Metabolic Engineering of Arabidopsis for Production of Punicic Acid Using Different Promoters

by

Ziliang Song

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

Master of Science

in

Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Fatty acids with conjugated double bonds play a special role in determining both the nutritional and industrial uses of plant oils. Punicic acid (18:3$\Delta^{9cis,11trans,13cis}$), a conjugated fatty acid naturally enriched in the pomegranate (*Punica granatum*) seeds, has gained increasing attention from the biotechnology community towards its production in metabolically engineered temperate oilseed crops because of its significant health benefits. The present study focused on selecting the best of four heterologous promoters to drive the over-expression of the *P.granatum FATTY ACID CONJUGASE* (*PgFADX*) cDNA as a means of producing punicic acid in the seed oil of transgenic Arabidopsis. Among the four promoters for genes encoding seed storage proteins from different oilseed crop species, the linin promoter resulted in the highest level of punicic acid (19.5% of total fatty acids in Arabidopsis seed oil). Analysis of *PgFADX* in developing seeds further confirmed that the linin promoter was the best choice to achieve a high level of seed-specific expression of the transgene. The native promoter of the *PgFADX* gene was also investigated. A 2027-bp upstream fragment of the *PgFADX* gene isolated from the pomegranate genome was fused to the *β-GLUCURONIDASE* (*GUS*) reporter gene. GUS activity, however, was not detected in transformed Arabidopsis. A conserved profile of *cis*-regulatory elements in the four heterologous promoters were identified and discussed in relation to their possible role in regulating gene expression during plant development. The *PgFADX* promoter was found deficient in several *cis*-regulatory elements that may be crucial for seed specific gene expression in Arabidopsis. This might indicate that the transcription factors in Arabidopsis could not recognize the special *cis*-regulatory elements in the *PgFADX* promoter as those in pomegranate could to achieve a uniquely high expression level.

Preface

This thesis is original work by Ziliang Song. No part of the thesis has been previously published. While the majority of experiments were done by myself, this work was built on the basis of some preliminary experiments by the courtesy of others as specified in the following.

The pomegranate was grown in the greenhouse by Dr. Elzbieta Mietkiewska. She isolated the total RNA from the pomegranate seeds, and synthesized the *PgFADX* cDNA by reverse transcription. As PCR templates, plasmids carrying the following three heterologous promoters: napin, phaseolin and linin, were used and kindly provided by Dr. Mietkiewska. Plasmid carrying the conlinin promoter used as a PCR template was provided by Dr. Qiu, the corresponding author of an article cited (Truksa et al. 2003). The modified binary vector pPZP-RCS1 (Goderis et al. 2002) carrying the *NPTII* expression cassette for plant screening was developed earlier by Dr. Mietkiewska (unpublished data).

Acknowledgements

First and foremost, my sincere gratitude goes to my supervisor, Dr. Randall Weselake, for his guidance, support and inspiration throughout my graduate research. I am very thankful for the opportunity he offered me to be part of his group within the Alberta Innovates Phytola Centre. His broad and profound erudition has brought me to a variety of interdisciplinary areas including lipid biochemistry and plant genetics. I appreciate his patience and diligence during the review of my thesis draft which in turn improved my academic writing. I also especially thank my committee member, Dr. Stephen Strelkov, for his constructive advice on my thesis project, and Dr. Natarj Kav for his critical role as my external examiner.

This work would not have been possible without the technical guidance of Dr. Elzbieta Mietkiewska, who taught me in the lab skills of molecular biology, spent much time on troubleshooting of my experiments and provided feedback on my writing through editing. In addition, I also acknowledge the following lecturers of my graduate courses in Biological Sciences. Drs. Janice Cooke and Enrico Scarpella provided me with a comprehensive background in plant science.

The Phytola Centre has provided me with a positive environment for research. I thank all the staffs and students, especially Chris Kazala and Annie Wong, for their responsible management of our lab, Dr. Michael Greer (Scott) for his golden touch in many aspects of my project, and Yang (Sarena) Xu, Yingyu (Rain) Lin and Kristian Caldo for being my "all-weather" friends who provided me with help and comfort as I moved forward in my research. I also thank Nikki-Karyssa Scott and Arlene Oatway for their kind assistance in using the facilities.

Furthermore, my appreciation goes to my house mate Luke McNulty for improving my English which greatly facilitated my writing process. I owe a debt of gratitude to my loved ones: my parents, my old friends in China, my new friends in Canada, particularly those contributed to my spiritual growth as a Christian. Finally, I am mostly grateful to God for every blessing and His everlasting love in me.

Table of Contents

List of Figures

List of Tables

Abbreviations

AAT: acetyl-CoA-ACP transacetylase

ABRE: abscisic acid response element

ACCase: acetyl-CoA carboxylase

ACP: acyl carrier protein

ANOVA: one-way analysis of variance

CaMV: Cauliflower Mosaic Virus

cDNA: complementary DNA

CDP: cytidine diphosphate

CFA: conjugated fatty acid

CLNA: conjugated linolenic acid

CoA: coenzyme A

DAG: diacylglycerol

DAP: days after pollination

DGAT: diacylglycerol acyltransferase

DHA: docosahexaenoic acid

DNA: deoxyribonucleic acid

DPA: docosapentaenoic acid

EAR: enoyl-ACP reductase

EPA: eicosapentaenoic acid

ER: endoplasmic reticulum

FA: fatty acid

FAD: fatty acid desaturase

FAE: fatty acid elongase

FAT: acyl-ACP thioesterase

FAX: fatty acid export

G3P: *sn*-glycerol-3-phosphate

GPAT: *sn*-glycerol-3-phosphate acyltransferase

GUS: β-glucuronidase

HSD: honestly significant difference

LACS: long-chain acyl-CoA synthetase

LC-MUFA: long-chain monounsaturated fatty acid

LPA: *lyso*phosphatidic acid

LPC: *lyso*phosphatidylcholine

LPAAT: *lyso*phosphatidic acid acyltransferase

LPCAT: *lyso*phosphatidylcholine acyltransferase

KAR: ketoacyl-ACP reductase

KAS: β-ketoacyl-ACP synthase

MAT: malonyl-CoA-ACP transferase

PA: phosphatidic acid

PAP: phosphatidic acid phosphatase

PC: phosphatidylcholine

PDAT: phospholipid: diacylglycerol acyltransferase

PDCT: phosphatidylcholine: diacylglycerol cholinephosphotransferase

PgFADX: *Punica granatum* fatty acid conjugase

PLA: phospholipase A

PKC: protein kinase C

PUFA: polyunsaturated fatty acid

qRT-PCR: quantitative real-time polymerase chain reaction

RNA: ribonucleic acid

TAG: triacylglycerol

TF: transcription factor

TSS: transcription starting site

WT: wild type

1. Introduction

Higher plants accumulate storage products such as proteins and oils in their seeds as a source of nutrients and energy for seed germination. The largest proportion of plant oils is produced from domesticated oilseed crops for human consumption as food and feed (Dyer et al. 2008). The properties of oils depend greatly on their fatty acid (FA) composition. While only five common FAs are found in most domesticated oilseed crops, there are a wide variety of unusual FAs in the seed oil of many wild plant species (Aitzetmuller et al. 2003). These include variations in the number and positional arrangement of double bonds such as those found in polyunsaturated fatty acids (PUFAs) and conjugated fatty acids (CFAs), as well as addition of diverse functional groups such as hydroxyl and epoxy groups. The structural specialties confer unusual FAs outstanding functionalities with high value for nutritional and industrial applications that are distinct from the basic dietary role of the common FAs (Dyer et al. 2008). However, plant species that naturally accumulate such unusual FAs are mostly unsuitable for modern agricultural practices (Jaworski and Cahoon 2003).

To overcome this limitation, metabolic engineering is applied to enable organisms to produce desired unusual FAs through introducing exogenous genes and/or modifying the endogenous genes involved in the biosynthetic pathways of target products (Yoon et al. 2013). With considerable efforts directed towards identification and over-expression of the genes encoding the unusual FA biosynthetic enzymes, the “designer oil” has been on the horizon, but numerous limiting factors are yet to be solved (Napier and Graham 2010). One significant bottleneck resides in the sub-optimal strength or improper spatiotemporal specificity of particular promoters (Rossak et al. 2001). While an array of promoters have been discovered from various oilseed crops and fully characterized over the past few decades, there is a need for a side-by-side

comparison of the relative effectiveness in driving the expression of the particular genes among different promoters (Boothe et al. 2010).

The present study focuses on punicic acid as an example of an unusual FA. Punicic acid is naturally accumulated to ~80% of total FAs in the seed oil of pomegranate (*Punica granatum*). It has recently attracted scientific interest because of its conjugated double bonds which are associated with extraordinary health benefits (Aruna et al. 2016). As pomegranate's growth is restricted to sub-tropical and tropical climates and the production of seed oil is modest, engineering existing temperate oilseed crops to produce punicic acid would be a preferable strategy to improve the nutritional value of vegetable oil and the bioavailability of punicic acid. Identification of the cDNA encoding the *P. granatum* fatty acid conjugase, PgFADX, which catalyzes the critical step in the biosynthesis of punicic acid represents an early breakthrough (Iwabuchi et al. 2003). It was later reported that combined seed-specific expression of *PgFADX* and *P. granatum FATTY ACID DESATURASE* (*PgFAD2*) in *Arabidopsis thaliana* (Arabidopsis) could lead to levels of about 20% (w/w) punicic acid in the seed oil (Mietkiewska et al. 2014b). Further enhancement of punicic acid content in the seed oil of Arabidopsis, and temperate oil crops, could benefit from use of more effective promoters. Based on the difference of expression pattern between genes encoding seed storage proteins and those related to FA-modifying enzymes, two hypotheses are presented in the case of production of punicic acid in the engineered oilseed crops: a. promoters of genes encoding different storage proteins vary in their effectiveness in driving the expression of *PgFADX* and subsequent production of punicic acid in Arabidopsis seeds; b. the promoter of *PgFADX* from pomegranate is very effective in driving the expression of *PgFADX* so as to result in relatively high levels of punicic acid. To test these hypotheses, experiments were designed to compare four promoters of genes encoding storage

proteins from different oilseed crops, as well as the promoter of *PgFADX* which was isolated and characterized from the pomegranate, tested in Arabidopsis for its ability to drive the heterologous expression of *PgFADX*.

2. Literature Review

2.1 Storage lipid biosynthesis in plants

Plant fatty acid (FA) and triacylglycerol (TAG) biosynthesis has been the subject of numerous detailed reviews, e.g. (Chapman and Ohlrogge 2012; Chen et al. 2015; Ohlrogge and Browse 1995; Ohlrogge and Jaworski 1997). Below a brief description of these processes is presented which is based upon these more detailed reviews.

2.1.1 Fatty acid biosynthesis

FA biosynthesis in the plastid and export from the plastid are shown in Figure 1. *De novo* synthesis of fatty acids (FAs) requires malonyl-coenzyme A (CoA) as the central carbon donor. The ATP-dependent formation of malonyl-CoA from acetyl-CoA and bicarbonate is the first committed step of the pathway, catalyzed by acetyl-CoA carboxylase (ACCase). The following sequential reactions in the synthetic process are catalyzed by FA synthase, a dissociable multi-subunit complex consisting of 8 monofunctional enzymes. Before the fatty acyl chain can start to grow, the malonyl group of malonyl-CoA is activated by being linked to the thiol group (-SH) of acyl carrier protein (ACP), catalyzed by malonyl-CoA-ACP transferase (MAT). Similarly, the acetyl group of acetyl-CoA is activated by acetyl-CoA-ACP transacetylase (AAT). The first reaction in the formation of a fatty acyl chain is condensation of the activated acetyl and malonyl groups to extend the acyl chain by two carbons, catalyzed by β-ketoacyl-ACP synthase (KAS) III (KAS III). The β-keto product is then reduced to acyl-ACP in three more reactions: reduction of β-keto group catalyzed by ketoacyl-ACP reductase (KAR), dehydration of the carbonyl group catalyzed by β-hydroxyacyl-ACP dehydratase and reduction of the double bond catalyzed by enoyl-ACP reductase (EAR). KAS III has been shown to catalyze the condensation of acetyl-CoA with malonyl-ACP to form 4:0-ACP, then another type: KASI is responsible for the

elongation of 4:0-ACP to 16:0-ACP. The final elongation of 16:0-ACP to 18:0-ACP is catalyzed by KASII.

A double bond can be introduced to stearoyl (18:0)-ACP by a stromal stearoyl-ACP desaturase to form oleoyl (18:1Δ^{9cis}; hereafter 18:1)-ACP (Cahoon et al. 2010). In temperate oil-forming species such as *Arabidopsis thaliana* (Arabidopsis) and *Brassica napus* (oilseed rape), 18:1 is the major FA formed in the plastid along with smaller amounts of stearic acid (18:0) and palmitic acid (16:0). After the acyl groups are hydrolyzed from ACP through the catalytic action of acyl-ACP thioesterases (FATA and FATB), the FAs are exported from the plastid via the involvement of a newly discovered protein fatty acid export 1 (FAX1) (Li et al. 2015). On the outer membrane of the chloroplast the FAs are then activated to CoA esters by long-chain acyl-CoA synthetase (LACS) (Zhao et al. 2010). Some acyl chains in the acyl-CoA pool may be further elongated in the cytoplasm. The acyl-CoA pool serves as an acyl donor in the biosynthesis of glycerolipids, including TAG, in the endoplasmic reticulum in a process known as the "eukaryotic pathway". In addition, some of the FAs synthesized in the plastid remain in this organelle to support lipid biosynthesis in the plastidial "prokaryotic pathway".

2.1.2 Triacylglycerol biosynthesis

As the main constituent of storage lipids in plant seeds, TAGs are triesters of glycerol in which all three hydroxyl groups of the glycerol backbone are esterified by FAs. There are five predominant FA species in TAGs commonly found in oilseed crops: palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2$\Delta^{9cis,12cis}$; hereafter 18:2), and α-linolenic (18:3$\Delta^{9cis,12cis,15cis}$;

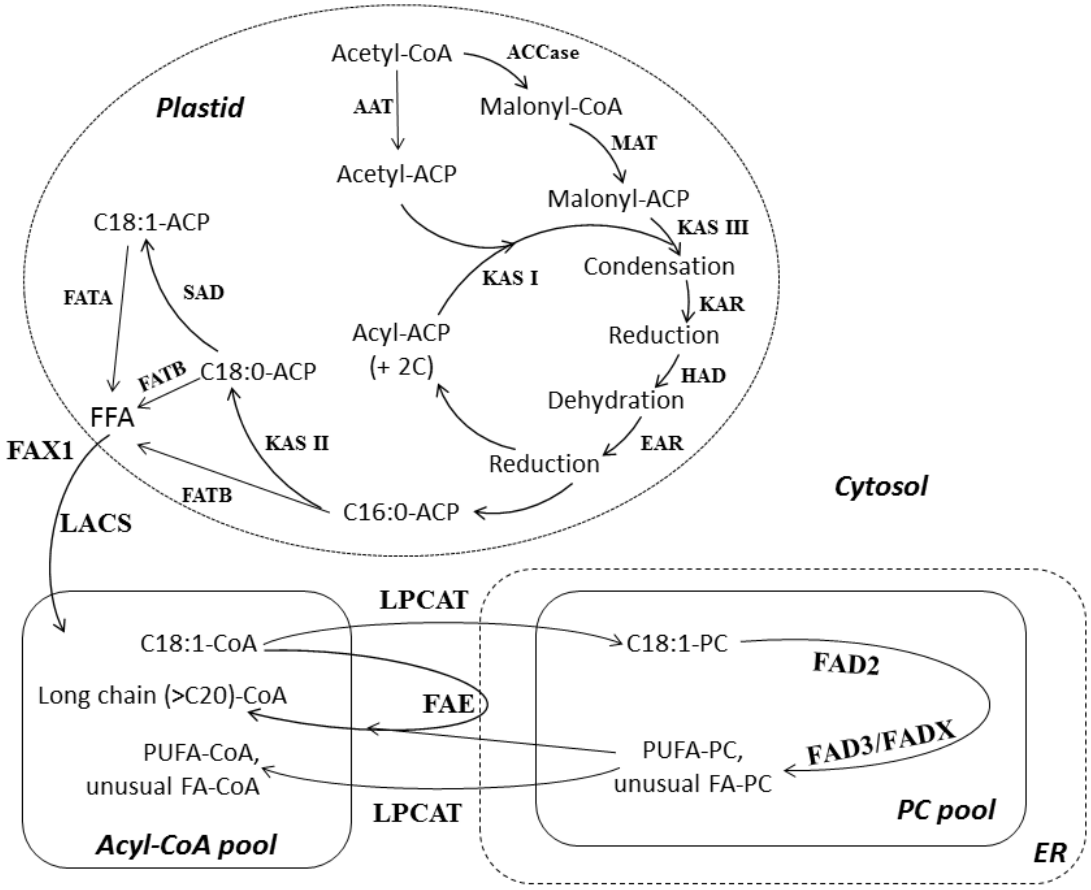


Figure 1. Biosynthesis of fatty acids (FAs) in plants. *De novo* synthesis of FAs takes place in the plastid, using acetyl-CoA as the building block. The activated acetyl and malonyl groups bound to acyl carrier protein (ACP) are condensed to form acetoacyl-ACP, which undergoes reduction, dehydration and again reduction to generate an acyl-ACP with two additional carbon atoms in the acyl chain. The FA chain grows to palmitoyl-ACP after seven cycles of the reactions. Palmitoyl-ACP can be elongated to stearoyl-ACP and then desaturated to oleoyl-ACP. The acyl chain is cleaved from ACP to release free fatty acid (FFA) which are exported to the cytosol. In the cytosol the FFA is assembled back to acyl-CoA for further modification or TAG synthesis in the ER. Modifications recruit oleoyl-CoA as the substrate which can be elongated to long chain FAs, desaturated to polyunsaturated FAs (PUFAs) or further modified to form

unusual FAs. Elongation occurs in the endoplasmic reticulum (ER) whereas other modifications rely on phosphatidylcholine (PC) as the intermediate. A series of desaturases and their divergent forms can edit the acyl chain esterified to PC. Enzymes involved in each reaction are abbreviations beside the arrows. AAT, acetyl-CoA-ACP transacetylase; ACCase, acetyl-CoA carboxylase; MAT, malonyl-CoA-ACP transferase; KAS, β-ketoacyl-ACP synthase; KAR, ketoacyl-ACP reductase; HAD, β-hydroxyacyl-ACP dehydratase; EAR, enoyl-ACP reductase; SAD, stearoyl-ACP desaturase; FATA/B, acyl-ACP thioesterases A,B; LACS, long-chain acyl-CoA synthetase; LPCAT, acyl-CoA: lysophosphatidylcholine acyltransferase; FAE, fatty acid elongase; FAD2, fatty acid desaturase 2 (Δ^{12} desaturase); FAD3, fatty acid desaturase 3 (Δ^{15} desaturase); FADX, divergent forms of FAD2.

hereafter α-18:3) acid (Singh et al. 2005).

