Uncovering the Molecular Mechanisms behind Triacylglycerol Biosynthesis in Flaxseed (*Linum usitatissimum* L.)

by

Xue Pan

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Flax (*Linum usitatissimum* L.) seed oil contains approximately 45%-65% of α-linolenic acid (ALA; 18:3$^{cis\Delta 9,12,15}$) and 15-29 % of linoleic acid (LA; 18:2$^{cis\Delta 9,12}$), and is one of the richest natural sources of these polyunsaturated fatty acids (PUFAs). Due to its high amounts of PUFAs, especially ALA, flax oil has a number of applications in the manufacture of food, feed and industrial products. In developing oilseeds, PUFAs are mainly synthesized on phosphatidylcholine (PC). To accumulate large amounts of PUFAs, flax may contain efficient mechanisms for transferring PUFAs from PC into triacylglycerols (TAGs), which are the main components of seed oils. Understanding these molecular mechanisms has the potential to develop new strategies to improve the quality and increase the value of flax seed oil. The present doctoral thesis includes four related studies, which aimed to use a multi-disciplinary approach, including biochemistry, molecular biology and bioinformatics, to investigate the TAG biosynthetic machinery in flax.

The first study isolated and functionally characterized enzymes involved in the final step of TAG synthesis from flax. Two copies of *ACYL COA:DIACYLGLYCEROL ACYLTRANSFERASE 1*(*DGAT1*), three copies of *DGAT2,* two copies of *DEFECTIVE IN CUTICULAR RIDGE* (*DCR*), and six copies of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes have been identified in the flax genome. Expression analyses along with functional characterization of candidate genes in yeast and Arabidopsis heterologous expression systems revealed that flax contains unique forms of PDATs that can efficiently channel ALA into TAG. Intrigued by the finding that multiple *PDAT* paralogs from flax are differentially expressed and encode enzymes with different TAG-forming capacities, the

second study was designed to investigate the evolutionary relationship of the *PDAT* gene family across green plants. The results showed that the *PDAT* gene family is widely present in plant species and can be divided into seven major clades. The study further revealed that there is a eudicot-wide *PDAT* gene expansion. The ancient gene duplication and divergent selective pressures may have led to the diversification of *PDAT* paralogs in the core eudicots. The third study investigated another mechanism for PUFA enrichment in TAG, which involves the biochemical coupling of the reverse reaction of acyl-CoA:lysophosphatidylcholine acyltransferase to the forward reaction of DGAT1. Data obtained from *in vivo* and *in vitro* experiments support that PUFAs can be transferred from PC into the acyl-CoA pool by the reverse reaction of LPCAT, and then be utilized by the DGAT1-catalyzed reaction for TAG synthesis. A higher amount of PUFA-containing TAGs was produced by the DGAT1 reaction in the presence than in the absence of LPCAT. The final study was designed to provide direct evidence showing the role of PDAT and DGAT1 in seed oil and ALA accumulation in native flax plants via an RNA interference approach. T_1 transgenic seeds were obtained. More analyses need to be carried out.

Overall, the present doctoral studies revealed important mechanisms for PUFA-rich TAG production in flax and discovered an interesting evolutionary pattern of the plant *PDAT* gene family. The knowledge obtained from these studies provides important insights into the seed oil biosynthesis in flax and will benefit future initiatives aimed at producing TAGs with increased PUFA content in plants.

Preface

This thesis is based on the findings presented in the following papers, referred to by Roman numerals in the text:

I. Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a Pair of Phospholipid:Diacylglycerol Acyltransferases from Developing Flax (*Linum usitatissimum* L.) Seed Catalyzing the Selective Production of Trilinolenin. *J Biol Chem* 288: 24173–24188

II. Pan X, Peng FY, Weselake RJ (2015) Genome-Wide Analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) Genes in Plants Reveals the Eudicot-Wide *PDAT* Gene Expansion and Altered Selective Pressures Acting on the Core Eudicot *PDAT* Paralogs. *Plant Physiol* 167: 887-904.

III. Pan X, Chen G, Kazachkov M, Greer MS, Zou J, Weselake RJ (2015) *In vivo* and *In vitro* Evidence for Biochemical Coupling of Reactions Catalyzed by Lysophosphatidylcholine Acyltransferase and Diacylglycerol Acyltransferase. *J Biol Chem* [Epub ahead of print]

IV. Pan X, Weselake RJ. Evaluating the Contribution of Phospholipid:Diacylglycerol Acyltransferases and Diacylglycerol Acyltransferase1 to α-Linolenic Acid Accumulation in Flax (*Linum usitatissimum* L.) Seed Oil through an RNA Interference Approach (Ongoing study).

The contribution of Xue Pan to the Papers I-IV included in this thesis is summarized as follows:

I. I designed the experiments, performed gene structure analysis, molecular cloning, Western blot analysis, recombinant protein expression in yeast and Arabidopsis system and data analysis. In addition, I wrote the first draft of the paper and further modified the paper with the co-authors.

II. I designed the experiments, performed all analyses except the InterPro analysis. In addition, I wrote the first draft of the paper and further modified the paper with the co-authors.

III. Designed and conducted all experiments and data analysis. In addition, I drafted the paper and further modified the paper with the co-authors.

IV. Designed the experiments and generated all RNAi constructs. In addition, I wrote the first draft of the manuscript.

Acknowledgements

It has always been my dream to earn a degree in Doctorate. This dream was made a reality thanks to a great number of people who have directly or indirectly helped me on this important journey. First, I would like to express my deep gratitude to my supervisor, Dr. Randall Weselake, for offering me this wonderful opportunity to peruse my PhD in his lab. I will always be grateful for his faith in me and for his support, patience, encouragement, mentorship and understanding through all this time. I would also like to say a sincere thank to my supervisory committee members, Dr. Jocelyn Ozga, Dr. Michael Gaenzle, and Dr. Michael Deyholos for their support and discussion on my PhD research. In addition, I acknowledge Dr. Stephen Strelkov for serving at my dissertation committee. I am also truly grateful to Dr. John B. Ohlrogge for agreeing to serve as my external examiner.

During my PhD study, I have worked in three different labs: Dr. Weselake' group, Dr. Jitao Zou' group (National Research Council, Saskatoon) and Dr. Xiao Qiu's group (University of Saskatchewan). I particularly thank Dr. Zou for his guidance and helpful criticisms on my research. I have nothing but great admiration and appreciation for him. I am also deeply grateful to Dr. Qiu for his support during my stay in his lab. He has always been respectful, approachable and really cares about my future and career. I also thank all past and present lab members of all three groups for their help, encouragement and friendship. A big thank you goes to Dr. Rodrigo Siloto, a former member of the Weselake group. His guidance and support at the beginning of my PhD study really opened my eyes to the wonders of plant molecular biology. I also want to give special thanks to Dr. Saleh Shah for performing flax transformation, to Dr. Elzbieta Mietkiewska for sharing experience in molecular cloning, to Dr. Michael Kazachkov and Dr. Guanqun (Gavin) Chen for teaching me enzymatic techniques, to Dr. Fred Peng for answering

my bioinformatics-related questions, to Michael S. Greer for the discussion on enzyme assays, to Dr. Stacy Singer for showing me the qRT-PCR technique, to Dr. Liping Wang and Dr. Peng Gao for giving me suggestions on generating RNAi constructs, and to Dr. Ida Lager for sharing the protocol for enzyme assays. I also thank Robin Miles and Annie Wong for their hard working, being the experts of troubleshooting, and keeping the lab organized, which greatly facilitated the completion of my thesis.

I especially appreciate Dr. Sten Stymne (Swedish University of Agricultural Sciences) for his patience answering my questions about enzyme assays through emails. His enthusiasm and interest in my project was energizing. Specially thanks to Dr. Karen Tanino (University of Saskatchewan) and Dr. Jorunn E. Olsen (Norwegian University of Life Sciences) who offered me the opportunity to take the microscopy training course in Norwegian University of Life Sciences. I also thank Jody Forsland for helping me with all administrative procedures for my PhD program. I also gratefully acknowledge the financial support of my Ph.D. program provided by Genome Alberta/Genome Canada, the Canada Research Chairs Program, the Canada Foundation for Innovation, the Research Capacity Program of Alberta Enterprise and Advanced Education, AVAC Ltd., Alberta Innovates Bio Solutions and the Natural Sciences and Engineering Research Council of Canada.

Above all, I thank my wonderful parents for their unconditional love and support. Finally, an extremely special thank to my loving husband who has given me more support than anyone truly deserves.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA: arachidonic acid

ACAT: acyl-CoA:cholesterol acyltransferase

ACCase: acetyl-CoA carboxylase

ACP: acyl carrier protein

ACS: acyl-CoA synthetase

AIC: akaike information criterion

ALA: α-linolenic acid

ASAT: acyl-CoA:sterol acyltransferase

BSA: bovine serum albumin

BLAST: basic local alignment search tool

CoA: coenzyme A

CTAB: hexadecyltrimethylammonium bromide

DAG: diacylglycerol

DCR: defective in cuticular ridge

DAcT: diacylglycerol acetyltransferase

DGAT: acyl CoA:diacylglycerol acyltransferase

DGLA: dihomo-γ-linolenic acid

DHA: docosahexaenoic acid

DPA: days post anthesis

EA: eicosenoic acid

EMS: ethyl methanesulfonate

EPA: eicosapentaenoic acid

ER: endoplasmic reticulum

ETA: eicosatrienoic acid

EST: expressed sequence tag

E-value: expectation value

FA: fatty acid

FAD: fatty acid desaturase

FAT: acyl-ACP thioesterase

FAS: fatty acid synthase

FAME: fatty acid methyl ester

FFA: free fatty acid

GAPDH: glyceraldehyde 3-phosphate dehydrogenase

GC: gas chromatography

G3P: glycerol-3-phosphate

GLA: γ-linolenic acid

GPAT: acyl-CoA: glycerol-3-phosphate acyltransferase

KAS: ketoacyl-ACP synthase

K_N: nonsynonymous substitution rate

K_S: synonymous substitution rate

LA: linoleic acid

LACS: long-chain acyl-CoA synthase

LPA: lysophosphatidic acid

LPAT: acyl-CoA: lysophosphatidic acid acyltransferase

LPC: lysophosphatidylcholine

LCAT: lecithin:cholesterol acyltransferase

LPCAT: acyl-CoA: lysophosphatidylcholine acyltransferase

Lyso-PAF: lyso-platelet-activating factor

LRT: likelihood ratio test

MCS: multiple cloning site

MCAT: malonyl-CoA:ACP acyltransferase

MEME: multiple expectation maximization for motif elicitation

ML: maximum likelihood

MS: mass spectrometry

MYA: million years ago

MUFA: monounsaturated fatty acid

OA: oleic acid

OD: optical density

ORF: open reading frame

PA: phosphatidic acid

PAP: phosphatidic acid phosphatase

PANTHER: protein analysis through evolutionary relationships

PC: phosphatidylcholine

PCR: polymerase chain reaction

PDAT: phospholipid:diacylglycerol acyltransferase

PDCT: phosphatidylcholine:diacylglycerol cholinephosphotransferase

PEP: phosphoenolpyruvate

3PGA: 3-phosphoglycerate

pI: isoelectric point

PLA: phospholipase A

PDH: pyruvate dehydrogenase

PLC: phospholipase C

PLD: phospholipase D

PPP: pentose phosphate pathway

PSAT: phospholipid:sterol acyltransferase

PUFA: polyunsaturated fatty acid

Pyr: pyruvate

Ru5P: ribulose 5-phosphate

RuBP: ribulose 1,5-bisphosphate

RuBisCO: ribulose-1,5-bisphosphate carboxylase/oxygenase

SAD: stearoyl-ACP desaturase

SDA: stearidonic acid

TAG: triacylglycerol

TLC: thin layer chromatography

TMD: transmembrane domains

VLC-ω-3-PUFA: very long chain omega-3 polyunsaturated fatty acids

UFA: unsaturated fatty acid

3'-UTR: 3'-untranslated region

UBI2: ubiquitin extension protein

WGD: whole-genome duplication

Chapter 1

Introduction

Seed oils with high content of polyunsaturated fatty acids (PUFAs), mainly linoleic (LA;18:2$^{cis\Delta 9,12}$) and α-linolenic acid (ALA; 18:3$^{cis\Delta 9,12,15}$), are the major plant-based sources of essential PUFAs for human diets, animal feeds and industrial uses. To meet the ever-increasing demands for edible oils and energy globally, there is an urgent need to develop new strategies to increase the supply of seed oils. A better understanding of seed oil biosynthesis has important implications for developing novel strategies to enhance the oil content and modify the fatty acid composition of the oil for improved food and industrial uses of seed oils.

Triacylglycerols (TAGs) are the major components of seed oils. For decades, it has been believed that the acyl-CoA-dependent Kennedy pathway is the dominant pathway for TAG biosynthesis in most organisms. The Kennedy pathway was first elucidated in the 1950s (Weiss, et al., 1960), and since then, considerable knowledge has been gained about the enzymes involved in this pathway. The pathway involves the sequential acylation of a glycerol backbone via the catalytic action of three acyl-CoA-dependent acyltransferases. The final and the only committed step for TAG synthesis is catalyzed by the enzyme known as acyl-CoA:diacylglycerol acyltransferase (DGAT), which utilizes acyl-CoA and diacylglycerol (DAG) as substrates for TAG formation. At least six distinct types of DGATs, namely DGAT1, DGAT2, DGAT3, defective in cuticular ridges, diacylglycerol acetyltransferase, and a bi-functional DGAT/wax synthase, have been reported in plants (Saha, et al., 2006; Hobbs, et al., 1999; Lardizabal, et al., 2001; Durrett, et al., 2010; Rani, et al., 2010; Li, et al., 2008). Much earlier work demonstrated that TAG biosynthesis in plants can be more complicated than the simple and rather linear Kennedy pathway. In 2000, Stymne and co-workers reported that TAG can also be formed via an

acyl-CoA independent reaction catalyzed by an enzyme called phospholipid:diacylglycerol acyltransferase (PDAT), which uses phospholipids as acyl donors and DAG as acceptor for TAG formation (Dahlqvist, et al., 2000). Multiple *PDAT* genes have been found across eudicots (Pan, et al., 2015). The PDAT orthologs from different species or paralogs within the same species can have different TAG-synthesizing capacities, as well as different substrate selectivity (Ståhl, et al., 2004; Zhang, et al., 2009; van Erp, et al., 2011; Kim, et al., 2011; Pan, et al., 2013).

In addition to diverse mechanisms for TAG synthesis, plants can use several potential routes to accumulate PUFAs in TAG. It is noteworthy that PC, serving as a major site for PUFA synthesis, plays a central role in determining the level of PUFAs in TAGs and membrane lipids. A wealth of studies has shed light on multiple possible pathways controlling the flux of PUFAs from PC into TAG. Besides the direct transfer of PUFAs from PC to TAG catalyzed by the PDAT reaction (Dahlqvist, et al., 2000; Lager, et al., 2013), PUFAs may be transferred into the acyl-CoA pool by the reverse reaction catalyzed by acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT) (Stymne and Stobart, 1984) or by the combined catalytic action of phospholipase A with acyl-CoA synthetase. Alternatively, PUFA-rich PC may be converted into PUFA-rich DAG via the catalytic action of phospholipase C, phospholipase D along with phosphatidic acid phosphatase, and/or the recently discovered phosphatidylcholine: diacylglycerol cholinephosphotransferase (Hu, et al., 2012). The resulting PUFA-rich DAGs or PUFA-CoAs may be further incorporated into TAGs via either or both of acyl-CoA-dependent and -independent pathways.

As discussed, TAG biosynthesis and PUFA accumulation are rather complicated processes with a number of alternative metabolic routes. The overall objective of this PhD project is to use flax (*Linum usitatissimum* L.) as a model system to study the molecular

mechanisms underlying the high PUFA accumulation in plants. I chose flax as a model system because its seed oils naturally contain high amount of PUFAs, especially ALA (~ 45-57%). Upon air exposure, ALA can rapidly react with oxygen in air and be polymerized into soft and stable films on sample surface. This polymerization process is also known as a drying process. Because of its drying properties, flaxseed oil has long been used in wood finishes, paints, coatings and other industrial supplies (Thiessen, 2004; Jhala, et al., 2009). In addition, ALA is a precursor for healthy long chain omega-3 PUFAs (LC-ω-3-PUFAs) (Abeywardena and Patten, 2011), adding a unique nutritional value to flaxseed oil. To accumulate a large proportion of ALA in its seed oils, it is likely that flax may have more efficient mechanisms to transfer ALA from PC into TAG as compared to other plant species. Understanding the mechanisms responsible for high ALA accumulation in flax will be useful for improving oil quality of flax and other oilseeds.

The present doctoral study, which used molecular biological, biochemical and bioinformatics approaches to study the TAG biosynthetic machinery in flaxseed, was guided by the following hypotheses:

1. At least one enzyme involved in the final step of TAG biosynthesis in flax can preferentially incorporate ALA-containing substrates into TAG.
2. Different selection pressures have acted on the duplicated *PDAT* genes during the evolution, leading to the diversification of the *PDAT* genes in eudicots.
3. The DGAT-catalyzed reaction can be coupled to the LPCAT-catalyzed reverse reaction for enhancing the amount of PUFA in TAG.
4. RNA interference-mediated knockdown of *PDATs* or *DGAT1s* in flaxseed can affect ALA and seed oil accumulation.

The thesis starts with a literature review followed by manuscripts based on the four hypotheses and a final chapter containing general discussion.

References

Abeywardena MY, Patten GS (2011) Role of ω3 longchain polyunsaturated fatty acids in reducing cardio-metabolic risk factors. Endocr Metab Immune Disord Drug Targets **11:** 232-246

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M (2010) A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in *Euonymus* and transgenic seeds. Proc Natl Acad Sci U S A **107:** 9464-9469

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Hu Z, Ren Z, Lu C (2012) The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic Arabidopsis. Plant Physiol **158:** 1944-1954

Jhala AJ, Weselake RJ, Hall LM (2009) Genetically engineered flax: Potential benefits, risks, regulations, and mitigation of transgene movement. Crop Sci **49:** 1943-1954

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S (2013)

Plant acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem **288:** 36902-36914

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol **148:** 97-107

Pan X, Peng FY, Weselake RJ (2015) Genome-wide analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes in plants reveals the eudicot-wide *PDAT* gene expansion and altered selective pressures acting on the core eudicot *PDAT* paralogs. Plant Physiol **167:** 887-904

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

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 Physiol **141:** 1533-1543

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Stymne S, Stobart AK (1984) Evidence for the reversibility of the acyl-CoA:lysophosphatidylcholine acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons and rat liver. Biochem J **223:** 305-314

Thiessen D (2004) Optimization of feed peas, canola and flaxseed for Auq feeds: the Canadian Prairie Perspective. In LE Cruz Suare, D Ricque Marie, MG Nieto Lopez, D Villarreal, Scholz, U. y Gonzalez, M., eds, Avances en Nutricion Acuicola VII. Memorias del VII symposium Internacional de Nutricion Acuicola.

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Chapter 2

Literature Review

In this chapter, background knowledge related to the present doctoral project will be presented in the following sections: basic knowledge of polyunsaturated fatty acids (PUFAs); main functions of triacylglycerols (TAGs) in different species; TAG biosynthesis in oilseeds; central role of phosphatidylcholine (PC) for the production of PUFA-containing TAG; and advances in oilseed metabolic engineering.

2.1 Basic knowledge of PUFAs

2.1.1 Nomenclature of fatty acids

Fatty acid (FA) is an aliphatic carboxylic acid consisting of an aliphatic chain and a carboxylic acid functional group at one end. FAs have two terminuses: the carboxyl terminus and the methyl terminus. The delta notation (Δ) designates the carbon position from the carboxyl terminus, whereas the omega notion (ω) or "n" designates the carbon position from the methyl terminus. Unsaturated FAs (UFAs) have one or more double bonds in the chain. FAs with one double bond are monounsaturated FAs (MUFAs) and those with more than one double bond are known as polyunsaturated FAs (PUFAs). The double bond exists in two stereoisomeric forms: the *cis* (or Z) configuration and the *trans* (or E) configuration. In nature, double bonds in FAs are mostly in *cis* configuration. FAs can be designated using the shorthand nomenclature, which indicates the number of carbon atoms, as well as the number and position of double bonds. For example, the notation $18{:}3^{cis\Delta 9,12,15}$ describes α-linolenic acid (ALA), where 18 represents the number of carbons; the number following the colon, 3 in this case, indicates the total number of double bonds present; and the superscript refers to the insertion of *cis* double bonds at the ninth, 12^{th} and 15^{th} carbons from the carboxyl terminus.

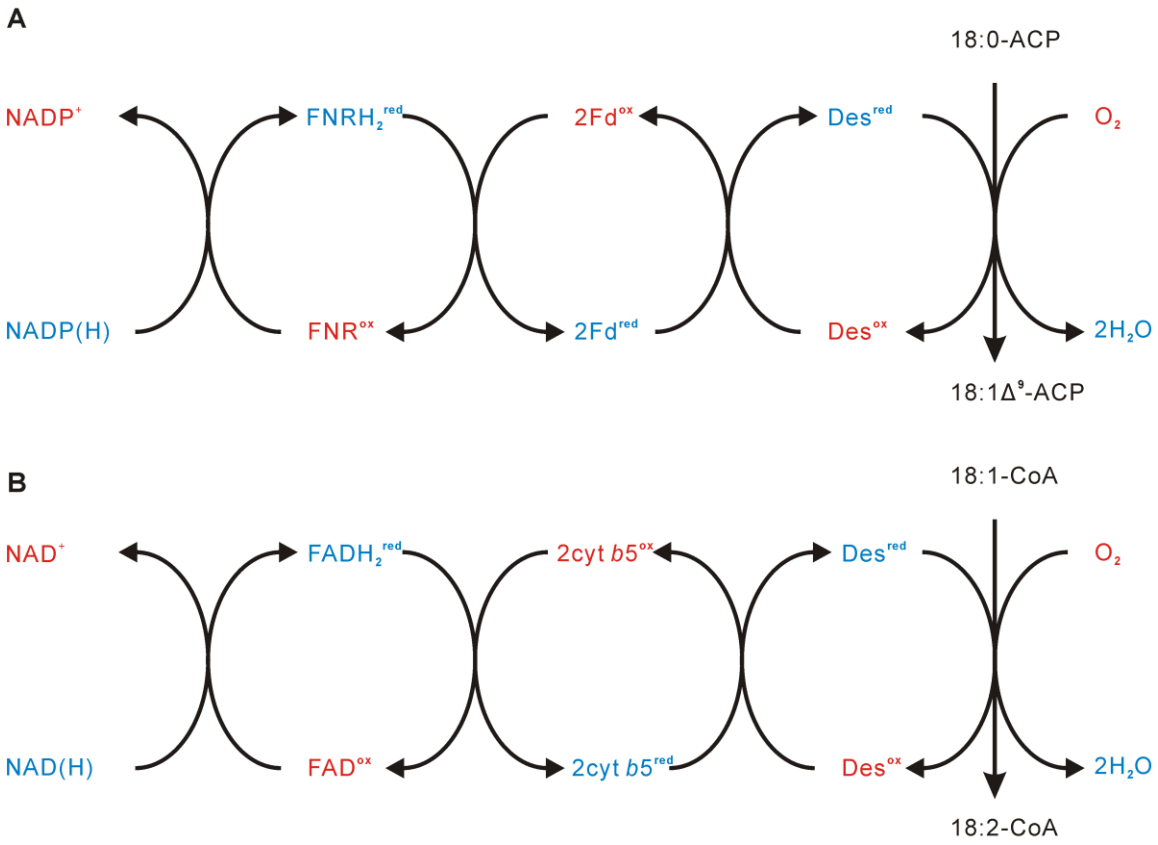
2.1.2 Mechanisms of FA desaturation

FA desaturation is a rather complicated process and involves two major pathways: the prokaryotic desaturation pathway that is confined in the plastid and the eukaryotic desaturation pathway that takes place in the endoplasmic reticulum (ER) (Shanklin and Cahoon, 1998). Linoleic acid (18:2$^{cis\Delta 9,}$12; LA) and ALA are two main PUFAs present in higher plants. The introduction of double bonds into FA chains is catalyzed by enzymes known as FA desaturases (FADs). The plastid and the ER contain analogous enzymes for catalyzing FA desaturation. In the plastid, MUFAs, are formed by the catalytic action of a solubale Δ9 acyl-acyl carrier protein (ACP) desaturase (Ohlrogge and Browse, 1995). In contast, PUFAs, are formed via the catalytic action of membrane-bound desaturase. Omega-6 (ω-6) desaturation, which introduces a double bond at the ω-6 position of oleic acid (18:1$^{cis\Delta 9}$; OA) to form LA, is catalyzed by FAD2 in the ER and FAD6 in the plastid, whereas ω-3 desaturation, which inserts a double bond at the ω-3 position of LA to produce ALA, is carried out by FAD3 in the ER and FAD7/FAD8 in the plastid (Wallis and Browse, 2002). ER-bound desaturases utilize acyl groups esterified to PC as substrates whereas plastid-localized desaturases utilize acyl groups esterified to galactolipids. In seeds, the production of LA and ALA take place almost exclusively in the ER via the enzymatic reactions of FAD2 and FAD3.

The basic mechanism of desaturation reaction is very similar across different types of desaturases. In plants, FA desaturases work through an aerobic mechanism, which requires molecular oxygen (O_2) for their enzymatic function (Shanklin and Cahoon, 1998). In addition to O_2, the desaturation reaction also requires the specific electron transport system and two electrons for the formation of each double bond. The plastid- and the ER-located desaturases have different electron transport systems with functionally equivalent components (Shanklin and

Cahoon, 1998). Both types of electron transport systems consist of five major components: electron donor, flavoprotein, carrier protein, desaturase domain and electron acceptor. Components of the electron transport systems involved in plastid-localized stearoyl-acyl carrier protein (ACP) desaturases (SAD) and ER-localized FAD2 are shown in Figure 2.1. In desaturation reaction, a pair of electrons originating from the electron donor is transferred at the same time to a flavoprotein that donates the electrons one at a time to a carrier protein, which is able to carry a single electron. The electrons are then sequentially passed from two reduced carrier proteins to the desaturase enzyme in order to fulfill the requirement for two-electron reduction of catalysis. In both systems, O_2 serves as the terminal electron acceptor and is completely reduced to water after desaturation. For plastid-located desaturases, the components involved in the electron transport system may depend on tissue type. In photosynthetically inactive tissues, the electron donor is NADPH; ferredoxin-NADP+ oxidoreductase (FNR) is a flavoprotein, and the electron carrier is the 2Fe-2S protein ferredoxin (Figure 2.1A) (Nagai and Bloch, 1966; Schmidt and Heinz, 1990). In photosynthetically active tissues, electrons are provided by photosystem I and directly passed to ferredoxin, which bypasses the FNR-mediated electron transfer. In the ER, the electron transport system is composed of NADH (electron donor), cytochrome *b*5 reductase (flavoprotein), heme protein cytochrome *b*5 (carrier protein), a desaturase domain, and O_2 (electron acceptor) (Dailey and Strittmatter, 1979; Hackett and Strittmatter, 1984; Spatz and Strittmatter, 1971). The insertion of each double bond requires one NAD(P)H as an electron donor and a molecule of O_2 as an electron acceptor for two electron pairs, one from NAD(P)H and the other from the substrate FAs.

Figure 2.1 The diagram illustrates the electron transport chain involved in the plastid-localized stearoyl- acyl carrier protein (ACP) desaturase (SAD)-(A) and the ER-localized fatty acid desaturase 2 (FAD2)-catalyzed reactions (B). The SAD reaction converts 18:0-ACP into 18:1-ACP in plastids. FAD2 catalyzes the conversion of 18:1-CoA into 18:2-CoA in the ER. Oxidized (^{ox}) and reduced (^{red}) states of the components are shown in red and blue, respectively. Black arrows indicate the directions of electron transport.



2.1.3 PUFA impact on plant physiology

PUFAs have many important biological functions in plants. First, trienoic FAs (TFA), ALA and hexadecatrienoic acid (HTA;16:3$^{cis\Delta7,10,13}$), represent the most abundant FAs in membrane lipids of higher plants (Douce and Joyard, 1982; Harwood, 1982). In particular, the thylakoid membrane, with 75-80 % of ALA or the combined ALA and HTA, is one of the most highly unsaturated biological membranes. The most abundant thylakoid lipids, monogalactosyldiacylglycerols, contain over 90% TFA (Douce and Joyard, 1982; Harwood, 1982). It is a widespread belief that UFAs play an essential role in maintaining membrane function at low temperature (Routaboul, et al., 2000). The cell membrane is the primary site for low-temperature (chilling or freezing) injuries. In response to low-temperature stress, cell membranes undergo a lipid phase transition from a fluid, liquid-crystalline phase to a more rigid gel phase (Huang, 2006). Unlike the disordered phospholipid acyl chains in the liquid-crystalline phase, acyl chains in a gel phase are very closely packed and more highly ordered, which affects the normal function of the cell membrane. UFAs can lower the transition temperature by introducing the rigid "kink" of the *cis* double bond in lipid acyl chains, which disrupts the lipid packing, and thereby maintains membrane fluidity under low-temperature stress. The capacity to adjust membrane fluidity can protect cell membranes from injuries caused by low temperature, thus enhancing the cold tolerance of plants. A number of studies reported that the membrane lipids become more unsaturated under the low-temperature stress (Latsague, et al., 1992; Palta, et al., 1993; Samala, et al., 1998; Cyril, et al., 2002). In keeping with this, accumulating evidence also indicated that there is a close relationship between the level of UFAs in membrane lipids and plant tolerance to cold stress (Cyril, et al., 2002; Murata, et al., 1992; Wolter, et al., 1992; Kodama, et al., 1994; Bertin, et al., 1998). In line with the importance of UFAs in cell

membranes during cold acclimation, characterization of the Arabidopsis triple mutant line, *fad3-2fad7-2fad8*, which contains no detectable TFA, revealed that TFAs are not absolutely required for growth and photosynthesis in Arabidopsis under normal growing conditions, however, they are required to maintain normal photosynthetic function during growth of plants at low temperatures (Routaboul, et al., 2000). In contrast, the relationship between TFA and plant response to high temperature is more complicated. Reductions in TFAs showed enhanced thermotolerance when plants were exposed to the high temperature in a short-term. Thermotolerance to long-term exposure at high temperature, however, requires the basal level of TFA (Routaboul, et al., 2012). Moreover, in higher plants, PUFAs, represented mainly by ALA, serving as a precursor for jasmonic acid (a lipid signal molecule) production (Rickauer, et al., 1997) and as a sink for reactive oxygen species, have an important role in defence reaction in plant-pathogen interactions (Mène-Saffrané, et al., 2009). In addition to its involvement in the adaptation to biotic and abiotic stresses, TFA, as a precursor for jasmonic acid, is required for final stages of pollen maturation and release of viable pollen. TFA depletion in Arabidopsis has been shown to result in male sterility (McConn and Browse, 1996).

2.1.4 Applications of PUFA-rich oils with a focus on flax oil

Seed oils with high amount of PUFAs have various food, feed and industrial applications. Flax seed oil with approximately 45-65%% ALA is one of the richest natural sources of omega-3 PUFAs. ALA with three double bonds is particularly susceptible to attack by atmospheric oxygen. Upon oxygen attack, ALA forms a series of products, which ultimately undergo the polymerization process. This polymerization process, also known as the drying process, is often visible by forming a soft and durable film on the sample surface. With a high content of ALA, flax oil is a typical oil with strong drying properties, which is exploited in most of its industrial

applications, such as in the manufacture of printing ink, paints, adhesives, varnishes, and the floor covering linoleum (Jhala and Hall, 2010). In addition, ALA is an essential FA for the human diet and the dietary precursor for the very long-chain omega-3 PUFAs, such as eicosapentaenoic acid (EPA; $20{:}5^{cis\Delta5,8,11,14,17}$) and docosahexaenoic acid (DHA; $22{:}6^{cis\Delta4,7,10,13,16,19}$), which are associated with many health benefits (Sinclair, et al., 2003; Das, 2006). In food applications, flax seed oil is marketed as a nutritional supplement. In addition, flax seed and meal can also be used in livestock feed to produce omega-3 enriched poultry meat products. However, it is important to note that the conversion from ALA to EPA and DHA in humans is very limited with a conversion rate of about 8% for EPA and only 0.05% for DHA (Burdge, 2006). Most research has focused on the health benefits of EPA and DHA. Scientific evidence supporting a beneficial effect of supplemental ALA is less convincing and unclear.

2.2 The main functions of TAGs in different species

TAGs are neutral lipids consisting of three FA molecules esterified to a glycerol backbone (Karantonis, et al., 2009). The occurrence of TAG is widespread in most eukaryotes, namely animal, plant and fungi, and in a few prokaryotes mainly belonging to the actinomycetes group, particularly *Mycobacterium*, *Rhodococcus*, *Streptomycetes*, and *Nocardia* (Alvarez and Steinbüchel, 2003; Wältermann, et al., 2007). As the highly reduced forms of carbon, TAGs serve as the major form of energy storage across eukaryotes. Upon oxidation, they release more energy per gram (over twice as much) than storage carbohydrates or proteins (Graham, 2008; Miquel and Browse, 1995). In addition, TAGs serve as a reservoir of FAs, and upon requirement, they can be mobilized and the released FAs can serve as building blocks for membrane biosynthesis. Besides the shared functions, TAGs have distinct roles in different species.

In mammals, TAGs have many important functions within the body, such as serving as energy storage in adipose tissue and skeletal muscle, protecting cells from FA-induced lipotoxicity, acting as a vehicle for energy transport and as precursors for membrane formation (Coleman and Lee, 2004; Listenberger, et al., 2003; Farese Jr., et al., 2000). Excessive accumulation of TAG in human adipose and non-adipose tissues due to fat-rich diets or sedentary lifestyles, however, is associated with a variety of pathological diseases such as obesity, type 2 diabetes, coronary heart disease and hypertriglyceridemia (Kopelman, 2000; Millar and Billheimer, 2005). With an increased occurrence of these diseases, there has been considerable interest in gaining a better understanding of TAG biosynthesis and metabolism, and using this knowledge to come up with potential pharmacological interventions.

In the case of plants, TAGs are major storage components in seeds of diverse oilseed plants (Stymne and Stobart, 1987) and serve as an important energy store to fuel the post-germinative seedling establishment until photosynthesis becomes efficient (Graham, 2008). These storage lipids are important renewable resources which are used for nutritional, therapeutic, pharmaceutical, fuel, and industrial purposes. In addition to seeds, many other non-seed tissues, such as pollen grains (Murphy and Vance, 1999), flower petals (Hudak and Thompson, 1997), stems and leaves (Durrett, et al., 2008), have been reported to have the ability to synthesize TAGs, albeit at a lower abundance. It has recently been shown that TAG synthesis is essential for normal pollen and seed development in Arabidopsis (Zhang, et al., 2009). Moreover, Fan et al. (2014) reported that phospholipid:diacylglycerol acyltransferase (PDAT)-mediated TAG synthesis is an important intermediate step for diverting free FAs from membrane lipid synthesis towards β-oxidation, and thus maintaining membrane lipid homeostasis in

Arabidopsis leaves (Fan, et al., 2014). Despite these recent findings, the physiological importance of TAG in non-seed tissues remains to be further clarified.

Similarly, TAGs serve as an energy store in microalgae. However, microalgae accumulate a considerable amount of TAGs only when they exposed to unfavorable environmental or stress conditions, such as nitrogen or phosphate starvation (Sharma, et al., 2012). In fungi and some bacteria, as depots of energy, TAGs are important for cell growth and development (Alvarez and Steinbüchel, 2003; Arabolaza, et al., 2008; Sandager, et al., 2002; Daum, et al., 2007).

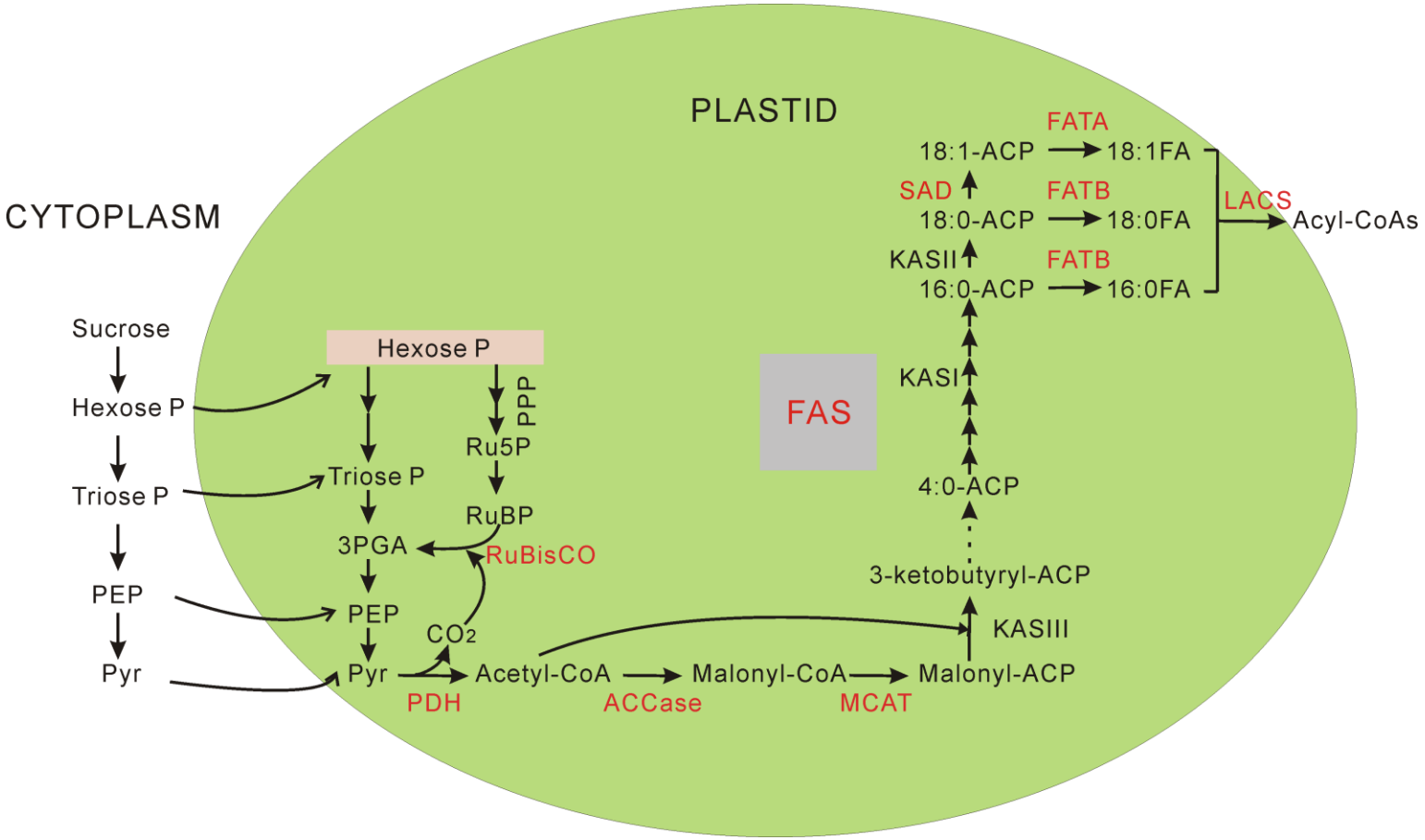
2.3 TAG biosynthesis in oilseeds

The biosynthesis of TAG is accomplished through three major events: *de novo* biosynthesis of FAs in plastids, the transfer of FAs from plastids into the cytoplasm and the packing of the nascent FAs into TAGs.

2.3.1 *De novo* biosynthesis of FAs

Seed oil biosynthesis starts with *de novo* biosynthesis of FAs, which involves four major steps: 1) the conversion of sucrose to acetyl-CoA, 2) the formation of malonyl-CoA, 3) the synthesis of FAs from acetyl-CoA and malonyl-CoA, and 4) the termination of FA synthesis. A simplified scheme is shown in Figure 2.2.

Figure 2.2 Simplified scheme of major pathways involved in sucrose metabolism and *de novo* FA biosynthesis in oilseeds. The compartment nature of the metabolism from sucrose to free FAs involves the cytoplasm and plastids. In most oilseeds, glycolytic breakdown of sucrose provides acetyl-CoAs for FA synthesis. Names of key enzymes are indicated in red. Re-drawn based on Chapman and Ohlrogge) 2012 with some modifications. Abbreviations: ACP, acyl carrier protein; ACCase, acetyl-CoA carboxylase; CoA, coenzyme A; FAS, fatty acid synthase; FAT, acyl-ACP thioesterase; Hexose P, hexose phosphate; KAS, ketoacyl-ACP synthase; LACS, long-chain acyl-CoA synthase; MCAT, malonyl-CoA:ACP acyltransferase; PDH, pyruvate dehydrogenase; PEP, phosphoenolpyruvate; 3PGA, 3-phosphoglycerate; PPP, pentose phosphate pathway; Pyr, pyruvate; Ru5P, ribulose 5-phosphate; RuBP, Ribulose 1,5-bisphosphate; RuBisCO, Ribulose-1,5-bisphosphate carboxylase/oxygenase; SAD, stearoyl-ACP desaturase; Triose P, triose phosphate.


PLASTID
CYTOPLASM
FATA
18:1-ACP
18:1FA
SAD
FATB
18:0-ACP
18:0FA
LACS
Acyl-CoAs
KASII
FATB
16:0-ACP
16:0FA
Sucrose
Hexose P
Hexose P
PPP
Ru5P
FAS
KASI
Triose P
Triose P
RuBP
4:0-ACP
3PGA
RuBisCO
3-ketobutyryl-ACP
PEP
PEP
KASIII
CO2
Pyr
Pyr
Acetyl-CoA
Malonyl-CoA
Malonyl-ACP
PDH
ACCase
MCAT

2.3.1.1 Carbon source and conversion of sucrose to acetyl-CoA

Acetyl-CoA plays a central role in *de novo* FA biosynthesis. In most seeds, photosynthate (mainly in the form of sucrose) provides the ultimate sources of carbohydrates for *de novo* biosynthesis of FAs. The conversion of sucrose to acetyl-CoAs can proceed through multiple pathways, taking place in either or both of two subcellular compartments, the cytoplasm and plastids. In general, sucrose is first converted into hexose phosphates in the cytoplasm. The resulting hexose phosphates are further converted into pyruvates via the glycolytic pathway, occurring in either or both cytoplasm and plastids. Pyruvate is then decarboxylated to acetyl-CoA by pyruvate dehydrogenase (PDH) and this step occurs exclusively in plastids. The reaction catalyzed by PDH decarboxylates a three-carbon pyruvate into a two-carbon acetyl-CoA, resulting in the loss of one carbon as CO_2. However, in green seeds, the presence of the photosystem allows the use of light to provide reductant and ATP for more efficient metabolic routes of transforming carbohydrate into oil (Schwender, et al., 2004). The CO_2 released from PDH action can be recaptured and channeled into FA synthesis via the ribulose-1,5-bisphophate carboxylase activity combined with the non-oxidative steps of the pentose phosphate pathway (Figure 2.2). This provides an alternative metabolic bypass for the conventional glycolytic pathway to oil and achieves increased carbon use efficiency.

2.3.1.2 Formation of malonyl-CoA

After the formation of acetyl-CoA, two enzyme systems, acetyl-CoA carboxylase (ACCase) and FA synthase (FAS), are utilized in plastids for catalyzing *de novo* FA synthesis (Figure 2.2). The first committee step in FA synthesis is catalyzed by ACCase, which converts acetyl-CoA into malonyl-CoA. In dicots, the plastidial ACCase is a multi-enzyme complex consisting of four subunits: biotin carboxylase (BC) (Shorrosh, et al., 1995), biotin carboxylase

carrier protein (BCCP) (Choi, et al., 1995), and α- and β-carboxyltransferases (α-CT and β-CT) (Sasaki, et al., 1993; Kozaki, et al., 2000). In higher plants, the genes encoding BC, BCCP and α-CT are nucleus-located, whereas the fourth gene encoding β-CT is plastid-located. The conversion of acetyl-CoA to malonyl-CoA actually takes place in a two-step reaction (Ohlrogge and Browse, 1995). In the first reaction, bicarbonate (HCO_3^-) is transferred to the nitrogen of a biotin prosthetic group by the catalytic reaction of the BC subunit of ACCase. This reaction is ATP-dependent. The second reaction is catalyzed by the CT subunit of ACCase. During the reaction, the activated carboxyl group is transferred from biotin to acetyl-CoA, yielding malonyl-CoA. Before entering the subsequent FA synthesis pathway, malonyl-CoA has to be transferred and attached covalently to the phosphopantetheine prosthetic group of an acyl carrier protein (ACP). The conversion of malonyl-CoA to malonyl-ACP is catalyzed by the enzyme known as malonyl-CoA:ACP acyltransferase.

2.3.1.3 Biosynthesis of FAs from acetyl-CoA and malonyl-CoA

Acetyl-CoA is the starting unit for FA synthesis and malonyl-ACP provides the activated donor of two-carbon units for FA elongation. The synthesis of FAs in plants from acetyl-CoA and malonyl-ACP is catalyzed by a type-II FAS, which is a multi-subunit complex comprising of monofunctional enzymes (Brown, et al., 2006). FAs are elongated by successive addition of two-carbon units through a repetitive process involving four successive reactions*:* condensation, reduction, dehydration and a second reduction, which are catalyzed by a condensing enzyme (β-ketoacyl-ACP synthase, KAS), first redutase (β-ketoacyl-ACP reductase), dehydrase (β-hydroxyacyl-ACP dehydrase) and a second reductase (enoyl-ACP reductase), respectively. Three types of condensing enzymes are found in plants, known as KASI, KASII and KASIII, which cooperate to elongate acyl chains on ACP up to 16 or 18 carbon atoms in length. KASIII, which

uses acetyl-CoA and malonyl-ACP as substrates, catalyzes the initial condensation reaction. KASI is responsible for the condensation of malonyl-ACP with acyl-ACPs containing acyl chain lengths ranging from 4 to 14 carbons. The final condensation, which is mainly responsible for the conversion of palmitoyl-ACP to stearoyl-ACP, is catalyzed by KASII. It noteworthy that stearoyl-ACP (18:0-ACP) produced by FAS is usually converted to oleoyl-ACP (18:1-ACP) via a nuclear-encoded, plastid-localized soluble Δ9-acyl-ACP desaturase. During the elongation cycle in FA synthesis, the growing FA chains are attached covalently to ACP.

2.3.1.4 Termination of FA synthesis

When FAs reach 16 or 18 carbon atoms long, FA synthesis is terminated by hydrolyzing the thioester bond between the fatty acyl chain and ACP through the catalytic action of acyl-ACP thioesterase. Acyl-ACP thioesterases are classified into two families, FAT A and FAT B (Jones, et al., 1995). FAT A thioesterase preferentially catalyzes the hydrolysis of 18:1-ACPs, while FAT B thioesterase preferably hydrolyzes acyl-ACPs with saturated acyl groups. The end products of FA synthesis are mostly palmitic acid (16:0), stearic acid (18:0) and OA, which are usually in the order of OA>palmitic acid>stearic acid.

2.3.2 Transfer of FAs from plastids into the cytoplasm

One of the major differences between the plastidial and ER lipid metabolism is that, within the plastid, lipid synthesis utilizes the acyl groups attached to the ACPs (acyl-ACPs) as substrates, whereas in the ER, acyl groups bound to CoAs (acyl-CoAs) are used as substrates. Before the newly synthesized FAs can be used for the lipid biosynthesis pathway operating at the ER, they need to be converted into the corresponding acyl-CoAs. In developing oilseeds, the newly synthesized FAs are almost entirely (> 95%) transported across plastids and used for cytoplasmic lipid metabolism (Ohlrogge and Browse, 1995). It is generally presumed that after being released from ACP, the newly synthesized FAs are activated to CoA esters by a long-chain

acyl-CoA synthase (LACS) located in the chloroplast outer envelope and then transported outside the plastid to be modified by the ER-bound enzymes or to be incorporated into complex lipid, such as TAG (Harwood, 1996). Nine genes encoding active LACS enzymes have been identified in Arabidopsis (Shockey, et al., 2002). Among these 9 isoforms, LACS9 has been identified as the major plastidial isoform (Schnurr, et al., 2002). However, a T-DNA disruption of this *LACS9* did not show compromised export of acyl groups from the plastid, suggesting other members of LACSs might have redundant functions to compensate for *LACS9* mutation (Schnurr, et al., 2002). In addition, a continued study showed that LACS4 has overlapping functions with LACS9 in lipid trafficking and both enzymes are involved in channelling lipids from the ER to the plastid instead of the other way around (Jessen, et al., 2015). Without more evidence, it is possible that additional mechanisms are used to incorporate plastid-derived FAs into the ER. A radiolabeling study conducted by Kjellberg et al. (2000) proposed that the newly synthesized FAs may be channelled into PC at the plastid envelop via the acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT) involved acyl-editing and then be transported from the plastid to the ER using PC as a carrier of acyl chains. Others have proposed that direct physical contact between chloroplast and ER membranes may allow lipid exchange between these two organelles (Andersson, et al., 2007). Recently, Li et al. (2015) provided evidences indicating that a novel plastid-localized protein, named as fatty acid export 1 (FAX1), is involved in FA export from plastids into the cytoplasm. Further knowledge on how acyl chains are transferred from the plastid to the ER and then incorporated into TAGs may allow the design of novel strategies to enhance the accumulation of TAGs.

2.3.3 Packing of the nascent FAs into TAGs

Once activated to acyl-CoAs, these acyl groups are available for TAG synthesis, which takes place in the ER. Multiple pathways have been proposed for TAG synthesis. In particular, different oilseeds may utilize different strategies for the assembly of TAG. A generalized scheme for TAG biosynthesis in oilseed plants is depicted in Figure 2.3.

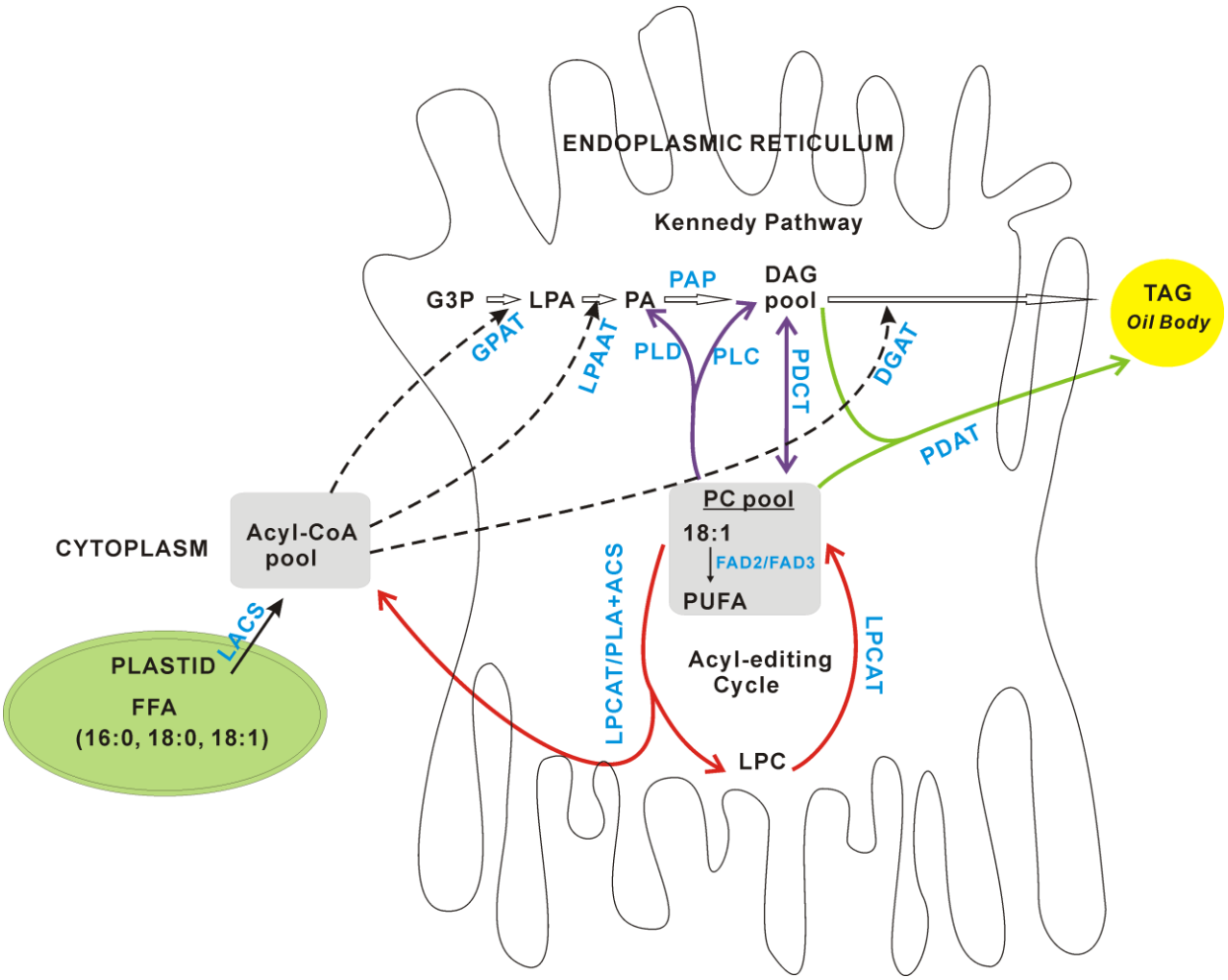
2.3.3.1 Conventional Kenney pathway

The conventional Kenney pathway, also known as an acyl-CoA dependent pathway, was first characterized in animals in the 1950s (Weiss, et al., 1960), and shortly thereafter was identified in plants (Barron and Stumpf, 1962). In this pathway, TAGs are synthesized through three sequential acyl-CoA dependent acylations of G3P and the four key enzymes involved in the pathway are: glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAAT), phosphatidic acid phosphatase (PAP) and DGAT. GPAT initiates TAG synthesis by catalyzing the first acylation of glycerol 3-phosphate (G3P) at the *sn*-1 position to produce lysophosphatidic acid (LPA). The resulting LPA is then acylated at the *sn*-2 position by LPAAT and forms phosphatidic acid (PA). Subsequently, the phosphate group of PA is removed by the catalytic action of PAP and yields diacylglycerol (DAG). The final step of the Kennedy pathway is catalyzed by diacylglycerol acyltransferase (DGAT), which utilizes DAG and acyl-CoA for TAG synthesis.

GPAT

Ten genes encoding GPAT activity have been identified in the Arabidopsis genome. One of these GPATs (At1g32200) is a plastid-localized enzyme with the *sn*-1 regiospecificity (Nishida, et al., 1993). The other eight GPATs, GPAT1-8, are land plant-specific and appear to possess the acylation of G3P at the *sn*-2 position that are almost certainly not responsible for

Figure 2.3 A schematic diagram showing the Kennedy pathway for TAG synthesis and the proposed mechanisms involved in transferring polyunsaturated fatty acids (PUFAs) from phosphatidylcholine (PC) into triacylglycerol (TAG). The Kennedy pathway is indicated by white arrows and acyl transfer reactions are shown in dashed lines. Phospholipid:diacylglycerol acyltransferase (PDAT)-catalyzed TAG synthesis is indicated by green arrows. Purple arrows indicate reactions involved in the conversion of PC to diacylglycerol (DAG). Red arrows indicate reactions involved in CoA:PC acyl exchange. Abbreviations: ACS, acyl-CoA synthetase; DGAT, acyl-CoA: diacylglycerol acyltransferase; FAD, fatty acid desaturase; FFA, free fatty acid; G3P, glycerol-3-phosphate; GPAT, acyl-CoA: glycerol-3-phosphate acyltransferase; LACS, long-chain acyl-CoA synthase; LPA, lysophosphatidic acid; LPAAT, acyl-CoA: lysophosphatidic acid acyltransferase; LPC, lysophosphatidylcholine; LPCAT, acyl-CoA:lysophosphatidylcholine acyltransferase; PA, phosphatidic acid; PAP, phosphatidic acid phosphatase; PDCT, phosphatidylcholine: diacylglycerol cholinephosphotransferase; PLA, phospholipase A; PLC, phospholipase C; PLD, phospholipase D.


ENDOPLASMIC RETICULUM
Kennedy Pathway
G3P
LPA
PA
PAP
DAG pool
TAG
Oil Body
GPAT
LPAAT
PLD
PLC
PDCT
DGAT
PDAT
PC pool
18:1
FAD2/FAD3
PUFA
CYTOPLASM
Acyl-CoA pool
LACS
PLASTID
FFA
(16:0, 18:0, 18:1)
LPCAT/PLA+ACS
Acyl-editing Cycle
LPCAT
LPC

TAG synthesis; instead, at least five of these eight GPATs are required for the synthesis of extracellular lipids, such as cutin or suberin (Yang, et al., 2012). Another isoform of GPAT, namely GPAT9 (At5g60620), is an ER-localized enzyme and a homolog of mammalian and yeast GPATs involved in TAG synthesis. GPAT9 may be an acyl-CoA dependent *sn*-1 GPAT (Gidda, et al., 2009), but this needs to be confirmed experimentally.

LPAAT

Genes encoding LPAAT, which catalyzes the second step of the Kennedy pathway, have been cloned and studied in many plant species (Maisonneuve, et al., 2010; Brown, et al., 1994; Knutzon, et al., 1995; Kim, et al., 2005). Five *LPAAT* genes have been identified in Arabidopsis. Among these five genes, only one gene (AtLPAAT1) encodes the plastidial LPAAT (At4g30560) and the other four encode cytoplasmic enzymes. Among the cytoplasmic *LPAATs*, *LPAAT3* (At1g51260) transcripts were detected predominately in pollen, and transcripts of *LPAAT4* (At1g75020) and *LPAAT5* (At3g18850) were ubiquitously detected in diverse tissues, albeit at very low levels. LPAAT2 has been confirmed to localize in the ER and is abundantly present in diverse tissues. A recent study showed that over-expression of a *Brassica napus LPAAT* isoform, which is homologous to *AtLPAAT2* (At3g57650), in Arabidopsis resulted in increased seed oil content and seed weight, indicating that BnLPAAT2 is likely to have an importance role in TAG synthesis (Maisonneuve, et al., 2010). However, genetic studies with AtLPAAT2 have been limited since the heterozygous mutant (*LPAAT2/lpaat2*) is female gametophyte lethal (Kim, et al., 2005). The direct role of LPAAT2 in TAG biosynthesis in Arabidopsis remains to be further explored.

PAP

PAP, the enzyme involved in the penultimate step in the Kennedy pathway, was first purified and characterized from the yeast *Saccharomyces cerevisiae* in 1989 (Lin and Carman, 1989). The gene encoding this enzyme, known as *PHOSPHATIDATE PHOSPHOHYDROLASE1* (*PAH1*), was identified only recently in 2006 (Han, et al., 2006) using a reverse genetics approach based on amino acid sequence information derived from the purified enzyme. The yeast *pah1*Δ mutant resulted in a significant decrease in TAG accumulation and an increase in total phospholipid content (Fakas, et al., 2011). The identification of the yeast *PAH1* gene led to the discovery of mammalian *Lipins* and multiple homologous genes in plants. Arabidopsis contains two orthologs of yeast PAH1, which are AtPAH1 and AtPAH2. The *pah1/pah2* double mutant only resulted in a 15% reduction of total FAs in seeds, suggesting that additional PAH enzymes can contribute to TAG synthesis (Eastmond, et al., 2010). It is noted that at least 11 genes have been annotated as potential PAPs in Arabidopsis genome (Bates, et al., 2013). The isoform(s) of PAP responsible for the majority of DAG produced from the Kennedy pathway has not yet been identified.

DGAT

DGAT, which catalyzes the final and only committed step in TAG synthesis in the Kennedy pathway, is a well-studied enzyme. Different types of DGATs that belong to unrelated protein families have been identified in plants. The first *DGAT* gene was identified and cloned from mouse in 1998 (Cases, et al., 1998) and its encoded enzyme was termed type1 DGAT (DGAT1), which belongs to the superfamily of membrane-bound O-acyltransferase (Hofmann, 2000). Subsequently, the first plant ortholog of the mouse *DGAT1* gene was cloned from Arabidopsis by four independent groups in 1999 (Hobbs, et al., 1999; Zou, et al., 1999;

Routaboul, et al., 1999; Bouvier-Navé, et al., 2000). The essential role of AtDGAT1 in seed oil biosynthesis was confirmed by multiple studies, in which an insertional mutation (AS11) (Zou, et al., 1999) or a frameshift mutation (ABX45) (Routaboul, et al., 1999) in this gene resulted in a significant reduction in seed oil levels. The orthologs of the *DGAT1* gene have also been identified and characterized in many other plant species, such as tobacco [*Nicotiana tabacum*] (Bouvier-Navé, et al., 2000), canola [*Brassica napus*] (Nykiforuk, et al., 2002), soybean [*Glycine max*] (Wang, et al., 2006), sunflower [*Helianthus annuus*] (Sun, et al., 2011), tung tree [*Vernicia fordii*] (Shockey, et al., 2006), castor [*Ricinus communis*] (He, et al., 2004), burning bush [*Euonymus alatus*] (Milcamps, et al., 2005), and nasturtium [*Tropaeolum majus*] (Xu, et al., 2008).

In 2001, Lardizabal et al. identified the first type 2 DGAT (DGAT2) from the oleaginous fungus *Umbelopsis* (formerly *Mortierella*) *ramanniana.* Despite their relevance in TAG synthesis, DGAT2 shows no DNA or protein sequence similarities with DGAT1 and belongs to a seven-member gene family, which also contains acyl-CoA:monoacylglycerol acyltransferase-1 (MGAT1), MGAT2, MGAT3, and wax monoester synthase. Furthermore, compared to DGAT1, DGAT2 is a smaller protein with less transmembrane domains (1 or 2 versus 6 or more). In the same study, Lardizabal et al. identified the first plant DGAT2 from Arabidopsis (AtDGAT2; At3g51520); however, the heterologous expression of this *AtDGAT2* in insect cells produced an enzyme with negligible activity. In addition, the heterologous expression of the native *AtDGAT2* failed to complement TAG synthesis mutation in yeast mutant (Zhang, et al., 2009; Liu, et al., 2012). The functionality of AtDGAT2 in TAG synthesis was not demonstrated until recently by expressing a codon-optimized version of *AtDGAT2* in yeast (Aymé, et al., 2014). *AtDGAT2* was also transiently functionally expressed in *Nicotiana benthamiana* leaves (Zhou, et al., 2013).

However, the *dgat2* mutant did not show any changes in seed oil content and adding the *dgat2* mutation in a *dgat1* mutant background did not lead to a further reduction of seed oil compared to the *dgat1* mutation alone, suggesting that DGAT2 does not play a substantial role in seed oil accumulation (Zhang, et al., 2009). The physiological function of DGAT2 in Arabidopsis remains to be further determined. Characterization of DGAT2 from tung tree and castor (Kroon, et al., 2006; Shockey, et al., 2006) revealed that DGAT2 exhibits a unique substrate preference for unusual FAs, such as α-eleostearic acids ($18:3^{\Delta cis9,trans11,trans13}$) and ricinoleic acids (18:1-OH;12-hydroxy-9-*cis*-octadecenoic acid), and thus plays an important role in incorporating these FAs into TAGs. It was confirmed that DGAT1 and DGAT2 exist in nearly all eukaryotes, with the exception of yeast *S. cerevisiae* and *Candida albicans* and Basidiomycetes fungi, which do not have a *DGAT1* gene in their genome. The yeast, *Yarrowia lipolytica*, has a DGAT1 encoded by *DGA2* (Beopoulos, et al., 2012; Zhang, et al., 2012). A recent study revealed that DGAT1 and DGAT2 from tung tree differ with regard to their expression patterns, substrate preferences as well as distributions on the ER, suggesting that tung tree DGAT1 and DGAT2 have non-redundant functions in TAG synthesis (Shockey, et al., 2006).

Unlike membrane-bound DGAT1 and DGAT2, the third type of DGAT (DGAT3) is a soluble cytosolic enzyme. The gene encoding DGAT3 was first identified from developing peanut (*Arachis hypogaea*) cotyledons (Saha, et al., 2006; Tumaney, et al., 2001). An ortholog of the peanut DGAT3 was later characterized in Arabidopsis (At1g48300) (Hernández, et al., 2012). The heterologous expression of *AtDGAT3* in *Nicotiana benthamiana* leaves confirmed its involvement in TAG synthesis. Further analyses showed that AtDGAT3-catalyzed TAG synthesis plays an important role in recycling of 18 carbon PUFAs (18:2 and 18:3) into TAGs when seed oil breakdown is impaired. Rani et al. reported another type of soluble DGAT

(At5g23940) in Arabidopsis, known as defective in cuticular ridges (DCR), which is involved in the cutin biosynthesis (Rani, et al., 2010). These findings revealed the presence of a cytosolic TAG biosynthetic pathway in plants. In addition to DGAT1, DGAT2 and soluble DGATs, other types of enzymes have been reported with DGAT activity. A novel bifunctional wax ester synthase (WS) /DGAT was identified from the bacterium *Acinetobacter calcoaceticus* (Kalscheuer and Steinbüchel, 2003). WS/DGAT-related proteins are not widely present among the prokaryotes and seem to be almost exclusively restricted to the Actinomycetes. A recent TBLASTN search, however, showed that a large number of bacterial WS /DGAT orthologs occur in plants. Characterization of the petunia and Arabidopsis WS /DGAT orthologs indicated that unlike its counterpart in *A. calcoaceticus*, the petunia and Arabidopsis WS /DGAT enzymes predominantly function as wax synthases involved in the synthesis of cuticular wax-esters (King, et al., 2007; Li, et al., 2008). Alternatively, TAG can also been synthesized by DAG: DAG transacylase, which uses two DAGs as substrates and forms TAG and monoacylglycerol. To date, a DAG: DAG transacylase gene has not been identified in plants (Mancha and Stymne, 1997; Stobart, et al., 1997).

2.4 The central role of PC for the production of PUFA-containing TAGs

The *sn*-2 position of PC is the primary site for the eukaryotic pathway for FA modifications, including desaturation. Therefore, alternative pathways, using PC either as direct substrates or as intermediates in the flux of FAs or DAGs, for TAG synthesis are crucial for the formation of PUFA-containing TAGs. As shown in Figure 2.3, three major mechanisms have been proposed for transferring PUFAs from PC into TAGs: phospholipid:diacylglycerol acyltransferase (PDAT)-mediated mechanism, acyl-editing mechanism and the conversion of PC to DAG.

2.4.1 PDAT-meditated mechanism for TAG synthesis

In addition to the conventional Kennedy pathway, it was discovered that TAG can also be formed by the enzymatic action of PDAT. PDAT catalyzes the acyl-CoA-independent transfer of an acyl group from the *sn*-2 position of the phospholipid to the *sn*-3 position of DAG, forming TAG and *sn*-1-lyso-phospholipid. The PDAT activity was first detected in microsomal preparations of castor, sunflower, and *Crepis palaestina* by Stymne and co-workers in 2000. In the same study, the first *PDAT* gene (*YNR008w, LRO1*) was cloned from the yeast *Saccharomyces cerevisiae. S. cerevisiae* PDAT, however, has been shown to exhibit low DAG:DAG transacylase activity (Ghosal et al. 2007). *LRO1 is* a homolog of lecithin:cholesterol acyltransferase (LCAT), which is a soluble enzyme catalyzing the formation of cholesteryl ester in blood plasma. LRO1 and DGA1 (the yeast homolog of the DGAT2 family) are key enzymes responsible for the majority of TAG synthesis in yeast (Oelkers, et al., 2002). A topology study (Choudhary, et al., 2011) showed that the active site of LRO1 is localized in the lumen of the ER, and this differs from the topological orientation of DGA1, which has the acyl-CoA binding domain facing the cytosolic side of the ER. Different topological orientations suggested that LRO1 and DGA1 are likely to catalyze TAG synthesis on different sides of the ER membrane. The relative contribution of LRO1 and DGA1 in TAG synthesis depends on growth phase. LRO1 appears to be a predominate contributor to TAG synthesis during logarithmic growth, whereas DGA1 exhibits major activity in stationary phase (Oelkers, et al., 2002; Sandager, et al., 2002).

Subsequently, two PDAT orthologs were identified in Arabidopsis, referred to as AtPDAT1 (At5g13640) and AtPDAT2 (At3g44830). Topology analysis showed that similar to yeast PDAT (LRO1), AtPDATs contain a short N-terminal cytoplasmic tail and a single

transmembrane domain with the rest of the protein sitting in the ER lumen. *AthPDAT1* expressed at higher levels in leaves than in seeds (Winter, et al., 2007; Pan, et al., 2015), whereas the opposite pattern was observed for *AthPDAT2*. Overexpression or knockout of *AtPDAT1* in Arabidopsis resulted in significant changes in oil content and FA composition in leaves but not in seeds (Mhaske, et al., 2005). A further study indicated that the AtPDAT1-mediated TAG synthesis is involved in the process of diverting FAs from membrane lipids towards peroxisomal β-oxidation, and thereby is important for maintaining membrane lipid homeostasis in Arabidopsis leaves (Fan, et al., 2014). The role of AtPDAT1 in TAG synthesis has also been studied in the absence of AtDGAT1 activity. Zhang, et al. (2009) reported that when the expression of *PDAT1* was suppressed by an RNAi strategy under a *dgat1* knockout background, the oil content was further reduced by 63% compared with *dgat1* control. This result suggested that AtPDAT1 is the major contributor for seed oil synthesis when AtDGAT1 activity is compromised. The embryonic lethality in the Arabidopsis *dgat1 pdat1* double mutant suggested that PDAT1 and DGAT1 have overlapping functions for TAG synthesis in pollens and seeds and their expression is essential for pollen and seed viability. Studies comparing the contribution of AtPDAT1 and AtDGAT1 to TAG synthesis in leaves have resulted in different conclusions. Fan et al. (2013) first compared TAG levels of the WT with that of *dgat1* and *pdat1* mutants. In young leaves, in contrast to the 57% reduction in the *pdat1* mutant, there was only a 31% decrease in TAG level in the *dgat1* mutant. They further compared TAG levels of the WT with that of 35S-*PDAT1* and 35S-*DGAT1* overexpression lines and the results showed that over-expressing *PDAT1* but not *DGAT1* resulted in significant increases in TAG accumulation in leaves. These results suggested that AtPDAT1 plays a more important role than AtDGAT1 in TAG synthesis in young Arabidopsis leaves. In contrast, pulse-chase labeling experiments

showed that [^{14}C]12:0 was incorporated into TAG by leaves of the *pdat1* mutant at a much higher rate than that of the *dgat1* mutant, indicating that DGAT1 is the predominate enzyme involved in TAG synthesis in young Arabidopsis leaves (Tjellström, et al., 2015). The relative contribution of AtPDAT1 and AtDGAT1 in lipid metabolism needs to be further explored. The *pdat2 dgat1* double mutant did not show a significantly further decrease in oil content beyond the decrease caused by the *dgat1* single mutant, which indicated that AtPDAT2 has no substantial role in TAG biosynthesis even though it is highly expressed in seeds.

In addition to Arabidopsis PDATs, three castor (Kim, et al., 2011; van Erp, et al., 2011), six flax (*Linum usitatissimum*) (Pan, et al., 2013) and a single green alga (*Chlamydomonas reinhardtii*) (Yoon, et al., 2012) PDATs have been, so far, functionally characterized. Overall, these characterization results support that multiple PDAT paralogs arising from the core eudicot-shared ancient genome duplication may have evolved with different TAG-synthesizing capacities and developed divergent expression patterns due to varied selection pressures (Pan, et al., 2015). It is also important to note that both *in vivo* and *in vitro* approaches revealed that some PDATs appear to have enhanced preferences for modified FAs, including PUFAs and unusual FAs. Dahlqvist et al. demonstrated that the microsomal PDAT from yeast has a preference for ricinoleoyl- or vernoloyl (12-epoxyoctadeca-*cis*-9-enoyl)-DAG over dioleoyl-DAG (Dahlqvist, et al., 2000). The PDAT activity in microsomal preparations of leaves from a *AtPDAT1* overexpresser showed that AtPDAT1 has a strong preference for PC containing acyl groups with several double bonds, epoxy, or hydroxyl groups (Ståhl, et al., 2004). Recently, Pan et al. (2013) identified a pair of flax *PDAT* genes, which are preferentially expressed in seeds and encode enzymes with the unique ability to efficiently channel ALAs into TAGs. Similarly, it appears that castor also contains a specialized PDAT for the selective transfer of hydroxy FAs into TAGs

(van Erp, et al., 2011; Kim, et al., 2011). Together, these results suggested that the contribution of PDAT to TAG synthesis in seeds might be significant in some oilseeds that are high in PUFAs or unusual FAs. These results also suggested that the contribution of PDAT1 in TAG synthesis might vary widely among plants and different tissues within the same plants. To provide a better understanding of role of PDAT, it is essential to extend the analyses to other plant tissues and species.

