

**University of Alberta**

**Molecular and functional characterization of *sn*-glycerol-3-phosphate  
acyltransferase of plants**

by

Xue Chen

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Plant Science

Department of Agricultural, Food and Nutritional Science

©Xue Chen  
Spring 2011  
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

**Examining Committee**

Dr. Randall Weselake (Supervisor)

Dr. Saleh Shah (Co-Supervisor)

Dr. Jocelyn Ozga

Dr. Enrico Scarpella

Dr. Walter Dixon

Dr. John Harada (External Examiner)

## Abstract

*sn*-Glycerol-3-phosphate acyltransferase (GPAT) catalyzes the acylation of *sn*-1 position of *sn*-glycerol-3-phosphate to produce lysophosphatidic acid and Coenzyme A. GPATs are involved in several lipid synthetic pathways and play important physiological roles in plant development. The present doctoral thesis includes three related studies, which aim to molecularly and functionally characterize several plant *GPAT* genes and the encoded enzymes.

The first study characterized three endoplasmic reticulum-bound GPAT4s encoded by three homologous *GPAT4* genes of *Brassica napus* (oilseed rape), focusing primarily on their functional divergence and physiological roles in plant development and lipid biosynthesis. The three homologous *GPAT4* genes exhibited different expression patterns and altered epigenetic features. Phenotypic rescue of a *gpat4 gpat8 Arabidopsis* double mutant and analysis of the *gpat4* RNAi *B. napus* lines suggested physiological roles for the GPAT4s in cuticle formation of the rosette leaves, early flower development, pollen development and storage lipid biosynthesis. The second study investigated stable internal reference genes for gene expression studies in *B. napus*. This project identified four reliable reference genes to be used in gene expression analysis of *BnGPAT4* homologues in both vegetative tissues and developing seeds. The third study focused on molecular cloning and biochemical characterization of a soluble plastidial GPAT isolated from a chilling-tolerant plant, western wallflower (*Erysimum asperum*). A truncated form of recombinant EaGPAT, with the putative transit peptide deleted, was functionally expressed in yeast. A series of

enzymatic assays were performed in order to determine the optimum *in vitro* reaction conditions for the recombinant EaGPAT. The recombinant EaGPAT was further assayed with different acyl-CoAs and exhibited a substrate preference for 18 carbon unsaturated acyl-CoAs. With this substrate preference, the EaGPAT could potentially be used as a biotechnological tool for improving plant chilling-tolerance or increasing unsaturated fatty acid content of seed oil. Overall, the present doctoral studies revealed the functional divergence and important physiological roles of the *GPAT4s* in *B. napus*, and biochemically characterized a plastidial GPAT from *E. asperum*. The knowledge obtained from these studies provides new insights into the role of GPAT in plants and will be useful for further development of biotechnological approaches to modify seed oil biosynthesis in oleaginous crops.

## Acknowledgements

Many people have helped me in my Ph.D. study. I would like to thank them for their generous help and support that make this doctoral thesis possible. First, I would like to sincerely thank my supervisor, Dr. Randall Weselake, for offering me the opportunity to pursue my Ph.D. degree and for his guidance, instruction, support and encouragement through my Ph.D. study. I am also very grateful to my co-supervisor, Dr. Saleh Shah, and my supervisory committee member Dr. Jocelyn Ozga for their guidance, support and discussion to my Ph.D. research. I would like to thank Dr. Enrico Scarpella for his generous help in my experiment and for serving as my examination committee member; and Dr. Walter Dixon for his support and advice in my Ph.D. program and for serving as my examination committee member. I am also grateful to Dr. John Harada for agreeing to serve as my external examiner.

I thank all other members of the Weselake group. This group has been a source of friendship as well as good advice and collaboration. Special thanks to Dr. Martin Truksa, a former member of the Weselake group, who taught me molecular biology of plant research and provided me guidance and instruction with tremendous patience. I am also very grateful to Crystal Synder for her encouragement and critical thinking and comments to my research and thesis/manuscript writing. Many thanks to Robin Hryniuk for her hard working that greatly facilitated the completion of my thesis. I also thank Dr. Rodrigo Siloto for his help and encouragement in my Ph.D. study.

I also thank Jody Forslund for her administrative support and the past and present members of the greenhouse of the Department of AFNS and the Microscopy unit of the Department of Biological Science for their countless technical support. I thank all my friends in the Department of AFNS for their help and encouragement. I also acknowledge the financial support of my Ph.D. program provided by Alberta Innovates-BioSolutions, the Alberta Canola Producers Commission, the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, and the Canada Research Chairs Program.

Finally, I thank my family for their everlasting love and support.

## Table of Contents

<b>Chapter 1. Introduction.....</b>	<b>1</b>
<b>Chapter 2. Literature Review .....</b>	<b>5</b>
2.1 Storage lipid biosynthesis in plants .....	5
2.1.1 Fatty acid synthesis .....	5
2.1.2 Triacylglycerol assembly .....	10
2.2 Seed oil modification in <i>Brassica napus</i> .....	14
2.3 The sn-glycerol-3-phosphate acyltransferase family in plants .....	17
2.3.1 Plastidial sn-glycerol-3-phosphate acyltransferase.....	19
2.3.2 Mitochondrial sn-glycerol-3-phosphate acyltransferase.....	21
2.3.3 ER-bound sn-glycerol-3-phosphate acyltransferase .....	21
2.3.4 A putative <i>Arabidopsis</i> sn-glycerol-3-phosphate acyltransferase 9 may play an important role in eukaryotic lipid metabolism.....	23
2.4 Plant lipid polymer matrices- cutin and suberin.....	24
2.4.1 Tissue distributions of cutin and suberin .....	26
2.4.2 Monomer compositions and chemical structures of cutin and suberin .....	28
2.4.3 Identification of genes for the biosynthesis of cutin and suberin in <i>Arabidopsis</i> .....	32
2.5 Polyploid genetic background of <i>Brassica</i> species .....	36
2.5.1 Evolution of <i>Brassica</i> species and <i>Arabidopsis</i> .....	36
2.5.2 Comparative mapping between <i>Brassica</i> and <i>Arabidopsis</i> .....	38
2.5.3 The possible fates of duplicated genes after polyploidy.....	38

2.5.4	International sequencing and genomics projects for <i>Brassica</i> species	39
2.6	Summary.....	40
2.7	Reference.....	40

**Chapter 3. Three homologous genes encoding *sn*-glycerol-3-phosphate acyltransferase 4 exhibit functional divergence in *Brassica napus* ..... 55**

3.1	Introduction .....	55
3.2	Results .....	58
3.2.1	Identification of <i>GPAT4</i> homologues in <i>Brassica napus</i> .....	58
3.2.2	Genome origins of the three <i>BnGPAT4</i> homologues.....	59
3.2.3	<i>BnGPAT4</i> homologues encode functional GPAT enzymes with different levels of polypeptide accumulation .....	63
3.2.4	The C-termini of the <i>BnGPAT4</i> enzymes are critical in regulating the protein metabolism .....	65
3.2.5	The three <i>BnGPAT4</i> homologues exhibit distinct expression patterns .....	68
3.2.6	Epigenetic variation among the <i>BnGPAT4</i> homologues .....	73
3.2.7	Phenotypic rescue of <i>Arabidopsis gpat4 gpat8</i> double mutant lines .....	74
3.2.8	<i>B. napus gpat4</i> RNAi lines exhibit alterations in cuticle load and stomatal structure resulting in increased water loss .....	78
3.2.9	<i>B. napus gpat4</i> RNAi exhibited abnormal inflorescence and severely reduced seed yield .....	84

3.2.10	Preliminary results indicated fatty acid composition changes in the <i>gpat4 B. napus</i> seeds .....	87
3.3	Discussion.....	93
3.3.1	Genome origins and evolutionary fates of <i>BnGPAT4</i> homologues... ..	93
3.3.2	Possible mechanisms involved in regulating the level of BnGPAT4 polypeptide accumulation in yeast .....	95
3.3.3	BnGPAT4 may be involved in suberin biosynthesis in root and seed coat.....	96
3.3.4.	The cuticle layer is closely related to stomatal development and structure. ....	97
3.3.5.	<i>GPAT4</i> plays important roles in regulating the development of inflorescence and pollen grain in <i>B.napus</i> .....	99
3.3.6.	<i>GPAT4</i> plays important roles in regulating the development of inflorescence and pollen grain in <i>B.napus</i> .....	102
3.3.7.	<i>GPAT4</i> may be involved in the storage lipid biosynthesis.....	103
3.4	Conclusion .....	104
3.5	Methods .....	105
3.5.1	Plant Materials .....	105
3.5.2	Cloning of <i>GPAT4</i> genes from <i>B. napus</i> .....	105
3.5.3	Heterologous Expression in Yeast and <i>in vitro</i> GPAT enzyme assay .....	105
3.5.4	Western blot.....	107

3.5.5	Promoter-GUS fusion analysis .....	107
3.5.6	TaqMan qRT-PCR.....	108
3.5.7	SYBR-green qRT-PCR.....	109
3.5.8	Genomic DNA cytosine methylation analysis.....	110
3.5.9	Constructs for expression of different <i>BnGPAT4</i> homologues in <i>Arabidopsis gpat4 gpat8</i> double mutant line.....	110
3.5.10	Preparation of RNAi construct for silencing <i>GPAT4</i> homologues in <i>B. napus</i> . .....	111
3.5.11	Generating <i>gpat4</i> RNAi <i>B. napus</i> lines .....	111
3.5.12	Toluidine blue test .....	111
3.5.13	Preparation of surface cast of live leaves and microscopy .....	112
3.5.14	Analysis of cutin monomers .....	112
3.5.15	Analysis of fatty acid composition of <i>B. napus</i> seed oil.....	113
3.6	Footnote .....	113
3.7	References .....	114

**Chapter 4. A survey of quantitative real-time PCR internal reference genes  
for expression studies in *Brassica napus* ..... 124**

4.1	Introduction .....	124
4.2	Results .....	126
4.2.1	Selection of candidate <i>B. napus</i> internal references.....	126
4.2.2	PCR specificity and amplification efficiency analysis.....	127
4.2.3	Expression stability of candidate reference genes .....	129
4.2.4	Validation of the candidate reference genes and the normalization	

	factors. ....	132
4.2.5	Comparison between the expression patterns of overall mRNA and the candidate reference genes .....	132
4.3	Discussion.....	135
4.4	Materials and methods.....	137
4.4.1	Plant material .....	137
4.4.2	RNA extraction .....	137
4.4.3	External reference and reverse transcription .....	138
4.4.4	QRT-PCR primer design and reaction.....	138
4.4.5	QRT-PCR data and stability of expression analysis.....	139
4.4.6	Quantification of overall mRNA in embryo via radio-labelling reverse transcription .....	140
4.5	Footnote .....	140
4.6	References .....	141

**Chapter 5. Molecular cloning and functional characterization of a  
plastidial sn-glycerol-3-phosphate acyltransferase from the chilling-  
tolerant plant *Erysimum asperum*.....138**

5.1	Introduction .....	144
5.2	Results .....	146
5.2.1	Cloning a full-length plastidial <i>GPAT</i> cDNA from <i>E. asperum</i> ...	146
5.2.2	Prediction of transit peptide .....	147
5.2.3	Transit peptide study using fluorescent protein as a fusion tag....	149
5.2.4	The EaGPAT without putative transit peptide exhibited GPAT	

	enzyme activity.....	152
5.3	Discussion.....	161
5.3.1	Prediction of the transit peptide of EaGPAT .....	161
5.3.2	Functional expression of the mature EaGPAT in yeast .....	162
5.3.3	Substrate preference of the mature EaGPAT and potential use of enzyme in the metabolic engineering of fatty acid composition in plants.....	162
5.4	Conclusion .....	164
5.5	Methods .....	164
5.5.1	Plant materials.....	164
5.5.2	Cloning of plastidial <i>GPAT</i> full-length cDNA from <i>E. asperum</i> .	165
5.5.3	Heterologous expression of the <i>EaGPAT</i> in yeast.....	166
5.5.4	Purification of the mature EaGPAT.....	166
5.5.5	<i>In vitro</i> GPAT assay .....	167
5.5.6	Western blot analysis .....	168
5.5.7	Preparation of <i>EaGPAT-eGFP</i> constructs .....	169
5.5.8	<i>Arabidopsis</i> transformation and screening the transgenic plants..	169
5.6	References .....	170
<b>Chapter 6. General discussion .....</b>		<b>174</b>
6.1	Molecular and functional characterization of three <i>GPAT4</i> homologues of <i>B. napus</i> .....	174
6.2	A survey of quantitative real-time PCR internal reference genes for expression studies in <i>B. napus</i> .....	179

6.3	Characterization of a recombinant plastidial GPAT from <i>E. asperum</i> .....	180
6.4	Concluding comments .....	184
6.5	References .....	184
<b>Appendix 1.....</b>		<b>187</b>
<b>Appendix 2.....</b>		<b>199</b>

## List of Tables

Table 4.1 Surveyed internal reference genes of <i>Brassica napus</i> .....	128
Table A1 Primers used in the Chapter 3.....	187
Table A2 Raw data of PCR efficiency, Ct, RQ and s.e. value of biological replicates for tested samples.....	191

## List of Figures

Figure 2.1. Simplified scheme of <i>de novo</i> fatty acid synthesis in plant plastids.....	8
Figure 2.2. Fatty acid synthesis termination, desaturation, and export from plastid .....	9
Figure 2.3. Sequential acyl-CoA-dependent acylation of the glycerol backbone to form triacylglycerol (TAG) in the ER.....	11
Figure 2.4. Triacylglycerol (TAG) synthesis pathways in the ER.....	15
Figure 2.5. Phylogenetic tree of the <i>sn</i> -glycerol-3-phosphate acyltransferase (GPAT) family of <i>Arabidopsis</i> .....	18
Figure 2.6. Phylogenetic analysis of the <i>sn</i> -glycerol-3-phosphate acyltransferases (GPATs) from <i>Arabidopsis</i> (At), <i>Homo sapiens</i> (Hs) and <i>Saccharomyces cerevisiae</i> (Sc).....	25
Figure 2.7. Structure of plant cuticle and suberized root cell wall.....	27
Figure 2.9. Structures of common cutin and suberin monomers, and typical ranges of composition values in plants.....	29
Figure 2.10. Structures of hypothetical monomer connectivity patterns.....	31
Figure 2.11. Possible reactions catalyzed by fatty acid hydroxylase (CYP86A), long-chain acyl-CoA synthetase (LACS) and <i>sn</i> -glycerol-3- phosphate acyltransferase (GPAT) in the lipid polyester biosynthesis pathway.....	35
Figure 3.1. Genome structure of the BnGPAT4 homologues.....	62
Figure 3.2. Sequence analysis of the <i>GPAT4</i> genes from <i>B. napus</i> and its diploid progenitors, <i>B. rapa</i> and <i>B. oleracea</i> .....	63

Figure 3.3. The three <i>BnGPAT4</i> homologues encode functional GPAT enzymes but with different activities.....	65
Figure 3.4. Comparison of the amino acid sequences of the three BnGPAT4 proteins.....	67
Figure 3.5. The C-termini of the BnGPAT4s are important for controlling the level of polypeptide accumulation.....	68
Figure 3.6. Histochemical analysis of the GUS activities directed by individual <i>BnGPAT4</i> promoters in transgenic <i>Arabidopsis</i> .....	72
Figure 3.7. Gene expression patterns of the three BnGPAT4 homologues.....	73
Figure 3.8. Cytosine DNA methylation patterns of the first exon regions of the three <i>BnGPAT4</i> homologues.....	76
Figure 3.9. Toluidine blue test of the cuticle permeability of <i>gpat4 gpat8</i> <i>Arabidopsis</i> transformed with individual <i>BnGPAT4</i> homologues.....	78
Figure 3.10. Silencing of <i>GPAT4</i> homologues in <i>B. napus</i> results in decreased cutin monomer load.....	80
Figure 3.11. Cuticle permeability of <i>B. napus gpat4</i> rosette leaves at different stages with the toluidine blue test.....	82
Figure 3.12. Comparison of stomatal structures between wild type and <i>gpat4 B.</i> <i>napus</i> lines.....	84
Figure 3.13. Abnormal inflorescence development in <i>gpat4</i> RNAi lines.....	86
Figure 3.14. Aborted seed development in <i>gpat4</i> RNAi lines.....	87
Figure 3.15. Abnormal pollen grains of the <i>gpat4</i> lines.....	90
Figure 3.16. Reduced seed yield of <i>gpat4 B. napus</i> .....	91

Figure 3.17. Fatty acid composition analysis indicated decreased content of 18:1 and increased content of 18:2 and 18:3 in the <i>gpat4 B. napus</i> T <sub>1</sub> seed oil.....	92
Figure 3.18. Gene expression pattern of the <i>napin</i> orthologous gene, At4g27150, in <i>Arabidopsis</i> .....	101
Figure 4.1. Gene expression stability evaluated by geNORM.....	131
Figure 4.2. Changes in mRNA content and expression levels of selected internal reference genes of the maturing embryos of <i>B. napus</i> .....	134
Figure 5.1 Amino acid sequence alignment of EaGPAT and other plastidial GPATs.....	148
Figure 5.2. Western blot showing the eGFP signal of protein samples prepared from different <i>Arabidopsis</i> lines.....	151
Figure 5.3. The yeast-expressing mature EaGPAT exhibited GPAT enzyme activity.....	153
Figure 5.4. Purification of mature EaGPAT.....	154
Figure 5.5. Dependence of EaGPAT activity on concentration and pH.....	156
Figure 5.6. Dependence of EaGPAT activity on glycerol-3-phosphate (G3P) concentration and lysophosphatidic acid (LPA) produced by the catalytic action of EaGPAT as a function of time.....	158
Figure 5.7. Acyl-CoA substrate specificity of the mature EaGPAT.....	160
Figure A1. PCR amplification of the 4th putative <i>BnGPAT4</i> homologue.....	193
Figure A2. Sequence alignment of the promoter regions of the three <i>BnGPAT4</i> homologues.....	193

Figure A3. Developing <i>B. napus</i> seeds at the stages from 20 to 40-DAP.....	193
Figure A4. Schematic diagram of the constructs used for expression of different <i>BnGPAT4</i> homologues in <i>Arabidopsis gpat4 gpat8</i> double mutant lines.....	194
Figure A5. Schematic diagram of the construct used for RNAi silencing <i>BnGPAT4</i> homologues in <i>B. napus</i> .....	194
Figure A6. Comparison of the <i>BnGPAT6</i> expression levels in the rosette leaves of wild type and RNAi <i>B. napus</i> lines. n=3.....	194
Figure A7. Water loss of excised rosette leaves recoded over 60 min measured as a percentage of the initial weight. n=3.....	195
Figure A8. Amino acid sequence alignment of EaGPAT and other plastidial GPATs.....	195
Figure A9. Analysis of the fatty acid composition of total lipids extracted from the <i>E. asperum</i> leaves.....	198

## List of abbreviations

ACCase	acetyl-CoA carboxylase
ACT	actin
ACP	acyl carrier protein
ATP	adenosine triphosphate
BSTFA	<i>N,O</i> -bis(trimethylsilyl)-trifluoroacetamide
BSA	bovine serum albumin
cDNA	complementary DNA
CoA	coenzyme A
CPT	CDP-choline: diacylglycerol cholinephosphotransferase
CTAB	cetyl-trimethyl-ammonium bromide
DAF	days after flowering
DAG	<i>sn</i> -1, 2-diacylglycerol
DAP	days after pollination
DGAT	diacylglycerol acyltransferase
DGDG	digalactosyldiacylglycerol
DTT	dithiothreitol

EAR	enoyl-ACP reductase
EDTA	ethylene diamine tetraacetate
EST	expression sequence tag
ER	endoplasmic reticulum
FA	fatty acid
FAME	fatty acid methyl ester
FAS	fatty acid synthase
FATB	acyl-ACP thioesterase B
FFA	free fatty acid
G3P	glycerol-3-phosphate
GC-MS	gas chromatography-mass spectrometry
gDNA	genomic DNA
GFP	green fluorescent protein
GPAT	glycerol-3-phosphate acyltransferase
HAD	hydroxyacyl-ACP dehydrase
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
KAR	ketoacyl-ACP reductase;

KAS	ketoacyl-ACP synthase
LACS	long-chain acyl-CoA synthetase
LPA	lysophosphatidic acid
LPAAT	lysophosphatidic acid acyltransferase
LPC	lysophosphatidylcholine
LPCAT	lysophosphatidylcholine acyltransferase
MGDG	monogalactosyldiacylglycerol
mRNA	messenger RNA
MYA	million years ago
PA	phosphatidic acid
PAP	phosphatidic acid phosphatase
PAT	phosphinothricin n-acetyl transferase
PC	phosphatidylcholine
PCR	polymerase chain reaction
PDAT	phospholipid: diacylglycerol acyltransferase
PG	phosphatidylglycerol
PLA2	phospholipase A2

PP2A	protein phosphatase 2A subunit A3
PUFA	polyunsaturated fatty acids
PVDF	polyvinylidene difluoride
RNAi	RNA interference
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SQDG	sulfoquinovosyldiacylglycerol (SQDG)
TAG	triacylglycerol
TLC	thin layer chromatography
UBC	ubiquitin-conjugating enzyme
UP	unknown protein
WT	wild-type
X-Gluc A	5-bromo-4-chloro-3-indolyl-b-glucuronic acid

## Chapter 1.

### Introduction

Plant lipids are composed of a broad group of fatty acids and their derivatives, such as glycerolipids, lipid polyesters and sterols. They are involved in a diverse range of metabolic reactions and play important physiological roles in plant development, as major components of cellular membranes, seed energy reserves, extracellular protective layers and signaling molecules.

In plants, the biosynthesis of these different types of lipids is controlled by a complex network of genes and proteins. In *Arabidopsis*, it has been estimated that more than 600 genes encoding enzymes or regulatory factors are involved in the lipid metabolic network, which includes at least 120 enzymatic reactions (Li-Beisson et al., 2010). The complexity of this metabolic network is caused in part by the fact that one enzymatic reaction is often not devoted to a single lipid metabolic pathway. The product of one enzymatic reaction can serve as a substrate for several different subsequent reactions, resulting in the formation of a variety of final products, often with diverse physiological roles in plant development.

As an example, *sn*-glycerol-3-phosphate acyltransferase (GPAT), the target enzyme of the current doctoral study, catalyzes the esterification of a fatty acyl moiety from acyl-CoA or acyl-ACP (where CoA is coenzyme A and ACP is acyl carrier protein) to the *sn*-1 position of *sn*-glycerol-3-phosphate (G3P), resulting in formation of lysophosphatidic acid (LPA) (Zheng et al., 2003). LPA is a substrate for the production of several important glycerolipid intermediates

including extracellular lipid polyesters, storage lipids, and membrane lipids. Recent studies of *Arabidopsis* ER-bound GPAT confirmed an important physiological role of this enzyme in the formation of extracellular lipid polyesters (i.e., cutin and suberin) (Li-Beisson et al., 2010). Additionally, because the biosynthesis of triacylglycerol (TAG) and membrane lipids mainly occurs in the ER, the ER-bound GPAT is also believed to be involved in the biosynthesis of these lipids (Gidda et al., 2009). Unlike the ER-bound GPAT, a soluble GPAT located in the plastidial stroma is involved in the biosynthesis of plastidial membrane phospholipids and plays an important role in regulation of plastidial lipid metabolism (Kunst et al., 1986; Xu et al., 2006).

GPAT has been considered a potential target to genetically manipulate the storage lipid synthesis pathway (Jain et al., 2000). For example, in oilseed crops such as *Brassica napus*, introducing a heterologous GPAT enzyme with a specific acyl substrate preference into the storage lipid synthesis pathway could lead to altered fatty acid composition in the seed oil.

*Brassica napus* is one of the most important oilseed crops in Canada. Numerous studies have focused on improving *B. napus* seed oil content and modifying fatty acid profiles for specific nutritional or industrial purposes (Verwoert et al., 1995; Voelker et al., 1996; Hawkins and Kridl 1998; Scarth and Tang, 2006). As an allotetraploid (AACC), *B. napus* and its progenitors (i.e., *B. rapa* and *B. oleracea*) evolved from multiple rounds of polyploidy along with extensive genome rearrangement. The complex genetic background, together with an incomplete understanding of the mechanisms involved in lipid biosynthesis,

are obstacles for genetically manipulating the key lipogenic pathways. Thus, for the success in seed oil engineering and improvement, it is fundamentally important to understand the mechanisms regulating the lipid biosynthesis within the complex polyploid genetic background of *B. napus*.

The present doctoral study includes three related projects, which aim to molecularly and functionally characterize *GPAT* genes and their encoding proteins from *B. napus* and *Erysimum asperum*.

1. The first project investigated three homologous *GPAT4* genes from *B. napus*, focusing primarily on their functional divergence, and physiological roles in plant development and lipid synthetic pathways. It was hypothesized that the three *BnGPAT4* homologues may have evolved through functional divergence, playing important roles in lipid polyester biosynthesis and seed development.
2. The second project investigated stable internal reference genes for gene expression studies in *B. napus*. This project identified four reliable reference genes to be used in gene expression analysis of *BnGPAT4* homologues in both vegetative tissues and developing seeds.
3. The third project functionally characterized a soluble recombinant plastidial GPAT enzyme from a chilling-tolerant plant, western wallflower (*Erysimum asperum*). It was hypothesized that a plastidial GPAT from a chilling-tolerant plant would exhibit preference for unsaturated fatty acyl-Coenzyme As (CoAs). This plastidial GPAT may be potentially useful in

biotechnological strategies aimed at increasing the unsaturated fatty acid content of seed oil.

The thesis begins with a literature review of relevant background information followed by manuscripts covering each of the three projects. The manuscript chapters are followed by a chapter on general discussion and future directions. The appendixes of this thesis contain supplemental research data (Appendix 1) and a review article (Appendix 2) on acyltransferase action in plant triacylglycerol biosynthesis to which I contributed as a co-author.

## Chapter 2.

### Literature Review

In this chapter, background knowledge related to the present doctoral project will be presented in the following five sections: storage lipid biosynthesis in plants, the current status of seed oil modification in *Brassica napus*, the *sn*-glycerol-3-phosphate acyltransferase (GPAT) family in plants, plant lipid polyester matrices and polyploid genetic background of *Brassica* species.

#### 2.1 Storage lipid biosynthesis in plants

In most plants, the major component of storage lipid is triacylglycerol (TAG), which is produced in seeds through fatty acid synthesis and subsequent acylation of the glycerol backbone with fatty acyl moieties.

##### 2.1.1 Fatty acid synthesis

Unlike animals, where fatty acids are synthesized in the cytosol (Numa, 1974), *de novo* synthesis of fatty acids in plants occurs in the plastids, where acetyl-CoA carboxylase (ACCase) and fatty acid synthase complex (FAS) are involved (Ohlrogge and Browse, 1995). ACCase catalyzes the ATP-dependent conversion of acetyl-CoA to malonyl-CoA. FAS catalyzes the elongation of the acyl chain by adding two-carbon units donated from malonyl-CoA in a series of repetitive reactions.

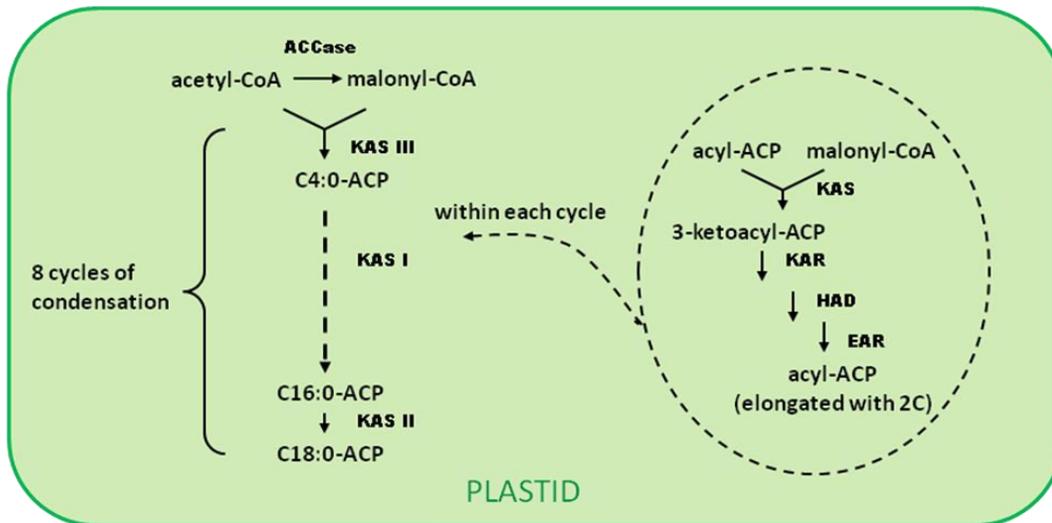
In plants, ACCase occurs in two structurally distinct forms. Type I ACCase, a large multifunctional enzyme (>200 kDa), is located primarily in the cytosol and is similar to the enzyme forms found in yeast and mammals. Type II ACCase, a four-subunit complex, is located predominately in plastids and is

similar to that found in *Escherichia coli*. In most plants, the plastidial ACCase produces the majority of the malonyl-CoA for fatty acid synthesis, whereas the cytosolic ACCase produces malonyl-CoA for acyl-chain extension beyond C18 and for the production of other metabolites (Schmidt and Ohlrogge, 2008).

There are also two types of FAS complex. Type I FAS complex is characterized by the large and multifunctional protein found in mammals and yeast, while Type II FAS complex has multiple subunits and is found in plants and prokaryotes. Plant FAS consists of at least 8 dissociable subunits catalyzing successive cycles of two-carbon unit addition (provided by malonyl-CoA) to the growing fatty acyl chain (Slabas et al., 2001). The first cycle is initiated by  $\beta$ -ketoacyl-ACP synthase III (KASIII, where ACP is acyl carrier protein), which results in the condensation of acetyl-CoA and malonyl-CoA to form a 4:0 acyl-chain. Subsequent condensation cycles from 4:0 to 16:0 and 16:0 to 18:0 are initiated by KAS I and KAS II, respectively, also using malonyl-CoA as a two-carbon unit donor (Ohlrogge and Jaworski, 1997). Within each two-carbon elongation cycle, the immediate product of KAS, 3-ketoacyl-ACP, is converted to a saturated acyl-ACP through three steps catalyzed sequentially by  $\beta$ -ketoacyl-ACP reductase (KAR),  $\beta$ -hydroxyacyl-ACP dehydrase (HAD), and enoyl-ACP reductase (EAR) (Ohlrogge and Jaworski, 1997) (Fig. 2.1).

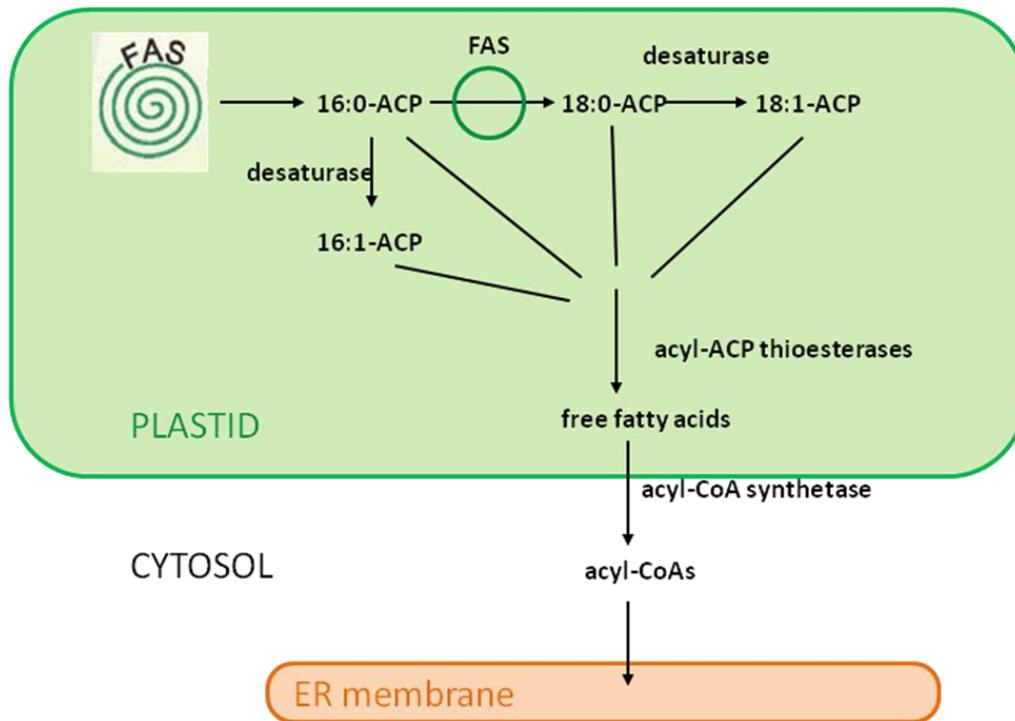
In most plants, plastidial fatty acid synthesis is terminated once the acyl chain reaches 16-18 carbons in length. A double bond may be introduced into palmitoyl-ACP (16:0) or stearoyl-ACP (18:0) by an acyl-ACP desaturase, resulting in the formation of palmitoleoyl-ACP (16:1 *cis*  $\Delta^9$ ) or oleoyl-ACP (18:1

*cis*  $\Delta^9$ ) (Shanklin and Cahoon, 1998). Within the plastid, a portion of these products will enter the so-called “prokaryotic pathway” in which the acyl-ACPs are used directly by plastidial acyltransferases to produce glycerolipids for the plastidial membranes (Ohlrogge and Browse, 1995). Alternatively, the nascent fatty acyl chains can be hydrolyzed from ACP by the catalytic action of acyl-ACP thioesterases. The resulting free fatty acids are exported outside the plastid and activated to acyl-CoAs by an acyl-CoA synthetase, which is located on the outer envelope of the plastid (Schnurr et al. 2002). The acyl-CoAs may be further modified and used for “eukaryotic” lipid biosynthesis in the endoplasmic reticulum (ER) (Fig. 2.2) (Ohlrogge and Browse, 1995).



**Figure 2.1. Simplified scheme of *de novo* fatty acid synthesis in plant plastids**

*De novo* fatty acid synthesis occurs in the plastids. The first committed step is the conversion of acetyl-CoA to malonyl-CoA, which is catalyzed by ACCase. Malonyl-CoA is used as a two-carbon donor in a series of repetitive reactions for the following elongation of the acyl chain. The FAS complex catalyzes the elongation cycle. Two-carbon unit (from malonyl-CoA) is first added to the growing acyl chain catalyzed by KAS. The immediate product of KAS, 3-ketoacyl-ACP, is converted to an acyl-ACP through additional three steps, reduction, dehydration, and reduction, which are catalyzed by KAR, HAD, and EAR, respectively (Li-Beisson et al., 2010). Abbreviations: ACCase, acetyl-CoA carboxylase; ACP, acyl carrier protein; EAR, enoyl-ACP reductase; HAD, hydroxyacyl-ACP dehydrase; KAR, ketoacyl-ACP reductase; KAS, ketoacyl-ACP synthase; CoA, coenzyme A.



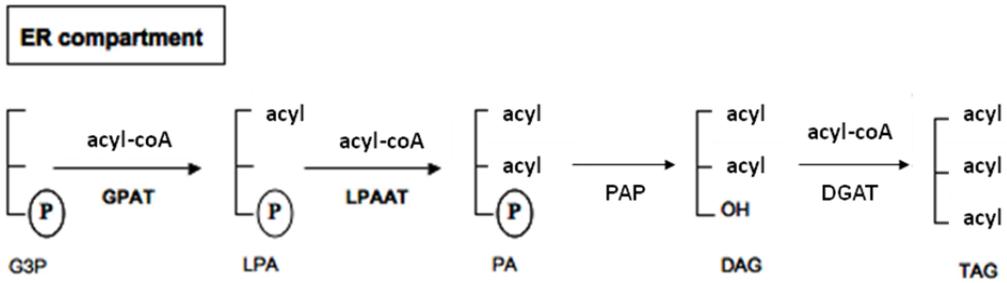
**Figure 2.2. Fatty acid synthesis termination, desaturation, and export from plastid.**

Acyl chain elongation is terminated once it reaches 16-18 carbons. Plastidial desaturase catalyzes the insertion of a double bond into 16:0 or 18:0-ACP. To be exported from the plastid, the acyl group is removed from ACP by the catalytic action of acyl-ACP thioesterase. The free fatty acid is then activated to acyl-CoA through the catalytic action of acyl-CoA synthetase. The nascent acyl-CoA is released to the cytosol for further modification and lipid biosynthesis in the ER. Re-drawn based on Li-Beisson et al., 2010.

Abbreviations: ACP, acyl carrier protein; ER, endoplasmic reticulum; FAS, fatty acid synthase; KAS, ketoacyl-ACP synthase.

### 2.1.2 Triacylglycerol assembly

Storage lipid assembly occurs in the ER of developing seeds. The classic Kennedy pathway (Fig. 2.3), first proposed by Eugene Kennedy and his colleagues in the 1950s, described TAG biosynthesis as a direct conversion of *sn*-glycerol-3-phosphate (G3P) to TAG by sequential acylations on the *sn*-1, 2, and 3 positions of the glycerol backbone (Ohlrogge and Browse, 1995; Weselake 2005). The first enzyme in this pathway, *sn*-glycerol-3-phosphate acyltransferase (GPAT), catalyzes the esterification of a fatty acyl moiety (in the form of acyl-CoA) to the *sn*-1 position of G3P to form lysophosphatidic acid (LPA). Lysophosphatidic acid acyltransferase (LPAAT) catalyzes the esterification of an acyl group from acyl-CoA to the *sn*-2 position of LPA to form phosphatidic acid (PA). PA is dephosphorylated by the catalytic action of phosphatidate phosphatase (PAP) to form *sn*-1, 2-diacylglycerol (DAG) which in turn can be acylated at the *sn*-3 position to form TAG via the catalytic action of diacylglycerol acyltransferase (DGAT). DAG may also serve as substrate in the synthesis of membrane glycerolipids.



**Figure 2.3. Sequential acyl-CoA-dependent acylation of the glycerol backbone to form triacylglycerol (TAG) in the ER.**

Kennedy pathway describes the TAG synthesis as sequential esterifications of acyl-CoA to the *sn*-1, 2, and 3 positions of the glycerol backbone (Kennedy et al., 1961).

Abbreviations: DAG: diacylglycerol; DGAT, diacylglycerol acyltransferase; G3P, glycerol-3-phosphate; GPAT, glycerol-3-phosphate acyltransferase; LPA, lysophosphatidic acid; LPAAT, lysophosphatidic acid acyltransferase; PAP, phosphatidate phosphatase; TAG, triacylglycerol.

In addition to the classic Kennedy pathway, a number of enzymes catalyzing alternative pathways for TAG synthesis have been identified and demonstrated to be more dominant in certain plant species.

For example, a recent study in the embryos of soybean (*Glycine max*) (Bates et al., 2009) demonstrated that about 60% of the newly synthesized fatty acids were first acylated to lysophosphatidylcholine (LPC) and underwent a subsequent acyl remodeling process (such as desaturation) at the *sn*-2 position of phosphatidylcholine (PC). A small portion of the remodeled acyl groups was released from PC back to the acyl-CoA pool for direct acylations on the glycerol backbone. The majority of PC, which carried the remodeled acyl groups, was converted to DAG for further TAG synthesis. These acyl remodeling and PC-DAG conversion processes involve enzymes that catalyze incorporation of fatty acyl groups into LPC, release of fatty acyl groups from PC, and conversion of PC to DAG. Lysophosphatidylcholine acyltransferase (LPCAT), which has been reported to have a high activity in developing seeds, catalyzes the reversible acyl acylation of LPC at the *sn*-2 position to form PC (Lands and Hart, 1965; Rochester and Bishop, 1984; Stymne and Stobart, 1984; Yurchenko et al., 2009). In addition, phospholipase A<sub>2</sub> (PLA<sub>2</sub>) catalyzes the hydrolysis of a fatty acyl group from the *sn*-2 position of PC (Das et al. 2001). Free fatty acids produced through the catalytic action of PLA<sub>2</sub> could in turn be used to form acyl-CoA via the catalytic action of acyl-CoA synthetase. For the conversion of PC to DAG, a PC:DAG cholinephosphotransferase (PDCT) catalyzes the exchange of the phosphocholine headgroup between PC and DAG. The absence of PDCT activity

in *Arabidopsis* mutant lines limited the conversion of polyunsaturated fatty acids (PUFA)-enriched PC to DAG, causing a marked decrease of 18:2 (*cis*  $\Delta^{9,12}$ ) and  $\alpha$ -18:3 (*cis*  $\Delta^{9,12,15}$ ) in TAG and an increased level of 18:3 in PC (Lu et al., 2009).

Furthermore, since the desaturation of acyl groups occurs primarily when they are esterified on the PC (Sperling et al., 1993; Sperling and Heinz, 1993), the PC-mediated alternative pathway for TAG biosynthesis has been proposed to be more dominant for the biosynthesis of high PUFA-containing TAG in developing seeds (Griffiths et al., 1988a; Bates et al., 2009). Previous studies with high PUFA-containing seeds, including flax (*Linum usitatissimum*), soybean (Slack et al., 1978; Bates et al., 2009), safflower (*Carthamus tinctorius*) (Griffiths et al., 1988b) and *Arabidopsis* (Lu et al., 2009) suggested that the DAG pool used for TAG synthesis in developing seeds is derived from PC. In contrast, a study of developing endosperm of oil palm kernel found that PC content was constantly at an extremely low level in comparison to DAG during the active stage of TAG formation (Wiberg and Bafor, 1995). This result suggests that the DAG-PC interconversion model does not apply to TAG synthesis in oil palm kernel.

A few other enzymes have also been reported to be involved in TAG synthesis. For example, a DAG: DAG transacylase (DGTA) was shown to catalyze two DAG molecules to produce TAG and monoacylglycerol (Stobart et al., 1997). A phospholipid: diacylglycerol acyltransferase (PDAT) catalyzes the transfer of an acyl group from the *sn*-2 position of PC to DAG, producing TAG and lysophosphatidylcholine (LPC) (Dahlqvist et al., 2000). An N-terminus-deleted yeast PDAT, de-anchored from the ER, has been shown to also exhibit

DGTA activity (Ghosal et al., 2007). A CDP-choline: diacylglycerol cholinephosphotransferase (CPT) catalyzes *de novo* synthesis of PC from DAG (Vogel and Browse 1996). A more comprehensive network of TAG synthesis pathways is illustrated in Fig. 2.4.

## **2.2 Seed oil modification in *Brassica napus***

The quality and utility of seed oil, determined mainly by its fatty acid composition and oil content, has been a major consideration in *Brassica* breeding projects worldwide. In the past two decades, both conventional and genetic engineering approaches have been explored to introduce new oil traits into *B. napus* (Scarath and Tang, 2006). Conventional breeding strategies have been successful in bringing new traits from natural germplasms and artificially induced mutants into adapted cultivars using traditional crossing. One of the major successes of conventional breeding is the development of low erucic acid (22:1 *cis*  $\Delta^{13}$ , less than 2%) *B. napus*, which was the product of backcrossing a low 22:1 *B. napus* mutant with adapted cultivars, which usually have 45% 22:1 in the seed oil (Stefansson and Downey, 1995). In addition to developing cultivars with low 22:1 content, oilseed researchers have also developed cultivars with high oleic acid (18:1 *cis*  $\Delta^9$ ), high erucic acid or low  $\alpha$ -linolenic acid (18:3 *cis*  $\Delta^{9,12,15}$ ) by using the conventional breeding approaches (Vilkki and Tanhuanpää, 1995; Scarth et al., 1995a; Scarth et al., 1995b; Wang et al., 2003).

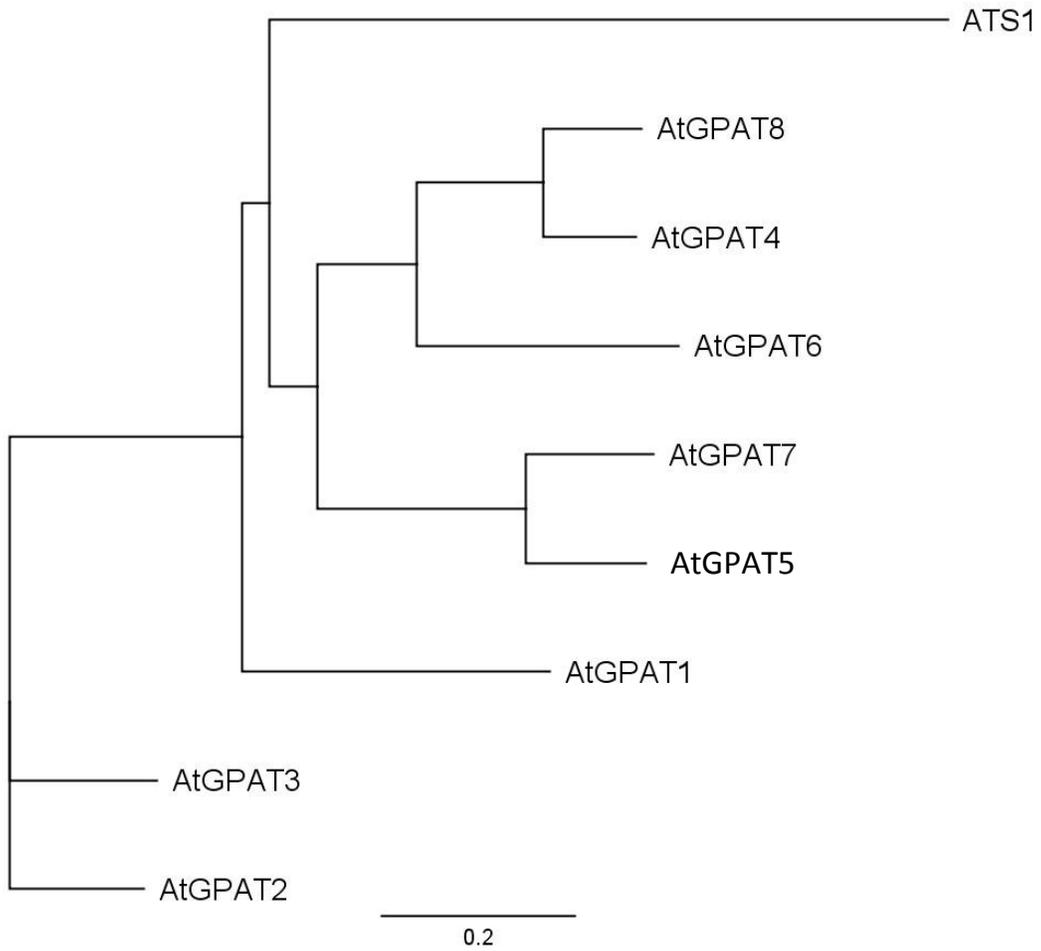


Although conventional breeding has proven useful for manipulating fatty acid composition in *Brassica* species, most *Brassica* oilseeds produce a limited range of fatty acids, namely palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2 *cis*  $\Delta^{9,12}$ ),  $\alpha$ -linolenic (18:3) and erucic acid (22:1) (Ackman, 1990). Other plants, however, exhibit considerable diversity in their fatty acid composition, producing more than 200 unusual fatty acids (van de Loo et al., 1993; Thelen and Ohlrogge, 2002; Jaworski and Cahoon, 2003; Scarth and Tang, 2006). Many of these fatty acids have great nutritional or industrial value, but their wide spread cultivation has been problematic because many of these plants do not possess suitable agronomic traits and are not amenable to domestication. In these cases, genetic engineering may be a powerful tool for introducing genes from plants producing novel fatty acids into highly productive oilseed crops such as *B. napus* to produce cultivars containing seed oil with modified fatty acid composition. For example, the development of high-lauric acid (12:0) *B. napus* was a pioneering achievement in genetic engineering of oilseeds. Lauric acid, commonly found in tropical oils, is used as a detergent in industry. The introduction of genes encoding a fatty acid thioesterase B (FATB) from California bay tree and a lysophosphatidic acid acyltransferase (LPAAT) from coconut resulted in accumulation of up to 67 mol% 12:0 in the seed oil of transgenic *B. napus* (Voelker et al., 1996; Knutzon et al., 1999). Increased dormancy of high laurate *B. napus* seeds under certain conditions (Linder, 1998), however, suggested a potential setback for developing Brassica oilseed species accumulating medium fatty acids. In many other cases, genetic engineering strategies have resulted in

limited accumulation of the desired fatty acid, well below the threshold required for commercialization (Scarath and Tang, 2006; Cahoon et al., 2007; Hildebrand, 2008). There are several possible reasons for this. Although the key steps in the biosynthetic pathways responsible for seed oil accumulation are largely defined, much less is understood about how they are regulated differently between species and how this influences the distinctive fatty acid profiles in various species. Progress in understanding and manipulating fatty acid composition and storage lipid accumulation in oilseeds will require detailed studies of the various lipid synthesis pathways and the associated enzymes.

### **2.3 The *sn*-glycerol-3-phosphate acyltransferase family in plants**

GPAT is an important enzyme in the glycerolipid synthetic pathway. It catalyzes the esterification of an acyl chain to the *sn*-1 position of G3P, and thus initiates the pathway of glycerolipid biosynthesis. In higher plants, there are three confirmed types of GPATs, compartmentalized in the plastid stroma, ER membrane and mitochondrial membrane (Zheng et al., 2003). In *Arabidopsis*, nine *GPAT* genes have been identified. One encodes a soluble GPAT in the plastid (annotation: *ATSI*), three encode membrane-bound GPATs in the mitochondria (annotation: *AtGPAT1-3*) and five encode membrane-bound GPATs in the ER (annotation: *AtGPAT4-8*) (Fig. 2.5) (Zheng et al., 2003; Xu et al., 2006). Based on the subcellular localizations, GPAT family members are involved in distinct or overlapping metabolic pathways.



**Figure 2.5. Phylogenetic tree of the *sn*-glycerol-3-phosphate acyltransferase (GPAT) family of *Arabidopsis*.**

ATS1, a soluble plastidial GPAT. AtGPAT1-3, the mitochondrial membrane-bound GPATs. AtGPAT4-8, the ER membrane-bound GPATs.

### 2.3.1 Plastidial *sn*-glycerol-3-phosphate acyltransferase

Glycerolipid biosynthesis within the plastid is usually referred as the “prokaryotic pathway” to distinguish it from the glycerolipid biosynthesis occurring in the ER, which is referred to as “eukaryotic pathway” (Ohlrogge and Browse, 1995). The prokaryotic pathway, which provides part of the glycerolipids for plastidial membranes, is mediated by plastid-localized enzymes, such as plastidial GPAT (Ohlrogge and Browse, 1995). Plastidial GPATs have been purified and cloned from several plant species, including *Arabidopsis* (Nishida et al., 1993), cucumber (*Cucumis sativus*) (Johnson et al., 1993), squash (*Cucurbita moschata*) (Nishida et al., 1987; Ishizaki et al., 1988), pea (*Pisum sativum*) (Bertrams and Heinz, 1981; Weber et al., 1991) and spinach (*Spinacia oleracea*) (Bertrams and Heinz, 1981; Ishizaki-Nishizawa et al., 1995). Most of the earlier studies of plastidial GPATs were focused on protein purification and enzymatic characterization of the acyl substrate specificity of these enzymes. Results from these studies revealed a close relationship between the substrate preference of plastidial GPAT for saturated vs unsaturated acyl-ACPs and the chilling-sensitivity of the plants.

It has been well-established that the level of unsaturation of membrane glycerolipids influences the fluidity of plant membranes, which is an important factor in enabling plants to survive at a low temperature (Nishida and Murata, 1996). As the major membrane phospholipid in chloroplast, phosphatidylglycerol (PG) plays an important role in controlling chloroplast membrane fluidity and hence affecting the chilling-sensitivity of the aerial tissues of plants (Nishida and

Murata, 1996). The fatty acid composition at the *sn*-1 position of the chloroplastidial PG has been found to be mainly determined by the acyl substrate preference of plastidial GPAT (Murata and Tasak, 1997). For example, plastidial GPATs isolated from chilling-tolerant plants such as *Arabidopsis*, spinach or pea, exhibited a substrate preference for 18:1-ACP, whereas GPAT isolated from chilling sensitive squash displayed similar activity with 16:0-ACP or 18:1-ACP (Murata and Tasak, 1997). A study of transgenic tobacco has shown that transformation with a cDNA encoding the squash (chilling-sensitive) plastidial GPAT results in decreased chilling tolerance, whereas transformation with cDNA encoding *Arabidopsis* (cold-tolerant) plastidial GPAT results in increased chilling tolerance (Murata et al. 1992).