TAG can be assembled in the endoplasmic reticulum (ER) via acyl-CoA-dependent and acyl-CoA-independent processes in a complex interplay with membrane metabolism (Figure 2). The Kennedy pathway or *sn*-glycerol-3-phosphate (G3P) pathway leading to TAG involves three acylation steps. The precursor of glycerol backbone, G3P, is first acylated at the *sn*-1 position by the catalytic action of *sn*-glycerol-3-phosphate acyltransferase (GPAT) to produce lysophosphatidic acid (LPA), which is further acylated at the *sn*-2 position by the catalytic action of lysophosphatidic acid acyltransferase (LPAAT) to become phosphatidic acid (PA). Then phosphatidic acid phosphatase (PAP) catalyzes the dephosphorylation of PA to *sn*-1, 2 diacylglycerol (DAG) and diacylglycerol acyltransferase (DGAT) catalyzes the final acylation of DAG to produce TAG. The Kennedy pathway is more prevalent in oil crops that produce TAG enriched in saturated FAs, such as coconut (*Cocos nucifera*), because polysaturated fatty acids (PUFAs) only represent two percent of the total FAs comprising this tropical/sub-tropical oil (Orsavova et al. 2015). Oilseed crops producing PUFAs or unusual FAs in their seed oils, however, rely on fatty acyl modifications which occur in the ER. A network of acyl editing reactions, head-group exchange reactions and acyl-CoA-independent-TAG formation appears to be involved in enriching PUFAs in seed TAG in temperate oil-forming species, such as Arabidopsis and *B. napus*. In these species, formation of PUFAs (18:2 and α-18:3) at the *sn*-2 position of phosphatidylcholine (PC) occurs through the sequential catalytic action of fatty acid desaturase (FAD) 2 and FAD3. PC substrate for this process is provided by the DAG skeleton produced in the Kennedy pathway. DAG formed *de novo* in this pathway can be converted to PC through the catalytic action of CDP-choline: DAG cholinephosphotransferase (CPT). In addition, phosphatidylcholine: diacylglycerol cholinephosphotransferase (PDCT) catalyzes the transfer of

the phosphocholine headgroup from PC with PUFA-containing acyl chains, to 18:1-enriched DAG produced *de novo* in the Kennedy pathway. This process allows modified FAs, such as 18:2 and α-18:3, to enter TAG via the Kennedy pathway and at the same time 18:1-enriched PC to be further modified by FAD action. PUFAs at the *sn*-2 position of PC can also be transferred to *sn*-1, 2-DAG to form TAG via an acyl-CoA-independent reaction catalyzed by phospholipid:diacylglycerol acyltransferase (PDAT), thus providing another channel for moving PUFA into TAG. *Lyso*phosphatidylcholine acyltransferase (LPCAT)-mediated acyl-exchange between acyl moieties at the *sn*-2 position of PC and the acyl-CoA pool represents another possibility for enriching PUFAs in TAGs. The forward reaction of LPCAT catalyzes the formation of PC from acyl-CoA and *lyso*phosphatidylcholine (LPC), whereas the reverse reaction can potentially lead to acyl-CoA formation. The Land's cycle may also contribute to PUFA enrichment in TAG. In the Land's cycle, PUFA would be released from PC through the catalytic action of phospholipase A_2 (PLA_2) followed by FA-activation via the action of LACS. The reverse reaction of LPCAT can also exchange the acyl group between the *sn*-2 position of PC and the acyl-CoA pool. The resulting LPC can then be re-acylated by the catalytic action of the forward reaction of LPCAT to complete the Land's cycle. The resulting PUFA-CoA, derived from LACS action, is now available as an acyl-donor for the Kennedy pathway. Similar processes to those described above appear to operate in oil-forming species which produce TAG containing unusual FAs such as puncic acid ($18{:}3\Delta^{9cis,11trans,13cis}$) (Mietkiewska et al. 2014a). In the case of pomegranate seed oil formation, the species utilizes the sequential action of FAD2 and a fatty acid conjugase to form punicoyl acyl chains on PC from *sn*-2-18:1-PC.

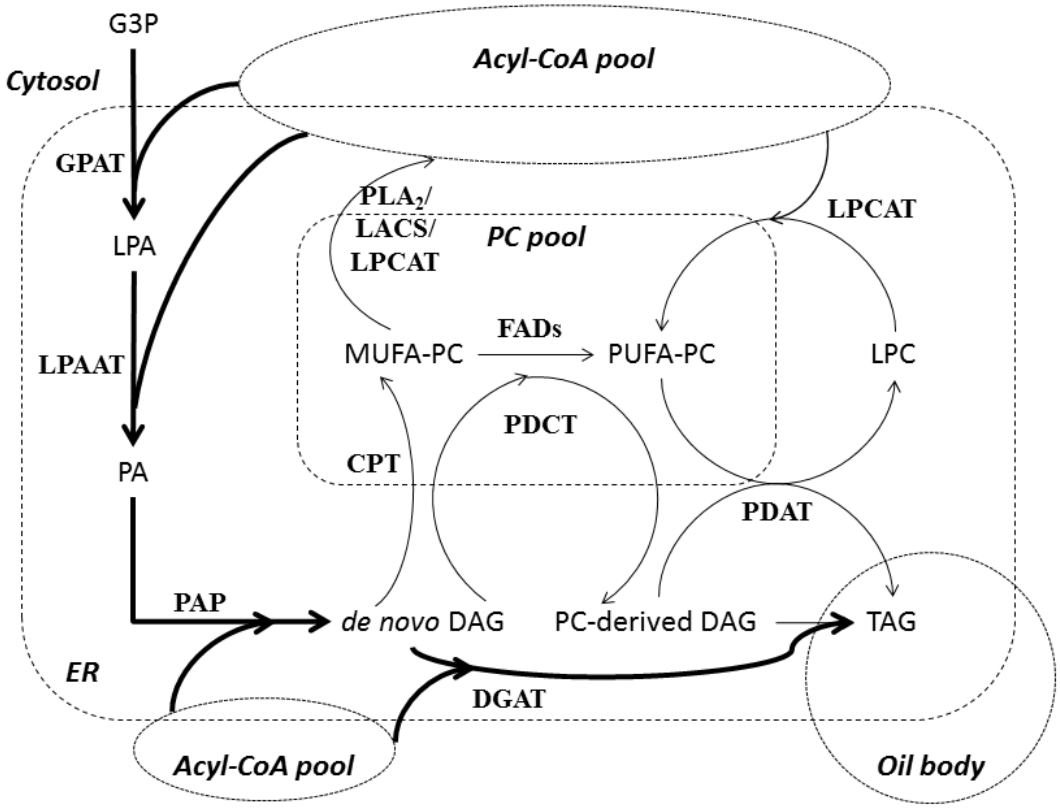


Figure 2. General scheme of triacylglycerol (TAG) assembly in oilseeds. The route in bold represents the Kennedy pathway, which starts with G3P as the glycerol backbone. The acyl-CoA pool serves as the direct acyl donor, stepwise acylating the glycerol backbone to produce TAG. TAG accumulates in the outer membrane of ER and eventually pinches off from the ER in the form of oil bodies. TAG can also be synthesized through acyl-CoA independent pathways involving the flux of acyl chain from PC to DAG. The DAG derived from PC carries varied types of FAs edited in PC to form TAG. This mechanism is particularly important in the case where plants accumulate abundant unusual FAs in their seed oil. Abbreviations: G3P, *sn*-glycerol-3-phosphate; LPA, *lyso*phosphatidic acid; PA, phosphatidic acid; DAG, *sn*-1,2 diacylglycerol; TAG, triacylglycerol; PC, phosphatidylcholine; MUFA, monounsaturated fatty

acid; PUFA, polyunsaturated fatty acid; LPC, *lyso*phosphatidylcholine; GPAT, *sn*-glycerol-3-phosphate acyltransferase; LPAAT, *lyso*phosphatidic acid acyltransferase; PAP, phosphatidic acid phosphatase; DGAT, diacylglycerol acyltransferase; FAD, fatty acid desaturase; PLA_2: phospholipase A_2; LPCAT, acyl-CoA: *lyso*phosphatidylcholine acyltransferase; CPT, CDP-choline: DAG cholinephosphotransferase; PDCT, PC: DAG cholinephosphotransferase; PDAT, phospholipid: DAG acyltransferase; ER, endoplasmic reticulum.

2.2 Production of conjugated fatty acids in plants

2.2.1 Mining the potential value of conjugated fatty acids

In contrast to animals, higher plants have evolved to collectively produce a wide variety of FAs. Besides the five major FAs common to most of the oilseeds, there is a rich diversity of FAs present in nature and classified as "unusual FAs" or "novel FAs". Most of the unusual FAs are derived from the five common FAs by elongation to a longer chain length (i.e., more than 18 carbons, examples: arachidic acid, 20:0; erucic acid, 22:1Δ^{13cis}) or other modifications including altering the position, configuration and number of double bonds (e.g., PUFAs) and adding new functional groups to the acyl chain (e.g., hydroxyl, epoxy groups). This wide array of unusual FAs found in the plant kingdom provides useful functionalities for both nutraceutic applications and industrial applications as chemical feedstocks (Dyer et al. 2008). For example, ricinoleic acid (12-OH 18:1Δ^{9cis}), a hydroxyl-FA produced in castor bean (*Ricinus communis*), is the feedstock for many industrial products including lubricants, nylon and cosmetics (Jaworski and Cahoon 2003).

Conjugated fatty acids (CFAs) are a typical class of structurally modified unusual PUFAs, with at least one pair of conjugated non-methylene interrupted double bonds (Mietkiewska et al. 2014a). Oils with CFAs exhibit higher rates of oxidation relative to other PUFAs with methylene-interrupted double bonds (Cahoon et al. 1999). This unique property makes CFAs valuable for specific industrial applications, exemplified by the well-known α-eleostearic acid (18:3$\Delta^{9cis,11trans,13trans}$) enriched in tung (*Vernicia fordii*) oil that has unique drying quality in formulations of paints, varnishes, coatings and resins (Biermann et al. 2011; Dyer et al. 2002). In addition to industrial value, significant health benefits have also been reported in some novel CFAs including calendic acid (18:3$\Delta^{8trans,10trans,12cis}$) (Li et al. 2013; Yasui et al. 2006), catalpic

acid (18:3$\Delta^{9trans,11trans,13cis}$) (Hontecillas et al. 2008) and jacaric acid (18:3Δ^{8}cis,10$trans,\ 12cis$) (Liu and Leung 2014; Shinohara et al. 2012). These CFAs are conjugated linolenic acid (CLNA) isomers with three double bonds in conjugated configuration. Structures of a few CLNAs found in various plant species are shown in Figure 3.

Punicic acid is a CLNA isomer and the main FA naturally produced at high concentrations (65-80%) in the seed oil of pomegranate (*P. granatum*), a fruit native to the Middle East which has been used as a functional food for centuries (Johanningsmeier and Harris 2011). A number of medical studies have associated punicic acid with the prevention and treatment of a wide range of diseases. It has been shown earlier that punicic acid has strong anti-cancer effects due to its ability to inhibit proliferation of various types of human cancer cells. For instance, intrinsic apoptosis of human breast cancer cells was induced by punicic acid via lipid peroxidation and the Protein Kinase C (PKC) pathway (Grossmann et al. 2010). Promising results wherein punicic acid was used to treat prostate cancer (Gasmi and Sanderson 2010; Wang et al. 2014; Wang and Martins-Green 2014) and colon cancer (Kohno et al. 2004) have been reported in either rat or human trials. Punicic acid has also been shown to be useful in treating metabolic syndrome which is characterized by a cluster of risk factors for cardiovascular disease and type 2 diabetes mellitus including high fasting glucose, abdominal obesity, high blood pressure, low fasting high-density lipoprotein cholesterol and high fasting serum TAG (Calton et al. 2014).

Consumption of pomegranate seed oil rich in punicic acid led to reduced weight gain and improved insulin sensitivity in the mice fed with a high-fat diet, which is an indication that punicic acid protects against diet-induced obesity and lowers the risk of type 2 diabetes (Banihani et al. 2013; Hontecillas et al. 2009; Vroegrijk et al. 2011). Punicic acid has been further found to exhibit an activity in preventing osteoporosis (Spilmont et al. 2013) and

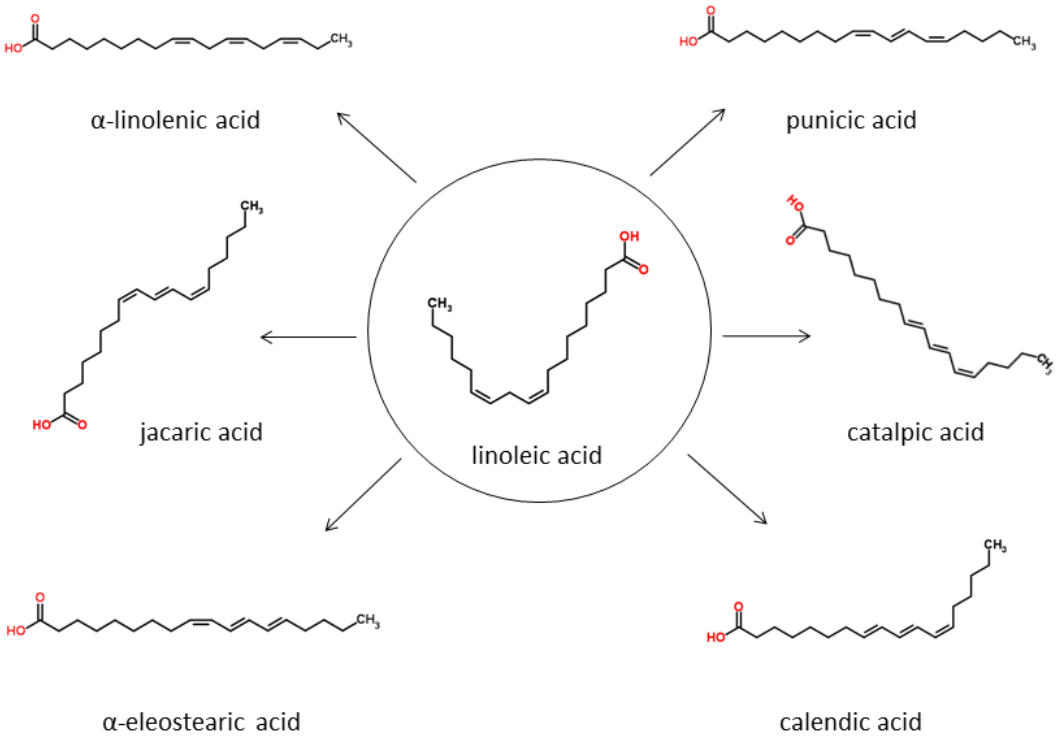


Figure 3. A few examples of conjugated linolenic fatty acids (CLNAs) derived from linoleic acid via the catalytic action of fatty acid conjugase (FADX). On the top left shows the common α-linolenic acid (not a CFA) formed via the catalytic action of fatty acid desaturase 2 (FAD2). The structures of fatty acids were retrieved from http://www.chemspider.com.

promoting skin regeneration, attributed to its anti-inflammatory and anti-oxidative properties (Johanningsmeier and Harris 2011). With these prominent health benefits, punicic acid has recently gained growing popularity as an ingredient in nutraceutic products (Yuan et al. 2014).

2.2.2 Metabolic engineering of oilseeds to produce conjugated fatty acids

Despite the diversity and high value of CFAs, most of the plants that accumulate these CFAs have limited agronomic potential and thus are not suited for large-scale cultivation. To address this issue, considerable effort has been directed towards modifying the fatty acid composition of commercial oilseed crops, with good agronomic traits, to improve the nutritional and industrial functionality of their seed oils. With advances in the basic understanding of the biosynthesis of CFAs in plant species naturally producing CFAs, metabolic engineering has made the accumulation of the desired CFAs possible in transgenic temperate oilseed crops (Mietkiewska et al. 2014a; Napier 2007).

Early research on the biosynthesis of CFAs began with α-eleostearic acid. The conjugated double bond in this FA arises from the modification of the existing Δ^{12} double bond of linoleic acid which serves as the precursor in the conversion to α-eleostearic acid (Liu et al. 1997). Additional corroborative evidence demonstrated the activity of a divergent form of Δ^{12} oleic acid desaturase (FAD2) enzyme termed "conjugase" or FADX involved in the biosynthesis of CLNA in *Momordica charantia* or *Impatiens balsamina* seeds, respectively (Cahoon et al. 1999). In fact, many of the enzymes that catalyze the formation of conjugated double bonds are related to the microsomal FAD2 family found in Arabidopsis and temperate oilseed crops like *B. napus*. Expression of the identified complementary deoxyribonucleic acid (cDNA) encoding conjugases

has resulted in the accumulation of α-eleostearic acid and its isomer, calendic acid in Arabidopsis and soybean (*Glycine max*) (Cahoon et al. 1999; Cahoon et al. 2006).

The understanding of the FAD2-like family has been further extended with the identification of a new conjugase (FADX) involved in the biosynthesis of punicic acid naturally occurring in the seeds of pomegranate (*P. granatum*) and featuring conjugated double bonds at the $\Delta^{9cis,}$ $\Delta^{11trans}$ and Δ^{13cis} positions. The *P. granatum* conjugase (designated as PgFADX in the following) also recruits linoleic acid as the substrate, and catalyzes the conversion of the Δ^{12cis} double bond of linoleic acid into two conjugated double bonds to form punicic acid (Hornung et al. 2002; Iwabuchi et al. 2003). The cDNA that encodes PgFADX was isolated from maturing seeds of pomegranate (GenBank# AY178446) and expressed in Arabidopsis, which resulted in the accumulation of punicic acid up to ~4.4% (w/w) of the total seed oil (Iwabuchi et al. 2003).

However, in contrast to ~80% in the pomegranate seeds, the above achievement were modest and limited by some metabolic bottlenecks, which have been recently reviewed (Jaworski and Cahoon 2003; Mietkiewska et al. 2014a; Napier et al. 2014). The first metabolic bottleneck in the biosynthesis of CLNAs resides in the poor availability of the substrate, with less than 27% linoleic acid present in wild-type (WT) Arabidopsis seeds. This can be overcome by using the Arabidopsis *fad3/fae1* mutant with a higher background of linoleic acid (Cahoon et al. 2006; Mietkiewska et al. 2014b). The challenge also comes from the transcriptional repression of other relevant enzymes in the biosynthetic pathway by the unusual FA product itself. This was suggested based on the observation of a reduced expression level of *FAD2* in Arabidopsis over-expressing *PgFADX* (Mietkiewska et al. 2014b). In addition to transcriptional repression by the FA product, another possible limiting factor is that the exogenous gene encoding the FAD2-diverged conjugase could be subjected to post-transcriptional gene silencing

as a result of high DNA sequence homology between the native *FAD2* and exogenous *FADX* (Mietkiewska et al. 2014b). To overcome this obstacle, the gene encoding FAD2 from pomegranate (*PgFAD2*) was over-expressed in combination with *PgFADX* in the *fad3*/*fae1* Arabidopsis background. The resulting Arabidopsis transgenic lines accumulated punicic acid with up to 21% of total FAs in the seed oil (Mietkiewska et al. 2014b).

Greater progress in metabolic engineering for the production of unusual fatty acids requires a more rational design of genetic approaches to optimize the expression of the transgenes. Another challenge that limits the production of unusual FAs in genetically engineered oilseeds involves the inefficient trafficking of these FAs from PC into TAG (Cahoon et al. 2006; Napier et al. 2014). The native plants which normally produce unusual FAs have evolved unique enzymes for an efficient metabolic flux to both synthesize the FAs and channel them from PC to TAG. Recent studies have focused on the identification of the native acyl-trafficking enzymes from plants naturally producing usual FAs (Chen et al. 2015; Mietkiewska et al. 2014a). Recently, it has been demonstrated that in the transgenic plants, the accumulation of unusual FAs can be inhibited by the competition between endogenous and the transgenically-introduced isozymes for substrates, and reducing this competition may be a useful strategy to further boost the level of target FAs in metabolically engineered oilseeds (van Erp et al. 2015).

2.3 The role of the promoter in gene expression and metabolic engineering of plants

The production of unusual FAs in the engineered plants largely depends on the expression of transgenes that encode FA-modifying enzymes, which can be limited by the differential temporal pattern of gene expression (Jaworski and Cahoon 2003). It has been shown earlier in Arabidopsis that genes related to the biosynthesis of different storage components differ in their

expression profile during seed development. Genes encoding core FA synthesis enzymes display a bell-shaped pattern between 4 and 10 days after pollination (DAP) and fall sharply from day 10. By contrast, expression of genes involved in FA modification and storage proteins increases at the later developmental stage and remains high (Baud and Lepiniec 2009; Ruuska et al. 2002). This contrapuntal pattern of gene expression implies a fine-tuned regulatory network behind the scene, where promoters come into play. Thus, it is interesting to study whether the promoters of genes encoding for seed storage proteins from the oilseed crops or those for FA modifications from the native plants have better contribution to the transgenic production of unusual FAs.

A strong expression system with an appropriate promoter is the key for successful metabolic engineering. By definition, a promoter is the region of DNA located upstream or 5' of the coding sequence of a gene, structurally comprising proximal and distal regions (Porto et al. 2014) (Figure 4). The proximal region covers approximately 250 bp (base pair) upstream of the transcription starting site (TSS), within which the closest part (70-80 bp) to the TSS refers to "core promoter". It is the minimal portion of the promoter that serves as the binding site for ribonucleic acid (RNA) polymerase II to initiate transcription (Buchanan et al. 2000). The core promoter features a conserved AT-rich region positioned at ~25-30 bp surrounding the TSS and called "TATA box" (Porto et al. 2014). Upstream of the core promoter various regulatory sequences are distributed in the proximal and distal regions that mediate the fine regulation of gene expression at the transcription level (Hernandez-Garcia and Finer 2014). A promoter controls when, where and to what level the gene is expressed. Functionally, promoters used in metabolic engineering are classified into three groups: constitutive (active in all tissues continuously), spatiotemporal (active in specific tissues or developmental stages) and inducible (regulated by the application of an external chemical or physical signal) (Peremarti et al. 2010).