2.4.2 Acyl-editing mechanism

Acyl editing involves a cycle of rapid acylation of LPC to PC and deacylation of PC to LPC, resulting in the dynamic exchange of acyl groups between PC and acyl-CoA pools. This process generates new molecular species of PC but it does not lead to net PC synthesis. The acyl editing cycle allows: 1) newly synthesized FAs from the plastid to enter PC for modifications, and 2) PC-modified FAs, including PUFAs, to be released into the acyl-CoA pool where they can be further incorporated into TAGs or used for other reactions. This acyl exchange was proposed to be a mechanism for PUFA enrichment in TAGs in oilseeds. Substantial progress has been made in understanding the molecular identify of enzymes involved in acyl editing.

2.4.2.1 The forward and reverse reactions of LPCAT

An acyl exchange between acyl-CoA and PC pools was demonstrated in soybean microsomal preparations over 30 years ago (Stymne and Glad, 1981). A few years later, it was suggested that the combined forward and reverse reactions of LPCAT are responsible for this acyl exchange (Stymne and Stobart, 1984). This LPCAT-involved mechanism for acyl editing is supported by the *in vivo* and *in vitro* metabolic analyses of acyl fluxes through PC in various plant tissues (Stymne and Stobart, 1987; Bates and Browse, 2012). LPCAT activity has been detected in chloroplasts (Bessoule, et al., 1995; Kjellberg, et al., 2000) and in microsomal

preparations from the developing seeds of several plant species (Rochester and Bishop, 1984; Stymne and Stobart, 1984; Bafor, et al., 1991; Demandre, et al., 1994).

Recent metabolic labeling experiments in multiple plant tissues showed that the majority of nascent FAs from the plastid are initially incorporated into PC largely through an acyl-editing mechanism rather than via the Kennedy pathway (Bates, et al., 2007; Bates, et al., 2009; Bates and Browse, 2012; Tjellström, et al., 2012). Subsequent studies showed that the Arabidopsis *lpcat1 lpcat2* double mutant was unable to incorporate nascent FAs into PC via acyl-editing in seeds, indicating that the forward reaction of LPCAT is responsible for the acyl-editing based initial flux of nascent FAs into PC (Bates, et al., 2012). The critical role of LPCAT2-mediated PC generation for PDAT-catalyzed TAG formation in Arabidopsis has been recently reported (Xu, et al., 2012). Despite substantial progress has been made in understanding of the LPC-acylation half of the acyl-editing cycle, the involvement of the reverse reaction of LPCAT for the PC-deacylation remains largely unknown. The equilibrium of the LPCAT-catalyzed reaction is normally far towards the exothermic forward reaction for the synthesis of PC. It was not until recently that the plant *LPCAT* genes were cloned and characterized (Ståhl, et al., 2008). This allows more rigorously approaches to evaluate the role of LPCAT in the acyl editing cycle. Now after more than two decades, Stymne and his co-workers have proven that plant LPCAT enzyme does in fact operate in a reversible fashion under *in vitro* conditions (Lager, et al., 2013).

2.4.2.2 Lands cycle

The second possible mechanism for acyl editing involves the PC-deacylation catalyzed by a combination of PLA (Chen, et al., 2011) and acyl-CoA synthetase (Shockey, et al., 2002) to produce LPC, followed by re-acylation of LPC to regenerate PC mediated by the forward reaction of LPCAT. This mechanism is also known as “the Lands cycle” (Lands, 1960).

Three major classes of PLAs have been identified in plants: DAD-like phospholipase A_1 (PLA_1), patatin like phospholipase (pPLA), and small molecular weight secretory phospholipase A2 ($sPLA_2$) (Chen, et al., 2011). Ten *pPLA* genes have been identified in Arabidopsis genome and can be classified as *pPLAI*, *pPLAII* (α, β, γ, δ, ε), and *pPLAIII* (α, β, γ, δ) (Scherer, et al., 2010). Four isoforms of $sPLA_2$s have been reported in Arabidopsis, which are names as $sPLA_2$α, -β, -γ, and –δ (Lee, et al., 2005). Recent studies suggested that the patatin-like PLA-IIIδ (pPLA-IIIδ) and the small molecular PLA_2α ($sPLA_2$α) may be involved in PC-deacylation half of the acyl editing cycle (Li, et al., 2014; Li, et al., 2013; Bayon, et al., 2015). pPLAIIIδ is a membrane bound protein which catalyzes the hydrolysis of PC at both *sn*-positions with a preference for the *sn*-2 position. This enzyme also has the acyl-CoA thioesterase activity, which hydrolyzes acyl-CoAs to free FAs. The overexpression of *pPLA-IIIδ*, resulted in enhanced seed oil content with an increased proportion of long chain (20- and 22- carbon) FAs, and its gene knockout plants had significantly decreased seed oil content compared with WT seeds (Li, et al., 2013; Li, et al., 2014). Bayon et al (2015) recently isolated and functionally characterized a $sPLA_2$α from castor. This *$sPLA_2$α* gene has the high relative expression in developing castor endosperm. *In vitro* enzyme assays further showed that this castor $sPLA_2$α is highly specific for ricinoleoyl moieties compared with that of Arabidopsis $sPLA_2$α. In addition, co-expression of castor *$sPLA_2$α* with castor hydroxylase (*RcFAH12*) in Arabidopsis resulted in a significant decrease in the hydroxy FA content not only in PC and but also in neutral lipid fraction of seeds, comparing to *RcFAH12* expression alone. Together, these results suggest that this castor $sPLA_2$α is involved in the de-acylation of ricinoleoyl moieties from the membrane lipids in transgenic Arabidopsis seeds.

2.4.3 PC-derived DAG production

In addition to the above mechanisms, PUFA-rich PC may be converted into PUFA-rich DAG, which may be further used for TAG synthesis and thus produce PUFA-rich TAG. Three alternative enzymatic routes may be involved in this PC to DAG conversion: 1) a lipase-based reaction catalyzed by phospholipase C or by the combined actions of phospholipase D and PAP; 2) the reverse reaction of CDP-choline:diacylglycerol cholinephosphotransferase (CPT) (Slack, et al., 1983; Slack, et al., 1985); and 3) the recently discovered phosphatidylcholine: diacylglycerol cholinephosphotransferase (PDCT) (Hu, et al., 2012). The PDCT enzyme was first discovered in Arabidopsis and is encoded by the *REDUCED OLEATE DESATURATION1* (*ROD1*) gene (At3g15820) (Hu, et al., 2012). PDCT belongs to the large family of lipid phosphatase/phosphotrasferase (LPT) proteins and catalyzes the transfer of the phosphocholine head group from PC to DAG. In contrast with the CPT reaction, which yields a net production of PC from DAG, the PDCT reaction produces new molecular species of PC and DAG and does not lead to net gain of PC or DAG. It is noteworthy to mention that due to the unfavorable nature of the reverse reaction, CPT might process mainly in the forward reaction for net synthesis of PC. The PDCT reaction allows: first, the flux of 18:1 into PC for further desaturation while second, the desaturated FAs (18:2 and 18:3) are channeled into the DAG pool, thus providing PUFA-rich DAG towards TAG synthesis. The Arabidopsis *rod1* mutant had substantially reduced levels of 18:2 and 18:3 with concomitant increase in the level of 18:1 in its seed oil. Recent evidence indicated that the main PC-DAG conversion occurs via the PDCT-catalyzed reaction (Lu, et al., 2009; Hu, et al., 2012). In addition, analyses of an Arabidopsis triple mutant *rod1/lpcat1/lpcat2* indicated that LPCAT-involved acyl editing and PDCT-based PC-DAG interconversion are

major mechanisms controlling the flux of PUFAs from PC into TAGs in Arabidopsis seeds (Bates, et al., 2012).

2.5 Advances in oilseed metabolic engineering

Nowadays, the main goal of oilseed biotechnology is to boost seed oil content and to produce a desirable FA composition. Over the last few decades, in part due to the technological advances in the field of molecular biology, coupled with advances in genomic technologies, huge strides have been made in our understanding of the genes involved in the biochemical pathways by which plants use for FA and TAG synthesis at the molecular level. Increasing knowledge about the genes involved in plant lipid metabolism makes it possible to enhance seed oil content and/or improve the composition of oils via regulating the expression level of a single gene or simultaneously manipulating the expression of specific combinations of multiple genes. Increased seed oil accumulation has been achieved in a number of studies by altering the expression levels of genes, which can either increase the supply from upstream sources to oil synthesis, or enhance the strength of 'demand' or 'sink' in the final steps of a pathway to draw FAs into TAGs, or both. For example, over-expressing the *WRINKLED1* which encodes a transcription factor affecting glycolysis and FA synthesis and *DGAT1* together with RNAi suppression of the *TAG LIPASE SUGAR-DEPENDENT1* resulted in a significant increase in seed oil content in Arabiodpsis (van Erp, et al., 2014). In addition to the seed oil content, silencing of endogenous genes and/or heterlogously overexpressing exogenous genes encoding the essential enzymes for FA or TAG synthesis has also been successfully used to increase the production of desirable FAs. A recent example is that co-expression of the *Euonymus alatus DIACYLGLYCEROL ACETYLTRANSFERASE* (*EaDAcT*) gene combined with RNAi suppression of endogenous *DGAT1* genes resulted in the accumulation of up to 85 mol % of

unusual 3-acetyl-1,2-diacyl-*sn*-glycerols (acetyl-TAGs) in the seed oil of the best (field-grown) transgenic camelina lines (Liu, et al., 2015).

Despite these important advances, there still remain many gaps in our understanding of all aspects of TAG synthesis, as examples, the precise molecular identities for many key enzymes (i.e., GPAT and PAP) and the relative contributions to TAG synthesis via the different metabolic pathways in different plant species or different tissues within the same plant need to be more fully investigated. A better understanding of these areas will be critical to future rational metabolic engineering approaches to produce high-value specialty oils for food, feed, biofuel and industrial applications.

2.6 References

Alvarez HM, Steinbüchel A (2003) Triacylglycerols in prokaryotic microorganisms. Appl Microbiol Biotechnol **60:** 367-376

Andersson MX, Goksör M, Sandelius AS (2007) Membrane contact sites: Physical attachment between chloroplasts and endoplasmic reticulum revealed by optical manipulation. Plant Signal Behav **2:** 185-187

Arabolaza A, Rodriguez E, Altabe S, Alvarez H, Gramajo H (2008) Multiple pathways for triacylglycerol biosynthesis in *Streptomyces coelicolor*. Appl Environ Microbiol **74:** 2573-2582

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):** e92237

Bafor M, Smith MA, Jonsson L, Stobart K, Stymne S (1991) Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor bean (*Ricinus communis*) endosperm. Biochem J **280:** 507-514

Barron EJ, Stumpf PK (1962) Fat metabolism in higher plants XIX. The biosynthesis of triglycerides by avocado-mesocarp enzymes. Biochim Biophys Acta **60:** 329-337

Bates PD, Browse J (2012) The significance of different diacylgycerol synthesis pathways on plant oil composition and bioengineering. Front Plant Sci **3:** 147-147

Bates PD, Durrett TP, Ohlrogge JB, Pollard M (2009) Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. Plant Physiol **150:** 55-72

Bates PD, Fatihi A, Snapp AR, Carlsson AS, Browse J, Lu C (2012) Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. Plant Physiol **160:** 1530-1539

Bates PD, Ohlrogge JB, Pollard M (2007) Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. J Biol Chem **282:** 31206-31216

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

Bayon S, Chen G, Weselake R, Browse J (2015) A small phospholipase A2-α from castor catalyzes the removal of hydroxy fatty acids from phosphatidylcholine in transgenic Arabidopsis seeds. Plant Physiol **167:** 1259-1270

Beopoulos A, Haddouche R, Kabran P, Dulermo T, Chardot T, Nicaud JM (2012) Identification and characterization of DGA2, an acyltransferase of the DGAT1 acyl-CoA:diacylglycerol acyltransferase family in the oleaginous yeast *Yarrowia lipolytica*. New insights into the storage lipid metabolism of oleaginous yeasts. Appl Microbiol Biotechnol **93:** 1523-1537

Bertin P, Bullens P, Bouharmont J, Kinet JM (1998) Somaclonal variation and chilling tolerance improvement in rice: changes in fatty acid composition. Plant Growth Regul **24:** 31-41

Bessoule JJ, Testet E, Cassagne C (1995) Synthesis of phosphatidylcholine in the chloroplast envelope after import of lysophosphatidylcholine from endoplasmic reticulum membranes. Eur J Biochem **228:** 490-497

Bouvier-Navé P, Benveniste P, Oelkers P, Sturley SL, Schaller H (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. Eur J Biochem **267:** 85-96

Brown AP, Affleck V, Fawcett T, Slabas AR (2006) Tandem affinity purification tagging of fatty acid biosynthetic enzymes in *Synechocystis*sp. PCC6803 and *Arabidopsis thaliana*. J Exp Bot **57:** 1563-1571

Brown AP, Coleman J, Tommey AM, Watson MD, Slabas AR (1994) Isolation and characterisation of a maize cDNA that complements a 1-acyl *sn*-glycerol-3-phosphate acyltransferase mutant of *Escherichia coli* and encodes a protein which has similarities to other acyltransferases. Plant Mol Biol **26:** 211-223

Burdge GC (2006) Metabolism of alpha-linolenic acid in humans. Prostaglandins Leukot Essent Fatty Acids **75:** 161-168

Cases S, Smith SJ, Zheng Y, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese Jr. RV (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A **95:** 13018-13023

Chapman KD, Ohlrogge JB (2012) Compartmentation of triacylglycerol accumulation in plants. J Biol Chem **287:** 2288-2294

Chen G, Snyder CL, Greer MS, Weselake RJ (2011) Biology and biochemistry of plant phospholipases. Crit Rev Plant Sci **30:** 239-258

Choi JK, Yu F, Wurtele ES, Nikolau BJ (1995) Molecular cloning and characterization of the cDNA coding for the biotin-containing subunit of the chloroplastic acetyl-coenzyme A carboxylase. Plant Physiol **109:** 619-625

Choudhary V, Jacquier N, Schneiter R (2011) The topology of the triacylglycerol synthesizing enzyme Lro1 indicates that neutral lipids can be produced within the luminal compartment of the

endoplasmatic reticulum: Implications for the biogenesis of lipid droplets. Commun Integr Biol **4:** 781-784

Coleman RA, Lee DP (2004) Enzymes of triacylglycerol synthesis and their regulation. Prog Lipid Res **43:** 134-176

Cyril J, Powell GL, Duncan RR, Baird WV (2002) Changes in membrane polar lipid fatty acids of seashore paspalum in response to low temperature exposure. Crop Sci **42:** 2031-2037

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Dailey HA, Strittmatter P (1979) Modification and identification of cytochrome *b*5 carboxyl groups involved in protein-protein interaction with cytochrome *b*5 reductase. J Biol Chem **254:** 5388-5396

Das UN (2006) Essential fatty acids - A review. Curr Pharm Biotechnol **7:** 467-482

Daum G, Wagner A, Czabany T, Athenstaedt K (2007) Dynamics of neutral lipid storage and mobilization in yeast. Biochimie **89:** 243-248

Demandre C, Bahl J, Serghini H, Alpha MJ, Mazliak P (1994) Phosphatidylcholine molecular species formed by lysophosphatidylcholine acyltransferase from soya bean microsomes. Phytochemistry **35:** 1171-1175

Douce R, Joyard J (1982) Plants galactolipids. In PK Stumpf, EE Conn, eds, The biochemistry of plants. Academic Press, New York, pp 331-332

Durrett TP, Benning C, Ohlrogge J (2008) Plant triacylglycerols as feedstocks for the production of biofuels. Plant J **54:** 593-607

Eastmond PJ, Quettier AL, Kroon JTM, Craddock C, Adams N, Slabas AR (2010) Phosphatidic acid phosphohydrolase 1 and 2 regulate phospholipid synthesis at the endoplasmic reticulum in Arabidopsis. Plant Cell **22:** 2796-2811

Fakas S, Qiu Y, Dixon JL, Han GS, Ruggles KV, Garbarino J, Sturley SL, Carman GM (2011) Phosphatidate phosphatase activity plays key role in protection against fatty acid-induced toxicity in yeast. J Biol Chem **286:** 29074-29085

Fan J, Yan C, Roston R, Shanklin J, Xu C (2014) Arabidopsis lipins, PDAT1 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward ß-oxidation, thereby maintaining membrane lipid homeostasis. Plant Cell **26:** 4119-4134

Fan J, Yan C, Zhang X, Xu C (2013) Dual role for phospholipid: diacylglycerol acyltransferase: Enhancing fatty acid synthesis and diverting fatty acids from membrane lipids to triacylglycerol in Arabidopsis leaves. Plant Cell **25:** 3506-3518

Farese Jr. RV, Cases S, Smith SJ (2000) Triglyceride synthesis: Insights from the cloning of diacylglycerol acyltransferase. Curr Opin Lipidol **11:** 229-234

Ghosal A, Banas A, Ståhl U, Dahlqvist A, Lindqvist Y, Stymne S (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyltransferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate specificities. Biochim Biophys Acta **1771:** 1457-1463

Gidda SK, Shockey JM, Rothstein SJ, Dyer JM, Mullen RT (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 Physiol Biochem **47:** 867-879

Graham IA (2008) Seed storage oil mobilization. Annu Rev Plant Biol **59:** 115-142

Hackett CS, Strittmatter P (1984) Covalent cross-linking of the active sites of vesicle-bound cytochrome *b*5 and NADH-cytochrome *b*5 reductase. J Biol Chem **259:** 3275-3282

Han GS, Wu WI, Carman GM (2006) The *Saccharomyces cerevisiae* Lipin homolog is a Mg^{2+}-dependent phosphatidate phosphatase enzyme. J Biol Chem **281:** 9210-9218

Harwood JL (1996) Recent advances in the biosynthesis of plant fatty acids. BBA-Lipid Lipid Met **1301:** 7-56

Harwood JL (1982) Plant acyl lipids. In PK Stumpf, EE Conn, eds, The biochemistry of plants. Academic Press, New York, pp 1-55

He X, Turner C, Chen GQ, Lin J-, McKeon TA (2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. Lipids **39:** 311-318

Hernández ML, Whitehead L, He Z, Gazda V, Gilday A, Kozhevnikova E, Vaistij FE, Larson TR, Graham IA (2012) A cytosolic acyltransferase contributes to triacylglycerol synthesis in sucrose-rescued Arabidopsis seed oil catabolism mutants. Plant Physiol **160:** 215-225

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Hofmann K (2000) A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. Trends Biochem Sci **25:** 111-112

Hu Z, Ren Z, Lu C (2012) The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic Arabidopsis. Plant Physiol **158:** 1944-1954

Huang B (2006) Plant-environment interactions. [electronic resource]. Boca Raton, FL : CRC/Taylor & Francis, 2006; 3rd ed,

Hudak KA, Thompson JE (1997) Subcellular localization of secondary lipid metabolites including fragrance volatiles in Carnation petals. Plant Physiol **114:** 705-713

Jessen D, Roth C, Wiermer M, Fulda M (2015) Two activities of long-chain acyl-coenzyme a synthetase are involved in lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis. Plant Physiol **167:** 351-366

Jhala AJ, Hall LM (2010) Flax (*Linum usitatissimum* L.): Current uses and future applications. Aust J Basic Appl Sci **4:** 4304-4312

Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell **7:** 359-371

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. J Biol Chem **278:** 8075-8082

Karantonis HC, Nomikos T, Demopoulos CA (2009) Triacylglycerol metabolism. Curr Drug Targets **10:** 302-319

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Kim HU, Li Y, Huang AHC (2005) Ubiquitous and endoplasmic reticulum-located lysophosphatidyl acyltransferase, LPAT2, is essential for female but not male gametophyte development in Arabidopsis. Plant Cell **17:** 1073-1089

King A, Nam J-, Han J, Hilliard J, Jaworski JG (2007) Cuticular wax biosynthesis in petunia petals: Cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. Planta **226:** 381-394

Kjellberg JM, Trimborn M, Andersson M, Sandelius AS (2000) Acyl-CoA dependent acylation of phospholipids in the chloroplast envelope. BBA-Mol Cell Biol L **1485:** 100-110

Knutzon DS, Lardizabal KD, Nelsen JS, Bleibaum JL, Davies HM, Metz JG (1995) Cloning of a coconut endosperm cDNA encoding a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase that accepts medium-chain-length substrates. Plant Physiol **109:** 999-1006

Kodama H, Hamada T, Horiguchi G, Nishimura M, Iba K (1994) Genetic enhancement of cold tolerance by expression of a gene for chloroplast ω-3 fatty acid desaturase in transgenic tobacco. Plant Physiol **105:** 601-605

Kopelman PG (2000) Obesity as a medical problem. Nature **404:** 635-643

Kozaki A, Kamada K, Nagano Y, Iguchi H, Sasaki Y (2000) Recombinant carboxyltransferase responsive to redox of pea plastidic acetyl-CoA carboxylase. J Biol Chem **275:** 10702-10708

Kroon JTM, Wei W, Simon WJ, Slabas AR (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry **67:** 2541-2549

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S (2013) Plant acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem **288:** 36902-36914

Lands WE (1960) Metabolism of glycerolipids. 2. The enzymatic acylation of lysolecithin. J Biol Chem **235:** 2233-2237

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Latsague M, Acevedo H, Fernández J, Romero M, Cristi R, Alberdi M (1992) Frost-resistance and lipid composition of cold-hardened needles of chilean conifers. Phytochemistry **31:** 3419-3426

Lee HY, Bahn SC, Shin JS, Hwang I, Back K, Doelling JH, Ryu SB (2005) Multiple forms of secretory phospholipase A_2 in plants. Prog Lipid Res **44:** 52-67

Li M, Wei F, Tawfall A, Tang M, Saettele A, Wang X (2014) Overexpression of patatin-related phospholipase AIIIδ altered plant growth and increased seed oil content in camelina. Plant Biotechnol J

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol **148:** 97-107

Li M, Bahn SC, Fan C, Li J, Phan T, Ortiz M, Roth MR, Welti R, Jaworski J, Wang X (2013) Patatin-related phospholipase pPLAIIId increases seed oil content with long-chain fatty acids in Arabidopsis. Plant Physiol **162:** 39-51

Li N, Gügel IL, Giavalisco P, Zeisler V, Schreiber L, Soll J, Philippar K. (2015) FAX1, a novel membrane protein mediating plastid fatty acid export. PLoS Biol **13:** e1002053

Lin YP, Carman GM (1989) Purification and characterization of phosphatidate phosphatase from *Saccharomyces cerevisiae*. J Biol Chem **264:** 8641-8645

Listenberger LL, Han X, Lewis SE, Cases S, Farese Jr. RV, Ory DS, Schaffer JE (2003) Triglyceride accumulation protects against fatty acid-induced lipotoxicity. Proc Natl Acad Sci U S A **100:** 3077-3082

Liu J, Rice A, Mcglew K, Shaw V, Park H, Clemente T, Pollard M, Ohlrogge J, Durrett TP (2015) Metabolic engineering of oilseed crops to produce high levels of novel acetyl glyceride oils with reduced viscosity, freezing point and calorific value. Plant Biotechnol J

Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ (2012) Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. Prog Lipid Res **51:** 350-377

Lu C, Xin Z, Ren Z, Miquel M, Browse J (2009) An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. Proc Natl Acad Sci U S A **106:** 18837-18842

Maisonneuve S, Bessoule J-, Lessire R, Delseny M, Roscoe TJ (2010) Expression of rapeseed microsomal lysophosphatidic acid acyltransferase isozymes enhances seed oil content in Arabidopsis. Plant Physiol **152:** 670-684

Mancha M, Stymne S (1997) Remodelling of triacylglycerols in microsomal preparations from developing castor bean (*Ricinus communis* L.) endosperm. Planta **203:** 51-57

McConn M, Browse J (1996) The critical requirement for linolenic acid is pollen development, not photosynthesis, in an Arabidopsis mutant. Plant Cell **8:** 403-416

Mène-Saffrané L, Dubugnon L, Chételat A, Stolz S, Gouhier-Darimont C, Farmer EE (2009) Nonenzymatic oxidation of trienoic fatty acids contributes to reactive oxygen species management in Arabidopsis. J Biol Chem **284:** 1702-1708

Mhaske V, Beldjilali K, Ohlrogge J, Pollard M (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (*At5g13640*). Plant Physiol Bioch **43:** 413-417

Milcamps A, Tumaney AW, Paddock T, Pan DA, Ohlrogge J, Pollard M (2005) Isolation of a gene encoding a 1,2-diacylglycerol-*sn*-acetyl-CoA acetyltransferase from developing seeds of *Euonymus alatus*. J Biol Chem **280:** 5370-5377

Millar JS, Billheimer JT (2005) Acyl CoA:diacylglycerol acyltransferases (DGATs) as therapeutic targets for cardiovascular disease. Lipids and Atherosclerosis. CRC Press, , pp 31-39

Miquel M, Browse J (1995) Lipid biosynthesis in developing seeds. In J Kigel, G Galili, eds, Seed development and germination. New York: Marcel Dekker, , pp 169-193

Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y, Nishida I (1992) Genetically engineered alteration in the chilling sensitivity of plants. Nature **356:** 710-713

Murphy DJ, Vance J (1999) Mechanisms of lipid-body formation. Trends Biochem Sci **24:** 109-115

Nagai J, Bloch K (1966) Enzymatic desaturation of stearyl acyl carrier protein. J Biol Chem **241:** 1925-1927

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 Mol Biol **21:** 267-277

Nykiforuk CL, Furukawa-Stoffer TL, Huff PW, Sarna M, Laroche A, Moloney MM, Weselake RJ (2002) Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of *Brassica napus* and sucrose-mediated induction of enzyme biosynthesis. BBA-Mol Cell Biol L **1580:** 95-109

Oelkers P, Cromley D, Padamsee M, Billheimer JT, Sturley SL (2002) The DGA1 gene determines a second triglyceride synthetic pathway in yeast. J Biol Chem **277:** 8877-8881

Ohlrogge J, Browse J (1995) Lipid biosynthesis. Plant Cell **7:** 957-970

Palta JP, Whitaker BD, Weiss LS (1993) Plasma membrane lipids associated with genetic variability in freezing tolerance and cold acclimation of Solanum species. Plant Physiol **103:** 793-803

Pan X, Peng FY, Weselake RJ (2015) Genome-wide analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes in plants reveals the eudicot-wide *PDAT* gene expansion and altered selective pressures acting on the core eudicot *PDAT* paralogs. Plant Physiol **167:** 887-904

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

Rickauer M, Brodschelm W, Bottin A, Véronési C, Grimal H, Esquerré-Tugayé MT (1997) The jasmonate pathway is involved differentially in the regulation of different defence responses in tobacco cells. Planta **202:** 155-162

Rochester CP, Bishop DG (1984) The role of lysophosphatidylcholine in lipid synthesis by developing sunflower (*Helianthus annuus* L.) seed microsomes. Arch Biochem Biophys **232:** 249-258

Routaboul JM, Benning C, Bechtold N, Caboche M, Lepiniec L (1999) The TAG11 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. Plant Physiol Biochem **37:** 831-840

Routaboul JM, Fischer SF, Browse J (2000) Trienoic fatty acids are required to maintain chloroplast function at low temperatures. Plant Physiol **124:** 1697-1705

Routaboul JM, Skidmore C, Wallis JG, Browse J (2012) Arabidopsis mutants reveal that short-and long-term thermotolerance have different requirements for trienoic fatty acids. J Exp Bot **63:** 1435-1443

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 Physiol **141:** 1533-1543

Samala S, Yan J, Baird WV (1998) Changes in polar lipid fatty acid composition during cold acclimation in 'Midiron' and 'U3' bermudagrass. Crop Sci **38:** 188-195

Sandager L, Gustavsson MH, Ståhl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H, Stymne S (2002) Storage lipid synthesis is non-essential in yeast. J Biol Chem **277:** 6478-6482

Sasaki Y, Hakamada K, Suama Y, Nagano Y, Furusawa I, Matsuno R (1993) Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant. J Biol Chem **268:** 25118-25123

Scherer GFE, Ryu SB, Wang X, Matos AR, Heitz T (2010) Patatin-related phospholipase A: Nomenclature, subfamilies and functions in plants. Trends Plant Sci **15:** 693-700

Schmidt H, Heinz E (1990) Involvement of ferredoxin in desaturation of lipid-bound oleate in chloroplasts. Plant Physiol **94:** 214-220

Schnurr JA, Shockey JM, De Boer G-, Browse JA (2002) Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. Plant Physiol **129:** 1700-1709

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature **432:** 779-782

Shanklin J, Cahoon EB (1998) Desaturation and related modifications of fatty acids. Annu Rev Plant Physiol Plant Mol Biol **49:** 611-641

Sharma KK, Schuhmann H, Schenk PM (2012) High lipid induction in microalgae for biodiesel production. Energies **5:** 1532-1553

Shockey JM, Fulda MS, Browse JA (2002) Arabidopsis contains nine long-chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism. Plant Physiol **129:** 1710-1722

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. Plant Cell **18:** 2294-2313

Shorrosh BS, Roesler KR, Shintani D, Van De Loo FJ, Ohlrogge JB (1995) Structural analysis, plastid localization, and expression of the biotin carboxylase subunit of acetyl-coenzyme A carboxylase from tobacco. Plant Physiol **108:** 805-812

Sinclair AJ, Attar-Bashi NM, Lib D (2003) What is the role of alpha-linolenic acid for mammals? Lipids **37:** 1113-1123

Slack CR, Campbell LC, Browse JA, Roughan PG (1983) Some evidence for the reversibility of the cholinephosphotransferasecatalysed reaction in developing linseed cotyledons *in vivo*. BBA-Lipid Lipid Met **754:** 10-20

Slack CR, Roughan PG, Browse JA, Gardiner SE (1985) Some properties of cholinephosphotransferase from developing safflower cotyledons. BBA-Lipid Lipid Met **833:** 438-448

Spatz L, Strittmatter P (1971) A form of cytochrome *b*5 that contains an additional hydrophobic sequence of 40 amino acid residues. Proc Natl Acad Sci U S A **68:** 1042-1046

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Ståhl U, Stålberg K, Stymne S, Ronne H (2008) A family of eukaryotic lysophospholipid acyltransferases with broad specificity. FEBS Lett **582:** 305-309

Stobart K, Mancha M, Lenman M, Dahlqvist A, Stymne S (1997) Triacylglycerols are synthesised and utilized by transacylation reactions in microsomal preparations of developing safflower (*Carthamus tinctorius* L.) seeds. Planta **203:** 58-66

Stymne S, Glad G (1981) Acyl exchange between oleoyl-CoA and phosphatidylcholine in microsomes of developing soya bean cotyledons and its role in fatty acid desaturation. Lipids **16:** 298-305

Stymne S, Stobart AK (1987) Triacylglycerol biosynthesis. In PK Stumpf, ed, The biochemistry of plants. Academic Press, New York, pp 175-214

Stymne S, Stobart AK (1984) Evidence for the reversibility of the acyl-CoA:lysophosphatidylcholine acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons and rat liver. Biochem J **223:** 305-314

Sun L, Ouyang C, Kou S, Wang S, Yao Y, Peng T, Xu Y, Tang L, Chen F (2011) Cloning and characterization of a cDNA encoding type 1 diacylglycerol acyltransferase from sunflower (*Helianthus annuus* L.). Z Naturforsch C **66(1-2):** 63-72

Tjellström H, Strawsine M, Ohlrogge JB (2015) Tracking synthesis and turnover of triacylglycerol in leaves. J Exp Bot

Tjellström H, Yang Z, Allen DK, Ohlrogge JB (2012) Rapid kinetic labeling of Arabidopsis cell suspension cultures: Implications for models of lipid export from plastids. Plant Physiol **158:** 601-611

Tumaney AW, Shekar S, Rajasekharan R (2001) Identification, purification, and characterization of monoacylglycerol acyltransferase from developing peanut cotyledons. J Biol Chem **276:** 10847-10852

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

van Erp H, Kelly AA, Menard G, Eastmond PJ (2014) Multigene engineering of triacylglycerol metabolism boosts seed oil content in Arabidopsis. Plant Physiol **165:** 30-36

Wallis JG, Browse J (2002) Mutants of Arabidopsis reveal many roles for membrane lipids. Prog Lipid Res **41:** 254-278

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:** 230-242

Wang HW, Zhang JS, Gai JY, Chen SY (2006) Cloning and comparative analysis of the gene encoding diacylglycerol acyltransferase from wild type and cultivated soybean. Theor Appl Genet **112:** 1086-1097

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE **2(8):** e718

Wolter FP, Schmidt R, Heinz E (1992) Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. EMBO J **11:** 4685-4692

Xu J, Carlsson AS, Francis T, Zhang M, Hoffman T, Giblin ME, Taylor DC (2012) Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. BMC Plant Biol **12:** 4

Xu J, Francis T, Mietkiewska E, Giblin EM, Barton DL, Zhang Y, Zhang M, Taylor DC (2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. Plant Biotech J **6:** 799-818

Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB (2012) A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, *sn*-2 preference, and evolution. Plant Physiol **160:** 638-652

Yoon K, Han D, Li Y, Sommerfeld M, Hu Q (2012) Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *chlamydomonas reinhardtii*. Plant Cell **24:** 3708-3724

Zhang H, Damude HG, Yadav NS (2012) Three diacylglycerol acyltransferases contribute to oil biosynthesis and normal growth in *Yarrowia lipolytica*. Yeast **29:** 25-38

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Zhou XR, Shrestha P, Yin F, Petrie JR, Singh SP (2013) AtDGAT2 is a functional acyl-CoA:diacylglycerol acyltransferase and displays different acyl-CoA substrate preferences than AtDGAT1. FEBS Lett **587:** 2371-2376

Zou J, Wei Y, Jako C, Kumar A, Selvaraj G, Taylor DC (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. Plant J **19:** 645-653

Chapter 3

Identification of a Pair of Phospholipid:Diacylglycerol Acyltransferases from Developing Flax (*Linum usitatissimum* L.) Seed Catalyzing the Selective Production of Trilinolenin

3.1 Introduction

Flax (*Linum usitatissimum* L.), which produces oil containing approximately 45% to 65% α-linolenic acid (ALA; $18:3^{cis\Delta 9,12,15}$), is considered one of the most abundant plant sources of ALA. With a low oxidative stability, ALA can react rapidly with oxygen and polymerize into a soft and durable film upon air exposure, which makes flax oil suitable for domestic and industrial coatings such as linoleum, varnishes and paints (Jhala and Hall, 2010). In addition, ALA is an essential fatty acid for human and a precursor for nutritionally beneficial very long chain omega-3 polyunsaturated fatty acids (VLC-ω-3-PUFA, ≥ 20 carbons) (Sinclair, et al., 2003; Das, 2006).

Efforts to characterize the molecular basis of high ALA content in flax have focused largely on the characterization of fatty acid desaturases (FAD). Two genes encoding Δ15 desaturases (*LuFAD3A* and *LuFAD3B*) have been identified in flax (Vrinten, et al., 2005). Ethyl methanesulfonate-generated point mutations in *LuFAD3A* and *LuFAD3B* led to a reduction in ALA content to approximately 1-2%, suggesting that these two genes encoded the main desaturases responsible for the synthesis of ALA in flax (Green, 1986; Rowland, 1991; Vrinten, et al., 2005). Generally, there is a good correlation between the expression of *FAD3* genes with the accumulation of ALA in flax and other plant species (Rao, et al., 2008). However, along with desaturases, many other enzymes can contribute to the flux of ALA to storage lipids. Essentially, ALA has to be efficiently transferred from the desaturation product (*sn*-2-linolenoyl-phosphatidylcholine, PC) to the substrates for triacylglycerol (TAG) synthesis and this process

can take different biochemical routes. ALA synthesized on PC can enter the acyl-CoA pool by either the reverse action of acyl-CoA: lysophosphatidylcholine acyltransferase (LPCAT) (Stymne and Stobart, 1984) or the combined action of phospholipase A_2 (PLA_2) and long-chain acyl-CoA synthetase. 2-Linolenoyl-*sn*-PC can also be potentially converted into 2-linolenoyl-*sn*-diacylglycerol (DAG) by the catalytic action of phospholipase C or phospholipase D (Lee, et al., 2011) together with phosphatidic acid phosphatase. In addition, the enzyme phosphatidylcholine:diacylglycerol cholinephosphotransferase (PDCT) catalyzes the transfer of a phosphocholine head group from PC to DAG in a symmetrical reaction (Lu, et al., 2009) which could potentially produce DAG enriched in ALA for the synthesis of TAG.

TAG can be formed via an acyl-CoA-dependent or acyl-CoA-independent process. The final step of the acyl-CoA dependent pathway, also known as Kennedy pathway (Weiss, et al., 1960), is catalyzed by acyl-CoA:diacylglycerol acyltransferase (DGAT) which uses acyl-CoA as acyl donor to convert DAG to TAG. At least four different types of DGAT have been identified in plants (Hobbs, et al., 1999; Saha, et al., 2006; Kalscheuer and Steinbüchel, 2003; Lardizabal, et al., 2001). DGAT1 has been proposed to exert dominant control over TAG accumulation in many oilseeds (Lung and Weselake, 2006; Zhang, et al., 2009; Katavic, et al., 1995; Zheng, et al., 2008). DGAT2 was first identified in the oleaginous fungus *Umbelopsis* (formerly *Mortierella*) *ramanniana* and shows no sequence homology to DGAT1 (Lardizabal, et al., 2001). DGAT1 and DGAT2 appear to localize to different subdomains of endoplasmic reticulum (ER) and have been suggested to have nonredundant functions in the production of TAG (Shockey, et al., 2006). The third type of DGAT (DGAT3) is a soluble enzyme isolated from developing peanut cotyledons and it differs from DGAT1 and DGAT2, which are membrane-bound (Saha,

et al., 2006). Another soluble enzyme with DGAT activity known as defective in cuticular ridge (DCR) has been identified in *Arabidopsis thaliana* (Rani, et al., 2010).

The acyl-CoA independent pathway of TAG synthesis is characterized by the enzymatic action of phospholipid:diacylglycerol acyltransferase (PDAT) that transfers the fatty acyl moiety from *sn*-2-position of a phospholipid to the *sn*-3-position of *sn*-1, 2-DAG (Dahlqvist, et al., 2000). A gene encoding a PDAT was first reported in yeast (Dahlqvist, et al., 2000). Two homologs of yeast *PDAT* have been identified in Arabidopsis named *PDAT1* (At5g13640) and *PDAT2* (At3g4480) (Ståhl, et al., 2004) and three putative *PDAT* genes have been identified in the castor genome (van Erp, et al., 2011).

The existence of specialized acyltransferases has been observed in other plant species. DGAT2 from tung tree (*Vernicia fordii*) and castor (*Ricinus communis*) display substrate preference for unusual fatty acids (FA) and are predominantly involved in the incorporation of these FAs into seed oils (Shockey, et al., 2006; Kroon, et al., 2006). Also, a ricinoleate-specific PDAT from castor (RcPDAT) has been reported (van Erp, et al., 2011). Co-expression of *RcPDAT1* with castor FA hydroxylase in Arabidopsis resulted in a relative increase of 58% hydroxy FAs in seeds. Considering the natural occurrence of evolved forms of acyltransferases that are selective for unusual FAs, it was hypothesized that similar mechanisms take place in flax. In spite of the wide range of applications for flax oil, many components involved in TAG biosynthesis have not been characterized at the molecular genetic level. Here, the data obtained from this study demonstrate that two pairs of *PDAT* genes encoding enzymes utilize preferentially substrates containing ALA, and more importantly, one of the pairs (*LuPDAT1* and *LuPDAT5*) has embryo-preferred expression pattern and appears to contribute mainly to the synthesis of trilinolenin in flax seeds.

3.2 Materials and Methods

3.2.1 Plant material and growth conditions

Flax (*Linum usitatissimum* L.) was grown in the growth chamber with 16 h light at 21^{o}C and 8 h dark at 18^{o}C. Seeds were planted in 1 Gallon pots containing Metro Mix soil (Greenleaf Products Inc., CA). Plants were fertilized with 17-5-19 (125g/100L) and 12-2-14 (83g/100L) once a week before flowering. From the flowering stage, plants were fertilized weekly with 17-5-19 (280g/100L) till harvest. Individual flowers were tagged at anthesis, and embryos at different stages of development were collected in liquid N_2 and stored at -80^{o}C. Arabidopsis wild-type Columbia and mutant line AS11 were obtained from the Salk Institute via the ABRC (Ohio State University, Columbus, OH). Arabidopsis seeds in plots or plates were cold-treated at 4^{o}C in the dark for 3 days and then placed into a controlled growth chamber with a constant temperature of 20^{o}C and 16 h photoperiod.

3.2.2 Identification and isolation of candidate genes

In general, all primers are summarized in Table 3.1. Cloning integrity was confirmed by sequencing at each step.

A BLAST analysis (Altschul, et al., 1990) was conducted against the flax genomic database (Wang, et al., 2012) by using AtDGAT1, AtPDAT1, AtDCR and tung tree VfDGAT2 as the protein query. The theoretical molecular weight and isoelectric point values were calculated using the Compute pI/Mw tool provided in the ExPASy server. To isolate the target genes, total RNA was extracted from the embryo of flax (cultivar AC Emerson) 12 days post anthesis (DPA) using Plant RNeasy plant mini kit (Qiagen, CA) as described by the manufacturer. The First stand of the cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen, CA) according to the protocol provided by the supplier. The target

genes were amplified from the resulting cDNA as the template for 30 cycles of PCR amplification using Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen, CA). PCR was performed under the following temperature cycle program: 95°C for 2 min; 30 cycles of denaturation (95°C, 20s), annealing (55°C, 15s), and extension (68°C, 2.5 min); and a final extension at 68°C for 2 min. To amplify *LuDGAT2-1*, *LuDGAT2-3*, *LuPDAT2*, and *LuPDAT6*, the forward primers contained a specific restriction site (underlined) and a Kozak translation initiation sequence (italic) in order to improve the translation of the protein (Table 3.1). A specific restriction site was also introduced in the reverse primer (underlined). The PCR products were cloned into the pYES vector collinear to the *GAL1* promoter inducible by galactose. The pYES is a modified pYES2.1/V5-HIS vector (Invitrogen, CA) constructed in our lab, which contains more restriction sites in its multiple cloning site (MCS). For *LuDCR1* and *LuPDAT1* that could not be amplified by these specially designed primers, the internal primers were used for amplification. For *LuDGAT2-2* and *LuPDAT5*, an internal forward primer and a reverse primer spanning the 3'-untranslated region (3'UTR) were used for amplification. The PCR products were subcloned into pYES2.1/V5-HIS vector using pYES2.1 TOPO kit (Invitrogen, CA) as demonstrated by the supplier. To construct the co-expression vectors, the *LuFAD2-1* (Krasowska, et al., 2007) and *LuFAD3B* (Vrinten, et al., 2005) genes were first amplified using PCR with appropriate primers that allowed to add specific restriction sites (underlined, Table 3.1) to the ends of amplified products and then inserted into MCS1 and MSC2 of the pESC-URA expression vector (Agilent Technologies, CA), yielding *LuFAD2-1-FAD3B*/pESC plasmid. The *ADH1* terminator: *LuFAD2-1*: Pro*GAL10*: Pro*GAL1*: *LuFAD3B*: *CYC1* terminator expression cassette of *LuFAD2-1-FAD3B*/pESC was then excised and subcloned into the recombinant pYES plasmids containing *LuPDAT1*, *LuPDAT2* or *LuDGAT1-1* through one-step, isothermal assembly

method described by Gibson (2011). The resulting plasmids were referred to as *LuFAD2-1-FAD3B-PDAT1*/pYES, *LuFAD2-1-FAD3B-PDAT2*/pYES and *LuFAD2-1-FAD3B-DGAT1-1*/pYES.

3.2.3 Real-time PCR quantification

Total RNA was isolated from vegetative tissues (stems, leaves, apexes, roots), reproductive tissues (flowers), developing embryos (8, 11, 14, 16, and 20 DPA) and seeds (4, 25 and 40 DPA) of flax cultivar CDC Bethune using RNeasy plant mini kits (Qiagen, CA). Due to technical difficulties in separating embryos from other seed components at 4, 25 and 40 DPA, the whole seeds at these three stages were used for gene expression analysis. The cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen, CA). Gene specific primers and probes were designed using Primer Express software (Applied Biosystems) and the sequences are listed in Table 3.1. 5'FAM/3'quencher-labelled probes were used to assay the genes involved in TAG synthesis, and probes for reference genes were labelled with 5'VIC/3'quencher. TaqMan-based qRT-PCR assays were performed using Fast Advance Master mix (Invitrogen, CA) on the ABI PRISM 7900HT Real-Time PCR system (Applied Biosystems). PCR efficiency for each amplicon was calculated using a dilution series of a single cDNA sample over several log concentrations. According to the recommendation of Huis et al. (Huis, et al., 2010), two internal reference genes, encoding glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and ubiquitin extension protein (UBI2), were used. The relative expression level was calculated using the comparative C*t* method after normalizing to controls using the reference genes.

Table 3.1 Primers and probes employed in the current study

Primer name	Oligonucleotide sequence
Primers used for pYES or pYES2.1/V5-HIS cloning of cDNAs (Yeast expression)	
LuDCR1-F	5'- ATGTCCTCCGCCGCCGTCAAAATCACT-3'
LuDCR1-R	5'- TTAATGAGAACGTTGAGCCATTTC-3'
LuDGAT2-1-F	5'-ATATGGATCC*TACACA*ATGTCACAGAAAGTAGAGGAGGAAGACCGG-3'
LuDGAT2-1-R	5'-ATATAAGCTTTCAAAGAATTTTGAGTTGAAGATCAGCAC-3'
LuDGAT2-2-F	5'-ATGGTTGAGAAAGCAGAGGAGGAGG-3'
LuDGAT2-2-R	5'-TTACGAAACACGTACAAATTTGTGC-3'
LuDGAT2-3-F	5'-ATATGGATCC*TACACA*ATGTCGGTACAGAAAGTAGAGGAGGAAAAC-3'
LuDGAT2-3-R	5'-ATATAAGCTTTTAAAGAATTTTGAGTTGAAGATCAGCAC-3'
LuPDAT1-F	5'-ATGTCTCTCTTGAGGCGGAGGTGG-3'
LuPDAT1-R	5'-TTAAAGCCGCAACTTAATCTTGTC-3'
LuPDAT2-F	5'-ATATGGTACC*TACACA*ATGTCAGTAGTCCGCCGCCGAAAACCTTC-3'
LuPDAT2-R	5'-ATATAAGCTTCTAGAGATTCAGTTTGATCCTATCAGACC-3'
LuPDAT5-F	5'-ATGTCTCTATTGAGGCGGAGGTGGAG-3'
LuPDAT5-R	5'-CATCATCATCATACAAACTTACCTAAAGC-3'
LuPDAT6-F	5'-ATATGGTACC*TACACA*ATGTCGCCTGGAATCGTCACCGGTGGTC-3'
LuPDAT6-R	5'-ATATCTCGAGTCACAGTTGAATGTTAATACGGTCCGAC-3'
LuFAD2-F	5'- ATAGGATCCACCATGGGTGCTGGTGGAAGAAT -3'
LuFAD2-R	5'- TATGGTACCTCACAGCTTGTTGTTGTACCA -3'
LuFAD3-F	5'- CCGGAATTCTACACAATGTCAATGAGCCCTCCAAACTCAATG -3'
LuFAD3-R	5'- TATGAGCTCTCAGCTGGATTTGGACTTGG -3'
Primers used for cloning LuFAD2-FAD3 cassette from pESC into recombinant pYES vector (Yeast expression)	
LuFAD2-FAD3-F	5'-GAGAGGCGGTTTGCGTATTGGGCGCGCTGAATTGGAGCGACCTCATGC-3'
LuFAD2-FAD3-R	5'-GTCAGTGAGCGAGGAAGCGGAAGACTGGATCTTCGAGCGTCCCAAAACC-3'
pYES-F	5'-GGTTTTGGGACGCTCGAAGATCCAGTCTTCCGCTTCCTCGCTCACTGAC-3'
pYES-R	5'-GCATGAGGTCGCTCCAATTCAGCGCGCCCAATACGCAAACCGCCTCTC-3'
Primers used for generating N-terminal-tagged construct for Western blot	
LuPDAT1-F	5' -TATAGGATCC*TACACA*ATGTCTCTCTTGAGGCGGAGGTG-3'
LuPDAT1-R	5' -TATACTCGAGTTAAAGCCGCAACTTAATCTTGTCAG-3'

Table 3.1 Continued

LuPDAT2-F	5’ - ATATGGTACC*ATACACA*ATGTCAGTAGTCCGCCGCCGAAAAC
LuPDAT2-R	5’ - TATAGCGGCCGCCTAGAGATTCAGTTTGATCCTATCAGACC
LuPDAT6-F	5’-TATAGGTACC*ATACACA*ATGTCGCCTGGAATCGTCACC-3’
LuPDAT6-R	5’-TATAGCGGCCGCTCACAGTTGAATGTTAATACG-3’
Primers used for pGreen cloning of cDNAs (Arabidopsis expression)	
LuDCR1-F	5’-ATATCTCGAG*AACACA*ATGTCCTCCGCCGCCGTCAAAATC-3’
LuDCR1-R	5’-TATATCTAGATTAATGAGAACGTTGAGCCATTTC-3’
LuDGAT1-F	5’-ATATCTCGAG*AACACA*ATGTCCGTGCTTGATACCCCTGAC-3’
LuDGAT1-R	5’-TATAGGATCCTTAGATACCATCTTTCCCATTCCTGTTCATC-3’
LuDGAT2-3-F	5’-TATACTCGAG*TACACA*ATGTCGGTACAGAAAGTAGAGG-3’
LuDGAT2-3-R	5’-TATAGGATCCTTAAAGAATTTTGAGTTGAAGATC-3’
LuPDAT1-F	5’-TATAAAGCTT*TACACA*ATGTCACTCTTGAGGCGGAGGTGG-3’
LuPDAT1-R	5’-TATAGGATCCTTAAAGCCGCAACTTAATCTTGTCAG-3’
LuPDAT2-F	5’-ATATCTCGAG*TACACA*ATGTCGGTAGTCCGCCGCCGAAAACC-3’
LuPDAT2-R	5’-TATATCTAGACTAGAGATTCAGCTTGATCCTATCAGACC-3’
LuPDAT6-F	5’-TATACCCGGG*TAAACA*ATGTCGCCTGGAATCGTCACC-3’
LuPDAT6-R	5’-TATACCCGGGTCACAGTTGAATGTTAATACGGTCCG-3’
Primers used for real-time PCR analysis	
LuDGAT1-1-F	5’-TCAAGGTTAAAGAGAGTCCGCTTAG-3’
LuDGAT1-1-R	5’-CCTGCATGGCTCTGCTTA-3’
LuDCR1-F	5’-GCATGCATCGCCCTTT-3’
LuDCR1-R	5’-CCCATTGCAAGTCCGTCTTC-3’
LuDGAT2-1-F	5’-CGCCTTCCGTTCTGATCGT-3’
LuDGAT2-1-R	5’-CCAATGGGAAGAACAGAATGC-3’
LuDGAT2-2-F	5’-CCTTCGGTCAGGTGGGTTT-3’
LuDGAT2-2-R	5’-GGTTTCCACCCCTTGT-3’
LuDGAT2-3-F	5’-ACTCGGAACTCCATTGCCATT-3’
LuDGAT2-3-R	5’-TGGGTCGGCCAATCACA-3’
LuPDAT1-F	5’-CAAAGCGGTGCTCATGTTGA-3’
LuPDAT1-R	5’-GCTGCAACTCTGATGATATCCTCTAT-3’

Table 3.1 Continued

LuPDAT2-F	5’-TGCTACATCCCGTTCCAGATC-3’
LuPDAT2-R	5’-TGTACACTCCGTCTTTCAGACAA-3’
LuPDAT3-F	5’-GATTTGTTAAAGCGCAAGTCG-3’
LuPDAT3-R	5’-CATTAGTCGAGTTTTGGTCAGCAA-3’
LuPDAT4-F	5’-TCTGAAAGATGGAGTGTACACAGTT-3’
LuPDAT4-R	5’-ACACATGTATCCTGCGCTCAAC-3’
LuPDAT5-F	5’-TGATGAACATAGGTGGACCTTTTTTA-3’
LuPDAT5-R	5’-ATATTGCTGTTGCCAGGGCT-3’
LuPDAT6-F	5’-GGTCTCCTGAAGATAACAACTATGCTT-3’
LuPDAT6-R	5’-GGTCATCCGGAGATTGCAAG-3’
GAPDH-F	5’-GACATCGTCTCCAACGCTAGCT-3’
GAPDH-R	5’-CATTGATAACCTTGGCCAAAG-3’
Ubi-F	5’-CGGACACCATCGACAATGT-3’
Ubi-R	5’-TCCGGCGGGATTCCTT-3’
Probes used for real-time PCR analysis	
LuDGAT1-1	5’-TCCGGCGCCATTT-3’
LuDCR	5’-TGGCTGTTCAGCTGACTA-3’
LuDGAT2-1	5’-CTTACGTTTTCGGGTACGAG-3’
LuDGAT2-2	5’-CTTGTTCCGTCAGATGTG-3’
LuDGAT2-3	5’-CAGCAACCGATGCAT-3’
LuPDAT1	5’-TAATGGGCAATTTTG-3’
LuPDAT2	5’-ATACATCAGCAGATGAAG-3’
LuPDAT3	5’-CAATCTCCGGTGGCAGT-3’
LuPDAT4	5’-ACGAAACGGTGCCC-3’
LuPDAT5	5’-CCAAAGTAGTCTCTGGACTT-3’
LuPDAT6	5’-ATGTGGTGATTTGTCGAAG-3’
GAPDH	5’-ACCACTAACTGCCTTGC-3’
Ubi2	5’-AAGGCCAAGATCCAG-3’

3.2.4 Heterologous expression in yeast mutant H1246

In general, the constructed recombinant plasmids were transformed into the quadruple mutant strain *S. cerevisiae* H1246 by using the LiAc/SS carrier DNA/PEG method (Gietz and Schiestl, 2007) and transformants were selected on minimal medium plates lacking uracil. The recombinant yeast cells were first grown in liquid minimal medium with 2% [w/v] raffinose overnight and then inoculated at a starting OD of 0.1 in minimum medium containing 2% [w/v] galactose and 1% [w/v] raffinose (referred to as induction medium) to induce gene expression.

For heterologous expression of single flax genes, yeast cells were grown at 30^{o}C and yeast transformed with pYES*LacZ* was used as a control. For the feeding experiments, free FAs, including oleic (OA; 18:1$^{cis\Delta 9}$), linoleic acid (LA; 18:2$^{cis\Delta 9,12}$), ALA, stearidonic acid (SDA; 18:4$^{cis\Delta 6,9,12,15}$), dihomo-γ-linolenic acid (DGLA; 20:3$^{cis\Delta 8,11,14}$), γ-linolenic acid (GLA; 18:3$^{cis\Delta 6,9,12}$), arachidonic acid (AA; 20:4$^{cis\Delta 5,8,11,14}$), eicosatrienoic acid (ETA; 20:3$^{cis\Delta 11,14,17}$), eicosapentaenoic acid (EPA; 20:5$^{cis\Delta 5,8,11,14,17}$) and docosahexaenoic acid (DHA; 22:6$^{cis\Delta 4,7,10,13,16,19}$), were dissolved at 0.5 M in ethanol. The FA solutions were first dissolved in 0.06% [v/v] tyloxapol (Sigma, CA) and then mixed with induction medium. Tyloxapol is a non-ionic surfactant which was used to disperse the FAs into the medium. Yeast was induced in induction medium with 100 μM FA supplementation.

To evaluate the concentration effect of exogenously provided ALA on total TAG content and composition in yeast expressing *LuPDAT1* and *LuPDAT2*, the recombinant yeast cells were cultivated in induction medium supplemented with different concentration of ALA (0 to 300 μM) and harvested at the same growth stage (OD_{600nm}= 6.5±0.05). Approximately 30 mg of dry yeast were weighted out and used for lipid analysis after 16 h freeze-drying. One hundred micrograms

of the TAG internal standard, triheptadecanoin (C17:0 TAG), were added to each sample before lipid extraction.

For growth curve construction, yeast cells were grown first in minimum medium containing 2% [w/v] raffinose, and then in induction medium supplemented with three different ALA concentrations (0, 100 μM and 1 mM). Aliquots (100 μl) were withdrawn from the culture at intervals of 6 or 12 h and diluted 1:10 with sterile water. One hundred microliters of the diluted cells were transferred to a 96-well clear, flat bottom, polystyrene microplate (UNIPLATE, Whatman). The yeast growth (OD_{600nm}) was measured using the Synergy H4 Hydrid multi-mode microplate reader (Biotek Instument, Inc., USA).

Growth phenotype were further analyzed by first culturing recombinant yeast cells in minimum medium containing 2% [w/v] raffinose, and then spotting 2.5μl of cell suspension serially diluted to OD_{600nm} values of 1, 10^{-1}, 10^{-2},10^{-3},and 10^{-4}, onto agar plates containing induction medium supplemented with different concentration of ALA (0, 100 μM and 1 mM). ALA was first dissolved at 0.5 M in ethanol, and then diluted in a warm medium containing 0.01% [v/v] tyloxapol immediately before plating. The same amount of ethanol and typloxapol was added to the plates without FA supplementation. Yeast strains BY4742 (wild-type) and H1246 transformed with empty pYES vector were used as positive and negative controls, respectively. Plates were incubated at 30 ^{o}C for 2 days.

For the co-expression study, yeast cells transformed with *LuFAD2-1-FAD3B-PDAT1*/pYES, *LuFAD2-1-FAD3B-PDAT2*/pYES or *LuFAD2-1-FAD3B-DGAT1-1*/pYES plasmid were inoculated in the same induction medium for 3 days at 20^{o}C before cell harvest. Yeast cells transformed with *LuFAD2-1-FAD3B/*pYES were used as a control.

3.2.5 Western blot analysis

To produce N-terminal-tagged constructs, the complete ORFs of three *LuPDATs* (*LuPDAT1*, *LuPDAT2* and *LuPDAT6*) were first amplified by the primers listed in Table 3.1 and then subcloned into the pYES2/NT-C expression vector (Invitrogen). For the generation of C-terminal-tagged constructs, the stop codons of the *LuPDATs* in the recombinant pYES plasmids were mutated by using the Quickchange mutagenesis kit (Stratagene). Yeast microsomes were prepared as previously described (Siloto, et al., 2009a). Briefly, the recombinant yeast cells were first cultivated in liquid minimal medium with 2% [w/v] raffinose and then inoculated at a starting OD of 0.4 for 12h. The cells were harvested, washed, and then resuspended in the homogenization buffer (20mM Tris-HCl pH 7.9, 10mM $MgCl_2$, 1mM EDTA, 5% (v/v) glycerol, 300mM ammonium sulphate, 2mM dithiotreitol). In the same buffer, the cells were broken mechanically with 1/3 volume of 0.5 mm glass beads by a bead beater (Biospec Bartlesvile, USA). To remove the cell debris and glass beads, the homogenate was centrifuged at 10,000 g at 4^{o}C for 20 min. The recovery supernatant was further centrifuged at 48,000 g at 4^{o}C for 1 h. The resulting microsomal pellets were dissolved in 3mM imidazole buffer (pH 7.4) containing 125mM sucrose. Equivalent amount of microsomal proteins (15μg, as determined by Bio-Rad Bradford assay) were separated by 4-12% gradient Tris-glycine gels and then blotted onto nitrocellulose membranes. The target N-terminal-tagged proteins were detected using anti-HisG-HRP antibody (Invitrogen). The bounded antibodies were detected using SuperSignal Fermo Chemiluminescent Substrate (Pierce). Equal protein loading was ensured by determining the level of constitutively expressed chaperone Kar2p using a rabbit polyclonal anti-Kar2p (Santa Cruz Biotechnology) as the primary antibody followed by incubation with HRP-goat anti-rabbit lgG(H+L) secondary antibody (Invitrogen).

3.2.6 Nile red fluorescence detection of neutral lipids in yeast

The Nile red fluorescence assay was conducted as described (Siloto, et al., 2009a; Siloto, et al., 2009b). Briefly, the recombinant yeast cells were cultivated in induction medium with or without ALA supplementation and harvested at the stationary growth stage. Cells with ALA feeding were washed three times with deionized water before Nile red measurement. One hundred microliters aliquots of induced yeast cell culture were transferred to a 96-well dark flat-bottom plate (UNIPLATE, whatman). The background fluorescence was measured using the Synergy H4 Hydrid multi-mode microplate reader (Biotek Instument, Inc., USA) with the emission and excitation filter setting at 485 and 538 nm, respectively. Five microliters of freshly prepared methanolic Nile red solution (0.1 mg/ml) was added first and mixed with yeast culture. The second fluorescence intensity value was measured under the same conditions. The final Nile red value was calculated by subtracting the first measurement from the second measurement (ΔF), and then dividing by the corresponding OD_{600nm} value.

3.2.7 Yeast lipid extraction and analysis

Yeast lipid extraction was performed using the method described previously (Siloto, et al., 2009a; Bligh and Dyer, 1959). Briefly, induced yeast cultures at the stationary growth stage were harvested by centrifuge, washed and resuspended in 1 ml of 0.9% [w/v] sodium chloride (NaCl). Cells were homogenized with an equal volume of glass beads (0.5 mm) by vigorously vortexing for 2 min. Lipids were extracted by the chloroform, methanol, 0.9% aqueous NaCl (2:1:0.9, v/v/v) method. The chloroform phase (lower phase) was collected, dried under nitrogen and resuspended in 30 µl of chloroform. The extracted lipids were resolved on the thin layer chromatography (TLC) plates (SIL G25, 0.25 mm, Macherey–Nagel, Germany) with the solvent system hexane/diethyl ether/acetic acid (80:20:1). The developed plate was visualized with 3%

cupric acetate [w/v] and 8% phosphoric acid [v/v] followed by charring at 280^{o}C for 20 min. The TAG bands were identified according to triolein and trilinolenin standards.

FA analysis was performed by gas chromatography (GC)/mass spectrometry (MS) as described previously (Mietkiewska, et al., 2011). Briefly, the extracted lipids were developed on TLC plate by using the same solvent system but visualized under UV after spraying with 0.05% primuline solution. The bands corresponding to TAG were scraped and transmethylated with 5% [w/v] sodium methoxide (NaOMe) at room temperature for 30 min. The resulting FA methyl esters (FAMEs) were extracted with two portions (2 ml each) of hexane. The hexane phases were pooled together and dried under a stream of nitrogen and immediately resuspended in 250 µl of iso-octane with 0.1 mg/ml C21:0 methyl ester standard. FAMEs were subjected to an Agilent 6890N GC equipped with DB-23 capillary column (30 m x 0.25 mm x 0.25 µm) and a 5975 inert XL Mass Selective Detector. The following temperature program was applied: 165^{o}C hold for 4 min, 10^{o}C/min to 180^{o}C, hold 5 min and 10^{o}C/min to 230^{o}C hold 5 min.

3.2.8 Overexpression of *LuDGATs* and *LuPDATs* in Arabidopsis

Agrobacterium tumefaciens strain GV3101 and pGreen/pSoup based dual binary vectors (Hellens, et al., 2000) were used for Arabidopsis transformation. The coding regions of target genes were amplified using recombinant pYES plasmids as template with specially designed primers for each gene (see Table 3.1). The amplicons were cloned into the pGreen vector under the control of the seed specific napin promoter. The resulting construct and the helper plasmid pSoup were co-transformed into *A. tumefaciens* GV3101 by electroporation (Weigel and Glazebrook, 2002). *Agrobacterium* strains containing the pGreen/pSoup dual binary vectors were used to transform the Arabidopsis wild-type (Columbia) and mutant line AS11 by the floral dipping method (Weigel and Glazebrook, 2002). Plants transformed with an empty vector

pGreen were used as controls. T_1 seeds of transgenic plants were selected on half-strength Murashige and Skoog (MS) agar plates supplemented with 80μM herbicide phosphinothricin. Transformants were then transferred to soil and grown to maturity to produce T_2 seeds. The presence of the target genes was confirmed by gene-specific PCR analysis using DNA extracted from T_2 young leaf tissue as template. T_2 seeds were collected and used for total lipid and FA analysis.

3.2.9 Analysis of Arabidopsis seed oil

Total lipid content and the FA composition of T_2 seeds were determined by GC. Approximately 10 mg seeds per replicate were weighted out after stabilizing seed moisture content in desiccators for 72h. One hundred micrograms triheptadecanoin (C17:0 TAG) were used as a TAG internal standard. Seeds were subjected to treatment with 2 ml of 3N methanolic-HCL and heated at 80^{o}C for 16 h. The extracted FAMEs were suspended in 1.5 ml of iso-octane with 0.1 mg/ml C21:0 methyl ester standard and analyzed by GC-MS using the same column and temperature gradient. Total lipid content was determined by multiplying the peak-area ratio of the total FA and the internal standard by the initial internal standard amount.

3.2.10 Flax seed oil and FA composition analysis

Flax embryos of cultivar CDC Bethune at different developmental stages (8, 11, 14, 16 and 20 DPA) and seeds (25, 30 and 40 DPA) were collected, measured for fresh weight, immersed in liquid nitrogen, and then placed at -80^{o}C for storage. Frozen embryos and seeds (approximately 15 to 75 mg fresh weight per biological replicate) were freeze-dried for four days and then homogenized in a mortar and pestle in the presence of liquid nitrogen. Homogenates were preceded for lipid extraction and FA analysis as described in Arabidopsis seed oil analysis. The oil and ALA content were calculated on a fresh weight basis.

3.2.11 Phylogenetic analysis

The multiple sequence alignments of PDAT, DCR, DGAT1 and DGAT2–like family proteins were generated using the ClustalW module (Thompson, et al., 1994) within MEGA5 (http://www.megasoftware.net/mega.html) (Tamura, et al., 2011) with the defaul parameters (gap penalty, 10.0; gap length penalty, 0.2; Gonnet matrix). The phylogenetic trees were constructed using the same software with the following parameters: neighbor-joining method, Poisson model, complete deletion and bootstrap (1000 replicates). Numbers above branches indicate the percentage of bootstrap values.

3.2.12 Accession number

Accession numbers of protein sequences from the Arabidopsis Genome Initiative or EMBL/GenBank database or flax genome database (http://www.phytozome.net/flax) used in the current study are given in Table 3.2.

3.3 Results

3.3.1 Identification and isolation of genes encoding TAG-synthesizing enzymes from flax

The recently completed flax genome sequence database (Wang, et al., 2012) provided the starting point for identifying flax genes homologous to known genes involved in the final step of TAG synthesis. Using basic local alignment search tool (BLAST), three *DGAT2s* (*LuDGAT2-1*, *LuDGAT2-2* and *LuDGAT2-3*), two *DCRs* (*LuDCR1* and *LuDCR2*) and six *PDATs* (*LuPDAT1*, *LuPDAT2*, *LuPDAT3*, *LuPDAT4*, *LuPDAT5* and *LuPDAT6*) were identified in the flax genome. In addition, the previous study from our group reported a *DGAT1* (here referred to as *LuDGAT1-1*) (Siloto, et al., 2009a) and now a second gene (*LuDGAT1-2*) was identified in the flax genome. General information about the identified genes is listed in Table 3.3 while the gene structures were represented in Figure 3.1.

Table 3.2 Accession number of protein sequences from Arabidopsis Genome Initiative or EMBL/GenBank database or flax genome database used in the current study.

Protein	Accession #	Organism
LuDCR1	KC437082	*Linum usitatissimum*
LuDCR2	Lus10039256	*Linum usitatissimum*
AtDCR	AT5G23940	*Arabidopsis thaliana*
BdDCR	XP_003571265.1	*Brachypodium distachyon*
CtDCR	BAF49303.1	*Clitoria ternatea*
GhDCR	AAL67994.1	*Gossypium hirsutum*
GmDCR1	XP_003533987.1	*Glycine max*
GmDCR2	XP_003548168.1	*Glycine max*
HvDCR	BAK01773.1	*Hordeum vulgare*
PtDCR1	XP_002318207.1	*Populus trichocarpa*
PtDCR2	XP_002322489.1	*Populus trichocarpa*
VvDCR1	XP_002271409.1	*Vitis vinifera*
VvDCR2	XP_002272435.1	*Vitis vinifera*
LuDGAT1-1	KC485337	*Linum usitatissimum*
LuDGAT1-2	Lus10021582	*Linum usitatissimum*
AtDGAT1	At2g19450	*Arabidopsis thaliana*
BnDGAT1	AAD45536.1	*Brassica napus*
RcDGAT1	XP_002514132.1	*Ricinus communis*
VfDGAT1	ABC94471.1	*Vernicia fordii*
OeDGAT1	AAS01606.1	*Olea europaea*
NtDGAT1	AAF19345.1	*Nicotiana tabacum*
GmDGAT1	NP_001237289.1	*Glycine max*
MtDGAT1	ABN09107.1	*Medicago truncatula*
JcDGAT1	ABB84383.1	*Jatropha curcas*
PtDGAT1	XP_002330510.1	*Populus trichocarpa*
EaDGAT1	AAV31083.1	*Euonymus alatus*
VvDGAT1	CAN80418.1	*Vitis vinifera*
VgDGAT1	ABV21945.1	*Vernonia galamensis*
TmDGAT1	AAM03340.2	*Tropaeolum majus*
PfDGAT1	AAG23696.1	*Perilla frutescens*
ZmDGAT1	ABV91586.1	*Zea mays*
LuDGAT2-1	KC437083	*Linum usitatissimum*
LuDGAT2-2	Lus10039136	*Linum usitatissimum*
LuDGAT2-3	KC437084	*Linum usitatissimum*

AtDGAT2	AT3G51520	*Arabidopsis thaliana*
CrDGAT2	XP_001693189.1	*Chlamydomonas reinhardtii*
MtDGAT2	ACJ84867.1	*Medicago truncatula*
OsDGAT2	NP_001057530.1	*Oryza sativa*
OtDGAT2	CAL58088.1	*Ostreococcus tauri*
PpDGAT2	XP_001777726.1	*Physcomitrella patens*
PsDGAT2	ABK26256.1	*Picea sitchensis*
PtDGAT2	XP_002317635.1	*Populus trichocarpa*
RcDGAT2	AAY16324.1	*Ricinus communis*
SpDGAT2	AAQ89590.1	*Spirodela polyrrhiza*
VfDGAT2	ABC94474.1	*Vernicia fordii*
VvDGAT2	CAO68497.1	*Vitis vinifera*
ZmDGAT2	ACG38122.1	*Zea mays*
LuPDAT1	KC437085	*Linum usitatissimum*
LuPDAT2	KC437086	*Linum usitatissimum*
LuPDAT3	Lus10019519	*Linum usitatissimum*
LuPDAT4	Lus10015639	*Linum usitatissimum*
LuPDAT5	Lus10017165	*Linum usitatissimum*
LuPDAT6	KC437087	*Linum usitatissimum*
AtPDAT1	AT5G13640	*Arabidopsis thaliana*
AtPDAT2	At3g44830	*Arabidopsis thaliana*
GmPDAT1-1	XP_003550622.1	*Glycine max*
GmPDAT1-2	XP_003548968.1	*Glycine max*
GmPDAT1-3	XP_003541296.1	*Glycine max*
GmPDAT1-4	XP_003528441.1	*Glycine max*
GmPDAT2	XP_003540780.1	*Glycine max*
JcPDAT	AEZ56255.1	*Jatropha curcas*
PtPDAT1	XP_002328081.1	*Populus trichocarpa*
PtPDAT2	XP_002303252.1	*Populus trichocarpa*
PtPDAT3	XP_002303252.1	*Populus trichocarpa*
RcPDAT1-1	XP_002521350.1	*Ricinus communis*
RcPDAT1-2	AEJ32006.1	*Ricinus communis*
RcPDAT2	AEJ32007.1	*Ricinus communis*
VvPDAT1-1	XP_002275795.1	*Vitis vinifera*
VvPDAT1-2	XP_002278397.1	*Vitis vinifera*
VvPDAT2	XP_002271001.1	*Vitis vinifera*
LuFAD2	KC 469054	*Linum usitatissimum*
LuFAD3	KC 469055	*Linum usitatissimum*

Both *LuDGAT1* genes contain 15 exons and like *DGAT1* from other plant species, the first exon is the longest and encodes a hydrophilic domain (Liu, et al., 2012). *LuDGAT2-1* and *LuDGAT2-3* are composed of 8 exons each and share a more similar gene structure than with *LuDGAT2-2*, which contains only 6 exons. Interestingly, *LuDGAT2-3* is located adjacent to *LuDGAT2-2* in the genome (Figure 3.2). Six LuPDATs can be divided into three groups based on polypeptide sequences (Figure 3.3A). The genes within each group have a similar gene structure, where *LuPDAT3* and *LuPDAT6* genes contain 5 exons, one less than in *LuPDATs* of the other two groups. Phylogenetic analyses of the predicted polypeptides for each class indicate a close relationship with known TAG-synthesizing enzymes from other species (Figures 3.3A to 3.3D). Therefore, it was concluded that the identified genes comprise the most probable candidates for encoding TAG-synthesizing enzymes from flax. In addition, the identification of gene pairs for all genes, except to *LuDGAT2-2*, is consistent with the findings that suggest whole-genome duplication in flax (Wang, et al., 2012). The two genes in a gene pair have a very high degree of sequence identity at both nucleotide and amino acid levels (Table 3.4). Further genomic analysis revealed that *LuDGAT1-1* and *LuPDAT5* are located close together on the same chromosome and the other member of the gene pair (*LuDGAT1-2* and *LuPDAT1*) are similarly arranged on a different chromosome (Figure 3.4).