Though the plastidial GPATs from different plant species have been biochemically characterized, the physiological functions of this enzyme were less understood until recently. The *Arabidopsis* plastidial GPAT (*ATS1*) deficient mutant lines (chemically mutagenized) were first isolated and characterized by Kunst et al (1988). In this study, the mutant *Arabidopsis* plants exhibited slightly reduced PG content (10–25%) in leaves. Xu et al (2006) further characterized the molecular and physiological defect of several *Arabidopsis ats1* T-DNA lines. Different from the earlier observation by Kunst et al (1988), here the authors discovered that the deficiency of *ATS1* did not affect PG synthesis in the plants. Nevertheless, when another plastidial acyltransferase (LPAAT, encoded by *ATS2*) catalyzing the acylation of the *sn*-2 position of PG was defective in the *ats1* mutant, the *Arabidopsis* double mutant line exhibited a delay in PG biosynthesis

and severe reduction of growth. ATS1 was thus confirmed to be involved in the coordinated regulation of plastidial PG synthesis, which is essential for plant development (Xu et al., 2006). Interestingly, an early study of the *Arabidopsis* *ats2* T-DNA lines demonstrated that the loss of the plastidial LPAAT also caused embryo-lethality (Yu et al., 2004). Nevertheless, the authors did not include experimental data related to the PG content in the *ats2* T-DNA lines.

### **2.3.2 Mitochondrial *sn*-glycerol-3-phosphate acyltransferase**

In *Arabidopsis*, three mitochondrial GPATs (AtGPAT1-3) together with five ER-bound GPATs (AtGPAT4-8) were first identified based on the amino acid sequence similarity to known GPAT enzymes from bacteria, yeast and mammals (Zheng et al., 2003). Among the three mitochondrial GPATs, only AtGPAT1 has been characterized in detail. By studying *gpat1* T-DNA *Arabidopsis* lines, defective tapetal differentiation was found to be associated with the deficiency of AtGPAT1 activity, which caused most of the pollen microspores to abort prior to maturation (Zheng et al., 2003). Therefore, AtGPAT1 was confirmed to play a pivotal role in pollen development and male fertility of *Arabidopsis* (Zheng et al., 2003).

### **2.3.3 ER-bound *sn*-glycerol-3-phosphate acyltransferase**

In the classical Kennedy pathway, ER-bound GPAT is considered to catalyze the first step in storage lipid synthesis (Fig. 2.3). In most mammalian cells, GPAT has been proved to be essential for TAG biosynthesis (Gimeno and Cao, 2008). In plant, the role of GPAT in TAG biosynthesis, however, has not been clarified to date (Snyder et al., 2009). The immediate product of the GPAT

catalyzed reaction, LPA, can be used for the synthesis of several products, including phospholipids, glycolipids, TAG and lipid polyesters (Ohlrogge and Browse, 1995; Pollard et al., 2008). Recent studies of T-DNA *Arabidopsis* mutant lines revealed that four ER AtGPAT members (i.e., AtGPAT4, 5, 6 and 8) are involved in the synthesis of extracellular lipid polyesters (i.e., cutin and suberin) by providing acyl-glycerol structure to the polyester matrices (Li-Beisson et al., 2010).

GPAT5 was demonstrated to be essential for suberin synthesis in roots and seed coats (Beisson et al., 2007). By analyzing the polyester monomer profiles in the seeds and roots of the *gpat5* T-DNA *Arabidopsis* mutant lines, the authors observed strong reductions in 22:0 and 24:0 fatty acids and their derivatives, and therefore proposed that the physiological role of GPAT5 is to provide acyl-glycerols containing 22-24 carbon groups to the suberin synthetic pathway. This substrate preference for C22 or C24 acyl chains of GPAT5 was further supported by the fact that *Arabidopsis* over-expressing *GPAT5* had increased accumulation of very long chain saturated fatty acids in suberin.

GPAT4 and GPAT8 exhibited functional redundancy in cutin synthesis in leaves and stems (Li et al., 2007). Although the single T-DNA mutant lines of *gpat4* or *gpat8* did not exhibit any obvious cuticle defect, the *gpat4 gpat8* double T-DNA lines of *Arabidopsis* exhibited a strong decrease in cutin content in the leaves and stems. Among all the cutin monomers,  $\alpha$ ,  $\omega$ -18:2 dicarboxylic acid exhibited the most significant decrease in these *gpat4 gpat8* lines. When *GPAT4* or *GPAT8* was overexpressed in *Arabidopsis*, however, the cutin monomers

exhibiting the most significant increases were,  $\alpha$ ,  $\omega$ -16:0 and 18:0 dicarboxylic acids. It is obvious that GPAT4 and GPAT8 are involved in the incorporation of C16 and C18 acyl chains into cutin. Modifications of these C16 and C18 monomers (i.e., desaturation, hydroxylation and carboxylation) have been confirmed to be regulated by other enzymes, such as desaturases and oxidases (Xiao et al., 2004; Li et al., 2007). Since the sequential order of these reactions is not clear, it is difficult to predict the natural substrates of GPAT4 and GPAT8.

GPAT6 is involved in cutin synthesis in flower petals (Li-Beisson et al., 2009). In the *gpat6* T-DNA mutant *Arabidopsis* lines, a strong reduction in cutin content was detected in the flowers. Further analysis revealed that the lack of cutin formation resulted in a lack of nanoridges on the petal surfaces. By analyzing the cutin monomer profiles of the *gpat6* T-DNA and *GPAT6* overexpression lines, the authors proposed that GPAT6 is involved in using 16:0 and its derivatives for petal cutin synthesis (Li-Beisson et al., 2009).

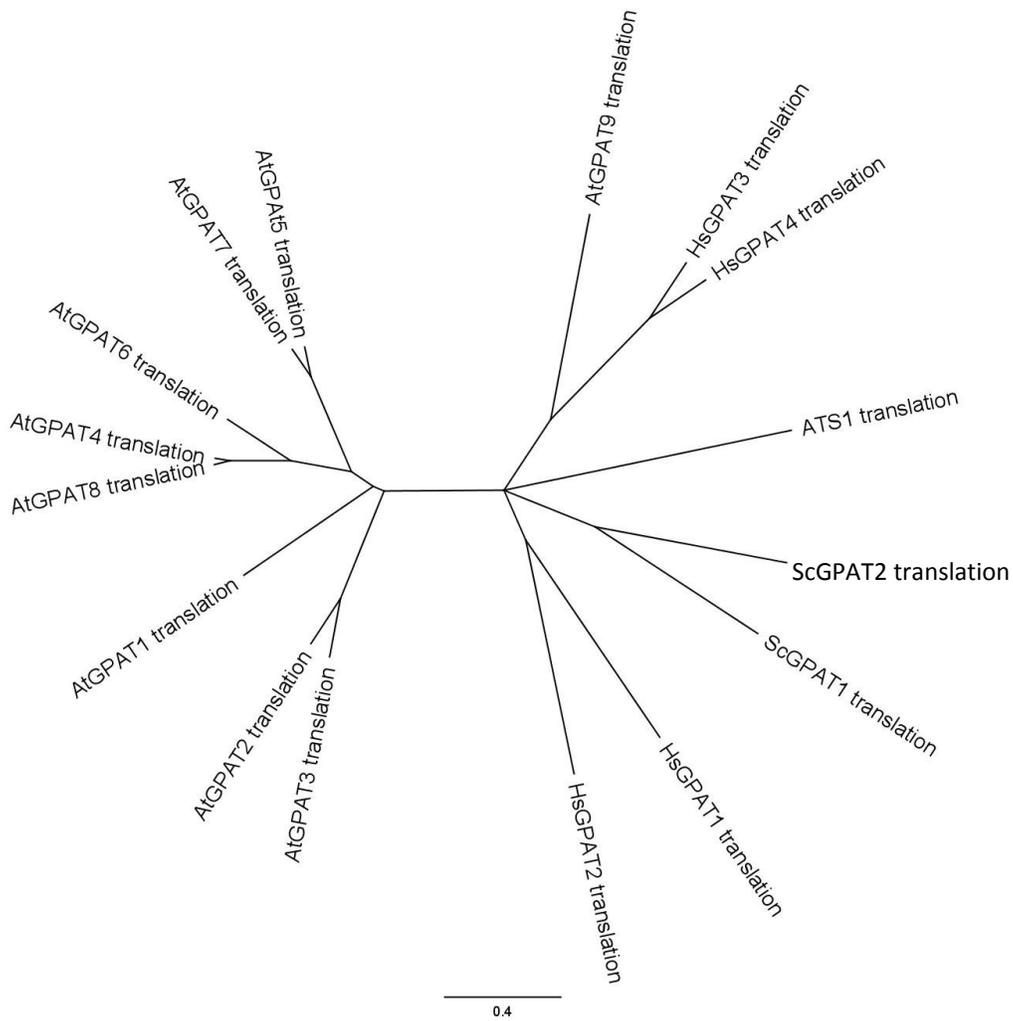
#### **2.3.4 A putative *Arabidopsis* sn-glycerol-3-phosphate acyltransferase 9 may play an important role in eukaryotic lipid metabolism**

It appears that the currently identified ER-bound GPAT family is important for extracellular lipid polyester synthesis. It remains unclear, however, whether any plant ER-bound GPAT(s) plays a key role in the eukaryotic membrane lipid and TAG biosynthetic pathways. Notably in mammalian cells, four membrane-bound GPATs, including two localized in the mitochondria (GPAT1-2) and two localized in the ER (GPAT3-4), have been confirmed to have important roles in storage lipid biosynthesis (Gimeno and Cao, 2008). The

phylogenetic analysis of the polypeptide sequences of the mammalian GPATs and *Arabidopsis* GPATs revealed that all the known membrane-bound AtGPATs have evolved quite distantly from mammalian GPAT1-4 (Fig. 2.6). Recently, a putative GPAT9 was identified in *Arabidopsis* by a bioinformatics approach, and exhibited a much closer evolutionary relationship with the mammalian GPATs (Gidda et al., 2009) (Fig. 2.6). Although the GPAT activity of this enzyme form has not been directly confirmed and its physiological function is unknown, polypeptide sequence alignment, phylogenetic analysis, conserved domain analysis and gene expression data suggested this AtGPAT9 may be a functional enzyme, playing an essential role in plant membrane lipid and TAG synthesis (Gidda et al., 2009).

#### **2.4 Plant lipid polymer matrices- cutin and suberin**

Cutin and suberin are quite distinct from storage or membrane lipids in the way that they are composed of polymerized acyl derivatives and glycerol groups. Cutin and suberin are hydrophobic layers that are closely associated with plant cell walls. The main functions of these impermeable layers include preventing water loss, protection against pathogen invasion and provision of mechanical support to plant organs (Pollard et al., 2008). In addition, cutin is also important in preventing cell wall fusions between the adjacent organs during early development (Nawrath, 2002).

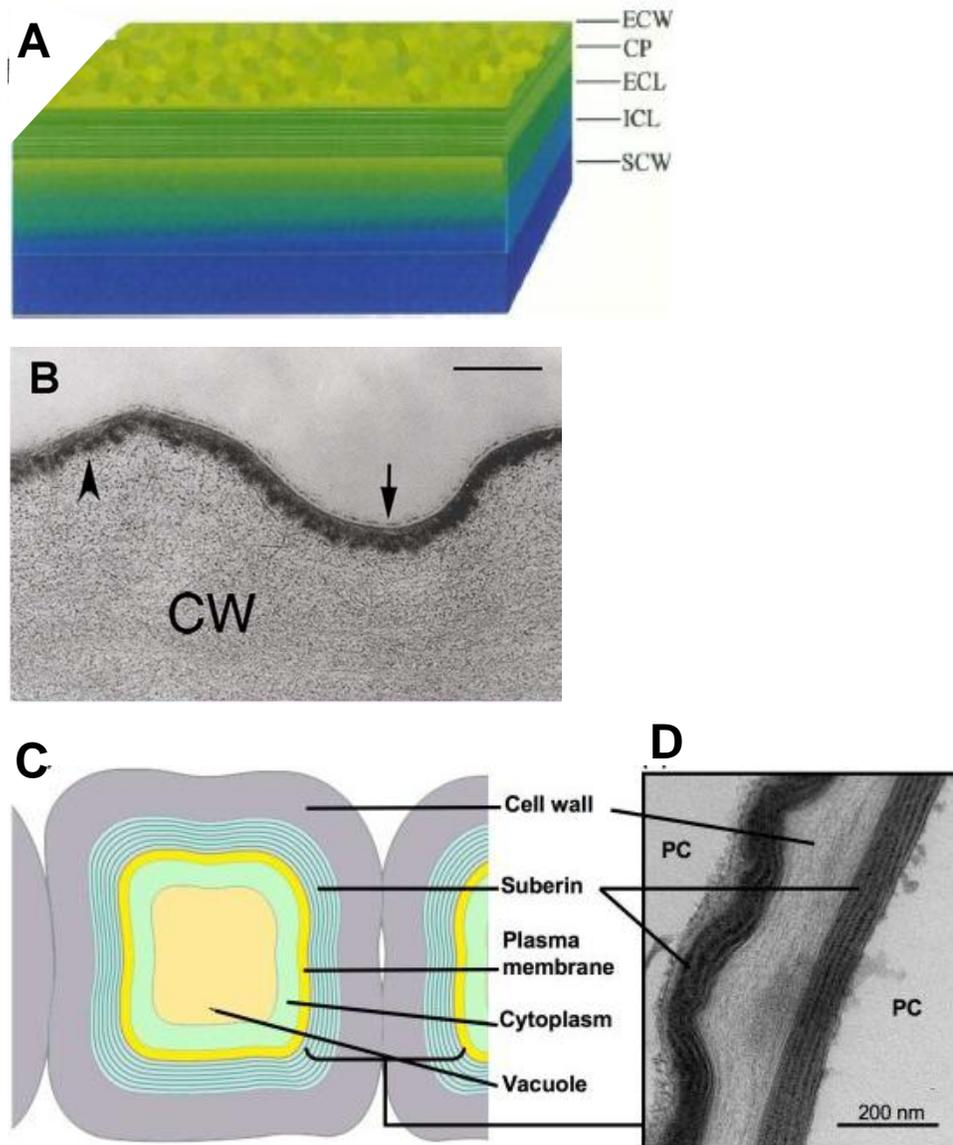


**Figure 2.6. Phylogenetic analysis of the *sn*-glycerol-3-phosphate acyltransferases (GPATs) from *Arabidopsis* (At), *Homo sapiens* (Hs) and *Saccharomyces cerevisiae* (Sc).**

Different from the other AtGPATs, the putative AtGPAT9 exhibits closer evolutionary relationship with the human ER-bound GPATs (HsGPAT3-4). AtGPAT1-3 and HsGPAT1-2 are mitochondrial membrane-bound proteins. AtGPAT4-9, HsGPAT3-4 and ScGPAT1-2 are ER membrane-bound proteins. ATS1 is a plastidial soluble protein.

#### **2.4.1 Tissue distributions of cutin and suberin**

Cutin and suberin are localized in different parts of the plant. Cutin is a key structural component of the cuticle, which is a heterogeneous layer (mainly containing cutin, wax, and polysaccharides) covering the epidermis of all aerial organs (Fig.2.7 A-B). The cuticle can have several different layers depending on the developmental stage of the plant tissue (Bird and Gray, 2003). Among the different cuticle layers, epicuticular wax is the outermost layer. Cuticle proper, which is the next layer beneath the epicuticular wax, is a mixture composed of cutin and polysaccharides. Between the cuticle proper and primary cell wall are two thick cutin layers (i.e., external and internal cuticular layers) mixed with wax and polysaccharide (Bird and Gray, 2003). Different from cuticle, which lies on the top surface of the primary cell wall, suberin is often deposited in the inner face of primary cell walls in both external and internal tissues (Höfer et al., 2008). Under transmission electron microscopy, suberin appears as multiple thickening bands within the cell wall (Fig. 2.7 C-D) (Franke et al., 2005). Suberin can be found in the external tissues of plants such as the periderms of cork and primary roots and seed coats. Additionally, suberin is also deposited in the internal tissues including root endodermis and the bundle sheaths of monocots (Höfer et al., 2008; Pollard et al., 2008).



**Figure 2.7. Structure of plant cuticle and suberized root cell wall.**

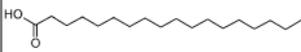
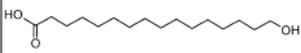
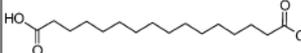
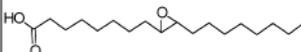
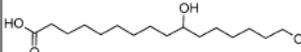
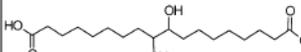
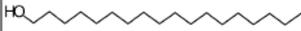
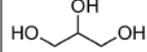
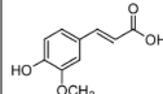
(A) Schematic depicting different layers of cuticle. ECW, epicuticular waxes; CP, cuticle proper; ECL, external cuticular layer; ICL, internal cuticular layer; PCW, primary cell wall; SCW, secondary cell wall. Source: Bird and Gray, 2003. Copy right (2010) with permission from John Wiley and Sons. (B) Cuticle ultrastructure of the leaf epidermis of *Arabidopsis*. The cuticle layer is indicated between the two arrows. CW, cell wall. Scale bar, 200 nm. Source: Franke et al., 2005. Copy right (2010) with permission from Elsevier. (C) Schematic depicting the root cell. (D)

Ultrastructure of the root cell wall. Source:  
<http://lipidlibrary.aocs.org/plantbio/polyesters/index.htm>; with permission.

#### **2.4.2 Monomer compositions and chemical structures of cutin and suberin**

Cutin and suberin are composed of polymerized acyl derivatives and glycerol groups, which make them insoluble in either water or organic solvents (Bonaventure et al., 2004). Because of these chemical properties, isolating and quantitatively analyzing these lipid polyesters from plants was technically challenging until the recent development of routine analytical methods for cutin and suberin (Bonaventure et al., 2004; Franke et al., 2005). These recent analytical developments have prompted studies of lipid polyesters in *Arabidopsis* and revealed the detailed monomer profiles and the regulatory roles of several key enzymes (such as GPAT) in polyester biosynthetic pathways.

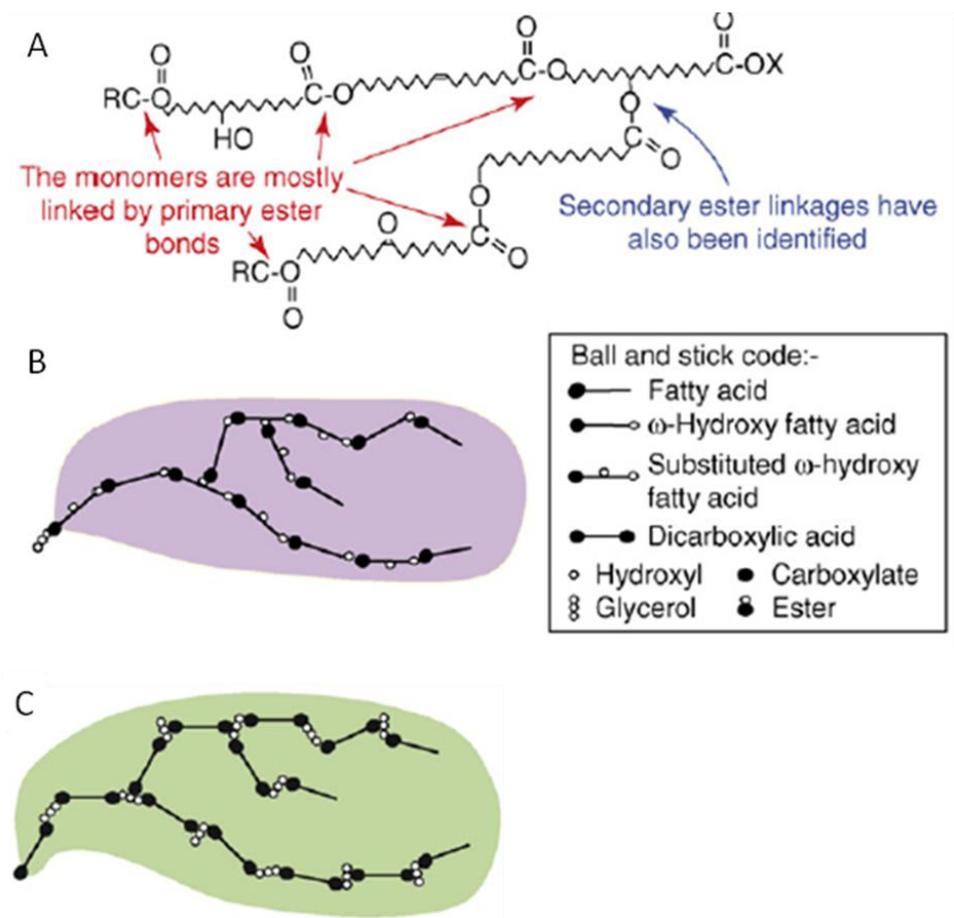
The dominant monomers in cutin are C16 and C18  $\omega$ -hydroxy fatty acids. In addition, glycerol and other types of fatty acids and their derivatives can also be found in cutin monomers. Notably, a great percentage (>50%) of  $\alpha$ ,  $\omega$ -C16 and C18 dicarboxyl fatty acids also exist in the cutin polyesters of *Arabidopsis* and *B. napus* (li et al., 2007). Suberin shares some common types of monomers with cutin, such as fatty acids,  $\omega$ -hydroxy fatty acids and glycerol, yet suberin also contains much higher proportions of aromatic hydroxycinnamic acid (predominantly ferulate) and longer chain length (>C20) of fatty alcohols / acids and their derivatives (Fig. 2.9) (Beisson et al., 2007).

(a) Monomer type	Abundance (%) and common monomers	
Example	Cutin	Suberin
<b>Unsubstituted fatty acids</b> 	<b>1-25%</b> C16:0, C18:0, C18:1, C18:2	<b>1-10%</b> C18:0 to C24:0
<b><math>\omega</math>-Hydroxy fatty acids</b> 	<b>1-32%</b> C16:0, C18:1, C18:2	<b>11-43%</b> C18:1, C16:0 to C26:0
<b><math>\alpha,\omega</math>-Dicarboxylic acids</b> 	<b>Usually &lt;5% but &gt;50% in Arabidopsis</b> C16:0, C18:0, C18:1, C18:2	<b>24-45%</b> C18:1, C18:2, C16:0 to C26:0
<b>Mid-chain functionalized monomers</b> <b>Epoxy-fatty acids</b> 	<b>0-34%</b> C18:0 (9, 10-epoxy)  C18:1 (9, 10-epoxy)	<b>0-30%</b> C18:1 (9, 10- epoxy-18- hydroxy) C18:0 (9, 10-epoxy-1, 18-diacid)
<b>Polyhydroxy-fatty acids</b> 	<b>16-92%</b> C16:0 (10, 16- dihydroxy) C18:0 (9,10, 18- trihydroxy)	<b>0-2%</b> C18:0 (9,10, 18-trihydroxy)
<b>Polyhydroxy <math>\alpha,\omega</math>-dicarboxylic acids</b> 	<b>Traces</b>	<b>0-8%</b> C18:0 (9,10- dihydroxy)
<b>Fatty alcohols</b> <b>Alkan-1-ols and alken-1-ols</b> 	<b>0-8%</b> C16:0, C18:1	<b>1-6%</b> C18:0 to C22:0
<b><math>\alpha,\omega</math>-Alkanediols and <math>\alpha,\omega</math>-alkenediols</b> 	<b>0-5%</b> C18:1	<b>0-3%</b> C22:0
<b>Glycerol</b> 	<b>1-14%</b>	<b>14-26%</b>
<b>Phenolics</b> 	<b>0-1%</b> Ferulate	<b>0-10%</b> Ferulate, smaller amounts of coumarate, sinapate, caffeate

**Figure 2.9. Structures of common cutin and suberin monomers, and typical ranges of composition values in plants.**

Source: Pollard et al., 2008. Copy right (2010) with permission from Elsevier.

Monomers in cutin and suberin are connected with each other through their hydroxyl and carboxyl groups via ester linkages. Analysis of isolated monomers and oligomers after depolymerization has indicated that the majority of ester linkages in linear polyesters are derived from the primary hydroxyl groups of  $\omega$ -hydroxy fatty acids, along with a few secondary hydroxyl groups in the linear chain to provide a branch point for cross-links or branching (Fig. 2.10 A-B) (Pollard et al., 2008). Nevertheless, the three-dimensional structures of these polyester matrices have not been resolved in detail (Kolattukudy 2001; Pollard et al., 2008). Additionally, little is known about the involvement of glycerol in constructing the polymeric structures of cutin or suberin. It is known that fatty acids can be esterified to glycerols at the three hydroxyl groups (representing *sn*-1, 2, 3 positions). Thus, it has been proposed that glycerol is likely to provide connections for building larger branched structures in cutin and suberin (Fig. 2.10 C) (Pollard et al., 2008). It is obvious that the combination of different monomers would give quite distinct structures of polyester domains (Fig. 2.10 B vs C). Thus, it remains to be determined if the polyester matrices have different types of domains or more homogeneous structures.



**Figure 2.10. Structures of hypothetical monomer connectivity patterns.**

(A) A small segment of hydroxy fatty acid-rich polyester is illustrated to show the dominant primary ester linkages that form chains, along with a secondary ester linkage to enable a branch point. (B) Organization of fatty acid and  $\omega$ -hydroxy fatty acid monomers to produce a dendrimer. (C) Organization of dicarboxylic acid and glycerol monomers to produce a dendrimer structure. Source: Pollard et al., 2008. Copy right (2010) with permission from Elsevier.

### **2.4.3 Identification of genes for the biosynthesis of cutin and suberin in *Arabidopsis***

Along with the development of routine and reliable analytical methods for lipid polyester analysis, a number of phenotypes have been confirmed to be related with lipid polyester defects in *Arabidopsis*. Phenotypes including fusion of cell walls from adjacent organs, increased permeability of cuticles to toluidine blue and altered susceptibility to pathogen infection are often caused by reduced lipid polyester deposition in plants (Nawrath, 2006; Pollard et al., 2008). Using forward and reverse-genetics approaches in combination with such phenotype analysis, several genes encoding enzymes related to lipid polyester synthesis have been identified (Bonaventure et al., 2004; Schnurr et al., 2004; Xiao et al., 2004; Franke et al., 2005; Molina et al., 2006; Beisson et al., 2007; Li et al., 2007).

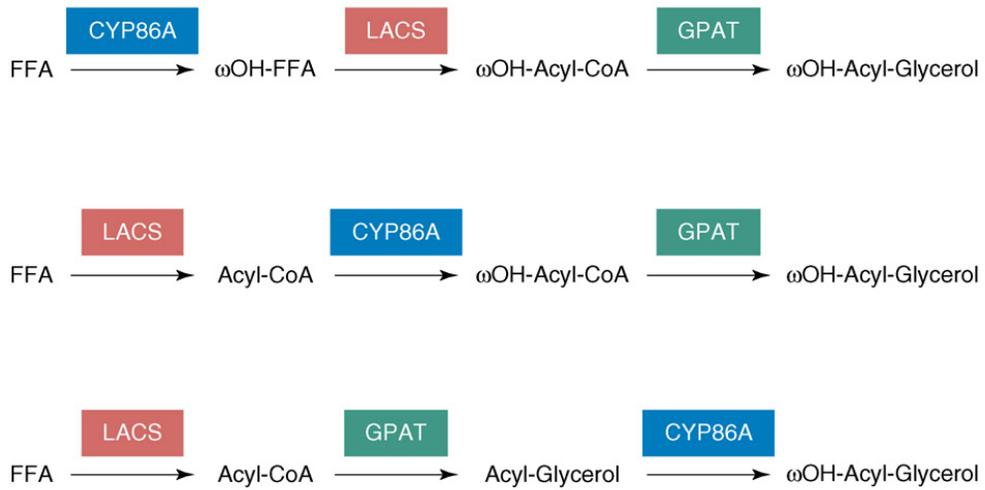
The *CYP86* and *CYP77* gene families, which encode cytochrome P450-dependent fatty acid hydroxylases, were found to play important roles in cutin and suberin biosynthesis in the formation of hydroxy fatty acids (Xiao et al., 2004; Beisson et al., 2007; Li et al., 2007; Compagnon et al., 2009; Li-Beisson et al., 2009). The *LACS* gene family encodes long-chain acyl-CoA synthetases, which convert the long-chain free fatty acids to acyl-CoA. Recent studies identified two enzymes encoded by *LACS1* and *LACS2* that are involved in the activation of long-chain free  $\omega$ -hydroxy fatty acids to the CoA groups for further cutin biosynthesis (Schnurr et al., 2004; Lü et al., 2009). The GPAT family catalyzes the formation of acyl-glycerols (Beisson et al., 2007; Li et al., 2007; Li-Beisson et al., 2009). It has been proposed that the function of GPAT for lipid polyester

biosynthesis is to introduce glycerol into the polyester matrices (Pollard et al., 2007). *RWPI* (also known as *ASFT*) encodes a  $\omega$ -hydroxyacid hydroxycinnamoyltransferase (also known as feruloyl-CoA transferase), which catalyzes the acyl-transfer from feruloyl-CoA to  $\omega$ -hydroxy fatty acids and fatty alcohols. *RWPI* was shown to be important for the deposition of ferulic acid in suberin synthesis (Gou et al., 2009; Molina et al., 2009). *WIN1* (also known as *SHINE1*), which encodes a transcriptional factor, was shown to be important in regulating cutin biosynthesis (Kannangara et al., 2007).

These studies of fatty acid hydroxylase, acyl-CoA synthetase and acyltransferases have provided researchers with insights into the processes of monomer formation and polymerization. Although the biochemical or physiological functions of these enzymes have been partially characterized, the exact substrates of these enzymes are still unclear. Using fatty acid hydroxylase as an example, a hydroxyl group could be added to a free fatty acid or to a fatty acid already being esterified to a glycerol. Thus, it is difficult to determine the sequential order of the reactions catalyzed by these enzymes in the biosynthetic pathways of cutin and suberin. As shown in Fig. 2.11, three possible pathways with the involvement of these enzymes have been proposed (Pollard et al., 2008).

In addition to these recently identified enzymes and their proposed pathways, a number of other key biosynthetic steps and the associated enzymes or proteins are still unknown. For example, the enzymes involved in catalyzing the formation of  $\alpha, \omega$ -dicarboxylic acids or formation of acyl-acyl ester are still not identified. It is also unclear what kinds of precursors are transported from the

cytoplasm to the cell walls. In addition, the processes of transport and assembling of precursors to form the final polyesters are still unknown. The synthetic pathways of lipid polyesters are complex and largely unknown. Based on the current knowledge, however, studies using genetic and biochemical approaches will expedite the progress of identifying enzymes and regulators associated with this pathway.



**Figure 2.11. Possible reactions catalyzed by fatty acid hydroxylase (CYP86A), long-chain acyl-CoA synthetase (LACS) and *sn*-glycerol-3-phosphate acyltransferase (GPAT) in the lipid polyester biosynthesis pathway.**

The formation of  $\omega$ -hydroxy-acyl-glycerol for polyester synthesis could occur through three possible synthetic pathways that are mediated by the same group of enzymes, but with different sequences of action. Abbreviations: FFA, free fatty acid.  $\omega\text{OH}$ ,  $\omega$ -hydroxyl group. Source: Pollard et al., 2008. Copy right (2010) with permission from Elsevier.

## **2.5 Polyploid genetic background of *Brassica* species**

Polyploidy refers to the existence of more than two sets of chromosomes in an organism. Polyploid species can be divided into allopolyploid and autopolyploid; allopolyploids were generated by inter-species crosses, where two or more sets of genetically distinct chromosomes are combined together within the nucleus. Autopolyploids were formed by multiplication of the chromosomes within the original diploid genome (Chen, 2007).

Polyploidy happened extensively among angiosperm species during their evolutionary process (Bennett and Leitch, 1997). It is recognized as a predominant factor in facilitating the evolution of flowering plants, such as forming new species, generating novel phenotypes, and providing genome buffering against lethal mutations (Wendel et al., 2000; Udall and Wendel., 2007). The evolution of allotetraploid *B. napus* involved genome fusion of two distinct but closely related genomes of *B. rapa* and *B. oleracea* (U, 1935). Consequently, such polyploidy has resulted in multiple homologous genes existing in the genome, encoding very similar proteins. Understanding the genetic background of *B. napus* is essential for studying the biological functions of the enzymes encoded by multiple homologous genes in the complex lipid synthetic pathways.

### **2.5.1 Evolution of *Brassica* species and *Arabidopsis***

Like many important crops, such as wheat, cotton, oat and sugar cane, *B. napus* is an allopolyploid (AACC; n=19), which has four sets of chromosomes from two different species, *B. rapa* (AA; n=10) and *B. oleracea* (CC; n=9) (U, 1935). The cultivated *Brassica* species are closely related to *Arabidopsis*, all of

which are members of the *Brassicaceae* tribe within the Brassicaceae family (Cruciferae) (Warwick and Black, 1991). *Arabidopsis* thus shares recent common ancestry with a diverse range of vegetable- and oil-producing *Brassica* crops (Parkin et al., 2005). Over the past few years, numerous studies on comparative mapping between *Arabidopsis* and *Brassica* species have provided many insights into the evolution process and structures of the *Brassica* genomes (Kowalski et al., 1994; Lagercrantz and Lydiate, 1996; Lysak et al., 2005; Parkin et al., 2005; Town et al., 2006; Cheung et al., 2009). The *Brassica* and *Arabidopsis* lineages diverged around 20 million years ago (MYA) (Yang et al., 1999). Within the *Brassica* species, *B. nigra* was estimated to diverge from *B. rapa* and *B. oleracea* lineages 8 MYA (Lysak et al., 2005), and between *B. rapa* and *B. oleracea* the lineages diverged 3.7 MYA (Inaba and Nishio, 2002). *B. napus* is normally believed to have arisen from a recent genome fusion (hybridization) event of the two diploid progenitors, *B. rapa* (AA) and *B. oleracea* (CC), followed by chromosome doubling (U, 1935), during human cultivation (i.e., <10,000 years ago) (Cheung et al., 2009). *Arabidopsis*, which has one of the smallest angiosperm genomes (157 Mbp) (Arabidopsis Genome Initiative, 2000; Bennett et al., 2003), has been estimated to have undergone multiple large-scale duplications followed by chromosome fusions and genome deletions over the last 300 million years, resulting in much of the *Arabidopsis* genome being represented by duplicated segments (Vision et al., 2000; Simillion et al., 2002; Blanc et al., 2003; Bowers et al., 2003). By comparison, *Brassica* species have much larger genomes that are even more extensively duplicated.

### **2.5.2 Comparative mapping between *Brassica* and *Arabidopsis***

Comparative mapping studies of the microstructure of targeted genome regions of *B. oleracea*, *B. rapa* and *B. nigra* in comparison to *Arabidopsis* revealed predominantly triplicated genome structures in these diploid *Brassica* species (O'Neill and Bancroft, 2000; Rana et al., 2004; Park et al., 2005; Parkin et al., 2005). On the basis of triplicated genome structure, a few single, duplicated and tetraplicated genome segments were also observed in the diploid *Brassica* species, which were possibly caused by gene-level deletions, additional segmental duplication and chromosomal-scale rearrangements occurring soon after the genome triplication (Parkin et al., 2005; Yang et al., 2006). Due to the fundamentally triplicated nature of the diploid *Brassica* genomes, *B. napus*, which formed from the hybridization of two *Brassica* diploid progenitors, *B. rapa* and *B. oleracea*, has on average six copies of the conserved genome regions for each corresponding region in *Arabidopsis*, with a number of exceptions of less or more than six copies due to genome-wide rearrangement after polyploidy (Parkin et al., 2005).

### **2.5.3 The possible fates of duplicated genes after polyploidy**

The existing replicated genes in the *B. napus* genome are the result of both short-term and long-term genome changes after the polyploidy event. In the short-term, two different genomes must rapidly adapt to a common nucleus through a series of biological processes, which facilitate the proper chromosome pairing for cell division and eventually lead to polyploid stabilization. In the long-term, the polyploid genome would undergo a series of changes at both gene and

genome levels to maintain the fitness of the new species. These changes could lead to genes with four evolutionary fates (Wendel 2000; Lukens et al., 2004; Blanc and Wolfe, 2004; Whittle and Krochko, 2009): (1) loss or silencing, which is believed to be especially prevalent shortly following polyploidy events; (2) subfunctionalization, a process in which the individual duplicated genes take on part of the functions of the ancestral gene; (3) neofunctionalization, a process in which redundant genes can develop new adaptive functions through beneficial mutations and positive Darwinian selection; (4) retention of original function.

In general, it has been proposed that loss or silencing will happen to the duplicated genes at a high rate after the genome fusion event.

Subfunctionalization is believed to happen predominantly to the surviving duplicated genes in polyploids (Force et al., 1999; Lukens et al., 2004).

#### **2.5.4 International sequencing and genomics projects for *Brassica* species**

The economic importance of *Brassica* crops has promoted the initiation of several multinational projects within the *Brassica* research community, such as “Associative Expression and Systems Analysis of Complex Traits in Oilseed Rape/canola” and the “Multinational *Brassica* Genome Project” (MBGP). These projects aim to develop various genomic resources for *Brassica* studies including genetic maps and markers, genotype diversity, cDNA libraries, genome sequencing, TILLING (Targeting Induced Local Lesions in Genomes), and transcriptional micro-arrays. The recent accomplishment of the *Brassica* A genome sequencing by MBGP (2010) will further facilitate the unraveling of the detailed genome structures of *Brassica* species.

## 2.6 Summary

Understanding the regulatory mechanisms in lipid biosynthetic pathways in *B. napus* is essential for further genetic engineering of this crop for seed oil modification. As a key lipogenic enzyme, GPAT, plays important roles in several lipid synthetic pathways, including lipid polyester synthesis, membrane lipid synthesis and storage lipid synthesis. One of the major goals of the present doctoral study is to understand the physiological functions and molecular divergence of the ER GPATs of *B. napus* under the allotetraploid genetic background. Another focus of this study is to enzymatically characterize a plastidial GPAT isolated from a cold tolerant plant *Erysimum asperium*. Results from these studies will contribute to further understanding the GPAT-associated lipid metabolism pathways and provide useful information for metabolic engineering of storage lipid synthesis in *B. napus*.

## 2.7 References

- Ackman RG** (1990) Canola fatty acids- an ideal mixture for health, nutrition, and food use. In F Shahidi, ed, Canola and rapeseed: Production, chemistry, nutrition, and processing technology. Van Nostrand Reinhold, New York, pp 81-89
- Arabidopsis Genome Initiative** (2000) Analysis of the genome of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796-815
- Bates PD, Durrett TP, Ohlrogge JB and Pollard M** (2009) Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. *Plant Physiol* **150**: 55-72

**Beisson F, Li Y, Bonaventure G, Pollard M and Ohlrogge JB** (2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *Plant Cell* **19**: 351-368

**Bennett MD and Leitch IJ** (1997) Nuclear DNA amounts in angiosperms: 583 new estimates. *Ann Bot (Lond)*. **80**: 169-196

**Bennett MD, Leitch IJ, Price HJ and Johnston JS** (2003) Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175Mb) using flow cytometry show genome size in *Arabidopsis* to be ~157 Mb and thus ~25 % larger than the *Arabidopsis* Genome Initiative of ~125 Mb. *Annals of Botany* **91**: 1-11

**Bertrams M and Heinz E** (1981) Positional specificity and fatty acid selectivity of purified *sn*-glycerol 3-phosphate acyltransferases from chloroplasts. *Plant Physiol* **68**: 653-657

**Bird SM and Gray JE** (2003) Signals from the cuticle affect epidermal cell differentiation. *New Phytol* **157**: 9-23

**Blanc G and Wolfe KH** (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**: 1679-1691

**Blanc G, Barakat A, Guyot R, Cooke R and Delseny M** (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12**: 1093-1101

**Bonaventure G, Beisson F, Ohlrogge J and Pollard M** (2004) Analysis of the aliphatic monomer composition of polyesters associated with *Arabidopsis* epidermis: Occurrence of octadeca-cis-6,cis-9-diene-1,18-dioate as the major component. *Plant J* **40**: 920-930

- Bowers JE, Chapman BA, Rong J and Paterson AH** (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**: 433-438
- Cahoon EB, Shockey JM, Dietrich CR, Gidda SK, Mullen RT and Dyer JM** (2007) Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr Opin Plant Biol* **10**: 236–244
- Chen ZJ** (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu Rev Plant Biol* **58**: 377-406
- Cheung F, Trick M, Drou N, Lim YP, Park, JY, Kwon SJ, Kim JA, Scott R, Pires JC, Paterson AH, Town C and Bancroft I** (2009) Comparative analysis between homoeologous genome segments of *Brassica napus* and its progenitor species reveals extensive sequence-level divergence. *Plant Cell* **21**: 1912-1928
- Compagnon V, Diehl P, Benveniste I, Meyer D, Schaller H, Schreiber L, Franke R, Pinot F** (2009) CYP86B1 is required for very long chain  $\omega$ -hydroxyacids and  $\alpha$ ,  $\omega$ -dicarboxylic acids synthesis in root and seed suberin polyester. *Plant Physiol* **150**: 1831-1843
- Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H and Stymne H** (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc Natl Acad Sci USA* **97**: 6487–6492
- Das S, Castillo C and Stevens T** (2001) Phospholipid remodeling/generation in *Giardia*: the role of the Lands cycle. *Trends Parasitol* **17**: 316-319

- Force A, Lynch M, Pickett FB, Amores A, Yan YL and Postlethwait J** (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531-1545
- Franke R, Briesen I, Wojciechowski T, Faust A, Yephremov A, Nawrath C and Schreiber L** (2005) Apoplastic polyesters in Arabidopsis surface tissues: a typical suberin and a particular cutin. *Phytochemistry* **66**: 2643–2658
- Ghosal A, Banas A, Stahl U, Dahlqvist A, Lindqvist Y and Stymne S** (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyl transferase (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 and 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
- Gimeno RE and Cao J** (2008) Mammalian glycerol-3-phosphate acyltransferases: New genes for an old activity. *J Lipid Res* **49**: 2079-2088
- Gou JY, Yu XH and Liu CJ** (2009) A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in Arabidopsis. *Proc Natl Acad Sci USA* **106**: 18855-18860
- Griffiths G, Stymne S and Stobart AK** (1988a) Phosphatidylcholine and its relationship to triacylglycerol biosynthesis in oil-tissues. *Phytochemistry* **27**: 2089–2093

- Griffiths G, Stymne S and Stobart AK** (1988b) The utilization of fatty-acid substrates in triacylglycerol biosynthesis by tissue slices of developing safflower (*Carthamus tinctorius* L) and sunflower (*Helianthus annuus* L) cotyledons. *Planta* **173**: 309-316
- Hawkins DJ and Kridl JC** (1998) Characterization of acyl-ACP thioesterases of mangosteen (*Garcinia mangostana*) seed and high levels of stearate production in transgenic canola. *Plant J* **13**: 743–752
- Hildebrand D** (2008) Biotechnology and crop improvement in agriculture. In R Weselake, ed, *Teaching innovations in Lipids Science*, CRC Press, IBoca Raton, pp 105-122.
- Höfer R, Briesen I, Beck M, Pinot F, Schreiber L and Franke R** (2008) The Arabidopsis cytochrome P450 CYP86A1 encodes a fatty acid omega-hydroxylase involved in suberin monomer biosynthesis. *J Exp Bot* **59**: 2347-2360
- Inaba R and Nishio T** (2002) Phylogenetic analysis of *Brassicaceae* based on the nucleotide sequences of the S-locus related gene, SLR1. *Theor Appl Genet* **105**: 1159-1165
- Ishizaki O, Nishida I, Agata K, Eguchi G and Murata N** (1988) Cloning and nucleotide sequence of cDNA for the plastid glycerol-3-phosphate acyltransferase from squash. *FEBS Lett* **238**: 424-430
- Ishizaki-Nishizawa O, Azuma M, Ohtani T, Murata N and Toguri T** (1995) Nucleotide sequence of cDNA from *Spinacia oleracea* encoding plastid glycerol-3-phosphate acyltransferase. *Plant Physiol* **108**: 1342

- Jain RK, Coffey M, Lai K, Kumar A and MacKenzie SL** (2000) Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. *Biochem Soc Trans* **28**: 958-961
- Johnson TC, Schneider JC and Somerville C** (1992) Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-phosphate acyltransferase from cucumber. *Plant Physiol* **99**: 771-772
- Kannangara R, Branigan C, Liu Y, Penfield T, Rao V, Mouille G, Höfte H, Pauly M, Riechmann JL and Broun P** (2007) The Transcription Factor WIN1/SHN1 Regulates Cutin Biosynthesis in *Arabidopsis thaliana* *Plant Cell* **19**: 1278-1294
- Kennedy EP** (1961) Biosynthesis of complex lipids. *Fed Proc Am Soc Exp Biol* **20**: 934-940
- Knutzon DS, Hayes TR, Wyrick A, Xiong H, Davies HM and Voelker TA** (1999) Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the sn-2 position of triacylglycerols in lauric rapeseed oil and can increase total laurate levels. *Plant Physiol* **120**: 739-746
- Kolattukudy PE** (2001) Polyesters in higher plants. *Adv Biochem Eng Biotechnol* **71**: 1-49
- Kowalski SP, Lan TH, Feldmann KA and Paterson AH** (1994) Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* **138**: 499-510

- Kunst L, Browse J and Somerville C** (1988) Altered regulation of lipid biosynthesis in a mutant of *Arabidopsis* deficient in chloroplast glycerol-3-phosphate acyltransferase activity. *Proc Natl Acad Sci USA* **85**: 4143-4147
- Lagercrantz U and Lydiate DJ** (1996) Comparative genome mapping in *Brassica*. *Genetics* **144**: 1903-1910
- Lands WEM and Hart P** (1965) Metabolism of glycerolipids VI. Specificities of acyl-coA: phospholipid acyltransferases. *J Biol Chem* **240**: 1905–1911
- Li YH, Beisson F, Koo AJK, Molina I, Pollard M and Ohlrogge J** (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc Natl Acad Sci USA* **104**: 18339-18344
- Li-Beisson Y, Pollard M, Sauveplane V, Pinot F, Ohlrogge J and Beisson F** (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proc Natl Acad Sci USA* **106**: 22008-22013
- Li-Beisson Y, Shorrosh B, Beisson F, Andersson M, Arondel V, Bates P, Baud S, Bird D, DeBono A, Durrett T, Franke R, Graham I, Katayama K, Kelly A, Larson T, Markham J, Miquel M, Molina I, Nishida I, Rowland O, Samuels L, Schmid K, Wada H, Welti R, Xu C, Zallot R and Ohlrogge J** (2010) Acyl lipid metabolism. In R Last, ed, *The Arabidopsis book*. American Society of Plant Biologists, Rockville. doi: 10.1199/tab.0133
- Linder CR** (1998) Potential persistence of transgenes: seed performance of transgenic canola and wild × canola hybrids. *Ecol Applic* **8**: 1180-1195

- Lu C, Xin Z, Ren Z, Miquel M and Browse J** (2009) An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of *Arabidopsis*. Proc Natl Acad Sci USA **106**: 18837-18842
- Lü S, Song T, Kosma DK, Parsons EP, Rowland O and Jenks MA** (2009) *Arabidopsis CER8* encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. Plant J **59**: 553-564
- Lukens L, Quijada PA, Udall J, Pires JC, Schranz ME and Osborn T** (2004) Genome redundancy and plasticity within ancient and recent Brassica crop species. Biol J Linn Soc **82**: 675-688
- Lysak MA, Koch MA, Pecinka A and Schubert I** (2005) Chromosome triplication found across the tribe *Brassicaceae*. Genome Res **15**: 516-525
- Molina I, Bonaventure G, Ohlrogge J and Pollard M** (2006) The lipid polyester composition of *Arabidopsis thaliana* and *Brassica napus* seeds. Phytochemistry **67**: 2597-2610
- Molina I, Li-Beisson Y, Beisson F, Ohlrogge JB and Pollard M** (2009) Identification of an *Arabidopsis* feruloyl-coenzyme A transferase required for suberin synthesis. Plant Physiol **151**: 1317-1328
- Murata N and Tasaka Y** (1997) Glycerol-3-phosphate acyltransferase in plants. Biochim Biophys Acta **1348**: 10-16
- Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y and Nishida I** (1992) Genetically engineered alteration in the chilling sensitivity of plants. Nature **356**: 710-713

- Nawrath C** (2002) The biopolymers cutin and suberin. In CR Somerville, EM Meyerowitz, editors, *The Arabidopsis Book*. American Society of Plant Biologists, Rockville, Doi: 10.1199/tab.0021
- Nawrath C** (2006) Unraveling the complex network of cuticular structure and function. *Curr Opin Plant Biol* **9**: 281-287
- Nishida I and Murata N** (1996) Chilling sensitivity in plants and cyanobacteria: the crucial contribution of membrane lipids. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 541-568
- Nishida I, Frentzen M, Ishizaki O and Murata N** (1987) Purification of isomeric forms of acyl-wacyl-carrier-protein: glycerol-3-phosphate acyltransferase from greening squash cotyledons. *Plant Cell Physiol* **28**: 1071-1079
- Nishida I, Tasaka Y, Shiraishi H and 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
- Numa S** (1974) Regulation of fatty-acid synthesis in higher animals. *Ergeb Physiol* **69**: 54-96
- Ohlrogge J and Browse J** (1995) Lipid biosynthesis. *Plant Cell* **7**: 957-970.
- Ohlrogge JB and Jaworski JG** (1997) Regulation of fatty acid synthesis. *Annu Rev Plant Physiol. Plant Mol Biol* **48**: 109-136
- O'Neill CM and Bancroft I** (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homeologous to

sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J.* **23**: 233-243

parameters influencing the accumulation of laurate in rapeseed. *Plant J* **9**:

**Park JY, Koo DH, Hong CP, Lee SJ, Jeon JW, Lee SH, Yun PY, Park BS, Kim HR, Bang JW, Plaha P, Bancroft I and Lim YP** (2005) Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kb gene-rich region of *Arabidopsis* chromosome 4 and partially duplicated on chromosome 5. *Mol Genet Genomics* **274**: 579-588

**Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC and Lydiate DJ** (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **171**: 765-781

**Pollard M, Beisson F, Li Y and Ohlrogge JB** (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci* **13**: 236-246

**Rana D, van den Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP and Bancroft I** (2004) Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. *Plant J* **40**: 725-733

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

**Scarth R and Tang J** (2006) Modification of *Brassica* oil using conventional and transgenic approaches. *Crop Sci* **46**: 1225-1236

**Scarth R, McVetty PBE and Rimmer SR** (1995a) Mercury high erucic acid low glucosinolate summer rape. *Can J Plant Sci* **75**: 205-206

- Scarth R, Rimmer SR and McVetty PBE** (1995b) Apollo low linolenic summer rape. *Can J Plant Sci* **75**: 203-204
- Schmid KM and Ohlrogge JB** (2008) Lipid Metabolism in Plants. In DE Vance, JE Vance, eds, *Biochemistry of Lipids, Lipoproteins and Membranes*, Ed 5 Vol 4. Elsevier Science, Amsterdam, The Netherlands, pp 97-130
- Schnurr J, Shockey J and Browse J** (2004) The acyl-CoA synthetase encoded by *LACS2* is essential for normal cuticle development in *Arabidopsis*. *Plant Cell* **16**: 629-642
- Schnurr JA, Shockey JM, De Boer GJ and 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
- Shanklin J and Cahoon EB** (1998) Desaturation and related modifications of fatty acids. *Annu Rev Plant Physiol Plant Mol Biol* **49**: 611-641
- Simillion C, Vandepoele K, Van Montagu MC, Zabeau M and Van de Peer Y** (2002) The hidden duplication past of *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **99**: 13627-13632
- Slabas AR, Simon JW and Brown AP** (2001) Biosynthesis and regulation of fatty acids and triglycerides in oil seed rape. Current status and future trends. *Eur J Lipid Sci Technol* **103**: 455-466
- Slack CR, Roughan PG and Balasingham N** (1978) Labeling of glycerolipids in cotyledons of developing oilseeds by [1-C-14] acetate and [2-H-3] glycerol. *Biochem J* **170**: 421-433

- Snyder CL, Yurchenko OP, Siloto RM, Chen X, Liu Q, Mietkiewska E and Weselake RJ** (2009) Acyltransferase action in the modification of seed oil biosynthesis. *New Biotechnol* **26**: 11-16
- Sperling P and Heinz E** (1993) Isomeric *sn*-1-octadecenyl and *sn*-2-octadecenyl analogues of lysophosphatidylcholine as substrates for acylation and desaturation by plant microsomes. *Eur J Biochem* **213**: 965-971
- Sperling P, Linscheid M, Stocker S, Muhlbach HP and Heinz E** (1993) In-vivo desaturation of cis-delta-9-monounsaturated to cis-delta-9,12-diunsaturated alkenylether glycerolipids. *J Biol Chem* **268**: 26935-26940
- Stefansson BR and Downey RK** (1995) Rapeseed. In AE Slinkard, DR Knott, eds, *Harvest of gold: The history of field crop breeding in Canada*. Univ Ext Press, Univ of Saskatchewan, Saskatoon, pp 140-152
- Stobart K, Mancha M, Lenman M, Dahlqvist A and 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 and 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
- Thelen JJ and Ohlrogge JB** (2002) Metabolic engineering of fatty acid biosynthesis in plants. *Metab Eng* **4**: 12-21

**Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortman JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M and Bancroft I** (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. *Plant Cell* **18**: 1348-1359

**U N** (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn J Bot* **7**: 389-452

**Udall JA and Wendel JF** (2006) Polyploidy and crop improvement. *The Plant Genome, a Supplement to Crop Science*. **46**: S3-14

**van de Loo FJ, Fox BG and Somerville C** (1993) Unusual fatty acids. In TS Moore Jr, ed, *Lipid metabolism in plants*. CRC, Boca Raton, FL. pp 91-126

**Verwoert IIGS, Van Der Linden KH, Walsh MC, Nijkamp HJJ and Stuitje AR** (1995) Modification of *Brassica napus* seed oil by expression of the *Escherichia coli fabH* gene, encoding 3-ketoacyl-acyl carrier protein synthase III. *Plant Mol Biol* **27**: 875-886

**Vilkki JP and Tanhuanpää PK** (1995) Breeding of high oleic acid spring turnip rape in Finland. In Proc 9th Int Rapeseed Cong, Cambridge, UK. pp 386-388

**Vision TJ, Brown DG and Tanksley SD** (2000) The origins of genomic duplications in *Arabidopsis*. *Science* **290**: 2114

**Voelker TA, Hayes TR, Cranmer AM, Turner JC and Davies HM** (1996) Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. *Plant J* **9**: 229-241

- Vogel G and Browse J** (1996) Cholinephosphotransferase and diacylglycerol acyltransferase (Substrate Specificities at a Key Branch Point in Seed Lipid Metabolism). *Plant Physiol* **110**: 923-931
- Wang YP, Sonntag K and Rudloff E** (2003) Development of rapeseed with high erucic acid content by asymmetric somatic hybridization between *Brassica napus* and *Crambe abyssinica*. *Theor Appl Genet* **106**: 1147-1155
- Warwick SI and Black LD** (1991) Molecular systematics of *Brassica* and allied genera (Subtribe Brassicinae, Brassiceae)-Chloroplast genome and cytodeme congruence. *Theor Appl Genet* **82**: 81–92
- Weber S, Wolter FP, Buck F, Frentzen M and Heinz E** (1991) Purification and cDNA sequencing of an oleate-selective acyl-ACP: *sn*-glycerol-3-phosphate acyltransferase from pea chloroplasts. *Plant Mol Biol* **17**: 1067-1076
- Wendel JF** (2000) Genome evolution in polyploids. *Plant Mol Bio* **42**: 225-249
- Weselake RJ** (2005) Storage production metabolism in microspore-derived cultures of Brassicaceae. In CE Don palmer, WA Keller, KJ Kasha, eds, Haploids in Crop Improvement II. Biotechnology in Agriculture and Forestry, Vol 56, Berlin Heidelberg: Springer-Verlag, pp 295-303
- Whittle CA and Krochko JE** (2009) Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* **21**: 2203-2219
- Wiberg E and Bafor M** (1995) Medium chain-length fatty acids in lipids of developing oil palm kernel endosperm. *Phytochemistry* **39**: 1325-1327

- Xiao F, Goodwin MS, Xiao Y, Sun Z, Baker D, Tang X, Jenks MA and Zhou JM** (2004) *Arabidopsis* CYP86A2 represses *Pseudomonas syringae* type III genes and is required for cuticle development. *EMBO J* **23**: 2903-2913
- Xu C, Cornish AJ, Froehlich JE and Benning C** (2006) Phosphatidylglycerol biosynthesis in chloroplasts of *Arabidopsis* mutants deficient in acyl-ACP glycerol-3-phosphate acyltransferase. *Plant J* **47**: 296-309
- Yang YW, Lai KN, Tai PY and Li WH** (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J Mol Evol* **48**: 597-604
- Yurchenko OP, Nykiforuk CL, Moloney MM, Ståhl U, Banaś A, Stymne S and Weselake RJ** (2009) A 10-kDa acyl-CoA-binding-protein (ACBP) from *Brassica napus* enhances acyl exchange between acyl-CoA and phosphatidylcholine. *Plant Biotechnol J* **7**: 602-610
- Zheng Z, Xia Q, Dauk M, Shen W, Selvaraj G and Zou J** (2003) *Arabidopsis* AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* **15**: 1872-1887

## Chapter 3.

### Three homologous genes encoding *sn*-glycerol-3-phosphate acyltransferase 4 exhibit functional divergence in *Brassica napus*

#### 3.1 Introduction

Polyploidy has long been considered a prominent evolutionary force for flowering plants (Udall and Wendel, 2006). Over 95% of the lineages of angiosperms have undergone at least one event of polyploidization over their evolutionary time, suggesting that most of the existing flowering plants evolved from ancient polyploids (Bennett and Leitch, 1997; Gaeta et al., 2007; Chen, 2007). Polyploidy, along with genomic segmental duplications, could benefit plants by increasing overall gene expression levels and cell sizes, and providing sources for novel variants and genome “buffering” of deleterious mutations (Udall and Wendel, 2006). Genes duplicated by such events could undergo three primary evolutionary fates over the long-term (Wendel 2000; Lukens et al., 2004; Blanc and Wolfe, 2004; Whittle and Krochko, 2009): (1) pseudogenization (loss or silencing), whereby duplicated genes with redundant functions accumulate deleterious mutations and are eventually lost without detrimental effects on plant fitness; (2) neofunctionalization, whereby some redundant genes develop new adaptive functions by positive Darwinian selection; (3) subfunctionalization, a process in which the ancestral gene functions become subdivided among the duplicated genes.

