J Plant Biotechnol (2019) 46:282–290
DOI:https://doi.org/10.5010/JPB.2019.46.4.282

ISSN 1229-2818 (Print)
ISSN 2384-1397 (Online)

Research Article

Overexpression of *Cuphea viscosissima CvFatB4* enhances 16:0 fatty acid accumulation in *Arabidopsis*

Jinouk Yeon · Jong-Sug Park · Sang Ho Lee · Kyeong-Ryeol Lee · Hankuil Yi

Received: 23 July 2019 / Revised: 9 December 2019 / Accepted: 9 December 2019

Abstract *Cuphea viscosissima* plants accumulate medium-chain fatty acids (MCFAs), i.e., those containing 8 ~ 14 carbons, in their seeds, in addition to the longer carbon chain fatty acids (≥16 carbons) found in a variety of plant species. Previous studies have reported the existence of three *C. viscosissima* MCFA-producing acyl-acyl carrier protein (ACP) thioesterases with different substrate specificities. In this study, *CvFatB4*, a novel cDNA clone encoding an acyl-ACP thioesterase (EC 3.1.2.14), was isolated from developing *C. viscosissima* seeds. Sequence alignment of the deduced amino acid sequence revealed that four catalytic residues for thioesterase activity are conserved and a putative N-terminal chloroplast transit peptide is present. Overexpression of *CvFatB4* cDNA, which was under the control of the cauliflower mosaic virus 35S promoter, in *Arabidopsis thaliana* led to an increase in 16:0 fatty acid (palmitate) levels in the seed oil at the expense of 18:1 and other non-MCFAs.

Keywords Acyl-acyl carrier protein thioesterase, *Cuphea viscosissima*, *FatB*, *CvFatB4*, 16:0 fatty acid

Introduction

The *de novo* synthesis of fatty acids in plants occurs in plastids, where the reducing chemicals and ATP required for fatty acid synthesis are generated by photosynthetic electron transport (Harwood 1988). The fatty acids are sequentially elongated through a series of two-carbon additions to a growing acyl chain in reactions catalyzed by acetyl-CoA carboxylase and fatty acid synthetase (Slabas and Fawcett 1992). Fatty acid chain elongation is terminated by acyl-acyl carrier protein (ACP) thioesterase, which catalyzes the hydrolysis of thioester bond and releases ACP and free fatty acid (Davies 1993; Stumpf 1987). Subsequently, free fatty acids can be exported outside plastids, reactivated to acyl-CoAs and utilized for the eukaryotic pathway of lipid metabolism. If the acyl-ACPs are used directly by the plastid acyltransferases, a prokaryotic pathway occurs (Somerville and Browse 1991; Thompson et al. 1986). Therefore, acyl-ACP thioesterases play a crucial role in determining the fatty acid chain length and allocating acyl chains between the prokaryotic and eukaryotic lipid biosynthesis pathways in plants. In plants, acyl-ACP thioesterases are divided into two different classes, termed FatA and FatB, which are responsible for the release of 18:1 and saturated fatty acids, respectively (Jones et al. 1995; Salas and Ohlrogge 2002). Medium-chain fatty acids (MCFAs) containing 8–14 carbons are used as raw materials for the industrial production of detergents, soap, and lubricants. They also have the potential to be hydrocarbon components in kerosene-type jet fuel along with the 16:0 fatty acid (palmitic acid) (Kim et al. 2015). To date, dozens of acyl-ACP thioesterases, which were isolated from seeds that accumulate high levels of MCFAs, have been functionally characterized (Beermann et al. 2007; Dehesh et al. 1996a, b; Filichkin et al. 2006; Jing et al. 2011; Jones et al. 1995; Kim et al. 2015; Leonard et al. 1997; Nam et al. 2019; Tjellström et al. 2013; Töpfer 1995; Voelker et al. 1992). For example, *UcFatB* cDNA from *Umbellularia californica* (California bay) seeds has been shown to encode a 12:0-ACP thioesterase

J. Yeon · H. Yi (✉)
Department of Biological Sciences, Chungnam National University, Daejeon, Republic of Korea
e-mail: hankuil.yi@cnu.ac.kr

J.-S. Park · K.-R. Lee (✉)
Department of Agricultural Biotechnology, National Institute of Agricultural Science, Rural Development Administration, Jeonju, Republic of Korea
e-mail: realdanny@korea.kr

S. H. Lee
Division of Biomedical Engineering & Health Science Management, Mokwon University, Daejeon, 35349, Republic of Korea

and produces lauric acid in the seeds when expressed in *Arabidopsis thaliana* (Arabidopsis) and *Brassica napus* (rapeseed) (Davies et al. 199; Knutzon et al. 1999; Pollard et al. 1991; Voelker et al. 1992). In addition, the transgenic expression of a *ChFatB2* cDNA from *Cuphea hookeriana*, a plant with 50 mol% caprylate (8:0) and 25 mol% caprate (10:0) in its seed oil, resulted in a dramatic increase of these two fatty acids (8:0 and 10:0) in rapeseed that normally does not accumulate any 8:0 and 10:0 (Dehesh et al. 1996; Voelker et al. 1996). However, unexpected or broader than expected substrate specificities in transgenic plants have also been reported when some thioesterase cDNAs were heterologously expressed. For instance, *CwFatB1* and *CwFatB2* from *C. wrightii* showed 10:0, 12:0, 14:0, and 16:0-ACP thioesterase activities in transgenic Arabidopsis (Leonard et al. 1997). These substrate specificities were broader than expected, considering *C. wrightii* seeds contain 29 mol% 10:0 and 54 mol% 12:0. It is possible that endogenous acyl-ACP synthetases reactivate MCFAs, which are generated by heterologous expression of ACP thioesterase (Tjellström et al. 2013; Voelker et al. 1997). Nonetheless, heterologous expression provides a useful way to determine fatty acid chain length preference of a given acyl-ACP thioesterase and to engineer fatty acid composition in transgenic plants.

