydrolysis step is critical for the nature of fatty acids produced by plastids and the specificities of Fats determine to a large extent the chain length and the saturated FA content of plant oil (Jones *et al*., 1995).

Based on sequence alignments and substrate specificities, Fats have been classified into two general families: FatA and FatB (Jones *et al*., 1995; Salas & Ohlrogge, 2002). FatAs show substrate preference towards 18:1-ACP, whereas FatBs primarily hydrolyze saturated acyl-ACPs with 8 to 18 carbons (Jones *et al*., 1995; Sánchez-García *et al*., 2010; Voelker & Davies, 1994). Till now, FatA orthologs from

diverse plants all strictly exhibit highest activity to 18:1-ACP, including Cuphea (*Cuphea lanceolata* Ait. and *Cuphea wrightii* A. Gray; Dörmann *et al*., 1993), mangosteen (*Garcinia mangostana*; Hawkins & Kridl, 1998), castor (*Ricinus communis* L.; Sánchez-García *et al*., 2010), oil seed rape (*Brassica napus*; Hellyer *et al*., 1992), safflower (*Carthamus tinctorius*; Knutzon *et al*., 1992). Meanwhile, numerous FatBs also have been identified and characterized from plants, especially from those species producing unusual fatty acids, such as medium chain fatty acids (MCFAs). California bay (*Umbellularia californica*; Davies *et al*., 1991; Pollard *et al*., 1991) and Cuphea (*Cuphea lanceolata* Ait, *Cuphea wrightii* A. Gray, *Cuphea hookeriana* and *Cuphea palustris*; Dehesh *et al*., 1996a, b; Dörmann *et al*., 1993) all possess at least one FatB showing substrate specificity towards MCFA-ACPs. Recently, the expression of *CnFatB3* from coconut (*Cocos nucifera*) endosperm mainly contributed to 12:0 production (34 mol%) in *E.coli* K27 (Jing *et al*., 2011); likewise, the transient expression of *EgFatB3* from oil palm (*Elaeis guineensis*) endosperm produced about 15% 14:0 and 1% 12:0 of total fatty acids in tobacco leaves (Dussert *et al*., 2013). Furthermore, 16:0/18:0-ACP preference FatBs have been identified from mangosteen (*Garcinia mangostana*; Hawkins & Kridl, 1998), *Arabidopsis thaliana* (Dörmann *et al*., 2000; Dörmann *et al*., 1995), Diploknema (*Madhuca butyracea*; Jha *et al*., 2006), *Brassica juncea* (Jha *et al*., 2010), and *Jatropha curcas* (Srikanta Dani *et al*., 2011; Wu *et al*., 2009). Therefore, the specificities and relative expression patterns of Fats in plants largely determine the final fatty acid composition of lipids from eukaryotic pathway, as 18:1 only can be further desaturated in the ER while saturated fatty acids will remain as such.

1.2.2 Regulation of fatty acid synthesis

Wrinkled1 (WRI1) is a transcription factor first found in *Arabidopsis* (Cernac & Benning, 2004), based on a wrinkled seed phenotype due to low accumulation of oil. It has been extensively studied in *Arabidopsis* and found to control transcription of genes coding for fatty acid synthesis enzymes as well as late glycolysis ones, such as

plastid pyruvate kinase. Other enzymes of central metabolism were found to be up-regulated in WRI1 overexpressing plants (Grimberg *et al*., 2015) but it is not clear if they are direct targets of WRI1. Many studies have shown that overexpression of *WRI1* together with *DGAT1* leads to a tremendous rise of lipid content of transgenic plants (van Erp *et al*., 2014; Vanhercke *et al*., 2014). Interestingly, WRI1 does not seem to control any enzyme involved in Kennedy and editing pathways leading to assembly of fatty acids into oil and very little is presently known on transcription regulation of those enzymes. In seeds, WRI1 is under the control of seed master regulators including FUSCA3 (FUS3), ABSCISIC ACID INSENSITIVE3 (ABI3) and LEAFY COTYLEDON 1 AND 2 (LEC1 and LEC2) (Baud *et al*., 2007). *Arabidopsis* contains other WRI-liked transcription factors that control fatty synthesis in lipid-producing tissues such as epidermis for surface lipid synthesis (To *et al*., 2012).

At post-translational level, it has been shown that fatty acids, when fed to cell cultures, would lead to down regulation of fatty acid synthesis in plastid (Shintani & Ohlrogge, 1995). This appears to be due to feedback regulation of ACCase by end products of fatty acid synthesis (oleoyl-ACP). Similarly, defective oil editing (see 1.3.2) in transgenic plants overexpressing castor bean fatty acid hydroxylase led also to downregulation of ACCase through a similar mechanism (Bates *et al*., 2014). As in yeast, animals and bacteria, plant ACCase is a key enzyme for regulation of fatty acid synthesis. It has been shown to be regulated by light and redox status, and molecular mechanisms involve phosphorylation and interation with a PII protein (Andre *et al*., 2012; Chapman & Ohlrogge, 2012; Bourrellier *et al*., 2010). Acyl-CoAs are likely to play an important role in regulating oil synthesis. *In vitro* data indicate that acyl-CoAs inhibit the plastidial glucose 6-phosphate transporter *in vitro*, therefore limiting carbon available to fatty acid synthesis in plastid, and this inhibition can be alleviated by addition of acyl-CoA binding proteins (ACBPs) (Fox *et al*., 2000). Also, adding ACBPs enhances acyltransferase activities assayed *in vitro* from microsomal membranes (Yurchenko *et al*., 2009). Another interesting control mechanism of overall oil synthesis was shown in transgenic lines of rapeseed that overexpress a

thioesterase, leading to accumulation of MCFAs in oil (Eccleston & Ohlrogge, 1998). It was found that newly synthesized fatty acids were directed to β-oxidation rather than integrated in oil, probably due again to a defective editing mechanism. This futile cycling was demonstrated to lead to highly significant losses of fatty acids.

1.2.3 Export of fatty acids to endoplasmic reticulum

The mechanisms by which fatty acids are exported to the ER remain unclear, while several ways have been proposed (Bates *et al*., 2007; Bates *et al*., 2013; Kim *et al*., 2013; Li-Beisson *et al*., 2013). It is considered that fatty acids are exported from plastid as acyl-CoA, which is then used for ER acyltransferases as substrate for lipid synthesis. The recent characterization of the *Arabidopsis FAX1* (FATTY ACID EXPORT 1) gene, encoding a novel chloroplast inner envelope membrane putative fatty acid transporter, shed new light on the mechanisms of plastid fatty acid export (Li *et al*., 2015). The *Arabidopsis FAX1* mutants showed that FAX1 is crucial for biomass production, male fertility, and the distribution of FA-derived compounds such as lipids, ketone waxes, and pollen cell-wall material (Li *et al*., 2015, 2016). When *FAX1* is absent, the levels of ER-derived lipids decrease, while the levels of plastid-produced lipid species increase. *FAX1* over-expressing plants showed an opposite phenotype, including a pronounced increase of TAGs in flowers and leaves (Li *et al*., 2015, 2016). The heterologous expression of *FAX1* from *Arabidopsis* in *fat1* mutant yeast (defective in fatty acid transporter) showed that it complements for FA transport with a specificity range of C16:0 > C18:1 ~ C18:0, and thus generally prefers C16 over C18 FAs *in vivo* (Li *et al*., 2015, 2016). It also indicated that FAX1 is acting only in membrane transfer of FAs, but not in FA-activation (Li *et al*., 2015, 2016).

Fatty acid activation, that is esterification to CoA, is catalyzed by long chain acyl-CoA synthetase (LACS). The two-step mechanism catalyzed by LACS has been found decades ago: the conversion of free FA and ATP to an enzyme-bound

acyl-AMP intermediate in the presence of Mg^{2+}, and the generation of acyl-CoA and free AMP (Shockey *et al*., 2002).

Nine isoforms of LACS have been identified from *Arabidopsis*, performing different functions including anabolic and catabolic pathways (Jessen *et al*., 2014; Schnurr *et al*., 2002; Shockey *et al*., 2002; Zhao *et al*., 2010). To carry out their function, these LACSs are localized at different cellular compartment. Two of them, LACS6 and LACS7, initiate the process of fatty acid *β*-oxidation in the peroxisome by activating the free fatty acids derived from lipolysis of TAG molecules during post-germinative growth (Fulda *et al*., 2002, 2004; Shockey *et al*., 2002). All other LACSs possess the ability to restore growth in the LACS-deficient strain of yeast (*Saccharomyces cerevisiae*) (Pulsifer *et al*., 2012; Shockey *et al*., 2002). LACS1, LACS4, and LACS8 are localized in the ER (Jessen *et al*., 2014; Zhao *et al*., 2010), while only LACS9 is in plastid, more precisely in envelop (Schnurr *et al*., 2002). As shown by mutant studies in *Arabidopsis*, few single mutants show obvious phenotypes, as most LACS have overlapping functions that can be detected in multiple mutants only. For example, LACS1 and LACS2 are involved in wax and cutin synthesis, and LACS2 is essential for normal cuticle development by catalyzing the synthesis of ω-hydroxy fatty acyl-CoA intermediates in *Arabidopsis* (Lü *et al*., 2009; Schnurr and Shockey, 2004; Weng *et al*., 2010). However, LACS1 also functionally overlaps with LACS9 to import acyl-CoA for TAG biosynthesis (Schnurr *et al*., 2002; Zhao *et al*., 2010). The most important impact on TAG synthesis was detected in *lacs4lacs9* double mutant, with oil content is reduced by 25%. More generally, it was found that LACS4 and LACS9 are both involved in lipid trafficking between ER and plastid and best evidence concerned transport from the ER to the plastid (Jessen *et al*., 2014). Besides, over-expression or disruption of *LACS8* did not affect the seed fatty acid content, whereas the inactivation of *LACS8* in the *lacs4 lacs9* background resulted in lethality (Jessen *et al*., 2014; Zhao *et al*., 2010).

How acyl groups are exported to the ER is still unclear. It has been hypothesized that newly synthesized FAs are incorporated into PC by lyso-PC acyltransferase

(LPCAT) in the plastid envelop and PC then transported to the ER (Bates *et al*., 2007, 2009, 2013; Chapman & Ohlrogge, 2012), though unknown mechanisms. The hypothesis has been proposed that direct physical contact with plastid-associated ER membranes or stromules might allow for lipid exchange (Andersson *et al*., 2007; Chapman & Ohlrogge, 2012; Hanson & Sattarzadeh, 2011). Also, lipid transfer proteins, including acyl-CoA-binding proteins (ACBPs), might transport the acyl-CoAs to the ER through the cytosol (Chapman & Ohlrogge, 2012). ACBP6, a 10 kDA cytosolic protein, was demonstrated to bind acyl-CoAs with best affinity towards 18:2-CoA, and to increase glycerol-3-phosphate acyl transferase (GPAT) activity *in vitro* (Brown *et al*., 1998) as well as that of LPCAT (Yurchenko *et al*., 2009). It has also been demonstrated to alleviate the inhibition of glucose 6-phosphate and ATP transport caused by acyl-CoA esters (Fox *et al*., 2000). The family of *Arabidopsis thaliana* ACBP consists of six members, which are different in structure, cellular localization, expression patterns and binding preferences (Xiao & Chye, 2011; Yurchenko & Weselake, 2011). Among these proteins, ACBP6 is the only ACBP with low molecular mass, and it is localized in the cytosol with ACBP4 and ACBP5 (Chen *et al*., 2008; Xiao *et al*., 2008). ACBP4 and ACBP5 both show binding preference towards 18:1-CoA (Xiao *et al*. 2008). ACBP1 and ACBP2 are transmembrane proteins mainly with the plasma membrane, ER and with small intracellular vesicles (Gao *et al*., 2009; Li & Chye, 2003, 2004). They have been hypothesized to be involved in the maintenance of a membrane-associated acyl-CoA pool (Chen *et al*., 2010; Gao *et al*., 2009; Kojima *et al*., 2007; Li & Chye, 2003; Xiao *et al*., 2009). ACBP3 is the only extracellular ACBP family member containing a predicted cleavable N-terminal signal sequence, and exhibits best affinity towards arachidonyl-CoA (Leung *et al*., 2006). Presently, involvement of ACBPs in lipid metabolism has not been demonstrated *in vivo* and *in vitro* data point to ACBP6 as best candidate of the family for acyl-CoA transport between plastid and ER. One also needs to point out that acyl-CoA are soluble molecules that can diffuse in cytosol, contrarily to "normal" lipids, and that simple diffusion might be actual mode of transfer.

Recently, a putative ER transporter of acyl-CoA has been described and hypothesized to control acyl-CoA import into the ER. However, biochemical evidence or complementation of yeast transporter mutant lacks (Kim *et al*., 2013).

1.3 Pathways for TAG assembly

1.3.1 Conventional Kennedy pathway

For a very long time, oil was considered to be directly synthesized from DAG produced by Kennedy pathway through a third acylation catalyzed by DGAT enzymes. Two sequential acylations of the *sn*-1 and *sn*-2 positions of glycerol-3-phosphate (G3P) with acyl-CoA lead to synthesis of lyso-phosphatidic acid (LPA) and subsequently phosphatidic acid (PA) by the acyl-CoA:glycerol-3 -phosphate acyltransferase (GPAT) and acyl-CoA:lyso-phosphatidic acid acyltransferase (LPAT), respectively (Fig.1-4). Then, PA is dephosphorylated to diacylglycerol (DAG) by PA phosphatase (PAP) (Fig.1-4). Phospholipids such as PC and PE are synthesized from DAG by action of cytidinediphosphate (CDP)-aminoalcohol:DAG aminoalcohol-phosphate transferases (AAPTs). In *Arabidopsis*, there are two AAPTs, each capable of using CDP-ethanolamine and CDP-choline as substrate (Dewey *et al*., 1994), but with distinct preferences. For headgroup synthesis, choline and ethanolamine are phosphorylated by choline and ethanolamine kinases, respectively. Choline phosphate and ethanolamine phosphate are converted to CDP-choline and CDP-ethanolamine, by ethanolamine or choline-phosphate cytidylyl transferases (ECT and CCT). CCT is considered a key target for regulation of PC synthesis. Another group of phospholipids, such as phosphatidylinositol (PI), are synthesized from PA. PA is converted to CDP-DAG by cytidine 5’-triphosphate (CTP):PA cytidylyl transferase (CDP-DAG enzyme in text or CDP-DAGS in Fig.1-4). Polar headgroup such as inositol is then directly transferred to CDP-DAG by a PI synthase to yield PI. There are other pathways that have been poorly studied in plants. PE can be synthesized by decarboxylation of phosphatidylserine (PS) by PS decarboxylase.

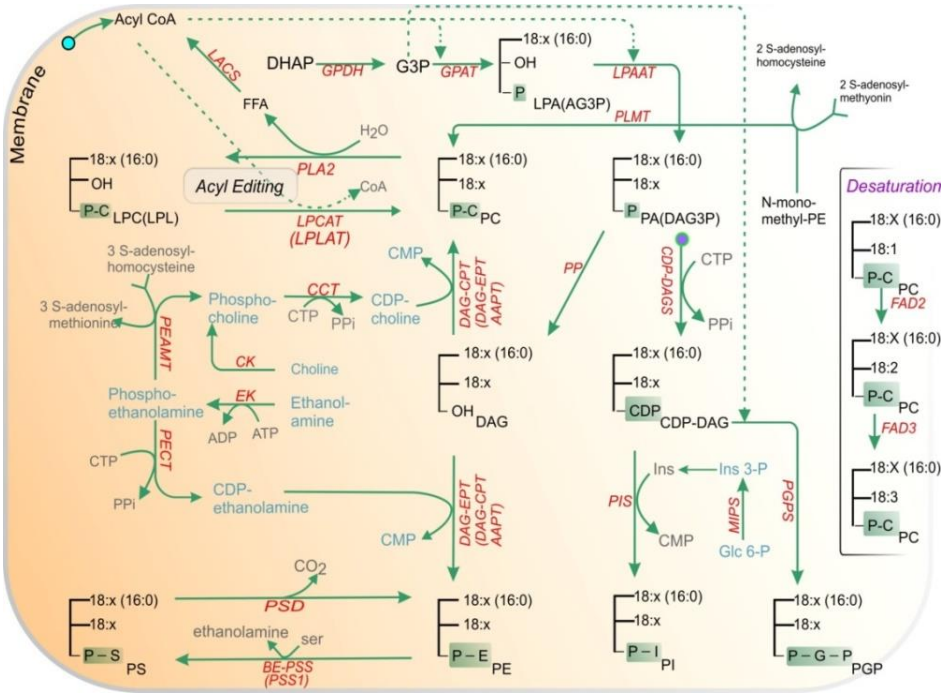

Fig.1-4 Eukaryotic phospholipid synthesis and editing (Fig.4 from Li-Beisson *et al*., 2013). All lipids originate either from diacylglycerol (DAG) or phosphatidic acid (PA), which are synthesized by two sequencial acylations from G3P and one dephosphorylation of phosphatidic acid. Phospholipids such as PC and PE are synthesized from DAG by action of AAPTs. For headgroup synthesis, choline and ethanolamine are phosphorylated by CCT and ECT enzymes, respectively. PI and PGP are synthesized from PA, which is converted to CDP-DAG by CDP-DAG synthase. Polar headgroup such as inositol is then directly transferred to CDP-DAG by a synthase (PIS for PI) to yield PI. PE can be synthesized by decarboxylation of phosphatidylserine (PS) by PS decarboxylase and a methylation pathway leads to PC synthesis from PE (see text for details). **Abbreviations**: AAPT, aminoalchoholphosphotransferase; BE-PSS, base-exchange-type phosphatydylserine synthase; CCT, CTP:phosphorylcholine cytidylyltransferase; CDP-DAGS, CDP-diacylglycerol synthetase; CK, choline kinase; DAG, diacylglycerol; DAG-CPT, CDP-choline:diacylglycerol cholinephosphotransferase; DAG-EPT, CDP-ethanolamine:diacylglycerol cholinephosphotransferase; DHAP, dihydroxyacetone phosphate; EK, ethanolamine kinase; FAD2, oleate desaturase; FAD3, linoleate desaturase; G3P, glycerol 3-phosphate; Glc 6-P, glucose 6-phosphate; GPAT, glycerol-3-phosphate acyltransferase; GPDH, glycerol-phosphate dehydrogenase; Ino 3-P, inositol 3-phosphate; LACS, long chain acyl-CoA synthetase; LPA, lysophosphatidic acid; LPAAT, lysophosphatidic acid acyltransferase; LPC, lyso-phosphatidylcholine; LPCAT: lysophosphatidylcholine acyltransferase; LPL, lysophospholipid; MIPS, myo-inositol-3-phosphate synthase; PA, phasphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PEAMT, phosphoethanolamine N-methyltransferase; PECT, CTP:phosphorylethanolamine cytidyltransferase; PG, phosphatidylglycerol; PGP, phosphatidyglycerophosphate; PGPS, phosphatidylglycerophosphate synthase; Pl, phosphatidylinositol; PIS, phosphatidylinositol synthase; PLMT, N-methylphospholipid methyltransferase; PLA2, phospholipase A2 (Cytosolic); PLD, phospholipase D; PP, phosphatidate phosphatase; PS, phosphatidylserine; PSD, phosphatidylserine decarboxylase.

In yeast, PC can be synthesized by three successive methylations (carried out by two distinct enzymes) of PE. In plants, this methylation pathway is slightly different as first methylation is carried out on ethanolamine-phosphate only; then, ethanolamine-P is transferred to PA through AAPTs to yield monomethyl-PE, which can be further methylated twice to PC. This pathway seems to be the only way for plants to synthesize choline, after hydrolysis of PC by phospholipase D. In yeast, PS is synthesized by a synthase, as for PI. The situation in plants remains unclear and it is thought that PS synthesis occurs mainly through a base-exchange mechanism rather than by a synthase reaction.

TAGs (Fig.1-5) are synthesized by a third acylation by acyl-CoA:diacylglycerol acyltransferase (DGAT) from DAG. For the sake of simplicity, the pathway concerning TAG alone is shown on Fig.1-5 but one needs to keep in mind that only last reaction is specific to TAGs and all other ones are shared with membrane glycerolipids because PA and DAG are also precursors to all phospholipids. This leads intuitively to consider that regulation mechanisms directing acyl-CoAs to such lipid and not to others are likely not to be simple.

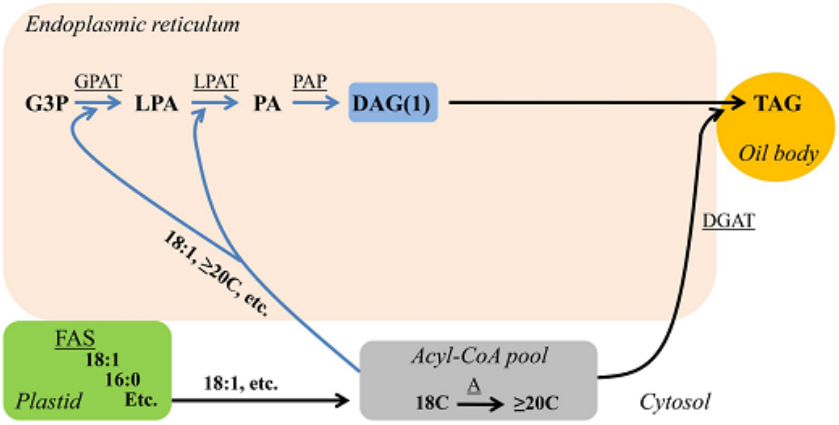

Fig.1-5 Kennedy pathway to oil synthesis (Fig.1 from Bates and Browse, 2012). Oil (TAG) is synthesized by two sequencial acylations from G3P, one dephosphorylation of phosphatidic acid (PA) to diacylglycerol (DAG) and one third acylation. Blue arrows indicate reactions involved in *de novo* DAG(1) synthesis. Abbreviations: DAG, diacylglycerol; G3P, glycerol-3-phosphate; LPA, lyso-phosphatidic acid; LPC, lyso-phosphatidylcholine; PA, phosphatidic acid; PC, phosphatidylcholine; TAG, triacylglycerol; DGAT, acyl-CoA:DAG acyltransferase; FAS, fatty acid synthesis; GPAT, acyl-CoA:G3P acyltransferase; LPAT, acyl-CoA:LPA acyltransferase; PAP, PA phosphatase. Enzymatic reactions are underlined: A, FA elongation.

A number of studies have been done on these key enzymes involved in Kennedy pathways in recent years. In *Arabidopsis*, one plastid-localized ATS1 and nine extraplastidial GPATs have been identified and characterized (Beisson *et al*., 2007; Gidda *et al*., 2009; Kunst *et al*., 1988; Li *et al*., 2007a, b; Nishida *et al*., 1993; Shockey *et al*., 2016; Singer *et al*., 2016; Yang *et al*., 2012; Yang *et al*., 2010; Zheng *et al*., 2003). ATS1 seems to be involved in producing major phospholipids for plastidial membranes using acyl-ACP as substrate (Kunst *et al*., 1988; Singer *et al*., 2016), while GPAT1-GPAT8 are a land plant-specific family of *sn*-2 GPATs using acyl-CoA as substrate and involved in cutin and suberin synthesis rather than membrane lipid and TAG biosynthesis (Beisson *et al*., 2007; Li *et al*., 2007a, b; Yang *et al*., 2012; Yang *et al*., 2010; Zheng *et al*., 2003). GPAT9 was identified previously as a candidate for an ER-associated GPAT required for TAG biosynthesis based on the substantial sequence identity to mammalian GPAT3 that is involved directly in the synthesis of TAG in lipid-rich mammalian tissues (Cao *et al*., 2006; Gidda *et al*., 2009). Recently, the function analysis of GPAT9 has been carried out and the critical role it plays in oil synthesis has been demonstrated by a series of *in vivo*, *in vitro*, and *in silico* experiments (Shockey *et al*., 2016; Singer *et al*., 2016). This gene is a highly conserved, single-copy, and essential gene that supplies the LPA necessary for the synthesis of membrane phospholipids in *Arabidopsis* and likely other plants as well (Shockey *et al*., 2016; Singer *et al*., 2016). It exhibited *sn*-1 acyltransferase activity with high specificity for acyl-CoA *in vitro*, and down-regulation of *GPAT9* led to decreased seed oil content while a *gpat9* knockout (KO) mutant exhibited both male and female gametophytic lethality phenotype (Shockey *et al*., 2016; Singer *et al*., 2016). Additionally, GPAT9 interacted with other enzymes involved in Kennedy pathway and acyl-editing cycle in the split-ubiquitin yeast two-hybrid system assay (Shockey *et al*., 2016). Therefore, GPAT9 plays a pivotal role in membrane lipid and TAG biosynthesis.

Previously, LPATs from special plants that can accumulate unusual acyl moieties at the *sn*-2 position of TAGs have been isolated and characterized, leading to

incorporation of 12:0 in coconut (*Cocos nucifera*; Knutzon *et al*., 1995, 1999) and 22:1 in meadowfoams (*Limnanthes alba* and *Limnanthes douglasii*; Brough *et al*., 1996; Brown *et al*., 1995; Lassner *et al*., 1995). These LPATs have been designated as the class B and display preference for distinct, unusual saturated or unsaturated acyl groups with high expression in storage organs. In contrast, the class A microsomal LPATs exhibit a substrate preference for 18:1-CoA typical of enzymes involved in membrane lipid synthesis and are ubiquitously expressed in the plant. Five *LPAT*s have been annotated in the *Arabidopsis* genome, and the mutational loss of the only plastidial *LPAT1* leads to the embryo death (Kim & Huang, 2004). However, *LPAT2* encodes the ubiquitous, abundant, and ER-located LPAT, and is necessary for female gametophyte development (Kim *et al*., 2005). LPAT3 seems to be essential for pollen development because of its exclusive and high expression in pollen (Kim *et al*., 2005). Moreover, overexpression of a *Brassica napus LPAT* isoform with homology to *LPAT2* resulted in the enhanced oil accumulation in *Arabidopsis* seeds (Maisonneuve *et al*., 2010). Much functional study remains to be done on the LPATs to understand their roles in oil biosynthesis.

PA and DAG are both intermediates in the biosynthesis of phospholipids and TAGs. PAPs hydrolyze phosphate group from PA generating DAG, which is used for the synthesis of storage lipid TAGs as well as the phospholipids PC and PE. As mentioned above, PA also serves as precursor for some phospholipids *via* the CDP-DAG pathway, including PI, PS, and PG (Li-Beisson *et al*., 2013). CDP-DAG synthase (CDP-DAGS) converts PA to CDP-DAG, which serves as a substrate for PI, PS, and PG biosynthesis. Phospholipids are the major component of most eukaryotic cell membranes in compartmentalizing various intracellular biochemistry actions, and their composition and quantity are tightly regulated so that membranes can maintain their structure and function in response to developmental and environmental changes (Eastmond *et al*., 2010; Ohlrogge & Browse, 1995). At last, PA is an important signaling molecule (Carman & Han, 2006), which function in the abscisic acid (ABA) signaling pathway has been extensively studied in plants (Wang *et al*., 2006). Thus,

PAPs, which hydrolyze PA to DAG are likely to be very important enzymes for overall plant homeostasis.

In yeast, there are two types of PAP enzyme, Mg^{2+}-dependent soluble PAP1 or phosphatidic acid phosphohydrolase (PAH) and Mg^{2+}-independent membrane-bound PAP2 or lipid phosphate phosphatase (LPP) (Carman & Han, 2006, 2009). Mutation of *pah1* in yeast leads to membrane lipid proliferation and decreased levels of DAG and TAG (Carman & Han, 2009; Han *et al*., 2006). So PAH1 commits PA to DAG conversion for TAG biosynthesis. In *Arabidopsis*, two homologous genes designated *PAH1* and *PAH2* have been identified, that encode the PAPs involved in producing DAG from ER phospholipids destinated to eukaryotic galactolipid biosynthesis in the plastid (Eastmond *et al*., 2010; Li-Beisson *et al*., 2013; Nakamura *et al*., 2009). The double KO of *pah1*/pah*2* in *Arabidopsis* resulted in disrupted membrane remodeling, increased PC coupled with overexpansion of the ER membrane, and a slight reduction in total seed FA content, consistent with a function in regulating membrane phospholipid accumulation as in animals and yeast (Chapman & Ohlrogge, 2012; Eastmond *et al*., 2010). Moreover, in this double mutant, the regulation system has been found: PA accumulation in the ER membrane directly activates the enzyme CCT, which catalyzes the key rate-limiting step of the nucleotide pathway of PC synthesis (Craddock *et al*., 2015), leading to higher PC synthesis. This regulation mechanism has also been demonstrated by overexpression of truncated constitutively active version of CCT1, which largely replicates the *pah1 pah2* phenotype (Craddock *et al*., 2015).

Another family of PAP, different from PAH (also designated as lipins), has been found and designated LPP for lipid phosphate phosphatase. Four isoforms of LPP have been found in *Arabidopsis* (Katagiri *et al*., 2005; Pierrugues *et al*., 2001). Both LPP1 and LPP2 can utilize PA and diacylglycerol pyrophosphate (DGPP) as substrates *in vitro*, similar to the LPPs from yeast and mammalian cells (Carman, 1997; Pierrugues *et al*., 2001; Waggoner *et al*., 1999). However, LPP1 preferred DGPP to PA as a substrate, while LPP2 showed no preference for either substrate

(Pierrugues *et al*., 2001). Additionally, LPP2 has been demonstrated to be a negative regulator involved in response to ABA signaling during seed germination (Katagiri *et al*., 2005). Although LPPs seem to be regulators of PA and/or DGPP signaling rather than of lipid biosynthesis, it has been hypothesized that they might also be involved in generating DAG for TAG synthesis because double mutation of PAH enzymes did not affect much oil synthesis, until a third family of PAP was found in *Arabidopsis* (Nakamura & Ohta, 2010; Nakamura *et al*., 2007). This new family resembles LPPs from prokaryotic origin, rather than ones from mammals. They are also designated LPPs, but isoforms are distinguished by a greek letter rather than a number, to underline their prokaryotic origin. This family comprises 5 members, three of which (LPP γ, LPPε1, and LPPε2) are predicted to be in plastid, and have been demonstrated to be involved in generating DAGs destined to synthesis of galactolipids that replace PC in extra-plastidial membranes during phosphate deprivation. Three of them (LPPβ, δ, γ) lack plastid transit peptide and LPPγ was found to be involved in sphingolipid synthesis (Nakagawa *et al*., 2012). Because LPPβ and δ showed transcription patterns compatible with function in oil production, it has been hypothesized that they might be involved in TAG synthesis (Bourgis *et al*., 2011; Yang *et al*., 2016).

Of the three acyltransferase, only DGATs are unique enzyme TAG synthesis and these enzymes have been therefore extensively studied. Two main classes of DGAT, designated DGAT1 and DGAT2, are ubiquitous in eukaryotes (Liu *et al*., 2012). They differ in sequence and topology, intracellular localization, and substrate specificity (Li-Beisson *et al*., 2013; Liu *et al*., 2012; Shockey *et al*., 2006). DGAT1s are predicted to contain between 8 and 10 transmembrane domains and belong to the membrane-bound O-acyltransferase (MBOAT) family, whereas DGAT2 polypeptides have only two to three predicted transmembrane domains and belong to the monoacylglycerol acyltransferase (MGAT) family (Chen *et al*., 2016; Liu *et al*., 2012). In addition, a third enzyme leading to TAG synthesis has been discovered 16 years ago (Dahlqvist *et al*., 2000). Phospholipid:DAG acyltransferase (PDAT1) belongs to a family of proteins that can also acylate sterols. It uses an acyl chain from

the *sn*-2 position of PC instead of the acyl-CoA molecule DGATs use (Bates & Browse, 2012; Dahlqvist *et al*., 2000; Zhang *et al*., 2009). Therefore, it generates also lyso-PC (LPC) that can be re-acylated by lyso-PC acyltransferases (LPCATs; Xu *et al*., 2012).

DGAT1 from *Arabidopsis* has been shown to be required for normal TAG accumulation in oil storing tissues, both by overexpression and mutation studies (Chapman & Ohlrogge, 2012; Jako *et al*., 2001; Katavic *et al*., 1995; Routaboul *et al*., 1999; Zhang *et al*., 2009). However, KO of *dgat1* mutant led to only 20-40% reduced TAG content in *Arabidopsis* seeds (Katavic *et al*., 1995; Routaboul *et al*., 1999; Zhang *et al*., 2009; Zou *et al*., 1999), suggesting it is not the only enzyme responsible for TAG synthesis in *Arabidopsis*. *Arabidopsis* also possesses a DGAT2, but the KO mutant does not show significant reduction of seed oil content, even when crossed with *dgat1* mutant (Liu *et al*., 2012). Thus, DGAT2 does not appear to be involved in TAG synthesis in developing seeds of *Arabidopsis*, and other mechanisms must contribute to the TAG accumulation in plants. While the *pdat1* mutant has no impact on TAG accumulation (Mhaske *et al*., 2005), the same mutant in which *DGAT1* has been silenced with RNAi shows 70 to 80% decrease in seed oil content and disruptions of embryo development (Zhang *et al*., 2009). These results suggest that DGAT and PDAT cooperate to synthesize the majority of TAG in oil-storing tissues of *Arabidopsis*. Thus, *DGAT1* and *PDAT1* acyltransferases have overlapping functions in oil biosynthesis of *Arabidopsis*. The double KO mutant of *dgat1*/*pdat1* is pollen lethal (Zhang *et al*., 2009).

The substrate specificity of DGATs from several plants has been reported. Most DGAT1 do not show striking specificities, except EgDGAT1-1 from oil palm, a kernel-specific enzyme, which shows preference for medium chain acyl-CoA substrates when expressed in yeast (Aymé *et al*., 2015). While DGAT2 does not appear to be much involved in *Arabidopsis* seed TAG synthesis, it is not the case in other plant species that accumulate oil with unusual fatty acids. DGAT2 from castor bean (*Ricinus communis*) endosperm, which accumulates a large amount of TAG

containing more than 90% ricinoleic acid (18:1 FA hydroxylated at the Δ^{12}-position), exhibits the highest expression levels in seeds and prefers acyl-CoA and DAG substrates containing hydroxylated FA (Burgal *et al*., 2008). Additionally, tung tree *(Vernicia fordii)* DGAT2 preferentially incorporates α-eleostearic acid into TAG in seed oil bodies (Shockey, 2006). Interestingly, DGAT1 and 2 localize to different areas of ER in Tung, in addition to synthesizing oil with distinct composition. Thus, it seems like DGAT2 plays a more prominent role than DGAT1 in the selective accumulation of unusual FAs in the seed oil of certain plants (Liu *et al*., 2012).

In those few prokaryotic organisms that accumulate TAGs, DGAT role is carried out by an unrelated family of enzymes (Wältermann *et al*., 2007). They can synthesize both wax ester and TAG in bacteria (Kalscheuer & Steinbüchel, 2003), and also produce TAG and fatty acid ethyl esters (FAEEs) as well as fatty acid isoamyl esters (FAIEs) in *Saccharomyces cerevisiae* (Kalscheuer *et al*., 2004). The genome of *Arabidopsis* contains 11 sequences that are homologous to the bifunctional *WS/DGAT*, and only one of them (*AtWSD1*) has so far been characterized (Li *et al*., 2008). While this enzyme shows, as its prokaryotic homologs do, bifunctional wax ester synthase/DGAT activities *in vitro*, *in vivo* expression studies in yeast failed to show any complementation of TAG-less phenotype of a yeast mutant. Similar results were obtained from a *Petunia* homolog *PhWS* (King *et al*., 2007).

In addition, a soluble cytosolic DGAT enzyme has been reported in peanut (*Arachis hypogaea*; Saha *et al*., 2006) and *Arabidopsis* (Hernandez *et al*., 2012). However, additional data are required to support a function for these proteins that were designated as DGAT3.

1.3.2 Impact of acyl editing pathways on oil synthesis

For a long time, people studying TAG biosynthesis in plants have been, with few exceptions (Stymne & Stobart, 1984), limiting their studies to conventional Kennedy pathway. However, several *in vivo* studies, carried out on organs that accumulate

important amounts of oil, showed that, when fatty acids are labeled with radioactive acetate, a very important label can be found in PC before it is found in TAGs (Cahoon & Ohlrogge, 1994). This does apparently make no sense because these organs accumulate oil, not PC. Similarly, feeding microsomes isolated from developing oil seeds with radioactive acyl-CoAs showed a similar pattern, that is, a fast important initial accumulation of label in PC rather than in TAGs (Stymne & Stobart, 1984). Most intriguingly, similar results could be obtained when labeling the glycerol-P backbone of glycerolipids by feeding oleaginous developing embryo with glycerol (Stobart & Stymne, 1985). These studies strongly suggest a role of PC in oil synthesis that would not be expected by looking at Kennedy pathway only. Furthermore, studies aiming at having transgenic *Arabidopsis* accumulate unusual fatty acids showed that defective incorporation of these fatty acids into oil could lead to decreased oil synthesis (van Erp *et al*., 2015). Taken together, these observations have recently led to reconsider TAG synthesis within a larger frame that included fatty acid editing. Fatty acid editing are these reactions that do not impact directly amount of oil synthesized, but that modify fatty acid composition of TAGs. As we will see, PC is a key player in these reactions. The term "*de novo*" is used to qualify lipids that have not been through the editing pathways and reactions. In case of fatty acids, this will designate those fatty acids molecules that are exported from plastid (16:0, 18:0 and 18:1) and in case of DAGs, those that are released by PA phosphatases following Kennedy pathway. The term "editing" is used to qualify those reactions that modify fatty acid composition of oil but do not directly affect amount of oil produced.

Acyl-CoAs from *de novo* fatty acid synthesis are either saturated (palmitic C16:0 and stearic C18:0 acids) or mono-unsaturated (oleic acid C18:1). Subsequent desaturation catalyzed by the ER-localized fatty acid desaturases including the oleate desaturase (FAD2) and the linoleate desaturase (FAD3), producing linoleic (C18:2) and linolenic (C18:3) acids occurs in the ER when fatty acid substrates are esterified to PC only (Sperling *et al*., 1993). Thus, any polyunsaturated fatty acid present in the oil must necessarily have transited by PC.

1.3.2.1 Acylation/deacylation of PC (Lands cycle)

The first editing mechanism (Fig.1-6) is a deacylation-reacylation cycle that does not lead to net PC synthesis. The cycle starts with the release of an acyl group from PC, generating LPC by reverse action of acyl-CoA:lyso-phosphatidylcholine acyltransferase (LPCAT) or phospholipase A2. Re-esterification of LPC by LPCAT generates PC and completes the cycle. Because PC might integrate 18:1 fatty acid that might be further desaturated into 18:2 and then released again from PC this process will leads to an enrichment of PUFAs in the acyl-CoA pool that is used for *de novo* DAG synthesis (Bates *et al*., 2009, 2012; Bates & Browse, 2012; Li-Beisson *et al*., 2013), allowing for synthesis of oil with PUFAs.

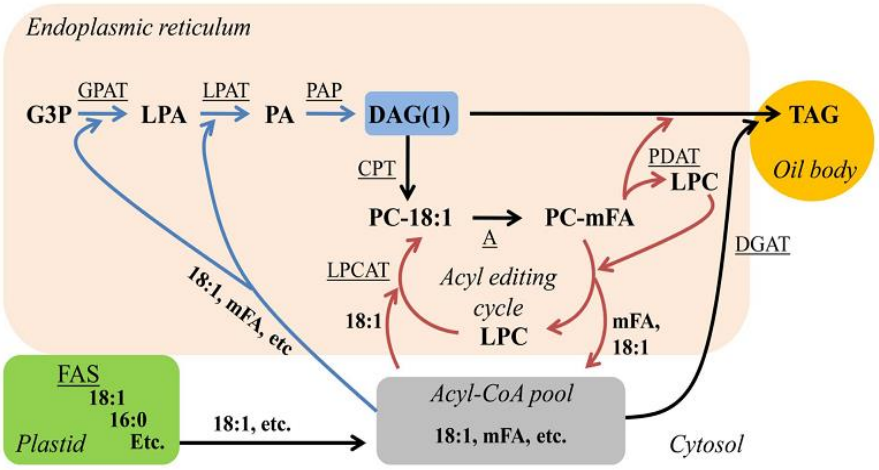

Fig.1-6 Kennedy pathway to oil synthesis and acyl editing cycle (Fig.2 from Bates and Brwose, 2009). The cycle starts with the release of an acyl group from PC, generating LPC by reverse action of LPCAT or phospholipase A2. Re-esterification of LPC by LPCAT generates PC and completes the cycle. Also, an enzyme called PDAT1 can lead to synthesis of TAG and lyso-PC from DAG and PC (see text for details). Blue arrows indicate reactions involved in *de novo* DAG(1) synthesis. Red arrows indicate reactions involved in acyl editing. **Abbreviations**: DAG, diacylglycerol; G3P, glycerol-3-phosphate; LPA, lyso-phosphatidic acid; LPC, lyso-phosphatidylcholine; PA, phosphatidic acid; PC, phosphatidylcholine; TAG, riacylglycerol; DGAT, acyl-CoA:DAG acyltransferase; FAS, fatty acid synthesis; GPAT, acyl-CoA:G3P acyltransferase; LPAT, acyl-CoA:LPA acyltransferase; LPCAT, acyl-CoA:LPC acyltransferase; PAP, PA phosphatase. Enzymatic reactions are underlined: A, modification of FA esterified to PC.

Analysis of *Arabidopsis lpcat1/lpcat2* double mutant showed increased amount of very-long-chain fatty acids and decreased PUFAs in seeds, indicating the involvement of these two enzymes in editing process of oil (Bates *et al*., 2012). Other possible

candidates are broad specificity lysophospholipid acyltransferase, such as lyso-phosphatidylethanolamine acyltransferase (LPEAT) (Stålberg *et al*., 2009). In human cells, LPEAT2 shows LPEAT, lyso-PG acyltransferase (LPGAT), lyso-PS acyltransferase (LPSAT) and LPCAT activities with 18:1-CoA or 20:4-CoA as acyl donors (Shindou *et al*., 2013). LPEAT2 from *Arabidopsis* also shows higher levels of reverse reaction (i.e. transfer of *sn*-2 fatty acid in acyl-CoA) towards PC and PE than *Arabidopsis* LPCAT1 does *in vitro* (Jasieniecka-Gazarkiewicz *et al*., 2016). LPC as part of the acyl editing cycle may be the major acceptor of nascent FA as they exit the plastid, regardless of whether the FA will be further modified or not (Tjellstrom *et al*., 2012). Indeed, labelling experiments indicate that fatty acid flux to PC is about 5-fold higher than that to DAGs (Bates *et al*., 2012; Bates *et al*., 2009). Besides, seven LPCATs from five different plant species, including species accumulating hydroxylated acyl groups in their seed oil, showed different specificities in their forward and reverse reactions *in vitro* (Lager *et al*., 2013).

As mentioned above, PDAT (Fig.1-6) catalyzes the synthesis of TAG from DAG using an acyl chain from the *sn*-2 position of PC thus generating LPC. It was shown that LPC released by PDAT is re-acylated by LPCAT through the acyl editing cycle (Xu *et al*., 2012).

1.3.2.2 Phosphatidylcholine as a direct intermediate to TAG synthesis

The second mechanism (Fig.1-7) that influences TAG fatty acid composition makes PC an intermediate for TAG synthesis (Bates & Browse, 2012). A newly discovered, plant specific, enzyme called PC:DAG cholinephosphotransferase (PDCT) plays a key role in this process. PDCT catalyzes the transfer of choline-phosphate head group of PC to DAG generating new molecular species of DAG and PC (Bates & Browse, 2012; Lu *et al*., 2009).

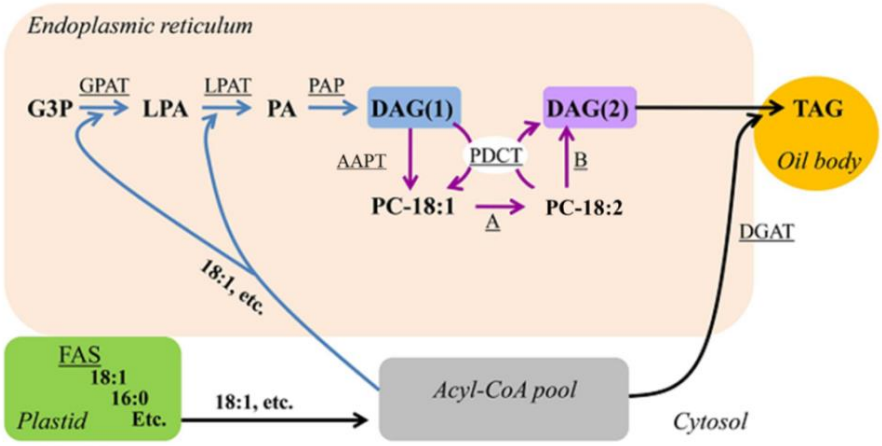

Fig.1-7 PC-derived DAG/TAG synthesis (Fig.3 from Bates and Browse, 2012). DAG from Kennedy pathway (DAG(1) can be converted to PC by AAPT and/or PDCT. Fatty acids can then be modified (desaturated) while esterified to PC. Then, PDCT converts these modified PC molecules to DAG (DAG(2)) molecules that will be used to synthesize TAGs. Blue arrows indicate reactions involved in *de novo* DAG(1) synthesis. Purple arrows indicate reactions involved in PC-derived DAG(2) synthesis. **Abbreviations**: DAG, diacylglycerol; G3P, glycerol-3-phosphate; LPA, lyso-phosphatidic acid; PA, phosphatidic acid; PC, phosphatidylcholine; TAG, triacylglycerol; AAPT, aminoalcoholphosphotransferase; DGAT, acyl-CoA:DAG acyltransferase; FAS, fatty acid synthesis; GPAT, acyl-CoA:G3P acyltransferase; LPAT, acyl-CoA:LPA acyltransferase; PAP, PA phosphatase; PDCT, PC:DAG cholinephosphotransferase. Enzymatic reactions are underlined: A, modification of FA esterified to PC; B, reversible AAPT or phospholipase C or phospholipase D/PAP DAG production.

The process starts with production of PC from *de novo* DAG (DAG(1) in Fig.1-7, directly generated by Kennedy pathway) by AAPT or PDCT. Once PC is formed, the FAs esterified to PC are available for FA modification to generate new molecular species of PC. Then PDCT transfers the phosphocholine headgroup from PC to DAG, generating new molecular species of DAG (DAG(2) in Fig.1-7) that can be then used by DGAT to synthesize TAGs. The interconversion of the PC and DAG pool by PDCT is thus a mechanism to efficiently channel *de novo* DAG(1) into PC for further FA modification, and to liberate PC-derived DAG(2) containing modified FAs for TAG synthesis. The *rod1* mutation of *Arabidopsis* (Lemieux *et al*., 1990), later found to affect PDCT (Lu *et al*., 2009), leads to a 40% decrease of PUFAs (18:2 and 18:3) in seed TAG, while the *Arabidopsis* triple mutant *pdct1/lpcat1/lpcat2* resulted in a drastic reduction of PUFA content in seed TAG, accumulating only one-third of the wild-type level (Bates *et al*., 2012). These results indicate LPCAT-based acyl editing

and DAG produced by PDCT together are responsible for at least 2/3 of the flux of PUFA from PC to TAG in wild *Arabidopsis* seeds (Bates *et al*., 2012, 2013). Moreover, PDCT can enhance the fluxes of fatty acids through PC and enrich TAG in modified or unusual fatty acids in *in planta* (Hu *et al*., 2012; Wickramarathna *et al*., 2015). For instance, the seed-specific co-expression of *PDCT* from castor (*Ricinus communis*) significantly increased hydroxyl fatty acid accumulation in transgenic *Arabidopsis* oil (Hu *et al*., 2012). Likewise, two PDCTs from Flax (*Linum usitatissimum* L.), which accumulates rich α-linolenic acid (18:3 $^{cis\Delta 9, 12, 15}$) in seeds, can substantially increase 18-PUFA levels in yeast and transgenic *Arabidopsis* seeds (Wickramarathna *et al*., 2015). Thus, discovery of PDCT enzyme helps to explain the considerable flux from PC back into DAG for TAG synthesis; one cannot exclude a contribution of AAPT catalyzing reverse reaction, which is synthesis of DAG and CDP-choline from PC (Chapman & Ohlrogge, 2012).

All plants have the capability to utilize each of these mechanisms to produce TAG with various FA compositions (Bates & Browse, 2012). However, the relative flux of TAG assembly from *de novo* DAG or PC-derived DAG may vary widely depending on plant species. Labeling experiments in soybean and *Arabidopsis* both showed that 90% of *de novo* DAG are initially utilized to synthesize PC during TAG synthesis (Bates & Browse, 2012; Bates *et al*., 2009). In such case, PC and DAG fatty acid compositions are very similar, while differences in DAG and PC composition in other plants might be explained by a minor contribution of PDCT pathway.

1.3.3 Conclusions on oil assmebly

To conclude, it now clearly appears that TAG and PC metabolisms are deeply intricate. The question now is how, upon a huge flux of fatty acids to TAGs, are cells capable of keeping unchanged the fatty acid composition of PC and other phospholipids, in order to maintain membrane properties? Based on labelling experiments carried out in soybean embryos (Fig.1-8), Bates *et al*. (2009) proposed a flux model to TAG synthesis. The simplest model that fitted best experimental data

(Fig.1-8) predicts two distinct pools of PC and three of DAGs. Because the pools are largely untractable by biochemical methods, we hypothesized that there might exist different multi-enzyme complexes that would form specific metabolons, explaining these different pools and preventing modified PC destined to TAG synthesis to spread into membranes. We chose oil palm as a model because it is possible, contrarily to *Arabidopsis*, to obtain large amounts of material. In addition, oil palm shows interesting features which are exposed below.

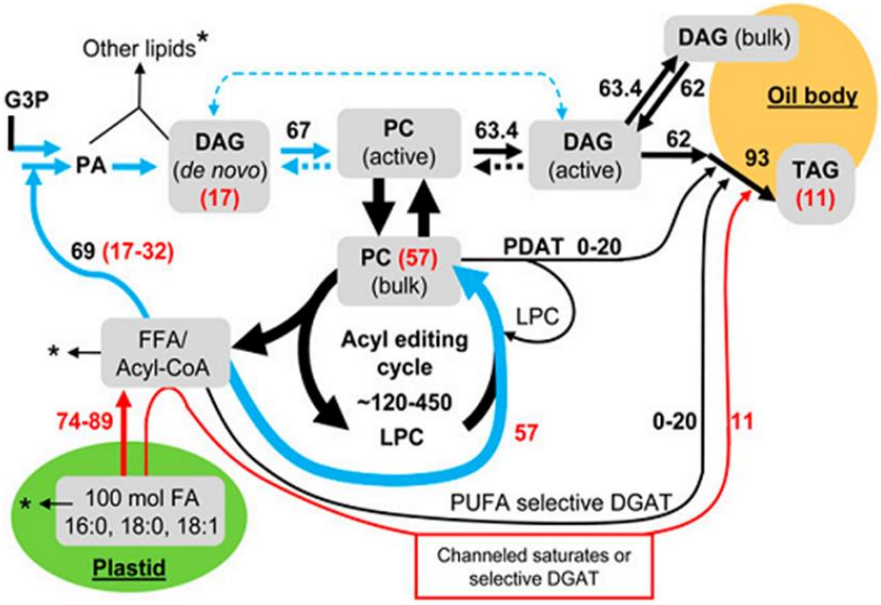

Fig.1-8 Soybean TAG synthesis acyl flux model (Fig.10 from Bates *et al*., 2009). The model describes fluxes and pools involved in TAG synthesis in soybean embryos. Thickness of arrows indicates relative importance of fluxes. Main fluxes are, by decreasing importance, to PC, to TAG through Kennedy plus PDCT pathway, and minor fluxes through Kennedy pathway (PUFA or saturate- selective DGATs) and through PDAT. Major metabolite pools proposed from kinetic analysis are highlighted in gray; there are three distinct pools of DAGs and two of PC. The red numbers represent the initial accumulation of nascent FA into the major labeled glycerolipids PC, DAG, and TAG, which together accumulate approximately 85% of initial FA. Approximately 15% of nascent FA initially accumulates in lipids other than PC, DAG, and TAG.Dashed arrows represent uncertainty for that reaction. The black numbers represent total net steady-state fluxes or flux ranges as calculated. The model follows 100 mol of newly synthesized FA in the plastid through the major fluxes of lipid metabolism such that the final accumulations of FAs are TAG, 93 mol; PC, 3.6 mol; DAG, 1.4 mol; and other lipids, 2 mol. The asterisks represent small fluxes of acyl groups to these other lipids, for which products and/or mass flux analysis have not been determined. **Abbreviations:** G3P, glycerol-3-phosphate; PA, phosphatidic acid; DAG, diacylglycerol; PC, phosphatidylcholine; TAG, triacylglycerol. PUFA, poly unsaturated fatty acid; DGAT, acyl-CoA:DAG acyltransferase; PDCT, PC:DAG cholinephosphotransferase; PDAT, phospholipid:DAG acyltransferase.

1.4 Oil metabolism in oil palm

Oil palm (*Elaeis guineensis* Jacq.), a monocotyledonous, perennial woody oil crop (Corley & Tinker, 2003), is the highest oil yielding plant in the world, accounting for approximately 40% of the total world vegetable oil production (Bourgis *et al*., 2011; Sambanthamurthi *et al*., 2000; Sundram *et al*., 2003). While the average yield of palm oil is about 4 tons/ha/yr, rapeseed produces 1 ton and soybean, second largest oil crop, 0.4 ton only. Thus, it takes ten times less acreage of palm to produce same amount of oil than soybean. Palm shows also much room for improvement as best genotypes in best agroecological conditions can produce up to 12 tons oil/ha/yr and it has been predicted that upper limit of yield should be near 17 tons oil/ha/yr (Corley & Tinker, 2003; Murphy, 2009; Rafii *et al*., 2012; Tan *et al*., 2009). Thus, oil palm has the potential to fulfill the large and growing world demand for vegetable oils that is estimated to reach 240 million tons by 2050 (Barcelos *et al*., 2015; Corley, 2009).

There are two major species in the genus *Elaeis*: *E. guineensis* and *Elaeis. oleifera* (Corley & Tinker, 2003). *E. guineensis* originates from West Africa and this species is the main (>95%) commercial planting material. *E. oleifera* from South American is a stumpy plant with a very low yield in oil but with higher oleic and linoleic acid content and lower content of palmitic and other saturated acids compared to the oil from *E. guineensis* mesocarp (Corley & Tinker, 2003; Sambanthamurthi *et al*., 2000). Such oil is more attractive to consumer than highly saturated (about 50% 16:0 and 18:0 fatty acids) palm oil from *E. guinensis*. *E. oleifera* shows other interesting traits such as low lipase activity in mesocarp, lower height increment, and resistance to diseases that affect *E. guineensis* and there are important breeding efforts to introgress these traits into *E. guineensis*. However, only interspecific hybrids are available now, with acceptable oil yields but they require hand-pollination which increases costs. One needs to note that oil palm belongs, as well as coconut, to the sub-family of Cocosoideae (Corley & Tinker, 2003), so these two species are closely related.

The oil palm is monoecious, producing separate male and female inflorescences in

leaf axils on the same palm. The oil palm fruit (Fig.1-9a) is a sessile drupe, which forms in a tight bunch containing about 1000-3000 fruitlets. The fruits on a bunch do not ripen simultaneously owing to slight variation in the time of pollination. Also, fruits at the end of a spikelet ripen first and those at the base last. Fruits on the outside of the bunch are large and deep orange when ripe while the inner fruits are smaller and paler (Sambanthamurthi *et al*., 2000). The fruit pericarp (Fig.1-9b) comprises three layers, including the exocarp (skin), mesocarp (outer pulp containing palm oil), and endocarp (Fig.1-9b). The endocarp is a hard shell enclosing the kernel, which contains a small embryo and large endosperm, both storing oil also.

Fig.1-9 (a) Crown of palm tree originating from intertropical Africa (*Elaeis guineensis*), showing several bunches at different stages of development. Trees produce continuously bunches all around the year. (b) Open fruit of oil palm: kernel, the central white part is endosperm, containing saturated medium-chain fatty acids oil (about 60% dry weight). It is surrounded by endocarp (black) **and mesocarp** (orange) part. Palm oil makes almost 90% of mesocarp dry weight; it is mostly composed of palmitic and oleic acids. Orange color is due to carotenoids, which accumulate in fruit chromoplast together with tocopherol and tocotrienol.

Based on the thickness of shell controlled by the *SHELL* gene, the oil palm fruits have been classified into three forms: *dura* (*Sh/Sh*, thick-shelled), *pisifera* (*sh/sh*, shell-less) and *tenera* (*Sh/sh*, thin-shelled), a hybrid between *dura* and *pisifera* (Corley & Tinker, 2003; Sambanthamurthi *et al*., 2000; Singh *et al*., 2013a). Because size of fruit remains similar, the room left due to reduced endocarp is occupied by oil-producing mesocarp, so that trees with thinnest endocarp produce highest amounts of oil. Unfortunately, the *pisifera* palm is mostly female sterile and exhibits fruit bunch abortion so that it is not used as a commercial planting material

(Sambanthamurthi *et al*., 2000; Ting *et al*., 2014). Therefore, *tenera* hybrids are the most productive trees and have been widely planted since 1961 (Sambanthamurthi *et al*., 2000; Ting *et al*., 2014).

Both mesocarp and kernel of oil palm fruits accumulate high amounts of oil (approximately 90% and 50% dry mass, respectively) and 90% of oil produced by oil palm is mesocarp oil (palm oil) and 10% kernel oil. Palm oil mainly contains palmitic acid (16:0, about 43%), oleic acid (18:1, about 45%), linoleic acid (18:2, about 8%) and stearic acid (18:0, about 4%). Kernel oil is very different: it contains large amounts of MCFAs, especially lauric acid (12:0, 50%) and myristic acid (14:0, 17%) (Corley & Tinker, 2003; Siew *et al*., 1995). This composition is very close to that of oil from coconut endosperm (Santoso *et al*., 1996). Palm oil is widely used in food industry, especially in applications that take advantage of its highly saturated fat content, such as cooking oil (high heat stability), shortenings and margarines (solid at room temperature), but it is also widely used as table oil in producing countries (South East asia, Africa). Solid fat might be removed by cold fractionation. More recently, palm oil has started to be used as biofuels as it is the cheapest plant oil to produce (Corley, 2009; Murphy, 2009), with production costs of about $300 a ton, very comparable to that of petroleum. Because of its high levels of unusual MCFAs, kernel oil is mostly destined to oleochemistry and is widely used to make detergents for example (Kok *et al*., 2013; Siew *et al*., 1995).

For a long time, research on oil palm has been almost exclusively carried out by breeders, or in those emerging countries (Malaysia) that developed large scale palm cultivation programs. These include mostly work carried out at CIRAD and IRD (Montpellier, France), MBOP (Malaysia) and by Unilever (mostly in the eighties); this work was largely destined to breeding applications. For about five years now, tremendous progresses have been made, mostly due to use of next generation sequencing techniques, providing resources that makes oil palm an attractive model for fundamental studies on oil synthesis.