---

A portation of this chapter has been published: Chen et al. (2011) Three homologous genes encoding *sn*-glycerol-3-phosphate acyltransferase 4 exhibit different expression patterns and functional divergence in *Brassica napus*. *Plant physiology* 155: 851-865.

*Brassica napus* (AACC; n=19) is an allotetraploid oilseed crop, which evolved from the hybridization of two diploid progenitors, *B. rapa* (AA; n=10) and *B. oleracea* (CC; n=9) during human cultivation (over 10,000 years ago) (U, 1935; Cheung et al., 2009). The *Brassica* species are closely related to the model plant *Arabidopsis*, all of which belong to the same tribe (*Brassicaceae*), and share a common recent ancestry (20 million years ago) (Yang et al., 1999). Comparative mapping studies of the genomic microstructures of *B. oleracea*, *B. rapa*, *B. nigra*, *B. napus* and *Arabidopsis* revealed extensive triplications in the genomes of the diploid *Brassica* progenitors and strongly suggested that the extant *Brassica* diploid species evolved from a common hexaploid ancestor (O'Neill and Bancroft, 2000; Rana et al., 2004; Park et al., 2005). A previous study also showed that the majority of the *Arabidopsis* conserved genomic regions could be mapped to six conserved segments within the allotetraploid genome of *B. napus* (Parkin et al., 2005). There are a few exceptions, however, where fewer or more copies of certain segments have been detected in the genome of *B. napus*. This could be caused by multiple rounds of duplication (either segmental or the result of polyploidy), along with genome-wide rearrangements and segmental deletions during the evolution process (Cheung et al., 2009). Thus, the complex genome structure of the diploid *Brassica* progenitors, together with the extensive genome rearrangements after speciation, have led to genes being represented as multiple homologues in the allotetraploid *B. napus*.

Although there are numerous studies comparing the genomic structures of the *Brassica* species and *Arabidopsis*, little is known about molecular and

functional variances among homologous genes arising from polyploidy and genomic segmental duplications in *B. napus*. In part, this is due to experimental challenges in distinguishing highly identical transcripts and polypeptides, which are inherited from two genomically similar and evolutionarily related progenitors. Such information, however, is fundamentally important for a better understanding of the complex mechanisms involved in variant biological pathways in the *Brassica* polyploid species. Additionally, in the genetic engineering of plants with polyploid backgrounds, knowledge of the transcriptional and functional behavior of individual homologues is essential to avoid pleiotropic effects. For example, in *Triticum aestivum* (bread wheat) three *WLHS1* homoeologous genes (originating from A, B and D genomes, respectively) are associated with different effects on flowering time (Shitsukawa et al., 2007). Such an understanding would allow for the manipulation of specific genes relevant to the targeted metabolic process, without compromising overall plant fitness.

The *sn*-glycerol-3-phosphate acyltransferases (GPAT; EC 2.3.1.15) are involved in catalyzing the initial step in the assembly of glycerolipids (Zheng et al., 2003). In *Arabidopsis*, ten genes have been identified as encoding GPAT enzymes located in various subcellular compartments, i.e., plastid (*ATS1*), mitochondria (*AtGPAT1-3*) and endoplasmic reticulum (ER) (*AtGPAT4-9*) (Zheng et al., 2003; Xu et al., 2006; Gidda et al., 2009). Recent studies in *Arabidopsis* have shown that several members in the ER-bound GPAT family are involved in lipid polyester synthesis (i.e., cutin and suberin) (Li et al., 2007; Beisson et al., 2007).

In the present study, we identified and characterized three *B. napus* *GPAT4* homologues. Phylogenetic analysis of the genomic DNA sequences of the *GPAT4* genes from *B. napus*, *B. rapa* and *B. oleracea* strongly suggested that two of the *BnGPAT4* homologues originated from the C genome, and the third originated from the A genome. Heterologous expression in yeast revealed that all three *BnGPAT4* homologues encoded functional GPAT enzymes with different levels of accumulation. The gene expression, epigenetic variations and phenotypic rescue of the *gpat4 gpat8 Arabidopsis* double mutant indicated that the three *BnGPAT4* genes have evolved through functional divergence and are involved in lipid polyester synthesis. In addition, analysis of the *gpat4* RNAi lines of *B. napus* revealed several important physiological roles of *BnGPAT4* in plant development, including cutin biosynthesis, early flower development, pollen development and storage lipid biosynthesis.

## **3.2 Results**

### **3.2.1 Identification of *GPAT4* homologues in *Brassica napus***

The cDNA sequence of *Arabidopsis GPAT4* (At1g01610) was used as a reference to query the *B. napus* EST database in GenBank (NCBI) for UniGene entries (Pontius et al., 2003). Overall, 43 UniGene entries were identified (UniGene accession No. Bna.684). After analyzing their alignment, the 43 UniGene sequences were condensed into three groups, which were denoted as three *BnGPAT4* homologues.

In *B. napus*, the A and C genomes of the diploid progenitors became homoeologous subgenomes. Due to this complex genome background, the three

identified *BnGPAT4* genes may represent paralogues (i.e., duplicated genes originating from the same subgenome) and homoeologues (i.e., orthologous genes originating from different subgenomes coexisting in the newly synthesized allotetraploid genome). As the homology between the subtypes is complex, the *GPAT4* genes from *B. napus* were referred to as homologues, according to the recommendation from Fitch (2000).

The full-length cDNAs and genomic DNA sequences of the three *BnGPAT4* homologues were cloned by PCR using homologue-specific primer pairs, which were designed based on single nucleotide polymorphisms in the sequences. The cDNA sequences of the three *BnGPAT4* homologues and *AtGPAT4* shared over 90% similarity. The deduced amino acid sequences shared even greater sequence similarity (over 94%). The major sequence differences between the three homologues were found in the two intron regions as shown in Fig. 3.1.

### **3.2.2 Genome origins of the three *BnGPAT4* homologues**

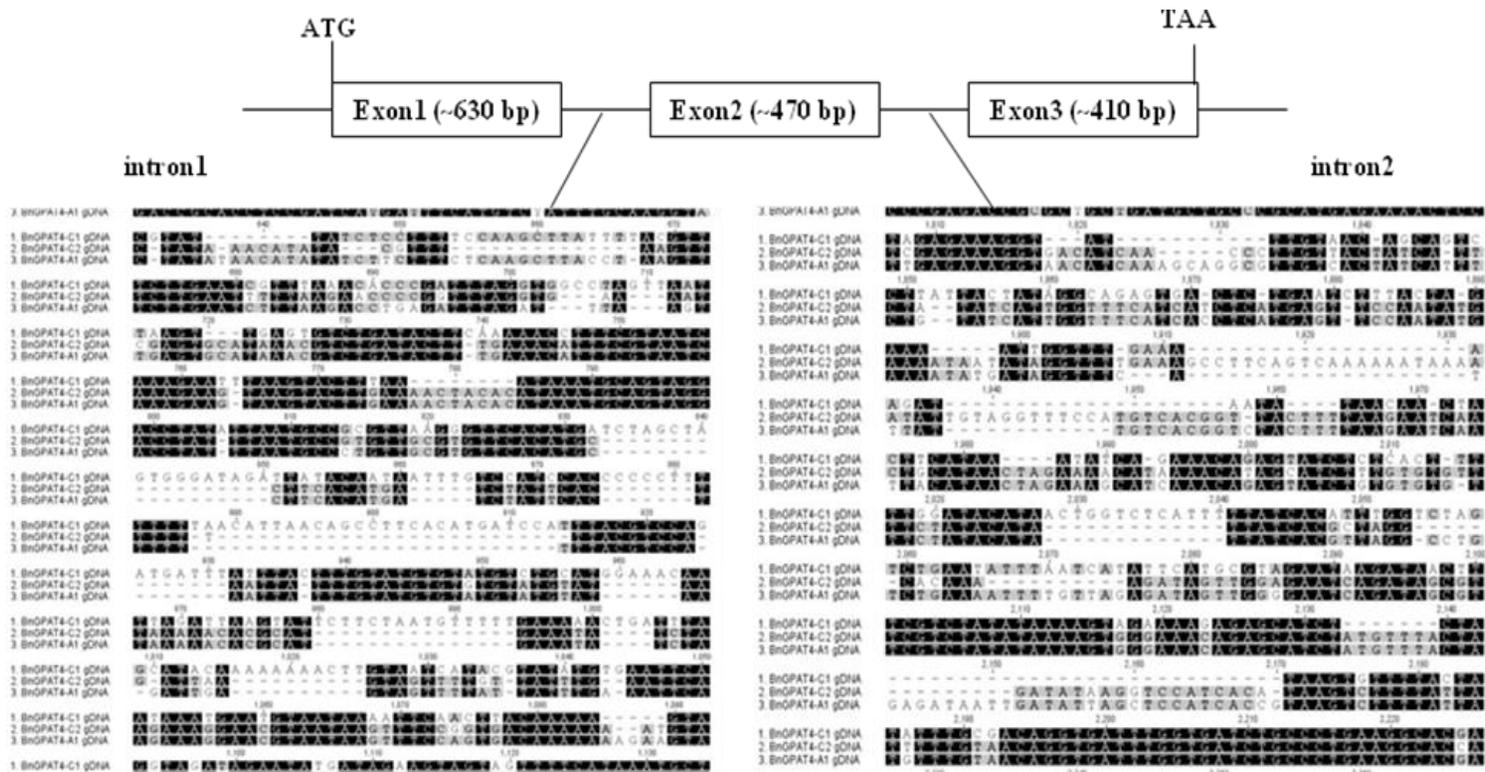
*B. napus* is an allotetraploid (AACC) formed from hybridization of two diploid progenitors, *B. rapa* (AA) and *B. oleracea* (CC) (U, 1935). To study the genome origins of the three *BnGPAT4* homologues, the genomic DNA sequences of *GPAT4* genes from *B. rapa* and *B. oleracea* were isolated and compared with the *BnGPAT4* homologues. Using primer pairs designed from the *BnGPAT4* sequences, two *GPAT4* genes were isolated from *B. rapa* and two were isolated from *B. oleracea*. A comparison of the genomic DNA sequences, particularly in the intron regions, revealed that two of the *BnGPAT4* homologues exhibited the

highest sequence similarity with two *B. oleracea* *GPAT4* genes (*CC\_GPAT4\_1* and 2), and were thus named *BnGPAT4-C1* and *C2*. The third *BnGPAT4* homologue exhibited the highest sequence similarity with one of the *GPAT4* genes isolated from *B. rapa* (*AA\_GPAT4\_1*) and was thus named *BnGPAT4-A1* (Fig. 3.2A). Phylogenetic analysis of the genomic DNA sequences further confirmed the homology among *GPAT4* genes from the three *Brassica* species and *Arabidopsis* (Fig. 3.2B).

Notably, the second *GPAT4* gene cloned from *B. rapa* (*AA\_GPAT4\_2*) did not have a counterpart within the three *BnGPAT4* homologues. To examine whether a fourth, unidentified *BnGPAT4* gene homologous to *AA\_GPAT4\_2* exists in *B. napus*, PCR was performed using *B. napus* genomic DNA as a template and *AA\_GPAT4\_2*- specific primers. No amplicon was generated from the PCR using *B. napus* genomic DNA as template; however, the target amplicon of ~420 bp was obtained using *B. rapa* genomic DNA as template (Fig. A1). This 420 bp amplicon was further sequenced and confirmed to be the target region of the *AA\_GPAT4\_2* gene. This result suggested that *AA\_GPAT4\_2* did not have a counterpart copy in the *B. napus* genome.

To further investigate whether there are other unidentified *GPAT4* homologues in *B. napus*, a genomic qRT-PCR approach was used. Genomic DNA extracted from a homozygous transgenic line of *B. napus* DH12075, which had a single copy of a transgene cassette (napin promoter-*BnDGAT1*- pea *rbcS* terminator). SYBR-green qRT-PCR was performed with two pairs of primers (Table A1), *BnGPAT4* primers (designed based on the highly conserved regions

of the identified *BnGPAT4* homologues) and *rbcS* terminator primers. The pea *rbcS* terminator, which has no sequence similarity to *B. napus* genes, was considered as a single copy internal control. The amplification quantities of the *GPAT4* homologues and *rbcS* terminator were then compared. The qRT-PCR result indicated that the amplification quantity of *BnGPAT4* homologues was three times of the quantity of *rbcS* terminator, which supported the conclusion of three *GPAT4* homologues existing in the *B. napus* genome.



Sequence identity of cDNA / genomic DNA of *BnGPAT4* homologues

	<i>BnGPAT4-C2</i>	<i>BnGPAT4-A1</i>	<i>AtGPAT4</i>
<i>BnGPAT4-C1</i>	93% / 87%	94% / 87%	90% / 82%
<i>BnGPAT4-C2</i>		97% / 96%	90% / 81%
<i>BnGPAT4-A1</i>			90% / 81%

Figure 3.1. Genome structure of the *BnGPAT4* homologues.

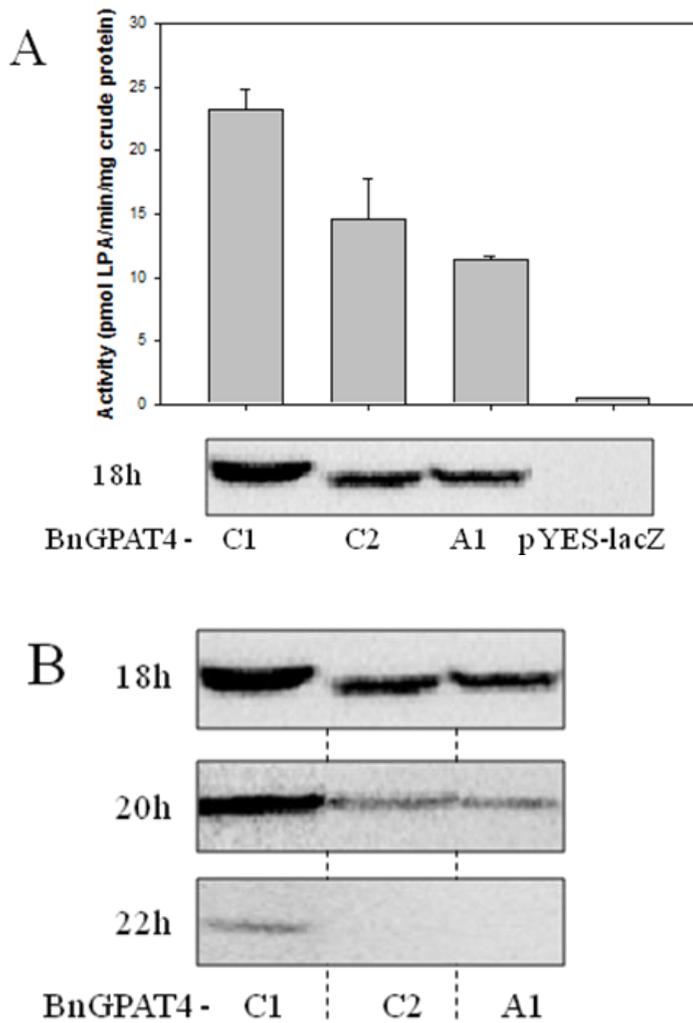
The major sequence differences among the homologues are in the intron regions as shown in the two blocks.



### **3.2.3 *BnGPAT4* homologues encode functional GPAT enzymes with different levels of polypeptide accumulation**

The cDNAs of each of the three *BnGPAT4* homologues were cloned in frame (without a stop codon) with the V5 epitope and polyhistidine tag of the yeast expression vector, pYES2.1/V5-His-TOPO (Invitrogen, Carlsbad, CA, USA). The expression of *BnGPAT4* homologues was controlled by an inducible promoter *GALI* in the presence of galactose in the growth medium. A yeast mutant strain *gat1Δ*, deficient in the major ER-bound GPAT activity (Zheng and Zou, 2001), was used for galactose-induced expression of the *BnGPAT4* recombinant proteins. The *gat1Δ* yeast strain has a very low GPAT activity background and has been commonly used as a test strain for putative *GPATs* cloned from a number of species (Zheng and Zou, 2001, Zheng et al., 2003). In *Arabidopsis*, *GPAT4* was confirmed to be an ER-bound enzyme (Zheng et al., 2003). Thus, we assayed GPAT activity in microsomal fractions of *gat1Δ* yeast expressing *BnGPAT4* cDNAs or a null vector. As shown in Fig. 3.3A, all three *BnGPAT4* proteins exhibited GPAT enzyme activity after 18h of galactose induction, but the levels of activity varied. Results from a Western blot of the same protein samples indicated that the higher enzyme activity of *BnGPAT4*-C1 was likely due to a higher abundance of the *BnGPAT4*-C1 protein in the transformed yeast. We further investigated the protein accumulation levels upon longer galactose induction time (up to 22h). The abundance of all *BnGPAT4* proteins decreased with the increasing time of induction (Fig. 3.3B). In addition,

the abundance of BnGPAT4-C1 was always higher than BnGPAT4-A1 and C2 in the yeast expression system.



**Figure 3.3. The three *BnGPAT4* homologues encode functional GPAT enzymes but with different activities.**

(A) The activities of the BnGPAT4 enzymes and the corresponding protein amounts detected by Western blot after 18h galactose induction in the *gat1Δ* yeast expression system. (B) The abundance of the BnGPAT4 polypeptides detected by Western blot after 18, 20 and 22h galactose induction in the *gat1Δ* yeast expression system.

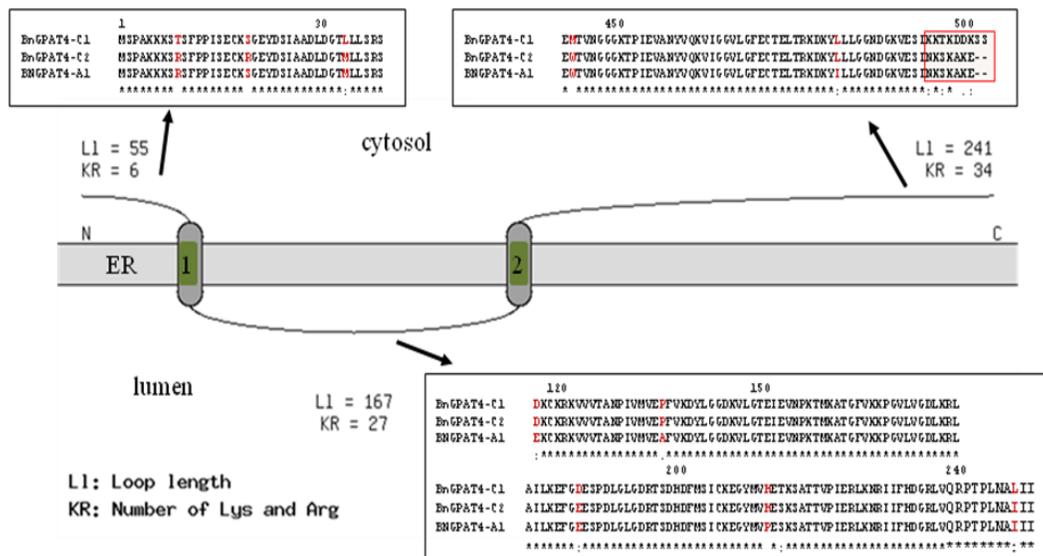
### **3.2.4 The C-termini of the BnGPAT4 enzymes are critical in regulating the protein metabolism**

The polypeptide sequences of the three forms of BnGPAT4 share up to 99% similarity (Fig. 3.4A). Protein topology prediction (TopPred, Claros and von Heijne, 1994) indicated that all three forms of BnGPAT4 proteins possessed two ER transmembrane domains (TMDs) at exactly the same positions (Fig. 3.4B). The major difference among the amino acid sequences of the three forms of BnGPAT4 is in the final 5-7 amino acid residues at the C-termini (Fig. 3.4B). BnGPAT4-C2 and A1 have the same C-terminal sequences but are different from BnGPAT4-C1, and both BnGPAT4-C2 and A1 exhibited lower protein abundance than BnGPAT4-C1. We hypothesized that these differences at the C-termini may contribute to the observed differences in BnGPAT4 protein abundance in yeast. To test this hypothesis, we made several modifications to the amino acid sequences of the three BnGPAT4 proteins including swapping and deletion of the C-termini. The cDNA sequences of the three *BnGPAT4* homologues were modified by PCR mutagenesis, subcloned into pYES2.1 and expressed in the yeast strain *gat1Δ*. The resulting mutated BnGPAT4s were named: BnGPAT4-C1<sup>M</sup>, -C2<sup>M</sup>, -A1<sup>M</sup>, -C1<sup>T</sup>, -C2<sup>T</sup> and -A1<sup>T</sup>, where superscript “M” refers to C-terminus exchanged with a different BnGPAT4 homologue and “T” refers to a truncated C-terminus (Fig. 3.5A).

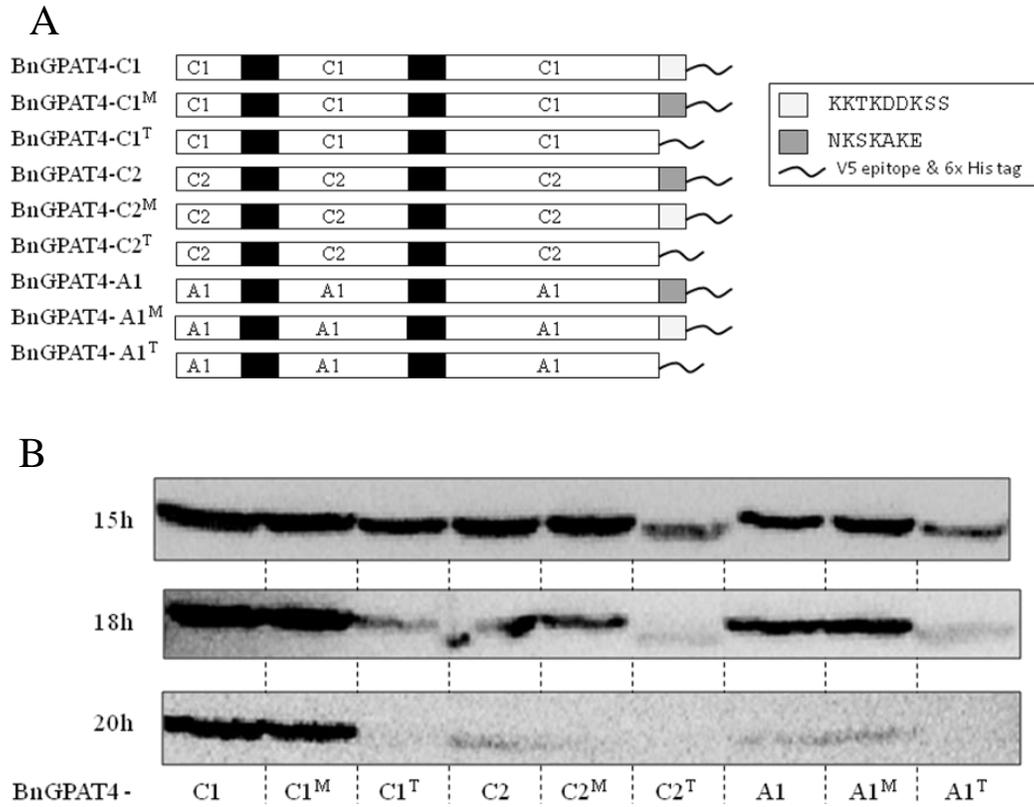
A

	BnGPAT4-C2	BnGPAT4-A1	AtGPAT4
BnGPAT4-C1	97%	97%	96%
BnGPAT4-C2		99%	95%
BnGPAT4-A1			94%

B



**Figure 3.4. Comparison of the amino acid sequences of the three BnGPAT4 proteins.**  
 (A) Amino acid sequence identity table of BnGPAT4 and AtGPAT4 proteins. (B) Predicted protein topology of the three BnGPAT4 enzymes and partial amino acid sequence alignment. Conserved residues are denoted by an asterisk, strongly similar amino acids by a colon, weakly similar amino acids by a period, and the different amino acids among the three BnGPAT4 proteins are shown in red.



**Figure 3.5. The C-termini of the BnGPAT4s are important for controlling the level of polypeptide accumulation.**

(A) The structures of native and chimeric BnGPAT4 proteins. Black bars represent the putative transmembrane domains. Light grey bar (KKTKDDKSS) represents the C-terminus of BnGPAT4-C1. Heavy grey bar (NKSKAKE) represents the C-terminus of BnGPAT4-C2 and A1. (B) Western blot analysis of the protein abundances of the native and chimeric BnGPAT4 proteins in transformed *gat1Δ* yeast with different galactose induction times.

Western blots were used to compare the quantity of the BnGPAT4 proteins in yeast after 15-20h of induction. As shown in Fig. 3.5B, each pair of the native and C-termini-exchanged BnGPAT4 proteins exhibited similar protein abundance, suggesting the exchange of C-termini had no effect on protein abundance. On the other hand, all BnGPAT4 proteins with truncated C-termini exhibited much lower abundance compared to proteins with unmodified or exchanged C-termini. Collectively, these results indicated that the C-termini of BnGPAT4 proteins play an important role in controlling the synthesis and/or degradation of the BnGPAT4 enzymes; however, the differences in protein abundance do not appear to be related to the sequence differences in the C-termini.

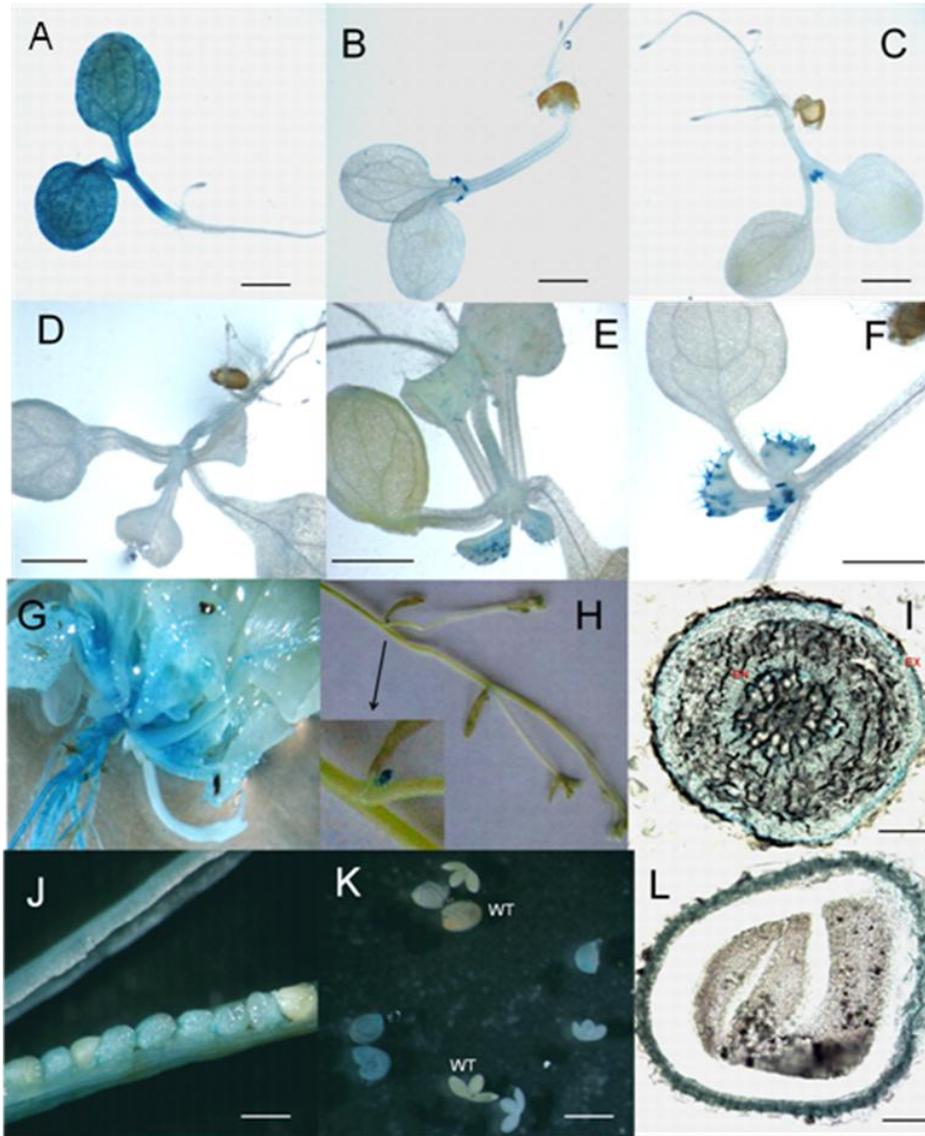
### **3.2.5 The three *BnGPAT4* homologues exhibit distinct expression patterns**

The promoter regions of the three *BnGPAT4* homologues have a much lower sequence similarity compared to their protein coding sequences (Fig. A2). To study the gene expression patterns of the three *BnGPAT4* homologues,  $\beta$ -glucuronidase (GUS) assay (using promoter-GUS fusion constructs) and Taqman qRT-PCR were performed. Fragments consisting of over 600 bp upstream of the first ATGs of the three *BnGPAT4* homologues were used to drive the expression of the  $\beta$ -glucuronidase (GUS) reporter gene in transgenic *Arabidopsis*. Three to five independent transgenic lines for each construct were analyzed for GUS activities. The results from histochemical staining revealed that the promoters of

different homologues directed GUS expression in various plant tissues and organs (Fig. 3.6). The differences in GUS expression patterns between different *BnGPAT4* promoters were most distinct in young seedlings (Fig. 3.6 A-F); at older stages, the expression patterns were more similar between all three promoters.

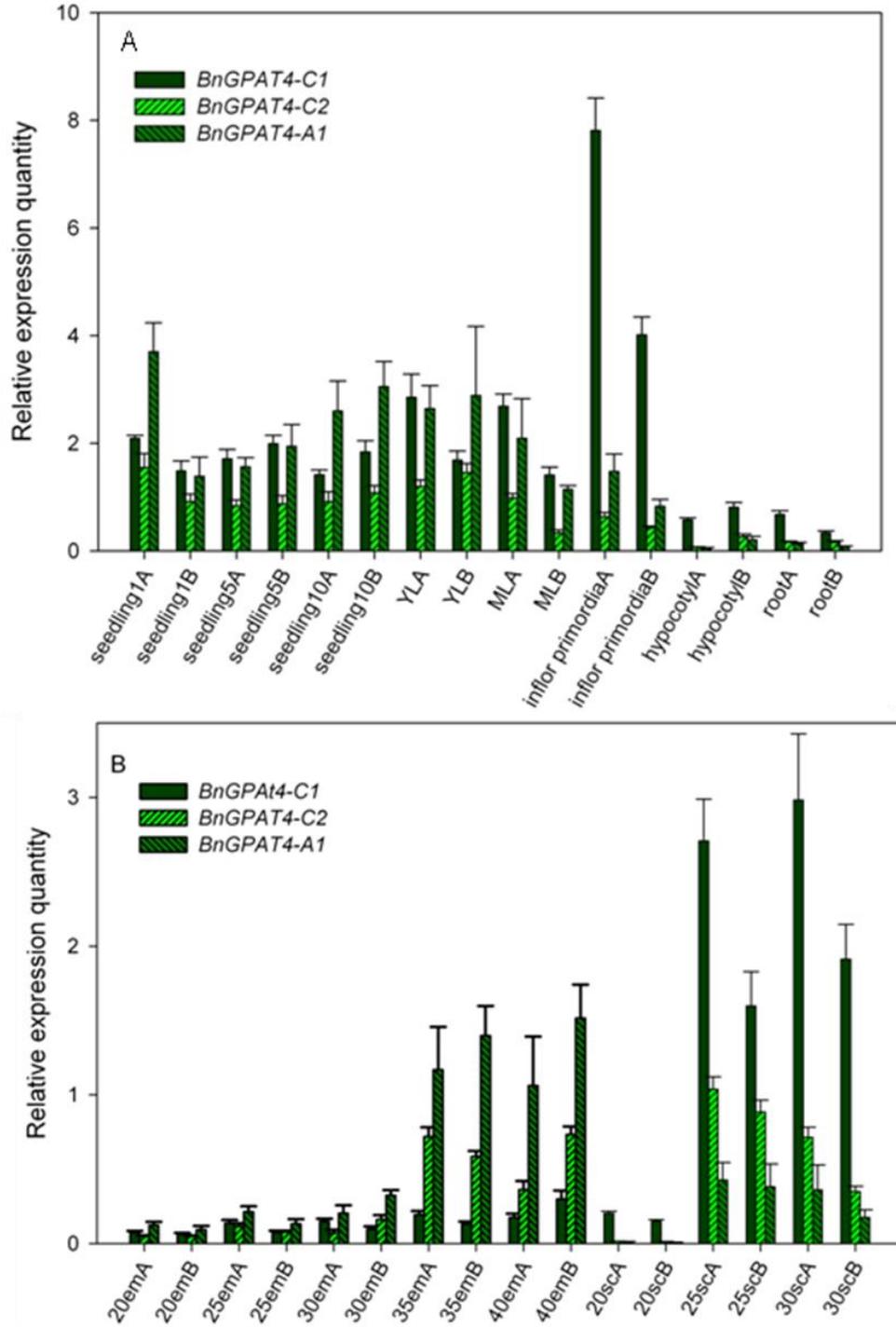
Considering that other regulatory factors and elements could also be involved in regulating gene expression *in vivo*, and the limitations of the promoter-GUS assay (such as the lack of accurate quantification and dosage effects of the transgene), we further performed Taqman qRT-PCR to quantify the transcript levels of the individual *BnGPAT4* homologues in wild type *B. napus*. Based on the preliminary results from the GUS assay, we tested several tissues and organs from *B. napus*, including tissues where the GUS assays indicated that the *BnGPAT4* homologues might be highly expressed. The qRT-PCR results (Fig. 3.7) were generally consistent with the earlier observations in GUS assay. In general, the *BnGPAT4* genes had higher expression levels in vegetative tissues than in seeds. In particular, *BnGPAT4-C1* was most highly expressed in inflorescence primordia (Fig. 3.7A). During embryo and seed coat development, *BnGPAT4-C1* and *A1* had quite distinct expression patterns. In the developing embryos, all *BnGPAT4* homologues exhibited increasing transcript abundance as the developing embryo matured, and *BnGPAT4-A1* was expressed at a much higher level than the other two homologues. In the developing seed coat, the transcript abundance of all *BnGPAT4* homologues was very low at the 20-DAP stage and increased substantially after the 25-DAP stage. Notably *BnGPAT4-C1*

was expressed at a very high level in the seed coat when compared to the other homologues; in contrast, its expression level was the lowest in embryos. We were only able to investigate the transcript abundance in seed coats excised from the developing seeds at 20, 25 and 30-DAP, since no RNA could be extracted from the samples at later developmental stages. This is due to the fact that as the seed matures, most of the seed coat cells undergo cell death and become compressed cell layers impregnated with brown pigments (Fig. A3) (Haughn and Chaudhury, 2005; Wan et al., 2002).



**Figure 3.6. Histochemical analysis of the GUS activities directed by individual *BnGPAT4* promoters in transgenic *Arabidopsis*.**

Similar GUS activities were observed in at least three independent transgenic lines for each construct. Only one transgenic line was shown in each figure. (A-C) 5-day old seedlings, GUS activities driven by the three promoters of *BnGPAT4-C1*, *C2* and *A1*, respectively. Bar 0.5 mm. (D-F) 10-day old seedlings, GUS activities driven by the three promoters of *BnGPAT4-C1*, *C2* and *A1*, respectively. Bar 5 mm. (G-L) Figures represent similar GUS activities observed with all three *BnGPAT4* promoters. (G) GUS activity driven by *BnGPAT4-C2* promoter. (H-L) GUS activities driven by *BnGPAT4-A1* promoter. (G) 20-day old whole plant. (H) inflorescence stem; left corner: enlarged figure of inflorescence primordia. (I) Cross section of the main root; EX, exodermis; EN, endodermis. Bar 0.3 mm. (J) Middle-late staged seeds in a silique. Bar 0.5 mm. (K) dissected embryos and seed coats. Bar 1 mm. (L) Cross section of a whole seed. Bar 0.1 mm.



**Figure 3.7. Gene expression patterns of the three *BnGPAT4* homologues.**

(A) Expression in vegetative tissues and organs. (B) Expression in embryos and seed coats.

Seedling 1, 5 and 10 refer to 1, 5 and 10-day old seedlings, respectively. YL, young leaves. ML, mature leaves. 20 to 40em, 20 to 40-DAP embryos. 20 to 30sc, seed coats excised from 20 to 30-DAP seeds. A and B after the sample names refer to the two biological replicates. Three technical replicates were performed for each sample.

### 3.2.6 Epigenetic variation among the *BnGPAT4* homologues

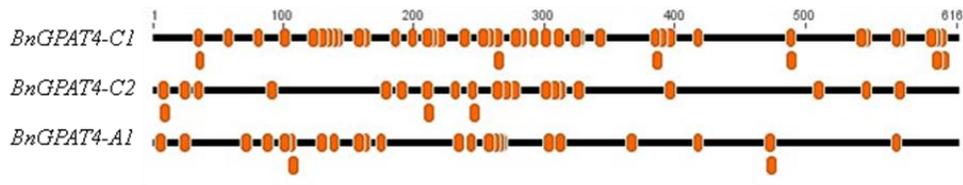
Cytosine DNA methylation has been recognized as an important mechanism for the regulation of gene expression and has been found to be a common epigenetic feature in most eukaryotes (Zilberman et al., 2007). Recent work in allotetraploid *Arabidopsis suecica*, allohexaploid *Triticum aestivum* (bread wheat) and resynthesized allotetraploid *B. napus* has shown that DNA methylation may play a crucial role in establishing homologue expression patterns in allopolyploid plants (Chen et al., 2007; Shitsukawa et al., 2007; Gaeta., 2007). In contrast to animals, in which methylation occurs primarily at CG dinucleotides, cytosine DNA methylations in plants occur at sites containing CG, CNG (where N is any nucleotide) and CHH (an asymmetric site, where H is A, C or T) (Chan et al., 2005). Another important feature of plant cytosine DNA methylation is that a large number of genes (~33.3%) are methylated in their coding regions (Zhang et al., 2006). Since the coding regions of *BnGPAT4* homologues share high sequence similarity (93-97%), it was of interest to investigate the methylation patterns within the coding regions of the three homologues. We investigated the methylation status of the first exon region (approx 600bp) of each *BnGPAT4* homolog, using sodium bisulfite-treated genomic DNA isolated from a mixed collection of plant tissues and organs as a PCR template. Twelve clones of each PCR amplicon were sequenced. The ratios of the methylated cytosines (presented as the number of clones having methylated cytosines / the number of total clones) were indicated by percentages. As shown in Fig. 3.8A, all the locations of methylated cytosines were labeled with orange dots, and only those with

methylation ratios over 17% were listed in Fig. 3.8B. In general, the three homologues had very different methylation patterns.

Notably, *BnGPAT4-C2* had more methylated cytosines with ratios over 15% than the other two homologues. This may be related to the lower gene expression levels observed for the *BnGPAT4-C2* homologue (Fig. 3.7). These very different methylation patterns again indicated that the homologous genes have evolved and are regulated differentially at molecular level.

### **3.2.7 Phenotypic rescue of *Arabidopsis gpat4 gpat8* double mutant lines**

In *Arabidopsis*, *GPAT4* and *GPAT8* are closely related (>80% sequence similarity between the cDNAs) and display functional redundancy in cutin synthesis; double T-DNA insertional mutations at both genes result in severe defects in cutin synthesis and increased cuticle permeability (Li et al., 2007). A BLAST search of the full-length cDNA sequence of *AtGPAT8* against the *B. napus* EST database showed that all the positive hits (i.e., those with the highest scores, sequence similarities greater than 80%) were from the sequences of *BnGPAT4* homologues. Thus, it is likely that *AtGPAT8* emerged as a duplicated copy of *AtGPAT4* during *Arabidopsis* evolution after the divergence of *Arabidopsis* and the *Brassica* species.

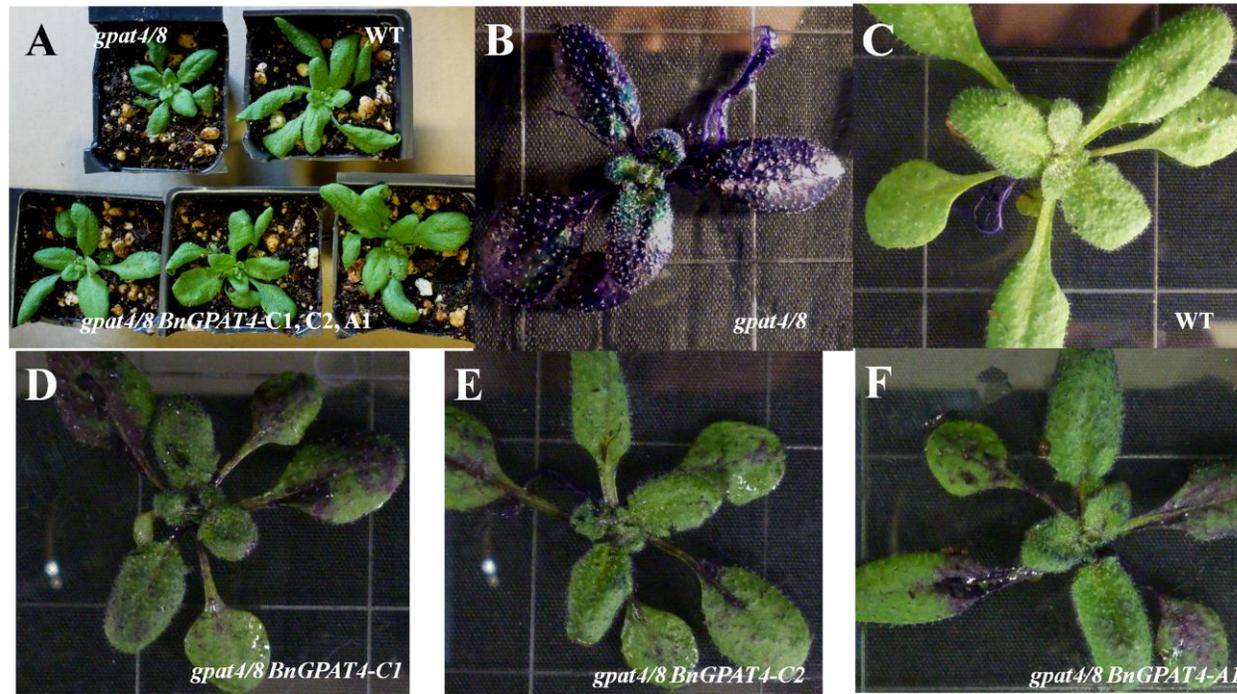


<i>BnGPAT4-C1</i>			<i>BnGPAT4-C2</i>			<i>BnGPAT4-A1</i>		
Location	Methylation ratio	Sequence	Location	Methylation ratio	Sequence	Location	Methylation ratio	Sequence
138	17%	<u>C</u> GC	7	17%	<u>C</u> CG	234	17%	<u>C</u> AT
141	25%	<u>C</u> AT	24	17%	<u>C</u> AG	264	17%	<u>C</u> GG
213	17%	<u>C</u> TT	178	17%	<u>C</u> TT	419	25%	<u>C</u> GT
257	17%	<u>C</u> CT	211	17%	<u>C</u> TC	472	33%	<u>C</u> CC
282	17%	<u>C</u> AT	246	17%	<u>C</u> AT	473	33%	<u>C</u> CA
342	17%	<u>C</u> AG	270	17%	<u>C</u> AA			
417	25%	<u>C</u> GT	300	17%	<u>C</u> GC			
542	17%	<u>C</u> CA	311	33%	<u>C</u> AC			
546	17%	<u>C</u> TT	326	17%	<u>C</u> GG			
			509	17%	<u>C</u> TG			
			572	17%	<u>C</u> GG			

**Figure 3.8. Cytosine DNA methylation patterns of the first exon regions of the three *BnGPAT4* homologues.**

(A) All methylated cytosine locations are indicated by orange color dots. (B) Summary of the methylated cytosines with methylation ratios over 17% in the three *BnGPAT4* homologues.