Cuphea, a member of the Lythraceae, is a dicot genus of 260 species, whose seed oils contain high levels of MCFAs (Graham et al. 1981). Substrate specificity or ranges of *Cuphea* seed acyl-ACP thioesterase were characterized using heterologous systems, such as *E. coli*, Arabidopsis, or rapeseed systems (Beermann et al. 2007; Dehesh et al. 1996a, b; Filichkin et al. 2006; Jing et al. 2011; Jones et al. 1995; Kim et al. 2015; Leonard et al. 1997; Tjellström et al. 2013). *Cuphea viscosissima*, known as blue waxweed, accumulates predominantly 8:0 and 10:0 fatty acids in its seed oil (Knapp and Tagliani 1991). To date, three thioesterase cDNAs (*CvFatB1*, *CvFatB2*, and *CvFatB3*) have been isolated from *C. viscosissima* (Jing et al. 2011). In *E. coli* strain K27, CvFatB1 has thioesterase activity that is specific for 8:0- and 10:0-ACP substrates, whereas CvFatB2 and CvFatB3 show 14:0/16:0/16:1-ACP and 14:0-ACP specific activity, respectively.

In this study, we isolated and characterized a new thioesterase cDNA, *CvFatB4*, from *C. viscosissima*. The biochemical characterization, using transgenic Arabidopsis, revealed that CvFatB4 is a thioesterase that prefers 16:0-ACP as a substrate, although CvFatB4 is expressed in developing *C. viscosissima* seeds that are accumulating MCFA. Identification of the fourth acyl-ACP thioesterase from *C. viscosissima* with different biochemical properties will help us understand the molecular basis underlying the different substrate preferences of the four acyl-ACP thioesterases in *C. viscosissima*.

Materials and Methods

Plant material

Seeds from *C. viscosissima* (PI 534726) were obtained from the North Central Regional Plant Introduction Station (NCRPIS, Ames, USA). Developing *C. viscosissima* seeds were isolated from plants grown under greenhouse conditions in Jeonju, Republic of Korea.

Cloning of acyl-ACP thioesterase cDNAs from *C. viscosissima*

A new acyl-ACP thioesterase cDNA from *C. viscosissima* was cloned using the PCR cloning strategy. Multiple sequence alignments of seven *Cuphea* Acyl-ACP thioesterase protein sequences (GenBank accession numbers: DQ238594, DQ238595, U17076, AF062399, AJ131739, AJ131740, and U56103) revealed two conserved regions with amino acid sequences NDGFGRTP and WNDLDVNQHV. Two primers, A14-F (5'- AATGACGGCTTTGGTCGTACTCCT-3') and A14-R (5'-ACGTGCTGATTGACATCCAAGTCATTCCAC-3') were designed based on these sequences. Total RNA was extracted from 100 mg of developing *C. viscosissima* seeds using Plant RNA Reagent (Invitrogen, San Jose, USA). The RNA was reverse-transcribed to cDNA using the SuperScript™ first-strand synthesis system (Invitrogen, Carlsberg, USA). A PCR was performed in a 50 μl reaction mixture containing 20 ng cDNA, 1X *Pfx* buffer, 1 mM $MgSO_4$, 0.3 mM dNTP, 1 pmole of each primer, and 0.5 U Platinum ™ *Pfx* polymerase (Invitrogen, Carlsbad, USA). The PCR reaction program was as follows: (1) 94°C for 5 min; (2) 35 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 60 s; and (3) 72°C for 10 min. The expected ~ 400 bp product was detected by agarose gel electrophoresis, recovered using a QiaQuick gel extraction kit (Qiagen, Valencia, USA), and cloned into the pGEM- T easy vector (Promega, Madison, USA) for sequence identification. The missing 5'- and 3'-parts of the coding sequences were obtained using a GeneRacer kit that utilized the gene-specific primers designed from the cloned 417 bp fragment sequences (Invitrogen, Carlsbad, USA). The full length (1,874 bp) *CvFatB4* cDNA sequence was assembled by MegAlign, which is part of the DNASTAR® software (DNAstar, Madison, USA).

Sequence analysis

Acyl-ACP thioesterase amino acid sequences from various plant species were obtained from the NCBI database (https://www.ncbi.nlm.nih.gov/protein/). The protein sequences were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), and the phylogenetic relationship was calculated by the maximum likelihood method based on the JTT matrix-based model in MEGA 6 with the default option, but without the bootstrap test (replication number = 1000) (Tamura et al. 2013).