1.4.1 Oil palm research related to oil synthesis

Very few studies have been carried out to characterize lipid-related enzymes either directly from palm fruit tissues or by heterologous expression. These include studies on a purified mesocarp Glycerol 3-phosphate acyltransferase (GPAT), which displayed a strong preference for palmitoyl-CoA compared to other acyl-CoA donors (Manaf & Harwood, 2000). Also an *acyl-ACP thioesterase* from oil palm was expressed in *E.coli* and the protein showed specificity towards 14:0-ACP and 16:1-ACP (Jing *et al*., 2011). *Thioesterases* were also transiently expressed in tobacco and impact on oil composition evaluated (Dussert *et al*., 2013). A kernel DGAT1-1 was found to show preference for MCFAs (Aymé *et al*., 2015) and a lipase, which physiological function remains unknown but which greatly impacts palm oil quality was characterized (Morcillo *et al*., 2013). As mentioned earlier, progress came with sequencing studies. First, 85% of genome (1.8 gigabases) was sequenced in 2013 (Singh *et al*., 2013b) and subsequent analysis led to annotation of about 35,000 gene models. The *SHELL* gene was cloned based mostly on association studies (Singh *et al*., 2013a). The molecular basis of an epigenetic modification affecting severely productivity of plant clones derived from *in vitro* cell cultures was identified, allowing early detection and elimination of anomalous plants (Ong-Abdullah *et al*., 2015). First genome-wide association study related to yield identified three candidates (Teh *et al*., 2016). Numerous transcriptomes have been recently published, using Sänger (Jouannic *et al*., 2005) and new generation (Bourgis *et al*., 2011; Dussert *et al*., 2013; Singh *et al*., 2013b; Tranbarger *et al*., 2011) technologies and there are now more than 300 million sequences available in GenBank. At last, other system's biology approaches have been published, including proteomic (Ooi *et al*., 2015) and metabolomics studies (Teh *et al*., 2013). We will now give details on oil accumulation in oil palm fruit upon ripening.

1.4.2 Characterization of oil accumulation during fruit ripening

Oil palm fruit is ripe about 22 to 25 weeks after pollination (WAP), depending on

climate (ripening is faster in South-East Asia than in Africa), and development pattern is illustrated in Fig.1-10 (Tranbarger *et al*., 2011).

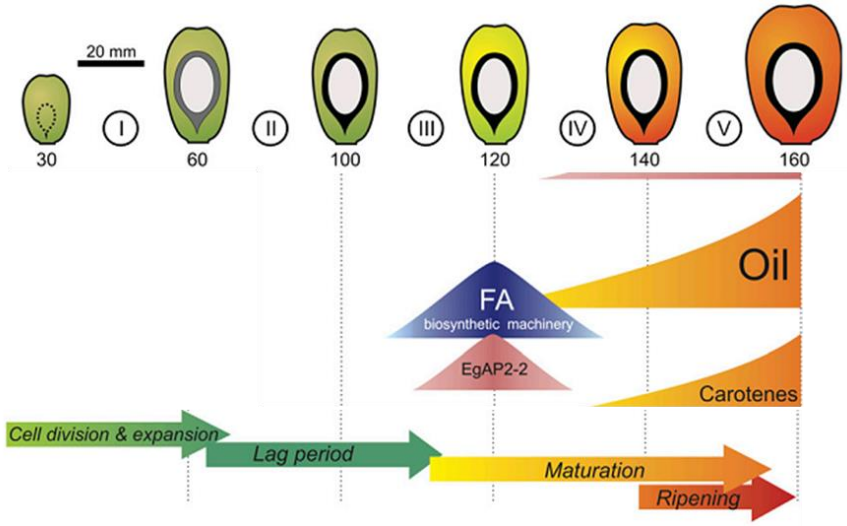

Fig.1-10 Oil accumulation in mesocarp during fruit development (Fig.6 from Tranbarger *et al*., 2011). Numbers of days after pollination are indicated below fruit drawing and it takes about 160 days for a fruit to ripen. Kernel develops first and reaches maturity about 120 days after pollination, when oil starts to accumulate in mesocarp. Then, mesocarp chloroplasts evolve in carotenoid-accumulating chromoplasts, concomitant to oil accumulation. Fatty acid (FA) synthesis is under the control of WRI1 transcription factor, which belongs to APETALA transcription factor family (EgAP2-2).

Kernel is liquid at 8 WAP, turns semi-gelatinous at 10 WAP and progressively hardens until maturity (14-20 WAP). Most of kernel is endosperm and embryo represents 0.4% (dry weight) of mature seed only. Oil accumulation rate in kernel are highest during one week or so between 9-12 WAP and slows down afterwards. At maturity, most of the oil (50-60% dry weight of endosperm) is made of MCFAs, and main fatty acids are 12:0 (about 49%), 14:0 (about 17%), 18:1 (about 17%), 16:0 (about 10%), rest been made of 18:0, 18:2, 6:0 and 8:0. Ripening is uneven in mesocarp, starting from proximal end and endocarp towards distal end and epidermis. Mesocarp is green initially and turns yellow at around 18 WAP. This indicates transformation of chloroplast into chromoplast and beginning of both oil and carotenoids synthesis. Highest rates of oil synthesis span over about two weeks between 19-21 WAP and continue at lower rates until fruit is ripe, about a week before fruit abscission. Ripening is uneven in bunches (about one week difference),

with external fruits ripening first. Oil palm is made of 16:0 (about 43.5%), 18:1 (about 40%), 18:2 (about 10%), 18:0 (about 4%), and 14:0 (about 1%). Embryo stores oil also (about 26% dry weight) and main fatty acids are 18:1 (about 37%), 18:2 (about 21%) and 16:0 (about 25%) (Dussert *et al*., 2013).

Several studies reporting on lipid composition of both oil palm mesocarp and kernel during ripening have been published (Bourgis *et al*., 2011; Dussert *et al*., 2013; Kok *et al*., 2013; Sambanthamurthi *et al*., 2000; Wiberg & Bafor, 1995; Prada *et al*., 2011). These studies include determination of total fatty acid content, main polar lipids and HPLC analyses of molecular species. The only LC-MS developmental study of molecular species content during ripening concern interspecific hybrid from *Elaeis guineensis* x *E. oleifera* (Mozzon *et al*., 2013). Analysis of fatty acid at position *sn*-2 of TAGs by pancreatic digestion of whole palm oil indicates that *sn*-2 position is predominantly occupied by 18:1 (65%) and 18:2 (20%).

Transcriptome analyses using 454 deep sequencing have been published by two groups: Bourgis *et al*. (2011) used about 700 000 reads per stage and Tranbarger *et al*. (2011) and Dussert *et al*. (2013) 100 000 to 200 000 reads per stage. Bourgis *et al*. (2011) compared oil palm mesocarp to leaf and to date palm mesocarp (which stores sugars and no oil). Tranbarger *et al*. (2011) worked on mesocarp and Dussert *et al*. (2013) on endosperm and embryo. Results and conclusions from these studies are about the same and are as follows.

Oil synthesis in mesocarp is associated to strong (at least 5-fold) up-regulation of genes coding for all enzymes of fatty acid synthesis, select plastid glycolytic enzymes (such as pyruvate kinase), plastid sugar transporters, WRI1 transcription factor and DGAT2. These data were confirmed by comparison with date palm data, and EgWRI1 was found to complement *Arabidopsis wri1* mutant (Ma *et al*., 2013). Other genes of central metabolism were found to be up-regulated such as those coding for cell wall invertases. Strangely, and contrarily to what occurs in seeds (Troncoso-Ponce *et al*., 2011), genes for oil assembly enzymes, of Kennedy pathway

were not up-regulated. However, comparison with date palm pointed out to genes more (LPCATs, some PA phosphatases) or less (PDCT, some PA phosphatases) expressed in oil palm mesocarp. A few other differences were noted with oil seeds (including oil palm kernel), especially concerning use of ATP-requiring enzymes. Seeds are known to show preference for using enzymes of central metabolism that work with pyrophosphate rather than ATP (Troncoso-Ponce *et al*., 2011). This is not the case for oil palm mesocarp, as transcripts coding for ATP-dependent enzymes of central metabolism are higher than those coding for enzymes using pyrophosphate. Interestingly, Dussert *et al*. (2013) identified three different genes homologous to WRI1. *WRI1-1* was predominantly transcribed in the mesocarp whereas *EgWRI1-2* was mainly expressed in endosperm. Both were able to increase fatty acid levels when transiently expressed in tobacco (*Nicotiana benthamiana*) leaves. *EgWRI1-3* transcript was mainly detected in endosperm, at lower levels than *EgWRI1-2*. Dussert *et al*. (2013) also identified three isoforms of FatB enzymes and one of them, expressed in kernel, showed preference for mid-chain acyl-ACP (14:0 mostly) when over-expressed in tobacco leaves.

Proteomic studies (Loei *et al*., 2013), carried out on fruit mesocarp from low- and high-yield palm trees, confirmed transcriptomic studies with regards to glycolytic and fatty acid synthesis enzymes. In addition, they pointed to differential expression of TCA (tricarboxylic acid) cycle enzymes and proteins linked to oxidative phosphorylation, such as a subunit of ATPase, as the major difference between high and low producing trees. At last, Ramli *et al*. (2009) conducted metabolic control analyses on oil palm cell cultures and found that fatty acid synthesis carried most control (about 2/3) on oil synthesis, twice as much as fatty acid assembly into oil (1/3) did. However, these studies have been carried out on material that produces 500-fold less oil than mesocarp cells do and they might not reflect was happens in mesocarp.

1.5 Thesis project

Because oil assembly pathway is complicated by editing processes involving PC metabolism, it is clear that oil assembly pathway in palm mesocarp must be tightly regulated to keep phospholipid content of membranes untouched by this high flux of fatty acids to TAGs. Transcriptomic studies have shown that genes coding for TAG assembly pathway exhibited very little up-regulation, which suggests that regulation should rather be searched at post-transcriptional levels. Flux model from soybean embryos pointed to importance of several distinct pools of lipid intermediates for efficient channeling of fatty acids to TAGs. A simple way of explaining at least some of these pools would be the existence of multi-enzyme complexes that would specifically channel fatty acids through metabolons to TAG synthesis, without leading to important modifications of membrane phospholipids. The main aim of my thesis project was to provide evidence for the existence of such complexes. Results are presented in three chapters. The first one (Chapter 3) concerns the use of lipidomics as a tool to better characterize oil accumulation in mesocarp and also in kernel. The second chapter (Chapter 4) concerns the study of interactions between main palm mesocarp enzymes and proteins involved in the oil assembly pathway. I used the split-ubiquitin yeast two-hybrid system to carry out this study. The third chapter (Chapter 5) concerns enzyme studies by heterologous expression in yeast, to characterize a putative acyltransferase possibly involved in oil assembly and the three major DGAT isoforms of palm mesocarp.

Chapter 2. Materials and methods

2.1 Materials

2.1.1 Yeast and *E. coli* strains and plasmids used in this study

Yeast (*Saccharomyces cerevisiae*) strain and plasmids (Table 2-1 and 2-2), specifically used in yeast two-hybrid experiments are indicated by (Y2H). The quadruple mutant yeast H1246 that cannot synthesize TAGs (Sandager *et al*., 2002) and plasmid pESC-URA containing two multiple cloning sites (Agilent) were used for testing gene function *in vivo*, with wild type SCY62 as positive control (Table 2-1 and 2-2). We used *E. coli* strain DH10B (New England BioLabs) for cloning experiments.

Table 2-1. Yeast and *E.coli* strains used in this study

Strain	**Relevant genotype**
THY.AP4 (Y2H)	*MATa leu2-3,112 ura3-52 trp1-289 lexA::HIS3 lexA::ADE2 lexA::lacZ*
SCY62 (wild type)	*MATa ADE2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*
H1246 (quadruple mutant)	*MATα ADE2-1 can1-100 ura3-1 are1-Δ::HIS3 are2-Δ::LEU2, dga1-Δ::KanMX4 lro1-Δ::TRP1*
E.coli DH10B	*Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)*

Table 2-2. Plasmids used in this study

Plasmid		**Auxotrophic marker (Yeast)**	**Resistance marker (*E. coli*)**
Bait vectors (Y2H)	pBT3-N	*LEU2*	Kanamycine
	pBT3-STE	*LEU2*	Kanamycine
Prey vectors (Y2H)	pPR3-N	*TRP1*	Ampicillin
	pPR3-STE	*TRP1*	Ampicillin
	pPR3-SUC	*TRP1*	Ampicillin
pESC-URA		*URA3*	Ampicillin

Y2H means yeast two-hybrid system.

2.1.2 Palm materials

All oil palm biological material was from the Center for Oil Palm Research (CEREPAH) of La Dibamba, harvested by Dr. G. Ngando-Ebongue or G. Ntsomboh-Ntsefong. Aside from diversity studies, material used was progenies

resulting from two self crosses of DA115D, a tree that belongs to Deli-dabou population. Most oil palm fruits used in this study have been flash-frozen in liquid nitrogen a few minutes after harvest and stored at -80 °C until use. After fast removal of pericarp, mesocarp (still frozen) has been dissected and powdered in liquid nitrogen, generating samples previously used in the laboratory for proteomic and metabolomics analyses, and that I used for lipidomics analyses. Likewise, oil palm leaves were also powdered in liquid nitrogen. Fresh fruits were also used, a few days (3-7) after harvest, for microsome isolation mostly.

2.2 Sequence analyses

CLC Main Workbench software has been used to carry out most bioinformatics analyses of DNA sequences including sequence analysis from Sanger sequencing, sequence comparisons, restriction mapping, defining oligonucleotides, etc. Also, we mapped trimmed RNAseq data from Bourgis *et al*. (2011) and Illumina unpublished reads onto 2015 oil palm genome release from NCBI using CLC Genomic workbench set at default settings.

Topology of proteins was analyzed by following methods: TOPCONS (Bernsel *et al*., 2009; http://topcons.cbr.su.se/), TMHMM (Krogh *et al*., 2001; http://www.cbs.dtu.dk/services/TMHMM/), HMMTOP (Tusnády & Simon, 2001; http://www.enzim.hu/hmmtop/html/submit.html), and Phobius (Käll *et al*., 2004; http://phobius.sbc.su.se/). SignalP 4.1 Server was used for predicting the presence of signal peptide and cleavage site in sequences of these proteins (Petersen *et al*., 2011; http://www.cbs.dtu.dk/services/SignalP/).

2.3 RNA extraction and cDNA synthesis

For RNA work, Tris-HCl (pH 8) was prepared using diethylpyrocarbonate (DEPC)-treated water, and all other solutions were treated with 0.1% DEPC. Total RNA from oil palm leaves and mesocarp were isolated using the method described by Chirgwin *et al*. (1979). Frozen powder was quickly homogenized in following buffer

(5 mL per gram): 5 M guanidine thiocyanate, 100 mM Tris-HCl (pH 8), 0.5% lauroyl sarcosine, 1% insoluble polyvinylpolypyrolidone (PVPP), and 1% β-mercaptoethanol. Debris were removed by centrifugation at 10 000 *g* for 10 min, and the supernatant (avoiding fat layer on top) was carefully overlaid onto a 9 mL cushion made of 5.7 M CsCl and 10 mM EDTA (pH 7.5). Ultracentrifugation was carried out in 36 mL tubes using a Sorvall AH629 rotor at -20 ℃ and centrifuged at 25 000 rpm (112 000 *gmax*) for 26 h. Supernatant was carefully removed and discarded. The RNA pellet was resuspended in 10 mM Tris-HCl (pH 8), 1 mM EDTA, and 0.1% SDS buffer. For kernel, RNA was successfully obtained using Meisel *et al*. (2005) protocol. Frozen powder was homogenized at 65 ℃ using an Ultraturax set at high speed into pre-warmed Tris-HCl 100 mM (pH 8), 25 mM EDTA, 2.5 M NaCl, 2% Hexadecyltrimethylammonium bromide, 2% PVPP, 0.05% spermidine, and 2% β-mercaptoethanol, using at least 10 mL of buffer per gram of frozen powder. The mixture was incubated for 15 min at 65 ℃, extracted twice with one volume of chloroform/isoamyl alcohol (24/1, v/v) and RNAs were precipitated overnight by 2M LiCl at 0 ℃. RNA were pelleted, re-suspended in Tris 10mM pH8, EDTA 1mM, 0.1% SDS, extracted again with chloroform/isoamyl alcohol, precipitated with ethanol and final pellet re-suspended in Tris 10mM pH8, EDTA 1mM. RNAs were quantified using a nanodrop spectrophotometer and stored at -80 ℃ until use.

First-strain cDNA was carried out using SuperScriptTM II reverse transcriptase (InVitrogen) according to manufacturer's instructions, using $Oligo(dT)_{12-18}$ oligonucleotide as primer.

2.4 PCR amplification and cloning methods in *E. coli*

PCR was performed using cDNA from oil palm leaves, mesocarp or kernel as template (or plasmids in case of subsequent cloning for yeast expression studies) and Q5 High-Fidelity (New England BioLabs) or Phusion High-Fidelity (Thermo Scientific) as DNA Polymerases under the condition: 98 ℃ for 30 sec for initial denaturation, then 35 cycles (25 cycles when using plasmids as template) of: 98 ℃ for

10 sec, optimal temperature for 30 sec, 72 ℃ for extension time (30s/kb). A final extension step at 72 ℃ for 5 min was then carried out. The templates, primers, optimal Tm, length, extension time and vectors are listed in Table S1 and S2 in the Appendix I. PCR success was assessed by loading an aliquot of the reaction on an agarose gel run in TAE buffer at 100V (Sambrook *et al*., 1989).

PCR products were purified using commercially available kits (Wizard SV Gel and PCR Clean-Up System; Promega) according to manufacturer's instructions, starting either with whole reaction mixture when a single band was detected or (rarely) with a band cut from a preparative agarose gel.

Digestion with restriction endonucleases (New England BioLabs) was carried out as follows: DNA (PCR product or plasmid) was incubated for 1-2 hours at enzyme optimal temperature (50 ℃ for *Sfi*I and 37 ℃ for most other restriction enzymes) in appropriate buffer (chosen according to manufacturer's instruction) with about 5 units of enzyme/μg of DNA for purified plasmid and twice that amount for PCR products. When plasmids were used as cloning vectors, digestion was assessed using an aliquot ran on agarose gel and, if successful, phosphatase was added to the mixture which was further incubated at 37 ℃ for 30 min. Restriction products were then purified using appropriate kit (Wizard SV Gel and PCR Clean-Up System; Promega) and final recovery was estimated by measuring DNA concentration with a nanodrop spectrophotometer.

Ligations were carried out mostly at 16 ℃ overnight in 10 μL reactions with about 100ng of vector and an equimolar amount of insert (in case of difficulties/strong background, the insert/vector ratio was 3 on a molar basis), following manufacturers instructions with regards to buffer composition and amount of enzyme (T4 DNA ligase; New England BioLabs). Transformation of *E. coli* was done by electroporation, using home-made electrocompetent cells obtained according to Sambrook *et al*. (1989). Briefly, cells are grown to exponential phase in LB medium at 37 ℃ (OD_{600} = 0.5-0.6); they are spun down and washed repeadly in ice-cold 10% glycerol to remove

salts. During final wash, they are concentrated to about 1/100 of initial culture medium, aliquoted in microtubes, flash-frozen in liquid nitrogen and stored at -80 ℃ until use. Electroporation is carried out using 50 µL of cells and 2 µL of ligation mixture, in a 1mm-gap cuvette; voltage is set at 1200 V (so that time constant is between 5 and 10 ms) and 500 µL of SOC medium are added immediately after the pulse. Cells are incubated for 1h at 37 ℃ and an appropriate amount is plated onto solid LB medium with appropriate antibiotic; plates are incubated at 37 ℃ overnight.

Minipreparations of *E. coli* plasmids are performed from 3 mL of a saturated LB culture of the clone of interest, using commercial kits (GenElute Plasmid Miniprep Kit; Sigma-Aldrich). Colonies can be screened prior to minipreparation by PCR made on *E. coli* colonies; this procedure is helpful if one expects a poor success rate of the transformation procedure.

Presence of insert is checked by digesting plasmid with restriction enzyme(s) used for cloning. Then, clones are sent for full sequencing (Beckman-Coulter Company) on both strands to verify if the clone sequence is correct.

2.5 Yeast work

2.5.1 Yeast transformation using the LiAc/single-stranded carrier DNA/PEG method

Yeast single or double transformation (for two hybrid system) was performed according to the manufacturer's instruction with some modifications (Gietz *et al*., 1995; Ito *et al*., 1983). Briefly, yeast cells were precultured overnight in 2 mL YPAD medium at 30 ℃ under cntiuous agitation (250 rpm). Then, the preculture was diluted with fresh medium to an OD_{546} of 0.2 in 50 mL, and grown again at 30 ℃ to OD_{546} at 0.6-0.8 (about 4 hours). Yeast cells were pelleted, washed once with sterilized water, gently re-suspended in 2 mL of 1 mM LiOAc, and divided into 5-10 microtubes (1.5 mL) equally. Then, cells were pelleted at 5000 rpm for 5 min, and supernatant was removed as much as possible. To cell pellets were added in following order 240 µL of

PEG (50%), 36 μL of 1M LiOAc, 25 μL of single-strand carrier DNA, 50 μL of sterilized water, and 1 μL (about 100 ng/μL) of (each) plasmid and the tube content was mixed by pipetting up and down thrice. Tubes were firstly incubated at 30 ℃ for 15 min, and transferred to water bath at 42 ℃ for 25 min. Then, the transformed cells were pelleted at 1500 *g* for 5 min, and were resuspended in 100 μL of ddH_2O. Cells from each transformation were then plated onto appropriate selective agar medium based on the auxotrophic growth marker(s) of transformed plasmid in yeast, and were incubated at 30 ℃ for 3-4 days.

2.5.2 Split-ubiquitin yeast two hybrid system

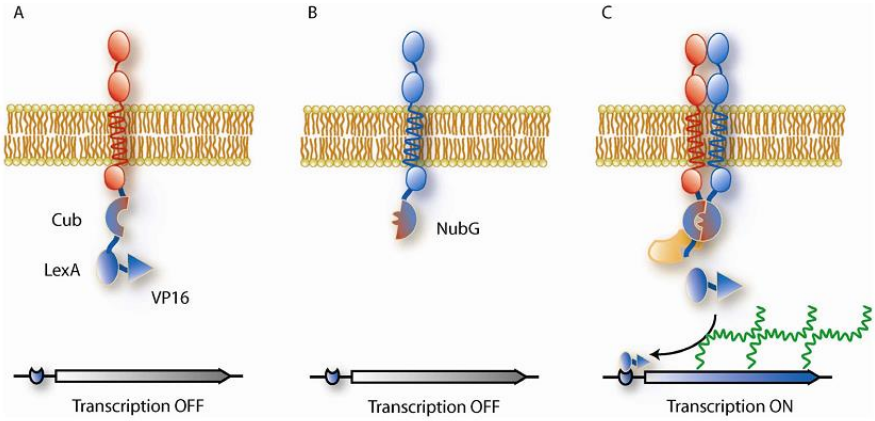

Fig.2-1 the principle of Y2H membrane system. Cub: the C-terminal of half of ubiquitin; NubG: the mutated N-terminal half of ubiquitin; LexA-VP16: the artificial transcription factor. UBPs: ubiquitin specific proteases.

The Y2H membrane (Stagljar *et al*., 1998; Thaminy *et al*., 2003) system is based on functional reconstitution of ubiquitin from two split parts. A first membrane protein of interest (the bait) is fused to the C-terminal half of ubiquitin (Cub) and the artificial transcription factor LexA-VP16 (Fig.2-1A). A second protein of interest (the prey) is fused to the mutated N-terminal half of ubiquitin (NubG) (Fig.2-1B). If bait and prey interact, NubG and Cub are forced into close proximity, resulting in functional reconstitution of split-ubiquitin. Reconstituted ubiquitin is immediately recognized by ubiquitin specific proteases (UBPs) which then cleave the polypeptide chain between Cub and LexA-VP16 (Fig.2-1C). As a result, the artificial transcription factor is released from the membrane and translocates to the nucleus where it activates the

expression of two auxotrophic growth markers (HIS3 and ADE2) that enables yeast to grow on defined minimal medium lacking histidine and/or adenine. It also allows expression of the lacZ gene which encodes β-galactosidase.

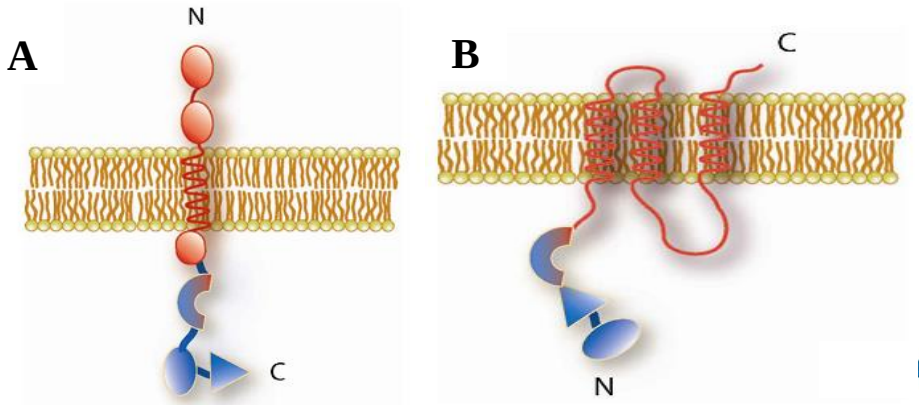

Fig.2-2 the different modules for proteins with different topologies. A: the integral membrane protein with N_{out}/C_{in} topology; the Cub-LexA-VP 16 cassette is fused to the C terminus of the bait. B: the integral membrane protein with N_{in}/C_{out} topology; the orientation of the module is reversed in pBT3-N (LexA-VP 16-Cub) to allow separation of the transcription factor from the bait *via* UBP cleavage.

The fundamental requirement of the Y2H membrane system is that the Cub-LexA-VP16 module of the bait and the NubG module of the prey must be located in the cytosol. Three different bait and prey vectors are made for protein with different topologies. As described in the protocol of DUAL membrane split-Ubiquitin system (Dualsystems Biotech), the vector pBT3-STE (C-terminal fusion) is used for those integral membrane proteins with a N_{out}/C_{in} topology (lumenal N-terminus, cytosolic C-terminus) (Fig.2-2A). If this kind of protein has a signal sequence, the N-terminal signal sequence will be removed from cDNA and the truncated cDNA will be cloned into the vector pBT3-SUC. As for the other type of proteins which have a N_{in}/C_{out} topology (cytosolic N-terminus, lumenal C-terminus), the vector pBT3-N (N-terminal fusion) is used for cloning since the Cub module must be located in the cytosol (Fig.2-2B).

The two *Sfi* I restriction sites with different overhangs are provided in the multiple cloning site of the bait and prey vectors, allowing directional cloning. Positive control is NubI, the wild type N-terminus of ubiquitin with isoleucine at position 3; it

interacts with Cub moiety no matter what protein it is fused to. Empty prey vector is used to check whether the bait may self-activate and NubG is used as negative control.

To test the activation of two auxotrophic growth markers (HIS3 and ADE2), a serial dilution assay is carried out on synthetic complete medium (SC) plates lacking tryptophan and leucine (-TL), and histidine (-HTL), and adenine (-AHTL). The first suspension has an OD_{600} value of 0.4, and three subsequent samples are prepared by 1:10 serial dilution from 1 to 10^{-3}. The transformed cells are washed once with ddH_2O before dilution to avoid trace amounts of medium -TL on selective plates. After 4 days incubation at 30 ℃, the yeast growth statuses on three plates are recorded. At least two independent clones from each double transformation were tested.

2.5.3 Yeast growth conditions for GAL promoter induction

Clones are grown in 2 mL SC-U medium with 2% raffinose growing at 30 ℃ overnight. Contrarily to glucose, raffinose does not repress GAL promoter, therefore allowing induction to start immediately. Precultures are spun down, washed once with sterile water and an aliquot is used to inoculate 20 mL of SC-U medium with 2% galactose in 250 mL flasks, so that final optical density (OD_{600}) of culture is 0.4. Flasks are then incubated for 24h at 30 ℃, under continuous agitation (250 rpm) and OD_{600} of culture is recorded at least three times to precisely determine amount of yeast cells as OD units; one OD unit is the amount of yeast cells that gives an OD_{600} of 1 in 1 mL. Then, a 2 mL aliquot is pelleted and an appropriate amount of fatty acid internal standard (C17) is added to quantify total fatty acids per OD unit. Other aliquots of appropriate size (carefully recorded), destinated to quantification of lipid species such as TAGs, DAGs and PC are transferred to 8 mL glass tubes with Teflon caps. Pellets are washed once with distilled water and used immediately for lipid extraction after adding appropriate internal standard or frozen at -20 ℃.

2.6 Lipid analysis

2.6.1 Lipid extraction methods from oil palm fruits and yeast

Plant tissues contain important amounts of phospholipase D and palm mesocarp a strong TAG lipase activity (Morcillo *et al*., 2013). These must be inactivated prior to lipid extraction, which is done in boiling isopropanol. Thus, about 50 mg of each frozen mesocarp or kernel powder are immediately dipped into 8 mL glass tubes with Teflon-lined caps containing 2 mL boiling isopropanol and incubated at 80 ℃ for 15 minutes. Then, 3 mL of hexane are added to bring hexane-isopropanol ratio to 3:2 (v/v) and total lipids are extracted according to the method described by Radin (1981). After addition of 2.5 mL of 6.5% (w/v) sodium sulfate, the mixture is centrifuged for 10 min at about 1000 *g*, and the upper organic phase transferred to a fresh tube. Lipids are dried down under a stream of nitrogen, re-suspended in chloroform (about 25 mg/mL) and stored at -20 ℃. When lipids need to be quantified on a dry weight basis, powdered material is lyophilized first to remove water inside, and appropriate amount of internal standard C17:0 is added at the beginning of lipid extraction procedure.

Lipids from transformed yeasts were extracted using the method described by Dittrich-Domergue *et al*. (2014). To a volume of packed yeast cells of about 300 µL are added 2 mL of chloroform/methanol (1/1, v/v) and glass beads (about 3 mL, up to meniscus). Tubes are shaken vigorously with vortex 4 times 30s. After centrifugation, organic phase is saved, beads are rinsed thoroughly once with 2 mL of chloroform/methanol (2/1, v/v) and once with 2 mL of chloroform; the two washes are combined with initial extract and 2 mL of NaCl 2.5% (w/v) in water are then added to allow phase separation. The aqueous phase is re-extracted with 2 mL chloroform and organic phase is combined with previous one and dried down. Finally, lipids are dissolved in an appropriate volume of chloroform (250 µL for 300 µL of cells).

2.6.2 Thin layer chromatography analyses

Thin layer chromatography (TLC) analyses are carried out on TLC silica Gel 60 plates (Merck), that have been pre-developed with chloroform/methanol (1/1, v/v), in order to remove possible pollutants. The solvent system methyl acetate/propanol /chloroform/methanol/KCl 0.25% (25:25:25:10:9, v/v/v/v/v) is used for separating polar lipids and hexane/diethyl ether/acetic acid (80:20:1, v/v/v) for neutral lipids. For a 7 mm wide lipid deposit, 10-20 µg of lipids are loaded, so that bands of interest represent at least 2 µg and can be easily visualized. When the plate is developed for preparative use, one aims that band of interest contains at least 5 µg lipids. Standard lipids are run side-by-side with samples to be able to identify bands of interest by co-migration. When TLC is run to further quantify lipid band, then a known amount of internal standard such as PC 19:0, DAG 15:0, TAG 17:0 is mixed to sample prior to loading on plate is it has not been already done during lipid extraction procedure. The amount of standard added is calculated so that it falls within 15-50% of the amount of band of interest. For example, we used 2 and 0.5µg for TAGs and DAG, respectively, to quantify these lipids from 10 yeast OD units, and 2 µg to quantify PC from 4 OD units. After separation, lipid bands are revealed by spraying plate with primuline and visualized under UV light at 366 nm. The spots of interest are marked with a soft pencil and scratched from the plate for further analysis by gas chromatography (GC).

2.6.3 Methylation of lipid samples

When carried out with an acid catalyst, all fatty acids (esterified or not) are methylated in presence of methanol (and in absence of water) to fatty acid methyl esters (FAMEs) which are suitable for GC analysis. About 1 mg oil from mesocarp or kernel is dissolved in 500 µL of toluene and 2 mL of methanol that contains 2.5% sulfuric acid (v/v) are added. The mixture is incubated at 80 ℃ for at least one hour (Ngando-Ebongue *et al.*, 2008). Two mL of 0.9% (w/v) sodium chloride (NaCl) are then added, together with 1 mL Hexane. After vigorous shaking and phase separation,

the upper phase is transferred to a vial for GC analysis. For yeast cells and silica powder, only 2 mL of methanol with 2.5% sulfuric acid (v/v) are used for methylation as the small amount of lipids is fully dissolved by methanol.

2.6.4 Gas chromatography (GC) analysis

FAMES can be analyzed and quantified by GC with either flame ionization (FID) or mass spectrometry (MS) detectors. For FID, we used a Hewlett-Packard 5890 Series II gas chromatograph (GC) with FID equipped with an HP-1 column (30m × 0.32 mm × 0.25 mm) as described previously (Maneta-Peyret *et al.*, 2014). The oven temperature is programmed as follows for long chain fatty acids (LCFAs): 160 ℃ for 1 min, raised to 190 ℃ at 20 ℃/min, held at 190 ℃ for 0.3 min, and then raised to 210 ℃ at 5 ℃/min, held for a further 0.3 min, subsequently raised to 230 ℃ at 20 ℃/min, and the final oven temperature was maintained for 10 min. For short chain fatty acids, the following settings were used: 100 ℃ for 1 min, raised to 190 ℃ at 20 ℃/min, held at 190 ℃ for 0.3 min, and then raised to 210 ℃ at 5 ℃/min, held for a further 0.3 min, subsequently raised to 230 ℃ at 20 ℃/min, and the final oven temperature was maintained for 6 min. The temperatures of injector and detector were 230 ℃ and 250 ℃, respectively. Helium was used as carrier gas at a flow rate of 3 mL/min (4 psi). FAMEs were identified by comparison of their retention time with those of standard FAMEs, and then were quantified according to their percentage area obtained and known amounts of internal standard (C17).

GC-MS analyses were performed using an Agilent 6850 gas chromatograph equipped with an HP-5MS column (30 m x 0.25 mm x 0.25 mm) and an Agilent 5975 mass spectrometric detector (70 eV, mass-to-charge ratio 50-750), with helium (1.5 mL/min) as carrier gas. For the resolution of FAMEs, the initial temperature of 50 ℃ was held for 1 min, increased at 25 ℃ min^{-1} to 150 ℃, held for 2 min at 150 ℃, then increased at 10 ℃ min^{-1} to 320 ℃, and held for 6 min at 320 ℃. Separation of the wax esters was achieved with the same program except that the final temperature (320 ℃) was held for 21 minutes.

2.6.5 Analysis of lipid molecular species by LC-MS

Most liquid chromatography/tandem mass spectrometry (LC-MS/MS) analyses were carried out by Laetitia Fouillen at the Bordeaux lipidomic platform and others were performed by F. Beisson and B. Légeret at the CEA center of Cadarache. The methods used by these two groups are described below.

2.6.5.1 Method used in Bordeaux by L. Fouillen

LC-MS analysis were carried out using using a QTRAP® 5500 (ABSciex) mass spectrometer coupled to a LC system (Ultimate 3000, Dionex) at the lipidomic facilities of the metabolomic platform of the CGFB (Centre de Génomique Fonctionnelle de Bordeaux, University of Bordeaux).

Triacylglycerol molecular species analyses An appropriate amount of internal lipid standard triacylglycerol (TAG) 17:0/17:0/17:0 was added to the lipid samples. Reversed phase chromatography was carried out at 30 °C using a Luna 3u C8 150×1 mm column, with 100 Å pore size, 3 µm particles (Phenomenex). Eluent A was acetonitrile/methanol/H_2O (19/19/2) +0.2% formic acid +0.028% NH_4OH and eluent B was isopropanol +0.2% formic acid +0.028% NH_4OH. The gradient elution program was 0-5 min, 15% B; 5-35 min, 15-40% B; 35-50 min, 55% B; 51-58 min, 80% B; 59-70 min, 15% B. The flow rate was 40 µL/min and injection volume was 3 µL. Detection was as multiple reaction monitoring (MRM) mode and list of TAG species was established based on previously published analyses of palm and palm kernel oils. MS Analyses (in LC-MS/MS) were achieved in positive mode and nitrogen was used as curtain gas (set to 30), gas1 (set to 25) and gas2 (set to 0). Needle voltage was at +5500 V without needle heating; the declustering potential was set at +40 V. The collision gas was also nitrogen; collision energy was adjusted at +34 V. The dwell time was set to 3 ms. Peak area was quantified using MultiQuant software (v2.1, ABSciex).

Diacylglycerol molecular species analyses Lipid extracts were resuspended in eluent A, and a synthetic internal lipid standard (DAG 17:0/17:0) from Nu-Check Prep was added. Reversed-phase separation was carried out at 30 ℃ on a Phenomenex Luna 3u C8 column (150 mm × 1 mm, 3-μm particle size, 100-Å pore size). Eluent A was acetonitrile/methonal/H_2O (19/19/2, vol/vol/vol) +0.2% formic acid +0.028% NH_4OH, and eluent B was isopropanol +0.2% formic acid +0.028% NH_4OH. The gradient elution program was 0–5 min, 15% (vol/vol) B; 5–35 min, 15–40% B; 35–40 min, 80% B; and 40–55 min, 15% B. The flow rate was 40 μL/min, and 3 μL sample volumes were injected. List of species used for MRM were determined based on oil fatty acid composition in addition to common fatty acids in plants (16:6, 1, 2, 3, 18:0, 1, 2, 3. Analyses (LC-MS/MS, MRM mode) were achieved in positive mode. Nitrogen was used as curtain gas (set to 20), gas1 (set to 25), and gas2 (set to 0). Needle voltage was at +5,500 V without needle heating; the declustering potential was set at +86 V. The collision gas was also nitrogen; collision energy was adjusted to +34 V. Dwell time was set to 3 ms. The relative levels of DAG species were determined by measuring the area under the peak, determined by using MultiQuant software (Version 2.1; ABSciex) and normalizing to the area of the DAG internal standard (Tabet *et al*., 2016).

Phospholipids (PE and PC) molecular species analyses Lipid extracts were re-suspended in 100 μL of Eluent A (isopropanol/methanol/H_2O + 0.2% formic acid + 0.028% NH_4OH). Synthetic internal lipid standards (PE 17:0/17:0 and PC 17:0 /17:0,) from Avanti Polar Lipids (USA) were added at different amounts adapted for the detection. Reversed phase separations were performed at 40 ℃ on a Luna C8 15031 mm column with 100-Å pore size and 5-mm particles (Phenomenex). The gradient elution program was as follows: 0min, 30% eluent B (isopropanol + 0.2% formic acid + 0.028% NH_3); 5min, 50% eluent B; 30 min, 80% eluent B; 31 to 41 min, 95% eluent B; 42 to 52 min, 30% eluent B. The flow rate was set at 40 mL/min, and 3-mL sample volumes were injected. Analyses were achieved in the negative (PE) and positive modes (PC) with fast polarity switching (50 ms); nitrogen was used for the curtain gas (set to 15), gas1 (set to 20) and gas2 (set to 0). Needle voltage was at −4,500 or +5,500

V without needle heating; the de-clustering potential was adjusted between -160 V or set at +40 V. The collision gas was also nitrogen; collision energy varied from −48 eV and +47 eV on a compound-dependent basis. The dwell time was set to 3 ms. The areas of LC peaks were determined using MultiQuant software (v2.1, ABSciex) for phospholipid relative quantification (Grison *et al*., 2015).

2.6.5.2 Analyses carried out in Cadarache by F. Beisson and B. Légeret

These analyses concerned TAG and DAGs only. Lipid molecular species were analyzed on an ultimate RS 3000 UPLC system (Thermo Fisher, Waltham, MA, USA) connected to a quadrupole-time-of-flight (QTOF) 5600 mass spectrometer (AB Sciex, Framingham, MA, USA) equipped with a duo-spray ion source operating in positive mode. Lipid extracts were first separated on a KinetexTM (Kinetex, Atlanta, GA, USA) C18 2.1 × 150 mm 1.7 μm column (Phenomenex, Torrance, CA, USA). Two solvent mixtures, acetonitrile-water (60:40, v/v) and isopropanol-acetonitrile (90:10, v/v), both containing 10 mM ammonium formate at pH 3.2, were used as eluent A and B respectively. The elution was performed with a gradient of 32 min; eluant B was increased from 27 to 97% in 20 min then maintained for 5 min, solvant B was decreased to 7% and then maintained for another 7min for column re-equilibration. The flow-rate was 0.3 mL/min and the column oven temperature was maintained at 45 ℃. Diacylglycerols (DAGs) and triacylglycerols (TAGs) identification was based on retention time, mass accuracy peaks from the MS survey scan compared with theoretical masses and fragment ions from MS/MS scan. Relative quantification was achieved with MultiQuant software (AB Sciex) on the basis of intensity values of $[M+NH4]^+$ adducts of TAGs and DAGs previously identified. TAG and DAG species were noted as follows: lipid class (total number of carbon atoms in fatty acids: total number of double bonds in fatty acids).

2.7 Protein analyses

2.7.1 Crude extract preparation for cell fractionation studies

The mesocarp from fresh fruits was dissected without exocarp, and chopped with razor blade in grinding buffer containing 0.1 M HEPES, pH7.2, 0.33 M sucrose, 5 mM $MgCl_2$, 10 mM KCl, 10 mM ascorbic acid, 5 mM L-cysteine, 1 mM DTT, 0.1% Bovine serum albumin (BSA; fatty acid free), 1% PVPP (insoluble), and protease inhibitor cocktail (EDTA-free). Then, the large debris was filtered out by 200 µm mesh, and transferred to a mortar with new grinding buffer and homogenized gently with pestle. The extract was combined with the first extract, and filtered on the 40 µm mesh. As an option, debris retained by 200 µm mesh were transferred to blender and ground in fresh grinding buffer in a waring blendor set at low speed 3 times for 3 sec each. The filtrates from 40 µm mesh were combined together, and centrifuged at 500 *g* for 10 min at 4℃. The pellet was discarded and the supernatant was used for fractionation by successive differential centrifugation steps. When washing steps were carried out, we used mostly a solution made of 100 mM HEPES, pH 7.2, 1 mM EDTA, and 1 mM DTT, with various sucrose concentrations.

If starting from frozen material (sunflower kernels or palm mesocarp), grinding is performed for 30 sec in liquid nitrogen in a tissue TissueLyser at a frequency of 1/30 sec^{-1}. After filtration, debris and oil pad were removed by two successive centrifugation steps at 500 *g* and 15 000 *g* and membranes were collected from mostly oil-free supernatant by centrifugation at 110 000 *g* for 70 minutes. The oil from fat pad was removed by using equal volume of diethyl ether for washing 10 times. Then, the proteins were precipitated by equal volume of acetone at -20℃ overnight. The concentration of protein was determined using QuantiPro BCA Assay Kit (SigmaAldrich).

2.7.2 DGAT activity assays

DGAT assay was performed in screw-capped glass tube in a final volume of 100 µL. The reaction mixture contained 0.05 M HEPES-KOH, pH 7.2, 5 mM $MgCl_2$, 0.1% BSA (fatty acid free), 0.3 mM *sn*-1,2- diolein or dipalmitin, 25 µM 58.2 mCi/mmol [1-^{14}C] palmitoyl-CoA or/and [1-^{14}C] oleoyl-CoA. The assay was initiated with appropriate microsome or oil body proteins. After incubation for 30 min at room temperature (24-25 ℃), the reaction was terminated by adding 10 µL of 5% SDS (w/v). Then, lipids were extracted from reaction mixtures using the adapted method from Bligh & Dyer (1959). About 0.5-1 mL of chloroform/methanol (1/1, v/v) was added to get a monophasic mixture, and thoroughly mixed. Then, 125-250 µL (1/4 vol) of KCl 0.88% (w/v) was added and the mixture was mixed by vortexing. The chloroform layer was separated by centrifuge and transferred to a fresh tube. The aqueous phase was re-extracted with an equal volume of chloroform. The combined chloroform extract was evaporated and re-suspended in 20 µL of chloroform, which then was applied to silica gel TLC plate using neutral lipid system hexane/diethyl ether/acetic acid (80:20:1, v/v/v) as described above (see 2.6.2). TAG spots were scraped, transferred to scintillation vials together with 2 mL of scintillation liquid. Radioactivity was quantified with a TRI-CARB 2100TR liquid scintillation counter.

Results and discussions

Chapter 3. Lipidomic analysis of oil palm fruits

We characterized lipids, especially TAGs, from oil palm and analyzed both fatty acid composition by GC-FID and molecular species by LC-MS. At first, we analyzed different populations of palm to search for TAG structural diversity. This study is presented as a draft manuscript that has already been submitted for publication. Then, we analyzed TAG, DAG, PC and PE molecular species throughout fruit development, during oil accumulation period of both mesocarp and kernel.

3.1 Variability of TAG composition in various populations of oil palm

Oil was extracted from mesocarp of ripe fruits samples of individual trees, from *Elaeis guineensis* populations presently used for commercial seed production. Fatty acid and TAG molecular species compositions were determined as well as fatty acid occupying *sn*-2 position of TAGs. If palmitic and oleic acids always represent the vast majority of palm oil fatty acids (about 43% palmitic, 43% oleic, 4% stearic and 10% linoleic), we show that variability occurs and palmitic acid content was found to range between 24.5% and 49%. This variability does not affect TAG structure as *sn*-2 position remains the same regardless of oil fatty acid composition: 18:1 and 18:2 fatty acids are the major fatty acids at *sn*-2 position. Specificity of *sn*-2 position for 18 carbon fatty acids of lipids synthesized by the Kennedy pathway in the ER is well known and attributed to specificity of LPAT enzymes towards acyl-CoAs (Sambanthamurthi *et al*., 2000). Also, major molecular species remain the same, largest variations concerning mostly those two species that contain one or two palmitic acid and two or one oleic acid (50:1 and 52:2), in addition to species with three palmitic acids (48:0) or three oleic acids (54:3). Major molecular species (above 10%) are: 48:0, 50:1, 50:2, 52:2, 52:3.

Major impact of study lies in possible breeding application: palm oil with less saturated fat is more fluid and preferred by the African consumer. In addition, such an oil might prove healthier if saturated fat in diet promotes indeed cardiovascular diseases, which is still controversial. Before this study, lowest reported level for palmitic acid in palm oil was 32%.

Short Communication

Oil-palm trees (*Elaeis guineensis*) that produce oil with substantially lower saturated fatty acids content

Yijun Yuan[a,b], Godswill Ntsomboh-Ntsefong[b,e,f], Marina Le Guédard[a,b,d], Joseph-Martin Bell[e], Emmanuel Youmbi[e,g], Hermine Bille-Ngalle[e], René Lessire[a,b], Georges-Frank Ngando-Ebongue[f], Vincent Arondel[*,a,b], and Laetitia Fouillen[*,a,b,c]

[a] *Univ. Bordeaux, Membrane Biogenesis Laboratory, UMR5200, 33140 Villenave d'Ornon, France*

[b] *CNRS, Membrane Biogenesis Laboratory, UMR5200, 33140 Villenave d'Ornon, France*

[c] *MetaboHUB - Metabolome Facility of Bordeaux, Functional Genomics Center of Bordeaux, 33140 Villenave d'Ornon, France*

[d] *LEB Aquitaine Transfert-ADERA, 33140 Villenave d'Ornon, France*

[e] *University of Yaounde I, Faculty of Science, Department of Plant Biology, PO Box 812 Yaounde, Cameroon*

[f] *Institute of Agricultural Research for Development, Specialized Centre for Oil Palm Research of La Dibamba, BP 243 Douala, Cameroon*

[g] *Tissue Culture Laboratory, African Centre for Research on Banana and Plantain (CARBAP), Njombe, Cameroon*

***Correspondence:** Vincent Arondel and Laetitia Fouillen, Membrane Biogenesis Laboratory, UMR5200, 71 Avenue Edouard Bourlaux, 33140 Villenave d'Ornon, France; **E-mail:** vincent.arondel@u-bordeaux.fr, laetitia.fouillen@u-bordeaux.fr

Running Title: Oil Palms with lower saturated oil

Author contributions: YY: lipid extraction, discussing results, and writing paper; GN: discussing results and writing paper; MLG: statistical treatment of data; JMB, EY, and HBG: mentoring GN; RL: conception of study and discussing results; GFNE: conception of study, proving fruit samples; VA: conception of study, data interpretation and writing paper; LF: conception of study, lipid analysis and MS spectra interpretation, writing paper.

Abstract

We have analyzed individual elite trees, from *Elaeis guineensis* populations presently used for commercial seed production, for oil composition in mesocarp fruit and found that some trees yield oil with as little as 24.5% palmitic acid versus usual 43%. These are the lowest values for palmitic content reported up to now in oil from *E. guineensis*, not far from 16-19% values recorded in *E. oleifera*, the American oil palm. Molecular species analysis and determination of fatty acid at *sn*-2 position of main triacylglycerol species indicated that variation of palmitic acid is not accompanied by major changes in triacylglycerol structure and species composition. These trees will provide a basis for breeding rapidly new palm varieties producing crude oil with significantly reduced saturated fat. Oil will be of higher economical value, preferred by consumer, especially in Africa, and likely to prove healthier.

Keywords: *Elaeis guineensis,* triacylglycerol composition, palmitic acid variability, oil palm quality

Abbreviations: TAG, triacylglycerol ; **PCA,** principal component analysis

1. Introduction

Crude palm oil (Corley & Tinker, 2003, Corley, 1979) contains a high level of saturated fat (about 45%, including 43% palmitic acid). As a result, about 40% of the fat is solid and sediments at the bottom, which is not appealing to the consumer. In South East Asia, refined palm oil is cold-fractionated (Kellens *et al*., 2007) into a liquid (olein, less saturated) and a solid fraction of lesser economical value (stearin, more saturated), is the latter being used mostly as shortening by food industry. The olein fraction, widely used as domestic oil in Southeast Asia, represents about 60% of initial volume of the crude oil it comes from (Zhang *et al*., 2013) and still contains a significant amount of saturated fat (about 35-42% saturated fatty acids, including 31-40% palmitic acid). In equatorial Africa, crude, unrefined palm oil constitutes the majority of palm oil available to consumers and is the main dietary source of vitamins A and E. It

is used in traditional cooking recipes and is preferred to imported refined olein (Cheyns, 2001), despite high level of crystallized fat. It has been reported (Poku, 2002) that the main reason small-scale planters were reluctant to adopt new high-producing varieties in the sixties was the higher level of solid fat in oil when compared to that extracted from traditional, low producing palms. It is clear that crude oil with lower solid fat would gain wide consumer preference in Africa. Equatorial Africa itself hosts presently a 400 million consumer-market which is poised to increase in the same trend as both population and *per capita* vegetable oil consumption increases.

Health impact of high saturated fat diet remains controversial, with most recent studies (de Souza *et al*., 2015) finding very little or no association between dietary intake of saturated fat and death due to heart disease and diabetes. However, it is still suggested that reducing saturated fat intake lowers the risk of developing coronary heart. More worrisome is a study linking increased cardiovascular diseases to increased palm oil consumption in developing countries (Chen *et al*., 2011). WHO forecasts predict an important increase in the incidence of cardiovascular diseases and obesity in developing countries for the next 20 years, especially in Africa (Mensah, 2008). Thus, lowering saturated fat in palm oil is an action that might prove beneficial for health reasons to those developing countries using exclusively palm oil for domestic purposes. Presently, attempts to lower saturated fat in oil palm are based on using the parent species *E. oleifera*, which produces oil with about 22% palmitic acid (16:0) and a total of saturated fatty acids below 25% (Barcelos *et al*., 2015). Because *E. oleifera* yields very little oil, the only material presently planted (mostly in South America) with lower saturated fat is an interspecific hybrid. It contains about 27-41% palmitic acid, 1-6% stearic acid and 43-49% oleic acid (Corley & Tinker, 2003) which is comparable to the composition of Malaysian oleins. However, this hybrid requires assisted pollination in order to provide acceptable yields, which significantly increases production costs. Here, we adopted another approach, that is, searching for fatty acid variability in oil produced by individual *E. guineensis* elite tree already used to produce commercial seeds in Cameroon, and were able to identify trees producing oil with as little as 24.5% palmitic

acid.

2. Materials and methods

2.1. Plant materials

All samples analyzed in this study were collected on the same day at CEREPAH La Dibamba, Cameroon, from trees that were at least 7 years-old. Fruits were from three populations of *Elaeis guineensis:* La Mé, Deli, and Deli x La Mé cross. Several fruits were collected from a single ripe bunch. After harvest, the fruits were frozen in liquid nitrogen and stored at -80 ℃. Mesocarp was fully and quickly dissected from individual frozen fruit and ground in liquid nitrogen to a fine powder. Biological replicates (at least 3) were individual fruits coming from the same bunch.

2.2. Extraction of lipids

About 50mg of each frozen mesocarp powder was immediately dipped for 15 min into 8 mL glass tubes with Teflon-lined caps containing 2 mL boiling isopropanol to inactivate mesocarp lipase (Ngando-Ebongue *et al*., 2008). Then, 3 mL of hexane were added and total lipids were extracted according to (Radin, 1981).

2.3. Fatty acid analyses

Analysis and quantification of fatty acids was performed by gas chromatography as in Ngando *et al.* (Ngando-Ebongue *et al*., 2008). To determine free fatty acid content, an appropriate amount of internal standard (heptadecaenoic acid = 17:0) was added to 1mg of oil. An aliquot was kept to quantify total fatty acids and the rest of the mixture was extracted according to Belfrage and Vaughan's procedure (Belfrage & Vaughan, 1969) and free fatty acids were quantified from the upper phase.

2.4. Triacylglycerol molecular species analyses

Analysis and quantification of triacylglycerol molecular species was realized by LC-MS/MS (MRM mode) analyses (Supporting Information). Peak area was quantified using MultiQuant software (v2.1, ABSciex). For the *regio*-localization was

performed by ESI-MS/MS and ESI-MS/MS/MS analyses using a QTRAP® 5500 (ABSciex) mass spectrometer. Settings were as described in the Supporting Information.

2.5. Statistical analyses

Data were submitted to statistical analysis using 2013 version of Xlstat including principal component analysis (PCA), using Pearson's correlation coefficient analysis.

3. Results and discussion

3.1. Fatty acid and TAG molecular species content and overall variability

We have determined the fatty acid composition of oil from material widely planted in Africa, resulting from the cross of trees from Deli and La Mé populations (Fig.1), and that from their parent populations. Palmitic (16:0) and oleic (18:1) fatty acids represent 80-88% of total fatty acids. Depending on the biological material, levels of 16:0 and 18:1 vary widely, from 28 to 49% and 32 to 56%, respectively. Relative environment impact on fatty acid composition variability should be limited, as fruits were collected the same day on the same site.

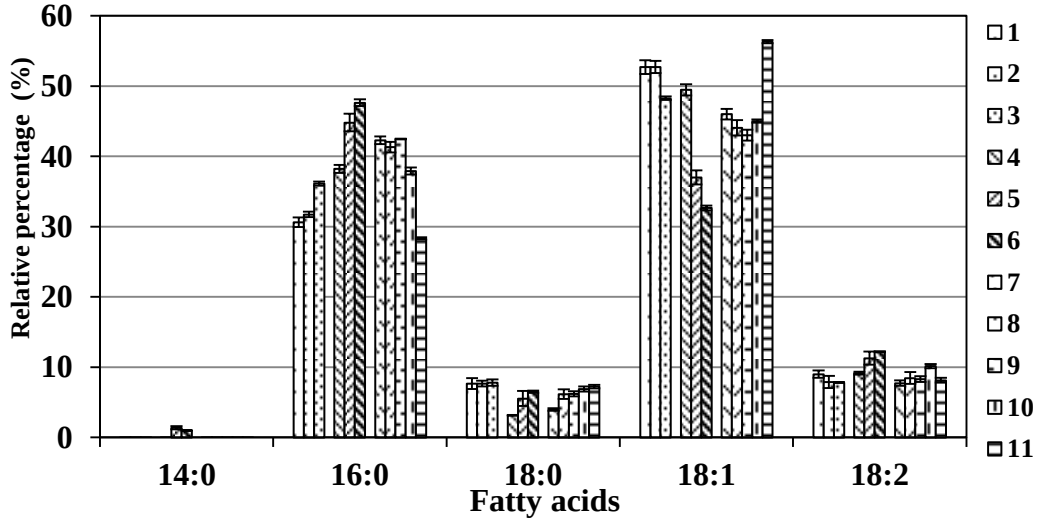

Fig.1 Fatty acid composition of mesocarp oil from La Mé (1-3), Deli (4-6) and Deli x La Mé (7-11) individual trees. Data are means of three biological replicates +/- standard error. 14:0, myristic acid; 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid. Relative composition is by mass.

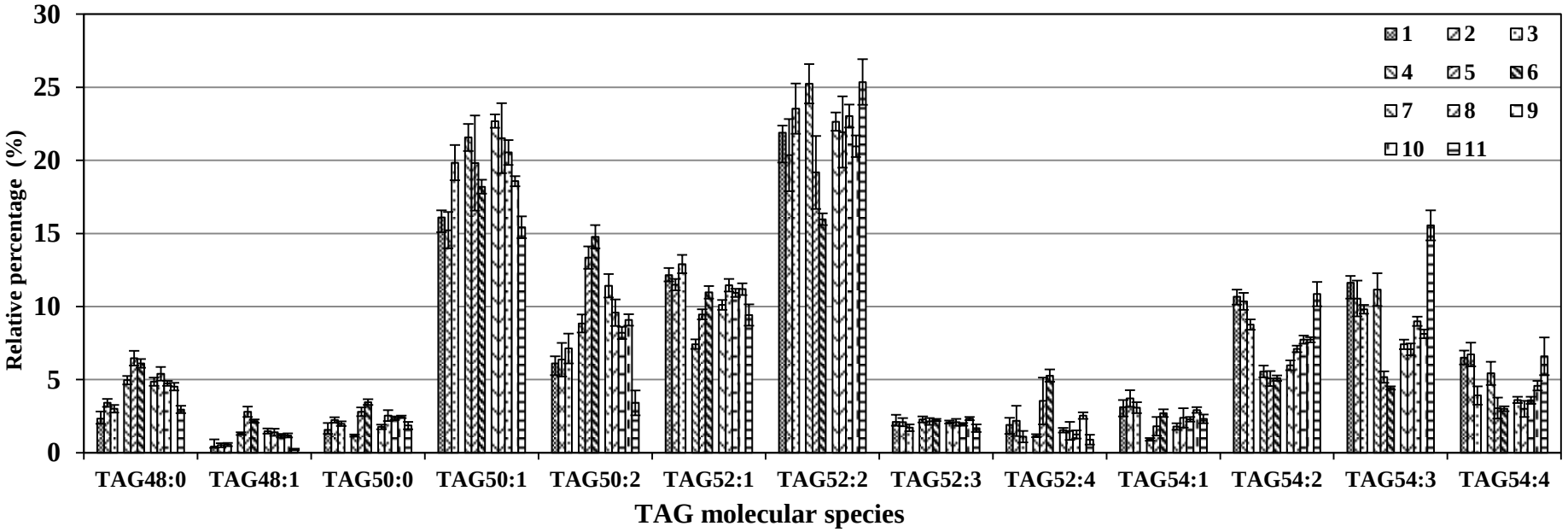

Fig.2 Main triacylglycerol molecular species of mesocarp oil from La Mé (1-3), Deli (4-6) and Deli x La Mé (7-11) individual trees. TAG species are designated by total number of fatty acid carbon and total number of double bonds, separated by a colon. Full list is in Table S2.

It has been shown previously that only little variations of fatty acid composition were observed between bunches from same tree and fruits from same bunch (Noiret & Wuidart, 1976). With regards to this last point, our data show indeed low standard deviation (1-3% on average for both 16:0 and 18:1 levels). Thus, it seems reasonable to presume that our fruit-based data reflect indeed the composition of oil from a given tree.