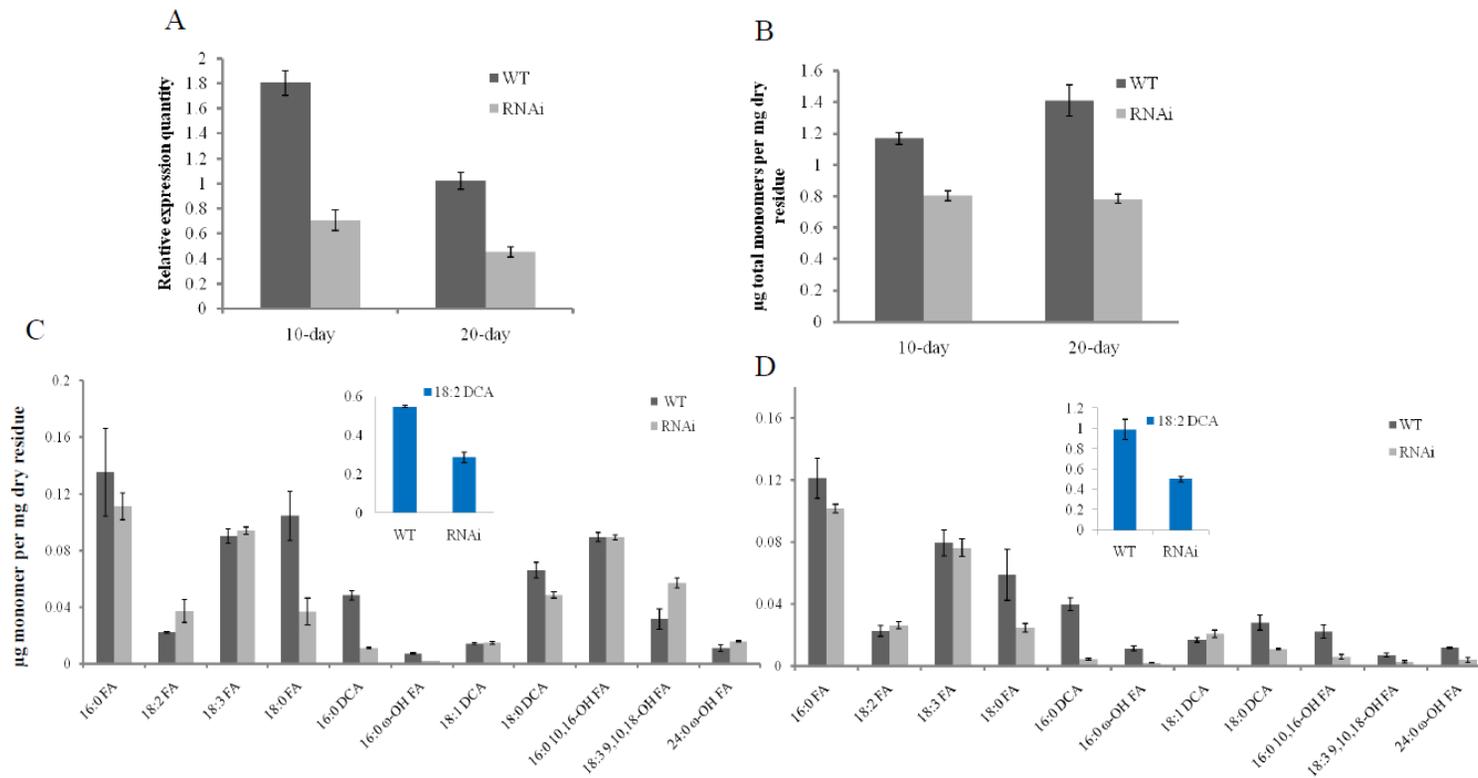
To determine whether the *BnGPAT4* homologues also function in epidermal cell lipid polymer biosynthesis, we expressed the three *BnGPAT4* cDNAs under the control of their native promoters in the *Arabidopsis gpat4 gpat8* double T-DNA lines (the schematic diagram of the binary vector is presented in Fig. A4). At least three independent T<sub>2</sub> transgenic lines for each construct were analyzed. In order to determine if the *BnGPAT4* homologues were able to rescue the defective cuticle phenotype of *Arabidopsis gpat4 gpat8* mutant plants, we examined the cuticle permeability of the transgenic *Arabidopsis* plants using a toluidine blue test. The toluidine blue test was developed by Tanaka et al (2004) and has been extensively used for the rapid visualization of cuticle defects of plant tissues. Varying degrees of phenotypic rescue were observed in the T<sub>2</sub> plants of *gpat4 gpat8* lines transformed with *BnGPAT4* homologues (Fig. 3.9). Segregation analysis indicated that the transgenic lines which had phenotypes closest to being fully restored had more than one copy of transgene; however, none of the transgenic lines displayed a complete restoration of the defective cuticle phenotype. These results suggest that *BnGPAT4* homologues are also involved in cutin synthesis.



**Figure 3.9. Toluidine blue test of the cuticle permeability of *gpat4 gpat8* *Arabidopsis* transformed with individual *BnGPAT4* homologues. (A) Plants before toluidine blue treatment. (B-F) Plants after toluidine blue treatment.**

### **3.2.8 *B. napus gpat4* RNAi lines exhibit alterations in cuticle load and stomatal structure resulting in increased water loss**

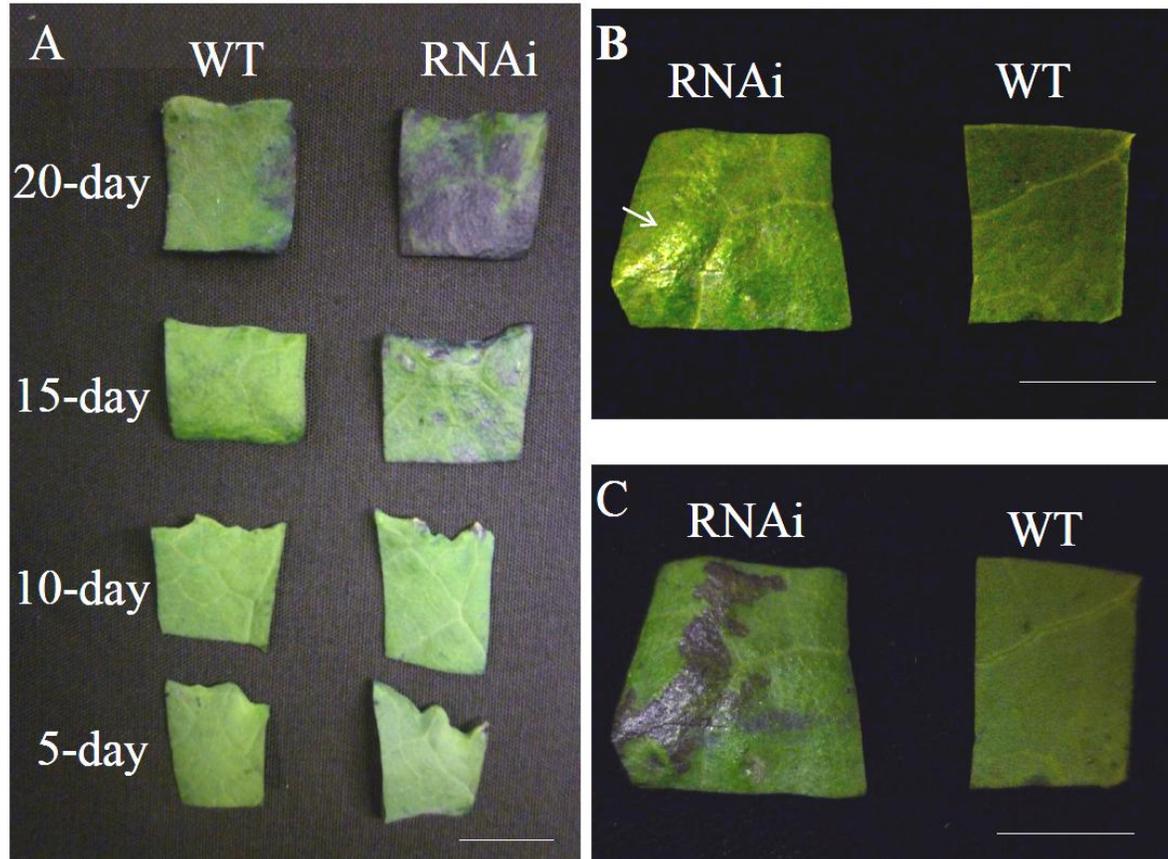
To further confirm the involvement of *BnGPAT4* homologues in cutin biosynthesis and explore other physiological functions of the *BnGPAT4* genes in *B. napus*, an RNAi approach targeting the expression of *BnGPAT4* homologues was used. Due to the high probability of functional overlap among the three homologues, silencing one particular homologue may not give a detectable phenotype; therefore the RNAi constructs (Fig. A5) were designed to silence all three *BnGPAT4* homologues. Two types of promoters, i.e., CaMV35S and napin, were used to direct the expression of RNAi construct at the vegetative and reproductive growth stages, respectively. The overall transcript abundance of the *BnGPAT4* homologues in the RNAi lines was determined using SYBR-green qRT-PCR (Fig. 3.10A). The overall expression levels of the *BnGPAT4* homologues were decreased by approximately 40% in both young and mature rosette leaves of the transgenic plants. To confirm that the RNAi silencing was specific to *GPAT4*, we also analyzed the expression level of *GPAT6* in the RNAi lines, because *GPAT6* shares the highest cDNA sequence similarity (approximately 60%) to *GPAT4* in the *GPAT* gene family. As shown in Fig. A6, the transcriptional levels of *GPAT6* gene were unchanged in both wild type and *gpat4*RNAi lines.



**Figure 3.10. Silencing of *GPAT4* homologues in *B. napus* results in decreased cutin monomer load.**

(A) Overall expression levels of *BnGPAT4* homologues in the rosette leaves of wild type and *gpat4* RNAi *B. napus* lines. Compared to wild type, the transcription levels of *GPAT4* were 43.3% and 35.8% lower in the rosette leaves at 10 and 20 days post emergence of the RNAi lines. n=3. (B) Total amount of monomers per mg dry residues. The rosette leaves of RNAi lines had 31.2% and 44.5% reductions in total cutin monomer load in comparison with the wild type lines at 10 and 20 days post emergence. n=3. (C-D) Cutin monomer profile of the rosette leaves at 10 (C) and 20 (D) days post emergence. FA, fatty acid. DCA, dicarboxylic acid. n=3.

In the CaMV35S promoter directed *gpat4* RNAi lines, young rosette leaves (5 and 10 days after emerging) did not exhibit visible morphological differences when compared to the wild type lines (Fig. 3.11A); however, in older leaves (~ 15 days post emergence), several small protruding areas with a glossy texture appeared on the adaxial leaf surface of RNAi lines (Fig. 3.11B), which exhibited increased cuticle permeability in the toluidine blue test (Fig. 3.11C). By the 20-day stage, increased cuticle permeability was observed over most of the leaf surface. Analysis of cutin monomers extracted from 10 and 20 day old leaves showed that the overall content of cutin aliphatic monomers was decreased by 31.2% and 44.5% in the 10 and 20-day old leaves of the *gpat4* RNAi lines, respectively (Fig. 3.10B). Interestingly, the cutin monomer compositions were quite different between 10 and 20- day old leaves (Fig. 3.10C and D), except the fact that the dominant cutin monomer was 18:2 (*cis*  $\Delta^{9,12}$ ) dicarboxylic acid (18:2 DCA) at both stages. In general, at the both stages, the monomer species exhibiting the greatest reductions were 18:2 DCA (~50%), 16:0 DCA (~80%) and 18:0 fatty acid (~58%).

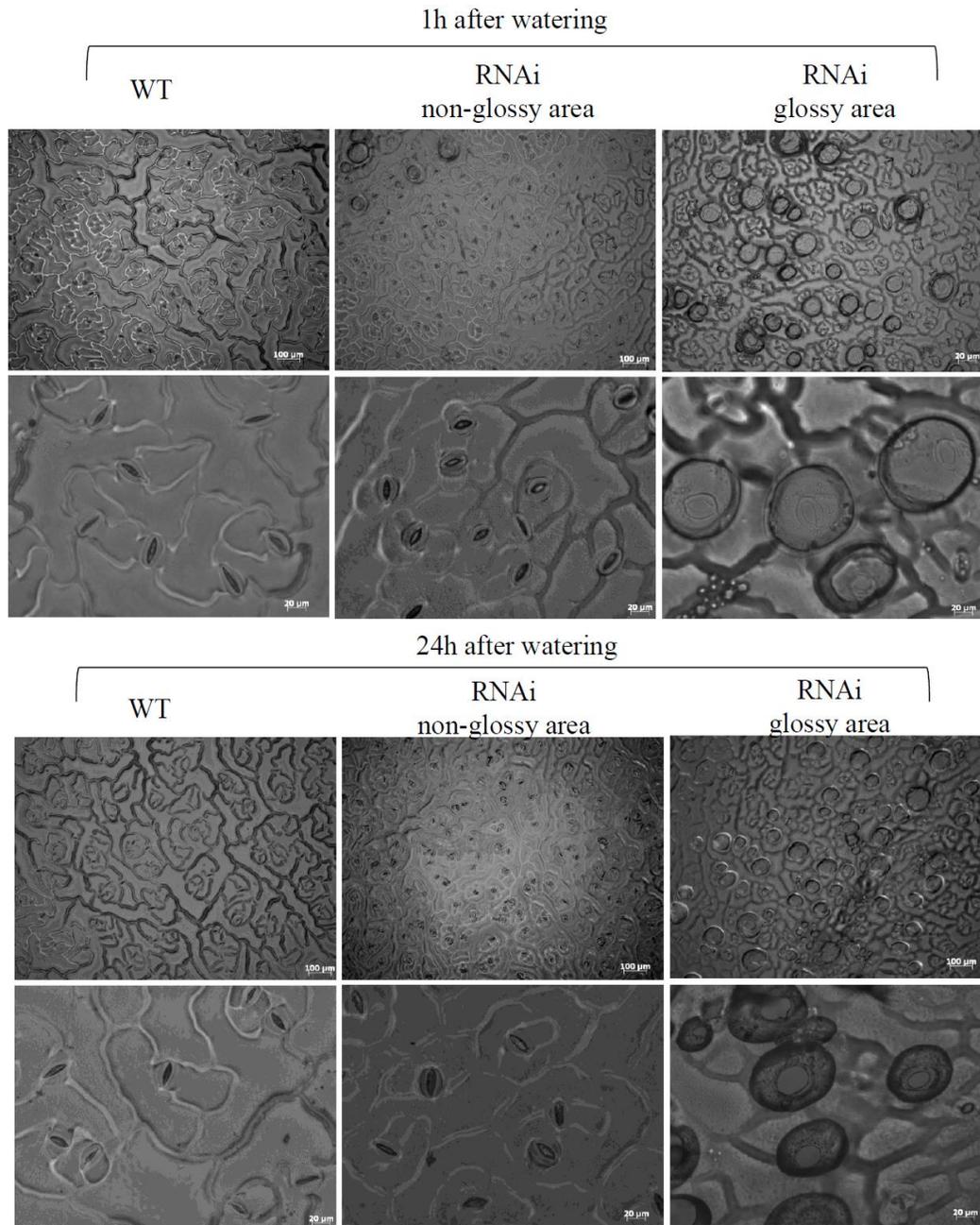


**Figure 3.11. Cuticle permeability of *B. napus gpat4* rosette leaves at different stages with the toluidine blue test.**

(A) Comparison of the rosette leaves of wide type and *gpat4* lines after treatment with the toluidine blue solution. (B) The rosette leaves at 15 days post emergence without toluidine blue treatment. The glossy area is indicated by an arrow. (C) The same leaf samples in (B) were treated with toluidine blue. Bar 1 cm.

Notably, even though there was a significant decrease (31.2%) in the overall level of cutin monomers in 10-day old rosette leaves, we did not observe any significant cuticle permeability defect using the toluidine blue test (Fig. 3.11A). One explanation could be that the younger staged leaves had sufficient cutin monomers per unit of leaf surface to maintain regular cuticle permeability, while at an older stage when the leaf size was much larger, the amount of cutin monomers per leaf surface unit was too low to maintain a normal level of cuticle permeability.

We also observed that the *gpat4 B. napus* lines exhibited increased water loss during the day time (Fig. A7). To further investigate the surface structure of *gpat4* RNAi lines exhibiting the glossy leaf phenotype, surface casts of the attached leaves were prepared at 1h and 24h after watering to observe the pavement cells. We observed that the stomata were widely open in these glossy areas of the rosette leaves of the RNAi plants (Fig. 3.12). In addition, the stomata opened much wider 24h after watering as compared to 1h after watering, and the surrounding guard cells were stretched much thinner than normal, which could possibly explain the observed increase of water loss in the RNAi lines. Our results suggested that an intact cuticle layer is essential for the guard cells to function properly and for maintaining a normal stomatal structure.



**Figure 3.12. Comparison of stomatal structures between wild type and *gpat4* *B. napus* lines.** The bubbles covering the widely opened stomata in the glossy area were likely caused by a high rate of water vapor release when the leaf surface was sealed. The same phenotype was observed in four RNAi lines.

### **3.2.9 *B. napus gpat4* RNAi exhibited abnormal inflorescence and severely reduced seed yield**

Different from the defective lipid polyester biosynthesis observed in the rosette leaves of the CaMV35S promoter directed RNAi lines, the RNAi lines under the direction of the napin promoter exhibited severely affected inflorescence development (Fig. 3.13). The development of floral buds was partially aborted, particularly on the lower portion of the inflorescence (Fig. 3.13B). Additionally, the development of axillary inflorescence primordia was also severely affected (Fig. 3.13C-D). Consequently, in comparison to the wild type plant, the number of open flowers in the *gpat4* line was much lower.

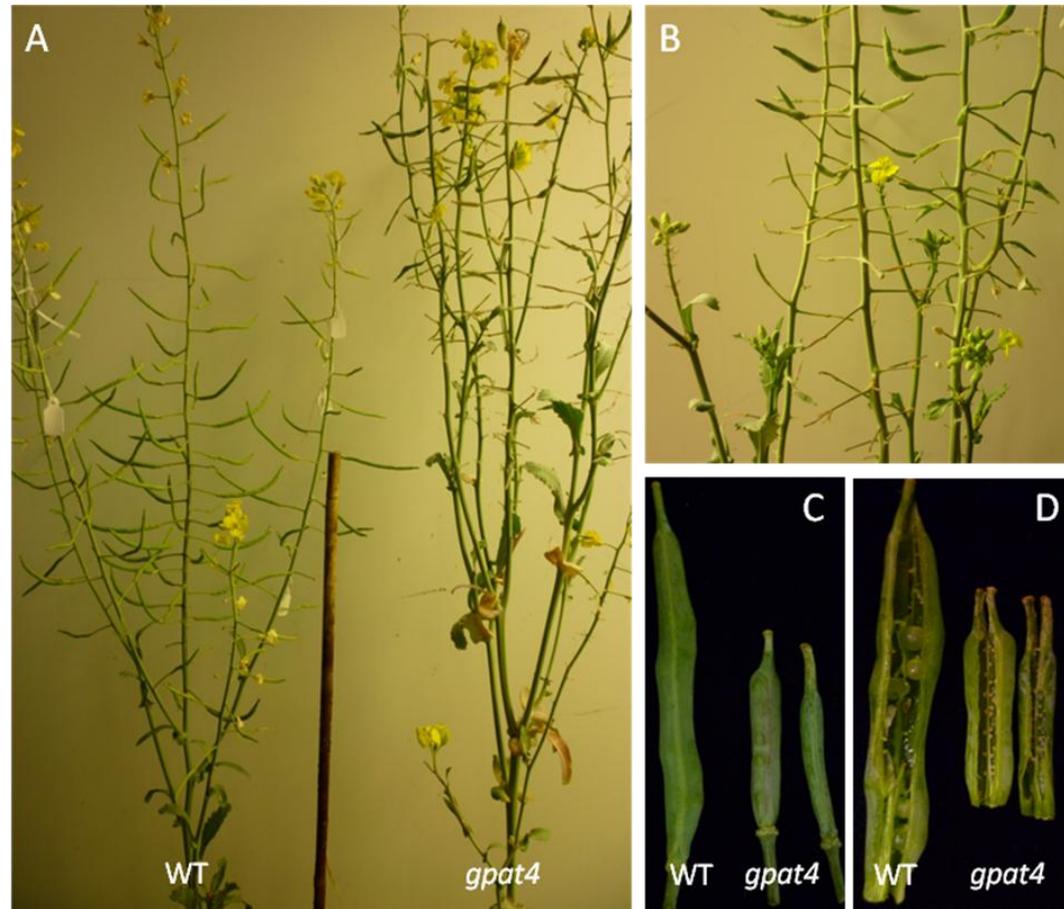
Although the individual open flowers of the *gpat4* lines did not exhibit any morphological difference in comparison to the wild type *B. napus* plants, the *gpat4* flowers did not generate normal siliques after self-pollination (Fig. 3.14). In several independent *gpat4* lines, most of the siliques did not contain seeds (Fig. 3.14C-D). The remaining siliques typically had less than 10 seeds per silique instead of more than 20 seeds normally found in a wild type silique.

To investigate the causes for the reduced seed yield in *gpat4* lines, pollen grains of several independent RNAi lines were examined under the microscope. As shown in Fig. 15C-D, some of the pollen grains of the *gpat4* lines were deformed and collapsed. This phenotype was observed with all *gpat4* RNAi lines examined. It is likely that these abnormal pollen grains account for the aborted seed development in the *gpat4* lines.



**Figure 3.13. Abnormal inflorescence development in *gpat4* RNAi lines.**

(A) Inflorescence comparison between a wild type plant and a *gpat4* RNAi line. (B) Close-up view of the *gpat4* developing inflorescence. The development of the floral buds on the lower portion of the inflorescence was aborted in the *gpat4* lines. (C) Close-up view of *gpat4* axillary inflorescence primordia. (D) Close-up view of aborted flower buds in later developmental stage of the axillary inflorescence. Similar phenotype was observed in all *gpat4* RNAi lines.



**Figure 3.14. Aborted seed development in *gpat4* RNAi lines.**

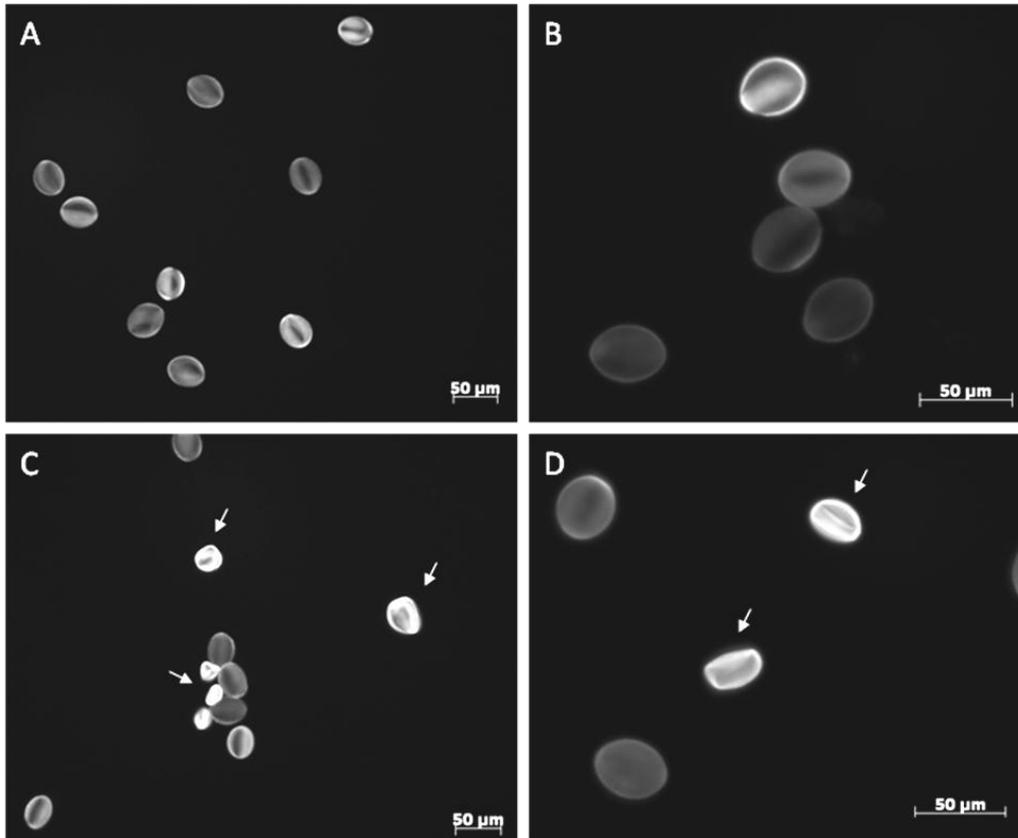
(A) Comparison between a wild type plant and a *gpat4* line during the silique developmental stage. Reduced number of siliques was observed in the *gpat4* lines. (B) Close-up view of the developing siliques of the *gpat4* line. (C-D) Before and after dissecting the developing siliques of wild type and *gpat4* lines. Most of the *gpat4* siliques had no or very few seeds.

To further confirm that the reduced seed yield in *gpat4* lines was caused by the defective pollen grains, reciprocal crosses were performed between *gpat4* and wild-type plants. The seed yield was fully restored when *gpat4* plants were pollinated with wild type pollen (Fig. 3.16B-E). On the other hand, pollination of wild-type plants with the *gpat4* pollen produced siliques with normal seed yield. Based on our current results, it is possible that not only the pollen but the stigma of *gpat4 B. napus* is also defective. In addition, it is likely that the defective changes of the pollen and (or) stigma in *B. napus* further affected the pollen-stigma interaction, which is important for a successful fertilization (Chapman and Goring, 2010). Further experiments need to be performed to confirm this hypothesis.

### **3.2.10 Preliminary results indicated fatty acid composition changes in the *gpat4 B. napus* seeds**

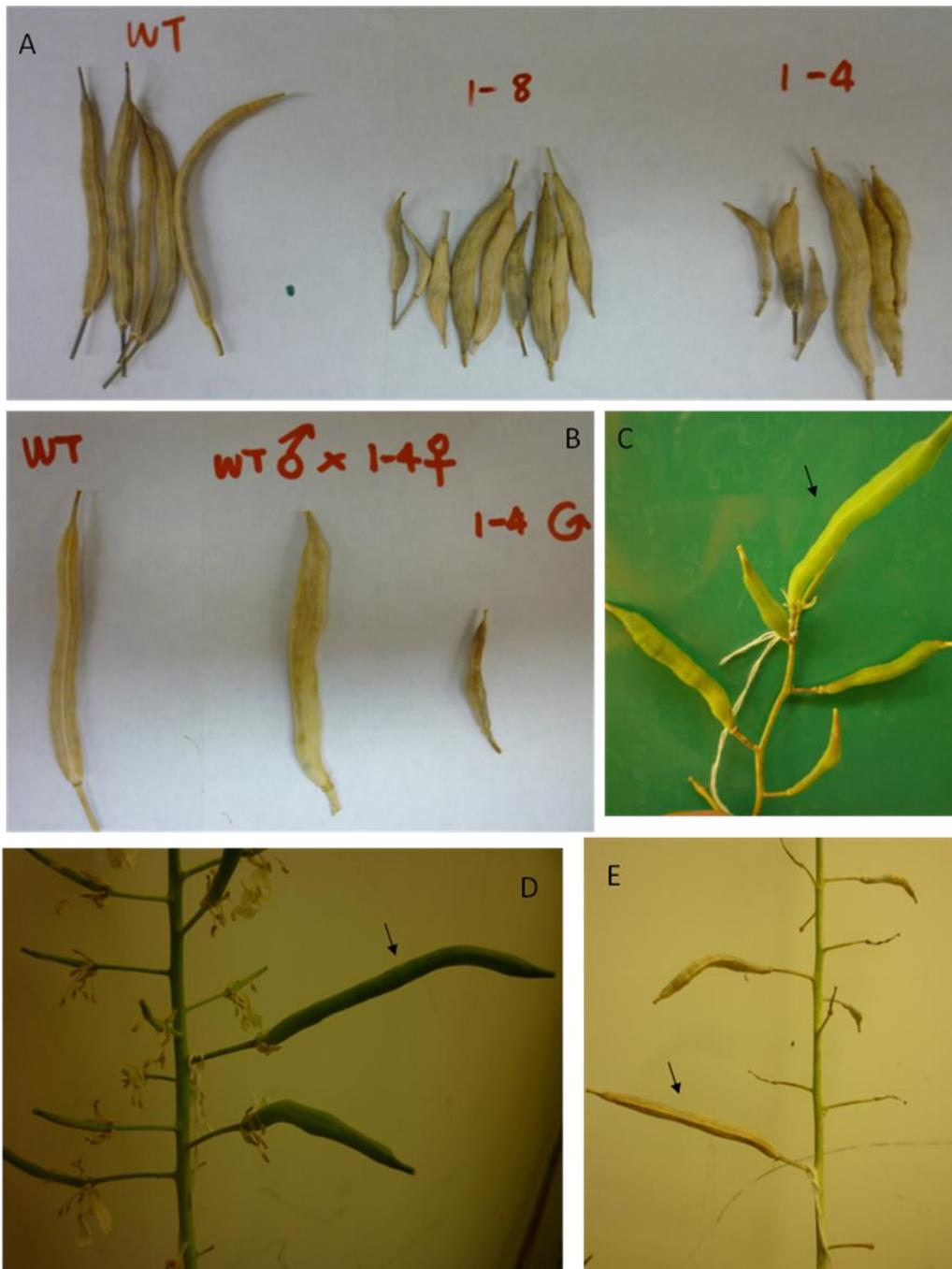
We analyzed the fatty acid composition of the seed oil from *gpat4* T<sub>1</sub> lines. Due to the reduced seed yield of the *gpat4 B. napus* lines, we were only able to analyze a small number of the *gpat4* T<sub>1</sub> seeds. Up to 10 seeds from each of the three independent *gpat4* lines and two wild type *B. napus* plants were selected for fatty acid composition analysis. The T<sub>1</sub> seeds from *gpat4* lines should have different genotypes caused by segregation, i.e., wild type, *gpat4* heterozygous and homozygous (with different copy numbers). Thus, instead of analyzing the fatty acid composition of the total seeds, we analyzed the fatty acid compositions of individual seeds. As shown in Fig. 3.17, in comparison to the wild type seeds, a portion of the seeds from the three *gpat4* lines had decreased 18:1 *cis*  $\Delta^9$  content

and increased content of 18:2 *cis*  $\Delta^{9,12}$  and 18:3 *cis*  $\Delta^{9,12,15}$  fatty acids in the storage lipid. The suppression of *GPAT4* genes in *B. napus* appeared to affect the incorporation of different 18 carbon unsaturated fatty acyl species into the TAG biosynthesis, nevertheless, statistically reliable analysis of the seed oil of the T2 and T3 generation needs to be performed to further explicate the role of *GPAT4* in TAG biosynthesis.



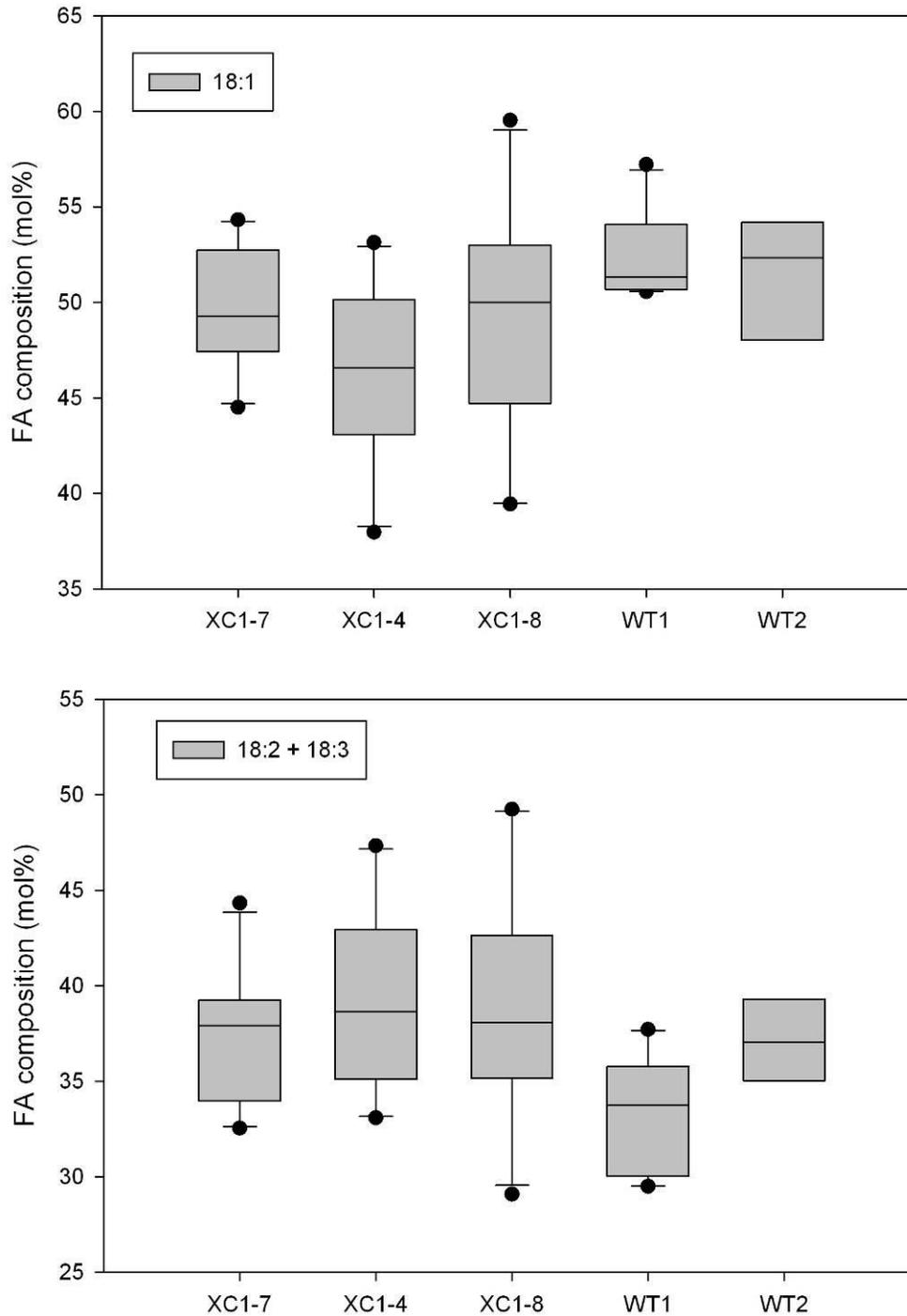
**Figure 3.15. Abnormal pollen grains of the *gpat4* lines.**

(A-B) Pollen grains from wild type plants. (C-D) Pollen grains from *gpat4* lines.



**Figure 3.16. Reduced seed yield of *gpat4* *B. napus*.**

(A) Comparison between the mature siliques from wild type (WT) and two *gpat4* lines (1-8 and 1-4). (B) Mature siliques from wild type, *gpat4* pollinated with wild type pollen, and *gpat4*. (C-E) Maturing siliques of *gpat4* pollinated with wild type pollen (as indicated with arrow).



**Figure 3.17. Fatty acid composition analysis indicated decreased content of 18:1 and increased content of 18:2 and 18:3 in the *gpat4* *B. napus* T<sub>1</sub> seed oil.**

Box plots were used to represent the 18:1 and 18:2+18:3 composition (mol%) of the individual seeds from each *B. napus* plant. (A) The 18:1 content and (B) the 18:2 and 18:3 content of three independent *gpat4* lines (XC1-7, XC1-4, and XC1-8) and two wild type plants (WT1 and WT2).

### 3.3 Discussion

#### 3.3.1 Genome origins and evolutionary fates of *BnGPAT4* homologues

In the present study, we identified three *GPAT4* homologues in *B. napus* and were able to determine the genome origin of each homologue based on their sequence identity with *GPAT4* genes from *B. rapa* and *B. oleracea*. Two *BnGPAT4* homologues showed greatest sequence similarity with *GPAT4* sequences from *B. oleracea* and are therefore believed to originate from the C-genome, while only one A-genome-derived *BnGPAT4* homologue (showing similarity to *GPAT4* from *B. rapa*) was identified. The recent sequencing of *B. rapa* (<http://brassicadb.org:8081/brad/index.php>) uncovered two *GPAT4* UniGene sequences (accession No. Bra033249 and Bra032643), which were highly similar (>99%) to the two *GPAT4* genes we isolated from *B. rapa*. Given the complex genetic background and incomplete genome sequence information of *B. napus*, we are unable to completely exclude the possible existence of other as yet unidentified *BnGPAT4* homologues (that may exist as a pseudogene or a functional gene). Genome sequencing of *B. napus* and *B. oleracea* is still ongoing; thus our identification of homologous *GPAT4* genes in *B. napus* is limited by the incomplete genome/EST sequence database. Nevertheless, our current results represent the most up-to-date information on expressed and functional *BnGPAT4* homologues.

The evolutionary fates of genes after polyploidy and genomic segmental duplications include loss or silencing, maintaining the ancestral function and functional divergence, which includes neofunctionalization and

subfunctionalization (Wendel 2000; Whittle et al., 2009). Neofunctionalization could result in the acquisition of new catalytic/ structural function or new expression domain; while subfunctionalization could result in the expression domain or the catalytic/structural functions of an ancestral gene being divided and shared by the existing duplicated genes (Force et al., 1999; Flagel and Wendel, 2009; Liu and Adams, 2010). In the case of *BnGPAT4* homologues, to answer the question of whether the current three homologues partition the ancestral expression domain or gain certain new expression domains, we need to examine the expression patterns of the *GPAT4* genes not only from the *Brassica* species, but also from out group plant species that diverged prior to the polyploidy event of the Brassicaceae family. Therefore, with our current results, because the expression domain of the ancestral GPAT4 gene is unclear, the different expression patterns of the *BnGPAT4* homologues could be the result of either subfunctionalization or neofunctionalization.

In addition, the different tissue/organ-specific expression patterns indicated that the *BnGPAT4* homologues are developmentally regulated. For example, in developing seeds, *BnGPAT4-A1* was expressed dominantly in maturing embryos, but was low in developing seed coats; in contrast, *BnGPAT4-C1* was expressed at low levels in the embryo, but at high levels in the seed coats (Fig. 3.7B). A previous study with allotetraploid cotton also discovered such developmental regulation (Adams et al., 2003). In that study, several pairs of homologous genes exhibited developmentally regulated reciprocal silencing, whereby one homologue was expressed in certain organs, but was low in others,

and its counterpart was expressed in a complementary way. This developmental regulation was predicted to be a frequent consequence of polyploidy (Adams et al., 2003). The possible mechanisms regulating the expression of homologous genes could include changes in DNA sequence, epigenetic modifications, *cis*- and *trans*-acting effects and RNA-mediated pathways (Chen, 2007). These mechanisms are also believed to be important in facilitating the adaptation of polyploid crops during the evolution and domestication processes (Chen, 2007).

### **3.3.2 Possible mechanisms involved in regulating the level of BnGPAT4 polypeptide accumulation in yeast**

In the present study, the three forms of BnGPAT4 proteins exhibited different levels of accumulation in the ER membrane of transformed yeast. We demonstrated that the last 5-7 amino acids at the C-termini of the BnGPAT4 proteins were important for regulating the level of polypeptide accumulation, but the amino acid sequence differences between the C-termini were not related to their differential accumulation in yeast. Protein abundance is determined by the opposing processes of synthesis and degradation; it remains unclear which of these processes, however, account for the observed differences in BnGPAT4 protein accumulation in transformed yeast. One possible mechanism behind our observations could be that yeast has different codon usage efficiencies for the coding sequences of the three *BnGPAT4* homologues. With this assumption, the BnGPAT4 proteins would be synthesized at different rates in the transformed yeast and thus exhibit different levels of accumulation. Another possible mechanism could be that the three forms of BnGPAT4 proteins are differentially

degraded. The three forms of BnGPAT4 proteins have a few amino acid differences (7 different amino acids between BnGPAT4-C1 and C2, Fig. 3.4B) in the remainder of their sequences, in addition to the differences at the C-termini. These differences might impart different half-lives to the BnGPAT4 polypeptides. A previous study comparing ER-bound fatty acid desaturases (Fad3) from tung (*Vernicia fordii*) and *B. napus* indicated that the BnFad3 contained a degradation signal in its N-terminus (approximately 60 amino acids), which was responsible for its higher protein turnover rate than the tung Fad3 (O'Quin et al., 2010). In the case of BnGPAT4, there are very minor differences at the N-termini (Fig. 3.4B), which are less likely to cause different protein half-lives. It also remains to be determined whether the observations on different levels of polypeptide accumulation for the BnGPAT4s produced in the yeast expression system reflect the *in planta* situation.

### **3.3.3 BnGPAT4 may be involved in suberin biosynthesis in root and seed coat**

Cutin and suberin are two types of insoluble fatty acid- and glycerol-based lipid polymers that are deposited in the cell walls of different plant tissues. They are different from each other based on their monomer compositions and deposition locations. In general, suberin deposits in the seed coat and root, whereas cutin is found mostly at the cutical layer that covers the aerial epidermis (Beisson et al., 2007). An earlier study in *Arabidopsis* revealed that GPAT5, another member of the ER-bound GPAT family, was involved in suberin synthesis in seed coat and root (Beisson et al., 2007). In the present study, we

found that all three *BnGPAT4* homologues were expressed (at different levels) in seed coat during the period from 20 to 30 DAP (Fig. 3.7), which corresponds to the timing of suberin deposition in the seed coat (Molina et al., 2008). In addition, the promoter-GUS assay also indicated the GUS gene was expressed in the periderm and endodermis parts of the root (Fig. 3.6I), where suberin is normally deposited to strengthen the cell wall and contributes to the control of water movement (Enstone et al., 2003). We therefore suggest that the *BnGPAT4* homologues (particularly *BnGPAT4-C1*, which exhibited the highest expression levels) are very likely to be involved in suberin synthesis.

#### **3.3.4. The cuticle layer is closely related to stomatal development and structure**

The stomata are tiny pores in leaves that regulate the exchange of gas for photosynthesis and the release of water vapor by transpiration. The guard cells surrounding the stomata control the opening and closing of the stomata via turgor-related shape changes (Kappen and Haeger, 1991). A number of environmental and internal factors, such as humidity, light, temperature, CO<sub>2</sub>, hormones and ion content, interact with the guard cells through a series of metabolic processes to control the stomatal apparatus (Zeiger, 1983).

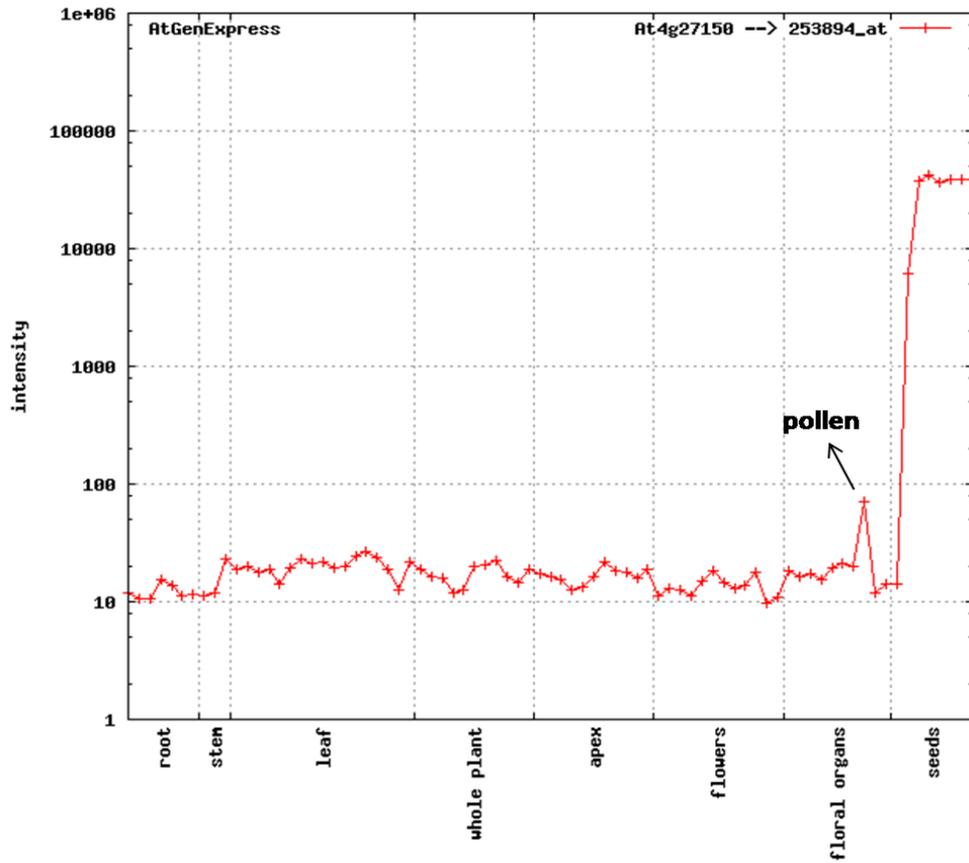
As the first line of contact with the environment, the cuticle is mainly composed of wax, cutin and polysaccharides (Bird and Gray, 2003). In addition to its function as a protective barrier, the cuticle has been shown to be important for stomatal development. It has been shown that defective cuticular wax biosynthesis affects stomatal development with altered stomatal densities in

*Arabidopsis* (Zeiger and Stebbins, 1972; Gray et al., 2000; Chen et al., 2003). The mechanism behind this was proposed to be that the environmental effects on stomatal development are significantly affected by changes in cuticular wax (Bird and Gray, 2003). In comparison to wax, cutin is another important lipid-derived component deposited in the cuticle layer. In the present study, instead of altered stomatal densities, we observed that the stomata were widely open in the glossy part of the mature rosette leaves of *gpat4* RNAi lines. This part of the leaf surface has been shown to exhibit a severe cuticle defect via the toluidine blue test. It is obvious that a cutin-defective cuticle surface has led to abnormal stomatal morphology. Li et al (2007) observed defective cuticular ledges on the guard cells in *Arabidopsis gpat4 gpat8* line. A previous study on *Vicia faba* leaves showed that stomatal opening could be caused by the reduction in epidermal turgor when the leaf was perturbed with a stream of dry air (Mott and Franks, 2001). In the case of *gpat4 B. napus*, it is possible that the defective cuticle layer altered the micro-environmental condition (i.e., lower humidity) for the epidermis, changing the turgor of the epidermal cells surrounding the guard cells and in turn caused the stomata to open. Nevertheless, given the complexity of the mechanisms involved in regulating the stomatal pores (Schroeder et al., 2001), a more comprehensive interpretation of how the defective cuticle layer affects stomatal aperture requires further study.

### **3.3.5. Napin promoter-directed RNAi silencing of *GPAT4* in the inflorescence primordia and pollen grain in *B.napus***

In the CaMV35S promoter-directed RNAi lines (10 independent lines), the major phenotype is the cuticle defect on the epidermis; however, there are two lines exhibiting abnormal inflorescence development and severely reduced seed yield. In the napin promoter-directed RNAi lines (12 independent lines), all the plants exhibited abnormal inflorescence and pollen development and severely reduced seed yield. It appears that the napin promoter-directed RNAi construct is more effective in down-regulating the expression of *GPAT4* in the reproductive organs of *B. napus*. The napin promoter used in the present study is originated from a *napA* gene (a member of the *napin* gene family, GenBank accession: J02798), which encodes a 1.7 S seed storage protein in *B. napus* (Josefsson et al., 1987). The expression patterns of the *napin* gene family or the corresponding promoters were mainly investigated within the developing seeds (Blundy et al., 1991; Kridl et al., 1991; Ellerstrom et al., 1996), thus, it is not clear whether *napin* genes could also be expressed in other plant organs. We further investigated the gene expression pattern of the corresponding orthologous *Arabidopsis* gene (At4g27150), which encodes a seed storage albumin 2 protein. As shown in Fig. 3.18, although At4g27150 gene is predominately expressed in developing seeds, it also exhibited a reasonable expression level in pollen. Nevertheless, this microarray data does not include the transcript quantity of At4g27150 in inflorescence primordia. Thus, it is likely that the napin promoter-directed RNAi construct is active during the pollen development. In addition, given the fact that

gene silencing regulated by RNAi can spread locally and systemically (Klahre et al., 2002; Kusaba, 2004), it is also possible that the napin promoter-directed RNAi silencing is unspecificly targeted to pollen and inflorescence primordia.



**Figure 3.18.** Gene expression pattern of the *napin* orthologous gene, At4g27150, in *Arabidopsis*.

At4g27150 is expressed predominately in developing seeds. It is also expressed in pollen at a lower level.

Data source: AtGenExpress Visualization Tool (Schmid et al., 2005)

### **3.3.6. *GPAT4* plays important roles in regulating the development of inflorescence and pollen grain in *B.napus***

Our earlier data on *BnGPAT4* gene expression patterns show that the *BnGPAT4-C1* homologue is highly expressed in the inflorescence primordia (Fig. 3.7A), which suggests that the *BnGPAT4-C1* may play an important role in the early flower development. The defective inflorescence primordia development in the *gpat4* lines further confirmed our hypothesis and strongly suggested that *GPAT4* is essential for the inflorescence development of *B. napus*.

The *B. napus gpat4* lines also produce abnormal pollen grains. The development of pollen is a complex process, which is known to be controlled coordinately by both the sporophytic and gametophytic tissues in the stamen (Ursin et al., 1989; Zheng et al., 2003). For example, the sporophytic tapetum, which is a layer of nutritive cells surrounding the developing microspores, undergoes a precise degeneration process to secrete lipids, proteins, and other nutrients to support the pollen development. Disruptions in either the tapetal development or the gametophytic microspores development can affect the final male fertility of the pollen grains (McCormick, 2004; Wang et al., 2008). The process of pollen development is also known to be closely connected to lipid metabolism within the tapetum and pollen grain itself (Zheng et al., 2003; Wang et al., 2008; Beaudoin et al., 2009). A previous study in *Arabidopsis* revealed that another GPAT family member, AtGPAT1, a mitochondrial membrane-bound isoform, is important for the degeneration process of tapetum and the formation of lipid bodies in pollen grains. The *gpat1 Arabidopsis* exhibited a similar phenotype

to that of the *gpat4 B. napus* in producing deformed pollen grains and having reduced seed set. In the case of BnGPAT4, further experiments need to be done to investigate the development of both tapetum and microspores to reveal the physiological role of BnGPAT4 to pollen development.

### **3.3.7 GPAT4 may be involved in the storage lipid biosynthesis**

The ER-bound GPAT was originally considered to be associated with the plant storage lipid biosynthesis for catalyzing the acylation of G3P at the *sn*-1 position in the Kennedy pathway (Weselake, 2005). Detailed studies of the physiological roles of ER-bound GPATs were only performed recently after the first ER-bound GPAT family was identified in *Arabidopsis* (Zheng et al., 2003). The studies in *Arabidopsis* revealed the essential roles of the ER-bound GPATs in extracellular lipid polyester biosynthesis (Beisson et al., 2007; Li et al., 2007; Li-Beisson et al., 2009). Nevertheless, there is little experimental data, so far, to demonstrate that the currently identified ER-GPATs are involved in the storage lipid biosynthesis.

Our preliminary fatty acid composition analysis of the *gpat4* seed oil indicated that the suppression of *GPAT4* may have resulted in an increased total amount of 18:2 and 18:3 and a decreased content of 18:1 (Fig. 3.17). It seems that more polyunsaturated fatty acids (PUFAs) were incorporated into the storage lipid with the suppression of GPAT4 homologues. A recent study of storage lipid biosynthesis in developing soybean embryos indicated that about 60% of the nascent fatty acyl moieties are directly incorporated into the *sn*-2 position of phosphatidylcholine (PC) through an acyl-remodeling process rather than being

sequentially esterified to the glycerol backbone to form triacylglycerol (see Fig. 2.4 in Chapter 2 Literature Review) (Bates et al., 2009). Notably, PUFAs for storage lipid biosynthesis are mainly formed on PC (Abbadi et al., 2004). Thus, it is possible that GPAT4 is involved in the storage lipid biosynthesis by catalyzing the incorporation of a small portion of the nascent fatty acyl moieties (such as 18:1) onto the G3P backbones; while the PC-mediated pathway channels the majority of nascent fatty acids into acyl-editing process. Therefore, it is possible that the suppression of *GPAT4* resulted in channeling of more nascent fatty acids into the acyl-editing process and hence resulted in an increased content of PUFAs in the seed oil. Further detailed seed oil analysis of the T2 and T3 *gpat4* RNAi generations will provide more information for understanding the role of GPAT4 in TAG biosynthesis.

### **3.4 Conclusion**

We have characterized three *GPAT4* homologous genes in the allotetraploid *B. napus* focusing primarily on their genomic origins and functional divergence. All three *BnGPAT4* genes were confirmed to encode functional GPAT enzymes but with different levels of polypeptide accumulation when expressed in yeast. The gene expression patterns, epigenetic variations and phenotypic rescue of the *gpat4 gpat8 Arabidopsis* double mutant indicated that the three *BnGPAT4* genes have evolved through functional divergence. Suppression of *GPAT4* expression in *B. napus* further revealed the important roles of BnGPAT4 in cutin biosynthesis, stomatal function, inflorescence development, male fertility, and seed oil biosynthesis.

## **3.5 Methods**

### **3.5.1 Plant Materials**

*Brassica napus* double haploid line (DH12075) plants were grown in a greenhouse under 16 h day / 8 h night at 23°C. *Arabidopsis thaliana* (ecotype Columbia) plants were grown in a growth chamber at 23°C under a photoperiod of 18h. Seeds of homozygous double T-DNA mutant *Arabidopsis gpat4 gpat8* plants were obtained from Dr. Yonghua Li-Beisson (CNRS, Bordeaux, France).

