Novel roles for acyl-synthetases and acyl-thioesterases in plant lipid metabolism.

By

Ian Pulsifer

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy:

Biology

Carleton University

Ottawa, Ontario, Canada

Abstract

Acyl lipids are a diverse and important class of biomolecules. In plants, acyl lipids include ubiquitous compounds such as phospholipids, galactolipids, and triacylglycerols, as well as more specialized compounds including cutin and suberin monomers, cuticular waxes, and oxylipins. All acyl lipids are derived from fatty acids, which generally require activation for metabolism to either a CoenzymeA (CoA) or an acyl carrier protein (ACP) thioester. Acyl-synthetases and acyl-thioesterases, enzymes that activate and deactivate fatty acids, respectively, provide important checkpoints in lipid metabolism. Historically, the activation/deactivation paradigm of acyl-synthetases and acyl-thioesterases has been quite simple. However, various more sophisticated roles of these enzymes classes have been suggested, which is in line with the fact that a typical plant genome encodes many more of these enzymes than a simple activation/deactivation model would require. In this thesis I explore some of these more specialized roles of fatty acyl-synthetases and fatty acyl-thioesterases using the model plant *Arabidopsis thaliana*. First, I present evidence that Arabidopsis long-chain acyl-CoA synthetases LACS1, LACS2, and LACS3 are involved in trans-membrane movement of fatty acids, possibly playing a role in directing flux of fatty acids into particular biosynthetic pathways. This finding has implications in our understanding of the trafficking of ubiquitous fatty acyl-CoA substrates in plant cells, the conclusions of which can be extrapolated to other LACS proteins. Second, I characterize a novel family of Arabidopsis acyl-ACP thioesterases called *ALT1*, *ALT2*, *ALT3*, and *ALT4.* These genes were found to be expressed in various tissues throughout the plant, each with a unique expression pattern. When

expressed heterologously in *E. coli*, each ALT protein generated a unique set of medium-chain fatty acids and/or medium-chain β-ketofatty acids. The occurrence of medium-chain fatty acids and β-ketofatty acids in various tissues of Arabidopsis plants provides clues regarding the biological roles of *ALT1-4*. Potential *ALT* homologues are found in nearly all plant genomes, and the encoded proteins are likely involved in specialized metabolism. This research opens many doors for future research, both in terms of basic research of plant lipid metabolism as well as applied uses of specialized high-value acyl lipids.

Preface

This thesis follows the format of the integrated thesis. The data and text from Chapters 2 and 3 are taken from peer-reviewed journal articles for which I am the primary author. Both articles have been reproduced in full for this thesis. Minor modifications have been made to the text and figures, mainly for re-formatting purposes, but also for updating in light of more recent publications. The data presented in Chapter 4 are not yet published.

Chapter 2 is derived from the research paper:

Pulsifer IP, Kluge S, Rowland O. (2012) Arabidopsis long-chain acyl-CoA synthetase 1 (LACS1), LACS2, and LACS3 facilitate fatty acid uptake in yeast. Plant Physiol Biochem. 51:31-9.

http://www.sciencedirect.com/science/article/pii/S0981942811003068

Statement of contribution

The experiments described in this chapter were conceived and designed by O. Rowland and me. All text and figures from this article are reproduced with permission of Elsevier publishers.

My contributions towards this paper:

- Construction of two plasmid vectors (for expression of *Sc*FAT1p and *Sc*FAA1p in *Saccharomyces cerevisiae*).
- All experiments were run and all data analysis was performed by me.

- The manuscript, including text and figures, was prepared by me, with guidance from O. Rowland.

Other author's contributions:

- Generation of four plasmid vectors (empty vector backbone, and plasmids for expressing *At*LACS1, *At*LACS2, *At*LACS3 in *Saccharomyces cerevisiae*).

Chapter 3 is derived from the research paper:

Pulsifer IP, Lowe C, Narayaran SA, Busuttil AS, Vishwanath SJ, Domergue F, Rowland O. (2014) *Acyl-lipid thioesterase1-4* from *Arabidopsis thaliana* form a novel family of fatty acyl-acyl carrier protein thioesterases with divergent expression patterns and substrate specificities. Plant Mol Biol. 84(4-5):549-63. http://link.springer.com/article/10.1007%2Fs11103-013-0151-z

Statement of contribution

The experiments described in this chapter were conceived and designed by C. Lowe, O. Rowland, and me. All data and text from this article is reproduced with permission of Springer Publishing.

My contributions towards this paper:

- Phylogenetic analysis of *ALT* genes, and potential homologues in other plant species (Figures 3.1-3.2)

- All work towards expression profiles of ALT genes using quantitative polymerase chain reaction (Figure 3.3)
- Staining and imaging of ALT promoter::GUS lines (Figure 3.4, panels G, H, I, L, and M)
- Generation of plasmids for expression of ALT2, ALT3, and ALT4 in *Nicotiana benthamiana*
- Imaging of all fluorescent proteins expressed in *Nicotiana benthamiana* (Figure 3.5)
- All work for expression of truncated ALT proteins in *Escherichia coli* (Figures 3.6-3.9)
- Purification of truncated ALT proteins, and *in vitro* acyl-ACP thioesterase assays (Figure 3.10)
- The manuscript, including text and figures, was prepared by me, with guidance from O. Rowland.

Other authors' contributions:

- Generation of ALT promoter::GUS lines
- Staining and imaging of some ALT promoter::GUS lines (Figure 3.4, panels A, D, E, F, J, K, N, O, and P)
- *In situ* hybridization of *ALT1* transcript (Figure 3.4, panels B and C)
- Generation of plasmids for expression of ALT1, ALT1-CTP only, and ALT1-ΔCTP in *Nicotiana benthamiana*

Chapter 4 presents data that is not yet published.

Statement of contribution

The experiments described in this chapter were conceived and designed by O. Rowland and me. The experiments and data analysis were performed entirely by me. The figures and text were prepared by me, with guidance from O. Rowland.

Table of contents

List of figures

List of tables

List of appendices

List of abbreviations

ACP	acyl carrier protein
ACS	acyl-CoA synthetase
BSA	bovine serum albumin
cDNA	complimentary DNA
CoA	coenzyme A
ER	endoplasmic reticulum
FA	fatty acid
FATP	fatty acid transport protein
GC-FID	gas chromatography-flame ionization detection
GC-MS	gas chromatography-mass spectrometry
LACS	long-chain acyl-CoA synthetase
LCFA	long-chain fatty acid
MCFA	medium-chain fatty acid
MK	methyl ketone
PBSE	phosphate buffered saline with EDTA
RT-PCR	reverse transcriptase polymerase chain reaction
TE	thioesterase
VLACS	very-long-chain acyl-CoA synthetase
VLCFA	very-long-chain fatty acid
qPCR	quantitative polymerase chain reaction

Chapter 1: Introduction to plant acyl lipids, acyl-synthetases, and acyl-thioesterases

Lipids are one of the four major classes of biological macromolecules. Acyl lipids are a large subset of lipids that are derived from fatty acids. Acyl lipids show immense diversity in both their structures and their functions. Certain classes of acyl lipids are fundamental to life and are found in every living cell. Other types are more specialized and are found in only specific cell types of certain organisms. This chapter will summarize the structures and functions of different acyl lipid classes important to plant cell functioning, and the metabolism of fatty acids to generate these various acyl lipids. The foci will be the roles of two classes of enzymes, acyl-synthetases and acyl-thioesterases, in the overall scheme of lipid metabolism. At the end of this chapter, the roles of acyl-synthetases and acyl-thioesterases from plants and other systems will be explored to highlight the theme of this thesis: the roles of acyl-synthetases and acyl-thioesterases in lipid metabolism of the model plant *Arabidopsis thaliana* are broad and are likely more sophisticated than is currently appreciated.

1.1 Acyl lipids have many important roles in plant cells

1.1.1 Phospholipids and galactolipids form cellular membranes.

Arguably the most fundamental role of acyl lipids is to form cellular membranes. These membranes encompass every living cell, separating it from its environment, and are paramount to the cellular compartmentalization that allows eukaryotic cells to attain their level of complexity. Cells modulate the fatty acid components of membranes to alter their basic properties, such as fluidity, curvature, and permeability to various solutes. Further, there are many membrane-associated

proteins that rely on interactions with specific lipids for correct folding and orientation (Hunte, 2005).

The principle components of cellular membranes are diacyl-glycerolipids. These complex lipids contain acyl groups esterified to two of the three hydroxyl groups of a glycerol backbone (*sn*-1[1] and *sn*-2). The *sn*-3 position contains one of various hydrophilic head groups, after which the molecules are generally named. In most cellular locations, the head group contains a phosphate group giving these lipids the name phospholipids. Plastids, however, have a unique set of membrane glycerolipids. Plastid membranes are comprised mostly of galactolipids, with galactose containing head groups, or sulfolipids, with sulfoquinovosyl head groups. In nearly all instances, the acyl groups of membrane lipids have 16 or 18 carbons, and up to three double bonds. The saturation level of acyl groups can be altered in response to different metabolic or environmental state, such as changing temperature (Martinière *et al*., 2011).

1.1.2 Triacylglycerols store energy in seeds

Another important function of acyl lipids is to act as an energy reserve. The highly reduced carbon atoms of acyl groups make fatty acids effective substrates for aerobic respiration. Further, the hydrophobic nature of lipids allows them to pack into a relatively small space, free of solvating water molecules. Combined, these properties make lipids a very efficient form of energy storage. Lipid energy stores are

[1] The *stereospecific numbering* (*sn*) system is used to denote specific hydroxyl groups of glycerol. *sn*-1 and *sn*-3 are the terminal OH-groups, and *sn*-2 is the internal OH-group.

generally in the form of triacylglycerols, which are found in all organisms. In oilseed plants, such as Arabidopsis or *Brassica napus*, triacylglycerols are relatively highly abundant in seeds (compared to protein and carbohydrate), where they act as the main source of carbon and cellular energy for developing seedlings before they begin photosynthesis. Smaller amounts of triacylglycerols are also found in pollen, fruit, and leaf tissues (Li-Beisson *et al*. 2013).

Triacylglycerols are a form of glycerolipid, in which all three hydroxyl groups of the glycerol backbone are esterified to acyl moieties. This makes triacylglycerols electrostatically neutral and almost completely non-polar. The acyl groups of Arabidopsis seed oil contain predominately 16:0[2], 18:(0-3), 20:(0-2), and 22:1 acyl groups (O'Neill *et al*., 2003). In certain cases, seed oils of other plant species may contain longer or shorter acyl chains, or acyl chain with additional functional groups (Dyer *et al*., 2008).

1.1.3 Extracellular lipid barriers protect plant cells from biotic and abiotic stresses

Like all organisms, plants must defend themselves against a myriad of environmental stresses. Plant use extracellular lipid barriers as a first line of defense against many of these stresses. The cuticle coats aerial surfaces of all terrestrial plants (Bernard & Joubès, 2012). Suberin is deposited, for example, in the cell walls of root endodermal and peridermal cells, in the seed coat, and at sites of wounding

[2] This nomenclature is used to describe the acyl group of a fatty acid. The number before the colon indicates the number of carbon atoms and the number after the colon indicates the number of double bonds in the acyl chain.

(Pollard *et al*., 2008). Sporopollenin is a component of the outer layer of pollen grains, the exine (Kim & Douglas, 2013). The hydrophobic nature of these barriers makes them very effective at preventing uncontrolled water loss, and individual constituents of these barriers mediate specific interactions with microbes, insects, and herbivores (Samuels *et al*., 2008).

Suberin, sporopollenin, and cutin (a component of cuticle) are cross-linked polymeric matrices deposited on the outer surfaces of their respective cell types (Pollard *et al*., 2008; Kim & Douglas, 2013). These polymers are built from fatty acid derivatives cross-linked with glycerol and various phenolic compounds (Pollard *et al*., 2008). Cutin acyl monomers are derived exclusively from long chain[3] fatty acids (16:0, 18:(0-3)), while suberin acyl monomer precursors also include saturated very-long-chain fatty acid (20:0-26:0). Prior to incorporation into cutin or suberin, these fatty acids undergo various oxidative modifications, including hydroxylation, carboxylation, and epoxidation (Pollard *et al*., 2008). The nature of sporopollenin monomers is less well characterized. The fatty acid precursors for sporopollenin include long- and medium-chain fatty acids, although the exact structures of the end products are unclear (de Azevedo Souza *et al.*, 2009; Morant *et al.*, 2007). Unlike cutin or suberin, sporopollenin also contains polyketides (Dobritsa *et al*., 2010).

[3] "Short-chain", "medium-chain", "long-chain", and "very-long-chain" describe the number of carbon atoms in a fatty acid's acyl chain. Short-chains have fewer than six carbons, medium-chains have 6-14 carbons, long-chains have 16-18 carbons, and very long-chains have 20 or more carbons.

Within the cuticle, cutin is both impregnated with and coated by cuticular waxes. Cuticular waxes are non-polymerized lipid molecules, including both acyl and terpenoid lipids. The acyl lipid component of cuticular waxes are derived from saturated very-long-chain fatty acids (20:0-34:0), and include such compounds as alkanes, ketones, and fatty alcohols (Samuels *et al*., 2008).

1.1.4 Sphingolipids have many diverse roles

Sphingolipids are a small, yet important, component of all eukaryotic cellular membranes (reviewed in Pata *et al*., 2010). The functions of sphingolipids in plant cells are numerous and varied. These roles include, but are not limited to, anchoring of certain proteins to membranes, modulating apoptosis, providing tolerance to cold and drought stress, and mediating cell-cell interactions (Sperling & Heinz, 2003; Pata *et al*., 2010). The structures of sphingolipids are highly varied, with over 150 different molecular species in Arabidopsis alone (Markham & Jaworski, 2007). Plant sphingolipids originate from saturated fatty acids (16:0 and 24:0) and the amino acid serine, though they undergo extensive and varied modifications.

1.1.5 Oxylipins are plant hormones and signalling agents

Oxylipins are another class of functionally and structurally diverse lipid molecules and are found in most aerobic organisms (reviewed in Mosblech *et al.*, 2009). Plants synthesize oxylipins by oxidation of polyunsaturated fatty acids, usually 18:3, followed by extensive and varied modifications. Oxylipins are used to regulate a variety of defense and developmental pathways. Acyclic oxylipins are usually made in response to cellular damage, either mechanical or pathogen/pest

induced. These compounds can act as signaling molecules, either intracellularly, between tissues, or even between individuals, or may be bacteriocides. Jasmonates, including jamsonic acid and derivatives thereof, are a class of cyclized oxylipins that act as plant homornes. Jasmonates have a wide variety of roles, mainly in regulation of defense and developmental pathways (Balbi & Devoto, 2008).

1.2 Fatty acid metabolism in plant cells

Plants do not acquire lipids from the environment, nor do they transport acyl lipids between cells or tissues. Because of this, every plant cell must synthesize and metabolize its own acyl lipids. For all cells, acyl lipid metabolism begins with *de novo* synthesis of fatty acids in the plastid, during which intermediates are activated with acyl carrier protein (ACP). Once long-chain acyl tails of fatty acids are synthesized, the acyl groups are partitioned into various pathways, which occur mainly at two cellular locations: the plastid and the endoplasmic reticulum (ER).

All cells will use a portion of these acyl groups to generate membrane lipids. Fatty acids that remain in the plastid are mostly converted into galacto- and sulfolipids to produce plastid membranes, in what is called the prokaryotic pathway. Fatty acids that are exported to the ER are synthesized into phospholipids to produce other cellular membranes, in what is known as the eukaryotic pathway. Although the prokaryotic and eukaryotic pathways are initiated in the plastid and the ER, respectively, glycerolipids are exchanged between the two organelles later in metabolism (Moreau *et al.*, 1998). In addition to the ubiquitous membrane lipids, other acyl lipids are generated in particular cell types and developmental stages.

Triacylglycerols and surface lipids are synthesized at the ER, while oxylipins are synthesized in the plastid.

1.2.1 Fatty acids are metabolically active as acyl-thioesters

Fatty acids can form thioesters with one of two activating groups: Coenzyme A (CoA) and acyl-carrier protein (ACP) (Figure 1.1). CoA is a ubiquitous metabolite found in all organisms and is used in many biochemical reactions involving activation or transfer of acyl groups. ACP is also found in all organisms, although not always as a distinct molecule, as it is found in plants. In animal and fungal cells, ACP is a domain of a large multi-functional protein. CoA and ACP are both used by all organisms to form acyl-thioesters, but they are used for different processes. ACP is the activating group for fatty acids that are undergoing *de novo* fatty acid synthesis in the plastid, whereas CoA is the activating group in essentially all other metabolic pathways.

The Arabidopsis genome encodes at least eight ACP isoforms, five of which are predicted to be localized to the plastid, and three to the mitochondria (Mekhedov *et al*., 2000). Gene expression analysis in Arabidopsis shows that multiple ACP isoforms are expressed in all tissues (Bonaventure & Ohlrogge, 2002). Although the various ACP isoforms have different expression patterns (Bonaventure & Ohlrogge, 2002), whether specific isoforms have distinct roles is not yet known.

The most important structural feature of CoA and ACP is the 4-phosphopantethine moiety (Figure 1.1). CoA is comprised of this group joined to an adenine nucleotide, whereas in ACP the 4-phosphopantetheine attaches to a serine residue of the protein. The end of this moiety has a free thiol group that forms the

thioester bond with a free fatty acid. The long and flexible 4-phosphopantetheine group acts as a tether, allowing the acyl group to interact with enzyme active sites, while remaining covalently bound to CoA or ACP. This is especially important for cyclic pathways, like *de novo* synthesis or oxidative degradation of fatty acids, where each protein within a complex acts on a given substrate molecule repeatedly.

Thioesterification of fatty acids serves various purposes. The first purpose is to increase their aqueous solubility. The exceedingly low water solubility of free fatty acids would make it very difficult for these substrates to interact with the various enzymes required for their metabolism, and to move around the cell. CoA and ACP are both hydrophilic, and therefore their acyl thioesters become more water soluble than the free acyl group. The solubility of 16:0-CoA is predicted to be more than 4000-fold greater than that of 16:0 fatty acid (Tetko *et al*., 2005).

A second purpose of thioesterification of fatty acids is thermodynamic in nature. It is very common for fatty acids to form oxygen esters, such as in glycerolipids. Since esterification of a free fatty acid to an oxygen ester is slightly energetically unfavourable (Mogelson *et al*., 1984) and the ester products (*e.g.* phospholipids within a membrane) accumulate to high levels, expenditure of energy is needed to drive these reactions forward. As discussed below, esterification of a fatty acid to CoA uses two ATP equivalents. The resulting CoA thioester is more reactive than the free fatty acid and the energetically favourable hydrolysis of the thioester bond is able to drive the oxygen esterification reaction.

Activation and deactivation of fatty acids by addition or removal of a thioester provide important checkpoints in the overall scheme of lipid metabolism in cells of plants and other organisms. The properties and roles of the enzymes that esterify free fatty acids and hydrolyze acyl-thioesters, acyl-synthetases and acyl-thioesterases, respectively, are discussed in more detail below (Sections 1.3 and 1.4).

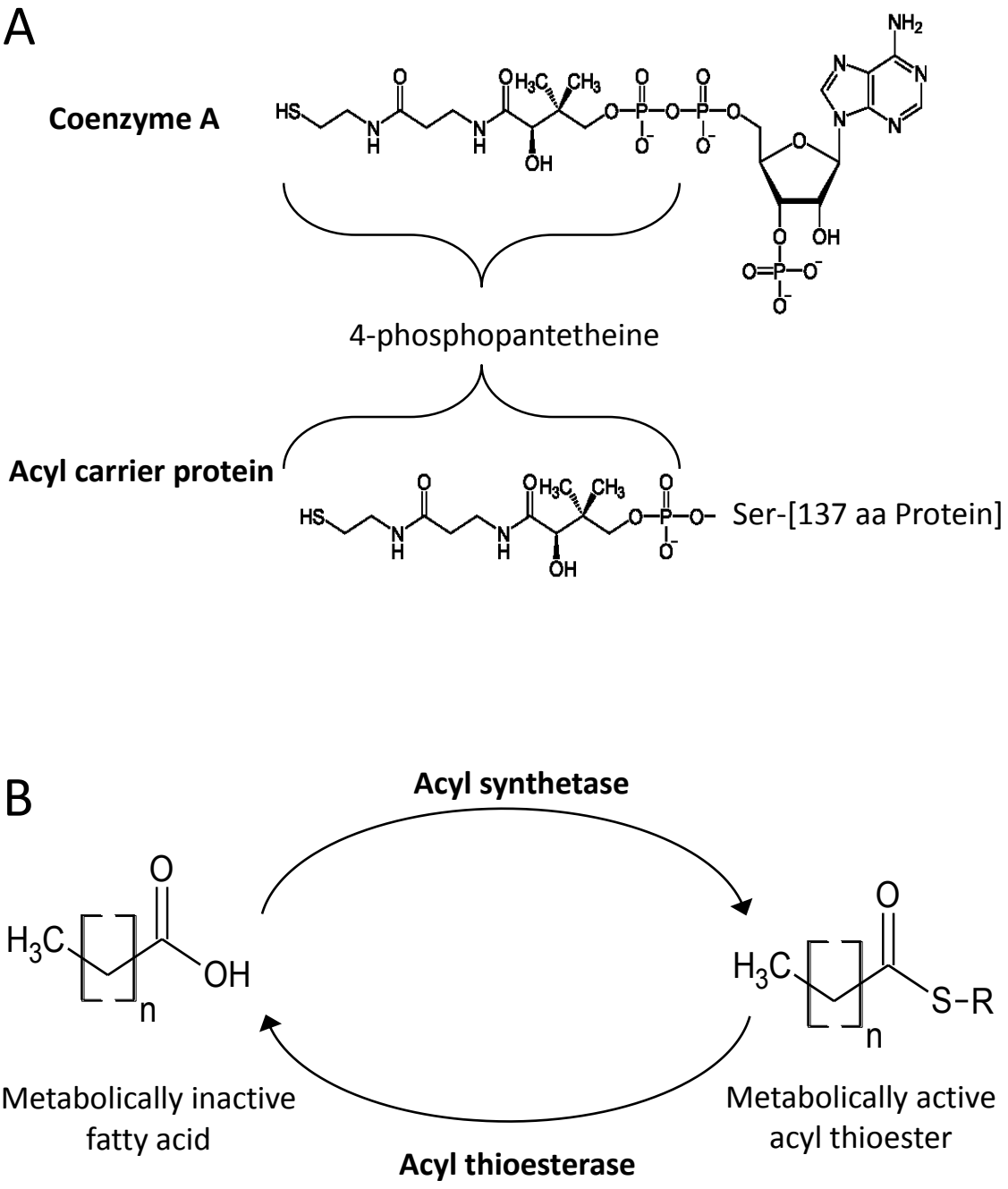


Figure 1.1 Fatty acids are metabolically active as acyl thioesters. (A) Structures of coenzyme A (CoA) and acyl carrier protein (ACP). The 4-phosphopantetheine group is attached either to an adenine nucleotide (CoA) or a protein molecule through a serine residue (ACP). (B) Fatty acids are activated to thioesters by acyl synthetases and deactivated to free fatty acids by acyl thioesterases. R = CoA or ACP.

`

1.2.2 Fatty acid synthesis in plants

Fatty acids are derived from acetyl-CoA. Acetyl-CoA is a ubiquitous metabolite that is used in a variety of metabolic pathways. A portion of the acetyl-CoA pool is dedicated to fatty acid synthesis by conversion to malonyl-ACP. First, the two-carbon acetyl-CoA is carboxylated to the three-carbon malonyl group. Next, the malonyl group is transacylated from CoA to ACP. Malonyl-ACP will be the carbon donor that sequentially builds fatty acyl chains for every turn of the fatty acid synthesis cycle. The only carbon atoms in an acyl chain that are not derived from malonyl-ACP are the two carbons used to initiate this process, which are derived from an acetyl group. It is unclear if acetyl-CoA is used directly, or if it is first transacylated to acetyl-ACP (Clough *et al*., 1992).

The fatty acid synthesis cycle (reviewed in Rawsthorne, 2002 & Li-Beisson *et al.,* 2013) repeatedly elongates a growing acyl chain by iterative additions of two-carbon groups (Figure 1.2). The first step is the condensation of malonyl-ACP to a fatty acyl-thioester. For the first turn of the cycle, this thioester is acetyl-ACP/CoA, whereas in all subsequent turns, the thioester is the fatty acyl-ACP generated by the previous turn of the cycle. Each cycle begins with production of a β-ketoacyl-ACP that is two carbons longer that the substrate thioester, accompanied by the release of CO_2 and free ACP. The β-ketoacyl-ACP intermediate is then reduced to a β-hydroxyacyl-ACP in an NADPH dependant reaction. Dehydration of the hydroxyl group results in a *trans*-β-enoylacyl-ACP. Finally, the double bond of this intermediate is reduced in another NADPH dependant reaction, producing a fully reduced acyl-ACP that has been elongated by two carbons.

This process is repeated, adding two-carbon groups each time derived from malonyl-ACP, until the acyl groups reach a chain length of 16 or 18 carbons. Three out of the four enzymatic activites involved in the fatty acid synthesis cycle are displayed by single proteins that participate in every round of elongation. In contrast, β-ketoacyl-ACP synthase activity (KAS, the first activity of the cycle), is displayed by three different isoforms, each active on different acyl chain lengths. KASIII is involved in only the first turn of the cycle, elongating 2:0-ACP/CoA to 4:0-ACP (Jaworski *et al.,* 1989). KASI is involved in the next six turns, elongating 4:0-ACP to 16:0-ACP (Shimakata & Stumpf, 1982). A portion of fatty acid exit the cycle at this stage, while the remainder are elongated to 18:0-ACP by a final turn of the cycle involving KASII (Shimakata & Stumpf, 1982). Essentially all C16 fatty remain fully saturated after fatty acid synthesis, whereas most of C18 fatty acids become monounsaturated (Li-Beisson *et al.,* 2013). In this case, a soluble stearoyl-ACP desaturase aerobically introduces a *cis*-double bond at the Δ9 position (Kachroo *et al.,* 2007).

The major end-products of fatty acid synthesis are 16:0-ACP and 18:1-ACP, while 18:0-ACP is a minor product (Browse & Somerville, 1991). At this point, the ACP group is removed from the acyl chain by one of two methods: transacylation of the acyl group onto a glycerol backbone or hydrolysis to a free fatty acid by an acyl-ACP thioesterase. As described in the preceding sections, the method of ACP removal dictates which subcellular site an acyl group will be metabolized, and the relative levels of these two activities are important in partitioning of acyl groups to various metabolic fates.

A minor level of fatty acid synthesis is also known to occur in mitochondria (Wada *et al*., 1997). The biochemistry is thought to be similar to that of plastid fatty acid synthesis. The major products of mitochondrial fatty acid synthesis are octanoyl-ACP (8:0), palmitoyl-ACP (16:0), and stearoyl-ACP (18:0) (Gueguen *et al*., 2000). The acyl moiety of octanoyl-ACP is directed towards lipoic acid (Wada *et al*., 1997), a sulfo-lipid cofactor utilized by several mitochondrial enzymes, by the action of an acyltransferase (Gueguen *et al*., 2000). The acyl groups of palmitoyl-ACP and stearoyl-ACP are believed to be directed towards *de novo* synthesis of phospholipids in the mitochondria through the action of acyltransferases (Li-Beisson *et al.,* 2013). It is thought that little, if any, acyl chains synthesised in the mitochondria are exported to other cellular locations.

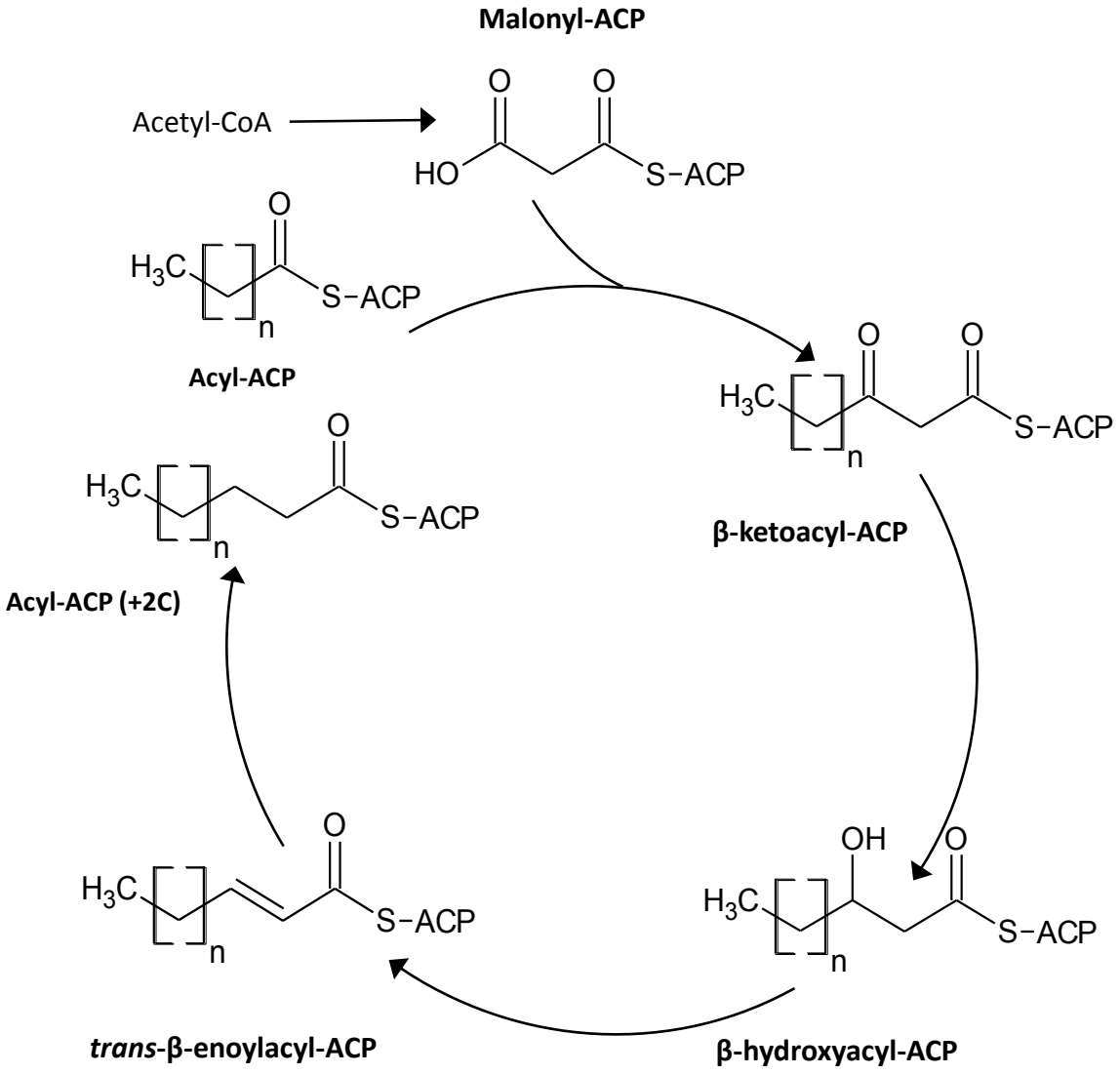


Figure 1.2 *De novo* synthesis of fatty acids is a four step cyclic pathway. Acetyl-CoA is the ultimate source of carbon atoms for generation of fatty acids, and is committed to this process by conversion to malonyl-ACP. Malonyl-ACP is condensed to either acetyl-CoA (initial cycle), or a fatty acyl-ACP (all subsequent cycles). The resultant β-ketoacyl-ACP is two carbon longer than the substrate, and is reduced to a fatty acyl-ACP by three additional reactions.

1.2.3 Further metabolism of fatty acids in the plastid

A portion of newly synthesized acyl groups are metabolized in the plastid (Figure 1.3). Plastids of all cells produce galacto- and sulfo-lipids (reviewed in Ohlrogge & Browse, 1995 & Li-Beisson *et al.,* 2013). Acyl groups are directed towards this pathway by acyltransferases that transfer fatty acid chains from ACP to the *sn*-1 and *sn*-2 positions of glycerol-3-phosphate. The diversity of plastid membrane lipids is achieved by desaturating acyl chains to varying degrees, and by addition of various headgroups onto the *sn*-3 phosphate (*e.g.* monogalactose, diagalactose, sulfoquinovose).

Oxylipins are signalling molecules involved in various developmental and pathogen defense pathways (reviewed in Mosblech *et al.*, 2008). Their biosynthesis begins with a lipase removing a trienoic fatty acid from a plastidal galactolipid, most often linoleic acid (18:3 Δ9, 12, 15). A double bond on this free fatty acid is then oxidized with a hydroperoxy group. In the synthesis of acyclic oxylipins, the fatty acid is cleaved at the hydroperoxy group, producing two shorter chain molecules that undergo various oxidative modifications. Jasmonic acid, a cyclic oxylipin, is generated by cyclising the hydroperoxy fatty acid, and shortening the acyl chain by three rounds of β-oxidation (see next section). Jasmonic acid can also undergo further modification, such as methylation or conjugation to an amino acid.

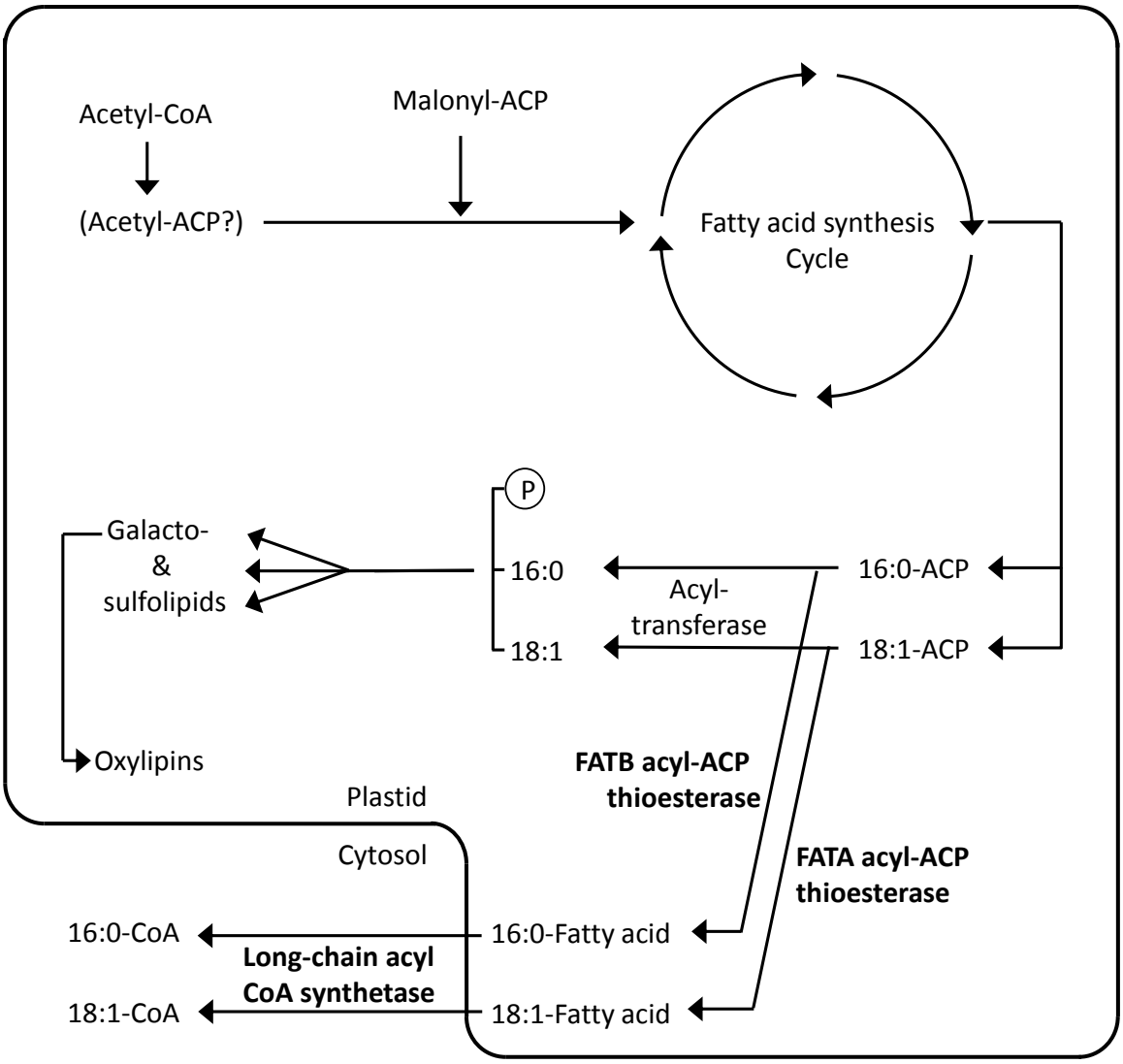


Figure 1.3 Fatty acid metabolism in the plastid. Fatty acids are synthesized *de novo* in all plant cells, producing mainly 16:0-ACP and 18:1-ACP. Mature acyl-ACP are partitioned into metabolic pathways, either within the plastid or exported to the cytosol. Some acyl chains are directed towards galacto- and sulfo-lipid synthesis within the plastid by the transfer from ACP onto a glycerol backbone. In some cells, fatty acids will be extracted from galactolipids for production of oxylipins. Acyl chains that will be exported from the plastid are hydrolyzed to free fatty acids by acyl-ACP thioesterases. FATA and FATB are acyl-ACP thioesterases that produce unsaturated and saturated fatty acids, respectively. As these fatty acids are exported, they are rapidly activated to CoA by a long-chain acyl CoA synthetase at the plastid outer envelope.

1.2.4 Export of fatty acids from the plastid for metabolism at the ER

Newly synthesized acyl chains that are destined for further metabolism at the ER are hydrolyzed from ACP into free fatty acids by plastid acyl-ACP thioesterases. These free fatty cannot be used as substrates to synthesize galacto- and sulfolipids, but instead are able to leave the plastid. The exact mechanism of exporting fatty acids is unclear, although it's known that the fatty acids are rapidly activated to CoA by a long-chain acyl-CoA synthetase upon reaching the cytosol, at least in part by an acyl-CoA synthetase localized at the plastid outer envelope (Schnurr *et al.*, 2004).

The majority of C16-18 fatty acyl-CoAs enter one of various metabolic pathways upon reaching the ER (Figure 1.4), including a portion that is further elongated to very-long-chains (≥C20) by the fatty acid elongation pathway (reviewed in Bernard & Joubès, 2013; Li-Beisson *et al*., 2013). This process is similar to fatty acid synthesis, but with a few key differences: mainly that the process occurs at the ER membrane and that substrates are acyl-CoA's and malonyl-CoA (rather than acyl-ACP's). Saturated or monounsaturated acyl-CoA's are repeatedly elongated by two-carbon units by the same reactions used in fatty acid synthesis. Products exit the fatty acid elongation cycle as 20:0-34:0 fatty acyl-CoA's, but the mechanisms for termination of this cycle and controlling of flux into downstream pathways are unknown. The ER-localized eukaryotic pathway generates phospholipids for most cellular membranes (reviewed in Ohlrogge & Browse, 1995; Li-Beisson *et al.*, 2013). The overall scheme of phospholipid synthesis is similar to that of galacto- and sulfo-lipids. Acyl chains are transacylated from CoA to the *sn*-1 and *sn*-2 positions of

glycerol-3-phophate, desaturated to varying degrees, and one of various head groups added to the *sn*-3 phosphate (*e.g.* choline, ethanalomine, inositol).

The compositions of acyl chains in phospholipids and the free acyl-CoA pool are dynamic. Acyl-editing is a process that deacylates and reacylates a phosphatidylcholine molecule, primarily at the *sn*-2 position (Bates *et al.,* 2009). This allows for rapid modification of phospholipid saturation levels, as well as generation of polyunsaturates in the acyl-CoA pool, which are then utilized for synthesis of other classes of complex lipids. Acyltransferase enzymes have been suggested to catalyze both reactions of this cycle, swapping acyl groups from a glycerol ester to CoA and back again (Lager *et al.*, 2013).

Synthesis of triacylglycerols (reviewed in Bates *et al.*, 2013; Li-Beisson *et al.,* 2013) also occurs at the ER and begins with metabolites of the eukaryotic phospholipid pathway. After acyl chains are transferred from CoA onto the *sn*-1 and *sn*-2 positions of glycerol-3-phosphate, the phosphate group is removed, and a third acyl group is transacylated to the *sn*-3 position. The diacylglycerol phosphate precursors are subject to acyl editing and the acyltransferases involved in generating triacylglycerols draw from the same acyl-CoA pool as those that synthesis phospholipids. However, acyltransferases involved in making triacylglycerols are often able to accept more diverse fatty acyl groups.