We have also analyzed triacylglycerol molecular species composition. Boiling mesocarp in isopropanol allowed us to inactivate endogenous lipase (Morcillo *et al*., 2013) and to limit free fatty acid content to less than 1% in most cases (Supporting Information Table S1). Main molecular species of TAGs are shown in Fig.2 (full results in Table S2). The two major species (about 20% each) are 50:1 and 52:2, which most abundant forms contain two 16:0, one 18:1 and one 16:0, two 18:1, respectively. Then, six TAG species account each for 4-10% of total; these are species containing two major fatty acids (16:0, 18:1) and one minor one (18:0, 18:2) and species with three 16:0 or three 18:1. Other species represent each on average less than 3% of total. These results are consistent with other published molecular species compositions of TAGs from *E. guineensis* (Che Man *et al*., 1999, Sambanthamurthi *et al*., 2000) and only a few differences can be noted concerning 52:3 and 54:2. Molecular species variability appears to be more important than that of fatty acids, but this might not only be due to genetic factors as fruit-to-fruit variation seems also to be higher (standard deviation about 6-10%).

For main fatty acids, abundance is best correlated to abundance of the molecular species that contains exclusively the considered fatty acid: TAG48:0 for 16:0 and TAG54:3 for 18:1, and also by abundance of those species with two of the considered fatty acid out of three (Fig.3A). The situation is less obvious for minor fatty acids, although 18:2 correlates with 54:4 and 18:0 with 54:1. Variations in palmitic acid content are not linked large changes in molecular species composition.

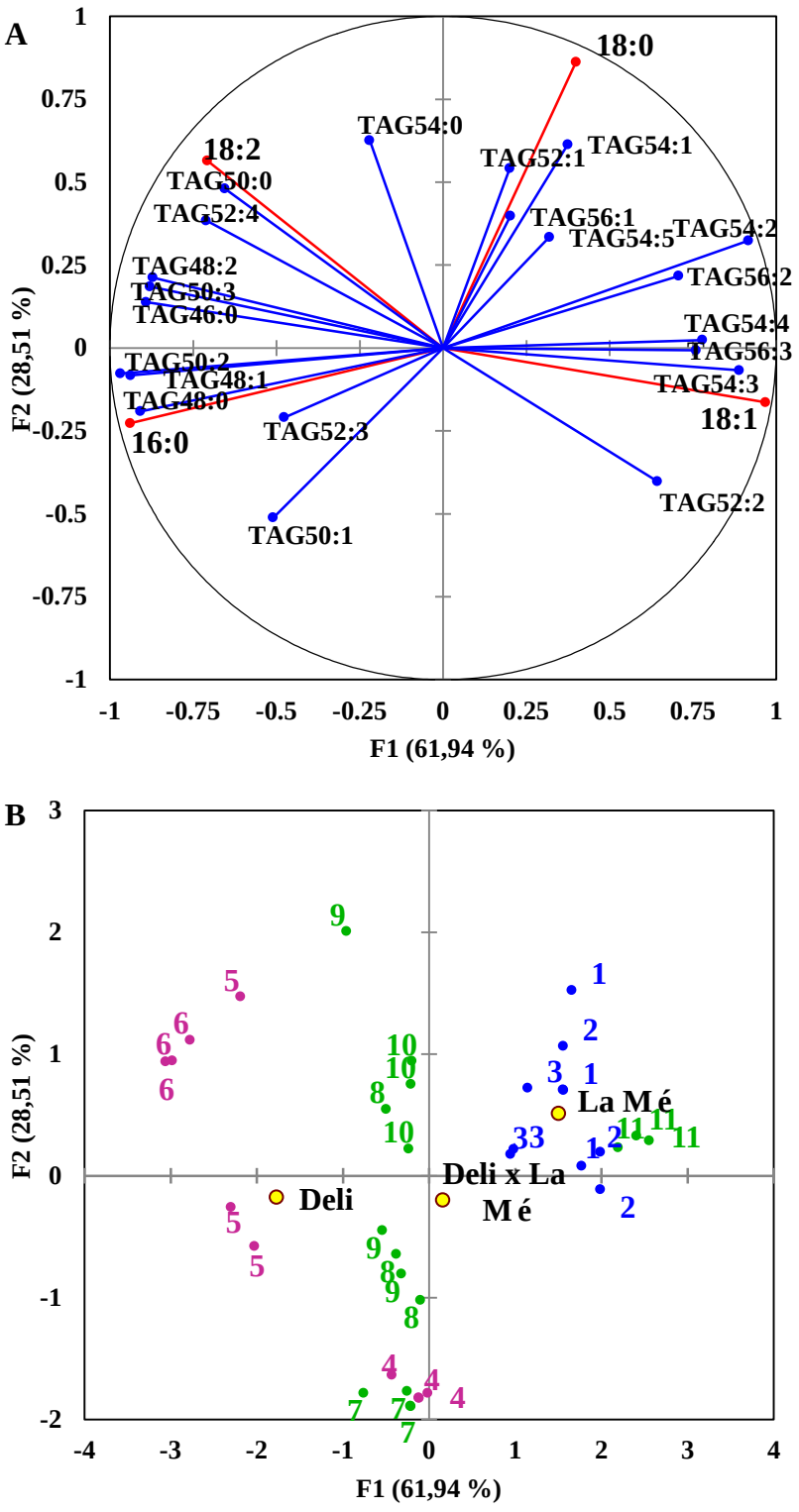

Fig.3 Principal component analysis based on fatty acid composition of mesocarp oil from La Mé (1-3), Deli (4-6) and Deli x La Mé (7-11) individual trees. PCA was built using fatty acids as active variables (vectors in red) using Xlstat software, based on Pearson's correlations. Triacylglycerol molecular species are projected as supplementary variables (vectors in blue). A: Correlation circle of variables calculated in the two first principal components. B: Scores plot of component 1 versus 2. Yellow dots: barycenter for each of the three populations.

We have determined the most abundant fatty acid at position *sn*-2 for each of the 11 major TAG species (Table 1). Each time a species contains either 18:1 or 18:2 fatty acid, one of these two fatty acids is at *sn*-2 position. When species contain both fatty acids, then *sn*-2 position is occupied sometimes by 18:1 (54:4), sometimes by 18:2 (52:3). The result was identical for the 11 different biological samples that were analyzed, suggesting that the variability of total fatty acid composition observed in these samples is not much related to acylation at position 2. These results are consistent with determination of fatty acid position carried out by other groups on whole oil rather than TAG species, showing that 18:1 represents about 65% of fatty acids present at *sn*-2 position (Mozzon *et al*., 2013, Sambanthamurthi *et al*., 2000).

Table 1 Fatty Acid at *sn*-2 Position of TAG.

Major TAGs	**TAG 50:0**	**TAG 50:1**	**TAG 50:2**	**TAG 52:1**	**TAG 52:2**	**TAG 52:3**	**TAG 52:4**	**TAG 54:1**	**TAG 54:2**	**TAG 54:3**	**TAG 54:4**
All fatty acids	16:0 16:0 18:0	16:0 16:0 18:1	16:0 16:0 18:2	16:0 18:0 18:1	16:0 18:1 18:1	16:0 18:1 18:2	16:0 18:2 18:2	18:0 18:0 18:1	18:0 18:1 18:1	18:1 18:1 18:1	18:1 18:1 18:2
Fatty acid at *sn*-2	***18:0***	***18:1***	***18:2***	***18:1***	***18:1***	***18:2***	***18:2***	***18:1***	***18:1***	***18:1***	***18:1***

Fatty acid composition of select molecular species are shown and major fatty acid at position *sn*-2 (italicized on table) was identified by ESI MS/MS/MS. TAG species are designated with total number of fatty acid carbon and total number of double bonds, separated by a colon. 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid. One replicate for each of the 11 samples was analyzed and results were identical for all samples.

PCA (Fig.3) indicates that 16:0 and 18:1 vary mostly according to first axis and are strongly negatively correlated (-0.92), while 18:0 varies mostly along the second axis and 18:2 along both axis 1 and 2. In addition, 18:2 correlates negatively with 18:1 (-0.696) and 18:0 correlates negatively with 16:0 (-0.49). These results are in agreement with those from Montoya *et al*. (2014) and can be commented on with regards to possible underlying mechanisms responsible for palm oil fatty acid composition. As already mentioned in Sambanthamurthi *et al*. (2000) and others, key steps are elongation of 16:0-ACP into 18:0-ACP by KASII-type enzyme, desaturation of 18:0-ACP to 18:1-ACP by stearoyl-ACP desaturases and cleavage of acyl-ACP

into free fatty acids by thioesterases. Two thioesterases of *Arabidopsis* (Salas & Ohlrogge, 2002) show distinct properties, FatA preferring unsaturated substrates (18:1-ACP) and FatB saturated substrates (16:0-ACP and 18:0-ACP). Because 18:0 does not correlate negatively with 18:1, it is unlikely that stearoyl-ACP desaturation step is a key determinant of variation of palm oil composition. Negative correlation between 16:0 and 18:0 might be mostly linked to the KASII-catalyzed elongation step, and the strong negative correlation between 16:0 and 18:1 to the thioesterase-catalyzed step. These three steps are likely to control relative amounts of 16:0, 18:0 and 18:1 leaving the plastid. Only 18:1 can be subsequently modified by the FAD2 desaturase to yield 18:2; therefore, the negative correlation between 18:1 and 18:2 is likely related to this desaturation step. Fatty acid assembly into TAGs might also impact final oil composition: mutations in acyl-editing enzymes lead to a strong reduction of polyunsaturated levels of *Arabidopsis* seeds (Bates *et al*., 2012). However, palm oil composition contains very little PUFAs, and our results show no changes in TAG structure and no large change in TAG species, which suggests that editing mechanisms are unlikely to show much impact on overall fatty acid unsaturation level (Bourgis *et al*., 2011, Tranbarger, *et al*., 2011). Genetic studies (Montoya *et al*., 2014, Ting *et al*., 2016) showed the existence of several loci controlling fatty acid composition in oil palm, and hypothetical candidate genes could be identified at these loci, including a thioesterase (Ting *et al*., 2016).

3.2. Variability of fatty acid composition between and within populations

Deli fruits show higher levels of 16:0 and lower levels of 18:1 than those of La Mé, while the Deli x La Mé population shows, on average, intermediate levels of both fatty acids. These data are in agreement with previous reports (Montoya *et al*., 2014, Noiret & Wuidart, 1976). PCA shows (Fig.3B) that most of the difference between the two populations, represented by their respective barycenter, is explained by the first and main component of variability (axis 1, about 62% of total variation), with Deli x La Mé population lying in between. In contrast, intra-population variability is explained mostly by the second axis (axis 2, about 28% of total variation), except for

two samples which group with another population than the one they belong to (Fig. 3B). Intra-population variability has already been reported for La Mé population but not for Deli (Noiret & Wuidart, 1976). Deli variability with regards to oil fatty acid composition might be explained by introgressions carried out after the Noiret and Wuidart study (Noiret & Wuidart, 1976). Fruits from trees that originate from Deli x La Mé crosses, that are commercial planting material, show a wide variability (28-42% 16:0, 44-56% 18:1) that extends at the lower limit for 16:0 level and which is comparable to that noted for *E. oleifera* x *E. guineensis* hybrids (27-41% 16:0, 43-59% 18:1) (Corley & Tinker, 2003).

Because of the limited number of samples analyzed, we could not exclude the fact that we may have, by chance, singled out exceptional trees with low saturated fatty acids. Therefore, we carried out a fast screen of an additional 22 trees representing most genotypes grown at CEREPAH; these fruit samples were harvested the same day as were the ones described above. The result (Table 2) indicated that trees with as little as 24% 16:0 in oil were not uncommon. One cannot exclude an environment impact on saturation level and this is the reason why we included two trees from *E. oleifera* in our screen. The latter still show lower levels of 16:0 (16-19% 16:0, Table 2), but the difference between both species appears to be much lower than initially thought.

Table 2 Fatty Acid Composition of Oil Extracted from Fruit Mesocarp of Different Trees Grown at CEREPAH.

Number	POPULATION	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0
12	Deli x La Mé	0.58	34.92	0.15	5.76	47.30	10.60	0.36	0.33
13	Deli x La Mé	0.39	40.10	0.17	4.98	46.31	7.45	0.27	0.32
14	Deli x La Mé	0.34	40.28	0.07	9.75	41.02	7.91	0.22	0.40
15	Deli x La Mé	0.73	31.62	0.00	5.33	45.51	15.45	0.45	0.91
16	Deli x La Mé	0.68	33.24	0.11	5.05	48.82	11.32	0.30	0.48
17	Deli x La Mé	0.49	33.95	0.17	4.64	48.49	11.73	0.24	0.31
18	Deli x La Mé	0.36	34.64	0.13	5.45	44.58	14.06	0.34	0.44
19	Deli x La Mé	0.14	24.50	0.10	6.63	58.16	9.58	0.41	0.49
20	La Mé	0.22	29.52	0.05	9.20	48.76	11.26	0.36	0.63
21	La Mé	0.18	25.56	0.05	8.44	51.99	12.88	0.36	0.55

22	La M é	0.15	24.60	0.06	8.22	53.01	13.05	0.34	0.56
23	La M é	0.34	36.51	0.08	7.66	46.27	8.43	0.24	0.47
24	Deli	0.97	36.13	0.24	3.48	44.80	13.62	0.43	0.33
25	Deli	1.10	40.48	0.17	4.53	42.29	10.65	0.36	0.42
26	*E. oleifera*	0.11	16.13	1.30	1.34	61.80	18.20	0.76	0.36
27	*E. oleifera*	0.12	18.85	1.23	1.39	61.97	15.57	0.69	0.18
28	Others	0.65	36.44	0.21	3.56	44.45	14.08	0.29	0.33
29	Others	0.79	44.22	0.29	3.71	37.70	12.59	0.29	0.42
30	Others	1.14	37.94	0.23	3.81	42.92	13.26	0.30	0.40
31	Others	1.05	36.62	0.14	4.95	46.12	10.41	0.36	0.34
32	Others	0.74	34.00	0.29	3.68	43.99	16.48	0.49	0.33
33	Others	0.42	37.72	0.99	2.60	34.93	22.50	0.44	0.40

Oil was extracted from the mesocarp of 3-5 fruits from the same bunch and an aliquot was methylated. Resulting methyl esters were analyzed by gas chromatography coupled to a flame ionization detector. Fruits were collected the same day. 14:0, myristic acid; 16:0, palmitic acid; 16:1, palmitoleic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid; 20:0, arachidic acid. Values are in percent by mass.

Taken together, these data strongly suggest that there is enough variability within *E. guineensis* to produce elite hybrids with an oil quality, with respect to fatty acid composition comparable to, or even better than that produced by best interspecific (*oleifera* x *guineensis*) hybrids, without the need for assisted pollination required by the interspecific hybrid. In addition, because the desirable trait is already present in some elite planting material, it should be much faster to produce commercial seeds than introgressing the trait from *E. oleifera* into a high yielding palm population. For this purpose, we are currently starting to determine fatty acid composition and TAG molecular species of all trees grown at the CEREPAH to identify as many appropriate genitors as possible.

References

Barcelos, E., Rios Sde, A., Cunha, R. N., Lopes, R., Motoike, S. Y., Babiychuk, E., Skirycz, A., & Kushnir, S. (2015). Oil palm natural diversity and the potential for yield improvement. *Frontiers in plant science, 6*, 190.

Bates, P. D., Fatihi, A., Snapp, A. R., Carlsson, A. S., Browse, J., & Lu, C. (2012). Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. *Plant physiology, 160*(3), 1530-1539.

Belfrage, P., & Vaughan, M. (1969). Simple liquid-liquid partition system for isolation of labeled oleic acid from mixtures with glycerides. *J Lipid Res, 10*(3), 341-344.

Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G. F., Drira, N., Ohlrogge, J. B., & Arondel, V. (2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc Natl Acad Sci U S A, 108*(30), 12527-12532.

Che Man, Y. B., Haryati, T., Ghazali, H. M., & Asbi, B. A. (1999). Composition and thermal profile of crude palm oil and its products. *Journal of the American Oil Chemists' Society, 76*(2), 237-242.

Chen, B. K., Seligman, B., Farquhar, J. W., & Goldhaber-Fiebert, J. D. (2011). Multi-Country analysis of palm oil consumption and cardiovascular disease mortality for countries at different stages of economic development: 1980-1997. *Global Health, 7*, 45.

Cheyns, E. (2001). La consommation urbaine de l'huile de palme rouge en Côte d'Ivoire : quels marchés ? *OCL, 8*(6), 641-645.

Corley, R. H. V. (1979). Palm oil composition and oil palm breeding. *Planter, Kuala Lumpur, 55*, 467-478.

Corley, R. H. V., & Tinker, P. B. (2003). *The oil palm* (4th ed.). Oxford ; Malden, MA: Blackwell Science.

de Souza, R. J., Mente, A., Maroleanu, A., Cozma, A. I., Ha, V., Kishibe, T., Uleryk, E., Budylowski, P., Schunemann, H., Beyene, J., & Anand, S. S. (2015). Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies. *BMJ, 351*, h3978.

Kellens, M., Gibon, V., Hendrix, M., & De Greyt, W. (2007). Palm oil fractionation. *European Journal of Lipid Science and Technology, 109*(4), 336-349.

Mensah, G. A. (2008). Ischaemic heart disease in Africa. *Heart, 94*(7), 836-843.

Montoya, C., Cochard, B., Flori, A., Cros, D., Lopes, R., Cuellar, T., Espeout, S., Syaputra, I., Villeneuve, P., Pina, M., Ritter, E., Leroy, T., & Billotte, N. (2014). Genetic architecture of palm oil fatty acid composition in cultivated oil palm (Elaeis guineensis Jacq.) compared to its wild relative E. oleifera (H.B.K) Cortes. *PLoS One, 9*(5), e95412.

Morcillo, F., Cros, D., Billotte, N., Ngando-Ebongue, G. F., Domonhedo, H., Pizot, M., Cuellar, T., Espeout, S., Dhouib, R., Bourgis, F., Claverol, S., Tranbarger, T. J., Nouy, B., & Arondel, V. (2013). Improving palm oil quality through identification and mapping of the lipase gene causing oil deterioration. *Nature communications, 4*, 2160.

Mozzon, M., Pacetti, D., Lucci, P., Balzano, M., & Frega, N. G. (2013). Crude palm oil from interspecific hybrid Elaeis oleiferaxElaeis guineensis: fatty acid regiodistribution and molecular species of glycerides. *Food Chem, 141*(1), 245-252.

Ngando-Ebongue, G. F., Koona, P., Nouy, B., Zok, S., Carrière, F., Zollo, P.-H. A., & Arondel, V. (2008). Identification of oil palm breeding lines producing oils with low acid values. *European Journal of Lipid Science and Technology, 110*(6), 505-509.

Noiret, J. M., & Wuidart, W. (1976). Possibilities for improving the fatty acid composition of palm oil. Results and prospects. *Oleagineux, 31*(11), 465-474.

Poku, K. (2002). Small-scale Palm oil processing in Africa. *FAO agricultural services bulletin, 148*.

Radin, N. S. (1981). Extraction of tissue lipids with a solvent of low toxicity. *Methods Enzymol, 72*, 5-7.

Salas, J. J., & Ohlrogge, J. B. (2002). Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Archives of biochemistry and biophysics, 403*(1), 25-34.

Sambanthamurthi, R., Sundram, K., & Tan, Y. (2000). Chemistry and biochemistry of palm oil. *Prog Lipid Res, 39*(6), 507-558.

Ting, N. C., Yaakub, Z., Kamaruddin, K., Mayes, S., Massawe, F., Sambanthamurthi, R., Jansen, J., Low, L. E., Ithnin, M., Kushairi, A., Arulandoo, X., Rosli, R., Chan, K. L., Amiruddin, N., Sritharan, K., Lim, C. C., Nookiah, R., Amiruddin, M. D., & Singh, R. (2016). Fine-mapping and cross-validation of QTLs linked to fatty acid composition in multiple independent interspecific crosses of oil palm. *BMC Genomics, 17*(1), 289.

Tranbarger, T. J., Dussert, S., Joet, T., Argout, X., Summo, M., Champion, A., Cros, D., Omore, A., Nouy, B., & Morcillo, F. (2011). Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, and functional specialization in lipid and carotenoid metabolism. *Plant physiology, 156*(2), 564-584.

Zhang, X., Li, L., Xie, H., Liang, Z., Su, J., Liu, G., & Li, B. (2013). Comparative analysis of thermal behavior, isothermal crystallization kinetics and polymorphism of palm oil fractions. *Molecules, 18*(1), 1036-1052.

Supplementary data

Table S1. The percentage of free fatty acids (FFA) in mesocarp oil from La Mé (1-3), Deli (4-6) and Deli x La Mé (7-11) individual trees.

Sample	FFA (%)
1	0.42 ±0.13
2	0.47 ±0.01
3	0.38 ±0.08
4	0.89 ±0.41
5	1.92 ±0.63
6	0.66 ±0.16
7	0.67 ±0.13
8	0.93 ±0.51
9	0.80 ±0.13
10	0.45 ±0.11
11	0.44 ±0.06

Table S2. Triacylglycerol molecular species of mesocarp oil from La Mé(1-3), Deli (4-6) and Deli x La Mé(7-11) individual trees.

TAG	1	2	3	4	5	6	7	8	9	10	11
TAG42:0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
TAG44:0	n.d.	n.d.	n.d.	n.d.	0.03±0.00	0.04±0.02	0.01±0.01	0.01±0.00	n.d.	n.d.	n.d.
TAG44:1	n.d.	n.d.	n.d.	n.d.	n.d.	0.01±0.00	0.02±0.02	n.d.	n.d.	n.d.	n.d.
TAG46:0	0.06±0.02	0.09±0.04	0.08±0.02	0.27±0.02	0.97±0.05	1.14±0.31	0.26±0.04	0.32±0.10	0.27±0.04	0.28±0.05	0.03±0.01
TAG46:1	0.01±0.01	0.02±0.01	0.01±0.01	0.02±0.00	0.08±0.01	0.09±0.05	0.04±0.03	0.03±0.01	0.02±0.01	0.02±0.00	n.d
TAG46:2	n.d.	n.d.	n.d.	n.d.	0.02±0.00	0.02±0.01	0.01±0.00	n.d.	n.d.	n.d.	n.d.
TAG46:3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
TAG48:0	2.34±0.36	3.41±0.27	3.01±0.27	4.96±0.29	6.11±0.30	6.47±0.51	4.85±0.28	5.41±0.47	4.75±0.17	4.52±0.26	2.96±0.26
TAG48:1	0.43±0.10	0.51±0.13	0.56±0.11	1.31±0.09	2.17±0.11	2.80±0.36	1.48±0.17	1.39±0.24	1.17±0.12	1.18±0.14	0.22±0.05
TAG48:2	0.05±0.02	0.07±0.03	0.06±0.02	0.13±0.01	0.83±0.09	0.72±0.38	0.17±0.03	0.18±0.07	0.14±0.03	0.18±0.02	0.02±0.01
TAG48:3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
TAG50:0	1.56±0.27	2.24±0.19	1.98±0.16	1.18±0.08	3.46±0.19	2.80±0.30	1.76±0.16	2.53±0.37	2.37±0.12	2.45±0.09	1.86±0.25
TAG50:1	16.09±1.01	15.22±1.24	19.84±1.20	21.57±0.93	18.20±0.47	19.81±3.25	22.69±0.46	21.52±2.39	20.54±0.84	18.58±0.34	15.42±0.74
TAG50:2	6.12±0.80	6.37±1.16	7.13±1.03	8.85±0.62	14.78±0.80	13.34±0.77	11.42±0.80	9.58±0.91	8.20±0.42	9.09±0.39	3.41±0.85
TAG50:3	0.08±0.03	0.09±0.04	0.06±0.02	0.14±0.01	0.46±0.04	0.48±0.22	0.14±0.02	0.16±0.06	0.12±0.03	0.21±0.02	0.04±0.01
TAG50:4	0.01±0.00	0.01±0.00	n.d.	0.01±0.01	0.06±0.02	0.05±0.00	0.01±0.00	0.01±0.00	0.01±0.00	0.02±0.00	n.d.
TAG52:0	0.35±0.13	0.55±0.17	0.46±0.08	0.57±0.03	0.92±0.08	0.82±0.10	0.24±0.04	0.51±0.20	0.89±0.03	0.53±0.05	0.75±0.06
TAG52:1	12.15±0.43	11.49±0.39	12.91±0.63	7.43±0.32	10.98±0.43	9.47±0.35	10.12±0.34	11.46±0.43	10.94±0.29	11.19±0.40	9.43±0.73
TAG52:2	21.89±2.04	20.36±2.47	23.54±1.72	25.24±1.35	15.97±0.4	19.17±2.5	22.64±0.63	21.94±2.44	23.03±0.79	20.97±0.74	25.36±1.56
TAG52:3	2.12±0.28	2.09±0.28	1.71±0.24	2.28±0.20	2.24±0.08	2.14±0.24	2.10±0.11	2.09±0.23	1.94±0.10	2.34±0.11	1.68±0.26
TAG52:4	1.90±0.61	2.18±1.03	1.09±0.39	1.17±0.11	5.28±0.42	3.54±1.59	1.54±0.16	1.49±0.61	1.24±0.28	2.54±0.23	0.90±0.34
TAG54:0	0.04±0.02	0.09±0.04	0.07±0.01	0.03±0.00	0.10±0.01	0.07±0.02	0.03±0.01	0.06±0.02	0.06±0.01	0.07±0.01	0.06±0.01
TAG54:1	3.12±0.63	3.72±0.57	3.08±0.37	0.91±0.10	2.71±0.26	1.81±0.64	1.79±0.22	2.38±0.66	2.28±0.16	2.92±0.19	2.32±0.31
TAG54:2	10.67±0.54	10.36±0.58	8.76±0.36	5.56±0.41	5.11±0.19	5.08±0.51	5.98±0.33	7.10±0.23	7.74±0.28	7.73±0.19	10.85±0.83

TAG54:3	11.62±1.08	10.54±1.22	9.81±0.30	11.16±1.12	4.42±0.12	5.18±0.39	7.41±0.33	7.08±0.41	8.99±0.32	8.13±0.30	15.55±1.03
TAG54:4	6.51±0.48	6.72±0.82	3.91±0.63	5.43±0.79	3.01±0.16	3.05±0.72	3.61±0.22	3.00±0.55	3.59±0.24	4.59±0.33	6.60±1.29
TAG54:5	1.16±0.36	1.39±0.68	0.44±0.15	0.65±0.06	0.84±0.11	0.70±0.37	0.47±0.06	0.32±0.11	0.41±0.05	0.88±0.06	0.80±0.30
TAG54:6	0.04±0.01	0.04±0.02	0.02±0.00	0.03±0.01	0.06±0.01	0.03±0.02	0.02±0.00	0.01±0.00	0.02±0.00	0.03±0.00	0.03±0.01
TAG56:0	0.01±0.01	0.03±0.02	0.03±0.00	0.04±0.00	0.04±0.00	0.04±0.01	0.03±0.00	0.03±0.01	0.02±0.00	0.03±0.00	0.03±0.00
TAG56:1	0.14±0.07	0.31±0.17	0.20±0.05	0.09±0.01	0.21±0.03	0.15±0.06	0.12±0.01	0.17±0.06	0.16±0.01	0.21±0.01	0.18±0.04
TAG56:2	1.00±0.33	1.46±0.32	0.81±0.17	0.50±0.07	0.55±0.06	0.58±0.24	0.65±0.05	0.77±0.24	0.71±0.07	0.86±0.08	1.01±0.16
TAG56:3	0.40±0.06	0.48±0.04	0.30±0.02	0.35±0.04	0.25±0.01	0.28±0.07	0.28±0.01	0.32±0.04	0.29±0.03	0.33±0.02	0.39±0.03
TAG56:4	0.06±0.01	0.06±0.00	0.05±0.00	0.06±0.00	0.06±0.00	0.06±0.01	0.03±0.00	0.04±0.00	0.04±0.00	0.06±0.00	0.05±0.00
TAG58:0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
TAG58:1	0.03±0.01	0.08±0.04	0.05±0.01	0.05±0.00	0.05±0.01	0.05±0.02	0.06±0.00	0.06±0.02	0.04±0.01	0.05±0.01	0.04±0.01
TAG58:2	0.01±0.00	0.01±0.01	0.01±0.00	0.01±0.00	0.02±0.00	0.02±0.01	0.01±0.00	0.01±0.00	0.01±0.00	0.01±0.00	0.01±0.00
TAG60:0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

n.d. : not detected. TAG species are designated by total number of fatty acid carbons and total number of double bonds, separated by a colon.

Supporting information for LC-MS/MS method:

An appropriate amount of internal lipid standard triacylglycerol (TAG) 17:0/17:0/17:0 was added to the lipid samples. LC-MS/MS (MRM mode) analyses were carried out using a QTRAP® 5500 (ABSciex) mass spectrometer coupled to a LC system (Ultimate 3000, Dionex). Analyses were achieved in positive mode and nitrogen was used as curtain gas (set to 30), gas1 (set to 25) and gas2 (set to 0). Needle voltage was at +5500 V without needle heating; the declustering potential was set at +40 V. The collision gas was also nitrogen; collision energy was adjusted at +34 V. The dwell time was set to 3 ms. Reversed phase chromatography was carried out at 30°C using a Luna 3u C8 150×1 mm column, with 100 Å pore size, 3 µm particles (Phenomenex). Eluent A was ACN/MeOH/H_2O (19/19/2) +0.2% formic acid +0.028% NH_4OH and eluent B was isopropanol +0.2% formic acid +0.028% NH_4OH. The gradient elution program was 0-5 min, 15% B; 5-35 min, 15-40 % B; 35-50 min, 55% B; 51-58 min, 80% B; 59-70 min, 15% B. The flow rate was 40 µL/min and injection volume was 3µL. Peak area was quantified using MultiQuant software (v2.1, ABSciex).

ESI-MS/MS and ESI-MS/MS/MS analyses for *regio*-localization were performed using a QTRAP® 5500 (ABSciex) mass spectrometer. Settings were as described above and the AF2 parameter varied from 0.08 to 0.3 depending on the compound for MS/MS/MS. Identification was made using Analyse software (v1.5.2., ABsciex). *Regio*-localization of fatty acid at position *sn*-2 was carried out using MS/MS/MS fragmentation. The first fragmentation leads to generation of a diacyl product ion and preferential loss of one fatty acid at *sn*-1/3 positions (Hvattum *et al*., 1998) and a second fragmentation of the diacyl product ion generates acylium and related ions that will reveal the identity of the fatty acid in *sn*-2 position by comparing the intensity of the ions $[FA_{1/3}CO]^+$ and $[FA_2\text{-}CO]^+$ (Hsu & Turk, 2010, McAnoy *et al*., 2005).

Hsu, F. F., & Turk, J. (2010). Electrospray ionization multiple-stage linear ion-trap mass spectrometry for structural elucidation of triacylglycerols: assignment of fatty acyl groups on the glycerol backbone and location of double bonds. *J Am Soc Mass Spectrom, 21*, 657-669.

Hvattum, E., Hagelin, G., & Larsen, A. (1998). Study of mechanisms involved in the collision-induced dissociation of carboxylate anions from glycerophospholipids using negative ion electrospray tandem quadrupole mass spectrometry. *Rapid Commun Mass Spectrom, 12*, 1405-1409.

McAnoy, A. M., Wu, C. C., & Murphy, R. C. (2005). Direct qualitative analysis of triacylglycerols by electrospray mass spectrometry using a linear ion trap. *J Am Soc Mass Spectrom, 16*, 1498-1509.

3.2 Lipidomic analysis of oil palm fruits during ripening

In order to understand acyl flux in TAG assembly occurring in oil palm fruits, quantitative analyses of total lipids, phospholipids, DAG and TAG from oil palm fruit (5 developmental stages) and kernel (5 developmental stages) during ripening have been made by gas chromatography (GC) and liquid chromatography/tandem mass spectrometry (LC-MS).

3.2.1 Time-course of oil accumulation in mesocarp and kernel during fruit ripening

We analyzed total fatty acid content of mesocarp and kernel during fruit development, from palms grown at La Dibamba (Cameroon). Oil palm kernel and mesocarp both accumulate high oil content, about 59% and 87.4%, respectively (Fig.3-1). However, the stages of the highest oil productivity of these two tissues are different. Oil started to accumulate from 10 weeks after pollination (WAP) and finished at 21 WAP in kernel, with the highest oil production rate at the beginning of oil accumulation (10-12 WAP). Oil synthesis in mesocarp starts later, when 80% of kernel oil has already been deposited: fatty acid content starts to increase from 17 WAP with highest rates of accumulation between 19-21 WAP in mesocarp. Strangely, oil decreases between 21 and 23 WAP.

Here, contrarily to previous reports that indicate highest levels are reached at 23 WAP (Bourgis *et al*., 2011; Dussert *et al*., 2013; Tranbarger *et al*., 2011), oil synthesis seems to peak at 21 WAP and decrease afterwards. This is likely to result from mislabeling of samples or from a recording error of pollination date. Work carried out in Southeast Asia shows that oil accumulation occurs from 15 WAP to 20 WAP in mesocarp (Sambanthamurthi *et al*., 2000), which is about 2-3 weeks faster than oil palm in Africa. Likewise, the kernel from Cameroon oil palm continues to accumulate oil till 21WAP, while oil deposition in kernel from Southeast Asia oil palm is almost

complete at 16WAP. For all these reasons, it is safer to determine ripening stages based on oil content rather than WAP.

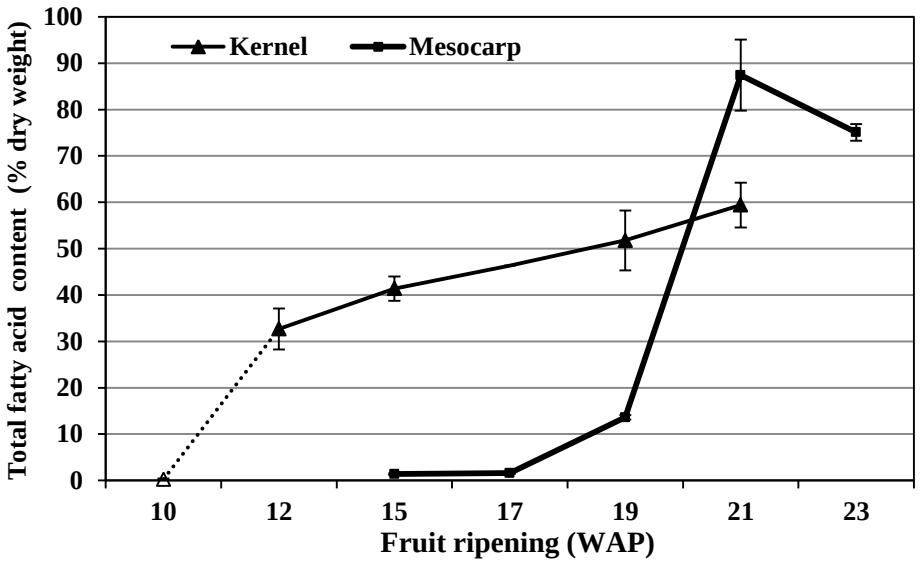

Fig.3-1 Total fatty acid content of kernel and mesocarp during fruit ripening. Data are means of four biological replicates (individual fruits from same bunch) +/- standard error. Each tissue was ground into powder in liquid nitrogen and total lipids were extracted from frozen powder to which an appropriate amount of internal standard C17:0 was added. Lipids could then be quantified after methyl ester analysis by GC-FID. Note: Kernel at 10 WAP is very tiny and still liquid and had therefore to be extracted together with surrounding endocarp. Because the shell is estimated to represent about 90-95% of total weight, the actual total fatty acid content of kernel at 10WAP is probably 2-6%.

3.2.2 Fatty acid composition of mesocarp during fruit ripening

Fatty acids in mesocarp are long-chain fatty acids (LCFAs), with 16 and 18 carbon fatty acids and composition of total fatty acid during development is shown in Fig.3-2. During two first stages, main fatty acids are the following, by decreasing level: 16:0, 18:2, 18:1, 18:3 and 18:0. During next stages of fruit development, the levels of 18:2 decrease markedly, 18:3 almost disappears, the sum of PUFAs decreases from 42% to 12% while the content of 18:1 is increased by twice (14 to 39%). Levels of saturated fatty acids remain about even (18:0) or show a slight increase (16:0 from 35 to 44%).

We have analyzed total fatty acids, which may come from oil and membrane lipids. Because oil is not abundant at 15 and 17 WAP, it is possible that most fatty acids come from polar, membrane lipids. This may explain the different fatty acid

composition of mesocarp from first two stages, especially concerning PUFAs.

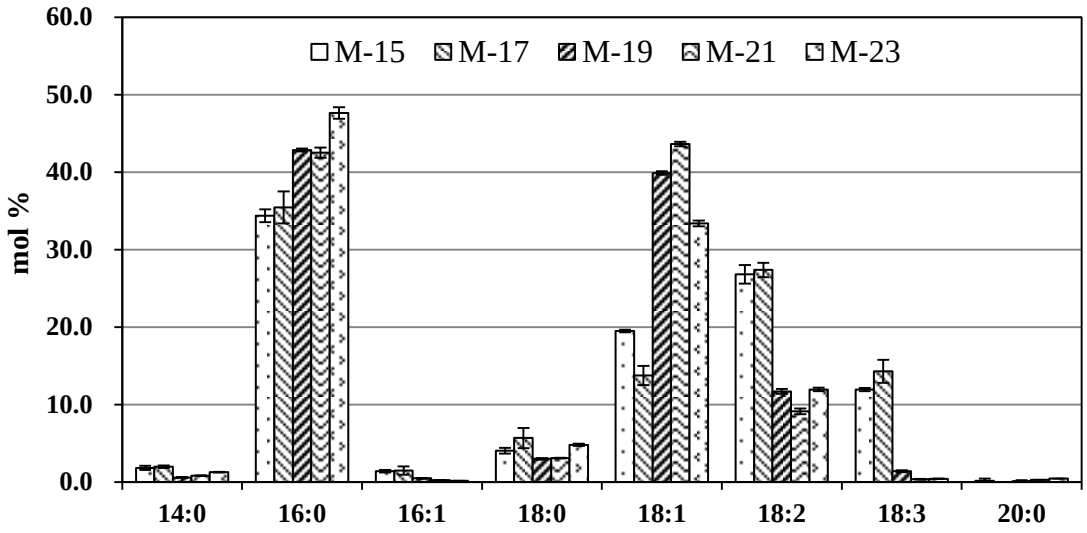

Fig.3-2 Total fatty acid composition of mesocarp from five developmental stages (15, 17, 19, 21, and 23 WAP) during fruit ripening. Data are means of four biological replicates +/- standard error. 14:0, myristic acid; 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid.

The fatty acid composition from mesocarp during development that we determined is about similar with results from others (Dussert *et al*., 2013; Sambanthamurthi *et al*., 2000), and the most striking difference is the amount of 16:0 from our results has about 5% more than other results (Dussert *et al*., 2013). Again, this is likely caused by the different origin of fruits.

3.2.3 Fatty acid composition of kernel during fruit development

When endosperm is still liquid (10 WAP), main fatty acids are 16:0, 18:2, 18:1 and 18:0 (Fig.3-3). Other minor (each below 5%, total: 10%) fatty acids are 12:0, 18:3 and 14:0. Afterwards, when oil starts to accumulate (stage 12 WAP), 18:1 remains about constant (at 27%) when 12:0 (38%) rises tremendously and other MCFAs (8:0, 10:0, 14:0) increase from 3 to 18%; 16:0 level decreases from 36 to 9%, and other 18-long fatty acid sum is about 6% only (versus 34% in previous stage) (Fig.3-3). In next stages, only a few slight changes are noted (increases of 12:0 and 14:0; decrease of 18:0 and 18:2). At maturity, kernel fatty acid composition is strikingly different from that of mesocarp as it contains a majority (about 75%) of saturated MCFAs, and the

content of 12:0 is up to 50% (Fig.3-3).

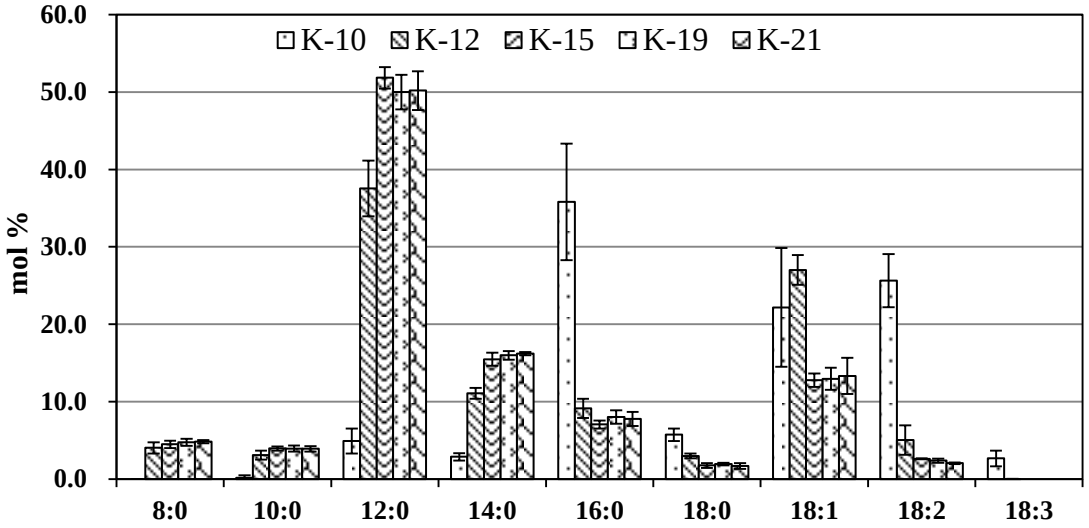

Fig.3-3 Total fatty acid composition of kernel from five developmental stages (10, 12, 15, 19, and 21 WAP) during fruit ripening. Data are means of four biological replicates +/- standard error. 8:0, caprylic acid; 10:0, capric acid; 12:0, lauric acid; 14:0, myristic acid; 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid.

The fatty acid profile of kernel at 10WAP resembles that of mesocarp from 15WAP (Fig.3-2 and 3-3). Again, this profile might reflect mostly that of polar lipids. During fruit ripening, the content of LCFAs is reduced obviously, whereas MCFAs (8:0 to 14:0) increases remarkably to reach about 75% total fatty acids (Fig.3-3).

Fatty acid composition from kernel during development is consistent with previous results (Crombie, 1956; Dussert *et al*., 2013). As for the liquid kernel at early stage, the same three major fatty acids have been detected, including 16:0 (43.7%), 18:1 (20.7%), and 18:2 (15.8%) (Crombie, 1956). The most striking difference is about 4.6% more 12:0 and 3.3% less 18:1 in kernel at 12 WAP we analyzed. Fatty acid composition of oil palm kernel is quite similar to that of endosperm from coconut (*Cocos nucifera* L.), which is grouped with *E. guineensis* in the subfamily Cocosoideae (Corley & Tinker, 2003; Santoso *et al*., 1996). Not only the fatty acid profile of oil palm kernel is like that of coconut endosperm, but also the profile change during fruits development resembles each other (Santoso *et al*., 1996). The mature endosperm of coconut (12-month after pollination) also accumulates about 78% MCFAs, including 49% 12:0 and 19% 14:0 (Santoso *et al*., 1996), just as mature

kernel from oil palm does (Fig.3-3).

3.2.4 TAG molecular species profile of mesocarp and kernel during fruit ripening

To investigate oil accumulation deeper, TAG molecular species compositions of mesocarp and kernel have been analyzed by LC-MS/MS (Fig. 3-4 and 3-5), using the same samples used for total fatty acid determination described above.

3.2.4.1 TAG molecular species profile of mesocarp

As noted before, the major TAG molecular species in ripe mesocarp are 50:1, 50:2 and 52:2 (about 15% for each) (Fig.3-4). Then, two TAG species 52:1 and 52:4 account each for 9% and 6% (Fig.3-4), respectively. Three TAG species 48:0, 48:1 and 52:3 are about 5%, and other species are just below 5% in ripe mesocarp, including all TAG species of 54 (Fig.3-4).

During ripening, main molecular species are the same and only a few of them show notable changes (Fig.3-4). Main species (48:0, 50:1, 50:2, 52:2) show variations (at most 20%) that do not follow a trend during ripening. TAG 46:0, 48:1 and 48:2 species are the major 14:0-containing species and sum of these represent about 10% at last stage, suggesting that total 14:0 content should be about 3.3%. This is not very consistent with total fatty acid pattern obtained from GC analyses as total 14:0 decreased from 1.8% to 1.3% (Fig.3-2) during ripening, while above-mentioned TAG species increased. Most striking differences that follow a trend during ripening are to be found with minor species that contain 18:0 fatty acids (52:1, 54:1, 54:2), the sum of which increase regularly from 4% at first stage to 16% at last stage. However, because total 18:0 fatty acid does not show changes, these increases are likely compensated for by slight decrease noted for 50:1 and un-noticed changes in those species that may contain either 18:0 plus 18:2 or 18:1 plus 18:1. Also, species that contain PUFAs (54:4, 54:5, 54:6) decrease from 9.6% to 3.8%. From MS data, these species do not appear to contain any 18:3 fatty acid and 18:2 is the only polyunsaturated fatty acid.

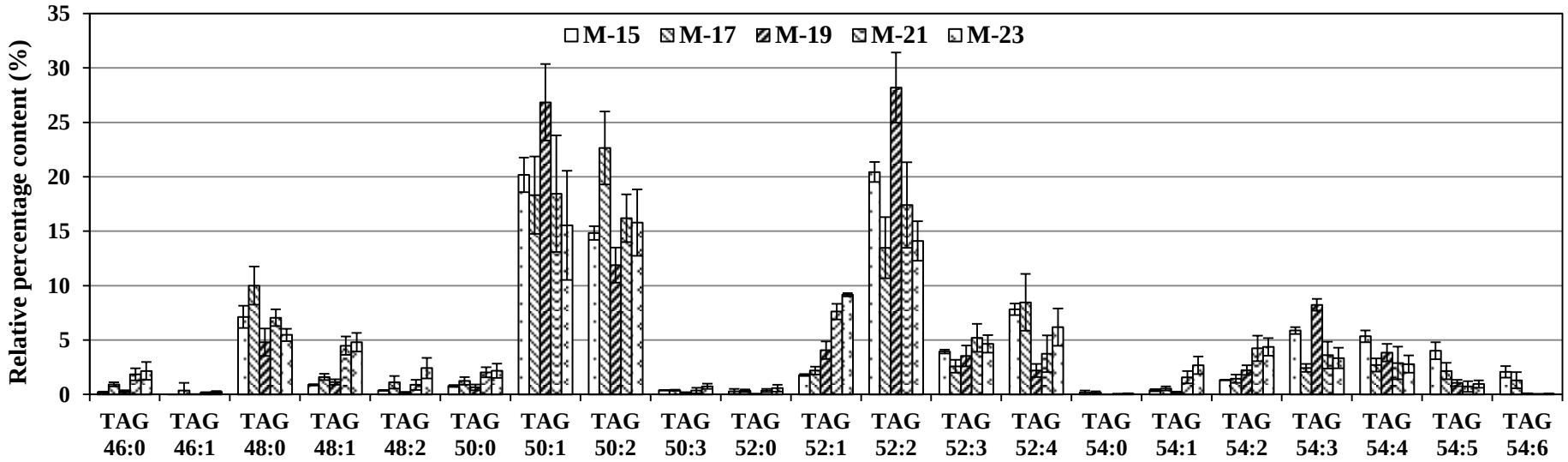

Fig.3-4 Relative quantification analysis of TAG from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (TAG 17:0/17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standards. TAG species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, TAG50:1 is likely to contain two palmitic acids (16:0) and one oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

Therefore, it is likely that TAGs contain more 18:2 fatty acids at early stages than at ripe stage; however, the differences in molecular species are too small to account for the large differences in total fatty acid content, especially with regards to 18:3.

Because of a contamination problem with DAG analysis (see 3.2.6), F. Beisson and B. Légeret in Cadarache kindly accepted to analyze the same samples on a different LC-MS system. Results are shown in Fig.S1 in the Appendix II and are quite similar to our data, with few differences though, notably with regards to species with 18:0 which fluctuation is narrower, from 9 to 15%. This strongly suggests that results from LC-MS are not to be over-interpreted. Mass spectrometry detects ions and molecules may be ionized more or less easily, and we use one standard only (TAG with 3 17:0) for quantification.

3.2.4.2 TAG molecular species profile of kernel

Mature kernel contains 3 types of molecular TAG species. The most abundant ones (55%) are 28-42 species with saturated fatty acids only; 38:0 is the highest species (13%) (Fig.3-5) and 36:0 and 40:0 each account for 10% while the level of 34:0 or 42:0 is separately represented about 9% and 5% (Fig.3-5). Then, there are about 27% of species (from 42 to 46) with at least one unsaturated and one medium-chain saturated fatty acids, major ones (>6% each) being 44:1 and 46:2. At last, species devoid of medium-chain fatty acids (50 and above) represent about 17% of total species. Thus, it seems that there are two separated types of TAGs, one with “normal” LCFAs only, one with medium-chain fatty acids only, and an “intermediate” group, which represents 14% of total species only.

Contrarily to mesocarp, TAG molecular species profile of kernel displayed a considerable change during fruit ripening (Fig.3-5). Species with LCFAs only decreased from 78% to 17%, species with mid-chain fatty acids only increased from 10% to 55% and intermediate species from 8% to 27%.

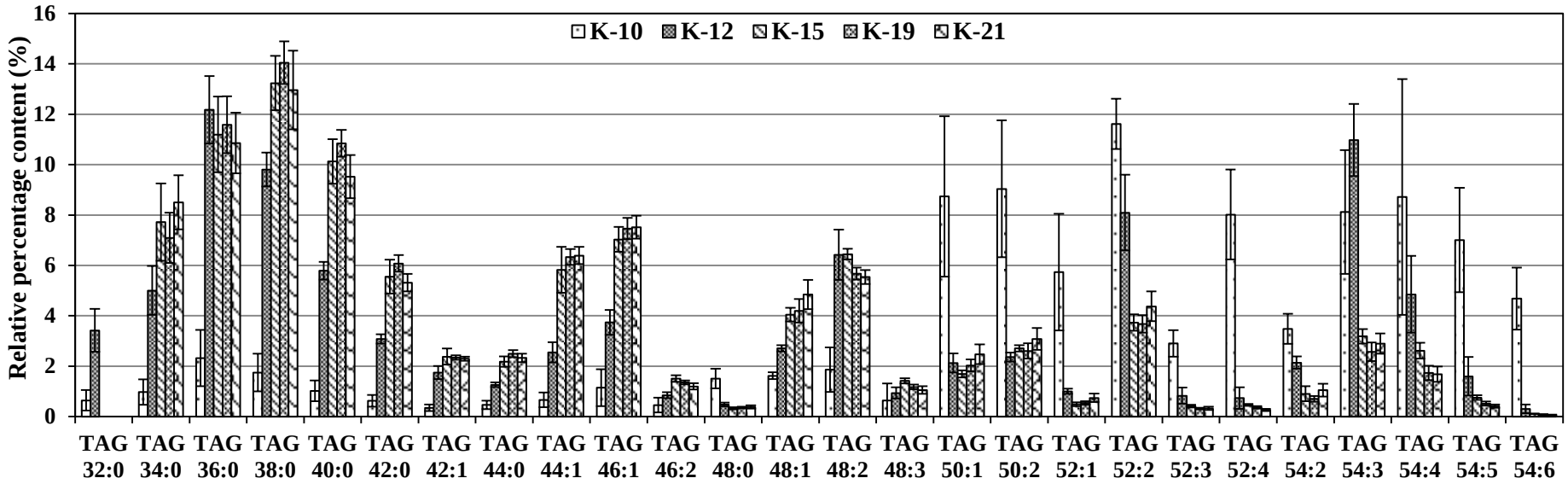

Fig.3-5 Relative quantification analysis of TAG from and kernel at five developmental stages (10, 12, 15, 19, and 21 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (TAG 17:0/17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standards. TAG species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, TAG50:1 is likely to contain two palmitic acids (16:0) and one oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

These results are consistent with total fatty acid composition from kernel (Fig.3-3), except for shorter chain fatty acids. When looking at the results from our colleagues in Cadarache (Fig.S2 in the Appendix II), we noted that our data missed 28:0, 30:0, and 32:0 species, which appeared to represent about 10% total species. Our data were obtained in MRM mode, which means that the machine detects only the molecules previously listed by the experimentator, and these species lacked in the list. This type of monitoring, powerful for routine analysis, points to another possible caveat of LC-MS analyses.

3.2.5 Phospholipid molecular species profile of mesocarp and kernel during fruit ripening

Phospholipid molecular species of mesocarp and kernel during fruit ripening have been analyzed by LC-MS/MS. Phosphatidic acid (PA) was barely detected in our samples and molecular species could not be quantified.

3.2.5.1 Phospholipid molecular species profile of mesocarp

PC and PE from mesocarp show very similar fatty acids species, with little changes upon ripening (Fig.3-6). In ripe fruit, the major components are 34:1 (34% in PC, 24% in PE) and 34:2 (41% in PC, 54% in PE). Total of 36 species is about 23% in PC and 19% in PE; 36:2, 36:3 and 36:4 are the most abundant, representing about 5-8% each. No 18:3 fatty acid could be detected, even in most unsaturated species. Because most of the species' fatty acid composition is not ambiguous, one can estimate the fatty acid composition of PC: it shows 38% 16:0, 3% 18:0, 24% 18:1 and 34% 18:2, which compares well with that of PC purified by TLC from a similar sample (30, 5, 24, 40) shown on Fig.3-7 (Bourgis *et al*., 2011) but shows significant differences for 16:0 and 18:2. During ripening, few changes can be noted, the most important one being the slight decrease of total 36 species (from 37% to 23% in PC, 26% to 19% in PE). The major species (34:1 and 34:2) in PC and PE show very similar change pattern during ripening, except that the amounts of 34:1 and 34:2 in PC are almost equal (around

30%), while the level of 34:2 is nearly twice that of 34:1 in PE (Fig.3-6).

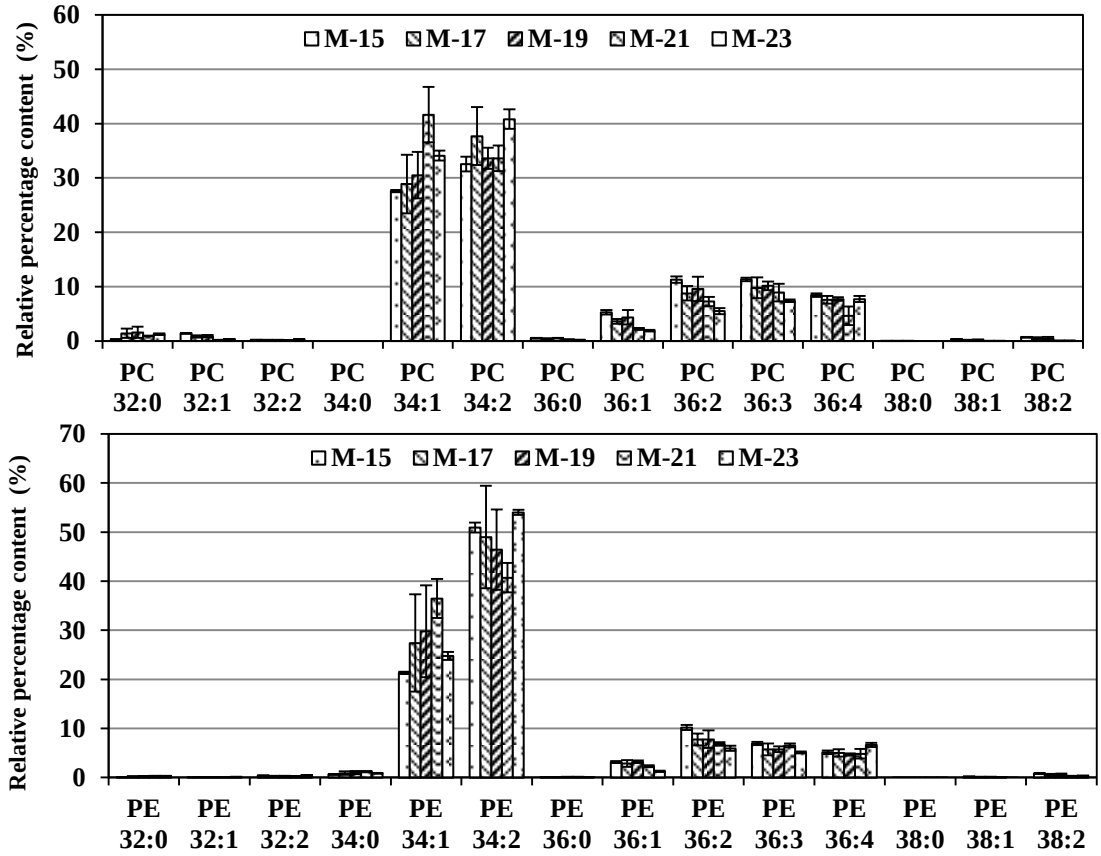

Fig.3-6 Relative quantification analysis of PC and PE from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standards (PE 17:0/17:0 and PC 17:0/17:0 from Avanti Polar Lipids (USA)) were used as internal standards. Species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, PC34:1 is likely to contain palmitic acid (16:0) and oleic acid (18:1).

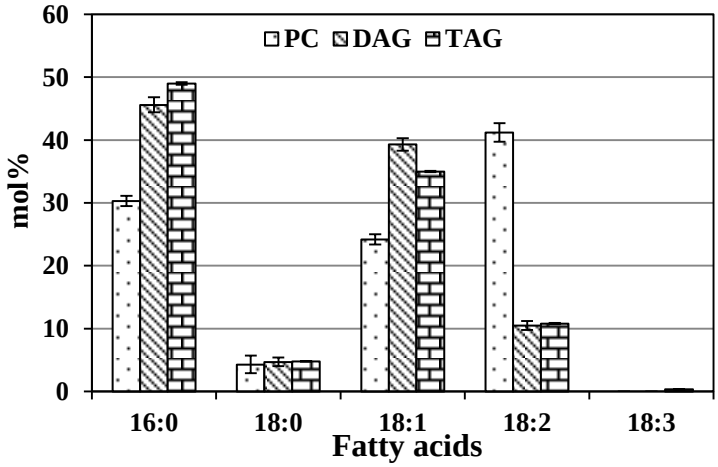

Fig.3-7 Fatty acid composition of PC, DAG and TAG in oil palm mesocarp at 23 WAP (Bourgis *et al*., 2011). 16:0, palmitic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid.

3.2.5.2 Phospholipid molecular species profile of kernel

PC and PE also showed almost identical molecular species composition in ripe kernel (Fig.3-8), and largest differences concern 34:2 (4% in PC, 10% in PE) and 36:2 (46% in PC, 40% in PE). Two main species are 36:2 (about 40%), 34:1 (about 30%) and all other species are below 10%; species above 4% are 32:1, 34:2, 36:1 and 36:3. Species with at least one medium-chain fatty acid (30:1, 32:1) represent about 10% of PC species and 8% of PE's. This is higher than previously reported (1-5%) by Wiberg & Bafor (1995) who used LC coupled to a UV detector rather than mass spectrometry.