Construction of the plant transformation vector

The *p35S::CvFatB4* vector expressing *CvFatB4* under the control of the *cauliflower mosaic virus* (*CaMV*) *35S* promoter was constructed as follows. The coding region of the *CvFatB4* cDNA was PCR-amplified with 5'-AACCATGG TGGCCACCGCTGCAAGTTCT-3' (underlined sequence: appended *Nco*I site for cloning) and 5'-CCGCGGCCGC TAAGAGTCTCCATGTGAGGTC-3' (underlined sequence: appended *Not*I site for cloning) primers. The PCR product was combined with the *CaMV 35S* promoter and the *NOS* terminator expression cassette using the *Nco*I and *Not*I restriction enzymes. The expression cassette was placed into the *Eco*RI and *Xba*I sites in the pCAMBIA 3300 vector (CAMBIA, Brisbane, Australia). The *p35S::CvFatB4* vector was introduced into *Agrobacterium tumefaciens* strain EHA105 using the freeze-thaw method.

Arabidopsis transformation

The Arabidopsis Columbia accession was transformed using the *Agrobacterium*-mediated floral dipping method (Clough and Bent 1998; Martinez-Trujillo et al. 2004). The modified floral dip inoculation medium consisted of 5% sucrose (w/v) and 0.05% triton-X 100 (v/v). The plants were inoculated when most of the secondary inflorescences, which emerged after the primary inflorescences had been clipped, were about 5 to 10 cm tall. The flower was inoculated using a micro-pipette and the drop-by-drop method. Herbicide-resistant plants were selected using BASTA™ (Bayer Crop Science Korea, Seoul, Republic of Korea).

Genomic Southern blot analysis

A Southern blot analysis was carried out to determine the copy numbers of the *CvFatB4* transgene in individual Arabidopsis transgenic lines. The genomic DNA was extracted from 1 g of 3-week-old Arabidopsis leaves using a urea-phenol extraction method, as described above (Yi and Richards, 2009). Two micrograms of genomic DNA was digested with *Xba*I, separated by electrophoresis at 30 V in a 0.8% agarose gel overnight, and transferred onto an Amersham Hybond N^+ nylon membrane (GE Healthcare, Chicago, USA) under alkaline conditions using 0.5 N NaOH. The probe containing the 1.3 kb *CvFatB4* coding sequence was generated by PCR using above mentioned gene-specific primers, and radioactively labeled with [α-^{32}P] dCTP and Ready-to-go DNA labeling beads (GE Healthcare, Chicago, USA) at 37°C for 1 h. After hybridization with the radioactive probe, the membrane was washed at 65°C with 2X SSC solution, which contained 0.1% SDS, and 0.2X SSC, which contained 0.1% SDS.

Fatty acid analysis

The fatty acid composition of the Arabidopsis seeds was analyzed by gas chromatography of the fatty acid methyl esters (FAMEs) using 15:0 FAME as the internal standard, as previously described (Wirasnita et al. 2013). Briefly, approximately 100 mg of crushed seeds were saponified in 1 ml 0.5 N NaOH-methanol in a glass test tube for 60 min at 50°C. The fatty acids were converted to FAMEs by adding 1 ml of 12.5% (v/v) boron trifluoride-methanol (Sigma-Aldrich, St. Louis, USA). They were then incubated for 30 min at 50°C. The FAMEs were extracted with 1 ml hexane and then mixed with 2 ml saturated NaCl solution. A 1 μl sample was injected into a Hewlett-Packard 5890 gas chromatograph (Agilent Technologies, Santa Clara, USA) equipped with a flame ionization detector. The fatty acids were separated using an Rt-2560 capillary column (Restek, State College, USA) and nitrogen as the carrier gas. The column was temperature-programmed at 3°C min^{-1} to 250°C from an initial temperature of 100°C. The fatty acid profiles were calculated using Hewlett-Packard Chemstation software.

Results

Isolation and sequence analysis of acyl-ACP thioesterases from *C. viscosissima*

Primers, which were designed based on the conserved regions of seven *Cuphea* FatB sequences (Supplementary Fig. 1), were used to clone a putative Acyl-ACP thioesterase sequence in *C. viscosissima* seeds that have 16 mol% 8:0 and 71 mol% 10:0 MCFAs. The full-length cDNA sequence

Fig. 1 Alignment of the amino acid sequences in CvFatB4 and representative plant FatB sequences. These were AtFatB1 from the *Arabidopsis* acyl-ACP thioesterase (GenBank accession no. Z36910), RcFatB1 from the *Ricinus communis* acyl-ACP thioesterase (GenBank accession No. EU000562), JcFatB from *Jatropha curcas* (GenBank accession No. GQ226036), and UcFatB from *Umbellularia californica* (GenBank accession no. Q41635). The predicted cleavage site of the transit peptide in CvFatB4 is indicated by the arrowhead. Important catalytic residues in UcFatB that are conserved in various plant FatB sequences are indicated by filled circles above the amino acid residues. Conserved amino acid residues are shaded in gray