Since triacylglycerols are a major energy source for seedlings, their catabolism is also an important aspect of lipid metabolism (reviewed in Theodoulou & Eastmond, 2012; Li-Beisson *et al*., 2013). Acyl groups are first hydrolyzed from their glycerol

backbone by lipases, producing free fatty acids and glycerol. Peroxisomes are the main site of β-oxidation of released fatty acids, although this process also occurs in the mitochondria of at least some plants (Masterson & Wood, 2001). At the site of β-oxidation, free fatty acids are activated to CoA thioesters by a peroxisomal long-chain acyl-CoA synthetase. Fatty acid oxidation is a cyclic pathway analogous to the fatty acid synthesis pathway turning in reverse. Each turn of the β-oxidation cycle generates acetyl-CoA and an acyl-CoA that has been shortened by two carbons. This core cycle is simply repeated for saturated fatty acids until they are fully oxidized to acetyl-CoA, while unsaturated acyl groups require auxiliary activities to isomerize or reduce their double bonds.

Sphingolipids are low abundance, but important components of cellular membranes. Their biosynthesis is complex and relatively poorly understood (reviewed in Sperling & Heinz, 2003; Li-Beisson *et al*., 2013). Synthesis of all sphingolipids begins with the condensation of a 16:0-CoA to serine at the ER. The resulting metabolite will then be modified by complex and branching pathways that derivatize the acyl chain by various means before an acyl group from either 16:0-CoA or 24:0-CoA is transferred onto the serine-derived amino group. These resulting ceramides can undergo further desaturation and/or hydroxylation, and finally addition and possibly exchange of a head group.

Cutin, suberin, and sporopollenin are extracellular lipid polymers (reviewed in Pollard *et al.,* 2008; Li-Beisson *et al*., 2013; Kim & Douglas, 2013). Although details of assembly of these polymers are scarce, the biosynthesis is well characterized for

suberin and cutin. Both cutin and suberin monomers are generated at the ER, although in different cell types. Cutin synthesis uses 16:0- and 18:(0-3)-CoAs as substrates, while suberin is derived from not only these long chain fatty acyl-CoAs but also very-long-chain 20:0-26:0-CoAs. Various enzymes oxidize these substrates, adding ω-hydroxyl, mid-chain hydroxyl, or mid-chain epoxyl groups; or oxidizing a ω-hydroxyl to a carboxyl group. Some of these modified acyl-CoA's are then transacylated to the *sn*-2 position of glycerol-3-phosphate (Li *et al*., 2007; Beisson *et al*., 2007) and phenolic compounds (Gau *et al*., 2009; Molina *et al.*, 2009). Long-chain acyl-CoA synthetase activity in important for production of cutin monomers (Lü *et al.,* 2009). Whether this enzyme simply activates free fatty acids prior to derivatization, of if they have more sophisticated roles is unclear. Little is known about the synthesis or even the composition of sporopollenin, although it's presumed that monomers are derived from similar substrates and by similar chemistry as for cutin and suberin.

Cuticular waxes are a major product of epidermal cells and together with cutin form the cuticle (reviewed in Bernard & Joubès, 2013; Li-Beisson *et al*., 2013). Cuticular waxes are synthesized in the ER from very-long-chain saturated acyl-CoA's. A portion of cuticular waxes are free fatty acids, which are presumably generated by a very long-chain acyl-CoA thioesterase. A very long-chain acyl-CoA synthetase has been implicated in the cuticular wax pathway (Lü et al., 2009), but since very-long-chain fatty acids are already in the form of a CoA thioester after elongation, its role is presently unclear. Two separate pathways convert very-long-chain acyl-CoA's into the various components of cuticular waxes, including primary alcohols, wax esters, aldehydes, alkanes, secondary alcohols and ketones (all very-long-chain).

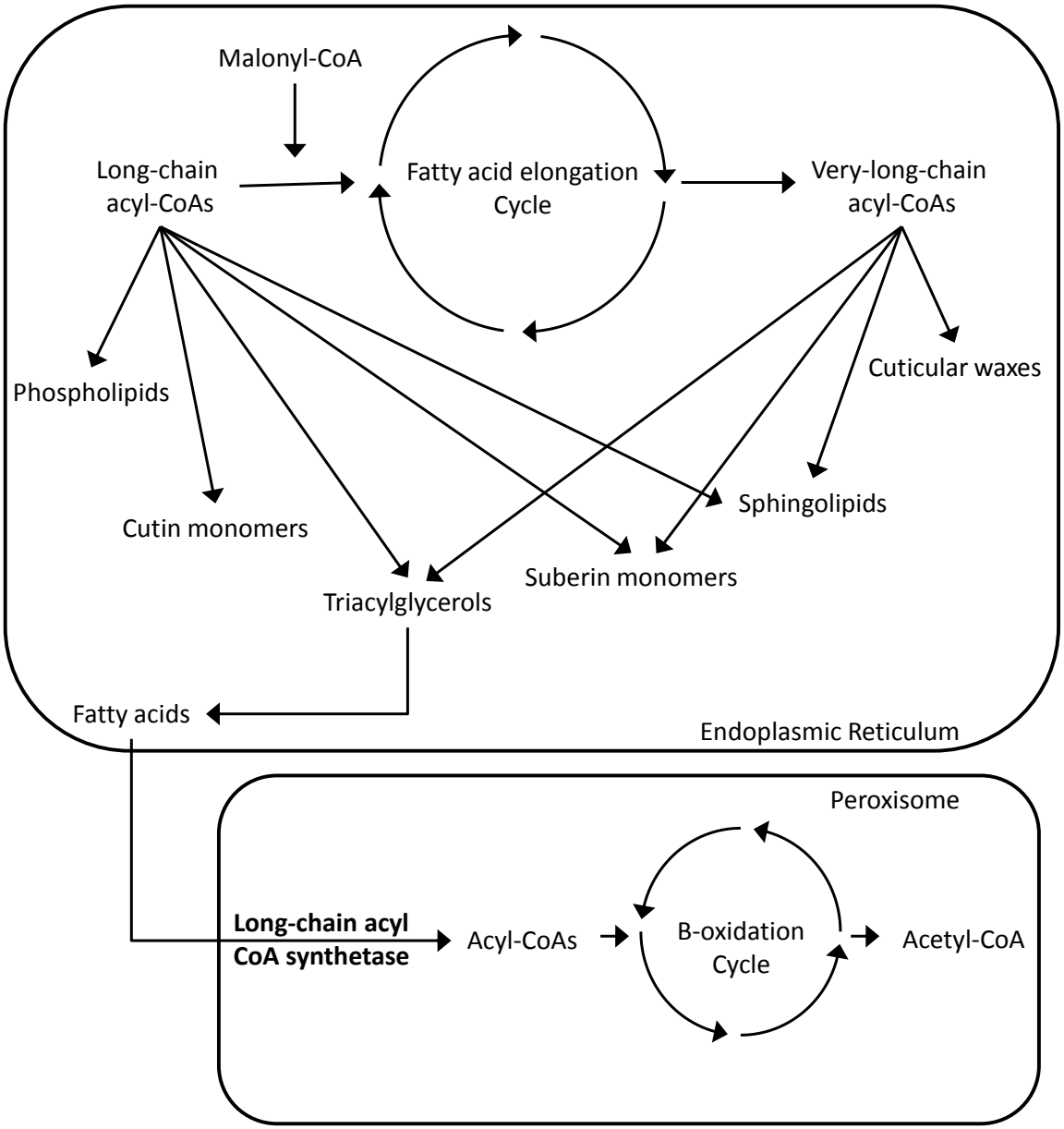


Figure 1.4 Fatty acid metabolism at the ER and in the peroxisome. Long-chain acyl CoAs are delivered to the ER, a portion of which are elongated to very-long-chain acyl-CoAs. These two pools of substrates feed into various downstream pathways (Note: not all of these pathways occur in all cell types). When they are to be oxidized, free fatty acids are released from triacylglycerols, and transported into the peroxisome. Here, they are activated to CoA, and oxidized to acetyl-CoA.

1.3 Acyl-synthetases

Acyl thioesters are formed from free fatty acids by an enzyme family known as acyl-synthetases. Fatty acid activation is a two-step process that uses energy from hydrolysis of ATP (Groot *et al*., 1976). The first step generates an acyl-AMP intermediate, accompanied by the release of pyrophosphate. The AMP moiety is then replaced by an activating group, usually CoA, although ACP is also possible. This process is relatively energetically neutral, as one high energy bond is consumed and one is formed, with subsequent hydrolysis of pyrophosphate driving this reaction forward.

The Arabidopsis genome encodes nine long-chain acyl CoA synthetases (Shockey *et al*., 2002). These proteins have been localized to various subcellular locations: the plastid, the ER, and the peroxisome (Schnurr *et al*., 2002; Weng *et al*., 2010; Fulda *et al*., 2004). Long-chain acyl CoA synthetases are invovled in various fatty acid biosynthetic pathways, including cuticular wax and cutin synthesis, synthesis of triacylglycerols, and β-oxidation of fatty acids in developing seeds (Lü *et al*., 2009; Zhao *et al*., 2010; Fulda *et al*., 2004). As discussed in Section 1.3.1 and in Chapter 2, many questions about the various roles of these enzymes in lipid metabolism remain unanswered.

Also present in the Arabidopsis genome are at least one plastid localized medium-chain acyl ACP synthetase (Koo *et al*., 2005) and at least one peroxisomal short/medium-chain acyl CoA synthetase (Wiszniewski *et al.*, 2009), neither of which has a known function. A broad search of the Arabidopsis genome identified 63

proteins with the characteristic amino acid motif of acyl synthetases (Shockey et al., 2003). This group of enzymes includes the fatty acyl synthetases mentioned above, and a group of CoA synthetases that activate carboxylic acid groups of non-fatty acid substrates. However, several of these 63 genes are presently uncharacterized and these are potentially additional fatty acyl synthetases.

The well described role of acyl synthetases in the grand scheme of lipid metabolism is quite simple. Newly synthesized fatty acids are exported from the plastid in an unesterified state, and are rapidly activated to acyl-CoAs by a long-chain acyl-CoA synthetase. The long-chain acyl-CoAs are then able to enter the various ER localized metabolic pathways (Figure 1.3). There are few known instances of free fatty acids downstream of this; free long-chain fatty acids are released from triacylglycerols for oxidative degradation as well as from galactolipids for synthesis of oxylipins. The fatty acids destined for β-oxidation are activated by another long-chain acyl-CoA synthetase, while those to be used for oxylipin synthesis remain unesterified.

1.3.1 Roles of long-chain acyl-CoA synthetases in Arabidopsis

Long chain acyl-CoA synthetase (LACS) enzymes have been implicated in various acyl lipid metabolic pathways in Arabidopsis. Phenotypic characterization of *lacs* mutants strongly suggests that particular LACS enzymes activate fatty acids for use in particular metabolic pathways. However, when the results from various experiments are compiled, many questions arise as to how LACS enzymes fit into the

grand scheme of acyl-lipid metabolism, and how a common pool of fatty acid substrates are divvied between several distinct pathways.

The above model suggests that all fatty acids that are metabolized outside of the plastid should flux through one or more long-chain acyl-CoA synthetases located at the plastid outer membrane. The major plastid LACS in Arabidopsis was identified to be *LACS9* (Schnurr *et al*., 2002). As expected, *LACS9* is expressed in all tissues and roughly correlates to amount of acyl lipids produced in a given organ (Schnurr *et al*., 2002). Plastids isolated from a *lacs9* knockout mutant have an 85-90% reduction in LACS activity compared to wild-type plastids, while the residual activity was attributed to an unidentified minor plastidal LACS activty (Schnurr *et al*., 2002).

Surprisingly, despite this significant reduction in LACS activity, the *lacs9* mutant displays no obvious phenotype. Mutant plants have the same appearance and growth rate as wild-type, and composition of glycerolipids across the cell was unaltered (Schnurr *et al*., 2002). Even in developing seeds, where huge amounts of fatty acids are directed towards the ER for seed oil synthesis, fatty acid composition between wild type and the *lacs9* mutant were similar (Zhao *et al.,* 2010). It's possible that the unidentified minor plastidal LACS is able to fully compensate for LACS9 in this mutant, although this would imply that wild type plants have at least a ten-fold excess of plastidal synthetase activity. *Why is no phenotypic effect seen with such a large decrease in plastidal LACS activity?*

Another important role of acyl-synthetases is to activate the free fatty acids liberated from triacylglycerols prior to β-oxidation. In Arabidopsis, this role is

accomplished by *LACS6* and *LACS7* (Fulda *et al.*, 2004). These two genes are highly expressed during seed germination, when tricacylglycerol stores are being consumed, and the encoded proteins localize to the peroxisome (Fulda *et al.*, 2002). Single knock out mutants of *LACS6* and *LACS7* have no discernable phenotype, whereas *lacs6 lacs7* double mutants are unable to utilize triacylglycerol stores (Fulda *et al.*, 2004). Free fatty acids accumulate in this mutant, and are not consumed, like they are in the wild type (Fulda *et al.*, 2004). Further, mutant seedlings die within one week of germination unless grown on media supplemented with sucrose as an alternate carbon source (Fulda *et al.*, 2004).

Interestingly, the *lacs6 lacs7* double mutant is sensitive to both the proherbicide 2,4-dichlorophenoxybutyric and the prohormone indole-3-butyric acid (Fulda *et al.*, 2004). These two compounds are used as indicators of β-oxidation in plants. Both compounds contain an acyl group with a free carboxylic acid, which in wild type plants are activated to CoA thioesters within the peroxisome and then shortened by one round of β-oxidation to generate active forms (Zolamn *et al.,* 2000; Wain & Wightman, 1954). The resulting products, indole-3-acetic acid and 2,4-dichlorophenoxyacetic acid, have hormonal or herbicidal activity, respectively, both of which halt plant growth. The sensitivity of the *lacs6 lacs7* mutant to these compounds indicates that they become activated by CoA in the peroxisome. However, despite this apparent CoA synthetase activity, fatty acids are not degraded in the *lacs6 lacs7* mutant. *Why are fatty acids not oxidized in the* lacs6 lacs7 *double mutant, despite the presence of a peroxisomal synthetase activity?*

Long-chain acyl-CoA synthase activity is also directly important for cuticle synthesis in Arabidopsis. A *lacs1* knock out mutant has diminished cuticular wax, while a *lacs2* knock out mutant has altered cutin monomer composition (Lü *et al*., 2009). Both mutants display a defective cuticle, displaying increased permeability (Shnurr *et al*., 2004; Lü *et al*., 2009; Weng *et al*., 2010), and in the case of *lacs2* altered susceptibility to pathogens (Tang *et al.,* 2007; Bessire *et al.,* 2007). The cuticle defects have an additive effect in the *lacs1 lacs2* double mutant (Lü *et al*., 2009; Weng *et al*., 2010). Both LACS1 and LACS2 protein localize to the ER (Weng *et al*., 2010). Although the *lacs1* and/or *lacs*2 mutant phenotypes clearly indicate that these proteins are involved in cuticle biosynthesis, their roles are paradoxical. Long chain acyl groups are delivered to the ER as CoA thioesters and the intermediates and products of fatty acid elongation remain esterified to CoA. Thus, ER localized LACSs seem unnecessary for generating very-long-chain acyl-CoA precursors used to generate cuticular waxes. *What are the roles of LACS proteins in synthesis of cuticular waxes and cutin monomers when the substrates for these pathways are already CoA activated?*

Another role of long-chain acyl-CoA synthetases in Arabidopsis is pollen coat development. A *lacs1 lacs4* double mutant is male sterile due to a defective pollen coat (Jessen *et al*., 2011). This mutant has reduced lipids, both in the sporopollenin layer and in cuticular waxes (Jessen *et al*., 2011). Less severe phenotypes are seen in the *lacs1* single mutant, but not in a *lacs4* single mutant, indicating that *LACS1* is the main synthetase required for these functions (Jessen *et al*., 2011).

Another acyl-synthetase, *ACOS5*, is thought to play a role in sporopollenin development as well. No chemical analysis of *acos5* sporopollenin has been reported, although this mutant produces no pollen and has structural defects in the anthers (de Azevedo Souza *et al.*, 2009). The ACOS5 protein has long-chain fatty acyl-CoA synthetase activity *in vitro* (de Azevedo Souza *et al.*, 2009), although it belongs to a different phylogenetic group than the other Arabidopsis LACSs (Shockey *et al*., 2003). This may be because ACOS5 also acts on medium- and long-chain hydroxylated fatty acid substrates (de Azevedo Souza *et al.*, 2009), which are presumed to be sporopollenin intermediates.

LACS activities also play a role in synthesis of triacylglycerols. A *lacs1 lacs9* double mutant shows a modest, though significant, reduction in seed triacylglycerol levels (Zhao *et al.*, 2010). As described previously, LACS1 localizes to the ER while LACS9 localizes to the plastid. Interestingly, *lacs1* and *lacs9* single mutants do not have a seed oil phenotype, implying that these two genes are genetically redundant (Zhao *et al.*, 2010). Further, LACS8, despite having the same or similar subcellular localization, expression levels in seeds, and substrate preferences as LACS1, is not able to substitute for LACS1 in triacylglycerol synthesis (Zhoa *et al.*, 2010; Shockey *et al*., 2002). *How are LACS1 and LACS9 genetically redundant, despite localizing to different parts of the cell? Why are LACS8 and LACS1 not redundant in seed oil production, despite significant similarities?*

1.3.3 Roles of acyl-CoA synthetases in other systems

It is evident that long-chain acyl-CoA synthetases play various important roles in plant lipid metabolism, although these roles are not currently clear. In addition to metabolic activation of fatty acids, these proteins are known to play more sophisticated roles in other non-plant organisms, and it is likely that some of these roles may apply to plants as well. In yeast and mammalian cells, acyl-CoA synthetases are involved in importing fatty acids from the extracellular space across the plasma membrane (see Chapter 2 for more details). There is building evidence that this role of acyl-synthetases is also involved in channeling of fatty acids throughout the cell, directing substrate towards specific locations and biochemical pathways (reviewed in Mashek *et al.,* 2007). It is currently unknown if acyl-CoA synthetases are involved in fatty acid trafficking in plant cells, but such a role may explain some of the currently unexplained aspects of the role of acyl-synthetases.

1.4 Acyl-thioesterases

Acyl-thioesterases catalyze the hydrolysis of acyl-thioesters, releasing free fatty acids. Acyl-ACP thioesterases play an important role in primary lipid metabolism, which is described in more detail in the following section. The Arabidopsis genome also encodes at least two acyl-CoA thioesterases, although their roles are less clear. The ACH2 protein is localized to the peroxisome and has 16:0-CoA thioesterase activity *in vitro* (Tilton *et al*., 2004). It has been suggested that this protein helps control the rate of fatty acid oxidation by modulating the available pool of acyl-CoA substrates. Acyl-CoA thioesterase activity is also displayed by a peroxisomal ABC transporter (De Marcos Lousa *et al*., 2013). It was originally though

that this transporter delivers acyl-CoAs to the peroxisome for β-oxidation, based on a mutant phenotype of reduce oxidation of exogenous fatty acids (Zolman *et al.,* 2001). However, the discovery that this protein also has acyl-CoA thioesterase activity, and that the transporter physically interacts with peroxisomal acyl-synthetases (De Marcos Lousa *et al*., 2013), suggests fatty acid must be in their free state before entering the peroxisome.

1.4.1 Roles of medium- and long-chain acyl-ACP thioesterases in Arabidopsis and other plants

The main role of acyl-ACP thioesterases is to release free fatty acids from mature acyl-ACPs, preventing them from being metabolised in the plastid and allowing them to be exported. This is a fundamental reaction in primary lipid metabolism, and as such, long chain acyl-ACP thioesterases are found in all plant species. These acyl-ACP thioesterases are divided into FATA and FATB types, which generate monounsaturated and saturated free fatty acids, respectively. The Arabidopsis genome encodes two FATAs and one FATB. Since FATA and FATB produce fatty acids used for ubiquitous acyl lipid classes, these enzymes are expressed in all cell types.

Certain oilseed species produce seed oils that contain saturated acyl groups shorter than the common C16 chain. These medium-chain fatty acids are generated by specialized FATB enzymes. These alternate FATBs show seed-specific expression and interrupt fatty acid synthesis, hydrolyzing acyl-ACPs to produce 8:0-14:0 fatty acids. For example, FATB2 from *Cuphea hookeriana* produces 8:0 and 10:0 fatty acids

(Dehesh *et al*., 1996). Without the ACP thioester, the medium-chain fatty acids cannot be elongated further, and are exported to the ER where they will be incorporated into triacylglycerols.

There is currently only one known example of an alternate acyl-ACP thioesterase (*i.e.* specificity less than C16) with a role other than seed triacylglycerol production. In the wild tomato *Solanum habrocaites* subspecies *glabratum*, an acyl-ACP thioesterase called METHYL KETONE SYNTHASE 2 (MKS2) is expressed in leaf trichome cells. *MKS2* also interrupts fatty acid synthesis, although instead of fully reduced fatty acyl-ACPs, the substrates for this enzyme are β-ketoacyl-ACPs, one of the intermediates of the fatty acid synthesis cycle (see Chapter 3 for more details). As will be detailed later, the Arabidopsis genome encodes several additional predicted acyl-ACP thioesterases, at least some of which are expressed in non-seed tissues. *The products of these thioesterase are not yet known nor the roles that they play in plant physiology.*

1.5 Goal of this thesis

It is evident that much remains unknown about the roles of acyl-synthetases and acyl-thioesterases in plant lipid metabolism. In this thesis, I will provide additional insights into how these two enzyme classes fit into the grand scheme of fatty acid metabolism. In Chapter 2, I explore the possibility that three Arabidopis long-chain acyl-CoA synthetase enzymes, LACS1, LACS2, and LACS3, aid in movement of fatty acids across membranes and play roles in directing fatty acids into specific metabolism pathways. I present in Chapter 3 the first characterization of a family of

medium acyl-ACP thioesterases called *ACYL LIPID THIOESTERASE 1* (*ALT1*), *ALT2*, *ALT3* and *ALT4*. The roles of these enzymes appear to be distinct from the well characterized FATA/FATB acyl-thioesterase roles described above. In Chapter 4, I extend the characterization of these novel thioesterases by characterization of mutant lines affected in these acyl-ACP thioesterase genes. Finally, in Chapter 5, I propose some future experiments to further advance these findings.

Chapter 2: Arabidopsis LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1), LACS2, and LACS3 facilitate fatty acid uptake in yeast

2.1 Abstract

The plant cuticle is a lipid-based barrier on the aerial surfaces of plants that plays a variety of protective roles. The cuticle is comprised largely of long-chain and very-long-chain fatty acids and their derivatives. In Arabidopsis, *LONG-CHAIN ACYL-COA SYNTHETASE1* (*LACS1*), *LACS2*, and *LACS3* are known or suspected cuticle biosynthetic genes. Very-long-chain acyl-coenzyme A (CoA) synthetase activity has been demonstrated for LACS1 and LACS2, although the role for such an activity in cuticle biosynthesis is currently unclear. In yeast and mammalian systems, some very-long-chain acyl-CoA synthetases are also called fatty acid transport proteins (FATPs) due to a second function of mediating transmembrane movement of fatty acids. We sought to determine if LACS1-3 also have this dual functionality. A yeast *fat1Δ* mutant is deficient in both very-long-chain acyl-CoA synthetase activity and exogenous fatty acid uptake. We demonstrate that heterologous expression of LACS1, 2, or 3 is able to complement both of these deficiencies. Furthermore, expression of each LACS enzyme in yeast resulted in uptake of the long-chain fatty acid analogue, C_1-BODIPY-C_{12}. Only expression of LACS1 resulted in uptake of the very-long-chain fatty acid analogue, BODIPY-C_{16}. These results demonstrate that LACS1, LACS2, and LACS3 have the dual functionality of yeast and mammalian FATP enzymes. These findings have implications in the transmembrane transport and intracellular trafficking of plant lipids destined for export to the cuticle.

2.2 Introduction

The plant cuticle is a protective, lipid-based coating on the aerial surfaces of land plants that is synthesized by epidermal cells (Samuels *et al.*, 2008). It consists of waxes embedded within and overlaying a cutin matrix. Cutin is mainly an esterified polymer of glycerol and C16/C18 fatty acid derivatives (*e.g.* omega hydroxy fatty acids, mid-chain hydroxylated fatty acids, and alpha-omega dicarboxylic fatty acids), whereas waxes are mainly derivatives of saturated very-long-chain fatty acids (VLCFAs, 20:0-34:0; *e.g.* primary and secondary alcohols, wax esters, aldehydes, alkanes, and ketones). Mutants with altered cuticular composition can show a variety of phenotypes, including altered epicuticular wax crystal morphology, increased non-stomatal water loss, and altered susceptibility to pathogens (Jenks & Ashworth, 1999). Two such mutants are in genes encoding LONG-CHAIN ACYL-COA SYNTHETASE1 (LACS1) and LACS2, which have overlapping functions in production of both long-chain cutin monomers and very-long-chain wax molecules in *Arabidopsis thaliana* (Schnurr *et al.*, 2004; Lü *et al*., 2009; Weng *et al*., 2010). Like other cuticle biosynthetic enzymes, LACS1 and LACS2 have been localized to the endoplasmic reticulum (ER) (Weng *et al.,* 2010; Zhao *et al*., 2010). LACS1 also contributes to triacylglycerol biosynthesis in the seed (Zhao *et al*., 2010). *LACS3*, another member of the nine Arabidopsis *LACS* genes (Shockey *et al*., 2002), is also expected to be involved in cuticle biosynthesis due to its highly enriched expression in epidermal cells (Suh *et al.*, 2005).

Acyl-CoA synthetases (ACS) are a large family of enzymes that catalyze thioesterification of free fatty acids to coenzyme A (CoA) (Watkins. 2008), which is generally a prerequisite for metabolic utilization. Activation of acyl chains with CoA is ATP-dependent and is a two-step process proceeding via an enzyme-bound acyl-AMP intermediate (Groot *et al*., 1976). Distinct classes of ACS enzymes exist, which can differ in their tissue distributions, subcellular locations, and substrate chain-length specificities (Watkins, 2008). A single ACS will typically act on a range of fatty acid chain lengths, but will have preference towards a subset. It has been demonstrated that LACS1 and LACS2 are broad-substrate ACSs with high activities for saturated fatty acids of chain-lengths C16 to C30 (Lü *et al*., 2009; Shockey *et al*., 2002). While the chain-length specificities of LACS1 and LACS2 correlate with the chain lengths of the cuticle components altered in the *lacs1* and *lacs2* mutants (Lü *et al*., 2009), the role of very-long-chain ACS (VLACS) activities in cuticle biosynthesis remains unclear, particularly in the case of cuticular waxes. Cuticular waxes are derived from very-long-chain fatty acyl substrates produced by the ER-associated fatty acid elongase (Samuels *et al.*, 2008). The elongated chains are already activated to CoA, and thus activation by a LACS seems unnecessary.

ACS enzymes that have activity towards very-long-chain fatty acids are often called Fatty Acid Transport Proteins (FATPs) because these were first identified as proteins that facilitate cellular uptake of exogenous fatty acids (Schaffer & Lodish, 1994; Coe *et al.*, 1999). A vectorial acylation model was proposed, based on a well characterized bacterial system (DiRusso & Black, 1999). In this model, a FATP functions together with a long-chain ACS (LACS) enzyme to facilitate transport of fatty

acids across the plasma membrane, with concomitant activation of fatty acids to acyl-CoAs that sequesters the lipids within the cell for subsequent metabolism. In *Saccharomyces cerevisiae*, the sole FATP protein, FAT1p, works in concert with either of two LACS proteins, FAA1p or FAA4p, to drive import of exogenous fatty acids (Faergeman *et al.*, 1997; Faergeman *et al.*, 2001). FAT1p has been reported to physically interact with FAA1p (or FAA4p), which is consistent with the vectorial acylation model (Zou *et al.*, 2003). The VLACS activity of FATP is not required for this mechanism and a recent study has shown a clear distinction of these two activities in the human protein FATP2 (Melton *et al.*, 2011). Various reports have demonstrated that FATPs are able to facilitate import of exogenous fatty acids even when localized to an internal membrane, rather than the plasma membrane (Milger *et al.*, 2006; Kage-Nakadi *et al.*, 2010). It has also been suggested that ACSs serve to traffic intracellular fatty acids by being localized to specific subcellular organelles (Digel *et al.*, 2009).

Since LACS1-3 have high VLACS activities, which is typically associated with FATP proteins in yeast and mammalian systems, we examined whether LACS1-3 can also facilitate fatty acid transport. We report here that Arabidopsis LACS1-3 can rescue the fatty acid uptake deficiencies of both yeast *fat1Δ* and *faa1Δfaa4Δ* mutants, suggesting that these proteins not only have LACS/VLACS activities, but also aid in transmembrane movement of fatty acids, which provides an additional model of how these proteins may contribute to plant cuticle biosynthesis.

2.3 Materials and methods

2.3.1 Phylogenetic Analysis

Complete protein sequences were aligned using the website "http://www.phylogeny.fr". The MUltiple Sequence Comparison by Log- Expectation (MUSCLE) algorithm was chosen to create the initial alignment, and maximum likelihood algorithm was use to infer evolutionary relationships (Dereeper *et al.*, 2008). The phylogenetic tree was rendered and formatted using TreeDyn software (Chevenet *et al.*, 2006).

2.3.2 Plasmid construction

For expression in yeast, a modified version of pYES2 (Invitrogen, Carlsbad, CA, USA) was generated. For this, the segment of pET-28a (EMD4Biosciences, Gibbstown, NJ, USA) that carries the coding region for producing fusion proteins in an N-terminal His6x-Tag/thrombin/T7-Tag configuration was amplified by PCR using the primers His-T7-tag-For and His-T7-tag-Rev (Appendix I). This DNA fragment was digested with *Hind*III and *Bam*HI and cloned between the corresponding sites of pYES2 to create pYES2- T7tag. The coding regions of Arabidopsis *LACS1*, *LACS2*, and *LACS3* as well as the coding regions of yeast *FAT1* and *FAA1* were amplified by PCR (see Appendix I for primer sequences) and then subcloned either between the *Bam*HI and *Not*I restriction sites (*LACS1*, *LACS2*, *LACS3,* and *FAT1*) or between the *Not*I and *Sph*I restriction sites (*FAA1*) of pYES2-His6x/T7tag. This allowed inducible expression by galactose in yeast of N-terminal tagged proteins that could be detected in Western blots using anti-T7 epitope antibodies.

2.3.3 Yeast strains and growth media

The 'wild-type' yeast strain YB332 (MATa ura3-52 his3Δ200 ade2-101 lys2-801 leu2-3112) and mutants in this background, fat1Δ (LS2020, MATa ura3-52 his3Δ200 ade2-101 lys2-801 leu2-3112 fat1::KanR) and *faa1Δfaa4Δ* (YB525, MATa *ura3-52 leu2-3,112 his3*D-*200 ade2-101 lys2-801 faa1*Δ*::HIS3 faa4*Δ*::LYS2*), were generously supplied by Dr. Paul Black (University of Nebraska-Lincoln). Another 'wild-type' yeast strain BY4742 (*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) and a mutant in this background, *fat1Δ* (BY4742 *fat1::kanMX*), were obtained from Open Biosytems (Huntsville, AL, USA). Plasmids were introduced into yeast using the lithium acetate method (Gietz & Schiestl, 2007). Transformed strains were maintained on synthetic media containing 0.67% yeast nitrogen base (without amino acids), 2% sugar (dextrose or galactose), and appropriate amino acid supplements. Cultures for all experiments were grown at 30°C with shaking at 250 rpm.

2.3.4 Heterologous protein expression in yeast monitored by Western blotting

After growth in galactose media for approximately 16 hours, cells were harvested and lysed using the post-alkaline method. which disrupts cells by incubation in NaOH followed by boiling in SDS-PAGE loading buffer (Kushnirov, 2000). This method was found to be faster and more consistent than methods that disrupt cells by physical means, such as glass bead beating. A phosphate-containing loading buffer (62.5 mM sodium phosphate buffer pH 7.0, 10% glycerol, 2% SDS, 5% β-mercaptoethanol, 0.0001% bromophenol blue) was used since it eliminated spontaneous proteolytic cleavage, which was observed when boiling in a Tris-based buffer. Proteins from the

yeast cell extracts were separated by SDS-PAGE and then transferred to a nitrocellulose membrane. The T7-epitope tagged proteins were detected with a mouse-raised α-T7 IgG primary antibody (EMD4Biosciences, Gibbstown, NJ, USA) and a sheep-raised α-mouse IgG peroxidase-conjugated secondary antibody (GE Healthcare, Buckinghamshire, England). The probed Western blots were imaged using the ECL-Advance western blotting detection kit (GE healthcare, Buckinghamshire, England) and a chemiluminescent imaging system.

2.3.5 Yeast fatty acid analysis

Yeast cells were inoculated into 50 ml of glucose-containing media and grown for 24 hours. The media was then swapped for galactose-containing media to induce ACS gene expression from the plasmids, and the cultures grown for an additional two days. Three ml of culture were harvested by centrifugation and the cell pellet was resuspended in 3 ml of 1N methanolic-HCl. 10 µg of heptadecanoic acid was added as an internal standard, and the mixture was incubated at 80°C for 90 minutes to transmethylate the fatty acids. One ml of 0.9% NaCl was added to decreased solubility of fatty acid methyl esters and promote their partitioning into the organic phase. Fatty acid methyl esters were extracted twice with 1 ml hexane. The hexane extractions were pooled and dried under a stream of nitrogen gas. Lipids were dissolved in 100 µl chloroform for analysis by reverse-phase gas chromatography using a flame ionization detector. One µl samples were injected into a HP-1 column (15m x 0.25mm inner diameter, 0.25µm thick film of dimethylpolysiloxane as stationary phase) at a 5:1 split ratio. Helium was used as a carrier gas, which flowed

at a constant rate of 2ml/min. The temperature program used was as follows: 150°C for 1 min, increased at 8°C/min up to 300°C, and 300°C for 10 min (for a total run time of 29.75 min). Fatty acid methyl-ester peaks were identified by comparison of retention times to standards, and quantified by comparison of peak areas to the internal standard. Percent complementation was calculated with the following formula: $(\mu g_{mutant} - \mu g_{proteinX})/(\mu g_{mutant} - \mu g_{WT}) * 100\%$.

2.3.6 Growth kinetics of yeast in the presence of oleate and/or cerulenin

Growth media supplemented with oleate and/or cerulenin was prepared using galactose media. Media was supplemented with 100µM oleate (from a 100mM stock dissolved in ethanol) and/or 45µM cerulenin (from a 45mM stock dissolved in acetone). 0.05% tergitol was included to solubilize the oleate. Yeast cultures were initially grown in glucose-containing media, and then inoculated into galactose- and supplement-containing media at an OD_{600} of 0.05. OD_{600} measurements were taken at the indicated times. For complementation of the *faa1Δfaa4Δ* double mutant, the wild type YB332 and isogenic *faa1Δfaa4Δ* strains were used, while for complementation of the *fat1Δ* mutant, the wild-type BY4742 and isogenic *fat1Δ* strains were used.

2.3.7 Fatty acid uptake monitored using the fluorescent fatty acid C_1-BODIPY-C_{12} and BODIPY-C_{16}

Cells were cultured in glucose-containing media for 24 hours followed by 16 hours of growth in galactose media. $1x10^6$ cells were harvested and resuspended in

50µl phosphate buffered saline plus EDTA (PBSE: 137mM NaCl , 2.7mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4, 50mM EDTA, pH 7.4). The assay was initiated by adding 50 µl PBSE supplemented with either 4,4-difluoro-5-methyl-4-bora-3a,4a-diaza-*s*-indacene-3-dodecanoic acid or 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-*s*-indacene-3-hexadecanoic acid (C_1-BODIPY-C_{12} or BODIPY-C_{16}, respectively, Molecular Probes, Carlsbad, CA, USA) and bovine serum albumin (BSA) to give final concentrations of 5µM and 2.5µM, respectively. BSA is required as it binds the hydrophobic BODIPY dyes, making them available to the yeast cells. The 2:1 molar ratio of BODIPY:BSA was determined empirically. The reactions were incubated for 10 minutes at room temperature. The assay was terminated by the addition of 1.5 ml ice-cold PBSE + 100µM BSA and placing the reactions on ice. Cells were harvested and washed twice more with 1.5 ml PBSE + 100µM BSA, washed once with 1.5 ml PBSE lacking BSA, and finally resuspended in 100 µl PBSE. The high level of BSA in these washes was to scavenge and sequester any remaining BODIPY molecules. Quantitative fluorescence and absorbance measurements were taken using a fluorescence plate reader (FLUOstar Optima, BMG Labtech, Offenburg, Germany) equipped with 485nm excitation and 520nm emission filters. OD_{600} was measured first on cell suspensions directly. Prior to fluorescence measurements, 100 µL of 4.2mM trypan blue (dissolved in PBSE and filtered) was added to reduce background fluorescence. These cells were also visualized by confocal microscopy (LSM 510 Meta, Carl Zeiss MicroImaging, Inc., Thornwood, New York). Samples were excited with a 488nm argon laser, and a 505-550nm emission filter was used. Microscope settings were optimized using the brightest sample, and were unchanged for the

remaining samples. To normalize values of C_1-BODIPY-C_{12} transport based on protein expression level, band intensities from the Western blot were quantified using digital imaging software (AlphaImager), and then divided by either the FAT1p or FAA1p band intensity, depending on the genetic background, to give a correction factor. Fluorescence values were scaled such that a mutant carrying empty vector (non-complemented) has a value of zero and a mutant complemented with the corresponding yeast protein (*i.e. fat1Δ*::FAT1p or *faa1Δfaa4Δ*::FAA1p) has a value of one. Dividing the scaled fluorescence value by the protein expression correction factor gives a relative indication of a given protein's ability to facilitate uptake of C_1-BODIPY-C_{12} taking into account the protein expression level.

2.4 Results

2.4.1 Comparison of Arabidopsis LACS1-3 amino acid sequences to yeast FAA1p and FAT1p

A previous study examining the predicted amino acid sequences of yeast and mammalian LACS proteins (typically termed ACSL proteins in the mammalian literature) and FATP proteins identified two highly conserved sequence motifs (Steinberg *et al.*, 2000). The first, an AMP binding motif found in both protein families, is also present in Arabidopsis LACS1, LACS2 and LACS3 (Shockey *et al.,* 2003). While no functional role for the second motif has been demonstrated, it has been used to distinguish LACS/ACSL proteins from VLACS/FATP proteins, as each protein family has a different consensus sequence for this region (Steinberg *et al.*, 2000). A multiple sequence alignment focusing on this region clearly shows high

sequence similarity between LACS1-3 and FAA1, with FAT1 having a distinct motif sequence (Figure 2.1A). Furthermore, a phylogram of proteins from several species separates LACS/ACSL and FATP proteins into distinct clades, with LACS1-3 being grouped with LACS/ACSL proteins (Figure 2.1B).

2.4.2 Heterologous expression of LACS1-3, FAT1p, and FAA1p in fat1Δ and faa1Δfaa4Δ yeast cells

The coding regions for LACS1-3, FAT1p, and FAA1p were cloned into a yeast expression vector downstream of the yeast *GAL1* promoter, allowing protein expression to be induced by the presence of galactose. Following growth in galactose media, heterologous protein expression was monitored in both *fat1Δ* and *faa1Δfaa4Δ* genetic backgrounds. The plant and yeast proteins contained an N-terminal T7 epitope tag that was detected using anti-T7 epitope antibody in Western blot experiments (Figure 2.2). All five proteins were successfully expressed in both genetic backgrounds, although the three Arabidopsis proteins were consistently expressed at higher levels than either of the two yeast proteins, especially FAT1p. All five proteins migrated at the expected sizes relative to molecular weight standards.

2.4.3 Genetic complementation analysis of very-long-chain acyl-CoA synthetase and fatty acid uptake defects of yeast fat1Δ with Arabidopsis LACS1-3

The lipid profile of wild type yeast cells has only one predominant VLCFA species, C26:0. The *fat1*Δ strain is characterized by accumulation of C20-C24 saturated fatty acids, as well as an increase in C26:0 levels (Watkins *et al*., 1998; Choi & Martin, 1999). This accumulation is relatively subtle for C20:0 and C26:0, showing accumulations of

3.8 and 2.2 fold increases over wild-type, respectively, but is much more pronounced for C22:0 and C24:0 species, showing 12.3 and 16.9 fold increases over wild-type, respectively (Figure 2.3). We tested the effect on VLCFA levels when Arabidopsis LACS1-3, yeast FAT1p, or yeast FAA1p were expressed in the *fat1*Δ strain (Figure 2.3). As expected, expression of yeast FAT1p in the *fat1Δ* background returned levels of C20:0-C26:0 VLCFAs to near wild-type levels, with VLCFA levels, showing 86%-113% complementation. Expression of FAA1p reduced the levels of C20:0 and C22:0 (88% and 55% complementation, respectively), but had no effect on C24:0 or C26:0 fatty acids (Figure 2.3). Like FAT1p, expression of LACS1 reduced levels of all VLCFAs to near wild-type levels (69%-102% complementation) (Figure 2.3). LACS3 expression had a similar, but less pronounced effect on all VLCFAs (44%- 67% complementation), while LACS2 had an effect only on C20:0 and C24:0 (39%-65% complementation).