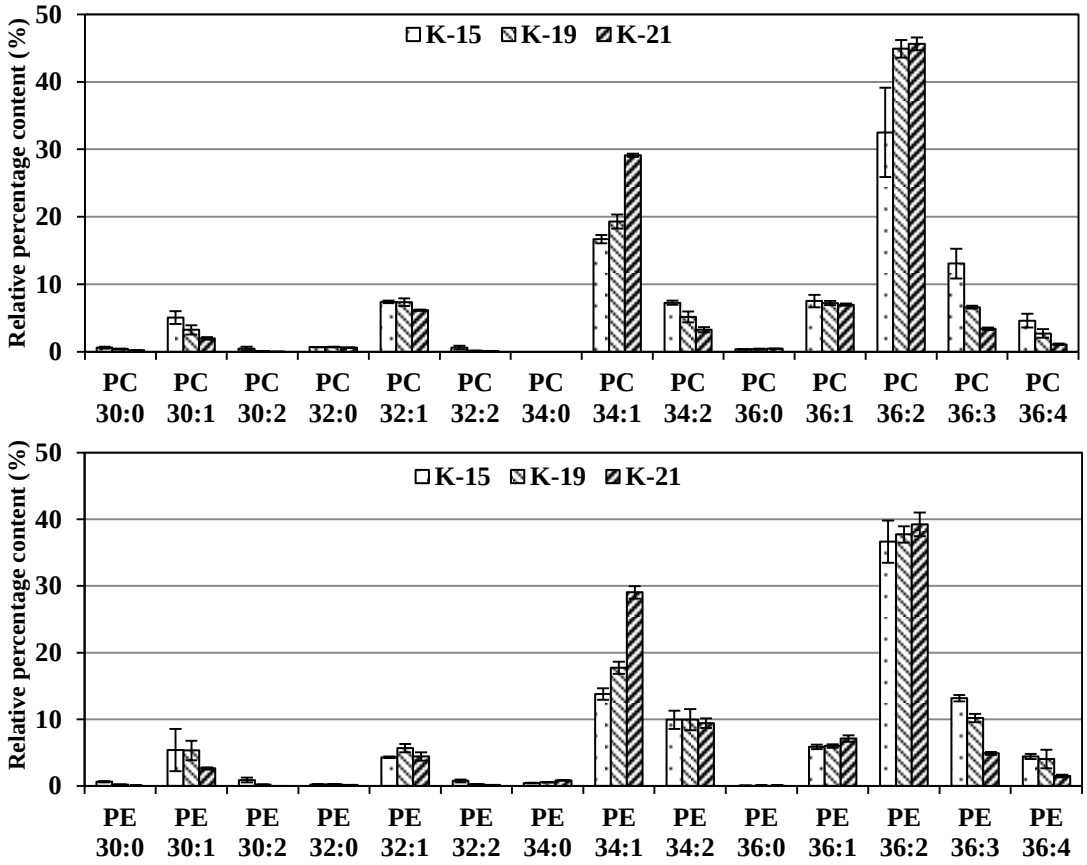

Fig.3-8 Relative quantification analysis of PC and PE from mesocarp at three developmental stages (15, 19, and 21 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standards (PE 17:0/17:0 and PC 17:0/17:0 from Avanti Polar Lipids (USA)) were used as internal standards.

During fruit ripening, trend is again about similar for 2 phospholipids: 34:1 in PC and PE are increasing, whereas polyunsaturated species (36:3 and 36:4) in PC and PE are decreasing. Main difference between PE and PC is noted at stage 15WAP and

concerns levels of 36:2, much lower in PC but standard deviation is very high. Another minor difference concerns levels of 34:2, decreasing in PC and remaining still in PE.

Phospholipids species differ between mesocarp and kernel. 34 species are the most abundant in mesocarp (about 75-80%) while they represent about one third of kernel's species. It is likely that 16:0 occupies *sn*-1 position of the vast majority of PE and PC in mesocarp while 18 carbon fatty acids are more abundant at this position than 16:0 is in kernel. In addition, kernel contains, contrarily to mesocarp, a significant amount of species with one mid-chain fatty acid (about 10% vs 1%).

PE 36:2 can contain two distinct species, namely 18:0-18:2 and 18:1-18:1. Even if MS analysis does not allow real separate quantification, it is sill possible to look at raw data to obtain an estimate of relative amounts of each species. For kernel PE, 18:1-18:1 is clearly the major species (two orders of magnitude difference), while mesocarp PE shows about equal representation of both species.

In conclusion, in each tissue, PE and PC compositions are very close. If phospholipid composition differs depending on tissues, PC and PE contain much more unsaturated fatty acids than TAGs do in both mesocarp and kernel.

3.2.6 DAG profile of mesocarp and kernel during fruit ripening

DAG profile of mesocarp and kernel during fruit ripening also has been analyzed by LC-MS/MS. We noted a very high level of DAG 30:0 (>50%) in mesocarp (Fig.S3 in the Appendix II). This is not consistent with previously-obtained fatty acid analyses of DAGs from ripe mesocarp, as no mid-chain fatty acid was detected (Fig.3-7; Bourgis *et al*., 2011). Dr. F. Beisson in Cadarache told us he had met with similar problem and found that most batches of isopropanol were contaminated by this compound. Dr. F. Beisson and B. Légeret kindly analyzed again our samples and found indeed that this peak was due to contamination, while no other striking difference could be noted. The two raw profiles (Bordeaux and Cadarache analyses)

are shown in supplementary data (Fig.S3 and S4 in the Appendix II); because we preferred to compare samples analyzed by the same machine, mesocarp DAG profile shown in Fig.3-9 comes from data acquired in Bordeaux from which DAG 30:0 has been subtracted.

3.2.6.1 DAG profile of mesocarp

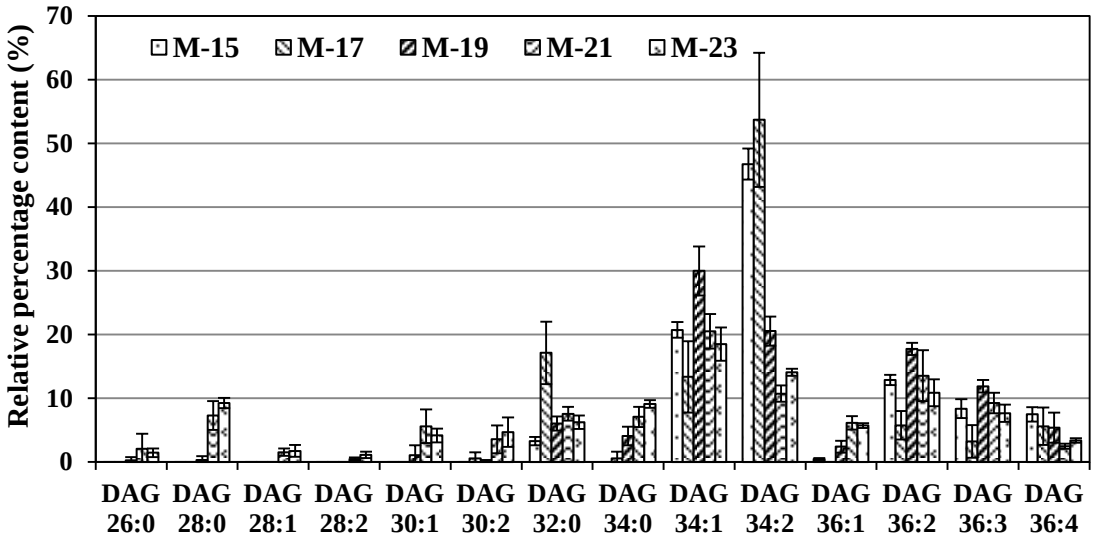

Fig.3-9 Revised relative quantification analysis of DAG from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (DAG 17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standard. Species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, DAG 34:1 is likely to contain palmitic acid (16:0) and oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

Unsaturated 34 and 36 species are the major components in DAG from mesocarp, (Fig.3-9). In ripe mesocarp, most abundant species are 34:1 (about 20%) followed by 34:2 (12%); 36:2, 36:3, 34:0 and 28:0 are represented at the level of about 10% each, and 32:0 and 36:1 account for about 6%. A few species that contain at least one mid-chain saturated fatty acid can be detected, including 28:0, 32:0 and unsaturated 28 and 32, for a total of about 25%, which means that there should be about 12% of MCFAs in DAGs. Again, we believe this is unlikely as no MCFA was detected by GC analysis of DAGs (detection limit of about 1%, Fig.3-7; Bourgis *et al*., 2011). Slight contamination of mesocarp samples by kernel is always possible (a hammer is being

used for breaking and “dissecting” frozen fruits) but this is unlikely to impact lipid composition of late stages of mesocarp because high oil content would tremendously dilute the contaminant. Actually, these MCFAs are not detected at the first stage and increase during ripening; mesocarp contamination by kernel would show an opposite trend. Alternative might be that MCFAs are poorly extracted from TLC plantes and/or evaporate together with hexane upon concentration of FAMEs and would be lost before GC analyses.

Several differences can be noted during ripening. Firstly, DAG 34:2 is the major species (about 50%) at the first two stages but represents about 12% only at last stage. DAG 34:0 and 36:1 are (almost) undetectable at first stage and represent about 10% and 6% of species at last stage, respectively. Levels of DAG 36:4 decrease also markedly (8% down to 3%). Thus, it seems that species with 18:2 decrease and that with 18:0 increase during ripening. This might be related to increases previously noted for 18:0-containing TAG species and decreases of polyunsaturated TAG species. Also, if DAG determined composition reflects that of DAG pool used for TAG synthesis, then sum of saturated 32:0 and 34:0 species (about 17%) can account for all saturated fatty acids found at *sn*-2 position of TAGs (about 15%, Sambanthamurthi *et al*., 2000). There are also a few differences from one stage to another that do not follow a trend during ripening.

3.2.6.2 DAG profile of kernel

The major DAG species in kernel is DAG 30:1 and it is represented about 80% in mature kernel (Fig.3-10). Then, DAG 28:0 and 30:0 account for 7% and 4%, respectively. Other minor DAG species all ranged 1-3%, including DAG 26:0, 28:1, 30:2, and 36:2.

During ripening (Fig.3-10), the level of DAG 30:1 (12:0 18:1) increases from 43% to 80% while DAG with mid-chain species, only decrease from 42% to 10%. The only detectable species with LCFAs only (36:2) remains below 2% total species.

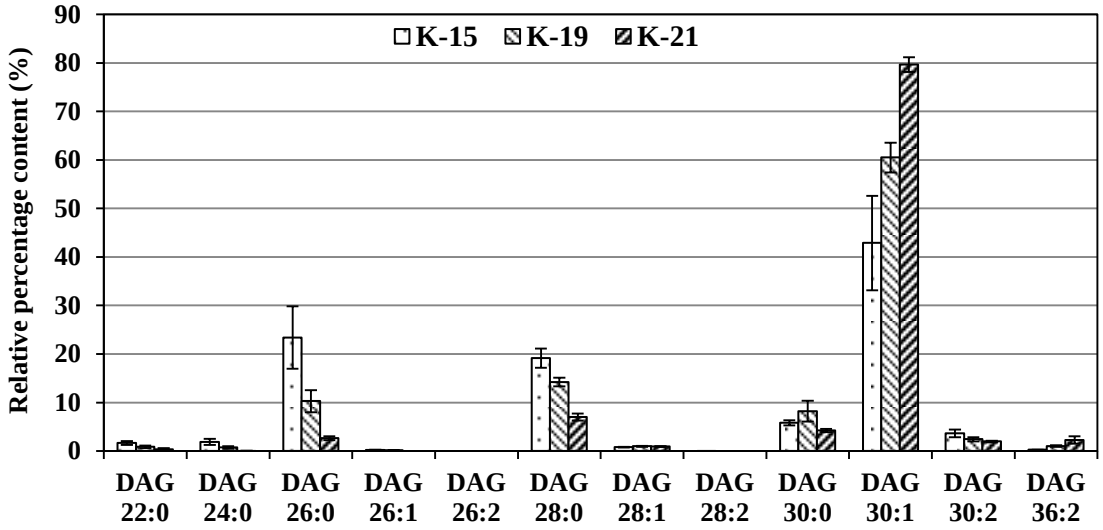

Fig.3-10 Relative quantification analysis of DAG from kernel at three developmental stages (15, 19, and 21 WAP), by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (DAG 17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standard. Species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, DAG 34:1 is likely to contain palmitic acid (16:0) and oleic acid (18:1).

DAGs are both precursor to TAGs and phospholipids and both metabolisms are active in kernel. In addition, there are at least two different TAG synthesizing systems and a comparison of DAG, phospholipids and TAGs might give us hints to possible fluxes. However, such analysis is restricted by lack of data for two first stages, lack of absolute quantification of DAGs and also by lack of knowledge about size of possible DAG pools, so we can just compare amounts of possible precursor DAGs to TAG and PC products. DAGs that are precursor to "intermediate" species of TAGs (mixed mid- and long chain fatty acids) increase from 53 to 88% of total DAGs, to be compared to 27% of "intermediate" TAGs. DAGs that are precursor to mid-chain-only TAGs decrease from 47 to 10%, to be compared to 55% of mid-chain TAGs. DAG 36:2 is potent precursor to majority of long-chain only TAG species (the ones with 52 and 54 carbons). They increase from traces to 2 percent, when possibly derived TAGs remain about constant during same stages (15-21 WAP). These DAGs are also precursors to main PC species (36:2, from 33 to 46%). The discrepencies between % precursors and products might suggest the existence of specific DAG pools of various sizes.

3.3 Conclusions

Lipid analyses in mesocarp show very few distinctive features.

First, variability in fatty acid content noted in different genotypes does not seem to show much impact on TAG structure as neither fatty acid at *sn*-2 position shows change, nor does emerge new major molecular species. ER-localized LPATs in plants are known to be selective for 18:1 fatty acid (Sambanthamurthi *et al*., 2000) except in rare cases (Class B LPAT; Lassner *et al*., 1995) but transcriptome analyses (see 4.1.1.1 in Chapter 4) indicate that LPATs in mesocarp are the same ones as those of leaf tissue. Most changes are actually linked to those TAG main species, 50:1 and 52:2, that contain one or two 16:0 and one or two 18:1, and also to those with three identical fatty acids.

This suggests, as expected, that TAG assembly and editing do not have much impact on final oil composition in these lines, which is likely determined by the nature of fatty acids exported from plastid, most likely under the control of thioesterases and KASIII (Dussert *et al*., 2013). Whatever the biochemical basis responsible for low levels of 16:0 of oil from select trees, it is clear that these might serve to produce palm oil with low saturate content, which is interesting for oil palm economy and possibly for health of consumers as discussed above.

Second, TAG molecular species show little changes during ripening. Main changes with a pattern that follows development concerns polyunsaturated species, which represent a few percent of total species only. Species 54:5 and 54:6 strongly decrease during ripening, which may suggest that they are not synthesized any more when oil accumulates and that they are likely diluted by newly synthesized TAGs that contain mostly 16:0 and 18:1. One may speculate that, at initial stages, TAG content is very low and serves mostly to recycle fatty acids from phospholipids, which contain more 18:2. At this time, DAGs are likely mostly used to synthesize phospholipids, which contain about 40% 18:2. This is consistent with major DAG species (34:2, about 50% total species) containing an 18:2 fatty acid. Then, when oil starts to accumulate, a new

system develops that use 16:0 and 18:1 fatty acids to synthesize storage TAG. DAGs, now mainly precursor to storage TAGs, show composition changes, with lower amount of 18:2 containing species (34:2 drops down to 12-14%). The storage oil synthesizing system is likely independent from phospholipid synthesis as composition of PC and PE molecular species remain unchanged throughout ripening. At the end of ripening, DAG fatty acid composition is nearly the same as that of TAGs, while that of PC contains more 18:2 fatty acids. Comparison of DAG and PC fatty acid composition is a criteria proposed by Bates and Browse (2012) to discriminate plants which use the PDCT pathway (PC direct intermediate to TAG synthesis) from those which do not. As for the 18:3 fatty acids detected at early stages, because they are found neither in oil nor in main phospholipids, they are likely part of plastid galactolipids which are known to usually contain mostly tri-unsaturated fatty acids.

Taken together, these data suggest that oil synthesis operates mostly through Kennedy pathway and that PC is not a significant intermediate in TAG synthesis in palm mesocarp. This hypothesis is consistent with low levels of PUFAs in oil which means that little PC-linked editing is required.

Analysis of kernel lipids showed, contrarily to mesocarp, striking differences. In ripe kernel, molecular species can be divided into three groups. The first group contains exclusively mid-chain saturated fatty acids and represents about 55% total species. The second group contains exclusively LCFAs and represents about 17% total species. The third group is mixed, with both mid- and long-chain fatty acids (about 27%). At first stage of development, about 80% of TAG species contain LCFAs only. Therefore, it is clear that at least two TAG synthesizing systems operate concomitantly and which relative importance fluctuates during kernel development. The intermediate pool might be due to the fact that both systems are not fully separated. This hypothetical lack of full separation might also explain the existence of PC and PE species with one mid-chain fatty acid (30:1, 32:1, about 10% total species). Interestingly, PC and PE show some changes during ripening, which means that their metabolism is still active. DAG main species, 30:1, is a “mixed” species, with 12:0

and 18:1; it represents 40 to 80% total species, increasing during last stages of ripening. Concomitantly, fully saturated species decrease from about 50% to 15% and 36:2, the only species that contain LCFAs, represents a few percent of total species only. This suggests that, during last stages, DAG synthesis for mid-chain system decreases and remains only a DAG species precursor to the “mixed” TAG population. Of course, there might also be active DAG pools of small size and overall DAG content might not reflect actual fluxes to TAG and phospholipids.

Chapter 4. Interaction study of oil assembly enzymes

4.1 Interaction study using yeast two-hybrid system (split-ubiquitin)

4.1.1 Choice of enzymes and cDNA cloning

I selected all enzymes of the Kennedy pathway to TAG synthesis (GPAT9, LPATs, PAPs, DGATs), enzymes involved in desaturation (FAD2) and acyl editing (LPCATs, LPEAT, PDCT, LACS), select enzymes of PC biosynthesis (AAPT and CCTs, a key enzyme for PC synthesis regulation (Craddock *et al*., 2015)), an enzyme has overlapping function with DGAT in TAG biosynthesis (PDAT1), a key enzyme of phospholipid (PI, PS) synthesis from PA (CDP-DAGs) and enzymes/proteins that may also be involved in pathway, such as ACBPs. Because oil palm may contain several isoforms for these enzymes, I looked at transcription data to select those with highest expression levels in mesocarp. In addition, I added four enzymes which physiological function is unknown but that might be candidates for oil synthesis in oil palm mesocarp as they show much higher transcript levels in oil palm mesocarp than in date palm mesocarp, which contains no oil. One of these belongs to bifunctional wax ester synthase/DGAT of the WSD1 family (WSD1-like); three others are possible PA phosphatase homologous to *Arabidopsis* LPPβ and LPPδ.

After enzymes were selected, new gene models of better quality were made available by GenBank in 2015 based on genome sequence (Singh *et al*., 2013b). Most of them showed medium expression levels, in the range of 10-70 reads per kilobase per million ESTs (RPKM). *CCT* and *ACBP3* showed transcript levels an order of magnitude above, while *ACBP6a* and *FAD2* mRNA represented 0.1 and 1% of total transcripts, respectively.

4.1.1.1 Transcription patterns of enzymes of Kennedy pathway to TAG synthesis

As noted in Bourgis *et al*. (2011), *EgGPAT9* and *EgLPATs* (Fig.4-1) showed flat transcript levels, comparable to that of leaves, a bit lower in mesocarp than in kernel. Also, *EgDGAT1* (Fig.4-1) showed about 5-fold difference with leaf and levels were

similar in medium ripening mesocarp and kernel. Only *EgDGAT2* isoforms 1 and 2 (Fig.4-1) were significantly up-regulated during mesocarp ripening (6 and 3 -fold, respectively) and were 4-fold higher than in kernel. On average, *EgPDAT1* levels (Fig.4-1) were similar in both mesocarp and kernel and slightly higher (about 2.5-fold) than in leaves. Putative *EgWSD1-like* candidate (Fig.4-1) showed transcript levels at a bit higher in ripe mesocarp than in leaves and were very low in kernel (30-fold less than in ripe mesocarp).

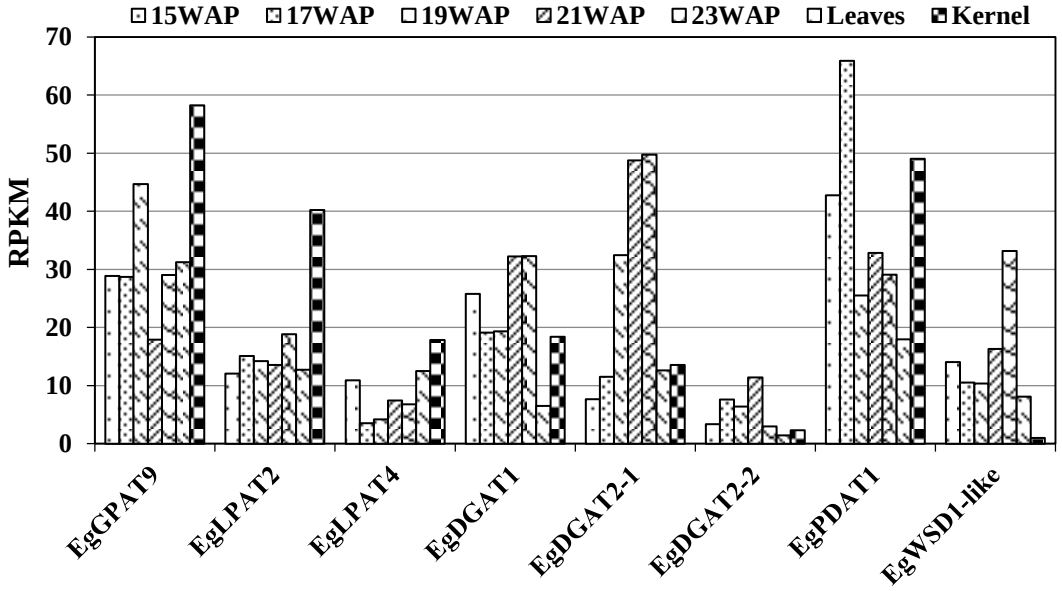

Fig.4-1 RNAseq-based expression patterns of enzymes (EgGPAT9, EgLPATs, EgDGATs, EgPDAT1, EgWSD1-like) in leaves, kernel, and mesocarp at different stages of fruit development. RPKM: reads per kilobase per million reads.

Most stricking differences concerned isoforms of PA phosphatases (Fig.4-2). Both *EgLPPβ 1* and *2* (Fig.4-2) were up-regulated in mesocarp (5 and 2 -fold, respectively). Their transcript levels were much higher in mesocarp than in leaf tissue (40 and 48 -fold, respectively), and sum of their transcripts was 15-fold higher than sum of *EgPAH* ones, making them great candidates for contributing to oil synthesis in mesocarp specifically, as kernel showed levels similar to that of leaf tissue. *EgLPPδ* transcript (Fig.4-2) remained about still with slight down-regulation (about 2-fold), and slightly higher in mesocarp than in leaf and kernel. *EgPAH* transcript levels are low in mesocarp. However, *EgPAH2* gene (Fig.4-2) is up-regulated in mesocarp

(about 8-fold) to reach levels 2.3-fold higher than in leaves, while levels of transcripts are much higher in kernel than in leaf tissue (10-fold), suggesting link to oil synthesis. In leaf tissue, *EgPAH1* is the most abundant form of PAPs with transcript levels higher than in both kernel (2-fold) and mesocarp (3-fold) and *EgPAH3* transcript levels are highest in kernel (10-fold) in these tissues. To summarize, mesocarp is mostly associated to *EgLPPβ* and *EgPAH2*, leaf to *EgPAH1* and kernel to *EgPAH2* and *3*.

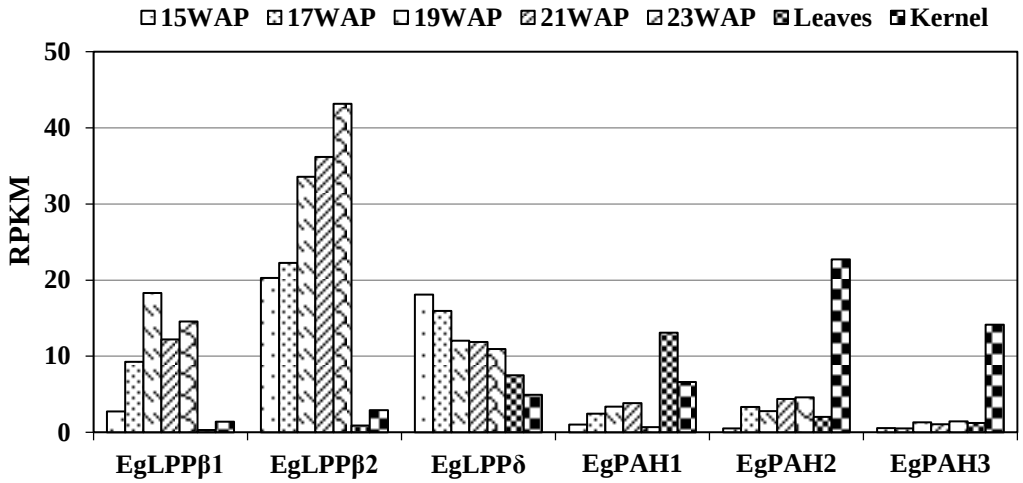

Fig.4-2 RNAseq-based expression patterns of PAPs in leaves, kernel, and mesocarp at different stages of fruit development. RPKM: reads per kilobase per million reads.

4.1.1.2 Phosphatidylcholine-related enzymes

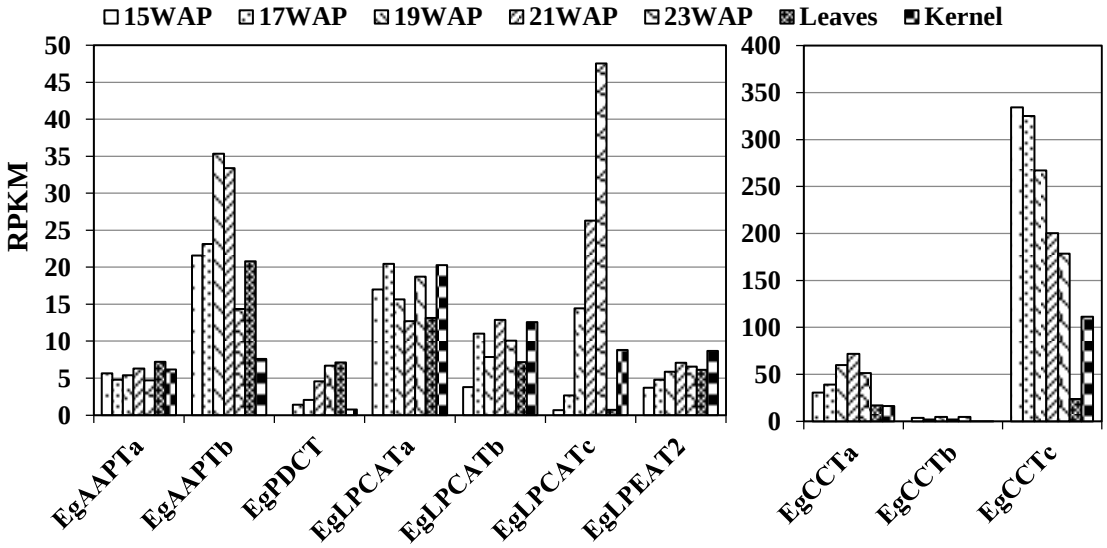

Fig.4-3 RNAseq-based expression patterns of PC-related enzymes in leaves, kernel, and mesocarp at different stages of fruit development. RPKM: reads per kilobase per million reads.

EgAAPTa (Fig.4-3) is constituvely expressed at low levels in all tissues tested (5-7 RPKM), while *EgAAPTb* transcripts (Fig.4-3) show a bell-shape curve in mesocarp (15-35 RPKM) with average levels similar than in leaves. *EgCCTb* (Fig.4-3) shows very few transcripts in all tissues tested, but the other isoforms show important differences. Both *EgCCT a* and *c* (Fig.4-3) have much higher transcript levels in mesocarp than in leaf tissue (5 and 10 -fold, respectively); *EgCCTa* shows slight increases (2-fold) while *EgCCTc* has slight decreases (2-fold) during mesocarp ripening. In both kernel and mesocarp, transcript levels of *EgCCTc* are about 7-fold higher than those of *EgCCTa*, while they are about even in leaves (Fig.4-3).

EgPDCT transcripts (Fig.4-3) are absent in first mesocarp stage and increase steadily during ripening (5-fold from 17WAP), to reach similar levels as that of leaves at last stage. Levels are very low in kernel (about 0.7 RKM, 10 times less than in leaves).

Most striking differences were noted for *EgLPCATc* isoform, with 60-fold up regulation in palm mesocarp; levels at last stage of mesocarp ripening are 60 and 5 -fold higher than in leaf tissue and kernel, respectively. By contrast, other *EgLPCATs* and *EgLPEAT* showed little differences, except possibly for a slight (3-fold) up-regulation of *EgLPCATb* in mesocarp (Fig.4-3).

EgFAD2 desaturase transcripts (Fig.4-4) showed highest differences: they were absent from kernel but represented almost 1% total mesocarp transcript at last stage, about 60 and 2200 -fold more than in leaves and kernel, respectively. Up-regulation during ripening was 32-fold.

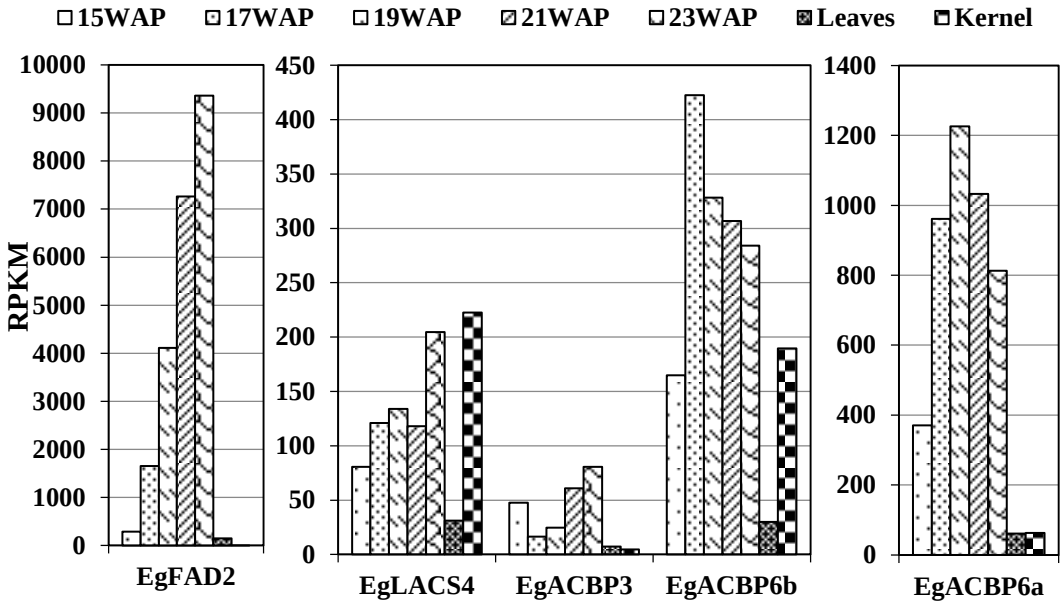

Fig.4-4 RNAseq-based expression patterns of enzymes (EgFAD2, EgLACS4, and EgACBPs) in leaves, kernel, and mesocarp at different stages of fruit development. RPKM: reads per kilobase per million reads.

4.1.1.3 Other enzymes

EgLACS4 (Fig.4-4) was up-regulated 3-fold during mesocarp ripening. Ripe mesocarp and kernel showed levels about 6-fold higher than in leaf tissue.

EgACBP6a (Fig.4-4) shows about 25-fold more transcripts in mesocarp (about 0.1% total transcripts) than in kernel or leaf tissue. It is up-regulated upon ripening (about 3-fold) and shows a bell-like curve. *EgACBP3* (Fig.4-4) is mesocarp-specific and transcript levels increase during ripening (about 3-fold) to reach about 80 RPKM at last stage.

4.1.1.4 Main finding from mesocarp transcriptome study

High oil synthesis in mesocarp is developmentally correlated to higher transcript levels for EgDGAT and EgPDAT1, EgLPPβ type of PA phosphatase and possibly EgPAH2, EgLPCATc isoform, EgPDCT, EgCCTa and c isoforms, EgFAD2 desaturase, EgLACS4, EgACBP3 and EgACBP6 and possibly EgWSD1-like. The two first acyltransferases of Kennedy pathway and EgLPAT show no change, and EgLPPδ very little. Enzymes of PC-editing (EgLPCAT and EgPDCT) are developmentally

regulated and EgCCT pattern is clearly different than that of leaf, suggesting significant editing might occur during mesocarp oil synthesis. Comparison with leaf and kernel point out to probable role of tissue-specific isoforms, especially for EgPAPs and EgLPCATs.

4.1.1.5 Cloning genes encoding chosen enzymes from oil palm mRNA

When I started the work, there were no good gene models so I selected all available ESTs from GenBank coding for these enzymes and reassembled then using different assembly softwares (CLC genomics and also a software included in the CLC suite that is used for sanger EST assembly). I obtained in most cases full-length cDNA models and I designed appropriate PCR primers including *Sfi* I cloning sites to clone open reading frames of these cDNAs; I used mostly RNA isolated from either 21WAP mesocarp or leaf as template for RT-PCR (Reverse transcription polymerase chain reaction). In most cases, cloning went fine, requiring just optimization of PCR annealing temperature, and sequences of clones matched that of cDNA models, with sometimes a silent mutation (that would not affect translated protein sequence). In a few cases, such as with *EgWSD1-like* and *EgPAH2*, deletions were noticed and these deletions were different depending on the clone. However, I finally managed to obtain all clones and their sequences were verified on both strands by Sanger sequencing.

EgAAPTb only was cloned, because strong similarity of *EgAAPTa* primers to *AAPTb* gene (one nucletide mismatches only (Table S2 in the Appendix I), NO amino acid change). True *EgAAPTb*, cloned as ubiquitin C-terminal fusion, is designated AAPTb1. *EgAAPTb* cloned with *EgAAPTa* primers was cloned as ubiquitin N-terminal fusion and designated *EgAAPTb2*.

EgLPCATa was only cloned, because of strong similarity of *EgLPCATb* primers to *EgLPCATa* gene (only two nucletides mismatch (Table S2 in the Appendix I), leading to ONE amino acid substitution at position 462). Both were cloned as ubiquitin N-terminal fusion. True *EgLPCATa* has lysine (AAG) at position 462 and is designated *EgLPCATa1*. *EgLPCATa* cloned with *EgLPCATb* primers has Glutamine

(CAG) instead and is designated *EgLPCATa2*.

All constructs with primer used and PCR conditions are listed in Table S1 and S2 in Appendix I. Then, 454 ESTs were then mapped to these clones to confirm that sequences were good indeed. These clonings were done in bait vector; they were used as template for PCR amplification destined to cloning in other vectors when necessary; again, sequencing was done on both strands prior to use these new clones.

4.1.1.6 Topology analyses of deduced proteins

Yeast two hybrid system requires that fused ubiquitin parts lie in the cytosol. Therefore, one needs to predict topology of proteins of interest prior to cloning, to choose appropriate vector. Deduced amino acid sequences from all clones were used for topology studies. I used four widely-used web sites/softwares to predict topology and possible presence of a signal peptide. Results were frequently different depending on software (Table S4 in Appendix I). However, in most cases, N-terminal part was predicted to be cytosolic and, in most constructions, I fused ubiquitin moieties to the N-terminal end of proteins. Later, each time a construction would not yield interactions, I made a new construction to fuse ubiquitin moiety to the other end of protein and check whether it would work, at least as a prey; while, in all the cases tested, I could not obtain better results. Only EgACBP3a was predicted to have a signal peptide, so it was cloned into PR3-SUC without N-terminal signal sequence, and the vector supplies a yeast signal sequence derived from the SUC2 (invertase) gene of *Saccharomyces cerevisiae*. Some known membrane proteins did not show transmembrane domains (EgPAHs, EgLACS4, EgCCTs); these maybe soluble enzymes that can be recruited to membrane after post-translational modification such as phosphorylation. These were used as prey only, as well as EgACBP6s, a known soluble protein.

4.1.2 Yeast two-hybrid results validation and controls

Thirty-two genes have been cloned from oil palm mRNA. Twenty-two of them

were cloned to both prey and bait vectors, whereas eleven genes coding for soluble proteins were tested as prey only. Thus, we tested all the possible pairs (more than 672) using double plasmids transformation in yeast strain THY.AP4. Serial dilutions (1:10, 5 μL, from 1 to 10^{-3}, starting from 0.4 OD units/mL) from each double transformation were plated on the selective mediums SC-HTL and SC-AHTL (adenine, histidine, tryptophan, and leucine). After 4 days incubation at 30 ℃, the yeast growth statuses on these selective plates were recorded.

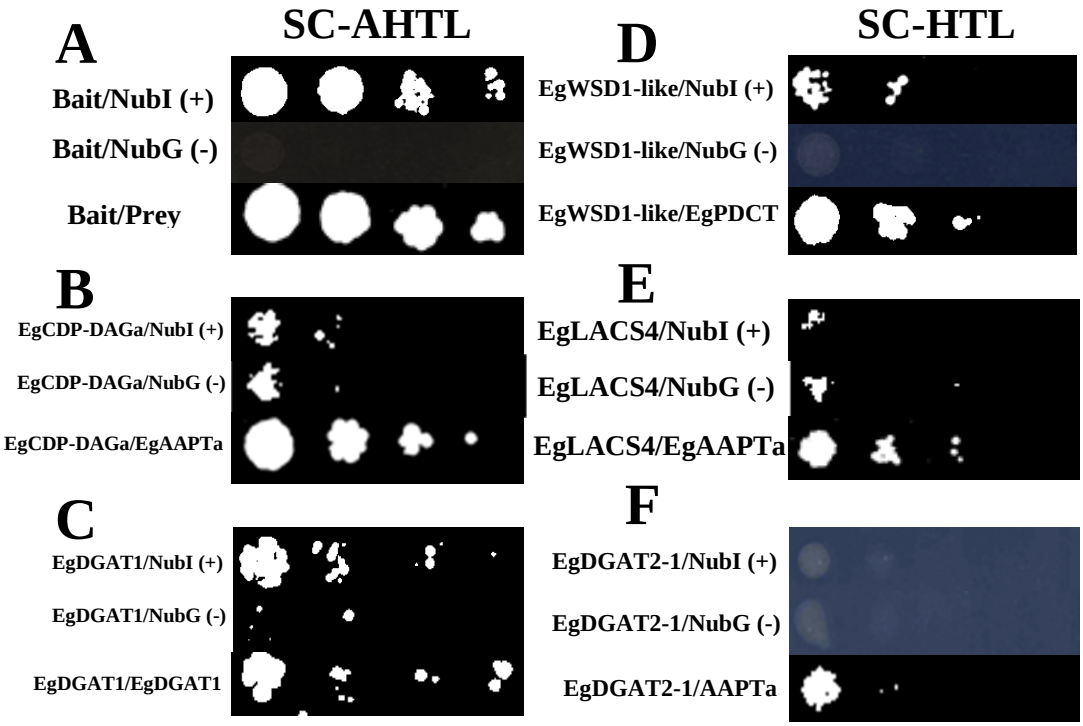

Fig.4-5 The different results from positive and negative controls of each bait and also one example of its positive interaction results. Bait/NubI is positive control (+), and Bait/NubG is negative control (-). Six different types of baits: (A) EgAAPTb2, EgAAPTb1, EgLPCATa1, EgLPCATa2, EgLPCATc, EgLPAT4 and EgFAD2; (B) EgCDP-DAGa and EgLPAT2; (C) EgDGAT1 and EgLPEAT2; (D) EgWSD1-like and EgDGAT2-2; (E) EgLACS4; (F) EgDGAT2-1, EgLPPδ, and EgPDAT1. Four different concentrations (OD_{600}=4×10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4}) of yeast dilutions were spotted on plates SC-TL, SC-HTL, and SC-AHTL.

Every bait was included a positive control (NubI prey construct that leads always to interaction) and two negative controls (empty prey vector and prey construct - mutated NubG - that never shows interactions). In most cases, both controls worked well (Fig.4-5A). In other cases, positive control did not work well (colonies on first

dilution only) but some clones to be tested showed colonies on 3 or 4 of dilutions (Fig.4-5C); these were considered as reflection true interactions. In other cases, negative control proved leaky, allowing cells to grow on first dilution but, again, some clones to be tested showed colonies on 3 or 4 dilutions (Fig.4-5B) and these were considered as reflecting also interaction, providing the difference was reproducible.

Most interactions could be demonstrated using -AHTL medium (Fig.4-5 A, B and C) but others could be detected when plating on -HTL only (Fig.4-5 D, E and F). We classified the positive results into three groups based on the degree of interaction, strong interaction (3-4 dots on -AHTL; Fig.4-5 A and B), medium interaction (1-2 dots on -AHTL, or 3-4 dots on -HTL; Fig.4-5 C and D), and weak interaction (1-2 dots on -HTL; Fig.4-5 E and F). In case of leaky negative control, only those clones growing on more dilutions than negative control were considered to be valid.

Most importantly, each interaction was tested on two independently obtained yeast clones. In most cases, same exact results were obtained. When this was not the case (less than 1.5 % of cases), the experiment was repeated with new yeast clones and we considered the interaction was valid if we could reproduce it 3 times out of 4.

4.1.3 Interaction results

4.1.3.1 General overview of yeast two hybrid results

Among 22 membrane-bound baits, 5 only could not work in the system. Four of them including EgPDCT, EgCDP-DAGb, EgLPPβ1 and EgLPPβ2 strongly interact with everything, and the last one EgGPAT9 shows the similar interaction with controls and proteins tested. However, EgGPAT9, EgPDCT and EgCDP-DAGb could work as prey in membrane Y2H system. EgLPPβ1 and EgLPPβ2 also can activate the system itself even when they work as prey. For the rest of preys we tested, most of them work well in the system except five of them EgACBP3a, EgPAH1, EgPAH2 and EgPAH3a, b, which show no interaction at all. This means that we were unable to test interaction with any of the PA phosphatases except for EgLPPδ.

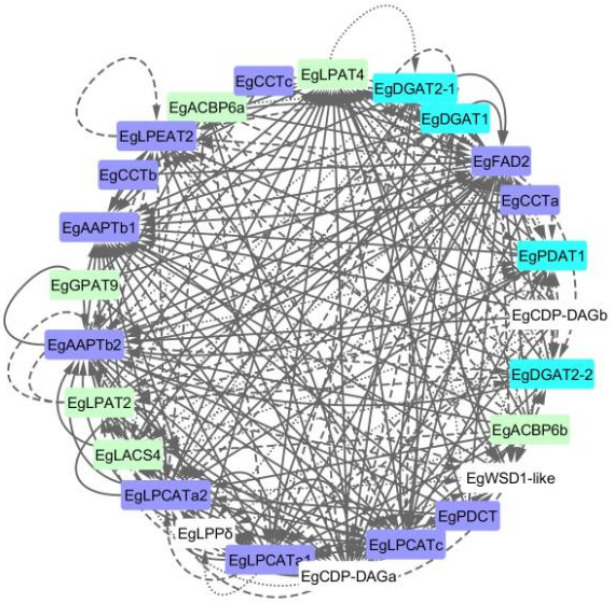

Fig.4-6 All the interactions among these proteins. Different lines mean different interaction degrees. Solid, dash, and dot lines mean strong medium, and week interactions between those two proteins, respectively. The arrows go from the prey to the bait. Enzymes involved in the common pathway for both phospholipids and TAGs are in **light green**, including EgGPAT9, EgLPAT2, EgLPAT4; Enzymes for TAG synthesis are in **blue**, including: EgDGAT1, EgDGAT2-1, EgDGAT2-2, and EgPDAT1; PC-related enzymes are in **purple**, such as EgCCTa, EgCCTb, EgCCTc, EgAAPTa, EgAAPTb, EgLPCATa, EgLPCATb, EgLPCATc, EgLPEAT, EgPDCT, and EgFAD2; Unknown enzymes such as EgCDP-DAGa, EgCDP-DAGb, EgLPPδ and EgWSD1-like have no color. These rules are also used for the following interaction results.

Totally, we showed 241 interactions, including 132 strong interactions, 73 medium interactions and 36 weak interactions. Forty of these interactions both ways, no matter what protein is prey or bait. Interaction results are shown as a network visualized using Cytoscape software (Fig.4-6). Clearly, this suggests that these enzymes may assemble into complexes that may comprise the whole pathway plus editing enzymes.

4.1.3.2 The interaction results of enzymes for TAG synthesis

A. All DGATs interact with both Kennedy pathway and editing enzymes

DGAT is the only enzyme of the Kennedy pathway that is specific for TAG assembly since all other enzymes tested are also involved in phospholipids synthesis. From the results showed in Fig.4-7, we can see that all 3 mesocarp EgDGATs can interact strongly with at least one (EgLPAT) enzyme of the Kennedy pathway. This was expected as lipid flux to TAGs uses this pathway. However, all DGATs interact

also with PC-related enzymes including EgAAPTbs and EgFAD2, which desaturates PC-linked 18:1. This is in agreement with recent *in vivo* studies showing that acyl editing cycle can be a major flux within acyl lipid metabolism of oil bearing tissues even in species that do not accumulate large amounts of PC-modified FA in TAG (Bates & Browse, 2012). Also, it has been shown that disruption of acyl-editing system in triple mutant *pdct1/lpcat1/lpcat2* (Bates *et al*., 2012) leads to reduced amount of oil in seeds. In addition, poor editing system in transgenic plants synthesizing unusual fatty acid leads also to reduced oil synthesis (Bates *et al*., 2014; Eccleston & Ohlrogge, 1998).

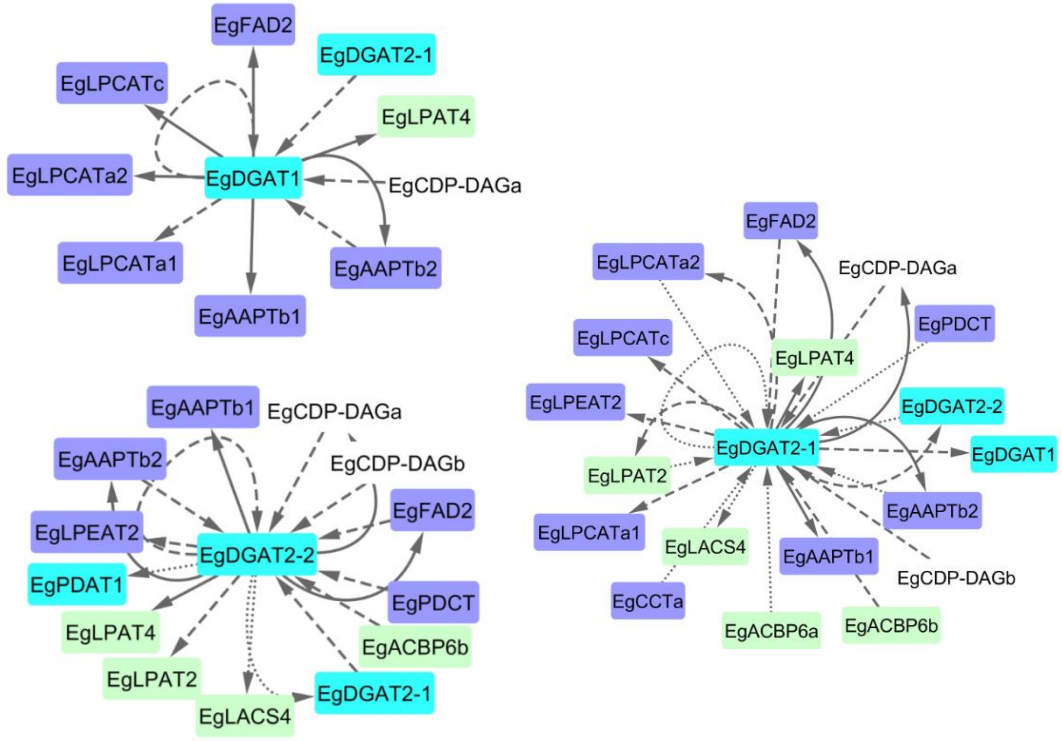

Fig.4-7 Interaction results of EgDGAT1, EgDGAT2-1, and EgDGAT2-2. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

If three EgDGAT enzymes share interactions partners, they also show different patterns. Contrarily to EgDGAT1, both EgDGAT2s interact with EgLPEAT2, a lysophospholipid acyltransferase with broad specificity, including PE and PC (Jasieniecka-Gazarkiewicz *et al*., 2016). They also show interactions with EgPDCT, contrarily to EgDGAT1; this will be further discussed in view of EgDGAT specificities. At last, again contrarily to EgDGAT1, EgDGAT2s interact with both

EgLACS4 and EgACBP6b. This may suggest that DGAT1 and 2, which do not share sequence homology, may play distinct roles. It has been shown in Tung tree that DGAT1 and 2 were located in distinct parts of the ER and synthesized an oil of distinct composition (Shockey *et al*., 2006). DGAT2 enzyme does not appear to play significant role in oil synthesis in *Arabidopsis* seeds (Li-Beisson *et al*., 2013). DGAT2 enzymes in plants are mostly linked to synthesis of oil with unusual fatty acids or oils high in saturated fat, as is the case of oil palm (Liu *et al*., 2012).

Other noticed results are that each EgDGAT interacts at least with another EgDGAT and that none of them interacts with EgGPAT9, contrarily to what has been reported in case of Tung (Gidda *et al*., 2009). At last, all three EgDGATs interact with EgCDP-DAG, an enzyme that is key to PS and PI synthesis (Li-Beisson *et al*., 2013), but which is not supposed to be involved in either PC or TAG synthesis.

B. EgPDAT1 interaction results

PDAT1 is also important to oil synthesis in *Arabidopsis* seeds (Zhang *et al*., 2009). Indeed, EgPDAT1 shows interaction pattern very similar to that of EgDGAT2-1, except that it does not interact with EgLPEAT2 (Fig.4-7 and 4-8). PDAT1 from *Ricinus communis* has been shown in *Arabidopsis* to interact with LPCAT, which is considered to be responsible for re-acylation of the LPC product of PDAT action (van Erp *et al*., 2011).

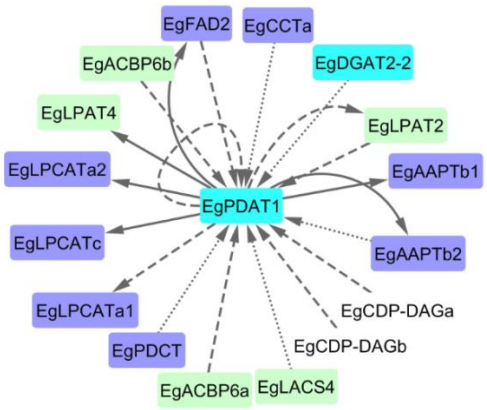

Fig.4-8 Interaction results of EgPDAT1. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

C. EgWSD1-like interacts with oil assembly enzymes

WSD1 from *Arabidopsis* is a member of the bifunctional wax ester synthase / diacylglycerol acyltransferase enzyme family and predominantly functions as a wax synthase (Li *et al*., 2008). Transcriptomic studies from Bourgis *et al*. (2011) indicated that this gene was upregulated in oil palm mesocarp, tissue which is not known to synthesize wax esters, and much higher than in date palm. However, better transcriptome (Fig.4-1) shows less convincing data. From our results, we can see WSD1 homolog from oil palm fruits showed numerous interactions with those enzymes involved in both phospholipid and oil syntheses (Fig.4-9) such as EgLPATs and EgLPCATs, enzymes which are not involved in wax synthesis. The interaction profile of EgWSD1-like is almost identical to that of EgPDAT1, except that WSD1-like does not interact with EgLACS4. These results might suggest that EgWSD1-like, thought to function as a wax ester synthase, might act as a DGAT and contribute to oil synthesis in oil palm mesocarp. Thus, the characterization analysis of *EgWSD1-like* in the mutant yeast H1246 has been carried out (see 5.2 in Chapter 5).

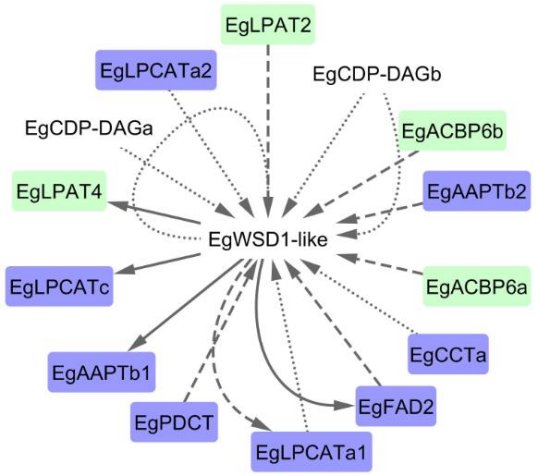

Fig.4-9 Interaction results of EgWSD1-like. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

4.1.3.3 The interaction results of PC related enzymes

A. EgPDCT

PDCT is an enzyme that works on PC but its physiological function has been almost exclusively studied in relation to its role in TAG biosynthesis; thus, it makes a nice transition between TAG- and PC-related enzymes. EgPDCT does not work as

bait, as it activates activate the Y2H membrane system by itself, but it works well as prey. It is just that interaction studies are limited to those enzymes that can be used as bait. The results indicate that EgPDCT have very strong interaction with EgAAPTs, EgLPCATs and EgFAD2 (Fig.4-10), which are all involved in PC metabolism and acyl editing pathways. Moreover, EgPDCT shows medium/weak interactions with known or putative TAG synthesizing enzymes such as EgDGAT2s, EgPDAT1 and EgWSD1-like (Fig.4-9). These data are consistent with the key role of PDCT in TAG synthesis and acyl editing.

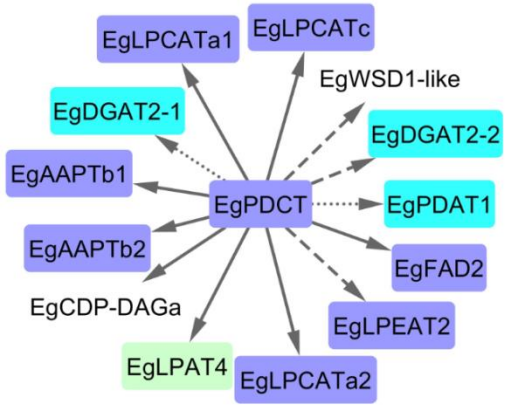

Fig.4-10 Interaction results of EgPDCT. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

B. EgAAPTs

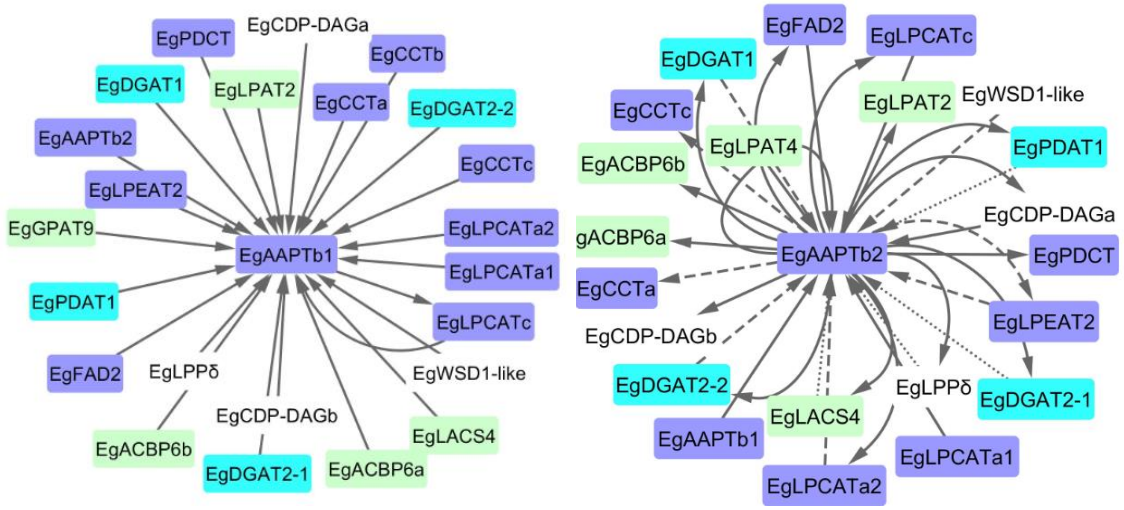

Fig.4-11 Interactions of EgAAPTs. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

EgAAPTa and EgAAPTb are supposed to be two isoforms of AAPT with high

identity. However, both pairs of primer amplified the same cDNA, so EgAAPTa and EgAAPTb are the same isoform EgAAPTb, renamed by EgAAPTb2 and EgAAPTb1, respectively. They were just cloned into two different vectors for Y2H membrane system, with ubiquitin fusion at either N- (EgAAPTb2) or C-terminus (EgAAPTb1). Although vast majority of interactions are identical for both constructs, slight differences can be noted. First is that EgAAPTb1 does not work well as prey (just one interaction with EgLPCATc as bait), contrarily to EgAAPTb2. In addition, while both EgAAPTs interact with almost all proteins tested that work in the system (Fig.4-11), exceptions are EgGPAT9 and EgCCTb for AAPTb2, EgLPAT4 for EgAAPTb1. This illustrates that yeast two-hybrid system results can vary easily just based on how is designed the experiment and it clearly indicates that lack of interaction can be due to experimental design and not to lack of real interaction. Otherwise, the large interaction network of EgAAPTs is very consistent with the key role of enzyme in PC synthesis, and the role it is likely to play to synthesize the PC molecules that will serve to generate DAG species for TAG synthesis. Taken together, these data are consistent with expected central role of AAPT is both TAG and PC synthesis.

C. EgCCTs have distinct interactions

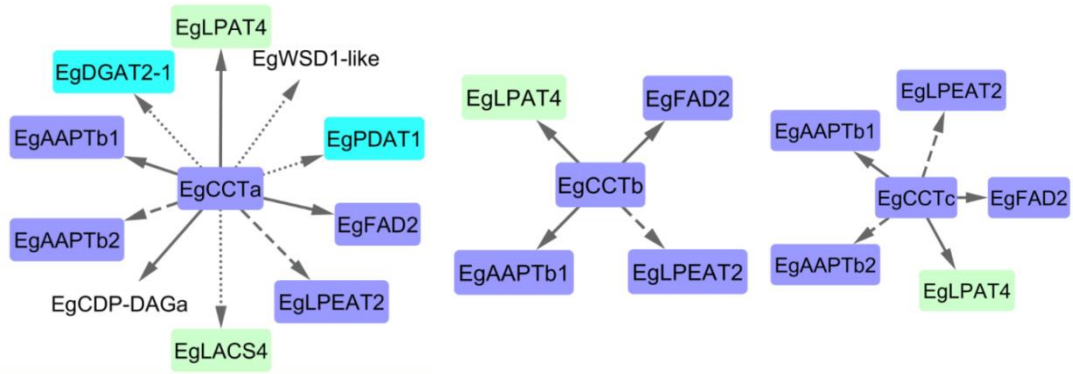

Fig.4-12 Interactions of EgCCTs. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

Choline:CTP cytidyltransferase (CCT) catalyses the synthesis of CDP-choline which will serve as substrate (Craddock *et al*., 2015), together with DAG, for PC synthesis by AAPT. In many biological systems (Cornell & Ridgway, 2015), it is considered as the rate-limiting enzyme of PC synthesis. Therefore, we felt it was

interesting to include it in our studies. We identified three different isoforms which were all tested as prey (CCT is a soluble enzyme that can bind to membranes). These three isoforms of EgCCT showed different interactions. EgCCTb and c just showed interactions with enzymes involved in PC metabolism (EgAAPT, EgFAD2, acyltransferases), while EgCCTa showed, in addition to above-mentioned enzymes, interactions with TAG-related enzymes such as EgDGAT2-1, EgPDAT1, EgWSD1-like, and EgLACS4 and EgCDP-DAG (Fig.4-12). Based on this result, it is quite tempting to speculate that EgCCTa might be specifically involved the biosynthesis of TAG-related PC and might specifically regulate synthesis of PC molecules involved in TAG synthesis.