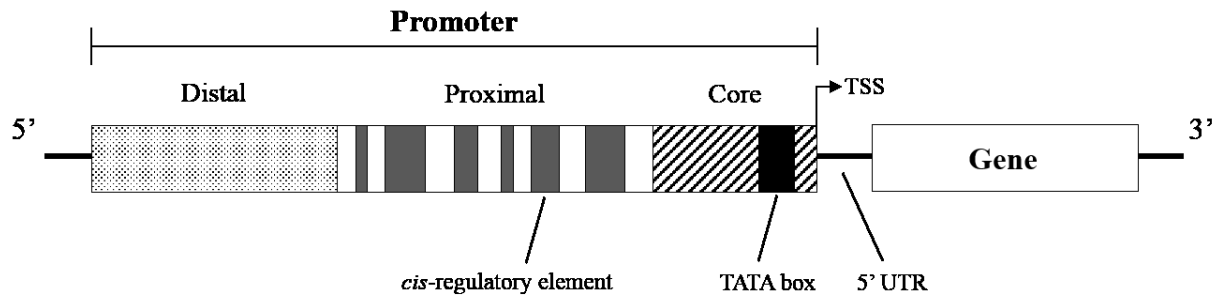


Figure 4. Diagrammatic model of a promoter linked to a gene. The TATA box is harbored in the core promoter and located 25-30 bp upstream of the transcription start site (TSS) of the gene. The proximal region contains *cis*-regulatory elements for the specific binding of transcription factors that contribute to the fine regulation of gene expression. The distal region following the proximal region contains additional regulatory elements such as enhancers and silencers that mediate transcription in other ways.

As mentioned before, seed development involves deposition of storage components such as proteins and oil (mostly TAG). However, with respect to metabolic engineering of unusual FAs in oilseeds, constitutive over-expression of the transgenes can have a negative effect on the engineered plant, including aberrant morphology and compromised growth (Hernandez-Garcia and Finer 2014). Therefore, spatiotemporal promoters with precise control of transgenes and restriction of gene expression to seeds or certain developmental stages merits the balance between enhancing the valuable crop traits and maintaining normal growth. Promoters of genes encoding seed storage proteins are attractive for this target and several of them have been studied in detail.

The promoter of the gene encoding Napin A (termed as napin promoter in the following) is a well-known example. The napin protein (2S) is one of the two major seed storage proteins in *B. napus* and encoded by a multigene family (Ellerstrom et al. 1996). Studies have revealed that the expression of the gene encoding Napin A is tightly regulated with expression targeted mainly to the embryo of the seed over most of the period of seed development, proving the seed specificity of the napin promoter (Ellerstrom et al. 1996). When the cDNA (*FAE*) encoding fatty acid elongase from nasturtium (*Tropaeolum majus*) was placed under the control of the napin promoter, exclusive expression of the gene in Arabidopsis seeds resulted in an 8-fold increase in erucic acid in the seed oil, which was much higher compared with using the CaMV 35S constitutive promoter (Mietkiewska et al. 2004). The napin promoter contains two complexes of *cis*-regulatory elements: a B-box and an RY/G complex, of which the roles in regulating gene expression have been elucidated (Ezcurra et al. 1999). Elements in the B-box constitute an abscisic acid (ABA)-responsive complex to mediate strong activity in seeds. The RY/G complex is located downstream of the B-box, composing of a G-box and its repeated flanking RY

elements. Both RY repeats and G-box have been shown to mediate transactivation through the transcriptional activator ABI3 regulated by ABA in the *NAPIN A* gene (Parcy et al. 1994).

Another seed-specific promoter is phaseolin, which drives the expression of the gene encoding β-phaseolin storage protein in green bean (*Phaseolus vulgaris*). Under the control of phaseolin promoter, the expression of the gene encoding β-phaseolin is completely silent during all phases of vegetative development, but is highly expressed in the seed only during embryogenesis (Bustos et al. 1989). Its activation requires the involvement of a number of transcription factors (TFs) that bind to over 20 *cis*-elements within the proximal 295 base pairs of the phaseolin promoter (Chandrasekharan et al. 2003). There are four RY elements each of which contributes to activating gene expression in different compartmentations of the Arabidopsis embryo. The G-box is the major ABRE of the promoter and has a positive effect on the promoter activity. The E-box functions as a mediator to complement the G-box. Other elements such as the CCAAAT box and the CACA element are also identified as positive elements. This set of regulatory systems has made phaseolin an indispensable candidate in the over-expression of transgenes. The phaseolin promoter was effectively used in a high-throughput screening process to over-express the cDNA encoding castor fatty acid hydroxylase in Arabidopsis to identify genes associated with the production of hydroxy FAs in the entire transcriptome of developing castor seeds (Lu et al. 2006).

Flax or linseed (*Linum usitatissimum*) is another commercially important oilseed crop. Known for its high α-linolenic acid content with health benefits, it offers a deep reservoir of seed-specific promoters of genes responsible for the production of storage components. It has two important seed storage protein fractions, namely linin as the major one corresponding to 12S fraction and conlinin corresponding to the 2S fraction with lower molecular weight (Truksa et al.

2003). The promoters of the genes encoding linin and conlinin have been isolated and protected by patents (Chaudhary et al. 2010; Qiu et al. 2008). Analysis of transgenic flax, *B. napus* and Arabidopsis containing the *β-GLUCURONIDASE* (*GUS*) reporter gene under the control of linin promoter showed strong expression in mature embryos prior to dessication. This late-stage seed specificity is ascribed to *cis*-regulatory elements including the ABRE and RY elements with a G-box in the centre (Chaudhary et al. 2010). Additionally, there is a special element called soybean embryo factor 3 (SEF3) element with a consensus sequence of AACCCA, which was initially identified in the upstream sequence of the gene encoding the α' subunit of β-conglycinin from immature soybean seeds. Early studies suggested its role as a primary recognition site for SEF3 TF to enhance transcription specifically in developing soybean seeds (Allen et al. 1989; Lessard et al. 1991). However, later studies produced different results indicating that the mutated SEF3 element had little effect on the activity of the promoter (Fujiwara and Beachy 1994). Two conlinin cDNAs have been identified as *cnl1* and *cnl2*. Their promoters have been isolated and found to be active in the flax developing seeds, driving maximum expression of the genes encoding conlinin at 20–35 DAP. A set of *cis*-regulatory elements similar to those found in the linin promoter have also been identified in both conlinin promoters (Qiu et al. 2008; Truksa et al. 2003). Interestingly, there is a modified G-box (TACGTG) located upstream from the original G-box and also a modified RY (CATGAA) element overlapping a complementarily reversed RY element farther upstream (Truksa et al. 2003). The linin and conlinin promoters have been widely utilized to optimize the production of omega-3 long chain-PUFAs such as eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) in transgenic oilseeds. Targeted expression of four genes respectively encoding three subunits of polyketide synthase and phosphopentetheinyl transferase under the control of the linin promoter has enabled the

engineering of Arabidopsis to generate 0.8% DHA with an additional 1.7% docosapentaenoic acid (DPA) in the seed oil (Weaver et al. 2011). The conlinin 1 promoter was also used with success in the metabolic engineering of Arabidopsis (Ruiz-Lopez et al. 2013), *B. carinata* (Cheng et al. 2009) and *Camelina sativa* (Ruiz-Lopez et al. 2014; Ruiz-Lopez et al. 2015) to produce omega-3 long chain-PUFAs.

The above four promoters of seed storage proteins from various oilseed crops were characterized almost three decades ago. However, the relative strength and application of these promoters in transgenic production of punicic acid has not been explored. Alteration of promoter action, which can influence the timing and intensity of gene expression, represents another potential way in which plant lipid biotechnologists can improve the effectiveness of heterologously-expressed genes (Mietkiewska et al. 2014a). This thesis project evaluated the four promoters described above (only conlinin 1 promoter was used in the case of conlinin, designated as conlinin promoter in the following) for their ability to drive the expression of the *PgFADX* gene during seed development in Arabdidopsis. The native promoter of *PgFADX* was also identified and characterized. This comparative investigation on promoter action led to the identification of the best promoter to facilitate the highest levels of punicic acid accumulation in the model plant Arabidopsis.

3. Materials and methods

3.1 Plant material

Pomegranate fruits were purchased at the local market. The seeds in the fruit were sown in the soil and grown to the pomegranate plants in the greenhouse at the University of Alberta. The growth environment was maintained at 20 °C under 16 h/8 h of natural light/dark period. For RNA extraction, fresh tissue samples were collected from the plants including leaves, stems and roots. Fresh pomegranate seeds for RNA extraction were obtained from the purchased fruit by peeling the arid covering the seed. The freshly collected tissue samples were quickly frozen in liquid nitrogen and stored at -80 °C.

Arabidopsis was grown to in the growth chamber at the University of Alberta. The growth conditions are as follows: 22 ºC, a photoperiod of 18 h light and 250 µmol $m^{-2}s^{-1}$ light intensity (Mietkiewska et al. 2014b). Fresh tissues were collected for RNA extraction by being quickly frozen in liquid nitrogen and stored at -80 °C.

3.2 Comparison of heterologous promoters of seed storage protein genes

3.2.1 Expression of *PgFADX* controlled by different heterologous promoters in Arabidopsis

The *PgFADX* open reading frame (ORF) (GenBank# AY178446) was amplified by polymerase chain reaction (PCR) using as a template the NCJ plasmid developed earlier by (Mietkiewska et al. 2014b) and primers: F1 and R1 (sequences of primers are shown in Appendix). The napin promoter was amplified with primers F2 and R2, while the termination fragment of *NOPALINE SYNTHASE* (*NOS*) gene was amplified with primers F3 and R3. Subsequently, napin promoter, *PgFADX* ORF and NOS termination fragment digested with the

appropriate endonucleases were ligated into AgeI and NotI sites of pAUX3131 vector resulting in SL1 plasmid. The cassette carrying napin promoter/*PgFADX*/NOS was excised from SL1 plasmid with I-SceI restriction enzyme and cloned into the corresponding sites of pPZP-RCS1 (Goderis et al. 2002) containing *NPTII* gene as a selection marker, resulting in the binary vector SL2. Likewise, other constructs for plant transformation were prepared by replacing the napin promoter with other promoters in SL1 plasmid, then excising the promoter/*PgFADX*/NOS cassette with the I-SceI enzyme and cloning into the corresponding sites of pPZP-RCS1. Other promoters were amplified with the following primers: phaseolin promoter: F4 and R4; linin promoter: F5 and R5; conlinin promoter: F6 and R6.

The four newly constructed binary plasmids plus the empty vector were introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation (Koncz and Schell 1986), and then transformed into Arabidopsis *fad3/fae1* mutant background (Smith et al. 2003) using the floral dip method as previously described (Clough and Bent 1998). T_1 seeds harvested from the transformed plants were screened on 1/2 MS selection media (Murashige and Skoog 1962) containing kanamycin (50 mg/L) as the selection antibiotic. The kanamycin-resistant seedlings were transferred to soil and grown to maturity in a growth chamber under the conditions described in 3.1. PCR analysis was performed to confirm the presence of the transgene in the genome of the growing T_1 lines with different constructed binary vectors, as well as the absence of the transgene in the genome of T_1 lines with empty vector. Leaf extract for PCR analysis was prepared using method as described by (Edwards et al. 1991). The T_2 segregating seeds harvested from each T_1 lines underwent segregation analysis to determine the number of transgenic loci, based on the segregation ratio of resistant seedlings and sensitive seedlings. For stable comparison of the promoter strength, lines with a single transgenic locus (segregation

ratio= 3:1) were selected and grown to obtain T_3 seeds. Segregation analysis was performed on T_3 seeds to categorize their parental lines as homozygous lines, heterozygous lines and null segregants.

3.2.2 Lipid analysis

FA composition of mature Arabidopsis seeds was analyzed as described earlier (Mietkiewska et al. 2014b). Briefly, two biological replicates of clean Arabidopsis seeds were sampled from each individual line. Total lipid extract was prepared by homogenizing 4-5 mg of seeds in a 1 mL mixture of chloroform and isopropanol (2:1, v/v) using a mortar and pestle. The homogenate was then dried under N_2 and transmethylated with 1 mL of 5% sodium methoxide ($Na\text{-}OCH_3$) in methanol solution at room temperature for 30 min. The reaction was stopped by adding 1 mL of 0.9% NaCl solution. The fatty acid methyl esters (FAMEs) were extracted twice with hexane. After being dried under N_2, FAMEs were resuspended in 1 mL of iso-octane containing 0.1 mg/mL methyl heneicosanoin (C21:0) as the internal standard.

The prepared FAMEs were analyzed by gas chromatography (GC)-mass spectrometry (MS) using an Agilent 7890A Gas Chromatograph (Agilent Technologies, Wilmington, DE, USA) interfaced with a 5977A Mass Selective Detector (Agilent Technologies, Wilmington, DE, USA). Samples were separated using a DB-23 capillary column (30 m×0.25 mm i.d., 0.25mm in film thickness, J&W Scientific/Agilent Technologies, Folsom, CA, USA) with a helium (He) flow at a constant rate of 1.2 mL/min. The oven temperature was programmed from 90 ^{o}C, at 10 ^{o}C /min, to 180 ^{o}C, held for 5 min, and increased at 5 ^{o}C /min to 230 ^{o}C and then held for 5 min. For identification of punicic acid, the mass spectra of FAMEs prepared from transgenic Arabidopsis

seeds were compared with those of FAMEs rich in punicic acid generated from pomegranate seed oil (Lot# B6132, Mountain Rose Herbs online store, Eugene, OR, USA).

3.2.3 Expression analysis of *PgFADX* in developing seeds of transgenic Arabidopsis

Total ribonucleic acid (RNA) was isolated from T_3 transgenic Arabidopsis seed lines and the wild type at the early stage of maturation phase (14 days after pollination when the accumulation of PUFAs accelerates) using the Spectrum Plant Total RNA kit (Sigma-Aldrich). First strand cDNA was synthesized using 400 ng of total RNA as template with the action of Superscript III reverse transcriptase (Invitrogen). Quantitative real-time PCR (qRT-PCR) was conducted in a 10-μL volume reaction containing 2 μL of 1/50 dilution of cDNA as template, 5 μL of 2× SYBR-Green Master Mix (Molecular Biology Service Unit, University of Alberta, Canada), 0.45 μL of forward primer (10 μM; final concentration 450 nM) and 0.45 μL of reverse primer (10 μM; final concentration 450 nM). Gene-specific primer pairs PF-683 and PR-768 were designed for *PgFADX* ORF. An Arabidopsis 18S rRNA gene (GenBank#NC_001284) was used as the internal standard with gene-specific primers 18S-F and 18S-R. The reaction was performed in a 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA) under the following thermal conditions: 40 cycles of 95°C for 15 s and 60°C for 1 min. The specificity of the amplification was verified by melting curve analysis. Triplicate C_t values were obtained and averaged for each sample. The relative expression level was calculated using the comparative C_t method after normalizing to the level of the internal standard (Livak and Schmittgen 2001).

3.2.4 Low resolution-nucleic magnetic resonance analysis of seed oil content

The seed oil content was evaluated by low resolution-nucleic magnetic resonance (LR-NMR) spectroscopy using a modification (Shah and Weselake, unpublished) of the method described by Samii-Saket et al. (Samii-Saket et al. 2011). The harvested Arabidopsis seeds were dried at room temperature in the desiccator for seven days to minimize the moisture. About 50 mg of intact seeds were weighed on an analytical balance and transferred into a 180 x 10 mm LR-NMR tube which was in turn transferred into a 180 x 17.7 mm LR-NMR tube containing 10.5 g of 4 mm diameter glass beads (Fisher Scientific, Ottawa, Canada). Two biological samples were prepared from each individual line in two different tubes. The wider LR-NMR tube containing the sample tube was loaded into a Bruker Minispec mq20 NMR analyzer (Bruker, Karlsruhe, Germany). The NMR signal was collected at 97dB, 100 scans, 1.50 s recycle delay, narrow band width and 40.0 °C. Start spectral width was set at 0.050 ms and stop spectral width was 0.060 ms. The signal area was proportional to the oil content, and the proportionality constant was estimated from a standard with known oil content. Therefore, the oil content was displayed as percentage of dry seed weight through calculation from the input seed weight. Each sample was measured twice and the average oil content from two biological samples was reported.

3.3 Isolation and characterization of native promoter of *PgFADX* gene from pomegranate

3.3.1 Expression pattern of *PgFADX* in different tissues of pomegranate

Total RNA from leaves and stems were extracted using RNeasy kit (Qiagen, Mississauga, ON, Canada) while Spectrum Plant Total RNA kit (Sigma-Aldrich, Oakville, ON, Canada) was used for seed and roots samples. The crude RNA preps were further purified and concentrated

with RNeasy MinElute™ Cleanup Kit (Qiagen, Mississauga, ON, Canada). First strand cDNA was synthesized using QuantiTect reverse transcription kit (Qiagen, Mississauga, ON, Canada) according to the manufacturer's instructions. Relative expression levels of *PgFADX* in different tissues were determined by qRT-PCR analysis under the same conditions as described in 3.2.3 and normalized to the level of *β-ACTIN* gene (GenBank# GU376750.1). The primers for the 3′ untranslated regions (3′ UTR) of *PgFADX* were F7 and R7. The primers for *β-ACTIN* were F8 and R8.

3.3.2 Isolation of 5' upstream DNA sequence of *PgFADX* from pomegranate genome

For a successful isolation of 5′ upstream sequence of the *PgFADX* gene, the genomic sequence of *PgFADX* with the exons and introns located should be obtained for accurate design of gene specific primers. Genomic DNA was extracted from fresh pomegranate leaves using the urea-phenol extraction method (Chen et al. 1992). The genomic *PgFADX* was amplified from the genomic DNA with primers F1 and R1. The genomic sequence of *PgFADX* was submitted to GenBank with the accession number # KU921416.

The 5′ upstream sequence of *PgFADX* gene was isolated from pomegranate genome using the Universal GenomeWalker 2.0 kit (Clontech, Mountain View, CA, USA). Genomic DNA (2.5 μg) in separate aliquots was digested with EcoRV that left blunt ends, and then ligated to the GenomeWalker Adaptors to construct the library. To amplify the specific promoter fragment, primary PCR was performed using the outer adaptor primer (AP1) and an outer gene-specific primer (GSP1), followed by a secondary "nested" PCR with the nested adaptor primer (AP2) and a nested gene-specific primer (GSP2). The amplified secondary PCR product was cloned into TA

cloning vector pCR 2.1 (Invitrogen) and sequenced. The isolated full length promoter element is a 2027-bp fragment and its sequence was submitted to GeneBank with the accession number KU921416.

3.3.3 Preparation of transgenic Arabidopsis harboring *GUS* reporter gene controlled by *PgFADX* native promoter

The *PgFADX* native promoter (a 2027-bp upstream fragment) was amplified from the TA cloning vector pCR 2.1 with primers F9 and R9. The *GUS* reporter gene (GenBank # AAL92040.1) was amplified with primers F10 and R10. The cassette *PgFADX* promoter/*GUS* was ligated into restriction sites AgeI and BamHI of the auxiliary vector pAUX3131 (Smith et al. 2003) to result in plasmid SL14. The cassette carrying *PgFADX* promoter/*PgFADX*/NOS was cut out from SL14 plasmid with I-SceI restriction enzyme and cloned into the respective sites of binary vector pPZP-RCS1 to generate a new plant transformation vector SL15. Similar cloning strategies were used to construct the vector with its *GUS* gene controlled by the napin promoter as a seed-specific control. The *GUS* gene was amplified with primers F11 and R11 and inserted between restriction site XhoI and XbaI in the auxiliary vector SL1 to result in SL12. The cassette carrying napin promoter/*GUS*/NOS was excised by restriction enzyme I-SceI and cloned to the corresponding sites of binary vector pPZP-RCS1 to generate vector SL13. The pBI 121 (previously available from Clontech Laboratories, USA) vector was used as the constitutive control in which the *GUS* gene was fused with the CaMV 35S promoter. These three constructs were introduced into Arabidopsis *fad3/fae1* mutant via Agrobacterium-mediated transformation as described in 3.2.1. After screening the T_1 seeds on the 1/2 MS medium containing kanamycin, seedlings which survived were grown to T_1 plants as the subjects for subsequent GUS assay.