A second phenotype of the *fat1Δ* mutant is a deficiency in fatty acid uptake. Cerulenin strongly inhibits *de novo* fatty acid synthesis, rendering cells dependent on the import of exogenous LCFA (*e.g.* oleate) for survival. Both FAT1p and FAA1p/FAA4p are needed for uptake of exogenous fatty acids; that is, both *fat1Δ* and *faa1Δfaa4Δ* mutants are unable to be rescued from cerulenin treatment by supplementation with oleate (Faergeman *et al*., 2001; Zou *et al*., 2003) While we observed this phenotype in the *faa1Δfaa4Δ* mutant derived from the parental strain YB332, the *fat1Δ* mutant in this parental background showed similar growth rates to wild type. However, we did observe this phenotype when using a *fat1Δ* mutant in a different parental background (BY4742). While the genetic basis for this difference

between strains is not known, additional differences were observed between the two strains: YB332 and its derivatives display a flocculation phenotype as well as a slightly slower growth rate compared to BY4742. It should be noted that in the literature, reports of the fatty acid uptake phenotype of the *fat1Δ* mutant either used a parental genetic background other than YB332 (Faergeman *et al*., 2001; Watkins *et al*., 1998; Dirusso *et al.*, 2000) or used *fat1Δ* in conjunction with *faa1Δ* (Zou *et al*., 2003; Dirusso *et al.*, 2005). Since we wished to clearly distinguish the effects of FAT1p and FAA1p/FAA4p, we opted to use BY4742 strains (wild-type and *fat1Δ*) for this experiment.

We tested if expression of LACS1, LACS2 or LACS3 restored growth in liquid media containing cerulenin and oleate to either the *fat1Δ* or the *faa1Δfaa4Δ* strains. All strains tested were viable when grown in galactose media supplemented with 100μM oleate alone, and each displayed similar growth kinetics (Figure 2.4A and B). When treated with 45μM cerulenin, growth was arrested in all strains (Figure 2.4C and D). As expected, when grown in media supplemented with both 100μM oleate and 45μM cerulenin, growth was restored in wild type cells, in *fat1Δ* cells expressing FAT1p, and in *faa1Δfaa4Δ* cells expressing FAA1p, albeit to lower levels than with cultures grown without cerulenin (Figure 2.4E and F). Growth was not restored in cells of either mutant carrying an empty vector, in *fat1Δ* expressing FAA1p, or *faa1Δfaa4Δ* expressing FAT1p. Consistent with a previous report (Shockey *et al*., 2002), expression of LACS1, LACS2, or LACS3 also restored growth to the *faa1Δfaa4Δ* mutant (Figure 2.4F). Similarly, growth of the *fat1Δ* strain was restored by expression of LACS1, LACS2, or LACS3 (Figure 2.4E).

2.4.4 Import patterns of fluorescent lipid analogues in fat1Δ and faa1Δfaa4Δ yeast cells heterologously expressing LACS1-3, Fat1p, or Faa1p

We further monitored the fatty acid import activities of the yeast strains using fluorescent lipid analogues. C_1-BODIPY-C_{12} mimics a LCFA (C18), whereas BODIPY-C_{16} mimics a VLCFA (C22). Fatty acid uptake in yeast requires both FAT1p and a LACS activity (either FAA1p or FAA4p). Therefore, the effects of expressing LACS1-3, FAT1p, and FAA1p were examined in both the *fat1Δ* and *faa1Δfaa4Δ* genetic backgrounds.

Using the LCFA analogue, C_1-BODIPY-C_{12}, quantitative fluorescence measurements showed a 3.7-fold greater signal for wild type cells than for either mutant (Figure 2.5A), consistent with an inability of these mutants to import exogenous fatty acids. Fluorescence signals of *fat1Δ* expressing LACS1, LACS2, or LACS3 were dramatically higher than the mutant (21.2, 9.7, and 18.4 fold increases, respectively) and even wild type cells, whereas *fat1Δ* expressing FAT1p had a similar fluorescence signal to that of wild type (Figure 2.5A). *fat1Δ* cells expressing FAA1p did not show a significant increase in signal over the mutant. In the *faa1Δfaa4Δ* mutant, expression of LACS1, LACS2, and LACS3 also increased the fluorescence signal, although to a more modest level (4.5, 4.0, 3.7 fold increases, respectively). Although *faa1Δfaa4Δ* cells expressing FAA1p showed only a small increase in signal (2.5 fold), this difference was found to be significant using a student's t-test (p-value = 0.002). Expression of FAT1p in *faa1Δfaa4Δ* cells had no effect on the fluorescence signal.

These results were corroborated when fluorescent signals from these cells were observed by confocal microscopy (Figure 2.5B). By examining the fluorescence signal overlaid on the differential interference contrast image, it is clear that the increased fluorescence of LACS1-3 expressing cells was intracellular, rather than associated with the cell wall, or floating in the media. The C_1-BODIPY-C_{12} taken up by wild type cells showed a relatively uniform distribution within the cells, perhaps having been incorporated into the plasma membrane, or bound by cytosolic fatty acid binding proteins (Figure 2.5B). The C_1-BODIPY-C_{12} distribution for cells expressing proteins off a plasmid was notably different. These cells showed areas of localized fluorescence intensity, perhaps within lipid bodies, and significant cell to cell variation (Figure 2.5B).

Using the VLCFA analogue, BODIPY-C_{16}, the wild type and *fat1Δ* cells did not show a significant difference from each other (Figure 2.6A). Neither cell type appeared to uptake BODIPY-C_{16} (Figure 2.6B). When LACS1, LACS2, LACS3, FAT1p, or FAA1p were expressed in *fat1Δ*, only LACS1 led to a significant increase in fluorescence signal (3.1 fold) over the mutant (Figure 2.6A). Confocal microscopy revealed fluorescent signal only within the LACS1-expressing cells (Figure 2.6B). As before, the BODIPY-C_{16} appeared in localized spots, possibly lipid bodies, rather than uniformly distributed. When *faa1Δfaa4Δ* cells were assayed for uptake of the VLCFA analogue (BODIPY-C_{16}), considerable background fluorescence was observed in the mutant alone and there was immense variability in the fluorescent measurements with the mutants expressing the ACSs (data not shown), thus preventing logical interpretation of these data.

Although expression of the Arabidopsis LACS proteins led to much greater C_1-BODIPY-C_{12} uptake than did expression of the yeast proteins, it is noted that the Arabidopsis proteins accumulate to higher levels in yeast, up to 32-fold higher (Figure 2.2). When the import values are corrected for protein levels, it was observed that all three Arabidopsis LACS proteins are roughly as efficient as FAA1p in facilitating uptake of C_1-BODIPY-C_{12} into the *faa1Δfaa4Δ* strain (Table 2.1). For import into the *fat1Δ* strain, LACS1 and LACS2 are ~17% as efficient as FAT1p, whereas LACS 3 is ~40% as efficient (Table 2.1). The cells assayed for import of BODIPY-C_{16} were subject to the same differences in expression levels, but in this case only LACS1 promoted an appreciable increase in fluorescence (Figure 2.6).

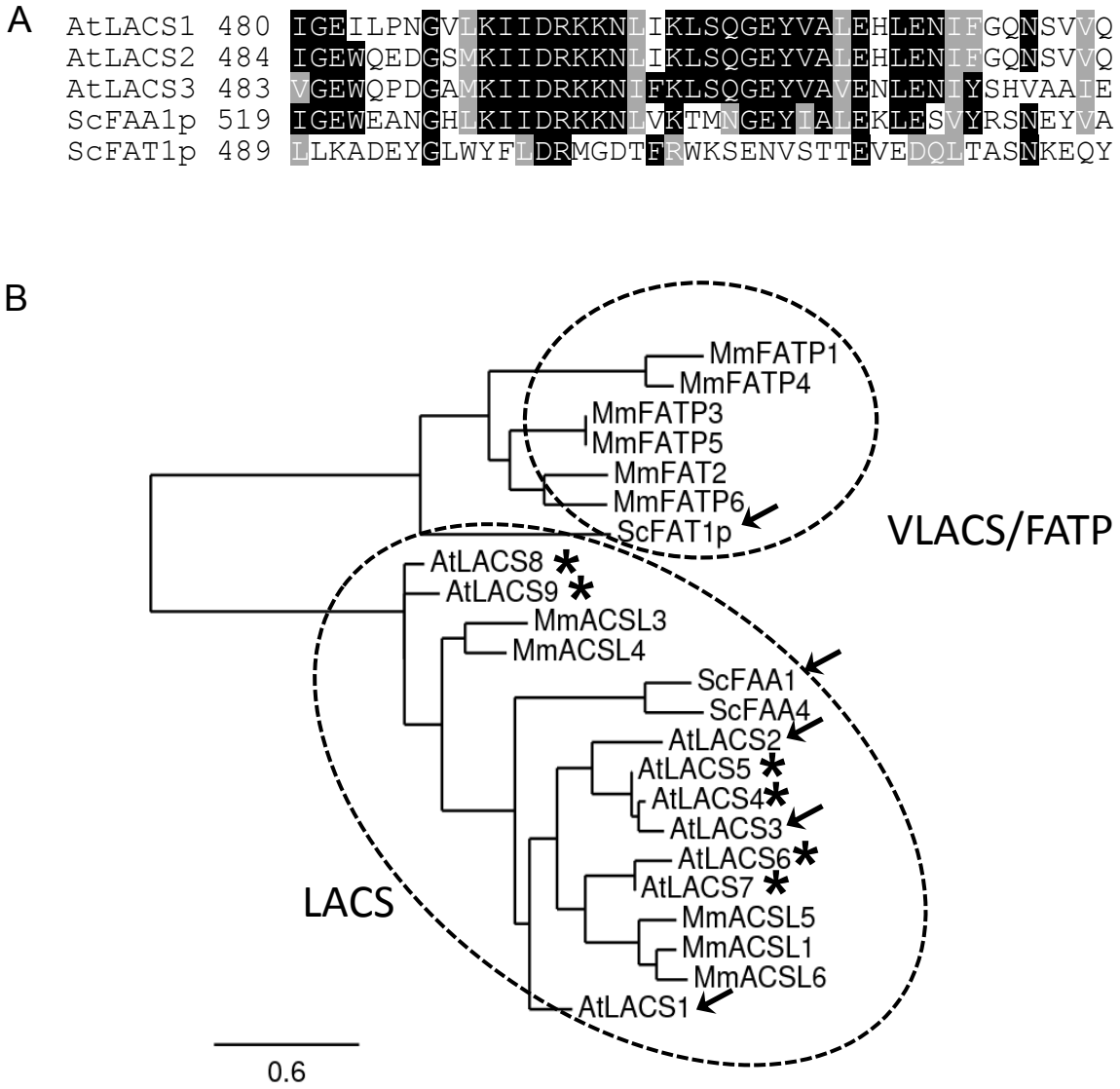


Figure 2.1 Sequence relationships between long-chain and very-long-chain acyl-CoA synthetases from Arabidopsis, yeast and mouse. (A) Multiple sequence alignment of acyl-CoA synthetase (ACS) motif II from Arabidopsis LACS1-3, yeast FAT1p, and yeast FAA1p. This motif, identified by Steinberg *et al.* (2000), differentiates the long-chain ACS family (*e.g.* FAA1p) from the very-long-chain ACS family (*e.g.* FAT1p). The position of the first amino acid of the motif within the respective open reading frame is noted at the left of each sequence. (B) Phylogram of LACS/ACSL/FAA and FATP/ACSVL proteins from Arabidopsis (At), yeast (Sc), and mouse (Mm). The complete protein sequences were aligned using MUSCLE and then evolutionary relationships inferred using the maximum likelihood method. The proteins characterized in this study are identified with arrows. Other Arabidopsis LACS's are identified with asterisks.

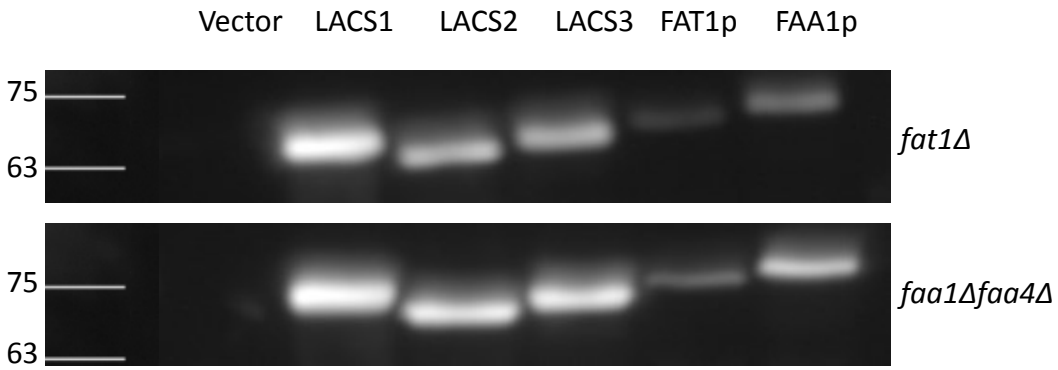


Figure 2.2 Expression of Arabidopsis LACS1-3, yeast FAT1p, and yeast FAA1p in *S. cerevisiae* strains *fat1Δ* and *faa1Δfaa4Δ*. Transformants of the strains were cultured in galactose to induce protein expression. Western blots of total cell lysates were performed using anti-T7 antibody to detect the amino-terminal T7 epitope in the protein fusions. The empty vector control was pYES2-T7tag. The positions of protein size markers (in kDa) are indicated at the left.

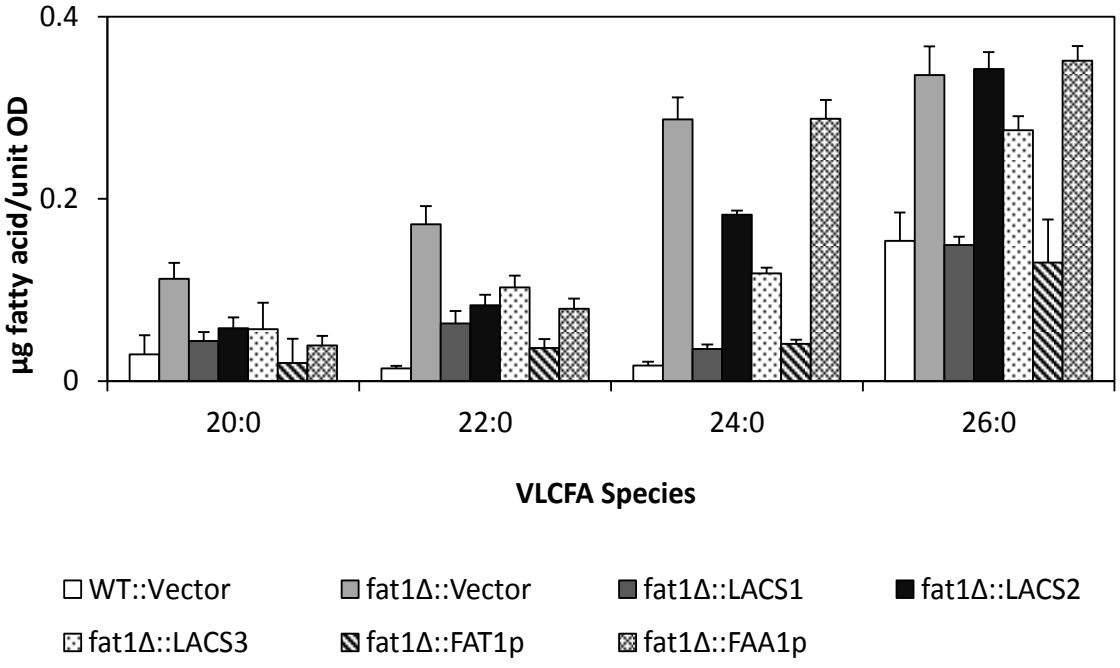


Figure 2.3 Elevated amounts of intracellular very-long-chain fatty acids due to deletion of the yeast *FAT1* gene are reduced by heterologous expression of Arabidopsis LACS1-3. Transformants of the strains were cultured in galactose to induce protein expression. Total fatty acids were extracted and quantified as their methyl esters (µg of fatty acid per unit OD of yeast). Each bar represents the average of three biological replicates, and error bars show standard deviation.

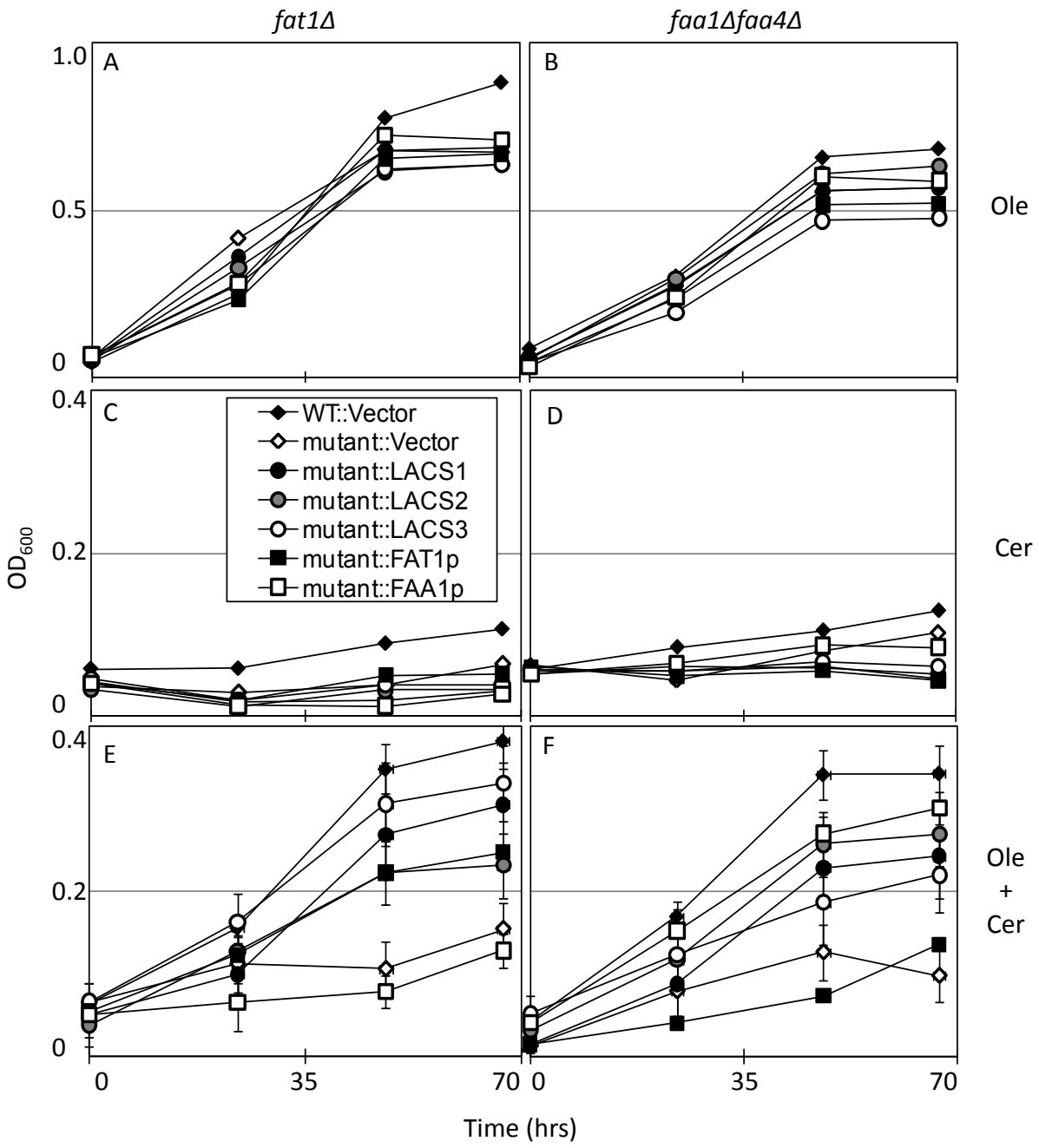


Figure 2.4 Growth impairment phenotype of *fat1Δ* (A, C, and E) and *faa1Δfaa4Δ* (B, D, and F) mutants due to a defect in fatty acid import is rescued by heterologous expression of Arabidopsis LACS1-3. The symbols representing each strain are depicted in the inset of panel C. The growth kinetics of each strain were analyzed under inducing conditions in the presence of oleate (Ole) alone (A and B), cerulenin (Cer) alone (C and D), and oleate plus cerulenin (Ole + Cer) (E and F). n=3 for each strain in panels A-D, and n=6 for each strain in panels E and F. Errors bars in panels E and F show standard deviations.

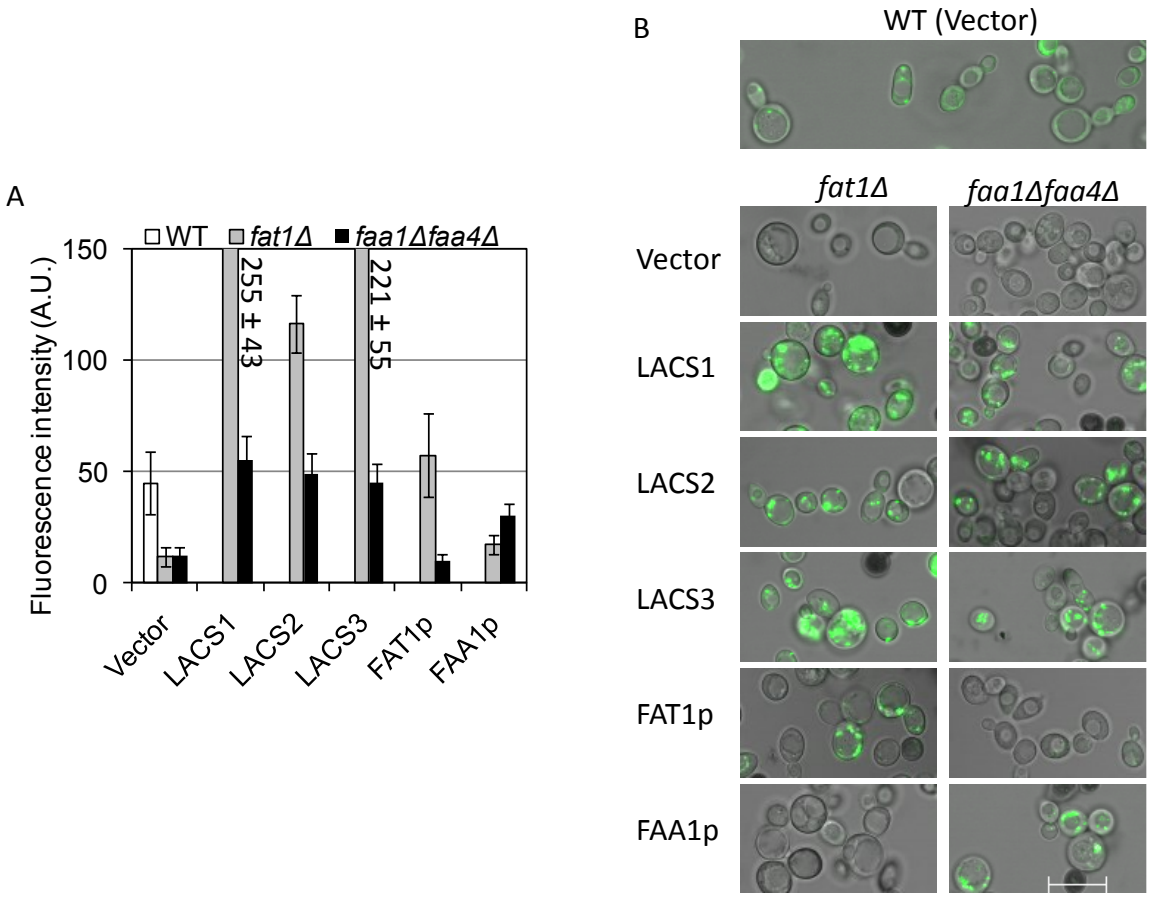


Figure 2.5 Import of the long-chain fatty acid fluorescent analogue C_1-BODIPY-C_{12} in yeast strains expressing Arabidopsis LACS1-3. (A) Quantification of C_1-BODIPY-C_{12} imported into yeast strains is indicated on the x-axis. Fluorescence intensity of extensively washed cells was measured using a fluorescence plate reader. n=4 and the error bars show standard deviations. (B) Representative images of C_1-BODIPY-C_{12} import into the same yeast strains were monitored by confocal microscopy. All pictures were taken at the same magnification; scale bar shows 10μm.

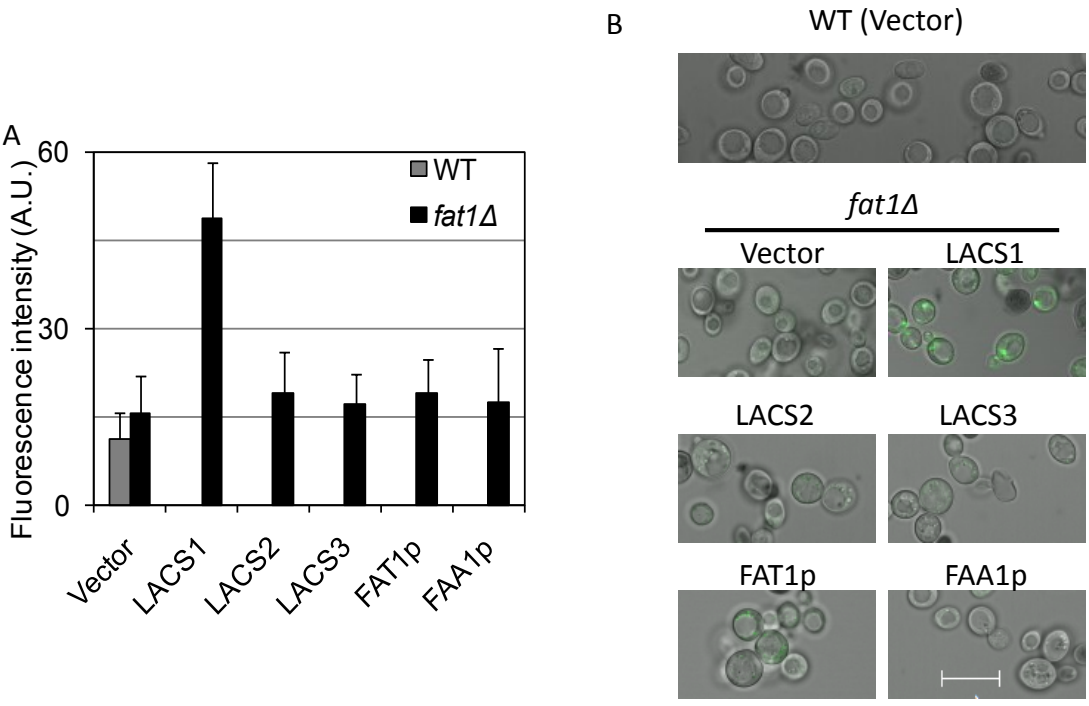


Figure 2.6 Import of the very-long-chain fatty acid fluorescent analogue BODIPY-C_{16} in yeast strains expressing Arabidopsis LACS1-3. (A) Quantification of BODIPY-C_{16} imported into yeast strains is indicated on the x-axis. Fluorescence intensity of extensively washed cells was measured using a fluorescence plate reader. n=4 and error bars show standard deviations. (B) Representative images of BODIPY-C_{16} import into the same yeast strains monitored by confocal microscopy. All pictures were taken at the same magnification; scale bar shows 10μm.

Table 2.1 Quantitative transport values of C_1-BODIPY-C_{12} into yeast cells, corrected for protein expression levels.

Protein	**Relative Transport (A.U.) ± S.D.**	
	fat1Δ	***faa1Δfaa4Δ***
LACS1	0.166 ± 0.029	0.897 ± 0.223
LACS2	0.168 ± 0.020	1.015 ± 0.261
LACS3	0.396 ± 0.103	0.961 ± 0.262
FAT1p	1.000 ± 0.419	-0.577 ± 0.813
FAA1p	0.035 ± 0.028	1.000 ± 0.308

2.5 Discussion

Arabidopsis LACS1-3 display broad substrate specificities, ranging from saturated and monounsaturated C16, C18 and C20 fatty acids (Shockey *et al*., 2002) to saturated very-long-chain fatty acids (Lü *et al*., 2009). For example, a comparison of the substrate preferences *in vitro* of LACS1 using saturated free fatty acids with chain lengths from C16 to C30 demonstrated that it had highest activity for C30 fatty acid, with C16 fatty acid being next in preference (Lü *et al*., 2009). This specificity is consistent with the *lacs1* mutant phenotype, which displays elevated C30 free fatty acids in cuticular wax and decreased levels of C16 cutin monomers (Lü *et al*., 2009). In contrast, yeast and mammalian LACS/ACSL enzymes, which are related in amino acid sequence to plant LACSs (Figure 2.1B), have very low activities towards saturated fatty acids with chain lengths greater than C18 (Knoll *et al.*, 1994; Soupene & Kuypers, 2008). In yeast and mammals, enzymes with VLACS activity are grouped into a distinct phylogenetic clade from LACS/ACSL enzymes (Figure 2.1B). Some VLACS proteins also play a second role, facilitating the uptake of exogenous long-chain fatty acids, which is why they are often termed fatty acid transport proteins (Schaffer & Lodish, 1994; Faergeman *et al*., 1997). Here we provide evidence that Arabidopsis LACS1-3 also have this transport functionality often associated with VLACS proteins.

To study whether LACS1-3 have the same dual functionality of yeast and mammalian FATPs, we tested for genetic complementation of two separate phenotypes of the *fat1Δ* mutant. Loss of VLACS activity prevents this mutant from catabolizing C20:0-C26:0 VLCFAs by β-oxidation, thereby causing these VLCFAs to

accumulate intracellularly (Watkins *et al.,* 1998; Choi & Martin, 1999). The full or partial reduction of VLCFA levels in the *fat1Δ* mutant due to expression of LACS1-3 demonstrates that these proteins are active VLACSs in the context of a heterologous yeast expression system. It should be noted, however, that the levels of C26:0 are not reduced in the *fat1Δ* mutant expressing LACS2, which suggests that this protein is not significantly active towards this substrate. This is in contrast to a previous report that LACS2 has significant activity towards C26:0, to a degree similar to LACS1 (Lü *et al*., 2009). This discrepancy may be due to the different experimental systems: Lü *et al*. (2009) assayed *E. coli* expressed proteins *in vitro*, whereas this report uses an *in vivo* yeast system.

Loss of FATP functionality renders *fat1Δ* cells unable to import exogenous fatty acids. When yeast cells are treated with cerulenin, *de novo* fatty acid synthesis is inhibited, leaving the cells reliant on import of exogenous fatty acids for growth. Since the LACS activity of FAA1p and FAA4p are also required for this process, we tested the effects of expression of LACS1-3 in both the *fat1Δ* mutant and the *faa1Δfaa4Δ* double mutant. Our results show that LACS1-3 are all able to complement this phenotype in both yeast strains tested, indicating that they can perform both roles involved in yeast fatty acid uptake. It is important to note that FAT1p only rescued *fat1Δ* and not *faa1Δfaa4Δ*, while FAA1p rescued only *faa1Δfaa4Δ* and not *fat1Δ*, supporting the model that two separate proteins, each with distinct activities, are required for fatty acid uptake.

The FATP functionality of LACS1-3 was further investigated using fluorescent fatty acid analogues. C_1-BODIPY-C_{12} mimics a C18 free fatty acid, and therefore it is

not surprising that assaying for import of this molecule closely mirrored the growth complementation results. This experiment allows a quantitative measure of fatty acid uptake. When expressed in *fat1Δ* mutant, LACS1-3 all show dramatically higher C_1-BODIPY-C_{12} import than wild type or FAT1p expressing mutant cells, and the levels of import are quite different between the three proteins, especially LACS1/3 versus LACS2. When corrected for protein expression levels, however, the normalized import levels of C_1-BODIPY-C_{12} of LACS1, LACS2, and LACS3 expressing yeast are similar to each other, but less than the normalized import levels of FAT1p expressing yeast. When expressed in the *faa1Δfaa4Δ* mutant, import driven by LACS1, 2, or 3 was similar to wild type levels and the import activities of the three Arabidopsis proteins were very similar to each other (including after correcting for protein expression levels). This highlights the different roles that the Arabidopsis LACS proteins are playing, depending on their heterologous context. In this study, it is difficult to speculate on the true "efficiency" of C_1-BODIPY-C_{12} transport since a proportion of highly expressed proteins may by misfolded or accumulating in various subcellular localizations, and thus not all protein may contribute to fatty acid uptake. Although quantitative conclusions regarding the rates or efficiencies of C_1-BODIPY-C_{12} transport into yeast cells cannot be made, the results presented clearly demonstrate that both a FATP and a LACS protein are required in this process, and that Arabidopsis LACS1, LACS2, and LACS3 are able to fulfill both roles.

Since the plant cuticle is comprised of both long-chain and very-long-chain fatty acid derivatives, we also asked if LACS1-3 could import BODIPY-C_{16}, a C22 fatty acid (VLCFA) mimic. In this case, only cells expressing LACS1 showed increased import.

However, the level of intracellular accumulation of this VLCFA mimic was notably lower than that of the LCFA analogue, C_1-BODIPY-C_{12}. While this may suggest that LACS1 is less able to transport VLCFAs than LCFAs, this result could possibly be due to use of this heterologous system. Since yeast cells do not typically import VLCFAs (note that neither wild-type nor FAT1p expressing cells showed increase fluorescence over the mutant), these cells may have intrinsic difficulties importing or storing the VLCFA analogue BODIPY-C_{16}. Taken together, the above results suggest that LACS1-3 are able to aid in transmembrane movement of fatty acids, but with distinct specificities.

The biochemical activities of fatty acid synthesis and fatty acid elongation are well characterized, and the enzymatic reactions that generate cutin monomers and wax molecules are becoming more clear (Samuels *et al*. 2008). However, details of how newly synthesized fatty acids are transported from the plastid to the ER for elongation, or how LCFAs and VLCFAs are channeled towards specific metabolic pathways (cuticle biosynthesis or otherwise) are notably lacking. After synthesis, LCFAs in the plastid are cleaved from their acyl carrier protein cofactor and a LACS activity associated with the plastid outer envelope is thought to rapidly generate cytosolic fatty acyl-CoAs (Koo *et al.*, 2004). In Arabidopsis, this role has been assigned to LACS9, since it is the major plastid localized LACS, and the *lacs9* mutant shows a 90% reduction in plastid-associated ACS activity (Schnurr *et al*., 2002). However, this deficiency surprisingly does not cause any apparent phenotype (Schnurr *et al*., 2002). To date, no other LACS have been localized to the plastid. This matter was further complicated when it was shown that the plastid-localized LACS9 functionally

overlaps with the ER-localized LACS1 in the generation of triacylglycerol in the seed (Zhao *et al*., 2010). This implies that simply activating fatty acids as they exit the plastid is not sufficient for effective metabolic utilization. A LACS at the target organelle may be needed for effective trafficking of specific fatty acids. It has also been suggested that LACS6 and LACS7 are involved in trafficking of fatty acids, in this case from triacylglycerol reserves into the peroxisome for catabolism in germinating seeds (Fulda *et al*., 2004). Cumulatively, various plant LACS proteins seem to be involved in metabolic channelling of cellular fatty acids, but it is apparent that our current understanding of this process is poor. Examples are seen in mammalian systems where the actions of LACS and/or FATP proteins direct fatty acid towards a specific metabolic fate. For example, overexpression of mouse FATP1 increased palmitate uptake and flux into β-oxidation, but had no effect on rates of incorporation into triacylglycerol (Holloway *et al.,* 2011), whereas down regulation of ACSL5 in rats affected flux of fatty acids into complex lipids, but not their rate of uptake (Bu & Mashek, 2010). It is likely that LACS1-3 act to traffic LCFAs/VLCFAs specifically towards cuticle biosynthesis by generating a pool of saturated acyl-CoAs of appropriate chain-lengths in the relevant subcellular location. Generation of spatially separated pools of phosphatidylinositol provides precedent of plant cells producing distinct lipid pools via differential localization and substrate specificities of the involved enzymes (Löfke *et al.*, 2008). Whether this phenomenon applies to cuticle biosynthesis awaits further study.

Before being metabolized into cuticle precursors, a large proportion of LCFAs are elongated by the fatty acid elongase, which occurs on the cytosolic face of the ER

(Joubès *et al*., 2008). Generation of cutin monomers and cuticular waxes is associated with the ER, but the specific sub-site is unknown. To our knowledge, the only data on this is an early report that at least one activity towards wax biosynthesis is physically separated from fatty acid elongation by a membrane, either within the ER lumen or an ER sub-domain (Bognar *et al.*, 1984). We propose that LACS1-3 are involved in trafficking of intracellular lipids towards cuticle biosynthesis through transmembrane movement of fatty acids. Notably, LACS1 plays the biggest role in generation of VLCFA derived cuticular waxes (Lü *et al*., 2009), and this protein is able to import a longer chain-length fatty acid analogue than LACS2 and LACS3 in yeast. Involvement of fatty acid transport functionality for LACS1, 2, or 3 would not preclude the acyl-CoA synthetase activities of these proteins also being utilized in cuticle biosynthesis. That is, both the FATP and VLACS activities of LACS1-3 may be utilized in cuticular lipid metabolism. This dual, yet separable, functionality is reminiscent of human FATP2, which acts as both a hepatic fatty acid transporter as well as a peroxisomal VLACS (Falcon *et al*., 2010).

Chapter 3: *ACYL-LIPID THIOESTERASE1-4* from *Arabidopsis thaliana* form a novel family of fatty acyl-acyl carrier protein thioesterases with divergent expression patterns and substrate specificities

3.1 Abstract

Hydrolysis of fatty acyl thioester bonds by thioesterases to produce free fatty acids is important for dictating the diversity of lipid metabolites produced in plants. We have characterized a four-member family of fatty acyl thioesterases from *Arabidopsis thaliana*, which we have called *ACYL LIPID THIOESTERASE1* (*ALT1*), *ALT2*, *ALT3*, and *ALT4*. The ALTs belong to the Hotdog fold superfamily of thioesterases. *ALT*-like genes are present in diverse plant taxa, including dicots, monocots, lycophytes, and microalgae. The four Arabidopsis *ALT* genes were found to have distinct gene expression profiles with respect to each other. *ALT1* was expressed specifically in stem epidermal cells and flower petals. *ALT2* was expressed specifically in root endodermal and peridermal cells as well as in stem lateral organ boundary cells. *ALT3* was ubiquitously expressed in aerial and root tissues and at much higher levels than the other *ALTs.* *ALT4* expression was restricted to anthers. All four proteins were localized in plastids via an N-terminal targeting sequence of about 48 amino acids. When expressed in *Escherichia coli*, the ALT proteins used endogenous fatty acyl-acyl carrier protein substrates to generate fatty acids that varied in chain length (C6-C18), degree of saturation (saturated and monounsaturated), and oxidation state (fully reduced and β-ketofatty acids). Despite their high amino acid sequence identities, each enzyme produced a different profile

of lipids in *E. coli*. The biological roles of these proteins are unknown, but they potentially generate volatile lipid metabolites that have previously not been reported in Arabidopsis.

3.2 Introduction

Fatty acids are generally activated for metabolism by thioester linkage to either acyl-carrier protein (ACP) or coenzyme A (CoA). The release of free fatty acids from their activating groups by fatty acyl-thioesterases (TEs) is important for regulating lipid metabolism and lipid trafficking. The most widely recognized function of fatty acyl-TEs in plants is to direct newly synthesized fatty acids towards specific metabolic fates. Fatty acid synthesis occurs in the plastids of plants. It is a cyclic pathway that iteratively elongates ACP-linked fatty acids by two carbons units (Li-Beisson *et al*., 2013). Nascent fatty acids are then partitioned into two cellular locations for further anabolism to produce a wide variety of lipids in the plastid-localized prokaryotic pathway and the endoplasmic reticulum (ER)-localized eukaryotic pathway (Benning, 2009). Fatty acyl groups enter the prokaryotic pathway via acyltransferases, which utilize metabolically active fatty acyl-ACPs for plastid membrane biogenesis. Fatty acyl-TEs direct substrates towards the eukaryotic pathway by cleaving the ACP group, thereby releasing metabolically inactive free fatty acids and allowing their export from the plastid.