4.1.3.4 The interaction results of enzymes common to all pathways

A. EgGPAT9 interacts with EgLPCATc and EgLPAT4

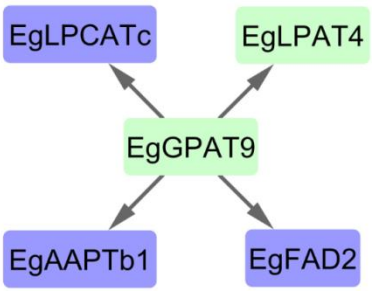

Fig.4-13 Interaction results of EgGPAT9. Solid line means strong interaction. The arrows go from the prey to the bait.

EgGPAT9 could not work as bait (tested in both C- and N- terminal fusion vectors) in Y2H membrane system, which makes it likely that some interactions were missed. However, being used just as bait, it showed strong interactions with EgFAD2, EgAAPTb, EgLPAT4, and EgLPCATc (Fig.4-13). Our results are consistent with a recent study by Shockey *et al*. (2016) in *Arabidopsis*, who showed interactions between AtGPAT9 and AtLPAT2, AtLPCAT2, but not with AtDGAT1, contrarily to a previous study in Tung (Gidda *et al*., 2009).

B. EgLPATs have different interactions

Two isoforms of EgLPAT showed different interaction results in Y2H membrane

system (Fig.4-14). EgLPAT4 works very well as bait and interacts strongly with all proteins we tested except EgAAPTb, but very poorly as prey (one interaction only); whereas EgLPAT2 showed medium and weak interactions as bait but strong one as prey. The fact that EgLPAT2 does not work well as bait may (or may not) explain why it carries less interactions than EgLPAT4. In any case, both EgLPAT interact with enzymes of TAG and PC pathways, and with EgCDP-DAG, consistent with their central function. It is interesting to note that EgLPAT4 interacts with EgLPPδ, the only (putative) PA phosphatase which could be used in our study.

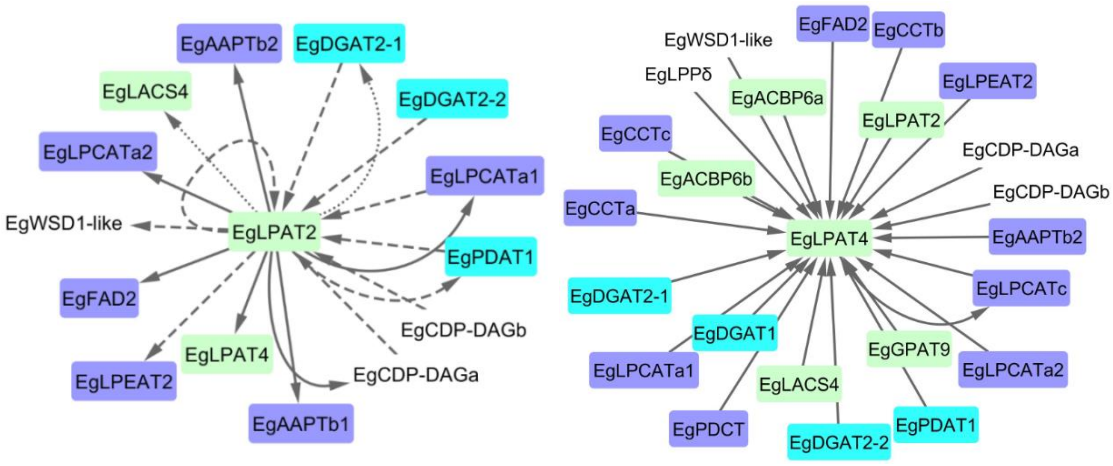

Fig.4-14 Interactions of EgLPATs. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

C. The interaction results of EgLPPδ

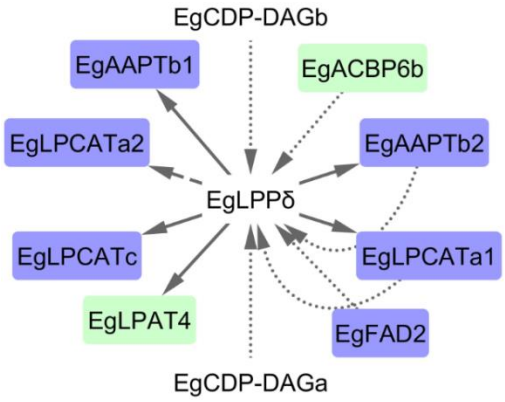

Fig.4-15 Interactions of EgLPPδ. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

Interestingly, EgLPPδ (Fig.4-15) interacts with several enzymes of PC metabolism, with EgCDP-DAGs but not with any TAG-specific enzyme which might suggest that

it might not be involved in TAG synthesis but rather in phospholipid synthesis.

4.1.3.5 Other interesting interaction results

EgLACS4 (Fig.4-16) interacts with all acyltransferases (except EgWSD1-like) and EgACBPs (except ACBP3). This is expected as LACS4 produces the substrate used by these enzymes/proteins. Interaction with EgFAD2 might allow quick acylation of of newly-desaturated fatty acid acid released by action of a phospholipase A2. In addition, EgLACS4 interacts with phospholipid-synthesizing such as EgAAPT, EgCDP-DAG and EgCCTa; it might be interesting to test whether acyl-CoAs can regulate the activity of these enzymes.

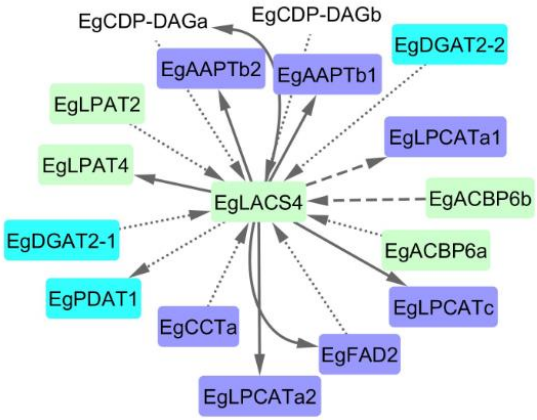

Fig.4-16 Interactions of EgLACS4. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

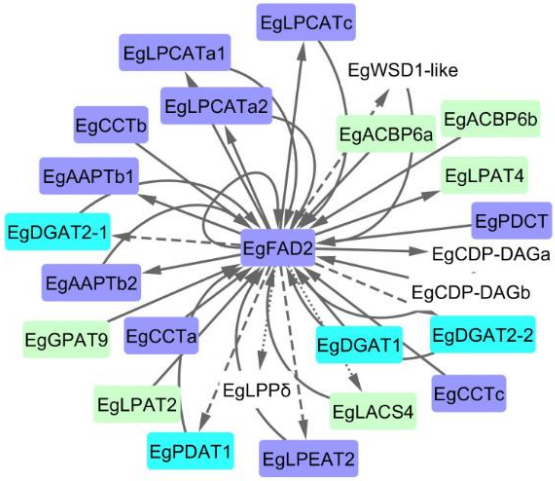

Fig. 4-17 Interactions of EgFAD2. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

EgFAD2 showed interactions with all the protein tested in Y2H membrane system

(Fig.4-17), except those not working in this system, such as EgPAHs, EgLPPβs, and EgACBP3a. This likely reflects its central role in lipid metabolism.

Other interaction networks (EgLPCATs, EgACBP6s, EgCDP-DAGs, and EgLPEAT2) not mentioned here are shown as supplementary data (Fig.S5-S8) in Appendix II.

4.1.4 Conclusions

We used the yeast two-hybrid system (split-ubiquitin) to test whether oil assembly enzymes were able to interact with each other. We selected 32 enzymes and tested 672 interactions. Results showed a total of 241 interactions, including 132 strong interactions, 73 medium and 36 weak ones. Therefore, it is clear from our result that all these enzymes can virtually form almost any possible combination of multi-enzyme complex.

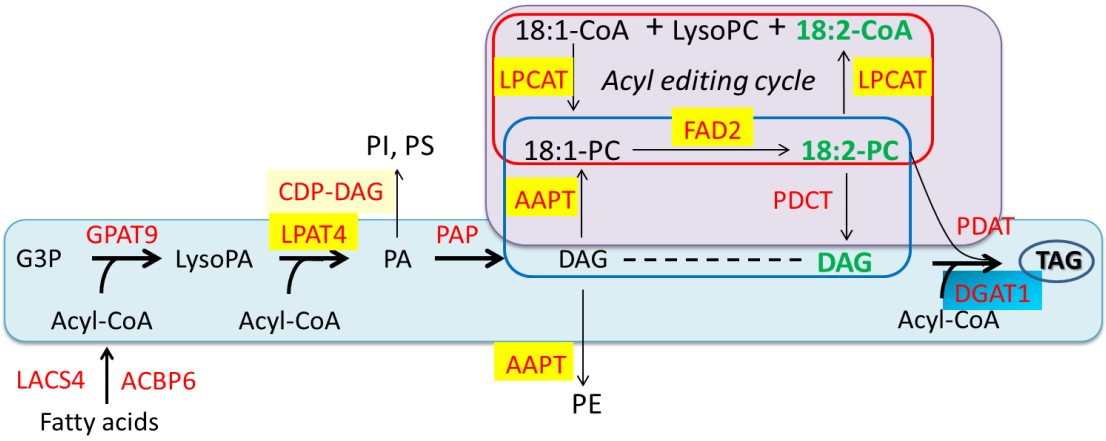

Fig.4-18 Interaction results of EgDGAT1. The pathway in blue background is Kennedy pathway, and the purple box shows acyl editing pathways. The enzymes interacted with EgDGAT1 are painted in yellow.

EgDGATs (ex: DGAT1 in Fig.4-18) and EgPDAT1, in addition of being capable of interacting with EgLPATs from the Kennedy pathway, all interacted with most key enzymes of the editing pathway, including EgPDCT, EgAAPTs, and EgLPCATs. Therefore, our results provide evidence for physical interactions between synthesizing

and editing pathways, in addition to functional evidence provided by flux (Bates & Browse, 2012) and mutant (Bates *et al*., 2012) studies previously published.

Of course, capability for interaction does not necessarily mean formation of stable multi-enzyme complexes. Also, some experimental data point to possible caveats of yeast split-ubiquitin system. For example, same protein cloned in different vector allowing either C- or N-terminus fusion of ubiquitin moiety lead to similar results but with slight differences. It is often difficult to verify a given interaction by permuting bait and prey and a few weak interactions might sometimes appear questionable, even if they are reproducible using independent clones. At last, not all enzymes could function as bait and several did not work at all in the system, including most PA phosphatases. For these reasons, it is necessary to use other approaches and we initiated studies aimed at demonstrating complexes by blue native electrophoresis. The principle is to submit membranes to increasing concentration of mild detergent to find appropriate concentration to release complex from membrane without dissociating it. The mixture is then submitted to gel electrophoresis without detergent to preserve native conformation and Coomassie blue is used to make charges uniform so that separation is based on molecular mass only. Then, bands are analyzed by proteomic, either directly or after running the gel in a second dimension with SDS, to separate each complex into its protein constituents. We carried out preliminary experiments to obtain suitable membrane fractions and we were able to obtain from mesocarp light microsomal fractions that were significantly enriched in peptides from oil assembly enzymes.

In most cases, analysis of individual interaction networks was very consistent with what we know/suspect about oil, PC and editing pathways. EgLPPδ, a putative PA phosphatase expressed at similar levels in all tissue tested showed interactions with most enzymes using either PA or DAG as substrates including EgAAPT, EgCDP-DAGS, EgLPAT4, but not with any TAG-specific enzymes, suggesting it might be related to PC synthesis rather than TAG. Highest number of interactions was noted for central enzymes of the pathways such as EgAAPT, EgLPAT4 (Fig.4-19),

EgFAD2. LPAT is a common enzyme for both phospholipid and TAG synthesis. It (Fig.4-19) shows strong interactions with all enzymes in the figure, enzymes for activating acyl-CoAs (LACS4 and ACBP6), enzymes for common pathway (GPAT9 and PAP), enzymes for oil synthesis (DGAT and PDAT), and editing pathway (AAPT, LPCAT, FAD2 and PDCT). So this is consistent with its being involved in synthesis of both phospholipids and TAGs. Some interesting patterns lead us to consider certain enzymes as good candidates for further studies. A putative bifunctional wax ester synthase/DGAT, designated EgWSD1-like, shows interactions with most enzymes of PC synthesis and editing which would be consistent with a DGAT function and further analyses are described below (5.1 in Chapter 5). Also, an isoform of EgCCT shows, contrarily to other isoforms, interactions with TAG synthesizing enzymes and might be a good candidate for specifically regulating PC destined to TAG synthesis as it also shows higher transcript levels in mesocarp than in leaves. Also, patterns of gene expression point out to other interesting isoforms, such as EgLPCATc which is tremendously up-regulated during mesocarp ripening and almost specific for this tissue. Patterns for PA phosphatase isoforms have been much improved since Bourgis *et al*. (2011) study and they strongly suggest that further investigations are heavily required to assess impact of some these on oil synthesis, especially EgLPPβ.

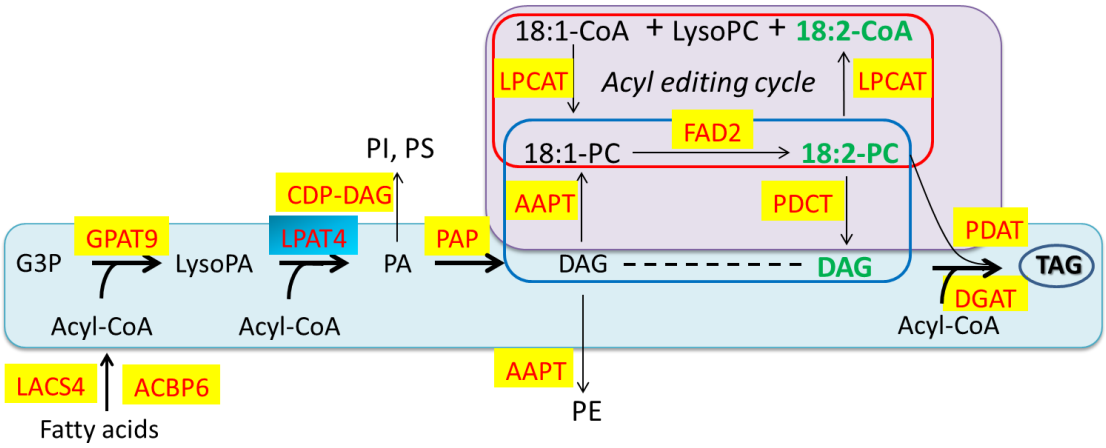

Fig.4-19 Interaction results of EgLPAT4. The pathway in blue background is Kennedy pathway, and the purple box shows acyl editing pathways. The enzymes interacted with EgLPAT4 are painted in yellow.

4.2 Cell fractionation studies from oil palm mesocarp

Membranes that contain important amounts of target enzymes are required to analyze successfully possible supra-molecular enzyme complexes by blue native electrophoresis. Therefore, it is important to develop an appropriate membrane purification protocol to obtain material that will be enriched in oil-assembly enzymes. I started to fractionate roughly frozen ripe mesocarp material using a post-10 000 *g* supernatant to isolate microsomal membranes by ultracentrifugation at 100 000 *g* for one hour. I started to assay DGAT activity from membrane and fat pad material using radiolabeled 18:1-CoA as substrate but had no success, while I could easily detect such activity from microsomal membranes isolated from frozen sunflower kernels using the same membrane fractionation protocol. I took the opportunity of fresh fruits availability to isolate new membranes but I was again unsuccessful to measure DGAT activity from membranes obtained from ripe fruits. It was only when I used unripe fruits that contained almost no oil in mesocarp that I could detect a weak but significant DGAT activity. Therefore, DGAT assay was not a good tool to track and characterize membrane fractions and I decided to detect the enzymes of interest by proteomics.

I used fresh fruits from a low lipase mutant to fractionate membranes, in order to avoid the oil being hydrolyzed during experimental procedures. Mesocarp slices were dissected and gently chopped with razor blades into a grinding buffer. Extract was then filtered successively through 200 µm and 40 µm nylon mesh and centrifuged at 500 *g* to remove debris. In these conditions, one may expect that organelles such as plastid should be intact, at least in part. The extract was centrifuged successively at 15 000 *g* (designated as sample A) and 100 000 *g* (Table 4-1). Pellets were gently re-suspended using a soft brush and centrifuged again to wash them and eliminate most contamination by soluble proteins. The huge fat pad, collected from post 15 000 *g* supernatant was also washed: it was mixed with sucrose and overlaid with buffer

that lacked sucrose and centrifuged again at 15 000 g; this centrifugation step led, in addition to floating fat pad, to a pellet which was collected. Centrifugation of post 15 000 *g* fat pad at 100 000 *g* led also to a pellet (designated as sample B) and a floating fat pad (designated as sample C), again collected. Centrifugation of 100 000 g supernatant at higher speeds (160 000 *g* for 1.5 hours) led to another pellet that was collected (designated as sample D). The samples from each step are listed in Table 4-1.

Table 4-1 Fractions from each centrifugation step

Sample	Centrifugation step
A	Pellet from centrifugation of 500 *g* supernatant at 15 000 *g*, 20min
B	Pellet from centrifugation of washed fat pad at 100 000 *g*, 1hour
C	Fat pat from centrifugation of washed fat pad at 100 000 *g*, 1hour
D	Pellet from centrifugation of 100 000 *g* supernatant at 160 000 *g*, 1.5hours

Table 4-2 Peptides of oil-assembly enzymes from fractions of palm mesocarp

Sample	A	B	C	D
EgLPAT2	1	4	2	6
EgLPCATa	0	0	1	2
EgLPCATc	0	0	0	4
EgCDP-DAGa	0	1	0	2
EgCDP-DAGb	0	1	0	1
EgGPAT	0	1	2	3
EgWSD1-like	1	1	2	3
EgDGAT1	0	1	0	1
EgDGAT2-2	0	0	0	1
Total	2	9	7	23
Total detected peptides	20548	25932	22910	24700
Percentage (%)	0.01	0.03	0.03	0.09

Samples were sent to proteomic facility to identify all enzymes of interest in these fractions (Table 4-2). As expected, few peptides could be identified in 15 000 *g* pellet (A), which should contain intact organelles and possibly heavier endomembrane fractions only. As expected, several peptides were found in microsomal fraction (B, 100 000 *g* pellet). Interestingly, largest number of peptides was in post 160 000 *g*

pellet. It has been shown earlier (Lacey & Hills, 1996) that TAGs were synthesized in lighter membranes of rapeseed. This might also be the case in oil palm mesocarp and this property will be of great help to obtain membrane fractions enriched in oil synthesizing enzymes for future blue native gel electrophoresis experiments in order to provide evidence for the possible existence of complexes. Interestingly, fat pad contains also peptides that belong to oil synthesizing enzymes. It might be that some of these light fractions are trapped with lipid droplets, or that lipid droplets themselves carry oil-synthesizing enzymes. These results, together with those expected from different other cell fractionation experiments from oil palm mesocarp will be of great help to optimize a protocol to obtain greatest possible enrichment.

Chapter 5. Yeast expression studies

5.1 Yeast expression studies of three palm DGAT isoforms

5.1.1 Triacylglycerol synthesis after expression in yeast

All three Oil Palm DGAT (EgDGAT1, 2-1 and 2-2) were expressed into TAG-deficient yeast mutant strain under the control of inducible GAL promoter. After induction, yeast cells were recovered and lipids were extracted and loaded onto a TLC plate. Results (Fig.5-1) indicate that all three EgDGATs lead to accumulation of TAGs at levels similar to that of wild type strain, thus complementing the mutation. However, they have no impact on steryl ester content.

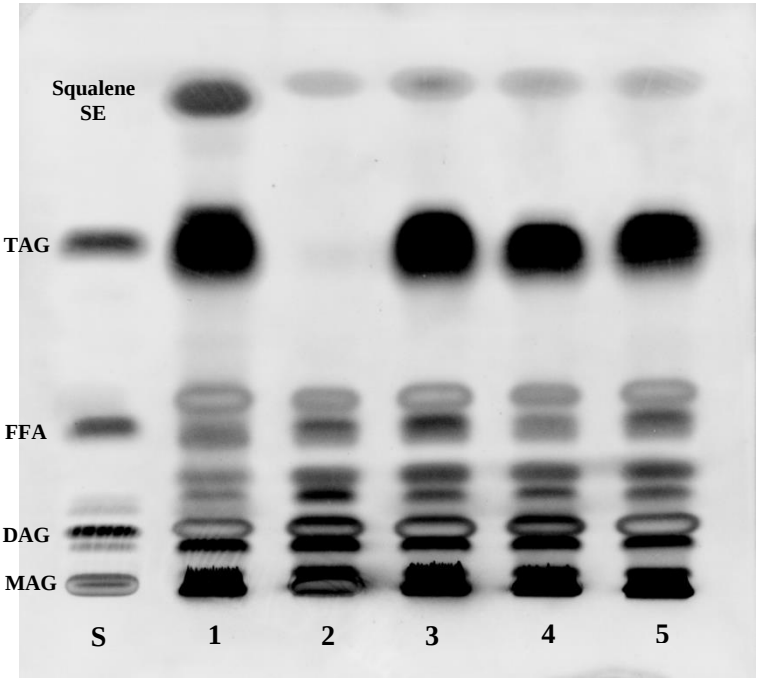

Fig.5-1 Analysis of neutral lipid content by TLC using hexane/diethyl ether/acetic acid (80:20:2) for development. S: 2.5 μg each (Oleic acid, monoolein, diolein, triolein). Lipid from 10 OD units of each yeast culture was loaded on TLC plate. 1: yeast strain SCY62 transformed with empty pESC-URA vector; 2: yeast strain H1246 transformed with empty pESC-URA vector; 3: yeast strain H1246 transformed with *EgDGAT1*-pESC-URA construct; 4: yeast strain H1246 transformed with *EgDGAT2-1*-pESC-URA construct; 5: yeast strain H1246 transformed with *EgDGAT2-2*-pESC-URA construct. MAG: monoacylglycerol; DAG: diacylglycerol; TAG: triacylglycerol; FFA: free fatty acid; SE: sterol ester.

5.1.2 Quantification of TFA, TAG, DAG, PC in complemented yeast

Lipid samples from yeast strains were mixed with appropriate amount of

appropriate lipid standards and run on TLC plates to isolate TAGs and DAGs (migration for neutral lipids) and PC (migration for polar lipids). FAMES were generated and ran on GC-MS and absolute quantification was done based on GC results. In addition, we quantified total fatty acid content of yeast strains and results are shown in Fig.5-2. PC can be taken as representative of total polar lipids as we found that PC % share of polar lipids did not appear to vary in strains tested (data not shown). WT parent and mutant strains contain about same amount of total lipids. Lack of TAGs in mutant strain appears to be fully compensated for by increase in polar lipid content (about 50% PC increase). Depending on EgDGAT isoform, TAG levels are slightly lower (EgDGAT2-1) or slightly higher (10 and 35% for EgDGAT2-2 and EgDGAT1, respectively) in complemented strains than in WT. Interestingly, PC levels remain slightly higher than in WT, with differences depending on EgDGATs, suggesting that palm DGATs do not fully complement side effects of mutation. Expression of EgDGAT1 leads to higher TAG levels than in WT while retaining mutant's PC levels. DAG amounts were slightly lower in strain complemented with each EgDGAT2s.

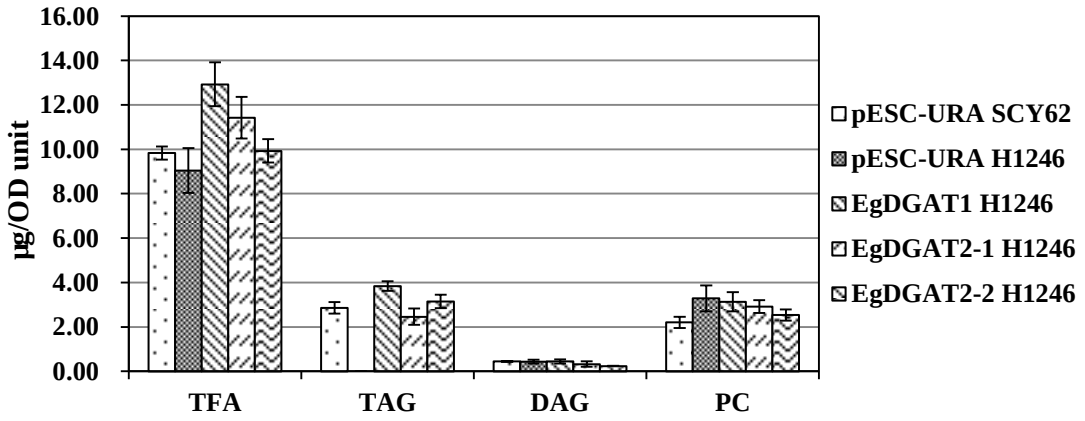

Fig.5-2 Amounts of different lipids in yeast mutant strain complemented with oil palm EgDGATs. TFA: total fatty acids determined by direct transmethylation from yeast pellets. TAG, DAG and PC were determined by gas chromatography of methyl esters obtained from spot scratched after separation from TLC plates. At least 3 replicates were used for analysis.

5.1.3 Effect on fatty acid composition

5.1.3.1 Total fatty acids

The total fatty acid composition of yeast strains showed very similar profiles except some difference on the levels of C16:0 and C18:1 (Fig.5-3). Mutant strain contains less 16:0 and more 18:1 than wild type parent strain. With this respect, fatty acid profile of mutant transformed with EgDGAT2-1 resembles mutant and the other two (EgDGAT1 and EgDGAT2-2) wild type parent strain. C26:0 levels are lower in mutant (complemented or not) than in wild type. Thus, there again, subtle associated phenotypes of mutant may not be fully complemented by some palm DGATs.

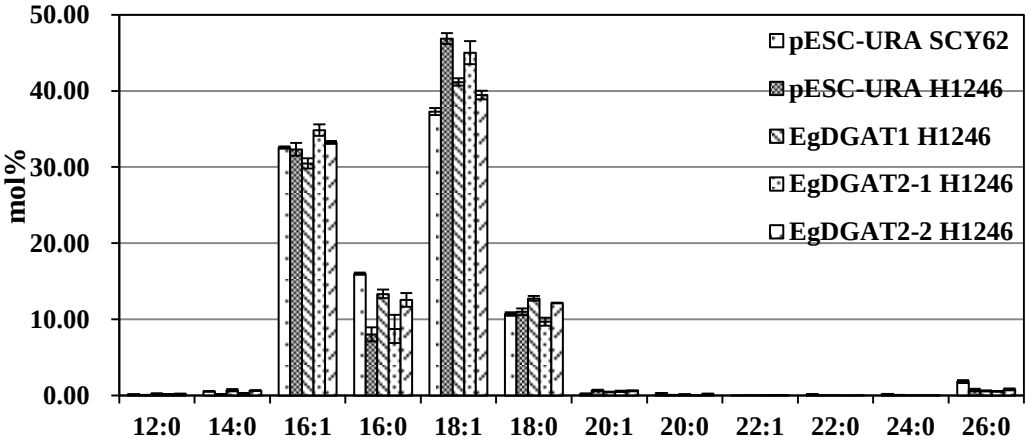

Fig.5-3 Total fatty acid composition of each yeast transformant. Transmethylation was carried out on yeast cell pellets and FAMEs analyzed by GC-MS.

5.1.3.4 Phosphatidylcholine

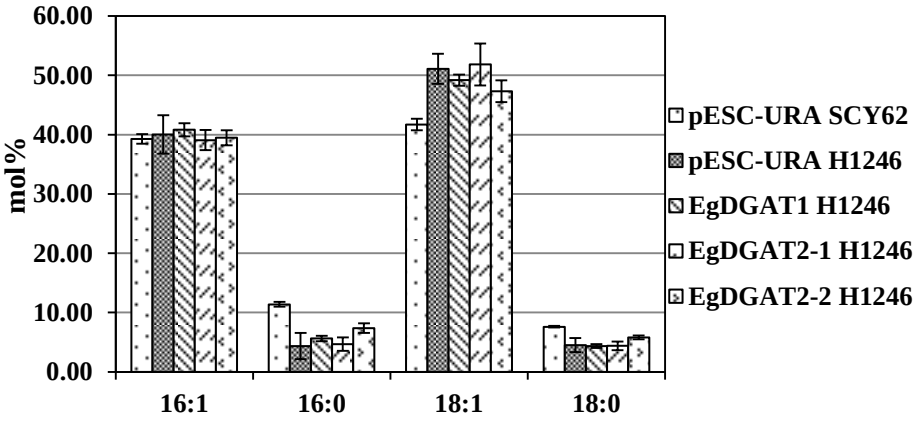

Fig.5-4 Phosphatidylcholine fatty acid composition of yeast strains. Polar lipids were separated by TLC and the phosphatidylcholine spot was scraped from the plate, transmethylated and FAMEs analyzed by GC-MS.

Wild type parent strain SCY62 transformed with empty vector contains more saturated fatty acids (C16:0 and C18:0) and less C18:1 than yeast mutant strain H1246

(Fig.5-4). Transformation of the mutant with oil palm DGAT does not appear to complement much this PC composition phenotype. We noted earlier that WT absolute amounts of PC were neither restablished by palm DGATs.

5.1.3.5 Triacylglycerol

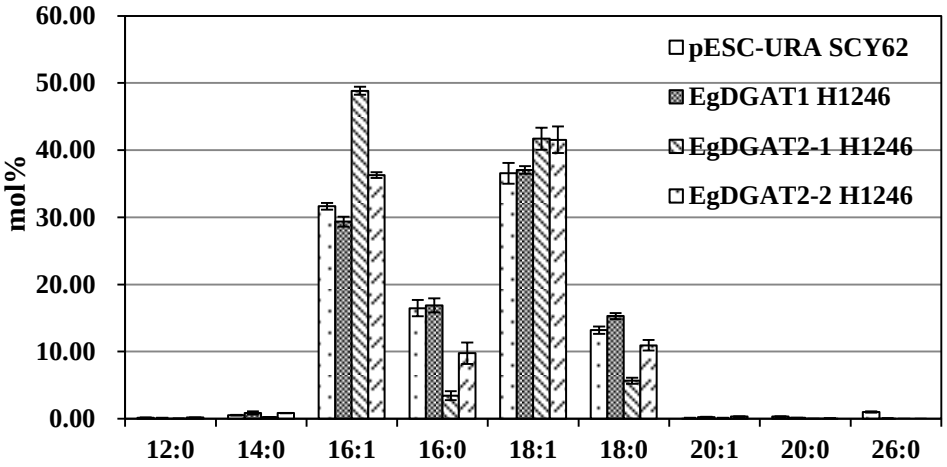

Fig.5-5 Fatty acid composition of triacylglycerol from different yeast strains. Neutral lipids were separated by TLC and the TAG spot was scraped from the plate, transmethylated and FAMES analyzed by GC-MS. Mutant TAGs were not analyzed as we verified once that endogenous TAGs in mutant was below 0.025 μg/OD unit and probably due to fatty acid contamination from TLC plate.

TAG fatty acid composition of complemented mutant differs markedly depending on the DGAT isoform used. This suggests that these enzymes show different selectivities towards acyl-CoA substrates.

5.1.3.6 Diacylglycerol

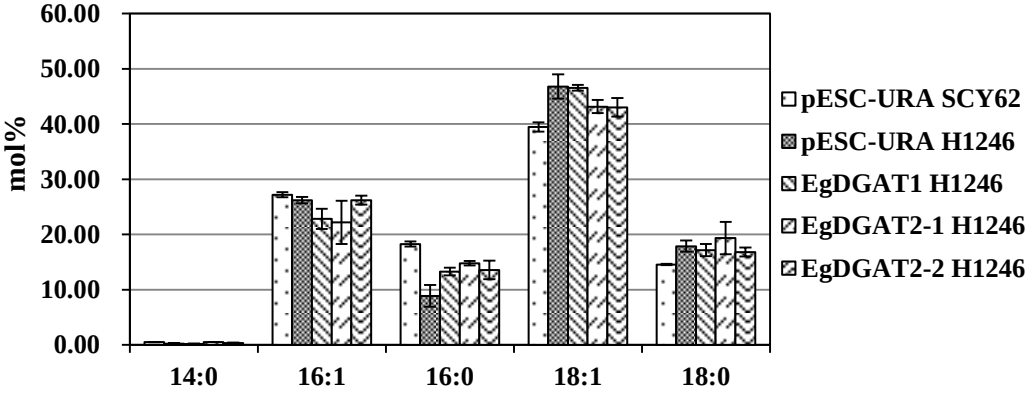

Fig.5-6 Fatty acid composition of diacylglycerol from different yeast strains. Neutral lipids were separated by TLC and the DAG spot was scraped from the plate, transmethylated and FAMEs analyzed by GC-MS.

DAG fatty acid profile is similar in all strains and composition does not seem to show any significant difference. This suggests that DAG substrate available to EgDGATs is similar in all strains, which conforts the hypothesis we made on distinct substrate selectivities of palm DGATs.

5.1.4 Conclusions

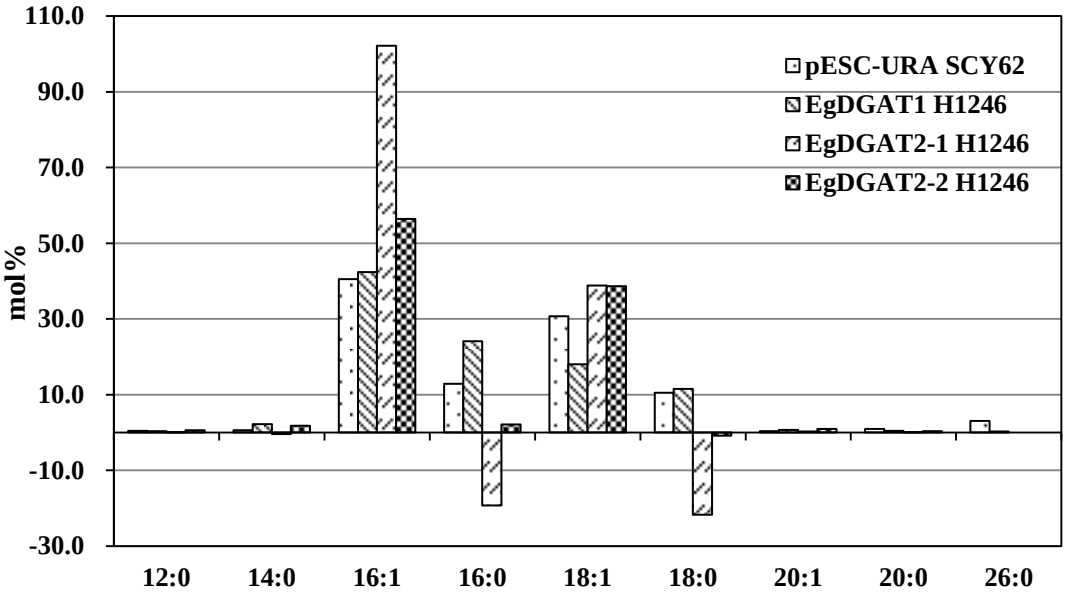

Fig.5-7 Theoretical calculation of fatty acid composition at *sn*-3 position. For each fatty acid and each yeast strain, we applied following formula, assuming that DGATs are not selective towards DAG fatty acid composition: % FA at *sn*-3 = ((% FA in TAG- 2/3(% FA in DAG)) x 3.

If one hypothesizes that EgDGATs are not specific with regards to DAG fatty acid content, then they should take any DAG species available equally and these are about the same in all strains analyzed (see 5.1.3.6). Therefore, we made the following calculations: DAG fatty acid composition, affected by a 2/3 coefficient, was subtracted from TAG fatty acid composition; fatty acid remaining (1/3) corresponds then to the fatty acid inserted at *sn*-3 position by DGAT. Results from this calculation are shown in Fig.5-7.

Results for EgDGAT1 are very close to that from WT parent strain; because putative fatty acid profile at *sn*-3 is not very different from total fatty acid content of yeast, it is likely that EgDGAT1 is not very selective towards yeast major acyl-CoA; it

just takes less 26:0 than endogenous yeast DGATs do.

Results from EgDGAT2-2 suggest that the enzyme shows a significant preference for unsaturated fatty acids (96%) versus saturated ones (4%). If we still assume that it is not selective for DAG substrates, then it means that it is almost exclusively selective for unsaturated acyl-CoAs.

Results from EgDGAT2-1 show an excess of unsaturated fatty acids (140%) at *sn*-3 that this cannot be explained alone by specificity for unsaturated acyl-CoA only but EgDGAT2-1 must have a preference for unsaturated DAG species (or may access in yeast to sub-pools of DAGs with higher unsaturation levels).

Our results are consistent with data recently obtained by Aymé *et al*. (2014) and (2015), who showed that *Arabidopsis* DGAT 1 and 2 had preferences for 16:0 and 16:1 in yeast, respectively.

5.2 EgWSD1-like expression in mutant yeast H1246

EgWSD1-like encodes a homologue of the Wax synthase/DGAT bifunctional enzyme from bacteria *Acinetobacter calcoaceticus* ADP1 (Kalscheuer *et al*., 2004; Kalscheuer & Steinbüchel, 2003). The expression profile of *EgWSD1-like* during fruit ripening is similar to that of EgDGAT1 and transcripts are higher in oil palm mesocarp than in date palm mesocarp, a tissue that stores no oil (Bourgis *et al*., 2011). Moreover, EgWSD1-like shows numerous interactions with other enzymes involved in oil synthesis in Y2H membrane system (see 4.1.3.2C in Chapter 4), all this making it a possible candidate with DGAT function for oil synthesis in oil palm mesocarp. To gain insights into its *in vivo* function, *EgWSD1-like* was expressed in the mutant yeast H1246, which is devoid of all four genes responsible for neutral lipids synthesis (Sandager *et al*., 2002). The result showed that EgWSD1-like is unable to restore TAG biosynthesis in mutant yeast H1246, but its expression led to synthesis of fatty acid isoamyl esters (FAIEs) with saturated long-chain and very-long-chain fatty acids (18 to 26), and to wax esters when fatty alcohols were provided, either by feeding yeast cultures or by co-expression of the fatty acyl reductase *TtFAR* from *Tetrahymena thermophile*. TtFAR was previously shown to produce 16OH and 18OH when expressed in yeast (Dittrich-Domergue *et al*., 2014). Our feeding experiment results indicated that EgWSD1-like shows preference for medium fatty alcohols (14 and 16), while coexpression with TtFAR showed that EgWSD1-like prefers saturated fatty acids (14 to 26) to unsaturated fatty acids. Interestingly, very similar results were found concerning a WSD1-like enzyme (PhWS) from *Petunia hydrida* (King *et al*., 2007) that shows high sequence identity to our oil palm enzyme. However, because levels of EgWSD1-like products are low in transformed yeast, further investigations might be required. Either this can be explained by expression very low levels of functional enzyme, or may be we did not find the true product of the enzyme, because substrate(s) are not available in yeast.

The whole EgWSD1-like study is presented on next page as a draft manuscript to be submitted for publication soon.

Characterization and heterologous expression of a wax ester synthase from oil palm (*Elaeis guineensis*) mesocarp in *Saccharomyces cerevisiae*

Yuan Y., Arondel V. and Domergue F.

Laboratoire de Biogenèse Membranaire, CNRS - Univ. Bordeaux - UMR 5200, Bâtiment A3 - INRA Bordeaux Aquitaine, 71 avenue Edouard Bourlaux - CS 20032, 33140 Villenave d'Ornon

Corresponding author: frederic.domergue@u-bordeaux.fr

Author contributions: V.A. and F.D. designed research; Y.Y. and F.D. performed research; Y.Y. and F.D. analyzed data; and Y.Y. and F.D. wrote the paper.

Abstract

Oil palm (*Elaeis guineensis*) can accumulate up to 88% oil in fruit mesocarp. A previous transcriptome study of oil palm fruits indicated that a gene coding for a putative bifunctional wax synthase (WS) / diacyl glycerol acyl transferase (DGAT), designated as *EgWSD1-like*, was specifically expressed in mesocarp during oil accumulation. In the present study, the full-length cDNA of *EgWSD1-like* was isolated, analyzed, and characterized by heterologous expression in the mutant yeast H1246, which is devoid of neutral lipid synthesis. Expression of *EgWSD1-like* could not restore TAG synthesis, but resulted in the synthesis of fatty acid isoamyl esters (FAIEs) with saturated long-chain and very-long-chain fatty acids. In the presence of exogenously supplied fatty alcohols, EgWSD1-like was able to produce wax esters, indicating that *EgWSD1-like* code for an acyltransferase with WS, but not DGAT, activity. The complete wax ester biosynthetic pathway was finally reconstituted in yeast by coexpression *EgWSD1-like* with a fatty acyl reductase from *Tetrahymena thermophila*. Altogether, our results indicated that the expression of *EgWSD1-like* in yeast produced wax esters, and that EgWSD1-like preferentially used medium chain fatty alcohols and saturated very-long chain fatty acids as substrates.

Keywords: *Elaeis guineensis*, wax ester synthase, *Saccharomyces cerevisiae*

1. Introduction

Oil in the form of triacylglycerols (TAGs) is a very common energy-dense storage compound found in all eukaryotes, and primarily accumulated in many plant seeds and fruits (Bates, 2016; Xu & Shanklin, 2016). The mesocarp of oil palm (*Elaeis guineensis*) fruits can accumulate up to 88% (dry weight) oil (Bourgis *et al*., 2011), with TAGs primarily made of 16:0 and 18:1 fatty acids. In addition, the kernel of palm fruits produces about 50% (dry weight) oil, with TAGs mostly composed of saturated medium chain fatty acid (MCFAs), especially lauric acid (12:0) (Corley & Tinker, 2003). Oil palm therefore represents one of the most productive oil crops in the plant kingdom. The percentage of TAG in oil palm is over 95%, the rest being composed of glycerolipids and phospholipids as well as sterols, triterpene alcohols, tocopherol, carotenoids and chlorophylls (Sambanthamurthi *et al*., 2000).

All modern oilseed crops, like rapeseed (*Brassica napus*) or sunflower (*Helianthus annuus*), accumulate storage lipids in the form of TAGs in their seeds. TAGs, which are composed of three fatty acids esterified to glycerol, will be hydrolyzed upon germination to produce the energy necessary to fuel post-germinative seedling growth before photosynthesis becomes effective (Baud & Lepiniec, 2010; Xu & Shanklin, 2016). In fruits, TAGs are believed to primarily serve as an effective strategy for attracting animal feeders in order to increase seed dispersal and plant habitats. Wax esters (WEs, esters of primary fatty alcohols and fatty acids) represent another kind of storage lipids synthesized in a wide range of microorganisms (Wältermann *et al*., 2007), such as the bacteria *Acinetobacter calcoaceticus* ADP1 (Kalscheuer & Steinbüchel, 2003) or the unicellular phytoflagellate protist *Euglena gracilis* (Teerawanichpan & Qiu, 2010). Jojoba (*Simmondsia chinensis*), a desert shrub from the south-western areas of North America, represents the only plant known to accumulate WEs rather than TAGs as seed storage lipids (Lardizabal *et al*., 2000; Ohlrogge *et al*., 1978). In contrast, the presence of WEs in the cuticle layer covering the aerial parts of land plants has been reported in many species (Lee & Suh, 2015).

This protective layer plays a critical role in protection from dessication, pathogens and insects herbivores (Bernard & Joubès, 2013).

The biosynthesis of TAGs has been widely studied in oilseeds, especially in *Arabidopsis thaliana*, and was shown to rely on different overlapping pathways. The well-known and simplest way to produce TAGs is known as the Kennedy pathway, where three fatty acids esterified to CoA are sequentially esterified to each position of the glycerol backbone. First, the *sn*-1 and *sn*-2 positions of glycerol-3-phosphate are respectively acylated by the acyl-CoA:glycerol-3-phosphate acyltransferase (GPAT) and the acyl-CoA:lyso-phosphatidic acid acyltransferase (LPAT), producing LPA and subsequently phosphatidic acid (PA). The phosphate of PA is then hydrolyzed by a PA phosphatase (PAP), yielding diacylglycerol (DAG), which is further acylated by the acyl-CoA:diacylglycerol acyltransferase (DGAT), leading to TAGs. DGAT is nevertheless the unique enzyme of the Kennedy pathway that is specific for TAG biosynthesis, since GPAT, LPAT and PAP are also involved in phospholipids biosynthesis. Similarly, PA and DAG are important intermediates for the production of both TAGs and phospholipids. Recent studies showed that phosphatidylcholine (PC) plays a central role in TAG biosynthesis, being the site of acyl-editing as well as the main provider of DAG backbones, suggesting that TAG assembly relies on a very complicated metabolic network rather than uniquely on the Kennedy pathway (recently reviewed in Bates, 2016).

The WE biosynthetic pathway has first been elucidated in jojoba (Lardizabal *et al*., 2000; Metz *et al*., 2000). It relies on two enzymatic activities catalyzed by the fatty acyl-CoA reductase (FAR) and the acyl-CoA:fatty alcohol acyltransferase or WE synthase (WS). FAR is responsible for the reduction of acyl-CoAs to fatty alcohols (Metz *et al*., 2000), while WS transfers the acyl group of another acyl-CoA onto the former fatty alcohol, yielding WEs (Lardizabal *et al*., 2000). Following this pioneering work in jojoba, the WE biosynthetic pathway has been characterized in many other species like in *Arabidopsis* or *Euglena*, a unicellular phytoflagellate which accumulates medium-chain WEs (Rowland *et al*., 2006; Li *et al*., 2010;

Teerawanichpan & Qiu, 2010). In *Arabidopsis*, both FAR and WS enzymes are associated with membranes, strongly suggesting that WEs are synthesized within the endoplasmic reticulum (ER) (Kunst & Samuels, 2003; Kunst & Samuels, 2009).

Interestingly, another type of WS, which exhibits both WS activity and DGAT activity has been identified in the bacteria *Acinetobacter calcoaceticus* ADP1 (Kalscheuer & Steinbüchel, 2003). This WS/DGAT bifunctional enzyme is unrelated to the other known WS or DGAT proteins, and is essential for the production of both TAGs and WEs in *A. calcoaceticus* (Kalscheuer & Steinbüchel, 2003). Heterologous expression of this bifunctional enzyme in a yeast mutant strain devoid of storage lipid biosynthesis restored TAGs (H1246), but not steryl esters, and yielded fatty acid ethyl esters (FAEEs) and fatty acid isoamyl ester (FAIEs) in the absence of long-chain (LC) fatty alcohols (Kalscheuer *et al*., 2004). The genome of *Arabidopsis* contains 11 sequences that are homologous to the *A. Calcoaceticus WS/DGAT*, and only one of them, *AtWSD1*, has so far been characterized. AtWSD1 was shown to play a key role in WE synthesis, using LC and very-long-chain (VLC) primary alcohols together with palmitoyl-CoA to production the WEs found in *Arabidopsis* stem waxes (Li *et al*., 2008). Interestingly, if *in vitro* assays showed that the WS activity of AtWSD1 was approximately 10-fold higher than its DGAT activity; its expression in the yeast mutant strain H1246 produced WEs, but not TAGs (Li *et al*., 2008).

In our previous study (Bourgis *et al*., 2011), we identified a gene coding for a bifunctional WS/DGAT protein, designated as *EgWSD1-like*, which was highly expressed during fruit maturation in oil palm (*Elaeis guineensis*) mesocarp but not in date palm (*Phoenix dactylifera*), a closely related species that accumulates almost exclusive sugars. In the present paper, we isolated the full-length cDNA of *EgWSD1-like* from oil palm mesocarp, and characterized the function by heterologous expression in yeast. Our data suggest that this enzyme may be involved in WE synthesis with preference towards saturated LC and VLC acyl-CoAs and medium chain fatty alcohols (14OH and 16OH).

2. Materials and methods

2.1 Materials

Oil palm fruits were collected at CEREPAH La Dibamba (Cameroon) and immediately frozen in liquid nitrogen upon harvest. Mesocarp was fully and quickly dissected from individual frozen fruit, ground in liquid nitrogen to a fine powder, and stored at -80 ℃ until used. All chemical reagents were from Sigma-Aldrich unless indicated otherwise.

2.2 RNA extraction and gene cloning

Total RNA from oil palm mesocarp (22 and 23 weeks after pollination) was extracted using the method described by Chirgwin *et al*. (1979). First-strain cDNA synthesis was carried out using SuperScript™ II reverse transcriptase (Invitrogen) according to manufacturer's instructions. The gene *EgWSD1-like* was then amplified using specific primers and the Q5 high-fidelity DNA polymerase (New England Biolabs) following manufacturer's protocol. The sequences of the forward and reverse primers containing specific restriction sites (indicated by underlines and brackets) were as follows: 5'-GAAGATCTATGGCTTCCAGAAGAAGGC-3' (*Bgl*II) and 5'-GGGGTACCCTATGATCTTGGAATTGCATCCT-3' (*Kpn*I). The PCR product was gel-purified, digested with *Bgl*II and *Kpn*I, and cloned into the multiple cloning site 2 of pESC-URA, which had been digested by *BamH*I and *Kpn*I. The positive construct *EgWSD1-like*-pESC-URA was verified by sequencing (Beckman Coulter Genomics, GENEWIZ). For co-expression experiments, *TtFAR* from *Tetrahymena thermophila* (Dittrich-Domergue *et al*., 2014) was cloned using the same strategy in the multiple cloning site 1 of pESC-URA and *EgWSD1-like*-pESC-URA using 5'-ATAAGAATGCGGCCGCATGGGAAAGGTTTTCCAATTC-3' (*Not*I) and 5'-GGACTAGTTTACTTGAATGGCTTTCCAGAAG-3' (*Spe*I) as forward and reverse primers, respectively. Positive constructs were confirmed by sequencing.

2.3 Bioinformatic analysis

The translated protein sequence from *EgWSD1-like* was analyzed by protein BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). HMMTOP (Tusnády & Simon, 2001; http://www.enzim.hu/hmmtop/html/submit.html) and Phobius (Käll *et al*., 2004; http://phobius.sbc.su.se/) were used for analyzing the protein topology. The phylogenic relationship of EgWSD1-like with WSs and acyl-transferases from other species was carried out using MEGA5 (Tamura *et al*., 2011).

2.4 Yeast transformation and growth conditions

The yeast (*Saccharomyces cerevisiae*) strain H1246 (*MATα ADE2-1 can1-100 ura3-1 are1-Δ::HIS3 are2-Δ::LEU2, dga1-Δ::KanMX4 lro1-Δ::TRP1*; Sandager *et al*., 2002), which is devoid of neutral lipid synthesis, was used in this study. pESC-URA and *TtFAR*-pESC-URA were respectively used as controls for *EgWSD1-like*-pESC-URA and *TtFAR-EgWSD1-like*-pESC-URA. All constructs were separately transformed into the yeast strain H1246 using the polyethylene glycol/lithium acetate method. Transformants were selected on synthetic complete medium lacking uracil (SC-U) supplemented with 2% (w/v) glucose. Three independent positive clones of each transformation were used for expression studies. Pre-cultures were made in 2 mL SC-U medium containing 2% (w/v) raffinose and were grown overnight at 30 ℃. For inducing the expression of the transgene(s), 100 µL of pre-culture was inoculated into 20 mL SC-U medium containing 2% (w/v) galactose in 250 mL flask. Flasks were incubated under continuous agitation (250 rpm) for 48 h at 30 ℃, and the OD_{600} of the yeast cultures was measured before harvest. When fatty alcohols were exogenously supplied, 500 µM of fatty alcohol was added to the growing cells 24 hours after induction and yeast cells were harvested 48 hours later.

2.5 Lipid extraction and separation

Usually, 2 (about 25 OD units) and 16 (about 200 OD units) mL of culture (one OD unit is the amount of yeast cells that gives an OD of 1 in 1 mL) were respectively used for total acyl-chain profiling and the preparation of lipid extracts. Cells were collected

by centrifugation in 8 mL glass tubes with Teflon caps, pellets were washed once with distilled water, and either immediately used for lipid extraction or frozen at -20 ℃.

For total fatty acyl-chain profiling, 40 µg of heptadecanoic acid (C17:0) was added to the yeast pellets and direct transmethylation with 2 mL of methanol containing 2.7% (v/v) H_2SO_4 was performed for 2 hours at 80 ℃. After cooling down, 3 ml of hexane was added for extracting fatty acid methyl esters (FAMEs) and fatty alcohols, and 2 ml of NaCl (2.5%, w/v) was used for phase separation. Following phase separation by centrifugation, the hexane phase was transferred to a fresh tube and washed with 2 mL of NaCl (0.09%, w/v) dissolved in 100 mM Tris (pH8.0). Following phase separation by centrifugation, the hexane phase was collected in a fresh tube and evaporated under a gentle stream of nitrogen. The residue was dissolved in 150 µL of *N*,*O*-bis(trimethylsilyl) trifluoroacetamide with trimethylchlorosilane (BSTFA), and free hydroxyl groups were derivatized at 110 ℃ for 20 min. The surplus BSTFA was evaporated under nitrogen and the samples was dissolved in 200 µL of hexane for GC-MS analysis.

For preparing the lipids extracts, 2 mL of chloroform/methanol (1:1; v/v) and glass beads (180 µm; about 3 mL, up to meniscus) were added to the yeast pellets (about 300 µL) and tubes were vigorously shaken 4 times for 30s. Following centrifugation, the organic phase was collected, and beads were rinsed thoroughly once with 2mL of chloroform/methanol (2:1; v/v) and once with 2 mL of chloroform. These two washes were combined with the initial extract, and 2 mL of NaCl 2.5% (w/v) was then added to allow phase separation. Following centrifugation, the organic phase was collected while the aqueous phase was re-extracted with 2 mL chloroform. This organic phase was combined with the previous one and dried down. Finally, lipid extracts were dissolved in 250 µL chloroform. For WE quantification, 8 µg heptadecanyl heptadecanoate (WE 17OH-17:0; NU-CHEK-PREP) was added as internal standard before starting the lipid extraction. Neutral lipids were isolated using solid phase extraction as described in Dittrich-Domergue *et al*. (2014), dried under nitrogen and resuspended in 200 µL chloroform for GC-MS analysis.

2.6 TLC and GC-MS analyses

For thin layer chromatography (TLC) analysis, total lipids equivalent to 5 or 10 OD units and lipid standards were loaded on HPTLC Silica Gel 60 plates (Merck) pre-washed with chloroform/methanol (1:1; v/v). Hexane/diethyl ether/acetic acid (80:20:1, v/v/v) and hexane/diethyl ether/acetic acid (90:7.5:1, v/v/v) were used as solvent systems for separating neutral lipids and isolating isoamyl esters, respectively. Following separation, lipids were visualized using 0.05% (w/v) primuline in acetone/water (4:1; v/v). Isoamyl esters were extracted from silica using hexane and diethyl ether (1:1; v/v). Oleic acid, monoolein, diolein, triolein and WE purchased from NU-CHEK-PREP were used lipid standards.

GC-MS analyses were performed using an Agilent 6850 gas chromatograph equipped with an HP-5MS column (30 m x 0.25 mm x 0.25 mm) and an Agilent 5975 mass spectrometric detector (70 eV, mass-to-charge ratio 50-750), with helium (1.5 $mL.min^{-1}$) as carrier gas. For the resolution of FAMEs, the initial temperature of 50 ℃ was held for 1 min, increased at 25 ℃ min^{-1} to 150 ℃, held for 2 min at 150 ℃, then increased at 10 ℃ min^{-1} to 320 ℃, and held for 6 min at 320 ℃. Separation of the WEs was achieved with the same program except that the final temperature (320 ℃) was held for 21 minutes.

3. Results

3.1 A putative bifunctional WS/DGAT acyltransferase from oil palm is steadily expressed in mesocarp during fruit ripening

Using the transcriptomic data from developing oil palm mesocarp (Bourgis *et al*., 2011), we identified a transcript, thereafter referred to as *EgWSD1-like*, which expression profile was very similar to that of *EgDGAT1*, one important enzyme involved in TAG assembly. Both transcripts showed high levels of expression during whole fruit development, whereas their expressions were very low in leaf (Fig.1;

Bourgis *et al*., 2011). For both genes, expression levels were high early in development (15WAP), slightly decreased during the following 4 weeks (till 19WAP), and then became highest during the fruit ripening phase. The expression level of *EgWSD1-like* was systematically 50 to 60% lower than that of *EgDGAT1* all along fruit development, except at 23WAP where both transcripts were nearly as abundant. However, *EgDGAT1* was highly expressed in kernel as in mesocarp at 17 or 19 WAP, while *EgWSD1-like* was barely expressed in kernel (Fig.1).

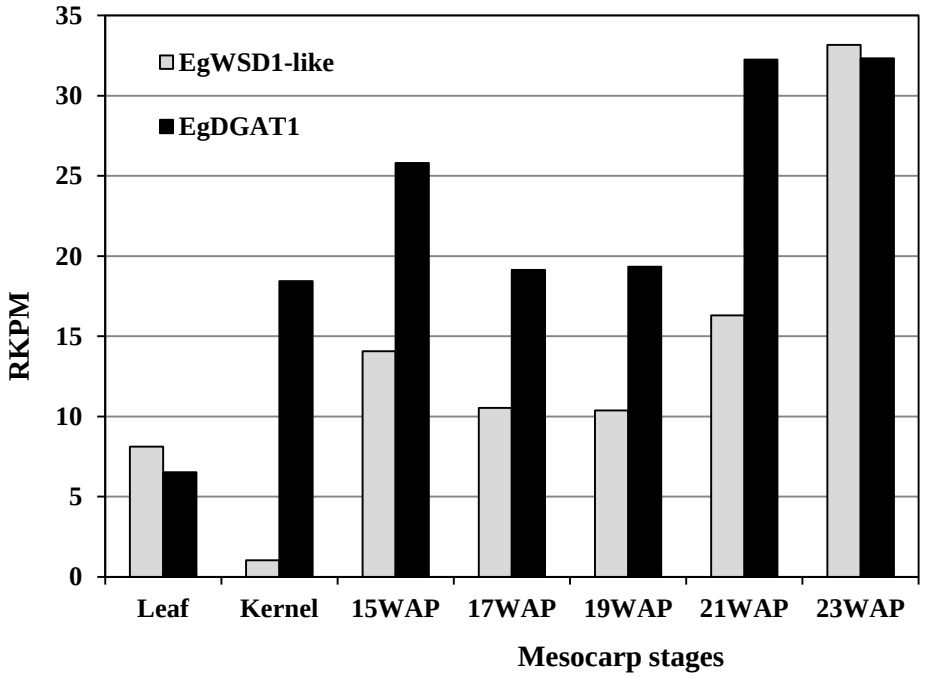

Fig.1 Temporal expression patterns of *EgWSD1-like* and *EgDGAT1* in oil palm leaf, kernel and mesocarp at five development stages. Expression levels were obtained using the transcriptome data from Bourgis *et al*. (2011). WAP, weeks after pollination. RPKM: reads per kilobase per million reads.