Table 3.3 General information on individual candidate genes and the encoded polypeptides (deduced from the DNA sequence information)

Genes	Gene length (bp)	Protein length	Molecular mass (Daltons)	Isoelectric point
DGAT1-1	1524	507	58038.19	8.98
DGAT1-2	1542	513	58793.08	8.92
DGAT2-1	1065	354	39405.12	9.49
DGAT2-2	1059	352	39004.68	9.19
DGAT2-3	1050	349	38951.64	9.42
DCR1	1137	378	41128.85	5.49
DCR2	1401	466	50887.96	5.10
PDAT1	2088	695	76891.89	8.23
PDAT2	2145	714	79072.61	6.4
PDAT3	1728	575	63093.05	6.19
PDAT4	2148	715	78923.51	6.72
PDAT5	2088	695	76817.64	8.28
PDAT6	1719	572	62792.54	6.19

Table 3.4 Sequence identity between the gene pairs and gene products

Gene Pair	Nucleotide (%)	Amino acid (%)
DGAT1-1 and *DGAT1-2*	85.9	97.7
DGAT2-1 and *DGAT2-3*	91.2	93.3
DCR1 and *DCR2*	90.2	89.3
PDAT1 and *PDAT5*	97	97.1
PDAT2 and *PDAT4*	95.6	97.1
PDAT3 and *PDAT6*	95.4	96

Figure 3.1 Genomic DNA structure of homologous genes encoding putative triacylglycerol (TAG)-synthesizing enzymes from flax. The thick lines indicate exons and the thin lines indicate introns. The numbers on top of each group denote the scale in base pairs.

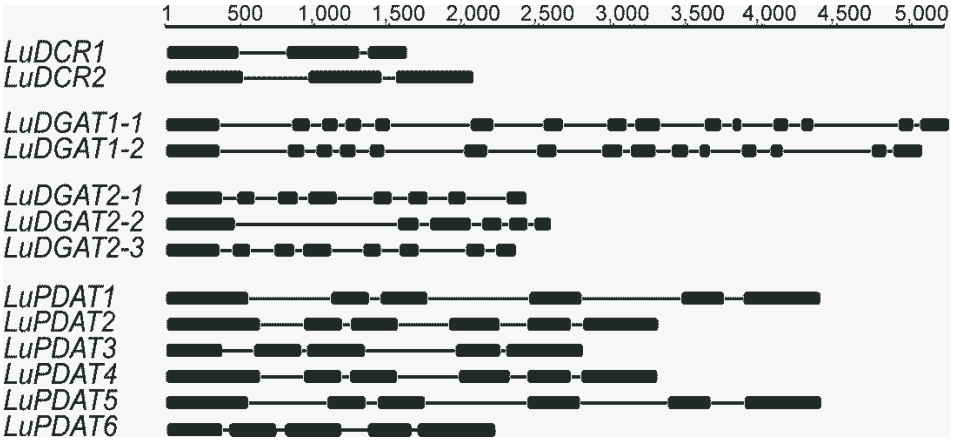


Figure 3.2 Genomic organization shows that *LuDGAT2-3* is located adjacent to *LuDGAT2-2* on the genome. The numbers on top of each group denote the scale in base pairs.

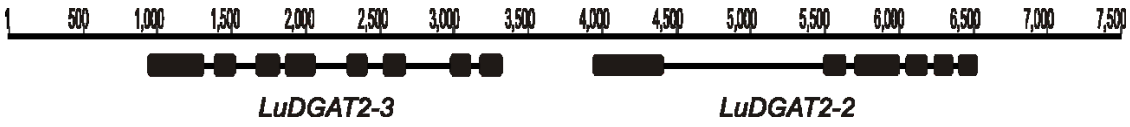


Figure 3.3 Phylogenetic relationship of PDAT (A), DCR (B), DGAT1 (C), and DGAT2 (D)–like family proteins from plants.

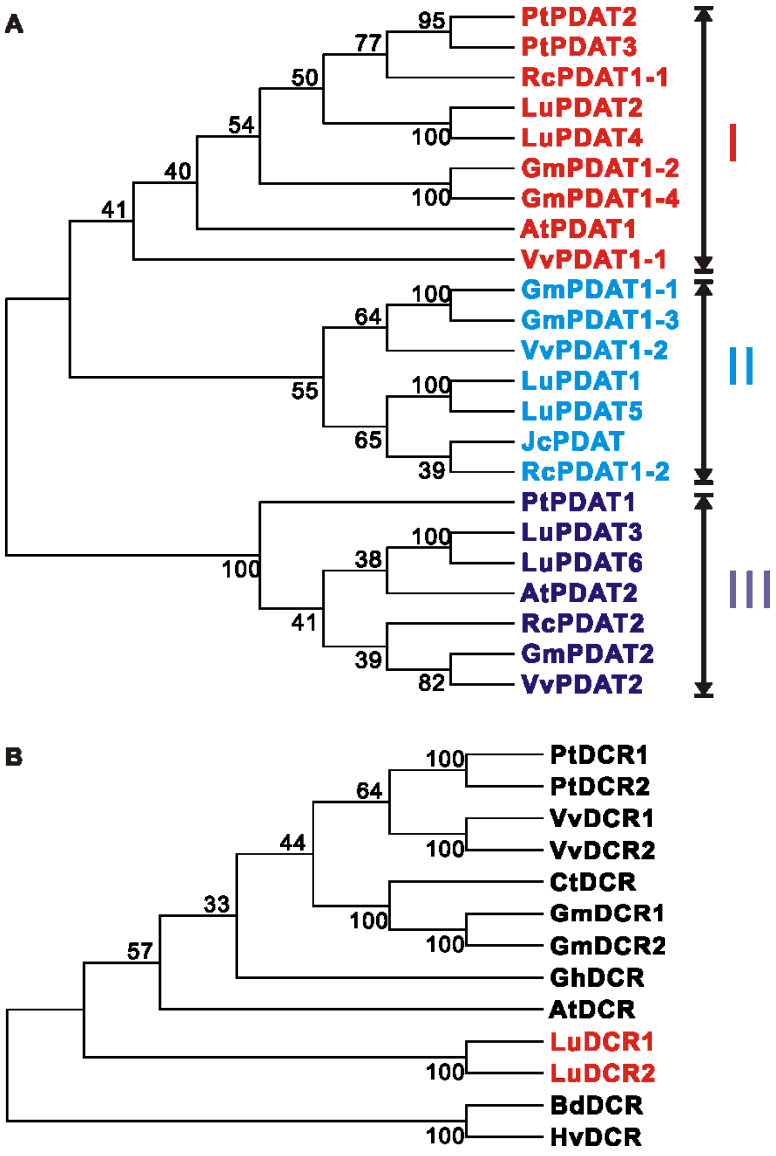


Figure 3.3 Continued

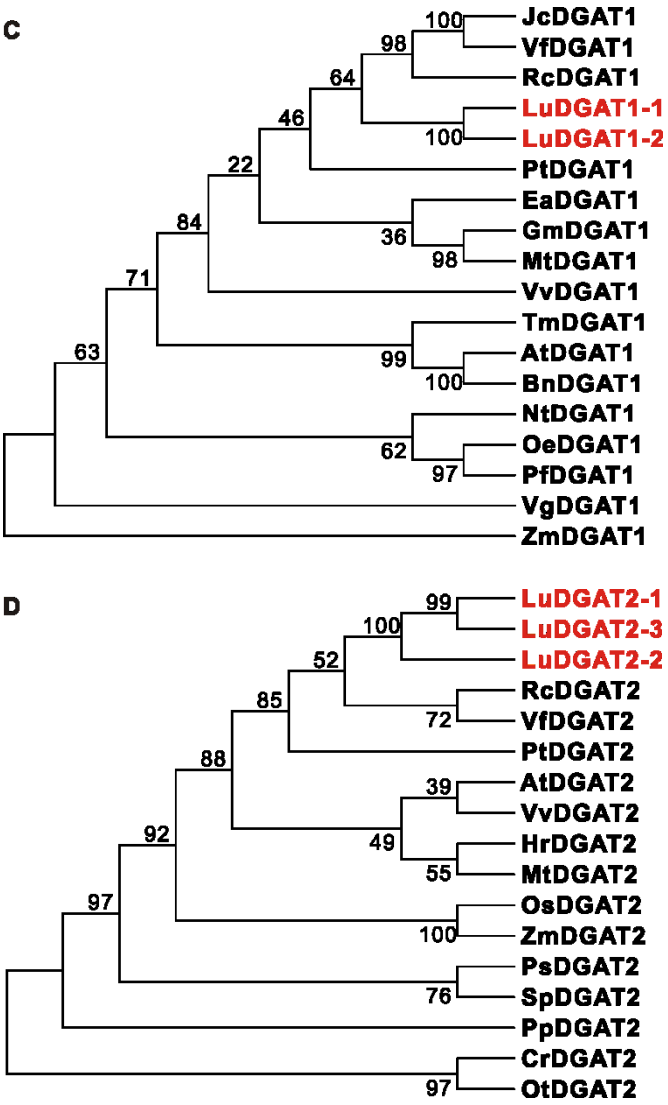
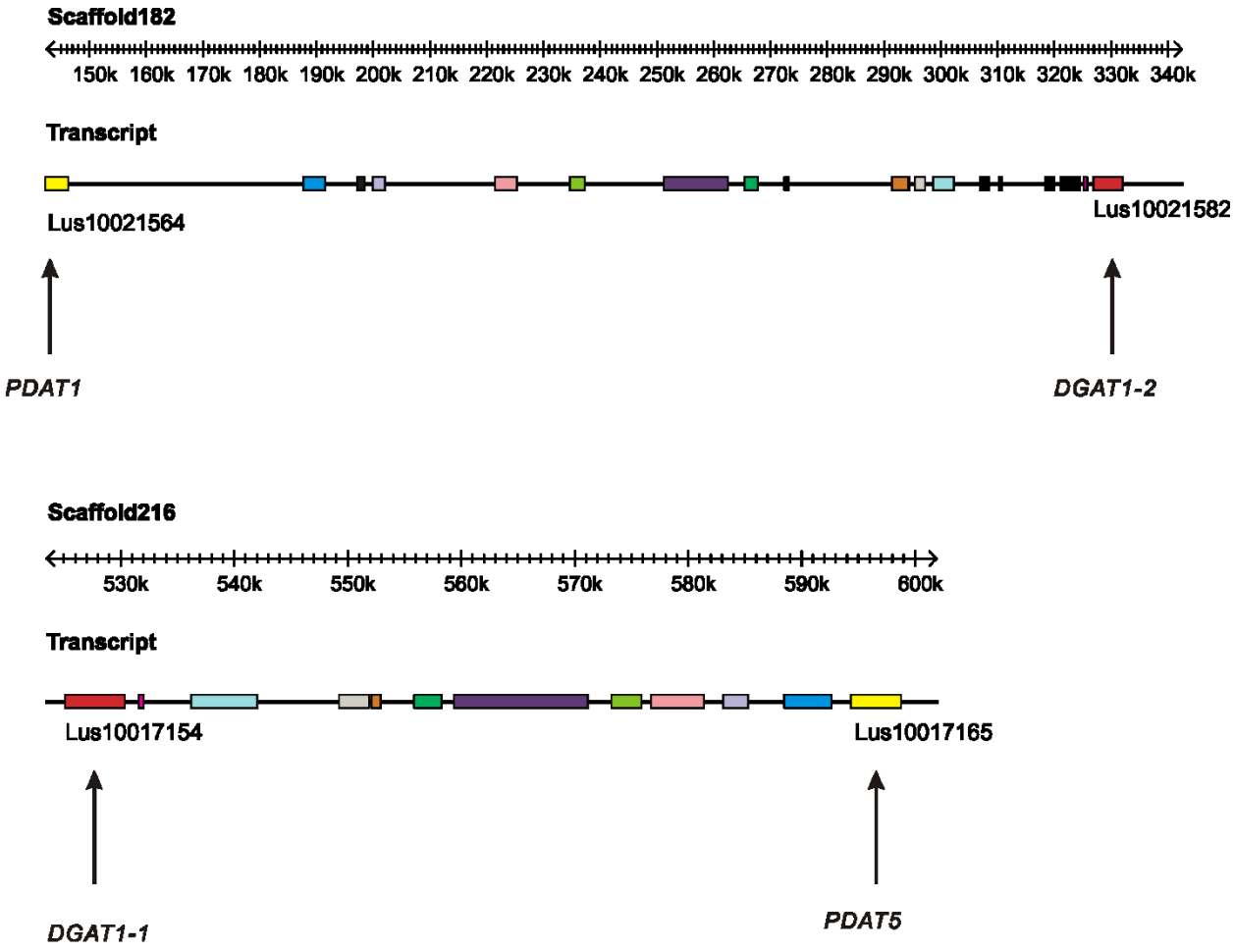


Figure 3.4 The relative locations of *LuPDATs* (*PDAT1* and *PDAT5*) and *LuDGAT1s* (*DGAT1-1* and *DGAT1-2*) on chromosomes (signified by the black lines). The colored boxes represent different genes as they are laid out on chromosome. Boxes with the same color on different chromosomes are a homologous gene pair. The linear scale denotes the distance along a chromosome at the scale of one thousand base pairs.



3.3.2 Several *DGAT1*, *DGAT2*, *PDAT* genes are expressed preferentially in seeds and correlate with oil accumulation

To investigate the potential physiological roles of all candidate genes, the expression patterns of *LuDGAT1, LuDGAT2, LuDCR* and *LuPDAT* were studied in a range of organs including vegetative tissues, reproductive tissues and mature seeds (40 DPA). Genes encoding LuDGAT1 or LuDGAT2 had significantly higher expression in seeds than in other vegetative tissues (Figures 3.5A to 3.5C). In the case of *LuDCR*, the expression level was significantly higher in vegetative tissues than that in seeds, suggesting that this gene might play a minor role in TAG biosynthesis (Figure 3.5D). *LuPDAT* genes displayed two distinct expression profiles. Transcripts of *LuPDAT1/LuPDAT5*, *LuPDAT3* and *LuPDAT6* (Figures 3.6A to 3.6C) were detected preferentially in seeds and to a lesser extent in other tissues, and the opposite was observed for *LuPDAT2* and *LuPDAT4* (Figures 3.6D and 3.6E). This result suggests that two groups of PDATs might have unrelated physiological function roles in flax. An attempt to individually detect the transcript of *LuPDAT5* failed, but I was able to isolate the cDNA of *LuPDAT5* using the reverse primer targeting 3'UTR. Due to the very high level of sequence identity between *LuPDAT1* and *LuPDAT5* (97%), the transcripts of *LuPDAT1* and *LuPDAT5* were detected together. The transcripts for *LuDGAT2-2* could not be detected in any of the tested plant tissues. Attempts to isolate the corresponding cDNAs were also unsuccessful, suggesting that *LuDGAT2-2* may not be expressed in the tested tissues.

These experiments indicated that *LuDGAT1*, *LuDGAT2-1*, *LuDGAT2-3*, *LuPDAT1/PDAT5*, *LuPDAT3* and *LuPDAT6* are preferentially expressed in seeds than other tissues. To further decipher the role of these seed-preferred genes in oil synthesis, the relationship between gene expression and seed oil and ALA accumulation patterns throughout

seed development was analyzed (Figures 3.7A to 3.7I). During seed development, the oil content on a fresh weight basis fit a sigmoidal curve (R^2 = 0.969) with the rapid phase of oil accumulation occurring between 8 and 20 DPA (Figure 3.7B). The ALA content on a fresh weight basis increased steadily until about 16 DPA (Figure 3.7C). The expression of *DGATs* and *PDATs* displayed dissimilar temporal regulation. The level of *DGATs* transcripts (Figures 3.7D to 3.7F) peaked during the late stages of seed development, when the rate of oil and ALA accumulation already reached a plateau. However, the highest expression of *PDATs* (Figures 3.7G to 3.7I) was found in the early stages of seed development, during the period of active seed oil and ALA accumulation. This result indicated that DGATs and PDATs might contribute differently to TAG synthesis during flax seed development.

3.3.3 DGATs and PDATs from flax are functional TAG-synthesizing enzymes

The ability of the polypeptides encoded by *LuDCR1*, *LuDGAT* and *LuPDAT* genes to catalyze the synthesis of TAG was investigated through a functional assay in yeast. Because genes in a gene pair show high sequence similarity, one gene of each pair, including *LuDGAT1-1, LuDGAT2-3, LuDCR1, LuPDAT1, LuPDAT2* and *LuPDAT6*, were chosen for this assay. The cDNAs of selected genes were isolated from 12 DPA flax developing embryos. The respective open reading frames (ORFs) were expressed in the quadruple mutant strain *S. cerevisiae* H1246, which lacks four genes *DGA1, LRO1, ARE1* and *ARE2* encoding DGAT1, PDAT1, ASAT1 (acyl-CoA:sterol acyltransferase 1) and ASAT2, respectively. With the disruption of these four genes, *S. cerevisiae* H1246 is unable to synthesize TAG and sterol esters (Sandager, et al., 2002), thus highlighting the production of TAG that results from the activity of recombinant flax enzymes. It has been previously demonstrated by our group that yeast H1246 expressing

Figure 3.5 Quantitative real-time PCR analysis of relative expression of flax *DGAT* genes in different tissues. The transcript levels of *LuDGAT1* (A) and *LuDGAT2* (B and C) genes were higher in seeds (40 DPA, black bar) than that in other tissues (gray bar) and the opposite was observed for *LuDCR1* (D). Data are given as means ± SE (n=3). S1: immature stem, S2: developing stem, S3: mature stem, L: leaves, A: apexes, R: roots, F: flowers, 40D: mature seeds.

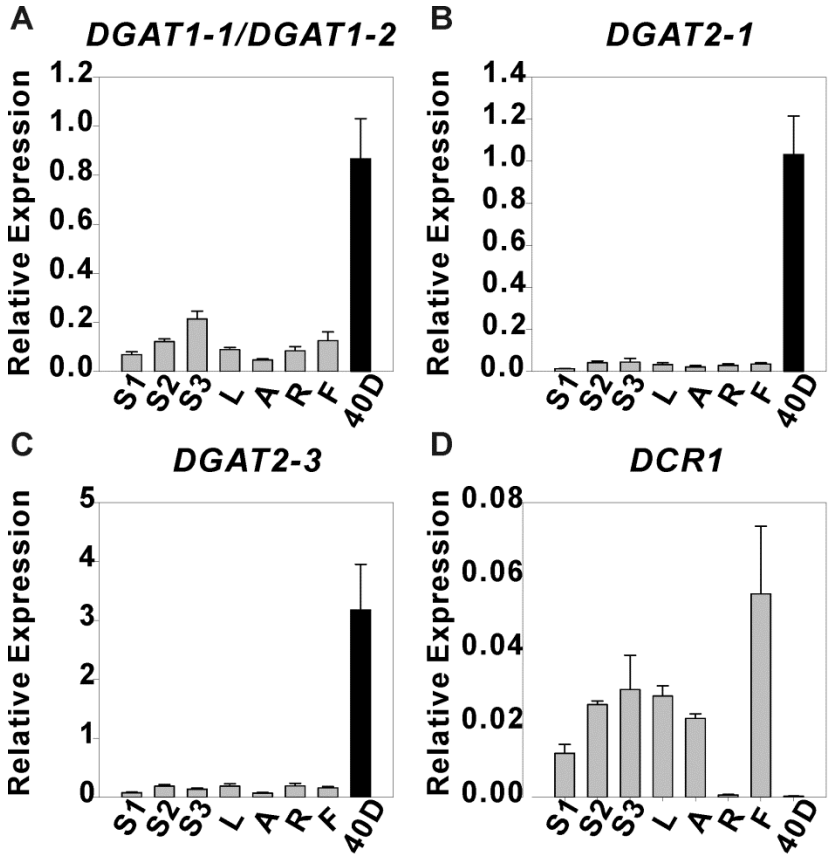


Figure 3.6 Quantitative real-time PCR analysis of relative expression of flax *PDAT* genes in different tissues. The *LuPDAT1/LuPDAT5* (A), *LuPDAT3* (B) and *LuPDAT6* (C) genes exhibited higher expression level in seeds (40 DPA, black bar) than that in other tissues (gray bar) and the opposite was observed for *LuPDAT2* (D) and *LuPDAT4* (E). (F) Phylogenetic analysis of six flax PDAT proteins. Genes within the same branch of the phylogenetic tree have the similar expression pattern. Data are given as means ± SE (n=3). S1: immature stem, S2: developing stem, S3: mature stem, L: leaves, A: apexes, R: roots, F: flowers, 40D: mature seeds.

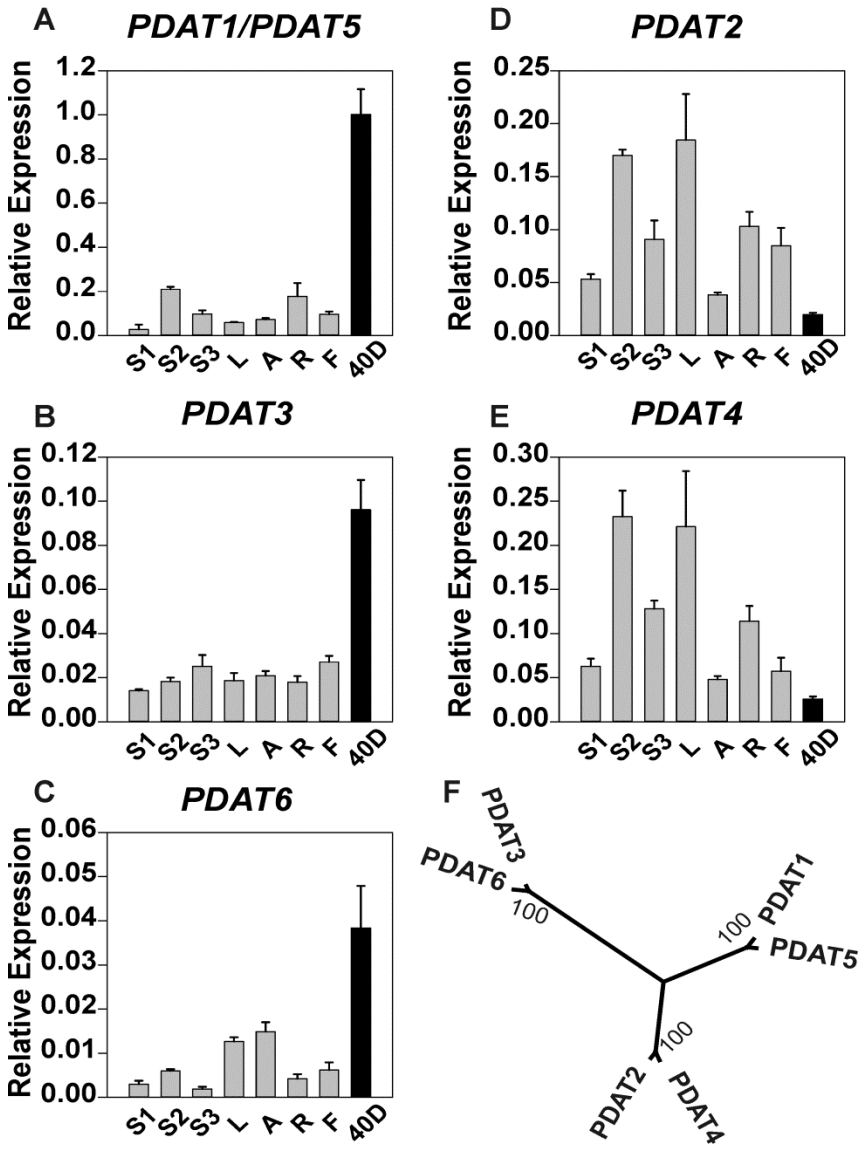
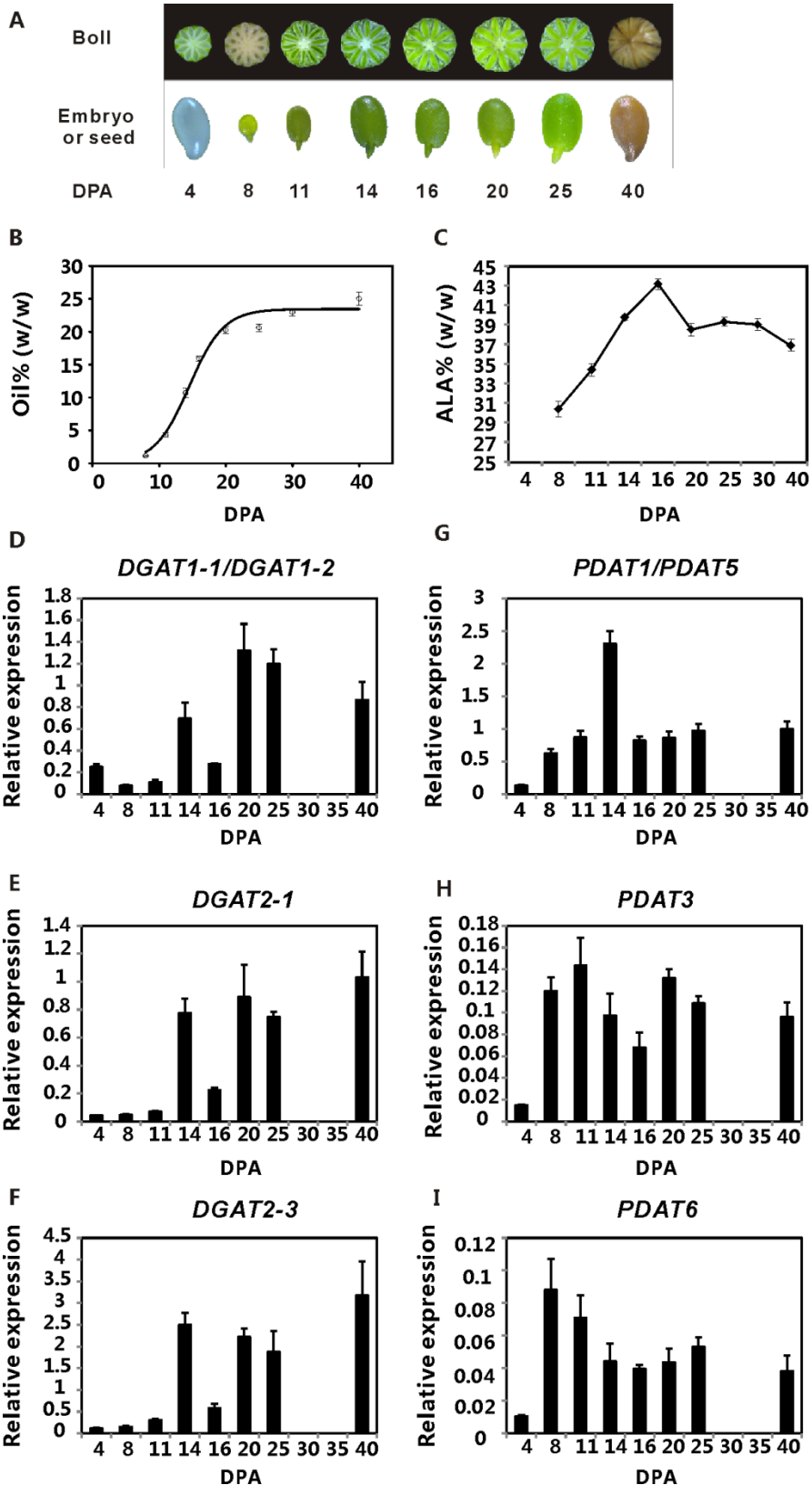
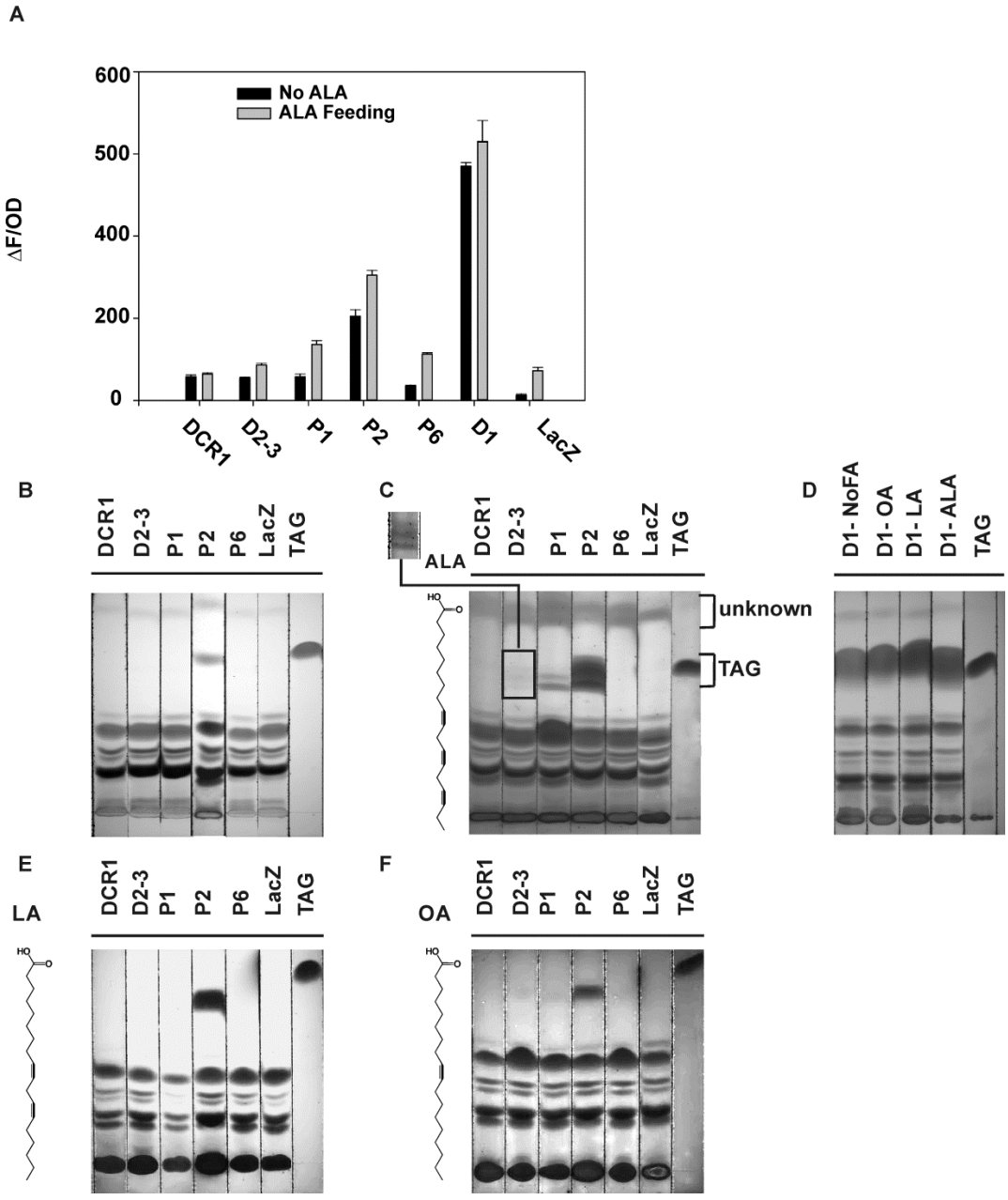


Figure 3.7 Oil and α-linolenic acid (ALA) accumulation correlate better with the expression of *LuPDATs* than with *LuDGATs*. (A) Flax embryos or seeds used for quantitative real-time PCR analysis. The top panel shows the cross-sections of the developing bolls at 4 to 25 DPA and the whole mature boll at 40 DPA from left to right. The bottom panel shows the cleared seed at 4 DPA, developing embryos at 8 to 25 DPA and the mature seed at 40 DPA. DPA, days post anthesis. (B) Oil deposition in seed on a fresh weight basis (% of fresh weight). The mean values of oil content (n=4) were fitted to a sigmoidal curve (R^2=0.969) using Sigmaplot 12.3. (C) ALA accumulation pattern during seed development. The data is presented on a fresh weight basis (% of fresh weight) and shown as mean ± SE (n=4). (D to I) Analysis of gene expression patterns of *LuDGATs* and *LuPDATs* during seed development. Data are shown as mean ± SE (n=3).


A
Boll
Embryo or seed
DPA
4 8 11 14 16 20 25 40
B
Oil% (w/w)
DPA
C
ALA% (w/w)
DPA
D
DGAT1-1/DGAT1-2
Relative expression
DPA
G
PDAT1/PDAT5
Relative expression
DPA
E
DGAT2-1
Relative expression
DPA
H
PDAT3
Relative expression
DPA
F
DGAT2-3
Relative expression
DPA
I
PDAT6
Relative expression
DPA

LuDGAT1-1 has the ability to synthesize TAG (Siloto, et al., 2009a), thus it was used as a positive control. Yeast transformed with the pYES vector coding for the bacterial protein LacZ was used as a negative control. Production of TAG in recombinant yeast was detected by the Nile red fluorescent assay (Siloto, et al., 2009a) as well as by TLC. The results of Nile red assay (Figure 3.8A) showed that besides LuDGAT1-1, LuPDAT2 was the only enzyme that had significant activity in yeast. Because flax oil contains a substantial amount of ALA, which is absent in *S. cerevisiae*, it was hypothesized that some of these enzymes might be selective for substrates containing ALA. Yeast is capable of importing FAs from the environment and converting them to their respective acyl-CoA derivatives (Færgeman, et al., 2001). Therefore, the expression of the genes was induced in the presence of exogenously added ALA. Nile red results indicated that all enzymes displayed an enhanced activity in the presence of ALA with the exception of LuDCR1 (Figure 3.8A). The Nile red assay is a high throughput technique for quantitative measurement of neutral lipids; however, it is unable to distinguish different neutral lipid classes. Therefore, TLC was used to further confirm the production of TAG in recombinant H1246 yeast cells. There was a good correlation between the results from Nile red and TLC analysis (Figures 3.8A to 3.8D). Two recombinant PDATs (LuPDAT1 and LuPDAT2) restored TAG synthesis in H1246 when culturing yeast in the presence of ALA (Figure 3.8C). The recombinant LuDGAT2-3 also produced TAG, but at significantly lower levels than those found in LuPDAT1, LuPDAT2 and LuDGAT1-1(Figure 3.8C and 3.8D). This analysis confirmed that most of the flax genes can be functionally expressed in yeast and their encoded enzymes are able to complement the TAG synthesis mutations in H1246.

Figure 3.8 Functional complementation assay in yeast strain H1246. Yeast cells expressing several flax genes show enhanced synthesis of triacylglycerol (TAG) in the presence of α-linolenic acid (ALA). (A) Nile red fluorescence assay was performed on the recombinant yeast mutant H1246 cultured in medium with ALA (gray bar) or without ALA supplementation (black bar). The value represents the levels of neutral lipids accumulated in the recombinant yeast, which is calculated by dividing Nile red fluorescence (ΔF) by the optical density (OD) at 600nm. Data are presented as means ± SE (n = 3). Total lipid extracts were prepared from these cultures and the individual lipid classes were separated by thin layer chromatography (TLC) (B and C). (B) TLC analysis of yeast lipid in the absence of ALA.(C) TLC analysis of yeast lipid in the presence of ALA. The square bracket indicates the position of TAG produced by yeast cells. The insert displays a high contrast image of TAG produced in yeast expressing *LuDGAT2-3* in the present of ALA. (D) TLC analysis of yeast cells expressing *LuDGAT1-1* in the absence or presence of different exogenously provided FAs. (E) and (F) TLC analysis of yeast lipid in the presence of linoleic acid (LA) and oleic acid (OA). Yeast cells transformed with pYES*LacZ* was used as the negative control. Although yeast cells transformed with pYES*LacZ* and pYES*LuPDAT6* also showed the higher Nile red value in the presence of ALA (A), TLC plate indicated that the increased value was very likely due to the production of the unknown compound (square bracket indicated) (C). The corresponding FA used for feeding is shown on the left of the figure with the chemical structure. D1– LuDGAT1-1; DCR1 – LuDCR1; D2-3 – LuDGAT2-3; P1– LuPDAT1; P2 – LuPDAT2; P6 – LuPDAT6 ; TAG – triolein TAG standard in (B, E and F), trilinolenin TAG standard in (C and D).


A
600
500
400
200
0
ΔF/OD
No ALA
ALA Feeding
DCR1
D2-3
P1
P2
P6
D1
LacZ
B
DCR1
D2-3
P1
P2
P6
LacZ
TAG
C
ALA
DCR1
D2-3
P1
P2
P6
LacZ
TAG
unknown
TAG
D
D1- NoFA
D1- OA
D1- LA
D1- ALA
TAG
E
LA
DCR1
D2-3
P1
P2
P6
LacZ
TAG
F
OA
DCR1
D2-3
P1
P2
P6
LacZ
TAG

3.3.4 LuPDAT1 and LuPDAT2 have unique ALA-selectivity

When ALA was exogenously added to induction medium, yeast cells expressing *LuPDAT1* and *LuPDAT2* produced multiple TAG bands on the TLC plate (Figure 3.8C). This was not observed in yeast expressing *LuDGAT1-1* (Figure 3.8D). Comparison of the migration rates of trilinolenin and triolein standards under our TLC conditions (Figure 3.9) suggested that the observed multiple TAG bands could be explained by the different composition of FAs in these TAGs. Therefore, it is plausible to assume that in the presence of ALA, LuPDAT1 and LuPDAT2 produce TAG with a specific FA composition. Indeed, GC analysis of TAG bands with different migration rates indicated that the TAG band with lower migration rate was composed uniquely of ALA, while the band with higher migration rate contained 65 mol% and 49 mol% ALA for LuPDAT1 and LuPDAT2, respectively (Figure 3.10). This result indicated that both LuPDAT1 and LuPDAT2 have the ability to synthesize trilinolenin, which is the major molecular species of TAG in flax oil (Ciftci, et al., 2012). This experiment also suggested that these two PDATs might be selective for substrates containing ALA, or have a preference for utilizing exogenously imported acyl moieties into yeast. To investigate the origin of this effect, the recombinant yeast strains were cultivated in the presence of OA or LA. As shown in Figure 3.8, yeast cells expressing *LuPDAT1* and *LuPDAT2* produced much weaker TAG bands in these medium conditions (Figure 3.8E and 3.8F) than in medium supplemented with ALA (Figure 3.8C). It is also worth noting that substantial amount of TAG was produced by LuDGAT1-1 under all conditions (Figure 3.8D). Overall, these results suggest that both LuPDAT1 and LuPDAT2 have a high preference for ALA-containing substrates. FAs commonly found in yeast, including stearic, palmitic, palmitoleic acids and OA are inefficient substrates for these PDATs.

Figure 3.9 Different migration rate of trilinolenin and triolein on the TLC plate.

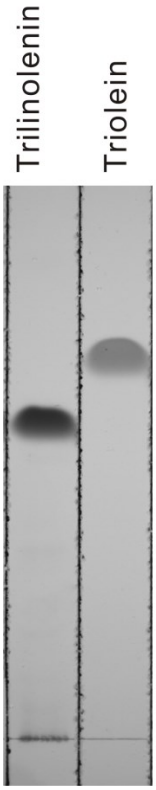
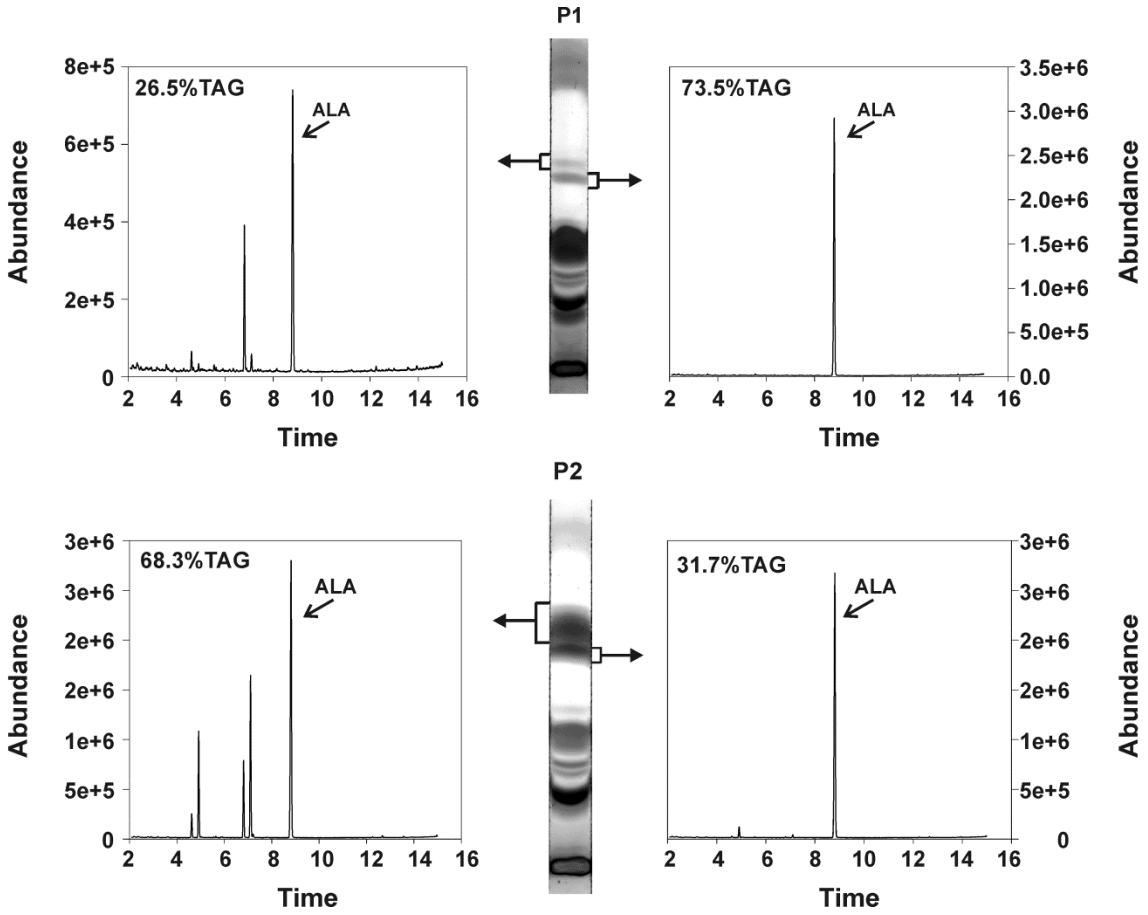


Figure 3.10 Gas chromatography-mass spectrometry chromatograms of yeast strain H1246 expressing *LuPDAT1* and *LuPDAT2* in the presence of α-linolenic acid (ALA).

The recombinant yeast cells were cultivated in the presence of ALA. The yeast lipids were extracted and separated by thin layer chromatography (TLC) plate. The compound corresponding to the upper and lower triacylglycerol (TAG) bands (marked by square brackets) was scrapped separately from the TLC plate, trasnsmethylated and analyzed through GC/MS. The ratio of TAG within each separated band to the total amount of TAG was calculated and the values are shown as percentage on the upper left corner of each chromatograph. GC profiles showed that yeast strain H1246 expressing *LuPDAT1* or *LuPDAT2* has the ability to produce TAG containing only ALA (trilinolenin), when ALA is added exogenously to the cell. P1– LuPDAT1; P2 – LuPDAT2.



3.3.5 LuPDAT1 and LuPDAT2 have the similar preference for ALA-containing substrates

The free FA feeding study revealed that both LuPDAT1 and LuPDAT2 are uniquely ALA-selective. It was also intriguing that, when ALA was provided exogenously, LuPDAT1 catalyzed the synthesis of a higher amount of trilinolenin with lower total TAG as compared with LuPDAT2 (Figure 3.10). To further compare these two enzymes, I first used Western blot to assess the protein expression levels of these PDATs in H1246. Both enzymes with C-terminal fusion tag had reduced TAG-forming ability. Therefore, N-terminal-tagged proteins were used for detection. Western blot analysis revealed that the band corresponding to PDAT2 was markedly stronger than that corresponding to PDAT1 (Figure 3.11A), suggesting that H1246 was able to accumulate much higher levels of PDAT2 than PDAT1. It should also be noted that PDAT6 was expressed at a similar protein level to PDAT2, but it failed to complement TAG synthesis in H1246 under all conditions.

It is likely that the higher level of PDAT2 produced in yeast may enhance the rate of channeling ALA into TAG, thus depriving the exogenously introduced ALA in the DAG/PC pool and leading to the formation of TAG with lower content of ALA as compared with PDAT1. To address this possibility, the concentration effect of exogenously provided ALA on total TAG content and composition in yeast expressing *LuPDAT1* and *LuPDAT2* was further investigated by culturing the recombinant cells in medium supplemented with 0 to 300 μM of ALA (Figure 3.11 B-E). As show in Figure 3.11C, both PDAT1 and PDAT2 were able to catalyze the production of TAG with up to 90% ALA, but PDAT2 reached the plateau at much higher concentration of ALA, suggesting that the increased abundance of recombinant PDAT2 would require more ALA-containing substrates to obtain TAG with the same level of ALA as produced by PDAT1. In addition, an increase in ALA concentration from 0 to 300 μM led to

approximately 168-fold and 44-fold increases in total TAG content for yeast expressing *LuPDAT1* and *LuPDAT2*, respectively (Figure 3.11D). Further quantifying the FA species in TAG revealed that elevated TAG content in these cells was mainly because of the increased amount of ALA (Figure 3.11E), suggesting that both enzymes may be more active in the presence of ALA. Taken together, a similar response of LuPDAT1 and LuPDAT2 to an increase of ALA concentration was observed, indicating that LuPDAT2 has similar ALA-selectivity as compared with LuPDAT1. The differential levels of protein accumulation, if not entirely, at least partially account for the observed differences in TAG-forming ability between PDAT1 and PDAT2.

3.3.6 Exogenously provided free FAs inhibit growth of the recombinant yeast mutants and this inhibitory effect can be rescued in a TAG accumulation-dependent manner

I observed that the recombinant yeast mutant displayed a different growth rate upon exposure to free FAs. To assess the effect of exogenously provided ALA on recombinant yeast cell growth, I evaluated the cell growth in induction medium supplemented with 0, 100 μM and 1mM of ALA by both growth curves analysis and the serial dilution plate assay. BY4742 (wild-type) and H1246 harboring the empty vector pYES were used as positive and negative controls, respectively. As shown in Figure 3.12, the presence of ALA produced a concentration-dependent inhibitory effect on recombinant yeast growth. Cells with higher TAG-synthesizing ability showed an earlier log phase and earlier appearance on plates, suggesting that inhibitory effect of ALA was rescued in a TAG accumulation-dependent manner. This result is in agreement with early findings, which reported that TAG biosynthesis plays a crucial role in detoxifying the excess imported unsaturated FAs (Siloto, et al., 2009a; Petschnigg, et al., 2009). Together, the

results implied that TAG-synthesizing enzymatic activity can be physiologically assessed by analyzing the growth of recombinant H1246 in the presence of ALA.

3.3.7 LuPDAT1 and LuPDAT2 can also preferably utilize substrates containing certain PUFAs

The previous results showed the unique selectivity properties of LuPDAT1 and LuPDAT2 for substrates containing ALA. To further examine the substrate preferences of flax PDATs, a similar experiment using other PUFAs that are not commonly found in plants was conducted, including SDA, DGLA, GLA, AA, ETA, EPA and DHA. LuPDAT1 and LuPDAT2 appeared to accept substrates containing SDA, GLA and EPA (Figure 3.13), while DGLA, AA, ETA and DHA were markedly poorer substrates for LuPDAT1 and LuPDAT2. Together, the data presented in Figures 3.8 and 3.13 indicated that LuDGAT1 are capable of utilizing a broad range of substrates containing different FAs. However, LuPDAT1 and LuPDAT2 had a clear preference for certain PUFAs, particularly those with at least three double bonds, such as ALA, GLA, SDA and EPA.

Feeding FAs to yeast provides an easy and flexible way to introduce novel acyl moieties in yeast endogenous lipids. This method is, however, limited as it does not allow a controlled distribution of exogenous FAs to specific glycerolipids as it occurs naturally. To circumvent this problem, *LuPDAT1*, *LuPDAT2* and *LuDGAT1-1* were individually co-expressed with *LuFAD2-1* and *LuFAD3B*, thus simulating a natural production of ALA in PC. In this experiment yeast cells expressing *LuPDAT1* and *LuPDAT2* produced TAG predominantly with ALA while cells expressing *LuDGAT1-1* produced TAG mainly with palmitoleic acid ($16{:}1^{cis\Delta 9}$) (Figure 3.14).

Figure 3.11 Yeast cell expressing LuPDAT1 and LuPDAT2 have similar responses to an increase of exogenously provided α-linolenic acid (ALA). (A) Representative Western blot analysis showing HisG-LuPDAT protein levels. Microsomes were prepared from yeast expressing HisG-PDATs (See Experimental Procedures). Cells transformed with empty pYES2/NT-C vector were used as a negative control, which was annotated as "-" in the image. Equal protein loading was ensured by measuring the constitutively expressed Kar2p in yeast. (B) Thin layer chromatography (TLC) analysis showing triacylglycerol (TAG)-forming ability of recombinant yeast in the presence of different concentration of ALA. The yeast cells expressing *LuPDAT1* and *LuPDAT2* were cultivated in the absence or presence of different ALA concentration and harvested at the same growth stage (OD_{600nm} = 6.5±0.05). The yeast lipids were extracted and separated by TLC and the corresponding TAG spots are shown in (B).(C-E) The ALA concentration effect on overall percentage of ALA in TAG (C), total TAG content (D) and amount of FAs in TAG (E) of recombinant yeasts. Total TAG content is presented on the basis of yeast dry weight. The amount of ALA and other FAs is presented as μmol of FAs per mg of yeast dry weight. Data are presented as means ± SE (n = 4). P1 – LuPDAT1; P2 – LuPDAT2; FA – fatty acid.

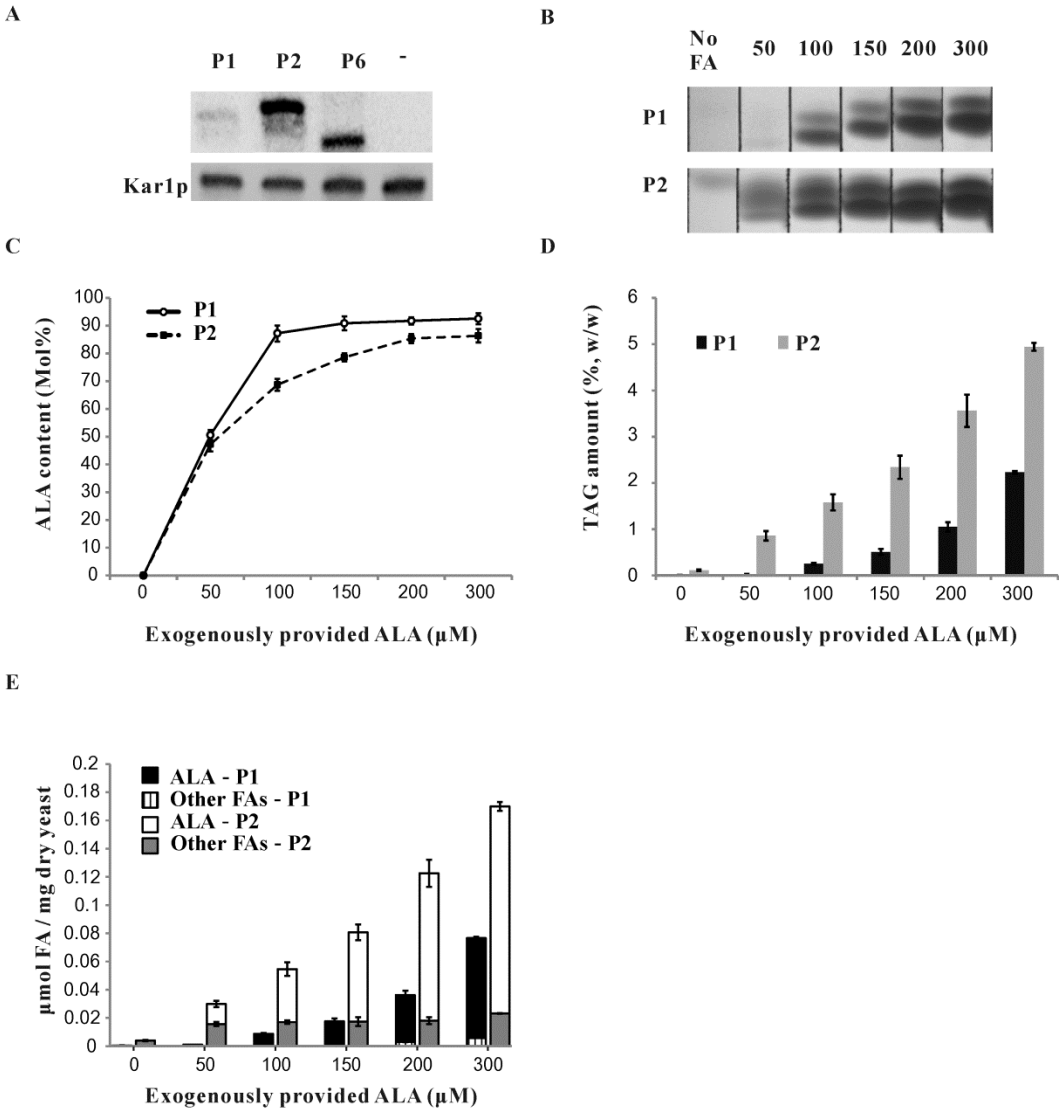

A
P1
P2
P6
-
Kar1p
B
No FA
50
100
150
200
300
P1
P2
C
ALA content (Mol%)
P1
P2
Exogenously provided ALA (µM)
D
TAG amount (%, w/w)
P1
P2
Exogenously provided ALA (µM)
E
µmol FA / mg dry yeast
ALA - P1
Other FAs - P1
ALA - P2
Other FAs - P2
Exogenously provided ALA (µM)

Figure 3.12 Rescue of lipotoxicity phenotype in H1246. Exogenous supplied α-linolenic acid (ALA) inhibits the growth of yeast H1246. This inhibition can be rescued by expressing genes encoding triacylglycerol (TAG)-synthesizing enzymes from flax and the rescuing effect is correlated with the ability to synthesize TAG as demonstrated with Nile red assay and TLC analysis. Growth curves of recombinant yeast H1246 cultured in induction medium with 0 (A), 100 μM (B) and 1mM (C) of ALA, were constructed by measuring the optical density (OD) of the cultures at 600 nm. Data are presented as means ± SE (n = 3 or 6). BY4742 (wild-type) and H1246 harboring the empty vector pYES were used as positive and negative controls, respectively. For negative controls, six independent colonies were analyzed and consistently showed delayed growth and low final density even in the absence of free FA. The panels on the right side, from top to bottom, show the growth of 2.5 μl of each tenfold dilution (1 to10^{-4}) of recombinant yeast H1246 spotted on to induction medium plates containing 0, 100 μM and 1 mM ALA. The negative controls were annotated as "-" in the image. WT– wild-type; D1-1 – LuDGAT1-1; DCR1 – LuDCR1; D2-3 – LuDGAT2-3; P1– LuPDAT1; P2 – LuPDAT2; P6 – LuPDAT6.

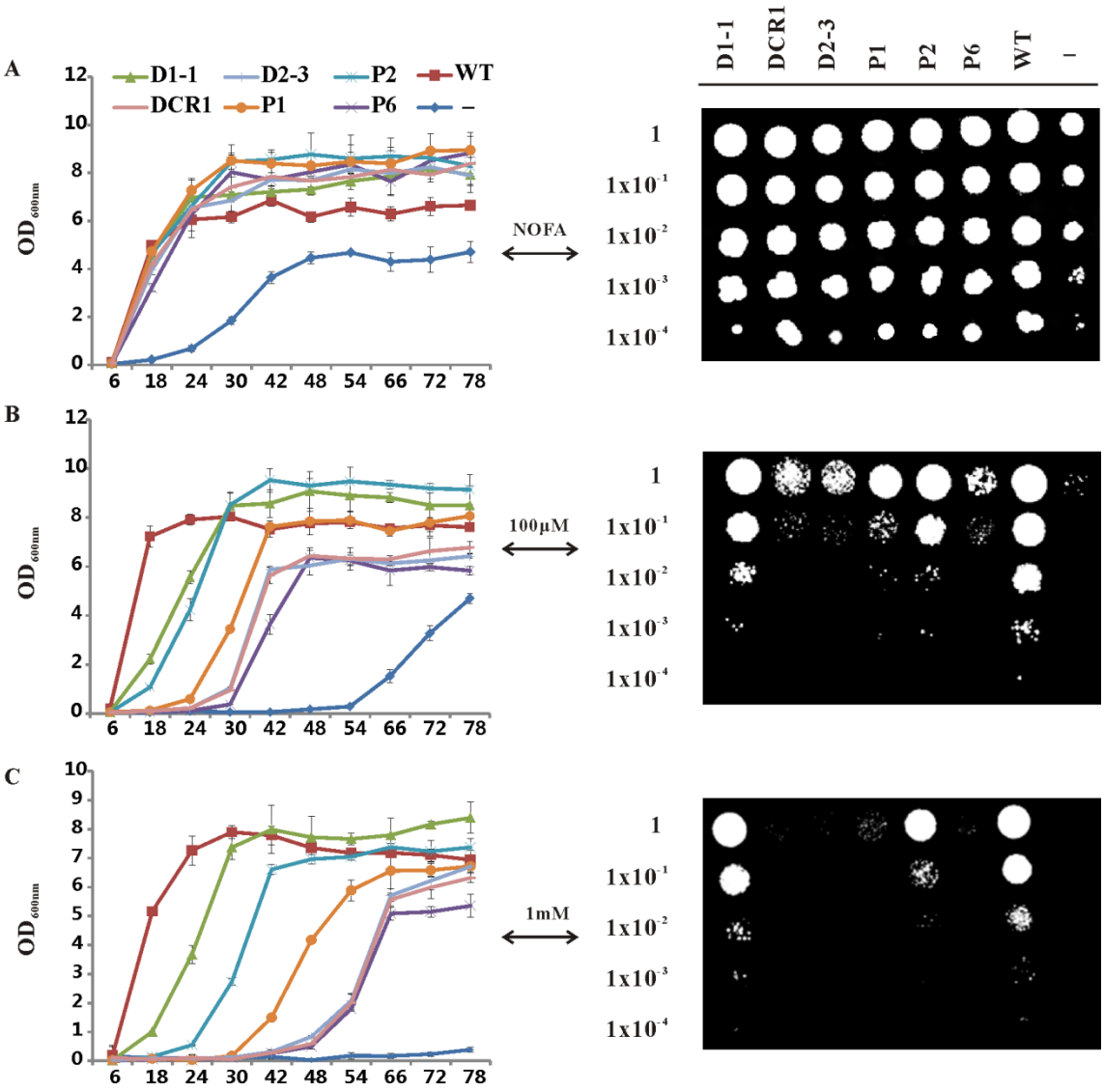

A
B
C
D1-1
D2-3
P2
WT
DCR1
P1
P6
–
OD600nm
NOFA
100μM
1mM
1
1x10-1
1x10-2
1x10-3
1x10-4
D1-1
DCR1
D2-3
P1
P2
P6
WT
–

Figure 3.13 Polyunsaturated FA (PUFA)-specific activity of LuPDAT1 and LuPDAT2 in yeast strain H1246. Yeast mutant H1246 transformed with *LuPDAT* genes was cultivated in the presence of different PUFAs fed individually to a final concentration of 100 μM. Yeast transformed with pYES*LacZ* was used as the negative control, which was annotated as "-" on the thin layer chromatography (TLC) plate. The corresponding FA used for feeding is shown on the left of the figure with the chemical structure. P1 – LuPDAT1; P2 – LuPDAT2; P6 – LuPDAT6 ; TAG – trilinolenin TAG standards; SDA – stearidonic acid ; GLA – γ-linolenic acid; DGLA – dihomo-γ-linolenic acid ; AA – arachidonic acid; EPA – eicosapentaenoic acid; ETA – eicosatrienoic acid; DHA – docosahexaenoic acid.

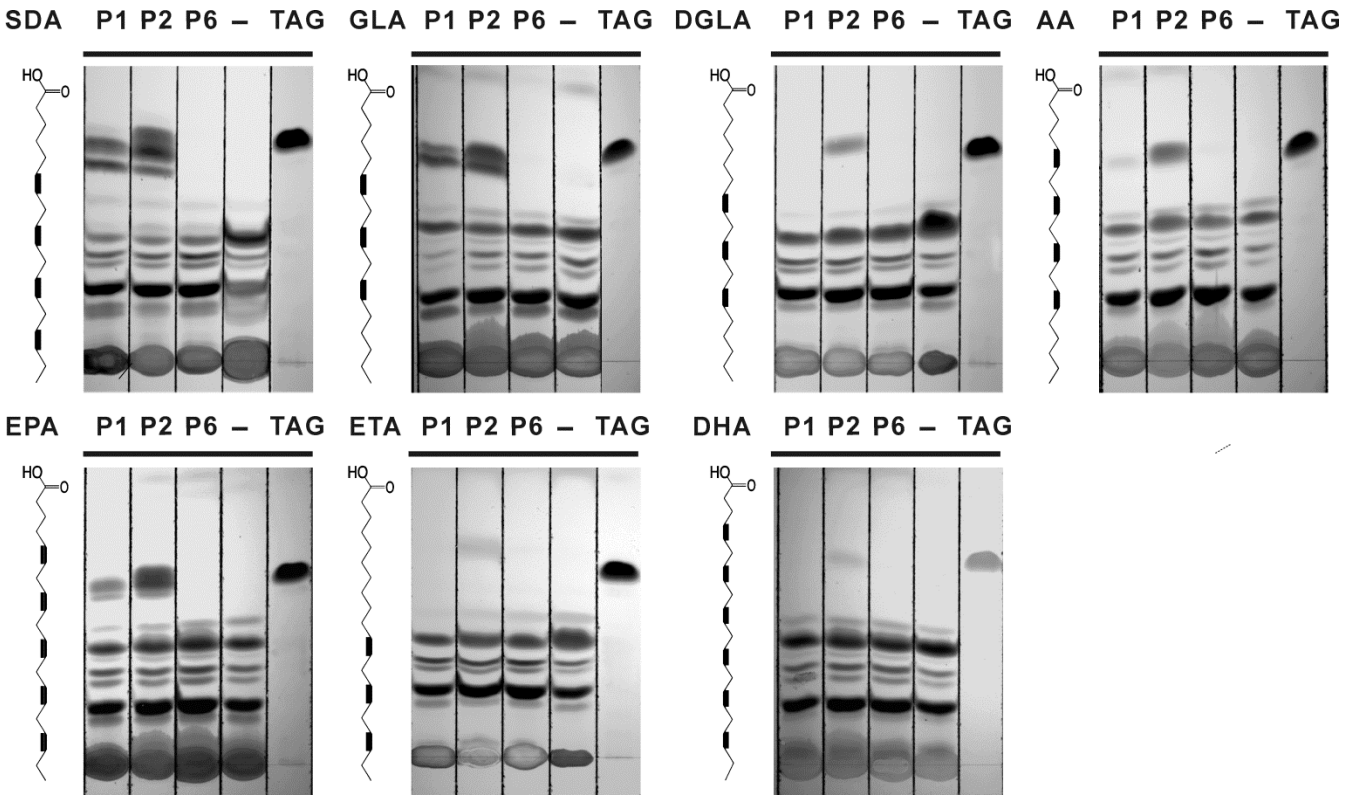
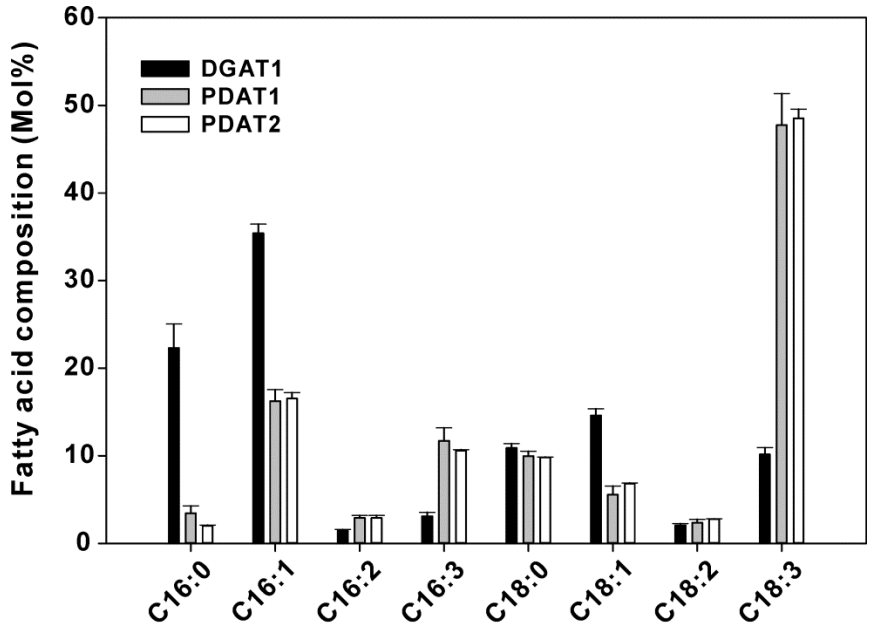

SDA P1 P2 P6 – TAG
GLA P1 P2 P6 – TAG
DGLA P1 P2 P6 – TAG
AA P1 P2 P6 – TAG
EPA P1 P2 P6 – TAG
ETA P1 P2 P6 – TAG
DHA P1 P2 P6 – TAG
HO
O

Figure 3.14 Comparison of FA methyl esters (FAME) profile of yeast strain H1246 co-expressing *LuPDAT1*, *LuPDAT2* or *LuDGAT1-1* individually with desaturases. When compared with the yeast cell co-expressing *LuDGAT1-1* with *LuFAD2-1* and *LuFAD3B* (black bar), co-expression of either *LuPDAT1* (gray bar) or *LuPDAT2* (white bar) with *LuFAD2-1* and *LuFAD3B* in yeast produces triacylglycerol (TAG) with α-linolenic acid (ALA, C 18:3) as the predominant FA. Yeast cultures were induced at 20^{o}C for three days before harvested. Data are presented as means ± SE (n = 3).



LuPDAT1 and LuPDAT2 were able to produce TAG with an ALA:LA ratio of 10:1, while LuDGAT1-1 produced TAG with an ALA:LA ratio of approximately 2:1. These results further confirmed that LuPDAT1 and LuPDAT2 have similar preference for ALA-containing substrate.

3.3.8 Seed-specific expression of *LuDGAT1-1*, *LuPDAT1* or *LuPDAT2* affects the FA profile and oil content in Arabidopsis

Complementation assays in yeast indicated that LuDGAT1-1, LuDGAT2-3, LuPDAT1 and LuPDAT2 are active TAG-synthesizing enzymes from flax. Though informative, this expression system is limited as yeast lacks the cellular and developmental complexity of multicellular higher plants. To investigate the functionality of flax TAG-synthesizing enzymes in plants, the ORFs of *LuDGAT1-1, LuDGAT2-3* and *LuDCR1* was expressed in Arabidopsis *dgat1* mutant AS11 under the regulation of the seed-specific napin promoter. AS11 has a reduced amount of TAG and altered FA profile (low OA and eicosenoic acid (EA, $20{:}1^{cis\Delta 11}$) and high ALA) (Katavic, et al., 1995). To minimize variation of genome position and copy number-dependent transgene expression (Tang, et al., 2003; Ahmad, et al., 2010), ten individual transgenic lines were analyzed for each construct. This strategy also helped to compensate for environmental effects that can significantly influence the accumulation of oil in the seeds of each plant even though they were cultivated side-by-side. Analysis of T_2 transgenic seed lines indicated that overexpression of *LuDGAT1-1* restored seed oil content of AS11 to levels comparable to the empty vector-transformed control wild-type plants (Figure 3.15A). LuDCR1 and LuDGAT2 did not complement the reduced TAG phenotype of the AS11 mutant (Figure 3.15A). GC analysis of the total lipid extract FA methyl esters (TLE-FAMEs) showed that LuDGAT1-1 could also alter the seed FA profile of AS11 making it similar to wild-type controls.

None of the *LuDGAT2-3* and *LuDCR1* transgenic lines could, however, complement this phenotype (Figure 3.15B).

It has been shown that the Arabidopsis *PDAT1* knockout mutant did not have any significant changes in either FA composition or oil content compared with wild-type (Mhaske, et al., 2005) while the attempt to obtain the Arabidopsis *dgat1 pdat1* double mutant failed due to the lethal effect of a *dgat1pdat1* genotype on pollen development (Zhang, et al., 2009). To study the biochemical properties of LuPDATs in plant oil biosynthesis, the constructs carrying the coding region of *LuPDAT*s (*LuPDAT1*, *LuPDAT2* or *LuPDAT6*) were transformed into Arabidopsis wild-type Columbia. As compared with the empty vector-transformed control wild-type plants, overexpression of the *LuPDAT1* and *LuPDAT2* increased the PUFA (LA and ALA) content, at the expense of mostly OA and EA (Figure 3.16A). This altered FA composition was not found in Arabidopsis lines overexpressing *LuPDAT6*. As shown in Figure 3.16A, the *LuPDAT1* lines exhibited a relative increase of LA in the range of 4.9% to 14.3% compared with the average exhibited by empty vector transformed wild-type controls. In addition, the ALA content was significantly increased in many lines, with the highest relative increase of 8.1%. The *LuPDAT2* transgenic T_2 seed lines showed a relative increase of LA by 4.4% to 21.8%. ALA content was generally increased, with the highest relative increase value observed being 10.5%. The oil content of *LuPDAT1, LuPDAT2* and *LuPDAT6* Arabidopsis transgenic lines were not significantly altered compared with controls (Figure 3.16B).

Figure 3.15 Lipid phenotype of *LuDGAT*-overexpressing seeds. Overexpressing *LuDGAT1-1* in Arabidopsis AS11 background restores the wild-type lipid phenotype. Ten individual transgenic lines were analyzed for each construct. Arabidopsis AS11 mutant and wild-type transformed with empty pGreen plasmid were used as controls (CTR). Data are presented as means ± SE (n = 3). OA – oleic acid; ALA – α-linolenic acid; EA – eicosenoic acid.(A) Seed oil content. Oil content is expressed as percentage of seed dry weight. (B) FA composition. Values were obtained from FA methyl esters (FAME) analysis of dry seeds.