Fatty acyl-TEs characterized to date mainly belong to one of two fold types: the α/β-hydrolase fold superfamily and the Hotdog fold superfamily (Cantu *et al.*, 2010). Both superfamilies are diverse groups of proteins found in all taxonomical

domains. Each fold type is characterized by its tertiary structure, and therefore proteins within the same superfamily can have highly divergent primary amino acid structures. α/β-hydrolase fold proteins are characterized by an eight-stranded β-sheet surrounded by α-helices (Nardini & Dijkstra, 1999). This superfamily is one of the largest known, with at least 30,000 known members, with widely ranging enzymatic and non-enzymatic roles (Lenfant *et al.,* 2013). The hallmark of a Hotdog fold domain is a 5-7 stranded antiparallel β-sheet wrapped around a central α-helix. Functionally characterized members of the Hotdog fold protein family have either thioesterase or dehydratase activities (Dillon & Bateman, 2004). Hotdog fold domains dimerize or form higher order oligomers *in vivo*, although there are examples of a single polypeptide containing two Hotdog fold domains (Pidugu *et al.*, 2009).

There are two types of highly conserved plant acyl-ACP TEs described to date that terminate fatty acid synthesis via hydrolysis of acyl-ACP thioester bonds: FATA- and FATB-type TEs. Both types are plastid localized enzymes with double Hotdog fold domains. FATA-type acyl-TEs are highly active towards monounsaturated acyl-ACP substrates, oleoyl-ACP (18:1Δ9) in particular, while FATB-type acyl-TEs act preferentially on saturated acyl-ACPs (Jones *et al.,* 1995). In most plant tissues, palmitic acid (16:0) is the major saturated fatty acid produced, although alternate FATB enzymes are expressed in seeds of certain plants (Voelker *et al*., 1992). Together, FATA and FATB TEs direct fatty acid precursors to the endoplasmic reticulum (ER) for generation of various lipid classes including phospholipids, triacylglycerols and sphingolipids, as well as specialized lipids such as cuticular waxes, cutin, and suberin.

In specific cases, highly specialized roles for plant fatty acyl-TEs have been demonstrated or proposed. The alternate FATB enzymes described above have a preference for saturated fatty acyl-ACPs of medium-chain lengths (C8-C14), leading to differing seed oil profiles between species (Voelker *et al*., 1992). In wild tomato, *Solanum habrocaites* subspecies *glabratum*, a fatty acyl-ACP TE called METHYL KETONE SYNTHASE2 (*Sh*MKS2) acts on C12-C16 β-ketoacyl-ACPs, which are intermediates of the fatty acid synthesis cycle (Ben-Israel *et al*., 2009; Yu *et al*., 2010). The resulting β-ketofatty acids are then converted to methyl ketones by the decarboxylase enzyme *Sh*MKS1 (Auldridge *et al.,* 2012). These methyl ketones accumulate in trichomes and are insecticides (Fridman *et al.,* 2005). Although *Sh*MKS2 is a single Hotdog fold protein, it is not similar at the primary amino acid level to the FATA or FATB double Hotdog fold proteins. Synthesis of cannabinoids by *Cannabis sativa* begins with a 6:0 fatty acid, which may be generated by a fatty acyl-ACP TE that interrupts fatty acid synthesis (Fellermeier *et al.,* 2001; Marks *et al.,* 2009), although no such enzyme has yet been identified. A portion of cuticular waxes on aerial surfaces are often very-long chain (*e.g.* 22:0-30:0) free fatty acids (Bernard & Joubès, 2012), presumably generated by an ER-localized very-long-chain fatty acyl-CoA TE (Lü *et al.,* 2009). Fatty acyl-CoA TEs have also been implicated in peroxisomal oxidation of fatty acids in mammals (Hunt *et al.,* 2012), which may apply to plants as well.

Four *Arabidopsis thaliana* fatty acyl-TEs have been experimentally characterized to date: two redundant FATA-type acyl-ACP TEs, one FATB-type acyl-ACP TE, and an acyl-CoA TE called ACH2. A *fata1 fata2* double mutant in which the

expression of each gene is reduced by approximately 50% has a modest decrease in lipid pools, most pronounced in seed triacylglycerols, and with a particular deficiency in 18:1 and 18:2 fatty acids (Moreno-Pérez *et al.,* 2012). A nearly complete knockdown mutant of *FATB* has a 20-50% reduction in saturated components of eukaryotic pathway derived products, accompanied by a dwarf phenotype and reduced viability (Bonaventure *et al.,* 2003). ACH2 is a peroxisomal acyl-CoA TE with *in vitro* activity towards 16:0-CoA, but its biological role is currently unknown (Tilton *et al.,* 2004).

Considering the diversity of plant lipid metabolic pathways, it is highly likely that Arabidopsis expresses many additional fatty acyl-TEs. Indeed, the Arabidopsis genome encodes several genes annotated as potential acyl-TEs. We report here on the characterization of a four member Arabidopsis gene family that we have called *ACYL-LIPID THIOESTERASE1 (ALT1)*, *ALT2*, *ALT3*, and *ALT4*, which are single Hotdog fold fatty acyl-ACP TEs related to wild tomato *ShMKS2*.

3.3 Materials and methods

3.3.1 Quantitative real-time PCR

Tissues were harvested from Col-0 wild-type plants, flash frozen in liquid N_2, and stored at -80°C. Most tissues (stems, rosette leaves, cauline leaves, flowers, and siliques) were harvested from 6-week-old plants, grown in long-day conditions (18 hours light/6 hours dark). Roots and whole seedlings were harvested after 15 days of growth on solid Murashige and Skoog (MS) media (no sucrose) and grown under

continuous light. RNA was extracted using a commercial kit and DNase treated (RNaqueous kit and Turbo DNase, Applied Biosystems, Carlsbad, CA, USA), following the manufacturer's instructions. RNA isolation aid (Applied Biosystems) was used during preparation of RNA from carbohydrate-rich silique tissues. Complimentary DNA (cDNA) was synthesized from 1 µg of total RNA with an 18-mer oligo-dT primer and SuperScript-reverse transcriptase III (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. Quantitative PCR was performed using a StepOnePlus instrument and PowerSyber qPCR mastermix (both from Applied Biosystems). Each 10 µl reaction contained 1x PowerSyber mix, 50 nM of each primer (Appendix I), and cDNA template (1% of the reverse transcriptase reaction). Transcripts were quantified using absolute standard curves, which were generated with a 10-fold dilution series using 10^{-5}-1 pg of plasmid DNA as template. These plasmids were generated by cloning either open reading frames plus 3' untranslated regions (for *ALT* genes) or the qPCR amplicon only (endogenous control genes) into pCR-Blunt II-TOPO (Invitrogen, Carlsbad, CA, USA). *ALT* transcripts were normalized with the endogenous controls *ACT2* (At1g49240), *eIF4A-1* (At3g13920), and *GAPC* (At3g04120). The transcript abundance of the three endogenous control genes were averaged, and *ALT1-4* expression is reported relative to this value. Each tissue had 2-3 biological replications, and every reaction was done as 3 technical replications. Due to the high level of sequence similarity between *ALT1-4*, we tested the specificity of qPCR primers using cloned transcripts as templates. After testing every primer and template combination, the highest level of cross hybridization was obtained with *ALT1* primers towards the *ALT2* template. This reaction had a Ct value of 11.7 cycles

above the reaction of *ALT1* primers towards the *ALT1* template. From this, we conclude that no more than 0.03% ($1/2^{11.7}$) of the amplification of a given target is due to cross hybridization of primers to different *ALT* genes.

3.3.2 Promoter::GUS analysis

Promoter sequences were amplified by PCR using bacterial artificial chromosome plasmids as templates (T9I1 for *ALT1* and *ALT2*; T22E19 for *ALT4*), with the primers ALT<1-4>-Prom-For-SalI and ALT<1-4>-Prom-Rev-BamHI (Appendix I). These fragments were inserted into pBI101.1 (Clontech, Mountain View, CA, USA) between the *Sal*I and *Bam*HI restriction sites and PCR amplified regions were verified by DNA sequencing. The resultant constructs contained ~1.8 kb of upstream sequence fused to the β-glucuronidase reporter gene. These constructs were transformed into *Agrobacterium tumefaciens* strain GV3101::pMP90 by electroporation. These Agrobacteria were used to transform Arabidopsis (Col-0) by the floral dip method (Clough & Bent, 1998). Tissues were prepared for histochemical staining by soaking tissues in heptane for 30 seconds, air drying for one minute, and then rinsing with GUS staining buffer (50mM NaH_2PO_4, 0.5mM $KFe(CN)_6$, 0.1% Triton X-100). The $KFe(CN)_6$ increases specificity of staining by preventing diffusion of the coloured precipitate. Staining was performed by incubation at 37°C for at least 2 hours and up to overnight in GUS staining buffer with 2 mM 5-bromo-4-chloro-indolyl-β-D-glucuronic acid (X-Gluc). Whole tissues were cleared by replacing the buffer with 70% ethanol and incubating at room temperature overnight, and then imaged with a Zeiss SteREO Discovery V20 stereomicroscope (Carl Zeiss MicroImaging, Inc., Thornwood, NY, USA).

Root cross sections were made from 10-day-old roots embedded in LR white resin (London Resin Company Ltd., London, UK) or from 5-week-old roots embedded in paraffin (Paraplast plus, Sigma-Aldrich, St. Louis, MO, USA) for examination of endodermal and peridermal cell layers, respectively. These sections were imaged with a Zeiss Axio Imager M2 compound microscope. To embed in LR white resin, roots were removed from the X-Gluc solution, soaked in 50% ethanol for 1 hour, and then fixed overnight in 25 mM $NaPO_4$ (pH 6.8) + 2% glutaraldehyde at 4^0C. Samples were dehydrated with an ethanol series (50%, 60%, 70%, 80%, 90%, 100%) for 2 times 30 min each, embedded in resin following the manufacturer's recommendations, cut to 10 µm sections with an ultra-microtome, and then imaged. To embed tissue in paraffin, tissues were fixed in FAA solution (3.7% formaldehyde, 50% ethanol, 5% acetic acid) for 3 hours, dehydrated with an ethanol series (70%, 80%, 90%, 95% for 30 min each), and then left in 100% ethanol overnight. Tissues were infiltrated by incubation in 25% *tert*-butanol in ethanol for 30 min, followed by 50% *tert*-butanol in ethanol for 30 min, then 100% *tert*-butanol overnight at 60^0C, overnight in 50% paraffin in *tert*-butanol at 60^0C, and finally in 100% paraffin at 60^0C for three days, changing the paraffin twice a day. The paraffin was allowed to solidify at 4^0C for a few hours, and then cut into 10 µm sections with a microtome. Sections were fixed to a glass slide at 42^0C overnight, dewaxed with *tert*-butanol, and then imaged.

3.3.3 In situ hybridization

An antisense probe specific for *ALT1* was amplified from the clone U14921 (obtained from the Arabidopsis Biological Resource Centre at Ohio State University)

containing the cDNA sequence for *ALT1* using the primers ALT1_Antisense_Fwd and ALT1_Antisense_Rev (Appendix I) that appends the T7 polymerase binding site to the 5' end of the template strand. As a negative control, a sense probe was also amplified from the same clone using the primers ALT1_sense_Fwd and ALT1_sense_Rev (Appendix I). Probe synthesis was performed as described previously (Hooker *et al.,* 2002). Stem tissue for *in situ* hybridization was harvested from 6-week-old wild-type plants and fixed in formal-acetic-alcohol (50% ethanol, 5% acetic acid, 3.7% paraformaldehyde). Tissues were embedded in paraffin as described previously (Samach *et al.,* 1997) and 8 µm cross sections of stem were sectioned with a rotary microtome and placed on frosted slides (Superfrost Plus, Fisher Scientific). The sections were bonded to the slides by incubating the slides at 42°C overnight on a slide warmer. Slide preparations were de-waxed and hybridized with probe as described previously (Samach *et al.,* 1997). The hybridized sections were imaged under 50% glycerol using a Zeiss Axio Imager M2 compound microscope.

3.3.4 Subcellular localization

Full length ALT1 and ALT2 open reading frames were amplified by PCR with the primers ALT<1-2>-ORF-For-XbaI and ALT<1-2>-ORF-RevNoStop-BamHI. Full length ALT3 and ALT4 open reading frames were amplified with the primers ALT<3-4>-ORF-For-BamHI and ALT<3-4>-ORF-RevNoStop-BamHI. The open reading frame for $ALT1_{1\text{-}47}$ was amplified with ALT1-ORF-For-XbaI and ALT1-ORF-Trunc-RevNoStop-BamHI, and the open reading frame for $ALT1_{48\text{-}189}$ was amplified with ALT1-ORF-Trunc-For-XbaI and ALT1-ORF-RevNoStop-BamHI (Appendix I). These fragments were ligated into pVKH18-35S::GFPC (Dean *et al.,* 2007), and verified by

DNA sequencing. The resultant constructs produced 35S promoter driven, C-terminal GFP fusion proteins (ALT-GFP). The constructs were transformed into *Agrobacteria* as described above, and the resultant strains were infiltrated into leaves of *Nicotiana benthamiana* as described previously (Sparkes *et al.,* 2006). GFP fused to the targeting peptides of *Nicotiana tabacum* rubisco small subunit and *Saccharomyces cerevisiae* cytochrome *c* oxidase IV, were used as localization markers for plastids and mitochondria, respectively, as described previously (Nelson *et al.,* 2007). Three days post-infiltration, leaves were imaged by confocal microscopy with a Zeiss LSM 510 Meta using a Plan Apochromat 63x objective (Carl Zeiss MicroImaging, Inc., Thornwood, NY). Samples were excited with a 488 nm argon laser and emission was recorded between 505 – 565 nm for GFP and LP 650 nm for chlorophyll autofluorescence.

3.3.5 Expression in E. coli

The coding sequences for ALT1-4 proteins lacking their N-terminal plastid targeting sequences were amplified using ALT<1-4>-ORF-Trunc-For-BamHI and ALT<1-4>-ORF-Rev-EcoRI primers (Appendix I), and inserted into pET-28a (EMD4Biosciences, Gibbstown, NJ, USA) using the indicated restriction sites. The coding sequence for a truncated FATB lacking the plastid targeting sequence (Mayer *et al.,* 2007) was amplified using FATB-ORF-trunc-For-SacI and FATB-ORF-Rev-HindIII, and ligated into the corresponding sites of pET28a. These constructs were transformed in *E. coli* K27, in which the long-chain acyl CoA synthetase gene, *FadD*, had been knocked out and the bacterial strain also modified for inducible expression of T7 RNA polymerase (Lü *et al.,* 2009). *E. coli* strains were grown in Lysogeny broth

(LB)-Miller at 37ºC to A_{600}=0.6, then protein expression was induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside and the cells grown for 16-18 hours at 18°C. Bacterial cells were pelleted by centrifugation at 4500 x g for five min in glass vials. One ml of the supernatant was mixed with 1 ml 20mM H_2SO_4 and incubated for 30 min at 75°C, then 30 min at 30°C. 50 µg tetracosane (24:0 alkane) was added as an internal standard, and then lipids were extracted into 500 µl hexane. One µl of the hexane fraction was used for reverse-phase gas chromatographic analysis. Lipids were quantified with a Varian GC3900 equipped with a HP-5MS column (30 m length, 0.25 mm inner diameter, 0.25 µm thick film of 5% Phenyl-95% dimethylpolysiloxane) and a flame ionization detector. Samples were injected in splitless mode. The carrier gas was helium with a constant flow rate of 2 mL/min. The column oven was initially at 50°C for 8 min, then ramped at 15°C/min up to 275°C, which was held for 5 min, giving a total run time of 28 min. Identification of novel products was performed using an Agilent 6850 gas chromatograph equipped with an HP-5MS column (same as above) and an Agilent 5975 mass spectrometric detector (70 eV, mass-to-charge ratio 50–750). The initial oven temperature of 33˚C was held for 5 min, then ramped at 2˚C/min up to 65˚C, which was held for 2 min, then ramped at 50˚C/min up to 320 ˚C, which was held for 1.9 min, giving a total run time of 30 min. The carrier gas was helium with a constant flow rate of 1.5mL/min.

3.3.6 Identification of novel compounds by mass spectrometry

All saturated FAs and MKs were identified by comparison of experimental spectra to those of commercial standards. Unsaturated FAs were also identified by comparison to standards, although the position of the double bond was not

investigated, but rather assumed based on the typical *E. coli* lipid profile. 11:1 and 13:1 MKs were identified by comparison to the spectra provided in Goh *et al.* (2012), and 15:1 MK was identified by the presence of the calculated molecular ion and the characteristic MK fragments (Appendix II), as well as the retention patterns of saturated versus unsaturated MKs.

3.3.7 In vitro thioesterase assay

The ALT1 catalytic mutant was made by site directed mutagenesis using overlap extension PCR. The pET28a::ALT1 clone described above was used as the template and the primers used were ALT1-ORF-Trunc-For-BamHI, ALT1-ORF-Rev-EcoRI, ALT1-mutant-DA-For, and ALT1-mutant-DA-Rev (Appendix I). This converted the aspartate residue at position 64 (of the full length protein) to alanine. This clone was transformed into the *E. coli* expression strain described above.

Bacterial strains carrying the wild-type and mutant *ALT*-containing plasmids were inoculated from overnight cultures into 250 ml of terrific broth (12 g/L tryptone, 24 g/L yeast extract, 0.4% glycerol, 89 mM K-phosphate buffer pH 6.2) and grown to A_{600}=0.4 and then protein expression was induced with 0.3 mM IPTG and cells were cultured at 10°C for approximately 40 hours. Cells were harvested by centrifugation, washed with phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4), and stored at -80°C until use. Proteins were purified using PrepEase Histidine-Tagged Protein Purification Midi Columns (Affymetrix, Santa Clara, CA, USA). Cell pellets resuspended in 5 ml lysis-binding buffer (50 mM Na phosphate, 300 mM NaCl, 10 mM imidazole, 0.5% TritonX-100, pH

8.0), and lysed by sonication. Lysates were clarified by centrifugation, and then loaded onto the commercial columns. Columns were washed 3 times with 2 ml lysis-binding buffer, and then histidine-tagged proteins eluted with 5 ml elution buffer (same composition as lysis-binding buffer, but with imidazole concentration raised to 250 mM). Eluates were dialyzed overnight against 5 L dialysis buffer (50mM Na phosphate, 500 mM NaCl, 1 M $(NH_4)_2SO_4$, 2 mM DTT, 15% glycerol, pH 6.8), divided into single use aliquots, flash frozen in liquid N_2, and stored at -80°C until use. Total protein concentrations were determined using the Bio-Rad protein assay kit (Bio-Rad, Hercules, CA, USA), using bovine serum albumin as a standard. Thioesterase concentrations were determined using Western blotting for the T7 epitope by comparing signal intensity to a commercial standard of known concentration. Purity of the ATL1, ALT1-D64A, and ALT2 preparations ranged from 10-12%, similar to that obtained previously for *Sh*MKS2 (Auldridge *et al.,* 2012). The ALT3 and ALT4 preparations had notably higher purity, 39% and 21%, respectively.

The radiolabeled [1-^{14}C]12:0-ACP substrate, prepared with recombinant spinach ACP, was generously provided by Dr. John Shanklin (Biosciences Department, Brookhaven National Laboratory). Assays were set up in 100 µl reactions containing 25 mM Na phosphate buffer, pH 8.0, 50 mM KCl, and 10 µl thioesterase preparation (diluted to about 40 ng TE/µl in dialysis buffer). The reaction was initiated by addition of 30 pmol acyl-ACP substrate, incubated at room temperature for 30 min, and then terminated by addition of 500 µl of 1M acetic acid in isopropanol. Free fatty acids were extracted 2 times with 500 µl hexane, pooled, and quantified by liquid scintillation counting. Control reactions in which 12:0-ACP substrate was extracted

directly, without incubation, gave a small background hydrolysis value that was subtracted from all the other reaction calculations.

3.4 Results

3.4.1 ALT1-4 form a highly related Arabidopsis gene family with homologues in many plant taxa.

The *ALT1-4* genes each consist of 5 exons and 4 introns and are present in pairs at two genomic loci on chromosome 1. *ALT1* (At1g35290) and *ALT2* (At1g35250) are separated by four genes, while *ALT3* (At1g68260) and *ALT4* (At1g68280) are separated by a single gene. The four *ALT* genes encode small proteins (188-192 amino acids) that are very similar to one another, sharing 66-83% amino acid sequence identity (76-88% similarity). These predicted proteins are 72-80% similar (50-62% identical) to wild tomato *Sh*MKS2. A multiple sequence alignment between ALT1-4 and *Sh*MKS2 illustrates the high sequence similarity between these proteins, especially in the C-terminal ~140 amino acids that corresponds to a single Hotdog fold domain (Figure 3.1). The amino acid sequence similarity of ALT1-4 in the Hotdog fold domain alone is 82-91% (72-89% identity). Although the plant protein sequences are very different from the bacterial enzyme, the experimentally characterized active site residues of the *E. coli* TE, YbgC, (aspartate, glycine, and valine in the DXXGXV motif) are conserved in the five plant proteins (Figure 3.1) (Benning *et al.,* 1998; Yu *et al.,* 2010). Each ALT protein contains a predicted plastid targeting sequence at the N-terminus. The N-terminal plastid targeting sequences are likely cleaved upon delivery of the proteins to their destination, as this is common for other plastid proteins.

A search of various genomes identified potential *ALT* gene homologues in a wide variety of plant species, but none outside the Plantae kingdom. These species include monocots, dicots, the lycophyte *Selaginella moellendorffii,* and the green microalgae *Chlamydomonas reinhardtii* and *Coccomyxa subellipsoidea*. The moss *Physcomitrella patens* does not contain an *ALT* homologue. None of the non-plant genomes searched contained homologues. The only ALT protein homologue to be characterized experimentally is *Sh*MKS2. Most plant genomes contain 4-5 homologues, although several contain only a single homologue (*e.g. Brassica rapa, Glycine max, Poplus trichocarpa*). The well characterized fatty acyl-ACP TEs FATA and FATB contain two sequential Hotdog fold domains, but these proteins are otherwise not similar to the ALTs. In a phylogenetic tree, the ALT proteins form a distinct clade from FATA and FATB proteins (Figure 3.2). Like in *Arabidopsis thaliana*, ALT-like genes are often found close to each other in the genome (*e.g.* the *Brachypodium distachyon* genes Bd5G17810, Bd5G17830, and Bd5G17840).

3.4.2 ALT1-4 have distinct gene expression patterns

The gene expression patterns of *ALT1-4* were examined by four complementary methods: analysis of publically available DNA microarray data, quantitative reverse transcriptase-PCR (qPCR), *in situ* transcript hybridization, and histochemical staining of promoter:GUS lines.

Examination of tissue-specific gene expression by qPCR revealed *ALT1* transcript levels were highest in stem and floral organs and at lower levels in siliques, whole seedlings, and cauline leaves (Figure 3.3). *ALT1* transcripts were not detected

in rosette leaves or in roots. Histochemical staining of *ALT1* promoter::GUS reporter lines confirmed expression in stems (Figure 3.4A). An *in situ* hybridization experiment showed that the *ALT1* transcript was restricted to epidermal cells in the stem (Figure 3.4B, C). There was GUS reporter gene expression in the top portion of the gynoecium (fused carpels) immediately beneath the stigma and in the petals (Figure 3.4D). These promoter::GUS lines also showed staining throughout the silique, with strongest staining at the abscission zone (Figure 3.4E). In cauline leaves, staining was observed in trichomes. When probing publically available DNA microarray data for genes that co-regulate with *ALT1*, we noted that several of the top ranked genes are involved in the biosynthesis of cuticular waxes by epidermal cells. These genes included *MAH1*, *CER4*, *WSD1*, *CER1*, and *CER3* (Table 3.1).

ALT2 was highly expressed in root and stem tissues, with lower expression in siliques and whole seedlings, as determined by tissue-specific qPCR (Figure 3.3). The magnitude of *ALT2* expression in qPCR experiments was comparable to that of *ALT1*. Staining of *ALT2* promoter::GUS roots was not continuous along the length of the root, but instead had a patchy pattern (Figure 3.4G, H). *ALT2* expression was limited to endodermal and peridermal cells in young and mature roots, respectively, as revealed in cross sections of stained roots (Figure 3.4K, L). GUS staining of siliques showed *ALT2* expression in the developing seeds, but not in the silique valves (Figure 3.4J). The stem expression of *ALT2* was localized to lateral organ boundaries (Figure 3.4I), and not along the stem length as was the case with *ALT1*. A cross section of a stem/petiole boundary showed darker GUS staining throughout all cell layers of the petiole, and lighter staining in the shoot, centered around the vascular cells (Figure

3.4M). *In silico* analysis of DNA microarray experiments identified several suberin biosynthetic genes as being co-expressed with *ALT2*. These included *FAR4*, *CYP86A1/HORST*, *CYP86B1/RALPH, ASFT,* and *GPAT5* (Table 3.1).

The tissue-specific qPCR experiments showed that *ALT3* was expressed in every tissue examined, both aerial and root (Figure 3.3). The transcript levels were much higher than *ALT1*, *ALT2*, or *ALT4.* These results are consistent with data from publically available DNA microarray experiments. Histochemical staining of *ALT3* promoter::GUS transgenic lines gave very different results between lines, and thus no informative data on the cell-specific expression pattern of *ALT3* could be generated using promoter-reporter gene fusions. It is possible that *cis*-acting DNA elements outside of the upstream ~1.8kbp region are important for conferring its gene expression pattern. No known lipid metabolic genes were found to have a high co-expression correlation with *ALT3*, likely due to its ubiquitous expression.

ALT4 was expressed in flowers, but the level of transcript was several orders of magnitude lower than the other genes tested as determined by qPCR (Figure 3.3). GUS staining showed specific expression in anthers of flowers (Figure 3.4N, O). An anther cross section revealed staining both in anther walls (endothecium) and in microspores (Figure 3.4P). This highly restricted expression of *ALT4* could explain the very low transcript abundance within the whole flower clusters used for our qPCR experiment. Co-expression analysis was not possible for *ALT4* because it is not represented on the 22K ATH1 gene chip.

3.4.3 ALT proteins localize to plastids

Subcellular localization prediction algorithms (Emanuelsson *et al.,* 2000; Nielsen *et al.,* 1997) indicated that ALT1-4 proteins contain N-terminal plastid targeting sequences of 47-51 residues. Although plastid localization was most likely for all four proteins, ALT2 and ALT4 also had high scores for mitochondrial targeting. We determined the subcellular localization of ALT1-4 by Agrobacterium-mediated transient expression of C-terminal green fluorescent protein (GFP) fusions (ALT-GFP) in leaf cells of *Nicotiana benthamiana* using confocal microscopy. GFP fusions of full-length ALT proteins all produced green fluorescent signal that co-localized with red plastid autofluorescence (Figure 3.5C-F), indicating plastid localization. The same fluorescence pattern was seen when the first 47 amino acids of ALT1 were fused to GFP, whereas a truncated ALT1 lacking these 47 residues was not plastid localized (Figure 3.5A-B). This indicates that the plastid targeting sequence of ALT1 is necessary and sufficient for plastid targeting of this protein, which is likely also the case for ALT2, ALT3, and ALT4.

3.4.4 ALT1-4 generate fatty acids and β-ketofatty acids when expressed in E. coli

We characterized the fatty acyl-TE activities of ALT1-4 by heterologous expression in *E. coli* using endogenous acyl-ACP substrates and analyzing lipid products secreted into the media. For these experiments, we used the *E. coli* mutant *fadD*, which lacks acyl-CoA synthetase activity (Overath *et al.,* 1969), causing free fatty acids to accumulate and get exported to the media, rather than being activated and degraded by β-oxidation. Due to the high sequence similarity between ALT1-4 and *Sh*MKS2, we analyzed not only unmodified fatty acids as potential enzymatic products, but also β-ketofatty acids. Since β-ketofatty acids are difficult to detect

directly, they were chemically decarboxylated to methyl ketones prior to extraction, as was done by Yu *et al.* (2010). The methyl ketones produced by this procedure are shortened by one carbon relative to the initial β-ketofatty acid.

Analysis of the medium of *E. coli* containing empty vector (negative control) revealed small amounts of saturated and unsaturated fatty acids ranging from 8-18 carbons (the most being 16:1 with 127.1 ng/OD unit) as well as traces of 13:0 and 15:1 methyl ketones (44.4 and 48.2 ng/OD unit, respectively) (Figures 3.6, 3.7). All heterologous proteins were successfully expressed and accumulated to similar levels, as observed by Western blotting (Figure 3.8). As a positive control, we expressed the well characterized Arabidopsis fatty acyl-ACP thioesterase FATB. Expression of FATB in *E. coli* resulted in large amounts of 16:1 fatty acid (5481.5 ng/OD unit), as previously reported (Mayer and Shanklin 2007). Expression of ALT1-4 each resulted in the production of fatty acids and/or β-ketofatty acids (the latter converted to methyl ketones for GC analysis) that were either absent or much less abundant in the empty vector negative control (Figures 3.6, 3.7). The structures of all these products were confirmed by GC-MS analyses (Appendix II). The major products of the ALT1-expressing *E. coli* strain was 12:0 (1250.3 ng/OD unit) with medium-to-long chain fatty acids 14:1, 14:0, 16:1, and 16:0 also being prominent (Figures 3.6, 3.7). ALT2 expression resulted in high levels of medium-chain methyl ketones 7:0 (1440.5 ng/OD unit) and 9:0 (445.3 ng/OD unit) and smaller levels of 8:0 fatty acid (Figures 3.6, 3.7). ALT3 expression generated predominantly 13:1 methyl ketone (891.5 ng/OD unit) along with smaller amounts of 7:0, 9:0, 11:0, 11:1, 13:0, and 15:0 methyl ketones. Fatty acid species were also produced by heterologous expression of ALT3,

the most prominent fatty acid being 16:1 (448.2 ng/OD unit) (Figures 3.6, 3.7). Finally, the ALT4-expressing strain generated high amounts of 6:0 (741.1 ng/OD unit) and 8:0 (1429.0 ng/OD unit) medium-chain fatty acids as well as small increases of 10:0 and 16:1 fatty acids and small amounts of 7:0, 13:0 and 15:1 methyl ketones (Figures 3.6, 3.7). For all proteins tested, both fatty acids and methyl ketones were generated, but with one of the two chemical classes usually dominating. Compounds of the non-dominant chemical class mirrored the relative levels of the corresponding products in the dominant class. For example, FATB produced mainly 16:1 and 14:0 fatty acids, with much smaller amounts of 15:1 and 13:0 methyl ketones, while ALT2 generated large amounts 7:0 methyl ketones, alongside smaller amounts of 8:0 fatty acids. The major products of ALT3 were 13:1, 11:0, and 13:0 methyl ketones, mirrored by 14:1, 12:0 and 14:0 fatty acids. ALT3 also produced notable amounts of 16:1 fatty acid (mirrored by a small amount of 15:1 methyl ketone), making it the only protein tested to have main products in both the fatty acid and methyl ketone classes.

In this experimental system, we were not able to observe β-ketofatty acids directly but rather assumed their presence based on detection of methyl ketones. To ensure that the methyl ketones detected were in fact chemically derived, rather than direct enzymatic products, extractions were performed without the heat and acid treatment used for decarboxylation of β-ketofatty acids. For this experiment, ALT3 was chosen as it generates both fatty acids and methyl ketones, as well as FATB because fatty acids are its only known biologically relevant products. As expected, this omission reduced the production of all methyl ketone species considerably, with no apparent correlation to chain length (75-95% reduction; Figure 3.9).

3.4.5 ALT1, 3, and 4 have lauroyl (12:0)-ACP thioesterase activity in vitro

To further examine the enzymatic activities of the ALT1-4, we prepared recombinant ALT proteins lacking the plastid targeting sequence (as described above) and assayed *in vitro* their activities towards 12:0-ACP (Figure 3.10). A background level of 12:0-ACP hydrolysis was observed in the absence of added protein (buffer alone). Inclusion of recombinant ALT1 increased the level of acyl-ACP hydrolysis above background levels (statistically significant, $p<0.005$). To further evaluate ALT1 activity and to eliminate any potential of co-eluting *E. coli* proteins contributing to acyl-ACP hydrolysis, we engineered a catalytically inactive ALT1 mutant to act as another negative control. This mutant was generated by replacing a conserved aspartate residue within the active site (position 64) to alanine (Figure 3.1). The corresponding residue in the *Pseudomonas sp.* enzyme, 4-hydroxybenzoyl-CoA thioesterase, was identified as an active site residue by its proximity to the thioester carbonyl group in the crystal structure and its conservation in other known TEs (Benning *et al.,* 1998). In agreement, the analogous aspartate-to-alanine substitution abolished *Sh*MKS2 thioesterase activity (Yu *et al.,* 2010). This ALT1 catalytically inactive mutant, ALT1-D64A, generated no fatty acids or β-ketofatty acids when expressed in the *E. coli fadD* expression strain (data not shown). In the *in vitro* thioesterase assay using 12:0-ACP substrate, ALT1-D64A did not increase hydrolysis above background levels. ALT3 and ALT4 preparations had relatively high 12:0-ACP thioesterase activities, while the ALT2 preparation had no activity above background.

Figure 3.1 Amino acid sequence alignment of ALT1-4 from *Arabidopsis thaliana* (At), MKS2 from the wild tomato species *Solanum habrochaitis* subspecies *glabratum* (Sh), and YbgC from *Escherichia coli* (Ec). The predicted plastid targeting sequences of the plant proteins are indicated with a dashed box. Each of the proteins contains a single Hotdog fold domain encompassing the rest of the predicted protein sequence. The experimentally characterized active site residues of the bacterial protein are marked with arrows. These active site residues are conserved in the plant proteins.

Figure 3.2 Phylogenetic tree of Arabidopsis ALT1-4 proteins, ALT-type proteins from a variety of plant species, and FATA- and FATB-type thioesterases from select plant species. ALT-type homologues were identified using BLASTP with ALT1 sequence as the query. The amino acid sequences were aligned with MUSCLE (Edgar 2004). The phylogenetic tree was generated using the maximum likelihood model (Guindon et al. 2010) and rendered with PhyloWidget (Jordan and Piel 2008). ALT1-4 are indicated with arrow heads, *Sh*MKS2 is indicated with a star, non-ALT-type *Arabidopsis thaliana* TEs are indicated with closed circles, and FATB TEs with specificity for shorter chain lengths are indicated with asterisks. *At- Arabidopsis thaliana, Al- Arabidopsis lyrata, Bd- Brachypodium distachyon, Br- Brassica rapa, Ch- Cuphea hookeriana, Cr- Chlamydomonas reinhardtii, Cs- Coccomyxa subellipsoidea, Gm- Glycine max, Mt- Medicago truncatula, Os- Oryza sativa, Pt- Poplus trichocarpa, Rc- Ricinus communis, Sb- Sorghum bicolour, Sh- Solanum habrochaites, Si- Setaria italica, Sl- Solanum lycopersicum, Sm-Selaginella moellendorffii, Uc- Umbelluria californica, Vv- Vinis vinifera, Zm- Zea mays.*

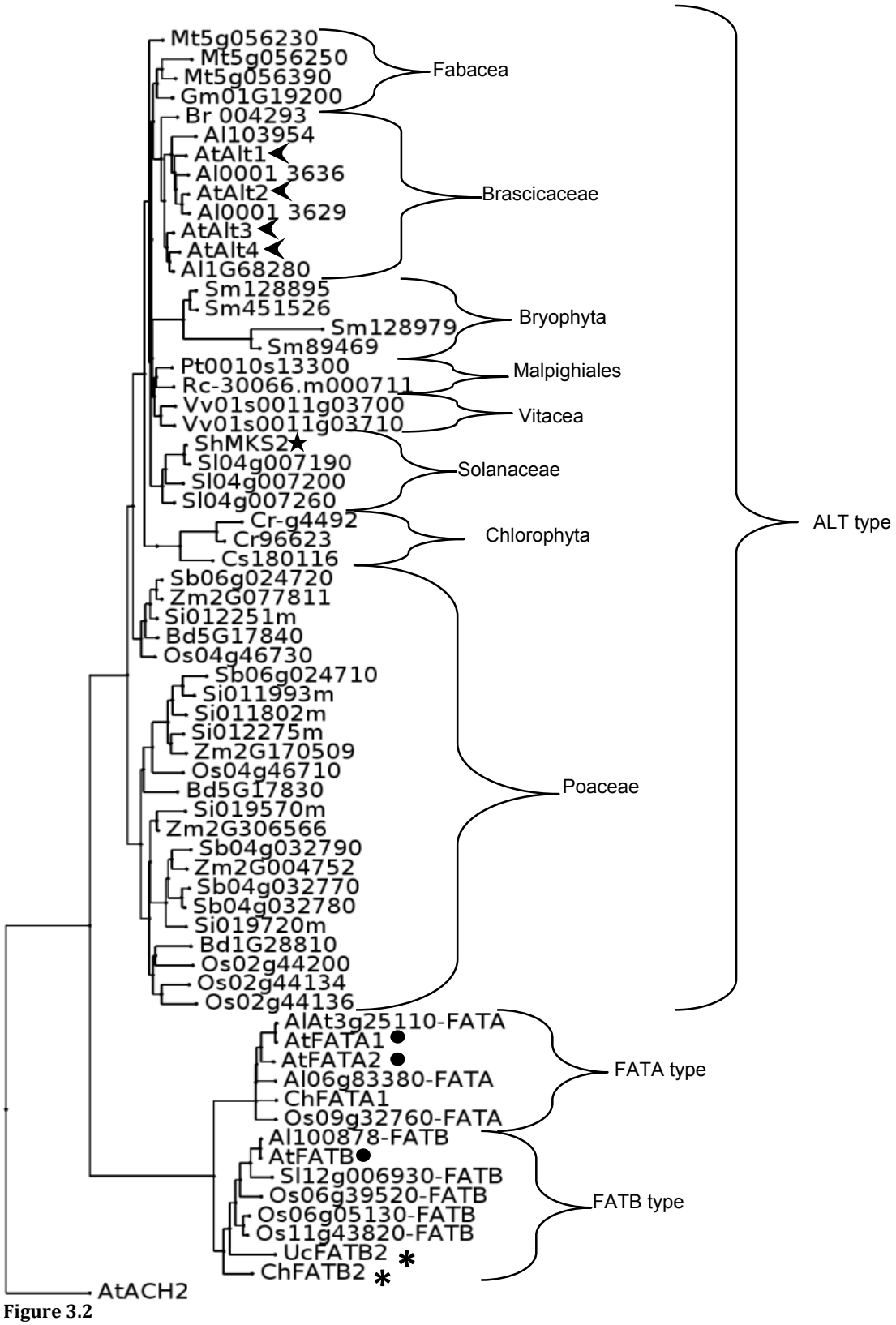

Mt5g056230
Mt5g056250
Mt5g056390
Gm01G19200
Br 004293
Al103954
AtAlt1
Al0001 3636
AtAlt2
Al0001 3629
AtAlt3
AtAlt4
Al1G68280
Sm128895
Sm451526
Sm128979
Sm89469
Pt0010s13300
Rc-30066.m000711
Vv01s0011g03700
Vv01s0011g03710
ShMKS2
Sl04g007190
Sl04g007200
Sl04g007260
Cr-g4492
Cr96623
Cs180116
Sb06g024720
Zm2G077811
Si012251m
Bd5G17840
Os04g46730
Sb06g024710
Si011993m
Si011802m
Si012275m
Zm2G170509
Os04g46710
Bd5G17830
Si019570m
Zm2G306566
Sb04g032790
Zm2G004752
Sb04g032770
Sb04g032780
Si019720m
Bd1G28810
Os02g44200
Os02g44134
Os02g44136
AlAt3g25110-FATA
AtFATA1
AtFATA2
Al06g83380-FATA
ChFATA1
Os09g32760-FATA
Al100878-FATB
AtFATB
Sl12g006930-FATB
Os06g39520-FATB
Os06g05130-FATB
Os11g43820-FATB
UcFATB2
ChFATB2
AtACH2
Fabacea
Brascicaceae
Bryophyta
Malpighiales
Vitacea
Solanaceae
Chlorophyta
Poaceae
ALT type
FATA type
FATB type