Using total RNA from the mesocarp of ripening fruits (22 and 23WAP) and gene-specific primers, we isolated the full-length cDNA (1548 bp) of *EgWSD1-like*. The open reading frame encoded for a protein of 515 amino acids with a molecular mass of 56.99 kD and a pI of 6.62. Bioinformatic analysis of EgWSD1-like indicated that it contains the conserved domain of WS-like acyl-CoA acyltransferase (pfam03007; amino acids 139-299) and the putative active-site motif (^{178}HH*XXX*DG^{185}), which was demonstrated to be essential for the catalytic activity of acyl-transferases involved in WEs and TAGs biosynthesis (Kalscheuer & Steinbüchel,

2003; Fig.S1). EgWSD1-like belonged to the bifunctional wax-ester synthase/ diacylglycerol acyltransferase (WS/DGAT) superfamily, and had 37% identity / 56% homology with AtWSD1, a WS/DGAT from *Arabidopsis thaliana* involved in stem WE biosynthesis (Li *et al*., 2008). Like the other members of this protein family, EgWSD1-like was predicted to be an integral membrane protein with one (according to Phobius; Käll *et al*., 2004) or up to 4 (according to HMMTOP; Tusnády & Simon, 2001) putative transmembrane domains (Fig.S1).

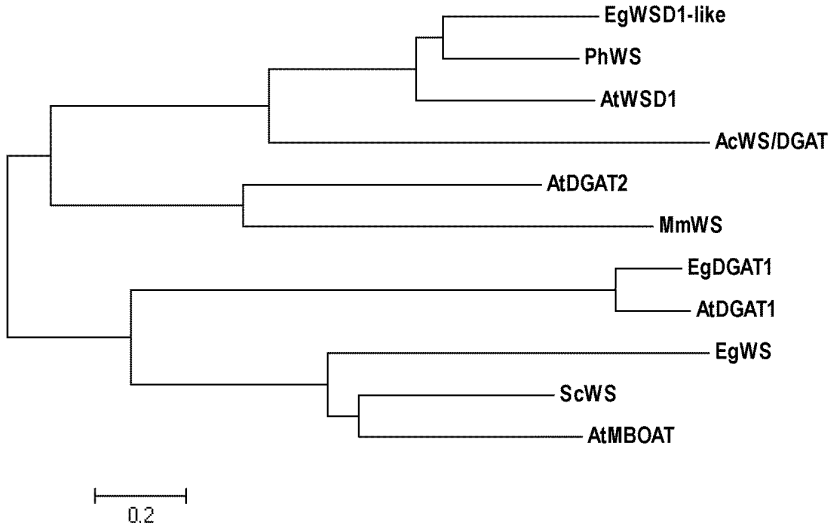

Fig.2 Phylogenetic analysis of EgWSD1-like and WS and/or DGAT from other species. Phylogenetic tree was generated by the neighbor-joining method in MEGA5. The GenBank accession numbers of the sequences showed are as follows: *Arabidopsis thaliana* AtWSD1 (NP_568547.1), *Simmondsia chinensis* ScWS (AAD38041.1), *Petunia hydrida* PhWS (AAZ08051.1), *Mus musculus MmWS* (NP_808414.2), *Acinetobacter calcoaceticus* ADP1 AcWS/DGAT (AAO17391.1), *Euglena gracilis* EgWS (ADI60058.1), *Arabidopsis thaliana* AtMBOAT (NP_174708), EgDGAT1 (*Elaeis guineensis*, unpublished), *Arabidopsis thaliana* AtDGAT1 (CAB45373.1) and AtDGAT2 (Q9ASU1.1).

Phylogenetic analysis of EgWSD1-like with various WS and/or DGAT from other species revealed that EgWSD1-like clusters a WS from petunia (*Petunia hydrida*; King *et al*., 2007) involved in the production of low molecular wax-esters in petals, as well as with *Arabidopsis* AtWSD1 (Li *et al*., 2008; Fig.2). It is also rather close to the bifunctional enzyme AcWS/DGAT from the bacterium *Acinetobacter calcoaceticus* ADP1, which is responsible for the production of both TAGs and WEs under so-called storage conditions (Kalscheuer & Steinbüchel, 2003). In addition, EgWSD1-like was more related to type-2 DGAT (AtDGAT2; Zhou *et al*., 2013) and

mammalian wax synthases (MmWS; Kawelke and Feussner, 2015) than to type-1 DGAT (AtDGAT1; Zou *et al*., 1999), membrane bound O-acyl transferase from *Arabidopsis thaliana* (AtMBOAT; Bouvier-Navé *et al*., 2010), or WS from jojoba (*Simmondsia chinensis*; Lardizabal *et al*., 2000) and *Euglena* (*Euglena gracilis;* Teerawanichpan & Qiu, 2010).

3.2 Heterologous expression of *EgWSD1-like* in the mutant yeast H1246 yields fatty acid isoamyl esters with saturated LC and VLC fatty acids.

In order to figure out whether EgWSD1-like could act as a DGAT, the heterologous expression of *EgWSD1-like* in the quadruple mutant yeast H1246, which is devoid of neutral lipids (Sandager *et al*., 2002), was first performed. Yeast transformed with an empty vector (pESC-URA) or with a construct containing oil palm DGAT1 (EgDGAT1, unpublished data) were respectively used as negative and positive controls. After 2 days of expression, lipids were extracted and neutral lipids were separated by TLC using hexane/diethyl ether/acetic acid (80:20:2; v/v/v) as mobile phase. As shown in Fig. 3A, while EgDGAT1 restored TAG synthesis, neither the yeast transformed with the empty vector nor that expressing *EgWSD1-like* produced TAGs in the mutant yeast strain H1246. Nevertheless, the expression of *EgWSD1-like* resulted in the appearance of a new compound migrating just below squalene in neutral lipids separation (Fig.3A). Using a more apolar solvent mixture (hexane/diethyl ether/acetic acid 90:7.5:0.1; v/v/v), this compound was separated from squalene, and shown to migrate between steryl esters (SE) and ethyl esters (EE; Fig. 3B). Using a larger volume of culture and the same TLC separation system, this new compound was then purified, extracted from silica using hexane and diethyl ether (1:1), and finally analyzed by GC-MS. MS fragmentation indicated that this new compound corresponded to FAIEs with saturated fatty acids ranging from 18:0 to 26:0 (Fig.3C; Fig.S2). Among the different molecular species, 20:0-, 22:0- and 26:0-IE were the most abundant, representing respectively 24.0, 30.8, and 26.7% of the total (Fig.3C). In contrast, the content of 18:0- and 24:0-IE was much lower, each representing less than 11 % of the total (Fig.3C).

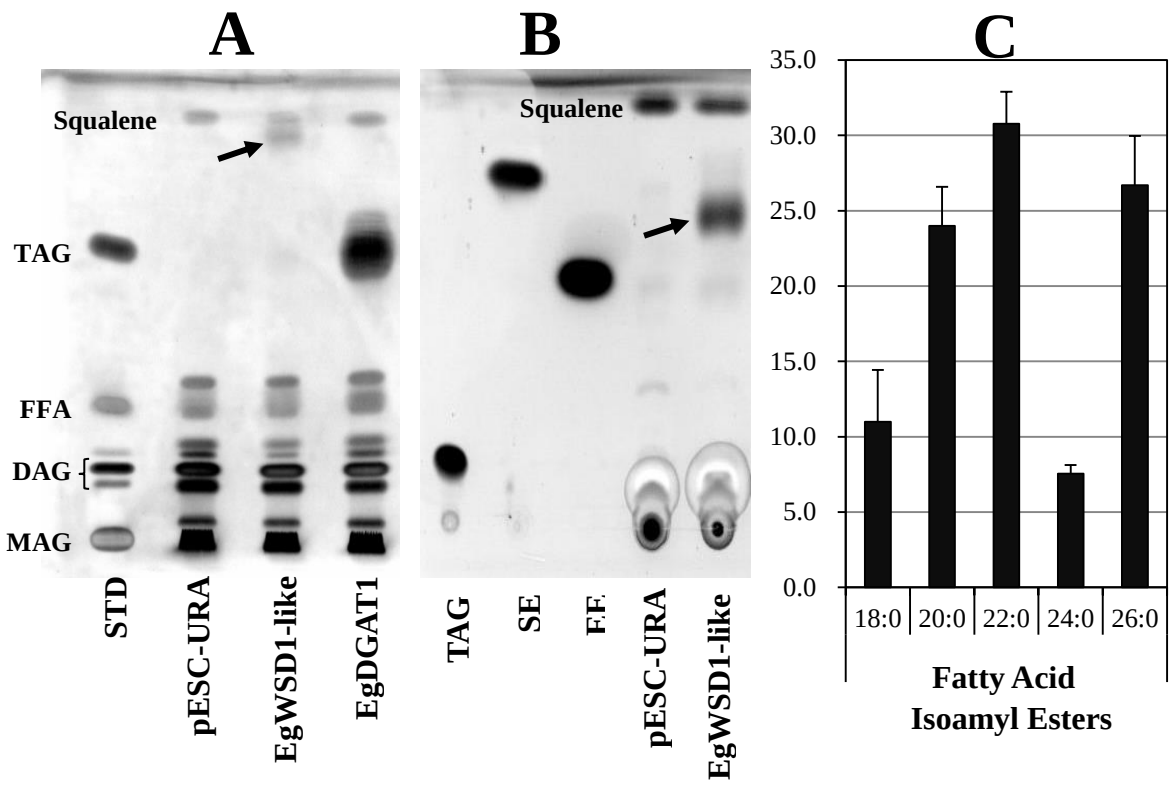

Fig.3 Expression of *EgWSD1-like* in the mutant yeast H1246 results in the production of fatty acid isoamyl esters. TLC analysis of neutral lipids synthesized in recombinant H1246 strains using (A) hexane/diethyl ether/acetic acid (80:20:2; v/v/v) or (B) hexane/diethyl ether/acetic acid (90:7.5:1; v/v/v)as solvent mixture. Lipids corresponding to an equivalent of 5 (A) or 10 OD units (B) of yeast culture cells were used. The new compound produced by heterologous expression of *EgWSD1-like* in the yeast mutant H1246 is indicated by an arrow (A, B) and was identified by GC-MS as a mixture of Fatty Acid Isoamyl Esters (FAIEs), which composition is showed in C. **pESC-URA**: yeast strain H1246 transformed with the empty vector ESC-URA ; **EgWDS1-like**: yeast strain H1246 expressing *EgWDS1-like*; **EgDGAT1**: yeast strain H1246 expressing *EgDGAT1* (unpublished). Standards: **STD**: standard mixture containing 2.5 µg of each oleic acid (FFA), monoolein (MAG), diolein (DAG), triolein (TAG); **TAG**: triolein (5µg); **SE**: cholestery oleate (10µg); **EE**: palmitoyl ethyl ester (10µg). This experiment was repeated three times with similar results.

3.3 Exogenous supply of different fatty alcohols to the yeast transformants

As *Saccharomyces cerevisiae* has no capacity to produce fatty alcohols, we then supplied exogenously several fatty alcohols to a culture of transgenic yeast expressing *EgWSD1-like* in order to determine if EgWSD1-like displayed WS activity. For all these feeding experiments, a culture of the H1246 yeast strain transformed with the empty vector was systematically grown in parallel and used for comparison (thereafter referred to as control). Three fatty alcohols with different chain-length (14OH, 18OH, 22OH) were first tested. However, significant levels of WEs could be detected in H1246 cultures transformed with a construct containing *EgWSD1-like* only in the presence of 14OH (data not shown). These results suggesting a preference of EgWSD1-like for medium chain-length fatty alcohols, 12OH, 14OH, 16OH were then used as exogenous substrates. If the control strain (H1246 transformed with the empty vector) could produce WEs in these conditions, supporting the existence of an endogenous WS activity in yeast, the expression of *EgWSD1-like* increased the WE levels by 50 to 150% (Fig.4A). Yeast cultures fed with 12OH showed the highest WE productivity, but the WE profiles of the control and the yeast strain expressing EgWSD1-like were very similar (Fig.4B). In the presence of 14OH or 16OH, the expression of *EgWSD1-like* resulted in much higher WE levels than in the control strain (150 and 100% increases, respectively; Fig.4A), and in the specific accumulation of WEs containing VLC fatty acids (20-26 carbons; Fig.4C and D).

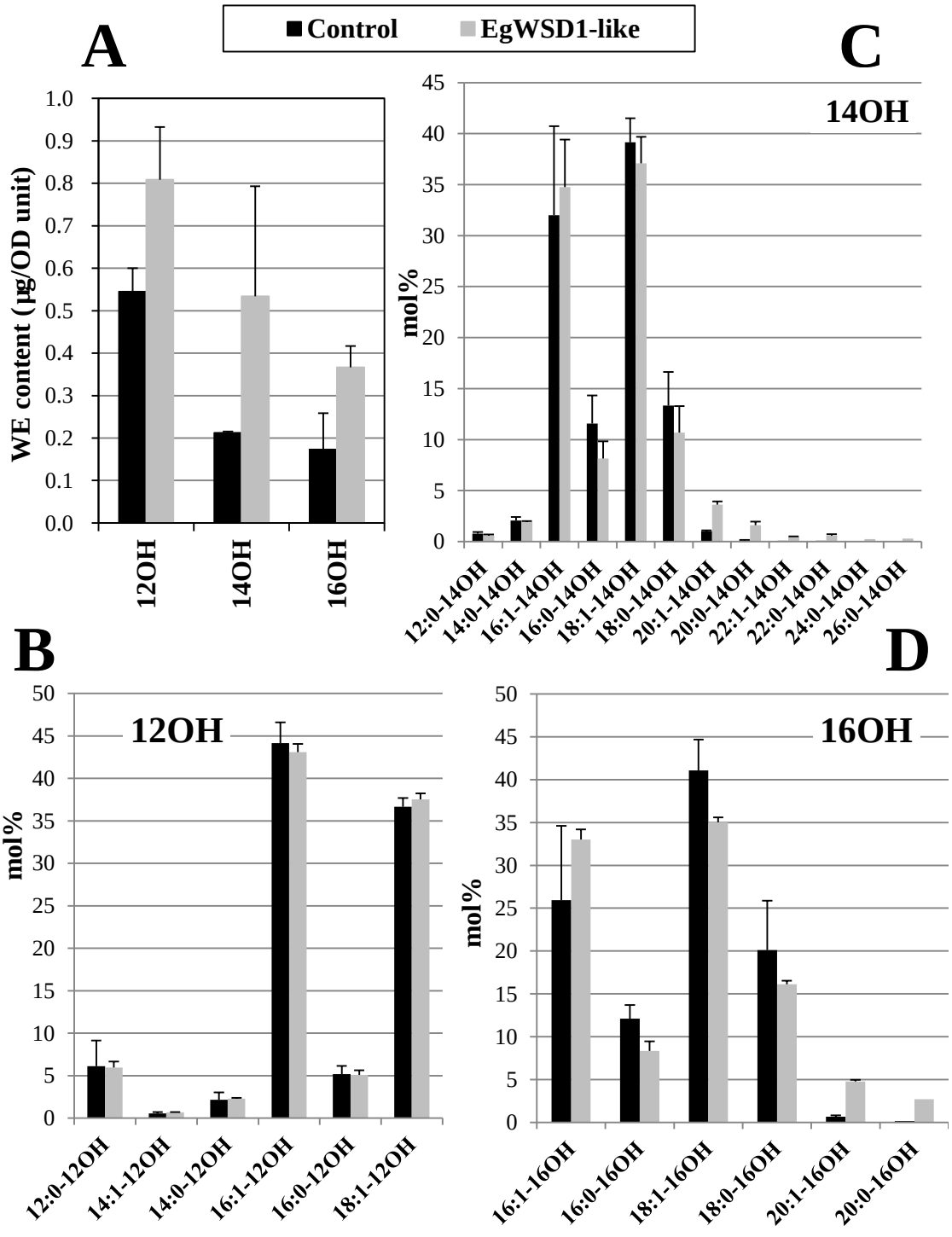

Fig.4 Wax esters produced by *EgWSD1-like* expression in the mutant yeast H1246 in the presence of exogenously supplied medium-chain fatty alcohols. Yeast transformed with the empty vector pESC-URA (Control) or with pESC-URA vector containing *EgWDS1-like* (EgWDS1-like) were grown for 48h in the presence of 12OH (A, B), 14OH (A, C) or 16OH (A, D) and their lipids were extracted and analyzed as indicated in the Material and Method section. The results combine three biological replicates. (A) Quantification of the amount of WEs produced. (B, C, D) Wax ester compositions in mol%.

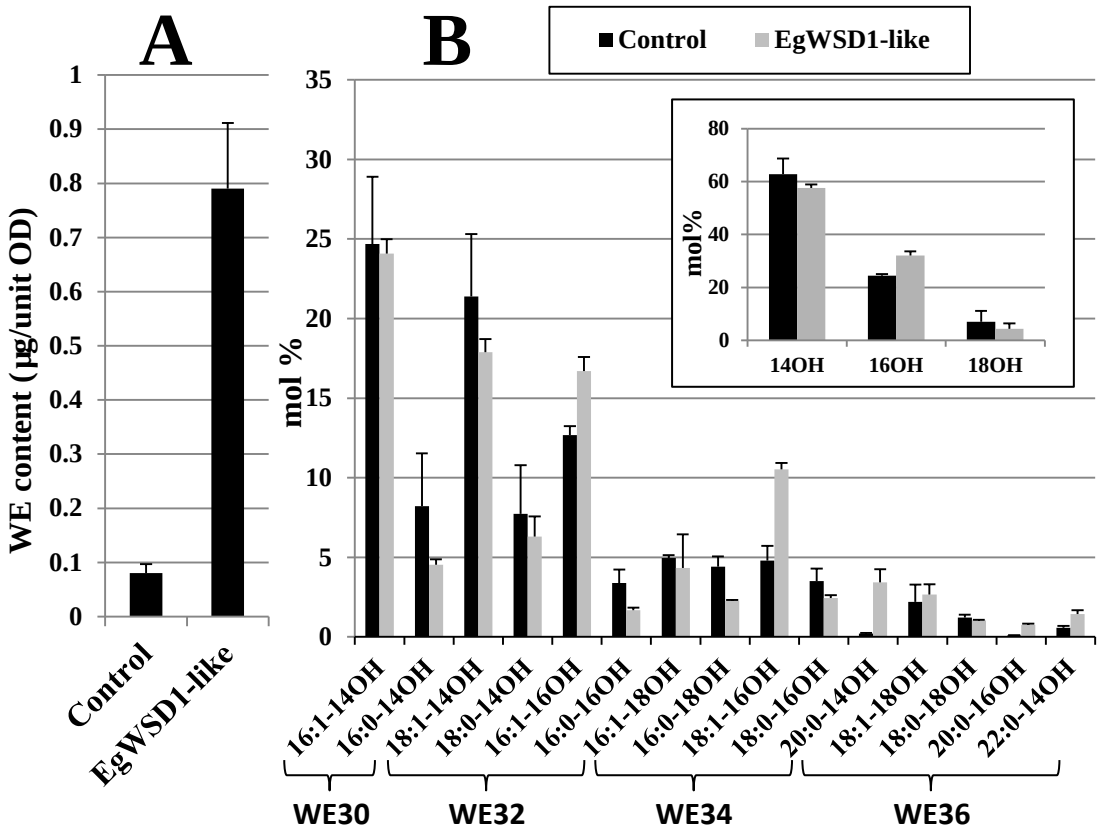

Fig.5 Wax esters produced by *EgWSD1-like* expression in the mutant yeast H1246 in the presence of an equimolar mixture of 14OH, 16OH and 18OH fatty alcohols. Yeast transformed with the empty vector pESC-URA (Control) or with pESC-URA vector containing *EgWDS1-like* (EgWDS1-like) were grown for 48h in the presence of an equimolar mixture of fatty alcohols (14OH, 16OH and 18OH) and their lipids were extracted and analyzed as indicated in the Material and Method section. The results combine three biological replicates. **(A)** Quantification of the amount of WEs produced. **(B)** Wax ester compositions in mol%. Insert, fatty alcohol composition of WE (in mol%).

To further study the substrate preference of EgWSD1-like, an equimolar mixture of fatty alcohols (14OH, 16OH and 18OH) was then supplied to the same H1246 transformants. Yeast expressing *EgWSD1-like* showed a WE productivity that was nearly 10 times that of the control (Fig.5A). Looking at the molecular species, WEs with 30 to 32 carbon atoms were most abundant, representing 70 to 80% of the total in both control and *EgWSD1-like* expressing strains (Fig.5B). In addition, these WEs mainly contained unsaturated fatty acids as for the same chain length and fatty alcohol (16:1/16:0-14OH, 18:1/18:0-14OH, 16:1/16:0-16OH), there was about three times

more WEs with the unsaturated fatty acids than with the saturated ones (compare 16:1-14OH, 18:1-14OH, 16:1-16OH with 16:0-14OH, 18:0-14OH and 16:0-16OH, respectively). Interestingly, WEs containing VLCFAs (i.e. 20:0-14OH, 22:0-14OH and 22:0-14OH) were overrepresented upon *EgWSD1-like* expression, confirming the preference of EgWSD1-like for these fatty acids. Concerning the fatty alcohol part, 14OH and 16OH were found in about 60 and 30% of the total WE, respectively, whereas WE containing 18OH represented less than 10% of the total, indicating the preference of EgWSD1-like for medium chain fatty alcohols (Fig.5B Insert).

3.4 Coexpression of *EgWSD1-like* with *TtFAR* in the mutant yeast strain H1246

We then reconstituted the complete WE biosynthetic pathway in the mutant yeast H1246 by coexpressing *EgWSD1-like* with the fatty acyl reductase *TtFAR* from *Tetrahymena thermophila* which was shown to produce 16OH and 18OH when expressed in yeast (Dittrich-Domergue *et al*., 2014). TLC analysis of lipids indicated that the heterologous expression of *EgWSD1-like* together with *TtFAR* in H1246 produced much more WEs than H1246 transformed with *TtFAR* alone (Fig.6A). Using 17:0-17OH WE as internal standard, the amount of WEs produced by coexpression in H1246 was about 7 times higher than that obtained upon *TtFAR* expression alone (Fig.6B). Concerning the WE composition, the proportion of WEs containing saturated fatty acids (14:0, 16:0 and 18:0) were systematically higher in yeast coexpressing *EgWD1-like* with *TtFAR* than in culture expressing *TtFAR* alone (Fig.6C). Conversely, the levels of WEs containing unsaturated fatty acid (16:1 and 18:1), such as 16:1-16OH, 18:1-16OH, 16:1-18OH, 18:1-18OH, were reduced in the coexpressing yeast compared to the control (expression *TtFAR* alone) (Fig.6C). Finally, WEs containing VLCFAs (20:0 to 26:0) were almost exclusively detected in the yeast coexpression *EgWSD1-like* together with *TtFAR* (Fig.6C insert).

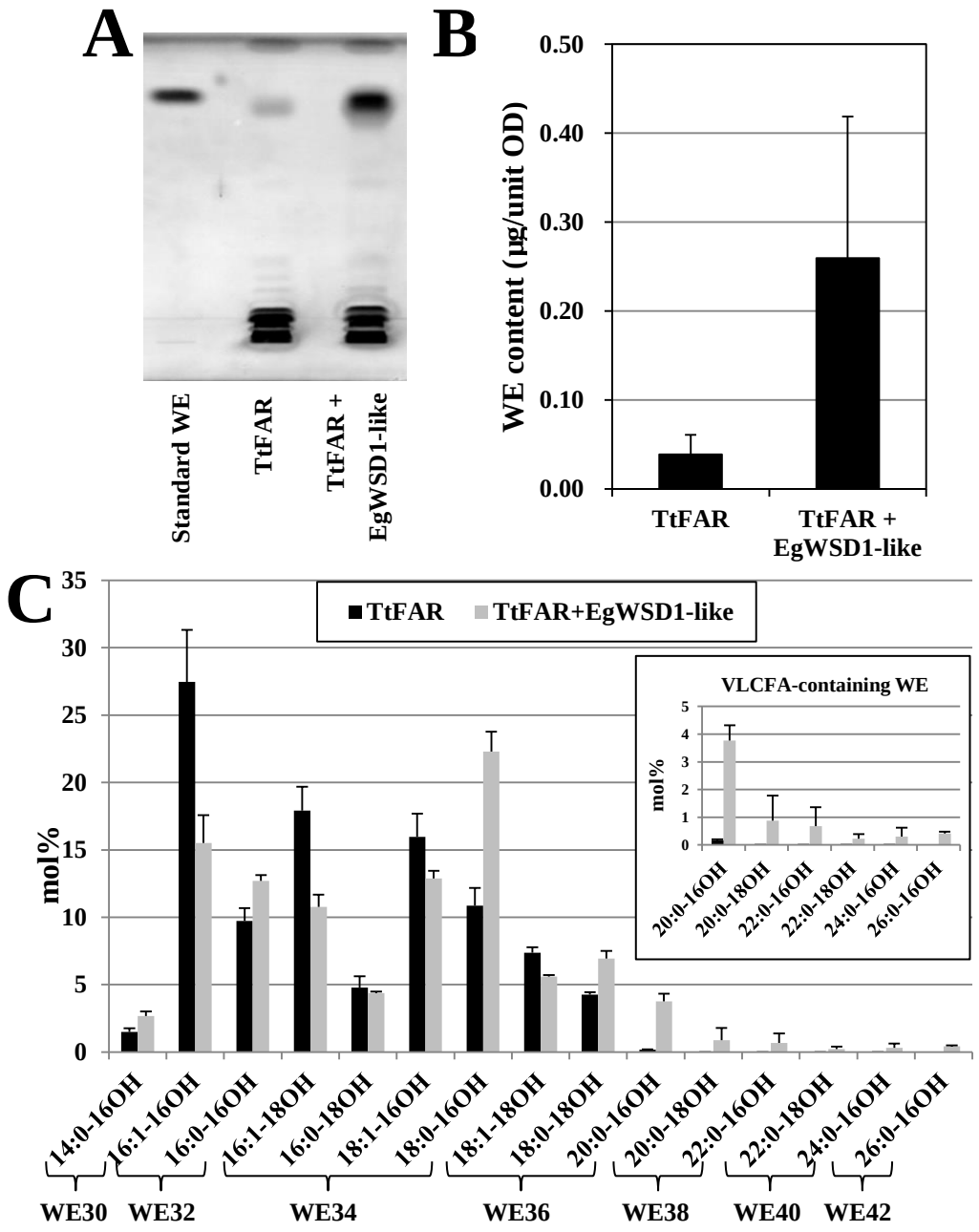

Fig.6 Wax esters produced by the mutant yeast H1246 expressing *TtFAR* alone or together with *EgWSD1-like*. (A) TLC analysis of lipids from yeast cultures using hexane/diethyl ether/acetic acid (90:7.5:1) solvent system. Lipids from 10 OD units of each yeast culture were loaded on TLC plate. **TtFAR**: yeast strain H1246 expressing TtFAR. **TtFAR+ EgWSD1-like**: yeast strain H1246 expressing TtFAR and EgWSD1-like. **Standard WE**: 10µg 17:0-17OH. **(B)** Quantification of the amount of WEs produced. **(C)** Wax ester compositions in mol%.

In order to get more insights on the substrate preference of EgWSD1-like, we then compared the fatty acid and fatty alcohol profiles of the WEs produced by coexpressing *TtFAR* and *EgWSD1-like* yeast to those of the corresponding total lipid

extracts (Fig.7). These later were obtained by directly transmethylating and silylating the lipid extracts as indicated in the Material and Method section. Using yeast expressing *TtFAR* alone or in combination with *EgWSD1-like*, very similar total fatty acid and fatty alcohol profiles were obtained (Fig.S3).

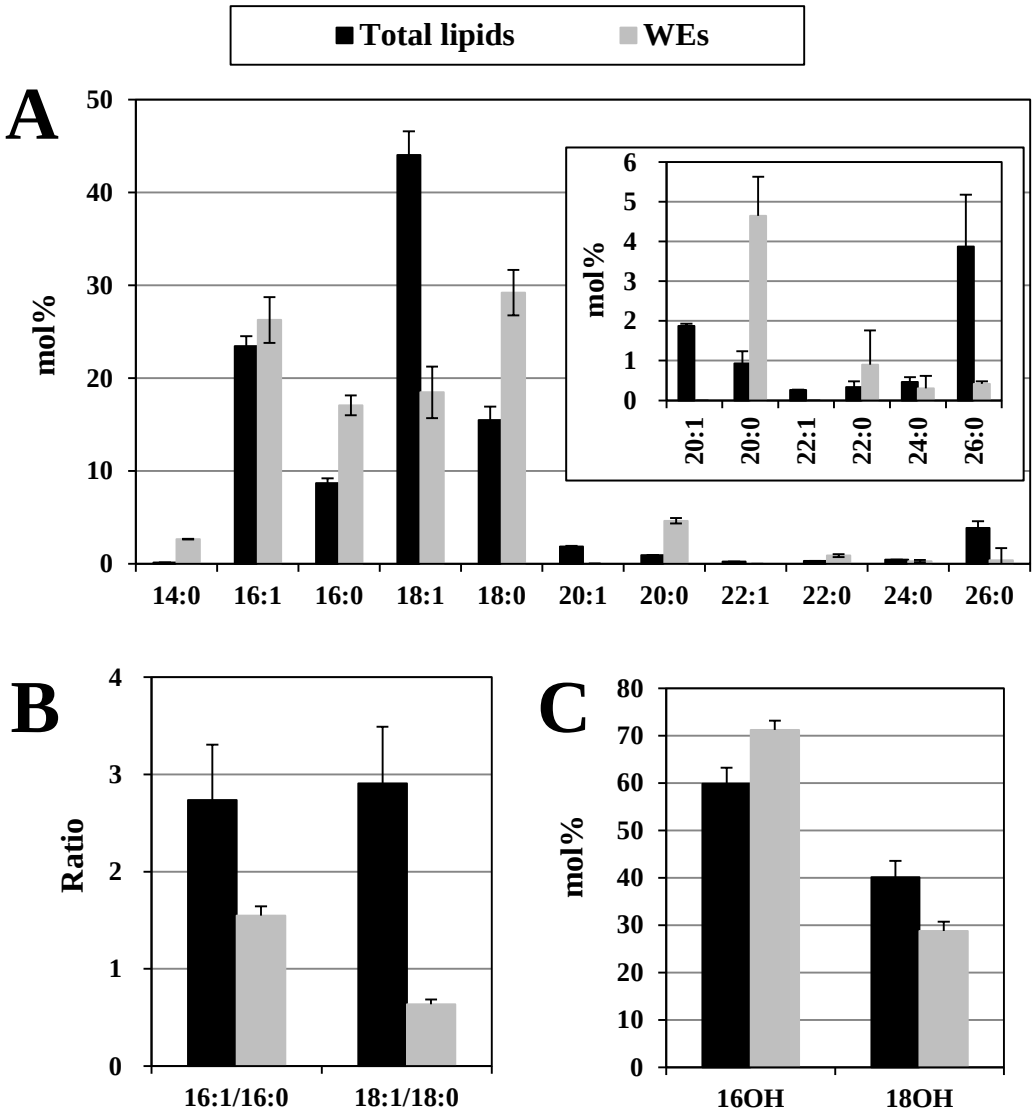

Fig.7 Comparative analysis of the total lipid and wax ester compositions of the mutant yeast H1246 expressing *TtFAR* together with *EgWSD1-like*. (A) Total fatty acid compositions; **(B)** Unsaturated fatty acid to saturated fatty acid ratios; **(C)** Fatty alcohol compositions.

As shown in Fig.7A, the proportions of most of the saturated fatty acid (14:0, 16:0, 18:0, 20:0 and 22:0) were higher in WEs than in total lipids. In contrast, the level of 18:1 was 2.3 times higher in the total lipids than in WEs, suggesting that 18:1 was severely discriminated by EgWSD1-like (Fig.7A). Calculation of 16:1 to 16:0 and

18:1 to 18:0 ratios clearly illustrated that saturated fatty acids are more likely incorporated by EgWSD1-like into WEs than monounsaturated ones (Fig.7B). Concerning VLCFAs, 20:0 appeared as the favorite substrate for EgWSD1-like (Fig.7A insert). In agreement with the feeding experiments previously described, analysis of the fatty alcohol profiles indicated that EgWSD1-like preferred 16OH over 18OH (Fig.7C).

4. Discussion

In the present study, we identified and characterized the first putative WS from oil palm mesocarp, *EgWSD1-like*. When heterologously expressed in yeast, it exhibited preference for 14OH and 16OH fatty alcohols and saturated LC and VLC fatty acids for WE synthesis. Although EgWSD1-like possesses the bifunctional WS/DGAT conserved domains with the proposed active site (HH*XXX*D), it did not function as a DGAT in yeast.

WSs have been classified in three distinct families based on sequence homologies. The first group is based on the jojoba WS, which was the first WS to be characterized (Lardizabal *et al*., 2000), and also contains the WS from *Euglena gracilis* (Teerawanichpan & Qiu, 2010; Fig.2). In *Arabidopsis*, there are twelve homologues of the jojoba WS, and 7 of them located side by side on chromosome 5. These sequences belong to the superfamily of Membrane-Bound O-Acyltransferase (MBOAT), and one of them (PSAT1) was characterized as a sterol acyltransferase (Bouvier-Navé *et al*., 2010; Fig.2). The second group contains the mammalian WSs, like *MmWS* from *Mus musculus* (Fig.2), and presents sequence similarity with several members of the acyltransferase family of enzymes involved in neutral lipid synthesis like acyl-CoA:cholesterol acyltransferase or acyl-CoA:monoacylglycerol acyltransferase (Cheng & Russell, 2004). Although this group shares some sequence similarity with AtDGAT2 from *Arabidopsis*, it has no absolute homologues in plants. The third type of WS is related to the WS/DGAT from the bacterium *Acinetobacter calcoaceticus* ADP1 (Kalscheuer & Steinbüchel, 2003; Wältermann *et al*., 2007). This

bifunctional protein showed both DGAT and WS activities *in vivo*, producing TAGs and WEs or FAIEs and FAEEs in bacteria and *S. cerevisiae* (Kalscheuer & Steinbüchel, 2003; Kalscheuer *et al*., 2004). *Arabidopsis* contains 11 genes annotated as *WS/DGAT*, including *AtWSD1*. AtWSD1 was shown to display both DGAT and WS activities *in vitro* using protein extracts from *E. coli* expressing *AtWSD1*, but only WS activity *in vivo* when heterologously expressed in *S. cerevisiae* (Li *et al*., 2008). *Arabidopsis wsd1* mutants displayed reduced WE levels in the stem waxes, supporting the key role played by AtWSD1 in WE synthesis *in planta* (Li *et al*., 2008). According to our phylogenetic analysis (Fig.2), *EgWSD1-like* from oil palm mesocarp is predicted to be one of this bifunctional type of WS.

The expression of *EgWSD1-like* in oil palm mesocarp was comparable with that of *EgDGAT1* (Fig.1), which is involved in TAG biosynthesis. Moreover, *EgWSD1-like* was at least 6 times more expressed in fruit mesocarp than in leaves, but not in kernel (Fig.1), suggesting that it plays a role in the accumulation of storage compounds in oil palm mesocarp. *In silico* analysis of the protein sequence predicted that EgWSD1-like contained at least one transmembrane domain, strongly suggesting that this protein is a membrane bound acyltransferase. This prediction is in agreement with the fact that WS and DGAT activities were measured using the insoluble fraction of crude extract of *E. coli* expressing the AcWS/DGAT from *Acinetobacter calcoaceticus* ADP1 (Kalscheuer & Steinbüchel, 2003). Since subcellular localization studies showed in addition that MmWS and AtWSD1 were associated with the ER (Cheng & Russell, 2004; Li *et al*., 2008), it is highly probable that EgWSD1-like is also a microsomal protein.

Unlike the bifunctional WS/DGAT enzyme from *Acinetobacter calcoaceticus* ADP1, which restored TAG biosynthesis when expressed in the quadruple yeast mutant H1246 (Kalscheuer *et al*., 2004), the expression of *EgWSD1-like* failed to produce any TAGs (Fig.3A), indicating that EgWSD1-like do not have any DGAT activity in these conditions. However, the lack of *in vivo* DGAT activity for WS/DGAT-like proteins from higher plants expressed in yeast has previously been

reported in the case of AtWSD1 and PhWS, from *Arabidopsis* and *Petunia hydrida* respectively (King *et al*., 2007; Li *et al*., 2008). Altogether these data suggest that in contrast to bacterial bifunctional WS/DGAT proteins, those of plant origin do not display any DGAT activity when expressed in *S. cerevisiae*. Nevertheless, the expression of *EgWSD1-like* in the mutant yeast H1246 resulted in the production of FAIEs with saturated LC and VLC fatty acids (Fig.3), which is consistent with results obtained with other WSs from the WS/DGAT family, including *AcWS/DGAT* from *A. calcoaceticus* and *PhWS* from *Petunia hydrida* (Kalscheuer *et al*., 2004; King *et al*., 2007). Interestingly, both EgWSD1-like and PhWS produced FAIEs with preferentially saturated LC and VLC acyl-groups when expressed in the mutant yeast H1246 (Fig.3; King *et al*., 2007).

When fatty alcohols were exogenously supplied to the transgenic yeast expressing *EgWSD1-like*, WEs were systematically detected, indicating that EgWSD1-like displayed WS activity. In these conditions, EgWSD1-like showed a slight preference for medium chain fatty alcohols (14:0 and 16:0), unsaturated fatty acid (16:1 and 18:1) as well as VLC fatty acids (Fig. 4 and 5). Nevertheless, these *in vivo* experiments showed that EgWSD1-like used as substrates a wide range of saturated and unsaturated acyl-CoAs from C14 to C26, like the jojoba WS (Lardizabal *et al*., 2000). In similar assays, AtWSD1 used palmitic acid and 18:0, 24:0, or 28:0 fatty alcohols to produce WEs with 34 to 44 carbon atoms, in line with the VLC-WEs found in *Arabidopsis* stem waxes (Li *et al*., 2008). When the WE biosynthetic pathway was reconstituted in yeast using coexpression of *TtFAR* together with *EgWSD1-like*, the most abundant WE produced was 18:0-16OH. In addition, WEs containing saturated VLC fatty acids were only detected in the presence of EgWSD1-like, indicating that the endogenous yeast activity generating WEs in these conditions did not accept VLC fatty acids as substrates (Fig.6). Conversely, these results confirmed the preference of EgWSD1-like for saturated VLC acyl-CoAs (Fig.6C). Furthermore, when comparing the fatty acyl compositions of the WEs and total lipids (Fig.7), EgWSD1-like appeared to preferentially use saturated fatty acids, especially 16:0, 18:0 and 20:0

(Fig.7A and B), and 16OH rather than 18OH (Fig.7C). Recently, the composition of the acyl-CoA pool from a culture of the yeast mutant strain H1246 exogenously supplied with 16OH was determined (Kawelke & Feussner, 2015). Interestingly, this pool was dominated by 16:1 and 18:1-CoAs, which together represented about 64% of the total, as well as by saturated VLC acyl-CoAs (>C18), which altogether represented about 24% of the total. In contrast, 16:0- and 18:0- CoAs only represented respectively only 8% and 4% of the total (Kawelke & Feussner, 2015). In our experiments, the composition of WEs synthesized by the yeast mutant strain H1246 co-expressing *EgWSD1-like* and *TtFAR* (which mainly produced 16OH) showed an extreme enrichment in saturated 16:0 and 18:0 fatty acids, which were present in respectively 17% and 29% of the WE molecular species (Fig.7A). With respect to VLC acyl-CoAs, 26:0- and 24:0-CoAs were the major species (24:0-CoA being about as abundant as 18:0), whereas 20:0- and 22:0- CoAs represented really minor species (Kawelke & Feussner, 2015). In contrast, in the yeast mutant strain H1246 co-expressing *EgWSD1-like* and *TtFAR*, WEs containing 22:0, 24:0 and 26:0 represented less than 1% of the total, whereas the amount of 20:0 was up to 4.6% (Fig.7A). These comparisons further support the idea that EgWSD1-like has a preference for saturated LC and VLC fatty acids, especially 18:0 and 20:0. Concerning the fatty alcohol substrates, EgWSD1-like preferentially used 14OH (Fig.5B), suggesting a slight substrate specificity for medium chain fatty alcohols.

Whatever the conditions (in the presence of exogenously supplied fatty alcohols or in coexpression assays with *TtFAR*), the expression of *EgWSD1-like* in the mutant yeast H1246 did not produce high amounts of WEs. The first reason could be that EgWSD1-like was not very stable when expressed in yeast resulting in low activity. Interestingly, the expression of the *WS* from jojoba in *S. cerevisiae* did not result in WE production (Lardizabal *et al*., 2000), suggesting that this system may not be suitable for expressing all WSs from plant. Another plausible reason may be that better substrates for EgWSD1-like, like for example unsaturated fatty alcohols, were not tested in our study. Nevertheless, PhWS which displays similar characteristics

with EgWSD1-like, did not show any activity for unsaturated fatty alcohols *in vitro* (King *et al*., 2007). Although EgWSD1-like showed a slight preference for VLC fatty acids, supplying yeast with saturated VLC fatty acids is difficult to perform because of the insolubility of these fatty acids and their low incorporation by yeast. Interestingly, when analyzing the fatty acyl composition of oil palm mesocarp during fruit ripening in detail, we could detect traces of VLCFAs, 20:0, 22:0 and 24:0 representing about 0.28, 0.03 and 0.02% of the total fatty acids. These data suggest that the preference of EgWSD1-like for VLCFAs we measured in yeast could play a role in oil palm mesocarp. Unfortunately, we could not confirm with certainty the presence of fatty alcohols in oil palm mesocarp, because traces of such compounds are often present in commercially available solvents. Nevertheless, the presence of aliphatic alcohols in palm oil has already been reported (Sambanthamurthi *et al*., 2000). Therefore, further studies are needed to demonstrate the exact function of EgWSD1-like in oil palm mesocarp.

In conclusion, this study presented the functional characterization of a new enzyme from oil palm mesocarp, EgWSD1-like, which belongs to the bifunctional WS/DGAT family. Although, EgWSD1-like lacked DGAT activity, it could produce FAIEs and use fatty alcohols for synthesizing WEs when expressed in yeast.

References

Bates, P. D. (2016). Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1861*(9), 1214–1225. https://doi.org/10.1016/j.bbalip.2016.03.021

Baud, S., & Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Progress in Lipid Research*, *49*(3), 235–249. https://doi.org/10.1016/j.plipres.2010.01.001

Bernard, A., & Joubès, J. (2013). Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. *Progress in Lipid Research*, *52*(1), 110–129. https://doi.org/10.1016/j.plipres.2012.10.002

Bourgis, F., Kilaru, A., Cao, X., Ngando-ebongue, G.-F., Drira, N., Ohlrogge, J. B., … Arondel, V.

(2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(30), 12527–12532. https://doi.org/10.1073/pnas.1106502108

Bouvier-Navé, P., Berna, A., Noiriel, A., Compagnon, V., Carlsson, A. S., Banas, A., … Schaller, H. (2010). Involvement of the phospholipid sterol acyltransferase1 in plant sterol homeostasis and leaf senescence. *Plant Physiology*, *152*(1), 107–19. https://doi.org/10.1104/pp.109.145672

Cao, X., Bourgis, F., Kilaru, A., Cao, X., Ngando-ebongue, G., Drira, N., & Ohlrogge, J. B. (2011). Correction for Bourgis et al., Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proceedings of the National Academy of Sciences*, *108*(44), 18186–18186. https://doi.org/10.1073/pnas.1115243108

Cheng, J. B., & Russell, D. W. (2004). Mammalian wax biosynthesis: II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family. *Journal of Biological Chemistry*, *279*(36), 37798–37807. https://doi.org/10.1074/jbc.M406226200

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, *18*(24), 5294–5299. https://doi.org/10.1021/bi00591a005

Corley, R. H. V., & Tinker, P. B. (Eds.). (2003). *The Oil Palm*. Oxford, UK: Blackwell Science Ltd. https://doi.org/10.1002/9780470750971

Dittrich-Domergue, F., Joubès, J., Moreau, P., Lessire, R., Stymne, S., & Domergue, F. (2014). The bifunctional protein TtFARAT from Tetrahymena thermophila catalyzes the formation of both precursors required to initiate ether lipid biosynthesis. *Journal of Biological Chemistry*, *289*(32), 21984–21994. https://doi.org/10.1074/jbc.M114.579318

Käll, L., Krogh, A., & Sonnhammer, E. L. . (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology*, *338*(5), 1027–1036. https://doi.org/10.1016/j.jmb.2004.03.016

Kalscheuer, R., Luftmann, H., & Steinbu, A. (2004). Synthesis of Novel Lipids in Saccharomyces cerevisiae by Heterologous Expression of an Unspecific Bacterial Acyltransferase Synthesis of Novel Lipids in Saccharomyces cerevisiae by Heterologous Expression of an Unspecific Bacterial Acyltransferase. *Applied and Env. Microbiol.*, *70*(12), 7119–7125. https://doi.org/10.1128/AEM.70.12.7119

Kalscheuer, R., & Steinbüchel, A. (2003). A novel bifunctional wax ester synthase/acyl-CoA:Diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1. *Journal of Biological Chemistry*, *278*(10), 8075–8082. https://doi.org/10.1074/jbc.M210533200

Kawelke, S., & Feussner, I. (2015). Two predicted transmembrane domains exclude very long chain fatty acyl-CoAs from the active site of mouse wax synthase. *PLoS ONE*, *10*(12), 1–19.

https://doi.org/10.1371/journal.pone.0145797

Kawelke, S., Feussner, I., Samuels, L., Kunst, L., Jetter, R., Butovich, I., … Li, W. (2015). Two Predicted Transmembrane Domains Exclude Very Long Chain Fatty acyl-CoAs from the Active Site of Mouse Wax Synthase. *PLOS ONE*, *10*(12), e0145797. https://doi.org/10.1371/journal.pone.0145797

King, A., Nam, J. W., Han, J., Hilliard, J., & Jaworski, J. G. (2007). Cuticular wax biosynthesis in petunia petals: Cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. *Planta*, *226*(2), 381–394. https://doi.org/10.1007/s00425-007-0489-z

Kunst, L., & Samuels, A. L. (2003). Biosynthesis and secretion of plant cuticular wax. *Progress in Lipid Research*, *42*(1), 51–80. https://doi.org/10.1016/S0163-7827(02)00045-0

Kunst, L., & Samuels, L. (2009). Plant cuticles shine: advances in wax biosynthesis and export. *Current Opinion in Plant Biology*, *12*(6), 721–727. https://doi.org/10.1016/j.pbi.2009.09.009

Lardizabal, K. D., Metz, J. G., Sakamoto, T., Hutton, W. C., Pollard, M. R., & Lassner, M. W. (2000). Purification of a jojoba embryo wax synthase, cloning of its cDNA, and production of high levels of wax in seeds of transgenic arabidopsis. *Plant Physiology*, *122*(March), 645–655. https://doi.org/10.1104/pp.122.3.645

Lee, S. B., & Suh, M. C. (2015). Advances in the understanding of cuticular waxes in Arabidopsis thaliana and crop species. *Plant Cell Reports*, *34*(4), 557–572. https://doi.org/10.1007/s00299-015-1772-2

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., … Kunst, L. (2008). Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. *Plant Physiology*, *148*(1), 97–107. https://doi.org/10.1104/pp.108.123471

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., … Zhang, D. (2010). Cytochrome P450 family member CYP704B2 catalyzes the {omega}-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *The Plant Cell*, *22*(1), 173–90. https://doi.org/10.1105/tpc.109.070326

Metz, J. G., Pollard, M. R., Anderson, L., Hayes, T. R., & Lassner, M. W. (2000). Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed. *Plant Physiology*, *122*(3), 635–44. https://doi.org/10.1104/PP.122.3.635

Ohlrogge, J. B., Pollard, M. R., & Stumpf, P. K. (1978). Studies on biosynthesis of waxes by developing jojoba seed tissue. *Lipids*, *13*(3), 203–210. https://doi.org/10.1007/BF02533401

Rowland, O., Zheng, H., Hepworth, S. R., Lam, P., Jetter, R., & Kunst, L. (2006). CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant Physiology*, *142*(3), 866–77. https://doi.org/10.1104/pp.106.086785

Sambanthamurthi, R., Sundram, K., & Tan YewAi. (2000). *Chemistry and biochemistry of palm oil. Progress in Lipid Research* (Vol. 39). https://doi.org/10.1016/S0163-7827(00)00015-1

Sandager, L., Gustavsson, M. H., Ståhl, U., Dahlqvist, A., Wiberg, E., Banas, A., … Stymne, S.

(2002). Storage lipid synthesis is non-essential in yeast. *The Journal of Biological Chemistry*, *277*(8), 6478–82. https://doi.org/10.1074/jbc.M109109200

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, *28*(10), 2731–9. https://doi.org/10.1093/molbev/msr121

Teerawanichpan, P., & Qiu, X. (2010). Fatty acyl-coA reductase and wax synthase from euglena gracilis in the biosynthesis of medium-chain wax esters. *Lipids*, *45*(3), 263–273. https://doi.org/10.1007/s11745-010-3395-2

Tusnády, G. E., & Simon, I. (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics (Oxford, England)*, *17*(9), 849–50. https://doi.org/10.1093/BIOINFORMATICS/17.9.849

Wältermann, M., Stöveken, T., & Steinbüchel, A. (2007). Key enzymes for biosynthesis of neutral lipid storage compounds in prokaryotes: Properties, function and occurrence of wax ester synthases/acyl-CoA:diacylglycerol acyltransferases. *Biochimie*, *89*(2), 230–242. https://doi.org/10.1016/j.biochi.2006.07.013

Xu, C., & Shanklin, J. (2016). Triacylglycerol Metabolism, Function, and Accumulation in Plant Vegetative Tissues. *Annual Review of Plant Biology*, *70*(3), 156–7. https://doi.org/10.1146/annurev-arplant-043015-111641

Zhou, X. R., Shrestha, P., Yin, F., Petrie, J. R., & Singh, S. P. (2013). AtDGAT2 is a functional acyl-CoA:Diacylglycerol acyltransferase and displays different acyl-CoA substrate preferences than AtDGAT1. *FEBS Letters*, *587*(15), 2371–2376. https://doi.org/10.1016/j.febslet.2013.06.003

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G., & Taylor, D. C. (1999). The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *The Plant Journal*, *19*(6), 645–653. https://doi.org/10.1046/j.1365-313x.1999.00555.x

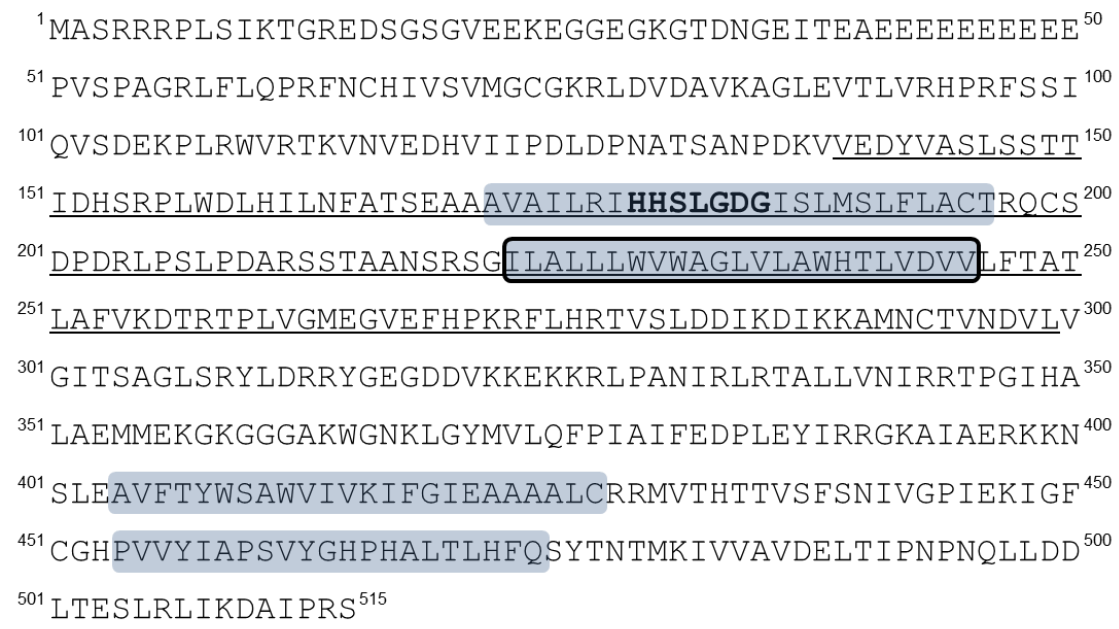

Fig.S1 Analysis of EgWSD1-like amino acid sequence. The wax ester synthase-like / acyl-CoA acyltransferase domain (pfam03007) is underlined while the essential motif for acyl-transferase catalytic activity (178HHXXXDG185) is bolded. Putative transmembrane domains identified with Phobius and HMMTOP are respectively surrounded and shaded in grey.

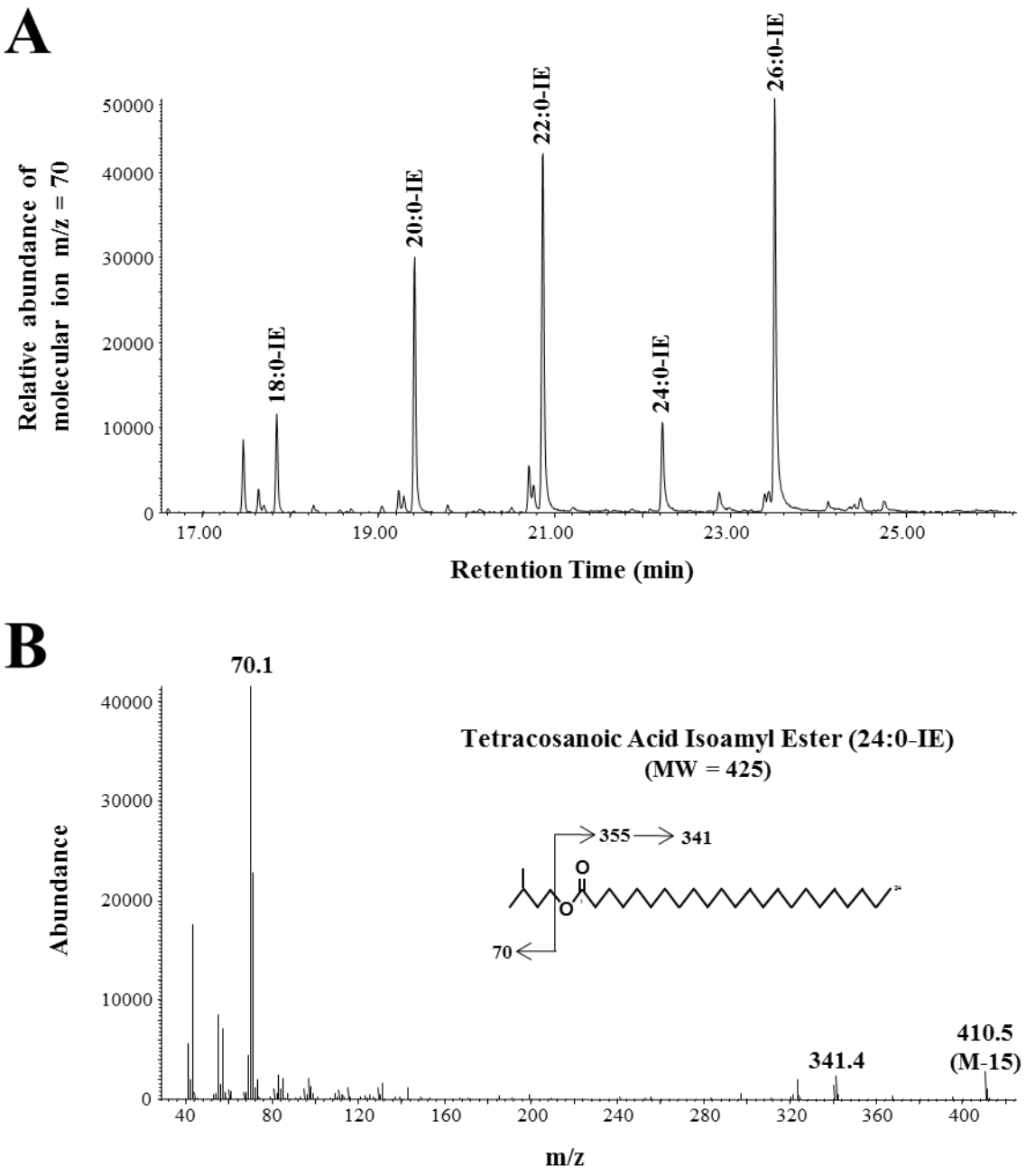

Fig.S2 Identification of the fatty acid isoamyl esters produced upon heterologous expression of *EgWSD1-like* in the mutant yeast H1246. (A) Representative gas chromatogram of the fatty acid isoamyl esters (FAIEs) produced when expressing *EgWSD1-like* in the mutant yeast H1246. Their specific detection was based on the molecular ion m/z = 70. **(B)** Mass spectra and drawing representing the fragmentation of tetracosanoic acid isoamyl ester (24:0-IE).

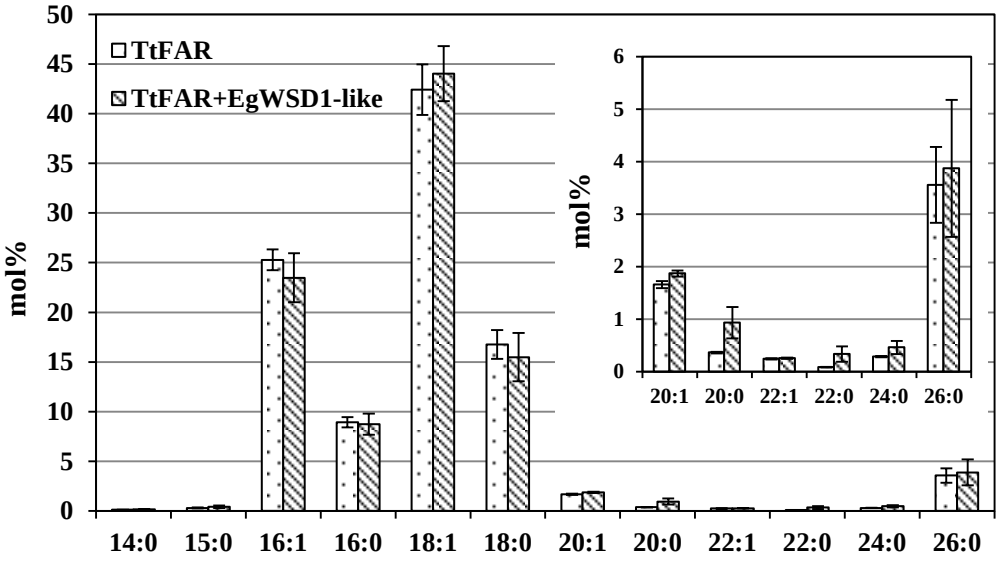

Fig.S3 Total fatty acid compositions of the mutant yeast H1246 expressing *TtFAR* alone (TtFAR) or in combination with *EgWSD1-like* (TtFAR+EgWSD1-like).