Various types and developing stages of tissues were collected including leaves, flowers, early-stage developing seeds (~7 DAP), middle-stage developing seeds (~12 DAP) and late-stage developing seeds (~17 DAP). The non-transformed plant background served as the negative control.

3.3.4 Histochemical GUS assay

GUS activity was assayed histochemically following the method of (Jefferson et al. 1987) using a staining solution containing: 0.5 mg/mL X-gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid), 10 mM EDTA, 100 mM NaH_2PO_4, 0.5 mM potassium ferrocyanide and 0.1% detergent Triton X-100.

Sections of freshly collected tissues of Arabidopsis plant were incubated in the X-gluc staining solution at 37 °C in the dark for 16-24 h. It is important to note that the developing seeds were longitudinally sectioned with sterile razor such that the interiors of the seeds were exposed to the staining solution. The staining solution was replaced with 95% ethanol to remove chlorophyll by overnight incubation. The treated tissues were stored in 70% ethanol solution for histochemical assay. The specimens were examined under the Olympus Stereo Light Microscope with OPLENIC Digital Eyepiece Camera (The Advance Microscopy Facility, University of Alberta, Canada) to observe and localize the blue staining where GUS was active.

3.4 Statistical analysis

One-way analysis of variance (ANOVA) was conducted to determine the significance level of difference in the mean fatty acid proportion and mean relative gene expression level among

different constructs. Null hypothesis was rejected when the p value was < 0.05. If there was an overall significant difference among different constructs, Tukey's honestly significant difference (HSD) post-hoc test was run afterwards to confirm where the differences occurred between any pair of constructs.

4. Results

4.1 Genetic constructs

To compare the efficiency of different promoters in driving the expression of *PgFADX*, four plant transformation vectors were constructed as illustrated in Figure 5. An empty vector without the constructed cassette was included as the control. Each test vector contained *PgFADX* cDNA encoding the pomegranate FA conjugase, which catalyzes the biosynthesis of punicic acid. These vectors differed with respect to the seed-specific promoter used to drive the expression of *PgFADX*: napin promoter in SL2, phaseolin promoter in SL4, linin promoter in SL6 and conlinin promoter in SL7. All four promoters are originally from genes encoding seed storage proteins of different oilseed crop species, as described in the literature review of this thesis. The cassette containing the *NPTII* gene near the left border (LB) of the T-DNA functioned as a selection marker with resistance to the antibiotic kanamycin. Agrobacterium-mediated transformation of the five vectors was performed to introduce the T-DNA into the genome of Arabidopsis mutant *fad3/fae1* which contained a high level of 18:2 (up to 51% of total FA in the seed oil) as the substrate of PgFADX. T_1 seeds of the putative transformed lines were screened on kanamycin-containing medium and lines exhibiting resistance were transplanted to soil. Upon the confirmation by PCR for a) the presence of the transgene in the genome of T_1 plants transformed with different constructs and, b) the absence of the transgene in the genome of T_1 plants transformed with the empty vector as the control, those transformants were grown to maturity to produce mature T_2 segregating seeds for oil analysis.

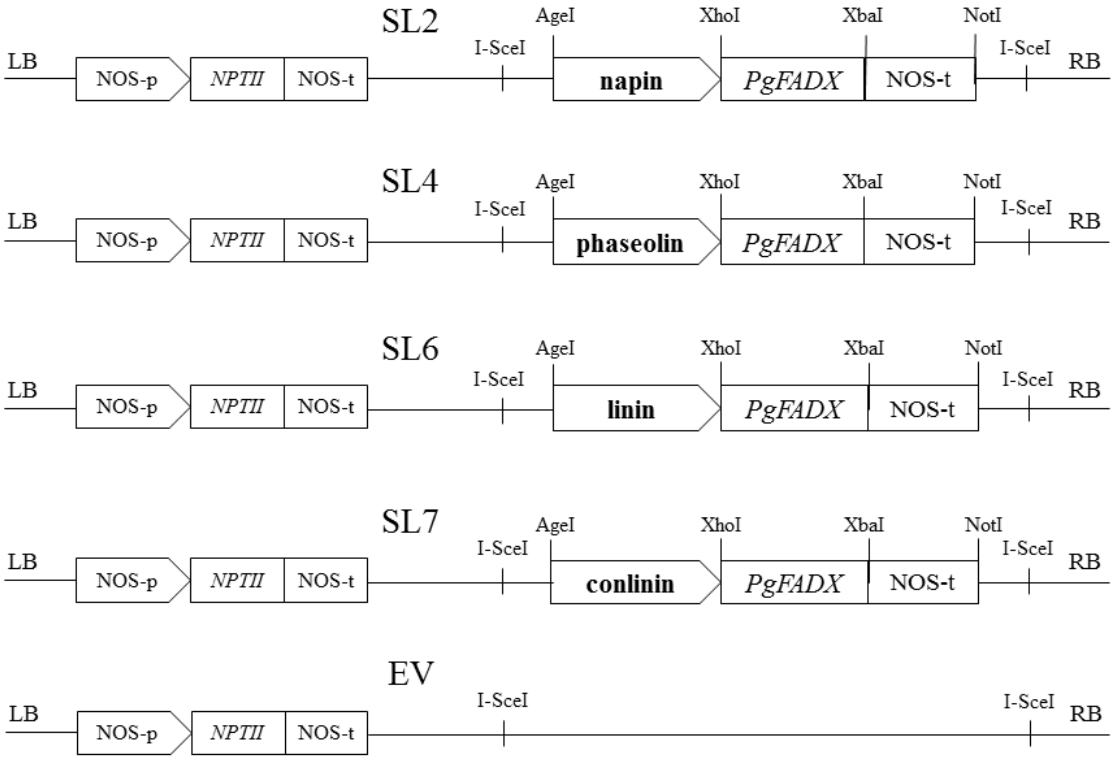


Figure 5. T-DNA structure of five binary vectors used for transformation of Arabidopsis plants. The left cassette contains *NPTII* gene which confers resistance to kanamycin for plant screening. Except for the empty vector (EV) shown at the bottom, the right constructed cassette comprises *PgFADX* gene under the control of different heterologous promoters as highlighted in bold, which is the only difference among the four constructs named SL2, SL4, SL6 and SL7, respectively. LB: left border; RB: right border; NOS-p: promoter sequence of the *Agrobacterium tumefaciens NOPALINE SYNTHASE* (*NOS*) gene; NOS-t: termination sequence of the *NOS* gene; *NPTII*: *NEOMYCIN PHOSPHORTRANSFERASE* II gene.

4.2 Comparison of promoter strength in driving the expression of *PgFADX*

The relative strengths of four promoters for seed storage protein genes were compared on the basis of punicic acid production and relative expression of *PgFADX* gene driven by the studied promoters in the seeds of engineered Arabidopsis. From the population of ~50 independent transgenic events per construct, representative lines of segregating T_2 seeds producing substantial levels of punicic acid were selected for further analysis. The proportions of punicic, oleic and linoleic acid in the seed oil of 18 transgenic lines, along with the lines carrying empty vector (EV) as controls are depicted in Figure 6. Arabidopsis seeds transformed with cDNA encoding PgFADX accumulated varied levels of punicic acid in contrast to EV controls, which were devoid of punicic acid. The highest level of punicic acid was 19.50% of total FA, found in the line SL6-16 carrying *PgFADX* driven by the linin promoter. The punicic acid levels of the construct SL2 with *PgFADX* driven by the napin promoter ranged from 7.74% to 14.9%. Lower amounts of punicic acid were produced in lines under the control of the phaseolin promoter (7.19% - 11.64%) and the conlinin promoter (4.32% - 9.84%), respectively. Changes in the proportions of the other two related FAs were consistent with those reported in the literature, indicating that Arabidopsis *FAD2* expression might be suppressed (Cahoon et al. 2006; Mietkiewska et al. 2014b). Observed changes in the proportions of other FAs in the seed oil of transgenic plants are presented in section 4.3.

The conversion of linoleic acid into punicic acid and other CLNAs is a reflection of the enzymatic activity of PgFADX. A linear, positive correlation was observed when the proportion of linoleic acid converted to CLNAs was plotted as a function of total CLNA content (Figure 7). Thus, CLNAs were synthesized at the expense of linoleic acid with progressively higher amounts

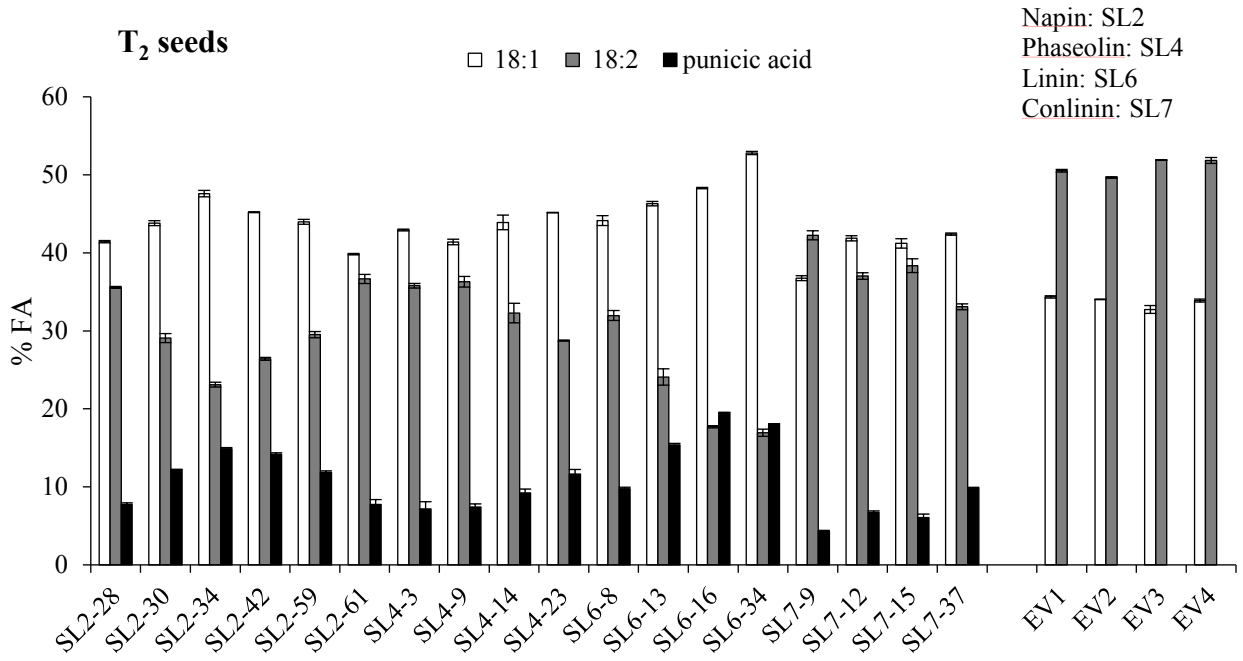


Figure 6. Oleic (18:1), linoleic (18:2) and punicic acid contents of T_2 segregating seeds of transgenic Arabidopsis lines overexpressing *PgFADX* under the control of four heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7), and four control lines carrying empty vector (EV). Values are represented as the mean of two biological replicates from each line. Error bars indicate the range.

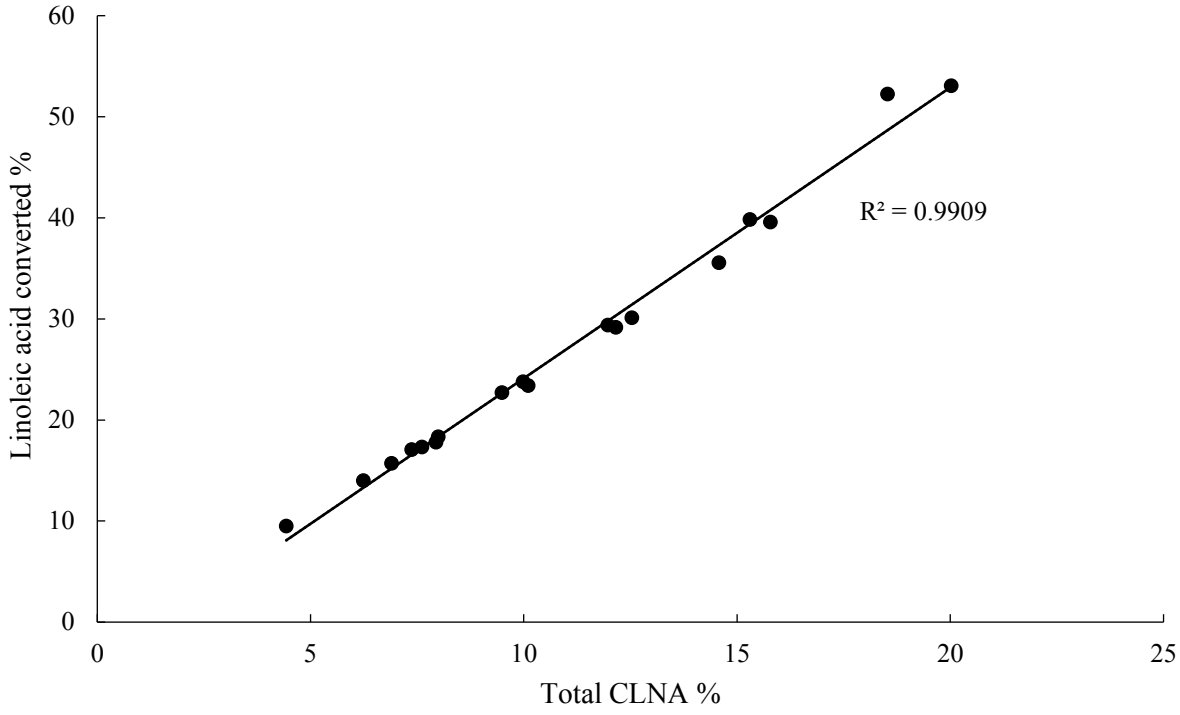


Figure 7. Relationship between the proportion of linoleic acid converted to conjugated linolenic acids (CLNAs) and the total CLNAs content of oil from the mature seed. This plot is based on the data from Figure 6. The proportion of linoleic acid converted to CLNAs was based on the ratio of the sum of the contents of punicic acid and two unidentified CLNAs to the linoleic acid content.

of conversion corresponding to higher proportions of CLNAs. The line SL6-16 producing ~20.0% CLNAs exhibited about 53.0% conversion while the line SL7-9 producing lowest level of CLNAs (4.43%) exhibited 9.50% conversion.

For a better comparison of the promoter strength, it is necessary to have stable performance among individual lines within each construct. Therefore, segregation analysis was performed on T_2 seeds of each line to determine the number of transgenic loci in each transgenic event which was reflected by the segregation ratio of seedlings resistant to those sensitive to the antibiotic selection culture. Lines with a segregation ratio of 3:1, which suggested a transgenic event of one insertion locus, were grown to obtain T_3 seeds for the comparison of promoter strength. Meanwhile, some lines with T_2 seeds with the highest accumulation of punicic acid were grown to next generation to evaluate the inheritability of their enhanced performance.

T_3 seeds were produced from parental lines that were confirmed by segregation analysis to harbor one insertion locus; the levels of punicic acid in the seed oil are shown in Figure 8. The lines were categorized into three groups based on the number of allelic loci of the transformed cassette by a second segregation analysis on T_3 seeds of each individual line. In general, the homozygous group accumulated punicic acid at a level approximately twice that of the heterozygous group; however, punicic acid was not found in the null segregants. Within the homozygous group, one line with *PgFADX* driven by the linin promoter produced the highest level of punicic acid (SL6-8-17, 13.27±0.30% of total FA). The level of punicic acid (13.02±0.12%) produced by SL6-8-7 was the lowest among the SL6 lines driven by the linin promoter but still higher than for any of the lines using napin, phaseolin or conlinin promoters. An analogous situation was found in the heterozygous group, where the highest level of punicic acid was produced by the linin promoter lines: SL6-8-2 (8.05±0.22%) and SL6-8-3

(8.07±0.21%). For a statistically valid comparison of the punicic acid levels among four constructs, one-way analysis of variance (ANOVA) was performed on both homozygous and heterozygous groups. The ANOVA of the homozygous group are shown in Table 1. There were significant differences in the punicic acid level among four constructs, F (3, 24) = 167.22, p< 0.01. Thereafter, Tukey's honestly significant difference (HSD) post-hoc test was applied to investigate possible significant differences in punicic acid levels in pairwise comparisons. As shown in Table 2, the SL6 construct driven by the linin promoter resulted in significantly higher punicic acid content than for other constructs driven by other promoters among which SL4 (phaseolin) and SL2 (napin) exhibited no significant difference but both were significantly lower than SL7 (conlinin). As for the heterozygous group shown in Table 3, significant differences were also present among the four constructs, F (3, 10)= 38.10, p< 0.01. While the linin promoter resulted in the punicic acid content which was significantly the highest, there was no significant difference in the punicic acid content among the other three promoters (Table 4). On combining the statistical analysis outcome of both homozygous and heterozygous T_3 seeds, it became evident that the punicic acid level driven by the linin promoter was significantly higher than those driven by other promoters.

To further confirm the promoter strength, relative expression levels of *PgFADX* in the developing T_3 homozygous seeds (13-16 DAP) of representative lines with a single insertion locus were evaluated by qRT-PCR analysis. Relative expression levels of *PgFADX* are compared to the punicic acid contents in the oil from mature seeds as shown in Figure 9. No expression of *PgFADX* was detected in the non-transformed control, ensuring the primer specificity for the *PgFADX* cDNA which has more than 65% sequence homology with the coding sequence of

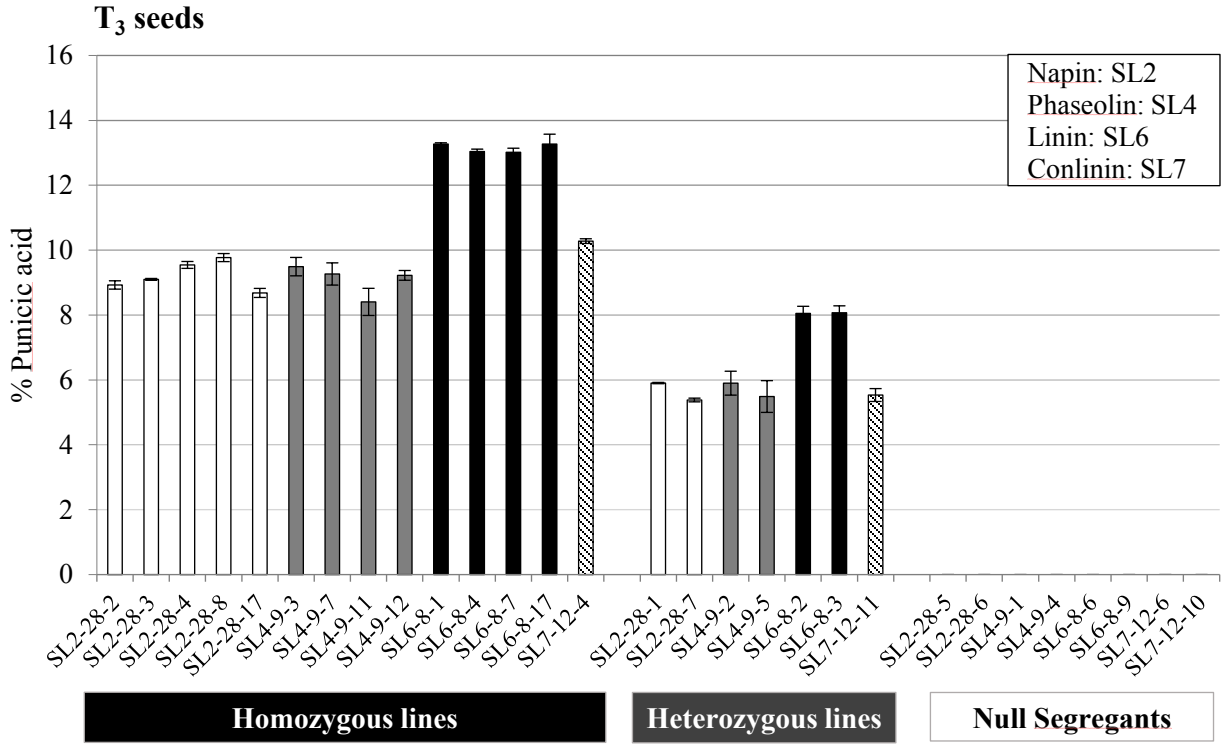


Figure 8. Punicic acid contents of T_3 Arabidopsis seeds of which the parental lines had a single insertion locus. A second segregation analysis on these T_3 seeds of each individual line categorized these lines into 3 groups: homozygous, heterozygous and null segregants. Values are represented as the mean of two biological replicates from each line. Error bars indicate the range.