Figure 3.2

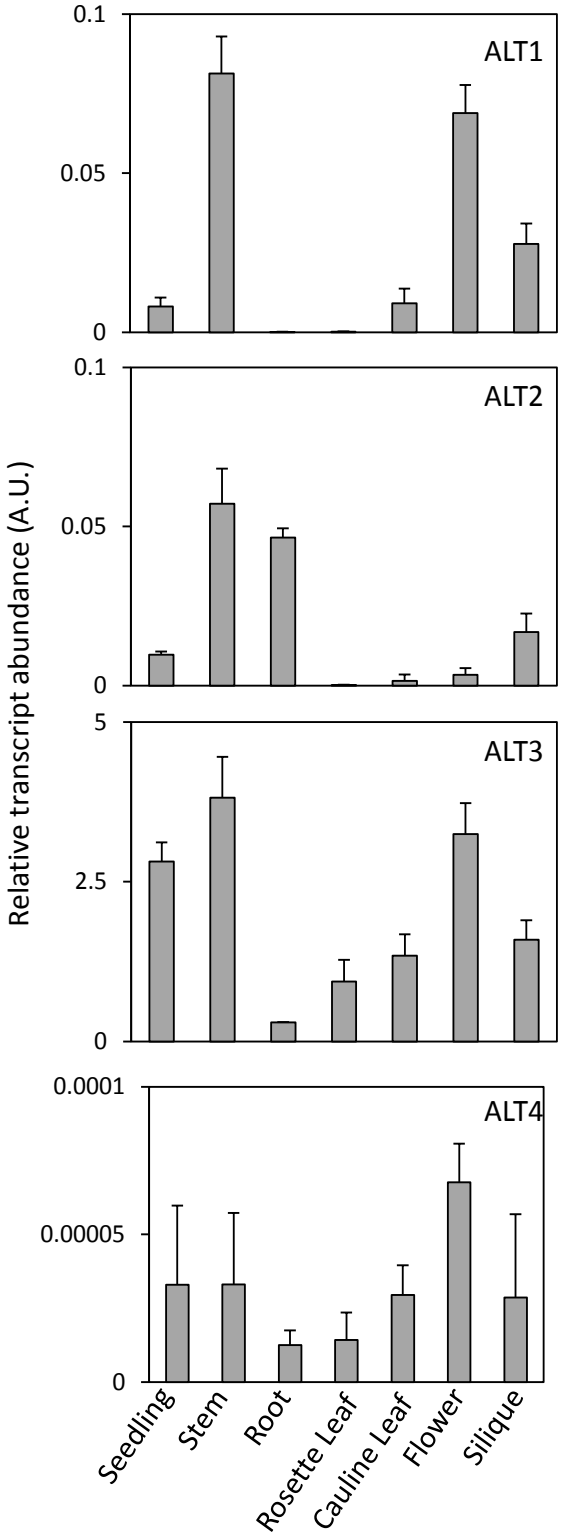


Figure 3.3 Tissue-specific gene expression patterns of Arabidopsis *ALT1-4* analyzed by quantitative reverse transcriptase-PCR. Transcript levels were quantified with absolute standard curves, and normalized with expression of the endogenous control genes *ACT2*, *eIFA4a*, and *GAPC.*

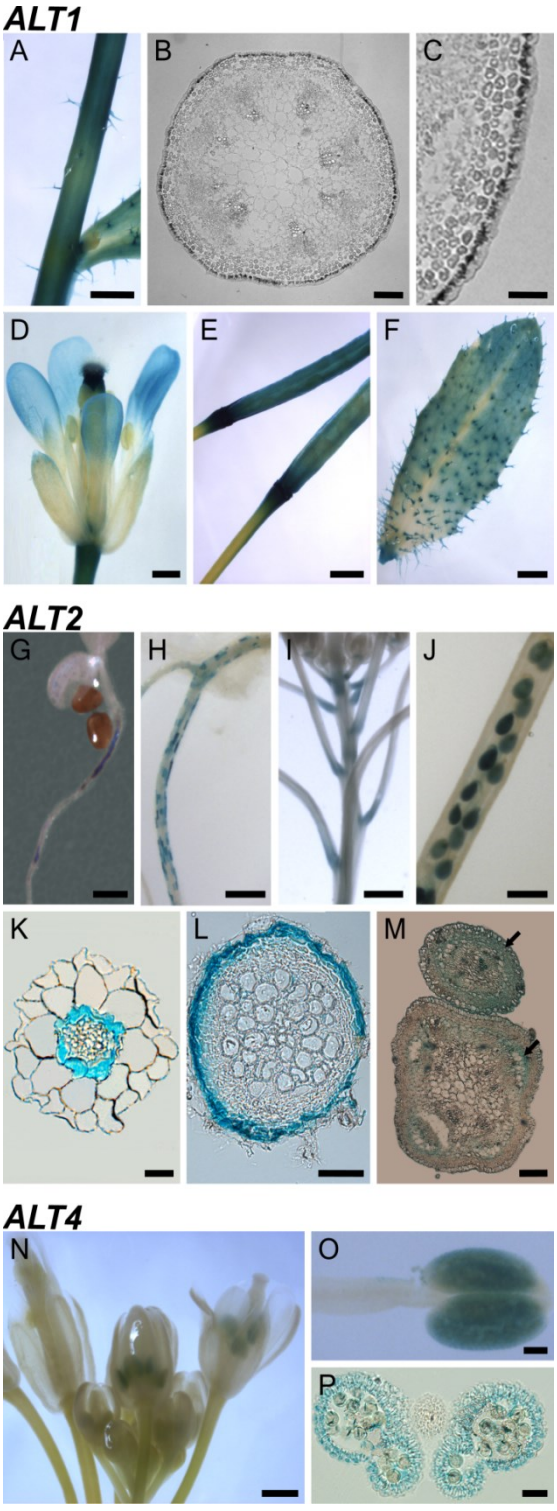


Figure 3.4 Expression patterns of *ALT1* (A-F), *ALT2* (G-M) and *ALT4* (N-P) examined by promoter::GUS analysis (A, D-P) and *in situ* hybridization (B, C). A, Stem; B-C, Stem cross section; D, Flower; E, Silique; F, Rosette leaf; G, Young root; H, Mature root; I, Boundary of shoot and petioles; J, Silique; K, Young root cross section; L, Mature root cross section; M, Shoot/petiole boundary cross section with arrows pointing to stained cells; N, Flower cluster; O, Stamen; P, Anther cross section. Scale bars: 2mm (A, E, F), 1mm (G, H, I, J, N), 0.5mm (D), 0.25mm (B, O), 0.1mm (C, P), 50µm (L), 25µm (K, M).

Table 3.1 List of cuticular wax- and suberin-associated genes that co-regulate with *ALT1* and *ALT2*, respectively.

Query	Co-regulated cuticular wax genes	R^2	Reference
	MAH1 (At1g57750)	0.712	Greer *et al.* 2007
	CER4 (At4g33790)	0.643	Rowland *et al.* 2006
	SHN2 (At5g25390)	0.595	Aharoni *et al.* 2004
ALT1	*WSD1* (At5g37300)	0.587	Li *et al.* 2008
	CER1 (At1g02205)	0.545	Aart *et al.*1995
	CER3 (At5g57800)	0.517	Kurata *et al.* 2003
	LACS1 (At2g47240)	0.512	Lü *et al.* 2009
Query	**Co-regulated suberin genes**	**R^2**	**Reference**
	FAR4 (At3g44540)	0.740	Domergue *et al.* 2010
	CYP86A1/HORST (At5g58860)	0.715	Höfer *et al.* 2008
ALT2	*CYP86B1/RALPH* (At5g23190)	0.709	Compagnon *et al.* 2009
	ASFT (At5g41040)	0.657	Molina *et al.* 2009
	GPAT5 (At3g11430)	0.607	Beisson *et al.* 2007
	FAR5 (At3g44550)	0.502	Domergue *et al.* 2010

In silico gene co-expression analysis was done using the Expression Angler tool at the Bio-Array Resource for Plant Biology (Toufighi *et al.* 2005). R^2 is the Pearson correlation coefficient that represents the similarity of the gene expression patterns over 392 DNA microarray experiments.

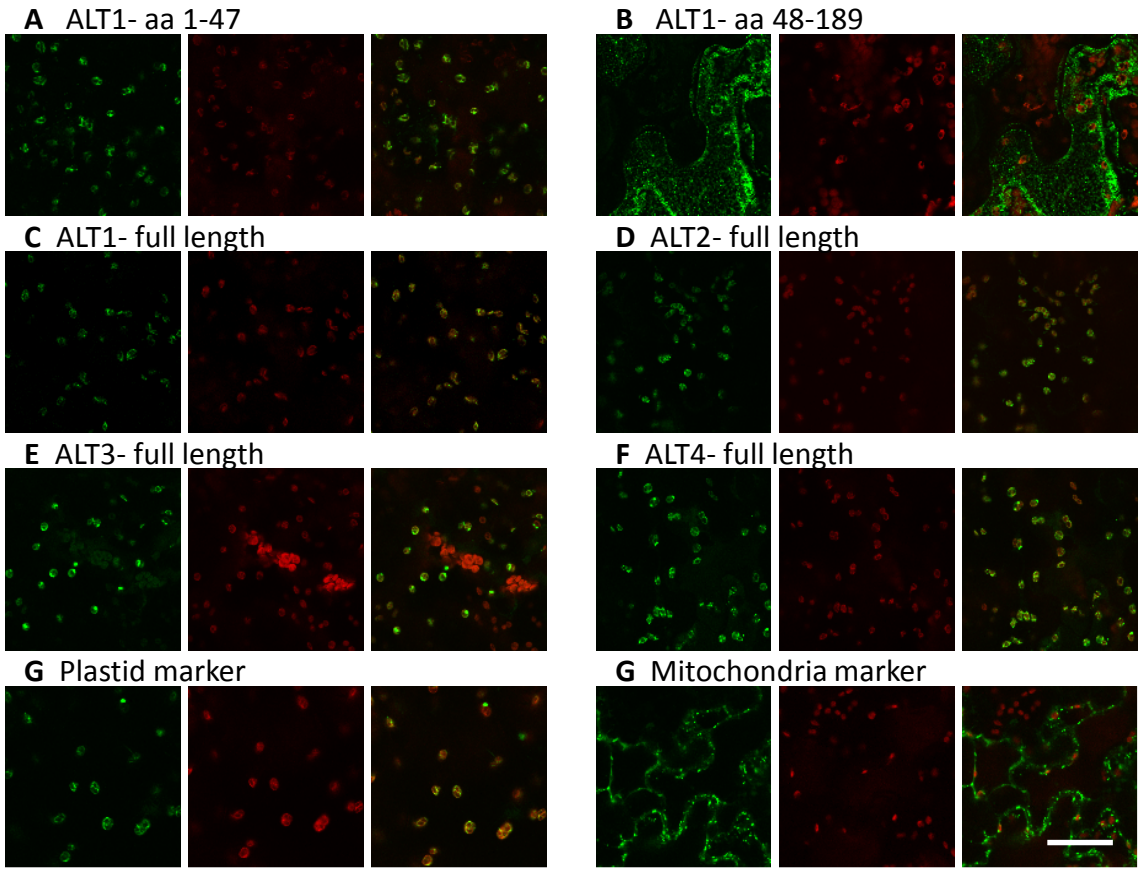


Figure 3.5 Subcellular localization of ALT-GFP fusions expressed transiently in leaves of *Nicotiana benthamiana* and observed with confocal microscopy. Fusions of full-length ALT1-4 were analyzed (C-F), as well as a fusion of only the predicted chloroplast targeting signal (CTP) of ALT1 (A) or a version of ALT1 lacking the predicted targeting signal (ΔCTP) (B). GFP signal is shown in green and plastid autofluoresence is in red. Plastid and mitochondrial organelle markers used were described previously (Nelson *et al*. 2007). The scale bar in panel F represents 50 μm and all panels are the same magnification.

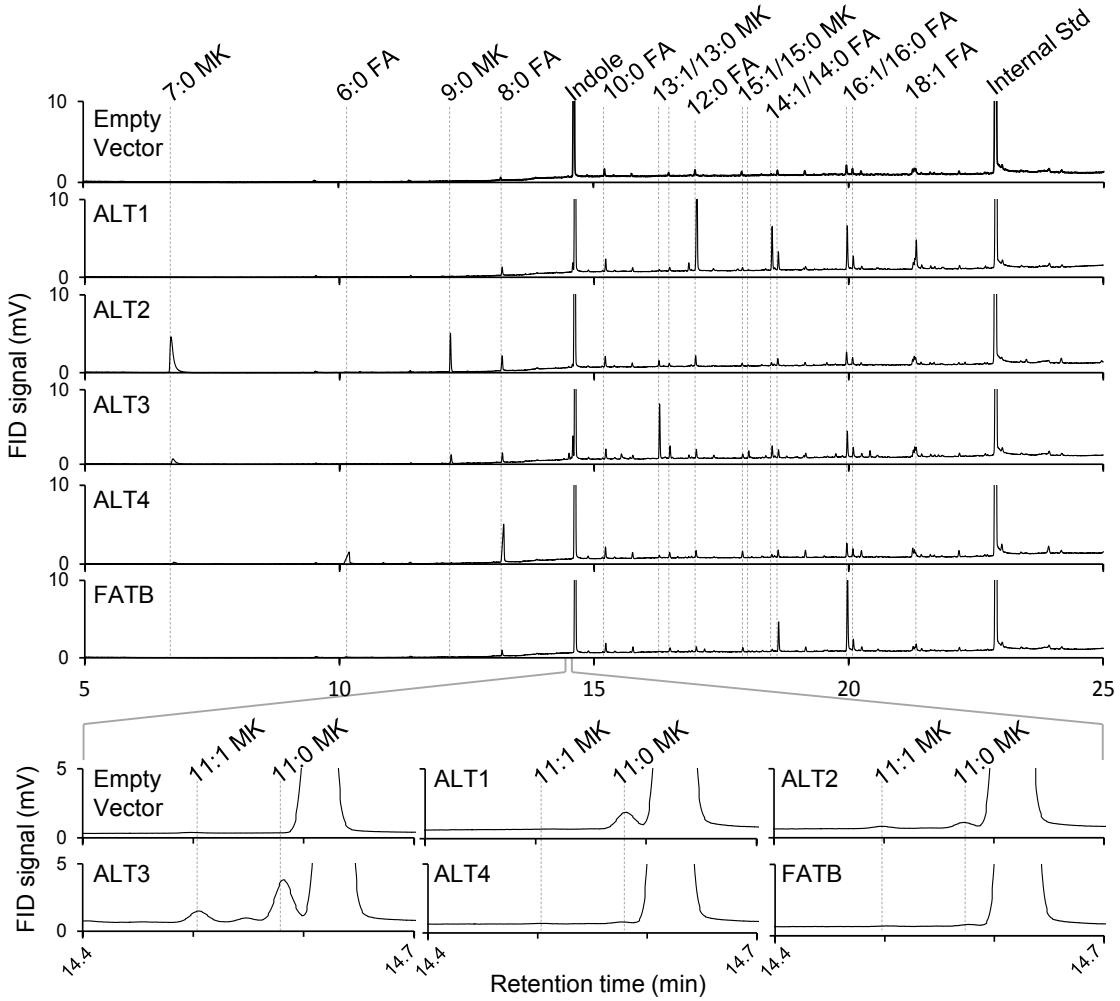


Figure 3.6 Gas chromatograph traces of lipids secreted into the media of *E. coli* cultures carrying an empty vector, or expressing ALT1-4, or FATB. The β-ketofatty acids produced were chemically decarboxylated by heat and acid for conversion to more stable methyl ketones prior to extraction and analysis by gas chromatography. Lipids were identified by GC-MS (see Appendix II) and are indicated above the traces (FA = fatty acid, MK = methyl ketone).

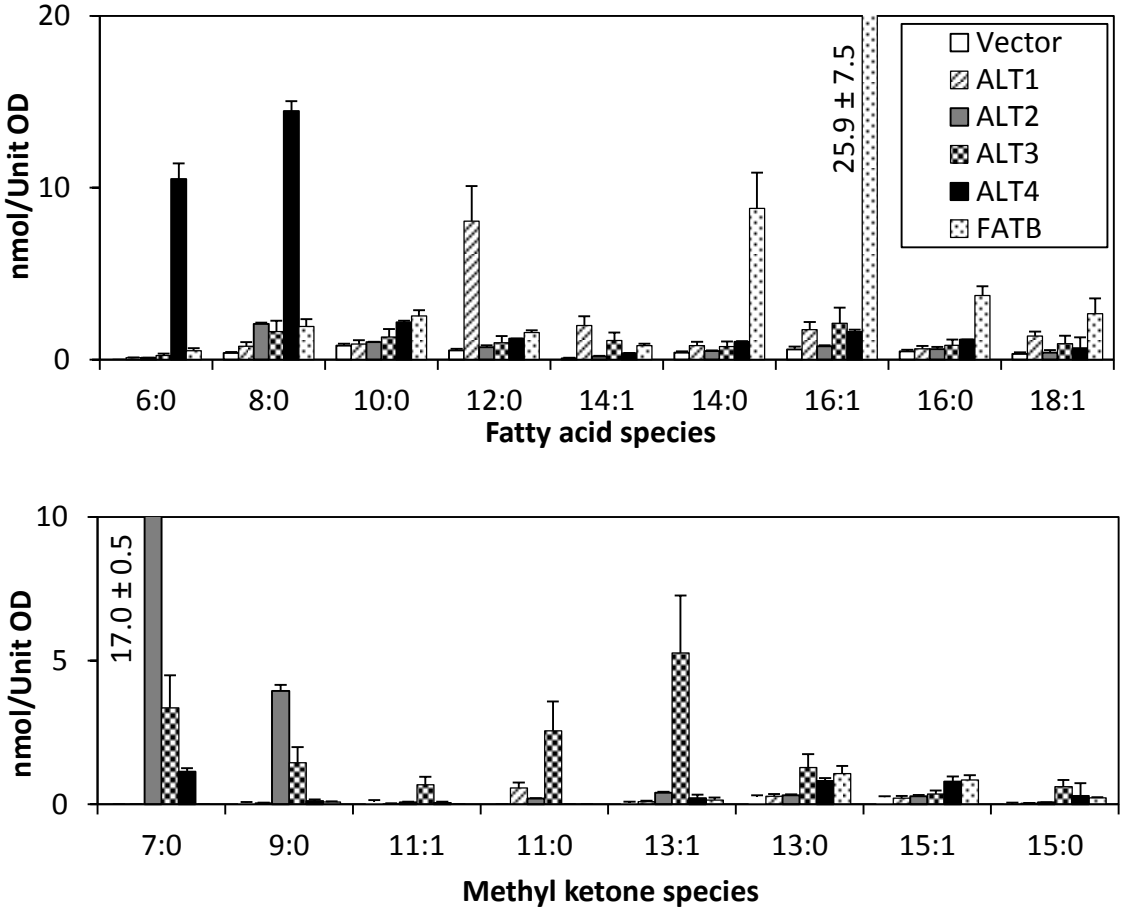


Figure 3.7 Quantification of lipids secreted into the media of *E. coli* cultures carrying an empty vector, or expressing ALT1-4, or FATB. A, fatty acids detected. B, methyl ketones detected, which were derived from β-ketofatty acids. The β-ketofatty acids produced by thioesterase activities were chemically decarboxylated by heat and acid for conversion to more stable methyl ketones prior to extraction and analysis by gas chromatography. Quantification was done by comparison of peak areas to that of a tetracosane internal standard using GC-FID.

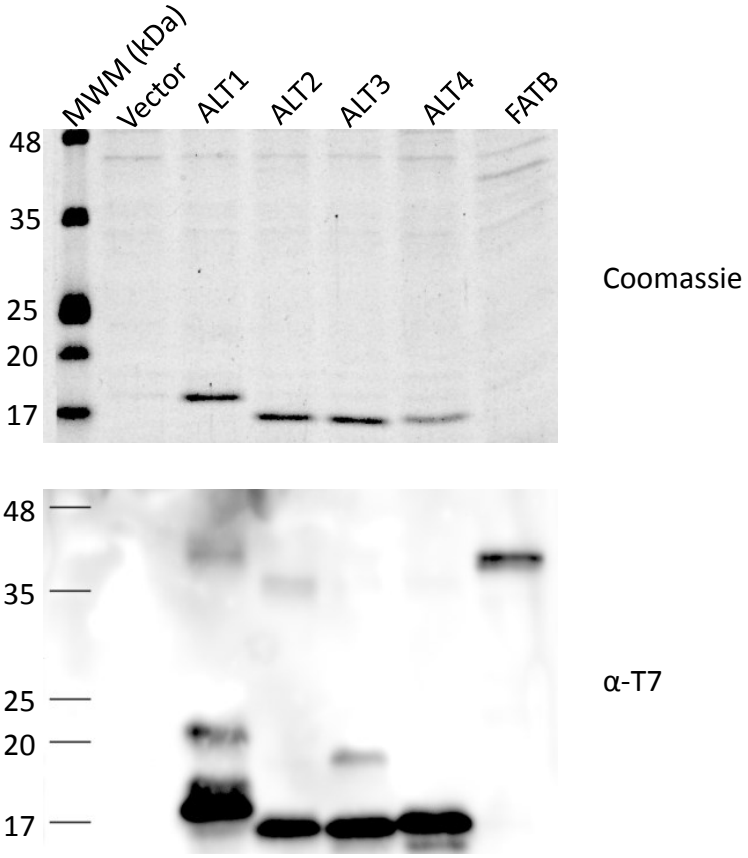


Figure 3.8 Expression of Arabidopsis ALT1-4 and FATB in *E. coli* mutant *fadD*. Transformants of the strains were cultured with IPTG to induce protein expression. Western blots of total cell lysates were performed using anti-T7 antibody to detect the amino-terminal T7 epitope in the protein fusions. The empty vector control was pET-28a. The positions of molecular weight markers (MWM) in kDa are indicated at the left.

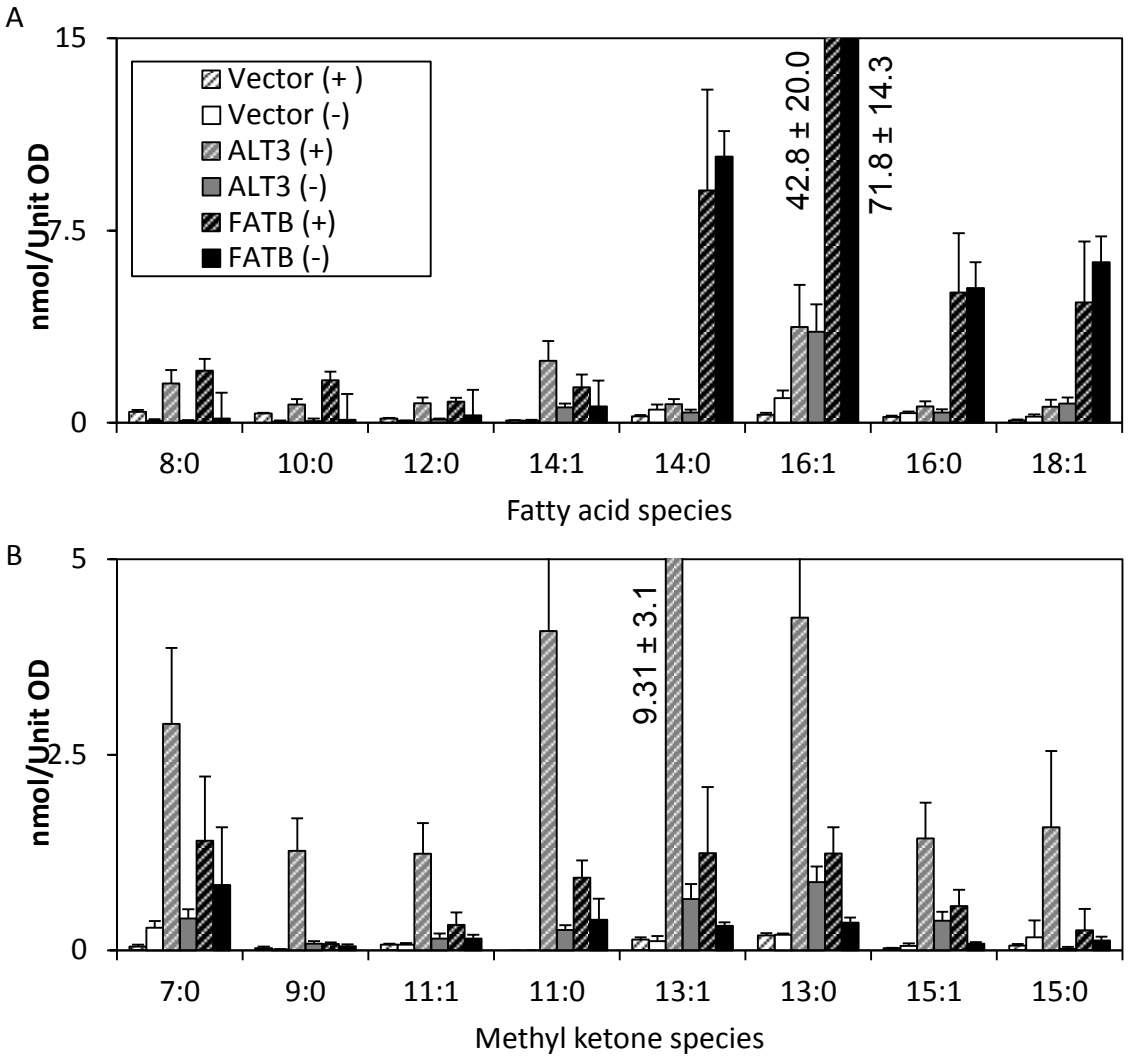


Figure 3.9 Effect of heat and acid during extraction of lipids in the spent media of *E. coli* cultures carrying an empty vector, or expressing ALT1-4, or FATB. (-) indicates no heat and acid treatment and (+) indicates heat and acid treatment. Quantification was done by comparison of peak areas to that of a tetracosane internal standard using GC-FID.

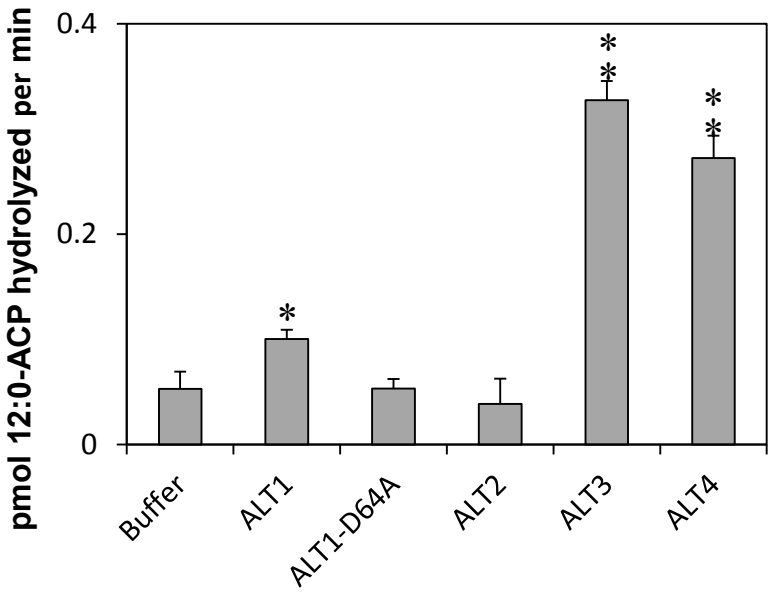


Figure 3.10 *In vitro* 12:0-ACP TE activities of recombinant ALT1-4 and catalytically inactive ALT1-D64A. 400 ng of each TE was incubated with 30 pmol ^{14}C-labeled 12:0-ACP at room temperature for 30 min. Hydrolyzed 12:0 fatty acid (free acid) was extracted with hexane and amounts determined by liquid scintillation counting. Errors bars are ± standard deviation, *n* = 3-5. A student's t-test comparing hydrolysis levels of ALT-containing reactions to the buffer only control gave p values of < 0.005 (*) for ALT1 or < 5 x 10^{-7} (**) for ALT3 and ALT4.

3.5 Discussion

3.5.1 ALT1-4 encode fatty acyl-ACP TEs with varied substrate preferences

The *ALT* gene family encodes four Hotdog fold proteins that are highly related to each other, as well as to *Sh*MKS2, for which β-ketoacyl-ACP TE activity has been demonstrated *in vitro* (Yu *et al.,* 2010). To investigate the hypothesis that ALT1, 2, 3, and 4 are acyl-ACP TEs, we expressed these four proteins in the *E. coli* mutant *fadD*. Monitoring acyl-ACP TE activity is often performed in this bacterial expression system, which has abundant acyl-ACP substrates, like the plastid where these enzymes are localized *in planta*. Acyl-ACP TEs previously tested in such experiments generate products that closely match their known or suspected biological products, both in terms of oxidation state and chain length (Dörmann *et al.,* 1995; Yu *et al.,* 2010). However, discrepancies are sometimes seen with regard to the saturation level of the fatty acids produced in these experiments. For example, the *in planta* product of FATB is known to be 16:0 fatty acid, but studies that measured fatty acids accumulating in FATB-expressing *E. coli* gave differing results. In one instance, 16:0 was the predominant product in *E. coli* (Dörmann *et al.,* 1995), whereas in another study, as well as in this report, 16:1 predominated in the *E. coli* media (Mayer and Shanklin 2007; Figures 3.6, 3.7). Therefore, the saturation level of substrates utilized by ALT1-4 in *E. coli* may be different from their true *in planta* substrates. This is especially evident when considering that unsaturated fatty acids of chains lengths shorter than C18, while present in *E. coli*, are not significant metabolites in plant cells. However, the deduced substrate specificities with regards to chain length and oxidation state are expected to be roughly in accordance with *in planta* substrates.

Despite the very high similarities between the ALTs in their protein sequences, the products they generated in *E. coli* were substantially different and this likely reflects differing substrate *in planta*. Fatty acids were the major products of ALT1 (C12, C14, and C16 chain lengths) as well as ALT4 (C6 and C8 chain lengths). Prior to this study, the shortest chain fatty acid reported to be a major product of a plant acyl-TE was 8:0 generated by *Cuphea* spp. FATB variants (Tjellström *et al.,* 2013). The 6:0 fatty acid generated by ALT4 is now the shortest known major fatty acid product of a plant fatty acyl-TE. The major products of ALT2 were 8:0 and 10:0 β-ketofatty acids. We chemically converted the inherently unstable β-ketofatty acids produced in *E. coli* to methyl ketones for detection and quantification by gas chromatography. In the plant, the β-ketofatty acids are expected to be further metabolized and methyl ketones are also the likely end product. In wild tomato, β-ketofatty acids produced by *Sh*MKS2, which is highly related to the ALTs, are converted to methyl ketones by a β-ketofatty acid decarboxylase (*Sh*MKS1). *Sh*MKS1 is an atypical member of the α/β–hydrolase superfamily (Auldridge *et al.,* 2012). Arabidopsis contains many genes encoding predicted α/β–hydrolases, but it is currently unknown whether any of these can act as β-ketofatty acid decarboxylases. Sequence similarity to *ShMKS1* identifies several of the 21 *METHYL ESTERASE* (*MES*) gene family members as potential orthologs in Arabidopsis, although no single candidate is obvious by sequence inspection or by co-regulation analysis. ALT3 appears to have broader substrate specificity, overlapping with the other ALT family members. All of the products generated by ALT1, ALT2, and ALT4 were produced at some level by ALT3. The

apparent broad specificity of ALT3 may indicate its involvement in more than one biological role and it may have functional overlap with the other ALTs.

Lauroyl-ACP TE activity was demonstrated *in vitro* for ALT1, 3, and 4, but not for ALT2. In our *E. coli* expression experiment, 12:0 fatty acid was the major product of the ALT1 expressing strain, with levels of this product elevated approximately 15-fold higher than in the vector only control. In contrast, 12:0 fatty acid was a minor product of ALT3 and 4, with levels only 1.8 and 2.3 times that of the vector only control. However, in our *in vitro* assay, ALT3 and ALT4 displayed significantly higher 12:0-ACP TE activity than did ALT1. This discrepancy may be explained by our difficulty in obtaining high quality enzyme preparations. We obtained a lower yield and purity for ALT1 than we did for ALT3 or ALT4, and our ALT1 preparation may contain a larger number of misfolded and/or inactive enzyme molecules. The levels of 12:0 fatty acid produced by the ALT2 strain were barely above the vector only control levels, which is consistent with our lack of 12:0-ACP TE activity *in vitro*. It should be noted that of the four ALT proteins, ALT2 was the only one to not produce any fully reduced fatty acids as a major product when expressed in *E. coli*. Future studies examining the preferred substrates of the ALT proteins need to be conducted by assaying *in vitro* different substrates and comparing relative enzyme kinetic activities.

ALT proteins likely form homodimers or heterodimers, as this is common for Hotdog fold enzymes (Pidugu *et al.,* 2009). Each ALT produced novel lipid products when expressed alone in *E. coli*, indicating that they can function as homodimers.

However, heterodimeric complexes have been reported for Hotdog fold proteins (Sacco *et al.,* 2007), and this could be the case for ALTs *in planta*. The ubiquitous expression of ALT3 would allow for the encoded protein to form heterodimers with ALT1, 2, or 4.

3.5.2 The ALT family spans broadly across plants, but functions might not be conserved

Homologues of the Arabidopsis ALTs are found in wide ranging plant taxa. In a phylogenetic tree, the ALT proteins cluster mainly with proteins from the same or related species. This is in contrast to FATA- and FATB-type thioesterases, where proteins cluster according to functional FATA-type and FATB-type clades, even across vast plant lineages (Mayer and Shanklin 2007, Figure 3.2). This suggests that there is an ancient *ALT*-type gene ancestor, but that gene duplication events leading to paralogues occurred relatively recently. The ubiquitous expression and the relatively broad substrate specificity of ALT3 suggest that it is most related in activity and function to an ancient, multifunctional enzyme. ALT1, ALT2 and ALT4, which are more restricted in both expression patterns and substrate preferences, may have evolved through sub-functionalization to have more specialized activities and biological roles. Atypical FATB enzymes with specificity for shorter chain lengths substrate (*e.g. Cuphea hookeriana* FATB2 and *Umbellularia californica* FATB, which produce 8:0-10:0 and 10:0-12:0, respectively), are evolutionarily distinct from ALT1 and ALT4, despite having similar substrate preferences, indicating that this specialization arose independently (Figure 3.2).

3.5.3 Possible biological roles for ALT1-4

The gene expression pattern of *ALT1* is suggestive of a role in cuticular wax synthesis. Expression in aerial tissues, in particular epidermal cells, is a characteristic of cuticle biosynthetic genes. This is consistent with public microarray data that shows expression of *ALT1* correlated with expression of genes involved in cuticular wax biosynthesis, a process that occurs late in cuticle development. The biochemistry of wax biosynthesis is well characterized (Bernard & Joubès, 2012). Wax precursors are delivered as very-long-chain acyl-CoAs to the ER where chemical modification occurs. It thus seems unclear how a plastid localized acyl-ACP TE that produces medium-chain fatty acids would be involved in this process. It is possible that ALT1 is involved in generating non-wax metabolites that accumulate in the cuticle. The high volatility of these compounds would cause them to be missed in typical cuticular wax extraction procedures that involve dipping of plant tissues in chloroform or hexane followed by evaporation under nitrogen gas.

The gene expression pattern of *ALT2* suggests a role in suberin biosynthesis. In addition to strong co-regulation coefficients with known suberin genes, *ALT2* is expressed at known sites of suberin deposition: endodermal and peridermal cell layers of roots as well as seeds. However, the plastid localization and production of short chain β-ketofatty acids in *E. coli* are inconsistent with a role in suberin synthesis, which uses long-chain fatty acids as starting material and occurs in the ER (Beisson *et al.,* 2012). A commonality with the sites of *ALT2* expression is that these are regions that are fortified against microbial pathogen and insect attack. ALT2 may generate β-ketofatty acids at these sites that are converted into methyl ketones for defense

against pests. It is interesting to note that 5:0 methyl ketone has been detected in volatile emissions from *Arabidopsis thaliana* after herbivore damage (van Poecke *et al.,* 2001), and ALT2 is a good candidate enzyme for producing this compound. It would be expected that a decarboxylase, likely a homolog of *Sh*MKS1 (Auldridge *et al.,* 2012), would catalyze the conversion of β-ketofatty acid to methyl ketone *in planta*.

As postulated earlier, *ALT3* function may be at least partially redundant to *ALT1, ALT2* and *ALT4* in tissues where gene expression overlap, but it likely also has a unique role. The main product of ALT3 in our *E. coli* experiment was uniquely 14:1 β-ketofatty acid (chemically converted to methyl ketone in our experimental procedure). It is possible that this is converted to a C13 methyl ketone *in planta*, like that produced in wild tomato, with the naturally produced β-ketofatty acid likely to be saturated. Metabolomic profiles of *Arabidopsis* indicate the presence of a 15:0 methyl ketone (Bais *et al.,* 2010) and ALT3 is a candidate for the production of this compound as it produced 16:0 β-ketofatty acids in *E. coli*. The presence of additional methyl ketones in *Arabidopsis* could easily have been overlooked as they are highly volatile, which may cause them to be lost during evaporation of solvents, a common step in extraction protocols. The shorter chain length methyl ketones would also elute too early to be detected by GC analysis using routine methods. ALT3 showed significant activity towards 16:1 fatty acyl-ACP in our *E. coli* experiments. If, like FATB, the degree of saturation of the preferred substrate is altered in this heterologous system, and ALT3 acts on 16:0 fatty acyl-ACPs *in planta*, the two proteins may be partially redundant. In a *fatb* mutant with less than 1% of wild-type transcript levels, saturated fatty acids and their derivatives are reduced by only 20-

50% depending on the tissue type and lipid species analyzed (Bonaventure *et al.,* 2003). Saturated fatty acids are reduced further, but not eliminated, in a *fatb act1* double mutant, which additionally lacks plastid glycerol-3-phosphate:acyl-ACP acyltransferase activity (Bonaventure *et al.,* 2003). These authors postulated that mitochondrial fatty acid synthesis may compensate for the low 16:0 acyl-ACP TE activity in plastids or that the low-level activity of FATA for saturated fatty acyl-ACPs provides some saturated fatty acids. We now propose that ALT3 generates a proportion of the saturated fatty acids that exit the plastid.

ALT4 is expressed in anthers, and generates 6:0 and 8:0 saturated fatty acids when expressed in *E. coli*. ALT4 may play a role in development of floral organs by generating short chain fatty acids to modulate ethylene sensing and production in anthers. Short chain saturated fatty acids are known to be produced by floral organs of certain ornamental flowers, and play roles in ethylene perception, for example, during the transition from pre-fruit maturation to fruit maturation and senescence (Whitehead and Halevy 1989; Whitehead and Vasiljevic 1993; Halevy *et al.,* 1996). Ethylene has been shown to promote anther dehiscence in tobacco, although jasmonic acid and not ethylene is thought to promote anther dehiscence in Arabidopsis (Rieu *et al.,* 2003). Another possibility is that ALT4 may be involved in production of odor compounds to attract insect pollinators or insect bodyguards. Although Arabidopsis generally self-pollinates, insect mediated cross-pollination is observed in wild populations (Tan *et al.,* 2005), and fatty acids and their derivatives are often part of floral scent mixtures (Knudsen *et al.,* 1993).

3.6 Conclusion

ACYL-LIPID THIOESTERASE1 (ALT1), ALT2, ALT3, and ALT4 represent a novel family of plastid-localized acyl-ACP thioesterases in *Arabidopsis thaliana.* These proteins produce different lipid products despite their high amino acid sequence similarities. They may be involved in generating specialized lipid compounds previously unreported in Arabidopsis. Predicted ALT-type proteins are found in a wide variety of plant species. Considering the enzymatic diversity within one plant species, the variety of potential activities of other ALT-type enzymes is immense.

Chapter 4: Towards characterization of the biological roles of *ACYL-LIPID THIOESTERASE1-4* from *Arabidopsis thaliana*

4.1 Abstract

Initial characterization of Arabidopsis *ACYL-LIPID THIOESTERASE1-4* indicates that these genes encode medium-chain acyl-ACP thioesterases with roles in specialized metabolism. To probe the biological roles of these genes, we attempted to identify the *in vivo* enzymatic substrates and products of the ALT1-4 enzymes. First, we examined the medium-chain acyl lipid profiles of various Arabidopsis tissues. 12:0 and 14:0 fatty acids were detected in all tissues, and their abundance roughly correlates with the transcript abundance of ALT1. 6:0 fatty acids were detected in flowers, specifically the stamens, which is the only tissue where ALT4 expression has been observed. A compound tentatively identified as 13:0 methyl ketone was observed, that is likely derived from a 14:0 β-ketofatty acid. This compound is detected in all tissues, consistent with the ubiquitous expression of ALT3. These observations help support and refine our hypotheses regarding enzymatic products of the ALTs that were initially presented in Chapter 2. Second, phenotypic characterization of *alt1*, *alt2*, *alt3*, *alt1 alt3*, and *alt2 alt3* loss-of-function mutants indicates that ALT1-3 all contribute to generation of 12:0 and 14:0 fatty acids. No differences in the levels of 13:0 methyl ketones were observed in any mutants. Several issues to consider during further analysis of these chemical and mutants are outlined.