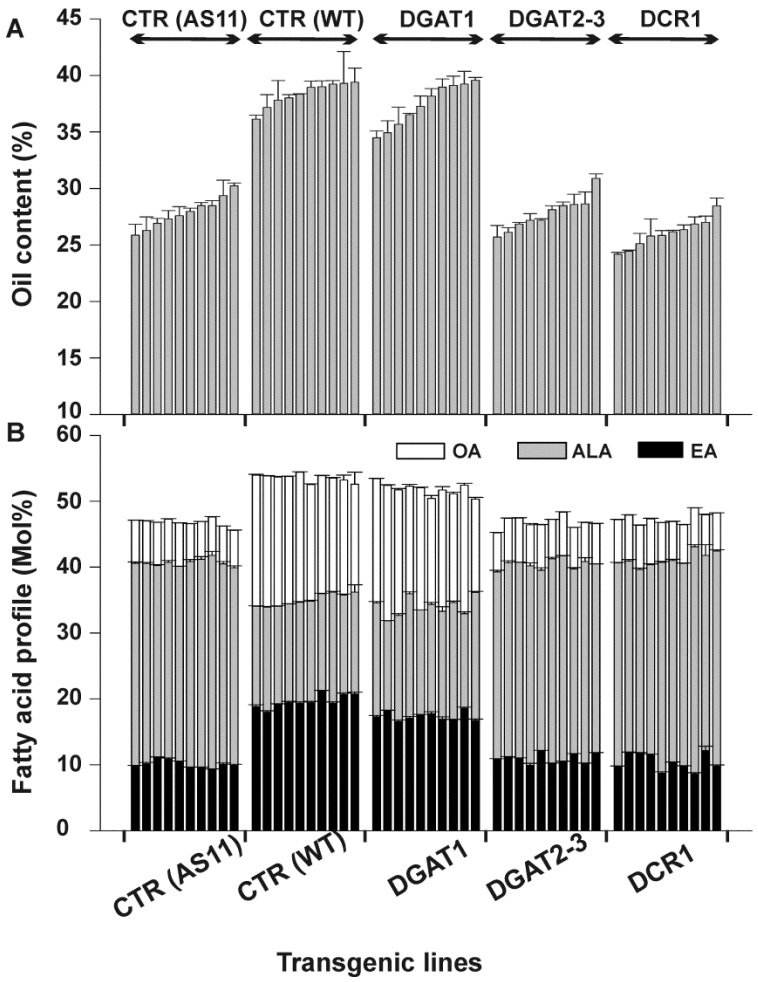
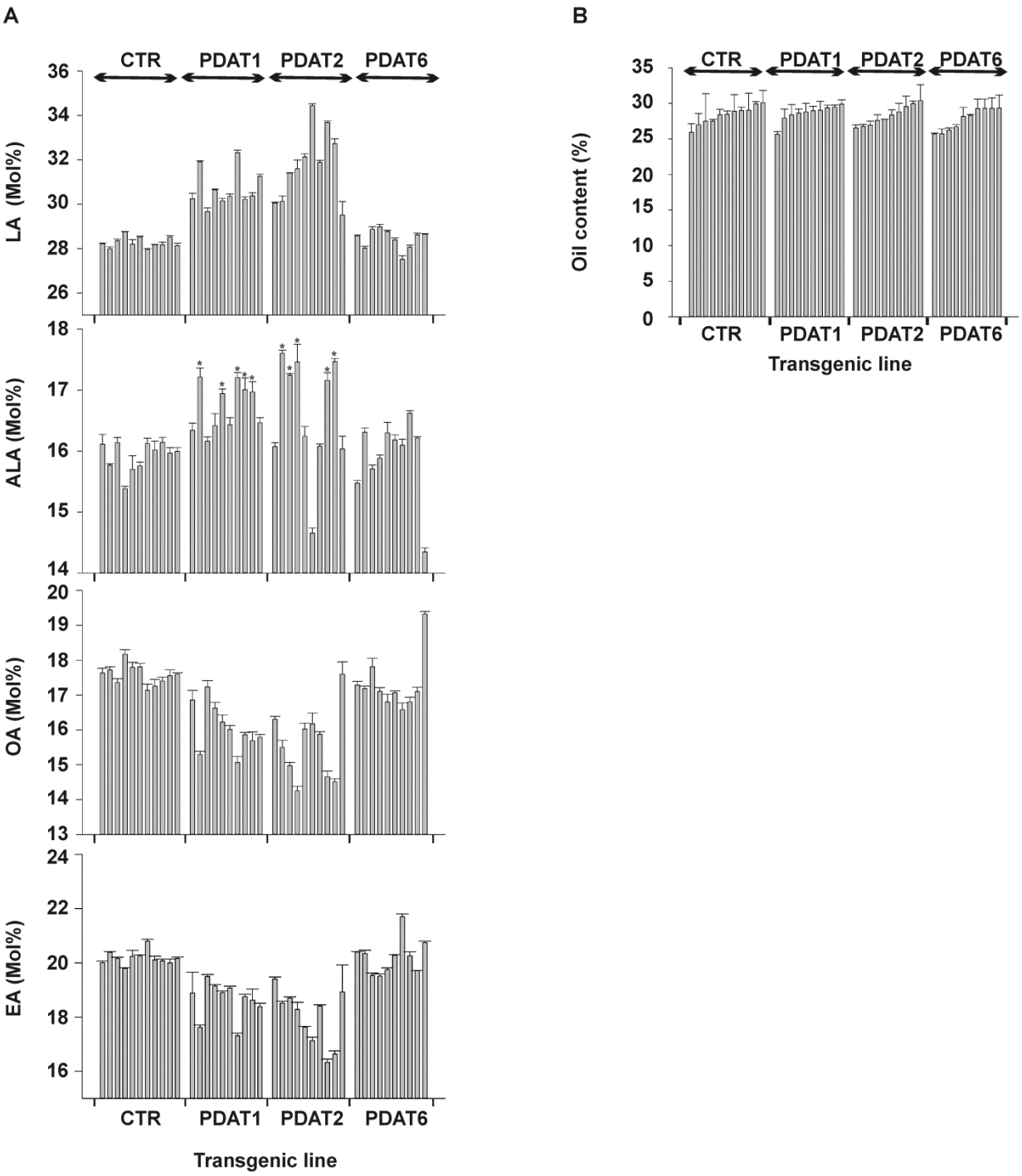

A
CTR (AS11)
CTR (WT)
DGAT1
DGAT2-3
DCR1
Oil content (%)
45
40
35
30
25
20
15
10
B
OA
ALA
EA
Fatty acid profile (Mol%)
60
50
40
30
20
10
0
CTR (AS11)
CTR (WT)
DGAT1
DGAT2-3
DCR1
Transgenic lines

Figure 3.16 Lipid phenotype of *LuPDAT*-overexpressing seeds. Ten individual transgenic lines were analyzed for each construct. Wild-type Arabidopsis transformed with empty pGreen plasmid were used as controls (CTR). (A) FA composition of *LuPDAT*-overexpressing seeds. Overexpression of *LuPDAT1* and *LuPDAT2* in Arabidopsis seeds results in an increased level of linoleic acid (LA) and α-linolenic acid (ALA). Data are presented as means ± SE (n = 3), with asterisks indicating $p<0.05$ (ANOVA, Duncan's multiple range test). OA – oleic acid; EA – eicosenoic acid. (B) Oil content of *LuPDAT*-overexpressing seeds. Data are presented as means ± SE (n = 3).


A
CTR
PDAT1
PDAT2
PDAT6
LA (Mol%)
ALA (Mol%)
OA (Mol%)
EA (Mol%)
Transgenic line
B
Oil content (%)
Transgenic line

3.4 Discussion

Flax seed is a prominent oilseed crop and one of the most important plant-based sources of vegetable oils rich in ALA. Understanding the metabolic pathway of oil synthesis is pivotal for a continuous improvement of flax as an oilseed crop. The goal of this study was to identify flax genes involved in TAG synthesis and functionally characterize the encoded enzymes.

In mainstream oilseed crops such as *Brassica napus* and *Glycine max*, DGAT1 has been proposed to be the dominant enzyme involved in the final step of TAG biosynthesis (Lock, et al., 2009; Li, et al., 2010). But in plant species that accumulate exotic FAs such as ricinoleic, vernolic and α-eleostearic acid, DGAT2 and PDAT appear to have a more predominant role in TAG synthesis although *DGAT1* is also expressed during seed development in these species (Shockey, et al., 2006; Kim, et al., 2011; Li, et al., 2010). In the case of flax, gene homologues encoding all major TAG-synthesizing enzymes were identified with the exception of DGAT3.

A previous study from our group cloned a DGAT1 from flax (here refereed to LuDGAT1-1) and has demonstrated that it can complement TAG synthesis in the yeast strain H1246 (Siloto, et al., 2009a). Although this enzyme was considerably active when produced recombinantly in yeast and plants (Figures 3.8A, 3.8D, 3.14 and 3.15), it did not appear to show a preference for substrates containing ALA (Figures 3.8D and 3.14). Like in other oilseeds, the expression of *LuDGAT1* was considerably up-regulated in seeds (Figure 3.5A), but a comprehensive analysis throughout seed development indicated that the transcript levels peaked at 20 DPA when the rate of oil accumulation reached a plateau and most ALA is already incorporated in TAG (Figures 3.7B to 3.7D). Together, these results suggested that flax DGAT1 appears to play a role in the non-specific transfer of acyl moieties to TAG during the final stages of seed development. During the time this manuscript was being prepared, I detected an

additional gene encoding DGAT1 (here referred to as *LuDGAT1-2*) in the final released version of flax genome sequence (Wang, et al., 2012). LuDGAT1-1 and LuDGAT1-2 share 97.7% identity and due to this high degree of homology, their expression was detected together (Figures 3.5A and 3.7D).

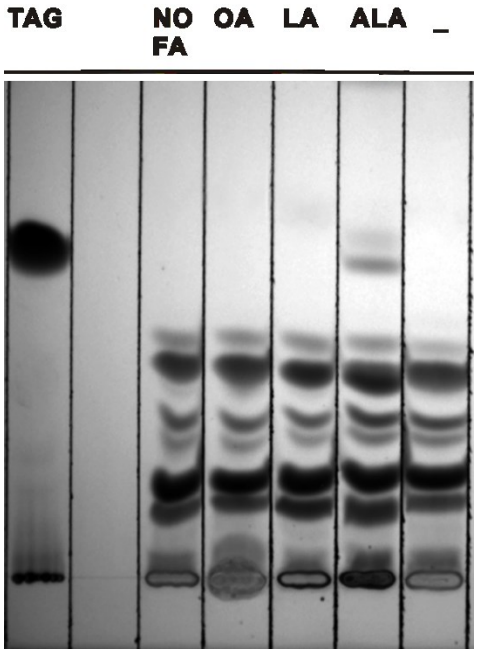
In contrast to DGAT1, flax DGAT2-3 failed to complement TAG synthesis in Arabidopsis AS11 mutant (Figure 3.15). Production of LuDGAT2-3 recombinant in yeast H1246 resulted in the production of small amounts neutral lipids in yeast (Figure 3.8C) and ineffectively rescued the FA lipotoxicity phenotype (Figure 3.12). The transcript of a third DGAT2 homologous gene (*LuDGAT2-2*) could not be detected or isolated. These results suggested that in flax DGAT2 might contribute minimally to TAG synthesis despite of being expressed specifically in seeds. Similarly, LuDCR1 did not show a significant effect on TAG biosynthesis in both yeast and Arabidopsis expression system (Figure 3.8 and 3.15).

In the case of flax PDATs, I analyzed one representative gene from each of the three groups of homologous genes (*LuPDAT1, LuPDAT2,* and *LuPDAT6*). Heterologous expression in yeast indicated that both LuPDAT1 and LuPDAT2 were active and able to produce TAG composed predominantly of ALA when ALA was supplemented as a free FA to the medium or endogenously produced through the action of desaturases LuFAD2-1 and LuFAD3B (Figures 3.8C and 3.14). Interestingly, LuPDAT2 has higher level of protein accumulation than LuPDAT1 in yeast (Figure 3.11A), which may account partially for its higher TAG-forming ability. The reason for poor expression levels of LuPDAT1 is currently unknown but may be due to the low mRNA stability or inefficient codon usage. Overexpression of *LuPDAT1* or *LuPDAT2* in Arabidopsis also resulted in an enhanced content of PUFAs (Figure 3.16A). Analysis of gene expression revealed that *LuPDAT1/LuPDAT5* transcripts accumulated substantially more in

seeds than in other vegetative tissues (Figure 3.6A). In fact, expression of *LuPDAT1/LuPDAT5* was closely correlated with oil and ALA accumulation in developing embryos, presenting a peak during the rapid phase of lipid accumulation at 14 DPA (Figure 3.7B, 3.7C and 3.7G). *LuPDAT2* was found to be mostly expressed in vegetative tissues (Figure 3.6D), indicating that TAG synthesized by the catalytic action of LuPDAT2 might have physiological functions in vegetative tissues of flax. *LuPDAT6* was also highly expressed during the initial stages of embryo development (Figure 3.7I) but the encoded enzyme had significantly lower activity in all conditions tested (Figures 3.8, 3.12, 3.13 and 3.16). Based on sequence homology (Table 3.4) and similarities in expression profile (Figure 3.6), it was suggested that the pairs *LuPDAT1/LuPDAT5*, *LuPDAT2/LuPDAT4*, and *LuPDAT3/LuPDAT6* may have similar functions in flax. The FA feeding experiment confirmed that LuPDAT5, similar to LuPDAT1, also has the higher preference for ALA-containing substrate (Figure 3.17). On the basis of these results, it was proposed that LuPDAT1/LuPDAT5 may play a more important role than LuDGAT1 in the production of ALA-containing TAG in developing flax seed. This hypothesis is in agreement with metabolic analyses of the acyl-CoA profile of developing flax seed which detected only small amounts of ALA-CoA in flax embryo (Ruiz-Lopez, et al., 2009).

The identified ALA-selective flax PDATs in this study could be considered as another example of a plant TAG-synthesizing enzyme that displays preference for certain substrates. The data obtained from this study indicates, however, that these PDATs not only favour substrates containing ALA, but also extend to other PUFAs that are not naturally formed in flax, including SDA, GLA and EPA (Figure 3.13). It appears that the activity of these PDATs is highly dependent on ALA availability. Increasing the concentration of exogenously provided ALA resulted in an enhanced amount of ALA-containing TAG and total TAG contents in yeast

Figure 3.17 Yeast cells expressing *LuPDAT5* show enhanced synthesis of triacylglycerol (TAG) in the presence of α-linolenic acid (ALA). Yeast cells transformed with pYES*LacZ* was used as the negative control, which was annotated as "-" on the thin layer chromatography (TLC) plate. TAG - trilinolenin TAG standard; FA – fatty acids; OA – oleic acid; LA – linolenic acid.



expressing *LuPDAT1* or *LuPDAT2* (Figure 3.11B to 3.11E). When ALA was added to the medium, both LuPDAT1 and LuPDAT2 were able to produce TAG with up to 90% ALA (Figure 3.11C). In contrast, when ALA was endogenously produced by the co-expression of *LuFAD2-1* and *LuFAD3B*, the total amount of ALA in TAG was only about 50% for LuPDAT1 or LuPDAT2 (Figure 3.14). This difference can be explained by two possible reasons. One possible explanation is that the ALA availability may limit the activity of these two PDATs. Inefficient desaturase activity in the heterologous yeast expression system has been recently reported by Dahmen et al. (2013) (Dahmen, et al., 2013). They found that yeast cytochrome *b*5, which transfers the electrons to the desaturase, poorly interacts with foreign desaturases, thus limiting desaturase enzymatic activity and resulting in the low production of unsaturated FAs. Future studies of the over-expression of flax cytochrome *b5* in the *LuFAD2-1-LuFAD3b-LuPDAT*-expressing yeast poses an opportunity to further enhance the production of ALA-containing TAG in yeast by metabolic engineering. The other possible reason is that to produce trilinolenin the flax PDATs require *sn*-2 ALA-PC and *sn*-1,2 ALA-DAG. Considering that the excess of ALA-CoA derived from the exogenous supplied ALA can be transferred indiscriminately into different glycerolipids, there is a better chance of obtaining *sn*-1,2 ALA-DAG. In contrast, ALA produced by a FAD3 is restricted to *sn*-2 PC, which can be utilized by PDAT thus limiting its redistribution to other lipid classes such as DAG. Therefore, the limited amount of *sn*-1,2 ALA-DAG may restrict the activity of LuPDAT1 and LuPDAT2. The same theory can be used to explain the small but significant increase of ALA content in the seed oil of *LuPDAT1*- and *LuPDAT2*-overexpressing Arabidopsis lines. It is possible that the activity of FAD3 in Arabidopsis might not be sufficiently high to supply the substrates required for LuPDAT1 and LuPDAT2. For instance, expression of *FAD3* in wild-type Arabidopsis is significantly lower

compared to the AS11 mutant (Xu, et al., 2012). Furthermore, labeling experiments indicated that Arabidopsis utilizes the PC-derived DAG pool as the major source for TAG biosynthesis (Bates and Browse, 2011). PC-derived DAG synthesis, however, may provide mostly *sn*-2-ALA-DAG. It is possible that when produced in Arabidopsis, LuPDAT1 and LuPDAT2 may rapidly transfer ALA from the *sn*-2 position of PC to the *sn*-3 position of the PC-derived DAG. This rapid conversion could limit the chance of rechanneling ALA from PC to acyl-CoA pool and subsequently incorporating it to the *sn*-1 position of DAG by the endogenous Arabidopsis enzymes, thus imposing a bottleneck for accumulating ALA in Arabidopsis seed oil.

A phylogenetic analysis of PDATs including the functionally tested enzymes from Arabidopsis (Zhang, et al., 2009) and castor (Kim, et al., 2011; van Erp, et al., 2011) indicate that LuPDAT1 and RcPDAT1-2 are closely related (Figure 3.3A). Interestingly, both polypeptides display remarkable substrate selectivity and the corresponding genes are preferentially expressed in seeds. The branch containing LuPDAT1 and RcPDAT1-2 lacks a corresponding polypeptide from Arabidopsis. LuPDAT2 is, however, more closely associated with AtPDAT1 and RcPDAT1-1. It is interesting that in Arabidopsis upon *AtDGAT1* disruption, AtPDAT1 is responsible for synthesizing the remaining 65-70% of TAG (Xu, et al., 2012) despite of the inability of AtPDAT1 or AtPDAT2 to complement TAG synthesis in H1246 (Zhang, et al., 2009). In addition, overexpression of *AtPDAT1* in wild-type Arabidopsis background did not produce a significant effect on oil phenotype (Ståhl, et al., 2004). Similarly, expression of *RcPDAT1-1* in Arabidopsis CL37 lines did not result in any substantial increase in hydroxy FAs (Kim, et al., 2011). Overexpression of *LuPDAT2* in wild-type Arabidopsis, however, led to a significantly increased PUFA level. This difference might be explained by the higher substrate selectivity of LuPDAT2. LuPDAT6 belongs to the same branch of AtPDAT2 and RcPDAT2.

Although *PDATs* in this branch are preferentially expressed in developing seeds, they do not have a significant effect on TAG accumulation. Overall, these findings suggest the existence of a class of plant PDATs that displays substrate selectivity properties.

One of the most remarkable advances in oilseed biotechnology is the successful reconstitution of VLC-PUFA synthesis in transgenic plants (reviewed by Ruiz-López et al. (Ruiz-López, et al., 2012). Flax would become an obvious candidate for this technology since it naturally contains substantial amounts of ALA that is the precursor for VLC-PUFAs. A previous report demonstrated that the introduction of a VLC-PUFA biosynthetic pathway in flax led to a high proportion of Δ6-desaturated C-18 FAs (GLA and SDA totaling up to 33% in some plants). The pathway, however, did not efficiently proceed further after the first desaturation step, resulting in very limited amount of AA and EPA (<5%) in flax oil (Abbadi, et al., 2004). As demonstrated before, the front-end desaturases and elongases involved in VLC-PUFA have different acyl carrier specificity (Domergue, et al., 2003). Desaturases act on PC, whereas elongases occur mainly within the acyl-CoA pool. According to the results presented in Figure 3.13, it is likely that LuPDAT1 contributes to a premature channeling of GLA and SDA to TAG, resulting in an insufficient supply of acyl-CoA substrates for elongation. The observation that GLA and SDA were nearly absent from the acyl-CoA pool of the transgenic flax lines (Abbadi, et al., 2004) further supports this hypothesis. To circumvent this problem, the metabolic pathway for VLC-PUFAs would have to be strategically modified to bypass other acyl moieties that could be utilized by LuPDAT1 until it reaches EPA. For example, the pathway involving the utilization of Δ8-desaturation pathway to convert LA to AA (Qi, et al., 2004) and then further desaturating AA to EPA by Δ17 desaturases (Pereira, et al., 2004), would contain intermediates (LA, DGLA and AA) which are poorly utilized by LuPDAT1 (Figures 3.8E and 3.13).

3.5 Conclusion

In conclusion, this study provided new insights into the biosynthesis of TAG in flax. The results demonstrate the existence of two pairs of flax PDATs that are highly selective for substrates containing ALA. Among them, one pair of *PDATs* (*LuPDAT1/LuPDAT5*) is expressed predominately in seeds. In addition, the oil and ALA accumulation during seed development closely correspond to the expression pattern of this seed-preferred *PDAT* gene pair, suggesting the critical role of these PDATs in seed oil biosynthesis in flax. The identified PDATs help in the understanding of the mechanisms involved in producing TAG with high PUFA content in plants. This characterization of TAG-synthesizing enzymes in flax will benefit future projects aimed at enriching PUFAs in plants and other organisms.

3.6 References

Abbadi A, Domergue F, Bauer J, Napier JA, Welti R, Zähringer U, Cirpus P, Heinz E (2004) Biosynthesis of very-long-chain polyunsaturated fatty acids in transgenic oilseeds: Constraints on their accumulation. Plant Cell **16:** 2734-2746

Ahmad A, Zhang Y, Cao XF (2010) Decoding the epigenetic language of plant development. Mol Plant **3:** 719-728

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Bates PD, Browse J (2011) The pathway of triacylglycerol synthesis through phosphatidylcholine in Arabidopsis produces a bottleneck for the accumulation of unusual fatty acids in transgenic seeds. Plant J **68:** 387-399

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Phys **37:** 911-917

Ciftci ON, Przybylski R, Rudzińska M (2012) Lipid components of flax, perilla, and chia seeds. Eur J Lipid Sci Tech **114:** 794-800

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Dahmen JL, Olsen R, Fahy D, Wallis JG, Browse J (2013) Cytochrome b5 Coexpression Increases Tetrahymena thermophila Delta6 Fatty Acid Desaturase Activity in Saccharomyces cerevisiae. Eukaryot Cell **12:** 923-931

Das UN (2006) Essential fatty acids - A review. Curr Pharm Biotechnol **7:** 467-482

Domergue F, Abbadi A, Ott C, Zank TK, Zähringer U, Heinz E (2003) Acyl carriers used as substrates by the desaturases and elongases involved in very long-chain polyunsaturated fatty acids biosynthesis reconstituted in yeast. J Biol Chem **278:** 35115-35126

Færgeman NJ, Black PN, Dan Zhao X, Knudsen J, DiRusso CC (2001) The Acyl-CoA synthetases encoded within FAA1 and FAA4 in Saccharomyces cerevisiae function as components of the fatty acid transport system linking import, activation, and intracellular utilization. J Biol Chem **276:** 37051-37059

Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. Methods Enzymol **498:** 349-361

Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc **2:** 31-34

Green AG (1986) A mutant genotype of flax (*Linum usitatissimum*L.) containing very low levels of linolenic acid in its seed oil. Can J Plant Sci **66:** 499-503

Hellens RP, Anne Edwards E, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: A versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol **42:** 819-832

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Huis R, Hawkins S, Neutelings G (2010) Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). BMC Plant Biol **10:** 71

Jhala AJ, Hall LM (2010) Flax (*Linum usitatissimum* L.): Current uses and future applications. Aust J Basic Appl Sci **4:** 4304-4312

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. J Biol Chem **278:** 8075-8082

Katavic V, Reed DW, Taylor DC, Giblin EM, Barton DL, Zou Jitao, MacKenzie SL, Covello PS, Kunst L (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. Plant Physiol **108:** 399-409

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Krasowska A, Dziadkowiec D, Polinceusz A, Plonka A, Łukaszewicz M (2007) Cloning of flax oleic fatty acid desaturase and its expression in yeast. J Am Oil Chem Soc **84:** 809-816

Kroon JTM, Wei W, Simon WJ, Slabas AR (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry **67:** 2541-2549

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Lee J, Welti R, Schapaugh WT, Trick HN (2011) Phospholipid and triacylglycerol profiles modified by PLD suppression in soybean seed. Plant Biotechnol J **9:** 359-372

Li R, Yu K, Hildebrand DF (2010) DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. Lipids **45:** 145-157

Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ (2012) Acyl-CoA:diacylglycerol acyltransferase: Molecular biology, biochemistry and biotechnology. Prog Lipid Res **51:** 350-377

Lock YY, Snyder CL, Zhu W, Siloto RMP, Weselake RJ, Shah S (2009) Antisense suppression of type 1 diacylglycerol acyltransferase adversely affects plant development in *Brassica napus*. Physiol Plantarum **137:** 61-71

Lu C, Xin Z, Ren Z, Miquel M, Browse J (2009) An enzyme regulating triacylglycerol composition is encoded by the *ROD1* gene of Arabidopsis. Proc Natl Acad Sci U S A **106:** 18837-18842

Lung SC, Weselake RJ (2006) Diacylglycerol acyltransferase: A key mediator of plant triacylglycerol synthesis. Lipids **41:** 1073-1088

Mhaske V, Beldjilali K, Ohlrogge J, Pollard M (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (*At5g13640*). Plant Physiol Bioch **43:** 413-417

Mietkiewska E, Siloto RMP, Dewald J, Shah S, Brindley DN, Weselake RJ (2011) Lipins from plants are phosphatidate phosphatases that restore lipid synthesis in a pah1δ mutant strain of *Saccharomyces cerevisiae*. FEBS J **278:** 764-775

Pereira SL, Huang YS, Bobik EG, Kinney AJ, Stecca KL, Packer JCL, Mukerji P (2004) A novel ω3-fatty acid desaturase involved in the biosynthesis of eicosapentaenoic acid. Biochem J **378:** 665-671

Petschnigg J, Wolinski H, Kolb D, Zelling G, Kurat CF, Natter K, Kohlwein SD (2009) Good fat, essential cellular requirements for triacylglycerol synthesis to maintain membrane homeostasis in yeast. J Biol Chem **284:** 30981-30993

Qi B, Fraser T, Mugford S, Dobson G, Sayanova O, Butler J, Napier JA, Stobart AK, Lazarus CM (2004) Production of very long chain polyunsaturated omega-3 and omega-6 fatty acids in plants. Nat Biotechnol **22:** 739-745

Rao S, Abdel-Reheem M, Bhella R, McCracken C, Hildebrand D (2008) Characteristics of high α-linolenic acid accumulation in seed oils. Lipids **43:** 749-755

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

Rowland GG (1991) An EMS-induced low-linolenic-acid mutant in McGregor flax (*Linum usitatissimum* L.). Can J Plant Sci **71:** 393-396

Ruiz-Lopez N, Haslam RP, Venegas-Caleron M, Larson TR, Graham IA, Napier JA, Sayanova O (2009) The synthesis and accumulation of stearidonic acid in transgenic plants: a novel source of 'heart-healthy' omega-3 fatty acids. Plant Biotechnol J **7:** 704-716

Ruiz-López N, Sayanova O, Napier JA, Haslam RP (2012) Metabolic engineering of the omega-3 long chain polyunsaturated fatty acid biosynthetic pathway into transgenic plants. J Exp Bot **63:** 2397-2410

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 Physiol **141:** 1533-1543

Sandager L, Gustavsson MH, Ståhl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H, Stymne S (2002) Storage lipid synthesis is non-essential in yeast. J Biol Chem **277:** 6478-6482

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. Plant Cell **18:** 2294-2313

Siloto RMP, Truksa M, He X, McKeon T, Weselake RJ (2009a) Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system. Lipids **44:** 963-973

Siloto RMP, Truksa M, Brownfield D, Good AG, Weselake RJ (2009b) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and characterization of *Brassica napus* DGAT1 mutagenized libraries. Plant Physiol Bioch **47:** 456-461

Sinclair AJ, Attar-Bashi NM, Lib D (2003) What is the role of alpha-linolenic acid for mammals? Lipids **37:** 1113-1123

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Stymne S, Stobart AK (1984) Evidence for the reversibility of the acyl-CoA:lysophosphatidylcholine acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons and rat liver. Biochem J **223:** 305-314

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. Mol Biol Evol **28:** 2731-2739

Tang J, Scarth R, Fristensky B (2003) Effects of genomic position and copy number of Acyl-ACP thioesterase transgenes on the level of the target fatty acids in Brassica napus L. Mol Breed **12:** 71-81

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

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

Vrinten P, Hu Z, Munchinsky MA, Rowland G, Qiu X (2005) Two FAD3 desaturase genes control the level of linolenic acid in flax seed. Plant Physiol **139:** 79-87

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, Lambert G, Galbraith DW, Grassa CJ, Geraldes A, Cronk QC, Cullis C, Dash PK, Kumar PA, Cloutier S, Sharpe AG, Wong GK, Wang J, Deyholos MK (2012) The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J **72:** 461-473

Weigel D, Glazebrook J (2002) Arabidopsis: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Xu J, Carlsson AS, Francis T, Zhang M, Hoffman T, Giblin ME, Taylor DC (2012) Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. BMC Plant Biol **12:** 4

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong GY, Tarczynski MC, Shen B (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat Genet **40:** 367-372

Chapter 4

Genome-wide Analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* Genes in Plants Reveals the Eudicot-wide *PDAT* Gene Expansion and Altered Selective Pressures Acting on the Core Eudicot *PDAT* Paralogs

4.1 Introduction

Plant oils with a broad variety of food and industrial applications are important agricultural commodities. Triacylglycerols (TAGs), glycerol esters of fatty acids (FAs), are the major components of plant oils in nearly all commercially important crops. In the last decade, the TAG biosynthetic pathway has received substantial attention from research community. In 1960, using chicken (*Gallus gallus*) liver as an enzyme source, Eugene P. Kennedy and his graduate student Samuel B. Weiss demonstrated the use of diacylglycerol (DAG) and fatty acyl-CoAs as substrates for TAG biosynthesis (Weiss et al., 1960). The enzyme catalyzing this reaction is named acyl-CoA:diacylglycerol acyltransferase (DGAT, EC 2.3.1.20). The elucidation of the DGAT reaction laid the groundwork for numerous studies on the role of DGAT in TAG biosynthesis. The molecular identity of DGAT, however, remained elusive until 1998, when Cases and colleagues isolated the first *DGAT* gene in mouse (*Mus musculus*) by homology searches of the expressed sequence tag (EST) databases using the coding sequences from acyl-CoA:cholesterol acyltransferase (ACAT, EC 2.3.1.26), an enzyme that catalyzes the synthesis of cholesteryl esters (Cases et al., 1998). At least six types of *DGAT* genes (*DGAT1*, *DGAT2*, *DGAT3*, *DEFECTIVE IN CUTICULAR RIDGES*, *DIACYLGLYCEROL ACETYLTRANSFERASE* and a bifunctional DGAT/wax synthase) have been identified in plants (Saha et al., 2006; Hobbs et al., 1999; Lardizabal et al., 2001; Durrett et al., 2010; Rani et al., 2010; Li et al., 2008). The biochemistry, molecular biology, evolutionary relationship, structure-function relationship, and

biotechnological applications of DGAT have been discussed in some recent publications (Liu et al., 2012; Cao, 2011; Turchetto-Zolet et al., 2011).

Recently, it has become obvious that TAG synthesis can also be catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT, EC 2.3.1.158). PDAT enzyme activity was first described in microsomal preparations from sunflower (*Helianthus annuus*), castor (*Ricinus communis*), and *Crepis palaestina* by Stymne and co-workers during their observation of the use of phospholipids as acyl donor and DAG as acceptor for TAG biosynthesis (Dahlqvist et al., 2000). They further found that PDAT activity is also present in yeast (*Saccharomyces cerevisiae*) and identified the first *PDAT* gene (*YNR008w, LRO1*) from yeast as a homolog of human lecithin:cholesterol acyltransferase (LCAT, *EC* 2.3.1.43). LCAT is a soluble acyltransferase that catalyzes cholesteryl ester synthesis in blood plasma. Knowledge of the yeast *PDAT* sequence led to the discovery of two *PDAT* orthologs in Arabidopsis (*Arabidopsis thaliana*), referred to as *AthPDAT1* (At5g13640) and *AthPDAT2* (At3g44830) (Ståhl et al., 2004). Although the RNA interference (RNAi)-based approach provides evidence that AthPDAT1 and AthDGAT1 have an overlapping function for TAG biosynthesis in both seed and pollen (Zhang et al., 2009), overexpression or knockout of *AthPDAT1* in Arabidopsis only led to significant changes in oil phenotype (oil content and FA composition) in developing leaves (Fan et al., 2013), but not in seeds (Mhaske et al., 2005). The other ortholog, AthPDAT2, has no role in TAG biosynthesis, even though the gene encoding PDAT2 is highly expressed in seeds. In castor, three *PDAT* orthologs have been identified (van Erp et al., 2011; Kim et al., 2011). One particular PDAT, RcoPDAT1A, appears to be ricinoleic-specific; seed-specific overexpressing this *PDAT* in Arabidopsis resulted in an enhanced proportion of hydroxy FAs in the seed oil. It was recently reported that flax (*Linum usitatissimum*) contains six PDATs (Pan et al., 2013). Four out of the

six PDATs (LusPDAT1/LusPDAT5 and LusPDAT2/LusPDAT4) have the unique ability to preferentially channel α-linolenic acid (ALA) into TAG, whereas another two PDATs (LusPDAT3/LusPDAT6) do not show TAG synthesizing ability. In addition to the PDATs from higher plants, a single PDAT with multiple catalytic functions has been characterized in the unicellular green alga *Chlamydomonas reinhardtii* (Yoon et al., 2012). It is worth noting that the PDAT-mediated TAG forming mechanism has also been detected in bacteria, *Streptomyces coelicolor* (Arabolaza et al., 2008), but it has no counterpart in mammals.

These previous studies reveal that: 1) *PDAT* can exist as multiple copies in plant genomes; 2) different *PDAT* gene paralogs can encode enzyme with different TAG-synthesizing ability; and 3) certain PDATs can have unique substrate selectivity. All these findings shed new light on TAG-biosynthetic mechanisms in plants and highlight the need for a deeper understanding of the complexity of plant PDATs. In this study, I have sought to provide further insights into the present-day diversity and ortholog/paralog relationship of plant *PDATs* via a genome-wide comparative analysis.

4.2 Materials and Methods

4.2.1 Identification of *PDAT* genes and their homologs in plants

To identify *PDAT* genes and their homologs, the TBLASTN search was performed using the Arabidopsis PDAT1 (AthPDAT1) and PDAT2 (AthPDAT2) protein sequences as queries against the Phytozome databases (http://www.phytozome.net/). All 39 listed plant species, including algae (*Chlamydomonas reinhardtii*, *Volvox carteri*, *Coccomyxa subellipsoidea C-169*, *Micromonas pusilla CCMP1545* and *Ostreococcus lucimarinus*), a lycophyte *Selaginella moellendorffii*, a moss *Physcomitrella patens*, monocots (*Sorghum bicolor* v1.4, *Zea mays*, *Setaria italica*, *Panicum virgatum*, *Oryza sativa* and *Brachypodium distachyon*), and eudicots

(*Manihot esculenta*, *Ricinus communis*, *Linum usitatissimum*, *Populus trichocarpa*, *Medicago truncatula*, *Phaseolus vulgaris*, *Glycine max*, *Cucumis sativus*, *Prunus persica*, *Malus domestica*, *Fragaria vesca*, *Arabidopsis lyrata*, *Capsella rubella, Brassica rapa*, *Thellungiella halophila*, *Carica papaya*, *Gossypium raimondii*, *Theobroma cacao*, *Citrus sinensis*, *Citrus clementina*, *Eucalyptus grandis*, *Vitis vinifera*, *Solanum tuberosum*, *Solanum lycopersicum*, *Mimulus guttatus* v1.1 and *Aquilegia coerulea Goldsmith*), were included in the analysis. Because the predicted transcripts for *Medicago truncatula* genome are not available in the Phytozome database, the LegumeIP database was used instead for this species (http://plantgrn.noble.org/LegumeIP/blast.do). The cDNA, genomic DNA and amino acid sequences corresponding to each PDAT or putative PDAT were downloaded from the Phytozome database. The theoretical molecular mass (*m)* and isoelectric point (pI) values were calculated using the Compute pI/Mw tool provided in ExPASy http://web.expasy.org/compute_pi/. For the InterPro domain analysis, all candidate sequences (without ending asterisk symbols) were scanned with InterProScan version 5 (Jones et al., 2014), which was installed locally in a 32-bit Red Hat Linux environment. The default parameters were used and its InterPro lookup option (iprlookup) was turned on to generate InterPro annotation. For protein classification, all sequences were subjected to Pfam (Punta et al., 2012; http://pfam.sanger.ac.uk/search) and PANTHER classification systems (Mi et al., 2013; http://www.pantherdb.org/). All *Taxa* were indicated by three-letter acronyms in which the first letter is the first letter of the genus and the next two letters are the first two letters of the species name (e.g., Ath corresponds to *Arabidopsis thaliana*). Extra numbers were added after taxon names to indicate individual gene copies. To avoid the confusion, the names for the previously

reported Arabidopsis (*A. thaliana)*, castor (*R. communis*), flax (*L. usitatissimum*) and alga (*C. reinhardtii*) *PDATs* followed the published names.

4.2.2 Construction of phylogenetic trees

The full-length nucleotide sequences were aligned based on their corresponding amino acid translations using TranslatorX server (Abascal et al., 2010; http://translatorx.co.uk/). Then, jModelTest 0.1.1 analysis (Posada, 2008) was carried out to select the best-fit model under the Akaike Information Criterion (AIC) framework (Akaike, 1974). The result of jModelTest indicates that the best-fit substitution model to determine the evolution for all data sets is the General Time Reversible (GTR) model with the shape of the gamma distribution (G) plus the proportion of invariable sites (I). According to the best-fit model, maximum likelihood (ML) phylogenetic analysis was constructed via the CIPRES Web Portal http://www.phylo.org/portal/Home using MrBayes 3.2.2 (Huelsenbeck and Ronquist, 2001) with 1,000,000 generations, four Markov chains and two runs. The first 25% of tree from all runs were discarded as burn-in. To verify the reliability of phylogenetic analysis, a maximum likelihood tree was also performed using online program RAxML (Randomized Axelerated Maximum Likelihood) (Stamatakis, 2006; http://www.trex.uqam.ca/index.php?action=raxml&project=trex) under the best-fit model with 100 bootstrap samples. The phylogenetic tree was visualized using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). The same methods were used to carry out all phylogenetic analyses included in this study. Sequence alignments for all analyses used in the phylogenetic construction are provided as Supplemental Datasets 1-3.

4.2.3 Gene structure analysis

Intron/exon distribution and intron phase patterns were analyzed using the online Gene Structure Display Server (Guo et al., 2007; http://gsds.cbi.pku.edu.cn).

4.2.4 Detection of transmembrane domains and conserved motifs

The potential transmembrane domains in PDATs were predicted using the TMHMM (Krogh et al., 2001) program provided by the CBS Prediction Servers (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Functional motifs of PDAT proteins were identified using the MEME (Multiple Expectation Maximization for Motif Elicitation) program (Bailey and Elkan, 1994; http://meme.nbcr.net/meme/cgi-bin/meme.cgi) with the following parameters: the distribution of motifs = any number of repetitions, maximum number of motifs = 100, and optimum motif width = 3 to 300 residues. The identified motifs were further subjected to Pfam analysis for protein classification (Punta et al., 2012; http://pfam.sanger.ac.uk/search).

4.2.5 Gene expansion pattern and selective pressure analysis

Tandem duplication was identified as multiple gene family members clustering within a 200kb region of a chromosome (Houb, 2001). The chromosomal locations of *PDAT* genes were determined using the Phytozome's GBrowse genome browser.

For calculating the synonymous substitution rates (*Ks*), amino acid sequences representing the duplicated *PDATs* were aligned using ClustalW (Thompson et al., 1994) implemented in Geneious Pro 5.3.6 (Drummond et al., 2013), and the obtained protein alignments were used to guide the conversion of the corresponding cDNA sequences into the codon alignments via PAL2NAL (Suyama et al., 2006; http://www.bork.embl.de/pal2nal/). The resulting codon alignments were imported into the codon substitution model (CodeML)

implemented in the PAML v4.4c software package (Yang, 1997) for *Ks* calculation. The Goldman and Yang maximum likelihood method and the F3x4 model were used in analyses.

The selective pressure operating on the core eudicot *PDATs* was estimated using the ratio (ω) of the nonsynonymous substitution rate (K_N) versus the synonymous substitution rate (K_S) as an indicator (Anisimova and Kosiol, 2009; Yang and Bielawski, 2000): $0< \omega <1$ corresponds to purifying selection, $\omega =1$ indicates neutral selection, and $\omega >1$ suggests positive selection. The estimation of the ω ratio was performed using the CodeML program within the PAML package. Nucleotide alignments were generated using the TranslatorX. Phylogenetic analyses of the core eudicot *PDATs* were performed using RAxML and the resulting trees without branch lengths were used as input trees for the simple one-ratio model (M0, model = 0 and NS site = 0) analyses. The trees with branch lengths generated by M0 analyses were further used to investigate functional divergence. To test divergent selective pressures among the core eudicot clades, I used the Clade model C (CmC, model = 3 and NS site = 2) of Bielawski and Yang (Bielawski and Yang, 2004) as modified by Yang, Wong, and Nielsen (Yang et al., 2005). The CmC assumes that the phylogeny can be divided into the foreground and background partitions. For each analysis, the clade of interest (all branches within the clade) was selected as the foreground partition and the remaining phylogeny was set as the background partition. The CmC contains three site classes across the entire phylogeny: site class 0 is under purifying selection ($0< \omega_0 <1$); site class 1 is under neutral selection ($\omega_1 =1$); site class 2 is the divergent site class where independent ω is estimated to the background ($\omega_2 > 0$) and foreground partitions ($\omega_3 > 0$). The null model M2a_rel hypothesizing the same ω between the foreground and background partitions also has three site classes. The first two site classes are the same as the ones in CmC, while a third site class is represented by a single ω ratio for all branches across the phylogeny (ω_2

> 0). Likelihood ratio tests (LRTs) were used to compare the fit of the CmC against the null model M2a_rel (Weadick and Chang, 2012). LRTs were performed by comparing twice the difference in *ln* likelihood scores of CmC and M2a_rel against a χ^2distribution with the degree of freedom equal to the difference in the number of parameters between the two models. The data set was run multiple times with different initial ω values to avoid local optima. Like CmC, extended clade model (Yoshida et al., 2011) also assumes three site classes. The first two site classes are the same as CmC model. The final class (site class 2) allows to model divergent selection for more than two phylogeny partitions, each with a separately estimated ratio. I specified three partitions in this analysis (Clade V, Clade VI and Clade VII) and three separate ω ratios were obtained for the three tree partitions (ω_2 for CladeV, ω_3 for Clade VI, and ω_4 for Clade VII). The null model has only two partitions, with ω_2 for Clade VII and ω_3 for both Clade V and Clade VI. LRTs were used to compare the fit of the extended clade model against the null model.

4.3 Results

4.3.1 Identification of the *PDAT* gene family in plants

The growing number of fully sequenced plant genomes makes it possible to perform a comparative genomic analysis of the *PDAT* gene family across a wide range of plant species. To identify PDATs in different plant species, a genome-wide search was performed using both Arabidopsis PDAT1 (AthPDAT1) and PDAT2 (AthPDAT2) amino acid sequences as queries to BLAST against 40 genomes listed in the Phytozome database. Candidate *PDAT* genes were found in all examined plant genomes, including algae, lowland plants (a moss and a lycophyte), monocots and eudicots. Multiple hits were identified in each of land plant genomes, with the exception of *Brachypodium distachyon*. Only one hit was identified in each of algal genomes. In

total, 139 sequences were identified and sequence information is provided in Table 4.1. Among the 139 sequences, six from five species (*Solanum tuberosum*, *Populus trichocarpa*, *Medicago truncatula*, *Malus domestica* and *Arabidopsis lyrata*) encode less than 200 amino acid residues, which is most likely due to genome annotation errors (Table 4.2). These short sequences were eliminated from further analysis. In addition, the predicted transcripts from *Malus domestica* genome have multiple stop codons and the predicted transcript from *Ostreococcus lucimarinus* genome does not start with a start codon, thus the sequences from these two species were excluded. In the end, a total of 128 sequences were included for the analysis. To verify the reliability of BLAST results, these 128 protein sequences were subjected to InterPro and Pfam analysis, and all of them were classified into the LCAT family (Pfam: 02450).

A previous study (Yoon et al. 2012) showed that the LCAT-like family proteins from plants can be divided into four major groups, including PDAT, LCAT, phospholipid:sterol acyltransferase (PSAT), and phospholipase A (PLA) proteins. Therefore, some candidate sequences identified by BLAST may not encode PDAT. To clarify if the sequences obtained from BLAST are *PDAT* genes, phylogenetic analysis of the 128 full-length LCAT-like gene sequences was carried out. The maximum likelihood (ML) trees (Figure 4.1 and Figure 4.2) shows that all algal candidates are grouped together into a single clade (algal group), whereas the land plant sequences are partitioned into four major clades, designated as Groups A, B, C and D, with 91, 3, 4 and 26 identified sequences, respectively. The sequences with an expectation value (E-value) < 1e-15 fall into Group A, while the remaining ones (E-value > 1e-15) branch into Groups B, C and D. Group B is more closely related to Group A than either Group C or Group D is. Concerning the genes already characterized, Group A contains all genes that are previously experimentally characterized as PDAT, including Arabidopsis, flax and castor PDATs (Ståhl et

al., 2004; Zhang et al., 2009; van Erp et al., 2011; Kim et al., 2011; Pan et al., 2013). The AthLCAT-like2 (AT1G04010) (Banas et al., 2005) from Group B was previously identified and experimentally characterized as PSAT. Two MtrLCAT-like sequences (Medtr7g080450 and Medtr4g083980) from Group C were previously predicted to be PLA while the AthLCAT-like1 sequence (AT1G27480) from Group D was classified as LCAT (Yoon et al. 2012).

Next, all protein sequences were classified using the InterPro and PANTHER (Protein Analysis Through Evolutionary Relationships) classification system. The results show that all sequences are classified as LCAT-related proteins (PTHR: 11440). PANTHER subfamily classification further reveals that the algal sequences and land plant sequences in Group A belong to the PDAT subfamily (PTHR:11440: SF4), while the land plant sequences in Group B and Group C are classified into PSAT subfamily (PTHR:11440:SF7) and LCAT-like 4-related subfamily (PTHR:11440:SF3), respectively. The sequences in Group D have no PANTHER subfamily classification.

The phylogeny result combined with the PANTHER classification suggests that the genes in Groups B, C and D very likely encode PSAT, PLA, and LCAT, respectively, rather than PDAT. Therefore, the sequences from Group A are named as PDAT while the sequences from Groups B, C and D are named as LCAT-like sequences. Only the sequences from Group A (PDAT group) and the algal clade were included for further analyses. To avoid incomplete sampling of PDAT paralogs within species, the species (*Solanum tuberosum*, *Populus trichocarpa, Medicago truncatula* and *Arabidopsis lyrata*) with short candidate PDATs were eliminated. In the end, 86 full-length PDAT candidates from 34 species were selected for further analyses. The obvious annotation errors, including incorrect stop codon predictions and splicing

errors, in six of the final selected sequences were manually corrected based on the EST database and the intron phases of closely related homologs (Table 4.2).

Table 4.1 Gene identifiers of protein sequences used in this study. Gene identifiers are obtained from Phytozome database or Arabidopsis Genome Initiative or EMBL/GenBank database.

Species	PDAT	Gene Identifier
Manihot esculenta	MesPDAT1	cassava4.1_003049m
	MesPDAT2	cassava4.1_002997m
	MesPDAT3	cassava4.1_002951m
	MesPDAT4	cassava4.1_034030m
	MesPDAT5	cassava4.1_002907m
	MesLCAT-like	cassava4.1_008172m
Ricinus communis	RcoPDAT1A	XP_002521350.1
	RcoPDAT1B	AEJ32006.1
	RcoPDAT2	HM807522
Linum usitatissimum	LusPDAT1	KC437085
	LusPDAT2	KC437086
	LusPDAT3	Lus10019519
	LusPDAT4	Lus10015639
	LusPDAT5	Lus10017165
	LusPDAT6	KC437087
	LusLCAT-like1	Lus10007728
	LusLCAT-like2	Lus10018664
Populus trichocarpa	PtrPDAT1	Potri.003G063100
	PtrPDAT2	Potri.001G171000
	PtrPDAT3	Potri.004G190100
	PtrPDAT4	Potri.009G150800
	PtrLCAT-like	Potri.014G014400
Medicago truncatula	MtrPDAT1	Medtr6g045280
	MtrPDAT2	Medtr8g104960
	MtrLCAT-like1	Medtr7g098010
	MtrLCAT-like2	Medtr4g083980
	MtrLCAT-like3	Medtr7g080450
Phaseolus vulgaris	PvuPDAT1	Phvul.011G088800
	PvuPDAT2	Phvul.003G133000

	PvuPDAT3	Phvul.010G132900
	PvuLCAT-like	Phvul.009G073200
Glycine max	GmaPDAT1	Glyma16g00790
	GmaPDAT2	Glyma07g04080
	GmaPDAT3	Glyma13g16790
	GmaPDAT4	Glyma17g05910
	GmaPDAT5	Glyma11g19570
	GmaPDAT6	Glyma12g08915
Cucumis sativus	CsaPDAT1	Cucsa.385510
	CsaPDAT2	Cucsa.360360
	CsaLCAT-like	Cucsa.325830
Prunus persica	PpePDAT1	ppa015987m
	PpePDAT2	ppa002500m
	PpeLCAT-like	ppa015352m
Malus domestica	MdoPDAT1	MDP0000537488
	MdoPDAT2	MDP0000311594
	MdoPDAT3	MDP0000293828
	MdoPDAT4	MDP0000575586
	MdoLCAT-like	MDP0000911376
Fragaria vesca	FvePDAT1	mrna14467.1-v1.0-hybrid
	FvePDAT2	mrna12713.1-v1.0-hybrid
	FveLCAT-like	mrna02603.1-v1.0-hybrid
Arabidopsis thaliana	AthPDAT1	AT5G13640
	AthPDAT2	AT3G44830
	AthLCAT-like1	AT1G27480
	AthLCAT-like2	AT1G04010
Arabidopsis lyrata	AlyPDAT1	488184
	AlyPDAT2	323168
	AlyPDAT3	935204
	AlyLCAT-like	921786
Capsella rubella	CruPDAT1	Carubv10002446m
	CruPDAT2	Carubv10019048m
	CruLCAT-like	Carubv10009090m
Brassica rapa	BraPDAT1	Bra008812
	BraPDAT2	Bra023426
	BraLCAT-like	Bra030052
Thellungiella halophila	ThaPDAT1	Thhalv10012888m
	ThaPDAT2	Thhalv10002852m
	ThaLCAT-like	Thhalv10007681m

Carica papaya	CpaPDAT1	evm.TU.supercontig_28.93
	CpaPDAT2	evm.TU.supercontig_25.98
	CpaPDAT3	evm.TU.supercontig_1971.1
Gossypium raimondii	GraPDAT1	Gorai.006G095300
	GraPDAT2	Gorai.001G118800
	GraPDAT3	Gorai.013G258900
	GraPDAT4	Gorai.004G284600
	GraPDAT5	Gorai.010G234300
	GraPDAT6	Gorai.010G086400
	GraLCAT-like	Gorai.009G221300
Theobroma cacao	TcaPDAT1	Thecc1EG029105
	TcaPDAT2	Thecc1EG037574
	TcaPDAT3	Thecc1EG041539
	TcaLCAT-like	Thecc1EG034671
Citrus sinensis	CsiPDAT1	orange1.1g005835m
	CsiPDAT2	orange1.1g005950m
	CsiPDAT3	orange1.1g005610m
Citrus clementina	CclPDAT1	Ciclev10027727m
	CclPDAT2	Ciclev10025070m
	CclPDAT3	Ciclev10025078m
	CclLCAT-like	Ciclev10001149m
Eucalyptus grandis	EgrPDAT1	Eucgr.J02832
	EgrPDAT2	Eucgr.J00412
	EgrPDAT3	Eucgr.C00841
Vitis vinifera	VviPDAT1	GSVIVG01031419001
	VviPDAT2	GSVIVG01016784001
	VviPDAT3	GSVIVG01015119001
	VviLCAT-like	GSVIVT01009315001
Solanum lycopersicum	SlyPDAT1	Solyc11g066710.1
	SlyPDAT2	Solyc07g041210.2
	SlyPDAT3	Solyc03g121960.2
	SlyPDAT4	Solyc06g062870.2
	SlyPDAT5	Solyc06g074680.2
	SlyLCAT-like	Solyc05g050710.2
Solanum tuberosum	StuPDAT1	PGSC0003DMG400002483
	StuPDAT2	PGSC0003DMG400000466
	StuPDAT3	PGSC0003DMG400007127
	StuPDAT4	PGSC0003DMG400020953
	StuPDAT5	PGSC0003DMG400004889

	StuLCAT-like	PGSC0003DMG400007000
Mimulus guttatus	MguPDAT1	mgv1a002396m
	MguPDAT2	mgv1a002483m
	MguPDAT3	mgv1a002508m
	MguPDAT4	mgv1a002273m
	MguLCAT-like	mgv1a006497m
Aquilegia coerulea	AcoPDAT1	Aquca_019_00083
	AcoPDAT2	Aquca_020_00522
	AcoLCAT-like	Aquca_060_00071
Sorghum bicolor	SbiPDAT	Sb02g026010
	SbiLCAT-like	Sb03g002720
Zea mays	ZmaPDAT1	GRMZM2G061885
	ZmaPDAT2	GRMZM2G095763
Setaria italica	SitPDAT	Si029118m
	SitLCAT-like	Si035645m
Panicum virgatum	PviPDAT1	Pavirv00024391m
	PviPDAT2	Pavirv00022530m
Oryza sativa	OsaPDAT	LOC_Os09g27210
	OsaLCAT-like1	LOC_Os03g52010
	OsaLCAT-like2	LOC_Os01g71800
Brachypodium distachyon	BdiPDAT	Bradi4g31540
Selaginella moellendorffii	SmoPDAT	164869
	SmoLCAT-like1	172410
	SmoLCAT-like2	85956
Physcomitrella patens	PpaPDAT	Pp1s336_57V6
	PpaLCAT-like1	Pp1s229_7V6
	PpaLCAT-like2	Pp1s324_50V6
Chlamydomonas reinhardtii	CrePDAT	XM_001699696.1
Volvox carteri	VcaPDAT	Vocar20014505m.g
Coccomyxa subellipsoidea C-169	CsuPDAT	fgenesh1_pm.11_#_66
Micromonas pusilla CCMP1545	MpuCMPPDAT	MicpuC2.EuGene.0000060413
Ostreococcus lucimarinus	OluPDAT	gwEuk.9.94.1
Saccharomyces cerevisiae	ScePDAT	NM_001183185

Table 4.2 Gene identifiers from Phytozome database for genes with annotation errors

Annotation Error	Species	Gene Identifier	Gene name used in this study
Short sequences (<200 amino acid residues)	*Populus trichocarpa*	Potri.009G150800	PtrPDAT4
	Medicago truncatula	Medtr6g045280	MtrPDAT1
	Medicago truncatula	Medtr8g104960	MtrPDAT2
	Malus domestica	MDP0000575586	MdoPDAT4
	Arabidopsis lyrata	935204	AlyPDAT3
	Solanum tuberosum	PGSC0003DMG400004889	StuPDAT5
Incorrect stop codon prediction	*Glycine max*	Glyma12g08915	GmaPDAT6
	Citrus clementina	Ciclev10027727m	CclPDAT1
	Vitis vinifera	GSVIVG01031419001	VviPDAT1
	Solanum lycopersicum	Solyc07g041210.2	SlyPDAT2
Splicing Error	*Fragaria vesca*	gene12713-v1.0-hybrid	FvePDAT2
	Solanum lycopersicum	Solyc11g066710.1	SlyPDAT1

Figure 4.1 Cladogram of 128 LCAT-like sequences from 38 plant species. The maximum likelihood tree was generated using the MrBayes program, and the support values along branches are Bayesian posterior probabilities (shown as percentage). The land plant LCAT-like sequences are grouped into four major clades, designated as Groups A, B, C and D. The scale bar denotes the number of nucleotide replacements per site.

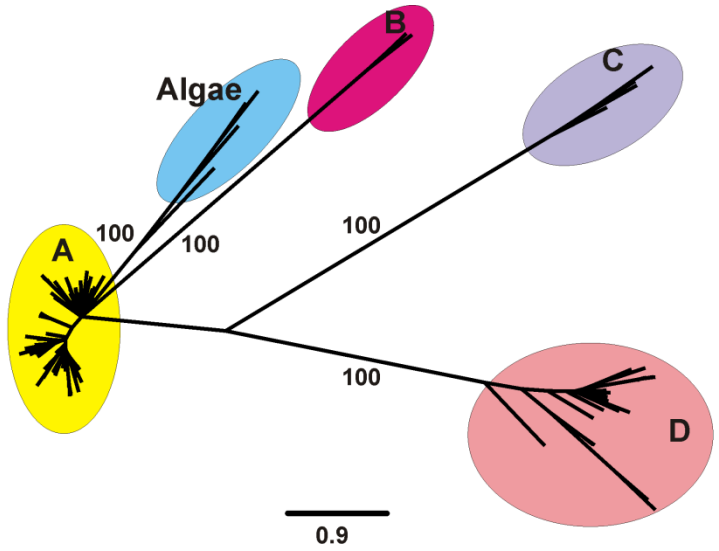
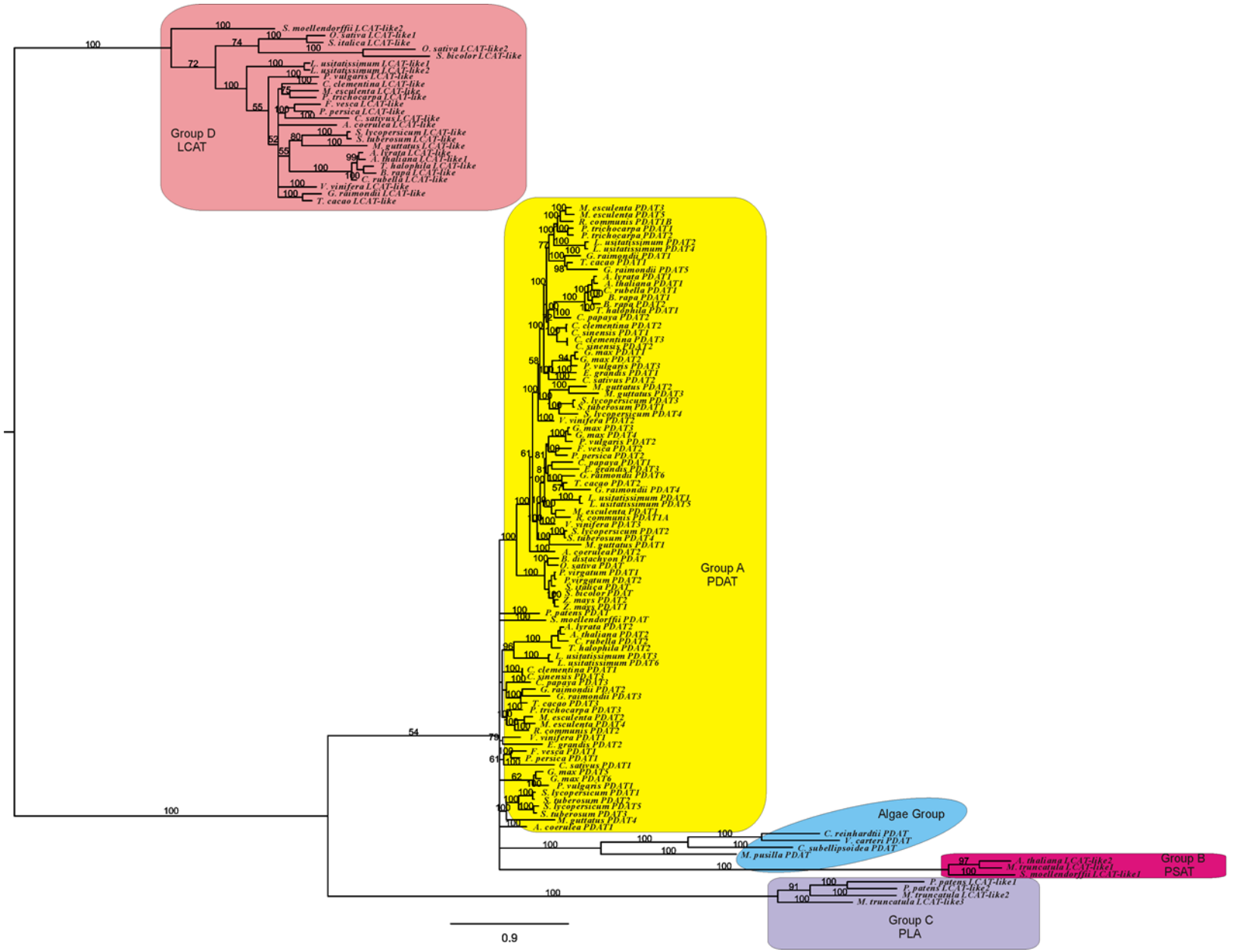


Figure 4.2 Phylogenetic relationship of 128 LCAT-like sequences from 38 plant species. The maximum likelihood tree was generated using the MrBayes program. The land plant LCAT-like sequences are grouped into four major clades, designated as Groups A, B, C and D. The support values (>50%) above branches are Bayesian posterior probabilities (shown as percentages). The scale bar represents the number of nucleotide replacements per site.


Group D
LCAT
Group A
PDAT
Algae Group
Group B
PSAT
Group C
PLA
0.9

In summary, algal and lowland plant species possess a single copy of the *PDAT* gene. In monocots, one duplication event appears to have occurred in *Zea mays* and *Panicum virgatum*, resulting in a duplicated gene pair, while remaining monocots contain only one copy of *PDAT*. The *PDAT* copy number varied from two to six among eudicots, suggesting that multiple duplication events may have occurred in eudicots. It is worth noting that the number of *PDAT* paralogs within Arabidopsis, castor, flax and alga *C. reinhardtii* identified in this study is consistent with previous studies (van Erp et al., 2011; Pan et al., 2013; Yoon et al., 2012; Ståhl et al., 2004).

4.3.2 Phylogenetic analysis divides plant *PDATs* into seven major clades

To explore the evolutionary relationship of the plant *PDAT* gene family, ML trees were further constructed using complementary DNAs (cDNAs) of 86 full-length candidate *PDATs*. A gene encoding PDAT was first identified in *S. cerevisiae*, therefore, this sequence was included as an outgroup in phylogenetic analyses. The ML trees were built by using two phylogenetic programs: MrBayes and RAxML. The trees based on both programs are topologically identical. Figure 4.3 shows the tree produced by MrBayes. Based on the topology and clade support values (≥85%), the *PDAT* gene family can be classified into seven clades designated as Clade I to Clade VII (Figure 4.3). The algal *PDATs* are phylogenetically divergent from the land plant *PDATs* and form a monophyletic group. Inside the land plants, *PDATs* from a moss, a lycophyte and monocots diverged from each other and form three distinct clades, assigned as Clade II to Clade IV. The eudicots can be divided into the basal and the core eudicots (Worberg et al., 2007). As shown in the species tree (Table 4.3), the basal eudicot (represented as *Aquilegia coerulea*) forms a paraphyly at the base of the core eudicots. Within the core eudicots, *PDATs* are grouped into three clades: Clades V, VI and VII. Clades V and VI are more closely related to each other than

they are to Clade VII. For the two *PDAT*s found in the basal eudicot *A. coerulea*, one (*AcoPDAT1*) is sister to the core eudicot Clade VII, while the other (*AcoPDAT2*) forms a sister clade to the core eudicot Clades V and VI. The number of *PDAT* paralogs in each species and their clade-distributions are shown in Table 4.3.

It is important to point out that taxa in the phylogenetic tree are very unevenly distributed among the clades, ranging from 1 to 28 sequences (Figure 4.3), which may have negative impact on phylogenetic accuracy (Heath et al., 2008). To exam the phylogenetic accuracy, I first compared the phylogenetic tree (Figure 4.3) with the species tree shown in Phytozome (Table 4.3). As shown, phylogenetic tree accords exactly with the evolutionary pathway from algae to basal eudicots. Algal *PDATs* are grouped at the base of the tree. The *PDATs* from *Physcomitrella patens* (moss) and *Selaginella moellendorffii* (lycophyte), two basal lineages of land plants, form monophyletic clades after the algal clade. Monocot *PDATs* form a monophyletic clade with the more related species being closer on the phylogenetic tree. The basal eudicot (*A. coerulea*) is placed as sister to the core eudicots. When it comes to the core eudicots, the existence of multiple *PDAT* copies makes the comparison between the species tree and the *PDAT* gene tree complicated. As shown in Figure 4.3, three core eudicot *PDAT* clades, each containing copies from a mixture of species, differ in topology from one another and from the species tree. To further test if the topological discordance between the core eudicot *PDAT* gene trees and the species tree is the result of the uneven taxonomic sampling, the dataset was pruned down and the phylogenetic tree was reconstructed with the sequences from Clades V, VI and VII, which have more balanced taxa sampling. The trees generated from pruned and complete datasets are topologically identical (Figure 4.3 and Figure 4.4), suggesting that the phylogenetic separation of core-eudicot *PDATs* was not affected by the very unevenly distributed taxa among the clades.

Figure 4.3 Phylogenetic relationship of the *PDAT* gene family. The maximum likelihood tree was generated using the MrBayes program. *PDATs* are grouped into 7 distinct clades (I to VII). Numbers above branches represent the support values (Bayesian posterior probabilities). The tree is rooted using PDAT sequences from *S. cerevisiae* as the outgroup. The scale bar represents the number of nucleotide replacements per site.

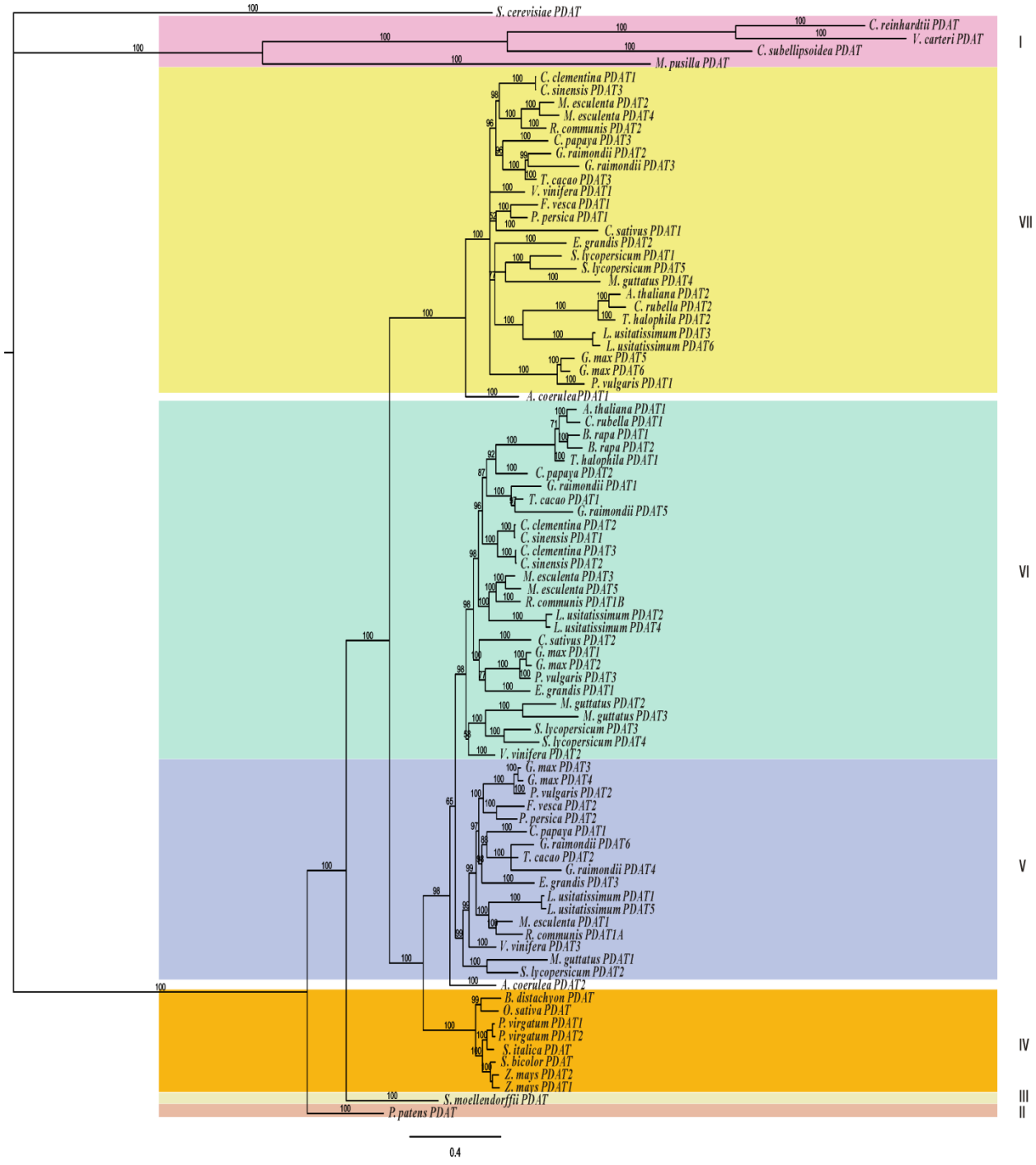
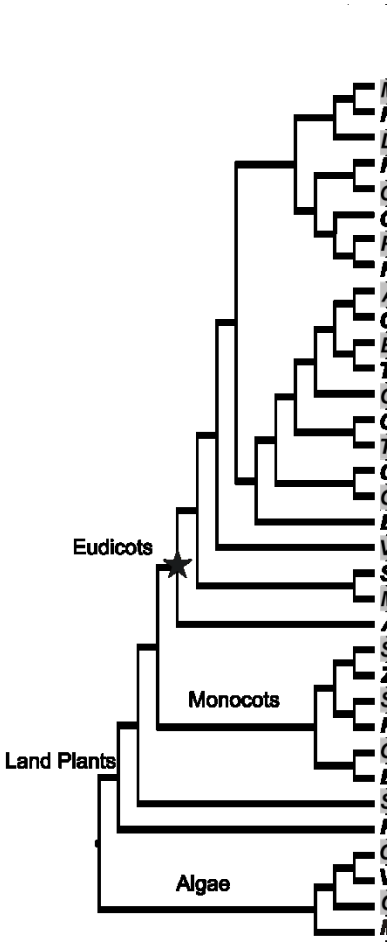
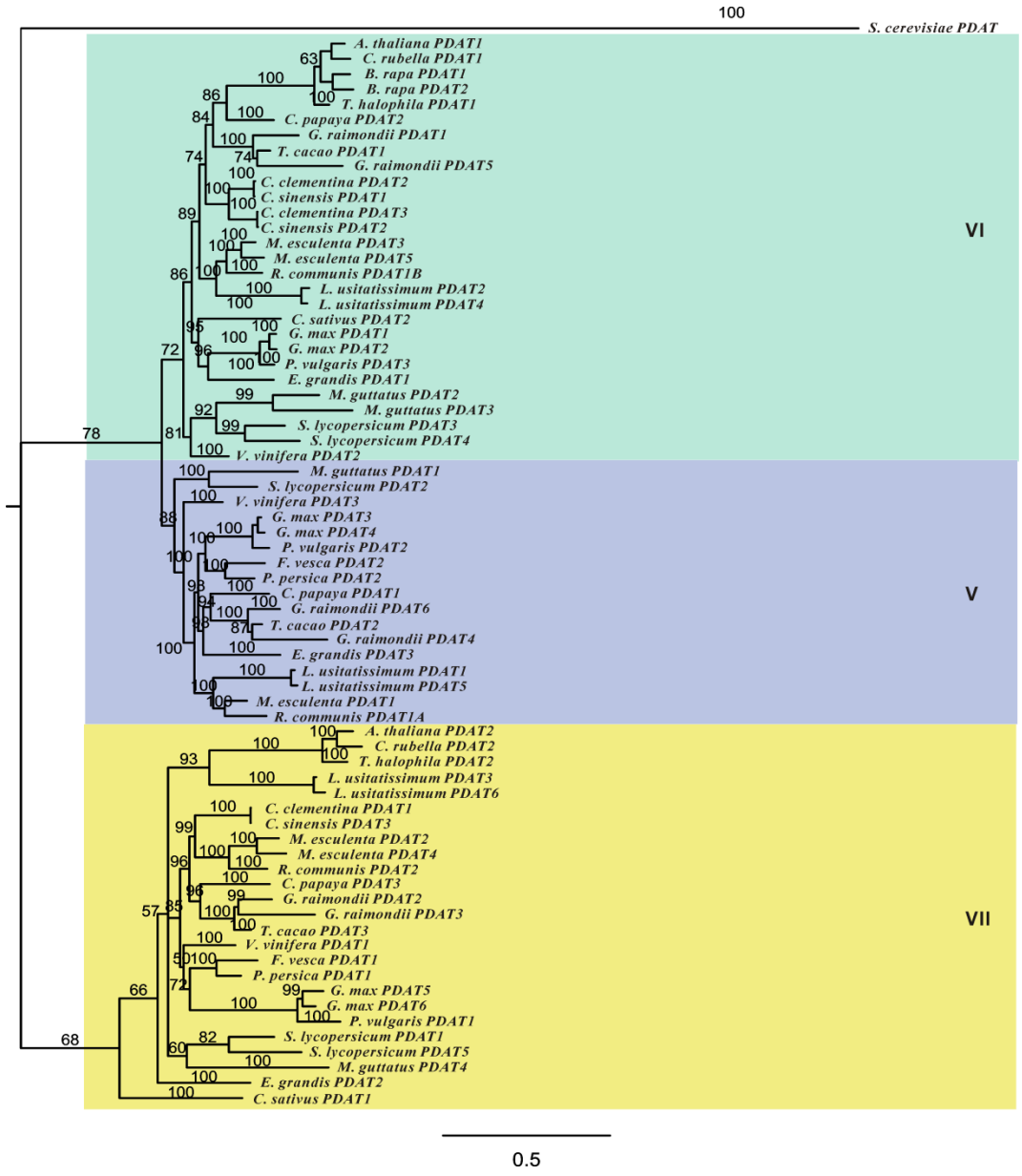

S. cerevisiae PDAT
C. reinhardtii PDAT
V. carteri PDAT
C. subellipsoidea PDAT
M. pusilla PDAT
I
C. clementina PDAT1
C. sinensis PDAT3
M. esculenta PDAT2
M. esculenta PDAT4
R. communis PDAT2
C. papaya PDAT3
G. raimondii PDAT2
G. raimondii PDAT3
T. cacao PDAT3
V. vinifera PDAT1
F. vesca PDAT1
P. persica PDAT1
C. sativus PDAT1
E. grandis PDAT2
S. lycopersicum PDAT1
S. lycopersicum PDAT5
M. guttatus PDAT4
A. thaliana PDAT2
C. rubella PDAT2
T. halophila PDAT2
L. usitatissimum PDAT3
L. usitatissimum PDAT6
G. max PDAT5
G. max PDAT6
P. vulgaris PDAT1
A. coeruleaPDAT1
VII
A. thaliana PDAT1
C. rubella PDAT1
B. rapa PDAT1
B. rapa PDAT2
T. halophila PDAT1
C. papaya PDAT2
G. raimondii PDAT1
T. cacao PDAT1
G. raimondii PDAT5
C. clementina PDAT2
C. sinensis PDAT1
C. clementina PDAT3
C. sinensis PDAT2
M. esculenta PDAT3
M. esculenta PDAT5
R. communis PDAT1B
L. usitatissimum PDAT2
L. usitatissimum PDAT4
C. sativus PDAT2
G. max PDAT1
G. max PDAT2
P. vulgaris PDAT3
E. grandis PDAT1
M. guttatus PDAT2
M. guttatus PDAT3
S. lycopersicum PDAT3
S. lycopersicum PDAT4
V. vinifera PDAT2
VI
G. max PDAT3
G. max PDAT4
P. vulgaris PDAT2
F. vesca PDAT2
P. persica PDAT2
C. papaya PDAT1
G. raimondii PDAT6
T. cacao PDAT2
G. raimondii PDAT4
E. grandis PDAT3
L. usitatissimum PDAT1
L. usitatissimum PDAT5
M. esculenta PDAT1
R. communis PDAT1A
V. vinifera PDAT3
M. guttatus PDAT1
S. lycopersicum PDAT2
A. coerulea PDAT2
V
B. distachyon PDAT
O. sativa PDAT
P. virgatum PDAT1
P. virgatum PDAT2
S. italica PDAT
S. bicolor PDAT
Z. mays PDAT2
Z. mays PDAT1
IV
S. moellendorffii PDAT
III
P. patens PDAT
II
0.4

Table 4.3 Number of *PDAT* paralogs in each species and their clade distributions. The species tree is based on information in Phytozome (http://www.phytozome.net). The star (★) on the branch point within eudicot species indicates the divergence point between a basal eudicot (*Aquilegia coerulea*) and the core eudicots. The tree is not scaled.