Table 1. One-way ANOVA of punicic acid levels of T_3 homozygous seeds from four constructs

	Sum of squares	Degrees of freedom	Mean square	F_s	p value
Among groups	88.38	3	29.46	167.22	3.24E-16
Within groups	4.23	24	0.18		
Total	92.61	27			

Table 2. Tukey's HSD test on the significance of means of punicic acid level from T_3 homozygous seeds between different pairs of constructs.

Promoter	Mean of punicic acid content (%)	Tukey's HSD
Napin	9.20	a
Phaseolin	9.10	a
Linin	13.15	b
Conlinin	10.28	c

Means with the same letter are not significantly different from each other, $p > 0.05$; those with different letters are significantly different from each other, $p < 0.05$.

Table 3. One-way ANOVA of punicic acid levels of T_3 heterozygous seeds from four constructs

	Sum of squares	Degrees of freedom	Mean square	F_s	*p* value
Among groups	16.75	3	5.58	38.10	8.81E-06
Within groups	1.47	10	0.15		
Total	18.22	13			

Table 4. Tukey's HSD test on the significance of means of punicic acid level from T_3 heterozygous seeds between different pairs of constructs.

Promoter	Mean of punicic acid content (%)	Tukey's HSD
Napin	5.64	a
Phaseolin	5.69	a
Linin	8.06	b
Conlinin	5.53	a

Means with the same letter are not significantly different from each other, $p > 0.05$; those with different letters are significantly different from each other, $p < 0.05$.

Arabidopsis *FAD2* (*AtFAD2*) (Mietkiewska et al. 2014b). One-way ANOVA followed by the Tukey's HSD test suggested significant difference in the relative expression level among the four promoters, F (3, 12) = 601.10, $p< 0.01$. The linin promoter resulted in the highest expression of *PgFADX* and this was also associated with the highest level of punicic acid in the mature seed oil. In contrast, use of the other three promoters resulted in lower expression of *PgFADX* than with the linin promoter. Overall, among the four selected lines representing each promoter, there was a positive correlation between the relative expression level of *PgFADX* and the punicic acid content, except in the case of conlinin promoter. The line based on the use of conlinin promoter accumulated 10.28% punicic acid in the oil of the mature seeds but exhibited low expression of *PgFADX*. Collectively, the expression levels of *PgFADX* in developing seeds and punicic acid content of the mature seed oil demonstrated that the linin promoter was most effective among the four promoters evaluated in driving the expression of *PgFADX* to produce punicic acid in Arabidopsis seed.

4.3 Fatty acid composition and punicic acid accumulation in advanced generation of Arabidopsis seeds

The over-expression of *PgFADX* driven by different promoters led to varied levels of accumulation of punicic acid in the Arabidopsis seeds. Stable expression of the transgene in advanced generations is important for a good inheritability for producing consistent levels of punicic acid. Thus, the FA composition as well as the punicic acid level between two generations of seeds with the *PgFADX* driven by different promoters was investigated. The comparison focused on top-performing lines producing the highest level of punicic acid. Total FA

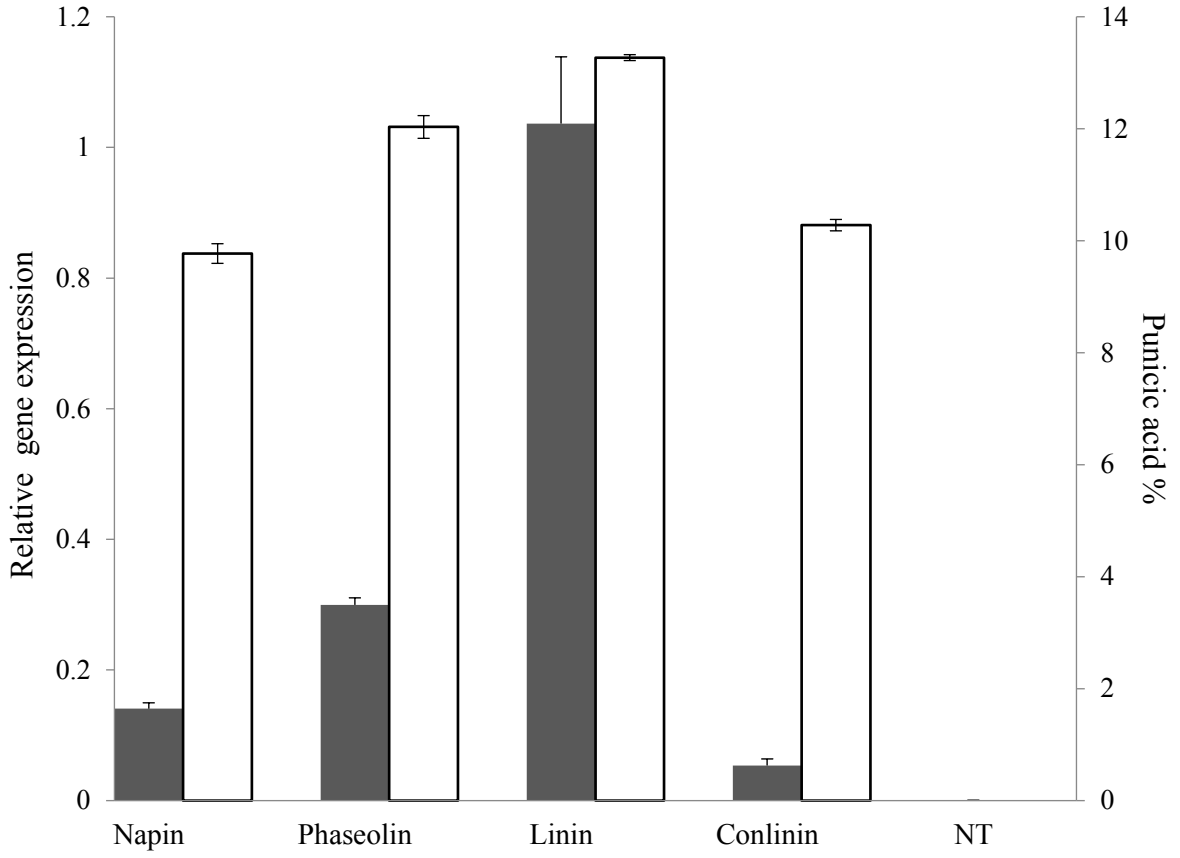


Figure 9. Relative expression of *PgFADX* driven by four different promoters in developing T_3 homozygous Arabidopsis seeds (grey bars) compared to the punicic acid content of oil from the mature T_3 seeds (white bars). The representative lines of which the parental lines carried a single insertion locus are as follows: SL2-28-8 for napin promoter, SL4-14-14 for phaseolin promoter, SL6-8-1 for linin promoter, SL7-12-4 for conlinin promoter. NT: non-transformed line.

compositions of the T_2 seed oil of lines from each construct producing the highest level of punicic acid are depicted in Table 5, along with four control lines carrying the empty vector (EV1-4). Marked differences in FA composition were found by comparing these top performers with the control lines. While all the control lines were devoid of punicic acid, they slightly differed in several FA levels: C18:1 ranges from 32.74% to 34.35%, with EV1 as the highest; C18:2 ranging from 49.66% to 51.89%, with EV3 as the highest. Elevations in punicic acid level were associated with a decrease in the level of palmitic acid (16:0) and increases in the levels of arachidic acid (20:0), behenic acid (22:0) and lignoceric acid (24:0). Changes in unsaturated FAs were also observed. The levels of linoleic acid ($18:2\Delta^{9cis,12cis}$) and α-linolenic acid ($18:3\Delta^{9cis,12cis,15cis}$) decreased considerably as a result of the accumulation of punicic acid. In contrast to the reduction in α-linolenic acid, over-expression of *PgFADX* indirectly increased the levels of a variety of unsaturated FAs: two unidentified CLNAs, eicosenoic acid ($20:1\Delta^{11cis}$) and nervonic acid ($24:1\Delta^{15cis}$). T_3 seeds exhibited a similar profile of change in total FA composition. As shown in Table 6, the elevation in punicic acid was accompanied by a concurrent increase in 18:1, two unknown CLNAs, eicosenoic acid and some long-chain saturated FAs including 20:0, 22:0 and 24:0. In contrast, there was a decrease in 16:0, 18:2 and α-linolenic acid. Interestingly, by associating the T_2 seeds (Table 5) with the T_3 seeds (Table 6), it was seen that a slight increase in the content of eicosenoic acid was accompanied by somewhat less of an increase in nervonic acid. More lines of T_2 seeds and homozygous T_3 seeds displayed a consistent trend of changes in these aforementioned FAs. Their total FA compositions are shown in Tables 7 and 8, respectively. The possible metabolic reasons for these FA changes are presented in the Discussion section.

Table 5. Total fatty acid (FA) compositions of the T_2 seed oils of lines from each construct producing the highest levels of punicic acid ($18{:}3\Delta^{9cis,11trans,13cis}$) in contrast to the controls carrying the empty vector (EV).

Construct	Fatty acid composition (wt %)														
	14:0	16:0	$16{:}1\Delta^{9}$	18:0	$18{:}1\Delta^{9cis}$	$18{:}2\Delta^{9cis,12cis}$	$18{:}3\Delta^{9cis,12cis,15cis}$	20:0	CLNA	$20{:}1\Delta^{11cis}$	$18{:}3\Delta^{9cis,11trans,13cis}$	22:0	CLNA	24:0	$24{:}1\Delta^{15cis}$
SL2-34	0.04	7.25	0.23	3.82	47.59	23.11	0.44	0.94	0.31	0.40	14.89	0.64	0.10	0.13	0.10
EV1	0.05	8.98	0.27	3.54	34.35	50.53	0.94	0.70	0.05	0.28	ND	0.16	0.02	0.05	0.07
SL4-23	0.04	7.59	0.21	3.59	45.17	28.76	0.53	0.90	0.23	0.43	11.64	0.58	0.09	0.15	0.11
EV2	0.04	9.43	0.24	4.10	34.03	49.66	0.90	0.85	0.02	0.32	ND	0.22	0.02	0.08	0.08
SL6-16	0.04	7.18	0.22	3.71	48.30	17.70	0.39	0.94	0.39	0.44	19.50	0.81	0.13	0.14	0.10
EV3	0.05	9.29	0.27	3.60	32.74	51.89	0.87	0.69	0.03	0.27	ND	0.15	0.02	0.05	0.05
SL7-37	0.04	7.54	0.24	3.84	42.40	33.08	0.68	0.90	0.19	0.39	9.84	0.51	0.07	0.15	0.11
EV4	0.04	8.54	0.23	3.31	33.89	51.84	0.96	0.66	0.02	0.26	ND	0.14	0.01	0.05	0.04

The means of duplicate analyses are reported. CLNA, conjugated linolenic acid. ND, not detected. Constructs: *PgFADX* driven by heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7).

Table 6. Total fatty acid (FA) compositions of the T_3 seed oils of lines from each construct producing the highest levels of punicic acid ($18:3\Delta^{9cis,11trans,13cis}$) in contrast to the corresponding null segregants (nt).

Construct	Fatty acid composition (wt %)														
	14:0	16:0	$16:1\Delta^{9}$	18:0	$18:1\Delta^{9cis}$	$18:2\Delta^{9cis,12cis}$	$18:3\Delta^{9cis,12cis,15cis}$	20:0	CLNA	$20:1\Delta^{11cis}$	$18:3\Delta^{9cis,11trans,13cis}$	22:0	CLNA	24:0	$24:1\Delta^{15cis}$
SL2-34-8	0.02	7.06	0.17	4.27	52.87	15.28	0.27	0.99	0.15	0.57	15.98	1.60	0.59	0.11	0.08
SL2-nt	0.05	9.10	0.26	3.49	33.73	51.09	0.95	0.69	0.05	0.28	ND	0.16	0.02	0.06	0.05
SL4-14-16	0.04	7.50	0.19	3.47	50.35	22.30	0.31	0.78	0.14	0.48	12.36	1.36	0.59	0.09	0.05
SL4-nt	0.05	9.28	0.27	3.62	33.53	50.98	0.91	0.70	0.04	0.28	ND	0.17	0.04	0.05	0.05
SL6-16-4	0.03	6.58	0.18	3.74	52.00	14.05	0.26	0.85	0.16	0.55	18.94	1.82	0.70	0.09	0.05
SL6-nt	0.04	8.84	0.27	3.61	32.14	52.72	1.10	0.73	0.02	0.25	ND	0.16	0.01	0.05	0.06
SL7-12-4	0.03	6.62	0.19	3.34	47.76	28.50	0.36	0.77	0.10	0.45	10.28	1.11	0.38	0.08	0.04
SL7-nt	0.04	8.69	0.23	3.37	32.89	52.71	0.95	0.63	0.02	0.23	ND	0.13	0.01	0.05	0.04

The means of duplicate analyses are reported. CLNA, conjugated linolenic acid. ND, not detected. Constructs: *PgFADX* driven by heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7).

Table 7. Total fatty acid (FA) compositions of the T_2 seed oils of lines from each construct.

Construct	Fatty acid composition (wt %)														
	14:0	16:0	16:1Δ^{9}	18:0	18:1Δ^{9cis}	18:2$\Delta^{9cis,12cis}$	18:3$\Delta^{9cis,12cis,15cis}$	20:0	CLNA	20:1Δ^{11cis}	18:3$\Delta^{9cis,11trans,13cis}$	22:0	CLNA	24:0	24:1Δ^{15cis}
SL2-28	0.05	8.01	0.27	3.81	41.43	35.57	0.80	0.91	0.16	0.42	7.77	0.49	0.06	0.13	0.12
SL2-30	0.05	7.64	0.24	3.77	43.81	29.07	0.77	0.89	0.25	0.43	12.21	0.56	0.07	0.12	0.11
SL2-42	0.04	7.25	0.21	3.75	45.21	26.41	0.43	0.92	0.29	0.38	14.21	0.59	0.08	0.12	0.11
SL2-59	0.04	7.60	0.24	3.79	43.98	29.51	0.61	0.90	0.24	0.37	11.85	0.58	0.07	0.12	0.10
SL2-61	0.05	8.45	0.27	4.00	39.83	36.65	0.86	0.90	0.15	0.35	7.73	0.47	0.06	0.11	0.11
SL4-3	0.05	7.86	0.22	3.38	42.93	35.79	0.62	0.78	0.13	0.36	7.19	0.43	0.05	0.12	0.11
SL4-9	0.05	8.24	0.25	3.57	41.39	36.30	0.76	0.83	0.14	0.34	7.41	0.46	0.05	0.10	0.10
SL4-14	0.04	7.94	0.22	3.60	43.90	32.28	0.57	0.85	0.18	0.37	9.24	0.52	0.07	0.12	0.10
SL6-8	0.04	7.46	0.22	3.65	44.13	31.97	0.57	0.86	0.18	0.38	9.74	0.50	0.05	0.14	0.11
SL6-13	0.04	7.36	0.22	3.47	46.31	24.09	0.53	0.86	0.32	0.43	15.37	0.67	0.10	0.13	0.11
SL6-34	0.04	6.03	0.15	2.96	52.80	16.94	0.29	0.74	0.35	0.50	18.05	0.77	0.13	0.14	0.11
SL7-9	0.06	8.69	0.32	4.32	36.75	42.25	1.18	0.96	0.08	0.35	4.32	0.42	0.03	0.14	0.14
SL7-12	0.05	7.76	0.24	3.64	41.86	37.03	0.63	0.86	0.12	0.38	6.74	0.41	0.04	0.13	0.11
SL7-15	0.05	7.74	0.25	3.58	41.21	38.36	0.71	0.83	0.11	0.37	6.09	0.40	0.04	0.14	0.12

The means of duplicate analyses are reported. Punicic acid (18:3$\Delta^{9cis,11trans,13cis}$). CLNA, conjugated linolenic acid. Constructs:

PgFADX driven by heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7).

Table 8. Total fatty acid (FA) compositions of the T_3 seed oils of homozygous lines from each construct.

Construct	Fatty acid composition (wt %)														
	14:0	16:0	16:1Δ^{9}	18:0	18:1Δ^{9cis}	18:2$\Delta^{9cis,12cis}$	18:3$\Delta^{9cis,12cis,15cis}$	20:0	CLNA	20:1Δ^{11cis}	18:3$\Delta^{9cis,11trans,13cis}$	22:0	CLNA	24:0	24:1Δ^{15cis}
SL2-28-2	0.05	7.35	0.21	3.57	50.35	25.70	0.40	0.83	0.15	0.52	8.93	1.24	0.62	0.06	0.04
SL2-28-3	0.04	7.18	0.20	3.56	50.31	25.70	0.40	0.85	0.16	0.50	9.10	1.26	0.61	0.08	0.05
SL2-28-4	0.04	7.30	0.21	3.58	49.54	25.90	0.39	0.86	0.13	0.49	9.54	1.27	0.61	0.08	0.05
SL2-28-8	0.04	7.30	0.21	3.56	49.49	25.71	0.42	0.85	0.14	0.48	9.78	1.26	0.60	0.09	0.06
SL2-28-17	0.02	7.16	0.12	4.10	52.99	23.01	0.25	1.02	0.12	0.60	8.68	1.24	0.52	0.10	0.06
SL4-9-3	0.03	7.97	0.18	3.96	48.31	26.56	0.37	0.94	0.11	0.47	9.49	1.07	0.37	0.11	0.06
SL4-9-7	0.02	7.83	0.16	3.96	49.42	25.87	0.33	0.95	0.09	0.49	9.27	1.08	0.37	0.10	0.07
SL4-9-11	0.05	8.17	0.21	3.65	47.69	28.26	0.37	0.81	0.13	0.45	8.41	1.16	0.53	0.07	0.05
SL4-9-12	0.04	8.04	0.22	3.61	47.36	27.91	0.43	0.83	0.13	0.45	9.22	1.14	0.50	0.07	0.05
SL6-8-1	0.04	7.08	0.21	3.86	49.62	21.97	0.33	0.87	0.14	0.50	13.27	1.40	0.57	0.10	0.05
SL6-8-4	0.04	7.00	0.20	3.71	50.14	22.07	0.30	0.84	0.14	0.49	13.04	1.36	0.54	0.09	0.05
SL6-8-7	0.04	6.61	0.19	3.55	51.21	21.57	0.29	0.83	0.14	0.52	13.02	1.34	0.54	0.10	0.06
SL6-8-17	0.03	6.61	0.17	3.39	50.49	22.28	0.28	0.76	0.11	0.48	13.27	1.42	0.59	0.07	0.04
SL7-12-4	0.03	6.62	0.19	3.34	47.76	28.50	0.36	0.77	0.10	0.45	10.28	1.11	0.38	0.08	0.04

The means of duplicate analyses are reported. Punicic acid (18:3$\Delta^{9cis,11trans,13cis}$). CLNA, conjugated linolenic acid. Constructs:

PgFADX driven by heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7).