### **3.5.2 Cloning of *GPAT4* genes from *B. napus***

The *Arabidopsis GPAT4* (At1g01610) cDNA sequence was used to query the *B. napus* EST database in GenBank (NCBI) using the megablast program. By analyzing the alignment of the positive hits, we identified three putative *GPAT4* homologues in *B. napus*. Based on single nucleotide polymorphisms in the cDNA sequences, homologue-specific primer pairs (Table A1) were designed to amplify the full-length cDNA and genomic DNA sequences. For cDNA amplification, total RNA prepared from young *B. napus* seedlings was used in RT-PCR. The promoter regions of the *BnGPAT4* homologue genes were cloned using the Universal Genome Walker kit (Clontech Laboratories, Palo Alto, CA, USA) according to the manufacturer's instructions. The corresponding primer pairs are listed in Table A1. All the generated PCR amplicons were subcloned into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA, USA) and sequenced.

### **3.5.3 Heterologous Expression in Yeast and *in vitro* GPAT enzyme assay**

The cDNAs of three *BnGPAT4* homologues were subcloned into the yeast expression vector pYES2.1/V5-His-TOPO (Invitrogen, Carlsbad, CA, USA) and

sequenced to confirm the PCR fidelity. The PCR primers used are listed in Table A1. The plasmids were then transformed into yeast (*Saccharomyces cerevisiae*) *GPAT* mutant strain *gat1* $\Delta$  (Zheng et al., 2003). Galactose-induced expression of the *BnGPAT4* cDNAs in yeast was performed according to the manual of pYES2.1/V5-His-TOPO TA expression kit (Invitrogen, Carlsbad, CA, USA). *LacZ*-transformed *gat1* $\Delta$  yeast strain was used as negative control. Yeast homogenates were prepared with glass beads in a lysate buffer as described previously (Zheng and Zou, 2001). The crude homogenates were first centrifuged at  $2500 \times g$  at  $4^{\circ}\text{C}$  for 10 min to pellet the cell debris; the supernatant was further centrifuged at  $100,000 \times g$  at  $4^{\circ}\text{C}$  for 1 h to pellet the microsomal fractions, which were then re-suspended in lysate buffer for enzyme assay. Protein concentration was determined using the Bradford assay (Bio Rad, Hercules, CA, USA) with bovine serum albumin as a standard.

The GPAT enzyme assay was performed at  $30^{\circ}\text{C}$  for 10 min in  $50 \mu\text{L}$  reaction mixture containing 40 mM Hepes (pH 7.0), 5 mM EDTA, 1 mM DTT, 2.5 mg/mL BSA,  $100 \mu\text{M}$  [ $^{14}\text{C}$  (U)]-glycerol-3-phosphate (30 Ci/mol; GE Healthcare),  $20 \mu\text{M}$  oleoyl-CoA (18:1-CoA), and  $40 \mu\text{g}$  of yeast crude microsomal protein. The reaction was quenched with 2 mL of chloroform/methanol (1:2, v/v) and 1 mL of 1 M KCl in 0.2 M  $\text{H}_3\text{PO}_4$ . The lower organic phase was extracted, dried under nitrogen, re-suspended in  $70 \mu\text{L}$  chloroform, and applied to a pre-coated thin layer chromatography plate (TLC plates, layer: 0.25 mm, SIL G-25, DC-Fertigplatten), which was subsequently developed in a solvent system containing chloroform/methanol/acetic acid/5%

aqueous sodium bisulfite (75:30:9:3, v/v). The radio-labelled final products ( $[^{14}\text{C}]$ -labeled glycerolipids) on the TLC plates were scraped and subjected to scintillation counting.

#### **3.5.4 Western blot**

The microsomal fractions (40  $\mu\text{g}$  protein) from yeast expressing *BnGPAT4* or *LacZ* were separated by SDS-PAGE and electrotransferred onto PVDF membranes following the manufacturer's instructions (Amersham, Arlington Heights, IL, USA). The membranes were then incubated in a blocking buffer (GE Healthcare, Piscataway, NJ, USA) for 1 h and followed by incubation with horseradish peroxidase (HRP) conjugated anti-His antibodies (Invitrogen, Carlsbad, CA, USA) at 1:10,000 dilution for 1 h. HRP activity was visualized by chemiluminescence using the ECL Advance Kit (GE Healthcare, Piscataway, NJ, USA).

#### **3.5.5 Promoter-GUS fusion analysis**

The upstream sequences of the first ATG in *BnGPAT4* homoeologs (615 bp for *BnGPAT4-C1*, 640 bp for *BnGPAT4-C2*, 656 bp for *BnGPAT4-A*) were cloned as a *HindIII-XbaI* fragment into the binary vector pBI121, in which the original GUS reporter gene was replaced with a *GUS-eGFP* bifunctional reporter gene fusion (Thilmony et al., 2006). The sequences of the primers used for amplification of the promoter sequences are listed in Table A1. The three constructs (*BnGPAT4-C1<sub>Pro</sub>:GUS-eGFP*, *BnGPAT4-C2<sub>Pro</sub>:GUS-eGFP*, *BnGPAT4-A1<sub>Pro</sub>:GUS-eGFP*) were introduced into wild type *Arabidopsis* plants by *Agrobacterium tumefaciens* mediated floral dip (Clough and Bent, 1998). The

T<sub>2</sub> and T<sub>3</sub> progeny from several individual transgenic plants were analyzed for GUS gene expression. Fresh tissues were immersed, vacuum-infiltrated, and incubated at 37°C overnight in a staining solution (10 mM sodium phosphate, pH 7.0, 10 mM EDTA, 0.5 to 2 mM potassium ferricyanide/potassium ferrocyanide, 1 mg/mL X-Gluc A, and 0.1% Triton X-100). Stained tissues were soaked in ethanol-acetic acid (3:1, v/v) to clear the chlorophyll and fixed in a mixture of ethanol/acetic acid/37% formaldehyde/H<sub>2</sub>O (50:5:27:18, v/v). Images of whole mount tissues were taken by a Wild M8 dissecting microscope and Zeiss Scope A1 microscope with digital cameras.

### **3.5.6 TaqMan qRT-PCR**

Total RNA was extracted from different *B. napus* samples at various developmental stages using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA), except for extraction of RNA from seed coat and root samples, where the CTAB method (Gambino et al., 2008) and TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) were used. First-strand cDNA synthesis was performed in a 20 µL reaction mixture with 1 µg of DNase (Qiagen, Valencia, CA, USA) -digested total RNA using the QuantiTect reverse transcription kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The 20 µL final cDNA product was then diluted to 500 µL, and 8.5 µL was used in each qRT-PCR reaction. Real-time PCR reactions were performed in a Fast Optical 96-Well Reaction Plate on ABI PRISM 7900 HT Real-Time PCR System (both from Applied Biosystems, Foster City, CA, USA). The 25 µL reactions contained 1× TaqMan Universal PCR Master Mix (Applied Biosystems, Foster City, CA, USA), 360 nM primers, 320

nM probes, and 8.5  $\mu$ L of cDNA. The following standard thermal profile was used: 50°C for 2 min, 95°C for 5 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min. Three technical replicate reactions were performed with each cDNA sample and individual primer pairs. According to the recommendation of Chen et al., (2010), three internal reference genes, *UBC21*, *TIP41* and *ACT7*, were used to normalize the qRT-PCR data. The analysis of the qRT-PCR data (including the calculation of standard error, SE) was performed according to the geNORM manual (Vandesompele et al., 2002). The sequences of the primers and probes used in this study are listed in Table A1.

### **3.5.7 SYBR-green qRT-PCR**

SYBR-green qRT-PCR was used to quantify the overall transcription level of the three *BnGPAT4* homologues in *gpat4* RNAi *B. napus* lines. The reactions were performed as described in Chen et al., (2010) with minor modifications. In brief, the 10  $\mu$ L reaction contained 1 $\times$ SYBR Green master mix (Molecular Biology Facility, University of Alberta), 450 nM primers, and 2.5  $\mu$ L of 50 $\times$  diluted cDNA sample (the cDNA samples were prepared as described above under “Taqman qRT-PCR”). The following standard thermal profile was used: 95°C for 2 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min. Amplicon dissociation (melting) curves were recorded thereafter by using the following cycle: 95°C for 15 s, 60°C for 15 s, and 95°C for 15 s with a ramp rate of 2%. Four technical replicate reactions were performed with each cDNA sample and individual primer pairs.

### 3.5.8 Genomic DNA cytosine methylation analysis

Genomic DNA was isolated from a mixture of different *B. napus* vegetative samples, including seedlings, leaves, stems, inflorescence primordia and flowers. The EpiTect Bisulfite Kit (Qiagen, Valencia, CA, USA) was used for bisulfite conversion and cleanup of DNA according to the manufacturer's handbook. For PCR amplification of the bisulfite converted genomic DNA, three pairs of homologue-specific primers were designed for each homologue to give three amplicons that would cover the 625 bp of target sequence. The sequences of the primers were listed in Table A1. The PCR products were then gel-purified and subcloned into the pCR4-TOPO TA cloning vectors (Invitrogen, Carlsbad, CA, USA). For each PCR amplicon, 12 clones were sequenced.

### 3.5.9 Constructs for expression of different *BnGPAT4* homologues in

#### *Arabidopsis gpat4 gpat8* double mutant line

The full-length cDNA and promoter of individual *BnGPAT4* homologues together with a NOS terminator were subcloned into a binary vector pGreenII0229 (Fig. A4) (Hellens et al., 2000). The sequences of the associated primers are listed in Table A1. The new plasmids, which were then named pGreenII0229- *BnGPAT4*-C1<sub>pro:cDNA</sub>, *BnGPAT4*-C2<sub>pro:cDNA</sub>, *BnGPAT4*-A1<sub>pro:cDNA</sub>, were transformed into *Arabidopsis gpat4 gpat8* double T-DNA mutant lines (provided by Dr. Yonghua Li-Beisson) via the floral dip method as described earlier.

### **3.5.10 Preparation of RNAi construct for silencing *GPAT4* homologues in *B. napus***

Two identical fragments of approximately 250 bp, which were amplified from the coding region of *BnGPAT4*, were subcloned into pKannibal (Helliwell and Waterhouse, 2003) at two orientations of sense (at the sites of *XhoI* and *KpnI*) and anti-sense (at the sites of *HindIII* and *XbaI*). The *XhoI-XbaI* cassette, which was then excised from the resulting pKannibal construct, together with a double CaMV35S or a napin promoter (with *SacI* and *XhoI* at the two ends) were subcloned into a binary plasmid RD400 (kindly provided by Dr. Elzbieta Mietkiewska) at the sites of *SacI* and *XbaI* (Fig. A4). The sequences of the primers used here were listed in Table A1.

### **3.5.11 Generating *gpat4* RNAi *B. napus* lines**

*B. napus* plants (DH12075) were transformed as described by Bondaruk et al (2007). Benzyl adenine (4.5 mg /L) and naphthalene acetic acid (0.1 mg /L) were used to induce the shoot and root formation from transformed callus. The resulting transgenic plants were transferred to soil once roots were established. Genomic DNA PCR and real-time PCR were then performed to confirm the presence of transgene cassette.

### **3.5.12 Toluidine blue test**

The toluidine-blue test was used according to the method described in Tanaka et al (2004). Plant tissues were incubated in an aqueous solution of 0.05% (w/v) toluidine-blue for 2 min and then rinsed with water.

### **3.5.13 Preparation of surface cast of live leaves and microscopy**

The surface casts were prepared by painting the adaxial surface of attached leaves with clear fingernail polish. After 10 minutes, the cast was pulled from the leaf surface using a piece of Scotch tape (Brewer 1992). Light microscopy was performed with a Zeiss Imager. M1 microscope coupled to a Zeiss AxioCam HRm camera.

### **3.5.14 Analysis of cutin monomers**

The method for analyzing cutin monomers was adapted from Bonaventure et al (2004) with slight modifications. Briefly, delipidated residues from 300mg of fresh leaves were treated with sodium methoxide for depolymerization and methanolysis. The products of methanolysis were derivatized with pyridine and BSTFA (*N,O*-bis(trimethylsilyl)-trifluoroacetamide) to prepare trimethylsilyl derivatives. The final monomer derivatives were dissolved in 200  $\mu$ L of heptanes:toluene (1:1, v/v), and then analyzed by GC-MS. GC-MS analysis was performed using an Agilent 6890N gas chromatograph with an Agilent 5975 Inert Mass Selective Detector. Chromatographic separation was achieved using an HP-5MS capillary column (30 m x 0.25 mm x 0.25  $\mu$ M; Agilent Technologies) with a constant helium flow rate of 1 mL/min and with temperature programmed from 140 to 300°C at 3 °C/min. The inlet was operated in split mode (10:1 split ratio, 1 $\mu$ L injection) at 310°C. For the mass spectra condition: the solvent delay was 4 min, ionization energy was 70 eV, and data was acquired in scan mode with a range from 35 to 500 amu.

### **3.5.15 Analysis of fatty acid composition of *B. napus* seed oil**

A single seed was boiled in 1 mL of isopropanol for 10 min at 80°C. After cool down the sample was dried down under N<sub>2</sub>. About 1 mL of methanolic HCl was added to the sample and incubated for 1 h at 80°C to prepare the fatty acid methyl esters (FAMES). FAMES were then extracted twice with 1 mL of hexane. The extracted FAMES were then resuspended in iso-octane for GC-MS analysis. GC-MS analysis was performed with the same gas chromatograph as the one used in the cutin monomer analysis. Chromatographic separation was achieved using a capillary DB-23 (30m) column (0.25 mm x 0.25 Sm x 30 m) with a constant helium flow rate of 1.2 mL/min and with temperature programmed from 90 to 180°C at 10°C/min. The inlet was operated in splitless mode at 290°C. For the mass spectra detection: the solvent delay was 4 min, ionization energy was 70 eV, and scan mode was used for data acquisition with a range from 30 to 350 amu.

### **3.6 Footnote**

A portion of this chapter has been published.

Xue Chen, Martin Truksa, Crystal L. Snyder, Aliaa El-Mezawy, Saleh Shah, Randall J. Weselake (2011) Three homologous genes encoding *sn*-glycerol-3-phosphate acyltransferase 4 exhibit different expression patterns and functional divergence in *Brassica napus*. *Plant Physiology* 155: 851-865.

Copyright (2011), with permission from American Society of Plant Biologists.

All co-authors approved the usage of the manuscript in the present thesis.

### 3.7 References

- Adams KL, Cronn R, Percifield R and Wendel JF** (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci USA* **100**: 4649-4654
- Beaudoin F, Wu X, Li F, Haslam RP, Markham JE, Zheng H, Napier JA and Kunst L** (2009) Functional characterization of the Arabidopsis beta-ketoacyl-coenzyme A reductase candidates of the fatty acid elongase. *Plant Physiol* **150**: 1174-1191
- Beisson F, Li Y, Bonaventure G, Pollard M and Ohlrogge JB** (2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *Plant Cell* **19**: 351-368
- Bennett MD and Leitch IJ** (1997) Nuclear DNA amounts in angiosperms: 583 new estimates. *Ann Bot (Lond)*. **80**: 169-196
- Bird SM and Gray JE** (2003) Signals from the cuticle affect epidermal cell differentiation. *New Phytol* **157**: 9-23
- Blanc G and Wolfe KH** (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**: 1679-1691
- Blundy KS, Blundy MA and Crouch ML** (1991) Differential expression of members of the napin storage protein gene family during embryogenesis in *Brassica napus*. *Plant Mol Biol* **17**: 1099-1104
- Bonaventure G, Beisson F, Ohlrogge J and Pollard M** (2004) Analysis of the aliphatic monomer composition of polyesters associated with *Arabidopsis*

epidermis: Occurrence of octadeca-*cis*-6, *cis*-9-diene-1, 18-dioate as the major component. *Plant J* **40**: 920-930

**Bondaruk M, Johnson S, Degafu A, Boora P, Bilodeau P, Morris J, Wiehler W, Foroud N, Weselake R and Shah S** (2007) Expression of a cDNA encoding palmitoyl-acyl carrier protein desaturase from cat's claw (*Doxantha unguis-cati* L.) in *Arabidopsis thaliana* and *Brassica napus* leads to accumulation of unusual unsaturated fatty acids and increased stearic acid content in the seed oil. *Plant Breed* **126**: 186-194

**Brewer CA** (1992) Responses by stomata on leaves to microenvironmental conditions. *Test Stud Laboratory Teach* **13**: 67-75

**Chan SW, Henderson IR and Jacobsen SE** (2005) Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet* **6**: 351-360

**Chapman LA and Goring DR** (2010) Pollen-pistil interactions regulating successful fertilization in the Brassicaceae. *J Exp Bot* **61**: 1987-1999

**Chen X, Goodwin SM, Boroff VL, Liu X and Jenks MA** (2003) Cloning and characterization of the WAX2 gene of *Arabidopsis* involved in cuticle membrane and wax production. *Plant Cell* **15**: 1170-1185

**Chen X, Truksa M, Shah S and Weselake RJ** (2010) A survey of quantitative real-time polymerase chain reaction internal reference genes for expression studies in *Brassica napus*. *Anal Biochem* **405**: 138-140

**Chen ZJ** (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev Plant Biol* **58**: 377-406

**Cheung F, Trick M, Drou N, Lim YP, Park, JY, Kwon SJ, Kim JA, Scott R, Pires JC, Paterson AH, Town C and Bancroft I** (2009) Comparative analysis between homoeologous genome segments of *Brassica napus* and its progenitor species reveals extensive sequence-level divergence. *Plant Cell* **21**: 1912-1928

**Claros MG and von Heijne G** (1994) TopPred II: an improved software for membrane protein structure predictions. *Comput Appl Biosci* **10**: 685-686

**Clough SJ and Bent AF** (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**: 735-743

**Ellerstrom M, Stalberg K, Ezcurra I and Rask L** (1996) Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. *Plant Mol Biol* **32**: 1019-1027

**Enstone DE, Peterson CA and Ma F** (2003) Root endodermis and exodermis: Structure, function, and responses to the environment. *J Plant Growth Regul* **21**: 335-351

**Fitch WM** (2000) Homology: a personal view on some of the problems. *Trends Genet* **16**: 227-231

**Flagel LE and Wendel JF** (2009) Gene duplication and evolutionary novelty in plants. *New Phytol* **183**: 557-564

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

- Gaeta RT, Pires JC, Iniguez-Luy F, Leon E and Osborn TC** (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* **19**: 3403-3417
- Gambino G, Perrone I and Gribaudo I** (2008) A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochem Anal* **19**: 520-525
- Gidda SK, Shockey JM, Rothstein SJ, Dyer JM and 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 Phys and Biochem* **47**: 867-879
- Gray JE, Holroyd GH, van der Lee FM, Bahrami AR, Sijmons PC, Woodward FI, Schuch W and Hetherington AM** (2000) The *HIC* signaling pathway links CO<sub>2</sub> perception to stomatal development. *Nature* **408**: 713-716
- Guindon S and Gascuel O** (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696-704
- Haughn G and Chaudhury A** (2005) Genetic analysis of seed coat development in *Arabidopsis*. *Trends Plant Sci* **10**: 472-477
- Hellens RP, Edwards EA, Leyland NR, Bean S and Mullineaux PM** (2000) pGreen: A versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol* **42**: 819-832
- Josefsson LG, Lenman M, Ericson ML and Rask L** (1987) Structure of a gene encoding the 1.7 S storage protein, napin, from *Brassica napus*. *J Biol Chem* **262**: 12196-12201

**Kappen L and Haeger S** (1991) Stomatal responses of *Tradescantia albiflora* to changing air humidity in light and in darkness. *J Exp Bot* **42**: 979-986

**Kennedy EP** (1961) Biosynthesis of complex lipids. *Fed Proc Am Soc Exp Biol* **20**: 934-940

**Klahre U, Crete P, Leuenberger SA, Iglesias VA and Meins F Jr** (2002) High molecular weight RNAs and small interfering RNAs induce systemic posttranscriptional gene silencing in plants. *Proc Natl Acad Sci USA* **99**: 11981–11986

**Kridl JC, McCarter DW, Rose RE, Scherer DE, Knutzon DS, Radke SE and Knauf VC** (1991) Isolation and characterization of an expressed napin gene from *Brassica rapa*. *Seed Sci Res* **1**: 209-219

**Kusaba M** (2004) RNA interference in crop plants. *Curr Opin Biotechnol* **15**: 139-143

**Li YH, Beisson F, Koo AJK, Molina I, Pollard M and Ohlrogge J** (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc Natl Acad Sci USA* **104**: 18339-18344

**Liu SL and Adams KL** (2010) Dramatic change in function and expression pattern of a gene duplicated by polyploidy created a paternal effect gene in the Brassicaceae. *Mol Biol Evol* **27**: 2817-2828

**Lukens L, Quijada PA, Udall J, Pires JC, Schranz ME and Osborn T** (2004) Genome redundancy and plasticity within ancient and recent Brassica crop species. *Biol J Linn Soc* **82**: 675-688

**McCormick S** (2004) Control of male gametophyte development. *Plant Cell* **16**: S142-153

**Molina I, Ohlrogge JB and Pollard M** (2008) Deposition and localization of lipid polyester in developing seeds of *Brassica napus* and *Arabidopsis thaliana*. *Plant J* **53**: 437-449

**Mott KA and Franks PJ** (2001) The role of epidermal turgor in stomatal interactions following a local perturbation in humidity. *Plant, Cell and Environ* **24**: 657-662

**O'Neill CM and Bancroft I** (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J* **23**: 233-243

**O'Quin JB, Bourassa L, Zhang D, Shockey JM, Gidda SK, Fosnot S, Chapman KD, Mullen RT and Dyer JM** (2010) Temperature-sensitive post-translational regulation of plant omega-3 fatty-acid desaturases is mediated by the endoplasmic reticulum-associated degradation pathway. *J Bio Chem* **285**: 21781-21796

**Park JY, Koo DH, Hong CP, Lee SJ, Jeon JW, Lee SH, Yun PY, Park BS, Kim HR, Bang JW, Plaha P, Bancroft I and Lim YP** (2005) Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kb gene-rich region of *Arabidopsis* chromosome 4 and partially duplicated on chromosome 5. *Mol Genet Genomics* **274**: 579-588

**Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC and Lydiate DJ** (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **171**: 765-781

**Pontius JU, Wagner L and Schuler GD** (2003) UniGene: a unified view of the transcriptome. In: *The NCBI Handbook*. Bethesda (MD): National Center for Biotechnology Information

**Rana D, van den Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP and Bancroft I** (2004) Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. *Plant J* **40**: 725-733

**Schroeder JI, Allen GJ, Hugouvieux V, Kwak JM and Waner D** (2001) Guard cell signal transduction. *Annu Rev Plant Physiol Plant Mol Biol* **52**: 627-658

**Shitsukawa N, Tahira C, Kassai K, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, Ogihara Y and Murai K** (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *Plant Cell* **19**: 1723-1737

**Tanaka T, Tanaka H, Machida C, Watanabe M and Machida Y** (2004) A new method for rapid visualization of defects in leaf cuticle reveals five intrinsic patterns of surface defects in *Arabidopsis*. *Plant J* **37**: 139-146

**Thilmony R, Guttman M, Chiniquy D and Blechl A** (2006) pGPro1, a novel binary vector for monocot promoter characterization. *Plant Mol Biol Rep* **24**: 57-69

- U N** (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn J Bot* **7**: 389-452
- Udall JA and Wendel JF** (2006) Polyploidy and crop improvement. *The Plant Genome, a Supplement to Crop Science*. **46**: S3-14
- Ursin VM, Yamaguchi J and McCormick S** (1989) Gametophytic and sporophytic expression of anther-specific genes in developing tomato anthers. *Plant Cell* **1**: 727-736
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F** (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: (7)
- Wan L, Xia Q, Qiu X, Selvaaraj G** (2002) Early stages of seed development in *Brassica napus*: a seed coat-specific cysteine proteinase associated with programmed cell death of the inner integument. *Plant J* **30**: 1-10
- Wang YX, Wu H and Yang M** (2008) Microscopy and bioinformatic analyses of lipid metabolism implicate a sporophytic signaling network supporting pollen development in *Arabidopsis*. *Mol. Plant* **1**: 667-74
- Wendel JF** (2000) Genome evolution in polyploids. *Plant Mol Bio* **42**: 225-249
- Weselake RJ** (2005) Storage lipids. In DJ Murphy, ed, *Plant Lipids: Biology, Utilization and Manipulation*. Blackwell Publishing, Oxford, pp 162-221

- Whittle CA and Krochko JE** (2009) Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* **21**: 2203-2219
- Xu C, Cornish AJ, Froehlich JE and Benning C** (2006) Phosphatidylglycerol biosynthesis in chloroplasts of *Arabidopsis* mutants deficient in acyl-ACP glycerol-3-phosphate acyltransferase. *Plant J* **47**: 296-309
- Yang YW, Lai KN, Tai PY and Li WH** (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J Mol Evol* **48**: 597-604
- Zeiger E** (1983) The biology of stomatal guard cells. *Annu Rev Plant Physiol* **34**: 441-475
- Zeiger E and Stebbins L** (1972) Developmental genetics in barley: a mutant for stomatal development. *Am J Bot* **59**: 143-148
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE and Ecker JR** (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126**: 1189-1201
- Zheng Z and Zou J** (2001) The initial step of the glycerolipid pathway: Identification of glycerol 3-phosphate/dihydroxyacetone phosphate dual substrate acyltransferases in *Saccharomyces cerevisiae*. *J Biol Chem* **276**: 41710-41716
- Zheng Z, Xia Q, Dauk M, Shen W, Selvaraj G and Zou J** (2003) *Arabidopsis AtGPAT1*, a member of the membrane-bound glycerol-3-phosphate

acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* **15**: 1872-1887

**Zilberman D, Gehring M, Tran RK, Ballinger T and Henikoff S (2007)**

Genome-wide analysis of *Arabidopsis* DNA methylation uncovers an interdependence between methylation and transcription, *Nat Genet* **39**: 61-69

## Chapter 4.

### A survey of quantitative real-time PCR internal reference genes for expression studies in *Brassica napus*

#### 4.1 Introduction

Assessing gene expression through quantification of transcripts is one of the most common experimental procedures in molecular biology. The instruments capable of monitoring DNA amplification in real time became widely accessible during the last decade, and Quantitative or Real-Time Reverse-Transcription Polymerase Chain Reaction (qRT-PCR) has been established as a standard method of quantifying the expression of individual genes. Compared to earlier methods, qRT-PCR has the advantage of high speed, high sensitivity, high degree of automation and reproducibility, and high throughput. Its unprecedented sensitivity alone, however, does not guarantee correct interpretation of data. Rather, a robust normalization procedure for the variations in the amount and quality of the starting RNA sample, as well as individual steps of the qRT-PCR assay, is paramount to accurate interpretation of the results (Vandesompele et al., 2002). Under certain experimental conditions the expression level of a gene or the number of gene-specific mRNA molecules, can be normalized to the number of cells used to prepare the RNA sample. It is, however, often impractical or impossible when working with tissues (Vandesompele et al., 2002; Kanno et al., 2006), and thus investigators have to rely on alternative normalization procedures

---

A version of this chapter was published: Chen et al. (2010) A survey of quantitative real-time polymerase chain reaction internal reference genes for expression studies in *Brassica napus*. *Analytical Biochemistry* 405: 138-140.

where the qRT-PCR data are normally viewed in relation to ribosomal RNA or stably expressed endogenous genes.

Careful selection and proper validation of the internal references becomes particularly important when gene expression is analyzed in cell types with varying transcriptional activity. It has also been demonstrated that many of the “traditional” house-keeping genes used for normalization of qRT-PCR data are only presumed to be stable, and their expression varies with different cell types or experimental conditions (Nicot et al., 2005; Waxman et al., 2007; Remans et al., 2008). Czechowski et al (2005) have used microarray-derived expression profiles to identify *Arabidopsis* genes superior in their expression stability to traditional “house-keeping” controls. The genetic proximity of *Arabidopsis* and *B. napus* enables researchers working with *Brassica* species to capitalize on the large body of data compiled for *Arabidopsis* genes. During the estimated 12.2-19.2 million years since the divergence of *Arabidopsis* and *Brassica* species (Arabidopsis Genome Initiative, 2000), closely related genes might have undergone functional divergence and thus their expression profiles could potentially differ between the two species. Consequently, *Brassica* internal reference genes selected on a basis of studies with *Arabidopsis* require independent validation in the new context.

Owing to the economic importance of oil from seeds of *B. napus* and the need to understand regulatory processes governing seed oil accumulation, the developing embryo of this oilseed species has been the object of numerous gene expression studies (Hays et al., 1999; Fourmann et al., 2001; Fei et al., 2007; Huang et al., 2009). The embryo undergoes substantial developmental changes in

its transition from the heart through to the late cotyledonary morphogenetic stage wherein large amounts of storage compounds are formed and deposited, followed by desiccation, and finally metabolic quiescence (Harada, 1997). These divergent processes are associated with the coordinated expression of gene subsets that are behind the qualitative and quantitative changes in embryo transcription (Comai and Harada, 1990; Ruuska et al., 2002).

In the present study, eight reference genes of *B. napus* were evaluated using qRT-PCR data, focusing on vegetative tissues and developing embryos. Analysis of expression stability indicated that *UPI1*, *UBC9*, *UBC21* and *TIP41* were the top four choices as stably expressed reference genes for vegetative tissues, and *ACT7*, *UBC21*, *TIP41* and *PP2A* were the top four choices for maturing embryos. In addition, radio-labeling of overall mRNA of maturing embryos indicated that the expression patterns of the top four ranked reference genes reflected the overall mRNA content changes in maturing embryos.

## **4.2 Results**

### **4.2.1 Selection of candidate *B. napus* internal references**

In the present study, eight candidate *B. napus* internal references were chosen on the basis of stable expression profiles of their *Arabidopsis* orthologs in developing seeds and across different tissues and organs. First, using *Arabidopsis* microarray data (Schmid et al., 2005), expression profiles of the top 15 reference genes from the developmental series ranking by Czechowski et al. (2005) were investigated. Out of these 15 genes, six candidate genes that exhibited the most stable expression during embryo maturation were selected. The genes encode:

protein phosphatase 2A subunit A3 (*PP2A*, *Arabidopsis* gene accession No. At1g13320), TIP41-like protein (*TIP41*, At4g34270), ubiquitin-conjugating enzyme 9 (*UBC9*, At4g27960), SAND-family protein (*SAND*, At2g28390), and two uncharacterized proteins (*UPI*, At4g33380; *UP2*, At4g26410) (Table 4.1). In addition, two “traditional” house-keeping genes often used as references were selected, namely actin 7 (*ACT7*, At5g09810) and ubiquitin-conjugating enzyme 21 (*UBC21*, At5g25760) (Table 4.1). The full-length cDNA sequences of the eight *Arabidopsis* genes were used to query *B. napus* ESTs using the nucleotide MEGABLAST algorithm (NCBI). Positive hits were aligned and evaluated for the number of putative orthologs.

#### **4.2.2 PCR specificity and amplification efficiency analysis**

The use of SYBR Green I dye, which can bind to any double-stranded DNA, requires a high specificity of amplification to avoid generation of false positive signals. To maximize the specificity of the designed primers, the pairs of sequences were checked against the *Arabidopsis* and *B. napus* EST databases using Primer-BLAST (NCBI). The specificity of primer pairs was further tested by inspecting the dissociation curves (Ririe et al., 1997). All tested primer pairs resulted in a single dominant peak on the dissociation curve with inflection point close to the calculated melting temperature of the amplicon. To further confirm the specificity of amplification, a more stringent test was performed by sequencing the amplicons generated from qRT-PCR. In all cases, the sequences of the PCR products matched that of the intended target.

**Table 4.1. Surveyed internal reference genes of *B. napus*.**

<b>Internal Reference</b>	<b><i>Arabidopsis</i> ortholog</b>	<b><i>Arabidopsis</i> annotation</b>	<b>Primers (F/R)</b>	<b>Amplicon length (bp)</b>	<b>Exon-exon junction</b>
<i>ACT7</i>	At5g09810	Actin	5'-TGGGTTTGCTGGTGACGAT 5'- TGCCTAGGACGACCAACAATACT	63	No
<i>UBC21</i>	At5g25760	Ubiquitin conjugating enzyme 21	5'- CCTCTGCAGCCTCCTCAAGT 5'- CATATCTCCCCTGTCTTGAAATGC	77	Yes
<i>PP2A</i>	At1g13320	Regulatory subunit of protein phosphatase 2A	5'- TGGCTTCAGTTATAATGGGAATGG 5'- GAAAGATTGGAAGGAGATGCTCAAT	75	No
<i>TIP41</i>	At4g34270	TIP41-like family protein	5'- AGAGTCATGCCAAGTTCATGGTT 5'- CCTCATAAGCACACCATCAACTCTAA	69	No
<i>UBC9</i>	At4g27960	Ubiquitin conjugating enzyme 9	5'- GCATCTGCCTCGACATCTTGA 5'- GACAGCAGCACCTTGGAATG	68	Yes
<i>SAND</i>	At2g28390	SAND-family protein	5'- GCTGGGTCACTCCAGATTTTG 5'- CCATCGCCTTGTCTGCAAG	63	Yes
<i>UPI</i>	At4g33380	Unknown protein	5'- AGCCTGAGGAGATATTAGCAGGAA 5'- ATCTCACTGCAGCTCCACCAT	87	Yes
<i>UP2</i>	At4g26410	Unknown protein	5'- AAATTCCTGGGAGGGAAGCTAT 5'- TTCTGTCTCAGGAGCGAAGTCAT	70	No

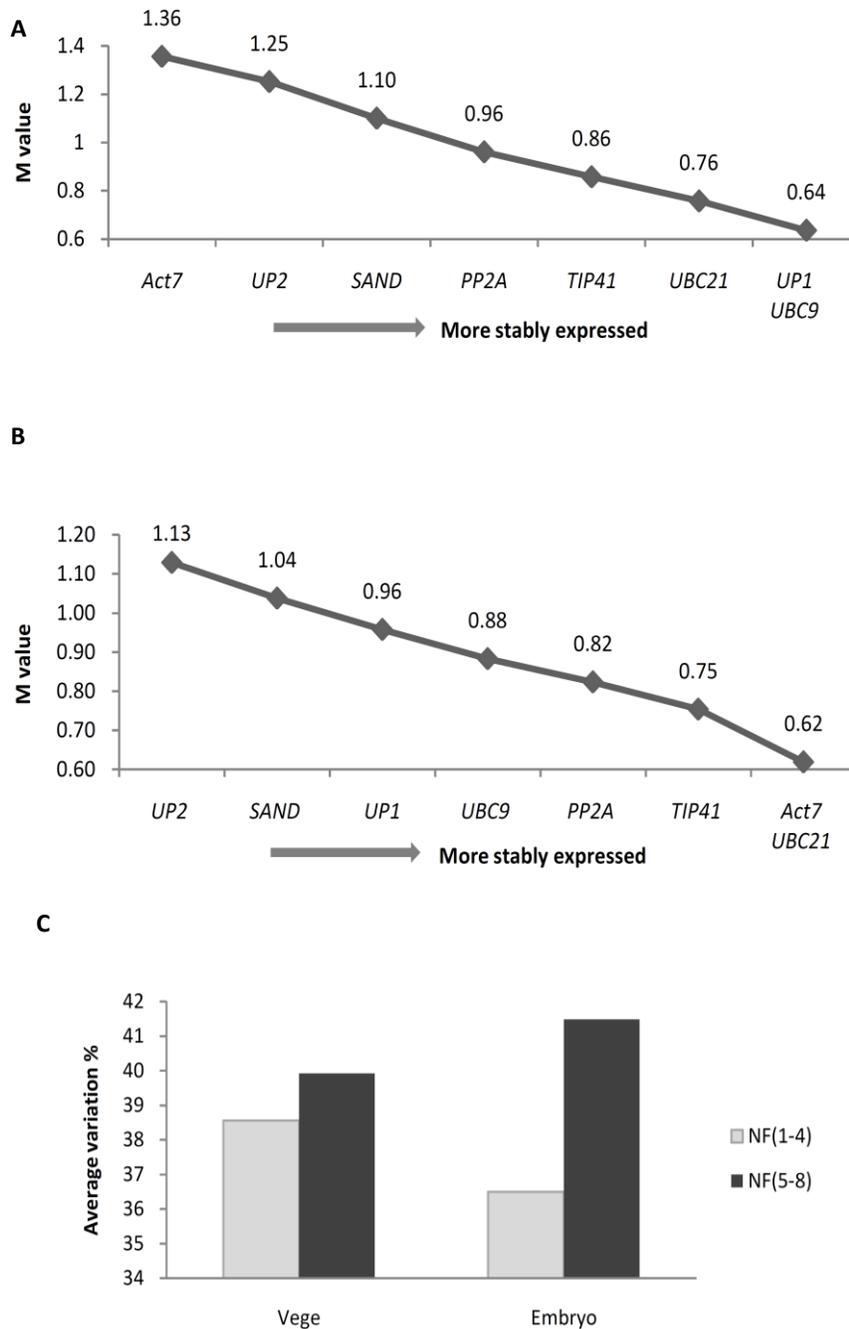
The most common way to analyze and interpret qRT-PCR data is the comparative Ct method (also known as  $2^{-\Delta\Delta C_t}$  method) which is very sensitive to variations in amplification efficiency. Typically, investigators rely on the assumption that the PCR efficiency ( $E$ ) of the analyzed amplicons is constant and equal 1.0. A few studies, however, have shown that this approach is inadequate due to small variations in  $E$  values which can cause the erroneous interpretation of expression levels (Ramakers et al., 2003).

To accurately evaluate the amplification efficiency of the surveyed internal references, we used LinRegPCR v.7.0 software (Ramakers et al., 2003), which calculates PCR efficiency based on linear regression between the log of the fluorescence data and the number of amplification cycles. The efficiency values were obtained for each reaction and their means are listed in Table A2. Of the total eight primer pairs, seven had efficiencies higher than 0.9, and one (*SAND*) had efficiencies above 0.85. For each pair of primers, the mean efficiency value of the three biological replicates for each sample was taken into account in all subsequent calculations.

#### **4.2.3 Expression stability of candidate reference genes**

The geNORM analysis assigns an expression stability value  $M$  to each evaluated reference gene, where smaller  $M$  value indicates more stable expression of a gene (Vandesompele et al., 2002). It is based on pair-wise comparisons and stepwise exclusion of candidates which results in their ranking according to expression stability. The maturing embryos, which undergo substantial developmental transitions, are considered as very different organs compared to

the vegetative organs. Reference genes suitable for vegetative tissues may not be appropriate for use with maturing embryos. Therefore, in order to obtain more accurate results for the expression stability analysis, two sets of data: qRT-PCR results derived from vegetative tissues, and those derived from the maturing embryos were analyzed by geNORM. As shown in Fig. 4.1A and B, in vegetative tissues, the top four internal references as ranked by geNORM: *UPI*, *UBC9*, *UBC21* and *TIP41* were recommended for calculating the normalization factor (geometric mean of the relative expression quantities of the selected reference genes). In maturing embryos, the top four references: *ACT7*, *UBC21*, *TIP41*, and *PP2A* were recommended for calculating the normalization factor. In both vegetative and embryo tissues, *TIP41* and *UBC21* were suggested to be stably expressed genes.



**Figure 4.1. Gene expression stability evaluated by geNORM.**  
 A: Reference gene ranking for vegetative samples; B: Reference gene ranking for maturing embryos; C: The variation coefficients of the relative expression quantities of surveyed reference genes according to the method described in Vandesompele et al., (2002).

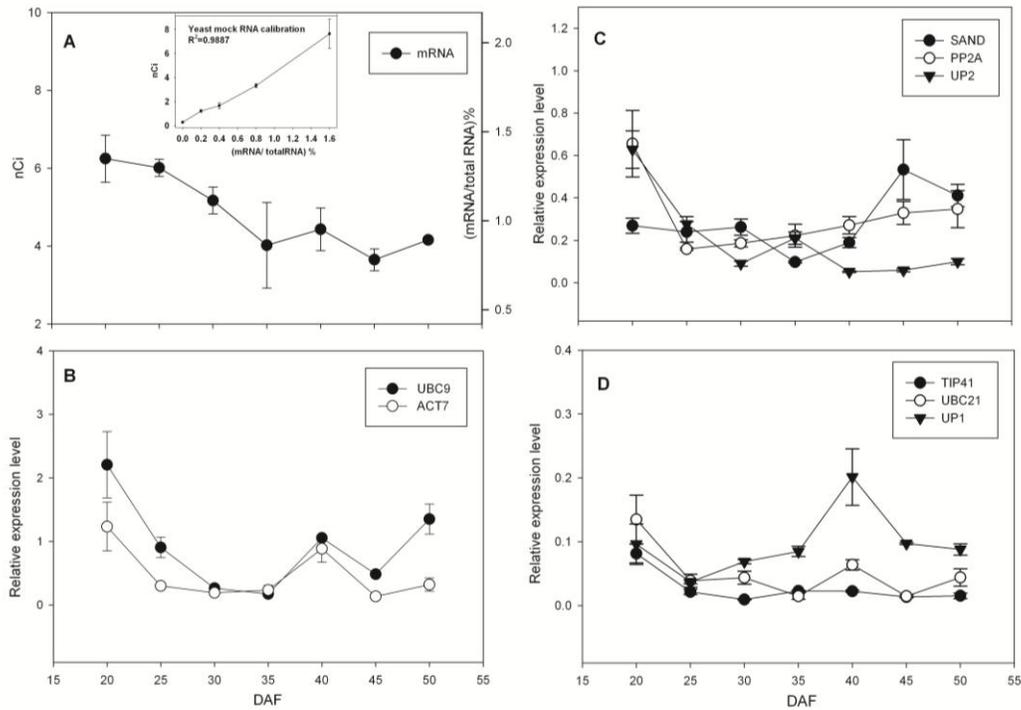
#### **4.2.4 Validation of the candidate reference genes and the normalization factors**

To confirm that the candidates with the lowest  $M$  values were indeed stably expressed, we performed a variation coefficient analysis on the relative expression quantities of surveyed reference genes according to the method described in Vandesompele et al (2002). In brief, two different normalization factors, NF1-4 (top four ranked genes by geNORM analysis) and NF5-8 (the other four genes), were calculated and used to normalize the relative expression quantities ( $RQ$ ) of the 8 reference genes, respectively. Thereafter, the two datasets of re-normalized expression quantities (by NF1-4 and NF5-8, respectively) were used to calculate the variation coefficients of each reference gene. For each dataset, the average gene-specific variation (determined as the average variation coefficient) of four reference genes, which exhibited the lowest variation coefficients, was calculated. This analysis was performed with both vegetative and embryo data. The final results suggested that the average gene-specific variation was much smaller when the datasets were normalized by NF1-4 than when normalized by NF5-8 (Fig. 4.1C), further confirming that the top four ranked reference genes were more stably expressed than the others.

#### **4.2.5 Comparison between the expression patterns of overall mRNA and the candidate reference genes**

To monitor the changes of the ratio of mRNA to total RNA in developing embryos, reverse transcription reactions were performed with anchored 20mer-dT primers, dNTP in the presence of [ $^{32}$ P] dCTP, and the same total RNA samples as

used in qRT-PCR studies. The assay was first validated using mock RNA samples consisting of yeast tRNA and increasing amounts of alien mRNA molecule to demonstrate the linear relationship between the number of polyadenylated transcripts and the amount of incorporated [<sup>32</sup>P] dCTP (coefficient of determination  $R^2=0.9887$ , Fig. 4.2A). Thereafter, multiple reverse transcription reactions were performed using 1 µg total RNA samples isolated from developing embryos at different stages, 20 to 50 DAF (days after flowering). As shown in Fig. 4.2A, the overall mRNA content was at the highest level at the beginning of the period of evaluation (20 DAF), and then decreased and reached a minimum at 35 DAF. As embryo maturation proceeded and seeds entered the desiccation stage, the relative proportion of mRNA increased slightly. Similar expression patterns were also observed in a few reference genes (i.e., *ACT7* and *UBC21*) (Fig. 4.2B-D).



**Figure 4.2. Changes in mRNA content and expression levels of selected internal reference genes of the maturing embryos of *B. napus*.**

A: Changes in overall mRNA content measured by reverse-transcription in the presence of radioactively labelled dCTP. Inside small figure: yeast mock RNA calibration. Left axis – radioactivity detected in purified cDNA fraction; right axis – approximate relative content of mRNA based on a calibration curve with mock RNA samples. B, C, D: The relative expression quantities of candidate reference genes (relative to the external control) represented as  $[(E+1)^{Ct}]_{target} / [(E+1)^{Ct}]_{external\ control}$ . DAF: days after flowering.

### 4.3 Discussion

One of the main advantages of using the qRT-PCR for studying gene expression is the relative ease with which a very sensitive and highly gene-specific assay can be set up. The outstanding capability of the technique does not, however, preclude misleading interpretation of the data if a robust normalization methodology is not used. Commonly, the expression level of the gene of interest is normalized to the level of an internal reference gene, which is subjected to the same conditions of reverse transcription and amplification, and suffers from the same experimental errors as the gene of interest. The internal reference gene is thus presumed to be representative of the quantity and quality of mRNA in the starting sample.

The importance of choosing the appropriate internal reference genes has been discussed extensively in the literature (Czechowski et al., 2005; Lefebvre et al., 2008; Udvardi et al., 2008; Guenin et al., 2009). The relatively slow acceptance of fully validated reference genes as a standard in plant sciences compared to the medical field (Guenin et al., 2008) is, in our opinion, a result of a wide range of organisms being studied by plant scientists, often with limited genomic information at hand. The effort required to select and validate novel reference genes undermines the aforementioned promptness and flexibility of qRT-PCR.

For calculation of the relative expression quantity ( $RQ$ ) of individual references ( $[(E+I)^{Ct}]_{target} / [(E+I)^{Ct}]_{external\ control}$ ), external control-“Alien RNA”-

was taken into account. The usage of a defined amount of foreign RNA molecule as an external control corrects variability among experiments, enables quantitative and qualitative assessment of the qRT-PCR process, and prevents unexpected errors (Czechowski et al., 2005; Udvardi et al., 2008). In the present work, the use of the alien RNA spike helped to detect qRT-PCR inhibitors in RNA samples prepared with RNeasy Plant Mini Kit (Qiagen, Valencia, CA) from entire seeds in mid-developmental stage (30-40 DAF). The RNA samples prepared from excised embryos were free of this inhibitor, while addition of isolated seed coats to leaf tissue before the RNA extraction substantially increased the observed external reference *Ct* values, thus confirming the source of the problem. The RNA samples from isolated embryos were used in all following experiments.

To compensate for variations in individual reference genes, measuring the expression levels of several reference genes, calculating their geometric mean and using it as a normalization factor is often preferable (Udvardi et al., 2008). In the present study, the use of the top four internal references as ranked by geNORM analysis were recommended for both vegetative tissues and embryo qRT-PCR studies.

It is difficult to find a gene that would serve as an ideal reference under all experimental conditions. In practice, the choice of the appropriate internal reference depends on multiple aspects of the experimental design including organs, tissues or cell types being analyzed as well as the relative expression level of the gene of interest. Stably expressed genes can be affected by experimental treatments and are subject to global changes in mRNA population. Consequently,

the reference genes used to normalize data have to be validated anew for each set of experimental conditions. The results of this study should provide researchers working with *Brassica* species, especially those interested in gene expression during the course of embryo maturation, with information that will expedite the choice and validation of appropriate reference genes for their qRT-PCR experiments.

#### **4.4 Materials and methods**

##### **4.4.1 Plant material**

*B. napus* double haploid line (DH12075) plants were grown under long-day conditions (16-h day/8-h night) at 25 °C. The following samples were collected: 2-day old seedlings, 5-day old seedlings, 10-day old seedlings, young rosette leaves, mature rosette leaves, anthers, flowers, and embryos at 20, 25, 30, 35, 40, 45, and 50 DAF. All samples were stored at -80 °C until RNA extraction.

##### **4.4.2 RNA extraction**

For each type of tissue/organ, three biological replicate samples were collected. Total RNA was extracted from each sample, and then reverse transcribed to cDNA samples. Total RNA was prepared using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNase-free DNase set (Qiagen) was used to eliminate genomic DNA contamination of all the RNA samples. RNA concentration was measured with a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop, Wilmington, DE, USA). Absorbance ratio 260/280 nm of all RNA samples was between 2.0 and

2.2. All samples were further tested by agarose gel electrophoresis to assess the quality and integrity of the RNA.

#### **4.4.3 External reference and reverse transcription**

Alien RNA (Alien qRT-PCR Inhibitor Alert, Stratagene, La Jolla, CA) was used as an external reference to monitor enzyme inhibition during reverse transcription, PCR, template degradation and to compare qRT-PCR experiments performed at different times. Before the reverse transcription step, each sample (1 µg of total RNA) was spiked with an equal quantity ( $10^6$  copies) of alien RNA. The 20 µL reactions were run using the QuantiTect reverse transcription kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions. Reverse transcription reactions were primed using the "RT primer mix" provided by the kit. The "RT primer mix" is an optimized blend of oligo-dT and random primers. The final 20 µL RT reactions were diluted 50 times, and 2.5 µL was used for each qRT-PCR reaction, and 10 µL was used for semi-quantitative PCR.

#### **4.4.4 qRT-PCR primer design and reaction**

The primer pairs for real-time PCR were designed using the PrimerExpress 2.0 program (Applied Biosystems) with default criteria. All amplicons are located close to the 3' ends of the full-length sequences and, when possible, at least one primer covers an exon-exon junction. The primer pairs were then checked against *Arabidopsis* and *B. napus* EST databases using Primer-BLAST (NCBI). For Alien RNA external control, an optimized primer pair provided by the manufacturer (Stratagene) was used. The qRT-PCR reactions were performed in a Fast Optical 96-Well Reaction Plate on ABI PRISM 7900

HT Real-Time PCR System (both from Applied Biosystems, Foster City, CA, USA) using SYBR Green I to monitor double-strand DNA synthesis. The 2xSYBR Green master mix “Dynamite” (MBSU, University of Alberta) contained SYBR Green I, ROX dye, Taq polymerase and dNTPs. The 10 µL reactions contained 1xSYBR Green master mix, 450 nM primers, and 2.5 µL of 50-times diluted RT reaction. The following standard thermal profile was used: 95°C for 2 min, 35 cycles of 95°C for 15 s, and 60°C for 1 min. Amplicon dissociation (melting) curves were recorded after 35 cycles by using the following cycle: 95°C for 15 s, 60°C for 15 s, and 95°C for 15 s with a ramp rate of 2%.

#### **4.4.5 QRT-PCR data and stability of expression analysis**

Data generated by qRT-PCR were analyzed using the SDS 2.2.1 software (Applied Biosystems). The  $C_t$  values were obtained with a baseline calculated from cycles 3 to 15 and a threshold set to 2.5. PCR efficiency ( $E$ ) was evaluated using LinRegPCR v.7.0 software (Ramakers et al., 2003). In this approach, the PCR amplification efficiency ( $E$ ) was estimated from the fluorescence data obtained from the exponential phase of each individual amplification plot. The relative expression quantities of individual reference genes was calculated as  $[(E+1)^{C_t}]_{target} / [(E+1)^{C_t}]_{external\ control}$  (Czechowski et al., 2005). The highest relative quantities for each gene were set to 1, thus the transformed data sets were used to evaluate the reference gene stability by geNORM applet (Vandesompele et al., 2002). The usage of the external control (a defined amount of Alien RNA, Alien qRT-PCR Inhibitor Alert, Stratagene, La Jolla, CA) in the above calculation was recommended to correct variability between experiments, prevent unexpected

errors and enables quantitative and qualitative assessment of the PCR reactions (Czechowski et al., 2005; Udvardi et al., 2008).