Species	Family	Clade							Copy Number
		I	II	III	IV	V	VI	VII	Total
Manihot esculenta	Euphorbiaceae	0	0	0	0	1	2	2	5
Ricinus communis	**Euphorbiaceae**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**3**
Linum usitatissimum	Linaceae	0	0	0	0	2	2	2	6
Phaseolus vulgaris	**Fabaceae**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**3**
Glycine max	Fabaceae	0	0	0	0	2	2	2	6
Cucumis sativus	**Cucurbitaceae**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**2**
Prunus persica	Rosaceae	0	0	0	0	1	0	1	2
Fragaria vesca	**Rosaceae**	**0**	**0**	**0**	**0**	**1**	**0**	**1**	**2**
Arabidopsis thaliana	Brassicaceae	0	0	0	0	0	1	1	2
Capsella rubella	**Brassicaceae**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**2**
Brassica rapa	Brassicaceae	0	0	0	0	0	2	0	2
Thellungiella halophila	**Brassicaceae**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**2**
Carica papaya	Caricaceae	0	0	0	0	1	1	1	3
Gossypium raimondii	**Malvaceae**	**0**	**0**	**0**	**0**	**2**	**2**	**2**	**6**
Theobroma cacao	Malvaceae	0	0	0	0	1	1	1	3
Citrus sinensis	**Rutaceae**	**0**	**0**	**0**	**0**	**0**	**2**	**1**	**3**
Citrus clementine	Rutaceae	0	0	0	0	0	2	1	3
Eucalyptus grandis	**Myrtaceae**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**3**
Vitis vinifera	Vitaceae	0	0	0	0	1	1	1	3
Solanum lycopersicum	**Solanaceae**	**0**	**0**	**0**	**0**	**1**	**2**	**2**	**5**
Mimulus guttatus	Scrophulariaceae	0	0	0	0	1	2	1	4
Aquilegia coerulea	**Ranunculaceae**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**
Sorghum bicolor	Poaceae	0	0	0	1	0	0	0	1
Zea mays	**Poaceae**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**2**
Setaria italic	Poaceae	0	0	0	1	0	0	0	1
Panicum virgatum	**Poaceae**	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**2**
Oryza sativa	Poaceae	0	0	0	1	0	0	0	1
Brachypodium distachyon	**Poaceae**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**
Selaginella moellendorffii	Selaginellaceae	0	0	1	0	0	0	0	1
Physcomitrella patens	**Funariaceae**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**1**
Chlamydomonas reinhardtii	Chlamydomonadaceae	1	0	0	0	0	0	0	1
Volvox carteri	**Vocaceae**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**
Coccomyxa subellipsoidea	Coccomyxaceae	1	0	0	0	0	0	0	1
Micromonas pusilla	**Mamiellaceae**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**1**

Figure 4.4 Phylogenetic relationship of core-eudicot *PDATs*. The maximum likelihood tree was generated using the MrBayes program. The support values (>50%) above branches are Bayesian posterior probabilities (shown as percentages). The tree is rooted using *PDAT* sequences from *S. cerevisiae* as the outgroup. The scale bar represents the number of nucleotide replacements per site.



The causes of discordance between the multi-copy *PDAT* gene trees and the species tree remain unknown, but it is a well-known phenomenon that gene trees do not necessarily agree with the species tree and this discordance can be the result of many evolutionary processes, such as gene duplication and loss, and incomplete lineage sorting (Page and Charleston, 1997; Maddison, 1997).

Overall, the high confidence of the phylogenetic separation of the *PDAT* gene family is achieved through the high bootstrap support obtained from multiple phylogenetic reconstruction methods, comparisons between the phylogenetic tree and species tree as well as consistent phylogenetic topologies inferred from complete and pruned datasets.

4.3.3 Gene structure analysis reveals highly conserved exon/intron structure and intron phase pattern throughout land plant *PDAT*s

To further investigate the structural diversity of plant *PDAT* genes, the exon/intron organization for each individual gene was analyzed (representative *PDATs* shown in Figure 4.5 and details shown in Figure 4.6). Diverse gene structure has been found in the algal *PDATs* (Clade I): Micromonas pusilla CCMP1545 *PDAT* (*MpuCMPPDAT*) has no intron, while the rest of the *PDAT* genes have 9 to 14 introns. By contrast, land plant *PDATs* (Clades II to VII) are remarkably well conserved in terms of exon/intron structure. Approximately 92% (76 out of 82) of the land plant *PDATs* have six exons and five introns. Six exceptions to this exon/intron pattern are *VviPDAT1*, *PvuPDAT1*, *GmaPDAT5* and *GraPDAT1* with seven exons and six introns, and *LusPDAT3* and *LusPDAT6* with five exons and four introns. In addition, intron phases across all *PDAT*s were investigated. Intron phase can be classified into three categories (0, 1, and 2) depending on the position of the intron relative to the codon: phase-0 intron does not interrupt the reading frame and lies between two consecutive codons; phase-1 intron inserts and

interrupts the reading frame between the first and second nucleotides; phase-2 intron inserts and interrupts the reading frame between the second and third nucleotides. The analysis shows that the intron phase pattern (2, 0, 2, 0, 2) is strikingly conserved across 75 out of 82 land plant *PDATs* (Figure 4.5 and Figure 4.6).

4.3.4 Evaluation of PDAT protein properties reveals that PDATs that belong to the core eudicot Clade VI had a tendency to maintain acidic isoelectric points (pI values) during evolution

After evaluation of gene structure, I continued the analysis with a focus on protein properties of 86 PDATs, including protein length, molecular mass and pI values. According to my analyses (Table 4.4 and Table 4.5), the length and molecular mass of PDATs from Clade I (algae clade) varied substantially. CrePDAT with 1041 amino acid residues and 104.5 kDa is the longest and largest PDAT, while CsuPDAT with 509 amino acid residues and 56.9 kDa is the shortest and smallest PDAT of all 86 PDATs. In contrast, the variation of protein length and molecular mass is small in land plant PDATs, ranging from 572 to 716 amino acid residues and 62.8 kDa to 80.3 kDa, with a mean of 671 amino acid residues and 74.7 kDa. For the pI values, MpuCMPPDAT has the highest value of 9.53. Except MpuCMPPDAT, PDATs within Clades I, II, III and IV, have very close pI values, ranging from 5.96 to 6.5, with an average of 6.21. Interestingly, PDATs from Clade VI (except GraPDAT5) have maintained acidic pI values with an average of 6.35, while more alkaline pI values (>7) have been observed in 31 out of 42 PDATs belongs to Clades V and VII.

Figure 4.5 Schematic diagram of gene structures of representative *PDATs*. The thin lines represent introns and thick bars represent exons. The numbers above the gene structure indicate intron phases. A scale bar with a unit of base pair (bp) is graphed on the top. TMD denotes transmembrane domain. Gene identifiers and abbreviations of listed *PDATs* can be found in the Table 4.1.

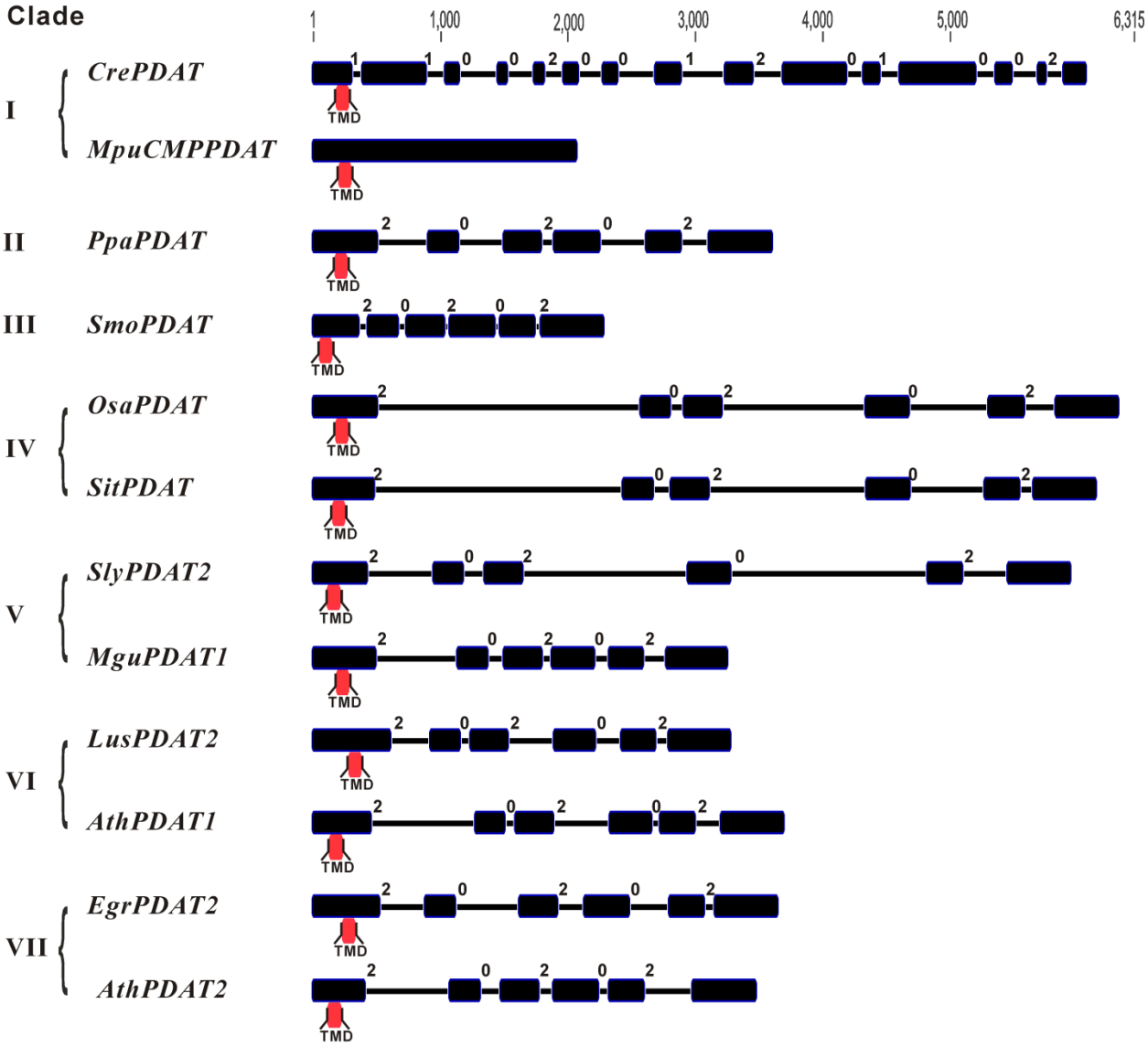


Figure 4.6 Schematic diagram of gene structures of 86 plant *PDATs*. The thin lines represent introns and thick bars represent exons. The numbers above the gene structure indicate intron phases. A scale bar with a unit of base pair (bp) is graphed on the top. Gene identifiers and abbreviations of listed PDATs can be found in the Table 4.1.

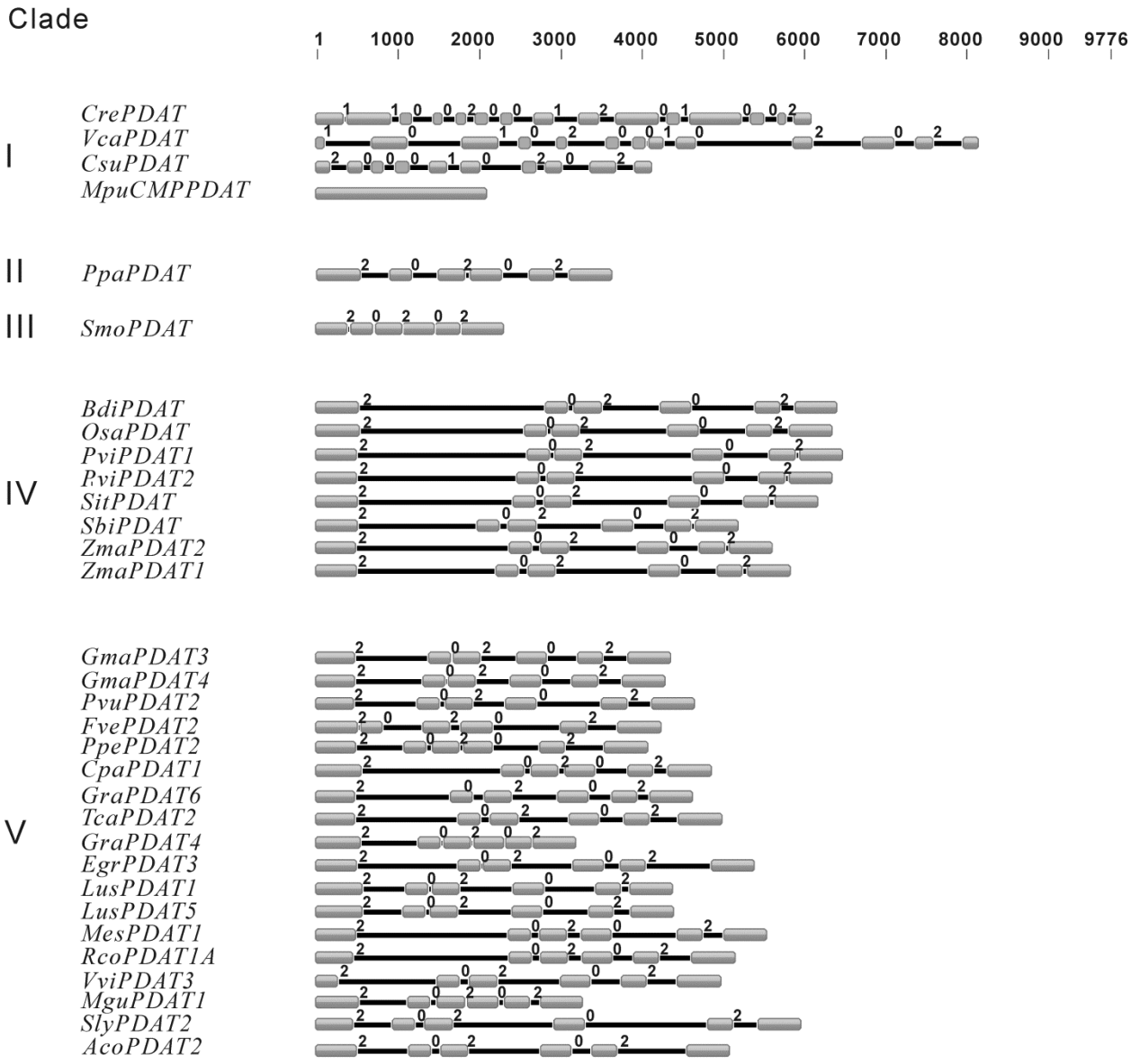


Figure 4.6 Continued

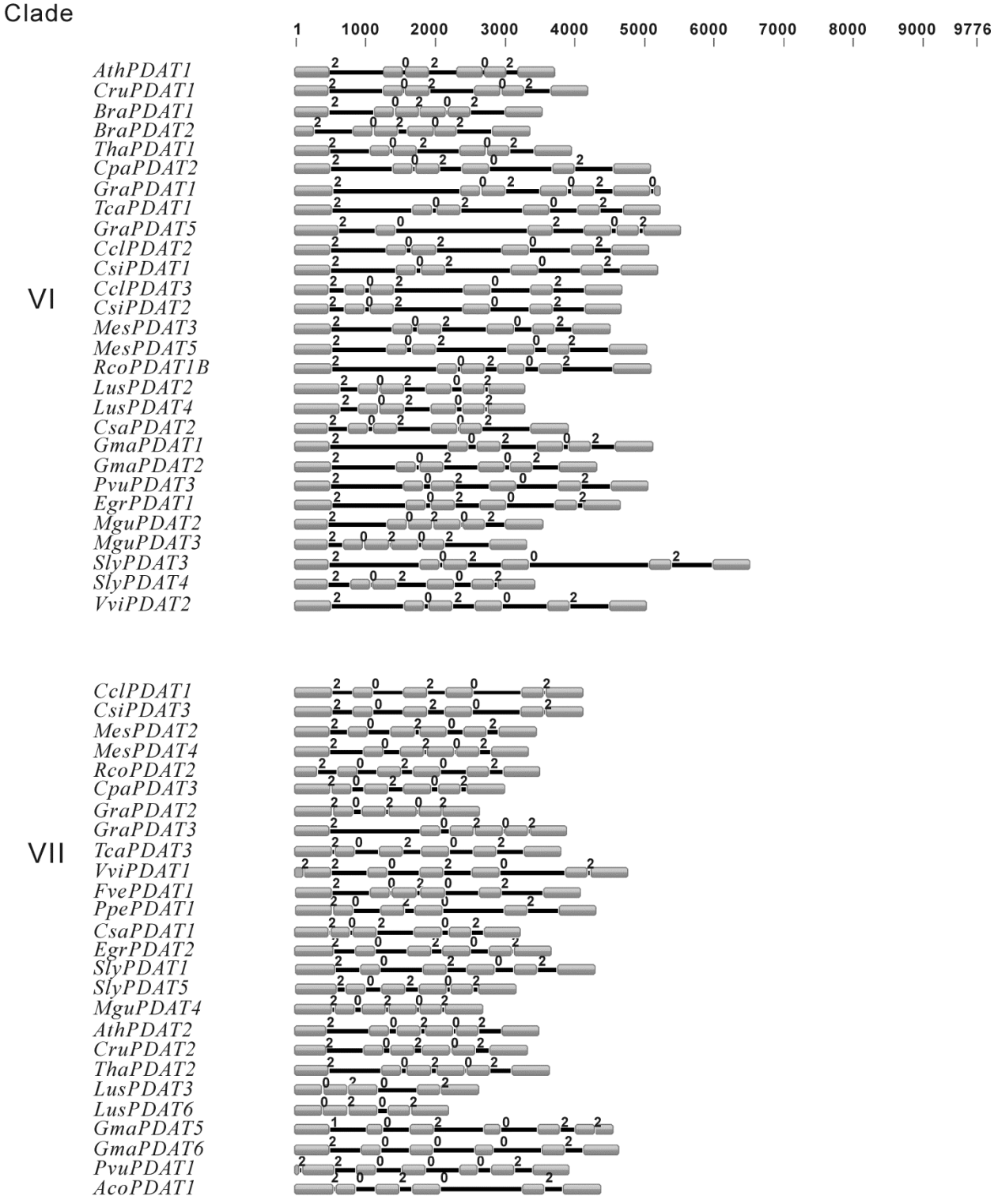


Table 4.4 Summary of PDAT protein properties

Clade	Total No. of PDATs	Protein length (amino acid residues)	Molecular mass (kDa)	Isoelectric point (pI)	No. of PDATs with pI>7.0
I	4	757±223	79.6±19.6	6.90±1.76	0
II	1	693	76.9	6.50	0
III	1	647	71.8	6.21	0
IV	8	682±4.87	75.4±0.4	6.25±0.13	0
V	17	670±21.96	74.5±2.53	7.65±0.94	12
VI	28	676±21.66	75.1±2.52	6.34±0.61	1
VII	25	663±40.7	74.1±4.76	7.94±0.97	19
Min		509	56.9	5.27	
Max		1041	104.5	9.53	

Table 4.5 PDAT protein properties and their clade-distributions.

Species	PDAT	Protein length (amino acid residues)	Molecular mass (Da)	pI	Clade
Manihot esculenta	MesPDAT1	672	74827.57	8.67	V
	MesPDAT2	677	75890.18	8.19	VII
	MesPDAT3	680	75648.30	6.08	VI
	MesPDAT4	677	76030.33	6.92	VII
	MesPDAT5	685	76524.33	6.80	VI
Ricinus communis	RcoPDAT1A	660	73401.82	7.22	V
	RcoPDAT1B	685	76547.17	5.77	VI
	RcoPDAT2	609	67723.74	8.20	VII
Linum usitatissimum	LusPDAT1	695	76891.89	8.23	V
	LusPDAT2	714	79072.61	6.4	VI
	LusPDAT3	575	63093.05	6.19	VII
	LusPDAT4	715	78923.51	6.72	VI
	LusPDAT5	695	76817.64	8.28	V
	LusPDAT6	572	62792.54	6.19	VII
Phaseolus vulgaris	PvuPDAT1	614	68893.06	8.66	VII
	PvuPDAT2	664	74321.16	7.56	V
	PvuPDAT3	675	75479.12	6.67	VI
Glycine max	GmaPDAT1	668	74557.01	6.28	VI
	GmaPDAT2	676	75571.09	6.28	VI
	GmaPDAT3	668	74776.84	8.59	V
	GmaPDAT4	668	74833.89	8.59	V
	GmaPDAT5	582	65783.50	5.91	VII
	GmaPDAT6	628	70846.31	6.34	VII
Cucumis sativus	CsaPDAT1	670	75432.15	6.84	VII
	CsaPDAT2	661	73517.94	6.22	VI
Prunus persica	PpePDAT1	685	76688.07	8.53	VII
	PpePDAT2	665	73657.99	6.02	V
Fragaria vesca	FvePDAT1	689	77009.31	8.42	VII
	FvePDAT2	666	73949.46	6.34	V
Arabidopsis thaliana	AthPDAT1	671	74156.71	6.50	VI
	AthPDAT2	665	73652.61	8.69	VII
Capsella rubella	CruPDAT1	666	73852.34	6.54	VI
	CruPDAT2	666	74072.70	7.56	VII
Brassica rapa	BraPDAT1	666	73746.09	6.11	VI

	BraPDAT2	597	65699.06	6.38	VI
Thellungiella halophila	ThaPDAT1	671	74353.89	6.48	VI
	ThaPDAT2	676	75020.24	8.90	VII
Carica papaya	CpaPDAT1	695	77420.27	8.39	V
	CpaPDAT2	675	74623.31	6.9	VI
	CpaPDAT3	683	76602.07	8.76	VII
Gossypium raimondii	GraPDAT1	697	77376.95	5.86	VI
	GraPDAT2	691	77342.87	8.29	VII
	GraPDAT3	677	75722.97	8.93	VII
	GraPDAT4	690	76893.66	8.24	V
	GraPDAT5	716	80280.15	8.71	VI
	GraPDAT6	672	74563.08	6.37	V
Theobroma cacao	TcaPDAT1	685	76097.70	6.11	VI
	TcaPDAT2	670	74153.69	8.40	V
	TcaPDAT3	691	77598.45	8.88	VII
Citrus sinensis	CsiPDAT1	675	75282.76	6.06	VI
	CsiPDAT2	668	74099.33	5.77	VI
	CsiPDAT3	688	77301.78	7.78	VII
Citrus clementina	CclPDAT1	688	77301.78	7.78	VII
	CclPDAT2	675	75319.82	6.1	VI
	CclPDAT3	668	74114.30	5.83	VI
Eucalyptus grandis	EgrPDAT1	684	75755.38	5.84	VI
	EgrPDAT2	702	78444.89	8.20	VII
	EgrPDAT3	677	74739.17	7.54	V
Vitis vinifera	VviPDAT1	665	74346.75	8.78	VII
	VviPDAT2	680	75488.14	6.12	VI
	VviPDAT3	599	65938.39	6.52	V
Solanum lycopersicum	SlyPDAT1	694	77548.20	8.72	VII
	SlyPDAT2	662	73710.56	8.46	V
	SlyPDAT3	668	74279.90	6.59	VI
	SlyPDAT4	663	74049.99	6.99	VI
	SlyPDAT5	709	79277.82	8.31	VII
Mimulus guttatus	MguPDAT1	680	74857.29	6.66	V
	MguPDAT2	668	74323.96	6.24	VI
	MguPDAT3	665	73905.89	5.27	VI
	MguPDAT4	692	77624.11	8.59	VII
Aquilegia coerulea	AcoPDAT1	699	78216.85	8.17	-
	AcoPDAT2	677	75724.40	6.22	-
Sorghum bicolor	SbiPDAT1	682	75423.87	6.10	IV

Zea mays	ZmaPDAT1	678	75153.72	6.19	IV
	ZmaPDAT2	676	74790.26	6.22	IV
Setaria italica	SitPDAT1	684	75634.12	6.07	IV
Panicum virgatum	PviPDAT1	680	75354.90	6.41	IV
	PviPDAT2	681	75365.84	6.31	IV
Oryza sativa	OsaPDAT	691	76144.77	6.43	IV
Brachypodium distachyon	BdiPDAT	687	75650.16	6.28	IV
Selaginella moellendorffii	SmoPDAT	647	71752.23	6.21	III
Physcomitrella patens	PpaPDAT	693	76851.27	6.50	II
Chlamydomonas reinhardtii	CrePDAT	1041	104535.61	5.96	I
Volvox carteri	VcaPDAT	792	81521.82	6.10	I
Coccomyxa subellipsoidea C-169	CsuPDAT	509	56931.71	6.00	I
Micromonas pusilla CCMP1545	MpuCMPPDAT	685	75494.77	9.53	I
Saccharomyces cerevisiae	ScePDAT	661	75393.13	6.21	

pI-isoelectric point

4.3.5 The membrane topology of the PDAT proteins is well conserved among most land plants

Next, the membrane topology of plant PDATs was analyzed. The putative transmembrane domains (TMDs) of 86 PDATs were predicted using the TMHMM program. To provide a better comparison of the TMDs among PDATs, the polypeptides with the annotated TMD regions were aligned using ClustalW. The results (examples shown in Figure 4.7A and details shown Figure 4.8) show that two out of four algal PDATs (Clade I) have one putative TMD (CrePDAT and MpuCMPPDAT), while the other two contain no TMD. CrePDAT was previously predicted to be localized in chloroplasts (Yoon et al., 2012); therefore, MpuCMPPDAT might also be a chloroplast-localized protein. Because the endoplasmic reticulum (ER) is the major site for TAG biosynthesis in plants (Lung and Weselake, 2006), it was assumed that land plant PDATs are inserted into the ER and interpreted the topology results based on the ER structure. The results (examples shown in Figure 4.7A and details shown Figure 4.8) show that 73 out of 82 land plant PDATs (except BraPDAT2, CruPDAT2, GraPDAT2, GraPDAT3, LusPDAT3, LusPDAT6, RcoPDAT2, TcaPDAT3 and VviPDAT3) have a single putative TMD, with the short N-terminus facing the cytosol and the bulk of the C-terminus residing in the ER lumen. This result is consistent with the topology reported for yeast and Arabidopsis PDATs (Ghosal et al., 2007; Yoon et al., 2012). The alignment results further indicate that the position of the TMD is highly preserved among land plant PDATs (Figure 4.7A and Figure 4.8). Mapping the TMD region onto the genomic sequence reveals that the position of the TMD corresponds to the region within the first exon in 71 out of 73 TMD-containing PDATs (examples shown in Figure 4.5). The alignment of land plant PDATs only (Figure 4.7B) also shows that the hydrophilic N-terminal region preceding the TMD appears to be the most

divergent region, which carries the only common feature: a cluster of consecutive arginine residues. Interestingly, the N-termini of DGAT1 is also the most variable region and carries the arginine cluster (Liu et al., 2012). The role of these conserved arginine residues remains unclear, but it has been speculated that they are potentially an ER-localization signal (Liu et al., 2012).

4.3.6 Plant PDATs contain the conserved amino acids in LCAT

To gain more insights about the structure/function features of PDATs, multiple sequence alignment was further used to identify conserved amino acid residues. The alignment shows that besides the initial methionine residue, 39 amino acid residues are completely conserved in 86 PDATs. Among the completely conserved amino acid residues, nine of them are located at the C-terminal portion and the rest are concentrated within the 320 amino acid residues following the TMD.

It is known that PDAT belongs to the LCAT-like family. The first *PDAT* gene was isolated based on the homology to human LCAT, which is a soluble protein with no TMD. It was previously reported that human LCAT contains several structurally conserved elements (Peelman et al., 1998; Peelman et al., 1999), including a catalytic triad of Ser181-His377-Asp345, a salt bridge between Asp145 and Arg147, and a so-called "lid region". The Trp61 within the lid region was proposed to play an important role in binding the cleaved FA into the active site for optimal acylation process. Aligning plant PDATs with human LCAT reveals (examples shown in Figure 4.9 and details shown Figure 4.10) that the Trp61, Asp145, Arg147, Ser181, and Asp345 of LCAT are completely conserved in 86 PDATs, while His377 is conserved in 83 out of 86 PDATs, with the exception of GmaPDAT5, GmaPDAT6 and PvuPDAT1. Since the available EST sequences of *GmaPDAT5*, *GmaPDAT6* and *PvuPDAT1* were not long enough to cover the

coding regions for His377, it is not sure if these mismatches are the result of genome sequencing errors.

4.3.7 Conservation and variation in the motif composition and arrangement of PDATs provides further support for the grouping of phylogenetic clades

The motifs in PDATs were further analyzed. InterPro search identified two signature protein motifs in all PDATs, which are IPR003386 for the LCAT family and IPR029058 for the alpha/beta hydrolase fold family. InterPro, however, is limited to the known motifs present in PDATs.

In order to further identify the conservation and variation in the motif arrangements among PDATs, all PDATs were subjected to a MEME analysis. A total of 51 distinct motifs were identified. The occurrences of the motifs in representative PDATs from seven major clades are shown in Figure 4.11. More detailed information is provided in Figure 4.12 and 4.13. The analysis shows that the motif composition of PDATs in algae is very different from that in land plants, which corresponds to their divergent gene structure. Land plant PDATs were found to share many of the motifs.

Among 51 motifs, 11 motifs are classified into the LCAT family. All PDATs contain four LCAT-like motifs occurring after the TMD, with first three sequentially arranged and the fourth one separated. Six major LCAT motif patterns (Figure 4.11 and Figure 4.13) have been identified based on the MEME combined block diagram. These patterns include: motifs 4, 50, 33, and 51 in Clade I; motifs 1, 12, 10, and 5 in Clade III; motifs 1, 15, 10 and 5 or motifs 1, 12, 10, 5 in Clade VI; motifs 1, 12, 7 and 5 or motifs 1, 12, 7 and 3 in Clade VII; and motifs 1, 15, 10 and 5 shared among Clades II, V and VI.

Figure 4.7 Alignment of transmembrane domains (TMDs) in representative plant PDATs. The putative TMDs are annotated as red arrows. The alignment was generated using ClustalW implemented in Geneious software and represented as thick lines (aligned characters) and thin lines (gaps). Overall alignment identity of 86 final selected plant PDATs (A) and a scale bar indicating the numbers of amino acid residues are graphed on the top. (B) Overall alignment identity of 80 land plant PDATs. Gene identifiers and abbreviations for listed PDATs can be found in Table 4.1.

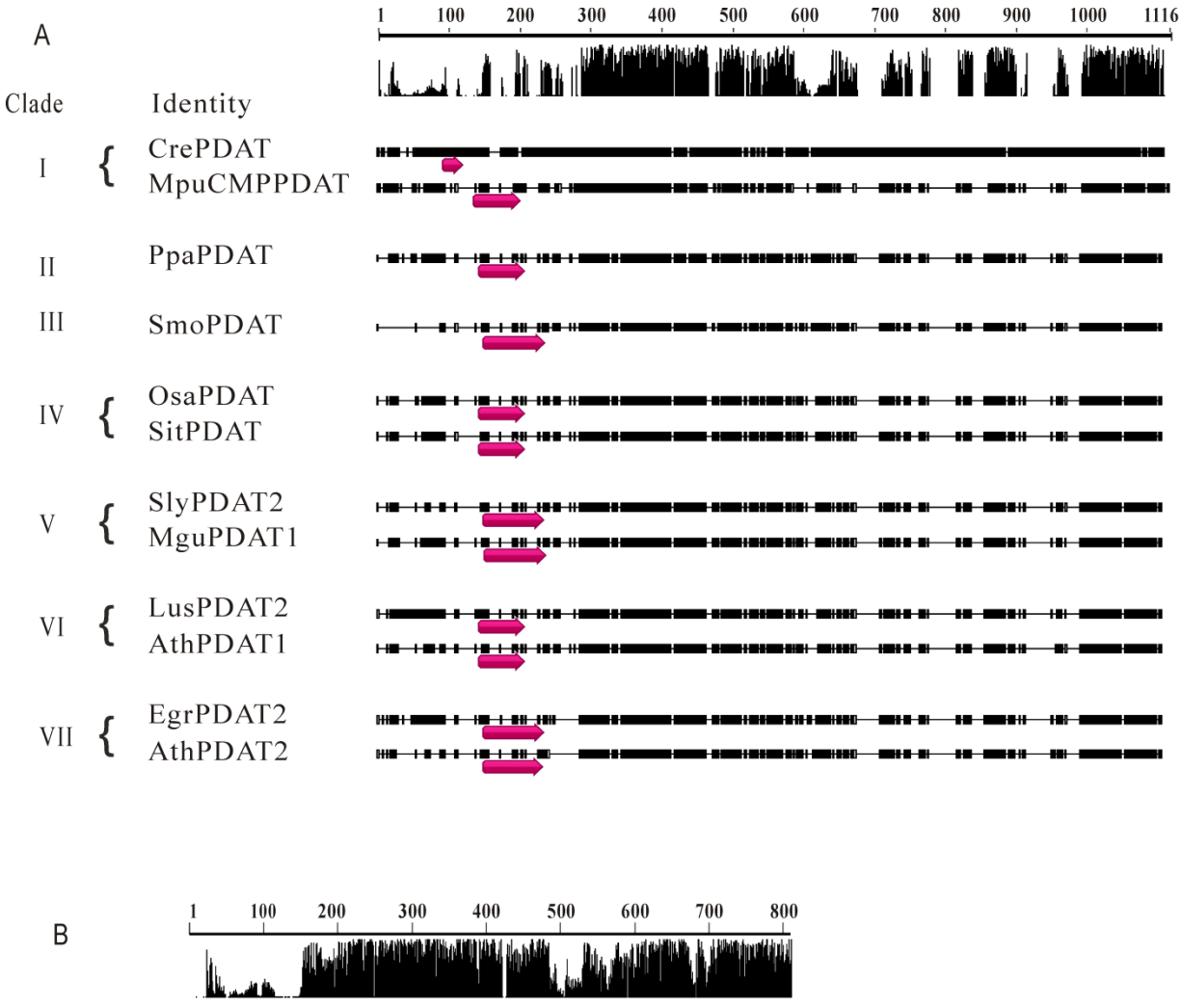


Figure 4.8 Alignment of transmembrane domains (TMDs) in 86 plant PDATs. The putative TMDs are annotated as red arrows. The alignment was generated using ClustalW implemented in Geneious software and represented as thick lines (aligned characters) and thin lines (gaps). Overall alignment identity and a scale bar indicating the numbers of amino acid residues are graphed on the top. Gene identifiers and abbreviations of listed PDATs can be found in Table 4.1.

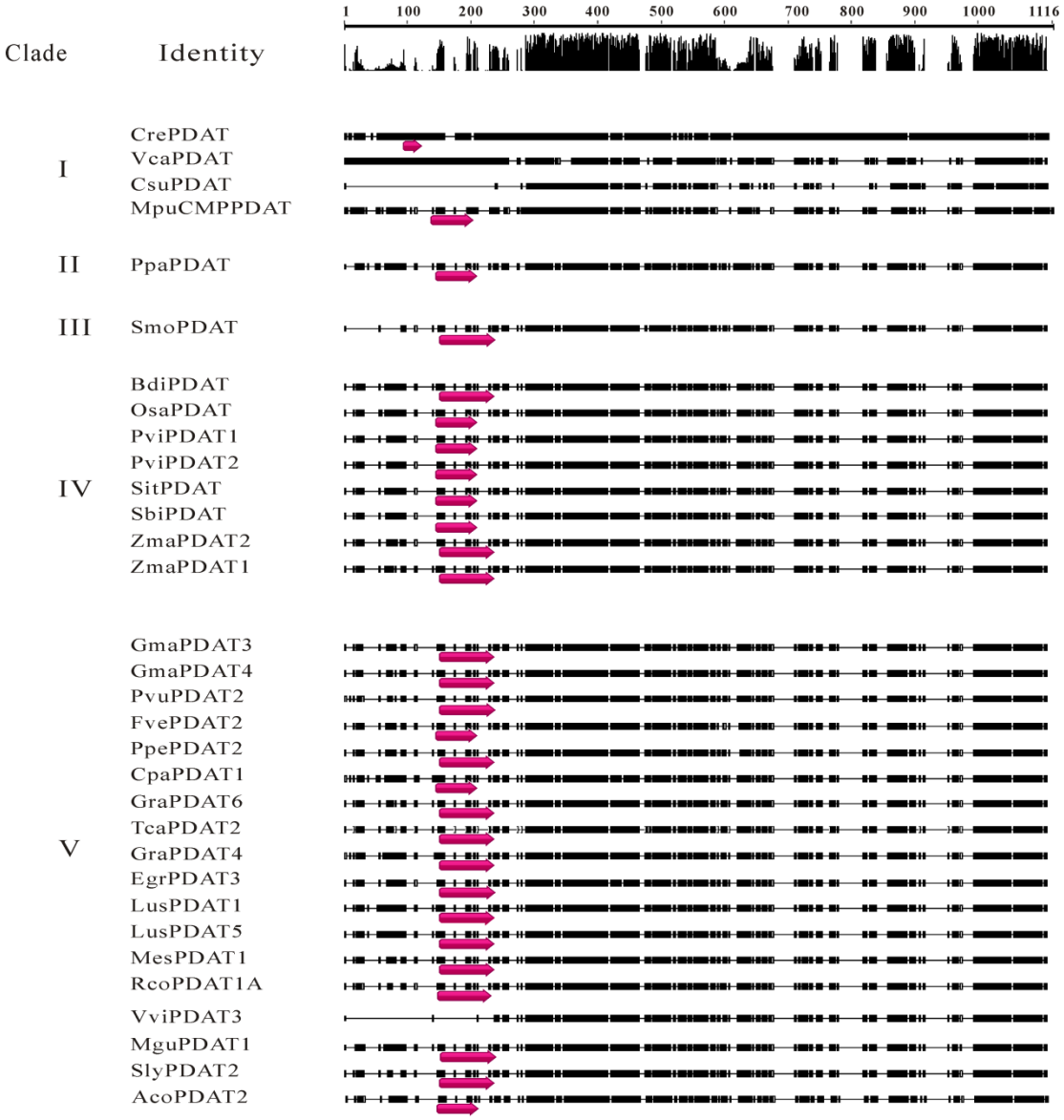


Figure 4.8 Continued

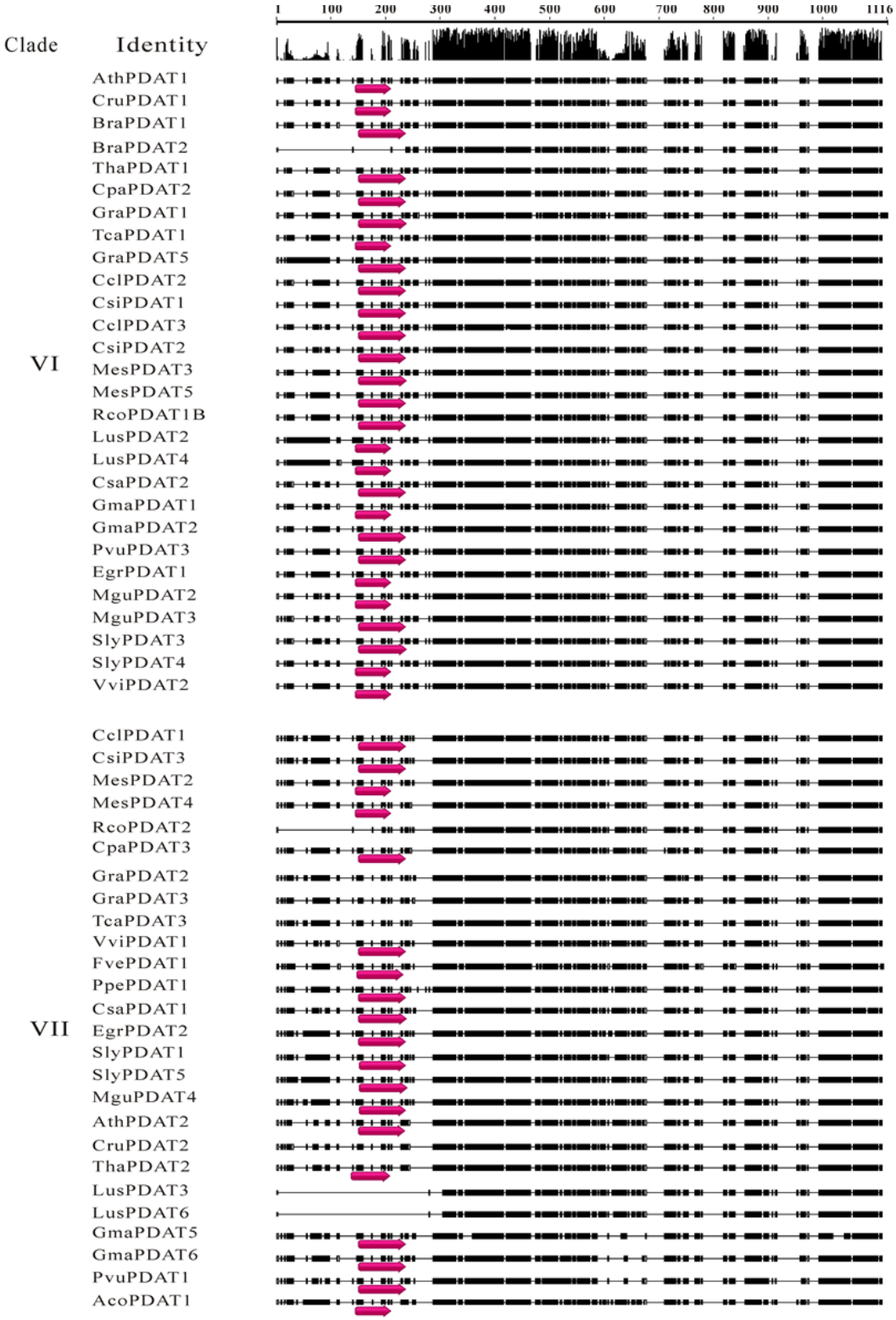


Figure 4.9 Alignment of representative PDAT polypeptides with the human lecithin:cholesterol acyltransferase (HsLCAT; Genbank Accession No. AAB34898.1).

The thick line represents the LCAT polypeptide. The positions of the structurally conserved elements in LCAT, including a catalytic triad of Ser181-His377-Asp345, a salt bridge between Asp145 and Arg147, and a so-called "lid region" (annotated as blue arrow) containing Trp 61, are indicated above. Stars (★) indicate the positions of the conserved amino acids of LCAT in PDAT polypeptides. Gene identifiers and abbreviations for the listed PDATs can be found in Table 4.1.

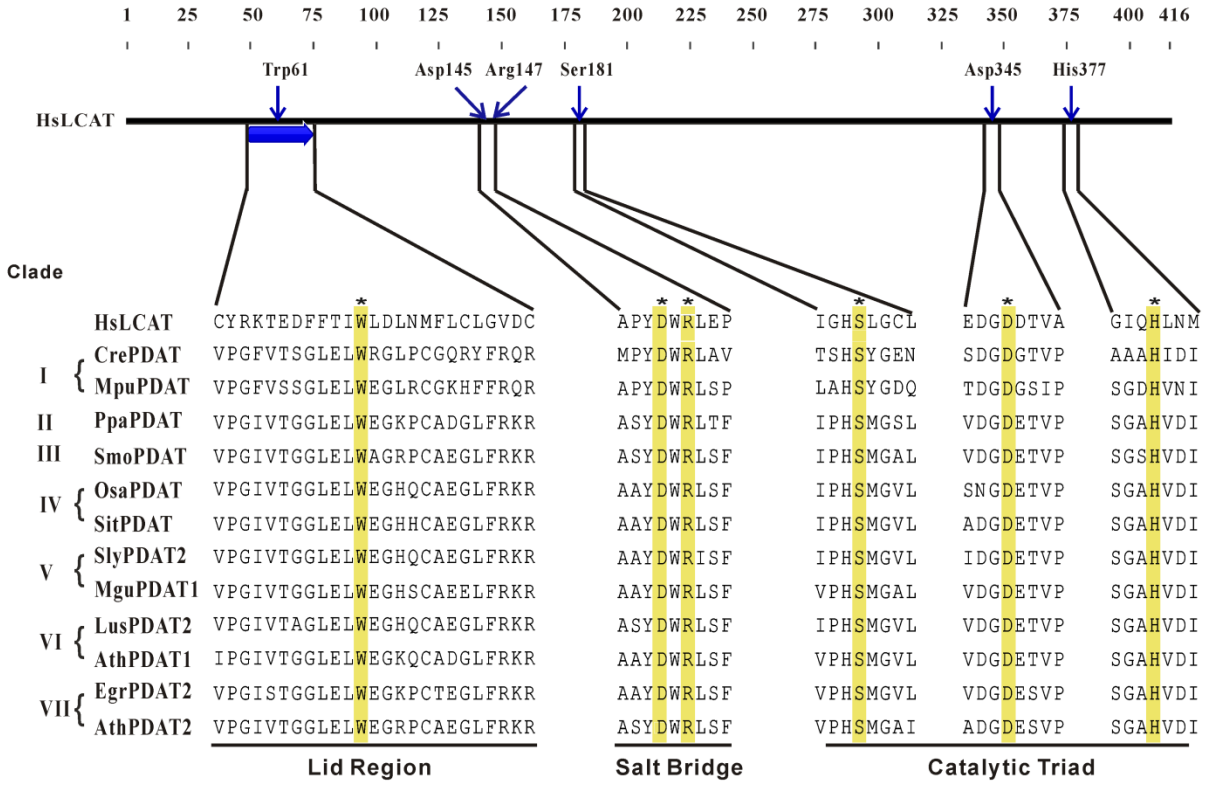


Figure 4.10 Alignment of 86 plant PDAT polypeptides with the human lecithin:cholesterol acyltransferase. Stars (★) indicate the positions of the structurally conserved elements in LCAT, including a so-called "lid region" containing Trp61, a salt bridge between Asp145 and Arg147, and a catalytic triad of Ser181-His377-Asp345. Gene identifiers and abbreviations for the listed PDATs can be found in Table 4.1.

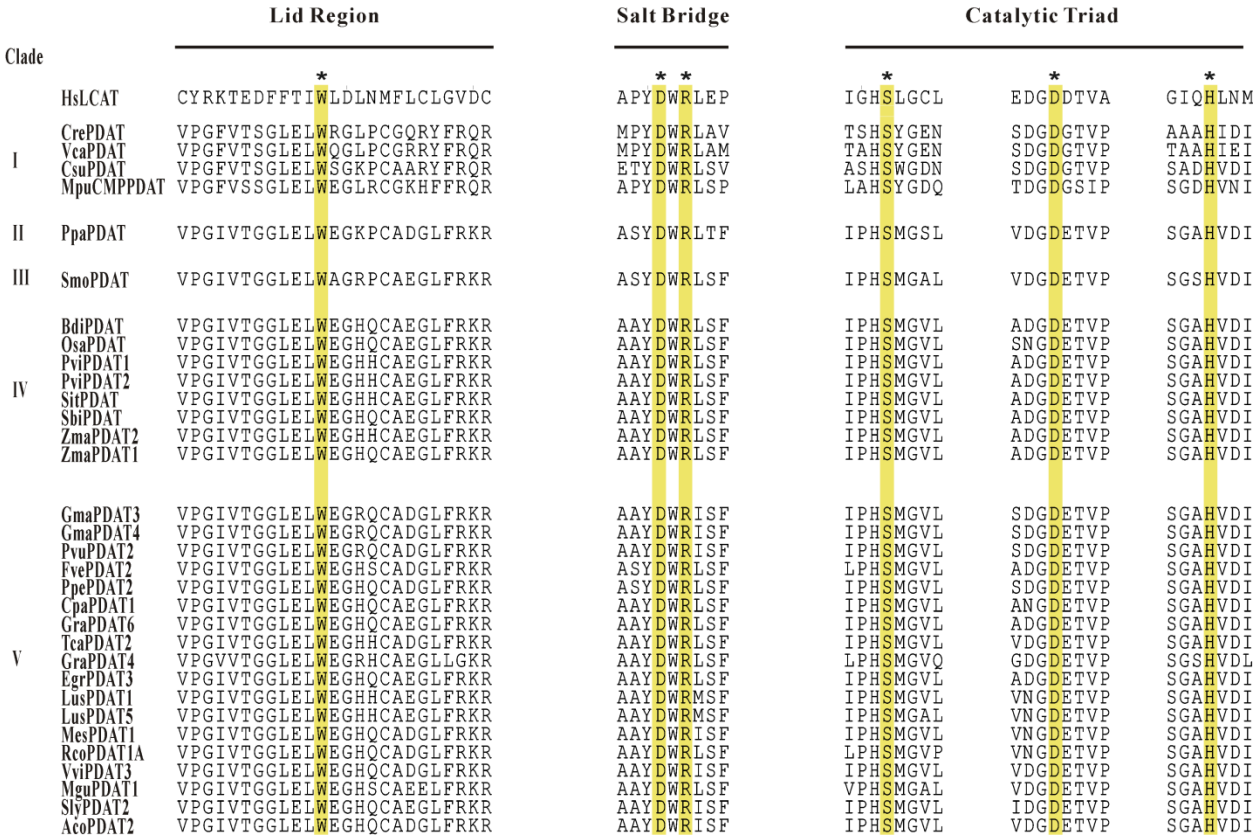


Figure 4.10 Continued

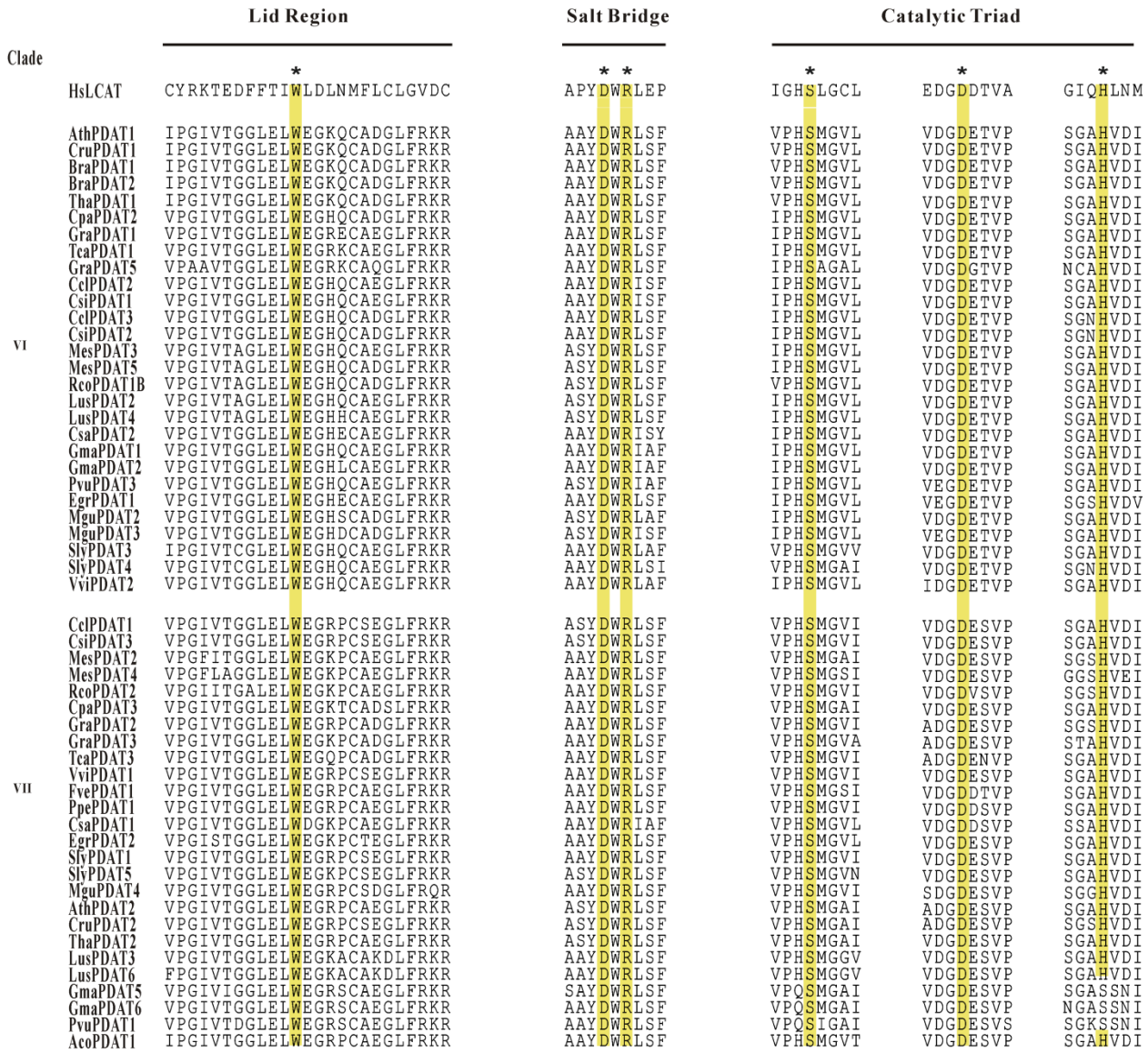


Figure 4.11 Motif patterns in representative PDATs. Motif occurrences were predicted using MEME program and the polypeptides were aligned using ClustalW implemented in Geneious software. Thick lines represent aligned characters and thin lines represent gaps. Different colored and numbered boxes represent separate and distinct motifs. The size of the box does not correspond to the size of the motif due to the alignment. The putative transmembrane domains (TMDs) are annotated as red arrows. Stars (★) indicate the LCAT-like motifs. Based on the position of the TMD and LCAT-like motifs, the PDAT polypeptide was further divided into four regions. Gene identifiers and abbreviations for PDATs can be found in Table 4.1.

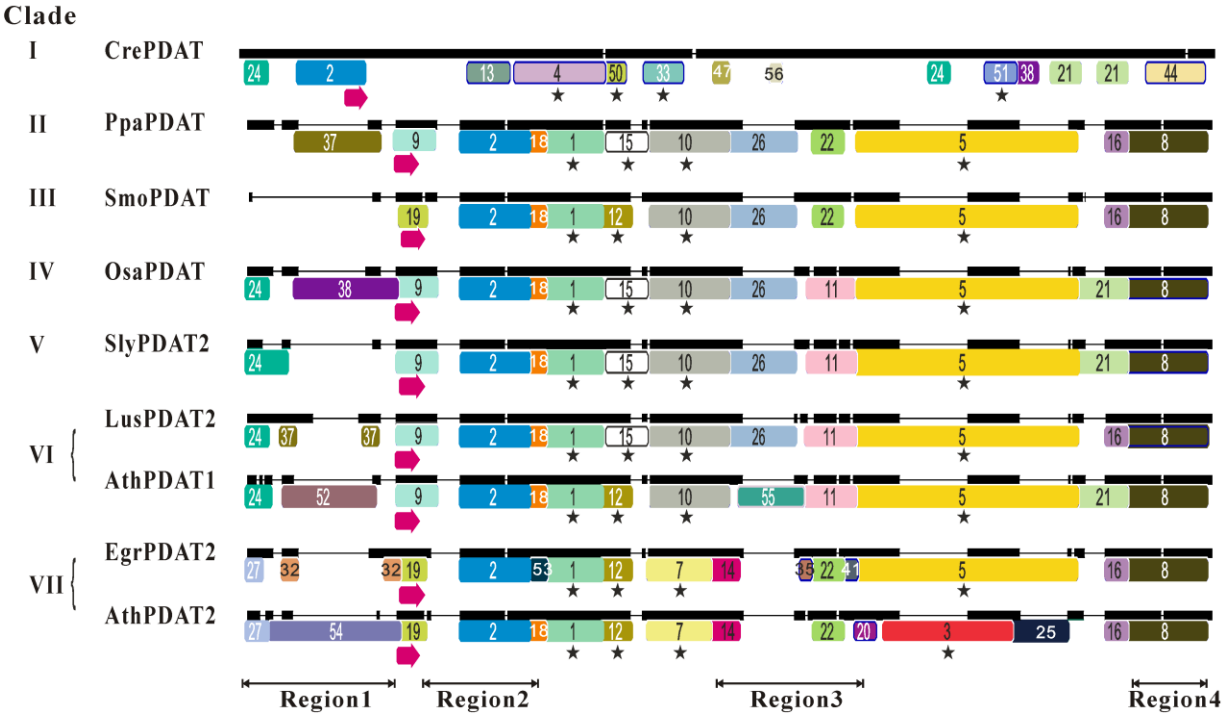


Figure 4.12 Sequence logos derived from the MEME analysis. Amino acid residues are presented by one-letter abbreviations. The letter height is proportional to the relative frequency of the corresponding amino acid in the aligned column. The E-value calculated by MEME for each identified motif and the number of motif occurrences (sites) across 86 PDATs is indicated next to the sequence logo.

Figure 4.12 Continued

Figure 4.13 Combined motif diagram of 86 PDAT polypeptides. The combine block shows the order and spacing of the best non-overlapping motif matches on the protein sequences. If two motif matches overlap each other, only the most significant one is displayed. Therefore, the occurrences (sites) of the motif are not exactly the same as those shown in Figure 4.12. Thick lines represent the PDAT polypeptides. Different colored and numbered boxes represent separate and distinct motifs identified using MEME program (See Materials and Methods for parameter settings). A scale bar indicating the numbers of amino acid residues is shown on the top. Motifs are drawn approximately to scale as boxes. Gene identifiers and abbreviations of listed PDATs can be found in Table 4.1.

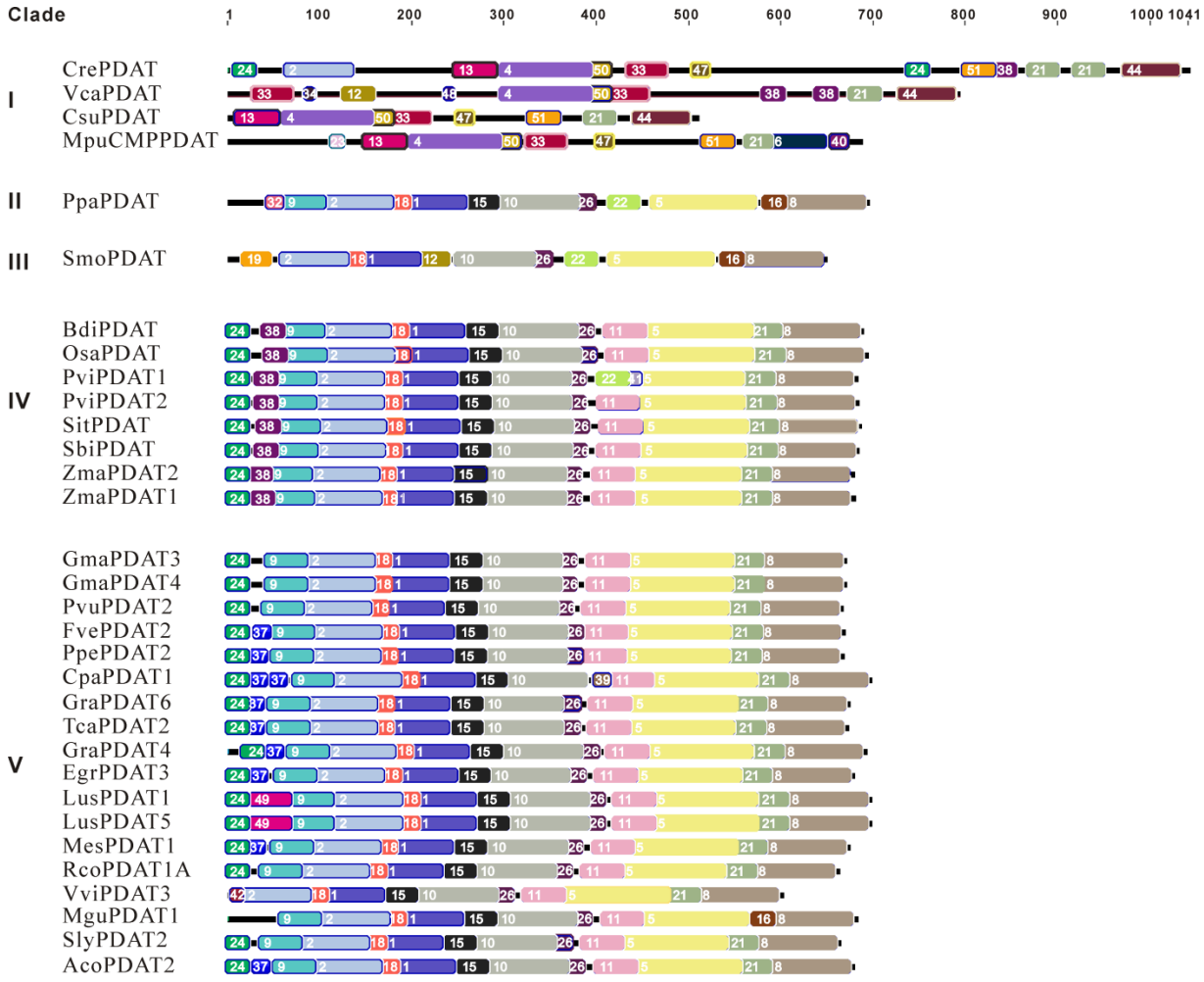
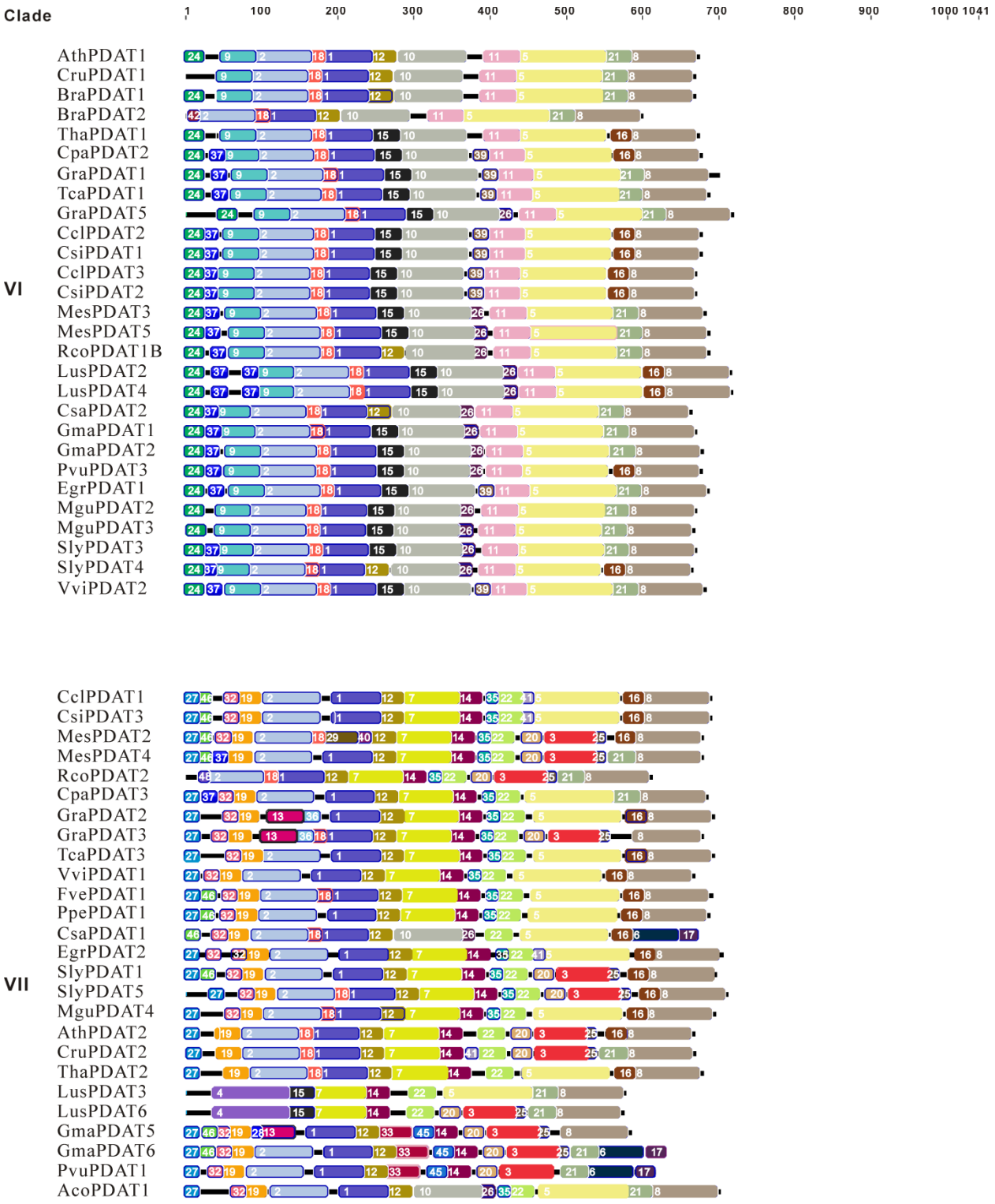

Clade
1
100
200
300
400
500
600
700
800
900
1000 1041
I
CrePDAT
VcaPDAT
CsuPDAT
MpuCMPPDAT
II
PpaPDAT
III
SmoPDAT
IV
BdiPDAT
OsaPDAT
PviPDAT1
PviPDAT2
SitPDAT
SbiPDAT
ZmaPDAT2
ZmaPDAT1
V
GmaPDAT3
GmaPDAT4
PvuPDAT2
FvePDAT2
PpePDAT2
CpaPDAT1
GraPDAT6
TcaPDAT2
GraPDAT4
EgrPDAT3
LusPDAT1
LusPDAT5
MesPDAT1
RcoPDAT1A
VviPDAT3
MguPDAT1
SlyPDAT2
AcoPDAT2

Figure 4.13 Continued



Next, the non-LCAT motif composition in land plant PDATs was examined. Based on the position of the TMD and LCAT-like motifs, PDATs were further divide into four regions (Figure 4.11): Region 1 covers the segment before the TMD; Region 2 spans the segment between the TMD and the first LCAT-like motif; Region 3 defines the region between the third and the fourth LCAT-like motifs; Region 4 corresponds to the C-terminal segment. Among these four regions, regions 2 and 4, mainly made up of motifs 2 and 8, are highly conserved within the land plants. Region 3 is less conserved, mainly composed of motif 22 in Clades II and III, motif 11 in Clades IV, V and VI, and motifs 14 and 22 in Clade VII. Region 1 appears to be the most divergent region, in which the clade-specific motifs, including the motif 38 in Clade IV and the motif 27 in Clade VII, were found. It is worth mentioning that individual MEME motifs are gapless (no insertions or deletions), which means that the motifs containing gaps can still be discovered but they will be split into multiple ungapped motifs. It will be interesting to find out which motifs are really different, and which motifs may be associated with specific functions.

Taken together, the identified LCAT and non-LCAT motif patterns match the cladding pattern in the phylogenetic tree.

4.3.8 Eudicot-wide *PDAT* gene expansion arose mainly from the eudicot-shared ancient gene duplication followed by species-specific segmental duplications

The existence of multiple *PDAT* gene copies across eudicots suggests that the *PDAT* gene family expanded in eudicots. Gene copy number expansions can occur via three major evolutionary events: segmental duplication, tandem duplication, and transposition events (Kong et al., 2007). In this study, I focused on segmental and tandem gene duplications.

The phylogenetic analysis clearly divides the core eudicot *PDATs* into three distinct clades (Figure 4.3). Each clade contains sequences from taxa across the core eudicots, including

both the Asterids and the Rosids, indicating that *PDAT* paralogs among different clades were produced by a core eudicot-shared ancient gene duplication that predated the split of the two major clades of core eudicots. Based on this data, I cannot say with certainty whether this ancient gene duplication was shared with the basal eudicots. This will have to be confirmed in the future when additional genome sequences of basal eudicots are available. The genes derived from the ancient gene duplication are named as "paleo-duplicated genes". Due to the fact that the ancient gene duplication was followed by specie-specific gene duplication, gene loss and chromosome rearrangements, only six species (*Vitis vinifera*, *Theobroma cacao*, *R. communis*, *Carica papaya*, *Eucalyptus grandis*, and *Phaseolus vulgaris*) have maintained the triplicated "paleologous" *PDAT* genes, each of which is present in one of the three core eudicot clades (Clades V to VII) (Figure 4.3 and Table 4.3).

Besides paleo-duplicated *PDAT* paralogs among different clades, some species contain duplicated gene pairs within the clades. These include three *PDAT* gene pairs in *L. usitatissimum*, *Glycine max* and *Gossypium raimondii*, two in *Solanum lycopersicum* and *Manihot esculenta*, one in *Mimulus guttatus* and *Brassica rapa* (Figure 4.3). To determine whether these within-clade gene pairs were derived from segmental duplication events, I analyzed 10 protein-coding genes from upstream and downstream of each *PDAT* gene pair. The results show that the genes flanking each *PDAT* gene pair are highly conserved in all species but *G. raimondii*, indicating that these *PDAT* gene pairs were formed via the segmental duplication event. The difference found in *G. raimondii* may be partly explained by the fact that the *Gossypium* genome has a unique evolutionary history. The lineage-specific whole genome multiplication event(s) has occurred approximately 60 million years ago in *Gossypium* genomes (Paterson et al., 2012).

The within-clade gene pairs tend to have higher sequence identity than the between-clade gene pairs (Tables 4.6 and 4.7). Thus, it was speculated that the within-clade gene pairs were derived from more recent duplication events. To confirm this hypothesis, we used the synonymous substitution rates (*K*s) as a proxy for time to compare the date of gene duplications. Judging from the *K*s values, the within-clade gene pairs have a much lower *K*s than between-clade gene pairs, suggesting more recent duplications (Tables 4.6 and 4.7). Because most of the *K*s values for between-clade gene pairs are saturated (>2), such data can only provide a rough estimate.

Next, the role of tandem duplication in the evolution of the eudicot *PDAT* gene family was investigated. The previous literature indicated that a chromosome region consisting of two or more copies of a gene within 200kb can be viewed as a gene cluster (Houb, 2001). Chromosome location analysis shows that the majority of the *PDAT* genes are located along scattered sites throughout the genome and a single tandem duplication cluster consisting of two genes has only been found in species of *Citrus clementine* (*CclPDAT2* and *CclPDAT3*) and *Citrus sinensis* (*CsiPDAT1* and *CsiPDAT2*). This suggests that tandem duplication does not play a dominant role in the expansion of the *PDAT* gene family in eudicots. Taken together, these analyses reveal that the eudicot-shared ancient gene duplication followed by specie-specific segmental duplication primarily contributes to the expansion of the *PDAT* gene family in eudicots.

Table 4.6 Sequence identity and synonymous substitution rates (*K*s values) of within-clade duplicated *PDAT* gene pairs

Species	Gene Pair	Clade	Sequence Identity (%)	*K*s	Mean *K*s
Glycine max	*GmaPDAT3* vs *GmaPDAT4*	V	96.1	0.08996 ± 0.01655	0.125
	GmaPDAT1 vs *GmaPDAT2*	VI	95.9	0.12586 ± 0.01946	
	GmaPDAT5 vs *GmaPDAT6*	VII	92.0	0.15779 ± 0.02457	
Gossypium raimondii	*GraPDAT4* vs *GraPDAT6*	V	79.1	0.70247 ± 0.06247	0.794
	GraPDAT1 vs *GraPDAT5*	VI	76.4	0.80740 ± 0.07259	
	GraPDAT2 vs *GraPDAT3*	VII	76.9	0.87113 ± 0.07781	
Solanum lycopersicum	*SlyPDAT3* vs *SlyPDAT4*	VI	81.4	0.71665 ± 0.06707	0.755
	SlyPDAT1 vs *SlyPDAT5*	VII	79.7	0.79243 ± 0.06856	
Manihot esculenta	*MesPDAT3* vs *MesPDAT5*	VI	91.0	0.29142 ± 0.03067	0.316
	MesPDAT2 vs *MesPDAT4*	VII	87.6	0.34022 ± 0.03314	
Linum usitatissimum	*LusPDAT1* vs *LusPDAT5*	V	97.1	0.08181 ± 0.01528	0.100
	LusPDAT2 vs *LusPDAT4*	VI	95.8	0.12603 ± 0.01866	
	LusPDAT3 vs *LusPDAT6*	VII	96.2	0.09326 ± 0.01564	
Mimulus guttatus	*MguPDAT2* vs *MguPDAT3*	VI	76.7	1.35287 ± 0.14409	1.353
Brassica rapa	*BraPDAT1* vs *BraPDAT2*	VI	91.5	0.3522	0.352

Table 4.7 Sequence identity and synonymous substitution rates (*K*s values) of between-clade duplicated *PDAT* gene pairs. The *K*s values for individual gene pair can be found in Table 4.5.