4.1 Introduction

Initial characterization of *ALT1-4* indicated that these genes encode acyl-ACP thioesterases (TEs) with activities towards a range of medium-chain acyl-ACP and β-ketoacyl-ACP substrates (Chapter 3; Pulsifer *et al.*, 2014). To date, the biological roles of other plant acyl-ACP thioesterases have been mostly focussed on a different family of thioesterases. The ubiquitous, plastid localized FATA and FATB TEs hydrolyze newly synthesized C16 and C18 acyl-ACPs to allow export from the plastid, where they enter into various metabolic pathways generally localized at the endoplasmic reticulum (ER) (Li-Beisson *et al*., 2013; also see Section 1.2.2). The genomes of certain plant species (*e.g. Cuphea* spp. and *Umbellularia californica* or California bay tree) also encode alternate FATBs with activities towards saturated medium-chain acyl-ACP chains (Voelker et al., 1992). These specialized FATBs are expressed in developing seeds, and direct medium-chain fatty acids (MCFAs) into seed oil synthesis. Only one plant acyl-ACP TE has been characterized with a role separate from these FATA/FATB activities. *Solanum habrochaitis* subspecies *glabratum* (wild tomato) *MKS2* is expressed in trichome cells, where it generates 12:0-16:0 β-ketofatty acids, which are further metabolized into insecticidal methyl ketones (Ben-Isreal *et al.*, 2009; Yu *et al*., 2010). ALT1-4 are phylogenetically closely related to tomato MKS2, and not closely related to the FATA/FATB-type thioesterases (Figures 3.1-3.2; Pulsifer *et al.*, 2014).

Like *MKS2*, *ALT1-4* likely have specialized roles, distinct from the general lipid metabolic roles of typical *FATA* or *FATB* genes. The enzymatic products of heterologously expressed ALT proteins are MCFAs and medium-chain β-ketofatty acids (Figure 3.7), suggesting that the functions of these proteins are distinct from

that of the ubiquitous FATA/FATB-type acyl-ACP TEs, which generate long-chain fatty acids (LCFAs). The functions of *ALT1-4* are also unlikely to be related to seed oil production, like that of the specialized *FATB* genes. Arabidopsis seed oil composition is well characterized (O'Neill *et al*., 2003; Tjellström *et al.,* 2013), and MCFA have not been reported to be present. Further, the expression of *ALTs* in non-seed tissues is inconsistent with such a role (Figures 3.3, 3.4).

A critical first step towards characterizing the biological roles of *ALT1-4* is to identify their *in planta* substrates and products. Towards this goal, we examined the acyl lipid profiles of various Arabidopsis tissues. By comparing these data to the *ALT* gene expression patterns and predicted substrates preferences (Figures 3.3, 3.4, 3.7), we refined our hypotheses regarding the biologically relevant products of *ALT1-4*. We then used publically available *alt1*, *alt2*, and *alt3* loss-of-function mutants to test some of these hypotheses.

4.2 Materials and methods

4.2.1 Fatty acid profiles

Arabidopsis thaliana ecotype Col-0 were grown for six-weeks under a 16 hour light / 8 hour dark photoperiod at 21°C with a light intensity of 100-125 μmol/m^2s. For most tissues, 50-150 mg of plant material was harvested and submerged in 3 ml of methanol with 2% (v/v) H_2SO_4, and 5 μg heptadecanoic acid (17:0) was added as an internal standard. For individual floral organs, flowers were dissected under a dissecting microscope, and individual organs were collected in methanol. After

collection was finished, the methanol was topped up to a volume of 2 ml, and 1 ml methanol with 6% (v/v) H_2SO_4 was added. Samples were incubated for 90 min at 80 °C. One ml of 0.9% NaCl was added to decrease solubility of fatty acid methyl esters, which were then extracted into 250 µl hexane.

One µl of the hexane fraction was used for gas chromatographic analysis. Lipids were quantified with a Varian GC3900 equipped with a HP-5MS column (30 m length, 0.25 mm inner diameter, 0.25 µm thick film of 5% Phenyl 95% dimethylpolysiloxane) and a flame ionization detector. Samples were injected in splitless mode. The carrier gas was helium with a constant flow rate of 2 mL/min. The column oven was initially at 50°C for 8 min, then ramped at 15°C/min up to 320°C, which was held for 4 min, giving a total run time of 30 min.

4.2.2 Reverse transcriptase PCR analysis of ALT transcripts

Total RNA was extracted from ~300 mg of flower tissue from Col-0 (wild-type), *alt1*, *alt2*, and *alt3,* using a commercial kit (RNaqueous kit, Applied Biosystems, Carlsbad, CA, USA) following the manufacturer's instructions. 500 ng of total RNA was used to synthesize complimentary DNA (cDNA) using an 18-mer oligo-dT primer and SuperScript-reverse transcriptase III (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. ALT specific primer pairs were chosen to amplify each of the entire open reading frames, from the start codon through to the 3' UTR (Appendix I).

4.2.3 Genotyping of alt T-DNA insertion mutants and generation of alt1 alt3 *and* alt2 alt3 *mutants*

T-DNA insertion mutant lines were obtained for *alt1*, *alt2*, and *alt3* from the Salk Institute for Genomic Analysis Laboratory and the Arabidopsis Biological Resource Center. The lines obtained were SALK_201400c, SALK_207758c, and SALK_206212c, for *alt1-1*, *alt2-1*, and *alt3-1*, respectively. Genomic DNA was prepared from leaf or stem tissue, following the protocol described in Edwards *et al*. (1991), and used as template for genotyping PCR. The absence of a given T-DNA (*i.e.* a wild type allele) was indicated by the presence of an amplicon using primers ALTX-RP and ALTX-LP (Appendix I), which anneal to genomic regions flanking the T-DNA insertion site. The presence of a given T-DNA (*i.e.* a mutant allele) was indicated by the presence of an amplicon using primers ALTX-RP and LBb1.3 (Appendix I), the latter of which anneals to the left border of the T-DNA sequence.

Double mutants were generated by crossing homozygous *alt1* and *alt3*, or *alt2* and *alt3* mutants. Seeds from these crosses were collected and grown as the F1 generation. F1 seeds were pooled and grown as the F2 generation. Individual F2 plants were genotyped, and seeds of homozygous double mutants were collected. The genotypes of four F3 individuals from each F2 line were verified by PCR genotyping.

4.3 Results

4.3.1 Medium-chain *acyl lipid profile varies between tissues.*

Previous characterization of the *ALT* genes and their encoded proteins indicated that these enzymes have divergent expression patterns and substrate

preferences (Chapter 3; Pulsifer *et al*., 2014). We sought to explore the *in planta* roles of these acyl-ACP thioesterases by looking for the predicted enzymatic products in different tissues. We began this analysis by determining the fatty acid profile of various Arabidopsis tissues, paying particular attention to medium-chain length fatty acids and derivatives.

MCFAs from 6:0-14:0 were detected in all tissues examined, although at significantly lower levels than the LCFAs that make up most of the cellular membrane lipids (Figure 4.1). The most abundant MCFA was 14:0 in flowers, with a level of 75.8 nmol/g fresh weight. In contrast, 16:0, a ubiquitous LCFA had levels of 800-3000 nmol/g fresh weight across various tissues. 6:0 and 10:0 fatty acids were most abundant in flowers (43.4 & 20.5 nmol/g fresh weight, respectively), with only trace amounts (<5nmol/ g fresh weight) of either of these fatty acids detected in stems, leaves, or siliques. 8:0 fatty acid was detected only at trace levels in leaves and flowers. 12:0 and 14:0 fatty acids were most abundant in flowers (45.9 & 75.8 nmol/g fresh weight, respectively), but were also present in stems, leaves and siliques (2.7-6.3 nmol/g fresh weight for 12:0, and 10.7-36.7 nmol/g fresh weight for 14:0).

Since *ALT* gene expression patterns varied among individual floral organs (Figure 3.4), we asked whether these organs show similar differences in MCFA profiles. It should be noted that due to the very small mass of Arabidopsis floral organs, an accurate measurement of fresh weight could not be taken. Thus, all MCFA levels are reported as percent of total extractable fatty acids (Figure 4.2). In all four floral whorls, the most abundant MCFA was 14:0 (0.4-2.0% of total extractable fatty

acids), followed by 12:0 (0.2-1.0%). In most cases, 6:0 and 8:0 represented no more than 0.2% of total extractable fatty acids. The exception was stamens, where 6:0 fatty acids accumulated to 0.5%.

A compound likely to be 2-tridecanone (13:0 methyl ketone) was also detected in all tissues examined (Figure 4.1). This compound elutes at the same retention time as a commercial standard of 2-tridecanone, although its identify remains to be confirmed by mass spectrometry. This compound was detected in stems, leaves, flowers, and siliques. No other methyl ketones were detected in any tissue.

4.3.2 Genetic characterization of alt1, alt2, *and* alt3 *T-DNA insertion line, and* alt1 alt3 *&* alt2 alt3 *double mutants*

T-DNA insertion mutant lines were obtained for *alt1*, *alt2*, and *alt3.* Only one line was available for each of these genes, while none were available for *alt4*. The genomic locations of the T-DNA insertions were confirm by DNA sequencing. The *alt1-1* T-DNA is located in the third exon of the gene. The *alt2-1* and *alt3-1* T-DNAs are located in the first and fourth introns of their respective genes (Figure 4.3). Analysis of *ALT* transcript levels by reverse transcriptase PCR revealed no detectable full length transcript for *ALT1*, *ALT2*, or *ALT3*, in their respective mutant lines, suggesting that all three lines are null mutants (Figure 4.4). Due to possible genetic redundancy, these mutants were crossed to make *alt1 alt3* and *alt2 alt3* double mutants. Before analysis, the genotypes were confirmed in four individuals from each double mutant line (Figure 4.5).

4.3.3 Medium-chain fatty acid and methyl ketone profiles in flowers and stems of alt *mutants*

The MCFA profiles of *alt* single and double mutants were analyzed in both flowers and stems. In the flowers, only subtle differences were seen between genotypes, and none proved to be statistically significant using a student's T-test with a cut off of p=0.05 (Figure 4.6). On the other hand, MCFA profiles in stems displayed statistically significant differences in *alt* mutants compared to wild-type (Figure 4.6). 12:0 was reduced by 44-52% in *alt1* and *alt1 alt3* mutants, and 32% in the *alt3* mutant. 14:0 was reduced 30-38% in *alt1*, *alt3*, *alt1 alt3*, and *alt2 alt3* mutants, and by 16% in *alt2* mutants. The compound tentatively identified as a 13:0 methyl ketone was detected in flowers and stems of all mutants, again at or near trace levels (Figure 4.6). The levels of this compound were not statistically different from wild-type compared to any of the *alt* mutants tested here.

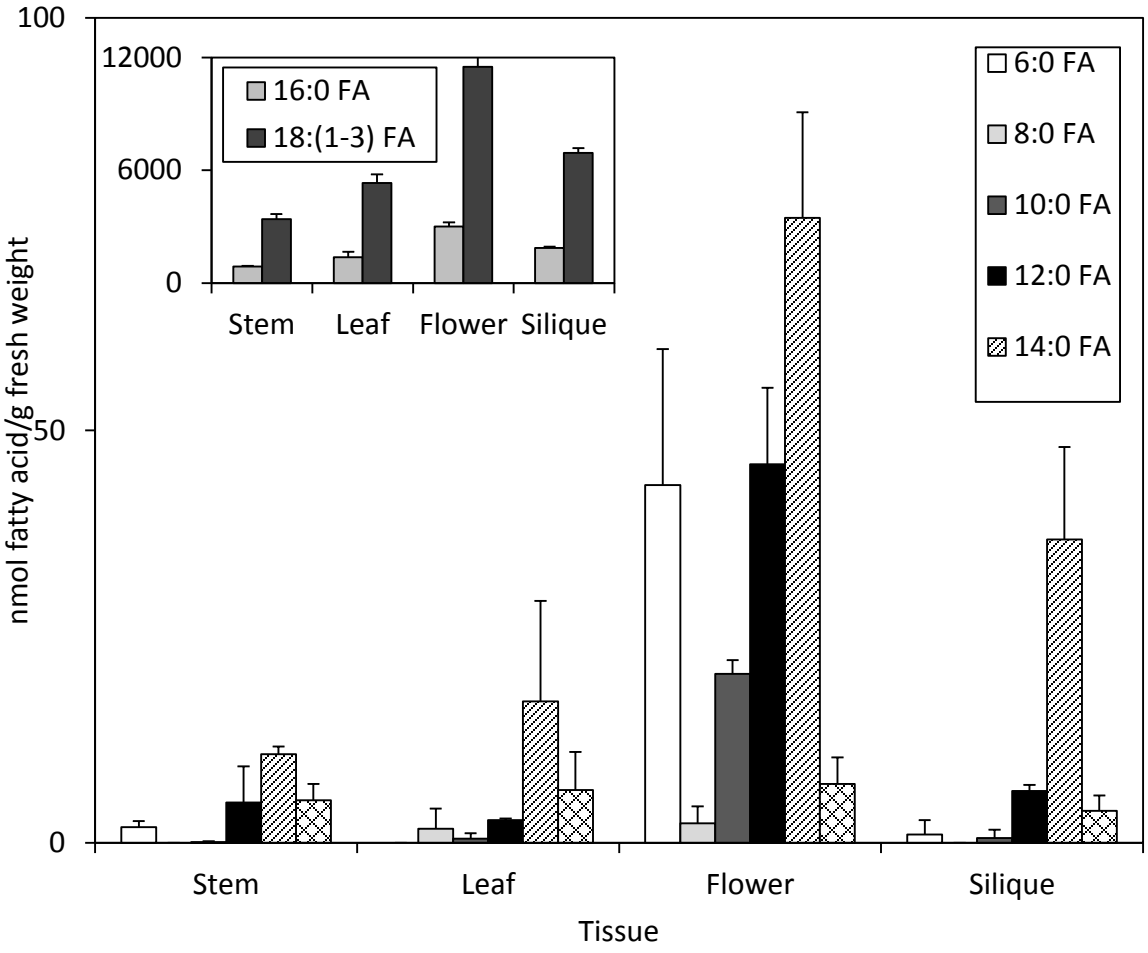


Figure 4.1 Medium-chain acyl lipid profiles of various Arabidopsis tissues. Fatty acids from whole tissues of wild-type plants were trans-methylated, extracted, and separated by GC-FID. Analytes were identified by comparing retention times to commercial standards. The identities of 6:0-14:0 fatty acids in flower extracts were confirmed by GC-MS. The identity of the 13:0 methyl ketone has yet to be confirmed. Compounds were quantified by comparison of GC-FID peak areas to that of a 17:0 fatty acid standard. Inset show quantifications of LCFAs for comparison (Y-axis is same units). FA= fatty acid; MK= methyl ketone.

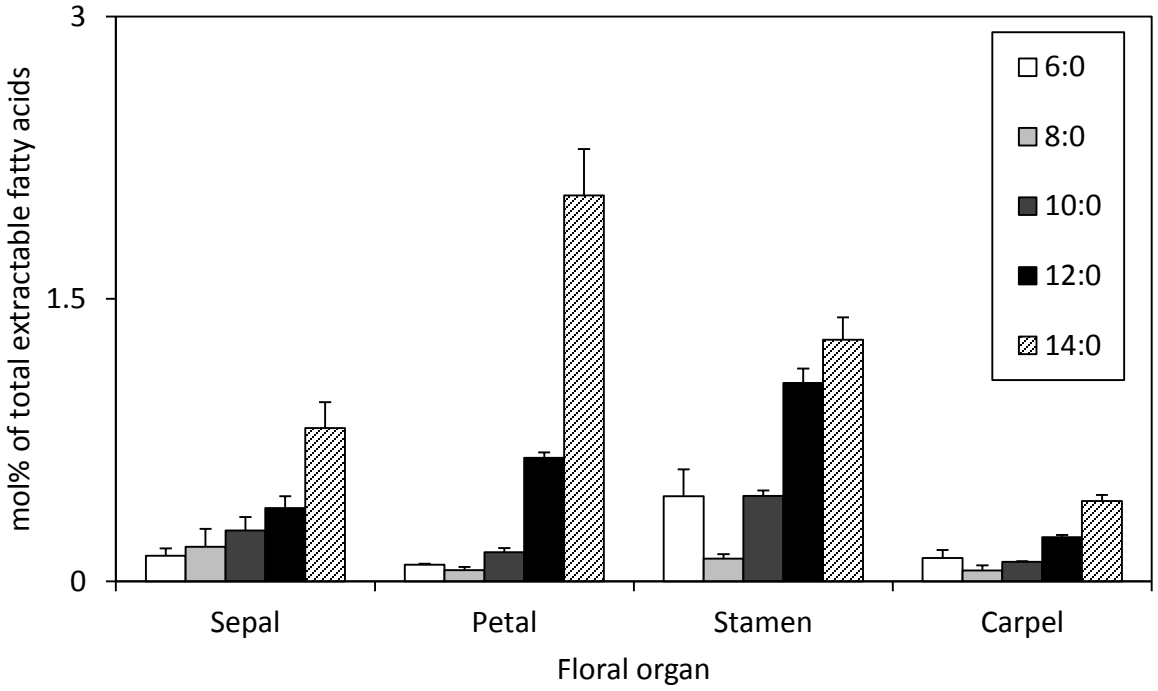


Figure 4.2 Medium-chain fatty acid profiles of individual floral organs. MCFAs were measured in dissected organs of Arabidopsis flowers, similar to that described for whole tissues in Figure 4.1. However, accurate masses could not be obtained, so MCFAs are reported as a proportion of total extractable fatty acids (mol %).

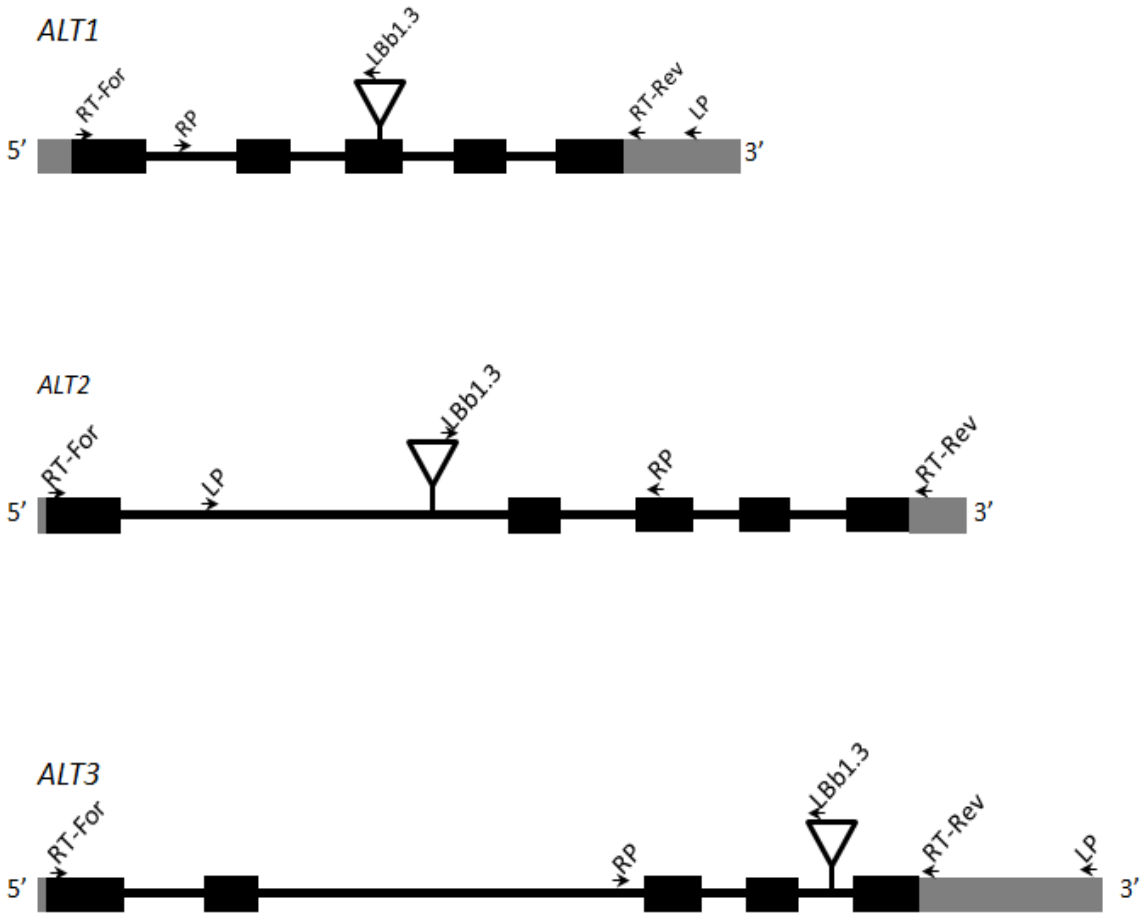


Figure 4.3 Genomic positions of T-DNA insertions in *alt1*, *alt2*, and *alt3* mutants. Commercially available mutant lines were obtained from the Arabidopsis Biological Resource Centre. The positions of the T-DNA insertions were confirmed by DNA sequencing. Wide grey bars represent untranslated regions, wide black bars represent exons, narrow black bar represent introns, triangles represent T-DNA insertions, and arrows indicate binding sites of primers used for genotyping or RT-PCR.

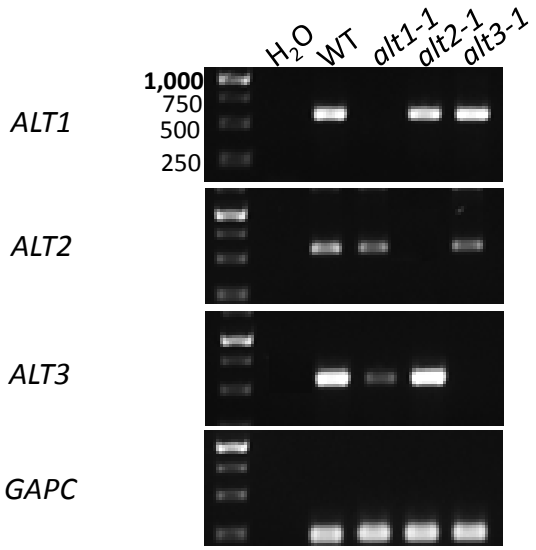


Figure 4.4 Transcript analysis of *alt* mutants by reverse transcriptase PCR. Total RNA was extracted from flowers and used to make cDNA. *ALT* specific primers anneal to ends of genes such that the entire open reading frames of their respective transcripts were amplified. PCR templates are labeled along the top, and transcripts being amplified are labeled down the left side. Molecular weight markers are show on the left on each panel; band sizes (in bp) are shown on the top panel.

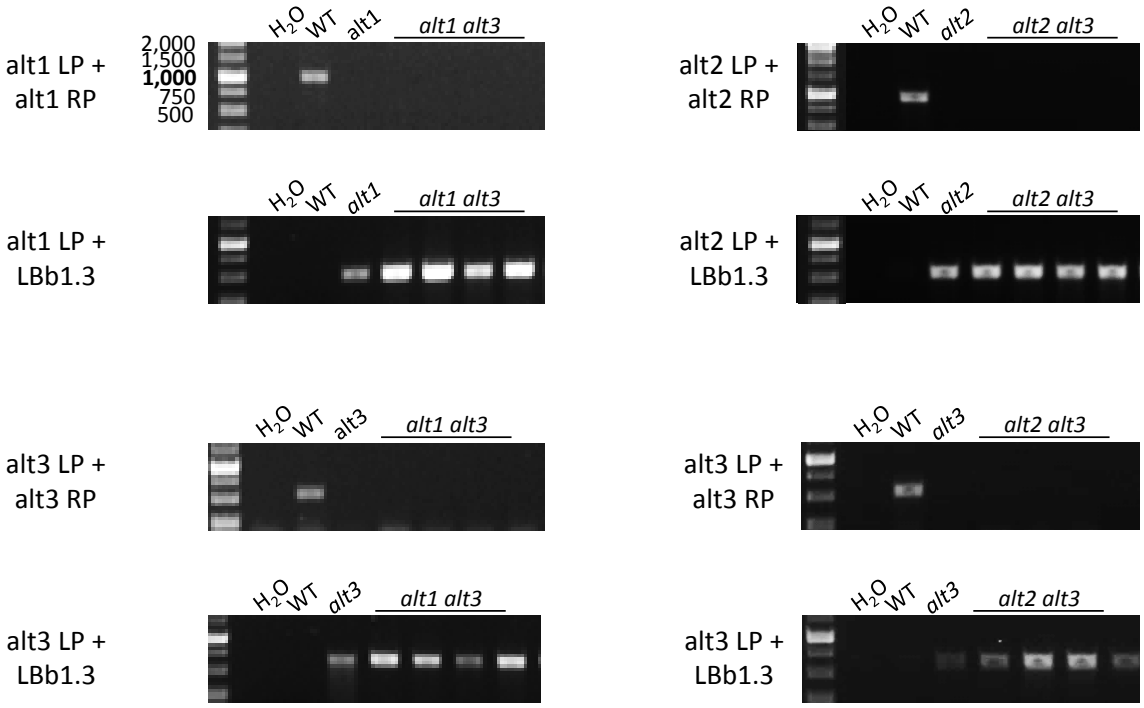


Figure 4.5 Genotyping of *alt1*, *alt2*, and *alt3* single mutants, and *alt1 alt3* and *alt2 alt3* double mutants. Four F3 individuals of each line were genotyped, which were derived from a pure-breeding F2 parent. A band using ALTX-LP + ALTX-RP primers indicates a wild-type allele, while a band using ALTX-RP + LBb1.3 primers indicates a mutant allele (see Figure 4.4 for primer binding sites). Molecular weight markers are show on the left on each panel; band sizes (in bp) are shown on the top left panel.

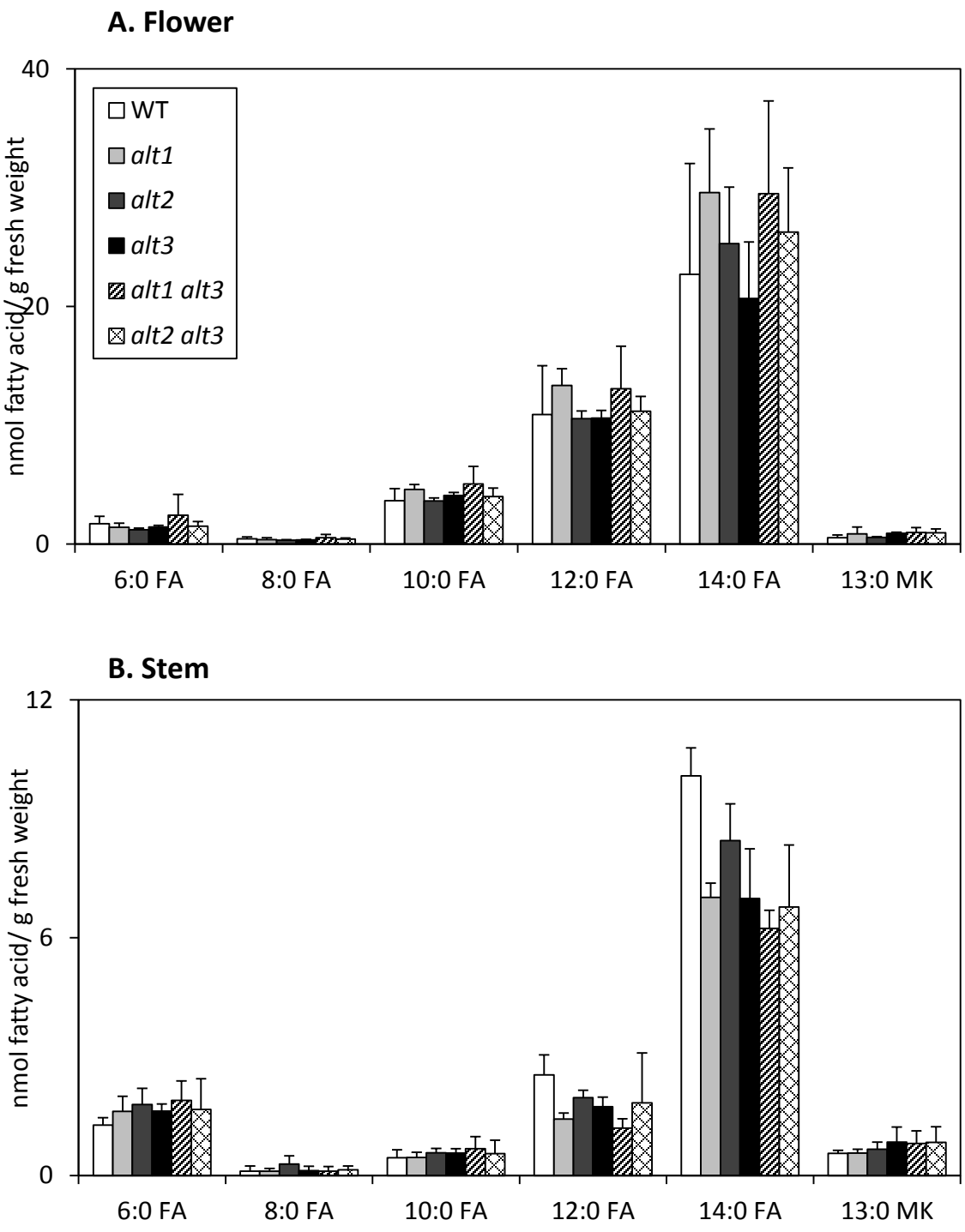


Figure 4.6 Medium-chain acyl lipid profiles of flowers (A) and stems (B) in *alt* mutants. Sample preparation and analysis techniques were the same as in Figure 4.1. FA= fatty acid; MK= methyl ketone.

4.5 Discussion

4.4.1 Predictions regarding the biological products of ALT1-4

Predications regarding the biological products ALT1-4 were previously made, based on the results of expressing these proteins in *E. coli* (Section 3.5.3). By combining these results with those from the *ALT1-4* gene expression patterns and the MCFA profiles in various tissues, these predictions can now be refined. *ALT1* is expressed in aerial tissues, most highly in epidermal cells of the stem and in flower petals (Figures 3.3, 3.4). When expressed in *E. coli*, ALT1 generated mainly 12:0 fatty acid, but also generated fatty acids of other chain lengths (Figure 3.7). 12:0 and 14:0 fatty acids were detected in all aerial tissues (Figure 4.1). Levels of these MCFAs were highest in flowers and 14:0 was particularly high in flower petals (Figures 4.1, 4.2). Both 12:0 and 14:0 fatty acids showed significant reductions in stems of *alt1* and *alt1 alt3* mutants (Figure 4.6). Based on these results, we modify our original hypothesis to include 14:0 as a predicted product of *ALT1* in addition to 12:0.

ALT2 is expressed in root endodermis and peridermis as well as aerial organ boundaries (Figures 3.3, 3.4). When expressed in *E. coli*, ALT2 produces 8:0 and 10:0 β-ketofatty acids. If these compounds were to be produced *in planta*, it's possible that they would be enzymatically decarboxylated to 7:0 and 9:0 methyl ketones. Even if this reaction doesn't occur *in vivo*, the extraction protocol used in these experiments would provide conditions for spontaneous decarboxylation. As neither 7:0 nor 9:0 methyl ketones were detected in any tissues tested, we do not have any evidence to support the hypothesis that ALT2 produces 8:0 and 10:0 β-ketofatty acids. As *ALT2*

is highly expressed in roots, this tissue should be tested for the presence of 7:0 and 9:0 methyl ketones.

The most abundant compound produced when *ALT3* was expressed in *E. coli* was a C14 β-ketofatty acid, although this enzyme displayed broad substrate preference, also producing fatty acids and β-ketofatty acids of other chain lengths (Figure 3.7). *ALT3* was found to be expressed in all tissues examined (Figure 3.3), and a compound that is potentially a 13:0 methyl ketone was detected in all aerial tissues (Figure 4.1). If this compound is a 13:0 methyl ketone, it is very likely derived from a 14:0 β-ketofatty acid. Again, this decarboxylation reaction could either be enzymatic or spontaneous. From this, we hypothesize that ALT3 generates 14:0 β-ketofatty acids, which may be enzymatically converted to 13:0 methyl ketones.

Heterologous expression in *E. coli* of *ALT4* generated mostly 6:0 and 8:0 fatty acids (Figure 3.7). *ALT4* is expressed exclusively in flowers, specifically anthers (a sub-structure of stamens), although the absolute level of *ALT4* transcript is very low (Figures 3.3, 3.4). 6:0 and 8:0 fatty acids are detected only at trace levels in most tissues, with 6:0 accumulating to relatively higher levels in flowers (Figure 4.1). 6:0 represented a higher proportion of extractable fatty acids in the stamen, compared to other floral organs (Figure 4.2). These observations lead to the hypothesis that *ALT4* generates 6:0 fatty acids *in planta*. No *alt4* mutants are currently available for phenotypic characterization.

4.4.2 Mutant analysis gives clues of ALT products

When testing *alt* mutants for changes in their acyl lipid profiles relative to wild-type, we considered the possibility that these genes would show genetic redundancy. *ALT3* was found to be expressed at high levels in all tissues tested (Figure 3.3). Further, heterologous expression in *E. coli* indicated that the ALT proteins are active toward a range of substrates. The substrate preference of ALT3 was particularly broad, generating both fully reduced and β-ketofatty acids (Figure 3.7). Since *ALT1, ALT2,* and *ALT4* all overlap with *ALT3* in terms of both gene expression patterns and substrate preferences, it's likely that *ALT3* is at least partially redundant to the other ALT proteins. For this reason, in addition to testing the *alt1*, *alt2*, and *alt3* mutants, we also included *alt1 alt3* and *alt2 alt3* double mutants in our analysis.

Differences in acyl lipid profiles of *alt* mutants were seen in stem tissue, but not in flowers. Further, the differences seen in stem lipid profiles were less dramatic than expected if the ALTs are generating all MCFAs found in these tissues. The reductions of 12:0 and 14:0 fatty acids to some degree in all mutants tested suggests that *ALT1*, *ALT2*, and *ALT3* all contribute to the generation of these compounds. This is a higher level redundancy than we expected, but is consistent with the hypothesis proposed in Section 3.5.3 that the four *ALT* genes diverged from a multifunctional ancestor relatively recently in evolutionary history.

Several considerations can be made as to why the effects of acyl lipid profiles seen in *alt* mutants were less dramatic than expected. First is the possibility that

additional genetic redundancy remains, even in *alt* single and double mutants. Although the main product of FATB is 16:0 fatty acids (Dörmann *et al.*, 2000), it also generates shorter chain fatty acids when expressed in *E. coli* (with chain lengths as short as C6, Figure 3.7), and is expressed at some level in all tissues (Schmid *et al.,* 2005). FATB expression is higher in flowers than in stems (Schmid *et al.,* 2005), which may explain why we saw differences only in the latter. Crossing the *alt* mutants into a *fatb* background and re-analysing medium-chain acyl lipid profiles would test this possibility.

Further, some ALT activity may remain in the mutants tested. In each of the *alt2* and *alt3* mutants, the T-DNA insertion is located in an intron (Figure 4.3). It is possible that some amount of mRNA could be correctly spliced, although at low levels. The *alt1* mutant has the T-DNA located in an exon, and therefore it's highly unlikely that a full length transcript could be spliced. However, there is precedent for an exon contanining a T-DNA to be spliced out, leaving a shortened transcript that retains some activity (Lehti-Shiu *et al*., 2005). Arabidopsis mutants could be generated with directed mutations, perhaps in the start codon or the codon encoding the catalytic aspartate residue (Figure 3.1) (see Chapter 5 for more detail on the CRISPR-Cas directed mutation strategy).

Another consideration is that the predicted enzymatic products of the ALT proteins would not be expected to accumulate to high levels. As stated in Chapter 1, free fatty acids are rarely found as a metabolic end-point. The MCFAs detected in Arabidopsis tissues are likely metabolic intermediates, which is consistent with the

low levels at which they are measured (Figures 4.1, 4.2). If a downstream step in the metabolism of these fatty acids is rate limiting, then even a small amount of TE activity, whether it be from FATB or residual ALT activity, might be enough to maintain the detected level of these fatty acids. It's possible that the metabolic end-points of these fatty acids show significant reductions in *alt* mutants, which could be measured once the identities of these end-points are elucidated. These final products may not be amenable to analysis by the whole tissue trans-methylation technique used in these experiments, and therefore it's likely that other protocols will be required. For example, if a metabolite of interest forms part of a lipid polymer, like cutin or suberin, depolymerisation into individual monomers will be required before they are extractable.

Finally, shorter chain fatty acids and their derivatives are quite volatile, and it's possible that they are released into the atmosphere. In this case, a different analysis technique would be required, such as GC-FID with a solid phase microextraction (SPME) sampling system. SPME uses an adsorbent solid matrix to collect and concentrate gaseous analytes emitted from a sample (Vas & Vékey, 2004). The volatility of these compounds also poses issues during sample processing. Analysis of lipids usually involves total or partial evaporation of a solvent in which the analytes are dissolved. Some amounts of all compounds are likely lost during such steps, with the recovery of these highly volatile compounds being disproportionately low. In the experiments presented in this chapter, all lipid analyses employed protocols that were modified to eliminate all evaporation steps. This is possible since tissues are submerged in methanol, and then lipids are extracted into another solvent

(in this case, hexane). The volume of hexane used for extraction can be small, thereby avoiding the need for concentration by drying. A drawback of this modified technique is that small sample vials must be used to allow for extraction into a small volume, limiting the amount of plant material used. Other protocols would be more difficult to modify, as they involve submerging tissue in a relatively large volume of solvent that must be dried prior to any extractions or derivatizations.

4.6 Conclusion

Hypotheses regarding the *in vivo* enzymatic products of *ALT1-4* were refined with data from MCFA and methyl ketone profiles from wild-type plants and *alt* mutants. *ALT1* is hypothesized to generate 12:0-14:0 fatty acids, although *ALT2* and *ALT3* also contribute to producing these fatty acids. *ALT3* is suggested to produce 14:0 β-ketofatty acids. This compound may be found in all tissues, which corresponds with the ubiquitous expression of *ALT3*. We proposed that ALT4 generates 6:0 fatty acids. Phenotypic characterization of *alt1, alt2, alt3, alt1 alt3,* and *alt2 alt3*, mutants has provided some evidence towards the roles of these genes, but there remain several complications that must be dealt with before further analyses. Genetic redundancy and residual ALT activity should be investigated. Also, dramatic decreases might only be observed for the downstream derivatives of *ALT* generated fatty acids and β-ketofatty acids. Such analysis will likely require alternate techniques to be used, and routine protocols to also be modified.

Chapter 5: Summary of findings and future directions

5.1 Long-chain acyl-CoA synthetases: A new look at their roles in plant lipid metabolism

5.1.1 Summary of findings

Long-chain acyl-CoA synthetases activate long-chain fatty acids to CoA thioesters. Although plant LACS proteins are implicated in a variety of lipid metabolic pathways, several questions remain about their exact roles. For example, Arabidopsis LACS1 and LACS8 are not redundant to each other in seed oil production, despite both being expressed in developing seeds, both being localized to the ER, and each having similar substrate preferences *in vitro* (Shockey *et al.,* 2003; Zhoa *et al.*, 2010). And yet, the lack of any notable phenotype in the Arabidopsis *lacs9* mutant implies another LACS is functionally redundant with LACS9, despite a large decrease in plastid LACS activity and the fact that no other LACS localizes to the plastid like LACS9 does (Schnurr *et al*., 2002). In many cases, LACSs seem unnecessary, as current models depict fatty acids being activated to CoA thioesters upstream of these enzymes (Lü *et al*., 2009; Zhoa *et al.*, 2010). These and other observations suggest that LACS proteins play more complicated roles *in vivo* than simply generating a pool of fatty acyl-CoA substrates.

In other organisms (*i.e.* yeast and mammals), LACS proteins are recognized to be involved in lipid trafficking, and therefore we asked if Arabidopsis LACS1-3 play such a role. LACS1-3 have acyl-CoA synthetase activities towards very-long-chain fatty acids, and in yeast and mammalian cells, such proteins have a second function: transmembrane movement of fatty acids. We explored the possibility that

Arabidopsis LACS1-3 have the same dual function by expressing these proteins in genetically defined yeast strains (Chapter 2; Pulsifer *et al.*, 2012).