Chapter 6. Conclusions and Perspectives

It has been long thought that oil assembly followed the straightforward Kennedy pathway, with two successive acylations of glycerol 3-P to phosphatidic acid, followed by dephosphorylation to DAG and a third acylation step to yield TAGs. Recent studies (Bates & Browse, 2012) have demonstrated that this process was complicated by editing mechanisms that could heavily influence the whole pathway, as defective editing can lead to tremendous changes in oil fatty acid composition (Bates *et al*., 2012; Lu *et al*., 2009) and to reduced overall TAG synthesis in oil-storing embryos (Bates *et al*., 2014; van Erp *et al*., 2011). There are two ways to edit oil fatty acid composition: first, an acylation-deacylation cycle mostly due to LPCAT leads to enrichment of acyl-CoA pools in PUFAs. Second, PC can be a direct intermediate in pathway, to produce DAG species with PUFAs to TAG synthesis. PDCT is likely the key enzyme for this pathway. Thus, PC and TAG biosynthetic pathways are very intricated and sophisticated regulation mechanisms are likely to occur to allow the cell to strictly control fatty acid content of PC (this is very important to maintain membrane functional properties), despite very strong fatty acid flux to oil synthesis in oil-storing tissues.

Flux models have been proposed by Bates *et al*. (2009), based on radiolabeling studies in soybean embryo, which is an organ that stores important amounts of polyunsaturated species in oil, therefore requiring extensive editing. Model that fits best with experimental data involve at least three distinct pools of DAGs and two of PC. Because there are no obvious biochemical techniques to easily access to these pools, indirect approaches based on reasonable hypotheses are required to further demonstrate the existence of these pools. We have made the simple hypothesis that synthesis of PC used for TAG and synthesis of PC used for membrane could be somehow physically separated, either through specific channeling through distinct supramolecular enzymes complexes forming metabolons, or by being localized in distinct areas of the ER membrane. Numerous metabolons have already been demonstrated concerning many biosynthetic pathways (Ratledge, 2002; Winkel, 2004) and few studies have pointed distinct localization of lipids at sub-organelle level; for

example, DGAT1 and 2 localize to distinct areas of ER in Tung, TAG and oleosin syntheses occur concomitantly in distinct sunflowersub fractions of ER (Lacey *et al*., 1999), sec14 protein (Alb *et al*., 1996) specifically regulates PC synthesis in yeast golgi apparatus, and a CCT isoform localizes specifically in nucleus envelop (Cornell & Ridgway, 2015). Thus, physical separation of such sort might explain kinetic evidence for distinct pools reported by Bates *et al*. (2009).

Therefore, we have initiated studies to possibly prove physical existence of multi-molecular complexes of oil-assembly enzymes. Because biochemical studies require lots of biological material, obvious models with very tiny oil storing organs, such as *Arabidopsis*, must be ruled out. We chose oil palm mesocarp for several reasons:

First, kilograms of mesocarp can be obtained throughout the year as palm continuously produces bunches; the only caveat is that it does not grow in France but fresh samples can brought from Cameroon within 24h (Ngando-Ebongue *et al*., 2006) and material can be flash-frozen immediately after harvest and shipped frozen to Bordeaux (Bourgis *et al*., 2011).

Second, palm is rather unique for several reasons. Mesocarp is the plant tissue with highest amount of stored oil (up to 90% dry weight, Bourgis *et al*., 2011), it is one of the rare tissue that uses, in addition to PDAT1, both DGAT1 and 2 isoforms without synthesizing unusual fatty acids. Also, palm stores high amounts of oil in kernel (Sambanthamurthi *et al*., 2000), with distinct fatty acid composition (high levels of medium-chain fatty acids), thus allowing intra-species comparative studies.

Third, palm accumulation has been extensively characterized, both in Bordeaux laboratory and also by other groups (MPOB in Malaysia, Sambanthamurthi *et al*., 2000; Sime Darby (Neoh *et al*., 2013; Ooi *et al*., 2015; Teh *et al*., 2013), CIRAD-IRD in Montpellier (Dussert *et al*., 2013; Tranbarger *et al*., 2011)). Genome (Singh *et al*., 2013) and transcriptomes are available and other systems biology type of studies have been published recently (Guerin *et al*., 2016; Teh *et al*., 2016b) and numerous

genotypes and segregating populations are available from breeders (Corley & Tinker, 2003). Our study (Chapter 3.1) demonstrates that genetic variability of fatty acid composition in palm mesocarp is more important than that noted for *Arabidopsis* (Hobbs *et al*., 2004; O'Neill *et al*., 2003).

At last, palm is world main oil crop and the most efficient one (Corley, 2009). It has, up to recently and contrarily to other major oil crops, poorly benefited from fundamental research studies and these are absolutely required to increase both yield and oil quality.

Our interaction studies clearly demonstrate that, virtually, any type of multi-enzyme complex devoted to fatty acid assembly into oil and PC might exist, as more than one third of enzyme couples tested by split-ubiquitin yeast double hybrid system showed interactions indeed. Because of experimental caveats inherent to system (interactions sometimes weak, difficulties in showing both-ways interactions) and the fact we probably missed some interactions as not all proteins tested could be used as baits, other approaches need to be used to confirm these results. One of these approaches is blue native electrophoresis to identify complexes released through the action of mild detergents. We have started to devise a cell fractionation protocol and showed that it is possible to obtain membrane fractions that contain oil-assembly enzymes detectable through proteomics. Further work is required to improve the protocol but these results look promising, even if DGAT activity is very difficult to measure from these fractions. Alternative approaches might be to enrich membranes that synthesize oil using antibodies specific to DGAT or PC-related enzymes, in pull-down experiments. Co-precipitating proteins can then be identified by proteomics analyses. Alternatives are to tag some of key enzymes and use these tags to enrich/purify complexes. However, such approach requires transgenic plants, which can virtually be ruled out in the case of oil palm, unless one uses a cultured cell system (Masani *et al*., 2014).

We also extensively characterized lipid content of mesocarp. Results proved a bit disappointing as very little variation could be noted. First, when studying material that shows important variability of fatty acid content, impact on molecular species structure (*sn*-2 position) and content appears to be very limited. Second, very few changes in molecular species can be noted throughout mesocarp development. Main information, that is, DAG and TAG showing similar fatty acid compositions distinct from that of PC, was obtained through "classical" TLC-GC analyses. This information might suggest that PDCT pathway does not operate much in oil palm. Also, oil assembly process does not affect much fatty acid variability amongst different genotypes, which was expected. This means that lipidomic analyses of palm mesocarp are not of much use to our studies as they bring very little information. The only very minor changes that correlate with oil accumulation concern increases in 18:0-containing species, while 18:0 total fatty acid content does not change; one might speculate that, as it correlates with increases of transcript levels for PDCT, LPPβ and DGAT2, one of these enzymes might show a selectivity that might be responsible for observed difference. Although this remains highly speculative, we feel it is worthwhile to determine the specificity and substrate preferences of these enzymes; preliminary heterologous expression studies of palm DGATs in yeast indicate that different isoforms produce oil with different fatty acid composition, suggesting that these enzymes possess distinct substrate preferences indeed.

Lipidomics analyses of kernel were initially carried out only to have a glimpse at the other lipid-accumulating system of oil palm, using fruit material that we had handy in freezer. However, results proved interesting as many changes could be noted throughout development. Although the kernel system still requires extensive characterization, it appears clearly that two distinct oil-synthesizing systems operate during kernel development, leading to synthesis of oils of very different composition: one contains mid-chain saturated fatty acids only and the other long-chain fatty acids, while mixed species can also be detected. PDCT transcripts in almost-ripe kernel are very low, suggesting that PC is unlikely to be intermediate to TAG synthesis.

However, because mixed TAG and PC species, with at least one mid-chain and one long chain fatty acid, can be detected, it is likely that that both oil synthesizing systems are not fully separated. Thus, because it is easy to track mid-chain, long-chain and intermediate species, lipidomics applied to kernel oil synthesizing system might prove a good model for our studies in the future. Hundreds of grams of kernel can be easily obtained for biochemical and complex studies. However, the kernel system will require extensive characterization, starting with new deep transcriptome studies, as data from Dussert *et al*. (2013) barely allow detection of TAG assembly transcripts.

Further investigations on palm systems will sooner or later require flux studies (Kruger & Ratcliffe, 2015). We labeled mesocarp slices with radioactive acetate and glycerol with very limited success, possibly due to age of fruits. Other investigators (Oo *et al*., 1985) reported success was such experiments in oil palm mesocarp, using fruits immediately after harvest. However, laboratory neighbors working on fluxes using other fruit models told us they had very low success (B. Beauvoit, personal communication). Thus, a possibility would be to construct transgenic cell lines (plant cell cultures and/or appropriate yeast strains and mutants) over-expressing large combinations of enzymes of pathway to study flux to high oil synthesis. These would be artificial models but the possibility of changing enzyme content would allow us to study their individual impact, as well as impact of the whole transgenic system on host transcriptome, lipidome, proteome and fluxome related to oil synthesis. Actually, studies have already started in our laboratory to clone enzymes to GoldenBraid system (Sarrion-Perdigones *et al*., 2013), which allows cloning of numerous genes in a single plasmid. Tests will be first carried out by transient expression in tobacco leaves to assess impact on oil synthesis.

Analysis of data gave us hints as to possible important genes to further study. Most interesting are PA phosphatases isoforms whose transcription pattern clearly points at distinct roles for phospholipid and TAG synthesis, which is consistent with double *pah1pah2* mutant studies in *Arabidopsis* (Craddock *et al*., 2015). Also, the spectacular expression pattern of LPCATc isoform points to a specific function of this enzyme for

oil editing in both mesocarp and kernel. CCTs are key enzymes for regulating PC synthesis in eukaryotic cells. Interestingly, CCTa is the only isoform that shows interactions with TAG-specific synthesizing enzymes and it is clearly expressed at higher levels in mesocarp. This suggests that it may play a role in keeping membrane PC levels adequate despite likely important flux, at least due to LPCATs, of fatty acids linked to oil accumulation in mesocarp. Transcriptome data from Bourgis *et al*. (2011) showed that a putative bifunctional wax ester synthase/DGAT, that we designated EgWSD1-like, exhibited higher transcript levels in oil palm mesocarp than in date palm, suggesting possible role in oil accumulation. Although our new transcriptome results do not point to significant up-regulation in mesocarp, interactome studies indicated a network very similar to that of EgDGAT2-1. However, yeast expression studies showed no complementation of TAG-less phenotype of yeast quadruple mutant. Only weak isoamyl and wax ester synthesis could be attributed to EgWSD1-like enzyme when expressed in yeast. Because it might lack a mesocarp-specific factor to carry out DGAT function, we co-expressed it with EgLPPβ2 in yeast TAG-less mutant, again with no success. Further studies are likely required to demonstrate a clear function for this enzyme.

Al last, It is clear that more fundamental studies are required for improving oil palm as a crop and we hope that we have brought our contribution. More specifically, it is clear that our study on variability of fatty acid content of palm genotypes might more or less rapidly allow breeders to select for new varieties of palm with lower saturated fat. This oil will appeal more to consumers and is likely to be healthier as well.

References

Alb, J. G., Kearns, M. A., & Bankaitis, V. A. (1996). Phospholipid metabolism and membrane dynamics. *Current Opinion in Cell Biology*, *8*(4), 534–541. https://doi.org/10.1016/S0955-0674(96)80032-9

Allen, D. K., Ohlrogge, J. B., & Shachar-Hill, Y. (2009). The role of light in soybean seed filling metabolism. *Plant Journal*, *58*(2), 220–234. https://doi.org/10.1111/j.1365-313X.2008.03771.x

Andersson, M. X., Goksör, M., & Sandelius, A. S. (2007). Optical manipulation reveals strong attracting forces at membrane contact sites between endoplasmic reticulum and chloroplasts. *Journal of Biological Chemistry*, *282*(2), 1170–1174. https://doi.org/10.1074/jbc.M608124200

Andre, C., Haslam, R. P., & Shanklin, J. (2012). Feedback regulation of plastidic acetyl-CoA carboxylase by 18:1-acyl carrier protein in Brassica napus. *Proceedings of the National Academy of Sciences*, *109*(25), 10107–10112. https://doi.org/10.1073/pnas.1204604109

Aymé, L., Baud, S., Dubreucq, B., Joffre, F., & Chardot, T. (2014). Function and localization of the Arabidopsis thaliana diacylglycerol acyltransferase DGAT2 expressed in yeast. *PLoS ONE*, *9*(3), 1–9. https://doi.org/10.1371/journal.pone.0092237

Aymé, L., Jolivet, P., Nicaud, J. M., & Chardot, T. (2015). Molecular characterization of the elaeis guineensis medium-chain fatty acid diacylglycerol acyltransferase DGAT1-1 by heterologous expression in yarrowia lipolytica. *PLoS ONE*, *10*(11), 1–21. https://doi.org/10.1371/journal.pone.0143113

Barcelos, E., Rios, S. de A., Cunha, R. N. V., Lopes, R., Motoike, S. Y., Babiychuk, E., … Kushnir, S. (2015). Oil palm natural diversity and the potential for yield improvement. *Frontiers in Plant Science*, *6*(March), 190. https://doi.org/10.3389/fpls.2015.00190

Bates, P. D. (2016). Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1861*(9), 1214–1225. https://doi.org/10.1016/j.bbalip.2016.03.021

Bates, P. D., & Browse, J. (2012). The significance of different diacylgycerol synthesis pathways on plant oil composition and bioengineering. *Frontiers in Plant Science*, *3*(July), 147. https://doi.org/10.3389/fpls.2012.00147

Bates, P. D., Durrett, T. P., Ohlrogge, J. B., & Pollard, M. (2009). Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. *Plant Physiology*, *150*(1), 55–72. https://doi.org/10.1104/pp.109.137737

Bates, P. D., Fatihi, A., Snapp, a. R., Carlsson, a. S., Browse, J., & Lu, C. (2012). Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. *Plant Physiology*, *160*(3), 1530–1539. https://doi.org/10.1104/pp.112.204438

Bates, P. D., Johnson, S. R., Cao, X., Li, J., Nam, J.-W., Jaworski, J. G., … Browse, J. (2014). Fatty acid synthesis is inhibited by inefficient utilization of unusual fatty acids for glycerolipid assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(3), 1204–9. https://doi.org/10.1073/pnas.1318511111

Bates, P. D., Ohlrogge, J. B., & Pollard, M. (2007). Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. *Journal of Biological Chemistry*, *282*(43), 31206–31216. https://doi.org/10.1074/jbc.M705447200

Bates, P. D., Stymne, S., & Ohlrogge, J. (2013). Biochemical pathways in seed oil synthesis. *Current*

Opinion in Plant Biology, *16*(3), 358–364. https://doi.org/10.1016/j.pbi.2013.02.015

Baud, S., & Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Progress in Lipid Research*, *49*(3), 235–249. https://doi.org/10.1016/j.plipres.2010.01.001

Baud, S., Mendoza, M. S., To, A., Harscoët, E., Lepiniec, L., & Dubreucq, B. (2007). WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant Journal*, *50*(5), 825–838. https://doi.org/10.1111/j.1365-313X.2007.03092.x

Baye, T., Becker, H. C., & Witzke-Ehbrecht, S. v. (2005). Vernonia galamensis, a natural source of epoxy oil: variation in fatty acid composition of seed and leaf lipids. *Industrial Crops and Products*, *21*(2), 257–261. https://doi.org/10.1016/j.indcrop.2004.04.003

Beisson, F., Li, Y., Bonaventure, G., Pollard, M., & Ohlrogge, J. B. (2007). The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *The Plant Cell*, *19*(1), 351–368. https://doi.org/10.1105/tpc.106.048033

Benning, C. (2009). Mechanisms of Lipid Transport Involved in Organelle Biogenesis in Plant Cells. *Annual Review of Cell and Developmental Biology*, *25*(1), 71–91. https://doi.org/10.1146/annurev.cellbio.042308.113414

Bernard, A., & Joubès, J. (2013). Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. *Progress in Lipid Research*, *52*(1), 110–129. https://doi.org/10.1016/j.plipres.2012.10.002

Bernsel, A., Viklund, H., Hennerdal, A., & Elofsson, A. (2009). TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Research*, *37*(Web Server issue), W465-8. https://doi.org/10.1093/nar/gkp363

Bligh, E. G., & Dyer, W. J. (1959). A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Canadian Journal of Biochemistry and Physiology*, *37*(8), 911–917. https://doi.org/10.1139/o59-099

Bourgis, F., Kilaru, A., Cao, X., Ngando-ebongue, G.-F., Drira, N., Ohlrogge, J. B., … Arondel, V. (2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(30), 12527–12532. https://doi.org/10.1073/pnas.1106502108

Bouvier-Navé, P., Berna, A., Noiriel, A., Compagnon, V., Carlsson, A. S., Banas, A., … Schaller, H. (2010). Involvement of the phospholipid sterol acyltransferase1 in plant sterol homeostasis and leaf senescence. *Plant Physiology*, *152*(1), 107–19. https://doi.org/10.1104/pp.109.145672

Brough, C. L., Coventry, J. M., Christie, W. W., Kroon, J. T. M., Brown, A. P., Barsby, T. L., & Slabas, A. R. (1996). Towards the genetic engineering of triacylglycerols of defined fatty acid composition: major changes in erucic acid content at the sn-2 position affected by the introduction of a 1-acyl-sn-glycerol-3-phosphate acyltransferase from Limnanthes douglasii int. *Molecular Breeding*, *2*(2), 133–142. https://doi.org/10.1007/BF00441428

Brown, A. P., Brough, C. L., Kroon, J. T. M., & Slabas, A. R. (1995). Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acyltransferase from Limnanthes douglasii. *Plant Molecular Biology*, *29*(2), 267–278. https://doi.org/10.1007/BF00043651

Brown, A. P., Johnson, P., Rawsthorne, S., & Hills, M. J. (1998). Expression and properties of acyl-CoA binding protein from Brassica napus. *Plant Physiology and Biochemistry*, *36*(9), 629–635. https://doi.org/10.1016/S0981-9428(98)80011-9

Buchanan, B. B., Gruissem, W., & Jones, R. L. (2000). *Biochemistry and Molecular Biology of Plants* (Eds.). American Society of Plant Biologists. Retrieved from http://www.wiley.com/legacy/wileychi/buchanan/

Burgal, J., Shockey, J., Lu, C., Dyer, J., Larson, T., Graham, I., & Browse, J. (2008). Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnology Journal*, *6*(8), 819–831. https://doi.org/10.1111/j.1467-7652.2008.00361.x

Cahoon, E. B., & Ohlrogge, J. B. (1994). Apparent Role of Phosphatidylcholine in the Metabolism of Petroselinic Acid in Developing Umbelliferae Endosperm. *Plant Physiology*, *104*(3), 845–855. https://doi.org/10.1104/PP.104.3.845

Cao, J., Li, J.-L., Li, D., Tobin, J. F., & Gimeno, R. E. (2006). Molecular identification of microsomal acyl-CoA:glycerol-3-phosphate acyltransferase, a key enzyme in de novo triacylglycerol synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(52), 19695–19700. https://doi.org/10.1073/pnas.0609140103

Cao, X., Bourgis, F., Kilaru, A., Cao, X., Ngando-ebongue, G., Drira, N., & Ohlrogge, J. B. (2011). Correction for Bourgis et al., Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proceedings of the National Academy of Sciences*, *108*(44), 18186–18186. https://doi.org/10.1073/pnas.1115243108

Carman, G. M. (1997). Phosphatidate phosphatases and diacylglycerol pyrophosphate phosphatases in Saccharomyces cerevisiae and Escherichia coli. *Biochimica et Biophysica Acta - Lipids and Lipid Metabolism*, *1348*(1–2), 45–55. https://doi.org/10.1016/S0005-2760(97)00095-7

Carman, G. M., & Han, G. S. (2006). Roles of phosphatidate phosphatase enzymes in lipid metabolism. *Trends in Biochemical Sciences*, *31*(12), 694–699. https://doi.org/10.1016/j.tibs.2006.10.003

Carman, G. M., & Han, G.-S. (2009). Phosphatidic acid phosphatase, a key enzyme in the regulation of lipid synthesis. *The Journal of Biological Chemistry*, *284*(5), 2593–2597. https://doi.org/10.1074/jbc.R800059200

Cernac, A., & Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant Journal*, *40*(4), 575–585. https://doi.org/10.1111/j.1365-313X.2004.02235.x

Chapman, K. D., & Ohlrogge, J. B. (2012). Compartmentation of triacylglycerol accumulation in plants. *Journal of Biological Chemistry*, *287*(4), 2288–2294. https://doi.org/10.1074/jbc.R111.290072

Chen, B., Wang, J., Zhang, G., Liu, J., Manan, S., Hu, H., & Zhao, J. (2016). Two types of soybean diacylglycerol acyltransferases are differentially involved in triacylglycerol biosynthesis and response to environmental stresses and hormones. *Scientific Reports*, *6*(June), 28541. https://doi.org/10.1038/srep28541

Chen, Q. F., Xiao, S., Qi, W., Mishra, G., Ma, J., Wang, M., & Chye, M. L. (2010). The Arabidopsis acbp1acbp2 double mutant lacking acyl-CoA-binding proteins ACBP1 and ACBP2 is embryo lethal. *New Phytologist*, *186*(4), 843–855. https://doi.org/10.1111/j.1469-8137.2010.03231.x

Chen, Q.-F., Xiao, S., & Chye, M.-L. (2008). Overexpression of the Arabidopsis 10-kilodalton acyl-coenzyme A-binding protein ACBP6 enhances freezing tolerance. *Plant Physiology*, *148*(1), 304–315. https://doi.org/10.1104/pp.108.123331

Cheng, J. B., & Russell, D. W. (2004). Mammalian wax biosynthesis: II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family. *Journal of Biological Chemistry*, *279*(36), 37798–37807. https://doi.org/10.1074/jbc.M406226200

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, *18*(24), 5294–5299. https://doi.org/10.1021/bi00591a005

Corley, R. H. V., & Tinker, P. B. (Eds.). (2003). *The Oil Palm*. Oxford, UK: Blackwell Science Ltd. https://doi.org/10.1002/9780470750971

Corley, R. H. V. (2009). How much palm oil do we need? *Environmental Science and Policy*, *12*(2), 134–139. https://doi.org/10.1016/j.envsci.2008.10.011

Cornell, R. B., & Ridgway, N. D. (2015). CTP:phosphocholine cytidylyltransferase: Function, regulation, and structure of an amphitropic enzyme required for membrane biogenesis. *Progress in Lipid Research*, *59*, 147–171. https://doi.org/10.1016/j.plipres.2015.07.001

Craddock, C. P., Adams, N., Bryant, F. M., Kurup, S., & Eastmond, P. J. (2015). PHOSPHATIDIC ACID PHOSPHOHYDROLASE Regulates Phosphatidylcholine Biosynthesis in Arabidopsis by Phosphatidic Acid-Mediated Activation of CTP:PHOSPHOCHOLINE CYTIDYLYLTRANSFERASE Activity. *The Plant Cell*, *27*(4), 1251–1264. https://doi.org/10.1105/tpc.15.00037

CROMBIE, W. M. (1956). Fat Metabolism in the West African Oil Palm (*Elaeis Guineensis*). *Journal of Experimental Botany*, *7*(2), 181–193. https://doi.org/10.1093/jxb/7.2.181

Dahlqvist, a, Stahl, U., Lenman, M., Banas, a, Lee, M., Sandager, L., … Stymne, S. (2000). Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(12), 6487–6492. https://doi.org/10.1073/pnas.120067297

Davies, H. M., Anderson, L., Fan, C., & Hawkins, D. J. (1991). Developmental induction, purification, and further characterization of 12:0-ACP thioesterase from immature cotyledons of Umbellularia californica. *Archives of Biochemistry and Biophysics*, *290*(1), 37–45. https://doi.org/10.1016/0003-9861(91)90588-A

Dehesh, K., Edwards, P., Hayes, T., Cranmer, a. M., & Fillatti, J. (1996). Two Nove1 Thioesterases Are Key Determinants of the Bimodal Distribution of Acyl Chain Length of Cuphea palustvis Seed Oil. *Plant Physiology*, *110*(1), 203–10. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=157710&tool=pmcentrez&rendertype=abstract

Dehesh, K., Jones, a., Knutzon, D. S., & Voelker, T. a. (1996). Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from Cuphea hookeriana. *Plant Journal*. https://doi.org/10.1046/j.1365-313X.1996.09020167.x

Dewey, R. E., Wilson, R. F., Novitzky, W. P., & Goode, J. H. (1994). The AAPT1 gene of soybean complements a cholinephosphotransferase-deficient mutant of yeast. *The Plant Cell*, *6*(10), 1495–1507. https://doi.org/10.1105/tpc.6.10.1495

Dittrich-Domergue, F., Joubès, J., Moreau, P., Lessire, R., Stymne, S., & Domergue, F. (2014). The bifunctional protein TtFARAT from Tetrahymena thermophila catalyzes the formation of both precursors required to initiate ether lipid biosynthesis. *Journal of Biological Chemistry*, *289*(32), 21984–21994. https://doi.org/10.1074/jbc.M114.579318

Dörmann, P., Spener, F., & Ohlrogge, J. B. (1993). Characterization of two acyl-acyl carrier protein thioesterases from developing Cuphea seeds specific for medium-chain- and oleoyl-acyl carrier protein. *Planta*, *189*(3), 425–432. https://doi.org/10.1007/BF00194441

Dörmann, P., Voelker, T. A., & Ohlrogge, J. B. (1995). Cloning and expression in Escherichia coli of a

novel thioesterase from Arabidopsis thaliana specific for long-chain acyl-acyl carrier proteins. *Archives of Biochemistry and Biophysics*. https://doi.org/10.1006/abbi.1995.1081

Dörmann, P., Voelker, T. a, & Ohlrogge, J. B. (2000). Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1. *Plant Physiology*, *123*(2), 637–644. https://doi.org/10.1104/pp.123.2.637

Dussert, S., Guerin, C., Andersson, M., Joet, T., Tranbarger, T. J., Pizot, M., … Morcillo, F. (2013). Comparative Transcriptome Analysis of Three Oil Palm Fruit and Seed Tissues That Differ in Oil Content and Fatty Acid Composition. *Plant Physiology*, *162*(3), 1337–1358. https://doi.org/10.1104/pp.113.220525

Dyer, J. M., Stymne, S., Green, A. G., & Carlsson, A. S. (2008). High-value oils from plants. *Plant Journal*, *54*(4), 640–655. https://doi.org/10.1111/j.1365-313X.2008.03430.x

Eastmond, P. J. (2006). SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating Arabidopsis seeds. *The Plant Cell*, *18*(3), 665–75. https://doi.org/10.1105/tpc.105.040543

Eastmond, P. J., Quettier, A.-L., Kroon, J. T. M., Craddock, C., Adams, N., & Slabas, A. R. (2010). PHOSPHATIDIC ACID PHOSPHOHYDROLASE1 and 2 regulate phospholipid synthesis at the endoplasmic reticulum in *Arabidopsis*. *The Plant Cell*, *22*(8), 2796–2811. https://doi.org/10.1105/tpc.109.071423

Ebongue, G. F. N., Koona, P., Nouy, B., Zok, S., Carrière, F., Zollo, P.-H. A., & Arondel, V. (2008). Identification of oil palm breeding lines producing oils with low acid values. *European Journal of Lipid Science and Technology*, *110*(6), 505–509. https://doi.org/10.1002/ejlt.200700263

Eccleston, V., & Ohlrogge, J. (1998). Expression of lauroyl-acyl carrier protein thioesterase in brassica napus seeds induces pathways for both fatty acid oxidation and biosynthesis and implies a set point for triacylglycerol accumulation. *The Plant Cell*, *10*(4), 613–622. https://doi.org/10.1105/tpc.10.4.613

Feria Bourrellier, A. B., Valot, B., Guillot, A., Ambard-Bretteville, F., Vidal, J., & Hodges, M. (2010). Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 502–507. https://doi.org/10.1073/pnas.0910097107

Fox, S. R., Rawsthorne, S., & Hills, M. J. (2000). Role of acyl-CoAs and acyl-CoA-binding protein in regulation of carbon supply for fatty acid biosynthesis. *Biochem Soc Trans*, *28*(6), 672–674. https://doi.org/10.1042/0300-5127:0280672

Fulda, M., Schnurr, J., Abbadi, A., & Heinz, E. (2004). Peroxisomal Acyl-CoA synthetase activity is essential for seedling development in Arabidopsis thaliana. *The Plant Cell Online*, *16*(February), 394–405. https://doi.org/10.1105/tpc.019646.Four

Fulda, M., Shockey, J., Werber, M., Wolter, F. P., & Heinz, E. (2002). Two long-chain acyl-CoA synthetases from Arabidopsis thaliana involved in peroxisomal fatty acid ??-oxidation. *Plant Journal*, *32*(1), 93–103. https://doi.org/10.1046/j.1365-313X.2002.01405.x

Gao, W., Xiao, S., Li, H. Y., Tsao, S. W., & Chye, M. L. (2009). Arabidopsis thaliana acyl-CoA-binding protein ACBP2 interacts with heavy-metal-binding farnesylated protein AtFP6. *New Phytologist*, *181*(1), 89–102. https://doi.org/10.1111/j.1469-8137.2008.02631.x

Garab, G., Ughy, B., & Goss, R. (2016). Role of MGDG and non-bilayer lipid phases in the structure and dynamics of chloroplast thylakoid membranes. *Lipids in Plant and Algae Development*. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-25979-6_6

Gidda, S. K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P., … Mullen, R. T. (2016). Lipid Droplet-Associated Proteins (LDAPs) Are Required for the Dynamic Regulation of Neutral Lipid Compartmentation in Plant Cells. *Plant Physiology*, *170*(4), 2052–71. https://doi.org/10.1104/pp.15.01977

Gidda, S. K., Shockey, J. M., Rothstein, S. J., Dyer, J. M., & Mullen, R. T. (2009). Arabidopsis thaliana GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: Functional divergence of the dilysine ER retrieval motif in plant cells. *Plant Physiology and Biochemistry*, *47*(10), 867–879. https://doi.org/10.1016/j.plaphy.2009.05.008

Gidda, S. K., Watt, S. C., Collins-Silva, J., Kilaru, A., Arondel, V., Yurchenko, O., … Dyer, J. M. (2013). Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. *Plant Signaling & Behavior*, *8*(11), e27141. https://doi.org/10.4161/psb.27141

Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, *11*(4), 355–360. https://doi.org/10.1002/yea.320110408

Graham, I. A. (2008). Seed Storage Oil Mobilization. *Annual Review of Plant Biology*, *59*(1), 115–142. https://doi.org/10.1146/annurev.arplant.59.032607.092938

Grimberg, Å., Carlsson, A. S., Marttila, S., Bhalerao, R., & Hofvander, P. (2015). Transcriptional transitions in Nicotiana benthamiana leaves upon induction of oil synthesis by WRINKLED1 homologs from diverse species and tissues. *BMC Plant Biology*, *15*(1), 192. https://doi.org/10.1186/s12870-015-0579-1

Grison, M. S., Brocard, L., Fouillen, L., Nicolas, W., Wewer, V., Dörmann, P., … Bayer, E. M. (2015). Specific membrane lipid composition is important for plasmodesmata function in Arabidopsis. *The Plant Cell*, *27*(4), 1228–50. https://doi.org/10.1105/tpc.114.135731

Guerin, C., Joët, T., Serret, J., Lashermes, P., Vaissayre, V., Agbessi, M. D. T., … Dussert, S. (2016). Gene coexpression network analysis of oil biosynthesis in an interspecific backcross of oil palm. *The Plant Journal*, *87*(5), 423–441. https://doi.org/10.1111/tpj.13208

Han, G. S., Wu, W. I., & Carman, G. M. (2006). The Saccharomyces cerevisiae lipin homolog is a Mg2+-dependent phosphatidate phosphatase enzyme. *Journal of Biological Chemistry*, *281*(14), 9210–9218. https://doi.org/10.1074/jbc.M600425200

Hanson, M. R., & Sattarzadeh, A. (2011). Stromules: recent insights into a long neglected feature of plastid morphology and function. *Plant Physiol*, *155*(April), 1486–1492. https://doi.org/10.1104/pp.110.170852

Hawkins, D. J., & Kridl, J. C. (1998). Characterization of acyl-ACP thioesterases of mangosteen (Garcinia mangostana) seed and high levels of stearate production in transgenic canola. *Plant Journal*, *13*(6), 743–752. https://doi.org/10.1046/j.1365-313X.1998.00073.x

Hay, J., & Schwender, J. (2011). Computational analysis of storage synthesis in developing Brassica napus L. (oilseed rape) embryos: Flux variability analysis in relation to 13C metabolic flux analysis. *Plant Journal*, *67*(3), 513–525. https://doi.org/10.1111/j.1365-313X.2011.04611.x

Hellyer, A., Leadlay, P. F., & Slabas, A. R. (1992). from developing seeds of oil seed rape (Brassica napus), 763–780.

Hernandez, M. L., Whitehead, L., He, Z., Gazda, V., Gilday, A., Kozhevnikova, E., … Graham, I. A. (2012). A Cytosolic Acyltransferase Contributes to Triacylglycerol Synthesis in Sucrose-Rescued Arabidopsis Seed Oil Catabolism Mutants. *Plant Physiology*, *160*(1), 215–225.

https://doi.org/10.1104/pp.112.201541

Hobbs, D. H., Flintham, J. E., & Hills, M. J. (2004). Genetic control of storage oil synthesis in seeds of Arabidopsis. *Plant Physiology*, *136*(2), 3341–9. https://doi.org/10.1104/pp.104.049486

Hu, Z., Ren, Z., & Lu, C. (2012). The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic Arabidopsis. *Plant Physiology*, *158*(4), 1944–54. https://doi.org/10.1104/pp.111.192153

Huang, A. H. (1996). Oleosins and oil bodies in seeds and other organs. *Plant Physiology*, *110*(4), 1055–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8934621

Ito, H., Fukuda, Y., & Murata, K. (1983). Transformation of intact yeast cells treated with alkali Transformation of Intact Yeast Cells Treated with Alkali Cations, *153*(1), 166–168.

Jako, C. (2001). Seed-Specific Over-Expression of an Arabidopsis cDNA Encoding a Diacylglycerol Acyltransferase Enhances Seed Oil Content and Seed Weight. *Plant Physiology*, *126*(2), 861–874. https://doi.org/10.1104/pp.126.2.861

Jasieniecka-Gazarkiewicz, K., Demski, K., Lager, I., Stymne, S., & Banaś, A. (2016). Possible Role of Different Yeast and Plant Lysophospholipid:Acyl-CoA Acyltransferases (LPLATs) in Acyl Remodelling of Phospholipids. *Lipids*, *51*(1), 15–23. https://doi.org/10.1007/s11745-015-4102-0

Jessen, D., Roth, C., Wiermer, M., & Fulda, M. (2014). Two activities of long-chain acyl-CoA synthetase are involved in lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis. *Plant Physiology*, *167*(2), 351–366. https://doi.org/10.1104/pp.114.250365

Jha, J. K., Maiti, M. K., Bhattacharjee, A., Basu, A., Sen, P. C., & Sen, S. K. (2006). Cloning and functional expression of an acyl-ACP thioesterase FatB type from Diploknema (Madhuca) butyracea seeds in Escherichia coli. *Plant Physiology and Biochemistry*, *44*(11–12), 645–655. https://doi.org/10.1016/j.plaphy.2006.09.017

Jha, S. S., Jha, J. K., Chattopadhyaya, B., Basu, A., Sen, S. K., & Maiti, M. K. (2010). Cloning and characterization of cDNAs encoding for long-chain saturated acyl-ACP thioesterases from the developing seeds of Brassica juncea. *Plant Physiology and Biochemistry*, *48*(6), 476–480. https://doi.org/10.1016/j.plaphy.2010.02.006

Jing, F., Cantu, D. C., Tvaruzkova, J., Chipman, J. P., Nikolau, B. J., Yandeau-Nelson, M. D., & Reilly, P. J. (2011). Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochemistry*, *12*(1), 44. https://doi.org/10.1186/1471-2091-12-44

Johnston, M. L., Luethy, M. H., Miernyk, J. A., & Randall, D. D. (1997). Cloning and molecular analyses of the Arabidopsis thaliana plastid pyruvate dehydrogenase subunits. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1321*(3), 200–206. https://doi.org/10.1016/S0005-2728(97)00059-5

Jones, A., Davies, H. M., & Voelker, T. a. (1995). Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. *The Plant Cell*, *7*(March), 359–371. https://doi.org/10.1105/tpc.7.3.359

Jouannic, S., Argout, X., Lechauve, F., Fizames, C., Borgel, A., Morcillo, F., … Tregear, J. (2005). Analysis of expressed sequence tags from oil palm (Elaeis guineensis). *FEBS Letters*, *579*(12), 2709–2714. https://doi.org/10.1016/j.febslet.2005.03.093

Jouhet, J., Maréchal, E., & Block, M. A. (2007). Glycerolipid transfer for the building of membranes in plant cells. *Progress in Lipid Research*, *46*(1), 37–55. https://doi.org/10.1016/j.plipres.2006.06.002

Käll, L., Krogh, A., & Sonnhammer, E. L. . (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology*, *338*(5), 1027–1036. https://doi.org/10.1016/j.jmb.2004.03.016

Kalscheuer, R., Luftmann, H., & Steinbu, A. (2004). Synthesis of Novel Lipids in Saccharomyces cerevisiae by Heterologous Expression of an Unspecific Bacterial Acyltransferase Synthesis of Novel Lipids in Saccharomyces cerevisiae by Heterologous Expression of an Unspecific Bacterial Acyltransferase. *Applied and Env. Microbiol.*, *70*(12), 7119–7125. https://doi.org/10.1128/AEM.70.12.7119

Kalscheuer, R., & Steinbüchel, A. (2003). A novel bifunctional wax ester synthase/acyl-CoA:Diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in Acinetobacter calcoaceticus ADP1. *Journal of Biological Chemistry*, *278*(10), 8075–8082. https://doi.org/10.1074/jbc.M210533200

Katagiri, T., Ishiyama, K., Kato, T., Tabata, S., Kobayashi, M., & Shinozaki, K. (2005). An important role of phosphatidic acid in ABA signaling during germination in Arabidopsis thaliana. *Plant Journal*, *43*(1), 107–117. https://doi.org/10.1111/j.1365-313X.2005.02431.x

Katavic, V., Reed, D. W., Taylor, D. C., Ciblin, E. M., Zou, J., Mackenzie, S., … Kunst, L. (1995). Alteration of Seed Fatty Acid Composition by an Ethyl Methanesulfonate-lnduced Mutation in Arabidomis fhaliana Affecting Diacylglycerol Acyltransferase Activity. *Plant Physiology*, *108*(1 995), 399–409.

Kawelke, S., & Feussner, I. (2015). Two predicted transmembrane domains exclude very long chain fatty acyl-CoAs from the active site of mouse wax synthase. *PLoS ONE*, *10*(12), 1–19. https://doi.org/10.1371/journal.pone.0145797

Kawelke, S., Feussner, I., Samuels, L., Kunst, L., Jetter, R., Butovich, I., … Li, W. (2015). Two Predicted Transmembrane Domains Exclude Very Long Chain Fatty acyl-CoAs from the Active Site of Mouse Wax Synthase. *PLOS ONE*, *10*(12), e0145797. https://doi.org/10.1371/journal.pone.0145797

Kim, H. U., & Huang, A. H. C. (2004). Plastid lysophosphatidyl acyltransferase is essential for embryo development in Arabidopsis. *Plant Physiology*, *134*(3), 1206–16. https://doi.org/10.1104/pp.103.035832

Kim, H. U., Li, Y., & Huang, A. H. C. (2005). Ubiquitous and endoplasmic reticulum-located lysophosphatidyl acyltransferase, LPAT2, is essential for female but not male gametophyte development in Arabidopsis. *The Plant Cell*, *17*(4), 1073–1089. https://doi.org/10.1105/tpc.104.030403

Kim, S., Yamaoka, Y., Ono, H., Kim, H., Shim, D., Maeshima, M., … Lee, Y. (2013). AtABCA9 transporter supplies fatty acids for lipid synthesis to the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(2), 773–8. https://doi.org/10.1073/pnas.1214159110

King, A., Nam, J. W., Han, J., Hilliard, J., & Jaworski, J. G. (2007). Cuticular wax biosynthesis in petunia petals: Cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. *Planta*, *226*(2), 381–394. https://doi.org/10.1007/s00425-007-0489-z

Knutzon, D., Hayes, T., Wyrick, A., Xiong, H., Maelor Davies H, & Voelker, T. (1999). Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the sn-2 position of triacylglycerols in lauric rapeseed oil and can increase total laurate levels. *Plant Physiology*, *120*(3), 739–46. https://doi.org/10.1104/pp.120.3.739

Knutzon, D. S., Bleibaum, J. L., Nelsen, J., & Thompson, G. A. (1992). Isolation and Characterization of Two Safflower Oleoyl-Acyl Carrier Protein Thioesterase cDNA Clones. *Plant Physiol.*, *100*(4), 1751–1758. https://doi.org/10.1104/pp.100.4.1751

Knutzon, D. S., Lardizabal, K. D., Nelsen, J. S., Bleibaum, J. L., Davies, H. M., & Metz, J. G. (1995). Cloning of a coconut endosperm cDNA encoding a 1-acyl-sn-glycerol-3-phosphate acyltransferase that accepts medium-chain-length substrates. *Plant Physiology*, *109*(3), 999–1006. https://doi.org/10.1104/pp.109.3.999

Kojima, M., Casteel, J., Miernyk, J. A., & Thelen, J. J. (2007). The effects of down-regulating expression of Arabidopsis thaliana membrane-associated acyl-CoA binding protein 2 on acyl-lipid composition. *Plant Science*, *172*(1), 36–44. https://doi.org/10.1016/j.plantsci.2006.07.009

Kok, S.-Y., Namasivayam, P., Ee, G. C.-L., & Ong-Abdullah, M. (2013). Biochemical characterisation during seed development of oil palm (Elaeis guineensis). *Journal of Plant Research*, *126*(4), 539–47. https://doi.org/10.1007/s10265-013-0560-8

Konishi, T., Shinohara, K., Yamada, K., & Sasaki, Y. (1996). Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant & Cell Physiology*, *37*(2), 117–22. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8665091

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. . (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567–580. https://doi.org/10.1006/jmbi.2000.4315

Kruger, N. J., & Ratcliffe, R. G. (2015). Fluxes through plant metabolic networks: measurements, predictions, insights and challenges. *Biochemical Journal*, *465*(1).

Kunst, L., Browse, J., & Somerville, C. (1988). Altered regulation of lipid biosynthesis in a mutant of Arabidopsis deficient in chloroplast glycerol-3-phosphate acyltransferase activity. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(12), 4143–7. https://doi.org/10.1073/pnas.85.12.4143

Kunst, L., & Samuels, A. L. (2003). Biosynthesis and secretion of plant cuticular wax. *Progress in Lipid Research*, *42*(1), 51–80. https://doi.org/10.1016/S0163-7827(02)00045-0

Kunst, L., & Samuels, L. (2009). Plant cuticles shine: advances in wax biosynthesis and export. *Current Opinion in Plant Biology*, *12*(6), 721–727. https://doi.org/10.1016/j.pbi.2009.09.009

Lacey, D. J., Beaudoin, F., Dempsey, C. E., Shewry, P. R., & Napier, J. a. (1999). The accumulation of triacylglycerols within the endoplasmic reticulum of developing seeds ofHelianthus annuus. *The Plant Journal*, *17*(October 1998), 397–405. https://doi.org/10.1046/j.1365-313X.1999.00387.x

Lacey, D. J., & Hills, M. J. (1996). Heterogeneity of the endoplasmic reticulum with respect to lipid synthesis in developing seeds of Brassica napus L. *Planta*, *199*(4), 545–551. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0029840509&partnerID=tZOtx3y1

Lager, I., Yilmaz, J. L., Zhou, X. R., Jasieniecka, K., Kazachkov, M., Wang, P., … Stymne, S. (2013). Plant Acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. *Journal of Biological Chemistry*, *288*(52), 36902–36914. https://doi.org/10.1074/jbc.M113.521815

Lardizabal, K. D., Metz, J. G., Sakamoto, T., Hutton, W. C., Pollard, M. R., & Lassner, M. W. (2000). Purification of a jojoba embryo wax synthase, cloning of its cDNA, and production of high levels of wax in seeds of transgenic arabidopsis. *Plant Physiology*, *122*(March), 645–655.

https://doi.org/10.1104/pp.122.3.645

Lassner, M. W., Levering, C. K., Davies, H. M., & Knutzon, D. S. (1995). Lysophosphatidic acid acyltransferase from meadowfoam mediates insertion of erucic acid at the sn-2 position of triacylglycerol in transgenic rapeseed oil. *Plant Physiology*, *109*(4), 1389–1394. https://doi.org/10.1104/pp.109.4.1389

Lee, S. B., & Suh, M. C. (2015). Advances in the understanding of cuticular waxes in Arabidopsis thaliana and crop species. *Plant Cell Reports*, *34*(4), 557–572. https://doi.org/10.1007/s00299-015-1772-2

Lemieux, B., Miquel, M., Somerville, C., & Browse, J. (1990). Mutants of Arabidopsis with alterations in seed lipid fatty acid composition. *Theoretical and Applied Genetics*, *80*(2), 234–240. https://doi.org/10.1007/BF00224392

Leung, K. C., Li, H. Y., Xiao, S., Tse, M. H., & Chye, M. L. (2006). Arabidopsis ACBP3 is an extracellularly targeted acyl-CoA-binding protein. *Planta*, *223*(5), 871–881. https://doi.org/10.1007/s00425-005-0139-2

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., … Kunst, L. (2008). Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. *Plant Physiology*, *148*(1), 97–107. https://doi.org/10.1104/pp.108.123471

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., … Zhang, D. (2010). Cytochrome P450 family member CYP704B2 catalyzes the {omega}-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *The Plant Cell*, *22*(1), 173–90. https://doi.org/10.1105/tpc.109.070326

Li, H. Y., & Chye, M. L. (2003). Membrane localization of Arabidopsis acyl-CoA binding protein ACBP2. *Plant Molecular Biology*, *51*(4), 483–492. https://doi.org/10.1023/A:1022330304402

Li, H. Y., & Chye, M. L. (2004). Arabidopsis Acyl-CoA-Binding Protein ACBP2 Interacts With an Ethylene-Responsive Element-Binding Protein, AtEBP, via its Ankyrin Repeats. *Plant Molecular Biology*, *54*(2), 233–243. https://doi.org/10.1023/B:PLAN.0000028790.75090.ab

Li, N., Gügel, I. L., Giavalisco, P., Zeisler, V., Schreiber, L., Soll, J., & Philippar, K. (2015). FAX1, a Novel Membrane Protein Mediating Plastid Fatty Acid Export. *PLoS Biology*, *13*(2), 1–37. https://doi.org/10.1371/journal.pbio.1002053

Li, N., Xu, C., Li-Beisson, Y., & Philippar, K. (2016). Fatty Acid and Lipid Transport in Plant Cells. *Trends in Plant Science*, *21*(2), 145–158. https://doi.org/10.1016/j.tplants.2015.10.011

Li, Y., Beisson, F., Koo, A. J., Molina, I., Pollard, M., & Ohlrogge, J. (2007). Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(46), 18339–18344. https://doi.org/10.1073/pnas.0706984104

Li, Y., Beisson, F., Ohlrogge, J., & Pollard, M. (2007). Monoacylglycerols are components of root waxes and can be produced in the aerial cuticle by ectopic expression of a suberin-associated acyltransferase. *Plant Physiology*, *144*(3), 1267–77. https://doi.org/10.1104/pp.107.099432

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M. X., Arondel, V., Bates, P. D., … Ohlrogge, J. (2013). Acyl-lipid metabolism. *The Arabidopsis Book / American Society of Plant Biologists*, *11*, e0161. https://doi.org/10.1199/tab.0161; 10.1199/tab.0161

Lin, J., Turner, C., Liao, L. P., & McKeon, T. A. (2003). Identification and Quantification of the Molecular Species of Acylglycerols in Castor Oil by HPLC Using ELSD. *Journal of Liquid*

Chromatography & Related Technologies, *26*(5), 773–780. https://doi.org/10.1081/JLC-120018421

Liu, Q., Siloto, R. M. P., Lehner, R., Stone, S. J., & Weselake, R. J. (2012). Acyl-CoA:diacylglycerol acyltransferase: Molecular biology, biochemistry and biotechnology. *Progress in Lipid Research*, *51*(4), 350–377. https://doi.org/10.1016/j.plipres.2012.06.001

Loei, H., Lim, J., Tan, M., Lim, T. K., Lin, Q. S., Chew, F. T., … Chung, M. C. M. (2013). Proteomic Analysis of the Oil Palm Fruit Mesocarp Reveals Elevated Oxidative Phosphorylation Activity is Critical for Increased Storage Oil Production. *Journal of Proteome Research*, *12*(11), 5096–5109. https://doi.org/10.1021/pr400606h

Lu, C., Xin, Z., Ren, Z., Miquel, M., & Browse, J. (2009). An enzyme regulating triacylglycerol composition is encoded by the *ROD1* gene of *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(44), 18837–18842. https://doi.org/10.1073/pnas.0908848106

Lü, S., Song, T., Kosma, D. K., Parsons, E. P., Rowland, O., & Jenks, M. A. (2009). Arabidopsis CER8 encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant Journal*, *59*(4), 553–564. https://doi.org/10.1111/j.1365-313X.2009.03892.x

Ma, W., Kong, Q., Arondel, V., Kilaru, A., Bates, P. D., Thrower, N. A., … Ohlrogge, J. B. (2013). WRINKLED1, A Ubiquitous Regulator in Oil Accumulating Tissues from Arabidopsis Embryos to Oil Palm Mesocarp. *PLoS ONE*, *8*(7), 1–13. https://doi.org/10.1371/journal.pone.0068887

Maisonneuve, S., Bessoule, J.-J., Lessire, R., Delseny, M., & Roscoe, T. J. (2010). Expression of rapeseed microsomal lysophosphatidic acid acyltransferase isozymes enhances seed oil content in Arabidopsis. *Plant Physiology*, *152*(2), 670–84. https://doi.org/10.1104/pp.109.148247

Manaf, a M., & Harwood, J. L. (2000). Purification and characterisation of acyl-CoA: glycerol 3-phosphate acyltransferase from oil palm (Elaeis guineensis) tissues. *Planta*, *210*(2), 318–328. https://doi.org/10.1007/PL00008140

Maneta-Peyret, L., Lai, Y.-S., Stefano, G., Fouillen, L., Brandizzi, F., & Moreau, P. (2014). Phospholipid biosynthesis increases in RHD3-defective mutants. *Plant Signaling & Behavior*, *9*(9), e29657. https://doi.org/10.4161/psb.29657

Masani, M. Y. A., Noll, G. A., Parveez, G. K. A., Sambanthamurthi, R., & Prüfer, D. (2014). Efficient transformation of oil palm protoplasts by PEG-mediated transfection and DNA microinjection. *PloS One*, *9*(5), e96831. https://doi.org/10.1371/journal.pone.0096831

MEISEL, L., FONSECA, B., GONZÁLEZ, S., BAEZA-YATES, R., CAMBIAZO, V., CAMPOS, R., … SILVA, H. (2005). A Rapid and Efficient Method for Purifying High Quality Total RNA from Peaches (Prunus persica) for Functional Genomics Analyses. *Biological Research*, *38*(1), 83–88. https://doi.org/10.4067/S0716-97602005000100010

Metz, J. G., Pollard, M. R., Anderson, L., Hayes, T. R., & Lassner, M. W. (2000). Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed. *Plant Physiology*, *122*(3), 635–44. https://doi.org/10.1104/PP.122.3.635

Mhaske, V., Beldjilali, K., Ohlrogge, J., & Pollard, M. (2005). Isolation and characterization of an Arabidopsis thaliana knockout line for phospholipid : diacylglycerol transacylase gene (At5g13640), *43*, 413–417. https://doi.org/10.1016/j.plaphy.2005.01.013

Morcillo, F., Cros, D., Billotte, N., Ngando-Ebongue, G.-F., Domonhédo, H., Pizot, M., … Domonhe, H. (2013). Improving palm oil quality through identification and mapping of the lipase gene

causing oil deterioration. *Nature Communications*, *4*, 2160. https://doi.org/10.1038/ncomms3160

Mozzon, M., Pacetti, D., Lucci, P., Balzano, M., & Frega, N. G. (2013). Crude palm oil from interspecific hybrid Elaeis oleifera × Elaeis guineensis: Fatty acid regiodistribution and molecular species of glycerides. *Food Chemistry*, *141*(1), 245–252. https://doi.org/10.1016/j.foodchem.2013.03.016

Munnik, T. (2001). Phosphatidic acid: An emerging plant lipid second messenger. *Trends in Plant Science*, *6*(5), 227–233. https://doi.org/10.1016/S1360-1385(01)01918-5

Murphy, D. J. (2009). Oil palm: Future prospects for yield and quality improvements. *Lipid Technology*, *21*(11–12), 257–260. https://doi.org/10.1002/lite.200900067

Nakagawa, N., Kato, M., Takahashi, Y., Shimazaki, K., Tamura, K., Tokuji, Y., … Imai, H. (2012). Degradation of long-chain base 1-phosphate (LCBP) in Arabidopsis: functional characterization of LCBP phosphatase involved in the dehydration stress response. *Journal of Plant Research*, *125*(3), 439–449. https://doi.org/10.1007/s10265-011-0451-9

Nakamura, Y., Koizumi, R., Shui, G., Shimojima, M., Wenk, M. R., Ito, T., & Ohta, H. (2009). Arabidopsis lipins mediate eukaryotic pathway of lipid metabolism and cope critically with phosphate starvation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(49), 20978–83. https://doi.org/10.1073/pnas.0907173106

Nakamura, Y., & Ohta, H. (2010). Phosphatidic Acid Phosphatases in Seed Plants (pp. 131–141). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03873-0_9

Nakamura, Y., Tsuchiya, M., & Ohta, H. (2007). Plastidic phosphatidic acid phosphatases identified in a distinct subfamily of lipid phosphate phosphatases with prokaryotic origin. *Journal of Biological Chemistry*, *282*(39), 29013–29021. https://doi.org/10.1074/jbc.M704385200

Napier, J. A. (2007). The Production of Unusual Fatty Acids in Transgenic Plants. *Annual Review of Plant Biology*, *58*(1), 295–319. https://doi.org/10.1146/annurev.arplant.58.032806.103811

Neoh, B. K., Teh, H. F., Ng, T. L. M., Tiong, S. H., Thang, Y. M., Ersad, M. A., … Appleton, D. R. (2013). Profiling of Metabolites in Oil Palm Mesocarp at Different Stages of Oil Biosynthesis. *Journal of Agricultural and Food Chemistry*, *61*(8), 1920–1927. https://doi.org/10.1021/jf304561f

Ngando Ebongue, G. F., Dhouib, R., Carrière, F., Amvam Zollo, P. H., & Arondel, V. (2006). Assaying lipase activity from oil palm fruit (Elaeis guineensis Jacq.) mesocarp. *Plant Physiology and Biochemistry*, *44*(10), 611–617. https://doi.org/10.1016/j.plaphy.2006.09.006

Nishida, I., Tasaka, Y., Shiraishi, H., & Murata, N. (1993). The gene and the RNA for the precursor to the plastid-located glycerol-3-phosphate acyltransferase of Arabidopsis thaliana. *Plant Molecular Biology*, *21*(2), 267–277. https://doi.org/10.1007/BF00019943

O'Neill, C. M., Gill, S., Hobbs, D., Morgan, C., & Bancroft, I. (2003). Natural variation for seed oil composition in Arabidopsis thaliana. *Phytochemistry*, *64*(6), 1077–1090. https://doi.org/10.1016/S0031-9422(03)00351-0

Ohlrogge, J. B., Pollard, M. R., & Stumpf, P. K. (1978). Studies on biosynthesis of waxes by developing jojoba seed tissue. *Lipids*, *13*(3), 203–210. https://doi.org/10.1007/BF02533401

Ohlrogge, J., & Browse, J. (1995). Lipid Biosynthesis. *THE PLANT CELL ONLINE*, *7*(7), 957–970. https://doi.org/10.1105/tpc.7.7.957

Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S.-E., Kok, S.-Y., Sarpan, N., … Martienssen, R. A. (2015). Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*, *525*(7570), 533–537. https://doi.org/10.1038/nature15365

Oo, K. C., Teh, S. K., Khor, H. T., & 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*(4), 205–210. https://doi.org/10.1007/BF02534189

Ooi, T. E. K., Yeap, W. C., Daim, L. D. J., Ng, B. Z., Lee, F. C., Othman, A. M., … Kulaveerasingam, H. (2015). Differential abundance analysis of mesocarp protein from high- and low-yielding oil palms associates non-oil biosynthetic enzymes to lipid biosynthesis. *Proteome Science*, *13*, 28. https://doi.org/10.1186/s12953-015-0085-2

Ooi, T. E. K., Yeap, W. C., Daim, L. D. J., Ng, B. Z., Lee, F. C., Othman, A. M., … Panis, B. (2015). Differential abundance analysis of mesocarp protein from high- and low-yielding oil palms associates non-oil biosynthetic enzymes to lipid biosynthesis. *Proteome Science*, *13*(1), 28. https://doi.org/10.1186/s12953-015-0085-2

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, *8*(10), 785–786. https://doi.org/10.1038/nmeth.1701

Pierrugues, O., Brutesco, C., Oshiro, J., Gouy, M., Deveaux, Y., Carman, G. M., … Kazmaier, M. (2001). Lipid phosphate phosphatases in Arabidopsis. Regulation of the AtLPP1 gene in response to stress. *Journal of Biological Chemistry*, *276*(23), 20300–20308. https://doi.org/10.1074/jbc.M009726200

Pollard, M., Beisson, F., Li, Y., & Ohlrogge, J. B. (2008). Building lipid barriers: biosynthesis of cutin and suberin. *Trends in Plant Science*, *13*(5), 236–246. https://doi.org/10.1016/j.tplants.2008.03.003

Pollard, M. R., Anderson, L., Fan, C., Hawkins, D. J., & Davies, H. M. (1991). A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of Umbellularia californica. *Archives of Biochemistry and Biophysics*, *284*(2), 306–312. https://doi.org/10.1016/0003-9861(91)90300-8

Prada, F., Ayala-Diaz, I. M., Delgado, W., Ruiz-Romero, R., & Romero, H. M. (2011). Effect of fruit ripening on content and chemical composition of oil from three oil palm cultivars (Elaeis guineensis Jacq.) grown in Colombia. *Journal of Agricultural and Food Chemistry*, *59*(18), 10136–10142. https://doi.org/10.1021/jf201999d

Pulsifer, I. P., Kluge, S., & Rowland, O. (2012). Arabidopsis LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1), LACS2, and LACS3 facilitate fatty acid uptake in yeast. *Plant Physiology and Biochemistry*, *51*, 31–39. https://doi.org/10.1016/j.plaphy.2011.10.003

Radin, N. S. (1981). [1] Extraction of tissue lipids with a solvent of low toxicity. *Methods in Enzymology*, *72*, 5–7. https://doi.org/10.1016/S0076-6879(81)72003-2

Rafii, M. Y., Jalani, B. S., Rajanaidu, N., Kushairi, A., Puteh, A., & Latif, M. A. (2012). Stability analysis of oil yield in oil palm (Elaeis guineensis) progenies in different environments. *Genet Mol Res*, *11*(4), 3629–3641. https://doi.org/10.4238/2012.October.4.10

Ramli, U. S., Salas, J. J., Quant, P. A., & 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 (oilpalm). *New Phytologist*, *184*(2), 330–339. https://doi.org/10.1111/j.1469-8137.2009.02962.x

Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions*, *30*(6).