Species	Clades	Mean Sequence identity (%)	Mean *K*s
Glycine max	V and VI	74.750±0.5	1.686
	V and VII	61.825±0.443	6.944
	VI and VII	61.375±0.506	69.654
Gossypium raimondii	V and VI	67.825±1.9	2.521
	V and VII	59.575±0.78	8.845
	VI and VII	58.675±1.721	6.686
Solanum lycopersicum	V and VI	72.75 ±0.495	2.542
	V and VII	62±0.566	11.594
	VI and VII	61.775±0.189	17.506
Manihot esculenta	V and VI	75.4±0.424	1.551
	V and VII	61.6±0.424	7.608
	VI and VII	60.95±0.252	6.6
Linum usitatissimum	V and VI	69.4±0.294	15.008
	V and VII	62.675±0.608	63.257
	VI and VII	63.325±0.377	15.013
Mimulus guttatus	V and VI	68.55±1.485	36.623
	V and VII	60.3	5.655
	VI and VII	61.3±0.283	41.615

4.3.9 Different selection pressures have acted on the paleo-duplicated *PDAT* paralogs

It has long been thought that gene duplication plays a crucial role in the evolution of gene diversity (Ohno, 1970; Roth et al., 2007; Hughes, 1994). To explore the evolutionary fate of paleo-duplicated eudicot *PDATs*, selection pressure analyses were performed. The selective pressure acting on the core eudicot PDATs was estimated using the ratio (ω) of the nonsynonymous substitution rate versus the K_S as an indicator. Because two basal eudicot *PDATs* are sister to, rather than nested within the core eudicot clades, they were not included in the analyses. Specifically, I extracted the core eudicot sequences and constructed a phylogenetic tree using *SmoPDAT1* as an outgroup (Figure 4.14).

To address the possibility of functional divergence among the core eudicot clades, the data was fit to the clade model C (CmC) implemented in PAML. In CmC, the entire target clade is set as the foreground partition, while the rest of the phylogeny comprises the background partition. In view of the above phylogenetic trees (Figure 4.3 and Figure 4.14), Clade VII is more distant from Clades V and VI. Therefore, I first applied CmC with the entire Clade VII set as the foreground partition; Clades V, Clade VI along with the outgroup comprised the background partition. This analysis was named as "CmC VII". The CmC model assumes that different selection pressures have acted on the foreground and background partitions; while the null model, M2a_rel, hypothesizes that there is no significant difference in selection pressures between the foreground and background partitions. The likelihood ratio test (LRT) comparing CmC versus the M2a_rel null model shows that CmC VII fits the data significantly better than the null model ($P<0.001$, Table 4.8). Parameter estimates indicate that a larger set of sites (~56%) evolving under stronger purifying selection ($\omega_0 =0.02764$), and a smaller set of sites

(~43%) evolving under divergent selective pressures, with weaker purifying selection in Clade VII (ω_3=0.23272), and stronger purifying selection in the background (ω_2=0.15745).

Similar CmC analysis was further applied to Clade V (referred to as CmC V) and Clade VI (referred to as CmC VI). The LRT results show that CmC VI (P<0.001), but not CmC V (P>0.2), provides a significant better fit than the null model (Table 4.8). However, including both Clades VI and VII in the background partition in CmC V might be inappropriate, as the average of their ω ratio (Clade VI ω_3=0.14847; Clade VII ω_3=0.23272) is close to the ω ratio for Clade V (ω_3=0.18669). To evaluate this possibility, the extended clade model was employed (Yoshida et al., 2011), which allows more than two partitions (foreground and background). Three partitions were specified in this analysis: CladeV, Clade VI and Clade VII. This analysis was referred to as "Ex-CmC". The null hypothesis is that selection pressure is the same for Clade V and Clade VI. In null model testing, the phylogeny was divided into two partitions: Clade VII and the combined Clades V and VI. The null model for this test is named as "Ex-Null". The LRT result (Table 4.8) indicates that the null hypothesis is rejected, supporting that different selection pressures have acted on Clade V and Clade VI. To further confirm this result, the Clade VII was excluded from the analysis and the functional divergence was only compared between Clade V and Clade VI. Using this dataset, the results show that setting either Clade V or Clade VI as the foreground partition yields a significant LRT result (P<0.001, Table 4.9), indicating that selective constraint indeed differs between Clade V and Clade VI.

Taken together, these results indicate that three core eudicot clades have evolved under divergent selection pressures and *PDATs* in Clade VII experienced the lowest selection constraint compared to *PDATs* from the other two clades.

Figure 4.14 Phylogenetic relationship of the core eudicot *PDATs*. The tree was generated using the M0 model implemented in PAML and rooted using PDAT sequences from *S. moellendorffii* as the outgroup. Branch length is defined as number of nucleotide substitutions per codon.

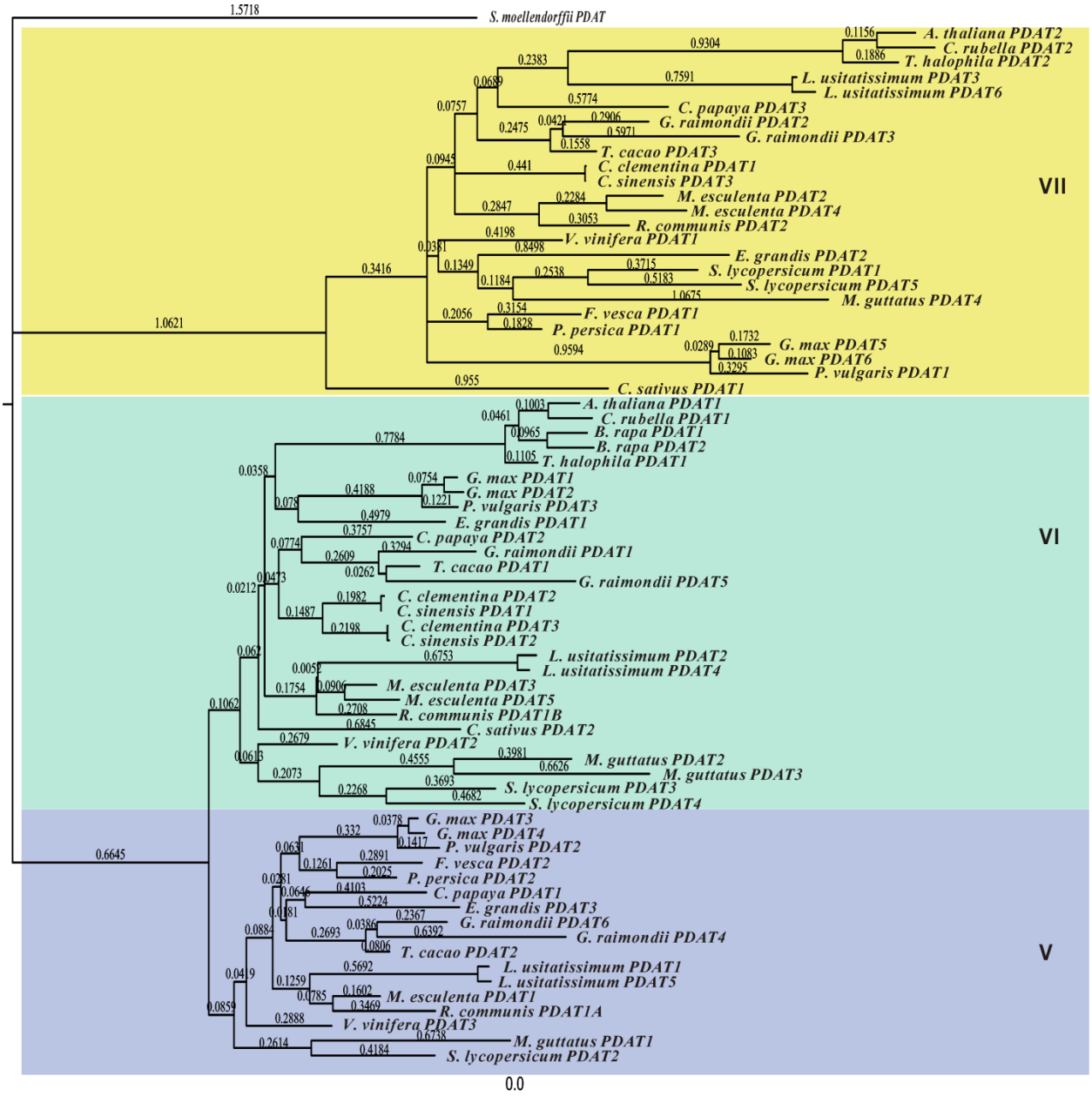


Table 4.8 Parameter estimates, likelihood values and likelihood ratio test (LRT) P values obtained from Clade model (CmC) and extended-Clade model (Ex-CmC) analyses of the 70 core eudicot *PDATs* data set

Model	*ln*L	0: Purifying		1: Neutral		2: Divergent		2Δℓ	*P*
		ω_0	p_0	ω_1	p_1	ω_2, ω_3,(ω_4)	p_2		
M2a_ref	-52695.098844	0.02847	0.57261	1.0000	0.01066	ω_2:0.18723	0.41673		
CmC V	-52695.097019	0.02846	0.57248	1.0000	0.01068	ω_2:0.18733 ω_3:0.18669	0.41684	0.00365	>0.2
CmC VI	-52669.267166	0.02896	0.57538	1.0000	0.00941	ω_2:0.21333 ω_3:0.14847	0.41521	51.663356	<0.001
CmC VII	-52662.297402	0.02764	0.56217	1.0000	0.01026	ω_2:0.15745 ω_3:0.23272	0.42757	65.602884	<0.001
Ex-Null	-52659.898615	0.02687	0.55770	1.0000	0.01064	ω_2:0.23108 ω_3:0.15854	0.43166		
Ex-CmC	-52653.785928	0.02732	0.56258	1.0000	0.00998	ω_2:0.18319 ω_3:0.14584 ω_4:0.23299	0.42743	12.225374	<0.001

*Clade model: ω_0 is the estimated ω value for site class 0, p_0 is the estimated proportion of sites in site class 0 (purifying selection); ω_1 is the estimated ω value for site class 1, p_1 is the estimated proportion of sites in site class 1 (neutral selection); ω_2 is the estimated ω value for divergent sites on the background partitions, ω_3 is the estimated ω value for divergent sites on the foreground partitions, p_2 is the estimated proportion of sites in site class 2 (divergent selection); *ln*L is the log likelihood value; 2Δℓ is the likelihood ratio test (LRT) statistic for comparing the CmC and M2a_ref (null) models; p is the p-value of the LRT.

*Extended-Clade model: The first two site classes (class 0 and class 1) are the same as CmC model. The final class (class 2) modeling divergent selection among three partitions, each with a separately estimated ω ratios (ω_2 for CladeV, ω_3 for Clade VI, and ω_4 for Clade VII). The Null-model has only two partitions and two estimated ω ratios (ω_2 for Clade VII and ω_3 for both Clade V and Clade VI). *ln*L is the log likelihood value; 2Δℓ is the likelihood ratio test (LRT) statistic for comparing the Ex-CmC and Ex-Null models; p is the p-value of the LRT.

Table 4.9 Parameter estimates, likelihood values, and likelihood ratio test (LRT) P values obtained from Clade model (CmC) analyses of 45 *PDATs* from the core eudicot Clades V and VI

Model	*ln*L	0: Purifying		1: Neutral		2: Divergent		2Δℓ	*P*
		ω_0	p_0	ω_1	p_1	ω_2, ω_3	p_2		
M2a_ref	-33307.980006	0.02467	0.6711	1.0000	0.02888	ω_2:0.20335	0.30002		
CmC V	-33298.090883	0.02432	0.66957	1.0000	0.02792	ω_2:0.18046; ω_3:0.24610	0.30252	19.778246	<0.001
CmC VI	-33301.896393	0.02506	0.67361	1.0000	0.02779	ω_2:0.23597; ω_3:0.18545	0.29860	12.167226	<0.001

*ω_0 is the estimated ω value for site class 0, p_0 is the estimated proportion of sites in site class 0 (purifying selection); ω_1 is the estimated ω value for site class 1, p_1 is the estimated proportion of sites in site class 1 (neutral selection); ω_2 is the estimated ω value for divergent sites on the background partitions, ω_3 is the estimated ω value for divergent sites on the foreground partitions, p_2 is the estimated proportion of sites in site class 2 (divergent selection); *ln*L is the log likelihood value; 2Δℓ is the likelihood ratio test (LRT) statistic for comparing the CmC and M2a_ref (null) models; p is the p-value of the LRT.

4.4 Discussion

Despite the fact that many studies have revealed the crucial role of PDATs in TAG biosynthesis, our knowledge on PDATs is still very limited. To advance our understanding of the involvement of PDATs in TAG biosynthesis, it is essential to first understand their evolution and diversity. The goal of this study was to provide an overall picture of plant PDATs, including their gene family members, evolutionary history, present-day diversity, and structural similarities and differences.

Evolution, conservation and variation of the *PDAT* gene family in plants

Taking advantage of publicly available sequenced plant genomes, 40 different plant species were surveyed and 139 LCAT-like sequences were identified. The results of phylogeny (Figure 4.1) and PANTHER classification analyses, along with the previous findings indicated that only the sequences with E-value < 1e-15 from land plants and the sequences from algae belong to the PDAT family and therefore were included for the further analyses. PDAT candidates exist in all plants analyzed, including algae, lowland plants (a moss and a lycophyte) and highland plants (monocots and eudicots). The evolutionary analysis shows that the *PDAT* gene family can be clearly divided into seven major clades (Figure 4.3).

Four algal *PDAT*s form a separate well-supported clade (Clade I) from land plant *PDATs*. This phylogenetic separation is supported by their different gene structure (Figure 4.5), protein properties (Table 4.4) and motif composition (Figure 4.11). The observed differences between algal and land pant PDATs might be associated with different biological functions. Consistent with this hypothesis, the study of the microalga *C. reinhardtii* PDAT (Yoon et al. 2012) revealed that it has some unique features that have not been reported in land plant PDATs. For instance, the algal PDAT appears to be a chloroplast-localized protein with a higher preference for

chloroplast membrane lipids (e.g., phosphatidylglycerol and phosphatidylinositol) over the other phospholipids (phosphatidylcholine and phosphatidylethanolamine). In addition, the algal PDAT is a multifunctional enzyme, which has not only PDAT and DAG:DAG acyltransferase function, but also galactolipid:DAG acyltransferase and lipase activities. It is important to note that the sequences within the algal clade (Clade I) have very long branches, suggesting that the algal *PDATs* are quite different from each other. Due to limited number of algal *PDATs*, this study did not further divide algal *PDATs* and considered them as a single clade in this study. It will be interesting to include more algal sequences to evaluate the extent of divergence among the algal *PDATs*.

Most land plant PDATs share four major structural features at both the gene and protein levels. First, the exon/intron structures (six introns/seven exons) and intron phase patterns (2, 0, 2, 0, 2) are remarkably conserved in most land plant *PDAT* genes (Figure 4.5 and Figure 4.6), suggesting that the *PDAT* gene structure in land plants has been established and retained after the divergence of land plants from algae. Second, a single TMD in the N-terminus has been preserved in most land plant PDATs (Figure 4.7 and Figure 4.8). Third, all LCAT-like motifs (Figure 4.11 and Figure 4.13) and LCAT-conserved amino acid residues (Figure 4.9 and Figure 4.10) are located at the C-terminal end of the TMD, suggesting that the active and/or binding sites of land plant PDATs are possibly facing the luminal side of the ER. Fourthly, the C-terminal portion and the region between the TMD and the first LCAT-like motif are highly conserved (Figure 4.11 and Figure 4.13). It has been reported that two Arabidopsis PDATs contain ER retrieval signals at their C-termini (McCartney et al., 2004); therefore, it is possible that the C-terminal is involved in assisting the association of PDAT with the ER.

Besides similarities, there are variations among land plant PDATs. The alignment of PDAT polypeptides shows that the hydrophilic N-terminus preceding the TMD is the most divergent region (Figure 4.7B). The motif occurrences in the N-terminus are quite unique for PDATs from each clade (Figure 4.11 and Figure 4.13). The Clade IV-specific motif 38 and Clade VII-specific motif 27 were found within this region. Therefore, this region could serve as a candidate target to study the functional and structural divergence among land plant PDATs from different clades. Although deletion of the TMD along with the N-terminus of yeast PDAT does not affect its catalytic activity and substrate selectivity (Ghosal et al., 2007), it is still possible that the N-terminus of land plant PDATs is associated with specific functions, such as sorting PDATs to the ER (Pelham, 2000) and forming a multimeric complex as demonstrated in DGAT1 from plants and animals (McFie et al., 2010; Weselake et al., 2006; Pelham, 2000; Cheng et al., 2001).

This study also reveals a eudicot-wide *PDAT* gene expansion. Combined with evidence from the phylogenetic (Figure 4.3) and *K*s analyses (Tables 4.6 and 4.7), the eudicot-shared ancient gene duplication followed by species-specific segmental duplications appears to be mainly responsible for the expansion of *PDAT* genes in eudicots. The duplicated core eudicot *PDATs* are grouped into three clades (Clades V, VI and VII). The MEME combined block reveals that the motif compositions of PDATs in Clade VII are quite different from those in Clades V and VI (Figure 4.11 and Figure 4.13). A tendency for the conservation of acidic pI values in Clade VI (Tables 4.4 and 4.5) adds another distinct characteristic for the separation among the core eudicot clades. In addition, Kim et al. (2011) observed different subcellular localizations for the proteins encoded by paleo-duplicated *PDAT* paralogs in castor. RcoPDAT1A and RcoPDAT1B were found to be ER-localized, whereas RcoPDAT2 was

proposed to be localized in the plasma membrane. It will be interesting to study the localization of PDAT protein paralogs in different species to determine whether the plasma membrane-localized RcoPDAT2 is an exclusive case.

Has ancient gene duplication led to functional and expression divergence among *PDAT* paralogs?

This study was motivated by the finding that some plant genomes contain multiple *PDAT* paralogs, which show evidence of diverging TAG-synthesizing function. The result of this study showed that there is a eudicot-wide *PDAT* gene expansion, but the questions remain: "why do eudicots contain multiple copies of *PDATs* in their genome?" and "Whether gaining functional divergence among *PDAT* paralogs is a general trend in the evolution of the eudicot *PDAT*s and how this may happen?" Now, I may be able answer these questions from an evolutionary perspective.

Gene duplication is believed to be one of the major driving forces for the evolutionary novelties, including neo-functionalization (Ohno, 1970; Force et al., 1999; Roth et al., 2007) and sub-functionalization (Li et al., 2005; Wang et al., 2012) at the level of expression or coding sequence. A central theory of molecular evolution states that most genes evolved primarily under strong purifying constraints for functional conservation, and gene duplication allows a gene to be free from this selection pressure and eventually accumulate mutations that can lead to new function or complete loss of function (Ohno, 1970; Lynch and Conery, 2000). The selection pressure analyses show that: (1) strong purifying selection is a primary evolutionary mode for the core eudicot *PDATs,* and (2) after ancient gene duplication, paleo-duplicated *PDAT* genes have been subjected to different selective constraints (Tables 4.8 and 4.9).

The observed heterogeneity in selection pressure among the core eudicot clades might enable the changes in the genes' functions and/or the development of expression-level divergence among duplicated genes. Consistent with this hypothesis, the previous studies (Ståhl et al., 2004; Zhang et al., 2009; van Erp et al., 2011; Kim et al., 2011; Pan et al., 2013) showed that PDATs, including Arabidopsis, flax and castor PDATs, from Clade VII (AthPDAT2, LusPDAT3, LusPDAT6, RcoPDAT2) do not have an apparent function in TAG biosynthesis. Clade VII under the weakest selection constraint (Table 4.8) seems to have evolved in a manner very different from Clades V and VI, and may have eventually lost the TAG-synthesizing function. In addition to the possible non-functionalization of Clade VII, previous studies (Ståhl et al., 2004; Zhang et al., 2009; van Erp et al., 2011; Kim et al., 2011; Pan et al., 2013) also provide some lines of evidence suggesting that expression divergence may have occurred between the *PDAT* paralogs from another two core eudicot clades (Clade V and Clade VI). More interestingly, many *PDATs* in CladeVI appear to have been sub-functionalized at the expression level into the non-seed tissues. Studies on flax (Pan et al., 2013) and castor *PDATs* (Kim et al., 2011) revealed that the *PDAT* paralogs sitting on Clade VI (*LusPDAT2*, *LusPDAT4* and *RcoPDAT1B*) are expressed at a high level in non-seed tissues and very low expression levels are detected in seeds. In this study, the RNA-sequencing (RNA-Seq) data was also used to examine the expression profile of *PDATs* in soybean (*G. max*) and common bean (*P. vulgaris*). The results show that soybean and common bean *PDAT* paralogs in Clade VI (*GmaPDAT1*,*GmaPDAT2* and *PvuPDAT3*) have significantly higher expression in leaves, flowers and roots than in developing seeds (Figures 4.15 and 4.16). Although similar expression levels of Arabidopsis *PDAT1* (from Clade VI) in leaves, roots, flowers and developing seeds were reported, a semi-quantitative reverse transcription-PCR approach with seeds at single

developmental stage (mid stage) was used in the study (Ståhl et al., 2004). To obtain a more detailed expression data, the microarray expression data for *AthPDAT1* was extracted from the AtGenExpress database (http://jsp.weigelworld.org/expviz/expviz.jsp). The result (Figure 4.17) shows that *AthPDAT1* does appear to have followed the trend of tissue-subfunctionalization and the expression level is generally higher in other tissues than in seeds across the different developmental stages. In addition, previous studies revealed that the TAG-synthesizing function of AthPDAT1 has only been detected in rapidly developing leaves rather than in seeds under both over-expression and RNAi approaches (Mhaske et al., 2005; Fan et al., 2013). A fairly recent study (Fan et al., 2014) indicated that the AthPDAT1-mediated TAG synthesis is involved in the process of diverting FAs from membrane lipids towards peroxisomal beta-oxidation, thereby maintaining membrane lipid homeostasis in Arabidopsis leaves. PDATs from Clade VI, which are closely related to AthPDAT1, may have a similar protective role in maintaining membrane integrity in leaf tissues. Furthermore, the studies on mammalian proteins suggest that the shifts in the pI values may be due to the functional divergence of proteins (Khaldi and Shields, 2011) or an adaptation to the changed subcellular localization or tissue compartmentalization (Alendé et al., 2011). Therefore, a tendency for the conservation of acidic pI values in Clade VI may be related to the tissue-subfunctionalization. At the same time, it must be noted that three *PDATs* from the third core eudicot clade (Clade V) characterized to date, including *LusPDAT1*, *LusPDAT5* and *RcoPDAT1A*, have seed-specific expression patterns and the encoded enzymes have unique substrate selectivity properties (Pan et al., 2013; Kim et al., 2011). The identified ALA-preferring flax PDATs (Pan et al., 2013) and hydroxy FA-selective castor PDAT (van Erp et al., 2011) support the speculation that: 1) the substrate selectivity of PDAT has co-evolved with the species' FA composition (Yoon et al., 2012) and 2) the

contribution of PDAT from Clade V to seed oil synthesis can be significant in some oilseeds that are high in polyunsaturated FAs or unusual FAs. It will be very exciting to learn whether other PDATs from Clade V have been specialized in seeds and developed unique substrate selectivity. In contrast, the *PDATs* found in algae, moss, lycophyte and monocots form monophyletic clades. Two *PDAT* paralogs from maize (*Z. mays*) have similar expression pattern (Figure 4.18). The *PDAT* expression profile in rice (*Oryza sativa*) did not show any significantly differential expression among tissues (Figure 4.18). The functional and expression divergence of *PDAT*s appears to be core eudicot-specific.

This study proposed answers to the above questions. Functional and expression divergence of *PDAT* paralogs appears to be a general trend in the evolution of the core eudicot *PDAT*s. Ancient gene duplication may enable one of paleo-duplicated *PDAT* paralogs to become non-functionalized and another two paralogs to develop divergent expression pattern. Since it has been well recognized that gain of functional diversification and expression-level divergence is a key process in promoting retention of duplicated genes in the genome (Lynch and Force, 2000; Li et al., 2005; Torgerson and Singh, 2004; Roth et al., 2007), non-functionalization and expression divergence among paralogs may account for the retention of multiple copies of *PDAT*s in the core eudicots.

Figure 4.15 RNA-Seq digital gene expression analysis of *PDAT* paralogs in soybean (*Glycine max*). The digital gene expression counts were obtained using RNA-Seq Atlas of *G. max* via the web http://soybase.org/soyseq/. The black bars show the expression counts of *GmaPDATs* in the non-seed tissues, including leaves (L), flowers (F), and roots (R), while the grey bars show the expression counts in seeds at various developmental stages, ranging from 10 to 42 days after anthesis. The expression of *GmaPDAT6* is undetected by RNA-seq.

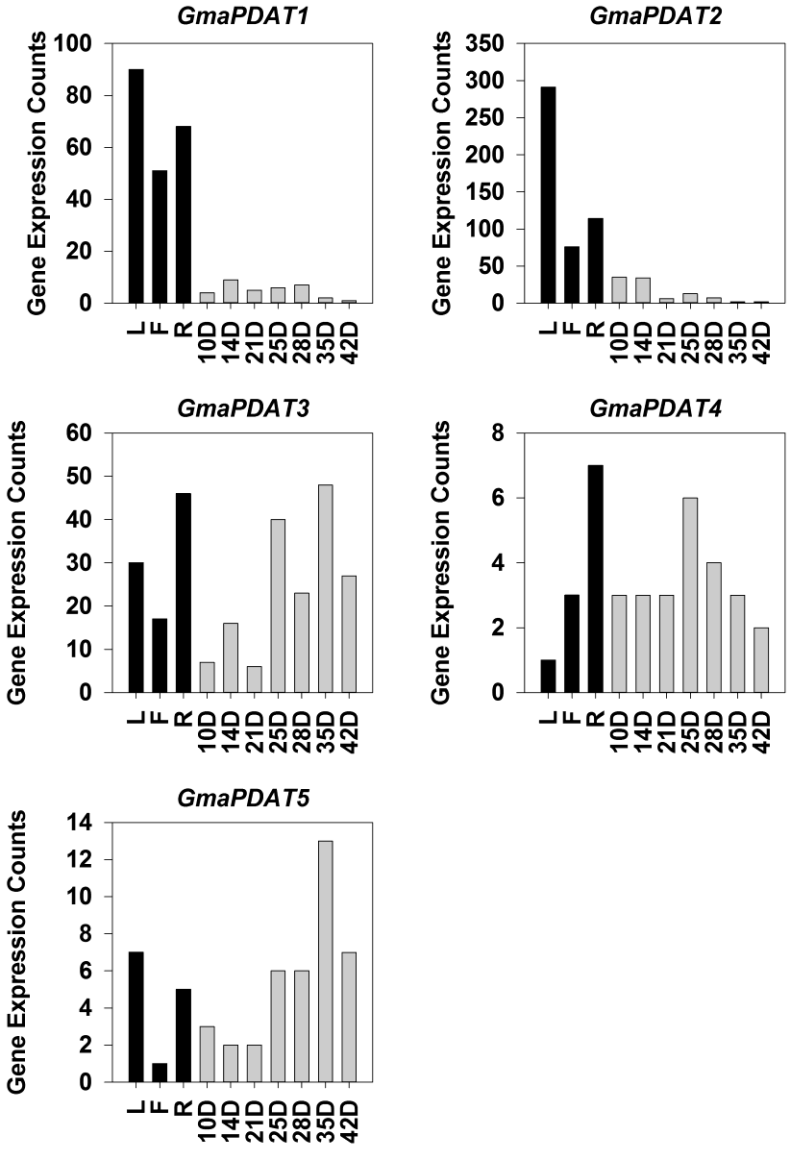


Figure 4.16 RNA-Seq digital gene expression analysis of *PDAT* paralogs in *Phaseolus vulgaris*. The digital gene expression counts were obtained using an RNA-Seq based gene expression atlas of the common bean *P. vulgaris* via the web http://plantgrn.noble.org/PvGEA/SearchVisual.jsp. The black bars show the expression counts of *PvuPDATs* in the non-seed tissues, while the grey bars show the expression counts in seeds, across different developmental stages. The expression of *PvuPDAT1* is undetected by RNA-seq. Gene expression profile from samples isolated from plants under ideal growth conditions is presented. YL- Fully expanded 2nd trifoliate leaf tissue; LF – Developing leaf tissue; FY- Young flowers, collected prior to floral emergence; YR- Whole roots collected at the 2nd trifoliate stage of development; RF- Whole roots; SH- Heart stage seeds, between 3 and 4 mm across and approximately 7 mg; S1- Stage 1 seeds, between 6 and 7 mm across and approximately 50 mg; S2- Stage 2 seeds, between 8 and 10 mm across and between 140 and 150 mg.

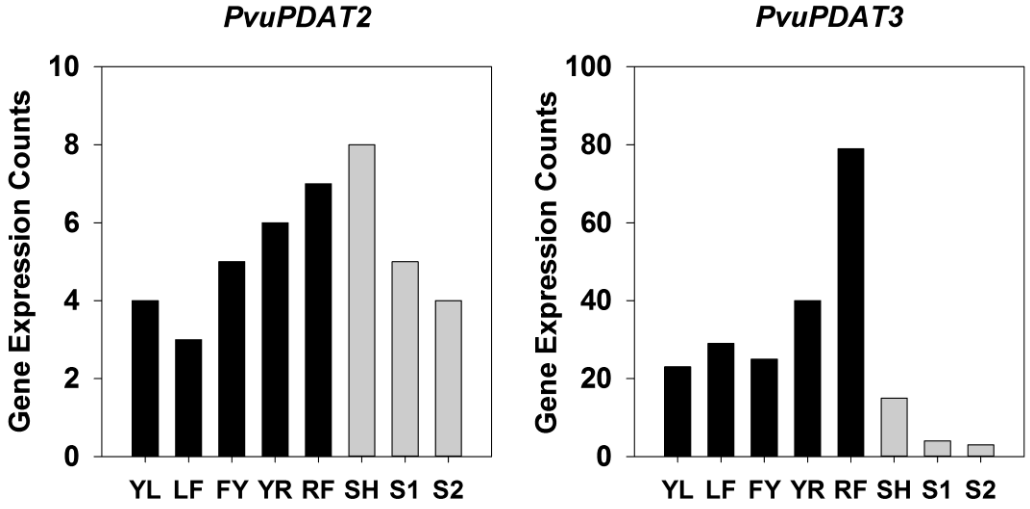


Figure 4.17 Gene expression profile of *PDAT* paralogs in *Arabidopsis thaliana*. The mean-normalized expression values were obtained from the AtGenExpress microarray database via the web http://jsp.weigelworld.org/expviz/expviz.jsp. The expression data was collected from different tissues, including leaves (L), flowers (F), roots (R) and seeds (S), across different developmental stages.

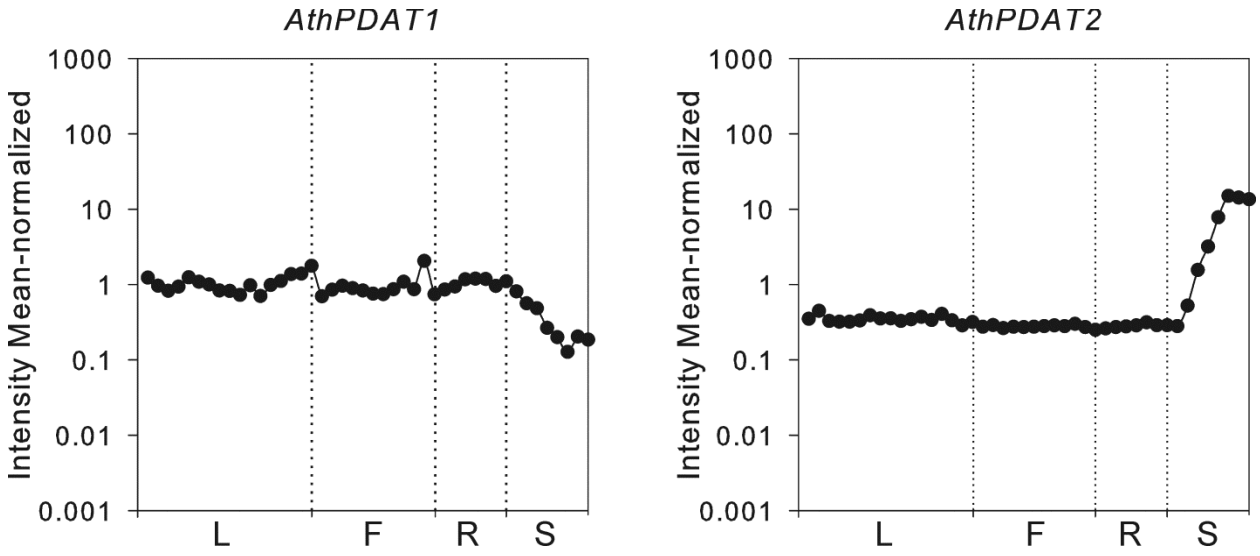
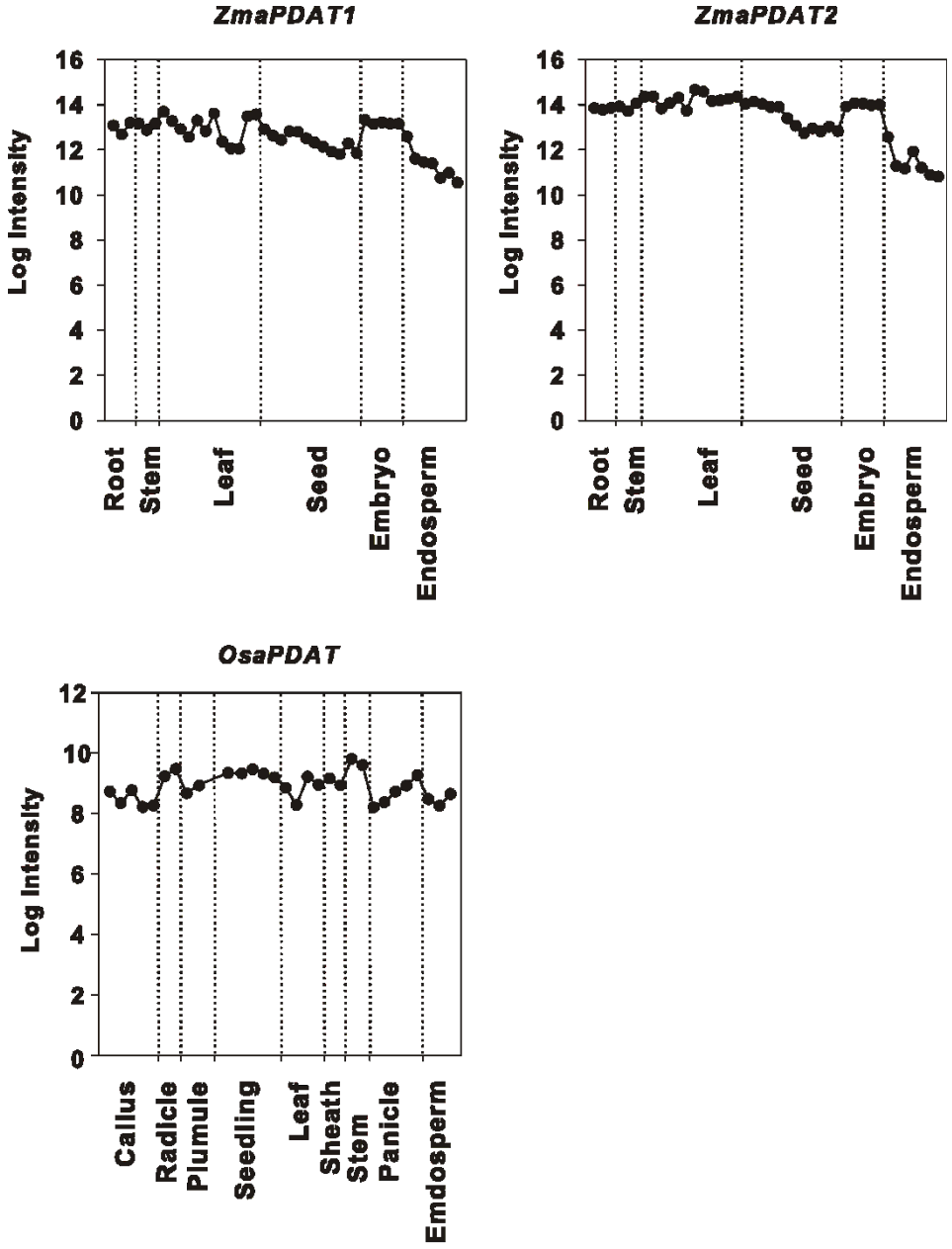


Figure 4.18 Gene expression profile of *PDAT(s)* in maize (*Zea mays*) and rice (*Oryza sativa*). The mean-normalized expression values were obtained from the Plant Expression Database (PLEXdb, http://www.plexdb.org/index.php). The expression data was collected from different tissues across different developmental stages.



4.5 Conclusion

In conclusion, this study provides a comprehensive genomic analysis of the *PDAT* gene family in plants, covering phylogeny, gene structure, protein properties, topology, critical amino acid identification, functional motifs and selection pressure analyses. Phylogenetic analysis indicates that plant *PDATs* can be clustered into seven distinct clades, which is further supported by conservation and variation in gene structure, protein properties, motif occurrences, and/or functional divergence among clades. In addition, selection pressure analyses demonstrate that paleo-duplicated core eudicot *PDATs* have evolved under different selection constraints. Combined with the insights of previous studies, the observed variation in selection constraints might have led to non-functionalization and expression divergence of paleo-duplicated *PDAT* paralogs. Our current knowledge regarding the functions of plant PDATs is limited to only four PDATs: one in the unicellular green alga *C. reinhardtii* and three in core eudicots (van Erp et al., 2011; Pan et al., 2013; Yoon et al., 2012; Ståhl et al., 2004). To obtain a more thorough understanding of the evolution of the plant *PDAT* family, further sampling, expression profile analyses and functional characterization of PDATs in more species will be necessary.

4.6 References

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res **38:** W7-W13

Akaike H (1974) New look at the statistical model identification. IEEE Trans Autom Control **AC-19:** 716-723

Alendé N, Nielsen JE, Shields DC, Khaldi N (2011) Evolution of the isoelectric point of mammalian proteins as a consequence of indels and adaptive evolution. Proteins: Struct , Funct , Bioinf **79:** 1635-1648

Anisimova M, Kosiol C (2009) Investigating protein-coding sequence evolution with probabilistic codon substitution models. Mol Biol Evol **26:** 255-271

Arabolaza A, Rodriguez E, Altabe S, Alvarez H, Gramajo H (2008) Multiple pathways for triacylglycerol biosynthesis in *Streptomyces coelicolor*. Appl Environ Microbiol **74:** 2573-2582

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings /International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology **2:** 28-36

Banas A, Carlsson AS, Huang B, Lenman M, Banas W, Lee M, Noiriel A, Benveniste P, Schaller H, Bouvier-Navé P, Stymne S (2005) Cellular sterol ester synthesis in plants is performed by an enzyme (phospholipid:sterol acyltransferase) different from the yeast and mammalian acyl-CoA:sterol acyltransferases. J Biol Chem **280:** 34626-34634

Bielawski JP, Yang Z (2004) A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J Mol Evol **59:** 121-132

Cao H (2011) Structure-function analysis of diacylglycerol acyltransferase sequences from 70 organisms. BMC Res Notes **4:** 249-249

Cases S, Smith SJ, Zheng Y, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese Jr. RV (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A **95:** 13018-13023

Cheng D, Meegalla RL, He B, Cromley DA, Billheimer JT, Young PR (2001) Human acyl-CoA:diacylglycerol acyltransferase is a tetrameric protein. Biochem J **359:** 707-714

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Drummond A, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2013) Geneious v5.3.6. Available from: http://www.geneious.com/

Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M (2010) A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in *Euonymus* and transgenic seeds. Proc Natl Acad Sci U S A **107:** 9464-9469

Fan J, Yan C, Roston R, Shanklin J, Xu C (2014) Arabidopsis lipins, PDAT1 ACYLTRANSFERASE, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward ß-oxidation, thereby maintaining membrane lipid homeostasis. Plant Cell **26:** 4119-4134

Fan J, Yan C, Zhang X, Xu C (2013) Dual role for phospholipid: diacylglycerol acyltransferase: Enhancing fatty acid synthesis and diverting fatty acids from membrane lipids to triacylglycerol in Arabidopsis leaves. Plant Cell **25:** 3506-3518

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151:** 1531-1545

Ghosal A, Banas A, Ståhl U, Dahlqvist A, Lindqvist Y, Stymne S (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyltransferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate specificities. Biochim Biophs Acta **1771:** 1457-1463

Guo AY, Zhu QH, Chen X, Luo JC (2007) GSDS: a gene structure display server. Yi Chuan **29:** 1023-1026

Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. J Syst Evol **46:** 239-257

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Houb EB (2001) The arms race is ancient history in Arabidopsis, the wildflower. Nat Rev Genet **2:** 516-527

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:** 754-755

Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond Ser B Biol Sci **256:** 119-124

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: Genome-scale protein function classification. Bioinformatics **30:** 1236-1240

Khaldi N, Shields DC (2011) Shift in the isoelectric-point of milk proteins as a consequence of adaptive divergence between the milks of mammalian species. Biology Direct **6:** 40-40

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, DePamphilis CW (2007) Patterns of gene duplication in the plant SKP1 gene family in angiosperms: Evidence for multiple mechanisms of rapid gene birth. Plant J **50:** 873-885

Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol **305:** 567-580

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol **148:** 97-107

Li WH, Yang J, Gu X (2005) Expression divergence between duplicate genes. Trends Genet **21:** 602-607

Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ (2012) Acyl-CoA:diacylglycerol acyltransferase: Molecular biology, biochemistry and biotechnology. Prog Lipid Res **51:** 350-377

Lung SC, Weselake RJ (2006) Diacylglycerol acyltransferase: A key mediator of plant triacylglycerol synthesis. Lipids **41:** 1073-1088

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science **290:** 1151-1155

Lynch M, Force A (2000) The probability of duplicate gene preservation by subfunctionalization. Genetics **154:** 459-473

Maddison WP (1997) Gene trees in species trees. Syst Biol **46:** 523-536

McCartney AW, Dyer JM, Dhanoa PK, Kim PK, Andrews DW, McNew JA, Mullen RT (2004) Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. Plant J **37:** 156-173

McFie PJ, Stone SL, Banman SL, Stone SJ (2010) Topological orientation of acyl-CoA:diacylglycerol acyltransferase-1 (DGAT1) and identification of a putative active site histidine and the role of the N terminus in dimer/tetramer formation. J Biol Chem **285:** 37377-37387

Mhaske V, Beldjilali K, Ohlrogge J, Pollard M (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (*At5g13640*). Plant Physiol Bioch **43:** 413-417

Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res **41:** D377-D386

Ohno S (1970) Evolution by gene duplication. Springer-Verlag, New York

Page RDM, Charleston MA (1997) From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. Mol Phylogenet Evol **7:** 231-240

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo MJ, Byers R, Chen W, Doron-Faigenboim A, Duke MV, Gong L, Grimwood J, Grover C, Grupp K, Hu G, Lee TH, Li J, Lin L, Liu T, Marler BS, Page JT, Roberts AW, Romanel E, Sanders WS, Szadkowski E, Tan X, Tang H, Xu C, Wang J, Wang Z, Zhang D, Zhang L, Ashrafi H, Bedon F, Bowers JE, Brubaker CL, Chee PW, Das S, Gingle AR, Haigler CH, Harker D, Hoffmann LV, Hovav R, Jones DC, Lemke C, Mansoor S, Rahman MU, Rainville LN, Rambani A, Reddy UK, Rong JK, Saranga Y, Scheffler BE, Scheffler JA, Stelly DM, Triplett BA, Van Deynze A, Vaslin MFS, Waghmare VN, Walford SA, Wright RJ, Zaki EA, Zhang T, Dennis ES, Mayer KFX, Peterson DG, Rokhsar DS, Wang X, Schmutz J (2012) Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature **492:** 423-427

Peelman F, Verschelde J, Vanloo B, Ampe C, Labeur C, Tavernier J, Vandekerckhove J, Rosseneu M (1999) Effects of natural mutations in lecithin:cholesterol acyltransferase on the enzyme structure and activity. J Lipid Res **40:** 59-69

Peelman F, Vinaimont N, Verhee A, Vanloo B, Verschelde J, Labeur C, Seguret-Mace S, Duverger N, Hutchinson G, Vandekerckhove J, Tavernier J, Rosseneu M (1998) A proposed architecture for lecithin cholesterol acyltransferase (LCAT): Identification of the catalytic triad and molecular modeling. Protein Sci **7:** 587-599

Pelham HRB (2000) Using sorting signals to retain proteins in endoplasmic reticulum. Methods Enzymol **327:** 279-283

Posada D (2008) jModelTest: Phylogenetic model averaging. Mol Biol Evol **25:** 1253-1256

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. Nucleic Acids Res **40:** D290-D301

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of Arabidopsis thaliana, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

Roth C, Rastogi S, Arvestad L, Dittmar K, Light S, Ekman D, Liberles DA (2007) Evolution after gene duplication: Models, mechanisms, sequences, systems, and organisms. J Exp Zool B Mol Dev Evol **308:** 58-73

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 Physiol **141:** 1533-1543

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:** 2688-2690

Suyama M, Torrents D, Bork P (2006) PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res **34:** W609-W612

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

Torgerson DG, Singh RS (2004) Rapid evolution through gene duplication and subfunctionalization of the testes-specific α4 proteasome subunits in drosophila. Genetics **168:** 1421-1432

Turchetto-Zolet AC, Maraschin FS, De Morais GL, Cagliari A, Andrade CMB, Margis-Pinheiro M, Margis R (2011) Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. BMC Evol Biol **11:** 263-263

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

Wang Y, Wang X, Paterson AH (2012) Genome and gene duplications and gene expression divergence: A view from plants. Ann N Y Acad Sci **1256:** 1-14

Weadick CJ, Chang BSW (2012) An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. Mol Biol Evol **29:** 1297-1300

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Weselake RJ, Madhavji M, Szarka SJ, Patterson NA, Wiehler WB, Nykiforuk CL, Burton TL, Boora PS, Mosimann SC, Foroud NA, Thibault BJ, Moloney MM, Laroche A, Furukawa-Stoffer TL (2006) Acyl-CoA-binding and self-associating properties of a recombinant 13.3 kDa N-terminal fragment of diacylglycerol acyltransferase-1 from oilseed rape. BMC Biochem **7:** 24-24

Worberg A, Quandt D, Barniske A, Löhne C, Hilu KW, Borsch T (2007) Phylogeny of basal eudicots: Insights from non-coding and rapidly evolving DNA. Org Divers Evol **7:** 55-77

Yang Z (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci **13:** 555-556

Yang Z, Bielawski JR (2000) Statistical methods for detecting molecular adaptation. Trends Ecol Evol **15:** 496-503

Yang Z, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol **22:** 1107-1118

Yoon K, Han D, Li Y, Sommerfeld M, Hu Q (2012) Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *chlamydomonas reinhardtii*. Plant Cell **24:** 3708-3724

Yoshida I, Sugiura W, Shibata J, Ren F, Yang Z, Tanaka H (2011) Change of positive selection pressure on HIV-1 envelope gene inferred by early and recent samples. PLoS ONE **6:** e18630-e18630

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Chapter 5

In vivo and *In vitro* Evidence for Biochemical Coupling of Reactions Catalyzed by Lysophosphatidylcholine Acyltransferase and Diacylglycerol Acyltransferase

5.1 Introduction

Seed oils are major plant-based sources of essential polyunsaturated fatty acids (PUFAs) for human diet and many industrial applications. Understanding the fundamental aspects of how plant fatty acids (FAs) are synthesized in seed oils is essential for developing new strategies to alter FA composition for improved nutritional quality and industrial uses of seed oils.

In oilseeds, FA synthesis occurs exclusively in the plastids and newly synthesized FAs, mostly as oleic acid (OA;18:1$^{cis\Delta 9}$), palmitic acid (16:0) and stearic acid (18:0), are then almost entirely (>95%) exported to the cytoplasm to enter into the acyl-CoA pool (Ohlrogge and Browse, 1995). In the cytoplasm, the 18:1 from the acyl-Coenzyme A (CoA) pool is quickly esterified to the phosphatidylcholine (PC) of the endoplasmic reticulum (ER). Desaturation of 18:1 to linoleic (LA;18:2$^{cis\Delta 9,12}$) and then α–linolenic acid (ALA; 18:3$^{cis\Delta 9,12,15}$) occurs through the catalytic action of ER-localized FA desaturases, the oleate desaturase (FAD2) and the linoleate desaturase (FAD3), producing PUFA-enriched PC (Sperling, et al., 1993). It has been shown that PC also is the major site to synthesize a number of unusual FAs, such as hydroxy- and conjugated FAs (Carlsson, et al., 2011).

As shown in Figure 5.1, various pathways have been proposed for transferring PUFAs from PC into triacylglycerol (TAG), which is the main component of seed oil. PUFAs may be directly transferred from PC into TAG via the catalytic action of phospholipid:diacylglycerol acyltransferase (PDAT) (Dahlqvist, et al., 2000). Some oilseeds that are high in PUFAs like flax (Pan, et al., 2013) or unusual FAs like castor (*Ricinus communis*) (Kim, et al., 2011; van Erp, et

al., 2011) appear to have unique PDAT(s) with high selectivity towards these FAs. Alternatively, the phosphocholine head group of PUFA-rich PC may be removed, producing PUFA-enriched DAG, which may be further incorporated into TAG via either or both of PDAT and diacylglycerol acyltransferase (DGAT)-catalyzed reactions. The conversion of PUFA-rich PC to PUFA-rich DAG may also involve the catalytic action of phospholipase C, phospholipase D along with phosphatidic acid phosphatase (PAP), and the recently discovered phosphatidylcholine: diacylglycerol cholinephosphotransferase (PDCT) (Hu, et al., 2012). In addition, PUFAs may be removed from PC and enter the acyl-CoA pool via acyl-exchange predominately between the acyl groups at the *sn*-2 position of PC and the acyl-CoA pool mediated by the combined forward and reverse reactions catalyzed by acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT) (Stymne and Stobart, 1984). PUFA-enriched acyl-CoA may also be generated through liberation of free FA from the *sn*-2 position of PC by phospholipase A (PLA) followed by activation of free FA to acyl-CoA through acyl-CoA synthetase (ACS) with LPCAT acting in the forward reaction to catalzye the re-acylation of lysophosphatidylcholine (LPC) to PC in a process known as the "Lands cycle" (Lands, 1958; Lands, 1960). The resulting PUFA-CoAs may be used by enzymes involved in the Kennedy pathway to form PUFA-enriched TAG. Recently, by using reverse genetics and metabolic labeling approaches, it was demonstrated that LPCAT-mediated acyl-editing together with PDCT-based PC-*sn*-1,2-diacylglycerol (DAG) interconversion control the major fluxes of PUFAs from PC to TAG (approximately 66%) in Arabidopsis seeds (Bates, et al., 2012). In addition, *in vitro* evidence has been reported that plant LPCAT does indeed operate in a reversible fashion, even though the reverse reaction is a thermodynamically unfavorable process

(Lager, et al., 2013). Despite recent progress, direct evidence of how LPCAT interacts with the enzymes involved in the Kennedy pathway is still lacking.

This study focused on the interaction between flax LPCAT1 and a type 1 DGAT. LPCAT can catalyze both forward and reverse reactions (Reaction 1, Figure 5.2). The reverse reaction of LPCAT uses PC as substrate and forms acyl-CoA and LPC. DGAT catalyzes the final step in the Kennedy pathway of TAG synthesis from the substrates acyl-CoA and diacylglycerol (DAG) (Reaction 2, Figure 5.2). The reverse reaction of LPCAT and the DGAT reaction share a common intermediate, acyl-CoA.

Based on reactions 1 and 2 (Figure 5.2), it was proposed that the PUFA synthesized on PC can be removed and transferred to acyl-CoA pool by the reverse reaction of LPCAT, and then further be utilized by the DGAT-catalyzed reaction for TAG biosynthesis. As highlighted in Figure 5.1, the overall reaction represents a route for transferring PUFAs from PC to the acyl-CoA pool, and then to TAG. This hypothesis was tested through heterologous expression of flax genes in yeast mutant background devoid of LPCAT activity and TAG-forming capacity. Flax oil naturally contains substantial amount of PUFAs, especially ALA. To accumulate a large proportion of PUFAs in seed oils, flax may have more efficient mechanisms to channel PUFAs from PC into TAG as compared to other plant species. Many TAG synthesis-related genes from flax are closely related to genes found in plant species that produce high amount of unusual FAs, such as castor and tung tree (*Vernicia fordii*) (Pan, et al., 2013). The PUFA-enriching mechanism found in flax may also represent a route for enriching unusual FAs in these species.

Figure 5.1 A schematic diagram showing the proposed mechanisms involved in transferring polyunsaturated fatty acids (PUFAs) from phosphatidylcholine (PC) into triacylglycerol (TAG). The Kennedy pathway for TAG synthesis is indicated by white arrows and acyl transfer reactions are shown in dashed lines. The transfer of PUFAs from PC into TAG catalyzed by coupled reactions of the LPCAT-catalyzed reverse reaction and the DGAT-catalyzed reaction is highlighted in black bold. Abbreviations: ACS, acyl-CoA synthetase; DGAT, acyl-CoA: diacylglycerol acyltransferase; ER, endoplasmic reticulum; FAD, fatty acid desaturase; FAX1, fatty acid export 1; FFA, free fatty acid; G3P, glycerol-3-phosphate; GPAT, acyl-CoA: glycerol-3-phosphate acyltransferase; LPA, lysophosphatidic acid; LPAAT, acyl-CoA: lysophosphatidic acid acyltransferase; LPC, lysophosphatidylcholine; LPCAT, acyl-CoA:lysophosphatidylcholine acyltransferase; PA, phosphatidic acid; PAP, phosphatidic acid phosphatase; PDCT, phosphatidylcholine: diacylglycerol cholinephosphotransferase; PLA, phospholipase A; PLC, phospholipase C; PLD, phospholipase D.

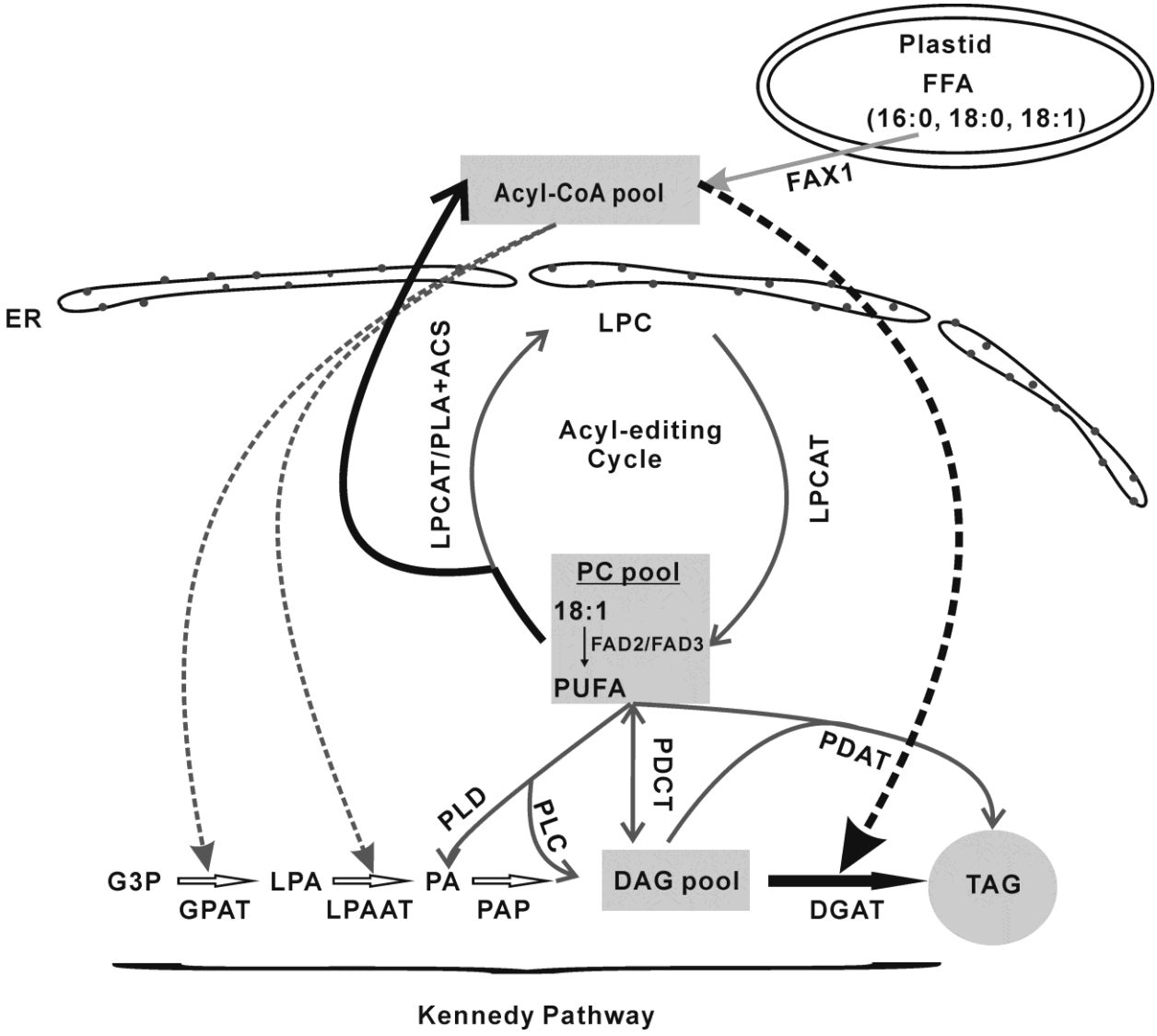
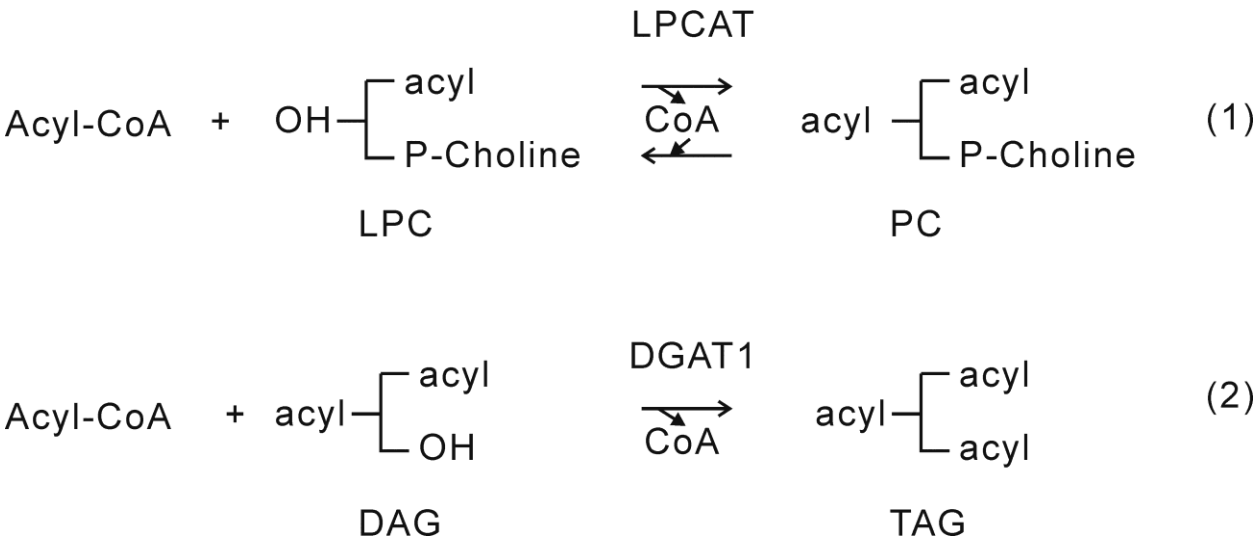

Plastid
FFA
(16:0, 18:0, 18:1)
FAX1
Acyl-CoA pool
ER
LPC
LPCAT/PLA+ACS
Acyl-editing
Cycle
LPCAT
PC pool
18:1
FAD2/FAD3
PUFA
PDCT
PDAT
PLD
PLC
G3P
GPAT
LPA
LPAAT
PA
PAP
DAG pool
DGAT
TAG
Kennedy Pathway

Figure 5.2 A schematic diagram showing the LPCAT-catalyzed forward and reverse reactions (Reaction 1) and the DGAT-catalyzed reaction for triacylglycerol (TAG) synthesis (Reaction 2). Abbreviations: CoA, coenzyme A; DAG, diacylglycerol; DGAT, acyl-CoA: diacylglycerol acyltransferase; LPCAT, acyl-CoA:lysophosphatidylcholine acyltransferase; LPC, lysophosphatidylcholine; PC, phosphatidylcholine.



Taken together with earlier reported *in vitro* data on plant LPCAT action (Lager, et al., 2013), data obtained from this study further confirm that plant LPCAT can act in a reversible fashion under *in vivo* condition. These data demonstrate for the first time *in vivo* evidence of the reversibility of the LPCAT. More interestingly, the *in vivo* and *in vitro* data support the hypothesis of biochemical coupling of the reverse reaction catalyzed by LPCAT with the DGAT-catalyzed reaction for the incorporation of PUFAs into TAG. These results contribute to our understanding of the mechanisms responsible for channeling PUFAs from PC into TAG in oleaginous plants that produce PUFA-enriched TAG during seed development.

5.2 Materials and Methods

5.2.1 Chemicals

[1-^{14}C]18:0, [1-^{14}C]18:1, [1-^{14}C]18:2, [1-^{14}C]18:3 and [^{14}C]18:1-LPC were purchased from PerkinElmer Life Sciences. Nonradioactive FAs were obtained from Nu-Chek Prep. *sn*-1-16:0 LPC and *sn*-1-16:1 LPC were purchased from Avanti Polar lipids. CoA trilithium salt, ACS and DAG were obtained from Sigma. [^{14}C]acyl-CoAs were prepared according to the method described by Taylor et al. (1990). *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC were prepared according to the assay used for measuring the forward reaction of LPCAT as described by Lager et al. (2013) (see the details under "*in vitro* enzyme assays" for LPCAT forward reaction).

5.2.2 Identification and isolation of *LPCAT* genes

BLAST (Altschul, et al., 1990) searches were performed against the flax genomic (Wang, et al., 2012) and expressed sequence tag (EST) (Venglat, et al., 2011) databases by using AtLPCAT1 (At1G12640) and AtLPCAT2 (At1G63050) as the protein queries. To isolate the *LPCAT* genes, total RNA was extracted from the developing embryo of flax (cultivar CDC Bethune) at 14 days post anthesis using the Plant RNeasy plant mini kit (Qiagen), according to

the manufacturer's instruction. First-strand cDNA was prepared using SuperScript III Reverse Transcriptase (Invitrogen) as described by the manufacturer. The resulting cDNA was used as a template in PCR with gene specific primer pairs (LF/ L1R for *LPCAT1*, LF/L2R for *LPCAT2*; Table 5.1) to amplify the *LPCAT* genes. The resulting PCR products were then subcloned into a pYES2.1/V5-HIS vector using pYES2.1 TOPO kit (Invitrogen) according to the protocol provided by the supplier. The plasmids are referred to as pYESTOPO-*LPCATs*.

5.2.3 Dating the segmental duplication events

To date the segmental duplication events, the K_S values (synonymous substitution rates) of the duplicated genes and 20 homologous gene pairs from their upstream and downstream flanking regions were used to calculate the average K_S. The full-length cDNA sequences of the homologous gene pair were first aligned based on their corresponding amino acid translations using Translator X server (Abascal, et al., 2010) (http://translatorx.co.uk/). Then, the resulting alignment was further subjected to PAL2NAL (Suyama, et al., 2006) (http://www.bork.embl.de/pal2nal/) to calculate the corresponding K_S value. Only K_S values between 0 and 1 were included. The approximate date of the duplication event (T) was calculated using the equation $T = K_S / 2\lambda$(Nei and Kumar, 2000), where the synonymous mutation rate per base (λ) of 1.5×10^{-8} or 8.1×10^{-9} is assumed (Wang, et al., 2012) and the mean K_S values from 21 homologous gene pairs is used.

5.2.4 Plasmid construction

Plasmids used in this study are listed in Table 5.2. The integrity of all constructs was confirmed by sequencing.

A detailed description of plasmid construction strategies used for generating p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1*, pCTR (control plasmid) is shown in Figure 5.3. A

Table 5.1 Primers used in this study

Primer name	Sequence
Primers used for cloning flax *LPCAT1* and *LPCAT2* into pYES2.1	
LF	5'-ATGGATTTAGACATGCAATCCATGGCGG-3'
L1R	5'-GAACCTTACTCTTCCTTTCGTGGCTTCG-3'
L2R	5'-AAACCTTATTCTTCCTTTCGGGGCTTCG-3'
Primers used for cloning flax *DGAT1-1* into pESC at MCS2	
ESCD1F1	5'- TATAGGATCCGCGGCCGCATGTCCGTGCTAGAC-3'
ESCD1R1	5'- TATAGGATCCGCGGCCGCTTAGATTCCATCTTTC-3'
Primers used for constructing p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1*, pCTR	
	Primers used to amplify the regions from ADH1 to CYC1 in pESC-*DGAT1-1* or pESC-URA
YESESC1	5'-GAGAGGCGGTTTGCGTATTGGGCGCGCTGAATTGGAGCGACCTCATGC-3'
YESESC2	5'-GTCAGTGAGCGAGGAAGCGGAAGACTGGATCTTCGAGCGTCCCAAAACC-3'
	Primers used to amplify pYESTOPO-*LPCAT1* or pYESBOP
YESESC3	5'-GGTTTTGGGACGCTCGAAGATCCAGTCTTCCGCTTCCTCGCTCACTGAC-3'
YESESC4	5'-GCATGAGGTCGCTCCAATTCAGCGCGCCCAATACGCAAACCGCCTCTC-3'
Primers used for constructing the plasmid containing the replacement cassette	
	Step 1: Insert flax *FAD3B* into pESC at MSC2 under GAL1 promoter using BamHI and KpnI restriction sites
ESCF3F	5'-TATAGGATCCATGAGCCCTCCAAAC-3'
ESCF3R	5'-TATAGGTACCTCAGCTGGATTTGG-3'
	Step 2: Insert P_{GAL1}-*FAD3*-T_{CYC1} cassette into pYESBOP using EcoRI restriction site
BOPF3F	5'-TATAGAATTCGTACGGATTAGAAGC-3'
BOPF3R	5'- TATAGAATTCGTCAGTGAGCGAGGA-3'
	Step 3: Insert Hygromycin cassette into pYESBOP using HindIII restriction site
BOPHgyF	5'- TATAAAGCTTCGTACGCTGCAGGTCGAC-3'
BOPHgy	5'- TATAAAGCTTATCGATGAATTCGAGCTCG-3'

R	
	Step 4: Insert flax *FAD2-1* into pESC at MSC2 using BamHI and KpnI restriction sites
ESCF2F	5'-ATAGGATCCACCATGGGTGCTGGTGGAAGAAT -3'
ESCF2R	5'-TATGGTACCTCACAGCTTGTTGTTGTACCA -3'
	Step 5: Insert P_{GAL1}-*FAD2*-T_{CYC1} cassette into pYESBOP using XhoI restriction site
BOPF2F	5'- TATACTCGAGGTACGGATTAGAAGC-3'
BOPF2R	5'- TATACTCGAGGTCAGTGAGCGAGGA-3'
Primers used for yeast mutant generation	
S1	5'-GACAAACCGCATACGCCAAGACAAACCGTGGTGATTTAATTCTGCTGCTGATCGCTTCCAACATGCGTACGCTGCAGGTCGAC-3'
S2	5'-AGGAACTGGAAAATAAGACAACAAGACTGTGACTTCCACACGCATCTGTCGTTTTTGGCCATCTAATCGATGAATTCGAGCTCG-3'
S3	5'-GACAAACCGCATACGCCAAGACAAACCGTGGTGATTTAATTCTGCTGCTGATCGCTTCCAACATGAGGGAATATTAAGCTCGCCC-3'
	Yeast *ale1* gene-specific (verification primers)
A	5'-TCTAACGGCTATTGCCAGCG-3'
B	5'-AACGTTGAATAGAAAAATGC-3'
C	5'-GCAATTAAACTAGCATTTGG-3'
D	5'-ACAGAAGTACCGACCGTACC-3'
	Disruption marker-specific (verification primers)
B-M	5'-GGATGTATGGGCTAAATG-3'
C-M	5'-CCTCGACATCATCTGCCC-3'
	Flax *FAD3B* gene-specific (verification primers)
FAD3R	5'-CTTGTGGCTAATTCTCC-3'

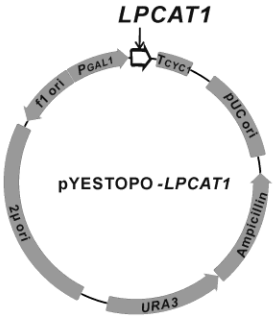
Table 5.2 Plasmids used in this study

Plasmid name	Description	Reference
pYESTOPO-*LPCAT1*	pYES-P_{GAL1}-*LPCAT1*-T_{CYC1}	This study
pYESTOPO-*LPCAT2*	pYES-P_{GAL1}-*LPCAT2*-T_{CYC1}	This study
pYES-*DGAT1-1*	pYES-P_{GAL1}-*DGAT1-1*-T_{CYC1}	Pan, et al., 2013
pECS-*DGAT1-1*	pESC-P_{GAL1}-*DGAT1-1*-T_{CYC1}	This study
pESC-*FAD2*	pESC-P_{GAL1}-*FAD2-1*-T_{CYC1}	This study
pESC-*FAD3*	pESC-P_{GAL1}-*FAD3B*-T_{CYC1}	This study
p*FAD2+FAD3+hph*	pYES-P_{GAL1}-*FAD2-1*-T_{CYC1}/P_{GAL1}-*FAD3B*-T_{CYC1}/ P_{TEF1}-*hphMX*	This study
p*LPCAT1+DGAT1-1*	pYES-P_{GAL1}-*LPCAT1*-T_{CYC1}/P_{GAL1}-*DGAT1-1*-T_{CYC1}	This study
p*LPCAT1*	pYES-P_{GAL1}-*LPCAT1*-T_{CYC1}/P_{GAL1}-T_{CYC1}	This study
p*DGAT1-1*	pYES-P_{GAL1}-T_{CYC1}/P_{GAL1}-*DGAT1-1*-T_{CYC1}	This study
pCTR	pYES-P_{GAL1}-T_{CYC1}/P_{GAL1}-T_{CYC1}	This study
pAG32	pAG32-P_{TEF1}-*hphMX*	Goldstein and McCusker, 1999

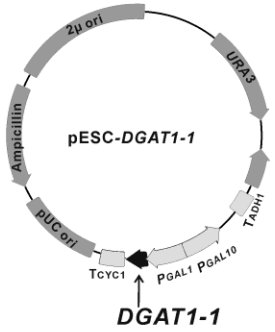
pair of duplicated *DGAT1* genes was previously identified in flax genome (Pan, et al., 2013). One of the two *DGAT1s*, named as *DGAT1-1*, was cloned into the pYESBOP vector as described in the previous study by Pan et al. (2013). In this study, the *DGAT1-1* gene was cloned into the multiple cloning site 2 (MCS2) of pESC-URA for expression from the GAL1 promoter (P_{GAL1}). The resulting plasmid is referred as pECS-*DGAT1-1*. Next, the region from CYC1 terminator (T_{CYC1}) to ADH1 terminator of pESC-URA or pESC-*DGAT1-1* was amplified by using the primer pair YESESC1/YESESC2 (Table 5.1). The PCR products were referred as fragment 1. At the same time, the entire plasmid, pYESTOPO-*LPCAT1* or pYESBOP, was amplified by using the primer pair YESESC3/YESESC4 (Table 5.1). The resulting PCR products were referred to as fragment 2. After the PCR amplification, fragment 1 was cloned into fragment 2 using the one-step, isothermal assembly method described by Gibson (2011). The primers used to amplify fragment 1 and fragment 2 were designed to provide 48 bp overlap regions between fragment 1 and fragment 2 (Figure 5.3). The sequence overlap between the two fragments allows for homologous recombination *in vivo* to clone fragment 1 into fragment 2. Four different combinations of fragment1 with fragment 2 resulted in four different plasmids (Figure 5.3), which are referred to as: 1) p*LPCAT1+DGAT1-1* contains both *LPCAT1* and *DGAT1-1* genes and both genes are under P_{GAL1}; 2) p*LPCAT1* contains only *LPCAT1* gene under P_{GAL1}; 3) p*DGAT1-1* contains only *DGAT1-1* gene under P_{GAL1}; and 4) pCTR is the control plasmid.