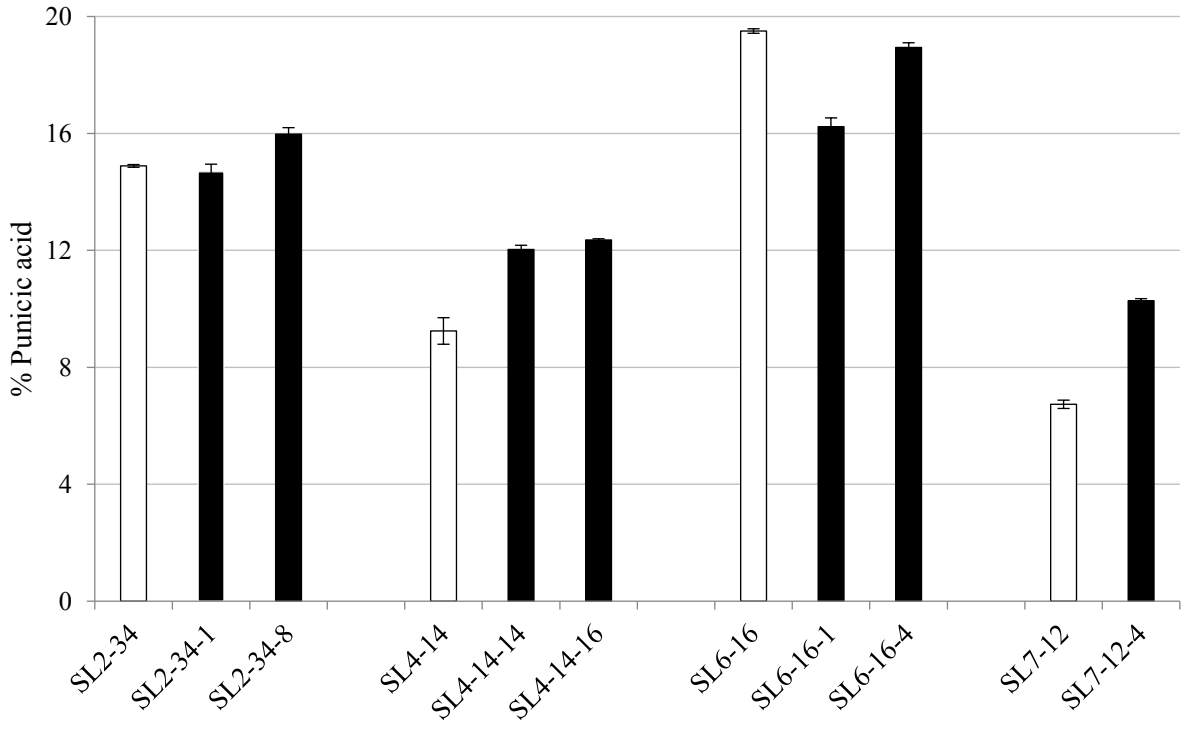


Figure 10. Comparison of punicic acid level in T_2 seeds (white bars) and T_3 seeds (dark bars) of the top-performing lines from each construct carrying *PgFADX* driven by heterologous promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7). Values are represented as the mean of two replicate samples from each line. Error bars indicate the range.

The stability of production of punicic acid over two generations of transgenic Arabidopsis was evaluated by making a comparison between T_2 and T_3 seeds. The changes in punicic acid levels from T_2 seeds to T_3 seeds of the top-performing lines of each construct are shown in Figure 10. SL2-34 driven by the napin promoter was the most stable in that T_3 seeds of SL2-34-1 exhibited a 1.3% relative decrease and those of SL2-34-8 showed a 7.3% relative increase in the punicic acid level. The linin promoter-driving SL6-16 resulted in 16.8% (SL6-16-1) and 2.9% (SL6-16-4) relative reduction in punicic acid content, respectively, in the advanced generation, although SL6 produced the highest level of punicic acid among other constructs. In contrast, the two constructs SL4 and SL7 with lower yields of punicic acid exhibited a substantial increase of punicic acid level in the advanced generation. SL4-14, controlled by the phaseolin promoter, exhibited a relative increase in the production of punicic acid by 32% on average and SL7-12, controlled by the conlinin promoter, also showed a 52.6% relative increase in punicic acid.

4.4 The top-performing line with high production of punicic acid exhibited a decrease in seed oil content

The seed oil content is another important factor that contributes to the total production of punicic acid. It is also hypothesized that different promoters may have different effects on the seed oil content. Therefore, the oil contents of T_3 homozygous seeds were analyzed non-destructively by LR-NMR. The null segregants in each transgenic event were used as controls for each construct. As shown in Figure 11, there was uniformity in the oil contents between the null segregant controls and the homozygous seeds of the lines carrying a single insertion locus. The oil contents of the null segregants from different transgenic events ranged from 37.39% (SL2-28-5) to 39.43% (SL7-12-6), while the oil contents of the T_3 homozygous seeds of lines

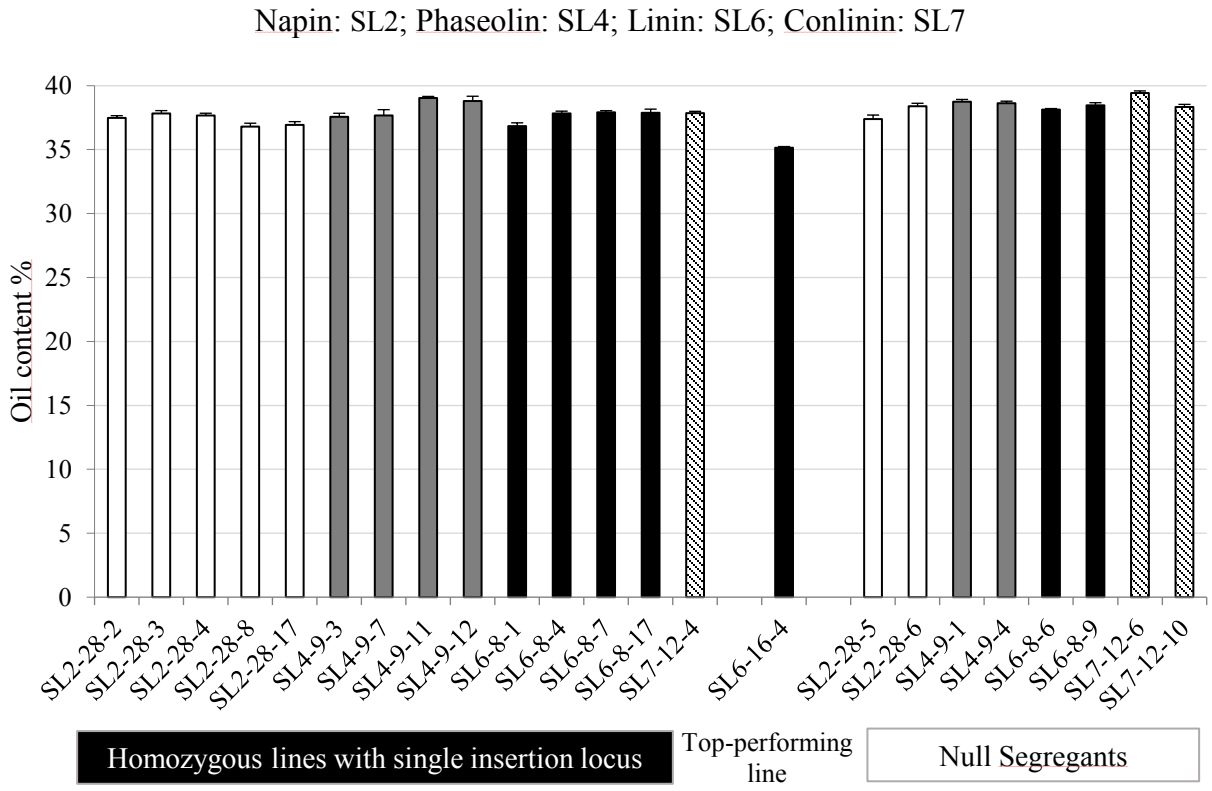


Figure 11. Oil contents of T_3 seeds. Homozygous seeds of lines with a single insertion locus were analyzed. The top-performing line SL6-16-4 which produced the highest level of punicic acid (18.94%) among the T_3 seeds was included in the analysis. Null segregants derived from each transgenic event were used as the controls. The homozygous lines carried the *PgFADX* gene driven by different promoters: napin (SL2), phaseolin (SL4), linin (SL6) and conlinin (SL7). Each column represents the mean oil content (% seed weight) of duplicate measurements of duplicate biological samples from the individual line. The standard deviation was displayed as the error bar on top of each column (n=4).

with a single insertion locus carrying different constructs were in the range between 36.80% (SL2-28-8) and 39.04% (SL4-9-11). Student's *t* test was performed to analyze the significance of difference between the T_3 homozygous group with a single transgenic locus and the control group of null segregants (t= 2.62, 20 degree of freedom; p= 0.016). Given the p value of 0.016, it suggests that the oil contents between the T_3 homozygous seeds of lines with a single insertion locus and the null segregants are indistinguishable at the significance level of 0.01 (p> 0.01); however, at the significance level of 0.05, the oil contents of T_3 homozygous seeds of lines with a single insertion locus are 1.87% lower than those of the controls (p< 0.05). This suggests in a general sense a single insertion locus of the transgenic constructs "heterologous promoter/*PgFADX*/NOS" has little influence on the seed oil production of Arabidopsis.

The effect of different heterologous promoters on the seed oil content was investigated by comparing different constructs within the T_3 homozygous group of lines with a single insertion locus. One-way ANOVA showed the oil contents of T_3 homozygous seeds were not significantly different among four constructs, F (3, 10)= 1.966, p> 0.05. Within each construct, the change in oil content was studied through comparison between transgenic lines and null segregants. No significant change was found among lines within each construct, as shown by the student's *t* tests: SL2: t= 1.27, 5 degree of freedom; p= 0.26; SL4: t= 0.74, 4 degree of freedom; p= 0.50; SL6: t= 1.69, 4 degree of freedom; p= 0.16; SL7: t= 1.86, 2 degree of freedom; p= 0.20.

Comparison of the top-performing line with the corresponding null segregants suggested that further increase in the accumulation of punicic acid had a considerable reduction in the seed oil content. The T_3 seeds of the top-performing line SL6-16-4 with its *PgFADX* driven by the linin promoter produced 18.94% punicic acid and 35.15% oil in the seed. The seed oil content was 3.15% lower than its controls containing 38.3% oil in average (t= 27.35, 6 degree of

freedom; $p < 0.01$). As segregation analysis showed that SL6-16-4 was a homozygous offspring of SL6-16 with two transgenic insertion loci, there were four transgenic loci in SL6-16-4. This indicated multiple insertions of the transgenic cassette "heterologous promoter/*PgFADX*/NOS" might affect the seed oil yield of transgenic Arabidopsis, as a trade-off of its high accumulation of punicic acid.

4.5 The *PgFADX* native promoter is not active in developing Arabidopsis seed

The activity of the *PgFADX* native promoter from pomegranate was also investigated. The relative expression of *PgFADX* in various tissues of pomegranate is shown in Figure 12. The results of qRT-PCR demonstrated that *PgFADX* is preferentially expressed in the pomegranate seed. The punicic acid content of the seed oil from this cultivar was about 80% of the total FAs. The strong seed-specific expression of *PgFADX* in pomegranate suggested that this plant has a strong and seed-specific native promoter which may be potentially useful in the metabolic engineering of Arabidopsis to drive transgene expression.

A 2027-bp fragment upstream of *PgFADX* was isolated from the pomegranate genome using genome walking. Alignment of the sequence of this fragment with the coding sequence of *PgFADX* (GenBank# AY178446.1) confirmed this isolated fragment flanks upstream of *PgFADX*. Two 5'UTRs compassing an intron were also found at the 3' end of the isolated fragment, followed by the putative promoter fragment located upstream. It was widely reported that 5'UTR, as well as the 5'UTR-flanking intron, plays an important role in enhancing gene expression, so it is useful to integrate not only the promoter region but also the whole 5'UTR up until the start codon to the upstream of the coding region of a reporter gene when studying the

promoter (Dan and Ow 2011). In the present study, the entire 2027-bp fragment containing the putative *PgFADX* promoter and 5'UTR was designated as the "*PgFADX* native promoter" in the following, unless otherwise specified. The ability to drive the expression of the reporter *GUS* gene in transgenic Arabidopsis was evaluated relative to two controls: the napin promoter and the CaMV 35S promoter. The histochemical assay revealed that three promoters directed *GUS* expression in various tissues and at different developmental stages (Figure 13). The *GUS* gene directed by the CaMV 35S promoter was expressed throughout the plant, including the leaf, the flower and seeds at three developmental stages (Figure 13a-e), which is in agreement with CaMV 35S being a constitutive promoter. In contrast, *GUS* expression driven by the napin promoter (Figure 13f–j) was restricted to developing seeds (Figure 13h-j), and its expression was particularly strong at the mid-developmental stage (Figure 13i). No visible expression was found in other tissues. These observations support the fact that the napin promoter regulates gene expression in a seed-specific manner which is temporally specific to the mid-developmental stage. Almost no GUS activity was detected in any tissues of Arabidopsis carrying a fusion of *GUS* to *PgFADX* native promoter (Figure 13k-o). There was only trace amount of staining in a few seeds at the late developmental stage (Figure 13o). In additional experiments, *PgFADX* cDNA was introduced into Arabidopsis under the control of the *PgFADX* native promoter. Punicic acid was not detected in the oil of T_2 seeds (data not shown). Taken together, these results indicate that the *PgFADX* native promoter is not active in Arabidopsis.

4.6 Bioinformatic analysis of *cis*-regulatory elements in four heterologous promoters and the *PgFADX* native promoter

The present study has been focused on a side-by-side comparison of the effectiveness of different promoters in driving the expression of *PgFADX* cDNA to achieve higher production of punicic acid in engineered Arabidopsis. It is also important to link these results with a characterization of the *cis*-regulatory elements embedded in these promoters. Bioinformatic analysis on the promoter sequences was performed to characterize *cis*-regulatory elements using PlantCARE software (Lescot et al. 2002). A series of *cis*-regulatory elements were revealed as illustrated in Figure 14. Details on the c*is*-regulatory elements associated with seed specificity and strength of the promoter that contains one or more of *cis*-regulatory elements are discussed in the next section.

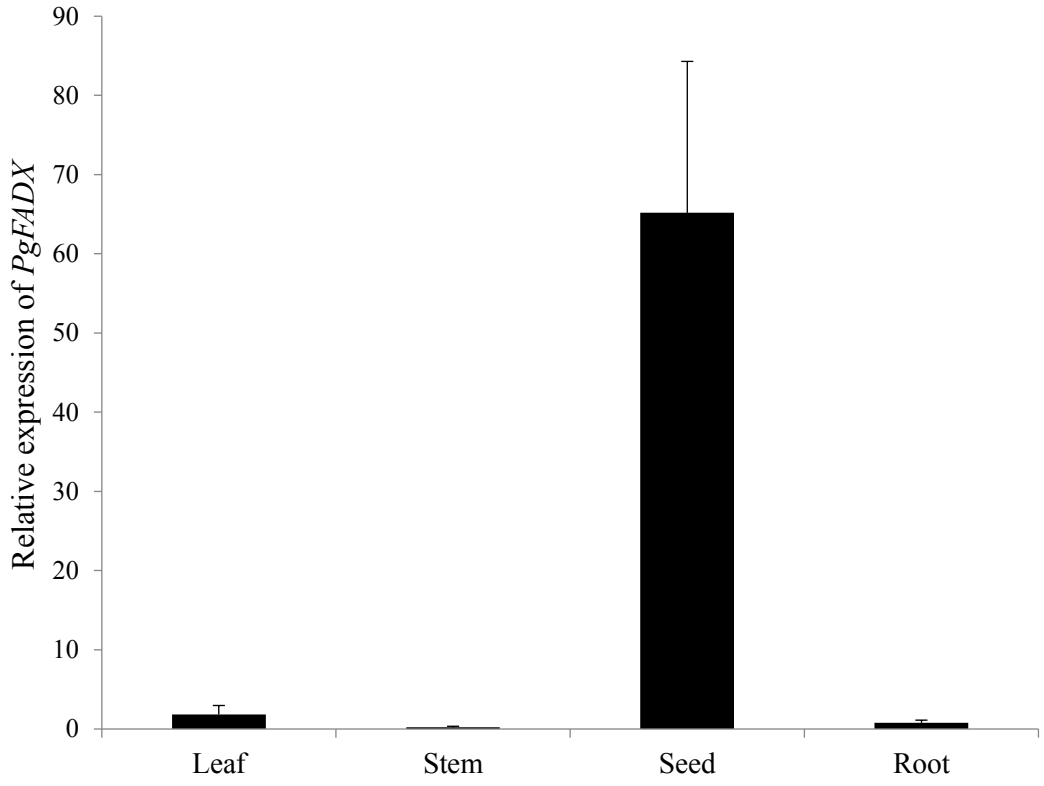


Figure 12. Expression pattern of *PgFADX* in different tissues of pomegranate. Relative expression levels of different tissues were measured by qRT-PCR. Data represent the mean ± SD for three replicates of each tissue type. The punicic acid content of the oil from the line used for this investigation was about 80%. The developing seeds were harvested from pomegranate fruit at its middle maturing stage purchased from the local grocery store.

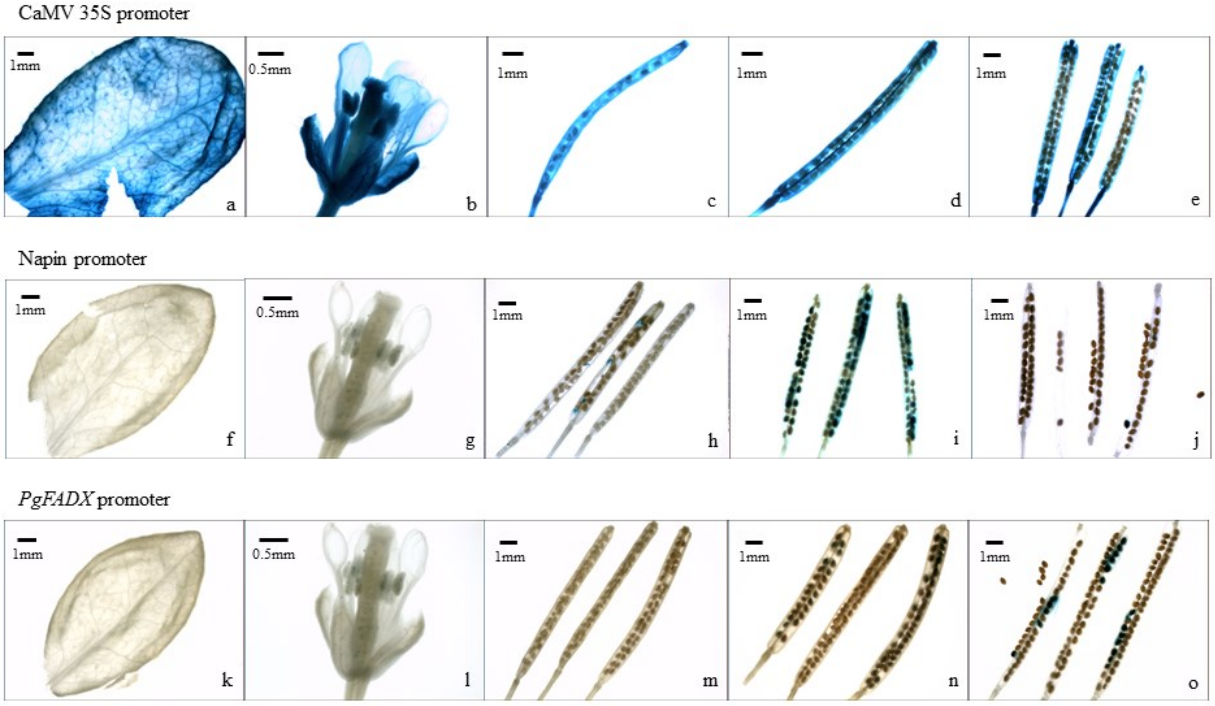
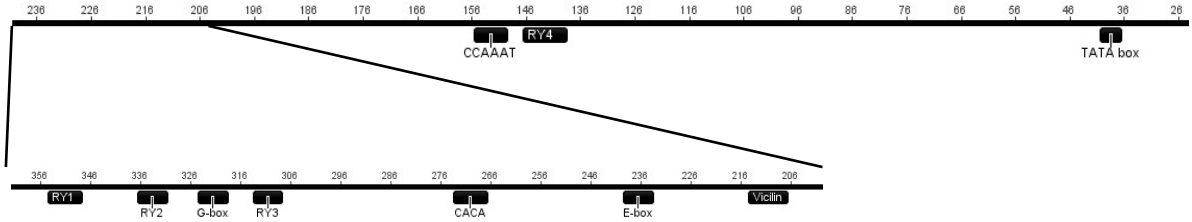


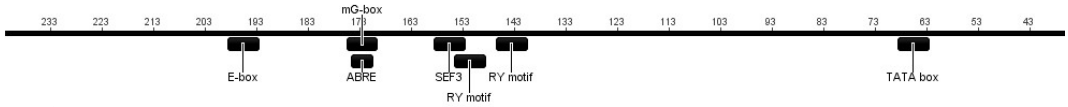
Figure 13. Histochemical β-glucuronidase (GUS) assay in various tissues of transgenic Arabidopsis carrying *GUS* gene under the control of different promoters: (a-e) CaMV 35S promoter; (f-j) napin promoter; (k-o) PgFADX promoter. a, f, k: leaf; b, g, l: flower; c, h, m: early-stage developing seeds; d, i, n: middle-stage developing seeds; e, j, o: late-stage developing seeds.

a. Napin promoter

b. Phaseolin promoter



c. Linin promoter



d. Conlinin promoter

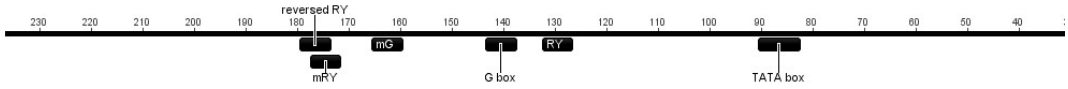


Figure 14. Schematic representation of *cis*-regulatory elements in the proximal regions of promoters. Base numbers starting from the 3' end are marked.