Routaboul, J. M., Benning, C., Bechtold, N., Caboche, M., & Lepiniec, L. (1999). The TAG11 locus of

Arabidopsis encodes for a diacylglycerol acyltransferase. *Plant Physiology and Biochemistry*, *37*(11), 831–840. https://doi.org/10.1016/S0981-9428(99)00115-1

Rowland, O., Zheng, H., Hepworth, S. R., Lam, P., Jetter, R., & Kunst, L. (2006). CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant Physiology*, *142*(3), 866–77. https://doi.org/10.1104/pp.106.086785

Saha, S., Enugutti, B., Rajakumari, S., & Rajasekharan, R. (2006). Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiology*, *141*(4), 1533–1543. https://doi.org/10.1104/pp.106.082198

Salas, J. J., & Ohlrogge, J. B. (2002). Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Archives of Biochemistry and Biophysics*, *403*(1), 25–34. https://doi.org/10.1016/S0003-9861(02)00017-6

Sambanthamurthi, R., Sundram, K., & Tan YewAi. (2000). *Chemistry and biochemistry of palm oil. Progress in Lipid Research* (Vol. 39). https://doi.org/10.1016/S0163-7827(00)00015-1

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning*. New York: Cold spring harbor laboratory press.

Sánchez-García, A., Moreno-Pérez, A. J., Muro-Pastor, A. M., Salas, J. J., Garcés, R., & Martínez-Force, E. (2010). Acyl-ACP thioesterases from castor (Ricinus communis L.): An enzymatic system appropriate for high rates of oil synthesis and accumulation. *Phytochemistry*, *71*(8–9), 860–869. https://doi.org/10.1016/j.phytochem.2010.03.015

Sandager, L., Gustavsson, M. H., Ståhl, U., Dahlqvist, A., Wiberg, E., Banas, A., … Stymne, S. (2002). Storage lipid synthesis is non-essential in yeast. *The Journal of Biological Chemistry*, *277*(8), 6478–82. https://doi.org/10.1074/jbc.M109109200

Santoso, U. . K. K. . O. T. . T. T. . M. A. (1996). Nutrient composition of kopyor coconuts. *Food Chemistry*, *51*(2), 299–304.

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., … Orzaez, D. (2013). GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiology*, *162*(3), 1618–31. https://doi.org/10.1104/pp.113.217661

Schnurr, J. (2004). The Acyl-CoA Synthetase Encoded by LACS2 Is Essential for Normal Cuticle Development in Arabidopsis. *The Plant Cell Online*, *16*(3), 629–642. https://doi.org/10.1105/tpc.017608

Schnurr, J. A., Shockey, J. M., de Boer, G.-J., & Browse, J. A. (2002). Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. *Plant Physiology*, *129*(4), 1700–9. https://doi.org/10.1104/pp.003251

Schwender, J., Goffman, F., Ohlrogge, J. B., & Shachar-Hill, Y. (2004). Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature*, *432*(7018), 779–782. https://doi.org/10.1038/nature03145

Shindou, H., Hishikawa, D., Harayama, T., Eto, M., & Shimizu, T. (2013). Generation of membrane diversity by lysophospholipid acyltransferases. *Journal of Biochemistry*, *154*(1), 21–28. https://doi.org/10.1093/jb/mvt048

Shintani, D., & Ohlrogge, J. (1995). Feedback inhibition of fatty acid synthesis in tobacco suspension cells. *Plant J*, *7*(4), 577–587.

Shockey, J. M. (2006). Tung Tree DGAT1 and DGAT2 Have Nonredundant Functions in Triacylglycerol Biosynthesis and Are Localized to Different Subdomains of the Endoplasmic

Reticulum. *The Plant Cell Online*, *18*(9), 2294–2313. https://doi.org/10.1105/tpc.106.043695

Shockey, J. M., Fulda, M. S., & Browse, J. A. (2002). Arabidopsis contains nine long-chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism. *Plant Physiology*, *129*(4), 1710–1722. https://doi.org/10.1104/pp.003269

Shockey, J., Regmi, A., Cotton, K., Adhikari, N., Browse, J., & Bates, P. D. (2016). Identification of Arabidopsis GPAT9 (At5g60620) as an Essential Gene Involved in Triacylglycerol Biosynthesis. *Plant Physiology*, *170*(1), 163–79. https://doi.org/10.1104/pp.15.01563

Siew, W.-L., Chong, C.-L., & Tan, Y.-A. (1995). Composition of the oil in palm kernel fromElaeis guineensis. *Journal of the American Oil Chemists' Society*, *72*(12), 1587–1589. https://doi.org/10.1007/BF02577859

Singer, S. D., Chen, G., Mietkiewska, E., Tomasi, P., Jayawardhane, K., Dyer, J. M., & Weselake, R. J. (2016). Arabidopsis GPAT9 contributes to synthesis of intracellular glycerolipids but not surface lipids. *Journal of Experimental Botany*, *67*(15), 4627–4638. https://doi.org/10.1093/jxb/erw242

Singh, R., Low, E.-T. L., Ooi, L. C.-L., Ong-Abdullah, M., Ting, N.-C., Nagappan, J., … Martienssen, R. A. (2013). The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature*, *500*(7462), 340–344. https://doi.org/10.1038/nature12356

Singh, R., Ong-Abdullah, M., Low, E.-T. L., Manaf, M. A. A., Rosli, R., Nookiah, R., … Sambanthamurthi, R. (2013). Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature*, *500*(7462), 335–9. https://doi.org/10.1038/nature12309

Sperling, P., Linscheid, M., Stöcker, S., Mühlbach, H. P., & Heinz, E. (1993). In vivo desaturation of cis-delta 9-monounsaturated to cis-delta 9,12-diunsaturated alkenylether glycerolipids. *The Journal of Biological Chemistry*, *268*(36), 26935–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8262928

Srikanta Dani, K. G., Hatti, K. S., Ravikumar, P., & Kush, A. (2011). Structural and functional analyses of a saturated acyl ACP thioesterase, type B from immature seed tissue of Jatropha curcas. *Plant Biology*, *13*(3), 453–461. https://doi.org/10.1111/j.1438-8677.2010.00410.x

Stagljar, I., Korostensky, C., Johnsson, N., & Heesen, S. Te. (1998). A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Genetics*, *95*, 5187–5192.

Stålberg, K., Ståhl, U., Stymne, S., & Ohlrogge, J. (2009). Characterization of two Arabidopsis thaliana acyltransferases with preference for lysophosphatidylethanolamine. *BMC Plant Biology*, *9*(X), 60. https://doi.org/10.1186/1471-2229-9-60

Stobart, A. K., & Stymne, S. (1985). The interconversion of diacylglycerol and phosphatidylcholine during triacylglycerol production in microsomal preparations of developing cotyledons of safflower (Carthamus tinctorius L.). *Biochemical Journal*, *232*(1).

Stymne, S., & Stobart, A. K. (1984). The biosynthesis of triacylglycerols in microsomal preparations of developing cotyledons of sunflower (Helianthus annuus L.). *Biochemical Journal*, *220*(2).

Sundram, K., Sambanthamurthi, R., & Tan, Y. A. (2003). Palm fruit chemistry and nutrition. *Asia Pacific Journal of Clinical Nutrition*, *12*(3), 355–362.

Tabet, R., Moutin, E., Becker, J. A. J., Heintz, D., Fouillen, L., Flatter, E., … Moine, H. (2016). Fragile X Mental Retardation Protein (FMRP) controls diacylglycerol kinase activity in neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(26), E3619-28. https://doi.org/10.1073/pnas.1522631113

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular

evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, *28*(10), 2731–9. https://doi.org/10.1093/molbev/msr121

Tan, K. T., Lee, K. T., Mohamed, A. R., & Bhatia, S. (2009). Palm oil: Addressing issues and towards sustainable development. *Renewable and Sustainable Energy Reviews*, *13*(2), 420–427. https://doi.org/10.1016/j.rser.2007.10.001

Teerawanichpan, P., & Qiu, X. (2010). Fatty acyl-coA reductase and wax synthase from euglena gracilis in the biosynthesis of medium-chain wax esters. *Lipids*, *45*(3), 263–273. https://doi.org/10.1007/s11745-010-3395-2

Teh, C.-K., Ong, A.-L., Kwong, Q.-B., Apparow, S., Chew, F.-T., Mayes, S., … Kulaveerasingam, H. (2016a). Genome-wide association study identifies three key loci for high mesocarp oil content in perennial crop oil palm. *Scientific Reports*, *6*(January), 19075. https://doi.org/10.1038/srep19075

Teh, C.-K., Ong, A.-L., Kwong, Q.-B., Apparow, S., Chew, F.-T., Mayes, S., … Kulaveerasingam, H. (2016b). Genome-wide association study identifies three key loci for high mesocarp oil content in perennial crop oil palm. *Scientific Reports*, *6*, 19075. https://doi.org/10.1038/srep19075

Teh, H. F., Neoh, B. K., Hong, M. P. L., Low, J. Y. S., Ng, T. L. M., Ithnin, N., … Appleton, D. R. (2013). Differential Metabolite Profiles during Fruit Development in High-Yielding Oil Palm Mesocarp. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0061344

Teh, H. F., Neoh, B. K., Hong, M. P. L., Low, J. Y. S., Ng, T. L. M., Ithnin, N., … Loewus, F. (2013). Differential Metabolite Profiles during Fruit Development in High-Yielding Oil Palm Mesocarp. *PLoS ONE*, *8*(4), e61344. https://doi.org/10.1371/journal.pone.0061344

Testerink, C., & Munnik, T. (2005). Phosphatidic acid: A multifunctional stress signaling lipid in plants. *Trends in Plant Science*, *10*(8), 368–375. https://doi.org/10.1016/j.tplants.2005.06.002

Thaminy, S., Auerbach, D., Arnoldo, A., & Stagljar, I. (2003). Identification of novel ErbB3-interacting factors using the split-ubiquitin membrane yeast two-hybrid system. *Genome Research*, *13*(7), 1744–53. https://doi.org/10.1101/gr.1276503

Ting, N.-C., Jansen, J., Mayes, S., Massawe, F., Sambanthamurthi, R., Ooi, L. C.-L., … Marra, M. (2014). High density SNP and SSR-based genetic maps of two independent oil palm hybrids. *BMC Genomics*, *15*(1), 309. https://doi.org/10.1186/1471-2164-15-309

Tjellstrom, H., Yang, Z., Allen, D. K., & Ohlrogge, J. B. (2012). Rapid Kinetic Labeling of Arabidopsis Cell Suspension Cultures: Implications for Models of Lipid Export from Plastids. *Plant Physiology*, *158*(2), 601–611. https://doi.org/10.1104/pp.111.186122

To, A., Joubès, J., Barthole, G., Lécureuil, A., Scagnelli, A., Jasinski, S., … Baud, S. (2012). WRINKLED transcription factors orchestrate tissue-specific regulation of fatty acid biosynthesis in Arabidopsis. *The Plant Cell*, *24*(12), 5007–23. https://doi.org/10.1105/tpc.112.106120

Tranbarger, T. J., Dussert, S., Joet, T., Argout, X., Summo, M., Champion, A., … Morcillo, F. (2011). Regulatory Mechanisms Underlying Oil Palm Fruit Mesocarp Maturation, Ripening, and Functional Specialization in Lipid and Carotenoid Metabolism. *Plant Physiology*, *156*(2), 564–584. https://doi.org/10.1104/pp.111.175141

Troncoso-Ponce, M. A., Cao, X., Yang, Z., & Ohlrogge, J. B. (2013). Lipid turnover during senescence. *Plant Science*, *205*, 13–19. https://doi.org/10.1016/j.plantsci.2013.01.004

Troncoso-Ponce, M. A., Kilaru, A., Cao, X., Durrett, T. P., Fan, J., Jensen, J. K., … Ohlrogge, J. B. (2011). Comparative deep transcriptional profiling of four developing oilseeds. *Plant Journal*, *68*(6), 1014–1027. https://doi.org/10.1111/j.1365-313X.2011.04751.x

Tusnády, G. E., & Simon, I. (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics (Oxford, England)*, *17*(9), 849–50. https://doi.org/10.1093/BIOINFORMATICS/17.9.849

van Erp, H., Bates, P. D., Burgal, J., Shockey, J., & Browse, J. (2011). Castor Phospholipid:Diacylglycerol Acyltransferase Facilitates Efficient Metabolism of Hydroxy Fatty Acids in Transgenic Arabidopsis. *Plant Physiology*, *155*(2), 683–693. https://doi.org/10.1104/pp.110.167239

van Erp, H., Kelly, A. A., Menard, G., & Eastmond, P. J. (2014). Multigene engineering of triacylglycerol metabolism boosts seed oil content in Arabidopsis. *Plant Physiology*, *165*(1), 30–6. https://doi.org/10.1104/pp.114.236430

van Erp, H., Shockey, J., Zhang, M., Adhikari, N. D., & Browse, J. (2015). Reducing Isozyme Competition Increases Target Fatty Acid Accumulation in Seed Triacylglycerols of Transgenic Arabidopsis. *Plant Physiology*, *168*(1), 36–46. https://doi.org/10.1104/pp.114.254110

Vanhercke, T., El Tahchy, A., Liu, Q., Zhou, X.-R., Shrestha, P., Divi, U. K., … Petrie, J. R. (2014). Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. *Plant Biotechnology Journal*, *12*(2), 231–239. https://doi.org/10.1111/pbi.12131

Voelker, T. a, & Davies, H. M. (1994). Alteration of the specificity and regulation of Alteration of the Specificity and Regulation of Fatty Acid Synthesis of Escherichia coli by Expression of a Plant Medium- Chain Acyl-Acyl Carrier Protein Thioesterase, *176*(23), 7320–7327.

Waggoner, D. W., Xu, J., Singh, I., Jasinska, R., Zhang, Q. X., & Brindley, D. N. (1999). Structural organization of mammalian lipid phosphate phosphatases: implications for signal transduction. *Biochimica et Biophysica Acta*, *1439*(2), 299–316. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10425403

Wallis, J.G., Watts, J.L., Browse, J. (2002). Polyunsaturated fatty acid synthesis: what will they think of next? *Trends in Biochemical Sciences*, *27*(9), 467–473.

Wältermann, M., Stöveken, T., & Steinbüchel, A. (2007). Key enzymes for biosynthesis of neutral lipid storage compounds in prokaryotes: Properties, function and occurrence of wax ester synthases/acyl-CoA:diacylglycerol acyltransferases. *Biochimie*, *89*(2), 230–242. https://doi.org/10.1016/j.biochi.2006.07.013

Wang, X., Devaiah, S. P., Zhang, W., & Welti, R. (2006). Signaling functions of phosphatidic acid. *Progress in Lipid Research*, *45*(3), 250–278. https://doi.org/10.1016/j.plipres.2006.01.005

Weng, H., Molina, I., Shockey, J., & Browse, J. (2010). Organ fusion and defective cuticle function in a lacs1 lacs2 double mutant of Arabidopsis. *Planta*, *231*(5), 1089–1100. https://doi.org/10.1007/s00425-010-1110-4

Wiberg, E. V. A., & Bafor, M. (1995). Pergamon M E D I U M CHAIN-LENGTH FATTY ACIDS IN LIPIDS OF DEVELOPING OIL PALM KERNEL E N D O S P E R M, *39*(6), 1325–1327.

Wickramarathna, A. D., Siloto, R. M. P., Mietkiewska, E., Singer, S. D., Pan, X., & Weselake, R. J. (2015). Heterologous expression of flax PHOSPHOLIPID:DIACYLGLYCEROL CHOLINEPHOSPHOTRANSFERASE (PDCT) increases polyunsaturated fatty acid content in yeast and Arabidopsis seeds. *Bmc Biotechnology*, *15*, 1–15. https://doi.org/ARTN 63\r10.1186/s12896-015-0156-6

Winkel, B. S. J. (2004). METABOLIC CHANNELING IN PLANTS. *Annual Review of Plant Biology*, *55*(1), 85–107. https://doi.org/10.1146/annurev.arplant.55.031903.141714

Wu, P. Z., Li, J., Wei, Q., Zeng, L., Chen, Y. P., Li, M. R., … Wu, G. J. (2009). Cloning and functional characterization of an acyl-acyl carrier protein thioesterase (JcFATB1) from Jatropha curcas. *Tree Physiology*, *29*(10), 1299–1305. https://doi.org/10.1093/treephys/tpp054

Xiao, S., Chen, Q. F., & Chye, M. L. (2009). Light-regulated Arabidopsis ACBP4 and ACBP5 encode cytosolic acyl-CoA-binding proteins that bind phosphatidylcholine and oleoyl-CoA ester. *Plant Physiology and Biochemistry*, *47*(10), 926–933. https://doi.org/10.1016/j.plaphy.2009.06.007

Xiao, S., & Chye, M. L. (2011). New roles for acyl-CoA-binding proteins (ACBPs) in plant development, stress responses and lipid metabolism. *Progress in Lipid Research*, *50*(2), 141–151. https://doi.org/10.1016/j.plipres.2010.11.002

Xiao, S., Li, H. Y., Zhang, J. P., Chan, S. W., & Chye, M. L. (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*(6), 571–583. https://doi.org/10.1007/s11103-008-9392-7

Xu, C., & Shanklin, J. (2016). Triacylglycerol Metabolism, Function, and Accumulation in Plant Vegetative Tissues. *Annual Review of Plant Biology*, *70*(3), 156–7. https://doi.org/10.1146/annurev-arplant-043015-111641

Xu, J., Carlsson, A. S., Francis, T., Zhang, M., Hoffman, T., Giblin, M. E., & Taylor, D. C. (2012). Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. *BMC Plant Biology*, *12*(1), 4. https://doi.org/10.1186/1471-2229-12-4

Yang, W. L., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M., & Ohlrogge, J. (2010). A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(26), 12040–5. https://doi.org/10.1073/pnas.0914149107

Yang, W., Simpson, J. P., Li-Beisson, Y., Beisson, F., Pollard, M., & Ohlrogge, J. B. (2012). A Land-Plant-Specific Glycerol-3-Phosphate Acyltransferase Family in Arabidopsis: Substrate Specificity, sn-2 Preference, and Evolution. *Plant Physiology*, *160*(2), 638–652. https://doi.org/10.1104/pp.112.201996

Yang, Z., Ji, H., & Liu, D. (2016). Oil Biosynthesis in Underground Oil-Rich Storage Vegetative Tissue: Comparison of Cyperus esculentus Tuber with Oil Seeds and Fruits. *Plant & Cell Physiology*. https://doi.org/10.1093/pcp/pcw165

Yurchenko, O. P., Nykiforuk, C. L., Moloney, M. M., Ståhl, U., Banaś, A., Stymne, S., & Weselake, R. J. (2009). A 10-kDa acyl-CoA-binding protein (ACBP) from brassica napus enhances acyl exchange between acyl-CoA and phosphatidylcholine. *Plant Biotechnology Journal*, *7*(7), 602–610. https://doi.org/10.1111/j.1467-7652.2009.00427.x

Yurchenko, O. P., & Weselake, R. J. (2011). Involvement of low molecular mass soluble acyl-CoA-binding protein in seed oil biosynthesis. *New Biotechnology*, *28*(2), 97–109. https://doi.org/10.1016/j.nbt.2010.09.011

Zhang, M., Fan, J., Taylor, D. C., & Ohlrogge, J. B. (2009). DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *The Plant Cell*, *21*(12), 3885–3901. https://doi.org/10.1105/tpc.109.071795

Zhao, L., Katavic, V., Li, F., Haughn, G. W., & Kunst, L. (2010). Insertional mutant analysis reveals that long-chain acyl-CoA synthetase 1 (LACS1), but not LACS8, functionally overlaps with

LACS9 in Arabidopsis seed oil biosynthesis. *Plant Journal*, *64*(6), 1048–1058. https://doi.org/10.1111/j.1365-313X.2010.04396.x

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G., & Zou, J. (2003). Arabidopsis AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *The Plant Cell*, *15*(8), 1872–87. https://doi.org/10.1105/tpc.012427

Zhou, X. R., Shrestha, P., Yin, F., Petrie, J. R., & Singh, S. P. (2013). AtDGAT2 is a functional acyl-CoA:Diacylglycerol acyltransferase and displays different acyl-CoA substrate preferences than AtDGAT1. *FEBS Letters*, *587*(15), 2371–2376. https://doi.org/10.1016/j.febslet.2013.06.003

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G., & Taylor, D. C. (1999). The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. *The Plant Journal*, *19*(6), 645–653. https://doi.org/10.1046/j.1365-313x.1999.00555.x

Appendix I

Table S1. The PCR condition of each gene from oil palm.

Gene	cDNA	Primers	Tm (℃)	Length (bp)	ET (s)	Vectors
EgAAPTa	leaf	YY23, YY24	P 58	1170	50	BT3-N, PR3-N
EgAAPTb	leaf	YY25, YY26	P 58	1170	50	BT3-STE
EgAAPTb	leaf	YY27, YY26	P 58	1170	50	PR3-STE
EgLPCATa	leaf	YY28, YY29	P 60	1398	60	BT3-N, PR3-N
EgLPCATb	leaf	YY30, YY31	P 60	1398	60	BT3-N, PR3-N
EgLPCATc	leaf	YY32, YY33	P 60	1398	60	BT3-N, PR3-N
EgPDAT1	leaf	YY34, YY35	P 66	2055	90	BT3-N, PR3-N
EgPDCT	leaf	YY36, YY37	P 63	957	40	BT3-N, PR3-N
EgWSD1	22W	YY38, YY39	Q 63.5	1548	60	BT3-N, PR3-N
EgDGAT1	23W	YY40, YY41	P 63	1623	60	BT3-N, PR3-N
EgDGAT1	23W	YY163, YY41	Q 66	1710	60	BT3-N, PR3-N
EgGPAT9	leaf	YY42, YY43	Q 69.7	1116	50	BT3-N, PR3-N
EgGPAT9	leaf	YY42, YY164	Q 69	1116	50	BT3-STE
EgGPAT9	leaf	YY165, YY164	Q 69	1116	50	PR3-STE
EgPAH1	leaf	ET40, YY48	Q 67	2649	90	BT3-N, PR3-N
EgPAH1	leaf	YY140, YY153	Q 69	2649	90	PR3-STE
EgPAH2	leaf	ET42, ET43	Q 61	3621	120	BT3-N, PR3-N
EgPAH2	leaf	YY142, YY154	Q 66	3621	120	PR3-STE
EgPAH3	22W	YY81, YY82	Q 59	3654	120	PR3-N
EgPAH3	22W	YY81, YY82	Q 59	3357	120	PR3-N
EgCDP-DAGa	leaf	YY125, YY126	Q 61	1278	45	BT3-N, PR3-N
EgCDP-DAGb	leaf	YY127, YY128	Q 66	1278	45	BT3-N, PR3-N
EgLPEAT2	leaf	YY144, YY145	Q 66	1566	50	BT3-N, PR3-N
EgACBP6a	leaf	YY129, YY130	Q 68	273	15	PR3-N
EgACBP6b	leaf	YY131, YY130	Q 63	273	15	PR3-N
EgACBP3a	22W	YY123, YY124	Q 67	1086	40	PR3-SUC
EgCCTc	leaf	YY132, YY133	Q 57	834	30	PR3-N
EgCCTb	leaf	YY134, YY135	Q 64	914	30	PR3-N
EgCCTa	leaf	YY136, YY137	Q 59	948	30	PR3-N
EgLACS4	leaf	YY106, YY107	Q 61	1969	60	BT3-N, PR3-N
EgFAD2	leaf	YY118, YY119	Q 72	1176	45	BT3-N, PR3-N
EgDGAT2-1	leaf	ET3, YY110	Q 65	999	30	BT3-N, PR3-N
EgDGAT2-2	kernel	YY108, YY109	Q 58	978	30	BT3-N, PR3-N
EgLPAT4	22W	ET6, ET7	P 71	1200	45	BT3-STE
EgLPAT4	22W	ET8, ET7	P 71	1200	45	PR3-STE
EgLPAT2	22W	ET1, ET2	P 69	1125	45	BT3-N, PR3-N
EgLPPδ	22W	ET16, ET17	P 72	1224	45	BT3-N, PR3-N
EgLPPβ1	22W	ET12, ET13	P 71	633	30	BT3-N, PR3-N
EgLPPβ2	22W	ET14, ET15	P 72	651	30	BT3-N, PR3-N

Q and P are Q5 High-Fidelity DNA Polymerase (BioLabs) and Phusion High-Fidelity DNA

Polymerase (Thermo Scientific), respectively. ET is extension time for PCR.

Table S2. Primers used in the PCR.

Number	Primer name	Sequence
YY23	EgAAPTa BT3-N FW	ATTAACAAGGCCATTACGGCCATGGGATATATAGGGCACC
YY24	EgAAPTa BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAAGCTTCTTTCCGAGTT
YY25	EgAAPTb BT3-STE FW	ATTAACAAGGCCATTACGGCCATGGGATACATAGGGCACC
YY26	EgAAPTb BT3-STE RV	AACTGATTGGCCGAGGCGGCCCCAGCTTCTTTCCGAGTTATCC
YY27	EgAAPTb PR3-STE FW	ATTAACAAGGCCATTACGGCCTTATGGGATACATAGGGCACC
YY28	EgLPCATa BT3-N FW	ATTAACAAGGCCATTACGGCCATGGAGCTCGAAATGGC
YY29	EgLPCATa BT3-N RV	AACTGATTGGCCGAGGCGGCCCTATTGGCTCTTCTTAGCTTTAGG
YY30	EgLPCATb BT3-N FW	ATTAACAAGGCCATTACGGCCATGGAGCTCGAGATGGC
YY31	EgLPCATb BT3-N RV	AACTGATTGGCCGAGGCGGCCCTATTGGCTCTTCTGAGCTTTAG
YY32	EgLPCATc BT3-N FW	ATTAACAAGGCCATTACGGCCATGGAGCTGGAGATGGA
YY33	EgLPCATc BT3-N RV	AACTGATTGGCCGAGGCGGCCTCACTGGCTTTTCTGAGTT
YY34	EgPDAT1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGTCGATGCTGCGACGG
YY35	EgPDAT1 BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAGAGTCGTAACTTGACCTTTTCT
YY36	EgPDCT BT3-N FW	ATTAACAAGGCCATTACGGCCATGCGCGACAAAGTTGG
YY37	EgPDCT BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAACTGCTGCCGTTGC
YY38	EgWSD1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGCTTCCAGAAGAAGGC
YY39	EgWSD1 BT3-N RV	AACTGATTGGCCGAGGCGGCCCTATGATCTTGGAATTGCATCCT
YY40	EgDGAT1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGCAATTCCGAGCG
YY41	EgDGAT1 BT3-N RV	AACTGATTGGCCGAGGCGGCCCTATTCGGTCCCCACTTTCC
YY163	EgDGAT1 BT3-N FW (New)	ATTAACAAGGCCATTACGGCCATGCGCTCCTCCCCTCT
YY42	EgGPAT9 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGTTGAGCTGCGGTCATCG
YY43	EgGAPAT9 BT3-N RV	AACTGATTGGCCGAGGCGGCCTTATTTTTCCTCCAACCGCTGCAAG
YY164	EgGPAT9 BT3-STE RV	AACTGATTGGCCGAGGCGGCCCCTTTTTCCTCCAACCGCTGCAA
YY165	EgGPAT9 PR3-STE FW	ATTAACAAGGCCATTACGGCCTTATGGTTGAGCTGCGGTCATCG
ET40	EgPAH1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGAACGTGGTGGGCAAGG
YY48	EgPAH1 BT3-N RV	AACTGATTGGCCGAGGCGGCCCTAGCATTCAATGTCAGGCAATGG
YY140	EgPAH1 PR3-STE FW	ATTAACAAGGCCATTACGGCCTTATGAACGTGGTGGGCAAGGT
YY153	EgPAH1 PR3-STE RV	AACTGATTGGCCGAGGCGGCCCCGCATTCAATGTCAGGCAATGGC
ET42	EgPAH2 BT3-N FW	ATTAACAAGGCCATTACGGCCATGTATGCAGTGGAGAAGCTG
ET43	EgPAH2 BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAAACATCAACTTCAGGAAG
YY142	EgPAH2 PR3-STE FW	ATTAACAAGGCCATTACGGCCTTATGTATGCAGTGGAGAAGCTG
YY154	EgPAH2 PR3-STE RV	AACTGATTGGCCGAGGCGGCCCCAACATCAACTTCAGGAAGAGGC
YY81	EgPAH3 BT3-N FW	ATTAACAAGGCCATTACGGCCATGTATGCGGTGGAGAAG
YY82	EgPAH3 BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAAACATCAACTTCAGGA
YY125	EgCDP-DAGa BT3-N FW	ATTAACAAGGCCATTACGGCCATGCAAAAGGATAATAGCTCC
YY126	EgCDP-DAGa BT3-N RV	AACTGATTGGCCGAGGCGGCCTTAGGCGAGCATAGATTGTC
YY127	EgCDP-DAGb BT3-N FW	ATTAACAAGGCCATTACGGCCATGCAAAAGGATAGCAGCTCC
YY128	EgCDP-DAGb BT3-N RV	AACTGATTGGCCGAGGCGGCCTTAGGGTGGCATAGCCTGTC
YY144	EgLPEAT2 PBT3-N FW	ATTAACAAGGCCATTACGGCCATGGAGGAACCCAATGGC
YY145	EgLPEAT2 PBT3-N RV	AACTGATTGGCCGAGGCGGCCTCAACTTTCGTACTTCATACAGGC

YY129	EgACBP6a PR3-N FW	ATTAACAAGGCCATTACGGCCATGGGTCTCAAGGAGGAGTTTG
YY130	EgACBP6a PR3-N RV	AACTGATTGGCCGAGGCGGCCTCAGGAAGCAGCAGCAGC
YY131	EgACBP6b PR3-N FW	ATTAACAAGGCCATTACGGCCATGGGTCTCAAGGAAGAATTTG
YY123	EgACBP3a PR3-SUC FW	ATTAACAAGGCCATTACGGCCCTCGGCGGCGTGGA
YY124	EgACBP3a PR3-SUC RV	AACTGATTGGCCGAGGCGGCCCCTAAATAGAAACTCACGAATGTCACTCTGG
YY132	EgCCTc PR3-N FW	ATTAACAAGGCCATTACGGCCATGGCAGATCCGAC
YY133	EgCCTc PR3-N RV	AACTGATTGGCCGAGGCGGCCCTACTCTTCTTCATCCTCATA
YY134	EgCCTb PR3-N FW	ATTAACAAGGCCATTACGGCCATGGCCATCATGCCCTC
YY135	EgCCTb PR3-N RV	AACTGATTGGCCGAGGCGGCCTTATTTGCTTGACTGCCTCTTC
YY136	EgCCTa PR3-N FW	ATTAACAAGGCCATTACGGCCATGGCGCGCGTGC
YY137	EgCCTa PR3-N RV	AACTGATTGGCCGAGGCGGCCCTATCCATCAAATGCCATTG
YY106	EgLACS4 BT3-N FW	ATTAACAAGGCCATTACGGCCATGAGGTATCTTATAGAGGTGGA
YY107	EgLACS4 BT3-N RV	AACTGATTGGCCGAGGCGGCCTTACTTCGCACTTTTGTACAG
YY118	EgFAD2 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGGCGCGGGCGGAC
YY119	EgFAD2 BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAGAACTTGTGCCGGTACCAGA
ET3	EgDGAT2-1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGACCACGGTAACGGAGAC
YY110	EgDGAT2-1 BT3-N RV	AACTGATTGGCCGAGGCGGCCTTATAGAACTCTCAAACGAAGATCAGG
YY108	EgDGAT2-2 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGAGGAACCTGAGAGG
YY109	EgDGAT2-2 BT3-N RV	AACTGATTGGCCGAGGCGGCCTCAAACAATTCTTAATTCAATGTC
ET6	EgLPAT4 BT3-STE FW	ATTAACAAGGCCATTACGGCCATGAATGGTTCTTCTTCCAATGAAGGAAAATCATT
ET7	EgLPAT4 BT3-STE RV	AACTGATTGGCCGAGGCGGCCCCGCATGCTTTCTTGCCACAGAATAAG
ET8	EgLPAT4 PR3-STE FW	ATTAACAAGGCCATTACGGCCTTATGAATGGTTCTTCTTCCAATGAAGGAAAATCATT
ET1	EgLPAT2 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGCGATCGTGGCGG
ET2	EgLPAT2 BT3-N RV	AACTGATTGGCCGAGGCGGCCATCTGTGTGTGTTCGTCTGCG
ET12	EgLPPβ1 BT3-N FW	ATTAACAAGGCCATTACGGCCATGTCCGCCGCCGCC
ET13	EgLPPβ1 BT3-N RV	AACTGATTGGCCGAGGCGGCCCTAGAAGTGGAAGAAATAGAACACCATGAAAGCT
ET14	EgLPPβ2 BT3-N FW	ATTAACAAGGCCATTACGGCCATGGCCGCCGCCGG
ET15	EgLPPβ2 BT3-N RV	AACTGATTGGCCGAGGCGGCCTTAGAACAAACCATCCCCCAAGCAAAAAATATTAAAAACCGG
ET16	EgLPPδ BT3-N FW	ATTAACAAGGCCATTACGGCCATGGAAGGAGGCATGCCGTTCT
ET17	EgLPPδ BT3-N RV	AACTGATTGGCCGAGGCGGCCCTACAAATTTAGATGAGAGAAGAGTGATGGAACCAAGT

Table S3. Primers for construction of *S. cerevisiae* expression plasmids

Primers	RE sites	Sequences
EgDGAT1-pESC-URA F	BamHI (site 2)	CGCGGATCCATGCGCTCCTCCCCTCT
EgDGAT1-pESC-URA R	KpnI (site 2)	GGGGTACCCTATTCGGTCCCCACTTTCC
EgDGAT2-1-pESC-URA F	BamHI (site 2)	CGGGATCCATGGACCACGGTAACGGAGAC
EgDGAT2-1-pESC-URA R	KpnI (site 2)	GGGGTACCTTATAGAACTCTCAAACGAAGATCAGG
EgDGAT2-2-pESC-URA F	BamHI (site 2)	CGGGATCCATGGAGGAACCTGAGAGG
EgDGAT2-2-pESC-URA R	SpeI (site 2)	GGACTAGTTCAAACAATTCTTAATTCAATGTC
EgLPPβ2-pESC-URA F	NotI (site 1)	ATAAGAATGCGGCCGCATGGCCGCCGCCG
EgLPPβ2-pESC-URA R	SpeI (site 1)	CGGACTAGTTTAAAACCTGTAAAAGTTGAAGAAATAGAACAC

Table S4. The topology of protein encoded by genes from oil palm

Protein	TOPCONS		TMHMM		HMMTOP		Phobius		SignalP
	N	C	N	C	N	C	N	C	4.1
EgAAPTa	in	in	out	in	in	in	in	in	no
EgAAPTb	in	in	out	in	out	in	out	in	no
EgLPCATa	in	in	in	in	in	in	in	in	no
EgLPCATb	in	in	in	in	in	in	in	in	no
EgLPCATc	in	in	in	in	in	in	in	in	no
EgPDAT1	in	out	in	out	in	out	in	out	no
EgPDCT	in	out	in	in	in	out	in	in	no
EgWSD1-like	in	out	No TMD		in	in	in	out	no
EgDGAT1	in	out	in	out	in	in	in	out	no
EgGPAT9	in	in	in	out	in	out	in	out	no
EgPAH1	No TMD		No TMD		No TMD		No TMD		no
EgPAH2	No TMD		No TMD		No TMD		No TMD		no
EgPAH3	No TMD		No TMD		No TMD		No TMD		no
EgCDP-DAGa	in	out	out	in	out	out	out	in	no
EgCDP-DAGb	in	out	out	in	out	out	out	in	no
EgLPEAT	in	out	in	out	in	out	out	out	no
EgACBP6a	No TMD		No TMD		No TMD		No TMD		no
EgACBP6b	No TMD		No TMD		No TMD		No TMD		no
EgACBP3a	No TMD		in	out	out	in	N-region	out	yes
EgCCTc	No TMD		No TMD		No TMD		No TMD		no
EgCCTb	No TMD		No TMD		No TMD		No TMD		no
EgCCTa	No TMD		No TMD		No TMD		No TMD		no
EgLACS4	No TMD		No TMD		No TMD		No TMD		no
EgFAD2	in	in	in	in	in	out	in	in	no
EgDGAT2-1	in	in	out	out	in	in	in	in	no
EgDGAT2-2	in	in	out	out	in	in	out	in	no
EgLPAT4	in	in	in	out	in	out	out	in	no
EgLPAT2	out	out	out	in	out	in	N-region	in	no
EgLPPδ	out	out	out	out	out	out	N-region	in	no
EgLPPβ1	in	out	out	out	out	out	out	out	no
EgLPPβ2	in	out	out	out	out	in	out	out	no

N and C mean the location of N-terminal and C-terminal of the predicted protein; in and out mean the location of protein terminal is cytosolic and luminal, respectively. N-region is the N-terminal of the predicted protein has a signal peptide. No TMD means the predicted protein has no transmembrane domain in the sequence. Yes or no means the predicted protein has a signal peptide or not.

Appendix II

Supplementary Figs:

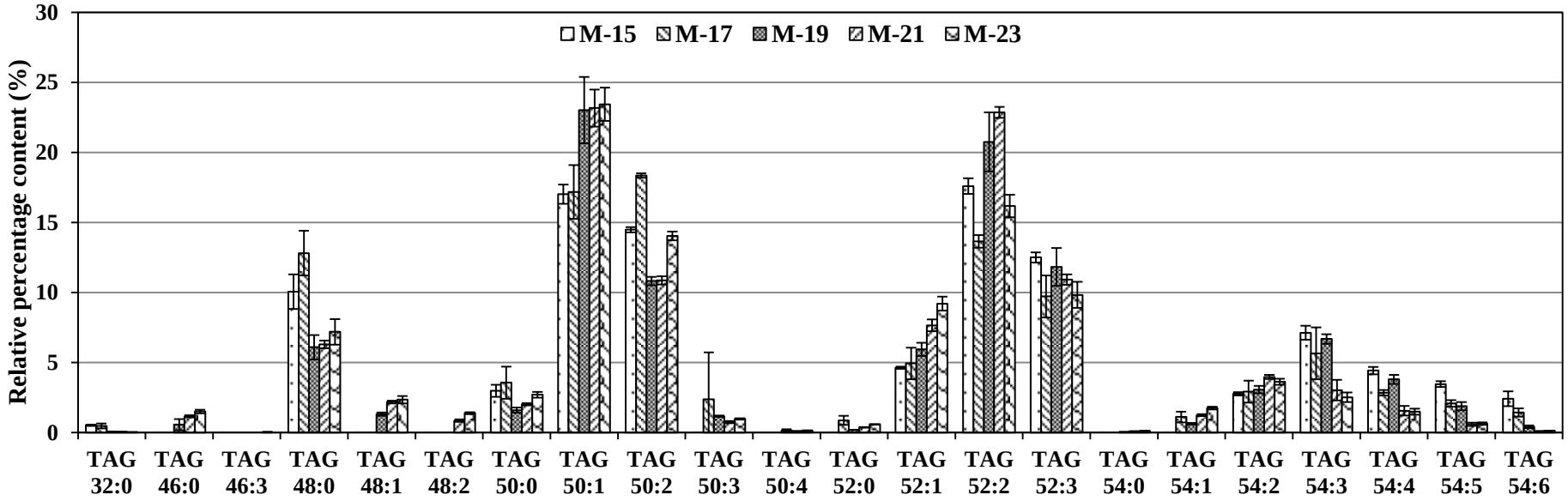

Fig.S1 Relative quantification analysis of TAG from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (TAG 17:0/17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standards. TAG species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, TAG50:1 is likely to contain two palmitic acids (16:0) and one oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

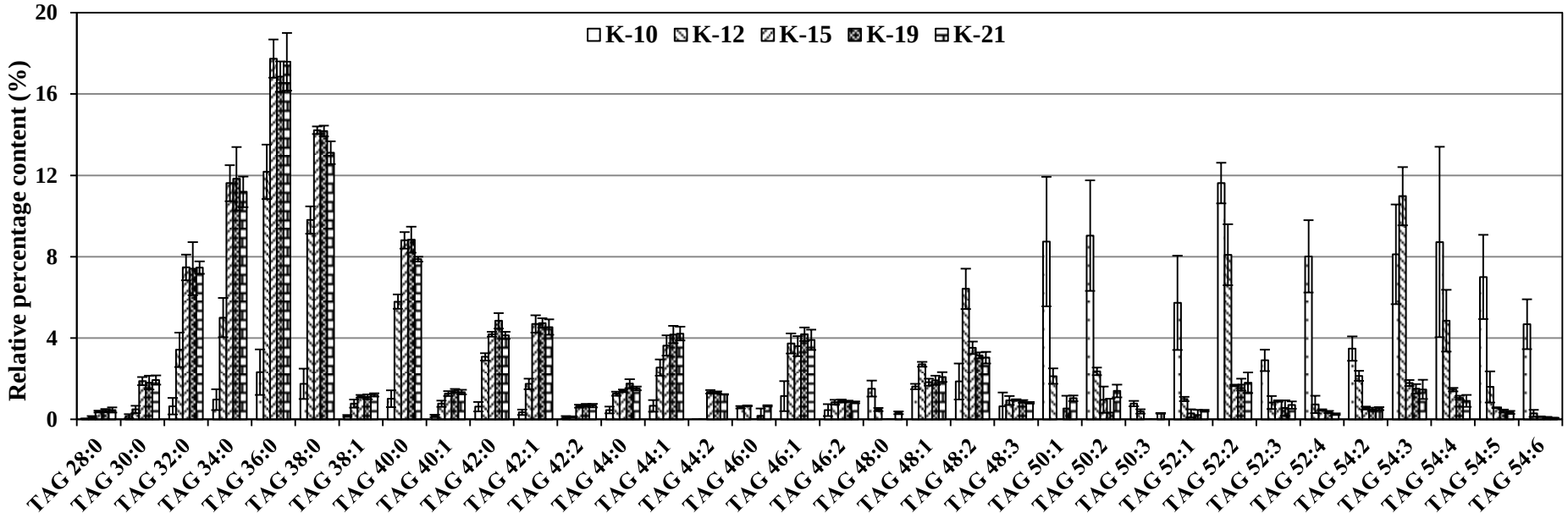

Fig.S2 Relative quantification analysis of TAG from and kernel at five developmental stages by LC-MS/MS (10 and 12 WAP are the same with Fig. 3-1-5). Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (TAG 17:0/17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standards. TAG species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, TAG50:1 is likely to contain two palmitic acids (16:0) and one oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

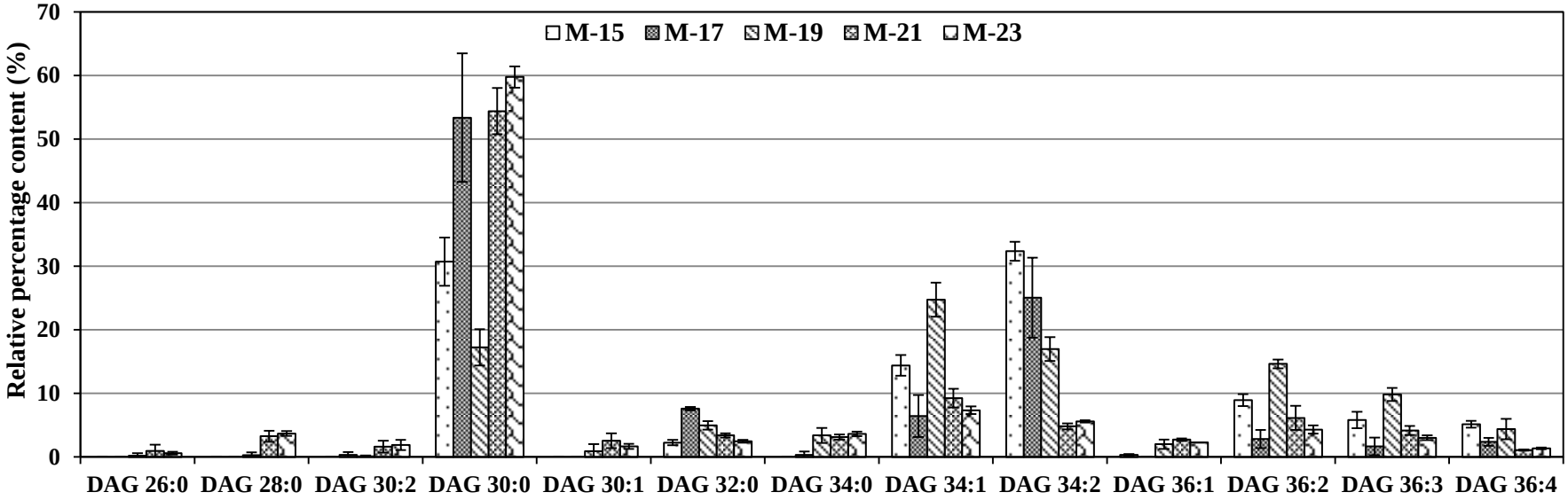

Fig.S3 Relative quantification analysis of DAG from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (DAG 17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standard. Species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, DAG 34:1 is likely to contain palmitic acid (16:0) and oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

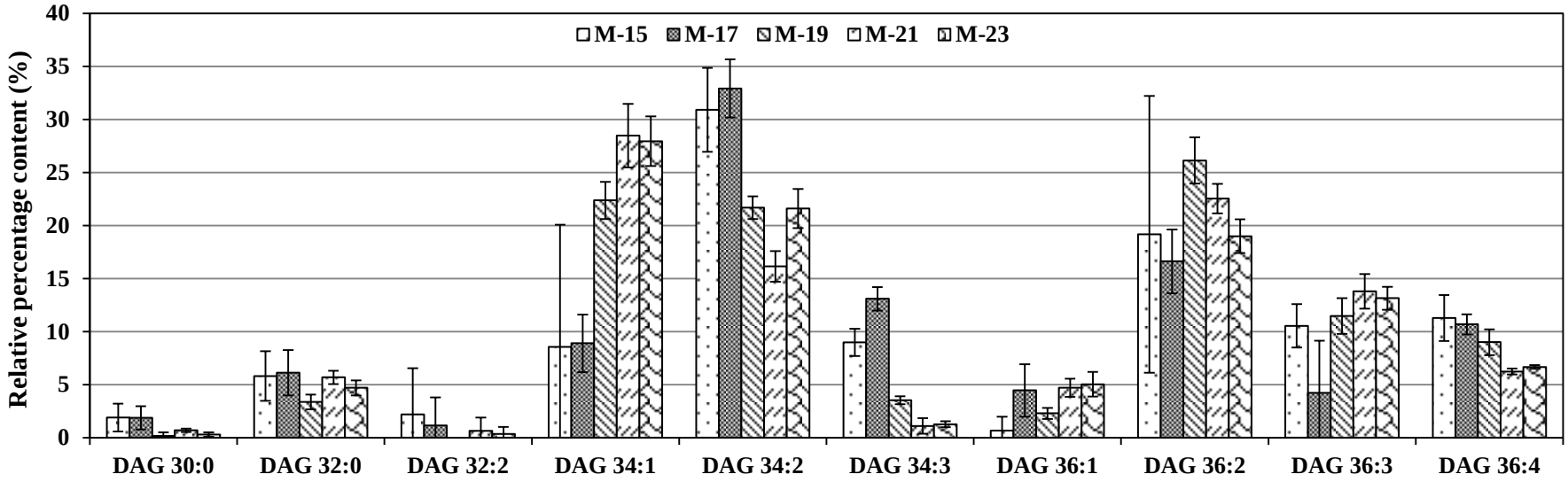

Fig.S4 Relative quantification analysis of DAG from mesocarp at five developmental stages (15, 17, 19, 21, and 23 WAP) by LC-MS/MS. Analyses were carried out on 4 biological replicates. Synthetic internal lipid standard (DAG 17:0/17:0 from Avanti Polar Lipids (USA)) was used as internal standard. Species are designated by the number of carbon atoms of all fatty acids and the number of double bonds separated by a colon. For example, DAG 34:1 is likely to contain palmitic acid (16:0) and oleic acid (18:1). Only species representing at least 1% of total lipids are shown.

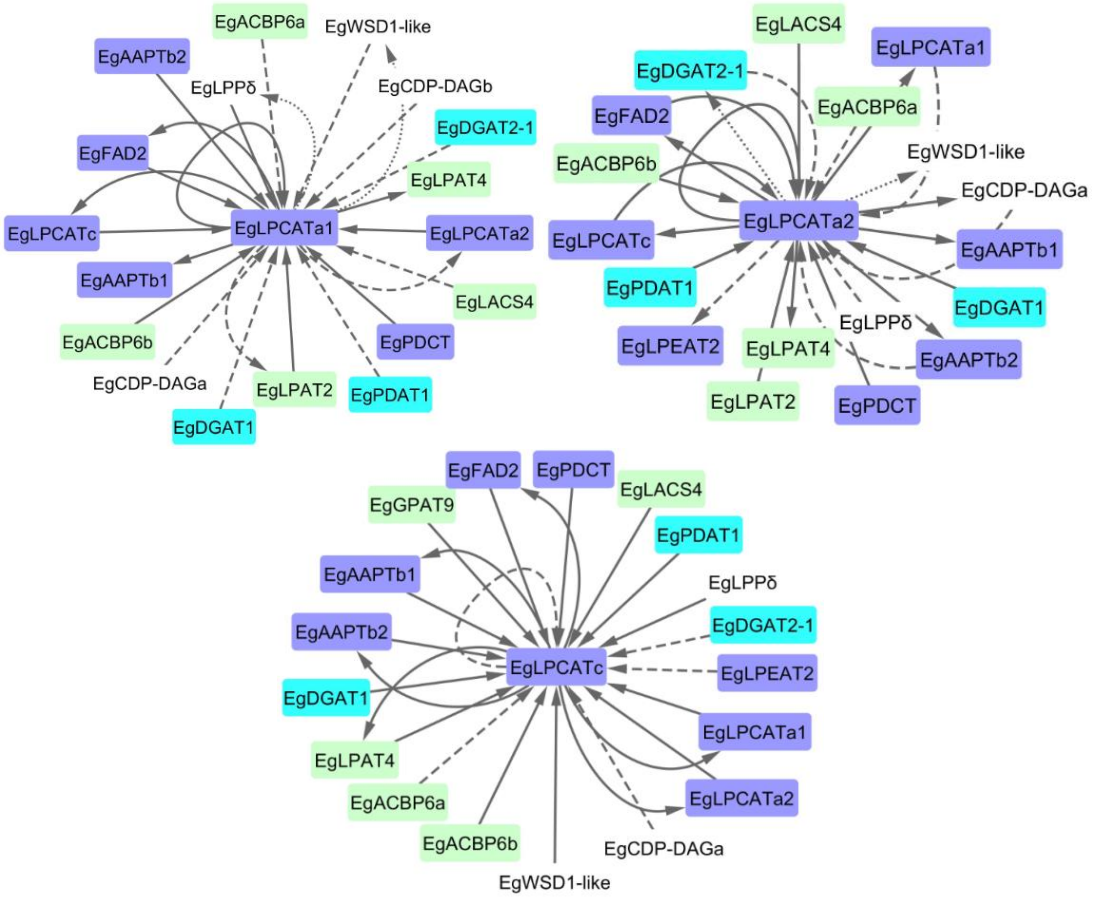

Fig.S5 Interactions of EgLPCATs. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

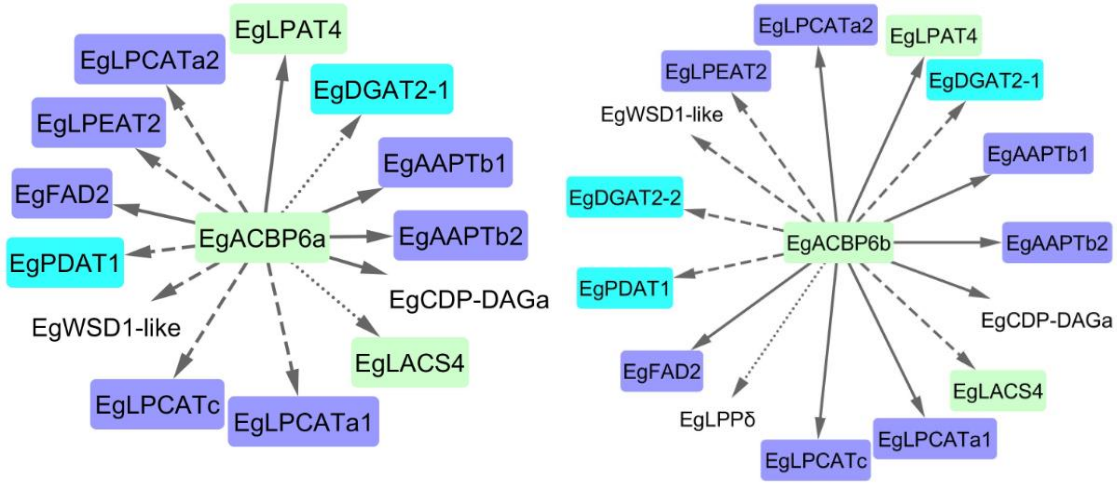

Fig.S6 Interactions of EgACBP6s. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

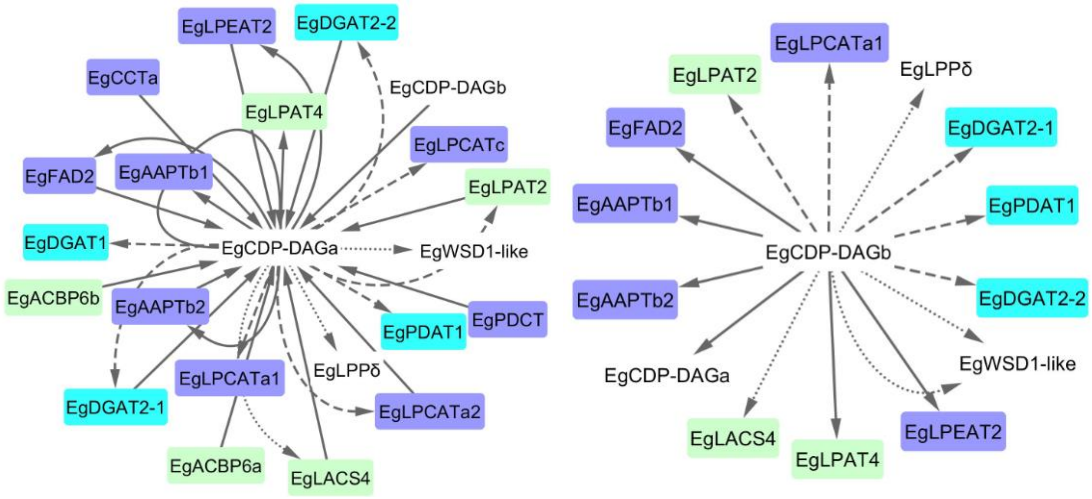

Fig.S7 Interactions of EgCDP-DAGs. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

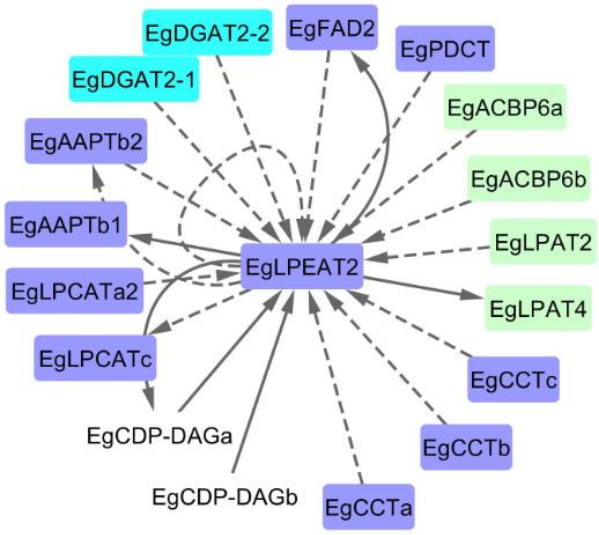

Fig.S8 Interactions of EgLPEAT2. Solid, dash, and dot lines mean strong, medium, and weak interactions, respectively. The arrows go from the prey to the bait.

Titre : Etude de l'assemblage des acides gras en huile chez le palmier à huile (*Elaeis guineensis* Jacq.)

Résumé : Le palmier à huile est la première culture oléagineuse, avec environ 40% de la production mondiale, et son fruit accumule deux huiles de composition très différente dans le mesocarpe et l'amande. Chez les plantes, les acides gras sont assemblés en huile dans le réticulum endoplasmique, ceci par la voie dite de Kennedy à laquelle s'ajoutent des mécanismes d'édition impliquant le métabolisme de la phosphatidylcholine. Nous avons utilisé les outils de la lipidomique pour analyser la variabilité au sein de différentes populations de palmier ainsi que pour caractériser l'accumulation d'huile durant le développement du mésocarpe et de l'amande. Puis, nous avons entrepris de tester, dans le système du double hybride de levure, les interactions entre toutes les enzymes de la voie de Kennedy et celles responsables des mécanismes d'édition, et mis en évidence 241 interactions, dont 132 sont fortes, 73 moyennes et 36 faibles. Ces résultats suggèrent que ces enzymes pourraient s'assembler en complexes supra-moléculaires susceptibles de former des métabolons. Certaines isoformes d'une même enzyme ont des profils d'interaction distincts, ce qui ouvre des perspectives pour de futures recherches. De plus, nous avons caractérisé, par expression fonctionnelle dans un mutant de levure dépourvu de TAG, une acyltransférase présumée (EgWSD1-like) ainsi que les trois formes majeures de diacylglycérol acyltransférases du mésocarpe. EgWSD1-like ne restaure que l'activité de synthèse d'esters de cire dans le mutant, tandis que les trois DGAT complémentent toutes la déficience en TAG du mutant, avec d'apparentes spécificités distinctes vis-à-vis des acides gras.

Mots clés : *Elaeis guineensis* Jacq., assemblage des acides gras en huile, analyses lipidomiques, interactions proteine/proteine, diacylglycerol acyltransférase, synthase d'esters de cires, *Saccharomyces cerevisiae*

Title : Functional study of oil assembly pathway in oil palm (*Elaeis guineensis* Jacq.) fruits

Abstract : Oil palm is the highest oil-yielding crop-plant, accounting for approximately 40% of the total world vegetable oil production. The fruit accumulates oil, made of triacylglycerol (TAG) molecules, in both mesocarp and kernel with totally different fatty acid profiles. Fatty acids are assembled into oil through Kennedy pathway in the endoplasmic reticulum, which is complicated by editing processes involving phosphatidylcholine metabolism. To investigate oil assembly in oil palm, we use lipidomics as a tool to analyze different populations of palm to search for TAG structural diversity, and to further characterize changes in lipid content and composition in mesocarp and kernel during fruit ripening. We used yeast two-hybrid system to test protein-protein interactions for almost all the enzymes involved in oil assembly pathway, and we demonstrated 241 interactions, including 132 strong interactions, 73 medium interactions and 36 weak interactions. Our results suggest that all enzymes might assemble into one or several complexes that may form metabolons. In addition, different isoforms of enzymes showed distinct interaction profiles, providing hints for future studies. Moreover, we also characterized the *in vivo* function of a putative acyltransferase (designated EgWSD1-like) possibly involved in oil assembly and the three major diacylglycerol acyltransferase (DGAT) isoforms of palm mesocarp in the mutant yeast H1246, which is devoid of neutral lipid synthesis. EgWSD1-like only shows wax ester synthase activity in yeast, while three EgDGATs all can restore TAG biosynthesis in yeast with different substrate specificities.

Keywords : *Elaeis guineensis* Jacq., oil assembly, lipidomic analysis, protein/protein interaction, diacylglycerol acyltransferase, wax ester synthase, *Saccharomyces cerevisiae*

Laboratoire de Biogénèse Membranaire

[CNRS, UMR 5200, Bat. A3, Campus INRA de Bordeaux, 71 avenue E. Bourlaux, CS 20032, 33140 Villenave d'Ornon, France]