Figure 5.3 A schematic diagram showing the strategy used for constructing p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1* and p*CTR* plasmids. The white arrow indicates the flax *LPCAT1* gene. The black arrow indicates the flax *DGAT1-1* gene. The overlapping sequences between fragment1 and fragment 2 are highlighted in white boxes.

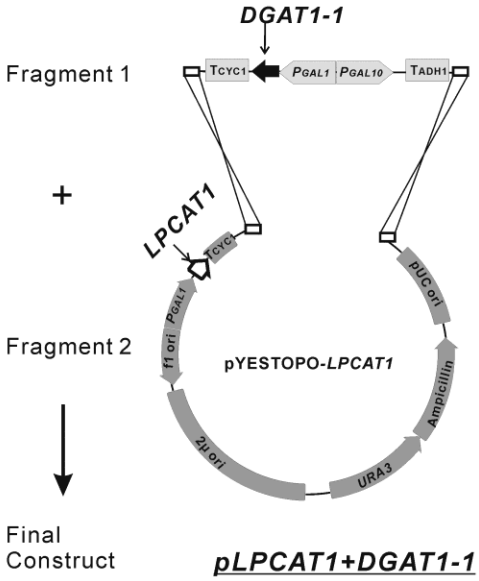
Step 1: Clone *LPCAT1* gene into pYESTOPO

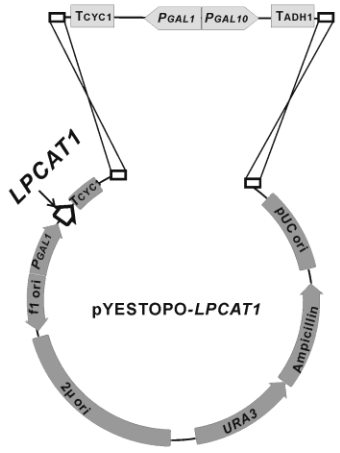


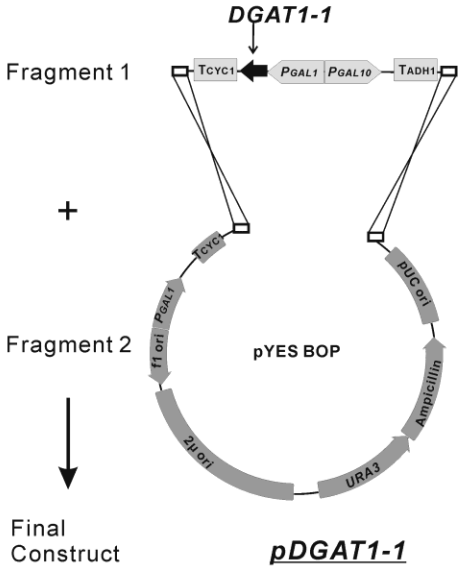
Step 2: Clone *DGAT1-1* gene into pESC-URA

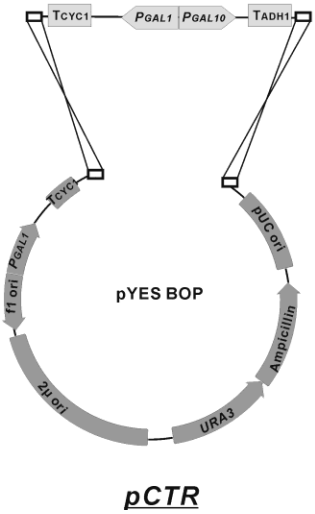


Step 3: Using one-step ISO cloning method to construct p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1* and p*CTR* plasmids.









To construct the plasmid (p*FAD2+FAD3+hph*) containing the replacement cassette for yeast mutant generation, previously isolated flax *FAD2-1* and flax *FAD3B* genes (Pan, et al., 2013) were first amplified using PCR with the gene specific primer pairs (ESCF2F/ESCF2R for *FAD2-1* and ESCF3F/ESCF3R for *FAD3B*; Table 5.1) and then inserted into MSC2 of the pESC-URA, yielding pESC-*FAD2* and pESC-*FAD3*. Next, the cassettes P_{GAL1}-*FAD2*-T_{CYC1} of pESC-*FAD2* and P_{GAL1}-*FAD3*-T_{CYC1} of pESC-*FAD3* were amplified using primer pairs (BOPF2F/BOPF2R for *FAD2* and BOPF3F/BOPF3R for *FAD3*), and then inserted into the same pYESBOP plasmid using XhoI and EcoRI restriction sites, respectively. In addition, the hygromycin B resistance gene *hph* (hygromycin B phosphotransferase gene) was cloned from the plasmid pAG32(Goldstein and McCusker, 1999) and then inserted into the same pYESBOP plasmid using the HindIII restriction site. The resulting plasmid contains flax *FAD2-1* and *FAD3B* genes under the GAL1 promoter and the selectable marker gene *hph*.

5.2.5 Yeast strains

Yeast strains used in this study are listed in Table 5.3.

Two types of yeast mutant strains (H1246 Δ *ale1* and H1246 *ale1 Δ::FAD2+FAD3*) were constructed from the quadruple mutant strain *S. cerevisiae* H1246, in which all four TAG biosynthesis genes (*dga1*, *lro1*, *are1*, and *are2*) are disrupted (Sandager, et al., 2002). H1246 Δ *ale1 is* a quintuple mutant strain. Besides the disrupted four genes in H1246, the *YOR175c* gene (also referred as *ale1*) encoding LPCAT (Chen, et al., 2007) was further deleted by using a disruption cassette containing a gene conferring resistance to Hygromycin B (*hph*). Another yeast mutant strain, H1246 *ale1 Δ::FAD2+FAD3,* was created by replacing the *ale1* coding region with the gene replacement cassette containing genes encoding flax FAD2-1, flax FAD3B and hygromycin B phosphotransferase.

Table 5.3 Yeast Strains used in this study

Strain name	Genotype	Description	Reference
Y02431 (*ale1Δ*)	*Mat a, ale1*Δ, *his3*Δ*1, leu2*Δ*0, met15*Δ*0, ura3*Δ*0, YOR175c::KanMX4*	The *YOR175c* gene (also referred as *ale1*) encoding LPCAT is disrupted.	(Chen, et al., 2007)
H1246	*Mat α, are1-Δ::HIS3, are2-Δ::LEU2, dga1-Δ::KanMX4, lro1-Δ::TRP1, ADE2*	Four genes *dga1*, *lro1*, *are1*, and *are2* encoding DGAT1, PDAT1, ASAT1 and ASAT2 are disrupted.	(Sandager, et al., 2002)
H1246 Δ *ale1*	*Mat α, are1-Δ::HIS3, are2-Δ::LEU2, dga1-Δ::KanMX4, lro1-Δ::TRP1, ale1-Δ::hph, ADE2*	H1246 in which *ale1* is deleted by using a destruction cassette containing a gene conferring resistance to Hygromycin B (*hph*).	This study
H1246 *ale1 Δ::FAD2+FAD3*	*Mat α, are1-Δ::HIS3, are2-Δ::LEU2, dga1-Δ::KanMX4, lro1-Δ::TRP1, ale1Δ::hph+FAD2-1+FAD3B, ADE2*	H1246 in which *ale1* is replaced by flax *FAD2-1* and *FAD3B* genes using a replacement cassette harboring genes encoding FAD2-1, FAD3B and hph.	This study

A PCR-mediated gene disruption strategy (Baudin, et al., 1993) was used for yeast mutant construction. Using pAG32 and p*FAD2+FAD3+hph* as templates, the *ale1* gene disruption and replacement cassettes were amplified by PCR with primer pairs S1/ S2 and S2/S3, respectively. The resulting PCR products were purified by gel extraction and then introduced into H1246 yeast cells using a standard Lithium-Acetate/PEG/single-strand DNA yeast transformation method (Gietz and Schiestl, 2007) with slight modifications. Transformants were selected on YAPD plates containing Hygromicin B (300 μg/ml) (Invitrogen). The correct disruption or replacement of the *ale1* gene in the mutants was verified by PCR with the *ale1* gene-specific and disruption cassette-specific/replacement cassette-specific verification primers (Table 5.1) designed according to Hegemann and Heick (2011). The disruption of the *ale1* gene was further confirmed by performing the lyso-platelet-activating factor (lyso-PAF) sensitivity tests (the detailed procedure can be found under *"Lyso-PAF sensitivity test"* section).

5.2.6 Heterologous expression in yeast

In general, all yeast transformations were performed according to the standard Lithium-Acetate/PEG/single-strand DNA yeast transformation method (Gietz and Schiestl, 2007). Transformants were screened on minimal medium plates lacking uracil and confirmed by colony PCR. The recombinant yeast cells were first cultivated in liquid minimal medium with 2% (w/v) raffinose and then grown in minimal medium containing 2% (w/v) galactose and 1% (w/v) raffinose (referred as induction medium) for induction of gene expression.

5.2.7 Lyso-PAF sensitivity test

Lyso-PAF sensitivity test was performed using the method described in Chen et al. (2007), with slight modifications. Yeast mutant strain Y02431 (*ale1Δ*) transformed with flax *LPCATs* was first grown in liquid minimal medium (lacking uracil) supplemented with 2%

glucose overnight at 30°C with shaking. After washing with sterile water for three times, the overnight cultures were transferred to induction medium for 12h. Cultures were then serial diluted 1:10 from OD_{600nm} (optical density at 600 nm)=2. A measure of 5 µl from each dilution was inoculated on minimal medium plates (lacking uracil) containing 0, 5, 10, 20 µg/ml lyso-PAF. Yeast H1246 and Y02431 mutant transferred with empty vector pYES-*LacZ* were used as positive and negative controls, respectively. The plates were incubated at 28°C for two days.

5.2.8 *In vivo* experiments

The yeast mutant strain (H1246 *ale1 Δ::FAD2+FAD3*) was used for *in vivo* studies. Yeast mutant cells transformed with p*LPCAT1+DGAT1-1* or p*DGAT1-1* were grown at 20°C in induction medium. After 3 days, the cells were harvested and used for yeast lipid analysis. Yeast lipid extraction and analysis was performed according to the method described in Pan et al. (2013).

5.2.9 *In vitro* experiments

Microsomal preparations were performed according to the method described in Pan et al. (2013). The crude protein concentration in the microsomal preparations was measured using the Bio-Rad Bradford assay (Bio-Rad).

5.2.9.1 DGAT enzyme assays

The DGAT enzyme assays were performed using the method described by Byers et al. (1999), with some modifications. Microsomal preparations from H1246 cells transformed with pYES-*DGAT1-1* were used in the assays. Control microsomes were prepared from H1246 yeast transformed with pYES-*LacZ*. The reaction mixture contained 200 mM HEPES-NaOH (pH 7.4), 3.2 mM $MgCl_2$, 333 µM *sn*-1,2-diolein and 15 µM [^{14}C]acyl-CoA in a total volume of 60 µl. The reaction was initiated with 5 µg of the microsomal preparation, allowed to proceed at 30 °C for 6

min, and then quenched by the addition of 10 µl of sodium dodecyl sulfate 10% (w/v). The reaction was then spotted on a silica thin layer chromatography (TLC) plate (SIL G25, 0.25 mm, Macherey-Nagel, Germany). The plate was developed with hexane/diethyl ether/acetic acid (80:20:1, v/v/v), and then dried and exposed to a phosphoimager screen overnight. The resulting phosphoimager screen was visualized using a Typhoon scanner (GE Healthcare). Corresponding TAG spots were scraped from the plate, and radioactivity was measured by a LS 6500 multi-purpose scintillation counter (Beckman-Coulter).

5.2.9.2 LPCAT enzyme assays

The LPCAT enzyme assays were performed according to the method previously described by Lager et al. (2013). Microsomes were prepared from the yeast mutant strain Y02431 (*ale1*Δ) transformed with pYESTOPO-*LPCAT1*. Microsomes prepared from yeast containing pYES-*LacZ* were used as a negative control. For testing the forward activity, the reaction mixture contained 200 µM [^{14}C]18:1-LPC, 100 µM acyl-CoA and 0.1 M potassium phosphate buffer, pH 7.2, in a total volume of 50 µl. The reaction was initiated by adding 0.3 µg of microsomal protein, incubated at 30 ^{o}C for 5 min, and then terminated by the addition of 170 µl of 5% acetic acid and 500 µl of chloroform/methanol (1:1, v/v). The mixture was vortexted and then centrifuged. The chloroform phase (lower phase) was collected, dried under nitrogen, resuspended in 40 µl of chloroform, and then applied to a TLC plate. The TLC plate was then developed with solvent system chloroform/methanol/formic acid (75:25:10, v/v/v), followed by phosphor-imaging analysis using the same method as described under "DGAT enzyme assays". Radioactivity of the PC spots was determined by scintillation counting.

For measuring the reverse activity, an aliquot of microsomal protein (43µg) was lyophilized overnight. The substrate, 9 nmol of *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC (2.5nci/nmol), was

dissolved in 14 µl of benzene and applied directly to the freeze-dried microsomes. The benzene was dried immediately under N_2 at 37 ^oC, after which 100 µl of reaction mixture containing 200 nmol of free CoA, 10 nmol of 18:1-CoA and 1 mg of bovine serum albumin (BSA) in 0.1 M potassium phosphate buffer, pH 7.2, was added. The assay was incubated at 30 ^oC for 60 min and then quenched with 100 µl of 0.15M acetic acid and 500 µl of chloroform/methanol (1:2, v/v). The mixture was then vortexed, centrifuged, and extracted with the Bligh and Dyer method (Bligh and Dyer, 1959). Subsequently, the lower chloroform phase was removed and the upper phase containing [^{14}C]acyl-CoAs produced by the reverse reaction of LPCAT was washed with 2.5 ml of chloroform for three times. The [^{14}C]acyl-CoAs in the upper phase was further hydrolyzed to free FAs with 500 µl of 4M KOH at 90^oC for 15 min and then neutralized using 500 µl of 6M HCl. After neutralization, the free FAs were extracted into chloroform using the Bligh and Dyer method (Bligh and Dyer, 1959). Radioactivity of [^{14}C]-labeled free FAs was determined by scintillation counting.

5.2.9.3 LPCAT1 and DGAT1-1 coupling enzyme assays

The yeast quintuple mutant strain H1246 Δ *ale1* was used to test the coupling hypothesis under *in vitro* conditions. Microsomes were prepared from the cells transformed with p*LPCAT1+DGAT1-1*, p*LPCAT1* or p*DGAT1-1*. Control microsomes were prepared from yeast cells transformed with p*CTR*. A 43µg portion of microsomal preparations was lyophilized overnight. The substrates of 9 nmol *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC (2.5 nci/nmol) and 32 nmol *sn*-1,2-diolein were dissolved in 14 µl of benzene and applied directly to the freeze-dried microsomes. The benzene was dried immediately under N_2 at 37 ^oC. The assay was started by adding 100 µl of reaction mixture containing 200 nmol of free CoA, 10 nmol of 18:1-CoA and 1 mg of BSA in 0.1 M potassium phosphate buffer, pH 7.2. After incubation for 60 mins at 30 ^oC,

the assay was terminated by the addition of 100 µl of 0.15M acetic acid and 500 µl of chloroform/methanol (1:2, v/v). Lipids were extracted into chloroform by adding chloroform/methanol according to the method of Bligh and Dyer (Bligh and Dyer, 1959). The chloroform phase was dried under nitrogen gas and applied to the TLC plate. The resulting TLC plate was first developed in a solvent system of chloroform/methanol/acetic acid/H_20 (60:30:3:1, v/v/v/v) to half plate. After drying in a N_2-filled chamber, the plate was then developed with a solvent system of hexane/diethyl ether/acetic acid (80:20:1, v/v/v) to full plate. The developed plate was then air-dried, followed by phosphor-imaging analysis. The corresponding TAG spots were scraped and radioactivity was measured by scintillation counting. The coupling activities were determined as the amount of [^{14}C] TAG produced.

5.3 Results

As a prelude to investigating possible coupling of the LPCAT-catalyzed reaction to the DGAT1-catalyzed reaction, the properties of individual flax enzymes were first examined.

5.3.1 Identification and isolation of *LPCAT* genes from flax

Through exploring flax genome and EST databases, a pair of duplicated *LPCAT* genes was identified in flax. Both *LPCAT* genes have a coding region of 1380 bp and encode proteins with 460 amino acid residues. These two LPCATs share 96.3% and 96.5% identity at nucleotide and amino acid level, respectively. Both *LPCAT* genes contain 8 exons and 7 introns (Figure 5.4A). Further genomic analysis revealed that the protein-coding genes flanking the *LPCAT* genes are highly conserved, suggesting that the duplicated *LPCAT* genes were derived from a segmental duplication event. In addition, the mean K_S value of 21 homologous gene pairs (including the *LPCAT* gene pair) within the segmental duplication blocks was used as a proxy of time to estimate the date of the segmental duplication event. The result presented in Figure 5.4B

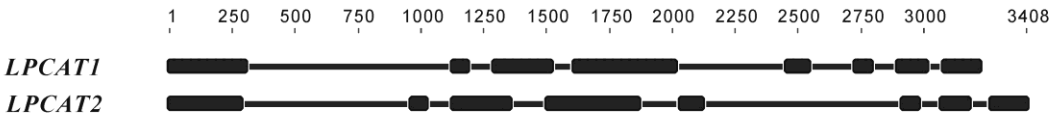
shows that a segmental duplication event occurred approximately 5.1 to 9.5 million years ago (MYA), which is consistent with the time of a recent (5-9 MYA) whole-genome duplication (WGD) in flax (Wang, et al., 2012), suggesting that the segmental duplication is the result of a recent WGD event.

5.3.2 *LPCAT* cDNAs from flax encode functionally active LPCAT enzymes

To determine whether *LPCAT* cDNAs from flax encode proteins possessing LPCAT activity, the cDNAs of both *LPCATs* were amplified from flax embryos and cloned into the yeast expression pYES2.1TOPO vector. The resulting recombinant plasmids were transformed into the yeast mutant strain Y02431 (*ale1*Δ), which has a very low LPCAT background activity (Chen et al., 2007). After transformation, a lyso-PAF sensitivity test was performed to confirm the proper expression of *LPCAT* genes in yeast mutant Y02431. Yeast H1246 and Y02431 mutants containing pYES-*LacZ* were used as positive and negative controls, respectively. The result (Figure 5.4C) showed that there was a severe growth defect in yeast mutant Y02431 transformed with pYES-*LacZ* when the plate was supplemented with 10 μg/ml lyso-PAF. No growth was observed for the negative control when 20 μg/ml lyso-PAF were added into the plates. However, yeast cells expressing *LPCAT1* or *LPCAT2* were able to overcome the growth inhibition of lyso-PAF and exhibited a growth even better than the vector-only H1246 transformants in the presence of 5, 10 and 20 μg/ml lyso-PAF. These results confirmed that these recombinant yeast mutant Y02431 can be further used for enzyme assays. Given the previous finding that the recently duplicated genes in flax, such as *DGAT2-1/DGAT2-3* and *PDAT1/PDAT5*, exhibit similar gene expression patterns and encode proteins with similar substrate selectivity (Pan, et al., 2013), only one of the duplicated *LPCAT* genes, *LPCAT1*, was used for subsequent analyses.

Figure 5.4 Results obtained from gene structure analyses, estimation of gene duplication dates and lyso-PAF sensitivity tests. A) Genomic DNA structure of flax *LPCATs*. The thick lines represent exons and thin lines represent introns. The numbers on the top denote the scale in base pairs; B) The mean K_S values and estimated dates for segmental duplication events corresponding to flax *LPCAT* and *DGAT1* genes. C) Lyso-PAF sensitivity tests of yeast mutant Y02431 expressing flax *LPCATs*. Yeast H1246 and Y02431 mutant strains containing empty vector (pYES-*LacZ*) were used as positive (+) and negative (-) controls, respectively. Cultures were serial diluted 1:10 from OD_{600nm}=2. Four concentrations of lyso-PAF (0, 5, 10, 20 µg/ml) were tested. A detailed procedure is described in the "Experimental Procedures" section.

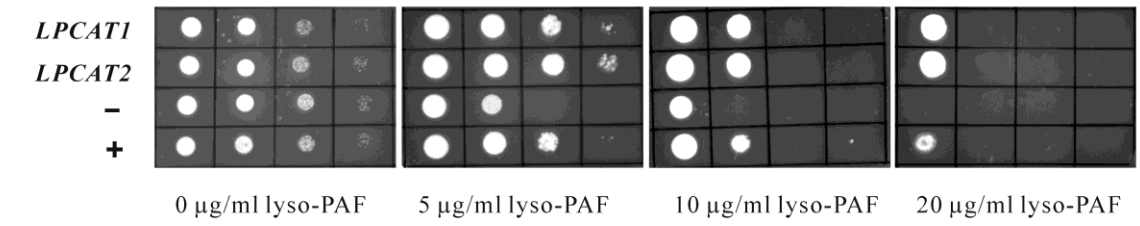
A



B

Segment Pairs	Number of gene Pairs included	*Ks* (Mean±SE)	Estimated time (MYA)
LPCAT	21	0.154±0.077	5.1 - 9.5
DGAT1	21	0.148±0.109	4.9 - 9.2

C



5.3.3 Flax LPCAT1 shows high acyl substrate specificity towards PUFA-containing substrates in both the forward and reverse reactions

To further assess LPCAT enzymatic properties, the acyl substrate specificities of LPCAT were evaluated in the forward and reverse reactions. First, the preference of the LPCAT1-catalyzed reaction towards various acyl-CoA donors was assessed (Figure 5.5A). LPCAT1 exhibited a higher preference for unsaturated fatty acyl-CoAs, including 18:1, 18:2 and 18:3, over the saturated fatty acyl-CoAs. The highest activity was observed with PUFA-CoAs (18:2- and 18:3-CoAs). To measure the acyl specificity in the reverse reaction, different *sn*-2-[^{14}C]acyl-PC species (*sn*-2-[^{14}C]-labeled 18:1-, 18:2- and 18:3-PC) were used as substrates. Among the tested substrates, the reverse reaction of LPCAT showed significantly higher specificity towards *sn*-2-[^{14}C]-labeled 18:2- and 18:3-PC than towards 18:1-PC (Figure 5.5B). In contrast to the previously reported Arabidopsis LPCAT (Lager, et al., 2013) with relatively low preference for *sn*-2-18:2-PC, LPCAT from flax transferred 18:2 acyl groups from PC at the highest rate. Together, these results suggested that both forward and reverse reactions prefer PUFA-containing substrates.

5.3.4 Flax DGAT1-1 displayed small but significantly enhanced specificity for 18:3-CoA

A pair of *DGAT1* genes with 97.7 % sequence identity was previously identified in the flax genome (Pan, et al., 2013). K_S analysis showed that similar to *LPCATs*, the duplicated *DGAT1*s were derived from a recent WGD (Figure 5.4B). It was previously (Pan, et al., 2013; Siloto, et al., 2009) shown that flax DGAT1-1 was able to complement the TAG synthesis deficiency in yeast mutant strain H1246, in which all four TAG biosynthesis genes are disrupted. Experiments of *in vivo* FA feeding indicated that DGAT1-1 produced a similar amount of TAG in the presence of diverse, exogenously provided FAs, suggesting that DGAT1-1 may have a broad

Figure 5.5 Acyl substrate specificities of flax LPCAT1 in the forward (A) and reverse (B) reactions. Microsomal preparations from the yeast mutant strain Y02431 (*ale1*Δ) expressing flax *LPCAT1* were used in assays. For measuring acyl-specificities in the forward reaction, different unlabeled acyl-CoAs and [^{14}C]18:1-LPC were used as substrates. For measuring acyl-specificities in the reverse reaction, different *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC species were provided as substrates. Detailed procedures for measuring the forward and reverse activities of LPCAT are described in "Experimental Procedures" section. Data are represented as mean of duplicate assays ± S.E. The asterisks indicate significant differences to 18:1-CoA or 18:1-PC (ANOVA, LSD test) at $p < 0.05$ level. Abbreviations: LPC, lysophosphatidylcholine.

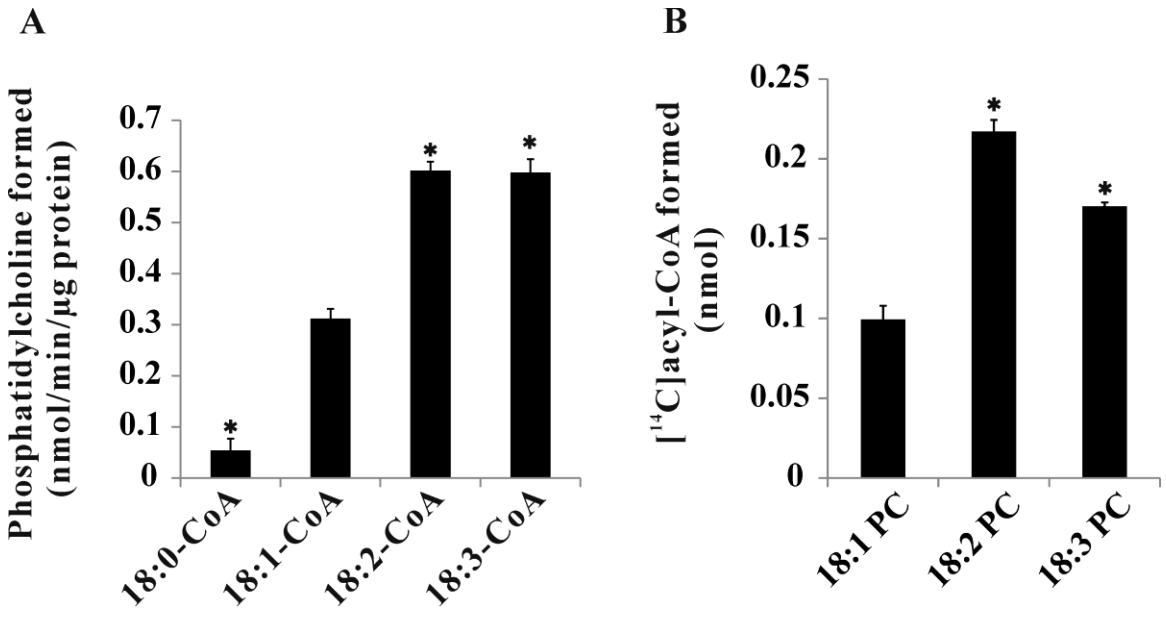
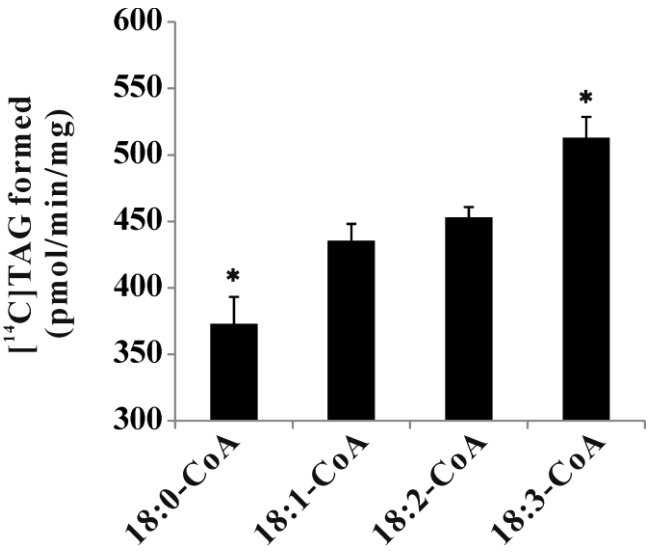


Figure 5.6 Acyl substrate specificities of flax DGAT1-1. Microsomal preparations from the yeast mutant H1246 expressing flax *DGAT1-1* were incubated with different [^{14}C]acyl-CoAs and unlabeled *sn*-1,2-diolein. Data are represented as mean of triplicate assays ± S.E. The asterisks indicate significant differences to 18:1-CoA (ANOVA, LSD test) at $p < 0.05$ level. Abbreviations: DGAT, acyl-CoA: diacylglycerol acyltransferase; TAG, triacylglycerol.



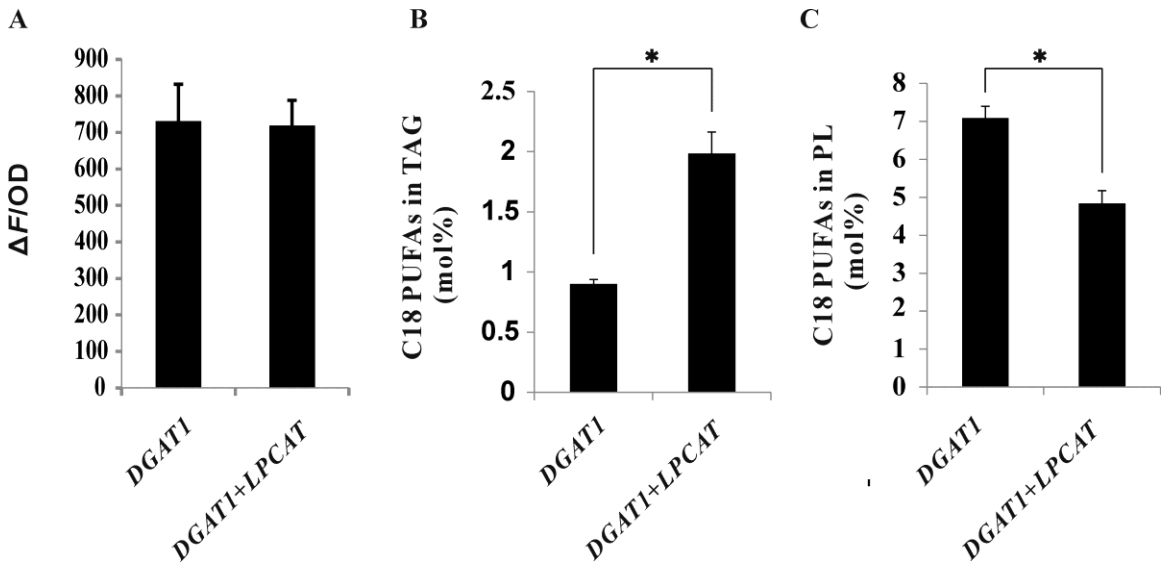
substrate preference (Pan, et al., 2013). In this study, fatty acyl-CoA substrate specificity for flax DGAT1-1 was further examined under *in vitro* conditions. The same yeast mutant strain, H1246, was used since it provides a clean background to detect [^{14}C]TAG produced from the introduced flax DGAT1-1. The assays were carried out with microsomal preparations from H1246 mutant expressing flax *DGAT1-1*. Different [^{14}C]acyl-CoAs were provided along with non-radiolabeled *sn*-1,2-dioleoylglycerol. As shown in Figure 5.6, DGAT1-1 exhibited small but significant changes of substrate specificity towards the acyl-CoA substrates and the highest specificity was observed with 18:3-CoA, which is in close agreement with substrate specificity data for overall microsomal DGAT activity from developing flax embryos (Sørensen, et al., 2005).

5.3.5 *In vivo* evidence of biochemical coupling of LPCAT1 and DGAT1-1-catalyzed reactions

After enzymatic characterization of individual flax LPCAT1 and DGAT1-1 enzymes, the coupling hypothesis was first tested in a yeast system under *in vivo* conditions. To avoid interference from endogenous TAG-synthesizing and LPCAT activities, a new yeast mutant strain was constructed from H1246. In the newly generated mutant strain, referred as H1246 *ale1 Δ::FAD2+FAD3* mutant strain (Table 5.3), the endogenous *ale1* gene was replaced by genes encoding FAD2-1 and FAD3B from flax; thus, no significant TAG-synthesizing and LPCAT activities were present in the mutant. The rationale behind inserting *FAD2* and *FAD3* into the yeast chromosome is that: 1) 18:2 and 18:3 do not naturally occur in the H1246 mutant, and 2) FAD2 and FAD3 catalyze the synthesis of 18:2 and 18:3, respectively, on the *sn*-2 position of PC, which is also the major substrate for the LPCAT reverse reaction. The newly created yeast strain containing flax *FAD2-1* and *FAD3B* genes can synthesize 18:2 and 18:3 on the *sn*-2 position of PC. Thus, this study avoided exogenous FA feeding, which usually gives a random

distribution of FAs over all *sn* positions in all types of lipid classes. Constructs containing flax *DGAT1-1* alone or both *DGAT1-1* and *LPCAT1* were then individually introduced into this newly generated yeast strain. It was hypothesized that the reverse action of LPCAT can catalyze the transfer of PUFAs produced on PC directly into the acyl-CoA pool, making the PUFA-CoAs available for the DGAT1-catalyzed reaction for TAG production. Consistent with this hypothesis, whereas no significant differences in the total amount of TAG were found between yeast cells expressing flax *DGAT1-1* and yeast cells co-expressing flax *DGAT1-1* and *LPCAT1* (Figure 5.7A), co-expressing flax *DGAT1-1* and *LPCAT1* significantly increased 18-carbon PUFAs (18:2 and 18:3) in TAG with a concomitant decrease of 18-carbon PUFAs in phospholipid, as opposed to those expressing *DGAT1-1* only (Figure 5.7B and 5.7C).

Figure 5.7 *In vivo* evidence supporting the biochemical coupling of the flax LPCAT1-catalyzed reverse reaction with the flax DGAT1-1-catalyzed reaction. A) Nile red assays performed with yeast cells co-expressing flax *DGAT1-1* with *LPCAT1* or expressing *DGAT1-1* alone. The values represent the amount of triacylglycerol (TAG) accumulated in the recombinant yeast cells, which is calculated by dividing Nile red fluorescence (ΔF) by the optical density (OD) at 600 nm. B) Mole percentage of 18C polyunsaturated fatty acids (PUFAs) in TAG of the recombinant yeast cells. C) Mole percentage of 18C PUFAs in phospholipid (PL) of the recombinant yeast cells. Instead of exogenous FA feeding, 18C PUFAs were synthesized by co-expressing flax *FAD2-1* and *FAD3B* genes in yeast mutant strains. Data are represented as mean (five biological replicates) ± S.E. The asterisks indicate significant differences in 18C PUFA levels of the TAG and PL fractions from yeast cells co-expressing flax *DGAT1-1* with *LPCAT1* versus expressing *DGAT1-1* alone (ANOVA, LSD test) at $p < 0.05$ level. Abbreviations: FAD, fatty acid desaturase.

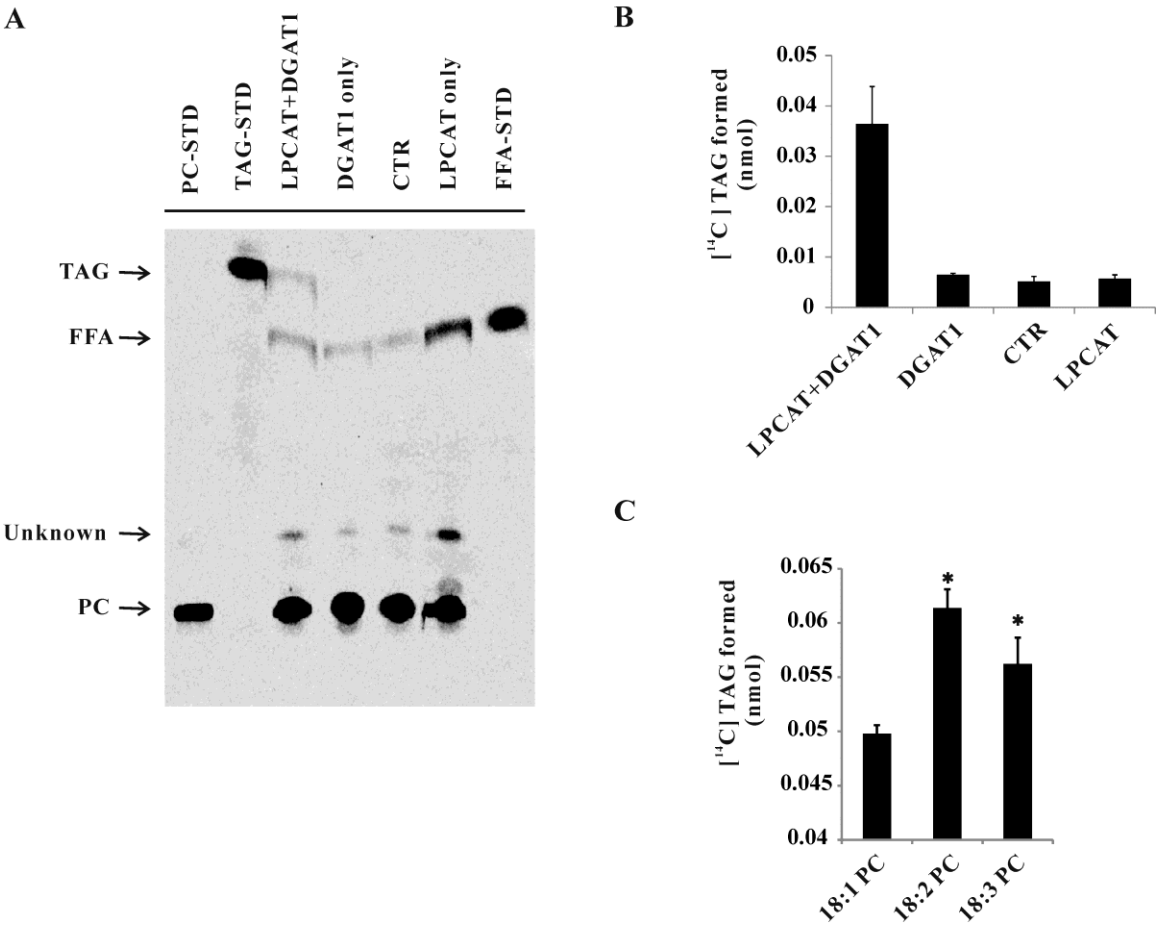


5.3.6 *In vitro* evidence of biochemical coupling of LPCAT1 and DGAT1-1 reverse reactions

Next, the coupling hypothesis was tested under *in vitro* conditions. For this purpose, another quintuple yeast mutant, referred as H1246 Δ *ale1* (Table 5.3), was constructed by disrupted the endogenous *ale1* gene of H1246. The newly generated quintuple mutant was unable to synthesize TAG and had a low background LPCAT activity. The mutant was then transformed with p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1*, or pCTR, individually. After induction, microsomes of the recombinant yeast cells were isolated and fed with PC radiolabeled with [^{14}C] at the *sn*-2 position. It was hypothesized that the biochemical coupling can enhance the transfer of [^{14}C]-labeled acyl groups from PC into TAG via the reverse reaction of LPCAT1 coupled to the DGAT1-1-catalyzed reaction. As shown in Figures 5.8A and 5.8B, [^{14}C]TAG was formed when *sn*-2-[^{14}C]-labeled acyl-PC (*sn*-2-[^{14}C]-labeled 18:1-PC, *sn*-2-[^{14}C]-labeled 18:2-PC or *sn*-2-[^{14}C]-labeled 18:3-PC) was incubated with yeast microsomes containing recombinant flax LPCAT1 and flax DGAT1-1. In contrast, no [^{14}C]TAG was detected following incubation of the radiolabeled PC with microsomes containing either flax LPCAT1 or flax DGAT1-1, alone. In addition, no [^{14}C]TAG could be detected when incubations were conducted with control microsomes. This result confirmed the hypothesis. It should be noted that, as shown in Figure 5.8A, the assays also produced [^{14}C]-free FAs, which may arise from the hydrolysis of [^{14}C]acyl-CoAs catalyzed by the endogenous acyl-CoA thioesterase (Lager, et al., 2013).

The acyl-specificities in overall LPCAT1 and DGAT1-1 coupling process was further tested by presenting *sn*-2-[^{14}C]-labeled 18:1-, 18:2- or 18:3- PC as single substrates to the microsomes containing both flax LPCAT1 and flax DGAT1-1. As shown in Figure 5.8C, yeast microsomes containing flax LPCAT1 and flax DGAT1-1 were able to transfer the [^{14}C]-labeled

Figure 5.8 *In vitro* evidence supporting the biochemical coupling of the flax LPCAT1-catalyzed reverse reaction with the flax DGAT1-1-catalyzed reaction. (A) A representative thin layer chromatography (TLC) plate of flax LPCAT1 and DGAT1-1 coupling enzyme assay. Microsomal preparations from the yeast quintuple mutant H1246 Δ *ale1* individually transformed with p*LPCAT1+DGAT1-1*, p*LPCAT1*, p*DGAT1-1*, or p*CTR* were incubated with *sn*-1-16:0-*sn*-2-[^{14}C]18:1-PC and unlabeled *sn*-1,2-diolein. Followed by the enzyme assay, the lipid fractions were extracted and separated by TLC plate. Radiolabeled [^{14}C] triacylglycerol (TAG) was only formed with yeast microsomes containing both flax LPCAT1 and DGAT1-1; B) The amount of [^{14}C] TAG produced in the recombinant yeast cells. Data are represented as mean of triplicate assays ± S.E.; C) The specific activity of overall DGAT1-1 and LPCAT1 coupling process. Different *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC species along with unlabeled *sn*-1,2-diolein were used as substrates. Data are represented as mean of triplicate assays ± S.E. The asterisks indicate significant differences to 18:1-PC (ANOVA, LSD test) at $p < 0.05$ level. Abbreviations: PC-phosphatidylcholine; FFA-free fatty acid; STD-standard; CTR-control (yeast cells transformed with pCTR plasmid).


A
PC-STD
TAG-STD
LPCAT+DGAT1
DGAT1 only
CTR
LPCAT only
FFA-STD
TAG
FFA
Unknown
PC
B
[14C] TAG formed (nmol)
0.05
0.04
0.03
0.02
0.01
0
LPCAT+DGAT1
DGAT1
CTR
LPCAT
C
[14C] TAG formed (nmol)
0.065
0.06
0.055
0.05
0.045
0.04
*
*
18:1 PC
18:2 PC
18:3 PC

18:2 or 18:3 moiety at a higher rate than 18:1 moiety from the *sn*-2 position of PC to TAG, which was consistent with the *in vivo* data.

5.4 Discussion

Using acyl-CoAs as donors, two major biochemical pathways, the Kennedy pathway (*de novo* pathway) (Weiss, et al., 1960) and the Lands cycle (remodelling pathway) (Lands, 1958), contribute to PC biosynthesis (Figure 5.1). Both pathways were proposed in 1950s. PC serves not only as a major component of cellular membrane, but also as the primary site for PUFA or unusual FA biosynthesis. The accumulation of a large proportion of PUFAs in flax seed oils necessitates efficient mechanisms to channel PUFAs from PC into TAG. Previous studies have shown that PDAT-mediated TAG biosynthesis (Pan, et al., 2013) and PDCT-based PC-DAG interconversion (Wickramarathna, et al., 2015) play critical roles in regulating PUFA content in flax seed oil. The Kennedy pathway and the Lands cycle are connected by LPCAT-catalyzed reactions. Here, it was hypothesized that the reverse reaction of LPCAT can be coupled to the DGAT1 reaction to produce the PUFA-containing TAG.

This hypothesis is supported by the following observations: 1) Acyl-specificity tests showed that the reverse reaction of flax LPCAT1 has the highest preference towards PUFA-containing PC (Figure 5.5B). In addition, the flax DGAT1-1 forward reaction showed a small but significant preference for 18:3-CoAs (Figure 5.6). 2) Under *in vivo* conditions, DGAT1-1 produced a higher amount of PUFA-containing TAG and a lower amount of PUFA-containing PC in the presence than in the absence of LPCAT1 (Figure 5.7B and 5.7C). 3) Under *in vitro* conditions, when *sn*-1-16:0-*sn*-2-[^{14}C]acyl-PC was included as substrate, radiolabeled [^{14}C] TAG was formed only in the reaction with the yeast microsomes containing both flax LPCAT1 and flax DGAT1-1 (Figure 5.8A and 5.8B). In contrast, under the same assay conditions, no

[^{14}C]TAG was detected in reactions with microsomes containing either flax LPCAT1 or flax DGAT1-1 alone. These results provided strong evidence supporting the hypothesis. 4) The acyl group specificity of overall LPCAT1 and DGAT1-1 coupling process under *in vitro* conditions showed that yeast microsomes containing DGAT1-1 and LPCAT1 were able to transfer the [^{14}C]-labeled linoleoyl or linolenoyl moiety at a higher rate than the oleoyl moiety from the *sn*-2 position of PC to TAG (Figure 5.8C).

Given that PUFA groups could be de-acylated by the action of PLA on PC and previous work (Liping, et al., 2012) indicated that the expression of *PLA$_2$* can be affected by altered expression of *LPCAT*, there was the possibility that expression of *LPCAT* triggered the expression of *PLA$_2$*, which could contribute to the increased PUFA content in TAG. It is important to note that PLA-catalyzed de-acylation of PC produces free FAs. Before they can be used by the DGAT reaction, free FAs need to be activated to acyl-CoAs. The formation of the acyl-CoAs is catalyzed by ACS, which requires ATP for its action. However, assay mixture used in this study did not contain ATP. More importantly, the inability to detect any [^{14}C]TAG products in the reaction with microsomes containing flax DGAT1-1 alone (Figure 5.8A and 5.8B) indicates that under this assay conditions, there was little or no interference from PLA. Interference may also arise from PC to DAG conversion catalyzed by the PLC reaction or a combined reaction of PLD/PLP. The resulting PUFA-containing DAG may be further incorporated into TAG by the DGAT-catalyzed reaction, and lead to an increased content of PUFAs in TAG. For the same reason described above, this interference was negligible because I did not detect any [^{14}C]TAG when microsomes contained only recombinant flax DGAT1-1. Taken together, the data in this study support the hypothesis that the transfer of PUFAs formed on PC to the acyl-CoA pool and further being incorporated into TAG is at least to some extent

catalyzed through coupling between the reverse reaction of flax LPCAT1 and the forward reaction of flax DGAT1-1.

It was proposed nearly 30 years ago that LPCAT can catalyze both forward and reverse reactions in plants (Stymne and Stobart, 1984). However, the equilibrium of the LPCAT-catalyzed reaction favors the exothermic forward reaction. It was not until recently that Stymne and his co-workers proved that plant LPCAT enzyme does in fact operate in a reversible fashion under *in vitro* conditions (Lager, et al., 2013). An intriguing question remains: how does the endothermic reverse reaction take place *in vivo*? It was previously suggested that in order for LPCAT reverse reaction to take place *in vivo*, the PC content and free CoA has to be very high in compared with the LPC content and available acyl-CoA within the cell (Lager, et al., 2013). Based on the data obtained from this study, an alternative scenario was proposed as follows: under certain metabolic conditions, the endothermic LPCAT-catalyzed reverse reaction can be pulled via coupling it to the DGAT1-catalyzed forward reaction, which is strongly exothermic and has a shared intermediate, acyl-CoA. The effective removal of acyl-CoAs from the acyl-CoA pool catalyzed by DGAT1 can lower the concentration of acyl-CoA, and thus stimulate the LPCAT-catalyzed reaction to operate in the reverse direction. It is important to mention that the forward reaction of flax LPCAT1 (Figure 5.5A) has highest specificities towards the PUFA-CoAs, indicating that once transferred to the acyl-CoA pool, PUFA groups can be quickly re-acylated into PC by the forward action of LPCAT. In order to maintain a sustained reverse reaction, the re-acylation of the formed LPC by the forward reaction of LPCAT has to occur simultaneously. In addition, the specific activity of flax LPCAT1-catalyzed forward reaction is about 1000 times higher than that of flax DGAT1-catalyzed reaction when the expressed recombinant enzymes were assessed in yeast microsomes under *in vitro* enzyme assays (Figure

5.5A and Figure 5.6). This forward reaction of LPCAT has the potential to compete with the DGAT1-1-catalyzed reaction for PUFA-CoA substrates. Therefore, in order to produce PUFA-enriched TAG, especially ALA-enriched TAG, the DGAT reaction has to be highly selective towards ALA-CoAs. It should be noted that even though flax DGAT1-1 showed enhanced substrate specificity toward ALA-CoAs, the differences in substrate specificity were fairly small (Figure 5.6). Previous *in vivo* FA feeding experiments showed that recombinant flax DGAT2 appears to have a stronger preference than flax DGAT1-1 towards ALA-containing substrates (Pan, et al., 2013). The activity of flax DGAT2 in the yeast system, however, was very low, which limits its potential use in testing the coupling hypothesis in this study. In the future, it will be very interesting to find out whether using codon optimized flax DGAT2 and flax LPCAT1 can further enhance the ALA content in TAG. The enhanced specificity of DGAT2 for ALA-CoA may provide a further operational advantage in the coupling process and the extent of this coupling process could potentially be influenced by the fatty acyl composition of the acyl-CoA pool. Furthermore, LPCAT and DGAT enzymes might physically interact with each other in the membrane. If LPCAT is physically associated with DGAT1 *in vivo*, the products of the LPCAT-catalyzed reverse reaction, PUFA-CoAs, can be potentially channeled directly into DGAT, providing an enhanced efficiency of PUFA-CoA transfer and establishing a favorable operating condition for the reverse reaction of LPCAT. Thus, substrate selectivity and substrate channeling may work collectively to promote the incorporation of PUFAs into TAG, while overcoming a formidable thermodynamic barrier.

It is well-known that PC serves not only the primary site for PUFA biosynthesis, but also the major site for biosynthesis of a number of unusual FAs. The results of this study show that the reverse reaction of LPCAT1 can be coupled with the DGAT1 reaction to enhance PUFA

content in TAG. The biochemical coupling mechanism found in flax might also come into play in other plant systems, such as castor oil seed, which accumulates oil enriched in the unusual FA known as ricinoleic acid. It may be difficult, however, to test the hypothesis of biochemical coupling of castor LPCAT-catalyzed and DGAT-catalyzed reactions using a yeast system because of potential interference by endogenous thioesterase activity, which appears to readily catalyze the hydrolysis of hydroxy-acyl-CoA (Lager, et al., 2013). If the hydroxy-acyl-CoA generated by the castor LPCAT-catalyzed reaction, however, is immediately used by the castor DGAT-catalyzed reaction and castor DGAT exhibits a high affinity for hydroxy-acyl-CoA, then the endogenous thioesterase activity may not be an issue. Indeed, as indicated in the previous study, castor DGAT2 exhibits enhanced preference for hydroxy-acyl-CoAs (Burgal, et al., 2008). The high activity of castor DGAT2 on hydroxy-acyl-CoA may keep the free concentration of hydroxy-acyl-CoA low enough to effectively pull LPCAT in the reverse direction. Further experiments are needed to confirm possible biochemical coupling of LPCAT and DGAT in other plant species.

5.5 Conclusion

In conclusion, seed oils producing enhanced levels of PUFAs or unusual FAs have a wide range of food, feed and industrial applications. Increasing our insight into the incorporation of PC-modified FAs into TAG can provide new strategies to effectively modify seed oils for different uses. This study provide direct *in vivo* and *in vitro* evidence for biochemical coupling of the flax LPCAT1-catalyzed reverse reaction to the flax DGAT1-1-catalyzed reaction as a route for enhancing the amount of PUFA in TAG. A similar mechanism may apply to other plant species which accumulate PUFAs or unusual FAs in their oils. These results provide an increased understanding of PC metabolism in relation to the Kennedy pathway for generating

PUFA-enriched TAG, and will benefit future initiatives aimed at enhancing the PUFA content of TAG in plants.

5.6 References

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res **38:** W7-W13

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Bates PD, Fatihi A, Snapp AR, Carlsson AS, Browse J, Lu C (2012) Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. Plant Physiol **160:** 1530-1539

Baudin A, Ozier-Kalogeropoulos O, Denouel A, Lacroute F, Cullin C (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res **21:** 3329-3330

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Phys **37:** 911-917

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 Biotech J **6:** 819-831

Byers SD, Laroche A, Smith KC, Weselake RJ (1999) Factors enhancing diacylglycerol acyltransferase activity in microsomes from cell-suspension cultures of oilseed rape. Lipids **34:** 1143-1149

Carlsson AS, Yilmaz JL, Green AG, Stymne S, Hofvander P (2011) Replacing fossil oil with fresh oil - with what and for what? Eur J Lipid Sci Tech **113:** 812-831

Chen Q, Kazachkov M, Zheng Z, Zou J (2007) The yeast acylglycerol acyltransferase LCA1 is a key component of Lands cycle for phosphatidylcholine turnover. FEBS Lett **581:** 5511-5516

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. Methods Enzymol **498:** 349-361

Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc **2:** 31-34

Goldstein AL, McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast **15:** 1541-1553

Hegemann JH, Heick SB (2011) Delete and repeat: A comprehensive toolkit for sequential gene knockout in the budding yeast *saccharomyces cerevisiae*. Methods Mol Biol **765:** 189-206

Hu Z, Ren Z, Lu C (2012) The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic Arabidopsis. Plant Physiol **158:** 1944-1954

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S (2013) Plant Acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem **288:** 36902-36914

Lands WE (1960) Metabolism of glycerolipids. 2. The enzymatic acylation of lysolecithin. J Biol Chem **235:** 2233-2237

Lands WE (1958) Metabolism of glycerolipides: a comparison of lecithin and triglyceride synthesis. J Biol Chem **231:** 883-888

Liping W, Shen W, Kazachkov M, Chen G, Chen Q, Carlsson AS, Stymne S, Weselake RJ, Zou J (2012) Metabolic interactions between the Lands cycle and the Kennedy pathway of glycerolipid synthesis in Arabidopsis developing seeds. Plant Cell **24:** 4652-4669

Nei M, Kumar S (2000) Molecular evolution and phylogenetics [electronic resource] / Masatoshi Nei, Sudhir Kumar. Oxford ; New York : Oxford University Press, 2000,

Ohlrogge J, Browse J (1995) Lipid biosynthesis. Plant Cell **7:** 957-970

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Sandager L, Gustavsson MH, Ståhl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H, Stymne S (2002) Storage lipid synthesis is non-essential in yeast. J Biol Chem **277:** 6478-6482

Siloto RMP, Truksa M, He X, McKeon T, Weselake RJ (2009) Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system. Lipids **44:** 963-973

Sørensen BM, Furukawa-Stoffer TL, Marshall KS, Page EK, Mir Z, Forster RJ, Weselake RJ (2005) Storage lipid accumulation and acyltransferase action in developing flaxseed. Lipids **40:** 1043-1049

Sperling P, Linscheid M, Stöcker S, Mühlbach HP, Heinz E (1993) *In vivo* desaturation of *cis*-Δ9-monounsaturated to *cis*-Δ9,12-diunsaturated alkenylether glycerolipids. J Biol Chem **268:** 26935-26940

Stymne S, Stobart AK (1984) Evidence for the reversibility of the acyl-CoA:lysophosphatidylcholine acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons and rat liver. Biochem J **223:** 305-314

Suyama M, Torrents D, Bork P (2006) PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res **34:** W609-W612

Taylor DC, Weber N, Hogge LR, Underhill EW (1990) A simple enzymatic method for the preparation of radiolabeled erucoyl-CoA and other long-chain fatty acyl-CoAs and their characterization by mass spectrometry. Anal Biochem **184:** 311-316

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, Alting-Mees M, Nowak J, Cloutier S, Deyholos M, Bekkaoui F, Sharpe A, Wang E, Rowland G, Selvaraj G, Datla R (2011) Gene expression analysis of flax seed development. BMC Plant Biol **11:** 74

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, Lambert G, Galbraith DW, Grassa CJ, Geraldes A, Cronk QC, Cullis C, Dash PK, Kumar PA, Cloutier S, Sharpe AG, Wong GK, Wang J, Deyholos MK (2012) The

genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J **72:** 461-473

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Wickramarathna AD, Siloto RMP, Mietkiewska E, Singer SD, Pan X, Weselake RJ (2015) Heterologous expression of flax phosphotidylcholine:diacylglycerol cholinephosphotransferases (PDCTs) increases polyunsaturated fatty acid content in yeast and Arabidopsis seeds. BMC Biotech **15:** 63

Chapter 6

Evaluating the Contribution of Phospholipid:Diacylglycerol Acyltransferases and Diacylglycerol Acyltransferase 1 to α-Linolenic acid Accumulation in Flax (*Linum usitatissimum* L.) Seed Oil through an RNA Interference Approach

6.1 Introduction

Flax (*Linum usitatissimum* L.) seed oil is highly enriched in α-linolenic acid (ALA; 18:3$^{cis\Delta 9,12,15}$). Because of its high ALA content, flax seed oil is valued for the potential health benefits as an essential omage-3 fatty acid (FA) and for its industrial applications. Understanding the molecular mechanisms responsible for the high ALA accumulation is essential to further enhance the nutritional and industrial value of flax seed oil.

Triacylglycerol (TAG) is the main component of seed oil. Two major pathways, acyl-CoA dependent (also referred as the Kennedy pathway) and acyl-CoA-independent pathways, contribute to TAG biosynthesis. Acyl-CoA:diacylglycerol acyltransferase (DGAT) catalyzes the final step of the acyl-CoA-dependent pathway for TAG biosynthesis by using *sn*-1,2-diacylglycerol (DAG) and acyl-CoA as its substrates. DGAT activity was first detected in particulate enzyme fractions prepared from chicken (*Gallus gallus*) liver in 1960 (Weiss, et al., 1960). At least six types of genes encoding enzymes with DGAT activity have been identified so far in plants, and they are *DGAT1*, *DGAT2*, *DGAT3*, *DCR*, *DAcT* and a bifunctional DGAT/wax synthase (Saha, et al., 2006; Hobbs, et al., 1999; Lardizabal, et al., 2001; Rani, et al., 2010; Durrett, et al., 2010; Li, et al., 2008). The type 1 *DGAT* gene (*DGAT1*) was the first cloned *DGAT* gene. It belongs to the superfamily of membrane-bound O-acyltransferase (Hofmann, 2000) and has high sequence similarity with sterol:acy-CoA acyltransferase (ACAT), an enzyme

catalyzing the synthesis of cholesteryl ester (Cases, et al., 1998). In 2001, the second family of DGAT (DGAT2), which shares almost no sequence similarity with DGAT1, was first cloned from the oleaginous fungus *Umbelopsis ramanniana.* DGAT1 and DGAT2 have received a lot of interest in recent years and their homologous genes have been cloned and studied in a number of plant species (Sun, et al., 2011; Ilaiyaraja, et al., 2008; Kroon, et al., 2006; Banilas, et al., 2011; Shockey, et al., 2006; Wang, et al., 2006; Hobbs, et al., 1999; Nykiforuk, et al., 2002; Bouvier-Navé, et al., 2000; He, et al., 2004) .

The acyl-CoA-independent pathway is mediated by an enzyme called phospholipid: diacylglycerol acyltransferase (PDAT), which utilizes phospholipids as acyl donors and DAG as an acyl acceptor for TAG formation. PDAT activity was first detected in microsomal preparations from sunflower (*Helianthus annuus*), castor (*Ricinus communis*), and *Crepis palaestina* (Dahlqvist, et al., 2000). The first *PDAT* gene (YNR008w, LRO1) was cloned from yeast as a homolog of human lecithin:cholesterol acyltransferase, an enzyme catalyzing the esterification of free cholesterol in blood plasma. Later, *PDAT* homologous genes have been cloned from *Arabidopsis* (Ståhl, et al., 2004), castor (Kim, et al., 2011; van Erp, et al., 2011), flax (Pan, et al., 2013) and green microalga *Chlamydomonas reinhardtii* (Yoon, et al., 2012). A recent genome-wide analysis of the plant *PDAT* gene family showed that *PDAT* genes are present in all examined plant species, including algae, lowland plants, monocots and eudicots, and multiple within-species *PDAT* paralogs exist in all eudicots (Pan, et al., 2015). Previous cloning and functional studies showed that PDATs can differ not only across species, but also within the same species with respect to expression patterns and substrate selectivity (Pan, et al., 2015; Pan, et al., 2013; Kim, et al., 2011; van Erp, et al., 2011; Zhang, et al., 2009).

To understand the molecular mechanisms behind TAG biosynthesis in flax, I previously performed a detailed functional analysis on enzymes responsible for the last step in TAG synthesis in flax through a sequence homology-based approach (Pan, et al., 2013). By exploiting the publically available flax genome database (www.phytozome.net), I identified two genes encoding DGAT1, three genes encoding DGAT2, two genes encoding DCR and six genes encoding PDATs in flax genome. Gene expression data together with the data obtained from the heterologous expression systems using yeast and Arabidopsis provide several lines of evidence in support of flax possessing a unique pair of *PDAT*s (*PDAT1* and *PDAT5*) which exhibits higher expression levels in developing seeds than in other tissues and encodes enzymes capable of efficiently channeling ALA into TAG. Heterologous expression of the flax *DGAT1* was able to complement not only the TAG synthesis mutations in yeast *Saccharomyces cerevisiae* mutant H1246 but also the FA and reduced oil phenotype of Arabidopsis mutant line AS11. In addition, *in vivo* free FA feeding experiments showed that flax DGAT1 appears to have a relatively broad substrate selectivity. In contrast, flax DGAT2 failed to complement the AS11 mutant lipid phenotype and displayed very low TAG-synthesizing activity only in the presence of exogenous ALA in yeast. Flax *DCR* did not show any significant TAG-forming activity when over-expressed in either yeast or Arabidopsis. The findings of this work provide substantial insights into TAG synthesis in flax. However, the role of candidate genes in seed oil and ALA accumulation has not been evaluated directly in native flax plants.

The main goal of the present study is to provide definitive evidence for the contribution of PDATs (PDAT1 and PDAT5) and DGAT1s to ALA accumulation in flax seed oils through the use of an RNA interference (RNAi) approach.

6.2 Materials and Methods

6.2.1 Plant materials and growth conditions

Flax plants were grown under sterile conditions on MS medium (Murashige and Skoog, 1962) during transformation and tissue culture. Transgenic T1 flax plants will be grown in the growth chamber with 16 h photoperiod and 21^{o}C day/ 18^{o}C night.

6.2.2 Cloning of 3'-UTR of *PDAT1* and *PDAT5* genes

Our previous study (Pan, et al., 2013) found that flax contains six genes encoding PDAT. Two of these six *PDATs*, *PDAT1* and *PDAT5*, are dominantly expressed in seeds and encode proteins with a strong preference for ALA-containing substrates. These two genes were chosen as targets for silencing. To minimize the potential off-target knockdown of the other four *PDATs*, the 3'untranslated regions (3'UTRs) of targeted *PDAT1* and *PDAT5* genes were used to generate the gene silencing constructs.

The 3'UTRs of *PDAT1* and *PDAT5* mRNAs were cloned using the 3'-rapid amplification of cDNA ends (3'-RACE) technique according to the manufacturer's protocol described in 3'RACE System kit (Life Technologies). For this purpose, total RNA was isolated from developing embryos of flax (cultivar CDC Bethune) at 12 days post anthesis using Plant RNeasy plant mini kit (Qiagen, CA). First-strand cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen, CA) with the adapter primer (AP) listed in Table 6.1. After first-strand cDNA synthesis, RNase H was used to remove the original mRNA template. The 3'UTR regions of *PDAT1* and *PDAT5* were then separately amplified by PCR from the resulting cDNA using a gene-specific primer along with the Abridged universal amplification primer (AUAP) (Table 6.1). The 3'-RACE PCR was performed under the following thermal cycling program: 94^{o}C for 3 mins; 30 cycles of desaturation (94^{o}C, 45s), annealing (55^{o}C, 30s), and extension (72^{o}C, 2min);

and a final extension at 72^{o}C for 10 mins. The PCR product (a diffused band around 1000 bp) was gel purified and then used as a template to re-amplify the specific 3'UTR region in a nested PCR using a nested gene specific primer together with AUAP. The setup for the nested PCR is the same as the one used in the 3'-RACE PCR. The final PCR product was cloned into pGEM-T vector (Promega, Madison, Wis.). The resulting vectors are referred as pGEM-P1 and pGEM-P5. The integrity of the cloned fragment was verified by sequencing. Sequence comparison of the cloned fragment with the corresponding open reading frame (ORF) showed that *PDAT1* and *PDAT5* contain 230-bp and 200-bp of 3'UTR, respectively. BLAST analyses of 3'UTR sequences against the flax genome database showed that the 3'UTR regions are highly specific regions, which will maximally avoid off-target gene silencing.

6.2.3 Gene-silencing constructs

All primers are listed in Table 6.1. The integrity of all constructs was verified by sequencing at each step.

Colinin promoter isolation – Genomic DNA was extracted from leaves of flax (cultivar CDC Bethume) according to the hexadecyltrimethylammonium bromide (CTAB) extraction method described by Tamari et al. (2013), with slight modifications. To extract genomic DNA, leaf tissues (~50mg) were ground to a fine powder in liquid nitrogen. The powder was transferred into a 1.5 ml Eppendorf microcentrifuge tube, followed by the addition of 500 µl of CTAB extraction buffer (2% CTAB, 1.4 M NaCl, 0.2% β2-ME, 20 mM EDTA, 100 mM Tris-HCl, pH 8*).* After vortexing, the sample was incubated at 65 ^{o}C for 30 min, and then 500 µl of chloroform was added to each sample. The sample was vortexed briefly and then centrifuged for 10 min at 13000 rpm. The supernatant was removed and transferred into a new

Table 6.1 Primers used in this study

Primer name	Primer sequence
3' RACE primers	
AP (ADAPTER PRIMER)	5'-GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT T-3'
AUAP (Abridged Universal Amplification Primer)	5'-GGC CAC GCG TCG ACT AGT AC -3'
	Gene Specific Primers (GSP)
GSP-P1	5'-GGGTAGTCACCCAGGAAATCACTG-3'
GSP-P5	5'-AGTTAAGGGAAGTCACCCAGGAAAC-3'
	Nested Gene Specific Primers (NGSP)
NGSP – P1	5'-GTCGGCTGCTCTTGGGATGCAAAG-3'
NGSP – P5	5'-GGTGCTTTAGGTAAGTTTGTATGATGATGATG-3'
	Colinin promoter isolation
ColininF	5'-CAACGGTTCCGGCGGTATAG-3'
ColininR	5'-TTTTTGGTGGTGATTGGTTC-3'
***pdat*RNAi construct**	
Overlapping PCR primers for fusing 3'UTRs of *PDAT1* with *PDAT5*	
OL-P1F	5'-CGGCTGCTCTTGGGATGCAAAGTTATG-3'
OL-P1R	5'-GTCAGACCATTTAAAACCCAAACTATATAG-3'
OL-P5F	5'-CTATATAGTTTGGGTTTTAAATGGTCTGAC-3'
OL-P5R	5'-AGAGGCGCACTATCCCATTTATTCAGC-3'
Overlapping PCR primers for fusing Pro-Colinin with 3'UTRs of *PDAT1 +PDAT5*	
OL-CPF	5'-TATAGAGCTCCAACGGTTCCGGCGGTATAG-3'
OL-CPR	5'-TCCCAAGAGCAGCCGTTTTTGGTGGTGATT-3'
OL-3UPPF	5'-AATCACCACCAAAAACGGCTGCTCTTGGGA-3'
OL-3UPPR	5'-TATACTCGAGAGAGGCGCACTATCCCAT-3'
Overlapping PCR pimers for fusing Term-NOS with 3'UTRs of *PDAT1 +PDAT5*	

OL-NPF	5’-AATGTTTGAACGATCCGGCTGCTCTTGGGA-3’
OL-NPR	5’-TATATCTAGACCCGATCTAGTAACATAG-3’
OL-3UPNF	5’-TATAATCGATAGAGGCGCACTATCCCATT-3’
OL-3UPNR	5’-TCCCAAGAGCAGCCGGATCGTTCAAACATT-3’
***dgat1* RNAi construct**	
Overlapping PCR Primers for fusing Pro-Colinin with 3’UTRs of *DGAT1*	
OL-CDF	5’-TATAGAGCTCCAACGGTTCCGGCGGTATAG-3’
OL-CDR	5’-TATGTGGAGGGAAACTTTTTGGTGGTGATTG-3’
OL-3UDPF	5’-CAATCACCACCAAAAAGTTTCCCTCCACATA-3’
OL-3UDPR	5’-TATACTCGAGTACACCAAGCTCTTGAAG-3’
Overlapping PCR Primers for fusing Term-NOS with 3’UTRs of *DGAT1*	
OL-NDF	5’-TATAAAGCTTTACACCAAGCTCTTGAAG-3’
OL-NDR	5’-AATGTTTGAACGATCGTTTCCCTCCACATA-3’
OL-3UDNF	5’-TATGTGGAGGGAAACGATCGTTCAAACATT-3’
OL-3UDNR	5’-TATATCTAGACCCGATCTAGTAACATAG-3’

Pro-Colinin: Colinin promoter; Term-NOS: Nopaline synthase (NOS) terminator; UTR-untranslated region

1.5 ml Eppendorf microcentrifuge tube to which an equal volume of ice-cold isopropanol was added, and the tube was inverted gently five times. The sample was then placed at -20 ^{o}C for 30 min, centrifuged at 13000 rpm for 10 min and the supernatant was discarded. The DNA pellet was washed twice with 750 µl of 70% ethanol, air-dried for 10 min and then resuspended in 150 µl of sterile water. The colinin promoter was amplified from genomic DNA using *Taq* DNA Polymerase (Life technologies, CA) with gene-specific primers ColininF/ ColininR. PCR was performed under the following thermal cycling program: 94^{o}C for 3 mins; 30 cycles of desaturation (94^{o}C, 45s), annealing (58^{o}C, 30s), and extension (72^{o}C, 1min 45s), and a final extension at 72^{o}C for 7 mins. The resulting amplicon was cloned into pGEM-T vector (Promega, Madison, Wis.) to produce pGEM-Colinin vector.

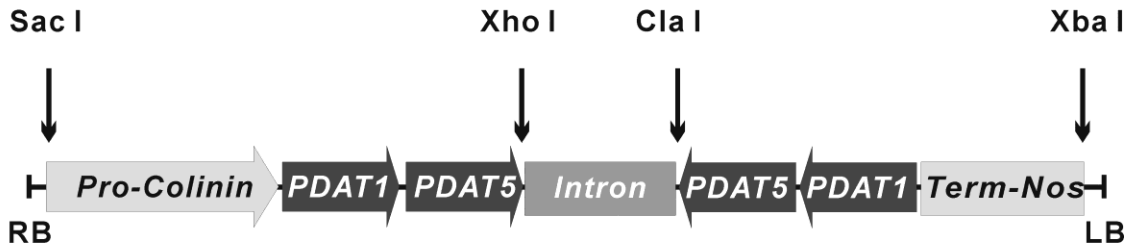
*pdat*RNAi construct- A single RNAi construct containing two gene fragments was designed and used to simultaneously knock down the expression of both *PDAT1* and *PDAT5* genes (Figure 6.1). First, the 3'UTR region of *PDAT1* (230 bp in length) was fused to a 230 bp fragment from the *PDAT5* gene (30 bp of CDS and 200 bp of 3'UTR) using the method of overlap extension PCR (http://openwetware.org/wiki/PCR_Overlap_Extension). Overlap extension PCR requires two pairs of overlapping primers and employs a two-step PCR strategy. The first step involves two independent PCRs to generate two PCR products. The reverse primer of the first primer pair was designed to be complementary to the forward primer of the second primer pair. These two primers with complementary sequences are referred to as the internal primers. The forward primer of the first primer pair and the reverse primer of the second primer pair have no complementary sequence but may contain restriction enzyme sites for inserting the PCR product into the destination vector. These two primers are referred to as the flanking primers. Accordingly, the 3'UTR region of *PDAT1* was amplified with the first overlapping

primers, OL-P1F/OL-P1R (Table 6.1), using pGEM-P1 as a template. In parallel, a 230 bp fragment of *PDAT5* was amplified with the second primer pair, OL-P5F/OL-P5R (Table 6.1), using pGEM-P5 as a template. The resulting PCR products were purified using the GeneJET PCR Purification Kit (Thermo scientific, CA). In the second step, the purified DNA fragments were used as templates for an overlap extension PCR with the outermost flanking primer pair OL-P1F/ OL-P5R (Table 6.1). The resulting overlapping PCR product (430 bp in length), namely the target sequence, was cloned into pGEM-T vector (Promega, Madison, Wis.) to produce pGEM-P1/P5.

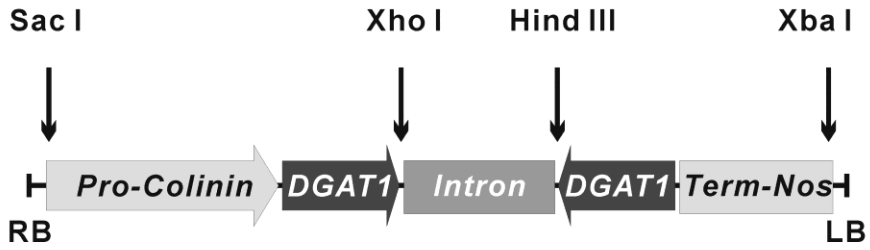
Using the same overlap extension PCR strategy, the target sequence was further fused in the sense direction to the colinin promoter. The resulting extended fragment was digested by *SacI* and *XhoI* and then inserted between the *SacI* and *XhoI* sites of pKannibal (Varsha Wesley, et al., 2001), resulting in pKannibal/A plasmid. Subsequently, the target sequence was fused in the anti-sense direction to a nopaline synthase (NOS) terminator and the extended fragment was inserted into the *ClaI* and *XbaI* sites of pKannibal/A plasmid, yielding pKannibal/A-B plasmid. As shown in Figure 6.1, the resulting pKannibal/A-B plasmid is comprised of a seed-specific colinin promoter and a NOS terminator between which an inverted repeat of 450 bp target sequence (sense and anti-sense) is separated by the intron of *pdk*. The *SacI-Xbal* cassette of pKannibal/A-B was then excised and subcloned into the *SacI* and *Xbal* restriction sites of a binary plasmid pRD400. The resulting plasmid was designated as *pdat*RNAi.