The budding yeast *Saccharomyces cereviseae* uses one FATP/VLACS protein (called FAT1p) and two redundant LACS proteins (called FAA1p & FAA4p) in uptake of exogenous fatty acids (Faergeman *et al.*, 1997; Faergeman *et al.*, 2001). Complementation of the synthetase deficient phenotypes of both *fat1Δ* and *faa1Δ faa4Δ* by expression of LACS1, LACS2, or LACS3 indicates that these proteins are active in yeast, and confirms their synthetase activity towards both long-chain and very-long-chain fatty acids. These proteins also complement the fatty acid transport deficient phenotype of *fat1Δ*, supporting our hypothesis that LACS1-3 have dual functionality. All three Arabidopsis proteins aid in uptake of a long-chain fatty acid analogue, while LACS1 also aids in uptake of a very-long-chain fatty acid analogue.

5.1.2 Future directions

Reviewing the FATP/VLACS literature in mammalians systems, it becomes clear that the dual functions of these proteins are difficult to separate conceptually. Some researchers are skeptical of the vectorial acylation model, proposing instead a "metabolic trapping" model (Füllekrug *et al.*, 2012; Zhan *et al.*, 2012). Under this model, free fatty acids diffuse free across membranes, and become trapped when they are activated to hydrophilic CoA. Free fatty acids are able to continue to diffuse into the cell, but provided there is high LACS activity, they cannot diffuse out as fatty acyl-CoA molecules. In other words, the fatty acid transport functionality of LACS proteins

has been challenged as an artifactual phenomenon, and the apparent increase in fatty acid uptake is due to the LACS activities of the FATP proteins.

Various pieces of evidence support the vectorial acylation model over the metabolic trapping model. In the yeast and mammalian literature, the strongest of such evidence is seen in reports that present mutant or splice variant proteins that affect these two functions differentially (Zou *et al.*, 2002; DiRusso *et al.*, 2008; Melton *et al*., 2011). Engineering such variants of Arabidopsis LACS1-3 may support our model of dual functional proteins. The clearest distinction between VLACS and FATP activity is seen in a naturally occurring splice variant of human FATP2, which retains the FATP function but displays no measurable VLACS activity (Melton *et al*., 2011). In this variant, a portion of a highly conserved AMP binding domain is missing. Based on this, the AMP binding domain of LACS1-3 would make an attractive region to start engineering mutants with the aim of potentially separating these two functions.

To date, no phenotype due to a *lacs3* mutation has been reported. Involvement in cuticle biosynthesis has been inferred by enriched expression in stem epidermal cells over stem ground tissues, which is a trademark of cuticle biosynthetic genes (Suh *et al.*, 2005; Pulsifer *et al.*, 2012). Phenotypic characterization of loss-of-function *lacs3* mutants might provide further evidence to support this model. Such characterization should include quantification of cuticular waxes and cutin monomers, and possibly also assessment of cuticle physiological properties, such as water retention and permeability. Since LACS1 and LACS2 are also involved in the

process (Lü *et al*., 2009), double and triple mutants should be generated and characterized.

Further evidence of the biological roles of LACS1-3, or even other Arabidopsis LACS proteins, may be gained through their over-expression. It would be interesting to test whether over expression of a LACS protein leads to increased flux into multiple downstream pathways, or if these effects are limited to fewer pathways. This would be particularly interesting for LACS proteins that localize to the ER, since multiple lipid metabolic processes occur together in this organelle. For example, LACS1 and LACS2 are known to contribute to cutin biosynthesis (Schnurr *et al.*, 2002; Lü *et al*., 2009), presumably by activating long-chain fatty acids for entry into the appropriate metabolic pathway. LACS1 and LACS2 localize to the ER (Weng *et al.*, 2009), where the synthesis of cutin monomers is known to occur (Pollard *et al*., 2008). Synthesis of phospholipids and triacylglycerols also occurs in the ER, and both pathways utilize the same long-chain acyl-CoA substrates as cutin biosynthesis (Li-Beisson *et al*., 2013). If LACS1 and LACS2 act by simply contributing to the pool of long-chain acyl-CoA substrates at the ER, increased flux into the alternate pathways might be expected as well. However, if LACS1 and LACS2 are directing substrates towards specific metabolic fates, these other pathways should not be affected.

Identifying proteins that LACS proteins interact with physically might provide additional evidence to their roles in lipid metabolism. Physical interaction with downstream enzymes might be one mechanism in which LACS proteins direct acyl-CoA products towards specific metabolic pathways. For instance, LACS1, LACS2, and

LACS3 may interact with enzymes that generate cuticular wax and cutin monomers, while peroxisomal-localized LACS6 and LACS7 might interact with the enzymes of fatty acid oxidation. Differences in interacting partners may be the reason that LACS1 contributes to seed oil production, but LACS8 does not. These interactions could be studied through various complementary approaches. The split-ubiquitin system was developed specifically to identify protein-protein interaction of membrane bound proteins through heterologous expression (Johnsson & Varshavsky, 1994). This system allows for pairwise testing of predicted interactions, or the screening of a given protein against a library. Tandem-affinity-purification tagging is a biochemical method developed for studying protein interactions without a heterologous system (Puig *et al*., 2001), and can be used for membrane bound proteins (Qi & Katagiri, 2009). This system allows for isolation of entire protein complexes, which may include additional proteins that are located in close proximity to a LACS but does not interact directly.

5.2 Acyl-ACP thioesterases: Getting more out of fatty acid synthesis

5.2.1 Summary of findings

Acyl-ACP thioesterases are a family of enzymes that cleave acyl carrier protein from an acyl-ACP, generating a free fatty acid. These proteins have widely recognized roles in primary lipid metabolism, specifically terminating fatty acid synthesis. Certain plant species also have specialized acyl-ACP thioesterases to generate medium-chain fatty acids that accumulate in seed oils. There is currently one example of an acyl-ACP thioesterase being involved in a specialized metabolic

pathway other than seed oil and that is MKS2, which is involved in production of medium-chain methyl ketones in wild tomato (Yu *et al.*, 2010).

ACYL LIPID THIOESTERASE1-4 were identified as a novel family of acyl-ACP thioesterases in *Arabidopsis thaliana* (Chapter 3; Pulsifer *et al.*, 2014). These genes are expressed in various tissues and each has a unique gene expression profile. When the ALT proteins were expressed in an *E. coli* expression system, they produced a range of medium-chain fatty acids and β-ketofatty acids. These β-ketofatty acids are easily decarboxylated to methyl ketones, a process that is also likely to occur *in planta*. Despite their high sequence similarities, each ALT protein generated a distinct profile of fatty acyl products. The expression profiles and expected enzymatic products highly suggest that *ALT1-4* are further examples of acyl-ACP TEs involved in specialized metabolism.

ALT1 is expressed most highly in stem epidermal cells and flower petals and its predicted enzymatic products are 12:0-14:0 fatty acids. Consistent with these observations, fatty acids of these chain lengths are found in all tissues, but are particularly abundant in flowers.

ALT2 is expressed in roots and in aerial lateral organ boundaries, and is predicted to generate 8:0 and 10:0 β-ketofatty acids. Although it's unknown if β-ketofatty acids are decarboxylated enzymatically to methyl ketones in Arabidopsis, the extraction protocol used in this thesis provides conditions for this reaction to occur spontaneously. The expected 7:0 and 9:0 methyl ketones were not detected in

Arabidopsis tissues, although the sampling technique used may not be appropriate for detecting these compounds.

ALT3 is expressed in all tissues and to significantly higher levels than the other *ALTs*. Although ALT3 displays broad substrate preference in the *E. coli* heterologous system, the most abundant product is a 14:0 β-ketofatty acid. A compound was detected in all Arabidopsis tissues that we have putatively identified as a 13:0 methyl ketone, which would likely have been derived from a 14:0 β-ketofatty acid, either enzymatically or chemically during the extraction process. Since *ALT3* overlaps with the other *ALTs* both in term of expression and substrate preference, it is possible that *ALT3* is at least partially redundant to the other *ALTs*.

ALT4 is expressed exclusively in anthers, a substructure of stamens, which are in turn a substructure of flowers. Heterologous expression suggests that ALT4 produces 6:0 and 8:0 fatty acids. 6:0 and 8:0 fatty acids are found only at trace levels in most tissues, although 6:0 fatty acids are found at relatively higher levels in flowers, stamens in particular. The lack of an available *alt4* loss-of-function mutant prevents the testing of the hypothesis that these 6:0 fatty acids are generated by *ALT4*.

Publically available *alt* mutants (T-DNA insertions) were used to test the hypotheses regarding the biologically relevant products of ALT1, ALT2 and ALT3 (Chapter 4). 12:0 and 14:0 fatty acids were decreased, though not eliminated in the stems of *alt1* and *alt1 alt3* mutants. This supports the hypothesis that *ALT1* generates the products and that *ALT3* is partially redundant, however further redundancy or residual *ALT* activity appears to be present. The partial reduction of 14:0 fatty acids

in *alt2* suggests that this gene also contributes to producing this metabolite. This was not expected, although significant functional overlap is not surprising considering that these genes are likely to have diverged fairly recently.

Known or predicted proteins with significant sequence similarities to ALT1-4 are encoded in the genomes of many plant species. With the exception of the moss *Physcomitrella patens*, all plant genomes searched had one to six potential *ALT* homologs. Phylogenetic analysis suggests that paralogs within a given species diverged from a single ancestral gene relatively recently in evolutionary history. Given the broad expression pattern and apparent substrate preferences of *ALT3*, it's likely that this is the ancestral gene, and that *ALT1, 2*, and *4* have more specialized functions. The recent divergence of paralogs suggests that these genes are involved in specialized metabolism, and the enzymatic activities could be quite varied.

5.2.2 Future directions

Further invetigation of the biological roles of ALT1-4

Further experiments are required to elucidate the biological substrates and products of ALT1-4. To date, no mutant lines are available for *ALT4*. Such a mutant would be crucial in characterizing the biological function of this gene. CRISPR-Cas is a relatively new technology that allows for targeted changes to an organism's genome (Belhaj *et al.*, 2013). Recently, this system has been shown to be useful for mutational studies in Arabidopsis, including generating stable mutant lines (Feng *et al*., 2013; Jiang *et al.*, 2014). This system works by transgenic expression of the CAS9 nuclease, which targets invading DNA via a guide RNA molecule (Belhaj *et al.*, 2013). By

engineering this guide RNA, the CAS9 nuclease can be directed to cleave the host's genomic DNA. Imperfect repair of the double stranded breaks leads to insertion or deletions in a target gene to create a mutant.

As discussed in Section 3.5.1, it's possible that *ALT3* is at least partially redundant with *ALT4*. Thus an *alt3 alt4* double mutant may be required to observe any potential phenotypes. *ALT3* and *ALT4* are located in close proximity to each other on chromosome 1. These genes are estimated to be separated by 0.02 map units based on the recombination map provided by Singer *et al.* (2006), and are therefore very tightly genetically linked. Because of this, generating an *alt3 alt4* double mutant by crossing single mutant parents would be near impossible. Using the CRISPR-Cas system to engineer an *alt4* mutation in both a wild type and an *alt3* background might prove to be an easier strategy to acquire both single and double mutants. The fatty acid profile of these mutants could then be determined to test the hypothesis that *ALT4* generates 6:0 fatty acids in flowers. As mentioned previously, this fatty acid is highly volatile, and thus an experimental approach that utilizes a solid-phase microextraction sampling method would likely be beneficial.

In light of the higher than expected redundancy in *ALT* gene functions (Section 4.4.2. Figure 4.6), further combinations of multiple *alt* mutations should be generated (*e.g.* an *alt1 alt2* double mutant, or *alt* triple/quadruple mutants). Like *ALT3* and *ALT4*, *ALT1* and *ALT2* are predicted to be tightly genetically linked (0.04 map units). The CRISPR-Cas system would also need to be used in any line where both of these genes are mutated.

As discussed in Section 4.4.2, the ALT generated fatty acids and β-ketofatty acids may not be metabolic end-points. Identifying the downstream products of these metabolites would be essential in determining the biological roles of *ALT1-4*. Lipidomic experiments with particular attention to medium-chain acyl compounds would be of great interest to such research. As mentioned previously, an experiment towards this goal, lipidomic or otherwise, would need to take into consideration the potentially high volatility of these analytes. Wherever possible, evaporation of solvent should be avoided. Where evaporation cannot be avoided, trapping vapours with a solid phase adsorbent might be possible.

Once the metabolic endpoints of *ALT1-4* derived fatty acids are identified, experiments can be conducted to determine the physiological roles of these genes. As these genes evolved recently in evolutionary history, they likely function in specialized metabolism, that is, for specific plant-environment interactions including with pathogens or pests. As such, *alt* mutants are unlikely to have obvious developmental or morphological phenotypes. The *ALT* genes, as well as several homologs in other species likely arose through tandem gene duplication; *ALT1* and *ALT2* are located very close to each other on chromosome 1, as are *ALT3* and *ALT4*. It has been suggested that genes that arise through this mechanism are more likely to be involved in biotic or abiotic stress responses (Rizzon *et al*., 2006; Hanada *et al.*, 2008). This is consistent with the functions of the *ALT* genes proposed in Section 3.5.3 (*e.g.* insecticidal methyl ketone production).

As discussed in Section 3.5.1, single hot-dog fold proteins are expected to dimerize. Although the ALT proteins are predicted to be active as homodimers (*i.e.* in a heterologous system with no other ALTs), it's possible they form heterodimers *in planta*. Heterodimers between proteins of the hot-dog fold superfamily have been documented previously (Sacco et al., 2007), and expression of *ALT3* overlaps with those of *ALT1, ALT2*, and *ALT4* (Figures 3.3-3.4). The formation of these heterodimers could be tested using a yeast-2-hybrid system (Miller & Stagljar, 2004) or by tandem-affinity-purification, for example. Enzymatic activity of ALT heterodimers could be determined by expression of pairwise combinations of ALT proteins in the *E. coli* expression system used in Chapter 3.

Biotechnological applications of ALT1-4

Medium-chain fatty acids and medium-chain β-ketofatty acids have industrial value. MCFAs have anti-corrosive properties (Kuznetsov & Ibatullin, 2002) and their derivatives can be used for various industrial purposes, including lubricants (Korlipara *et al*., 2011), antimicrobials (Molly & Bruggemann, 2008), bioplastics (Srivastava & Tripathi, 2013), and biodiesel (Knothe, 2008). Methyl ketones are also used or have potential to be used for various commercial applications, including fragrances and flavourings (Goh *et al.*, 2012), insect repellants (Roe, 2002), anesthetics (Papachristoforou *et al*., 2012), and also have potential for use in biofuels (Goh *et al.*, 2012). Some of these compounds are currently extracted from plant sources, but the yield is often low, and as such, there would be significant interest in alternative means of production.

Much effort has already been put into generating medium-chain fatty acids and β-ketofatty acids by heterologous expression of acyl-ACP thioesterases in oilseed crops or microbial systems, as a renewable and efficient means to produce these important chemicals. In addition to characterizing the enzymatic activities of various acyl-ACP thioesterases (Jones *et al.*, 1995; Leonard *et* al., 1998; Jing *et al*., 2011), research has focused on engineering microbial systems for increased yield (Goh *et al.,* 2012; Fan *et al.*, 2013; Xu *et al*., 2013; Cao *et al*., 2014) or for improved utilization of abundant carbon sources (Steen *et al.*, 2010; Wu *et al*., 2014). Transgenic expression of acyl-ACP thioesterases in oilseeds has been performed for close to 20 years, and remains a current research focus (Eccleston *et al*., 1996; Sun *et al.*, 2014). I have already generated constructs and transformed plants to express *ALT1-4* in seeds of Arabidopsis. Analysis of these transgenic lines will provide evidence towards the feasibility of expressing these proteins in commercial oil seed crops.

As shown in Figure 3.2, many potential ALT homologues are found throughout the genomes of the plant kingdom, and as discussed in Section 3.6, their substrate preferences are likely quite varied. There would be significant interest in cataloguing the enzymatic products of these various ALT-like enzymes. The *E. coli* expression system used in section Chapter 3 allows for fast and simple determination of enzymatic products of these enzymes. Once particular acyl-ACP thioesterases have been identified as generating a product of interest, they could be transferred to the desired expression system. For example, monounsaturated MCFAs are a particularly desirable starting material for biodiesel, due to their low viscosities and high oxidative stabilities (Knothe, 2008). Although they may not be biologically relevant

products, ALT1 generates these products when expressed in *E. coli*, making ALT1 an attractive choice for an enzyme in the production of biodiesel.

Characterization of the enzymatic activities of the vast array of potential ALT homologs would also be useful for the engineering of custom thioesterases. Using Arabidopsis ALT1-4 as examples, these proteins share 72-89% amino acid sequence identity in their catalytic domains. Considering the small size of these proteins (141-142 residues in their catalytic domains), only a relatively small number of amino acid differences lead to significantly different compounds being produced in *E. coli*, in terms of chain length, oxidation state, and degree of saturation. This makes these enzymes ideal candidates for directed mutational studies, aimed at engineering amino acid changes that alter enzymatic specificity. Inclusion of the various ALT homologues from other plant species would make this source of knowledge even richer. Amino acid substitution could be even more rationally directed if three-dimensional structures of one or more ALT or ALT-like proteins were solved. Conceivably, this approach would allow for relatively rapid engineering of designer acyl-ACP thioesterases that generate specific products of industrial value.

Reference List

Aarts MG, Keijzer CJ, Stiekema WJ, Pereira A (1995) Molecular characterization of the *CER1* gene of Arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. Plant Cell 7(12):2115-27

Aharoni A, Dixit S, Jetter R, Thoenes E, van Arkel G, Pereira A (2004) The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. Plant Cell 16(9):2463-80

Auldridge ME, Guo Y, Austin MB, Ramsey J, Fridman E, Pichersky E, Noel JP (2012) Emergent decarboxylase activity and attenuation of α/β-hydrolase activity during the evolution of methyl ketone biosynthesis in tomato. Plant Cell 24(4):1596-607

Bais P, Moon SM, He K, Leitao R, Dreher K, Walk T, Sucaet Y, Barkan L, Wohlgemuth G, Roth MR, Wurtele ES, Dixon P, Fiehn O, Lange BM, Shulaev V, Sumner LW, Welti R, Nikolau BJ, Rhee SY, Dickerson JA (2010) PlantMetabolomics.org: a web portal for plant metabolomics experiments. Plant Physiol 152(4):1807-16

Balbi V, Devoto A. (2008) Jasmonate signalling network in Arabidopsis thaliana: crucial regulatory nodes and new physiological scenarios. New Phytol. 177(2):301-18.

Bates PD, Durrett TP, Ohlrogge JB, Pollard M. (2009) Analysis of acyl fluxes through multiple pathways of triacylglycerol synthesis in developing soybean embryos. Plant Physiol. 150(1):55-72.

Bates PD, Stymne S, Ohlrogge J. (2013) Biochemical pathways in seed oil synthesis. Curr Opin Plant Biol. 16(3):358-64.

Beisson F, Li Y, Bonaventure G, Pollard M, Ohlrogge JB (2005) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. Plant Cell 19(1):351-68

Beisson F, Li-Beisson Y, Pollard M (2012) Solving the puzzles of cutin and suberin polymer biosynthesis. Curr Opin Plant Biol 15(3):329-37

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

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods. 9(1):39.

Ben-Israel I, Yu G, Austin MB, Bhuiyan N, Auldridge M, Nguyen T, Schauvinhold I, Noel JP, Pichersky E, Fridman E (2009) Multiple biochemical and morphological factors underlie the production of methyl ketones in tomato trichomes. Plant Physiol 151(4):1952-64

Benning C (2009) Mechanisms of lipid transport involved in organelle biogenesis in plant cells. Annu Rev Cell Dev Biol 25:71-91

Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D, Holden HM (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. Strain CBS-3. J Biol Chem 273(50):33572-9

Bernard A, Joubès J (2012) Arabidopsis cuticular waxes: Advances in synthesis, export and regulation. Prog Lipid Res 52(1):110-129

Bessire M, Chassot C, Jacquat AC, Humphry M, Borel S, Petétot JM, Métraux JP, Nawrath C.. (2002) A permeable cuticle in Arabidopsis leads to a strong resistance to *Botrytis cinerea*. EMBO J. 26(8): 2158–2168.

Bognar AL, Paliyath G, Rogers L, Kolattukudy PE. (1984) Biosynthesis of alkanes by particulate and solubilized enzyme preparations from pea leaves (*Pisum sativum*). Arch Biochem Biophys. 235(1):8-17.

Bonaventure G, Ohlrogge JB. (2002) Differential regulation of mRNA levels of acyl carrier protein isoforms in Arabidopsis. Plant Physiol. 28(1):223-35.

Bonaventure G, Salas JJ, Pollard MR, Ohlrogge JB (2003) Disruption of the *FATB* gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. Plant Cell 15(4):1020-33

Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN. (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. Science 318(5851):801-6.

Brown AP, Affleck V, Fawcett T, Slabas AR. (2006) Tandem affinity purification tagging of fatty acid biosynthetic enzymes in *Synechocystis* sp. PCC6803 and Arabidopsis thaliana. J Exp Bot. 57(7):1563-71.

Browse J, Somerville C. (1991). GLYCEROLIPID SYNTHESIS: Biochemistry and regulation. Annu. Rev. Plant Physiol Plant Mol Biol. 42:467-506.

Bu SY, Mashek DG. (2010) Hepatic long-chain acyl-CoA synthetase 5 mediates fatty acid channeling between anabolic and catabolic pathways. J Lipid Res. 51(11):3270-80.

Cantu DC, Chen Y, Reilly PJ (2010) Thioesterases: a new perspective based on their primary and tertiary structures. Protein Sci 19(7):1281-95

Cao Y, Liu W, Xu X, Zhang H, Wang J, Xian M. (2014) Production of free monounsaturated fatty acids by metabolically engineered Escherichia coli. Biotechnol Biofuels. 7(1):59.

Chevenet F, Brun C, Bañuls AL, Jacq B, Christen R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics. 7:439.

Choi JY, Martin CE. (1999) The *Saccharomyces cerevisiae* FAT1 gene encodes an acyl-CoA synthetase that is required for maintenance of very long chain fatty acid levels. J Biol Chem. 274(8):4671-83.

Clough RC, Matthis AL, Barnum SR, Jaworski JG. (1992) Purification and characterization of 3-ketoacyl-acyl carrier protein synthase-III from spinach: a condensing enzyme utilizing acetyl-coenzyme-A to initiate fatty acid synthesis. J Biol Chem. 267(29):20992–20998.

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J 16(6):735-43

Coe NR, Smith AJ, Frohnert BI, Watkins PA, Bernlohr DA. (1999) The fatty acid transport protein (FATP1) is a very long chain acyl-CoA synthetase. J Biol Chem. 274(51):36300-4.

Compagnon V, Diehl P, Benveniste I, Meyer D, Schaller H, Schreiber L, Franke R, Pinot F (2009) *CYP86B1* is required for very long chain omega-hydroxyacid and alpha, omega -dicarboxylic acid synthesis in root and seed suberin polyester. Plant Physiol 150(4):1831-43

de Azevedo Souza C, Kim SS, Koch S, Kienow L, Schneider K, McKim SM, Haughn GW, Kombrink E, Douglas CJ. (2009) A novel fatty Acyl-CoA Synthetase is required for pollen development and sporopollenin biosynthesis in Arabidopsis. Plant Cell. 21(2):507-25.

De Marcos Lousa C, van Roermund CW, Postis VL, Dietrich D, Kerr ID, Wanders RJ, Baldwin SA, Baker A, Theodoulou FL. (2013) Intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding

cassette transporter is required for transport and metabolism of fatty acids. Proc Natl Acad Sci USA. 110(4):1279-84

Dean GH, Zheng H, Tewari J, Huang J, Young DS, Hwang YT, Western TL, Carpita NC, McCann MC, Mansfield SD, Haughn GW (2007) The Arabidopsis *MUM2* gene encodes a beta-galactosidase required for the production of seed coat mucilage with correct hydration properties. Plant Cell 19(12):4007-21

Dehesh K, Jones A, Knutzon DS, Voelker TA. (1996) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of ChFatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J. 9(2):167-72.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36(Web Server issue):W465-9

Digel M, Ehehalt R, Stremmel W, Füllekrug J. (2009) Acyl-CoA synthetases: fatty acid uptake and metabolic channeling. Mol Cell Biochem. 326(1-2):23-8.

Dillon SC, Bateman A (2004) The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. BMC Bioinformatics 5:109

DiRusso CC, Black PN. (1999) Long-chain fatty acid transport in bacteria and yeast. Paradigms for defining the mechanism underlying this protein-mediated process. Mol Cell Biochem. 192(1-2):41-52

Dirusso CC, Connell EJ, Faergeman NJ, Knudsen J, Hansen JK, Black PN. (2000) Murine FATP alleviates growth and biochemical deficiencies of yeast *fat1*Δ strains. Eur J Biochem. 267(14):4422-33.

DiRusso CC, Darwis D, Obermeyer T, Black PN. (2008) Functional domains of the fatty acid transport proteins: studies using protein chimeras. Biochim Biophys Acta. 2008 Mar;1781(3):135-43.

DiRusso CC, Li H, Darwis D, Watkins PA, Berger J, Black PN. (2005) Comparative biochemical studies of the murine fatty acid transport proteins (FATP) expressed in yeast. J Biol Chem. 280(17):16829-37.

Dobritsa AA, Lei Z, Nishikawa S, Urbanczyk-Wochniak E, Huhman DV, Preuss D, Sumner LW. (2010) LAP5 and LAP6 encode anther-specific proteins with similarity to chalcone synthase essential for pollen exine development in Arabidopsis. Plant Physiol. 153(3):937-55.

Domergue F, Vishwanath SJ, Joubès J, Ono J, Lee JA, Bourdon M, Alhattab R, Lowe C, Pascal S, Lessire R, Rowland O (2010) Three Arabidopsis fatty acyl-coenzyme A reductases, *FAR1*, *FAR4*, and *FAR5*, generate primary fatty alcohols associated with suberin deposition. Plant Physiol 153(4):1539-54

Dörmann P, Voelker TA, Ohlrogge JB (1995) Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* specific for long-chain acyl-acyl carrier proteins. Arch Biochem Biophys 316(1):612-8

Dörmann P, Voelker TA, Ohlrogge JB. (2000) Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1. Plant Physiol. 123(2):637-44.

Dyer JM, Stymne S, Green AG, Carlsson AS. (2008) High-value oils from plants. Plant J. 54(4):640-55.

Eccleston VS, Cranmer AM, Voelker TA, Ohlrogge JB. (1996) Medium-chain fatty acid biosynthesis and utilization in *Brassica napus* plants expressing lauroyl–acyl carrier protein thioesterase. Planta 198: 46–53.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113

Edwards K, Johnstone C, Thompson C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucl. Acid. Res. 19(6): 1349

Emanuelsson O, Nielsen H, Brunak S, von Heijne G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005-1016

Faergeman NJ, Black PN, Zhao XD, Knudsen J, DiRusso CC. (2001) The Acyl-CoA synthetases encoded within FAA1 and FAA4 in *Saccharomyces cerevisiae* function as components of the fatty acid transport system linking import, activation, and intracellular Utilization. J Biol Chem. 276(40):37051-9.

Faergeman NJ, DiRusso CC, Elberger A, Knudsen J, Black PN. (1997) Disruption of the Saccharomyces cerevisiae homologue to the murine fatty acid transport protein impairs uptake and growth on long-chain fatty acids. J Biol Chem. 272(13):8531-8.

Falcon A, Doege H, Fluitt A, Tsang B, Watson N, Kay MA, Stahl A. (2010) FATP2 is a hepatic fatty acid transporter and peroxisomal very long-chain acyl-CoA synthetase. Am J Physiol Endocrinol Metab. 299(3):E384-93.

Fan L, Liu J, Nie K, Liu L, Wang F, Tan T, Deng L. (2013) Synthesis of medium-chain length fatty acid ethyl esters in engineered Escherichia coli using endogenously produced medium-chain fatty acids. Enzyme Microb Technol. 53(2):128-33.

Fellermeier M, Eisenreich W, Bacher A, Zenk MH (2001) Biosynthesis of cannabinoids. Incorporation experiments with (13)C-labeled glucoses. Eur J Biochem 268(6):1596-604

Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK. (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. 2013 Oct;23(10):1229-32.

Finlayson SA, Dennis DT. (1983) Acetyl-coenzyme A carboxylase from the developing endosperm of *Ricinus communis*. I. Isolation and characterization. Arch Biochem Biophys. 225(2):576-85.

Fischer K, Weber A. (2002) Transport of carbon in non-green plastids. Trends Plant Sci. 7(8):345-51.

Fridman E, Wang J, Iijima Y, Froehlich JE, Gang DR, Ohlrogge J, Pichersky E (2005) Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species *Lycopersicon hirsutum* identify a key enzyme in the biosynthesis of methyl ketones. Plant Cell 17(4):1252-67

Fulda M, Schnurr J, Abbadi A, Heinz E, Browse J. (2004) Peroxisomal Acyl-CoA synthetase activity is essential for seedling development in *Arabidopsis thaliana*. Plant Cell. 16(2):394-405.

Fulda M, Shockey J, Werber M, Wolter FP, Heinz E. (2002) Two long-chain acyl-CoA synthetases from *Arabidopsis thaliana* involved in peroxisomal fatty acid beta-oxidation. Plant J. 32(1):93-103.

Füllekrug J, Ehehalt R, Poppelreuther M. (2012) Outlook: membrane junctions enable the metabolic trapping of fatty acids by intracellular acyl-CoA synthetases. Front Physiol. 3:401.

Gau, J-Y, Yu, X-H and Liu, C-J (2009). A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in Arabidopsis. Proc. Natl. Acad. Sci. USA 106: 18855-18860.

Gietz RD, Schiestl RH. (2007) Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2(1):35-7.

Goh EB, Baidoo EE, Keasling JD, Beller HR (2012) Engineering of bacterial methyl ketone synthesis for biofuels. Appl Environ Microbiol 78(1):70-80.

Greer S, Wen M, Bird D, Wu X, Samuels L, Kunst L, Jetter R (2007) The cytochrome P450 enzyme *CYP96A15* is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of Arabidopsis. Plant Physiol 145(3):653-6

Groot PH, Scholte HR, Hülsmann WC. (1976) Fatty acid activation: specificity, localization, and function. Adv Lipid Res. 14:75-126.

Gueguen V, Macherel D, Jaquinod M, Douce R, Bourguignon J. (2000) Fatty acid and lipoic acid biosynthesis in higher plant mitochondria. J Biol Chem. 275(7):5016-25.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3):307-21

Halevy AH, Porat R, Spiegelslein H, Borochov A, Botha L, Whitehead CS (1996) Short-chain saturated fatty acids in the regulation of pollination induced ethylene sensitivity of Phalaenopsis flowers. Physiologia Plantarum 97(3):469-47

Hellyer A, Leadlay PF, Slabas AR. (1992) Induction, purification and characterisation of acyl-ACP thioesterase from developing seeds of oil seed rape (*Brassica napus*). Plant Mol Biol. 20(5):763-80.

Höfer R, Briesen I, Beck M, Pinot F, Schreiber L, Franke R (2008) The Arabidopsis cytochrome P450 *CYP86A1* encodes a fatty acid omega-hydroxylase involved in suberin monomer biosynthesis. J Exp Bot 59(9):2347-60

Holloway GP, Chou CJ, Lally J, Stellingwerff T, Maher AC, Gavrilova O, Haluzik M, Alkhateeb H, Reitman ML, Bonen A. (2011) Increasing skeletal muscle fatty acid transport protein 1 (FATP1) targets fatty acids to oxidation and does not predispose mice to diet-induced insulin resistance. Diabetologia. 54(6):1457-67.

Hooker TS, Millar AA, Kunst L (2002) Significance of the expression of the CER6 condensing enzyme for cuticular wax production in Arabidopsis. Plant Physiol 129:1568-80

Hunt MC, Siponen MI, Alexson SE (2012) The emerging role of acyl-CoA thioesterases and acyltransferases in regulating peroxisomal lipid metabolism. Biochim Biophys Acta 1822(9):1397-410

Hunte C (2005) Specific protein-lipid interactions in membrane proteins. Biochem Soc Trans. 33(Pt 5):938-42. Baldet P, Alban C, Axiotis S, Douce R. (1993) Localization of free and bound biotin in cells from green pea leaves. Arch Biochem Biophys. 303(1):67-73.

Jaworski JG, Clough RC, Barnum SR. (1989) A cerulenin insensitive short chain 3-ketoacyl-acyl carrier protein synthase in *Spinacia oleracea* Leaves. Plant Physiol. 90(1):41-4.

Jenks MA, Ashworth EN (1999) Plant epicuticular waxes: function, production, and genetics. Horticultural Rev. 23: 1-68

Jessen D, Olbrich A, Knüfer J, Krüger A, Hoppert M, Polle A, Fulda M. (2011) Combined activity of LACS1 and LACS4 is required for proper pollen coat formation in Arabidopsis. Plant J. 68(4):715-26.

Jiang W, Yang B, Weeks DP. (2014) Efficient CRISPR/Cas9-Mediated Gene Editing in Arabidopsis thaliana and Inheritance of Modified Genes in the T2 and T3 Generations. PLoS One. 9(6):e99225.

Jing F, Cantu DC, Tvaruzkova J, Chipman JP, Nikolau BJ, Yandeau-Nelson MD, Reilly PJ. (2011) Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. BMC Biochem. 12:44.

Johnsson N, Varshavsky A. (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A. 91(22):10340-4.

Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7(3):359-71

Jordan GE, Piel WH (2008) PhyloWidget: web-based visualizations for the tree of life. Bioinformatics. 24(14):1641-2

Joubès J, Raffaele S, Bourdenx B, Garcia C, Laroche-Traineau J, Moreau P, Domergue F, Lessire R. (2008) The VLCFA elongase gene family in *Arabidopsis thaliana*: phylogenetic analysis, 3D modelling and expression profiling. Plant Mol Biol. 67(5):547-66.

Kachroo A, Shanklin J, Whittle E, Lapchyk L, Hildebrand D, Kachroo P. (2007) The Arabidopsis stearoyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. Plant Mol Biol. 63(2):257-71.

Kage-Nakadai E, Kobuna H, Kimura M, Gengyo-Ando K, Inoue T, Arai H, Mitani S. (2010) Two very long chain fatty acid acyl-CoA synthetase genes, acs-20 and acs-22, have roles in the cuticle surface barrier in *Caenorhabditis elegans*. PLoS One. 5(1):e8857

Kim SS, Douglas CJ. (2013) Sporopollenin monomer biosynthesis in Arabidopsis. J Plant Biol. 56(1):1-6

Knoll LJ, Johnson DR, Gordon JI. (1994) Biochemical studies of three Saccharomyces cerevisiae acyl-CoA synthetases, Faa1p, Faa2p, and Faa3p. J Biol Chem. 269(23):16348-56.

Knothe G. (2008) "Designer" Biodiesel: Optimizing Fatty Ester Composition to Improve Fuel Properties. *Energy Fuels*. 22(2): 1358-64

Knudsen JT, Tollsten L, Bergstrom LG (1993) Floral scents: A checklist of volatile compounds isolated by headspace techniques. Phytochemistry 33(2):253–280

Koo AJ, Fulda M, Browse J, Ohlrogge JB. (2005) Identification of a plastid acyl-acyl carrier protein synthetase in Arabidopsis and its role in the activation and elongation of exogenous fatty acids. Plant J. 44(4):620-32.

Koo AJ, Ohlrogge JB, Pollard M. (2004) On the export of fatty acids from the chloroplast. J Biol Chem. 279(16):16101-10.

Korlipara VP, Bhamidipati VSKR, Potula BP, Rachapudi BNP, Singh AK. Polyol esters of medium-chain fatty acids and process for preparation thereof. U.S. Patent US 20110263885 A1, Oct 27, 2011.

Kuhn DN, Knauf M, Stumpf PK. (1981) Subcellular localization of acetyl-CoA synthetase in leaf protoplasts of *Spinacia oleracea*. Arch Biochem Biophys. 209(2):441-50

Kurata T, Kawabata-Awai C, Sakuradani E, Shimizu S, Okada K, Wada T (2003) The *YORE-YORE* gene regulates multiple aspects of epidermal cell differentiation in Arabidopsis. Plant J 36(1):55-66

Kushnirov VV. (2000) Rapid and reliable protein extraction from yeast. Yeast. 16(9):857-60.

Kuznetsov YI, Ibatullin KA. (2002) On the inhibition of the carbon dioxide corrosion of steel by carboxylic acids. Prot. Met. 38: 439–43

Lager I, Yilmaz JL, Zhou XR, Jasieniecka K, Kazachkov M, Wang P, Zou J, Weselake R, Smith MA, Bayon S, Dyer JM, Shockey JM, Heinz E, Green A, Banas A, Stymne S. (2013). Plant acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J Biol Chem. 288(52):36902-14.

Lehti-Shiu MD, Adamczyk BJ, Fernandez DE. (2005) Expression of MADS-box genes during the embryonic phase in Arabidopsis. Plant Mol Biol. 58(1):89-107.

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013) ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41(Database issue):D423-9

Leonard JM, Knapp SJ, Slabaugh MB. (1998) A Cuphea beta-ketoacyl-ACP synthase shifts the synthesis of fatty acids towards shorter chains in Arabidopsis seeds expressing Cuphea FatB thioesterases. Plant J. 13(5):621-8.

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase *WSD1* required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol 148(1):97-107

Li, Y., Beisson, F., Koo, A.J.K., Molina, I., Pollard, M. and Ohlrogge, J. (2007) Identification of acyltransferases required for cutin synthesis and production of cutin with suberin-like monomers. Proc. Natl. Acad. Sci. U. S. A. 104: 18339–18344.

Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, DeBono A, Durrett TP, Franke RB, Graham IA, Katayama A, Kelly AA, Larson T, Markham JE, Miquel M, Molina I, Nishida I, Rowland O, Samuels L, Schmid KM, Wada H, Welti R, Xu C, Zallot R, and Ohlrogge J. (2013) Acyl-Lipid Metabolism. The Arabidopsis Book 11:e0161. doi:10.1199/tab.0161

Löfke C, Ischebeck T, König S, Freitag S, Heilmann I. (2008) Alternative metabolic fates of phosphatidylinositol produced by phosphatidylinositol synthase isoforms in Arabidopsis thaliana. Biochem J. 2008 413(1):115-24.

Lü S, Song T, Kosma DK, Parsons EP, Rowland O, Jenks MA (2009) Arabidopsis *CER8* encodes *LONG-CHAIN ACYL-COA SYNTHETASE 1* (*LACS1*) that has overlapping functions with LACS2 in plant wax and cutin synthesis. Plant J 59(4):553-64

Markham JE, Jaworski JG. (2007) Rapid measurement of sphingolipids from Arabidopsis thaliana by reversed-phase high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry. Rapid Commun Mass Spectrom. 21(7):1304-14.

Marks MD, Tian L, Wenger JP, Omburo SN, Soto-Fuentes W, He J, Gang DR, Weiblen GD, Dixon RA (2009) Identification of candidate genes affecting delta9-tetrahydrocannabinol biosynthesis in *Cannabis sativa*. J Exp Bot 60(13):3715-26

Martinière A, Shvedunova M, Thomson AJ, Evans NH, Penfield S, Runions J, McWatters HG. (2011) Homeostasis of plasma membrane viscosity in fluctuating temperatures. New Phytol. 192(2):328-37.

Mashek DG, Li LO, Coleman RA. (2007) Long-chain acyl-CoA synthetases and fatty acid channeling. Future Lipidol. 2(4):465-476.

Masterson C, Wood C (2001) Mitochondrial and peroxisomal beta-oxidation capacities of organs from a non-oilseed plant. Proc Biol Sci. 268(1479): 1949–1953.

Mayer KM, Shanklin J (2007) Identification of amino acid residues involved in substrate specificity of plant acyl-ACP thioesterases using a bioinformatics-guided approach. BMC Plant Biol 7:1

Mekhedov S, de Ilárduya OM, Ohlrogge J. (2000) Toward a functional catalog of the plant genome. A survey of genes for lipid biosynthesis. Plant Physiol. 122(2):389-402.

Melton EM, Cerny RL, Watkins PA, DiRusso CC, Black PN. (2011) Human fatty acid transport protein 2a/very long chain acyl-CoA synthetase 1 (FATP2a/Acsvl1) has a preference in mediating the channeling of exogenous n-3 fatty acids into phosphatidylinositol. J Biol Chem. 286(35):30670-9.

Milger K, Herrmann T, Becker C, Gotthardt D, Zickwolf J, Ehehalt R, Watkins PA, Stremmel W, Füllekrug J. (2006) Cellular uptake of fatty acids driven by the ER-localized acyl-CoA synthetase FATP4. J Cell Sci. 119(22):4678-88.

Miller J, Stagljar I. (2004) Using the yeast two-hybrid system to identify interacting proteins. Methods Mol Biol. 261:247-62.