#### **4.4.6 Quantification of overall mRNA in embryo via radio-labelling reverse transcription**

The reverse transcription reactions were performed using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA). The 20 µL reaction mix was prepared according to the manufacturer's instructions and contained 1 µg of plant or mock RNA, 2.5 µM anchored 20mer-dT primer (IDT, Coralville, IA, USA) and 5 µCi [<sup>32</sup>P] dCTP (PerkinElmer, Waltham, MA, USA). After 50 min incubation at 50°C, the reactions were stopped by heating the mix at 85°C for 5 min and then the RNA was eliminated by treating with *E. coli* RNase H (2 U for 20 min at 37°C). The volume of the reaction was then adjusted to 100 µL, and 90 µL of the reaction were applied onto an equilibrated Illustra NICK column (GE Healthcare Bio-Sciences, Piscataway, NJ, USA) for elution. To eliminate the contamination of free radiolabeled dNTP, 380 µL of the elution buffer was first applied to column. And then 50 µL of the elution buffer was used to elute the cDNA. To validate the reverse transcription assay and obtain an approximate calibration, a series of mock samples were prepared consisting of the purified yeast tRNA (Sigma-Aldrich Oakville, ON, Canada) and increasing quantities (0.2% - 1.6% mass) of Alien RNA spike control (Alien qRT-PCR Inhibitor Alert, Stratagene).

#### **4.5 Footnote**

A modified version of this chapter has been published.

Xue Chen, Martin Truksa, Saleh Shah, Randall J. Weselake (2010) A survey of quantitative real-time polymerase chain reaction internal reference genes for expression studies in *Brassica napus*. *Analytical Biochemistry* 405: 138-140.

Copyright (2010), with permission from Elsevier.

All co-authors approved the usage of the manuscript in the present thesis.

#### **4.6 References**

**Arabidopsis Genome Initiative** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796-815

**Comai L and Harada JJ** (1990) Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. *Proc Natl Acad Sci U S A* **87**: 2671-2674

**Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR** (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* **139**: 5-17

**Fei H, Tsang E and Cutler A** (2007) Gene expression during seed maturation in *Brassica napus* in relation to the induction of secondary dormancy. *Genomics* **89**: 419-428

**Fourmann M, Chariot F, Froger N, Delourme R and Brunel D** (2001) Expression, mapping, and genetic variability of *Brassica napus* disease resistance gene analogues. *Genome* **44**: 1083-1099

**Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C and Gutierrez L** (2009) Normalization of qRT-PCR data: the necessity of adopting a

systematic, experimental conditions-specific, validation of references. *J Exp Bot* **60**: 487-493

**Harada JJ** (1997) Seed maturation and control of germination. In BA Larkins, IK Vasil, eds, *In cellular and molecular biology of plant seed development*. Vol 4. Dordrecht: Kluwer Academic Publishers, pp 545-592

**Hays DB, Wilen RW, Sheng C, Moloney MM and Pharis RP** (1999) Embryo-specific gene expression in microspore-derived embryos of *Brassica napus*. An interaction between abscisic acid and jasmonic acid1, 2. *Plant Physiol* **119**: 1065-1072

**Huang Y, Chen L, Wang L, Vijayan K, Phan S, Liu Z, Wan L, Ross A, Xiang D, Datla R, Pan Y and Zou J** (2009) Probing the endosperm gene expression landscape in *Brassica napus*, *BMC Genomics* **10**: 256

**Kanno J, Aisaki K, Igarashi K, Nakatsu N, Ono A, Kodama Y and Nagao T** (2006) "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays. *BMC Genomics* **7**: 64

**Lefebvre J, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C and Van Wuytswinkel O** (2008) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* **6**: 609-618

**Nicot N, Hausman JF, Hoffmann L and Evers D** (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J Exp Bot* **56**: 2907-2914

**Ramakers C, Ruijter J, Deprez R and Moorman A** (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**: 62-66

**Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J and Cuypers A** (2008) Normalization of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta* **227**: 1343-1349

**Ririe K** (1997) Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Anal Biochem* **245**: 154-160

**Ruuska S, Girke T, Benning C and Ohlrogge J** (2002) Contrapuntal Networks of Gene Expression during *Arabidopsis* Seed Filling. *Plant Cell* **14**: 1191-1206

**Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M,**

**Schölkopf B, Weigel D and Lohmann JU** (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* **37**: 501-506

**Udvardi M, Czechowski T and Scheible W** (2008) Eleven Golden Rules of Quantitative RT-PCR. *Plant Cell* **20**: 1736-1737

**Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F** (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: RESEACH0034

**Waxman S and Wurmbach E** (2007) De-regulation of common housekeeping genes in hepatocellular carcinoma. *BMC Genomics* **8**: 243

## Chapter 5.

### Molecular cloning and functional characterization of a plastidial *sn*-glycerol-3-phosphate acyltransferase from the chilling-tolerant plant *Erysimum asperum*

#### 5.1 Introduction

*sn*-Glycerol-3-phosphate acyltransferase (GPAT, EC 2.3.1.15) catalyzes the first step of glycerolipid assembly by esterifying an acyl chain to the *sn*-1 position of *sn*-glycerol-3-phosphate (G3P). Three types of GPAT exist in higher plant cells, localized in the plastid stroma, ER membrane and mitochondrial membrane (Zheng et al., 2003). In different cell compartments, the natural acyl substrates for GPAT are distinct: in the ER and mitochondria, the natural acyl-donor is acyl-CoA, while in the plastid, it is acyl-ACP (where ACP is acyl-carrier protein) (Murata and Tasaka, 1997). The plastidial GPAT-mediated glycerolipid assembly pathway provides various types of glycerolipids for building the subcellular membranes within the plastid (Schmid and Ohlrogge, 2008). The typical membrane lipids found in plastid are monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG), phosphatidylglycerol (PG), and sulfoquinovosyldiacylglycerol (SQDG) (Ohlrogge and Browse, 1995). Among these lipids, PG is predominantly synthesized within the plastid, while the others can be synthesized either by the plastidial or ER glycerolipid synthetic pathways, depending on the plant species (Ohlrogge and Browse, 1995).

As the major phospholipid in the thylakoid membrane, PG plays a critical role in mediating the chilling-sensitivity of plants (Wolter et al., 1992). Previous

studies revealed that chilling-tolerant plants normally have a higher proportion of unsaturated fatty acids in PG in comparison to chilling-sensitive plants (Murata and Tasaka, 1997). Additionally, it was also confirmed that the fatty acid composition of PG is largely determined by the substrate selectivity of plastidial GPAT. For example, the plastidial GPATs isolated from chilling-tolerant plants such as *Arabidopsis* (Nishida et al., 1993), spinach (*Spinacia oleracea*) (Bertrams and Heinz, 1981; Ishizaki-Nishizawa et al., 1995), and pea (*Pisum sativum*) (Bertrams and Heinz, 1981; Weber et al., 1991) exhibited a substrate preference for unsaturated acyl substrates, resulting in higher levels of unsaturation in PG; in contrast, the plastidial GPATs isolated from chilling-sensitive plants, squash (*Cucurbita moschata*) (Tamada et al., 2004) and *Amaranthus lividus* (Cronan and Roughan, 1987), exhibited substrate preference for saturated acyl substrates. This correlation between plastidial GPAT substrate selectivity and chilling-sensitivity was further confirmed by two studies expressing plastidial GPATs with or without substrate preference for unsaturated fatty acids in transgenic tobacco (Murata et al. 1992) and *Arabidopsis* (Wolter et al., 1992) resulting in changes in the chilling-sensitivities of the plants.

Plastidial GPATs have been purified and cloned from several plants, including *Arabidopsis* (Nishida et al., 1993), cucumber (*Cucumis sativus*) (Johnson et al., 1992), squash (Nishida et al., 1987; Ishizaki et al., 1988), pea (Bertrams M and Heinz E, 1981; Weber et al., 1991) and spinach (Bertrams M and Heinz E, 1981; Ishizaki-Nishizawa et al., 1995). Aside from differences in the amino acid sequence of the N-terminal transit peptides, the mature GPAT proteins

shared a high level of sequence similarity (Murata and Tasaka 1997; Weber et al., 1991). Xu et al. (2006) characterized the biochemical and molecular defect of several *Arabidopsis* plastidial *gpat* (*ats1*) mutant lines and discovered that AT5G10330, along with the plastidial lysophosphatidic acid acyltransferase (LPAAT), plays a role in coordinated regulation of plastidial PG biosynthesis. Their results further revealed the important role of plastidial GPAT in regulating plastid lipid metabolism and plant development.

The present study focuses on molecular cloning and biochemical characterization of a plastidial GPAT isolated from *Erysimum asperum*, which is a wild, chilling-tolerant plant from the Brassicaceae family found in western Canada. The truncated EaGPAT with deletion of the putative transit peptide was functionally expressed in yeast and purified by immobilized metal ion affinity chromatography. Enzymatic assays were performed to determine the optimum *in vitro* reaction conditions for the EaGPAT. The truncated EaGPAT was further assayed with different acyl-CoAs, including 16:0, 16:1 (*cis*  $\Delta^9$ ), 18:0, 18:1 (*cis*  $\Delta^9$ ), 18:2 (*cis*  $\Delta^{9,12}$ ),  $\alpha$ -18:3 (*cis*  $\Delta^{9,12,15}$ ) and 22:1 (*cis*  $\Delta^{13}$ )-CoAs, and exhibited a substrate preference for 18 carbon unsaturated fatty acids. With this substrate preference, the EaGPAT could potentially be used as a biotechnological tool for improving plant chilling-tolerance and increasing unsaturation of seed oil.

## 5.2 Results

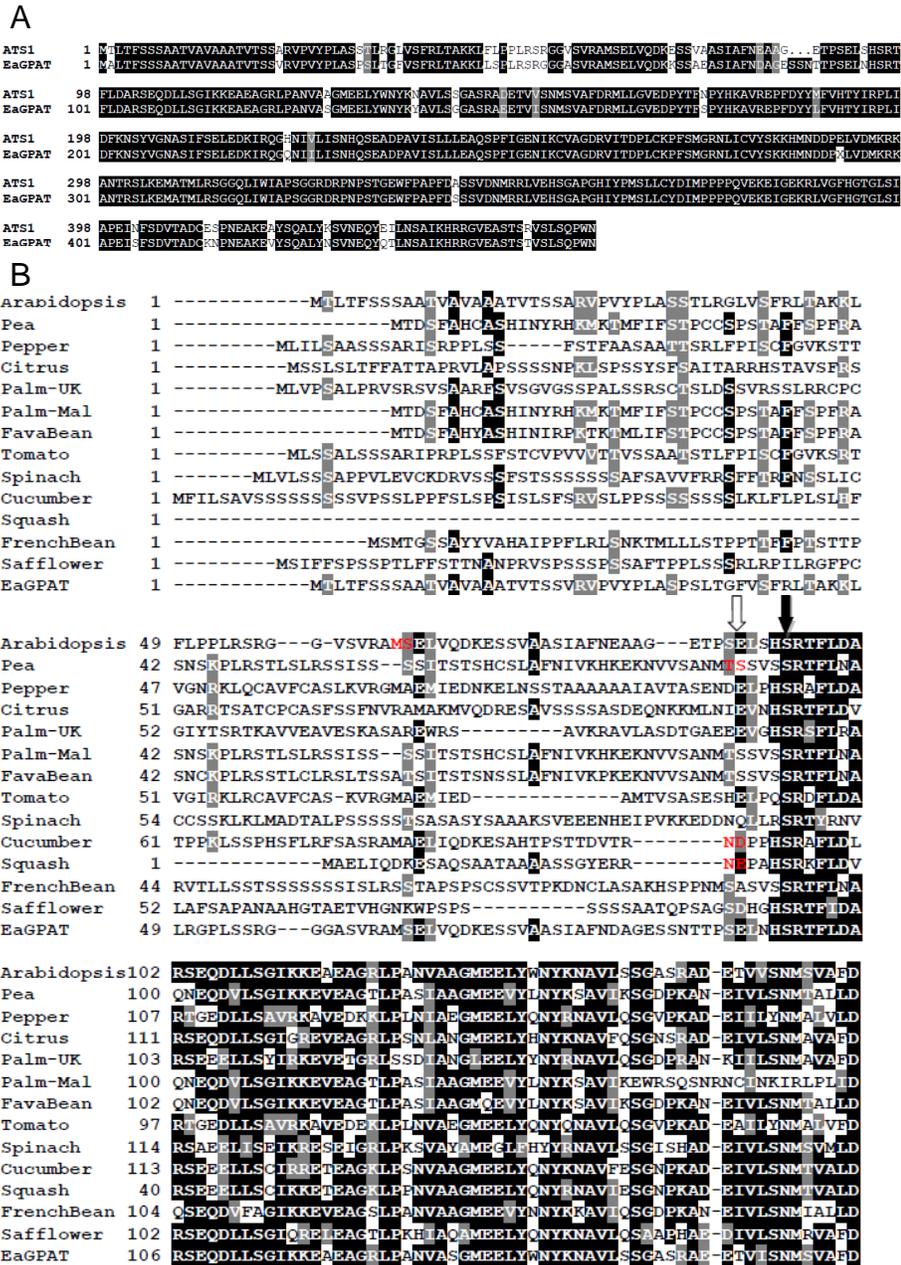
### 5.2.1 Cloning a full-length plastidial GPAT cDNA from *E. asperum*

To isolate the plastidial *GPAT* gene from *E. asperum*, special attention was given to the ESTs obtained from species within the Brassicaceae family.

Based on the sequences of the conserved regions of several known plant plastidial GPAT genes/ESTs, an approximately 350 bp fragment was isolated, which exhibited over 90% sequence similarity to the *Arabidopsis* plastidial GPAT (*ATS1*). The RACE method was used to amplify the 5' and 3' ends for the full-length cDNA. The full-length plastidial *EaGPAT* cDNA was then amplified based on the sequences at the 5' and 3' ends. The full-length *EaGPAT* cDNA is 1389 bp and encodes 462 amino acids, exhibiting 90.5% sequence identity to the amino acid sequence of *ATS1* (Fig. 5.1A).

### **5.2.2 Prediction of transit peptide**

Plastidial proteins encoded by nuclear genes are imported to the plastid under the direction of a transit peptide, which is a short polypeptide located at the N-terminus of the precursor protein. The transit peptide will be cleaved from the mature protein during the import process. The precise cleavage site has not been determined for any of the known plastidial GPAT transit peptides. In previous studies, Ishizaki et al. (1988), Weber et al. (1991) and Nishida et al. (1993) predicted the tentative cleavage sites of transit peptide sequences in the plastidial GPATs from squash, pea and *Arabidopsis*, respectively (Fig. 5.1B, indicated by red color), based on comparison of the estimated molecular masses of GPAT precursors with the corresponding purified mature proteins from the plants.



**Figure 5.1 Amino acid sequence alignment of EaGPAT and other plastidial GPATs.**

(A) Amino acid sequence alignment between ATS1 (*Arabidopsis* plastidial GPAT) and EaGPAT. (B) The N-termini of the plastidial GPATs from different plant species share much lower sequence similarity compared to the rest of their sequences (see full-length sequence alignment in Fig. A8). The red color amino acids and white arrow indicate the previously predicted cleavage sites of transit peptides for the plastidial GPATs from squash, pea, cucumber, and *Arabidopsis*, respectively (Webber et al., 1988; Johnson et al., 1992; Ishizaki et al., 1988; Nishida et al., 1993). The black arrow indicates the putative cleavage site for the EaGPAT transit peptide.

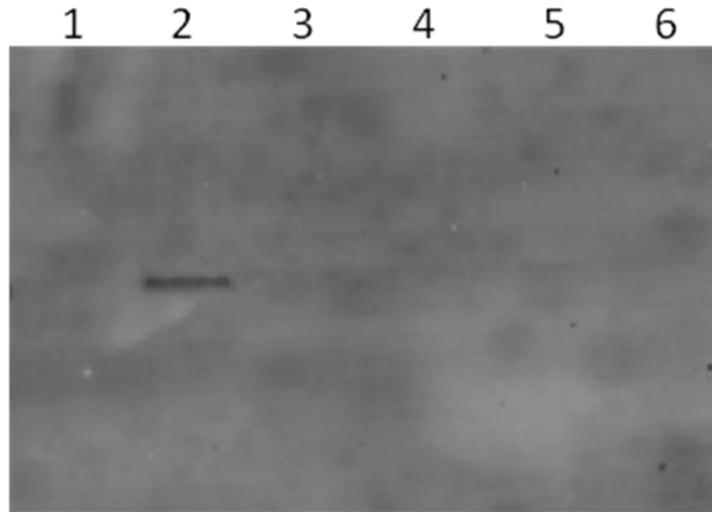
To determine the transit peptide sequence of the EaGPAT, we first compared the precursor amino acid sequences between EaGPAT and a number of plastidial GPATs from different plant species. As shown in Fig. 5.1B and Fig. A8, there is a very limited sequence similarity at the N-termini (the first 80-95 amino acids) between the different plastidial GPAT precursors. This is consistent with a previous study suggesting that the transit peptides of orthologous plastidial proteins tend to have a low conservation (Keegstra et al., 1989). In contrast, after the Ser<sup>98</sup> of EaGPAT (indicated by the black arrow), the EaGPAT shares a much higher sequence similarity to all other GPAT proteins (Fig. 5.1B). Notably, the Ser<sup>98</sup> of EaGPAT is very close to the previously predicted transit peptide processing sites (indicated by white arrow) of the plastidial GPATs from pea (Weber et al., 1991), cucumber (Johnson et al., 1992) and squash (Ishizaki et al., 1988). Thus, it is possible that the first 80-95 amino acids at the N-termini are the transit peptides. We further analyzed the sequence of EaGPAT precursor with a web-based program, WoLF PSORT (Horton et al., 2007), to predict the transit peptide length for the EaGPAT. WoLF PSORT gave a prediction of an N-terminal sequence of 90 amino acids to be the transit peptide. Based on these results, we concluded that the transit peptide processing site was likely located in the vicinity of Thr<sup>90</sup> to Ser<sup>98</sup>.

### **5.2.3 Transit peptide study using fluorescent protein as a fusion tag**

To study the subcellular localization of EaGPAT, we used enhanced green fluorescent protein (eGFP) as a fusion tag to monitor the subcellular localization of the full-length and several N-terminal truncated EaGPATs in transgenic

*Arabidopsis* plants. The purpose of this study was to test the subcellular localizations of the EaGPAT before and after the transit peptides being removed.

The eGFP was fused to the C-termini of the EaGPAT proteins. We expected to observe the eGFP signal in the chloroplasts of lines expressing the full-length EaGPAT-eGFP, and eGFP signal in the cytosol of lines expressing EaGPAT-eGFP without the the transit peptide. Four truncated forms of *EaGPAT* with 5' end deletions of various lengths were generated by PCR and named *EaGPATΔ258*, *EaGPATΔ291*, *EaGPATΔ348*, and *EaGPATΔ432*. The number in each truncated *EaGPAT* referred to the number of the deleted nucleotides. Thus, the corresponding deduced proteins had 86, 97, 116, and 144 amino acids deleted from the N-termini of the full-length EaGPAT precursor, respectively. Several (7-10) independent transgenic lines for each construct were obtained and confirmed by genomic PCR. The expression of different *EaGPAT-eGFP* constructs in transgenic *Arabidopsis* was confirmed by RT-PCR. Seedlings of the transgenic *Arabidopsis* were examined via confocal microscopy. Except for the positive controls that expressed only eGFP, eGFP signal could not be detected in any *EaGPAT-eGFP* transgenic plants. Western blots of crude protein from the transgenic *Arabidopsis* lines using anti-GFP antibody also showed signal only with protein samples prepared from positive control *Arabidopsis* lines (Fig. 5.2).



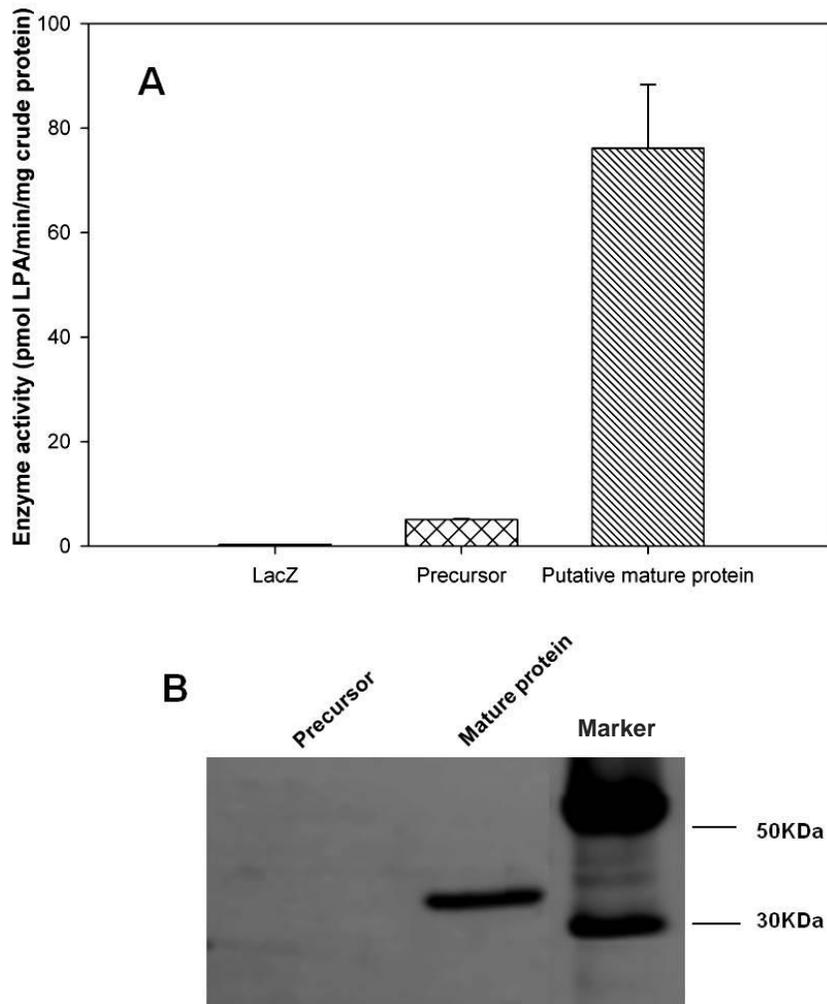
**Figure 5.2. Western blot showing the eGFP signal of protein samples prepared from different *Arabidopsis* lines.**

Lane 1, wild type *Arabidopsis*. Lane 2, transgenic *Arabidopsis* expressing only eGFP (positive control). Lane 3-6, transgenic *Arabidopsis* lines expressing constructs of *EaGPAT1258*, *EaGPAT1291*, *EaGPAT1348*, and *EaGPAT1432*-eGFP, respectively.

#### **5.2.4 The EaGPAT without putative transit peptide exhibited GPAT enzyme activity**

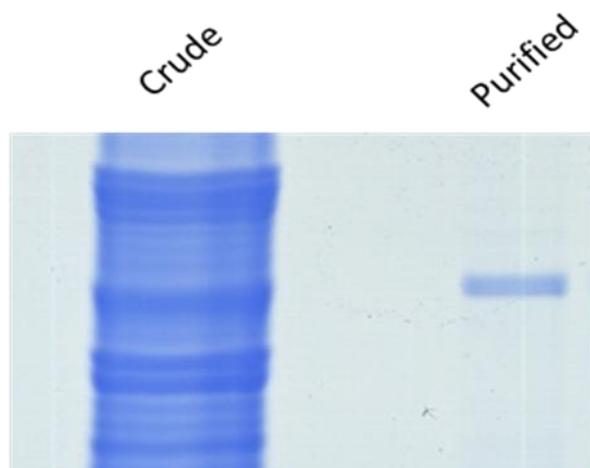
Based on our earlier analysis of different plastidial GPAT sequences, we estimated that the first 97 amino acids at the N-terminus may function as the transit peptide. To test the enzymatic activity of the EaGPAT, two forms of this protein, with and without the putative transit peptide (referred to as precursor and mature protein, respectively), were cloned in frame (without a stop codon) with the V5 epitope and polyhistidine tag of the yeast expression vector, pYES2.1/V5-His-TOPO and over-expressed in yeast strain *gat1Δ*. *LacZ*-transformed *gat1Δ* yeast strain was used as a negative control. The GPAT assays were performed with 100 μM [<sup>14</sup>C (U)]-G3P and 25 μM 18:1-CoA as substrates. As shown in Fig. 5.3, only the mature EaGPAT exhibited a substantial GPAT activity, while the EaGPAT precursor had much lower activity. Furthermore, the mature EaGPAT was detected on Western blot with anti-His antibody as a ~37 KDa band, but the EaGPAT precursor could not be detected (Fig. 5.3).

To further enzymatically characterize the mature EaGPAT without potential interference from other endogenous yeast proteins, we purified the mature EaGPAT from yeast using nickel-chelating resin (ProBond, Invitrogen, Carlsbad, CA, USA). The purity and enzymatic activity of the purified EaGPAT protein were confirmed by SDS-PAGE (Fig. 5.4) and *in vitro* GPAT enzyme assay (Fig. 5.5), respectively.



**Figure 5.3. The yeast-expressing mature EaGPAT exhibited GPAT enzyme activity.**

(A) The GPAT enzyme activities of negative control (LacZ), EaGPAT precursor (full-length protein) and mature EaGPAT (with deletion of the putative transit peptide) after being over-expressed in yeast. The assays were performed with 30  $\mu$ g yeast crude soluble protein, 100  $\mu$ M [ $^{14}$ C (U)] G3P, 25  $\mu$ M 18:1-CoA and HEPES reaction buffer at pH 7.0 for 10 min. (B) Western blots of the yeast-expressed EaGPAT precursor and mature protein.



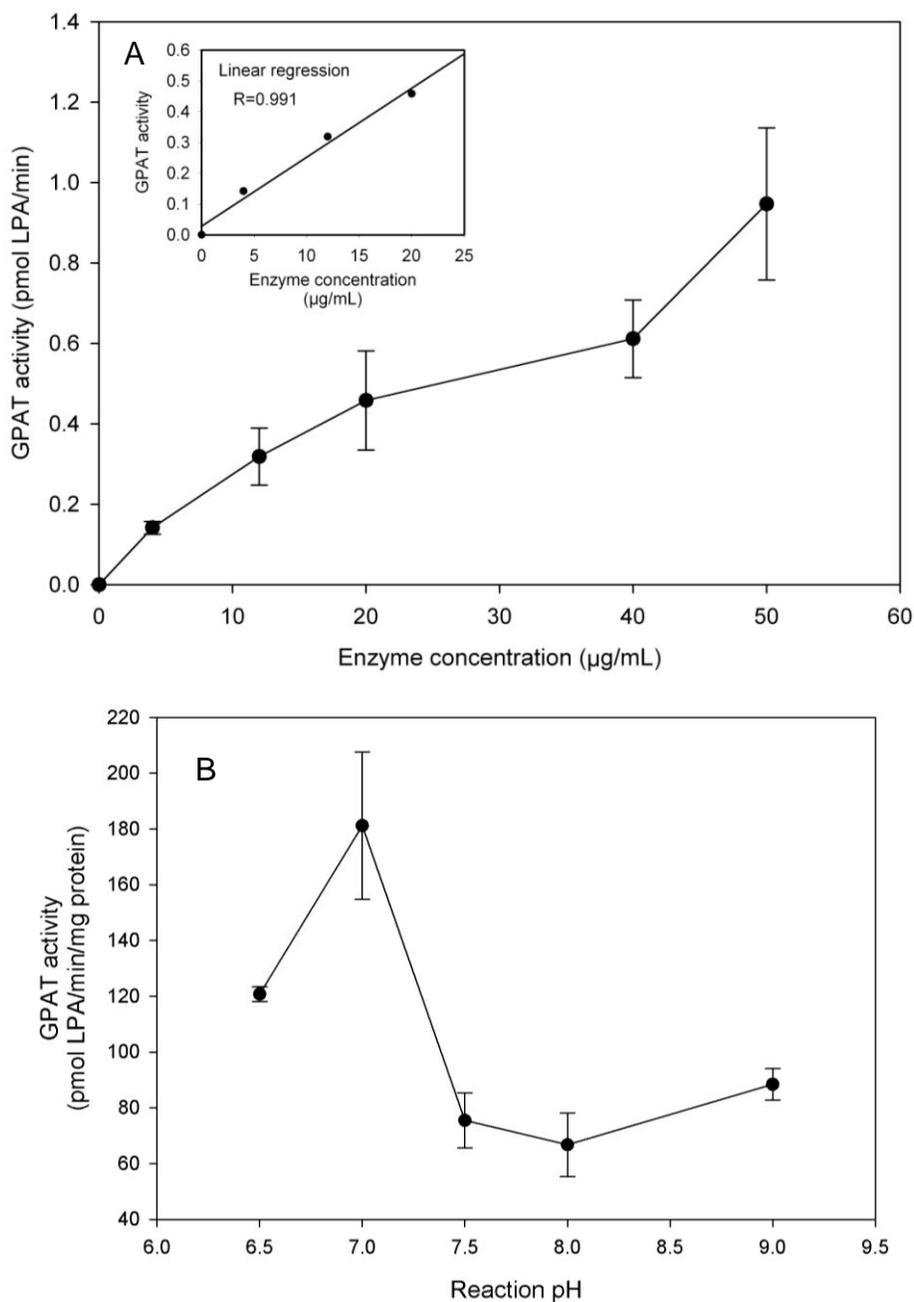
**Figure 5.4. Purification of mature EaGPAT**

SDS-PAGE showing the purity of the purified mature EaGPAT. Left lane: 20  $\mu\text{g}$  of crude soluble protein extracted from the homogenate of yeast expressing the mature EaGPAT. Right lane: 1  $\mu\text{g}$  of purified mature EaGPAT from the homogenate of yeast expressing the mature EaGPAT.

The purified mature EaGPAT was then tested in several types of enzyme assays to determine its dependence on protein concentration, pH, G3P concentration, time, and acyl-CoA specificity.

To determine the linear range of the GPAT activities in relation to the enzyme concentrations, GPAT assays were performed with the purified mature EaGPAT enzyme at a concentration range from 0 to 50  $\mu\text{g/mL}$ . The assays were performed with 100  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 6  $\mu\text{M}$  18:1-CoA and Hepes reaction buffer at pH 7.0 for 10 min. As shown in Fig. 5.5A, the enzymatic activities (pmol LPA/min) of EaGPAT were linear in response to protein concentration ( $r=0.991$ ,  $P<0.05$ ) within the range of 0 to 20  $\mu\text{g/mL}$ . Thus, the specific activity (pmol LPA/min/mg protein) of EaGPAT can be reliably determined using enzyme concentrations within this linear range. An enzyme concentration of 10  $\mu\text{g/mL}$  was used in subsequent experiments.

To determine the optimum pH for EaGPAT activity, GPAT assays were performed in a reaction buffer of Bis-Tris Propane with pH range from 6.5 to 9.0. The assays were performed with 100  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 6  $\mu\text{M}$  18:1-CoA and 10  $\mu\text{g/mL}$  protein for 10 min. As shown in Fig. 5.5B, maximum activity of the mature EaGPAT was observed at pH 7.0. At higher or lower pH, the enzyme activities of EaGPAT were 40-70% of the activity observed at the optimum pH.

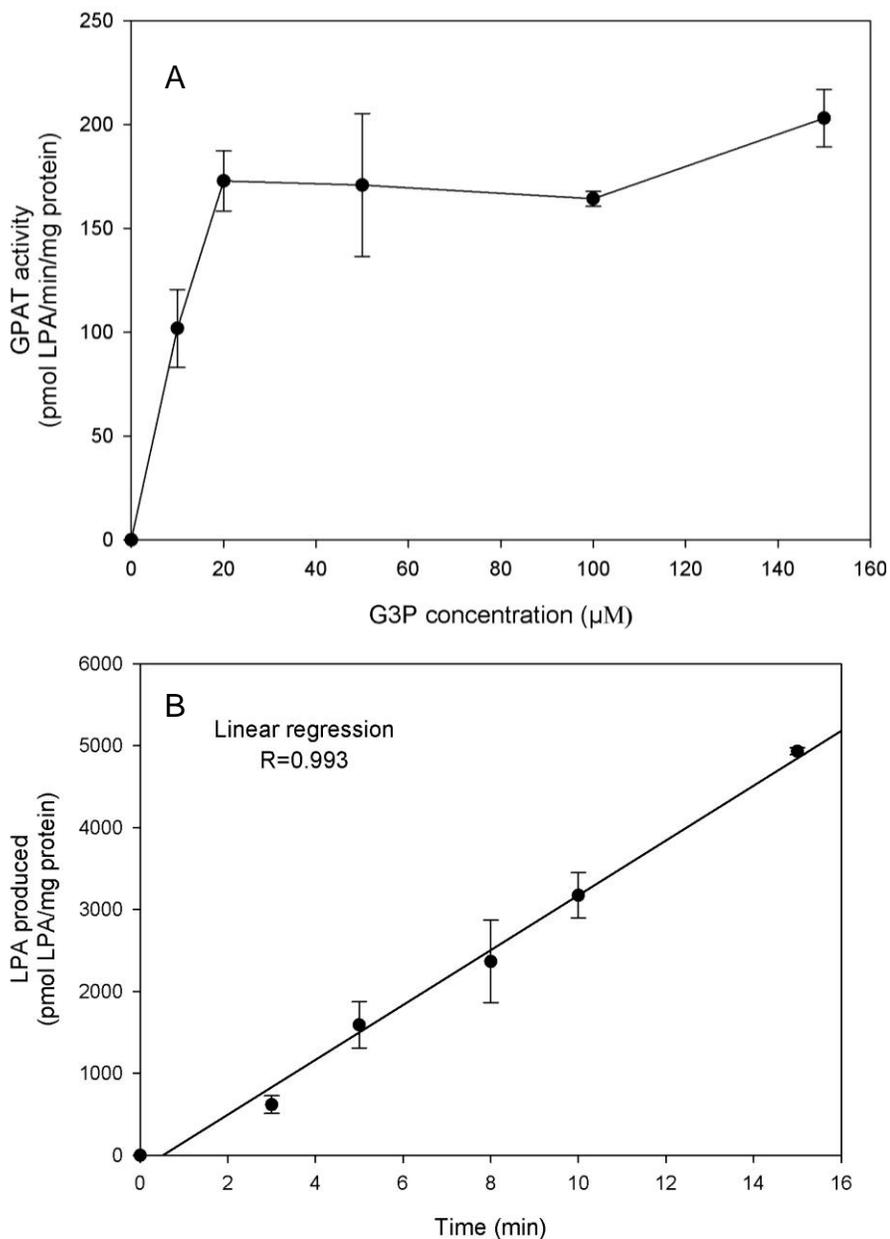


**Figure 5.5. Dependence of EaGPAT activity on concentration (A) and pH (B).**

(A) As shown in the small figure (upper left corner), within a protein concentration range of from 0 to 20 µg/mL, the enzymatic activities of EaGPAT were linear in response to the protein concentration. The assays were performed with 100 µM [<sup>14</sup>C (U)] G3P, 6 µM 18:1-CoA and HEPES reaction buffer at pH 7.0 for 10 min. (B) The optimum pH for EaGPAT enzymatic activity is 7. The assays were performed with 100 µM [<sup>14</sup>C (U)] G3P, 6 µM 18:1-CoA and 10 µg/mL protein for 10 min. Error bars represent standard deviation of triplicate assays.

To determine the mature EaGPAT activity in response to different G3P concentrations, we performed enzyme assays with [ $^{14}\text{C}$  (U)] G3P concentrations from 0 to 150  $\mu\text{M}$ . The assays were performed using 6  $\mu\text{M}$  18:1-CoA, 10  $\mu\text{g}/\text{mL}$  protein and HEPES reaction buffer at pH 7.0 for 10 min. As shown in Fig. 5.6A, when assaying with a G3P concentration of 10  $\mu\text{M}$ , the activity of mature EaGPAT was about 60% of the activities when assaying with G3P at concentrations of equal or higher than 20  $\mu\text{M}$ . Within the G3P concentration range from 20 to 150  $\mu\text{M}$ , the mature EaGPAT exhibited similar enzymatic activities. This result indicated a G3P concentration of more than 20  $\mu\text{M}$  is sufficient for determining the specific activity (pmol LPA/min/mg protein) of the mature EaGPAT.

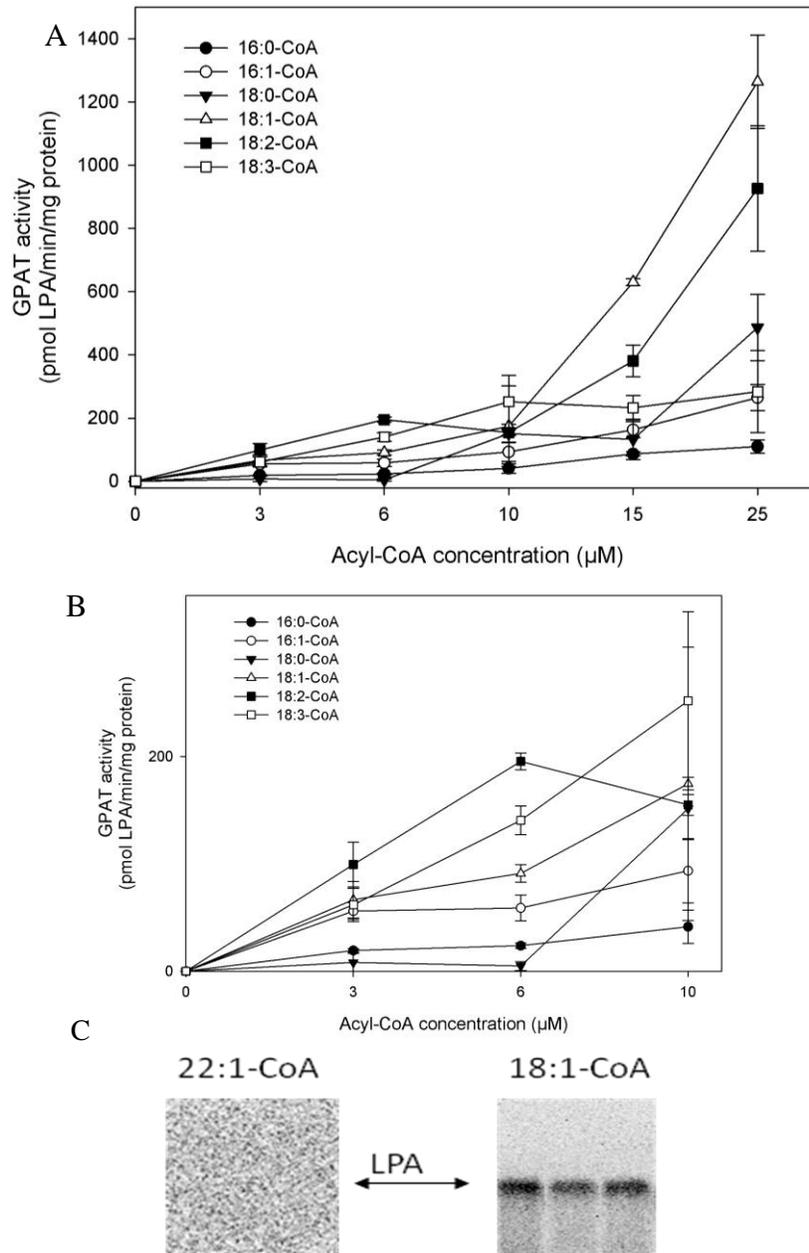
To determine the initial reaction velocity of the enzyme-catalyzed reaction, GPAT assays were performed under different reaction times ranging from 0 to 15min. The assays were performed using 50  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 6  $\mu\text{M}$  18:1-CoA, 10  $\mu\text{g}/\text{mL}$  protein and HEPES reaction buffer at pH 7.0. As shown in Fig. 5.6B, the EaGPAT-catalyzed reaction was linear for at least 15min.



**Figure 5.6. Dependence of EaGPAT activity on glycerol-3-phosphate (G3P) concentration (A) and lysophosphatidic acid (LPA) produced by the catalytic action of EaGPAT as a function of time (B).**

(A) The mature EaGPAT activity was substantially increased when the G3P concentration was increased from 10  $\mu\text{M}$  to 20  $\mu\text{M}$ . When assaying with the G3P concentration range from 20 to 150  $\mu\text{M}$ , the mature EaGPAT exhibited similar enzymatic activities. The assays were performed using 6  $\mu\text{M}$  18:1-CoA, 10  $\mu\text{g}/\text{mL}$  protein and HEPES reaction buffer at pH 7.0 for 10 min. (B) The activities of mature EaGPAT exhibited a linear relationship in response to the reaction time within 15 min. The assays were performed using 50  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 6  $\mu\text{M}$  18:1-CoA, 10  $\mu\text{g}/\text{mL}$  protein and HEPES reaction buffer at pH 7.0. Error bars represent standard deviation of triplicate assays.

Although acyl-ACP is the natural acyl donor in plastids, plastidial GPAT can also use acyl-CoA as substrate, which is often used in the *in vitro* assays of GPAT activity (Slabas et al., 2000; Nishida et al., 2004). To determine the acyl substrate preference of the EaGPAT, we performed enzyme assays with different types of acyl-CoAs including 16:0, 16:1, 18:0, 18:1, 18:2, 18:3 and 22:1-CoAs, at concentrations ranging from 3 to 25  $\mu\text{M}$ . The assays were performed using 50  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 10  $\mu\text{g}/\text{mL}$  protein and Hepes reaction buffer at pH 7.0 for 10 min. In general, EaGPAT exhibited an enhanced specificity for unsaturated acyl-CoAs with the highest activities observed using 18:1 or 18:2-CoA at concentrations greater than 10  $\mu\text{M}$  acyl-CoA. At 6  $\mu\text{M}$  acyl-CoA, the most effective substrates were 18:2 and 18:3-CoA (Fig. 5.7A). Notably, the EaGPAT exhibited no activity when assaying with 22:1-CoA (Fig. 5.7B). EaGPAT activity as a function of acyl-CoA concentration was not a typical hyperbolic response such as the response of activity to G3P concentration shown in Fig. 5.6 A.



**Figure 5.7. Acyl-CoA substrate specificity of the mature EaGPAT.**

In general, the mature EaGPAT had higher activities with unsaturated acyl-CoAs than with saturated acyl-CoAs, but EaGPAT was not able to utilize 22:1-CoA as a substrate. (A) The mature EaGPAT was assayed with different acyl-CoAs of various concentrations. (B) A close-up view of the enzyme activity of the mature EaGPAT with different acyl-CoAs at concentrations ranging from 0 to 10  $\mu\text{M}$ . (C) Phosphor-imaging of the TLC plates showing the lysophosphatidic acid (LPA) produced from GPAT enzyme assays with 15  $\mu\text{M}$  22:1-CoA (triplicates) or with 15  $\mu\text{M}$  18:1-CoA (triplicates) as substrates. The mature EaGPAT was assayed with 22:1-CoA at concentrations ranging from 3 to 25  $\mu\text{M}$  and exhibited no activities. The substrate specificity assays were performed using 50  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P, 10  $\mu\text{g/mL}$  protein and HEPES reaction buffer at pH 7.0 for 10 min. Error bars represent standard deviation of triplicate assays.

## 5.3 Discussion

### 5.3.1 Prediction of the transit peptide of EaGPAT

EaGPAT, like most plastidial proteins, is nuclear-encoded and synthesized in the cytosol as a precursor with an N-terminal transit peptide, which directs the protein across the chloroplast envelope into stroma. Although a number of plastidial *GPATs* have been cloned from different plant species, their transit peptides have not been studied in detail and the sequence information carried by transit peptide is not fully understood (Murata and Tasaka, 1997). In the present study, we used eGFP as a fusion tag to monitor the subcellular localization of full-length and four N-terminal truncated EaGPATs in transgenic *Arabidopsis*; however, we did not detect any eGFP signal in the transgenic plants. Similar to our results, attempts to detect *Arabidopsis* plastidial GPAT (ATS1) by a similar approach were also unsuccessful (Xu et al., 2006) (Personal communication with Xu). The reason is still not clear. Our RT-PCR results indicated the presence of the EaGPAT-eGFP transcripts in transgenic *Arabidopsis* lines; however, we could not detect the fusion protein of EaGPAT-eGFP by Western blot. Thus, it is likely that our current results are associated with the translational or post-translational regulation.

When comparing the plastidial GPAT amino acid sequences from different plant species, we noticed that the N-terminal sequences exhibited much lower similarity in comparison to the rest of the sequences. In addition, none of the common conserved domains of plastidial GPATs (Nishida et al., 1993) are found in the N-terminal region. This, together with the comparison of previously

predicted transit peptide cleavage sites and WoLF PSORT prediction, suggest that the N-terminus of EaGPAT functions as a transit peptide rather than a catalytic domain.

### **5.3.2 Functional expression of the mature EaGPAT in yeast**

Because no plastid exists in the yeast cells, a plastidial GPAT precursor may not be properly expressed in yeast due to the presence of its transit peptide. In the present study, when expressing the full-length EaGPAT in yeast, we could not detect the expression of this protein by Western blot or GPAT activity assay. In contrast, when expressing a truncated EaGPAT, in which the first 97 amino acids at the N-terminus were deleted, we detected the protein on Western blot and confirmed its GPAT enzyme activity. These results suggest that the EaGPAT precursor may not be processed into a stable mature protein in the presence of transit peptide and thus degraded quickly in yeast.

### **5.3.3 Substrate preference of the mature EaGPAT and potential use of enzyme in the metabolic engineering of fatty acid composition in plants**

The *in vitro* GPAT substrate specificity assays indicated that although in general, the mature EaGPAT exhibited higher activities when assayed with unsaturated acyl-CoAs, the mature EaGPAT exhibited no activity with 22:1-CoA. Our earlier analysis of the fatty acid composition of total lipids extracted from the *E. asperum* leaves suggested that only trace amounts of 22:1 were present in the leaf lipids (Fig. A9). Because 22:1 is not a constituent of plastidial membranes, it

is possible that EaGPAT has evolved substrate preferences toward acyl species more commonly found in the plastid.

Due to the substrate preference for unsaturated fatty acyl moieties, plastidial GPATs isolated from chilling-tolerant plants are of potential value for metabolic engineering of plants to improve their chilling-tolerance. By expressing plastidial *GPATs* from the chilling-tolerant plants, *Arabidopsis* and spinach, in rice, the transgenic rice seedlings exhibited higher rates of photosynthesis and growth at low temperature (17/14°C, day/night) in comparison to the wild type rice seedlings (Ariizumi et al., 2002). Additionally, the fatty acid compositions of PG from the wild type and transgenic rice seedlings were also different. There was a significant increase in unsaturated fatty acids in PG, mainly 18:1, 18:2 and 18:3, but not 16:1, in the transgenic rice seedlings. This result is consistent with our observation that the EaGPAT prefers to use 18 carbon unsaturated fatty acyl-CoAs in *in vitro* assays. The amino acid sequence of the *Arabidopsis* plastidial GPAT exhibited 90% similarity with the EaGPAT; thus, it is quite likely that the *Arabidopsis* plastidial GPAT has a very similar substrate preference to EaGPAT. Therefore, this preference for 18 carbon unsaturated acyl substrates may substantially contribute to the fatty acid composition changes of PG in the transgenic rice.

With a substrate preference for unsaturated acyl-CoAs, the EaGPAT could also be potentially useful in genetic engineering of seed oil to increase the incorporation of unsaturated fatty acids into triacylglycerol. A previous study of *Arabidopsis* transformed with mature plastidial GPAT from spinach, showed

relative increases of up to 22% in the seed oil content of transgenic *Arabidopsis* (Jain et al., 2000). This study demonstrated that a mature plastidial GPAT could provide LPA for the storage lipid synthesis pathway in the ER. The physiological concentration of acyl-CoAs in developing seeds of *Arabidopsis* and *Brassica napus* has been estimated to be in the range of 3 to 6  $\mu\text{M}$  (Larson and Graham, 2001). Under these conditions, EaGPAT displayed enhanced preference for 18:2 and 18:3-CoA.

## **5.4 Conclusion**

In conclusion, we have cloned a plastidial *GPAT* from the chilling-tolerant plant *E. asperum*. The mature EaGPAT protein, in which the predicted transit peptide was deleted, was functionally expressed in yeast. The optimum conditions for assaying EaGPAT activity were determined. The mature EaGPAT exhibited a substrate preference for unsaturated acyl-CoAs. This substrate preference suggests that the enzyme may be useful in genetic engineering strategies aimed at improving plant chilling-tolerance or increasing the unsaturated fatty acid content of seed oil.

### **5.5.1 Methods**

#### **5.5.1 Plant materials**

*E. asperum* (western wallflower) plants were grown in a greenhouse under 16 h day/8 h night at 25°C. *Arabidopsis thaliana* (ecotype Columbia) plants were grown in a growth chamber under 18 h day/6 h night at 23 °C.

### 5.5.2 Cloning of plastidial *GPAT* full-length cDNA from *E. asperum*

Total RNA was extracted from the leaf tissues of the plant using RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). First-strand cDNA was thus synthesized using total RNA with QuantiTect reverse transcription kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Based on the sequences of the conserved regions of several known plant plastidial *GPAT* genes, a cDNA fragment of a plastidial *GPAT* gene was amplified with the following primer pair: WWF1F (5'-TCCTCTTTGTAAGCCGTTTCAGTATGGG-3') and WWF1R (5'-TCTTTCTCAACCTGGGGTGGTGGTG-3'). The 5' and 3' ends of the full-length cDNA were amplified by Rapid Amplification of cDNA Ends (RACE) using the GeneRacer Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The primers used for 5' and 3' ends amplification were: WWF5'RACE (5'-GAAGGATTAGGGCGGTCCCTTCCAC-3') and WWF3'RACE (5'-CGGAGAGAAAAG ATTAGTTGGGTTTCACGG-3'). For amplification of the full-length cDNA sequence of the *EaGPAT*, the following primer pair was used: WWF5' (5'-GTAAT GGCTCTCACATTTTCCTCCTCCGCC-3') and WWF3' (5'-ATTCCAGGGTTGTGACAA GGAGACCGTTG-3'). Multiple sequence alignment was performed using ClustalW software (Thompson et al., 1994) and was re-formatted using Boxshade software ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

### 5.5.3 Heterologous expression of the *EaGPAT* in yeast

The precursor (full-length) and mature (with deletion of the putative transit peptide) *EaGPAT* cDNAs were over-expressed in yeast. The cDNA fragment encoding the mature *EaGPAT* was amplified without the first 291 bp (encoding 97 amino acids at the N-terminus) at the 5' end. The full-length and the truncated cDNA fragments of *EaGPAT* were generated by PCR using the following primers: WWF5' (5'-GTAATGGCTCTCACATTTTCCTCCTCCGCC-3') (for full-length cDNA), TWWF5' (5'-GCCATGTCCCGTACCTTCTTGGA TGCGAG-3') (for truncated cDNA fragment) and WWF3' (5'-ATTCCAGGGT TGTGACAAGGAGACCGT TG-3'). Thereafter, the full-length and the truncated forms of *EaGPAT* were cloned in frame without stop codon with the V5 epitope and polyhistidine tag of the yeast expression vector pYES2.1/V5-His-TOPO (Invitrogen, Carlsbad, CA, USA). The constructs of pYES2.1-*EaGPAT*-V5-His were thus transformed into yeast *GPAT* deficient strain *gat1Δ* (Zheng and Zou, 2001; Zheng et al., 2003). The *gat1Δ* strain has only one functional endogenous *GPAT* giving low background activity in the *in vitro* *GPAT* enzyme assay (Zheng et al., 2003). The over-expression of the *EaGPAT* proteins in yeast was induced in galactose medium according to the manual of pYES2.1/V5-His-TOPO expression kit (Invitrogen, Carlsbad, CA, USA).

### 5.5.4 Purification of the mature *EaGPAT*

After growing in induction galactose medium for ~20 h, the yeast cells were harvested for protein extraction. Yeast cells were homogenized with glass beads in a lysate buffer (50mM Tris-HCl pH 8.0, 1mM EDTA, 1mM DTT, 10%

glycerol) as described previously (Zheng and Zou, 2001). The crude homogenate was first centrifuged at  $2500 \times g$ ,  $4^{\circ}\text{C}$  for 10 min to pellet the cell debris, and the supernatant was further centrifuged at  $150,000 \times g$ ,  $4^{\circ}\text{C}$  for 1.5 h to pellet the microsomal fraction. The supernatant was used for enzyme assays and Western blotting. For preparation of purified EaGPAT, yeast cells were homogenized with glass beads in a native binding buffer that was provided in the ProBond purification kit (Invitrogen, Carlsbad, CA, USA). The  $150,000 \times g$  supernatant of the yeast crude protein was used in the ProBond system for further protein purification according the manufacturer's instructions. In brief, the crude protein was first incubated with the nickel-chelating resin in the Native Binding Buffer (containing 10 mM imidazole, pH 8.0) for 60 min at  $4^{\circ}\text{C}$  with gentle shaking. After the resin settled by gravity, the resin was washed sequentially with Native Wash Buffer (containing 20 mM imidazole) at pH 8.0, 7.0 and 6.0, and then eluted with Native Wash Buffer (containing 250 mM imidazole, pH 8.0). Protein concentration was determined using Bradford assay (Bio Rad, Hercules, CA, USA) with bovine serum albumin (BSA) as standard.