5. Discussion

Oil crops represent the main source of oils and fats in the human diet. The nutritional properties depend on the FA composition of the seed oil. However, common oil crop species produce a limited diversity of FAs, relative to many non-agronomic plant species. Metabolic engineering offers an attractive approach to produce novel FAs by cloning genes related to the biosynthesis of novel FAs from other species, characterizing the metabolic pathways, and expressing the transgenes in suitable hosts (Rogalski and Carrer 2011). Early studies of simple over-expression of the *FADX* gene resulted in low levels of accumulation of CLNAs in the engineered plants compared with those in the original species (Cahoon et al. 1999; Cahoon et al. 2006). In the case of punicic acid, over-expression of *PgFADX* under the control of napin promoter in WT Arabidopsis has been shown to result in the accumulation of punicic acid with up to ~4.4% (w/w) of the total FAs (Iwabuchi et al. 2003). Several factors limiting the production of CLNAs have been discussed (Mietkiewska et al. 2014a) and one of the factors has to do with promoter effectiveness. Therefore, the current work has studied the effectiveness of different promoters in expressing *PgFADX* to accumulate punicic acid in transgenic Arabidopsis seeds. Four heterologous promoters for seed storage protein genes from different crop species were compared for their efficiency in driving the expression of *PgFADX* in engineered Arabidopsis.

Overexpression of *PgFADX* under the control of the four promoters resulted in moderate production (~4-20%) of punicic acid in Arabidopsis seeds. The highest level was achieved in the T_2 segregating seeds of SL6-16 carrying *PgFADX* driven by the linin promoter, accounting for 19.5% of total FAs. This value is very close to the highest production level (21%) previously reported (Mietkiewska et al. 2014b), where *PgFADX* and *PgFAD2* were co-expressed under the

direction of the napin promoter. Therefore, it is possible that co-expression of *PgFADX* and *PgFAD2* under the direction of the linin promoter may result in higher than 21% punicic acid in the Arabidopsis seed oil. T_3 seeds derived from T_2 seeds with a single transgenic locus accumulated more stable levels of punicic acid among lines carrying the same construct (i.e. driven by the same promoter). Promoter strength was evaluated based on a comparison associating the punicic acid level with the relative expression level of *PgFADX*. The linin promoter was the most effective in directing the expression of *PgFADX* to produce punicic acid. While the other three promoters had comparable strength, the conlinin promoter exhibited inconsistency with fairly high production of punicic acid (10.28%) but a low expression of *PgFADX*. It is possible that the conlinin promoter directed more efficient gene expression at an earlier or later stage of seed development. A recent study on the production of α-eleostearic acid by over-expressing cDNA encoding the tung tree (*Vernicia fordii*) FA conjugase (VfFADX) compared the effectiveness of promoters from the soybean *β-CONGLYCININ α' SUBUNIT* gene, soybean *GLYCININ G1 SUBUNIT* gene and *PHASEOLIN* gene. The promoter from soybean *GLYCININ G1 SUBUNIT* gene was shorter than its original form and other counterparts, but performed very well with comparable production of α-eleostearic acid to that of the *β-CONGLYCININ α' SUBUNIT* promoter. Thus, it is useful to optimize the linin promoter by examining the performance of its truncated versions in driving the expression of *PgFADX*.

Consistency in FA composition of both generations of seeds (T_2 and T_3) was found in all selected Arabidopsis lines carrying different constructs, indicating a stable expression of *PgFADX* under the control of these four promoters. In a previous study, the napin promoter used to drive the expression of *PgFADX* was shown to result in distinct changes in the proportions of several major FAs relative to the non-transformed lines (Mietkiewska et al. 2014b). In the

present work there were six representative lines of T_2 segregating seeds driven by the napin promoter (Figure 6). The proportions of those major FA species were consistent with the previous study: the punicic acid content was up to 11.26% in the previous NCJ line of T_2 segregating seeds while in the present study the highest punicic acid content produced by the T_2 segregating seeds of the napin promoter lines is 14.89% (SL2-34). Other major FAs exhibited similar proportions to those previously reported: 16:0 (7.25-8.45%), 18:0 (3.77-4.00%), 18:1 (39.83-47.60%), 18:2 (23.11-36.65%), 18:3$\Delta^{9cis,12cis,15cis}$ (0.44-0.86%).

Variations were present in total FA composition among the top-performing transgenic lines carrying different constructs. Among T_2 top performers shown in Table 5, SL6-16 driven by the linin promoter produced the highest level of punicic acid (19.5%), accompanied by the highest levels of 18:1, 20:0 (equal to SL2-34), 20:1, 22:0 and two unidentified CLNAs. However, it produced the lowest levels of 16:0, 18:2 and 18:3. A somewhat modified FA composition was apparent in SL7-37 driven by the conlinin promoter, which had the lowest production of punicic acid (9.84%) among the top performers. SL7-37 accumulated the lowest levels of 18:1, 20:0 (same as SL4-23), 20:1, 22:0 and two unidentified CLNAs, but the highest levels of 18:0, 18:2 and 18:3. Similar changes in FA composition were generally observed in the T_3 top performers. While the levels of those FA species closer to punicic acid in the FA biosynthetic network appeared to have a similar amplitude of change, those more distant from punicic acid in the FA biosynthetic network, such as 16:0 and 20:0, exhibited more scattered changes from punicic acid. These changes in FA composition indicate promoter strength from another perspective: as higher production of punicic acid resulted from higher expression of *PgFADX* driven by the stronger promoter, those FA species that are closer to punicic acid in the biosynthetic network exhibited changes in their proportions that mirrored the changes in punicic acid content.

Lines accumulating increased punicic acid evidently exhibited changes in FA metabolism. Over-expression of *PgFADX* driven by heterologous promoters from seed storage protein genes probably led to down-regulation of native *FAD2* expression in the transgenic Arabidopsis as described by (Mietkiewska et al. 2014b). Lower *FAD2* expression was assumed to result in lower FAD2 enzymatic activity. The increased proportion of 18:1 due to reduced native FAD2 activity could have provide more substrate available for metabolic processes utilizing 18:1. As the long-chain FAs (>C18) are synthesized via the elongation of C18 FA catalyzed by FA elongase (FAE), it is noteworthy that the Arabidopsis mutant *fad3/fae1* selected, as the background for transformation in this work, is deficient in long-chain monounsaturated FAs (LC-MUFAs) due to the inactivation of FAE1 compared with its original Columbia WT (Smith et al. 2003). In particular, the Columbia WT contains ~20% 20:1 (Li et al. 2006) whereas in the *fad3/fae1* mutant there is only 0.2-0.3% 20:1 (Table 5 and 6). By comparing the FA composition among Columbia WT, the *fad3/fae1* mutant background and the *PgFADX*-over-expressed lines driven by heterologous promoters, it is speculated that the excess of 18:1 produced due to down-regulation of native *FAD2* might have partially restored the levels of 20:1 and 24:1, albeit at much lower levels than those of the Columbia WT due to the inhibited FAE1. Moreover, as described in the Results section, the joint increase in 20:1 and 24:1 reflected their biosynthetic connection: that 24:1 was accumulated via indirect elongation from 20:1. The lower level of α-linolenic acid in the transgenic lines versus the mutant background is the result of reduction in the 18:2 as its substrate. The considerable decrease of 16:0 upon the over-expression of *PgFADX* was in marked contrast to the increase of the long-chain saturated FAs, namely 20:0, 22:0 and 24:0 (Table 5 and 6). This may have been due to the stepwise elongation of 16:0 to generate other FAs with longer chain length; however, 18:0 content remained stable without obvious

fluctuation. One possible explanation is that part of the 18:0 was used in the biosynthesis of C18 FAs such as 18:1, which balanced out the increased portion of 18:0. Over-expression of *PgFADX* resulted in concomitant generation of two unidentified CLNAs along with punicic acid. Previously, (Cahoon et al. 1999) reported that expression of *M. charantia FADX* cDNA in soybean embryos resulted in accumulation of α-eleostearic acid along with a small amount of α-parinaric acid, but only α-eleostearic acid is naturally found in the seeds of *M. charantia*. The proposed mechanism was that the soybean embryo accumulated α-linolenic acid as the substrate for biosynthesis of α-parinaric acid, whereas *M. charantia* lacked this substrate in its seeds. By analogy, in the current study, it is speculated that PgFADX is able to act on other FAs similar to linoleic acid as the substrate, thus producing these two CLNAs yet to be identified.

The production of punicic acid was generally stable over two generations of Arabidopsis, with some fluctuations that varied from promoter to promoter. The changes in the punicic acid levels may have been due to the effects of different numbers of insertion loci in different transgenic events. Segregation analysis showed different segregation ratios in the four top-performing lines as follows: SL2-34 segregated 15:1, SL4-14 segregated 3:1, SL6-16 segregated 15:1 and SL7-12 segregated 3:1. The segregation ratios correspond to the numbers of insertion loci as follows: SL2-34 and SL6-16 had two insertion loci, whereas SL4-14 and SL7-12 had a single insertion locus, respectively. Notably, the selected lines of the advanced generation were homozygous, meaning that the numbers of insertion loci were doubled in those lines relative to their parental lines. The number of insertion loci is associated with the stability of transgene expression, which could potentially explain the changes shown in Figure 10. The considerable increases in the punicic acid levels found in SL4-14 and SL7-12 with a single insertion locus resulted from the enhanced expression of *PgFADX* driven by double insertion loci in the

homozygous offspring. However, in the cases of SL2-34 and SL6-16 carrying two insertion loci respectively, there were four insertion loci in their homozygous offspring, respectively. It has been reported that multiple insertion loci of the transgene may have negative effects on gene expression induced by transgene silencing (Meyer and Saedler 1996). In the case of SL6-16, the offspring of which produced considerably lower levels of punicic acid than the parental line, it is possible that the expression of *PgFADX* was significantly affected due to multiple insertion loci in the offspring chromosome. Likewise, the milder increase in the punicic acid level of the offspring of SL2-34, relative to those of SL4-14 and SL7-12, may also have resulted from compromised transgene expression due to a multiple insertion loci effect, albeit not as serious as in SL6-16. Furthermore, the position where the transgenic loci were inserted may also have had an influence on transgene expression (Kohli et al. 2010). It is possible that in addition to the effect of multiple insertion loci, the insertion site of the transgene in SL6-16 further affected the expression of *PgFADX* in the advanced generation.

The over-expression of *PgFADX* by heterologous promoters had little influence on the seed oil content in the case where only a single insertion locus is introduced; whereas in the case of multiple insertion loci which may accumulate higher level of punicic acid, there is a considerable decrease in the seed oil content by 2-3%. As the total production of punicic acid depends on both the punicic acid proportion of total FAs and oil content, both transgenic events would result in similar total production of punicic acid. Alternatively, it is reasonable to combine biotechnological strategies for enhancing oil content with those for increasing the punicic acid level in attempts to achieve an elevated total production of punicic acid.

Promoter performance is closely related to the *cis*-regulatory elements that constitute the proximal promoter region. Bioinformatic analysis revealed that the four heterologous promoters

contained some conserved elements that are responsible for seed-specific expression. As shown in Figure 14a, the B-box is made of a distal part (DistB) (GCCACTTGTC) and a proximal part (ProxB) (CAAACAAC). The DistB functions similarly to abscisic acid responsive elements (ABREs) while the ProxB is a CA-rich element that controls the tissue specificity of the ABRE-like DistB by repressing the hormone response in vegetative tissues (Ezcurra et al. 1999). The synergistic interaction between the G-box and the repeated RY elements is crucial for seed-specific expression (Ezcurra et al. 1999). The phaseolin promoter implements stringent regulation that allows the gene to be highly expressed in seeds only during embryogenesis and completely silent in vegetative tissues (Chandrasekharan et al. 2003). As shown in Figure 14b, four RY elements are widely distributed in the proximal region of the phaseolin promoter. Notably, two RY elements (RY2 and RY3) encompass the G-box in the region after a series of positive elements including the CACA element, E-box and the CCAAAT box. However, the vicilin box spanning 207-214 bp has a negative role in gene expression to re-establish the repressive state during late stages of embryogenesis (Chandrasekharan et al. 2003). The *cis*-regulatory elements in the linin promoter also orchestrate seed-specific expression (Figure 14c). The repeated RY elements, a modified G-box as the ABRE, an E-box and a SEF3 element are present in the linin promoter. The functions of these elements were presented in the literature review of this thesis. For SEF3, it is still unknown whether or not the SEF3 element is directly involved in seed-specific transcriptional control and thus a fine structure-function study is needed to unravel its role. The *cis*-regulatory elements present in the conlinin promoter include an RY element, a G-box and interestingly, their modified versions (Figure 14d). This series of modified versions of *cis*-regulatory elements may alter slightly the temporal regulatory pattern,

which is presumably associated with the inconsistency relative to the other three promoters between the punicic acid level and the expression level of *PgFADX* gene.

Taking the above *cis*-regulatory elements found in each promoter together with the present experimental results, it is reasonable to credit the promoter strength in driving seed-specific expression of *PgFADX* gene to some *cis*-regulatory elements commonly found in the four promoters. The repeated RY element is critical for gene activation in the seed, and the G-box and B-box act as major ABREs for the seed-specific response to ABA which is a gene regulator during seed maturation. In addition, synergistic interaction among different elements may enhance the promoter strength and regulatory effect. Variations in localization, combination and modification of these elements may differentially modulate gene expression, resulting in varied levels of punicic acid.

The native promoter of *PgFADX* was first isolated from pomegranate and characterized using a reporter gene assay. As opposed to the constitutive expression driven by the CaMV 35S promoter and the seed-specific expression driven by the napin promoter, the *PgFADX* promoter was not able to drive the expression of the *GUS* reporter gene and *PgFADX* transgene in Arabidopsis. While this might be the first case of the promoter for the gene encoding FA conjugase, similar observations of poor activity of promoters for other genes encoding FA-modifying enzymes in directing seed-specific gene expression have been recently reported. A comparative analysis between the promoter isolated from the sesame *2S ALBUMIN* (*2Salb*) gene and the promoter of the soybean *FAD3C* (*Gmfad3C*) gene demonstrated that the expression level of the *GUS* gene in tobacco seeds was 5% higher under the control of the *2Salb* promoter than of the *Gmfad3C* promoter (Bhunia et al. 2014). In another case, where the tung tree *FA CONJUGASE* (*VfFADX*) gene was over-expressed to produce α-eleostearic acid in transgenic

Arabidopsis seeds, the Arabidopsis *FA ELONGASE 1* (*AtFAE1*) promoter had considerably inferior performance to other promoters of storage protein genes, including *β-CONGLYCININ α'-SUBUNIT*, soybean *GLYCININ G1 SUBUNIT*, *PHASEOLIN*, Arabidopsis *2S-3* and *NAPIN (Shockey et al. 2015)*.

Sequence analysis of the 1000-bp upstream region of the isolated 2027-bp upstream flanking sequence of *PgFADX* gene revealed the presence of only two *cis*-regulatory elements which have been reported to confer seed-specific expression to the family of *Brassicaceae*: a single RY element 296 bp upstream of the transcription start site (TSS), a G-box 625 bp upstream of the TSS. These two elements are too distant from each other. Compared with the four heterologous promoters for seed storage protein genes, the *PgFADX* promoter lacks a RY repeats/G-box complex that is essential for activity in Arabidopsis. As Arabidopsis is a representative species of the Brassicaceae family, a shortage of *cis*-regulatory elements conserved in promoters of seed storage protein genes might account for the lack of activity of the *PgFADX* promoter in Arabidopsis. In contrast, significant *cis*-regulatory elements are conserved and enriched in the promoters for seed storage protein genes from the Brassicaceae and Fabaceae families (Fauteux and Stromvik 2009). This has been proven by the identification of *cis*-regulatory elements in the napin promoter from oilseed rape (a member of the Brassicaceae) and the phaseolin promoter from green bean (a *Fabaceae* species). The repeated RY element and G-box have been identified to be highly conserved and localized in these two promoters as well as the linin promoter and the conlinin promoter from flax (Figure 14). The phylogenic difference between pomegranate from the Lythraceae family and Arabidopsis from the Brassicaceae family may imply the presence of other *cis*-regulatory elements in the *PgFADX* promoter that are

recognizable by the native TFs in pomegranate to confer high level of seed-specific expression of *PgFADX*.

On the other hand, with respect to the category of gene products, it has been known that genes encoding storage proteins show a distinct temporal expression pattern from those encoding enzymes involved in FA biosynthesis (Baud and Lepiniec 2009; Ruuska et al. 2002), which suggests that different transcriptional networks control storage protein accumulation and FA biosynthesis. Coordination of the genome-wide response to seed development requires a set of transcriptional machineries to specifically bind to their corresponding *cis*-regulatory elements in the promoter. In Arabidopsis, there are four master regulators of seed maturation, including LEAFY COTYLEDON 1 (LEC1) and 3 TFs of the B3 DNA-binding domain family, namely ABSCISIC ACID INSENSITIVE 3 (ABI3), FUSCA3 (FUS3) and LEAFY COTYLEDON 2 (LEC2). Both repeated RY element and the G-box serve as the binding sites for the B3 TFs (Vicente-Carbajosa and Carbonero 2005). ABI3 is an important transcriptional activator related to ABA-dependent dormancy and gene expression in the maturing Arabidopsis embryos (Parcy et al. 1994). The ability of binding the promoter of genes involved in the biosynthesis of seed storage compounds (maturation-specific) to trigger the accumulation of seed-specific transcripts has also been found in LEC1, LEC2 and FUS3. By recruiting specific TFs, the composition of *cis*-regulatory elements in the promoter mediates transcription in a specific temporal and spatial pattern. Therefore, it is possible that promoters for genes encoding FA-modifying enzymes deliver more precise control over gene expression to achieve fine tuning of the temporal appearance of enzymatic activity, whereas promoters for genes encoding seed storage proteins drive the expression of genes in a more direct and sustained way which makes them more broadly useful in transgenic expression.

6. Conclusions and future perspectives

In summary, the present study demonstrated that varied levels of punicic acid were accumulated in Arabidopsis over-expressing the *PgFADX* gene driven by four heterologous promoters for seed storage protein genes from different oilseed crop species. Analyses of the FA composition of the seed oil and relative expression level of *PgFADX* in developing seeds showed that the linin promoter is most efficient. The highest content of punicic acid was 19.5% of total FAs, which was detected in the T_2 seeds of the line SL6-16 carrying *PgFADX* controlled by the linin promoter. Punicic acid levels were maintained generally stable in two generations of the transgenic seeds. Changes in FA composition were consistent among all the transgenic lines, indicating stable expression of *PgFADX* driven by these heterologous promoters. Conserved *cis*-regulatory elements were identified in the four evaluated promoters, with the repeated RY elements and the G-box being shown before to be crucial for seed-specific expression. The side-by-side comparison of promoter strength, associated with the conserved pattern of *cis*-regulatory elements, suggested the linin promoter as the best choice for the production of punicic acid by over-expressing foreign genes in Arabidopsis and possibly in other oilseed crops such as oilseed rape. Metabolic engineering strategies may benefit from using a diversity of promoters with various strengths, tissue specificities and temporal patterns combined with modifications in *cis*-regulatory elements. This requires a complete characterization of *cis*-regulatory elements in the promoter with the assistance of computational biology.