Figure 6.1 Schematic diagrams of *pdat*RNAi- and *dgat1*RNAi- expression cassettes used for the knockdown of the expression of the *PDAT1*/*PDAT5* genes and a pair of *DGAT*1 genes, respectively. Both expression cassettes are driven by a seed-specific colinin promoter (Pro-Colinin) and a nopaline synthase terminator (Term-NOS). The restriction enzyme sites used for cloning are indicated above the fragments.

***pdat*RNAi cassette**



***dgat1* RNAi cassette**



*dgat1*RNAi construct – Flax contains a pair of *DGAT1* genes with 97% sequence identity. The part of the shared ORF with 100% sequence identity (254 bp) was used for constructing the *dgat1*RNAi vector. A 245bp *DGAT1* mRNA-specific fragment was amplified by PCR using the previously constructed plasmid pYESDGAT1 as a template (Pan, et al., 2013). Subsequently, this PCR product was fused in the sense direction to the colinin promoter by the overlap extension PCR and the extended fragment was inserted into pKannibal using *SacI* and *XhoI* restriction sites giving pKannibal/C plasmid. In parallel, the 245 bp *DGAT1* mRNA-specific fragment was fused in the anti-sense direction to the NOS teminator using the overlap extension PCR and the extended fragment was further cloned into pKannibal/C at the restriction sites of *HindIII* and *XbaI*, resulting in pKannibal/C-D plasmid. Then the *SalI-XbaI* cassette carrying a colinin promoter, an intron-interrupted inverted repeat of a 245 bp fragment corresponding to the ORF of the *DGAT1* gene, and a NOS terminator was excised from pKannibal/C-D and subsequently cloned into the *SacI* and *Xbal* restriction sites of pRD400. The resulting plasmid was named *dgat1*RNAi.

The final binary vectors, *pdat*RNAi and *dgat1*RNAi, were individually transformed into the *Agrobacterium tumefaciens* strain GV3101 containing the helper plasmid pMP90. The sequence integrity of the inserts of both constructs was verified by sequencing after its re-isolation from *A. tumefaciens* and re-transformation into *Escherichia coli.*

6.2.4 Plant transformation and verification

A. tumefaciens strain harboring the binary RNAi construct was used for transformation. The vector contained a selection marker gene *NPTII* (neomycin phosphotransferase II gene conferring resistance to kanamycin) for transformed plant selection. Flax was transformed by using hypocotyls as explants. The procedure of flax transformation involved four major steps

(Mlynarova, et al., 1994): hypocotyl inoculation with *A. tumefaciens*, callus induction, shoot initiation and root initiation. To confirm the transformation events, embryo assays and leaf callusing tests were carried out in putative kanamycin resistant transgenic plants. In addition, PCR will be performed with genomic DNA from T1 plants to detect integration of T-DNA into the plant genome.

6.2.5 Real-time RT-PCR quantification

The expression level of *PDAT1/PDAT5* and *DGAT1s* in T_1 transgenic and control plants will be examined at mid developing seeds by performing real time RT-PCR analyses according to the method described by Pan et al. (2013). Similarly, genes encoding glyceraldehyde 3-phosphate dehydrogenase and ubiquitin extension protein will be used as internal reference genes. The relative expression level will be calculated according to the comparative C*t* method.

6.2.6 Lipid analyses

Total lipid content and FA composition of T_1 and T_2 transgenic seeds will be determined by gas chromatography of the fatty acid methyl esters (FAMEs) with 17:0 FAME as an internal standard as previously described (Pan, et al., 2013). The lipid class separation and FA composition in both phospholipid and TAG fractions will be performed according to the method described by Mietkiewska et al. (2008).

6.3 Results and Discussion

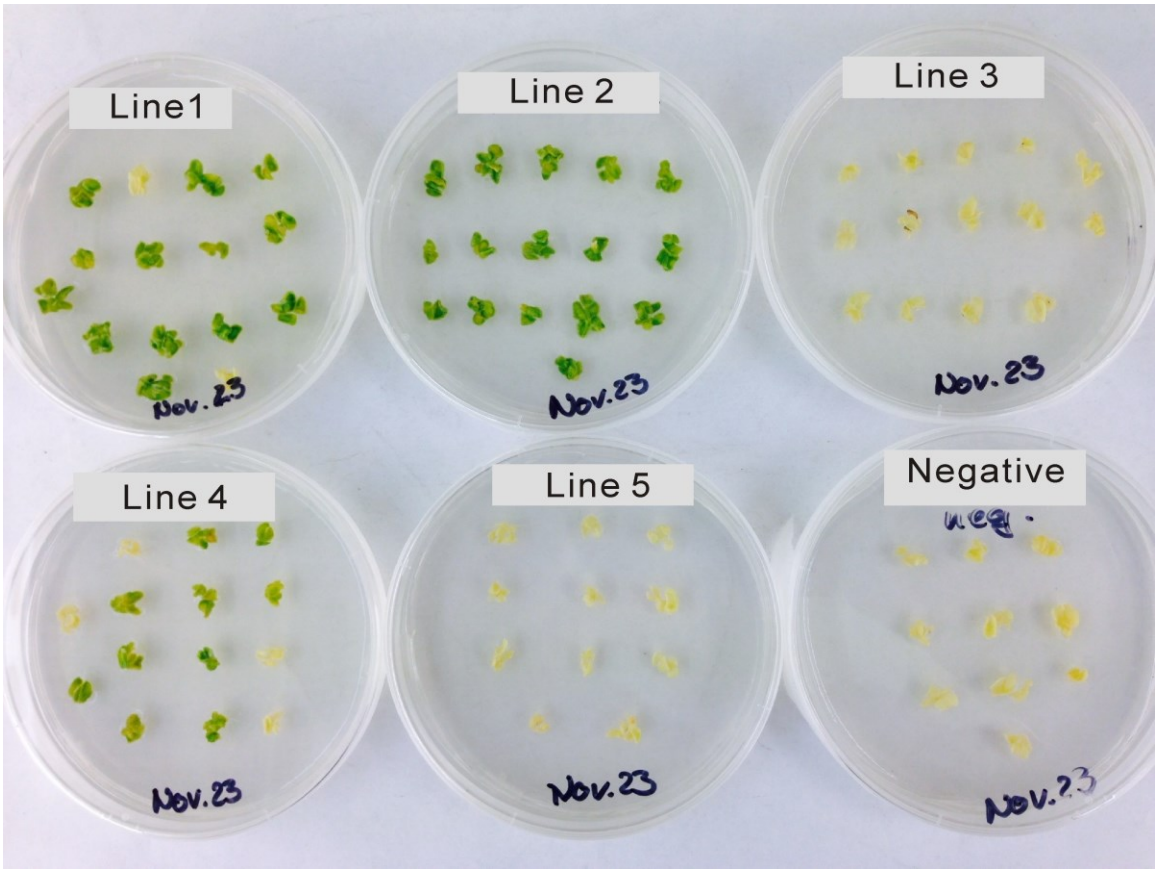
PDAT and DGAT1 are two major enzymes involved in the final step of TAG synthesis in oilseeds. In Arabidopsis, DGAT1 appears to be the dominant enzyme controlling seed oil biosynthesis, while in the absence of DGAT1 activity, PDAT 1 is responsible for the majority of TAG synthesis in seeds and essential for normal seed development (Zhang, et al., 2009). Previous findings suggested that the contribution of DGAT1- and PDAT-catalyzed TAG

synthesis might be different in flaxseed. Functional characterization of flax genes in heterologous expression systems showed that flax contains unique PDATs (PDAT1 and PDAT5) capable of preferentially synthesizing ALA-containing TAGs. In contrast, flax DGAT1 exhibits a relatively broad substrate selectivity. In addition, the expression pattern of *PDAT1/PDAT5* correlates better with the oil and ALA accumulation during seed development than the expression of *DGAT1* does. Together, these results suggest that PDATs (PDAT1/PDAT5) might play a more important role than DGAT1s in seed oil synthesis in flax. However, the limitation of this previous study is that whilst all observations are indicative of the importance of PDATs in ALA-containing TAG synthesis, the function of flax genes was evaluated in the recombinant systems and there was no direct evidence to prove the role of PDAT in its native plant. The obvious and crucial experiment to prove the hypothesis is to individually knock down the expression of *PDATs* and *DGAT1s* in flax and carry out a comprehensive analysis of neutral lipids and phospholipids of flax seed oil.

In order to determine the role of PDATs (PDAT1 and PDAT5) and DGAT1s in TAG and ALA synthesis in flax, two RNAi constructs (*pdat*RNAi and *dgat1*RNAi) were generated. The *pdat*RNAi construct was designed to simultaneously knock down the expression of both *PDAT1* and *PDAT5* genes, while the *dgat1*RNAi construct was used for targeted knockdown of the expression of the *DGAT1* gene pair at the same time. Both RNAi constructs were generated using the seed-specific colinin promoter, which silences the expression of targeted genes mainly in seeds. In this way, the possible influence of altered vegetative growth due to gene silencing on seed oil accumulation will be maximally minimized. Two conventional flax cultivars (CDC Bethune and McGregor) and a high ALA cultivar (F6B) were used for transformation. Embryo assays and leaf callusing test were performed to screen for transgenic plants. As shown in Figure

6.2, transgenic embryos grew green on the selection medium [medium composition: MS medium with 2 mg/L 2,4-D and 100 mg/L kanamycin (Jordam and McHughen, 1988)] while “escape” or untransformed ones turned yellow. In parallel, the leaf sections of transgenic plants formed callus on selection medium [medium composition: MS medium with 2 mg/L 2,4-D, 0.5 mg/L Benzyl Adenine and 50 mg/L kanamycin] while leaves from non-transgenic plants did not. Several batches of transformation were performed; however, the overall transformation efficiency was very low. Only a few transgenic plants were obtained. Oil content and FA composition in neutral and phospholipids of flax seed oil will be analyzed for both T_1 and T_2 transgenic seeds. These results will provide important insights into the role of PDAT and DGAT1 in seed oil synthesis in flax.

Figure 6.2 Representative image showing embryo assays used for screening five independent lines of McGregor transformed with the *pdat*RNAi construct. McGregor transformed with empty vector was used as a negative control. Approximately 15 developing embryos of each transgenic line were placed in a Petri dish on solid embryo development and maturation medium supplemented with kanamycin. Under such a selection pressure for 10 days, embryos of transformed plants with green color are considered positive while embryos with yellow color are negative.



6.4 References

Banilas G, Karampelias M, Makariti I, Kourti A, Hatzopoulos P (2011) The olive DGAT2 gene is developmentally regulated and shares overlapping but distinct expression patterns with DGAT1. J Exp Bot **62:** 521-532

Bouvier-Navé P, Benveniste P, Oelkers P, Sturley SL, Schaller H (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. Eur J Biochem **267:** 85-96

Cases S, Smith SJ, Zheng Y, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese Jr. RV (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A **95:** 13018-13023

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M (2010) A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in *Euonymus* and transgenic seeds. Proc Natl Acad Sci U S A **107:** 9464-9469

Mietkiewska E, Hoffman TL, Brost JM, Giblin EM, Barton DL, Francis T, Zhang Y, Taylor DC (2008) Hairpin-RNA mediated silencing of endogenous *FAD2* gene combined with heterologous expression of *Crambe abyssinica FAE* gene causes an increase in the level of erucic acid in transgenic *Brassica carinata* seeds. Mol Breeding **22:**619–627

He X, Turner C, Chen GQ, Lin J-, McKeon TA (2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. Lipids **39:** 311-318

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Hofmann K (2000) A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. Trends Biochem Sci **25:** 111-112

Ilaiyaraja N, Rajarani AP, Santha IM (2008) Cloning and characterization of diacylglycerol acyltransferase (DGAT) cDNA sequence from *Brassica juncea* cv. Pusa Bold. Indian J Biochem Biophys **45:** 30-36

Jordan MC, McHughen A (1988) Glyphosate tolerant flax plants from Agrobacterium mediated gene transfer. Plant Cell Rep **7:**281-284

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Kroon JTM, Wei W, Simon WJ, Slabas AR (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry **67:** 2541-2549

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol **148:** 97-107

Mlynarova L, Bauer M, Nap JP, Pretova A (1994) High efficiency Agrobacterium-mediated gene transfer to flax. Plant Cell Rep **13:**282–285

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant **15:** 493-497

Nykiforuk CL, Furukawa-Stoffer TL, Huff PW, Sarna M, Laroche A, Moloney MM, Weselake RJ (2002) Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of *Brassica napus* and sucrose-mediated induction of enzyme biosynthesis. BBA-Mol Cell Biol L **1580:** 95-109

Pan X, Peng FY, Weselake RJ (2015) Genome-wide analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes in plants reveals the eudicot-wide *PDAT* gene expansion and altered selective pressures acting on the core eudicot *PDAT* paralogs. Plant Physiol **167:** 887-904

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

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 Physiol **141:** 1533-1543

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. Plant Cell **18:** 2294-2313

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Sun L, Ouyang C, Kou S, Wang S, Yao Y, Peng T, Xu Y, Tang L, Chen F (2011) Cloning and characterization of a cDNA encoding type 1 diacylglycerol acyltransferase from sunflower (*Helianthus annuus* L.). Z Naturforsch C **66(1-2):** 63-72

Tamari F, Hinkley CS, Ramprashad N (2013) A comparison of DNA extraction methods using Petunia hybrida tissues. Journal of Biomolecular Techniques **24:** 113-118

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

Varsha Wesley S, Helliwell CA, Smith NA, Wang M, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant Journal **27:** 581-590

Wang HW, Zhang JS, Gai JY, Chen SY (2006) Cloning and comparative analysis of the gene encoding diacylglycerol acyltransferase from wild type and cultivated soybean. Theor Appl Genet **112:** 1086-1097

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Yoon K, Han D, Li Y, Sommerfeld M, Hu Q (2012) Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *chlamydomonas reinhardtii*. Plant Cell **24:** 3708-3724

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Chapter 7

General Discussion

The present doctoral thesis was guided by four hypotheses: 1) At least one enzyme involved in the final step of triacylglycerol (TAG) biosynthesis in flax (*Linum usitatissimum* L.) can preferentially incorporate α-linolenic acid (ALA; 18:3$^{cis\Delta 9,12,15}$)-containing substrates into TAG; 2) Different selection pressures have acted on the duplicated *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes during the evolution, leading to the diversification of the *PDAT* genes in eudicots; 3) The acyl CoA:diacylglycerol acyltransferase (DGAT)-catalyzed reaction can be coupled to the reverse reaction of acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT) for the incorporation of polyunsaturated fatty acids (PUFAs) into TAG; 4) RNA interference (RNAi)-mediated knockdown of *PDATs* or *DGAT1s* in flaxseed can affect ALA and seed oil accumulation. The overall objective of this project was to achieve a better understanding of TAG biosynthesis in flaxseed and especially aimed at uncovering the molecular mechanisms underlying the high PUFA accumulation in its seed oils. Results from these studies provide substantial insights into the TAG biosynthetic machinery in flax seed, which will be useful for improving oil quality of flax and other oilseeds. In the following sections the results of evaluating each hypothesis is summarized and discussed in terms of the key findings and their significance.

Systematic characterization of enzymes involved in the final step of TAG biosynthesis in flax

In general, very limited knowledge on seed oil biosynthesis in flax is available. I started this project started by performing a systematic characterization of enzymes involved in the last step of TAG synthesis in flax. By exploiting the publically available flax genome database

(www.phytozome.net), seven genes encoding DGATs and six genes encoding PDATs were identified in the flax genome. Expression of the associated transcripts was analyzed by reverse transcription-PCR to identify the predominant genes during the period of active seed oil and ALA accumulation. The results indicate that two *DGAT1s*, two *DGAT2s* and four *PDATs* are preferentially expressed in developing flax seed embryos than in other tissues. In addition, the accumulation of seed oil and ALA correlates better with the expression of seed-preferred *PDAT* genes than with the expression of *DGAT1* genes. Functional expression of the corresponding cDNAs in *S. cerevisiae* (mutant strain H1246) identified several enzymes that were able to restore TAG synthesis to H1246. When yeast cells were cultured in the presence of ALA, specific PDATs appeared to preferentially catalyze the synthesis of trilinolenin, which is the major molecular species of TAG in flax seed oil (Ciftci, et al., 2012). Moreover, the substrate selectivity of the identified PDATs was not limited to ALA, but extended to other PUFAs, including stearidonic acid ($18{:}4^{cis\Delta 6,9,12,15}$), γ-linolenic acid ($18{:}3^{cis\Delta 6,9,12}$) and eicosapentanoic acid ($20{:}5^{cis\Delta 5,8,11,14,17}$). Furthermore, the overexpression of *DGAT1* in the Arabidopsis mutant AS11 restored the fatty acid (FA) profile and oil content to the wild-type levels. In addition, over-expressing *PDAT1* or *PDAT2* in wild-type (WT) Arabidopsis resulted in an increase in ALA content in seed oils. Overall, this study provided several lines of evidence in support of flax containing a unique pair of *PDATs* (*PDAT1*/*PDAT5*), which has a seed-preferred expression profile and encodes enzymes with a preference for catalyzing the synthesis of TAGs containing ALA acyl moieties. The findings of this work provide new insights into seed oil biosynthesis in flax. The identified ALA-selective PDATs will benefit future engineering efforts aimed at enhancing the accumulation of ALA-containing TAGs in plants. In addition, this study points out that the direct transfer of ALA from PC into TAG catalyzed by the PDAT activity may represent

a substrate-dichotomy bottleneck for the production of very-long chain omega-3 PUFAs (VLC-ω-3-PUFAs) in transgenic flax. This bottleneck may be alleviated by the use of the Δ8-desaturation pathway together with the Δ17 desaturation to bypass the step of ALA synthesis on PC. Further experiments are needed to test this new enzyme combination for metabolic engineering of VLC-ω-3-PUFAs in flax. In the future, it will also be interesting to find out how the ALA-selective PDAT interplays with other enzymes involved in the TAG biosynthesis for ALA accumulation.

Comprehensive genome-wide analysis of the *PDAT* gene family across green plants

Intrigued by the finding from the first part of this thesis that multiple PDAT paralogs with diverse expression pattern and different TAG-forming capacity exist in the flax genome, a comprehensive genome-wide analysis of the *PDAT* gene family was conducted to study the evolution and diversity of PDATs in plants. The results show that the *PDAT* gene family is present across green plants, including algae, lower land plants, monocots and dicots. After careful screening, a total of 86 full-length PDAT sequences from 34 species were used for further analyses. Phylogenetic analyses divide the *PDAT* gene family into seven major clades. This phylogenetic separation is supported by evolutionary conservation and variation in gene structure, protein properties, motif occurrences, and/or selection constraints. In addition, the study revealed that a eudicot-wide *PDAT* gene expansion has occurred. It appears that the eudicot-shared ancient gene duplication and subsequent species-specific segmental duplications are mainly responsible for the expansion of *PDAT* genes in eudicots. Selection pressure analyses demonstrate that the paleo-duplicated core eudicot *PDATs* have undergone different selection constraints during evolution. Based on previous functional and expression analyses of PDATs from different plant species, it was proposed that, during evolution, the observed variation in

selection constraints might have led to non-functionalization and expression divergence of paleo-duplicated *PDAT* paralogs. More interesting, it appears that the PDAT paralogs from specific evolutionary clade may have developed unique substrate selectivity and play an important role for seed oil biosynthesis in plant species, which have high levels of modified FAs, including PUFAs and unusual FAs, in their seed oils. PDATs from this clade may represent potential targets for biotechnological approaches to produce unusual FAs in oilseeds. However, it is important to note that, by now, functional characterization of plant PDATs has been limited to four species: one in the unicellular green alga *C. reinhardtii* and three in core eudicots. Further sampling, functional characterization and expression profile analyses are needed to confirm this evolutionary theory. Together, these data provide an in-depth look at evolution, conservation and variation of the *PDAT* gene family in plants and acts as a first step towards a comprehensive understanding of the biological functions of the *PDAT* genes in plants.

Evaluation of the possible biochemical coupling of the LPCAT-catalyzed reverse reaction with the DGAT-catalyzed reaction

In developing seeds, PUFAs are mainly synthesized on phosphatidylcholine (PC). The involvement of LPCAT in the formation of PUFA-rich TAGs has been suggested nearly 30 years ago (Stymne and Stobart, 1984). There are two essential steps involved in this mechanism: first, PUFAs have to be released from PC into the acyl-CoA pool via the reverse reaction of LCPAT and second, the resulting PUFA-CoAs have to be subsequently incorporated into TAGs. Recently, Stymne and his co-workers proved that plant LPCAT enzyme does in fact operate in a reversible fashion under *in vitro* conditions (Lager, et al., 2013). However, the key question remains: can the thermodynamically unfavorable reverse reaction of LPCAT occur *in vivo*? More importantly, there is no direct evidence showing that PUFA-CoAs produced from the

reverse reaction of LPCAT can be incorporated into TAGs. The main focus of this study was to evaluate the possible biochemical coupling of the reverse action of LPCAT with the diacylglycerol acyltransferase1 (DGAT1)-catalyzed reaction for the incorporation of PUFAs into TAGs. This study proved for the first time that LPCAT does in fact act in a reversible fashion *in vivo*. More interestingly, several lines of evidence from both *in vivo* and *in vitro* approaches were provided supporting the fact that PUFAs can be transferred from PC into the acyl-CoA pool by the reverse reaction of LPCAT, and then be utilized by the DGAT1-catalyzed reaction for TAG synthesis. A higher amount of PUFA-containing TAGs was formed by the DGAT1-catalyzed reaction in the presence than in the absence of LPCAT. It seems that the effective removal of acyl-CoAs from the acyl-CoA pool catalyzed by DGAT1 may lower the concentration of acyl-CoAs, and thus stimulate the LPCAT-catalyzed reaction to operate in the reverse direction, which in turn generates the PUFA-rich acyl-CoA pool for the formation of TAGs in the reaction catalyzed by DGAT1. Together, the results indicate that the reverse reaction of LPCAT can be coupled to the DGAT1-catalyzed reaction for PUFA-containing TAG production. The findings of this study provide further insights of PC acyl editing in relation to the Kennedy pathway for generating PUFA-enriched TAGs. Further studies are needed to elucidate whether there is a physical interaction between DGAT and LPCAT enzymes. Moreover, it remains to be tested whether the use of the codon optimized DGAT2 with LPCAT can further enhance the ALA accumulation in TAG. Furthermore, it will be interesting to find out whether the DGAT-LPCAT coupling mechanism exists in other plant species that accumulate high amount of unusual FAs in their seeds, such as castor and tung tree, and if so, co-expressing specific *DGAT* along with *LPCAT* may represent a promising strategy for unusual FA production in oilseeds.

Examination of the contribution of PDAT and DGAT1 to seed oil and ALA accumulation in flaxseed

The existence of unique ALA-selective PDAT enzymes (PDAT1/PDAT5) in flaxseed has been suggested in the first part of this thesis using heterologous expression systems. However, the data do not provide direct evidence comparing the contribution of PDAT or DGAT1 to seed oil and ALA accumulation in flaxseed. As an extended study, the last part of the thesis research was designed to evaluate the role of PDAT and DGAT1 in seed oil and ALA accumulation in native flax plants through the use of an RNA interference (RNAi) approach. To achieve this goal, two RNAi constructs driven by the seed-specific colinin promoter were generated. One construct was designed to simultaneously knock down the expression of *PDAT1* along with *PDAT5* genes, while the other construct was built to knock down the expression of a pair of *DGAT1* genes at the same time. The transgenic RNAi flax lines have been obtained. Detailed analyses of seed oil content and FA composition in both the phospholipid and TAG fractions of the T_1 and T_2 transgenic seeds will be carried out. The results obtained from this study will allow a direct comparison of the contribution of PDAT (PDAT1/PDAT5) and DGAT1 to seed oil and ALA accumulation in flaxseed. In addition, the findings of this study will provide a clearer picture of the biosynthetic pathways responsible for producing ALA-enriched seed oils in flax. If the results confirm that PDAT is the dominant enzyme contributing to ALA accumulation in flax seed, in the future, the applicability of using the ALA-selective flax PDAT1 (via over-expression) will be examined to generate transgenic flax with ultra-high ALA levels. If successful, in the future, attempt will also be made to improve the activity of LuPDAT1 through directed evolution as done previously for DGAT (Siloto, et al., 2009). If a high activity form of LuPDAT1 with a

single amino acid substitution is identified, then the gene encoding the enzyme would be a good target for genome editing to specifically introduce the single point mutation.

Concluding comments

Flax oils with a substantial amount of PUFAs, especially ALA, have various food, feed and industrial uses. The research work presented in this thesis explored the molecular mechanisms regulating TAG biosynthesis in flax. Two important mechanisms for PUFA enrichment in TAG, the PDAT-mediated mechanism and the LPCAT-DGAT coupling mechanism, have been uncovered in flax. In addition, a comprehensive genome-wide analysis of the *PDAT* gene family generates a rich repertoire of knowledge for detailed functional characterization of the *PDAT* genes in plants in the future. The knowledge obtained from this thesis research will benefit the future development of novel biotechnological strategies to produce value-added seed oils in flax and other oilseed crops.

References

Ciftci ON, Przybylski R, Rudzińska M (2012) Lipid components of flax, perilla, and chia seeds. Eur J Lipid Sci Tech **114:** 794-800

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S (2013) Plant acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem **288:** 36902-36914

Siloto RMP, Truksa M, He X, McKeon T, Weselake RJ (2009a) Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system. Lipids **44:** 963-973

Stymne S, Stobart AK (1984) The biosynthesis of triacylglycerols in microsomal preparations of developing cotyledons of sunflower (*Helianthus annuus* L.). Biochem J **220:** 481-488

Bibliography (All sources used)

Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res **38:** W7-W13

Abbadi A, Domergue F, Bauer J, Napier JA, Welti R, Zähringer U, Cirpus P, Heinz E (2004) Biosynthesis of very-long-chain polyunsaturated fatty acids in transgenic oilseeds: Constraints on their accumulation. Plant Cell **16:** 2734-2746

Abeywardena MY, Patten GS (2011) Role of ω3 longchain polyunsaturated fatty acids in reducing cardio-metabolic risk factors. Endocr Metab Immune Disord Drug Targets **11:** 232-246

Ahmad A, Zhang Y, Cao XF (2010) Decoding the epigenetic language of plant development. Mol Plant **3:** 719-728

Akaike H (1974) New look at the statistical model identification. IEEE Trans Autom Control **AC-19:** 716-723

Alendé N, Nielsen JE, Shields DC, Khaldi N (2011) Evolution of the isoelectric point of mammalian proteins as a consequence of indels and adaptive evolution. Proteins: Struct , Funct , Bioinf **79:** 1635-1648

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Alvarez HM, Steinbüchel A (2003) Triacylglycerols in prokaryotic microorganisms. Appl Microbiol Biotechnol **60:** 367-376

Andersson MX, Goksör M, Sandelius AS (2007) Membrane contact sites: Physical attachment between chloroplasts and endoplasmic reticulum revealed by optical manipulation. Plant Signal Behav **2:** 185-187

Anisimova M, Kosiol C (2009) Investigating protein-coding sequence evolution with probabilistic codon substitution models. Mol Biol Evol **26:** 255-271

Arabolaza A, Rodriguez E, Altabe S, Alvarez H, Gramajo H (2008) Multiple pathways for triacylglycerol biosynthesis in *Streptomyces coelicolor*. Appl Environ Microbiol **74:** 2573-2582

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):** e92237

Bafor M, Smith MA, Jonsson L, Stobart K, Stymne S (1991) Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor bean (*Ricinus communis*) endosperm. Biochem J **280:** 507-514

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings /International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology **2:** 28-36

Banas A, Carlsson AS, Huang B, Lenman M, Banas W, Lee M, Noiriel A, Benveniste P, Schaller H, Bouvier-Navé P, Stymne S (2005) Cellular sterol ester synthesis in plants is performed by an enzyme (phospholipid:sterol acyltransferase) different from the yeast and mammalian acyl-CoA:sterol acyltransferases. J Biol Chem **280:** 34626-34634

Banilas G, Karampelias M, Makariti I, Kourti A, Hatzopoulos P (2011) The olive DGAT2 gene is developmentally regulated and shares overlapping but distinct expression patterns with DGAT1. J Exp Bot **62:** 521-532

Barron EJ, Stumpf PK (1962) Fat metabolism in higher plants XIX. The biosynthesis of triglycerides by avocado-mesocarp enzymes. Biochim Biophys Acta **60:** 329-337

Bates PD, Browse J (2011) The pathway of triacylglycerol synthesis through phosphatidylcholine in Arabidopsis produces a bottleneck for the accumulation of unusual fatty acids in transgenic seeds. Plant J **68:** 387-399

Bates PD, Browse J (2012) The significance of different diacylgycerol synthesis pathways on plant oil composition and bioengineering. Front Plant Sci **3:** 147-147

he accumulation of unusual fatty acids in transgenic seeds. Plant J **68:** 387-399

Bates PD, Durrett TP, Ohlrogge JB, Pollard M (2009) Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. Plant Physiol **150:** 55-72

Bates PD, Fatihi A, Snapp AR, Carlsson AS, Browse J, Lu C (2012) Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. Plant Physiol **160:** 1530-1539

Bates PD, Ohlrogge JB, Pollard M (2007) Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. J Biol Chem **282:** 31206-31216

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

Baudin A, Ozier-Kalogeropoulos O, Denouel A, Lacroute F, Cullin C (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res **21:** 3329-3330

Bayon S, Chen G, Weselake R, Browse J (2015) A small phospholipase A2-α from castor catalyzes the removal of hydroxy fatty acids from phosphatidylcholine in transgenic Arabidopsis seeds. Plant Physiol **167:** 1259-1270

Beopoulos A, Haddouche R, Kabran P, Dulermo T, Chardot T, Nicaud JM (2012) Identification and characterization of DGA2, an acyltransferase of the DGAT1 acyl-CoA:diacylglycerol acyltransferase family in the oleaginous yeast *Yarrowia lipolytica*. New insights into the storage lipid metabolism of oleaginous yeasts. Appl Microbiol Biotechnol **93:** 1523-1537

Bertin P, Bullens P, Bouharmont J, Kinet JM (1998) Somaclonal variation and chilling tolerance improvement in rice: changes in fatty acid composition. Plant Growth Regul **24:** 31-41

Bessoule JJ, Testet E, Cassagne C (1995) Synthesis of phosphatidylcholine in the chloroplast envelope after import of lysophosphatidylcholine from endoplasmic reticulum membranes. Eur J Biochem **228:** 490-497

Bielawski JP, Yang Z (2004) A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J Mol Evol **59:** 121-132

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Phys **37:** 911-917

Bouvier-Navé P, Benveniste P, Oelkers P, Sturley SL, Schaller H (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. Eur J Biochem **267:** 85-96

Brown AP, Affleck V, Fawcett T, Slabas AR (2006) Tandem affinity purification tagging of fatty acid biosynthetic enzymes in *Synechocystis*sp. PCC6803 and *Arabidopsis thaliana*. J Exp Bot **57:** 1563-1571

Brown AP, Coleman J, Tommey AM, Watson MD, Slabas AR (1994) Isolation and characterisation of a maize cDNA that complements a 1-acyl *sn*-glycerol-3-phosphate

acyltransferase mutant of *Escherichia coli* and encodes a protein which has similarities to other acyltransferases. Plant Mol Biol **26:** 211-223

Burdge GC (2006) Metabolism of alpha-linolenic acid in humans. Prostaglandins Leukot Essent Fatty Acids **75:** 161-168

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 Biotech J **6:** 819-831

Byers SD, Laroche A, Smith KC, Weselake RJ (1999) Factors enhancing diacylglycerol acyltransferase activity in microsomes from cell-suspension cultures of oilseed rape. Lipids **34:** 1143-1149

Cao H (2011) Structure-function analysis of diacylglycerol acyltransferase sequences from 70 organisms. BMC Res Notes **4:** 249-249

Cases S, Smith SJ, Zheng Y, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese Jr. RV (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A **95:** 13018-13023

Carlsson AS, Yilmaz JL, Green AG, Stymne S, Hofvander P (2011) Replacing fossil oil with fresh oil - with what and for what? Eur J Lipid Sci Tech **113:** 812-831

Chapman KD, Ohlrogge JB (2012) Compartmentation of triacylglycerol accumulation in plants. J Biol Chem **287:** 2288-2294

Chen G, Snyder CL, Greer MS, Weselake RJ (2011) Biology and biochemistry of plant phospholipases. Crit Rev Plant Sci **30:** 239-258

Chen Q, Kazachkov M, Zheng Z, Zou J (2007) The yeast acylglycerol acyltransferase LCA1 is a key component of Lands cycle for phosphatidylcholine turnover. FEBS Lett **581:** 5511-5516

Cheng D, Meegalla RL, He B, Cromley DA, Billheimer JT, Young PR (2001) Human acyl-CoA:diacylglycerol acyltransferase is a tetrameric protein. Biochem J **359:** 707-714

Choi JK, Yu F, Wurtele ES, Nikolau BJ (1995) Molecular cloning and characterization of the cDNA coding for the biotin-containing subunit of the chloroplastic acetyl-coenzyme A carboxylase. Plant Physiol **109:** 619-625

Choudhary V, Jacquier N, Schneiter R (2011) The topology of the triacylglycerol synthesizing enzyme Lro1 indicates that neutral lipids can be produced within the luminal compartment of the endoplasmatic reticulum: Implications for the biogenesis of lipid droplets. Commun Integr Biol **4:** 781-784

Ciftci ON, Przybylski R, Rudzińska M (2012) Lipid components of flax, perilla, and chia seeds. Eur J Lipid Sci Tech **114:** 794-800

Coleman RA, Lee DP (2004) Enzymes of triacylglycerol synthesis and their regulation. Prog Lipid Res **43:** 134-176

Cyril J, Powell GL, Duncan RR, Baird WV (2002) Changes in membrane polar lipid fatty acids of seashore paspalum in response to low temperature exposure. Crop Sci **42:** 2031-2037

Dahlqvist A, Ståhl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne S (2000) Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci U S A **97:** 6487-6492

Dahmen JL, Olsen R, Fahy D, Wallis JG, Browse J (2013) Cytochrome b5 Coexpression Increases Tetrahymena thermophila Delta6 Fatty Acid Desaturase Activity in Saccharomyces cerevisiae. Eukaryot Cell **12:** 923-931

Dailey HA, Strittmatter P (1979) Modification and identification of cytochrome *b*5 carboxyl groups involved in protein-protein interaction with cytochrome *b*5 reductase. J Biol Chem **254:** 5388-5396

Das UN (2006) Essential fatty acids - A review. Curr Pharm Biotechnol **7:** 467-482

Daum G, Wagner A, Czabany T, Athenstaedt K (2007) Dynamics of neutral lipid storage and mobilization in yeast. Biochimie **89:** 243-248

Demandre C, Bahl J, Serghini H, Alpha MJ, Mazliak P (1994) Phosphatidylcholine molecular species formed by lysophosphatidylcholine acyltransferase from soya bean microsomes. Phytochemistry **35:** 1171-1175

Domergue F, Abbadi A, Ott C, Zank TK, Zähringer U, Heinz E (2003) Acyl carriers used as substrates by the desaturases and elongases involved in very long-chain polyunsaturated fatty acids biosynthesis reconstituted in yeast. J Biol Chem **278:** 35115-35126

Douce R, Joyard J (1982) Plants galactolipids. In PK Stumpf, EE Conn, eds, The biochemistry of plants. Academic Press, New York, pp 331-332

Drummond A, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2013) Geneious v5.3.6. Available from: http://www.geneious.com/

Durrett TP, Benning C, Ohlrogge J (2008) Plant triacylglycerols as feedstocks for the production of biofuels. Plant J **54:** 593-607

Durrett TP, McClosky DD, Tumaney AW, Elzinga DA, Ohlrogge J, Pollard M (2010) A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in *Euonymus* and transgenic seeds. Proc Natl Acad Sci U S A **107:** 9464-9469

Eastmond PJ, Quettier AL, Kroon JTM, Craddock C, Adams N, Slabas AR (2010) Phosphatidic acid phosphohydrolase 1 and 2 regulate phospholipid synthesis at the endoplasmic reticulum in Arabidopsis. Plant Cell **22:** 2796-2811

Fakas S, Qiu Y, Dixon JL, Han GS, Ruggles KV, Garbarino J, Sturley SL, Carman GM (2011) Phosphatidate phosphatase activity plays key role in protection against fatty acid-induced toxicity in yeast. J Biol Chem **286:** 29074-29085

Fan J, Yan C, Roston R, Shanklin J, Xu C (2014) Arabidopsis lipins, PDAT1 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward ß-oxidation, thereby maintaining membrane lipid homeostasis. Plant Cell **26:** 4119-4134

Fan J, Yan C, Zhang X, Xu C (2013) Dual role for phospholipid: diacylglycerol acyltransferase: Enhancing fatty acid synthesis and diverting fatty acids from membrane lipids to triacylglycerol in Arabidopsis leaves. Plant Cell **25:** 3506-3518

Farese Jr. RV, Cases S, Smith SJ (2000) Triglyceride synthesis: Insights from the cloning of diacylglycerol acyltransferase. Curr Opin Lipidol **11:** 229-234

Færgeman NJ, Black PN, Dan Zhao X, Knudsen J, DiRusso CC (2001) The Acyl-CoA synthetases encoded within FAA1 and FAA4 in Saccharomyces cerevisiae function as components of the fatty acid transport system linking import, activation, and intracellular utilization. J Biol Chem **276:** 37051-37059

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151:** 1531-1545

Ghosal A, Banas A, Ståhl U, Dahlqvist A, Lindqvist Y, Stymne S (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyltransferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate specificities. Biochim Biophys Acta **1771:** 1457-1463

Gidda SK, Shockey JM, Rothstein SJ, Dyer JM, Mullen RT (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 Physiol Biochem **47:** 867-879

Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. Methods Enzymol **498:** 349-361

Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc **2:** 31-34

Goldstein AL, McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast **15:** 1541-1553

Graham IA (2008) Seed storage oil mobilization. Annu Rev Plant Biol **59:** 115-142

Green AG (1986) A mutant genotype of flax (*Linum usitatissimum*L.) containing very low levels of linolenic acid in its seed oil. Can J Plant Sci **66:** 499-503

Guo AY, Zhu QH, Chen X, Luo JC (2007) GSDS: a gene structure display server. Yi Chuan **29:** 1023-1026

Hackett CS, Strittmatter P (1984) Covalent cross-linking of the active sites of vesicle-bound cytochrome *b*5 and NADH-cytochrome *b*5 reductase. J Biol Chem **259:** 3275-3282

Han GS, Wu WI, Carman GM (2006) The *Saccharomyces cerevisiae* Lipin homolog is a Mg^{2+}-dependent phosphatidate phosphatase enzyme. J Biol Chem **281:** 9210-9218

Harwood JL (1996) Recent advances in the biosynthesis of plant fatty acids. BBA-Lipid Lipid Met **1301:** 7-56

Harwood JL (1982) Plant acyl lipids. In PK Stumpf, EE Conn, eds, The biochemistry of plants. Academic Press, New York, pp 1-55

He X, Turner C, Chen GQ, Lin J-, McKeon TA (2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. Lipids **39:** 311-318

Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. J Syst Evol **46:** 239-257

Hegemann JH, Heick SB (2011) Delete and repeat: A comprehensive toolkit for sequential gene knockout in the budding yeast *saccharomyces cerevisiae*. Methods Mol Biol **765:** 189-206

Hellens RP, Anne Edwards E, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: A versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol **42:** 819-832

Hernández ML, Whitehead L, He Z, Gazda V, Gilday A, Kozhevnikova E, Vaistij FE, Larson TR, Graham IA (2012) A cytosolic acyltransferase contributes to triacylglycerol synthesis in sucrose-rescued Arabidopsis seed oil catabolism mutants. Plant Physiol **160:** 215-225

Hobbs DH, Lu C, Hills MJ (1999) Cloning of a cDNA encoding diacylglycerol acyltransferase from *Arabidopsis thaliana* and its functional expression. FEBS Lett **452:** 145-149

Hofmann K (2000) A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. Trends Biochem Sci **25:** 111-112

Houb EB (2001) The arms race is ancient history in Arabidopsis, the wildflower. Nat Rev Genet **2:** 516-527

Hu Z, Ren Z, Lu C (2012) The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic Arabidopsis. Plant Physiol **158:** 1944-1954

Huang B (2006) Plant-environment interactions. [electronic resource]. Boca Raton, FL : CRC/Taylor & Francis, 2006; 3rd ed,

Hudak KA, Thompson JE (1997) Subcellular localization of secondary lipid metabolites including fragrance volatiles in Carnation petals. Plant Physiol **114:** 705-713

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:** 754-755

Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond Ser B Biol Sci **256:** 119-124

Huis R, Hawkins S, Neutelings G (2010) Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.). BMC Plant Biol **10:** 71

Ilaiyaraja N, Rajarani AP, Santha IM (2008) Cloning and characterization of diacylglycerol acyltransferase (DGAT) cDNA sequence from *Brassica juncea* cv. Pusa Bold. Indian J Biochem Biophys **45:** 30-36

Jessen D, Roth C, Wiermer M, Fulda M (2015) Two activities of long-chain acyl-coenzyme a synthetase are involved in lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis. Plant Physiol **167:** 351-366

Jhala AJ, Hall LM (2010) Flax (*Linum usitatissimum* L.): Current uses and future applications. Aust J Basic Appl Sci **4:** 4304-4312

Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell **7:** 359-371

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: Genome-scale protein function classification. Bioinformatics **30:** 1236-1240

Jordan MC, McHughen A (1988) Glyphosate tolerant flax plants from Agrobacterium mediated gene transfer. Plant Cell Rep **7:**281-284

Jhala AJ, Weselake RJ, Hall LM (2009) Genetically engineered flax: Potential benefits, risks, regulations, and mitigation of transgene movement. Crop Sci **49:** 1943-1954

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. J Biol Chem **278:** 8075-8082

Karantonis HC, Nomikos T, Demopoulos CA (2009) Triacylglycerol metabolism. Curr Drug Targets **10:** 302-319

Katavic V, Reed DW, Taylor DC, Giblin EM, Barton DL, Zou Jitao, MacKenzie SL, Covello PS, Kunst L (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. Plant Physiol **108:** 399-409

Khaldi N, Shields DC (2011) Shift in the isoelectric-point of milk proteins as a consequence of adaptive divergence between the milks of mammalian species. Biology Direct **6:** 40-40

Kim HU, Lee KR, Go YS, Jung JH, Suh MC, Kim JB (2011) Endoplasmic reticulum-located PDAT1-2 from castor bean enhances hydroxy fatty acid accumulation in transgenic plants. Plant Cell Physiol **52:** 983-993

Kim HU, Li Y, Huang AHC (2005) Ubiquitous and endoplasmic reticulum-located lysophosphatidyl acyltransferase, LPAT2, is essential for female but not male gametophyte development in Arabidopsis. Plant Cell **17:** 1073-1089

King A, Nam J-, Han J, Hilliard J, Jaworski JG (2007) Cuticular wax biosynthesis in petunia petals: Cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. Planta **226:** 381-394

Kjellberg JM, Trimborn M, Andersson M, Sandelius AS (2000) Acyl-CoA dependent acylation of phospholipids in the chloroplast envelope. BBA-Mol Cell Biol L **1485:** 100-110

Knutzon DS, Lardizabal KD, Nelsen JS, Bleibaum JL, Davies HM, Metz JG (1995) Cloning of a coconut endosperm cDNA encoding a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase that accepts medium-chain-length substrates. Plant Physiol **109:** 999-1006

Kodama H, Hamada T, Horiguchi G, Nishimura M, Iba K (1994) Genetic enhancement of cold tolerance by expression of a gene for chloroplast ω-3 fatty acid desaturase in transgenic tobacco. Plant Physiol **105:** 601-605

Kong H, Landherr LL, Frohlich MW, Leebens-Mack J, Ma H, DePamphilis CW (2007) Patterns of gene duplication in the plant SKP1 gene family in angiosperms: Evidence for multiple mechanisms of rapid gene birth. Plant J **50:** 873-885

Kopelman PG (2000) Obesity as a medical problem. Nature **404:** 635-643

Kozaki A, Kamada K, Nagano Y, Iguchi H, Sasaki Y (2000) Recombinant carboxyltransferase responsive to redox of pea plastidic acetyl-CoA carboxylase. J Biol Chem **275:** 10702-10708

Krasowska A, Dziadkowiec D, Polinceusz A, Plonka A, Łukaszewicz M (2007) Cloning of flax oleic fatty acid desaturase and its expression in yeast. J Am Oil Chem Soc **84:** 809-816

Kroon JTM, Wei W, Simon WJ, Slabas AR (2006) Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. Phytochemistry **67:** 2541-2549

Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol **305:** 567-580

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S (2013) Plant Acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem **288:** 36902-36914

Lands WE (1960) Metabolism of glycerolipids. 2. The enzymatic acylation of lysolecithin. J Biol Chem **235:** 2233-2237

Lands WE (1958) Metabolism of glycerolipides: a comparison of lecithin and triglyceride synthesis. J Biol Chem **231:** 883-888

Lardizabal KD, Mai JT, Wagner NW, Wyrick A, Voelker T, Hawkins DJ (2001) DGAT2 is a new diacylglycerol acyltransferase gene family. Purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity. J Biol Chem **276:** 38862-38869

Latsague M, Acevedo H, Fernández J, Romero M, Cristi R, Alberdi M (1992) Frost-resistance and lipid composition of cold-hardened needles of chilean conifers. Phytochemistry **31:** 3419-3426

Lee HY, Bahn SC, Shin JS, Hwang I, Back K, Doelling JH, Ryu SB (2005) Multiple forms of secretory phospholipase A_2 in plants. Prog Lipid Res **44:** 52-67

Lee J, Welti R, Schapaugh WT, Trick HN (2011) Phospholipid and triacylglycerol profiles modified by PLD suppression in soybean seed. Plant Biotechnol J **9:** 359-372

Li M, Wei F, Tawfall A, Tang M, Saettele A, Wang X (2014) Overexpression of patatin-related phospholipase AIIIδ altered plant growth and increased seed oil content in camelina. Plant Biotechnol J

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol **148:** 97-107

Li M, Bahn SC, Fan C, Li J, Phan T, Ortiz M, Roth MR, Welti R, Jaworski J, Wang X (2013) Patatin-related phospholipase pPLAIIId increases seed oil content with long-chain fatty acids in Arabidopsis. Plant Physiol **162:** 39-51

Li R, Yu K, Hildebrand DF (2010) DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. Lipids **45:** 145-157

Li N, Gügel IL, Giavalisco P, Zeisler V, Schreiber L, Soll J, Philippar K. (2015) FAX1, a novel membrane protein mediating plastid fatty acid export. PLoS Biol **13:** e1002053

Li WH, Yang J, Gu X (2005) Expression divergence between duplicate genes. Trends Genet **21:** 602-607

Liping W, Shen W, Kazachkov M, Chen G, Chen Q, Carlsson AS, Stymne S, Weselake RJ, Zou J (2012) Metabolic interactions between the Lands cycle and the Kennedy pathway of glycerolipid synthesis in Arabidopsis developing seeds. Plant Cell **24:** 4652-4669

Lin YP, Carman GM (1989) Purification and characterization of phosphatidate phosphatase from *Saccharomyces cerevisiae*. J Biol Chem **264:** 8641-8645

Listenberger LL, Han X, Lewis SE, Cases S, Farese Jr. RV, Ory DS, Schaffer JE (2003) Triglyceride accumulation protects against fatty acid-induced lipotoxicity. Proc Natl Acad Sci U S A **100:** 3077-3082

Liu J, Rice A, Mcglew K, Shaw V, Park H, Clemente T, Pollard M, Ohlrogge J, Durrett TP (2015) Metabolic engineering of oilseed crops to produce high levels of novel acetyl glyceride oils with reduced viscosity, freezing point and calorific value. Plant Biotechnol J

Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ (2012) Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. Prog Lipid Res **51:** 350-377

Lock YY, Snyder CL, Zhu W, Siloto RMP, Weselake RJ, Shah S (2009) Antisense suppression of type 1 diacylglycerol acyltransferase adversely affects plant development in *Brassica napus*. Physiol Plantarum **137:** 61-71

Lu C, Xin Z, Ren Z, Miquel M, Browse J (2009) An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. Proc Natl Acad Sci U S A **106:** 18837-18842

Lung SC, Weselake RJ (2006) Diacylglycerol acyltransferase: A key mediator of plant triacylglycerol synthesis. Lipids **41:** 1073-1088

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science **290:** 1151-1155

Lynch M, Force A (2000) The probability of duplicate gene preservation by subfunctionalization. Genetics **154:** 459-473

Maddison WP (1997) Gene trees in species trees. Syst Biol **46:** 523-536

Maisonneuve S, Bessoule J-, Lessire R, Delseny M, Roscoe TJ (2010) Expression of rapeseed microsomal lysophosphatidic acid acyltransferase isozymes enhances seed oil content in Arabidopsis. Plant Physiol **152:** 670-684

Mancha M, Stymne S (1997) Remodelling of triacylglycerols in microsomal preparations from developing castor bean (*Ricinus communis* L.) endosperm. Planta **203:** 51-57

McCartney AW, Dyer JM, Dhanoa PK, Kim PK, Andrews DW, McNew JA, Mullen RT (2004) Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. Plant J **37:** 156-173

McConn M, Browse J (1996) The critical requirement for linolenic acid is pollen development, not photosynthesis, in an Arabidopsis mutant. Plant Cell **8:** 403-416

McFie PJ, Stone SL, Banman SL, Stone SJ (2010) Topological orientation of acyl-CoA:diacylglycerol acyltransferase-1 (DGAT1) and identification of a putative active site histidine and the role of the N terminus in dimer/tetramer formation. J Biol Chem **285:** 37377-37387

Mène-Saffrané L, Dubugnon L, Chételat A, Stolz S, Gouhier-Darimont C, Farmer EE (2009) Nonenzymatic oxidation of trienoic fatty acids contributes to reactive oxygen species management in Arabidopsis. J Biol Chem **284:** 1702-1708

Mhaske V, Beldjilali K, Ohlrogge J, Pollard M (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (*At5g13640*). Plant Physiol Bioch **43:** 413-417

Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res **41:** D377-D386

Mietkiewska E, Hoffman TL, Brost JM, Giblin EM, Barton DL, Francis T, Zhang Y, Taylor DC (2008) Hairpin-RNA mediated silencing of endogenous *FAD2* gene combined with heterologous expression of *Crambe abyssinica FAE* gene causes an increase in the level of erucic acid in transgenic *Brassica carinata* seeds. Mol Breeding **22:**619–627

Mietkiewska E, Siloto RMP, Dewald J, Shah S, Brindley DN, Weselake RJ (2011) Lipins from plants are phosphatidate phosphatases that restore lipid synthesis in a pah1δ mutant strain of *Saccharomyces cerevisiae*. FEBS J **278:** 764-775

Milcamps A, Tumaney AW, Paddock T, Pan DA, Ohlrogge J, Pollard M (2005) Isolation of a gene encoding a 1,2-diacylglycerol-*sn*-acetyl-CoA acetyltransferase from developing seeds of *Euonymus alatus*. J Biol Chem **280:** 5370-5377

Millar JS, Billheimer JT (2005) Acyl CoA:diacylglycerol acyltransferases (DGATs) as therapeutic targets for cardiovascular disease. Lipids and Atherosclerosis. CRC Press, , pp 31-39

Miquel M, Browse J (1995) Lipid biosynthesis in developing seeds. In J Kigel, G Galili, eds, Seed development and germination. New York: Marcel Dekker, , pp 169-193

Mlynarova L, Bauer M, Nap JP, Pretova A (1994) High efficiency Agrobacterium-mediated gene transfer to flax. Plant Cell Rep **13:**282–285

Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y, Nishida I (1992) Genetically engineered alteration in the chilling sensitivity of plants. Nature **356:** 710-713

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant **15:** 493-497

Murphy DJ, Vance J (1999) Mechanisms of lipid-body formation. Trends Biochem Sci **24:** 109-115

Nagai J, Bloch K (1966) Enzymatic desaturation of stearyl acyl carrier protein. J Biol Chem **241:** 1925-1927

Nei M, Kumar S (2000) Molecular evolution and phylogenetics [electronic resource] / Masatoshi Nei, Sudhir Kumar. Oxford ; New York : Oxford University Press, 2000,

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 Mol Biol **21:** 267-277

Nykiforuk CL, Furukawa-Stoffer TL, Huff PW, Sarna M, Laroche A, Moloney MM, Weselake RJ (2002) Characterization of cDNAs encoding diacylglycerol acyltransferase from cultures of *Brassica napus* and sucrose-mediated induction of enzyme biosynthesis. BBA-Mol Cell Biol L **1580:** 95-109

Oelkers P, Cromley D, Padamsee M, Billheimer JT, Sturley SL (2002) The DGA1 gene determines a second triglyceride synthetic pathway in yeast. J Biol Chem **277:** 8877-8881

Ohlrogge J, Browse J (1995) Lipid biosynthesis. Plant Cell **7:** 957-970

Ohno S (1970) Evolution by gene duplication. Springer-Verlag, New York

Page RDM, Charleston MA (1997) From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. Mol Phylogenet Evol **7:** 231-240

Palta JP, Whitaker BD, Weiss LS (1993) Plasma membrane lipids associated with genetic variability in freezing tolerance and cold acclimation of Solanum species. Plant Physiol **103:** 793-803

Pan X, Peng FY, Weselake RJ (2015) Genome-wide analysis of *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE* (*PDAT*) genes in plants reveals

the eudicot-wide *PDAT* gene expansion and altered selective pressures acting on the core eudicot *PDAT* paralogs. Plant Physiol **167:** 887-904

Pan X, Siloto RMP, Wickramarathna AD, Mietkiewska E, Weselake RJ (2013) Identification of a pair of phospholipid: diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. J Biol Chem **288:** 24173-24188

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo MJ, Byers R, Chen W, Doron-Faigenboim A, Duke MV, Gong L, Grimwood J, Grover C, Grupp K, Hu G, Lee TH, Li J, Lin L, Liu T, Marler BS, Page JT, Roberts AW, Romanel E, Sanders WS, Szadkowski E, Tan X, Tang H, Xu C, Wang J, Wang Z, Zhang D, Zhang L, Ashrafi H, Bedon F, Bowers JE, Brubaker CL, Chee PW, Das S, Gingle AR, Haigler CH, Harker D, Hoffmann LV, Hovav R, Jones DC, Lemke C, Mansoor S, Rahman MU, Rainville LN, Rambani A, Reddy UK, Rong JK, Saranga Y, Scheffler BE, Scheffler JA, Stelly DM, Triplett BA, Van Deynze A, Vaslin MFS, Waghmare VN, Walford SA, Wright RJ, Zaki EA, Zhang T, Dennis ES, Mayer KFX, Peterson DG, Rokhsar DS, Wang X, Schmutz J (2012) Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature **492:** 423-427

Peelman F, Verschelde J, Vanloo B, Ampe C, Labeur C, Tavernier J, Vandekerckhove J, Rosseneu M (1999) Effects of natural mutations in lecithin:cholesterol acyltransferase on the enzyme structure and activity. J Lipid Res **40:** 59-69

Peelman F, Vinaimont N, Verhee A, Vanloo B, Verschelde J, Labeur C, Seguret-Mace S, Duverger N, Hutchinson G, Vandekerckhove J, Tavernier J, Rosseneu M (1998) A proposed

architecture for lecithin cholesterol acyltransferase (LCAT): Identification of the catalytic triad and molecular modeling. Protein Sci **7:** 587-599

Pelham HRB (2000) Using sorting signals to retain proteins in endoplasmic reticulum. Methods Enzymol **327:** 279-283

Pereira SL, Huang YS, Bobik EG, Kinney AJ, Stecca KL, Packer JCL, Mukerji P (2004) A novel ω3-fatty acid desaturase involved in the biosynthesis of eicosapentaenoic acid. Biochem J **378:** 665-671

Petschnigg J, Wolinski H, Kolb D, Zelling G, Kurat CF, Natter K, Kohlwein SD (2009) Good fat, essential cellular requirements for triacylglycerol synthesis to maintain membrane homeostasis in yeast. J Biol Chem **284:** 30981-30993

Posada D (2008) jModelTest: Phylogenetic model averaging. Mol Biol Evol **25:** 1253-1256

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. Nucleic Acids Res **40:** D290-D301

Qi B, Fraser T, Mugford S, Dobson G, Sayanova O, Butler J, Napier JA, Stobart AK, Lazarus CM (2004) Production of very long chain polyunsaturated omega-3 and omega-6 fatty acids in plants. Nat Biotechnol **22:** 739-745

Rao S, Abdel-Reheem M, Bhella R, McCracken C, Hildebrand D (2008) Characteristics of high α-linolenic acid accumulation in seed oils. Lipids **43:** 749-755

Rani SH, Krishna THA, Saha S, Negi AS, Rajasekharan R (2010) Defective in Cuticular Ridges (DCR) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. J Biol Chem **285:** 38337-38347

Rickauer M, Brodschelm W, Bottin A, Véronési C, Grimal H, Esquerré-Tugayé MT (1997) The jasmonate pathway is involved differentially in the regulation of different defence responses in tobacco cells. Planta **202:** 155-162

Rochester CP, Bishop DG (1984) The role of lysophosphatidylcholine in lipid synthesis by developing sunflower (*Helianthus annuus* L.) seed microsomes. Arch Biochem Biophys **232:** 249-258

Roth C, Rastogi S, Arvestad L, Dittmar K, Light S, Ekman D, Liberles DA (2007) Evolution after gene duplication: Models, mechanisms, sequences, systems, and organisms. J Exp Zool B Mol Dev Evol **308:** 58-73

Routaboul JM, Benning C, Bechtold N, Caboche M, Lepiniec L (1999) The TAG11 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. Plant Physiol Biochem **37:** 831-840

Routaboul JM, Fischer SF, Browse J (2000) Trienoic fatty acids are required to maintain chloroplast function at low temperatures. Plant Physiol **124:** 1697-1705

Routaboul JM, Skidmore C, Wallis JG, Browse J (2012) Arabidopsis mutants reveal that short-and long-term thermotolerance have different requirements for trienoic fatty acids. J Exp Bot **63:** 1435-1443

Ruiz-Lopez N, Haslam RP, Venegas-Caleron M, Larson TR, Graham IA, Napier JA, Sayanova O (2009) The synthesis and accumulation of stearidonic acid in transgenic plants: a novel source of 'heart-healthy' omega-3 fatty acids. Plant Biotechnol J **7:** 704-716

Ruiz-López N, Sayanova O, Napier JA, Haslam RP (2012) Metabolic engineering of the omega-3 long chain polyunsaturated fatty acid biosynthetic pathway into transgenic plants. J Exp Bot **63:** 2397-2410

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 Physiol **141:** 1533-1543

Samala S, Yan J, Baird WV (1998) Changes in polar lipid fatty acid composition during cold acclimation in 'Midiron' and 'U3' bermudagrass. Crop Sci **38:** 188-195

Sandager L, Gustavsson MH, Ståhl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H, Stymne S (2002) Storage lipid synthesis is non-essential in yeast. J Biol Chem **277:** 6478-6482

Sasaki Y, Hakamada K, Suama Y, Nagano Y, Furusawa I, Matsuno R (1993) Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant. J Biol Chem **268:** 25118-25123

Scherer GFE, Ryu SB, Wang X, Matos AR, Heitz T (2010) Patatin-related phospholipase A: Nomenclature, subfamilies and functions in plants. Trends Plant Sci **15:** 693-700

Schmidt H, Heinz E (1990) Involvement of ferredoxin in desaturation of lipid-bound oleate in chloroplasts. Plant Physiol **94:** 214-220

Schnurr JA, Shockey JM, De Boer G-, Browse JA (2002) Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. Plant Physiol **129:** 1700-1709

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature **432:** 779-782

Shanklin J, Cahoon EB (1998) Desaturation and related modifications of fatty acids. Annu Rev Plant Physiol Plant Mol Biol **49:** 611-641

Sharma KK, Schuhmann H, Schenk PM (2012) High lipid induction in microalgae for biodiesel production. Energies **5:** 1532-1553

Shockey JM, Fulda MS, Browse JA (2002) Arabidopsis contains nine long-chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism. Plant Physiol **129:** 1710-1722

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. Plant Cell **18:** 2294-2313

Shorrosh BS, Roesler KR, Shintani D, Van De Loo FJ, Ohlrogge JB (1995) Structural analysis, plastid localization, and expression of the biotin carboxylase subunit of acetyl-coenzyme A carboxylase from tobacco. Plant Physiol **108:** 805-812

Sinclair AJ, Attar-Bashi NM, Lib D (2003) What is the role of alpha-linolenic acid for mammals? Lipids **37:** 1113-1123

Siloto RMP, Truksa M, He X, McKeon T, Weselake RJ (2009a) Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system. Lipids **44:** 963-973

Siloto RMP, Truksa M, Brownfield D, Good AG, Weselake RJ (2009b) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and characterization of *Brassica napus* DGAT1 mutagenized libraries. Plant Physiol Bioch **47:** 456-461

Slack CR, Campbell LC, Browse JA, Roughan PG (1983) Some evidence for the reversibility of the cholinephosphotransferasecatalysed reaction in developing linseed cotyledons *in vivo*. BBA-Lipid Lipid Met **754:** 10-20

Slack CR, Roughan PG, Browse JA, Gardiner SE (1985) Some properties of cholinephosphotransferase from developing safflower cotyledons. BBA-Lipid Lipid Met **833:** 438-448

Sørensen BM, Furukawa-Stoffer TL, Marshall KS, Page EK, Mir Z, Forster RJ, Weselake RJ (2005) Storage lipid accumulation and acyltransferase action in developing flaxseed. Lipids **40:** 1043-1049

Spatz L, Strittmatter P (1971) A form of cytochrome *b*5 that contains an additional hydrophobic sequence of 40 amino acid residues. Proc Natl Acad Sci U S A **68:** 1042-1046

Sperling P, Linscheid M, Stöcker S, Mühlbach HP, Heinz E (1993) *In vivo* desaturation of *cis*-Δ9-monounsaturated to *cis*-Δ9,12-diunsaturated alkenylether glycerolipids. J Biol Chem **268:** 26935-26940

Ståhl U, Carlsson AS, Lenman M, Dahlqvist A, Huang B, Banaś W, Banaś A, Stymne S (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from Arabidopsis. Plant Physiol **135:** 1324-1335

Ståhl U, Stålberg K, Stymne S, Ronne H (2008) A family of eukaryotic lysophospholipid acyltransferases with broad specificity. FEBS Lett **582:** 305-309

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:** 2688-2690

Stobart K, Mancha M, Lenman M, Dahlqvist A, Stymne S (1997) Triacylglycerols are synthesised and utilized by transacylation reactions in microsomal preparations of developing safflower (*Carthamus tinctorius* L.) seeds. Planta **203:** 58-66

Stymne S, Glad G (1981) Acyl exchange between oleoyl-CoA and phosphatidylcholine in microsomes of developing soya bean cotyledons and its role in fatty acid desaturation. Lipids **16:** 298-305

Stymne S, Stobart AK (1987) Triacylglycerol biosynthesis. In PK Stumpf, ed, The biochemistry of plants. Academic Press, New York, pp 175-214

Stymne S, Stobart AK (1984) Evidence for the reversibility of the acyl-CoA:lysophosphatidylcholine acyltransferase in microsomal preparations from developing safflower (*Carthamus tinctorius* L.) cotyledons and rat liver. Biochem J **223:** 305-314

Sun L, Ouyang C, Kou S, Wang S, Yao Y, Peng T, Xu Y, Tang L, Chen F (2011) Cloning and characterization of a cDNA encoding type 1 diacylglycerol acyltransferase from sunflower (*Helianthus annuus* L.). Z Naturforsch C **66(1-2):** 63-72

Suyama M, Torrents D, Bork P (2006) PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res **34:** W609-W612

Tamari F, Hinkley CS, Ramprashad N (2013) A comparison of DNA extraction methods using Petunia hybrida tissues. Journal of Biomolecular Techniques **24:** 113-118

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. Mol Biol Evol **28:** 2731-2739

Tang J, Scarth R, Fristensky B (2003) Effects of genomic position and copy number of Acyl-ACP thioesterase transgenes on the level of the target fatty acids in Brassica napus L. Mol Breed **12:** 71-81

Taylor DC, Weber N, Hogge LR, Underhill EW (1990) A simple enzymatic method for the preparation of radiolabeled erucoyl-CoA and other long-chain fatty acyl-CoAs and their characterization by mass spectrometry. Anal Biochem **184:** 311-316

Thiessen D (2004) Optimization of feed peas, canola and flaxseed for Auq feeds: the Canadian Prairie Perspective. In LE Cruz Suare, D Ricque Marie, MG Nieto Lopez, D Villarreal, Scholz, U. y Gonzalez, M., eds, Avances en Nutricion Acuicola VII. Memorias del VII symposium Internacional de Nutricion Acuicola.

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

Tjellström H, Strawsine M, Ohlrogge JB (2015) Tracking synthesis and turnover of triacylglycerol in leaves. J Exp Bot

Tjellström H, Yang Z, Allen DK, Ohlrogge JB (2012) Rapid kinetic labeling of Arabidopsis cell suspension cultures: Implications for models of lipid export from plastids. Plant Physiol **158:** 601-611

Torgerson DG, Singh RS (2004) Rapid evolution through gene duplication and subfunctionalization of the testes-specific α4 proteasome subunits in drosophila. Genetics **168:** 1421-1432

Tumaney AW, Shekar S, Rajasekharan R (2001) Identification, purification, and characterization of monoacylglycerol acyltransferase from developing peanut cotyledons. J Biol Chem **276:** 10847-10852

Turchetto-Zolet AC, Maraschin FS, De Morais GL, Cagliari A, Andrade CMB, Margis-Pinheiro M, Margis R (2011) Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. BMC Evol Biol **11:** 263-263

van Erp H, Bates PD, Burgal J, Shockey J, Browse J (2011) Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol **155:** 683-693

van Erp H, Kelly AA, Menard G, Eastmond PJ (2014) Multigene engineering of triacylglycerol metabolism boosts seed oil content in Arabidopsis. Plant Physiol **165:** 30-36

Varsha Wesley S, Helliwell CA, Smith NA, Wang M, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant Journal **27:** 581-590

Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, Alting-Mees M, Nowak J, Cloutier S, Deyholos M, Bekkaoui F, Sharpe A, Wang E, Rowland G, Selvaraj G, Datla R (2011) Gene expression analysis of flax seed development. BMC Plant Biol **11:** 74

Vrinten P, Hu Z, Munchinsky MA, Rowland G, Qiu X (2005) Two FAD3 desaturase genes control the level of linolenic acid in flax seed. Plant Physiol **139:** 79-87

Wallis JG, Browse J (2002) Mutants of Arabidopsis reveal many roles for membrane lipids. Prog Lipid Res **41:** 254-278

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:** 230-242

Wang HW, Zhang JS, Gai JY, Chen SY (2006) Cloning and comparative analysis of the gene encoding diacylglycerol acyltransferase from wild type and cultivated soybean. Theor Appl Genet **112:** 1086-1097

Wang Y, Wang X, Paterson AH (2012) Genome and gene duplications and gene expression divergence: A view from plants. Ann N Y Acad Sci **1256:** 1-14

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, Lambert G, Galbraith DW, Grassa CJ, Geraldes A, Cronk QC, Cullis C, Dash PK, Kumar PA, Cloutier S, Sharpe AG, Wong GK, Wang J, Deyholos MK (2012) The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J **72:** 461-473

Weadick CJ, Chang BSW (2012) An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. Mol Biol Evol **29:** 1297-1300

Weigel D, Glazebrook J (2002) Arabidopsis: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem **235:** 40-44

Weselake RJ, Madhavji M, Szarka SJ, Patterson NA, Wiehler WB, Nykiforuk CL, Burton TL, Boora PS, Mosimann SC, Foroud NA, Thibault BJ, Moloney MM, Laroche A, Furukawa-Stoffer TL (2006) Acyl-CoA-binding and self-associating properties of a recombinant 13.3 kDa N-terminal fragment of diacylglycerol acyltransferase-1 from oilseed rape. BMC Biochem **7:** 24-24

Wickramarathna AD, Siloto RMP, Mietkiewska E, Singer SD, Pan X, Weselake RJ (2015) Heterologous expression of flax phosphotidylcholine:diacylglycerol cholinephosphotransferases

(PDCTs) increases polyunsaturated fatty acid content in yeast and Arabidopsis seeds. BMC Biotech **15:** 63

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE **2(8):** e718

Wolter FP, Schmidt R, Heinz E (1992) Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. EMBO J **11:** 4685-4692

Worberg A, Quandt D, Barniske A, Löhne C, Hilu KW, Borsch T (2007) Phylogeny of basal eudicots: Insights from non-coding and rapidly evolving DNA. Org Divers Evol **7:** 55-77

Xu J, Carlsson AS, Francis T, Zhang M, Hoffman T, Giblin ME, Taylor DC (2012) Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. BMC Plant Biol **12:** 4

Xu J, Francis T, Mietkiewska E, Giblin EM, Barton DL, Zhang Y, Zhang M, Taylor DC (2008) Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (DGAT1) gene from *Tropaeolum majus*, and a study of the functional motifs of the DGAT protein using site-directed mutagenesis to modify enzyme activity and oil content. Plant Biotech J **6:** 799-818

Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB (2012) A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, *sn*-2 preference, and evolution. Plant Physiol **160:** 638-652

Yang Z (1997) PAML: A program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci **13:** 555-556

Yang Z, Bielawski JR (2000) Statistical methods for detecting molecular adaptation. Trends Ecol Evol **15:** 496-503

Yang Z, Wong WSW, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol **22:** 1107-1118

Yoon K, Han D, Li Y, Sommerfeld M, Hu Q (2012) Phospholipid:diacylglycerol acyltransferase is a multifunctional enzyme involved in membrane lipid turnover and degradation while synthesizing triacylglycerol in the unicellular green microalga *chlamydomonas reinhardtii*. Plant Cell **24:** 3708-3724

Yoshida I, Sugiura W, Shibata J, Ren F, Yang Z, Tanaka H (2011) Change of positive selection pressure on HIV-1 envelope gene inferred by early and recent samples. PLoS ONE **6:** e18630-e18630

Zhang H, Damude HG, Yadav NS (2012) Three diacylglycerol acyltransferases contribute to oil biosynthesis and normal growth in *Yarrowia lipolytica*. Yeast **29:** 25-38

Zhang M, Fan J, Taylor DC, Ohlrogge JB (2009) DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. Plant Cell **21:** 3885-3901

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong GY, Tarczynski MC, Shen B (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat Genet **40:** 367-372

Zhou XR, Shrestha P, Yin F, Petrie JR, Singh SP (2013) AtDGAT2 is a functional acyl-CoA:diacylglycerol acyltransferase and displays different acyl-CoA substrate preferences than AtDGAT1. FEBS Lett **587:** 2371-2376

Zou J, Wei Y, Jako C, Kumar A, Selvaraj G, Taylor DC (1999) The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. Plant J **19:** 645-653

APPENDIX

List of a book chapter, two research manuscripts, a review article and a patent to which I contributed as a co-author or a co-inventor.

I. Singer SD, Greer MS, Mietkiewska E, **Pan X,** Weselake RJ. Genetic engineering of lipid biosynthesis in the seeds of crucifer species. *Biotechnology of Crucifers* Gupta S.K. ed. July 31 2013

- For this work, I contributed to writing the section of engineering plants for polyunsaturated fatty acids and fatty acids with functional groups.

II. Wickramarathna AD, Siloto RMP, Mietkiewska E, Singer SD, **Pan X**, Weselake RJ. (2015) Heterologous expression of flax *PHOSPHOTIDYLCHOLINE:DIACYGLYCEROL CHOLINEPHOSPHOTRANSFERASES* (*PDCTs*) increases polyunsaturated fatty acid content in yeast and Arabidopsis seeds. BMC Biotechnology. (In press)

- For this work, I performed RNA isolation and cDNA synthesis for *PDCT* gene cloning, and assisted the first author in literature research, manuscript preparation, and submission of phylogenetic tree.

III. Chen G, Woodfield HK, **Pan X,** Harwood JL, Weselake RJ. (2015) Control of carbon flow and acyl-trafficking in plant oil accumulation. Lipids. (In preparation)

- For this work, I contributed to writing the sections of cholinephosphotransferase, phosphatidylcholine diacylglycerol cholinephosphotransferase and phospholipid:diacylglycerol acyltransferase.

IV. Greer MS, **Pan X,** Weselake RJ. (2015) Two clades of type-1 *Brassica napus* diacylglycerol acyltransferase exhibit differences in substrate preference. (In preparation)

- For this work, I assisted in phylogenetic analysis and manuscript preparation.

V. Full patent application on which I contributed as a co-inventor.

Weselake RJ, **Pan X**, Siloto RMP. Enrichment of oils with polyunsaturated fatty acids. Publication number: WO2013181761 A1

- For this work, I performed all experiments and data analyses. In addition, I wrote the first draft of the patent and further modified the patent with co-inventors.