was deposited in the NCBI Genbank database (Accession No. GU225692), and called as *CvFatB4* in this paper to avoid any confusion with the previously reported *CvFatB1* (Jing et al. 2011). The predicted open reading frame of the *CvFatB4* cDNA encodes a 414-amino acid protein, which had a calculated molecular weight of 45.7 kDa and a theoretical isoelectric point of 8.2. The CvFatB4 sequence is almost identical (410 out of 414 amino acids) to a putative *C. lanceolate* acyl-ACP thioesterase (Genbank accession No. CAA54060) with unknown biochemical activity (Töpper and Martini 1994). The CvFatB4 contains a 66-amino acid putative N-terminal chloroplast transit peptide according to TargetP 1.1 (Fig. 1) (Emanuelsson et al. 2007). Sequence alignment of the deduced amino acid sequence of *CvFatB4* and several plant FatB proteins with known biochemical activities revealed that four amino acid residues (Asp_{312}, Asn_{314}, His315, and Glu_{350} in CvFatB4) are conserved and localized in so-called hot-dog folds, suggesting that they are catalytically important and CvFatB4 catalyzes the hydrolysis of the thioester bond using the catalytic mechanism found in UcFatB (Dani et al. 2011; Dörmann et al. 1995; Feng et al. 2017; Sánchez-García et al. 2010).

Phylogenetic analysis of CvFatB4

A phylogenetic analysis of the fourteen *Cuphea* acyl-ACP thioesterases with known substrate preferences and California bay UcFatB, along with CvFatB4, indicated that CvFatB4 is most closely related to CcFatB2, CvFatB2, and ChFatB1 among the *Cuphea* FatBs (Fig. 2). When CcFatB2 sharing 97.6% amino acid sequence identity with CvFatB4 was expressed in Arabidopsis, a 2.8-fold increase in 16:0 fatty acid production was observed (Filichkin et al. 2006). In addition, significant accumulations of 12:0 and 14:0 fatty acids, which do not accumulate in Arabidopsis Col-0 plants, were also detected when CcFatB2 was overexpressed in a *fab2* mutant lacking a desaturase for 18:1. When substrate preference was investigated for CvFatB2 using *E. coli* K27, a > 2-fold increase in 14:0 fatty acid accumulation was observed, along with a small, but significant, accumulation of 8:0 fatty acid (Jing et al. 2011). CvFatB2, showing 92.8% amino acid sequence identity to CvFatB4, was also cloned from *C. viscosissima* like CvFatB4, and one of three biochemically characterized FatBs in *C. viscosissima*. Biochemical characterization of ChFatB1 using *E. coli* and rapeseed

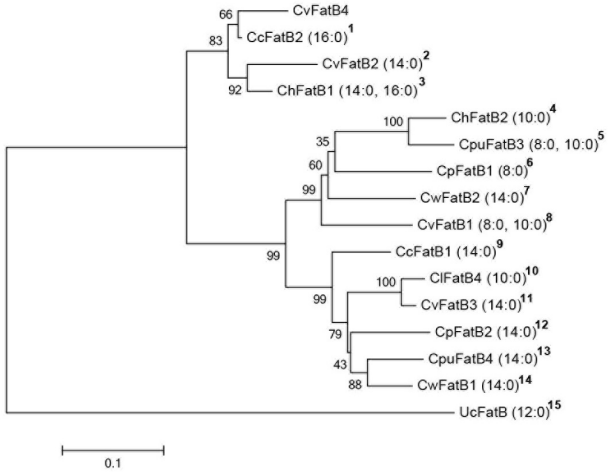

Fig. 2 Phylogenetic tree for CvFatB4, other *Cuphea* acyl-ACP thioesterases, and UcFatB. The percentage of the tree in which the associated taxa are clustered together in bootstrapping is shown next to the branches. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. The fatty acid most strongly accumulated via heterologous expression is indicated in the parentheses to the right of the protein name. Multiple fatty acids are listed in the parentheses when different fatty acids were dominant in separate experiments. Note that rather broad substrate use was observed in most cases and that fatty acids not indicated in the figure also showed higher accumulation levels compared to the control. NCBI accession numbers, organism(s) used in the heterologous enzyme expression analysis, and relevant references are as follows:[1] DQ238595 in *Arabidopsis* (Filichkin et al. 2006);[2] GU225693 in *E. coli* K27 (Jing et al. 2011);[3] U17076 in *E. coli* and rapeseed (Jones et al. 1995);[4] U39834 in *E. coli* (Dehesh et al. 1996[b]), *Arabidopsis* (Tjellström et al. 2013), and rapeseed (Dehesh et al. 1996[b]);[5] KC675178 in *Arabidopsis* (Tjellström et al. 2013) and *Camelina* (Kim et al. 2015);[6] U38188 in *E. coli* (Dehesh et al. 1996[a]);[7] U56104 in *Arabidopsis* (Leonard et al. 1997);[8] JF338906 in *E. coli* (Jing et al. 2011) and *Camelina* (Kim et al. 2015);[9] DQ238594 in *Arabidopsis* (Filichkin et al. 2006);[10] AJ131741 in rapeseed (Beermann et al. 2007);[11] JF338908 in *E. coli* (Jing et al. 2011);[12] U38189 in *E. coli* (Dehesh et al. 1996[a]) and *Arabidopsis* (Tjellström et al. 2013);[13] KC675179 in *Camelina* (Kim et al. 2015);[14] U56103 in *Arabidopsis* (Leonard et al. 1997);[15] and Q41635 in *Arabidopsis* (Voelker et al. 1992) and rapeseed (Knutzon et al. 1999). The scale bar shows the fractional differences in the amino acid sequences

demonstrated that ChFatB1 expression results in 14:0 and 16:0 fatty acid increase (Jones et al. 1995). Interestingly, a much greater increase in 14:0 fatty acid accumulation than 16:0 was observed when a fatty acid degradation mutant of *E. coli* was used. Studies using the other two previously reported *C. viscosissima* FatBs revealed that 8:0 and 10:0 fatty acids increase when CvFatB1 was expressed in both *E. coli* and *Camelina sativa* (Kim et al. 2015; Jing et al. 2011), while increase in 14:0 and 16:0 fatty acids was evident when CvFatB3 was expressed in *E. coli* (Jing et al. 2011).