#### **5.5.5 *In vitro* GPAT assay**

Unless indicated otherwise, the enzyme was assayed at  $30^{\circ}\text{C}$  for 10 min in 50  $\mu\text{L}$  reaction mixture containing 40 mM Hepes (pH 7.0), 5 mM EDTA, 1 mM DTT, 2.5 mg/mL BSA, 100  $\mu\text{M}$  [ $^{14}\text{C}$  (U)] G3P (60 Ci/mol; American Radiolabeled Chemicals) and 6  $\mu\text{M}$  acyl-CoA. The enzyme reaction was initiated with 40  $\mu\text{g}$  yeast crude soluble protein or 0.5  $\mu\text{g}$  purified protein. The reaction was quenched with 2 mL of chloroform/methanol (1:2, v/v) and 1 mL of 1M KCl

in 0.2 M H<sub>3</sub>PO<sub>4</sub>. The lower organic phase was extracted, dried under nitrogen, re-suspended in 100 µL chloroform, and applied to TLC plates (0.25 mm, SIL G-25, DC-Fertigplatten) which were developed in a solvent system containing chloroform/methanol/acetic acid/5% aqueous sodium bisulfite (75:30:9:3, v/v). The radio-labelled final products ([<sup>14</sup>C]-labeled LPA) on the TLC plates were scraped and subjected to scintillation counting.

#### **5.5.6 Western blot analysis**

Protein samples (40 µg of total protein) prepared from yeast expressing *EaGPAT* or *LacZ* were separated by SDS-PAGE and electrotransferred onto PVDF membranes (Amersham, Arlington Heights, IL, USA). The membranes were then incubated with horseradish peroxidase (HRP) conjugated anti-His antibody (Invitrogen, Carlsbad, CA, USA) at 1:10000 dilution. HRP activity was detected using ECL Plus Kit (GE Healthcare, Piscataway, NJ, USA) according to the manufacturer's instruction. For analysis of transgenic plants, crude protein prepared from transgenic *Arabidopsis* seedlings was prepared for Western blot. About 50-100 mg seedlings were ground in 200 µL of phosphate buffer containing 100 mM PO<sub>4</sub><sup>-</sup> (pH 7.0), 10 mM EDTA, 10% glycerol, 10mM DTT, and 1% Triton X-100. The homogenate was centrifuged at 17,000 g, 4°C for 10 min to pellet the tissue debris. The supernatant was then tested in Bradford assay (Bio Rad, Hercules, CA, USA) to determine its protein concentration, using BSA as standard. About 40 µg of this crude protein was separated by SDS-PAGE and electrotransferred onto PVDF membranes. The HRP conjugated anti-GFP antibody

(Novus, Littleton, CO, USA) at 1:10000 dilution was used for detecting the eGFP signal on the membrane.

### **5.5.7 Preparation of *EaGPAT-eGFP* constructs**

To study the transit peptide of EaGPAT protein, eGFP was used as a fusion tag to monitor the subcellular localization of full-length and four different N-termini truncated EaGPATs in transgenic *Arabidopsis*. Full-length and truncated *EaGPAT* cDNA fragments (with *XbaI* restriction site located at the 5' and 3' ends) were generated by PCR using the following primers: full\_5' (5'-GACTCTAGAAATGGCTCTCACATTTTCCTCC-3'), 258\_5'(5'-CGCTCTAGAAATGAGTAGTAATACGACGCCG-3'), 291\_5'(5'-GCCTCTAGAAATGTCCC GTACCTTCTTGG-3'), 348\_5'(5'-GCCTCTAGAAATG GAAGCTGAAGCTGGAAGG-3'), 432\_5'(5'-GTATCTAGAAATG GGCGGAGCTTCCAGGGC-3'), and EaGP\_3'(5'-GTATCTAGAAATTCCAGGGTTGTGACAAG-3'). The PCR fragments were sequenced and subcloned into the vector PBI121-*eGFP*, which was re-engineered by replacing the original *GUS* gene with *eGFP* in the vector.

### **5.5.8 *Arabidopsis* transformation and screening the transgenic plants**

The constructs of PBI121-*EaGPAT-eGFP* were transformed into wild type *Arabidopsis* plants via *Agrobacterium* mediated floral dip method (Clough and Bent, 1998). Transgenic progeny were selected on ½ MS medium (Murashige and Skoog, 1962) containing 50 mg/L kanamycin. Genomic PCR and RT-PCR were performed to confirm that the target genes were inserted in the plant genome and transcribed. The following primers were used for genomic PCR and RT-PCR:

wwf-3'race130F (5'-ACAATGCTAAGGTCTGGCGGTCAAC-3') and egfp-seqR (5'-CTCGCCGGACACGCTGAACTTGT-3'). The subcellular localization of EaGPAT-eGFP fusion protein was analyzed with a Leica TCS-SP2 multiphoton confocal laser scanning microscope.

## 5.6 References

**Ariizumi T, Kishitani S, Inatsugi R, Nishida I, Murata N, Toriyama K** (2002)

An increase in unsaturation of fatty acids in phosphatidylglycerol from leaves improves the rates of photosynthesis and growth at low temperatures in transgenic rice seedlings. *Plant Cell Physiol* **43**: 751–758

**Bertrams M and Heinz E** (1981) Positional specificity and fatty acid selectivity

of purified *sn*-glycerol 3-phosphate acyltransferases from chloroplasts. *Plant Physiol* **68**: 653-657

**Clough SJ and Bent AF** (1998) Floral dip: a simplified method for

*Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**: 735-743

**Cronan JE and Roughan PG** (1987) Fatty acid specificity and selectivity of the

chloroplast *sn*-glycerol-3-phosphate acyltransferase of the chilling sensitive plant, *Amaranthus lividus*. *Plant Physiol* **83**: 676-680

**Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ and**

**Nakai K** (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* **35**: 585-587

**Ishizaki O, Nishida I, Agata K, Eguchi G and Murata N** (1988) Cloning and nucleotide sequence of cDNA for the plastid glycerol-3-phosphate acyltransferase from squash. *FEBS Lett* **238**: 424-430

**Ishizaki-Nishizawa O, Azuma M, Ohtani T, Murata N and Toguri T** (1995) Nucleotide sequence of cDNA from *Spinacia oleracea* encoding plastid glycerol-3-phosphate acyltransferase. *Plant Physiol* **108**: 1342

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

**Johnson TC, Schneider JC and Somerville C** (1992) Nucleotide sequence of acyl-acyl carrier protein: glycerol-3-phosphate acyltransferase from cucumber. *Plant Physiol* **99**: 771-772

**Keegstra K, Olsen LJ and Theg SM** (1989) Chloroplastic precursors and their transport across the envelope membranes. *Annu Rev Plant Physiol Mol Biol* **40**: 471-501

**Larson TR, Graham IA** (2001) Technical advance: a novel technique for the sensitive quantification of acyl CoA esters from plant tissues. *Plant J* **25**: 115-125

**Murashige T and Skoog F** (1962) A revised medium for rapid growth and bioassays with tobacco cultures. *Physiol Plant* **15**: 473-497

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

**Murata N and Tasaka Y** (1997) Glycerol-3-phosphate acyltransferase in plants.

Biochim Biophys Acta **1348**: 10-16

**Nishida I, Frentzen M, Ishizaki O and Murata N** (1987) Purification of

isomeric forms of acyl-wacyl-carrier-protein: glycerol-3-phosphate

acyltransferase from greening squash cotyledons. Plant Cell Physiol **28**: 1071-

1079

**Nishida I, Tasaka Y, Shiraishi H and 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

**Nishida I, Tamada T, Feese MD, Ferri SR, Kato Y, Yajima R, Toguri T and**

**Kuroki R** (2004) Substrate recognition and selectivity of plant glycerol-3-

phosphate acyltransferases (GPATs) from *Cucurbita moscata* and *Spinacea*

*oleracea*. Acta Cryst **60**: 13-21

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

**Schmid KM and Ohlrogge JB** (2008) Lipid Metabolism in Plants. In DE Vance,

JE Vance, eds, Biochemistry of Lipids, Lipoproteins and Membranes, Ed 5 Vol 4.

Elsevier Science, Amsterdam, The Netherlands, pp 97–130

**Slabas AR, Simon WR, Schierer T, Kroon J, Fawcett T, Hayman M, Gilroy**

**J, Nishida I, Murata N, Rafferty J, Turnbull A and Rice D** (2000) Plant

glycerol-3-phosphate-1-acyltransferase (GPAT): structure selectivity studies.

Biochem Soc Trans **28**: 677-679

- Thompson JD, Higgins DG and Gibson TJ** (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence weighing, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680
- Tamada T, Feese MD, Ferri SR, Kato Y, Yajima R, Toguri T and Kuroki R** (2004) Substrate recognition and selectivity of plant glycerol-3-phosphate acyltransferases (GPATs) from *Cucurbita moscata* and *Spinacea oleracea*. *Acta Cryst* **60**: 13-21
- Weber S, Wolter FP, Buck F, Frentzen M and Heinz E** (1991) Purification and cDNA sequencing of an oleate-selective acyl-ACP: *sn*-glycerol-3-phosphate acyltransferase from pea chloroplasts. *Plant Mol Biol* **17**: 1067-1076
- Wolter FP, Schmidt R and Heinz E** (1992) Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. *EMBO J* **11**: 4685-4692
- Xu C, Cornish AJ, Froehlich JE and Benning C** (2006) Phosphatidylglycerol biosynthesis in chloroplasts of *Arabidopsis* mutants deficient in acyl-ACP glycerol-3-phosphate acyltransferase. *Plant J* **47**: 296-309
- Zheng Z and Zou J** (2001) The initial step of the glycerolipid pathway: Identification of glycerol 3-phosphate/dihydroxyacetone phosphate dual substrate acyltransferases in *Saccharomyces cerevisiae*. *J Biol Chem* **276**: 41710-41716
- Zheng Z, Xia Q, Dauk M, Shen W, Selvaraj G and Zou J** (2003) *Arabidopsis AtGPAT1*, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* **15**: 1872-1887

## Chapter 6.

### General discussion

The present doctoral thesis included three research projects: 1) molecular and functional characterization of three *GPAT4* homologues of *Brassica napus*; 2) survey of internal reference genes for gene expression studies in *B. napus*; 3) enzymatic characterization of a plastidial GPAT from *Erysimum asperum*. The main purpose behind these studies was to better understand the lipogenic enzyme GPAT from molecular, genomic, enzymatic and physiological perspectives. Results from these studies not only contribute to understanding the lipid biosynthetic pathways in plants, but also provide new insights into the physiological significance of the lipogenic enzymes in plant development. All of this knowledge will be valuable for further genetic engineering of oilseed crops. In the following sections each study is summarized and discussed in terms of limitations, future direction and potential applications in crop biotechnology.

#### **6.1 Molecular and functional characterization of three *GPAT4* homologues of *B. napus***

The goal of this study was to investigate the functional divergence of the three *BnGPAT4* homologues and understand their physiological roles in plant development. *B. napus* has one of the most complex plant genomes, mainly due to multiple genome duplications during the evolution of the *Brassica* species. Given that duplicated genes within a single polyploid genome have a high likelihood of evolving into divergent functions (Wendel 2000), it was hypothesized that the

three *BnGPAT4* homologues may play different physiological roles in plant development.

The first approach to test the hypothesis was to examine the gene expression patterns of the three *BnGPAT4* homologues. Taking advantage of the Taqman qRT-PCR system's ability to distinguish the very similar transcript sequences of the *BnGPAT4* homologues, it was demonstrated that the three *BnGPAT4* homologues have very different expression patterns. A common feature of the gene expression patterns of the *BnGPAT4* homologues is that one homologue is expressed predominantly in one tissue/organ, while a different homologue is expressed highly in another type of tissue/organ. For example, among the three homologues, *BnGPAT4-C1* is expressed at the highest level in the inflorescence primordia, while the *BnGPAT4-A1* is expressed at the highest level in the maturing embryos.

The different expression patterns of the *BnGPAT4* homologues strongly suggest that individual *BnGPAT4* homologues should have distinct and specific physiological functions in plant development. One direct approach to further confirm the hypothesis is to analyze the phenotypes of *B. napus* lines in which different individual *GPAT4* homologues were knocked down. Nevertheless, due to the very similar sequences of the *BnGPAT4* homologues, it was only possible to design RNAi constructs targeting suppression of all three *BnGPAT4* homologues. Although the physiological functions of individual *BnGPAT4* homologues could not be compared, the phenotypes of *gpat4* RNAi lines (with

suppression towards all *BnGPAT4* homologues) provided substantial information about the roles of *BnGPAT4* in plant development.

In the constitutive 35S promoter-directed RNAi lines, the major phenotype observed was the cuticle defect on the epidermal cells of the rosette leaves, which further caused the abnormal stomatal opening in certain leaf area that has severe cuticle defect. The conclusion that GPAT4 is involved in lipid polyester (i.e., cutin and suberin) biosynthesis is consistent with the previous study of the *Arabidopsis GPAT4* (Li et al., 2007). The majority of the 35S promoter-directed RNAi lines did not exhibit any morphological differences in comparison to the wild type plants during the later reproductive growth phase. On the other hand, in the RNAi lines directed by the seed-specific napin promoter, severe developmental defects in the inflorescence and siliques were observed, which were not reported in the *Arabidopsis* study. One of the striking phenotypes of these *B. napus gpat4* lines was the aborted development of the inflorescence primordia and young flower buds (Fig. 3.13). Gene expression data also indicated that one *BnGPAT4* homologue (*BnGPAT4-C1*) was expressed at very high levels in the wild type inflorescence primordia (Fig. 3.7A). Based on these results, we conclude that GPAT4 is important for early flower development in *B. napus*. Another notable phenotype of the *B. napus gpat4* lines is the reduced number of total mature siliques and seeds per silique, which were likely caused by defective male fertility imparted by abnormal pollen grains or affected pollen-stigma interaction (Fig. 3.15).

Preliminary seed oil analysis suggested that the suppression of *GPAT4* genes resulted in less 18:1 *cis*  $\Delta^9$  and more 18:2 *cis*  $\Delta^{9,12}$ /18:3 *cis*  $\Delta^{9,12,15}$  being incorporated into the seed oil. Thus, it is possible that the down-regulation of *GPAT4* altered the original seed oil synthesis pathway. Previous studies by other groups proposed that a phosphatidylcholine (PC)-mediated acyl-editing pathway is more dominant for the biosynthesis of high polyunsaturated fatty acids (PUFA)-containing TAG in developing seeds. This pathway channels 60% of the nascent acyl-CoAs to the *sn*-2 position of lysophosphatidylcholine (LPC) to form PC, where double bonds are introduced into the nascent fatty acids through the catalytic action of desaturases (Bates et al., 2009). After this acyl-editing process, some of the modified fatty acids are released from PC back to the acyl-CoA pool. The majority of PC is converted to DAG. Thus, this acyl-editing pathway produces PUFA-enriched acyl-CoA and diacylglycerol (DAG) pools for further TAG biosynthesis (see Literature review, section 2.2.2). The role of GPAT in this pathway or other possible pathways for TAG biosynthesis has not been clarified. GPAT catalyzes the synthesis of lysophosphatidic acid (LPA), which could be further used for TAG synthesis. Based the current results, it is proposed that *GPAT4* may be involved in incorporating the nascent acyl-CoAs (such as 18:1-CoA) into the glycerol backbone for TAG biosynthesis. Therefore, as a result of down-regulation of *GPAT4* enzymes in *B. napus*, more nascent acyl-CoAs may have been channeled to the PC-mediated acyl-editing process for PUFA-enriched TAG synthesis. In summary, this study investigated the evolution and genomic origins of *GPAT4* homologous genes and revealed their functional divergence

within the context of the complex polyploid background of *B. napus*. The physiological role of *GPAT4* has been further probed through investigation of the *B. napus* RNAi lines, which confirmed that the *BnGPAT4* homologues are involved in lipid polyester biosynthesis, early flower development, pollen grain development and possibly storage lipid biosynthesis.

Further work needs to be done to better understand the mechanisms regulating the functional divergence and physiological functions of the *BnGPAT4* homologues. Investigation of tissue/organ-specific cytosine DNA methylation patterns, in combination with gene expression analysis, will provide useful information for understanding how DNA methylation regulates the transcription of individual *BnGPAT4* homologues. Suppression of the individual *GPAT4* homologues in *B. napus* may reveal their specific physiological functions for plant development. Further detailed seed oil analysis of the *gpat4 B. napus* T<sub>2</sub> and T<sub>3</sub> generations will provide more information to confirm the involvement of *GPAT4* in storage lipid biosynthesis. If *GPAT4* has a role in diverting 18:1, from 18:1-CoA, into LPA, it is possible that the *GPAT4* enzyme variants have increased selectivity for 18:1-CoA compared to polyunsaturated acyl-CoAs. In this regard, *in vitro* substrate specificity assays of *BnGPAT* enzymes with 18:1-CoA and polyunsaturated acyl-CoAs may provide useful information. In addition, given that some *gpat4* lines also give deformed mature embryos, detailed examination of the developing embryos may reveal another physiological role of the *GPAT4* for embryo development.

## **6.2 A survey of quantitative real-time PCR internal reference genes for expression studies in *B. napus*.**

This study evolved with the study on gene expression of *BnGPAT4* homologues, providing a number of stable internal reference genes for qRT-PCR studies in *B. napus*.

Although qRT-PCR is a powerful technique for gene expression studies, its usefulness depends on good internal reference and accurate data interpretation. In the earlier phase of this research, it became apparent that the choice of reference genes used to normalize qRT-PCR data is particularly critical when dealing with samples from different organs or developmental stages of the plant. There have been a number of studies investigating internal reference genes for qRT-PCR applications with different plant species, including *Arabidopsis* (Czechowski et al., 2005), tomato (Expósito-Rodríguez et al., 2008), soybean (Libault et al., 2008), wheat (Paolacci et al., 2009) and rice (Jain et al., 2006), but not in *B. napus*. A program that has been extensively used in these studies for evaluating reference genes is geNORM. The geNORM algorithm has gained much popularity since being developed by Vandesompele et al. (2002), and is frequently applied in studies of evaluating reference genes for different species (<http://medgen.ugent.be/~jvdesomp/genorm/>). It was developed based on the principle that the expression patterns of ideal internal reference genes, assuming they are not co-regulated, are identical in all samples, no matter the sample type or experimental condition (Vandesompele et al., 2002). Relying on this principle, non-normalized gene expression data from the candidate reference genes are

compared with each other in a pairwise fashion by the geNORM algorithm. Thus, for each candidate reference gene, the pairwise variations with all other reference genes are determined as standard deviations. A calculated  $M$  value given by geNORM indicates the average standard deviations for individual candidate reference genes. Thus, a lower  $M$  value indicates more stable expression of a candidate reference gene (Vandesompele et al., 2002).

In current study, eight candidate reference genes were selected with the potential to serve as internal references in qRT-PCR, focusing primarily on genes with stable expression in the embryo during the period of oil synthesis and accumulation. The transcription levels of the candidate genes were also assessed in seedlings, young and mature leaves, flowers and anthers. The eight candidate genes were ranked according to the  $M$  values obtained by geNORM analysis using data from vegetative and embryo samples.

In summary, this study provided useful information for laboratories performing qRT-PCR gene expression studies in *B. napus*, and will expedite the adoption and validation of more suitable reference genes in other *Brassica* species.

### **6.3 Characterization of a recombinant plastidial GPAT from *E. asperum***

The chilling tolerance of a plant is partially determined by its cell membrane fluidity, which is controlled by the unsaturated fatty acids on the membrane glycerolipids (Wolter et al., 1992). A number of studies have shown that plastidial GPAT is involved in the biosynthesis of PG, a major phospholipid in the plastidial membrane (Murata and Tasaka, 1997). Thus, the substrate

preference of a plastidial GPAT could partially determine the fatty acid composition of PG, and hence affect the fluidity of the chloroplast membranes (Wolter et al., 1992; Murata et al., 1992).

Based on this knowledge, the rationale behind this study was to explore the possibility of using a plastidial GPAT with preference for unsaturated acyl substrates to decrease the saturation level in *B. napus* seed oil. Towards this goal, a plastidial *GPAT* from a chilling tolerant plant *E. asperum* was cloned and characterized. As a nuclear-encoded plastidial enzyme, the EaGPAT has a transit peptide at the N-terminus for directing the protein precursor to translocate to the chloroplasts. Thus, it was necessary to determine the transit peptide sequence of the EaGPAT before this GPAT could be applied to seed oil biosynthesis, which primarily takes place in the ER. Previous studies in *Arabidopsis* and other plants have determined the tentative transit peptide sequences by comparing the polypeptide sizes of the precursor and the plant purified mature protein (Murata and Tasaka, 1997). In this study, a different approach was applied wherein constructs of EaGPAT (with different length of truncations at the N-terminus)-eGFP were designed for monitoring the subcellular localization of the fusion proteins in transgenic *Arabidopsis* plants. Fusion proteins, however, could not be detected in these plants either by fluorescence microscopy or Western blotting. Similar results were obtained in another study of *Arabidopsis* plastidial GPAT by Xu et al. (2006) (unpublished data, personal communication with Dr. Xu). As a consequence, it was only possible to estimate the transit peptide sequence of the EaGPAT by comparing the amino acid sequences between EaGPAT and a

number of known plastidial GPATs from different plant species. We further confirmed the enzyme activity of recombinant EaGPAT following removal of the putative transit peptide.

Recombinant truncated EaGPAT was purified from yeast extract and characterized in terms of dependence of LPA production on protein concentration, dependence of GPAT activity on pH, time course of LPA production and acyl-CoA specificity. EaGPAT activity was linear from 2 to 20  $\mu\text{g/mL}$  protein in the reaction mixture. The enzyme exhibited a pH optimum of 7 and the reaction was linear for at least 15 min. At concentrations of 3 and 6  $\mu\text{M}$  acyl-CoA, EaGPAT exhibited enhanced substrate preference for 18:2 and 18:3-CoA in comparison to 16:0, 16:1, 18:0, and 18:1-CoAs (Fig. 5.7B). At concentrations above 10  $\mu\text{M}$  acyl-CoA, the most effective substrates were 18:1 and 18:2-CoAs (Fig. 5.7A). This result confirmed the hypothesis that a plastidial GPAT from a chilling-tolerant plant prefers unsaturated acyl-CoAs.

The substrate preference for 18:1 and 18:2-CoAs suggests a potential for applying this EaGPAT for future biotechnological applications, such as modification of the unsaturated fatty acid content in the seed oil. Further work could be focused on investigating how to incorporate EaGPAT into the storage lipid synthesis pathway of *B. napus* in order to increase the unsaturation level in seed oil. In *B. napus* seed oil, the content of saturated fatty acid (mainly 16:0) is about 7 mol%, half of which is believed to be located at the *sn*-1 position of TAG (Brockerhoff and Yurkowski, 1966; Booth and Gunstone, 2004). In one of the previous studies from this laboratory, targeting the conversion of 16:0 to 16:1

through manipulation of plastidial desaturase activity resulted in increased content of 16:1 and its downstream elongation products, but no decrease in the level of 16:0 in seed oil (Bondaruk et al, 2007). These results suggested that 16:0-CoA could not be eliminated from seed oil by the desaturase's action of converting 16:0 to 16:1. Since 50% of the 16:0 is enriched at the *sn*-1 position of TAG, a different strategy proposed by this team is to explore whether manipulating the acyl substrate preference (unsaturated vs saturated) of GPAT would regulate the incorporation of 16:0 to TAG. The EaGPAT characterized in the current thesis exhibits a preference for unsaturated 18 carbon acyl groups, which may be useful for increasing the overall level of unsaturated fatty acids in seed oil. In order to enable the EaGPAT to have a significant impact on the fatty acid composition at the *sn*-1 position of TAG, a few major issues need to be carefully considered. First, competition by enzymes in the endogenous biosynthetic pathways could substantially dilute the transgene's impact on the targeted pathway. Therefore, the endogenous enzymatic action for incorporating 16:0 to TAG has to be down-regulated seed-specifically. It is possible that an endogenous GPAT mediates the incorporation of 16:0 to the *sn*-1 position of TAG. Second, the plastidial EaGPAT needs to be expressed under the direction of a seed-specific promoter and be targeted to the ER. In addition to removing the plastidial transit peptide, a small polypeptide tag could be added to the C-terminus for the purpose of redirecting and attaching the EaGPAT to the ER. Third, the expression of the EaGPAT could be coupled with the suppression of the endogenous acyl-ACP thioesterase B (FATB), which releases 16:0 from plastidial fatty acid synthase (Buhr et al.,

2002), in order to decrease the 16:0 level in the acyl-CoA pool and hence to promote the usage of 18 carbon unsaturated acyl-CoAs by the acyl-CoA-dependent acyltransferases. Results from these studies will also provide valuable information for better understanding the role of GPAT in TAG biosynthesis.

#### **6.4 Concluding comments**

As an important oilseed crop, *B. napus* has become a research focus for modifying seed oil using modern biotechnology. Successful strategies for seed oil modification are developed based on a sound understanding of the mechanisms behind the lipid synthetic pathway. Results from the present doctoral thesis contribute to our knowledge of the physiological functions and biochemical properties of the lipid synthesis-related enzyme GPAT. This knowledge will be useful for further development of new biotechnological approaches to modify the seed oil of crops.

#### **6.5 References**

- Bates PD, Durrett TP, Ohlrogge JB and Pollard M** (2009) Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. *Plant Physiol* **150**: 55-72
- Bondaruk M, Johnson S, Degafu A, Boora P, Bilodeau P, Morris J, Wiehler W, Foroud N, Weselake R and Shah S** (2007) Expression of a cDNA encoding palmitoyl-acyl carrier protein desaturase from cat's claw (*Doxantha unguis-cati* L.) in *Arabidopsis thaliana* and *Brassica napus* leads to accumulation of unusual unsaturated fatty acids and increased stearic acid content in the seed oil. *Plant Breed* **126**: 186-194

- Booth EJ and Gunstone FD** (2004) Rapeseeds and rapeseed oil: agronomy, production, and trade. In FD Gunstone, ed, Rapeseed and canola oil: Production, processing, properties and uses. Blackwell Publishing, Oxford, UK, pp 1-15
- Brockerhoff H and Yurkowski M** (1966) Stereospecific analyses of several vegetable fats. *J Lipid Res* **7**: 62-64
- Buhr T, Sato S, Ebrahim F, Xing A, Zhou Y, Mathiesen M, Schweiger B, Kinney A, Staswick P and Clemente T** (2002) Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. *Plant J* **30**: 155-163
- Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR** (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* **139**: 5-17
- Expósito-Rodríguez M, Borges A, Borges-Pérez A and Pérez J** (2008) Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC plant biol* **8**: 131
- Jain M, Nijhawan A, Tyagi AK and Khurana JP** (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochem Biophys Res Commun* **345**: 646-651
- Li YH, Beisson F, Koo AJK, Molina I, Pollard M and Ohlrogge J** (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc Natl Acad Sci USA* **104**: 18339-18344

- Libault M, Thibivilliers S, Bilgin D, Radwan O, Benitez M, Clough S and Stacey G** (2008) Identification of four soybean reference genes for gene expression normalization. *Plant Genome* **1**: 44-54
- Murata N and Tasaka Y** (1997) Glycerol-3-phosphate acyltransferase in plants. *Biochim Biophys Acta* **1348**: 10-16
- Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y and Nishida I** (1992) Genetically engineered alteration in the chilling sensitivity of plants. *Nature* **356**: 710-713
- Paolacci A, Tanzarella O, Porceddu E and Ciaffi M** (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC mol biol* **10**: 11
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F** (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: RESEACH0034
- Wendel JF** (2000) Genome evolution in polyploids. *Plant Mol Bio* **42**: 225-249
- Wolter FP, Schmidt R and Heinz E** (1992) Chilling sensitivity of *Arabidopsis thaliana* with genetically engineered membrane lipids. *EMBO J* **11**: 4685-4692
- Xu C, Cornish AJ, Froehlich JE and Benning C** (2006) Phosphatidylglycerol biosynthesis in chloroplasts of *Arabidopsis* mutants deficient in acyl-ACP glycerol-3-phosphate acyltransferase. *Plant J* **47**: 296-309

## Appendix 1

**Table A1. Primers used in the Chapter 3**

<b>Gene-specific primers for full-length cDNA cloning</b>		
	Forward	Reverse (without stop codon)
<i>BnGPAT4-C1</i>	GAAATGTCTCCGGCGAAGAAGA	GCTGGATTTATCATCCTTTGTCTTCTTG
<i>BnGPAT4-C2</i>	GAGAGATCGAAAAATGTCTCCGGCA	CTCCTTGGCCTTGCTCTTATTGATAG
<i>BnGPAT4-A</i>	CCATCAAAGAGTGATGTCTCCGGCG	CTCCTTGGCCTTGCTCTTATTGATAG
<b>Cloning full-length cDNA into yeast expression vector <i>pYES2.1</i></b>		
	Forward	Reverse (without stop codon)
<i>BnGPAT4-C1</i>	GAAATGTCTCCGGCGAAGAAGA	GCTGGATTTATCATCCTTTGTCTTCTTG
<i>BnGPAT4-C2</i>	same as above*	CTCCTTGGCCTTGCTCTTATTGATAG
<i>BnGPAT4-A</i>	same as above	CTCCTTGGCCTTGCTCTTATTGATAG
* This primer included a single nucleotide mutation compared to the original sequence of C2 (the 15th nucleotide counted from 5', A was changed to G), yet the deduced amino acid did not change		
<b>Gene-specific primers for genomic DNA cloning</b>		
	Forward	Reverse
<i>BnGPAT4-C1</i>	GAAATGTCTCCGGCGAAGAAGA	GCTGGATTTATCATCCTTTGTCTTCTTG
<i>BnGPAT4-C2</i>	GAGAGATCGAAAAATGTCTCCGGCA	GCTGGATTTATTACTCCTTGGCCTTGC
<i>BnGPAT4-A</i>	CCATCAAAGAGTGATGTCTCCGGCG	GCTGGATTTATTACTCCTTGGCCTTGC

<b>Cloning individual promoters into PBI121-GUS-eGFP</b>		
	Forward ( <i>HindIII</i> restriction site was underlined)	Reverse ( <i>XbaI</i> restriction site was underlined)
<i>BnGPAT4-C1</i>	TACA <u>AAGCTT</u> GAATGCCACCCCTAACATCATCTATAA	ACT <u>TCTAGAT</u> TTTCTTTCTTTCTCTCTTGAGTC
<i>BnGPAT4-C2</i>	TACA <u>AAGCTT</u> CAGGGGATCAGTTATTTTCATCA	ACT <u>TCTAGAT</u> TTTTCGATCTCTCTTTGATGG
<i>BnGPAT4-A1</i>	TACA <u>AAGCTT</u> TTCATCTACATAGTTGGATACACAGC	ACG <u>TCTAGAC</u> ACTCTTTGATGGTAAACAG
<b>Gene-specific primers for cloning of the putative 4th GPAT4 in <i>B. napus</i></b>		
	Forward	Reverse
<i>BnGPAT4-X</i>	TTACCTCTTCGTCTCCGAGGCTA	CCTTGCAAATGGACATGAAGTCG
<b>SYBR-green qRT-PCR for analyzing the GPAT4 homologue copy number in <i>B. napus</i></b>		
Pea <i>rbcS</i> terminator	GGTTTTCGCTATCGAACTGTGA;	TGAGAATGAACAAAAGGACCATATCA
<i>BnGPAT4</i>	CTACGAGGCCACTTTCTTGGA	AGCCACCTCGATTGGAGTTTT
<b>Cytosine methylation analysis</b>		
	Forward	Reverse
<i>BnGPAT4-C1_1</i>	GAYAAAAAGAYTYAAGAGAGAAAGAAAG	CTTCACGAAARRCTCAACCATCACTAT
<i>BnGPAT4-C1_2</i>	GTGAAGGATTAYTTGGGAGGAGAYAAAGTTTTAG	CCAAGATCTGGTGATTCTGTCTCCAAACTCCTTT
<i>BnGPAT4-C1_3</i>	GAAAAGATTAGYYATYTTAAAGGAGTTTGGAG	ACAAATTATTRTATAATCTATCCCACTA
<i>BnGPAT4-C2_1</i>	TTGTAAAGAAGTYATATTGGATAAGGAG	ATCRAATCATACTCTCCACRTTTACACTC
<i>BnGPAT4-C2_2</i>	GCGAGTGTAACGTGGAGAGTATGATT	AARAACACCARRCTTCTTCACAAACCCC
<i>BnGPAT4-C2_3</i>	GATGAAAGYYAYGGGTTTTGTGAAGAAG	TTTATRCACCTCRATTTTCACCTAAACC
<i>BnGPAT4-A1_1</i>	ATTAGTAAATATGTYATATATTGGATAGGAA	ATARAATCATACTCTCCACTTTTACACTC

<i>BnGPAT4-A1_2</i>	GCGAGTGTAAAAGTGGAGAGTATGATT	CCTTAACGAAARCCTCAACCATCACTATC	
<i>BnGPAT4-A1_3</i>	GAYTGYTAATYYGATAGTGATGGTTGAGG	TTTATRCACTCAACTTAAATCTAAATCTC	
*note: Y represent CT and R represented AG			
<b>Taqman qRT-PCR for transcription analyses of individual homologues</b>			
	Forward	Reverse	Taqman MGB Probes
<i>BnGPAT4-C1</i>	CTACGAGGCCACTTTCTTGGA	AGCCACCTCGATTGGAGTTTT	(FAM)CCTGAAGAAATGACGG
<i>BnGPAT4-C2</i>	CCTGAAGGCACCACTTGTAGAGA	CGATCCGGTCGCTGAGCT	(TET)TATCTACTACGATTTAGCGCT
<i>BnGPAT4-A1</i>	CTGAAGGCACGACTTGTAGAGAAC	CGTACTGTGGTTCCGTTGAACA	(TET)TATCTACTACGATTCAGTGCT
<i>TIP41</i>	GCAGATTGATTTGGCTGCTCTT	CATCTTCATACAAAACCTCATCATAGA	(VIC)CACTTAAAGAACCTATTCTC
<i>ACT7</i>	TGGGTTTGCTGGTGACGAT	TGCCTAGGACGACCAACAATACT	(VIC)CTCCCAGGGCTGTGTT
<i>UBC21</i>	CAAGGGACCGTCV*GAGACT	TTCGGGAACAGCRAAB*GC	(VIC)CTTATGAAGGCGGAGTGTT
*V represented ACG and B represented CGT			
<b>Primers for RNAi construct targeting to <i>BnGPAT4</i> homologues</b>			
	Forward (restriction enzyme sites were underlined)	Reverse (restriction enzyme sites were underlined)	
<i>BnGPAT4_sense</i>	CAGT <u>CTCGAGTCG</u> TCTCTCGCGCCGTTCTTCCA	CAGCGGT <u>ACCGCTAATCTCTTCAAGTCTCCAAC</u>	
<i>BnGPAT4_anti-sense</i>	CAGT <u>TCTAGACG</u> TCTCTCGCGCCGTTCTTCCA	CAGCAAGCTT <u>GCTAATCTCTTCAAGTCTCCAAC</u>	
CaMV35S	CAGC <u>GAGCTCGG</u> TCCCCAGATTAGCCTTTTC	CAGC <u>CTCGAGG</u> TCCCCCGTGTCTCTCCAAATGA	
<b>SYBR-green qRT-PCR for analyzing the overall transcription level of <i>BnGPAT4</i> homologues</b>			
	Forward	Reverse	
<i>BnGPAT4</i>	CTACGAGGCCACTTTCTTGGA	AGCCACCTCGATTGGAGTTTT	

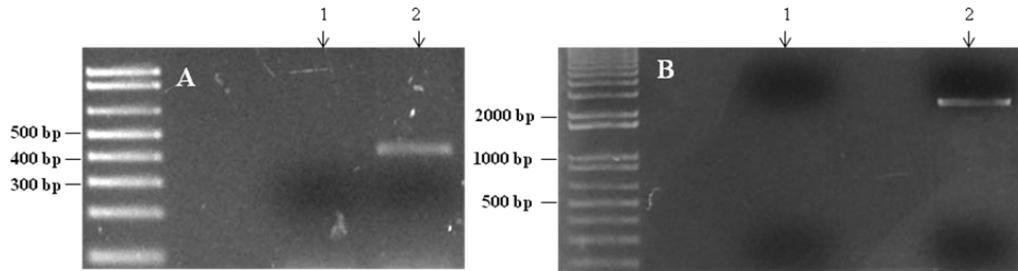
<i>UBC9</i>	GCATCTGCCTCGACATCTTGA	CGATAGCAGCACCTTGGAGATA
<i>UPI</i>	AGCCTGAGGAGATATTAGCAGGAA	ATCTCACTGCAGCTCCACCAT
<i>TIP41</i>	AGAGTCATGCCAAGTTCATGGTT	CCTCATAAGCACACCATCAACTCTAA

**Table A2. Raw data of PCR efficiency, Ct, RQ and s.e. value of biological replicates for tested samples.**

	Alien RNA			UBC9				ACT7				PP2A				UP2							
	1+E <sup>a</sup>	Ct <sup>b</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>
1-2 day seedlings	1.97	23.40	3	2.02	20.24	5.37	0.40	3	2.03	23.18	0.59	0.05	3	1.91	23.68	1.79	0.16	3	1.91	24.15	1.29	0.21	3
5-day seedlings	1.91	23.34	3	2.00	20.45	2.41	0.07	3	2.02	23.54	0.24	0.00	3	1.98	24.39	0.22	0.00	3	1.97	25.02	0.14	0.03	3
10-day seedlings	1.93	23.24	3	1.96	20.61	4.26	0.52	3	2.02	23.33	0.38	0.11	3	1.96	24.34	0.33	0.03	3	1.89	26.09	0.58	0.05	3
Young leaves	1.95	23.16	3	1.96	21.07	3.50	0.91	3	2.03	23.39	0.34	0.01	3	1.90	23.76	1.20	0.37	3	1.92	25.42	0.34	0.02	3
Mature leaves	1.95	23.42	3	1.98	21.30	3.04	0.76	3	2.01	24.26	0.28	0.09	3	1.98	23.35	0.76	0.19	3	1.91	24.48	0.77	0.06	3
Anthers	1.92	23.26	3	1.98	20.41	3.42	0.58	3	2.02	27.11	0.02	0.01	3	1.92	24.51	0.42	0.10	3	1.96	25.38	0.15	0.04	3
Flowers	1.87	23.20	3	1.99	20.70	1.35	0.19	3	2.00	24.26	0.10	0.04	3	2.05	24.39	0.05	0.02	3	1.83	24.31	0.82	0.08	3
20-day embryo	1.93	23.33	3	2.00	21.02	2.21	0.52	3	2.02	21.51	1.23	0.38	3	1.96	23.52	0.66	0.16	3	1.90	24.58	0.63	0.09	3
25-day embryo	1.88	23.26	3	1.95	22.00	0.91	0.16	3	2.00	22.90	0.30	0.01	3	1.95	24.60	0.16	0.01	3	1.85	25.91	0.28	0.03	3
30-day embryo	1.92	23.32	3	2.00	23.85	0.27	0.05	3	2.06	23.29	0.19	0.02	3	1.97	24.94	0.19	0.02	3	1.92	27.08	0.09	0.01	3
35-day embryo	1.93	23.35	3	2.05	23.81	0.18	0.01	3	2.00	24.17	0.24	0.07	3	1.94	25.50	0.22	0.05	3	1.86	27.27	0.21	0.03	3
40-day embryo	1.94	23.24	3	2.02	21.82	1.06	0.04	3	1.96	23.06	0.89	0.21	3	1.97	24.59	0.27	0.04	3	2.02	26.05	0.05	0.00	3
45-day embryo	1.87	23.43	3	2.03	21.77	0.48	0.01	3	2.01	23.93	0.14	0.03	3	1.91	24.48	0.33	0.05	3	1.93	26.59	0.06	0.01	3
50-day embryo	1.91	23.54	3	1.98	21.85	1.35	0.23	3	1.97	24.21	0.32	0.10	3	1.92	25.04	0.35	0.09	3	1.92	26.87	0.10	0.01	3

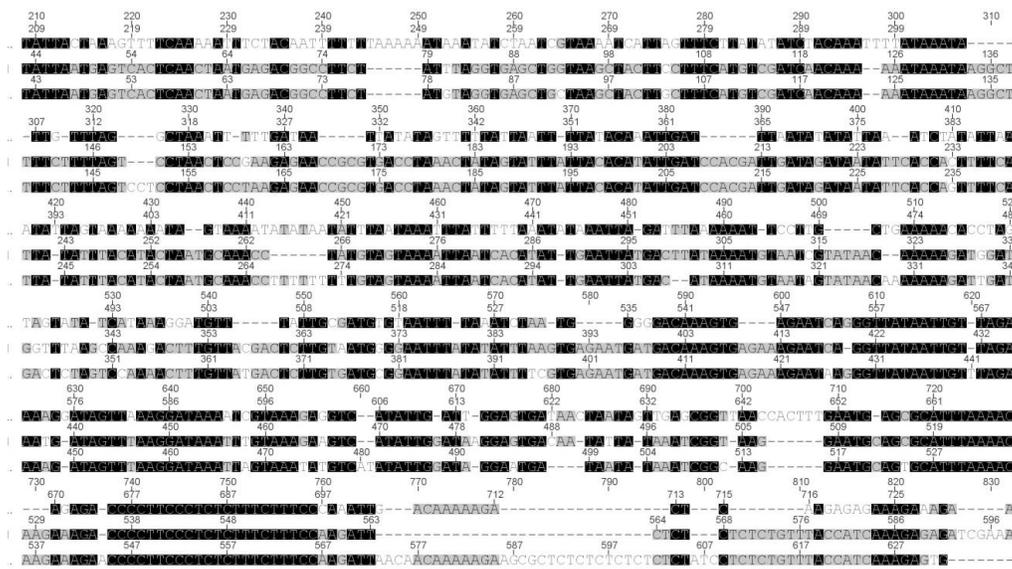
	Alien RNA			UBC21					SAND					UP1					TIP41				
	1+E <sup>a</sup>	Ct <sup>b</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>	1+E <sup>a</sup>	Ct <sup>b</sup>	RQ <sup>c</sup>	s.e. <sup>d</sup>	n <sup>e</sup>
1-2 day seedlings	1.97	23.40	3	1.95	24.34	0.70	0.10	3	1.92	25.64	0.41	0.01	3	1.91	25.18	0.63	0.07	3	1.96	25.80	0.22	0.02	3
5-day seedlings	1.91	23.34	3	1.94	25.85	0.11	0.01	3	1.87	26.11	0.28	0.01	3	1.93	26.39	0.10	0.01	3	1.98	26.73	0.05	0.00	3
10-day seedlings	1.93	23.24	3	1.96	25.75	0.13	0.01	3	1.82	25.96	0.77	0.13	3	1.93	25.66	0.21	0.04	3	2.03	26.72	0.03	0.00	3
Young leaves	1.95	23.16	3	1.94	25.30	0.24	0.05	3	1.80	26.01	1.03	0.03	3	1.92	25.71	0.25	0.04	3	1.98	26.22	0.09	0.01	3
Mature leaves	1.95	23.42	3	1.97	26.15	0.13	0.04	3	1.90	25.42	0.52	0.09	3	1.96	26.60	0.10	0.01	3	1.95	26.61	0.12	0.02	3
Anthers	1.92	23.26	3	1.92	24.42	0.47	0.11	3	1.82	24.59	1.53	0.24	3	1.92	26.49	0.12	0.01	3	1.96	27.21	0.04	0.01	3
Flowers	1.87	23.20	3	1.96	24.84	0.11	0.01	3	1.94	26.08	0.06	0.01	3	1.95	25.74	0.07	0.01	3	2.02	26.02	0.02	0.00	3
20-day embryo	1.93	23.33	3	1.98	25.43	0.13	0.04	3	1.89	26.14	0.27	0.04	3	1.97	25.73	0.10	0.03	3	1.95	26.72	0.08	0.01	3
25-day embryo	1.88	23.26	3	1.94	27.07	0.04	0.01	3	1.81	27.10	0.24	0.05	3	1.97	26.53	0.04	0.00	3	1.96	27.47	0.02	0.00	3
30-day embryo	1.92	23.32	3	1.92	28.24	0.04	0.01	3	1.81	27.91	0.26	0.04	3	1.92	27.47	0.07	0.00	3	2.03	28.11	0.01	0.00	3
35-day embryo	1.93	23.35	3	1.96	29.05	0.01	0.00	3	1.87	27.93	0.10	0.00	3	1.92	27.40	0.08	0.01	3	1.95	28.54	0.02	0.00	3
40-day embryo	1.94	23.24	3	1.94	27.36	0.06	0.01	3	1.88	26.99	0.19	0.02	3	1.91	26.27	0.20	0.04	3	2.00	27.74	0.02	0.00	3
45-day embryo	1.87	23.43	3	1.96	28.23	0.01	0.00	3	1.78	26.57	0.53	0.14	3	1.90	26.43	0.10	0.00	3	1.97	28.10	0.01	0.00	3
50-day embryo	1.91	23.54	3	1.93	28.07	0.04	0.01	3	1.82	26.85	0.41	0.05	3	1.94	26.61	0.09	0.01	3	1.99	28.29	0.02	0.00	3

Legend: a - average PCR efficiency value of the replicates; b - mean Ct value of the replicates; c - relative quantity,  $[(E+I)^{Ct}]_{target} / [(E+I)^{Ct}]_{external\ controls}$ , external control-"Alien RNA"; d - standard error; e - number of biological replicates.



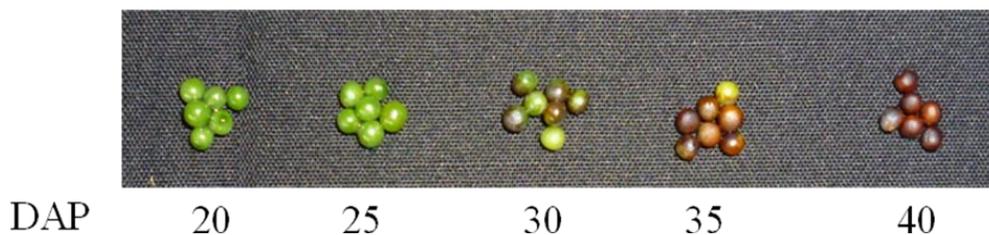
**Figure A1. PCR amplification of the 4th putative *BnGPAT4* homologue.**

(A) lane 1: gene-specific primers for *AA\_GPAT4\_2* + genomic DNA from *B. napus*; lane 2: gene-specific primers for *AA\_GPAT4\_2* + genomic DNA from *B. rapa*, positive control for the primers. (B) left 1: gene-specific primers for *AA\_GPAT4\_2* + genomic DNA from *B. napus*; lane 2: gene-specific primers for *BnGPAT4C1* + genomic DNA from *B. napus*, positive control for the genomic DNA of *B. napus*.



**Figure A2. Sequence alignment of the promoter regions of the three *BnGPAT4* homologues.**

Line 1, *BnGPAT4-C1*. Line 2, *BnGPAT4-C2*. Line 3, *BnGPAT4-A1*.



**Figure A3. Developing *B. napus* seeds at the stages from 20 to 40-DAP.**

pGreenII0229-*BnGPAT4*<sub>pro:cDNA</sub>

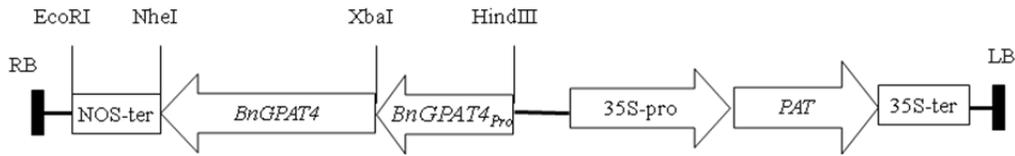
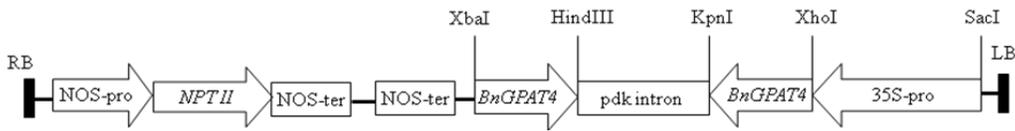


Figure A4. Schematic diagram of the constructs used for expression of different *BnGPAT4* homologues in *Arabidopsis gpat4 gpat8* double mutant lines.

RD400-35S-*BnGPAT4*



RD400-napin-*BnGPAT4*

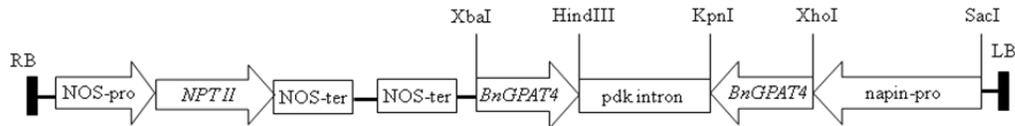


Figure A5. Schematic diagram of the construct used for RNAi silencing *BnGPAT4* homologues in *B. napus*.

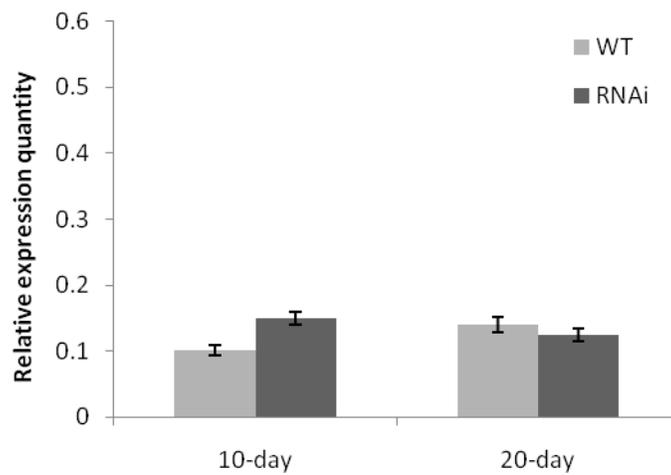


Figure A6. Comparison of the *BnGPAT6* expression levels in the rosette leaves of wild type and RNAi *B. napus* lines. n=3.