Mogelson S, Pieper SJ, Lange LG. (1984) Thermodynamic bases for fatty acid ethyl ester synthase catalyzed esterification of free fatty acid with ethanol and accumulation of fatty acid ethyl esters. Biochemistry. 1984 23(18):4082-7.

Molina I, Li-Beisson Y, Beisson F, Ohlrogge JB, Pollard M (2009) Identification of an Arabidopsis feruloyl-coenzyme A transferase required for suberin synthesis. Plant Physiol 151(3):1317-28

Molly K, Bruggemann G. Medium-chain fatty acids applicable as antimicrobial agents. U.S. Patent US 20030176500 A1, Sep 8, 2003

Morant M, Jørgensen K, Schaller H, Pinot F, Møller BL, Werck-Reichhart D, Bak S (2007) CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen. Plant Cell 19: 1473–1487.

Moreau P, Bessoule JJ, Mongrand S, Testet E, Vincent P, Cassagne C. (1998) Lipid trafficking in plant cells. Prog Lipid Res. 37(6):371-91.

Moreno-Pérez AJ, Venegas-Calerón M, Vaistij FE, Salas JJ, Larson TR, Garcés R, Graham IA, Martínez-Force E (2012) Reduced expression of *FatA* thioesterases in Arabidopsis affects the oil content and fatty acid composition of the seeds. Planta 235(3):629-39

Mosblech A, Feussner I, Heilmann I. (2009) Oxylipins: structurally diverse metabolites from fatty acid oxidation. Plant Physiol Biochem. 47(6):511-7.

Nardini M, Dijkstra BW (1999) Alpha/beta hydrolase fold enzymes: the family keeps growing. Curr Opin Struct Biol 9(6):732-7

Nelson BK, Cai X, Nebenführ A (2007) A multicolored set of *in vivo* organelle markers for co-localization studies in Arabidopsis and other plants. Plant J 51(6):1126-36

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10(1):1-6

Ohlrogge J, Browse J. (1995) Lipid biosynthesis. Plant Cell. 7(7):957-70.

O'Neill CM, Gill S, Hobbs D, Morgan C, Bancroft I. (2003) Natural variation for seed oil composition in Arabidopsis thaliana. Phytochemistry. 64(6):1077-90.

Overath P, Pauli G, Schairer HU (1969) Fatty acid degradation in *Escherichia coli*. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants. Eur J Biochem 7(4):559-74

Papachristoforou A, Kagiava A, Papaefthimiou C, Termentzi A, Fokialakis N, Skaltsounis AL, Watkins M, Arnold G, Theophilidis G. (2012) The bite of the honeybee: 2-heptanone secreted from honeybee mandibles during a bite acts as a local anaesthetic in insects and mammals. PLoS One. 7(10):e47432.

Pata MO, Hannun YA, Ng CK. (2010) Plant sphingolipids: decoding the enigma of the Sphinx. New Phytol. 185(3):611-30.

Pidugu LS, Maity K, Ramaswamy K, Surolia N, Suguna K. (2009) Analysis of proteins with the 'hot dog' fold: prediction of function and identification of catalytic residues of hypothetical proteins. BMC Struct Biol 9:37

Pollard M, Beisson F, Li Y, Ohlrogge JB. (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci. 3(5):236-46.

Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Séraphin B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods. 24(3):218-29.

Pulsifer IP, Kluge S, Rowland O. (2012) Arabidopsis long-chain acyl-CoA synthetase 1 (LACS1), LACS2, and LACS3 facilitate fatty acid uptake in yeast. Plant Physiol Biochem. 51:31-9.

Pulsifer IP, Lowe C, Narayaran SA, Busuttil AS, Vishwanath SJ, Domergue F, Rowland O. (2014) Acyl-lipid thioesterase1-4 from Arabidopsis thaliana form a novel family of fatty acyl-acyl carrier protein thioesterases with divergent expression patterns and substrate specificities. Plant Mol Biol. 84(4-5):549-63.

Qi Y, Katagiri F. (2009) Purification of low-abundance Arabidopsis plasma-membrane protein complexes and identification of candidate components. Plant J. 57(5):932-44.

Rawsthorne S. (2002) Carbon flux and fatty acid synthesis in plants. Prog Lipid Res. 41(2):182-96

Rieu I, Wolters-Arts M, Derksen J, Mariani C, Weterings K (2003) Ethylene regulates the timing of anther dehiscence in tobacco. Planta 217(1):131-7

Roe M. Method of repelling mosquitoes and ticks using 2-undecanone, 2-tridecanone and/or mixtures thereof. Canadian Patent CA 2441295 A1, Sep 19, 2002.

Rowland O, Zheng H, Hepworth SR, Lam P, Jetter R, Kunst L (2006) *CER4* encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. Plant Physiol 142(3):866-77

Sacco E, Covarrubias AS, O'Hare HM, Carroll P, Eynard N, Jones TA, Parish T, Daffé M, Bäckbro K, Quémard A (2007) The missing piece of the type II fatty acid synthase system from *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 104(37):14628-33

Salas JJ, Ohlrogge JB (2002) Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Arch Biochem Biophys 403(1):25-34

Samach A, Kohalmi SE, Motte P, Datla R, Haughn GW (1997) Divergence of function and regulation of class B floral organ identity genes. Plant Cell 9(4):559–570

Samuels L, Kunst L, Jetter R (2008) Sealing plant surfaces: cuticular wax formation by epidermal cells. Annu Rev Plant Biol. 59:683-707.

Sasaki Y, Nagano Y. (2004) Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. Biosci Biotechnol Biochem. 68(6):1175-84.

Schaffer JE, Lodish HF. (1994) Expression cloning and characterization of a novel adipocyte long chain fatty acid transport protein. Cell. 79(3):427-36.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU. (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet. 37(5):501-6.

Schnurr J, Shockey J, Browse J. (2004) The acyl-CoA synthetase encoded by LACS2 is essential for normal cuticle development in Arabidopsis. Plant Cell. 16(3):629-42.

Schnurr JA, Shockey JM, de Boer GJ, Browse JA. (2002) Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. Plant Physiol. 129(4):1700-9.

Shimakata T, Stumpf PK. (1982) Isolation and function of spinach leaf beta-ketoacyl-[acyl-carrier-protein] synthases. Proc Natl Acad Sci U S A. 79(19):5808-12.

Shockey JM, Fulda MS, Browse J. (2003) Arabidopsis contains a large superfamily of acyl-activating enzymes. Phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme a synthetases. Plant Physiol. 132(2):1065-76.

Shockey JM, Fulda MS, Browse JA. (2002) Arabidopsis contains nine long-chain acyl-coenzyme a synthetase genes that participate in fatty acid and glycerolipid metabolism. Plant Physiol. 129(4):1710-22.

Singer T, Fan Y, Chang HS, Zhu T, Hazen SP, Briggs SP. (2006) A high-resolution map of Arabidopsis recombinant inbred lines by whole-genome exon array hybridization. PLoS Genet. 2(9):e144.

Soupene E, Kuypers FA. (2008) Mammalian long-chain acyl-CoA synthetases. Exp Biol Med (Maywood). 233(5):507-21.

Sparkes IA, Runions J, Kearns A, Hawes C (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat Protoc 1(4):2019-25

Sperling P, Heinz E. (2003) Plant sphingolipids: structural diversity, biosynthesis, first genes and functions. Biochim Biophys Acta. 1632(1-3):1-15.

Srivastava SK, Tripathi AD. (2013) Effect of saturated and unsaturated fatty acid supplementation on bio-plastic production under submerged fermentation. 3 Biotech. 3(5): 389-97.

Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD. (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature. 463(7280):559-62.

Steinberg SJ, Morgenthaler J, Heinzer AK, Smith KD, Watkins PA. (2000) Very long-chain acyl-CoA synthetases. Human "bubblegum" represents a new family of proteins capable of activating very long-chain fatty acids. J Biol Chem. 275(45):35162-9.

Suh MC, Samuels AL, Jetter R, Kunst L, Pollard M, Ohlrogge J, Beisson F. (2005) Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. Plant Physiol. 139(4):1649-65.

Sun JY, Hammerlindl J, Forseille L, Zhang H, Smith MA. (2014). Simultaneous over-expressing of an acyl-ACP thioesterase (FatB) and silencing of acyl-acyl carrier protein desaturase by artificial microRNAs increases saturated fatty acid levels in Brassica napus seeds. Plant Biotech J. 12(5): 624-37.

Tan YY, Xu HH, Tao WJ, Hoffmann MH, Wang XF, Lu YT. (2005) Transgenic GFP as a molecular marker for approaches to quantify pollination mechanism and gene flow in *Arabidopsis thaliana*. Plant Biol (Stuttg) 7(4):405-10

Tang D, Simonich MT, Innes RW. (2007) Mutations in LACS2, a long-chain acyl-coenzyme A synthetase, enhance susceptibility to avirulent *Pseudomonas syringae* but confer resistance to *Botrytis cinerea* in Arabidopsis. Plant Physiol. 144(2):1093-103.

Tetko IV, Gasteiger J, Todeschini R, Mauri A, Livingstone D, Ertl P, Palyulin VA, Radchenko EV, Zefirov NS, Makarenko AS, Tanchuk VY, Prokopenko VV. (2005) Virtual computational chemistry laboratory--design and description. J Comput Aided Mol Des. 19(6):453-63.

Theodoulou FL, Eastmond PJ. (2012) Seed storage oil catabolism: a story of give and take. Curr Opin Plant Biol. 15(3):322-8.

Tilton GB, Shockey JM, Browse J (2004) Biochemical and molecular characterization of ACH2, an acyl-CoA thioesterase from *Arabidopsis thaliana*. J Biol Chem 279(9):7487-94

Tjellström H, Strawsine M, Silva J, Cahoon EB, Ohlrogge JB (2013) Disruption of plastid acyl:acyl carrier protein synthetases increases medium-chain fatty acid accumulation in seeds of transgenic Arabidopsis. FEBS Lett 587(7):936-42

Toufighi K, Brady SM, Austin R, Ly E, Provart NJ (2005) The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. Plant J 43(1):153-63

van Poecke RM, Posthumus MA, Dicke M (2001) Herbivore-induced volatile production by *Arabidopsis thaliana* leads to attraction of the parasitoid *Cotesia rubecula*: chemical, behavioral, and gene-expression analysis. J Chem Ecol 27(10):1911-28

Vas G, Vékey K. (2004). Solid-phase microextraction: a powerful sample preparation tool prior to mass spectrometric analysis. J Mass Spectrom. 39(3):233-54.

Voelker TA, Davies HM. (1994) Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. J Bacteriol. 176(23):7320-7.

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE, Davies HM (1992) Fatty acid biosynthesis redirected to medium-chains in transgenic oilseed plants. Science 257(5066):72-4

Wada H, Shintani D, Ohlrogge J. (1997) Why do mitochondria synthesize fatty acids? Evidence for involvement in lipoic acid production. Proc Natl Acad Sci USA. 94(4):1591-6.

Wain RL, Wightman F. (1954) The growth regulating activity of certain omega-substituted alkyl carboxylic acids in relation to their beta-oxidation within the plant. Proc R Soc Lond B Biol Sci. 142(909):525-36.

Watkins PA, Lu JF, Steinberg SJ, Gould SJ, Smith KD, Braiterman LT. (1998) Disruption of the *Saccharomyces cerevisiae* FAT1 gene decreases very long-chain fatty acyl-CoA synthetase activity and elevates intracellular very long-chain fatty acid concentrations. J Biol Chem. 273(29):18210-9.

Watkins PA. (2008) Very-long-chain acyl-CoA synthetases. J Biol Chem. 283(4):1773-7.

Weaire PJ, Kekwick RG. (1975) The synthesis of fatty acids in avocado mesocarp and cauliflower bud tissue. Biochem J. 146(2):425-37.

Weng H, Molina I, Shockey J, Browse J. (2010) Organ fusion and defective cuticle function in a *lacs1 lacs2* double mutant of Arabidopsis. Planta. 231(5):1089-100.

Whitehead CS, Halevy AH (1989) Ethylene sensitivity: The role of short-chain saturated fatty acids in pollination-induced senescence of *Petunia hybrida* flowers. Plant Growth Regulation 8(1):41-54

Whitehead CS, Vasiljevic D (1993) Role of short-chain saturated fatty acids in the control of ethylene sensitivity in senescing carnation flowers. Physiologia Plantarum 88(20):243-250

Wiszniewski AA1, Zhou W, Smith SM, Bussell JD. (2009) Identification of two Arabidopsis genes encoding a peroxisomal oxidoreductase-like protein and an acyl-CoA synthetase-like protein that are required for responses to pro-auxins. Plant Mol Biol. 69(5):503-15.

Wu H, Karanjikar M, San KY. (2014) Metabolic engineering of Escherichia coli for efficient free fatty acid production from glycerol. Metab Eng. doi: 10.1016/j.ymben.2014.06.009. [Epub ahead of print]

Xu P, Gu Q, Wang W, Wong L, Bower AG, Collins CH, Koffas MA. (2013) Modular optimization of multi-gene pathways for fatty acids production in E. coli. Nat Commun. 2013;4:1409.

Yu G, Nguyen TT, Guo Y, Schauvinhold I, Auldridge ME, Bhuiyan N, Ben-Israel I, Iijima Y, Fridman E, Noel JP, Pichersky E (2010) Enzymatic functions of wild tomato METHYL KETONE *SYNTHASES 1* and *2*. Plant Physiol 154(1):67-77

Zhan T, Poppelreuther M, Ehehalt R, Füllekrug J. (2012) Overexpressed FATP1, ACSVL4/FATP4 and ACSL1 increase the cellular fatty acid uptake of 3T3-L1 adipocytes but are localized on intracellular membranes. PLoS One. 7(9):e45087

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(6):1048-58.

Zolman BK, Silva ID, Bartel B. (2001) The Arabidopsis pxa1 mutant is defective in an ATP-binding cassette transporter-like protein required for peroxisomal fatty acid beta-oxidation. Plant Physiol. 127(3):1266-78.

Zolman BK, Yoder A, Bartel B. (2000) Genetic analysis of indole-3-butyric acid responses in Arabidopsis thaliana reveals four mutant classes. Genetics. 156(3):1323-37.

Zou Z, DiRusso CC, Ctrnacta V, Black PN. (2002) Fatty acid transport in Saccharomyces cerevisiae. Directed mutagenesis of FAT1 distinguishes the biochemical activities associated with Fat1p. J Biol Chem. 277(34):31062-71.

Zou Z, Tong F, Faergeman NJ, Børsting C, Black PN, DiRusso CC. (2003) Vectorial acylation in *Saccharomyces cerevisiae*. Fat1p and fatty acyl-CoA synthetase are interacting components of a fatty acid import complex. J Biol Chem. 278(18):16414-22.

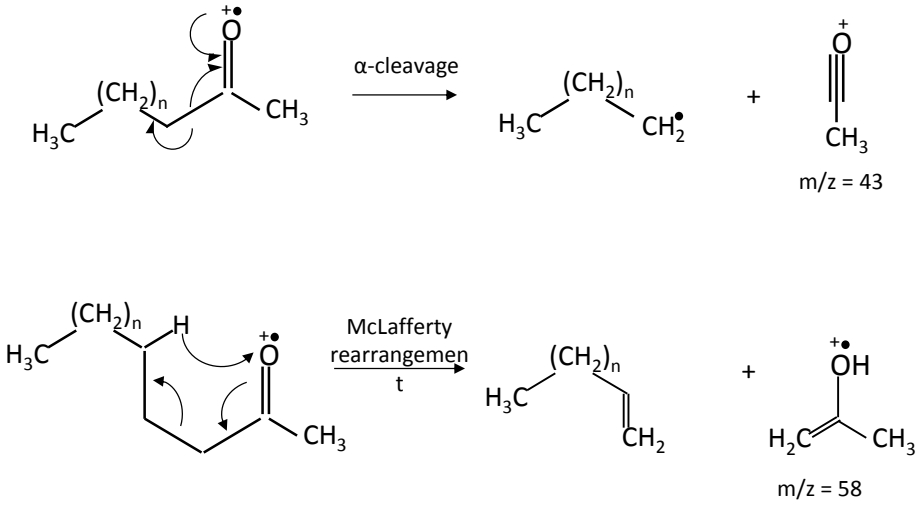
Appendix I List of oligonucleotide primers in this thesis

Chapter 2	
Yeast expression	
His-T7-tag-For	AGAAAGCTTCCATGGGCAGCAGCCATC
His-T7-tag-Rev	GCTAGTTATTGCTCAGCGG
LACS1-For-BamHI	AGAGGATCCATGAAGTCTTTTGCGGCTAAG
LACS1-Rev-NotI	TCTGCGGCCGCTCAGATTTTCTTTGAGGCCAAT
LACS2-For-BamHI	AGAGGATCCATGGCGACTGGTCGATACAT
LACS2-Rev-NotI	TCTGCGGCCGCCTATGCCATGGACCTCTTTG
LACS3-For-BamHI	AGAGGATCCATGGCGACTGGTCGATACAT
LACS3-Rev-NotI	TCTGCGGCCGCTTACACTCGTAGCTGCACTTC
FAT1p-For-BamHI	AGAGGATCCATGTCTCCCATACAGGTTG
FAT1p -Rev-NotI	TCTGCGGCCGCTCATAATTTAATTGTTTGTGCATCGA
FAA1p-For-NotI	TTTGCGGCCGCTATGGTTGCTCAATATACCGTTC
FAA1p -Rev-SphI	AAAGCATGCTTAAGACGAACTATAAACGGCG
Chapter 3	
Promoter::GUS	
ALT1-Prom-For-SalI	AAAGTCGACATGCATCGTTCACCATATCCAC
ALT1-Prom-Rev-BamHI	AAAGGATCCTGTGCCGGTAGCTTTAAGCAT
ALT2-Prom-For-SalI	AAAGTCGACCACCTTCTCGTAATACAAAGTTTCTC
ALT2-Prom-Rev-BamHI	AAAGGATCCCGTGCTGGTAGCTTGAAACA
ALT4-Prom-For-SalI	AAAGTCGACCATTTAGAGCGATCCCTTACTCATG
ALT4-Prom-Rev-BamHI	AAAGGATCCTGAAGTCGGAAACACCACAGACATA
qPCR	
ALT1-qPCR-For	GAATGGCTGTGTGGCTTG
ALT1-qPCR-Rev	TTTGTCGAGCCCTAGATCAAT
ALT2-qPCR-For	TGTGTGGCTTGACAACAAGTACC
ALT2-qPCR-Rev	TTGTGATGCCCTAGAATGTGG
ALT3-qPCR-For	CTGTTCGCATCCCATCTTCT
ALT3-qPCR-Rev	GCTAATGACAGGCTTATTAAGTAGATTC
ALT4-qPCR-For	CCTGTTCGTATCCCATCTTCG
ALT4-qPCR-Rev	ACGATGAGAGAAATACGATGTTGC
ACT2-qPCR-For	CCGAGCAGCATGAAGATTAAG
ACT2-qPCR-Rev	CATACTCTGCCTTAGAGATCCACA
eIF4a-qPCR-For	CTGATTTTGACCCGTCGTCT
eIF4a -qPCR-Rev	AAGACAAACAACAAAGCCGAAT
GAPC-qPCR-For	TCAGACTCGAGAAAGCTGCTAC
GAPC-qPCR-Rev	CGAAGTCAGTTGAGACAACATCATC
GFP fusions	
ALT1-ORF-For-XbaI	AAATCTAGAATGCTTAAAGCTACCGGCACAG
ALT1-ORF -RevNoStop -BamHI	AAAGGATCCATATTCGACAACGTGTTGACGT
ALT1-ORF-Trunc-For-XbaI	AAATCTAGAATGAATGGAGTCCATGAGATTGAA
ALT1-ORF-trunc-RevNoStop-BamHI	AAAGGATCCTCCTTTGCCTCCTTGCTGTT
ALT2-ORF-For-XbaI	AAATCTAGAATGTTTCAAGCTACCAGCAC
ALT2-ORF-RevNoStop -BamHI	AAAGGATCCATCGACCAAGTGTTGACA
ALT3-ORF-For-BamHI	AAAGGATCCATGTTTCTTCAGGTTACCGGCA
ALT3-ORF-RevNoStop -BamHI	AAAGGATCCAACGGCGTCGTCTTGGCG
ALT4-ORF-For-BamHI	ATAGAATTCATGATTCGGGTTACCG
ALT4-ORF-RevNoStop -BamHI	AAAGGATCCAACTGTGTCGTTTTGGC

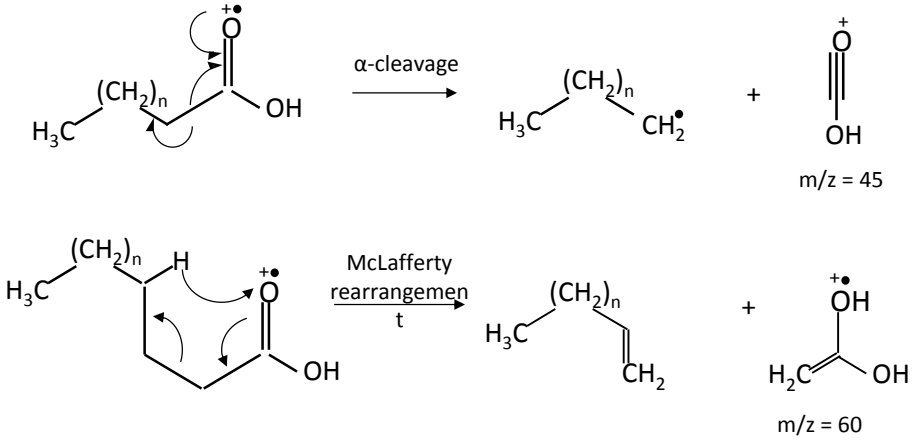
Appendix I

***E. coli* expression**	
ALT1-ORF-Trunc-For-BamHI	ATAGGATCCATGAATGGAGTCCATGAGATTGAA
ALT1-ORF-Rev-EcoRI	AAAGAATTCTCAATATTCGACAACGTGTTGAC
ALT2-ORF-Trunc-For-BamHI	ATAGGATCCATCCATTGGTGGGTTCCATGAGATTGAA
ALT2-ORF-Rev-EcoRI	AAAGAATTCTCAATCGACCAACTGTTGAC
ALT3&4-ORF-Trunc-For-BamHI	ATAGGATCCATGAGTGAGTTCCATGAGGTTGA
ALT3-ORF-Rev-EcoRI	AAAGAATTCTCAAACGGCGTCGTCTTGG
ALT4-ORF-Rev-EcoRI	AAAGAATTCTCAAACTGTGTCGTTTTGGCGT
FATB-ORF-trunc-For-SacI	AAGAGCTCTTACCTGACTGGAGCATGCTTCTTGC
FATB-ORF-Rev-HindIII	GCAAGCTTGGTAGTAGCAGATATAGTT
ALT1-mutant-DA-For	TTATGAATTAGCCCAATTTGGTG
ALT1-mutant-DA-Rev	CACCAAATTGGGCTAATTCATAA
Chapter 4	
PCR genotyping	
ALT1 -LP	AAAAACAAAAACTGTCCACCTC
ALT1 -RP	GTCTATAATGGCCAAGAGGGC
ALT2- LP	TGACACTGGAAGTGACCA
ALT2- RP	ACATTTACCTTCTCACGGC
ALT3- LP	TCATCCAAATATTTGTTTTATTTTATGG
ALT3- RP	GCTAATGACAGGCTTATTAAGTAGATTC
LBb1.3	ATTTTGCCGATTTCGGAAC
RT-PCR	
ALT1-RT-For	AAAGGATCCATGCTTAAAGCTACCGGCACAG
ALT1-RT-Rev	TTTGTCGAGCCCTAGATCAAT
ALT2-RT-For	AAAGGATCCATGTTTCAAGCTACCAGCAC
ALT2-RT-Rev	TTGTGATGCCCTAGAATGTGG
ALT3-RT-For	AAAGGATCCATGTTTCTTCAGGTTACCGGCA
ALT3-RT-Rev	GCTAATGACAGGCTTATTAAGTAGATTC
GAPC-RT-For	TCAGACTCGAGAAAGCTGCTAC
GAPC-RT-Rev	CGAAGTCAGTTGAGACAACATCATC

Appendix II Mass spectra of lipids extracted from spent media of *E. coli* cultures carrying an empty vector, or expressing ALT1-4, or FATB. The first page describes some fragmentations mechanisms. The subsequent pages are the mass spectra of compounds identified in this study.

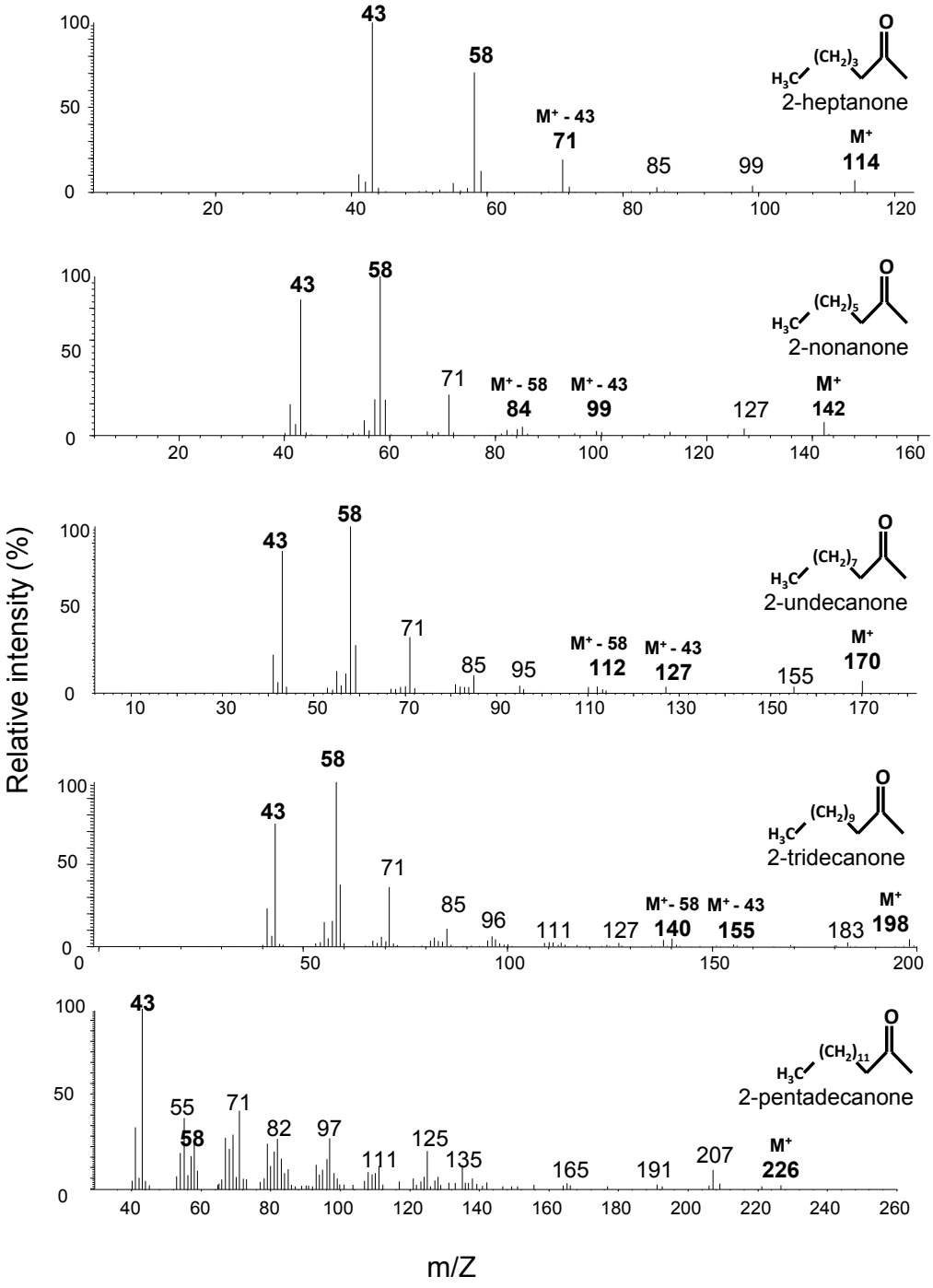


Fragmentation mechanisms showing diagnostic ions of methylketones

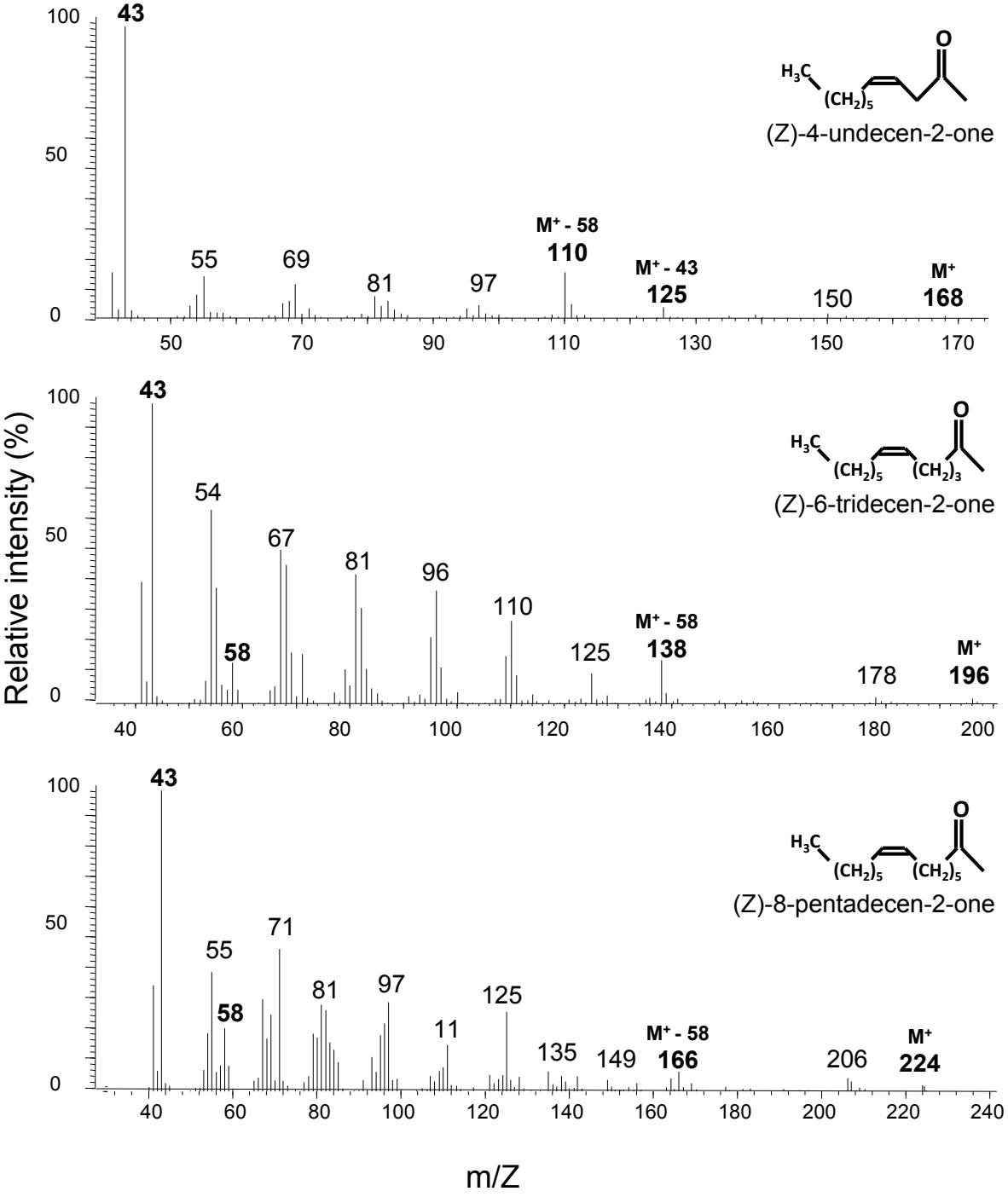


Fragmentation mechanisms showing diagnostic ions of fatty acids

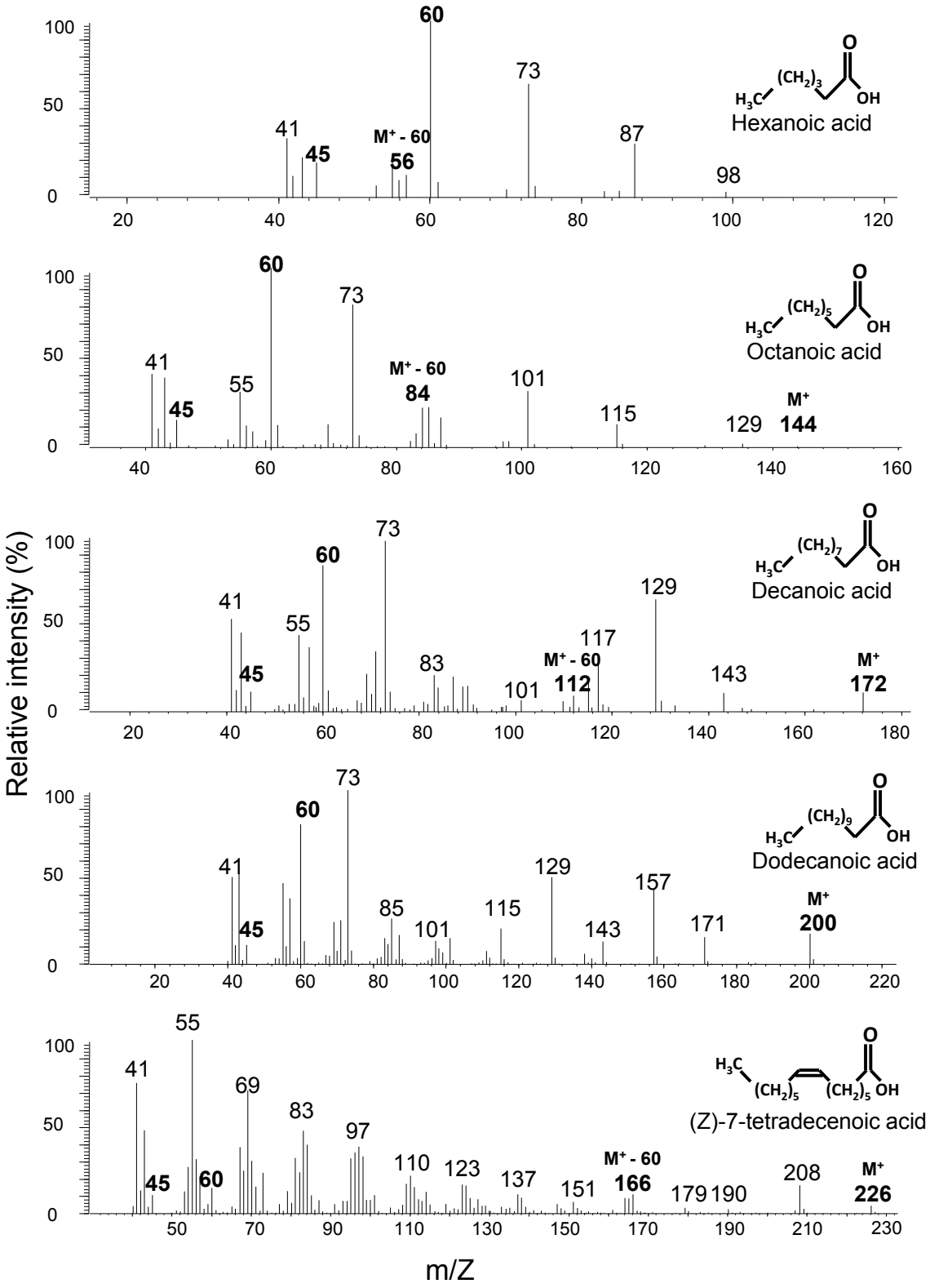
Appendix II


Relative intensity (%)
m/Z
2-heptanone
2-nonanone
2-undecanone
2-tridecanone
2-pentadecanone
43
58
71
85
99
114
127
142
155
170
183
198
207
226
M+ - 43
M+ - 58
M+

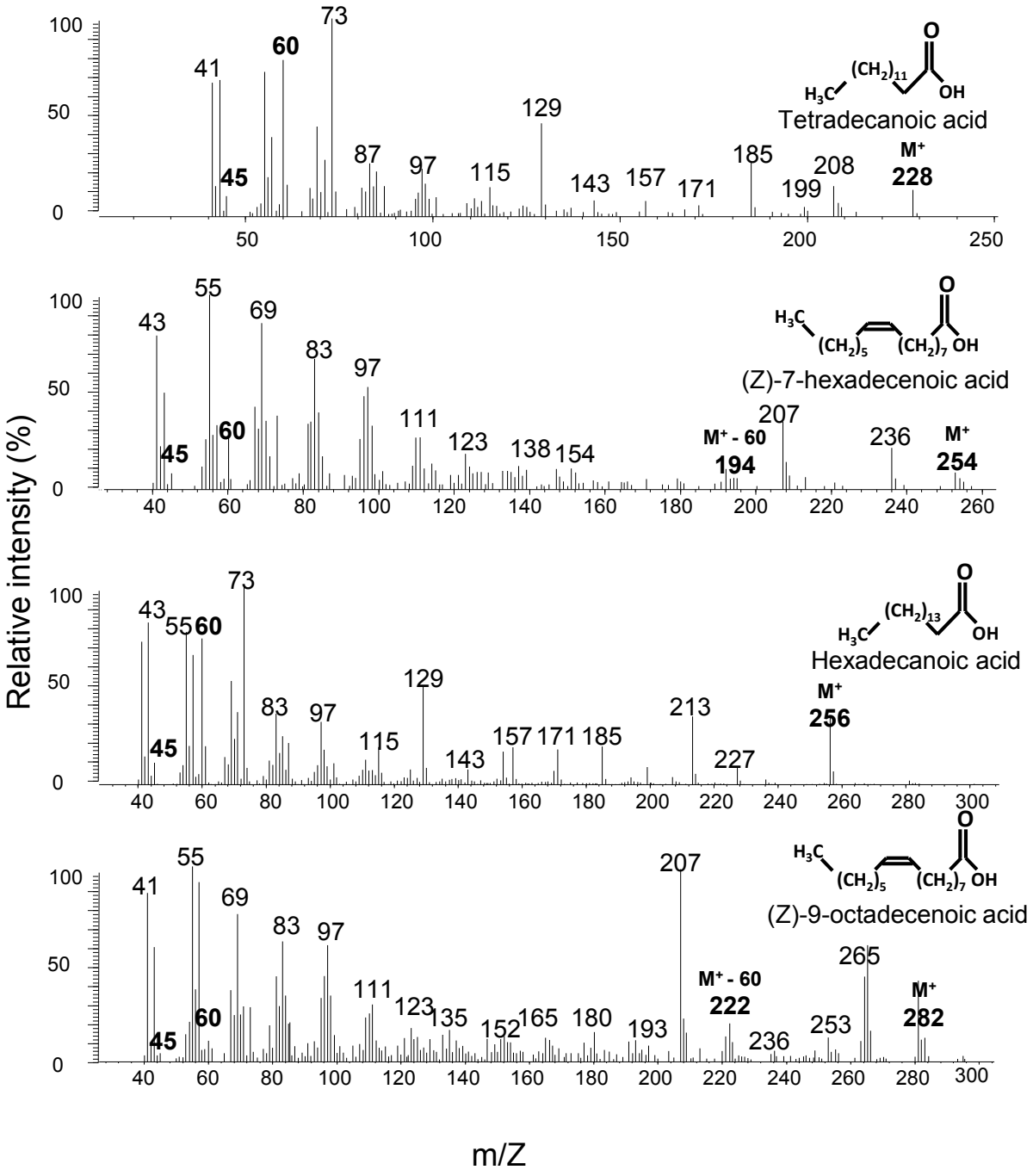
Appendix II


43
55
69
81
97
M+ - 58
110
M+ - 43
125
150
M+
168
(Z)-4-undecen-2-one
43
54
58
67
81
96
110
125
M+ - 58
138
178
M+
196
(Z)-6-tridecen-2-one
43
55
58
71
81
97
11
125
135
149
M+ - 58
166
206
M+
224
(Z)-8-pentadecen-2-one
Relative intensity (%)
m/Z

Appendix II


Hexanoic acid
Octanoic acid
Decanoic acid
Dodecanoic acid
(Z)-7-tetradecenoic acid
Relative intensity (%)
m/Z

Appendix II


Relative intensity (%)
m/Z
Tetradecanoic acid
M+ 228
(Z)-7-hexadecenoic acid
M+ - 60 194
M+ 254
Hexadecanoic acid
M+ 256
(Z)-9-octadecenoic acid
M+ - 60 222
M+ 282