Transgenic over-expression of CvFatB4 results in the increase of 16:0 fatty acid in Arabidopsis T1 plants

To determine the substrate specificity or preference of CvFatB4, we generated transgenic Arabidopsis plants that overexpressed *CvFatB4* cDNA under the control of the *CaMV* 35S promoter. Fatty acid profiles were determined using T_2 seeds (seeds harvested from T_1 plants) by gas chromatography (Table 1; Fig. S2). All *CvFatB4* transgenic lines had higher proportions (9.0 ~ 12.1 mol%) of 16:0 palmitate than the control (7.9 mol%), but had slightly smaller oleic acid (18:1). However, no significant differences were noted for medium-chain fatty acids between the wild-type and *CvFatB4* transgenic plants, although *CvFatB4 was isolated from C. viscosissima* seeds in which 8:0 and 10:0 fatty acids are predominant seed oil (Knapp and Tagliani 1991). The *CvFatB4* transgene copy number was determined by Southern blot analysis (Fig. 3). The results showed that transgenic Arabidopsis lines #6 and #9 had one copy of the *CvFatB4* transgene, along with Arabidopsis *FatB*1 which cross-hybridized with the *CvFatB4* probe (indicated by an arrow in Fig. 3A). Transgenic line #1 and #10 had two and three transgene copies, respectively, while more than four copies of the *CvFatB4* transgene were identified in transgenic lines #3 and #16.

Table 1 Total fatty acid composition in seeds from wild-type and *CvFatB4*-expressing transgenic *Arabidopsis* (T1) plants

Plants	Fatty acid (mol%)						
	16:0	18:0	18:1	18:2	18:3	20:0	20:1
Wild-type	7.9	3.1	15.1	30.3	18.5	2.1	20.1
Transgenic line #1	9.4	3.0	15.8	29.6	18.2	2.0	19.4
Transgenic line #3	10.7	3.2	14.0	29.6	18.3	2.1	19.7
Transgenic line #6	12.1	3.0	13.8	30.0	17.4	2.0	19.2
Transgenic line #9	11.3	3.1	13.2	29.4	18.9	2.2	19.3
Transgenic line #10	9.0	3.1	14.5	28.8	19.2	2.0	20.7
Transgenic line #16	12.1	3.3	13.5	28.3	18.5	1.9	19.9

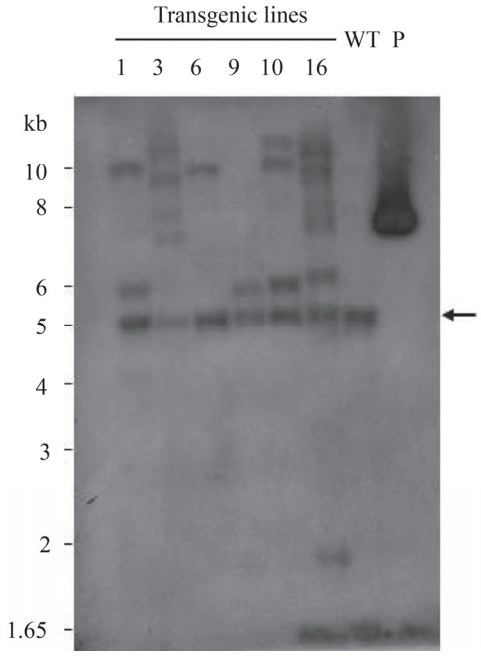

Fig. 3 Transgene copy numbers for each *CvFatB4* transgenic line and 16:0 fatty acid accumulation among the T_2 progenies produced by each line. Genomic Southern blot results, with the positions of the DNA molecular weight standards indicated on the left. The arrow indicates the endogenous *AtFatB* gene with a 70% homology to *CvFatB4*, which was used as a probe. WT, genomic DNA of wild-type plants; P, pUC19-*CvFatB4* expression cassette DNA

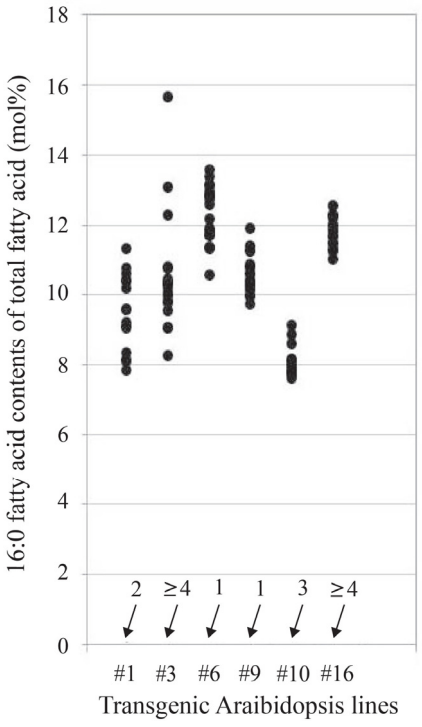

Fig. 4 16:0 fatty acid levels in seeds from individual T_2 plant seeds resulting from transgene segregation in each transgenic line. The numbers above the arrows show the *CvFatB4* transgene copy number in each transgenic line as determined by the genomic Southern blot analysis shown in Fig. 3. Seeds from at least 10 individual T_2 plants were used for the analyses. Wild-type plants had a 16:0 fatty acid molar fraction of about 8 mol% (Table 1)