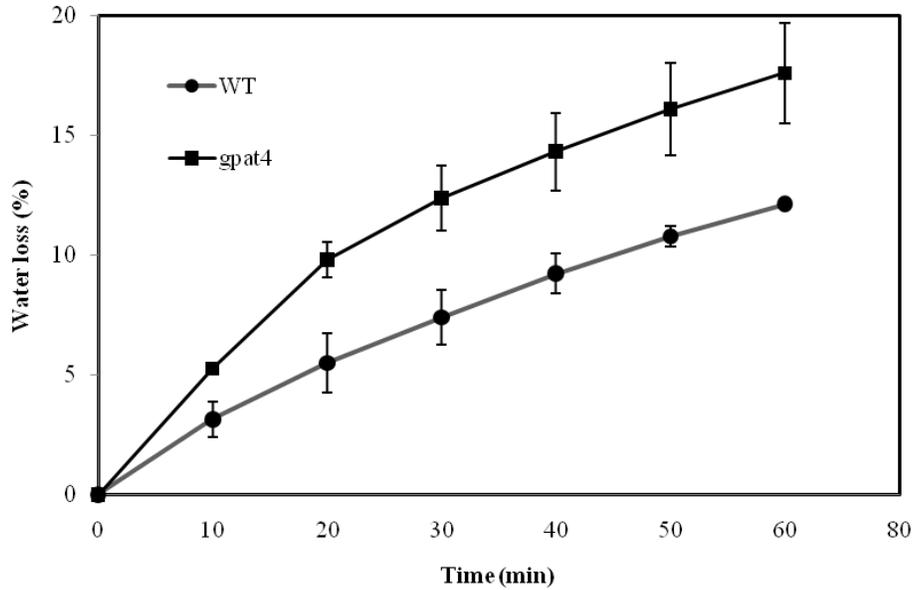


Figure A7. Water loss of excised rosette leaves recorded over 60 min measured as a percentage of the initial weight. n=3.

```

Arabidopsis 1 -----MLTFSSAAVVAATVTSSARVPVYPLASSTLRGLVSFRLTAKRL
Pea 1 -----MTDSEPHCASHINYRHKMKMFIFSTPCCSPSTAFESPPFRA
Pepper 1 -----MLILSAASSSARISRPPLSS-----FSTFAASAATISRLFPISCGVVKSTT
Citrus 1 -----MSSLSLTFFATTAPRVLPSSSSNPKLSPSSYSFSAITARRHSTAVSFERS
Palm-UK 1 -----MLVPSALPRVRSRVSFAREFSVSGVSSPALSSRSOSLDSVRSLSLRRCCPC
Palm-Mal 1 -----MTDSEPHCASHINYRHKMKMFIFSTPCCSPSTAFESPPFRA
FavaBean 1 -----MTDSEPHYASHINIRPKTKMLIFSTPCCSPSTAFESPPFRA
Tomato 1 -----MLSSALSSSARIPRELSSSFSTCVFVVVTVSSAAITLFPISCGVVKSR
Spinach 1 -----MLVLSSAPPVLEVCKDRVSSSFSTSSSSSSSAFSAVVFRSFFRFRNSLIC
Cucumber 1 MFILSAVSSSSSSSSSVPSLPPFSLSPSISLSFSRVSLPSSSSSSSSSLKLFPLSLHF
Squash 1 -----
FrenchBean 1 -----MSMTGSSRYVAHAIPFFLRISNKTMLLLSTPPTFFFFSTTP
Safflower 1 -----MSIFFSPSSPTLFFSTTNANPRVSPSSSPSAFTPPLSSSRLRPIILRGFPC
EaGPAT 1 -----MLTFSSAAVVAATVTSSRVVPYPLASPSLTGFVVSFRLTAKRL

Arabidopsis 49 FLFPLRSRG---G-VSVRAMSEVQDKESSVFAASIAFNEAAG---ETPSELSSHRTFLDA
Pea 42 SNSKPLRSTLSLRSSISS--SSITSTSHCSLAFNIVKHKEKNVVSANMTSSVSRTFFLNA
Pepper 47 VGNRKLQCAVFCASLVRGMAEVEDNKNELNSSTAAAAAIAVTASENDELPHSRAFFLDA
Citrus 51 GARRTSATPCASFSSFNVRAMAKMVQDRESAVSSSSASDEQNKKMLNIEVNHSTRFLDV
Palm-UK 52 GIYTSRTRKAVVEAVESKASARWRSS-----AVKRAVLASDTGAEVEVGHRSRFLRA
Palm-Mal 42 SNSKPLRSTLSLRSSISS--SSITSTSHCSLAFNIVKHKEKNVVSANMTSSVSRTFFLNA
FavaBean 42 SNCKPLRSSTLCLRSLTSSAITSITSTSNSSLAFNIVKPKKKNVVSANMTSSVSRTFFLNA
Tomato 51 VGIKRLRCAVFCAS-KVRGMAEVED-----AMTVSASESHELPCSRDFLDA
Spinach 54 CCSSKCLKMADTALPSSSSSSTASASYSAAAKSVEEENHEIPVKKEDDNOQLRSRTYRNV
Cucumber 61 TPEKLSPPHSFLRFASRAMAEIQDKESAHTPSTTDVTR-----NDPPHSRAFDL
Squash 1 -----MAELIQDKESAQAATAAFASSGYERR-----NEPAHSRKFLDV
FrenchBean 44 RVTLLSSTSSSSSSSISLRSSTAPSPSCSSVTFKDNCLASAKHSPNMSASVSRFFLNA
Safflower 52 LAFSAPANAAHGTAETVHGKWPSPS-----SSSAATQPSAGSDHGHSTRFFIDA
EaGPAT 49 LRGPLSSRG---GGASVRAMSEVQDKESSVFAASIAFNDAAGSSNTTPESELNHSTRFLDA

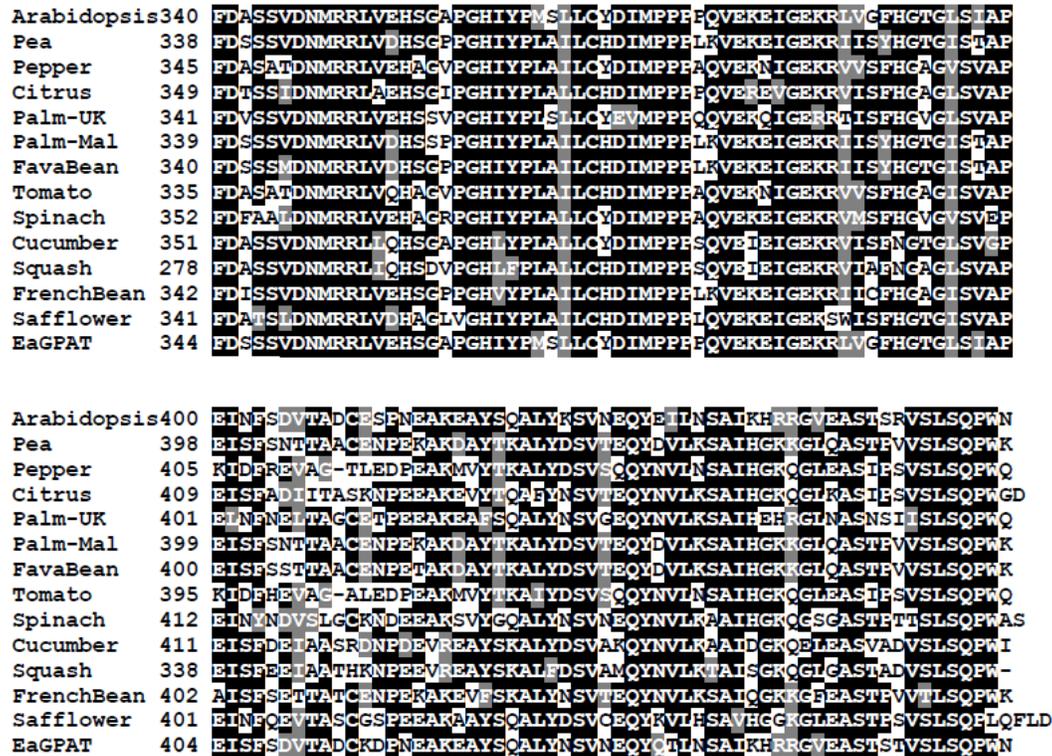
```

Arabidopsis	102	RSEQDLLSGIKKEAEAGRLPANVAAGMEELYNYKNAVLSGASRAD-ETVLSNMSVAED
Pea	100	QNEQDVLSGIKKEVEAGTLPASIAAGMEEVYNYKSAVTKSGDPKAN-EIVLSNMTALLD
Pepper	107	RUGEDLLSAVRKAVEDKRLPLNIAEGMEELYNYRNAVLCSGVPKAD-EIILYNNMALVLD
Citrus	111	RSEQDLLSGICREVEAGRLPSNLAN GMEELYNYKNAVFCSGNSRAD-EIVLSNMAVAED
Palm-UK	103	RSEBELLSYIRKEVEAGRLSSDIANGLEEELYNYRNAVLCSGDPRAN-KIILSNMAVAED
Palm-Mal	100	QNEQDVLSGIKKEVEAGTLPASIAAGMEEVYNYKSAVTKSWRSQSNRNCINKIRLPLLD
FavaBean	102	QNEQDVLSGIKKEVEAGTLPASIAAGMEEVYNYKSAVTKSGDPKAN-EIVLSNMTALLD
Tomato	97	RUGEDLLSAVRKAVEDKRLPLNVAEGMEELYNYQNAVLCSGVPKAD-EAIIYNNMALVED
Spinach	114	RSABELLISEIKRESEIGRLPKSVYAMEGLEHYRNAVLSGISHAD-EIVLSNMSVMLD
Cucumber	113	RSEBELLSQIRRETEAGRLPSNVAAGMEELYNYKNAVFESGNPKAD-EIVLSNMTVALD
Squash	40	RSEBELLSQIKKETEAGRLPNSVAAGMEELYNYRNAVIESGNPKAD-EIVLSNMTVALD
FrenchBean	104	QSEQDVFAGIKKEVEAGSLPANVAAGMEEVYNYKSAVTKSGDPKAN-EIVLSNMTALLD
Safflower	102	RSEQDLLSGIQRELEAGTLPKHIAQAMEELYNYKNAVLCSPAFAE-DIVLSNMRVAED
EaGPAT	106	RSEQDLLSGIKKEAEAGRLPANVASGMEELYNYKNAVLSGASRAE-ETVLSNMSVAED

Arabidopsis	161	RVLGVEDPFTFENFYHKAAREPFDYIMFVHTYIRPLIDFENSYVGNASIESELEDKIRQG
Pea	159	RVELDVKEPFVFEAHHKAKREPFDYIMFGQNYIRPLVDEETSIVYGNMELFIQMEEQLKQG
Pepper	166	RVEVDVKDPPEFSPMHKATREPFDYIMFGQNYIRQLVDFESSYVGNISVFGEMEEKLKQG
Citrus	170	RVLIDIEEPFTFSSYHKSMPREPFDYIMFGQNYIRPLVDFESSYVGNVSLFEMEEKLNQGG
Palm-UK	162	RVLIDVEDPFTFSPHQAIREPFDYIMFGQNYIRPLIDFERSYIGNISFESDMEEKLQGG
Palm-Mal	160	RVELDVKEPFVFEAHHKAKREPFDYIMFGQNYIRPLVDEETSIVYGNMELFIQMEEQLKQG
FavaBean	161	RVELDVKEPFVFEAHHKAKRCPFDYIMFGQNYIRPLVDEETSIVYGNMELFIQMEEQLKQG
Tomato	156	RVEVDVKDPPEFSPMHKATREPFDYIMFGQNYIRQLVDFESSYVGNISVFSSEMFEKLGQG
Spinach	173	RVLIDIEDPFTFEPHKAIREPDIYISFGQDYIRPLVDFGNSYVGNIAIEFQEMEEKLKQG
Cucumber	172	RVLIDVEDPFTFSPHKAIREPFDYIMFGQNYIRPLIDFENSFVGNLSLFKDIIEEKLHQG
Squash	99	RVLIDVEDPFTFSSHKAIREPFDYIMFGQNYIRPLIDFENSFVGNLSLFKDIIEEKLHQG
FrenchBean	163	RVELDVTDPFVFEQPHHAKRREPFDYIMFGQNYIRPLVDFENAYVGNMELFIEMEEKLKQG
Safflower	161	RVELDVKEPFVFEAHHKATREPFDYIMFGQNYIRPLVDFESSYVGNISVFGVMEEQLKQG
EaGPAT	165	RVLGVEDPFTFSPMHKAVREPFDYIMFVHTYIRPLIDFENSYVGNASIESELEDKIRQG

Arabidopsis	221	HNIVLISNHQSEADPAVISLLEAQSFEI GENIKCVAGDRVITDPLCKPFSMGRNLCVY
Pea	219	HNIIILMSNHQSEADPAIIALLLEMRLEHIAENLIYVAGDRVITDPLCKPFSIGRNLCVY
Pepper	226	DNVVILMSNHQSEADPAIIALLLELKHEDIAENLIYVAGDRVITDPLCKPFSMGRNLCVY
Citrus	230	HNIVLISNHQTEADPAIIALLLESTNEHIAENLIYVAGDRVITDPLCKPFSMGRNLCVY
Palm-UK	222	HNIVLMSNHQTEADPAIIALLLEBRTNSHIAETIMVAVAGDRVITDPLCKPFSMGRNLCVY
Palm-Mal	220	HNIIILMSNHQSEADPAIIALLLEMRLEHIAENLIYVAGDRVITDPLCKPFSIGRNLCVY
FavaBean	221	HNIIILMSNHQSEADPAIIALLLEMRLEHIAENLIYVAGDRVITDPLCKPFSIGRNLCVY
Tomato	216	DNVVILMSNHQSEADPAIIALLLESKLEDIAENLIYVAGDRVITDPLCKPFSMGRNLCVY
Spinach	233	DNIILMSNHQSEADPAIIALLLEKTNSTIAENLIYVAGDRVITDPLCKPFSMGRNLCVY
Cucumber	232	HNIVLISNHQTEADPAIIALLLEKTNFYIAENLIYVAGDRVITDPLCKPFSIGRNLCVY
Squash	159	HNIVLISNHQTEADPAIIALLLEKTNFYIAENLIYVAGDRVITDPLCKPFSIGRNLCVY
FrenchBean	223	HNIIILMSNHQTEADPAIIALLLETRLFYIAENLIYVAGDRVITDPLSKPFSIGRNLCVY
Safflower	221	DKVVVLSNHQTEADPAIIALLLETTNEHISENLIYVAGDRVITDPLCKPFSMGRNLCVY
EaGPAT	225	HNIIILISNHQSEADPAVISLLEAQSFEI GENIKCVAGDRVITDPLCKPFSMGRNLCVY

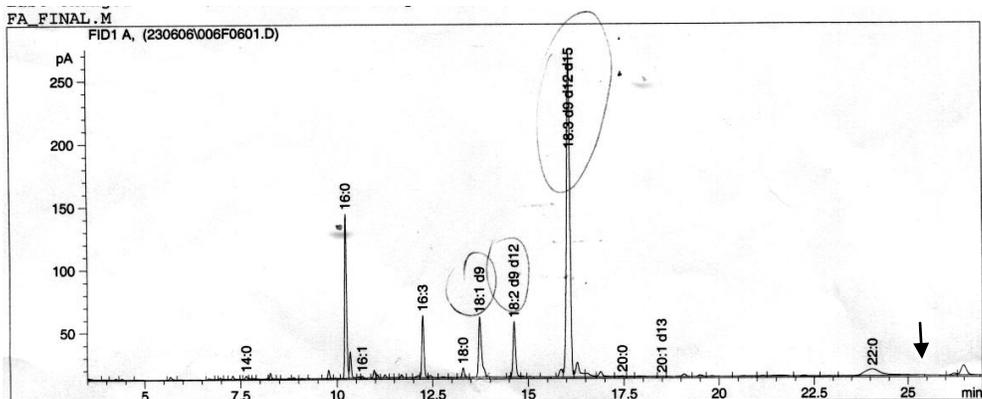
Arabidopsis	281	SKKHMDDPELVDMKRRKANTRSLK-EMAMLLRSGGQLIWIAPSGGRDRPDESIGEWFPAP
Pea	279	SKKHMDDNPPELVDMKRRKANTRSRK-EMAMLLRSGSQIWIWIAPSGGRDRPVANSIGEWFPAP
Pepper	286	SKKHMDDPELADMKRRKANTRSLK-EMAMLLRGGSKLIWIWIAPSGGRDRPDEVIGEWFPAP
Citrus	290	SKKHMDDVPELIEMKRRKANTRSLK-EMAMLLRGGSQIWIWIAPSGGRDRPDEVIGEWFPAP
Palm-UK	282	SKKHMDDVPELIEMKRRKANTRSLK-EMAMLLRGGSQIWIWIAPSGGRDRPDESIGEWFPAP
Palm-Mal	280	SKKHMDDNPPELVDMKRRKANTRSRK-EMAMLLRSGSQIWIWIAPSGGRDRPVANSIGEWFPAP
FavaBean	281	SKKHMDDNPPELIDMKRRKANTRSLK-EMAMLLRSGSQIWIWIAPSGGRDRPVANSIGEWFPAP
Tomato	276	SKKHMDDPELAEMKRRKANTRSLK-EMAMLLRGGSKLIWIWIAPSGGRDRPDEVIGEWFPAP
Spinach	293	SKKHMDDPELVDMKRRKANTRSLK-ELVLLRGGSKLIWIWIAPSGGRDRPDAVIGEWFPCT
Cucumber	292	SKKHMDDPELAEMKRRKANTRSLK-EMAMLLRGGSQIWIWIAPSGGRDRPDESIGEWFPAP
Squash	219	SKKHMDDPELIEMKRRKANTRSLK-EMAMLLRGGSQIWIWIAPSGGRDRPDESIGEWFPAP
FrenchBean	283	SKKHMDDPELIVEMKRRKANTRSLK-EMAMLLRNGSQIWIWIAPSGGRDRPDACTREWFPAP
Safflower	281	SKKHMDDVPELAEMKRRKANTRSLK-EMAMLLRGGSKLIWIWIAPSGGRDRPDEVIGEWFPAP
EaGPAT	285	SKKHMDDPELVDMKRRKANTRSLK-EMAMLLRSGGQLIWIWIAPSGGRDRPDESIGEWFPAP



**Figure A8. Amino acid sequence alignment of EaGPAT and other plastidial GPATs.**

The N-termini of the plastidial GPATs from different plant species share much lower sequence similarity compared to the rest of their sequences. The red color amino acids and white arrow indicate the previously predicted cleavage sites of transit peptides for the plastidial GPATs from squash, pea, cucumber, and *Arabidopsis*, respectively (Webber et al., 1988; Johnson et al., 1992; Ishizaki et al., 1988; Nishida et al., 1993). The black arrow indicates the putative cleavage site for the EaGPAT transit peptide.

NCBI gene accession No.: Arabidopsis: NP\_174499; Pea: CAA41769; Pepper: AAP79443; Citrus: BAB79529; Oil Palm (UK): AAF64066; Oil Palm (Malaysia): CAB75874; Fava bean: AF090734; Tomato: AAD05164; Spinach: CAA88913; Cucumber: Q39639; Squash: P10349; French bean: Q43822; Safflower: Q42713.



=====  
 Normalized Percent Report  
 =====

Sorted By : Signal  
 Calib. Data Modified : 6/22/06 4:27:09 PM  
 Multiplier : 1.0000  
 Dilution : 1.0000

Signal 1: FID1 A,

RetTime [min]	Type	Area [pA*s]	Amt/Area	Norm %	Grp	Name
6.125		-	-	-		
6.973	VV	6.26412	1.00000	0.219158		
7.648	VV	10.84516	1.00000	0.379431		14:0
10.246	VV	433.87198	1.00000	15.179546		16:0
10.662	VP	4.80373	1.00000	0.168065		16:1
11.438	VV	2.19180	1.00000	0.076683		
11.572	VV	2.34838	1.00000	0.082161		
12.253	VV	186.43227	1.00000	6.522563		16:3
12.970	VP	1.09270	1.00000	0.038229		
13.298	VP	30.06013	1.00000	1.051691		18:0
13.737	VB	259.38852	1.00000	9.075027		18:1 d9
13.921		-	-	-		18:1 d11
14.637	VV	200.90587	1.00000	7.028939		18:2 d9 d12
15.414	VP	9.70091e-1	1.00000	0.033940		
15.521		-	-	-		
16.095	VV	1553.58435	1.00000	54.354063		18:3 d9 d12 d15
17.455	VV	7.73675	1.00000	0.270680		20:0
18.161		-	-	-		20:1 d11
18.494	VV	1.57598	1.00000	0.055138		20:1 d13
24.054	VB	156.19537	1.00000	5.464688		22:0
25.212		-	-	-		22:1 d13 ←

Totals : 100.000000

**Figure A9. Analysis of the fatty acid composition of total lipids extracted from the *E. asperum* leaves.**

Black arrows indicate the retention time (25.212 min) of 22:1 *cis* Δ<sup>13</sup>.

## Appendix 2

**Snyder CL, Yurchenko OP, Siloto RMP, Chen X, Liu Q, Mietkiewska E and Weselake RJ** (2009) Acyltransferase action in the modification of seed oil biosynthesis. *New Biotechnology* **26**: 11-16

My contribution to this review paper is providing information on the current understanding of the functions of ER-bound GPATs in lipid biosynthesis of *Arabidopsis*.

Copyright (2010), with permission from Elsevier.

All co-authors approved the usage of the manuscript in the present thesis.



# Acyltransferase action in the modification of seed oil biosynthesis

Crystal L. Snyder, Olga P. Yurchenko, Rodrigo M.P. Siloto, Xue Chen, Qin Liu, Elzbieta Mietkiewska and Randall J. Weselake

Department of Agricultural, Food and Nutritional Science, University of Alberta, 4-10 Agriculture/Forestry Centre, Edmonton, Alberta, Canada T6G 2P5

Seed oils represent a major source of dietary lipid and an increasingly valuable feedstock for industrial applications. There have been several attempts to modify seed oil content and composition through biotechnological approaches, resulting in the identification of several ‘bottlenecks’ limiting the accumulation of unusual fatty acids in storage lipids of oilseed crops. It has been suggested that the substrate preferences of endogenous acyltransferases play an important role in the utilization of unusual fatty acids in transgenic oilseeds, and there is increasing evidence that mechanisms of ‘acyl-editing’ via phospholipids are also involved in substrate trafficking and utilization. In this review, we will examine acyltransferase substrate specificity and selectivity in the context of designing strategies to maximize the accumulation of unusual fatty acids using biotechnological approaches.

## Contents

Introduction . . . . .	11
Acyltransferases involved in plant storage lipid biosynthesis . . . . .	12
Substrate specificity properties of the Kennedy pathway enzymes. . . . .	12
<i>sn</i> -Glycerol-3-phosphate acyltransferase. . . . .	12
Lysophosphatidic acid acyltransferase . . . . .	13
Diacylglycerol acyltransferase . . . . .	13
The contribution of lysophosphatidylcholine acyltransferase and phospholipid:diacylglycerol acyltransferase to acyl-trafficking via phosphatidylcholine . . . . .	13
Future outlook . . . . .	14
Acknowledgements . . . . .	14
References. . . . .	14

## Introduction

Soaring global demand for plant oils for edible and industrial applications has sparked a considerable amount of research focused on biotechnological approaches to oil modification [1] (Table 1). Despite steady progress in the cloning and characterization of genes involved in plant storage lipid metabolism, attempts

to produce unusual fatty acids at commercially viable levels in oilseed crops have largely fallen short. Several potential bottlenecks have been identified, including substrate dichotomy between acyl-CoA and phospholipid pools and the inability of endogenous acyltransferases to efficiently incorporate the desired acyl groups into storage lipids ([2]; reviewed by [1,3]). In this review, we will discuss the role of acyltransferase substrate preferences in the modification of seed oils and discuss strategies for overcoming these obstacles through biotechnological approaches.

Corresponding author. Weselake, R.J. (randall.weselake@ualberta.ca)

TABLE 1

## Biotechnological approaches to modifying seed oil content and composition through the use of acyltransferases in transgenic plants

Host species	Source species	Upregulated Gene	Phenotype	Refs
<i>Brassica napus</i> , <i>Tropaeolum majus</i>	<i>B. napus</i> , <i>Arabidopsis thaliana</i>	DGAT1	Increased seed oil content	[52,53,69]
<i>A. thaliana</i> , <i>N. tabacum</i>	<i>A. thaliana</i>	DGAT1	Increased seed oil content, seed weight	[51,70]
<i>Glycine max</i>	<i>Umbelopsis ramanniana</i>	DGAT2	Increased seed oil content	[54]
<i>A. thaliana</i>	<i>Ricinus communis</i>	FAH12 DGAT2	Up to 30% hydroxyl fatty acids	[48]
<i>B. napus</i>	<i>Cocos nucifera</i>	LPAAT 12:0 ACP thioesterase <sup>a</sup>	Increased accumulation of 12:0 at the <i>sn</i> -2 position from 5 to 30%	[30]
<i>B. napus</i>	<i>Limnanthes</i> spp., yeast	LPAAT	Increased incorporation of 22:1 in <i>sn</i> -2 position	[31–33,35]
<i>A. thaliana</i>	<i>E. coli</i> ; <i>C. tinctorius</i> <sup>b</sup>	GPAT	Increased seed oil content	[71]

<sup>a</sup> *Umbellularia californica* ACP-thioesterase.

<sup>b</sup> *C. tinctorius* (safflower) plastidial GPAT.

### Acyltransferases involved in plant storage lipid biosynthesis

The pathways involved in plant storage lipid biosynthesis have been recently reviewed in detail elsewhere [4,5] and are summarized in Fig. 1. For the purposes of this review, we will consider acyltransferases involved in two pathways leading to triacylglycerol (TAG). The classical *sn*-glycerol-3-phosphate or Kennedy pathway involves the sequential acyl-CoA-dependent acylation of *sn*-glycerol-3-phosphate catalyzed by *sn*-glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAAT) and diacylglycerol acyltransferase (DGAT), respectively, with phosphatidic acid phosphatase catalyzing the dephosphor-

ylation of PA before the final acylation. In a second pathway, acyl groups can be channeled into phosphatidylcholine (PC) via the activity of lysophosphatidylcholine acyltransferase (LPCAT), and then subsequently transferred from PC to TAG via the activity of phospholipid:diacylglycerol acyltransferase (PDAT) [6,7]. This latter pathway represents one of several possible mechanisms for 'acyl-editing', permitting acyl groups which are modified at the level of PC (i.e. by desaturation or hydroxylation) to be channeled into storage lipids. Understanding these acyl-editing mechanisms and the substrate preferences of the enzymes involved is important for developing strategies for overcoming the problem of substrate dichotomy in the accumulation of unusual fatty acids in transgenic plants.

### Substrate specificity properties of the Kennedy pathway enzymes

#### *sn*-Glycerol-3-phosphate acyltransferase

Of the three known types of plant GPATs (plastidial, mitochondrial and endoplasmic reticulum), only the ER-bound GPAT is known to be involved in the Kennedy pathway. To date, five *Arabidopsis* genes encoding ER-bound GPAT have been reported [8–10], mostly with roles in suberin and cutin synthesis. The role of specific GPAT genes in TAG synthesis has not been clarified to date. While there are few published reports on the substrate specificity of ER-bound GPAT enzymes, positional analyses of several plant oils have suggested that plant GPATs have a relatively broad specificity, utilizing saturated, monounsaturated and polyunsaturated moieties [11–13]. Dutta *et al.* [14] demonstrated that microsomal GPATs from safflower and rapeseed utilized oleoyl-CoA (18:1<sup>cis</sup> $\Delta^9$ ) more efficiently than petroselinoyl-CoA (18:1<sup>cis</sup> $\Delta^6$ ). Subsequent studies of transgenic *Arabidopsis*, however, showed that petroselinic acid was incorporated into TAG as effectively as oleic acid [15]. In *B. napus* engineered to produce medium chain saturated fatty acids, up to 80% of the acyl groups at the *sn*-1 position were medium chain saturated fatty acids [16–18], whereas *B. napus* normally contains about 7% saturated fatty acids, mainly occurring at the *sn*-1 position. These observations support the view that GPAT may not exhibit a strong substrate preference *in vivo*, rather, acyl-CoA availability may play a more pivotal role in determining the acyl composition at the *sn*-1 position of TAG [16].

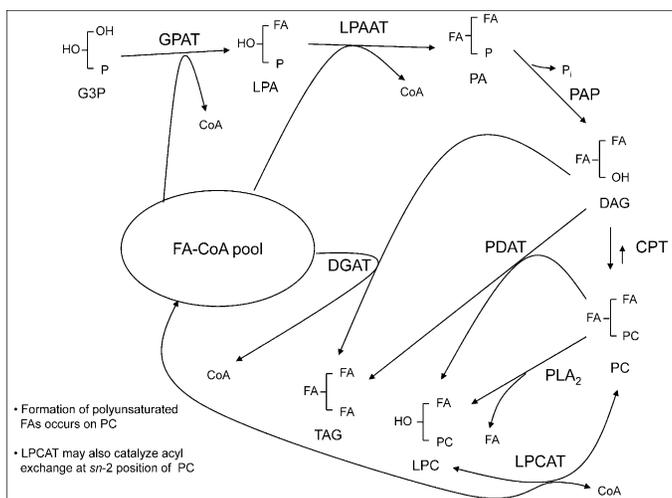


FIGURE 1

Generalized scheme for triacylglycerol (TAG) assembly in developing seeds of oleaginous plants. Abbreviations: CoA, coenzyme A; CPT, cholinephosphotransferase; DAG, *sn*-1,2-diacylglycerol; FA, fatty acid; FA-CoA, fatty acyl-coenzyme A; G3P, *sn*-glycerol-3-phosphate; GPAT, *sn*-glycerol-3-phosphate acyltransferase; LPA, lysophosphatidic acid; LPAAT, lysophosphatidic acid acyltransferase; LPC, lysophosphatidylcholine; LPCAT, lysophosphatidylcholine acyltransferase; PA, phosphatidic acid; PAP, phosphatidic acid phosphatase; PC, phosphatidylcholine; PDAT, phospholipid:diacylglycerol acyltransferase; PLA<sub>2</sub>, phospholipase A<sub>2</sub>. Based on [72,4]. PDAT may also catalyze the synthesis of TAG and monoacylglycerol from two molecules of DAG [73].

### Lysophosphatidic acid acyltransferase

Microsomal LPAAT has the highest substrate stringency of the three Kennedy pathway acyltransferases [19]. Plants producing unusual FAs often have microsomal LPAAT activities with enhanced preferences for these FAs [20–26]. In addition to its role in storage lipid assembly, LPAAT also plays a role in the *de novo* synthesis of PC and this may be reflected in the acyl-donor specificity of LPAATs containing high levels of polyunsaturated fatty acids. For example, it has been shown that flax microsomal LPAAT prefers 18:2<sup>cisΔ9,12</sup>-CoA > 18:1<sup>cisΔ9</sup>-CoA > 18:3<sup>cisΔ9,12,15</sup>-CoA [27]. This specificity is consistent with a possible role for LPAAT in channeling substrates toward the *sn*-2 position of PC for further desaturation to α-18:3, which constitutes more than 60% of the FA in mature flax seed. Brown *et al.* [28] have shown that microsomal LPAATs from several plant species (including *B. napus*) display little selectivity between 18:1- and 18:2-CoA donor substrates, whereas 16:0- and 18:0-CoA are generally discriminated against. Such discrimination against saturated acyl groups in the *sn*-2 position of TAG is relatively common among major oilseed crops [11,12], and hampered early progress toward engineering laurate (12:0) accumulation in transgenic *B. napus* to produce cocoa butter-like oils [20,29]. Incorporation of 12:0 at the *sn*-2 position of TAG was eventually accomplished by the expression of a microsomal laurate-specific LPAAT from coconut (*Cocos nucifera*) in *B. napus* expressing a California bay laurel (*Umbellularia californica*) 12:0-acyl carrier protein (ACP) thioesterase [30]. Similarly, LPAATs with enhanced specificity for erucoyl-CoA (22:1<sup>cisΔ13</sup>) were expressed in *B. napus*, resulting in substantial quantities of 22:1 at the *sn*-2 position of TAG [31–35], overcoming the tendency for *B. napus* LPAAT to exclude 22:1 from this position [23,36]. Oils with a high 22:1 content can serve as feedstocks in the preparation of detergents, plasticizers, surface coatings and slip-promoting/anti-blocking agents with trierucin serving as a high temperature lubricant [37].

### Diacylglycerol acyltransferase

DGAT has the lowest specific activity of the Kennedy pathway enzymes and is regarded as a 'bottleneck' in TAG synthesis [38]. As such, DGAT has been one of the more extensively studied enzymes both in terms of its substrate preferences and for the potential to increase seed oil content through DGAT overexpression. Early biochemical studies evaluated DGAT substrate preference through *in vitro* assays using microsomal fractions of developing seeds. In safflower, it was shown that DGAT can utilize a broad range of molecular species of DAG and acyl-CoA [39,40]. Other studies have indicated that in species such as *B. napus*, *Tropaeolum majus* (Nasturtium), *Ricinus communis* (Castor bean) and *Cuphea lanceolata*, endogenous DGATs exhibit preference for substrates with defined acyl chains [25,36,41–46].

It is important to note that early *in vitro* studies on DGAT substrate preference did not distinguish between type-1 DGATs (DGAT1) and the more recently identified type-2 DGATs (DGAT2). The discovery of genes encoding these two non-homologous DGAT polypeptides has allowed a more refined dissection of TAG biosynthesis through the study of individual genes in recombinant systems. In *Vernicia fordii* (tung tree), DGAT1 and DGAT2 display minor differences in substrate selectivity as determined by *in vitro* assays. *In vivo* selectivity, evaluated through a yeast expres-

sion system, has revealed that DGAT2 is able to catalyze the incorporation of higher amounts of exogenous eleostearate into trieleostearin, an abundant TAG found in tung tree [47]. Moreover, it has been shown that these two DGATs are present in distinctive ER domains, corroborating the hypothesis that they do not have overlapping functions. Similarly, in castor bean, DGAT2 is selective for substrates with ricinoleic acid and is responsible for the accumulation of triricinolein [48]. Bursal *et al.* [48] reported that coexpression of fatty acid hydroxylase (*FAH12*) and *DGAT2* from castor bean resulted in up to 30% hydroxy fatty acids, nearly double that attainable from the expression of *FAH12* alone. Such studies highlight the importance of using acyltransferases with appropriate substrate preferences for maximizing the accumulation of unusual fatty acids in transgenic plants.

Although these results suggest that DGAT2 may be more important than DGAT1 in plants, the actual contribution of each isoform to seed oil accumulation is still uncertain as only DGAT1 knockouts have been characterized in plants. In *Arabidopsis*, an EMS-induced mutation of *DGAT1* (AS11) resulted in a 15% decrease in seed lipids [49]. In maize, a high-oil quantitative trait locus *qHO6* was shown to be caused by a single residue insertion in DGAT1 polypeptide and *qHO6*<sup>-/-</sup> maize lines possessed lower oil content [50]. In both *Arabidopsis* and maize a decrease of DGAT1 activity also affected the fatty acid composition, resulting in an increase of α-18:3 content. This phenotype could be reversed by restoring the DGAT1 activity [50,51], suggesting that in these species DGAT1 might also influence certain acyl channeling to TAG.

Both *DGAT1* and *DGAT2* have been successfully overexpressed in crops as a means of increasing seed oil content. *B. napus* expressing *DGAT1* [52,53] and soybean expressing *DGAT2* [54] both exhibited similar gains in oil content when tested under field conditions, suggesting that both enzymes may play a crucial role in TAG accumulation despite differences in substrate preferences. The precise contribution of DGAT1 and DGAT2 to overall TAG content and composition remains an open area of investigation. Further biotechnological applications of DGAT in oilseed metabolic engineering will require a better understanding of the mechanisms governing the enzyme activity. Molecular genetics approaches such as site directed mutagenesis and directed evolution could potentially be used to increase the overall catalytic efficiency and modify substrate selectivity [55].

### The contribution of lysophosphatidylcholine acyltransferase and phospholipid:diacylglycerol acyltransferase to acyl-trafficking via phosphatidylcholine

It has been shown that a majority of the fatty acids synthesized in plastids enter PC through acyl-editing of PC rather than through *de novo* synthesis [56,57]. These observations are supported by experiments demonstrating that LPCAT is one of the most active acyltransferases in seed oil biosynthesis [58]. In soybean, safflower and flax, 18:1-CoA and 18:2-CoA are the preferred acyl donors and are incorporated into PC at similar rates [59–61]. Linolenate (α-18:3) is incorporated into PC at lower rates [60–62], and saturated acyl groups are not efficiently utilized [58]. Two recently cloned *Arabidopsis* LPCATs were shown to discriminate against 18:2 and 22:0, but utilized both saturated and monounsaturated C16 and

C18 acyl groups [63]. Similarly, LPCAT from microspore-derived cell suspension cultures of *B. napus* L. cv Jet Neuf could utilize 18:1, 18:0 and 16:0-CoA equally at concentrations around 20  $\mu\text{M}$ , but there was a preference for 18:1-CoA at higher concentrations [64]. Thus, the *in vitro* assay conditions should be considered when interpreting and comparing LPCAT specificity studies.

The reverse reaction of LPCAT is believed to be one mechanism by which acyl groups can be returned to the acyl-CoA pool from PC [61]. *In vitro* studies with safflower microsomes have shown that in the presence of CoA and an acyl-CoA binding protein (such as BSA), the radiolabel from *sn*-2-[ $^{14}\text{C}$ ]18:2-PC was transferred to the acyl-CoA pool [61]. Subsequent radioisotope feeding studies have suggested that acyl exchange occurs at both the *sn*-1 and *sn*-2 positions [57], although it is unclear whether the LPCAT acts at both positions and how much acyl-editing it accounts for. The reverse reaction of LPCAT is less efficient than the forward reaction and corresponds to about 5% of the acylation rate [58]. No data are currently available on substrate specificity of the LPCAT reverse reaction in plants, but experiments in rat liver microsomes have demonstrated that 20:4, 18:0 and 18:2-CoA were preferentially synthesized from endogenous lipids under ATP-independent conditions [65]. Thus, it is reasonable to expect that substrate specificities of both the forward and reverse reactions of LPCAT probably play a role in acyl-editing, and a more comprehensive understanding of these preferences may be key to overcoming substrate dichotomy in transgenic plants. The involvement of acyl-CoA binding proteins in mediating the reverse reaction of LPCAT may also indirectly influence the substrate specificities of the reaction, because the binding proteins may have varying affinities for different acyl-CoA species, which would influence the substrate availability for the LPCAT reaction.

PDAT substrate specificity has only been characterized in a few species but it appears to exhibit some preference for polyunsaturated or unusual acyl groups. Ricinoleoyl, vernoloyl, 18:3 and 20:4 groups at the *sn*-2 position of PC were more efficiently utilized by PDAT in *Arabidopsis* microsomes compared to 18:2, 18:1 and 10:0, while 18:0 and 22:1 were excluded from the reaction [7]. It appears that PDAT plays a role in membrane maintenance by channeling unusual acyl groups from phospholipids into storage lipids. In some plants that produce unusual fatty acids (*Ricinus communis*,

*Crepis palaestina*), PDAT has been shown to play a major role in removing these unusual acyl groups from membrane phospholipids, while others (*Crepis rubra*, *Euphorbia lagascae*) rely more on different mechanisms [66]. In *Arabidopsis*, it was demonstrated that PDAT is probably not a major determinant of TAG composition or content in developing seeds [67]. It is possible that other acyl-editing mechanisms, such as the Land's Cycle [68], or exchange of backbones between PC and diacylglycerol catalyzed by choline phosphotransferase contribute to trafficking of acyl groups between PC and TAG in these species. The relative contribution of various acyl-editing mechanisms to overall TAG composition, however, has yet to be completely dissected.

### Future outlook

Recent attempts to produce unusual fatty acids in transgenic plants have faced several challenges, which have the potential to drive major advances in our understanding of plant lipid metabolism. Earlier studies on the substrate preferences of various acyltransferases have often been limited to the use of crude microsomal preparations. Today, the ability to study acyltransferases in recombinant systems has revealed differences in substrate utilization between previously undistinguishable acyltransferase isoforms and is already providing the necessary insight to increase the levels of unusual fatty acids obtainable in transgenic plants. Similarly, an emerging understanding of the role of acyl-editing mechanisms and their associated specificities will be crucial to overcoming existing bottlenecks in acyl group trafficking between various substrate pools. Insights into the structure–function and evolutionary relationships between various acyltransferases may also reveal new targets for biotechnological modification of plant lipids.

### Acknowledgements

The authors acknowledge the support (to RJW) of the Natural Sciences and Engineering Research Council of Canada, the Alberta Agricultural Research Institute, AVAC, Ltd., the Alberta Crop Industry Development Fund, Alberta Advanced Education and Technology, Genome Alberta and the Canada Research Chairs Program in relation to his research on acyltransferase action. We thank Dr Disa Brownfield for her assistance in proofreading the manuscript.

### References

- Drexler, H. *et al.* (2003) Metabolic engineering of fatty acids for breeding of new oilseed crops: strategies, problems and first results. *J Plant Physiol* 160, 779–802
- Larson, T.R. *et al.* (2002) Acyl-CoA profiles of transgenic plants that accumulate medium-chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. *Plant J* 32, 519–527
- Cahoon, E.B. *et al.* (2007) Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr Opin Plant Biol* 10, 236–244
- Weselake, R.J. (2005) Storage lipids. In *Plant Lipids: Biology, Utilization and Manipulation* (Murphy, D.J., ed.), pp. 162–221, Blackwell Publishing
- Harwood, J.L. (2005) Fatty acid biosynthesis. In *Plant Lipids: Biology, Utilization and Manipulation* (Murphy, D.J., ed.), pp. 27–66, Blackwell Publishing
- Dahlqvist, A. *et al.* (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
- Stahl, U. *et al.* (2004) Cloning and functional characterization of a phospholipid:diacylglycerol acyltransferase from *Arabidopsis*. *Plant Physiol.* 135, 1324–1335
- Zheng, Z. *et al.* (2003) *Arabidopsis* AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* 15, 1872–1887
- Beisson, F. *et al.* (2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *Plant Cell* 19, 351–368
- Li, Y. *et al.* (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18339–18344
- Christie, W.W. *et al.* (1991) Stereospecific analysis of triacyl-*sn*-glycerols via resolution of diastereomeric diacylglycerol derivatives by high performance liquid chromatography on silica. *J. Am. Oil Chem. Soc.* 68, 695–701
- Brokerhoff, H. *et al.* (1966) Stereospecific analyses of several vegetable fats. *J. Lipid Res.* 7, 62–64
- Lisa, M. *et al.* (2008) Triacylglycerols profiling in plant oils important in food industry, dietetics and cosmetics using high-performance liquid chromatography–atmospheric pressure chemical ionization mass spectrometry. *J. Chromatogr. A* 1198–1199, 115–130

- 14 Dutta, P.C. *et al.* (1992) Utilization of petroselinic acid (C18:1Δ6) by glycerol acylation enzymes in microsomal preparations of developing embryos of carrot (*Daucus carota* L.), safflower (*Carthamus tinctorius* L.) and oil rape (*Brassica napus* L.). *Plant Sci.* 81, 57–64
- 15 Suh, M.C. *et al.* (2002) What limits production of unusual monoenoic fatty acids in transgenic plants? *Planta* 215, 584–595
- 16 Voelker, T. *et al.* (2001) Variations in the biosynthesis of seed storage lipids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 335–361
- 17 Voelker, T.A. *et al.* (1996) Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. *Plant J.* 9, 229–241
- 18 Wiberg, E. *et al.* (2000) The distribution of caprylate, caprate and laurate in lipids from developing and mature seeds of transgenic *Brassica napus* L. *Planta* 212, 33–40
- 19 Laurent, P. *et al.* (1992) Organ- and development-specific acyl coenzyme A lysophosphatidate acyltransferases in palm and meadowfoam. *Plant Physiol.* 99, 1711–1715
- 20 Sun, C. *et al.* (1988) Acyl Coenzyme A preference of the glycerol phosphate pathway in the microsomes from the maturing seeds of palm, maize, and rapeseed. *Plant Physiol.* 88, 56–60
- 21 Oo, K.C. *et al.* (1989) Lysophosphatidate acyltransferase activities in the microsomes from palm endosperm, maize scutellum, and rapeseed cotyledon of maturing seeds. *Plant Physiol.* 91, 1288–1295
- 22 Bafar, M. *et al.* (1990) Properties of the glycerol acylating enzymes in microsomal preparations from the developing seeds of safflower (*Carthamus tinctorius*) and turnip rape (*Brassica campestris*) and their ability to assemble cocoa-butter type fats. *J. Am. Oil Chem. Soc.* 67, 217–225
- 23 Bernerth, R. *et al.* (1990) Utilization of erucyl-CoA by acyltransferases from developing seeds of *Brassica napus* L. involved in triacylglycerol biosynthesis. *Plant Sci.* 67, 21–28
- 24 Cao, Y.Z. *et al.* (1990) Lysophosphatidate acyltransferase in the microsomes from maturing seeds of meadowfoam (*Limnanthes alba*). *Plant Physiol.* 94, 1199–1206
- 25 Lohden, I. *et al.* (1992) Triacylglycerol biosynthesis in developing seeds of *Tropaeolum majus* L. and *Limnanthes douglasii* R. Br. *Planta* 188, 215–224
- 26 Davies, H.M. *et al.* (1995) Lysophosphatidic acid acyltransferase from immature coconut endosperm having medium chain length substrate specificity. *Phytochemistry* 39, 989–996
- 27 Sorensen, B.M. *et al.* (2005) Storage lipid accumulation and acyltransferase action in developing flaxseed. *Lipids* 40, 1043–1049
- 28 Brown, A.P. *et al.* (2002) Substrate selectivity of plant and microbial lysophosphatidic acid acyltransferases. *Phytochemistry* 61, 493–501
- 29 Frentzen, M. (1993) Acyltransferases and triacylglycerols. In *Lipid Metabolism in plants* (Moore, Jr., T.S. (ed.)), pp. 195–230, CRC Press
- 30 Knutzon, D.S. *et al.* (1999) Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the sn-2 position of triacylglycerols in lauric rapeseed oil and can increase total laurate levels. *Plant Physiol.* 120, 739–746
- 31 Lassner, M.W. *et al.* (1995) Lysophosphatidic acid acyltransferase from meadowfoam mediates insertion of erucic acid at the sn-2 position of triacylglycerol in transgenic rapeseed oil. *Plant Physiol.* 109, 1389–1394
- 32 Brough, C.L. *et al.* (1996) Towards the genetic engineering of triacylglycerols of defined fatty acid composition: major changes in erucic acid content at the sn-2 position affected by the introduction of a 1-acyl-sn-glycerol-3-phosphate acyltransferase from *Limnanthes douglasii* into oil seed rape. *Mol. Breed.* 2, 133–142
- 33 Zou, J. *et al.* (1997) Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast sn-2 acyltransferase gene. *Plant Cell* 9, 909–923
- 34 Katavic, V. *et al.* (2001) Improving erucic acid content in rapeseed through biotechnology: What can the *Arabidopsis* FAE1 and the yeast SLC1-1 genes contribute? *Crop Sci.* 41, 739–747
- 35 Taylor, D.C. *et al.* (2002) Field testing of transgenic rapeseed cv. Hero transformed with a yeast sn-2 acyltransferase results in increased oil content, erucic acid content and seed yield. *Mol. Breed.* 8, 317–322
- 36 Taylor, D.C. *et al.* (1991) Triacylglycerol bioassembly in microspore-derived embryos of *Brassica napus* L. cv Reston. *Plant Physiol.* 97, 65–79
- 37 Princen, L.H. *et al.* (1984) Development of new crops for industrial raw materials. *J. Am. Oil Chem. Soc.* 61, 281–289
- 38 Perry, H.J. *et al.* (1993) Changes in the lipid content of developing seeds of *Brassica napus*. *Phytochemistry* 32, 1411–1415
- 39 Ichihara, K. *et al.* (1982) Lipid metabolism in safflower. 4. Some properties of diacylglycerol acyltransferase in a particulate fraction from maturing safflower seeds. *Phytochemistry* 21, 1895–1901
- 40 Ichihara, K. *et al.* (1988) Diacylglycerol acyltransferase in maturing safflower seeds – its influences on the fatty acid composition of triacylglycerol and on the rate of triacylglycerol synthesis. *Biochim. Biophys. Acta* 958, 125–129
- 41 Martin, B.A. *et al.* (1983) Properties of diacylglycerol acyltransferase from spinach leaves. *Lipids* 18, 1–6
- 42 Taylor, D.C. *et al.* (1992) Biosynthesis of acyl lipids containing very-long chain fatty acids in microspore-derived and zygotic embryos of *Brassica napus* L. cv Reston. *Plant Physiol.* 99, 1609–1618
- 43 Vogel, G. *et al.* (1996) Cholinephosphotransferase and diacylglycerol acyltransferase – substrate specificities at a key branch point in seed lipid metabolism. *Plant Physiol.* 110, 923–931
- 44 Yu, K.S. *et al.* (2006) Diacylglycerol acyltransferases from Vernonia and Stokesia prefer substrates with vernolic acid. *Lipids* 41, 557–566
- 45 Lung, S.C. *et al.* (2006) Diacylglycerol acyltransferase: a key mediator of plant triacylglycerol synthesis. *Lipids* 41, 1073–1088
- 46 He, X.H. *et al.* (2004) Cloning and characterization of a cDNA encoding diacylglycerol acyltransferase from castor bean. *Lipids* 39, 311–318
- 47 Shockey, J.M. *et al.* (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
- 48 Bursal, J. *et al.* (2008) Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnol. J.* 6, 819–831
- 49 Katavic, V. *et al.* (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
- 50 Zheng, P. *et al.* (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat. Genet.* 40, 367–372
- 51 Jako, C. *et al.* (2001) Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol.* 126, 861–874
- 52 Weselake, R.J. *et al.* (2008) Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. *J. Exp. Bot.* 59, 3543–3549
- 53 Taylor, D.C. *et al.* (2009) Molecular modification of triacylglycerol accumulation by over-expression of DGAT1 to produce canola with increased seed oil content under field conditions. *Botany* 87, 533–543
- 54 Lardizabal, K. *et al.* (2008) Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiol.* 148, 89–96
- 55 Siloto, R.M. *et al.* (2009) Directed evolution of acyl-CoA:diacylglycerol acyltransferase: development and characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiol. Biochem.* 47, 456–461
- 56 Williams, J.P. *et al.* (2000) The role of phosphatidylcholine in fatty acid exchange and desaturation in *Brassica napus* L. leaves. *Biochem. J.* 349, 127–133
- 57 Bates, P.D. *et al.* (2007) Incorporation of newly synthesized fatty acids into cytosolic glycerolipids in pea leaves occurs via acyl editing. *J. Biol. Chem.* 282, 31206–31216
- 58 Ichihara, K. *et al.* (1995) 1-Acylglycerophosphocholine O-acyltransferase in maturing safflower seeds. *Planta* 196, 551–557
- 59 Szymne, S. *et al.* (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
- 60 Szymne, S. *et al.* (1985) Oil synthesis in vitro in microsomal membranes from developing cotyledons of *Linum usitatissimum* L. *Planta* 164, 101–104
- 61 Szymne, S. *et al.* (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
- 62 Griffiths, G. *et al.* (1985) The acylation of sn-glycerol-3-phosphate and the metabolism of phosphatidate in microsomal preparations from the developing cotyledons of safflower (*Carthamus tinctorius* L.) seed. *Biochem. J.* 230, 379–388
- 63 Shen, W. *et al.* (2008) Functional study of lysophosphatidylcholine acyltransferase in *Arabidopsis*. Poster presented at 18th International Symposium on Plant Lipids, Bordeaux, France, July 20–25
- 64 Furukawa-Stoffer, T.L. *et al.* (2003) Properties of lysophosphatidylcholine acyltransferase from *Brassica napus* cultures. *Lipids* 38, 651–656
- 65 Sugiura, T. *et al.* (1995) Coenzyme A dependent modification of fatty acyl chains of rat liver membrane phospholipids – possible involvement of ATP independent acyl-CoA synthesis. *J. Lipid Res.* 36, 440–450
- 66 Banas, A. *et al.* (2000) The involvement of phospholipid:diacylglycerol acyltransferases in triacylglycerol production. *Biochem. Soc. Trans.* 28, 703–705

- 67 Mhaske, V. *et al.* (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid:diacylglycerol transacylase gene (At5g13640). *Plant Physiol. Biochem.* 43, 413–417
- 68 Lands, W.E.M. (1960) Metabolism of glycerolipids. 2. Enzymatic acylation of lysolecithin. *J. Biol. Chem.* 235, 2233–2237
- 69 Xu, J. *et al.* (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 Biotechnol. J.* 6, 799–818
- 70 Bouvier-Nave, P. *et al.* (2000) Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase. *Eur. J. Biochem.* 267, 85–96
- 71 Jain, R.K. *et al.* (2000) Enhancement of seed oil content by expression of glycerol-3-phosphate acyltransferase genes. *Biochem. Soc. Trans.* 28, 958–961
- 72 Stymne, S. *et al.* (1987) Triacylglycerol biosynthesis. In *The Biochemistry of Plants, Vol. 9 Lipids: Structure and Function* (Stumph, P.K., ed.), pp. 175–214, Academic Press
- 73 Ghosal, A. *et al.* (2007) *Saccharomyces cerevisiae* phospholipid:diacylglycerol acyl transferase (PDAT) devoid of its membrane anchor region is a soluble and active enzyme retaining its substrate specificities. *Biochim. Biophys. Acta-Mol. Cell Biol. Lipids* 1771, 1457–1463