A 2027-bp flanking sequence upstream of *PgFADX* gene was isolated from the pomegranate genome by the genome walking approach and designated as the *PgFADX* promoter. The GUS assay showed this promoter had little activity in the transformed Arabidopsis seeds. Bioinformatic analysis suggested the *PgFADX* promoter was deficient in some crucial *cis*-

regulatory elements compared with the four heterologous promoters of genes encoding seed storage proteins. Furthermore, the *PgFADX* promoter might contain *cis*-regulatory elements that are specific for pomegranate but not recognizable by the TFs of Arabidopsis. Additional investigation is required for a systematic analysis of the regulatory network of the *PgFADX* gene in pomegranate associated with the high level of seed-specific expression. A deeper understanding of the nature of accumulating unusual FAs, e.g. punicic acid, in the native species, is essential for the rational design of metabolic engineering strategies to further enhance the production of the value-added FAs in oilseed crops.

References

Aitzetmuller K, Matthaus B, Friedrich H (2003) A new database for seed oil fatty acids - the database SOFA. Eur J Lipid Sci Tech 105: 92-103

Allen RD, Bernier F, Lessard PA, Beachy RN (1989) Nuclear factors interact with a soybean beta-conglycinin enhancer. Plant Cell 1: 623-631

Aruna P, Venkataramanamma D, Singh AK, Singh RP (2016) Health benefits of punicic acid: a review. Compr Rev Food Sci Food Saf 15: 16-27

Banihani S, Swedan S, Alguraan Z (2013) Pomegranate and type 2 diabetes. Nutr Res 33: 341-348

Baud S, Lepiniec L (2009) Regulation of *de novo* fatty acid synthesis in maturing oilseeds of Arabidopsis. Plant Physiol Biochem 47: 448-455

Bhunia RK, Chakraborty A, Kaur R, Gayatri T, Bhattacharyya J, Basu A, Maiti MK, Sen SK (2014) Seed-specific increased expression of 2S albumin promoter of sesame qualifies it as a useful genetic tool for fatty acid metabolic engineering and related transgenic intervention in sesame and other oil seed crops. Plant Mol Biol 86: 351-365

Biermann U, Bornscheuer U, Meier MA, Metzger JO, Schafer HJ (2011) Oils and fats as renewable raw materials in chemistry. Angew Chem Int Ed Engl 50: 3854-3871

Boothe J, Nykiforuk C, Shen Y, Zaplachinski S, Szarka S, Kuhlman P, Murray E, Morck D, Moloney MM (2010) Seed-based expression systems for plant molecular farming. Plant Biotechnol J 8: 588-606

Buchanan BB, Gruissem W, Jones RL (2000) Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rockville, MD

Bustos MM, Guiltinan MJ, Jordano J, Begum D, Kalkan FA, Hall TC (1989) Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, *cis*-acting sequence found upstream of a French bean β -phaseolin gene. Plant Cell 1: 839-853

Cahoon E, Clemente T, Damude H, Kinney A (2010) Modifying vegetable oils for food and non-food purposes. In: Vollmann J, Rajcan I (eds) Oil Crops. Springer New York, pp 31-56

Cahoon EB, Carlson TJ, Ripp KG, Schweiger BJ, Cook GA, Hall SE, Kinney AJ (1999) Biosynthetic origin of conjugated double bonds: production of fatty acid components of high-value drying oils in transgenic soybean embryos. Proc Natl Acad Sci USA 96: 12935-12940

Cahoon EB, Carlson TJ, Ripp KG, Schweiger BJ, Cook GA, Hall SE, Kinney AJ (1999) Biosynthetic origin of conjugated double bonds: production of fatty acid components of high-value drying oils in transgenic soybean embryos. Proc Natl Acad Sci U S A 96: 12935-12940

Cahoon EB, Dietrich CR, Meyer K, Damude HG, Dyer JM, Kinney AJ (2006) Conjugated fatty acids accumulate to high levels in phospholipids of metabolically engineered soybean and Arabidopsis seeds. Phytochemistry 67: 1166-1176

Calton EK, James AP, Pannu PK, Soares MJ (2014) Certain dietary patterns are beneficial for the metabolic syndrome: reviewing the evidence. Nutr Res 34: 559-568

Chandrasekharan MB, Bishop KJ, Hall TC (2003) Module-specific regulation of the β-phaseolin promoter during embryogenesis. Plant J 33: 853-866

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

Chaudhary S, van Rooijen G, Moloney MM, Singh S (2010) Flax seed specific promoters. Patent US 7642346 B2

Chen G, Woodfield HK, Pan X, Harwood JL, Weselake RJ (2015) Acyl-trafficking during plant oil accumulation. Lipids 50: 1057-1068

Chen J, Greenblatt IM, Dellaporta SL (1992) Molecular analysis of Ac transposition and DNA replication. Genetics 130: 665-676

Cheng B, Wu G, Vrinten P, Falk K, Bauer J, Qiu X (2009) Towards the production of high levels of eicosapentaenoic acid in transgenic plants: the effects of different host species, genes and promoters. Transgenic Res 19: 221-229

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

Dan Y, Ow DW (2011) Historical technology developments in plant transformation. In: Plant transformation technology revolution in last three decades. Bentham Science Publishers

Dyer JM, Chapital DC, Kuan JC, Mullen RT, Turner C, McKeon TA, Pepperman AB (2002) Molecular analysis of a bifunctional fatty acid conjugase/desaturase from tung. Implications for the evolution of plant fatty acid diversity. Plant Physiol 130: 2027-2038

Dyer JM, Stymne S, Green AG, Carlsson AS (2008) High-value oils from plants. Plant J 54: 640-655

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19: 1349

Ellerstrom M, Stalberg K, Ezcurra I, 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

Ezcurra I, Ellerström M, Wycliffe P, Stålberg K, Rask L (1999) Interaction between composite elements in the *napA* promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol 40: 699-709

Fauteux F, Stromvik MV (2009) Seed storage protein gene promoters contain conserved DNA motifs in *Brassicaceae*, *Fabaceae* and *Poaceae*. BMC Plant Biol 9: 126

Fujiwara T, Beachy RN (1994) Tissue-specific and temporal regulation of a β-conglycinin gene: roles of the RY repeat and other *cis*-acting elements. Plant Mol Biol 24: 261-272

Gasmi J, Sanderson JT (2010) Growth inhibitory, antiandrogenic, and pro-apoptotic effects of punicic acid in LNCaP human prostate cancer cells. J Agric Food Chem 58: 12149-12156

Goderis IJ, De Bolle MF, Francois IE, Wouters PF, Broekaert WF, Cammue BP (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. Plant Mol Biol 50: 17-27

Grossmann ME, Mizuno NK, Schuster T, Cleary MP (2010) Punicic acid is an omega-5 fatty acid capable of inhibiting breast cancer proliferation. Int J Oncol 36: 421-426

Hernandez-Garcia CM, Finer JJ (2014) Identification and validation of promoters and *cis*-acting regulatory elements. Plant Sci 217-218: 109-119

Hontecillas R, Diguardo M, Duran E, Orpi M, Bassaganya-Riera J (2008) Catalpic acid decreases abdominal fat deposition, improves glucose homeostasis and upregulates PPAR alpha expression in adipose tissue. Clin Nutr 27: 764-772

Hontecillas R, O'Shea M, Einerhand A, Diguardo M, Bassaganya-Riera J (2009) Activation of PPAR gamma and alpha by punicic acid ameliorates glucose tolerance and suppresses obesity-related inflammation. J Am Coll Nutr 28: 184-195

Hornung E, Pernstich C, Feussner I (2002) Formation of conjugated Δ11Δ13-double bonds by Δ12-linoleic acid (1,4)-acyl-lipid-desaturase in pomegranate seeds. Eur J Biochem 269: 4852-4859

Iwabuchi M, Kohno-Murase J, Imamura J (2003) Delta 12-oleate desaturase-related enzymes associated with formation of conjugated trans-delta 11, cis-delta 13 double bonds. J Biol Chem 278: 4603-4610

Jaworski J, Cahoon EB (2003) Industrial oils from transgenic plants. Curr Opin Plant Biol 6: 178-184

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901-3907

Johanningsmeier SD, Harris GK (2011) Pomegranate as a functional food and nutraceutical source. Annu Rev Food Sci Technol 2: 181-201

Kohli A, Miro B, Twyman RM (2010) Transgene Integration, Expression and Stability in Plants: Strategies for Improvements. In: Kole C, Michler CH, Abbott AG, Hall TC (eds) Transgenic Crop Plants: Principles and Development. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 201-237

Kohno H, Suzuki R, Yasui Y, Hosokawa M, Miyashita K, Tanaka T (2004) Pomegranate seed oil rich in conjugated linolenic acid suppresses chemically induced colon carcinogenesis in rats. Cancer Sci 95: 481-486

Koncz C, Schell J (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Molecular and General Genetics MGG 204: 383-396

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30: 325-327

Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN (1991) Multiple nuclear factors interact with upstream sequences of differentially regulated β-conglycinin genes. Plant Mol Biol 16: 397-413

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

Li Q, Wang H, Ye S, Xiao S, Xie Y, Liu X, Wang J (2013) Induction of apoptosis and inhibition of invasion in choriocarcinoma JEG-3 cells by α-calendic acid and β-calendic acid. Prostaglandins Leukot Essent Fatty Acids 89: 367-376

Li Y, Beisson F, Pollard M, Ohlrogge J (2006) Oil content of Arabidopsis seeds: the influence of seed anatomy, light and plant-to-plant variation. Phytochemistry 67: 904-915

Liu L, Hammond EG, Nikolau BJ (1997) *In vivo* studies of the biosynthesis of α-eleostearic acid in the seed of *Momordica charantia L.* Plant Physiol 113: 1343-1349

Liu WN, Leung KN (2014) Apoptosis- and differentiation-inducing activities of jacaric acid, a conjugated linolenic acid isomer, on human eosinophilic leukemia EoL-1 cells. Oncol Rep 32: 1881-1888

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408

Lu C, Fulda M, Wallis JG, Browse J (2006) A high-throughput screen for genes from castor that boost hydroxy fatty acid accumulation in seed oils of transgenic Arabidopsis. Plant J 45: 847-856

Meyer P, Saedler H (1996) Homology-dependent gene silencing in plants. Annu Rev Plant Physiol Plant Mol Biol 47: 23-48

Mietkiewska E, Giblin EM, Wang S, Barton DL, Dirpaul J, Brost JM, Katavic V, Taylor DC (2004) Seed-specific heterologous expression of a nasturtium *FAE* gene in Arabidopsis results in a dramatic increase in the proportion of erucic acid. Plant Physiol 136: 2665-2675

Mietkiewska E, Lin Y, Weselake RJ (2014a) Engineering production of C18 conjugated fatty acids in developing seeds of oil crops. Biocatal Agric Biotechnol 3: 44-48

Mietkiewska E, Miles R, Wickramarathna A, Sahibollah AF, Greer MS, Chen G, Weselake RJ (2014b) Combined transgenic expression of *Punica granatum* conjugase (FADX) and FAD2 desaturase in high linoleic acid Arabidopsis thaliana mutant leads to increased accumulation of punicic acid. Planta 240: 575-583

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

Napier JA (2007) The production of unusual fatty acids in transgenic plants. Annu Rev Plant Biol 58: 295-319

Napier JA, Graham IA (2010) Tailoring plant lipid composition: designer oilseeds come of age. Curr Opin Plant Biol 13: 330-337

Napier JA, Haslam RP, Beaudoin F, Cahoon EB (2014) Understanding and manipulating plant lipid composition: metabolic engineering leads the way. Curr Opin Plant Biol 19: 68-75

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

Ohlrogge JB, Jaworski JG (1997) Regulation of fatty acid synthesis. Annu Rev Plant Physiol Plant Mol Biol 48: 109-136

Orsavova J, Misurcova L, Ambrozova JV, Vicha R, Mlcek J (2015) Fatty acids composition of vegetable oils and its contribution to dietary energy intake and dependence of cardiovascular mortality on dietary intake of fatty acids. Int J Mol Sci 16: 12871-12890

Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J (1994) Regulation of gene expression programs during Arabidopsis seed development: roles of the ABI3 locus and of endogenous abscisic acid. Plant Cell 6: 1567-1582

Peremarti A, Twyman RM, Gomez-Galera S, Naqvi S, Farre G, Sabalza M, Miralpeix B, Dashevskaya S, Yuan D, Ramessar K, Christou P, Zhu C, Bassie L, Capell T (2010) Promoter diversity in multigene transformation. Plant Mol Biol 73: 363-378

Porto MS, Pinheiro MP, Batista VG, dos Santos RC, Filho Pde A, de Lima LM (2014) Plant promoters: an approach of structure and function. Mol Biotechnol 56: 38-49

Qiu X, Truksa M, Hu Z (2008) Flax (*Linum usitatissimum* L.) seed-specific promoters. Patent EP 1925671 A1

Rogalski M, Carrer H (2011) Engineering plastid fatty acid biosynthesis to improve food quality and biofuel production in higher plants. Plant Biotechnol J 9: 554-564

Rossak M, Smith M, Kunst L (2001) Expression of the *FAE1* gene and *FAE1* promoter activity in developing seeds of *Arabidopsis thaliana*. Plant Mol Biol 46: 717-725

Ruiz-Lopez N, Haslam RP, Napier JA, Sayanova O (2014) Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. Plant J 77: 198-208

Ruiz-Lopez N, Haslam RP, Usher S, Napier JA, Sayanova O (2015) An alternative pathway for the effective production of the omega-3 long-chain polyunsaturates EPA and ETA in transgenic oilseeds. Plant Biotechnol J 13: 1264-1275

Ruiz-Lopez N, Haslam RP, Usher SL, Napier JA, Sayanova O (2013) Reconstitution of EPA and DHA biosynthesis in Arabidopsis: iterative metabolic engineering for the synthesis of n−3 LC-PUFAs in transgenic plants. Metab Eng 17: 30-41

Ruuska SA, Girke T, Benning C, Ohlrogge JB (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. Plant Cell 14: 1191-1206

Samii-Saket G, Boersma JG, Ablett GR, Falk DE, Fletcher R, Rajcan I (2011) Single soybean seed NMR calibration for oil measurement using commercial cooking oils. J Am Oil Chem Soc 88: 1795-1798

Shinohara N, Tsuduki T, Ito J, Honma T, Kijima R, Sugawara S, Arai T, Yamasaki M, Ikezaki A, Yokoyama M, Nishiyama K, Nakagawa K, Miyazawa T, Ikeda I (2012) Jacaric acid, a linolenic acid isomer with a conjugated triene system, has a strong antitumor effect *in vitro* and *in vivo*. Biochim Biophys Acta 1821: 980-988

Shockey J, Mason C, Gilbert M, Cao H, Li X, Cahoon E, Dyer J (2015) Development and analysis of a highly flexible multi-gene expression system for metabolic engineering in Arabidopsis seeds and other plant tissues. Plant Mol Biol 89: 113-126

Singh SP, Zhou XR, Liu Q, Stymne S, Green AG (2005) Metabolic engineering of new fatty acids in plants. Curr Opin Plant Biol 8: 197-203

Smith MA, Moon H, Chowrira G, Kunst L (2003) Heterologous expression of a fatty acid hydroxylase gene in developing seeds of *Arabidopsis thaliana*. Planta 217: 507-516

Spilmont M, Leotoing L, Davicco MJ, Lebecque P, Mercier S, Miot-Noirault E, Pilet P, Rios L, Wittrant Y, Coxam V (2013) Pomegranate seed oil prevents bone loss in a mice model of osteoporosis, through osteoblastic stimulation, osteoclastic inhibition and decreased inflammatory status. J Nutr Biochem 24: 1840-1848

Truksa M, MacKenzie Samuel L, Qiu X (2003) Molecular analysis of flax 2S storage protein conlinin and seed specific activity of its promoter. Plant Physiol Biochem 41: 141-147

van Erp H, Shockey J, Zhang M, Adhikari ND, Browse J (2015) Reducing isozyme competition increases target fatty acid accumulation in seed triacylglycerols of transgenic Arabidopsis. Plant Physiol 168: 36-46

Vicente-Carbajosa J, Carbonero P (2005) Seed maturation: developing an intrusive phase to accomplish a quiescent state. Int J Dev Biol 49: 645-651

Vroegrijk IO, van Diepen JA, van den Berg S, Westbroek I, Keizer H, Gambelli L, Hontecillas R, Bassaganya-Riera J, Zondag GC, Romijn JA, Havekes LM, Voshol PJ (2011) Pomegranate seed oil, a rich source of punicic acid, prevents diet-induced obesity and insulin resistance in mice. Food Chem Toxicol 49: 1426-1430

Wang L, Li W, Lin M, Garcia M, Mulholland D, Lilly M, Martins-Green M (2014) Luteolin, ellagic acid and punicic acid are natural products that inhibit prostate cancer metastasis. Carcinogenesis 35: 2321-2330

Wang L, Martins-Green M (2014) Pomegranate and its components as alternative treatment for prostate cancer. Int J Mol Sci 15: 14949-14966

Weaver CA, Zirkle R, Metz JG (2011) PUFA polyketide synthase systems and uses thereof. Patent US7902427B2

Yasui Y, Hosokawa M, Kohno H, Tanaka T, Miyashita K (2006) Growth inhibition and apoptosis induction by all-trans-conjugated linolenic acids on human colon cancer cells. Anticancer Res 26: 1855-1860

Yoon JM, Zhao L, Shanks JV (2013) Metabolic engineering with plants for a sustainable biobased economy. Annu Rev Chem Biomol Eng 4: 211-237

Yuan GF, Chen XE, Li D (2014) Conjugated linolenic acids and their bioactivities: a review. Food Funct 5: 1360-1368

Zhao L, Katavic V, Li F, Haughn GW, Kunst L (2010) Insertional mutant analysis reveals that long-chain acyl-CoA synthetase 1 (LACS1), but not LACS8, functionally overlaps with LACS9 in Arabidopsis seed oil biosynthesis. Plant J 64: 1048-1058

Appendix: Primers used in this project

F1	5′-ATATCTCGAGATGGGAGCTGATGGAACAAT-3′
R1	5′-ATATTCTAGATCAGAACTTGCTCTTGAACC-3′
F2	5'-ATATACCGGTAAGCTTTCTTCATCGGTGAT-3'
R2	5'-ATATCTCGAGGTCCGTGTATGTTTTTAATC-3'
F3	5'-ATATTCTAGAGATCGTTCAAACATTTGGCA-3'
R3	5'-ATATGCGGCCGCCGATCTAGTAACATAGATGA-3'
F4	5′-TATAACCGGTAATTCATTGTACTCCCAGTATC-3′
R4	5′-ATATCTCGAGGTAGAGTAGTATTGAATATGAG-3′
F5	5'- TATAACCGGTCTCAAGCATACGGACAAGG -3'
R5	5'- ATATCTCGAGTGGGTATCTTTTGGATGGATAA-3'
F6	5'- TATAACCGGTCAACGGTTCCGGCGGTATAG-3'
R6	5'- ATATCTCGAGTTTTTGGTGGTGATTGGTTC-3'
PF-683	5′-AGATATTCAACTTGAGAGAGCG-3′
PR-768	5′- GGCTAGCCGGTAGAGGATGT-3′
18S-F	5′-GCCAAAACGGCTCCGAAACA-3′
18S-R	5′-ACTGGCAGTCCCTCGTGAG-3′

F7	5′-CTTGAGAGAGCGGTTCTGGG-3′
R7	5′-ACTGGGACTCCGTAGATGCT-3′
F8	5′-TGCCCTTCCACATGCTATCC-3′
R8	5′-TTCCCGTTCAGCAGTAGTTGT-3′
GSP1	5′-TGAGGACATGAGGAGCGAGCGGTGGAAG-3′
GSP2	5′-GCGCTCCGGAGGTCGCTCAAGGTGAAG-3′
F9	5′-ATATACCGGTATCCGAAAGTCTAATTGAACTT -3′
R9	5′-TATAGGATCCGTTTTGGGTGTAGTAGTAGG-3′
F10	5′-ATATGGATCCATGTTACGTCCTGTAGAAACC-3′
R10	5′-ATATGGTACCTCATTGTTTGCCTCCCTGCT-3′
F11	5′-ATATCTCGAGATGTTACGTCCTGTAGAAACC-3′
R11	5′-ATATTCTAGATCATTGTTTGCCTCCCTGCT-3′

The underlied sequences indicate the restriction enzymes used for cloning.