Palmitate (16:0) contents in seeds from the T_2 progenies

The palmitate proportions in the seed oil were determined for T_2 progenies obtained from the self-pollination of six T_1 lines shown in Table 1. A range of 16:0 fatty acid contents in the T_2 progenies was observed, which is possibly due to transgene segregation in the T_2 generation (Fig. 4). The 16:0 fatty acid content in wild-type Arabidopsis is about 8% (Table 1). The highest 16:0 proportion (15.6 mol%) was detected in the T_2 population of transgenic line #3, which had the widest range (8.2 ~ 15.6 mol%). In genomic Southern analysis, this line was predicted to have more than four copies of the *CvFatB4* transgene in its genome. The T_2 progenies of another multi-copy line, transgenic line #16, also showed high levels of the 16:0 fatty acid, but had a much narrower concentration range (11 ~ 12.6 mol%) (Fig. 3 and 4). In addition, some T_2 progenies of other transgenic lines containing 1 ~ 2 transgene copies also showed significant increases in 16:0 fatty acid. In contrast, the T_2 progenies from line #10, which contained three copies of the transgene, seventeen progenies had wild-type palmitate levels and three progenies showed very low increases in 16:0 that ranged from 8.6 ~ 9.1 mol%.

Discussion

In this study, we biochemically characterized CvFatB4 expressed in seeds of *C. viscosissima* that contains high levels of MCFAs (8:0 or 10:0). Alignment of the deduced protein sequence and other previously characterized plant acyl-ACP thioesterase protein sequences suggested that the newly cloned *C. viscosissima* cDNA may encode FatB. The CvFatB4 showed over 95% amino acid sequence identity to CcFatB2, ChFatB1 and CvFatB2, which had been previously isolated from *Cuphea* species (Fig. 2). It has been reported that overexpression of *CcFatB2* in Arabidopsis increased 16:0 fatty acid accumulation, while overexpression of *ChFatB1* in rapeseed resulted in a dramatic increase in 16:0 fatty acid, along with 14:0 (Filichkin et al. 2006; Jones et al. 1995). Characterization of CvFatB2 using *E. coli*, which were also expressed in *C. viscosissima* seeds like CvFatB4, showed that CvFatB2 prefers 14:0 fatty acid as substrate (Jing et al. 2011). Because it was not clear whether CvFatB4 uses 14:0, 16:0 fatty acid or both as

substrate, based on sequence comparison alone, we over-expressed *CvFatB*4 in Arabidopsis seeds and found that 16:0 (palmitate) levels were elevated up to 12.1 mol% in T_1 plant seeds and up to 15.6 mol% in T_2 plant seeds, compared to 7.9 mol% in wild-type control (Table 1 and Fig. 4). Whereas, no significant change in 14:0 fatty acid accumulation was detected. Therefore, we conclude that CvFatB4 catalyzes the hydrolysis of 16:0-ACP.

We also investigated the relationship between transgene copy numbers and 16:0 fatty acid contents in the *CvFatB4* T_2 progenies (Fig. 3). There were complicated inheritance patterns for the transgenes. These patterns were likely to be related to the high copy numbers and unstable transgenic loci, which affect transgene expression (Tizaoui et al. 2012). For example, one T2 progeny in line #3 produced up to 15.6 mol% of 16:0, whereas another only produced the same amount of 16:0 as the wild-type plant (about 8 mol%). This indicated that transgene segregation was still happening in the T_2 generation of this line. In contrast, only a small variation in 16:0 accumulation was observed among the T_2 progenies of line #16, which also carried more than four copies of transgene. Individual T_2 progenies from line #10 carrying three copies of transgene showed little differences in 16:0 accumulation, compared to the wild-type (Table 1). It is possible that the multi-copy *CvFatB4* transgene in line #10 was silenced. Multiple copies of a transgene leading to high expression rates or silencing, as previously reported (Tang et al. 2007). It is worth noting that a transgenic line #6 harboring a single copy of the transgene can also produce high levels of 16:0 fatty acid, which is possibly related to the genomic location of the T-DNA integration (Fig. 4).

The CvFatB4 protein sequence revealed that four catalytically important residues for thioesterase activity are conserved (Fig. 1) (Feng et al. 2017). Previous structural and biochemical analyses have shown that four amino acids residues − Asp_{281}, Asn_{283}, His_{285}, and Glu_{319} in UcFatB play important roles in 12:0-ACP thioesterase, while Cys_{320}, which was previously assumed to have an important role in catalytic activity, has a relatively minor role (Feng et al. 2017; Yuan et al. 1996). In CvFatB4, Asp_{312}, Asn_{314}, His_{316}, and Glu_{350} correspond to these four catalytically important residues and show sequence conservation with UcFatB and other plant FatBs (Fig. 1 and Fig. S1). These results strongly suggest that *CvFatB4* catalyzes the hydrolysis of the thioester bond, which is similar to UcFatB and other plant FatBs (Feng et al. 2017). In FatBs, residues for substrate specificity are located at the N-terminal domain, whereas those for catalytic activities are located at the C-terminal domain (Feng et al. 2017; Mayer and Shanklin 2005). The characterization of the four FatBs with different substrate specificities or preferences in *C. viscosissima* will enable us to further understand the mechanism by which substrate preference for medium-chain fatty acyl-ACP is determined by different *C. viscosissima* FatBs (Jing et al. 2011; Kim et al. 2015).

In conclusion, we isolated and characterized *CvFatB4,* which encodes a novel acyl-ACP thioesterase from *C. viscosissima* seeds. Overexpression of *CvFatB4* under the control of the *CaMV* 35S promoter in Arabidopsis resulted in an increase in palmitate (16:0 fatty acid) at the cost of oleic acid (18:1 fatty acid) reduction. This result shows that *CvFatB4* has 16:0-ACP thioesterase activity.

Acknowledgements

We express our appreciation to the late Dr. Kyung Hee Roh, whose contribution to this work was of great significance. We also appreciate helpful discussion of Dr. Jeong-Won Nam (Department of Biological Sciences, Chungnam National University, Daejeon, Republic of Korea).

Funding

This study was funded by the Research Program for Agricultural Science & Technology Development (Project No. PJ01257102), the National Institute of Agricultural Science, Rural Development Administration, Republic of Korea, and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A3B04035080), Republic of Korea.

Compliance with Ethical Standards

Jinouk Yeon, Jong-Sug Park, Kyeong-Ryeol Lee, and Hankuil Yi declare that they have no conflict of interest.

References

Beermann C, Winterling N, Green A, Möbius M, Schmitt JJ, Boehm G (2007) Comparison of the structures of triacylglycerols from native and transgenic medium-chain fatty acid-enriched rape seed oil by liquid chromatography-atmospheric pressure chemical ionization ion-trap mass spectrometry (LC-APCI-

ITMS). Lipids 42:383-394. doi.org/10.1007/s11745-006-3009-1

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735-743. doi.org/10.1046/j.1365-313x.1998.00343.x

Dani KG, Hatti KS, Ravikumar P, Kush A (2011) Structural and functional analyses of a saturated acyl-ACP thioesterase, type B from immature seed tissue of *Jatropha curcas*. Plant Biol (Stuttg) 13:453-61.doi.org/ 10.1111/j.1438-8677.2010.00410.x

Davies HM (1993) Medium chain acyl-ACP hydrolysis activities of developing oilseeds. Phytochemistry 33: 1353-1356. doi.org/ 10.1016/0031-9422(93)85089-A

Davies HM, Anderson L, Fan C, Hawkins DJ (1991) Developmental induction, purification, and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Arch Biochem Biophys 290: 37-45. doi.org/10.1016/0003-9861(91)90588-a

Dehesh K., Edwards P, Hayes T, Cranmer AM, Fillatti J (1996[a]) Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol 110: 203-210. doi.org/10.1104/pp.110.1.203

Dehesh K, Jones A, Knutzon DS, Voelker TA (1996[b]) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by over-expression of ChFatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J 9: 167-172. doi.org/10.1046/j.1365-313X.1996.09020167.x

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. 316:612-618

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. Nature Protocols 2: 953-971. doi.org/10.1038/nprot.2007.131

Feng Y, Wang Y, Liu J, Liu Y, Cao X, Xue S (2017) Structural insight into acyl-ACP thioesterase toward substrate specificity design. ACS Chem Biol 12:2830-2836. doi.org/10.1021/acschembio.7b00641

Filichkin SA, Slabaugh MB, Knapp SJ (2006) New FATB thioesterases from a high-laurate Cuphea species: Functional and complementation analyses. Eur J Lipid Sci Technol 108:979-990. doi.org/10.1002/ejlt.200600158

Graham SA, Hirsinger F, Röbbelen G (1981) Fatty acids of Cuphea (Lythraceae) seed lipids and their systematic significance. Am J Bot 68:908-917. doi.org/10.1002/j.1537-2197.1981.tb07806.x

Harwood JL (1988) Fatty acid metabolism. Ann Rev Plant Biol 39: 101-138. doi.org/10.1146/annurev.pp.39.060188.000533

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. doi.org/10.1186/1471-2091-12-44

Jones A, Davies HM, Voelker TA (1995) Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7: 359-371. doi.org/10.1105/tpc.7.3.359

Kim HJ, Silva JE, Vu HS, Mockaitis K, Nam JW, Cahoon EB (2015) Toward production of jet fuel functionality in oilseeds: identification of FatB acyl-acyl carrier protein thioesterases and evaluation of combinatorial expression strategies in *Camelina* seeds. J Exp Bot 66:4251-4265. doi.org/10.1093/jxb/erv225

Knapp SJ, Tagliani LA (1991) Two medium-chain fatty acid mutants of *Cuphea viscosissima*: A source of medium-chain fatty acids. J Am Oil Chem Soc 68:515-517. doi.org/10.1111/j.1439-0523.1991.tb00520.x

Knutzon DS, Hayes TR, Wyrick A, Xiong H, Maelor Davies H, Voelker, TA (1999) Lysophosphatidic acid acyltransferase from coconut endosperm mediates the insertion of laurate at the sn-2 position of triacylgycerols in lauric rapeseed oil and can increase total laurate levels. Plant Physiol 120:739-746. doi.org/10.1104/pp.120.3.739

Leonard JL, Slabaugh MB, Knapp SJ (1997) *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Mol Biol 34:669-679.

Martinez-Trujillo M, Limones-Briones V, Cabrera-Ponce JL, Herrera-Estrella L (2004) Improving transformation efficiency of Arabidopsis thaliana by modifying the floral dip method. Plant Mol Biol Rep 22:63-70. doi.org/10.1007/BF02773350

Mayer KM, Shanklin J (2005) A structural model of the plant acyl-acyl carrier protein thioesterase FatB comprises two helix/4-stranded sheet domains, the N-terminal domain containing residues that affect specificity and the C-terminal domain containing catalytic residues. J Biol Chem 280: 3621-3627. doi.org/10.1074/jbc.M411351200

Nam JW, Yeon J, Jeong J, Cho E, Kim HB, Hur Y, Lee KR, Yi H (2019) Overexpression of acyl-ACP thioesterases, CpFatB4 and CpFatB5, induce distinct gene expression reprogramming in developing seeds of *Brassica napus.* Int J Mol Sci. 20(13), 3334. doi.org/10.3390/ijms20133334

Pollard MR, Anderson L, Fan C, Hawkins DJ, Davies HM (1991) A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of *Umbellularia californica*. Arch Biochem Biophys. 284: 306-312. doi.org/10.1016/0003-9861(91)90300-8

Salas JJ, Ohlrogge JB (2002) Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Arch Biochem Biophys 403: 25-34. doi.org/10.1016/S0003-9861(02)00017-6

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

Slabas AR, Fawcett T (1992) The biochemistry and molecular biology of plant lipid biosynthesis. Plant Mol Biol 19: 161-191. doi.org/10.1007/BF00015613

Somerville C, Browse J (1991) Plant lipids: Metabolism, mutants and membrane. Science 252: 80-87. doi.org/10.1126/science.252.5002.80

Stumpf PK (1987) The biosynthesis of saturated fatty acids. In: Stumpf PK, Conn EE. (ed) The Biochemistry of Plants, Academic press, New York, pp 121-136. Https://10.1016/B978-0-12-

675404-9.50013-8

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-2729. doi.org/10.1093/molbev/mst197

Tang W, Newton RJ, Weidner DA (2007) Genetic transformation and gene silencing mediated by multiple copies of a transgene in eastern white pine. J Exp Bot 58:545-554. doi.org/10.1093/jxb/erl228

Thompson GA, Roughan PG, Browse JA, Slack CR, Gardiner SE (1986) Spinach leaves desaturate exogenous [^{14}C] palmitate to hexadecatrienoate: evidence that *de novo* glycerolipid synthesis in chloroplasts can utilize free fatty acids imported from other cellular compartments. Plant Physiol 82: 357-362. doi.org/10.1104/pp.82.2.357

Tizaoui K, Kchouk ME (2012) Genetic approaches for studying transgene inheritance and genetic recombination in three successive generations of transformed tobacco. Genet Mol Biol 35:640-649. doi.org/10.1590/S1415-47572012000400015

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 587:936-942. doi.org/10.1016/j.febslet.2013.02.021

Töpfer R, Martini N, Schell J (1995) Modification of plant lipid synthesis. Science 268:681-686. doi.org/10.1126/science.268.5211.681

Töpper R, Martini N (1994) Molecular cloning of cDNAs or genes encoding proteins involved in de novo fatty acid biosynthesis in plants. J of Plant Physiol 143:416-425. doi.org/10.1016/S0176-1617(11)81801-8

Voelker TA, Hayes TR, Cranmer AM, Turner, JC, Davies HM (1996) Genetic engineering of a quantitative trait-metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. Plant J 9: 167-172. doi.org/10.1046/j.1365-313X.1996.09020229.x

Voelker TA, Jones A, Cranmer AM, Davies HM, Knuntzon DS (1997) Broad-range and binary-range acyl-acyl-carrier-protein thioesterases suggest an alternative mechanism for medium-chain production in seeds. Plant Physiol 114:669-677. doi.org/10.1104/pp.114.2.669

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE, Davies M (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257: 72-74. doi.org/10.1126/science.1621095

Wirasnita R, Hadibarata T, Novelina YM, Yusoff ARM, Yusop Z (2013) A modified methylation method to determine fatty acid content by gas chromatography. Bull Korean Chem Soc 34: 3239-3242. doi.org/10.5012/bkcs.2013.34.11.3239

Yi H, Richards EJ (2009) Gene duplication and hypermutation of the pathogen Resistance gene *SNC1* in the *Arabidopsis bal* variant. Genetics 183: 1227-1234. doi.org/10.1534/genetics.109.105569

Yuan L, Nelson BA, Cary G (1996) The catalytic cysteine and histidine in the plant acyl-acyl carrier protein thioesterases. J Biol Chem 271:3417-3419. doi.org/10.1074/jbc.271.7.3417