The effect of *AINTEGUMENTA-LIKE 7* over-expression on seed fatty acid biosynthesis, storage oil accumulation and the transcriptome in *Arabidopsis thaliana*

Stacy D. Singer[1,3*], Kethmi N. Jayawardhane[1], Chen Jiao[2], Randall J. Weselake[1], Guanqun Chen[1*]

[1] *Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada T6G 2P5*

[2] *Boyce Thompson Institute, Cornell University, Ithaca, New York, USA 14853*

[3] *Agriculture and Agri-Food Canada, Lethbridge Research and Development Centre, Lethbridge, Alberta, Canada T1J 4B1*

*Corresponding authors: Stacy D. Singer, stacy.singer@canada.ca; Guanqun Chen, gc24@ualberta.ca

ORCID numbers:

Stacy D. Singer: 0000-0002-6973-3881; Kethmi N. Jayawardhane: 0000-0003-4173-2497; Chen Jiao: 0000-0002-3327-0547; Randall J. Weselake: 0000-0001-8699-690X; Guanqun Chen: 0000-0001-5790-3903

Key Message: *AIL7* over-expression modulates fatty acid biosynthesis and triacylglycerol accumulation in Arabidopsis developing seeds through the transcriptional regulation of associated genes.

AUTHOR CONTRIBUTION STATEMENT

S.D.S. and G.C. were responsible for project conception; R.J.W. supervised experiments; S.D.S. and K.N.J. carried out all molecular, phenotypic and lipid-related studies; C.J. carried out bioinformatic analyses; S.D.S. wrote the article with contributions from all authors. S.D.S. and G.C. agree to serve as the authors responsible for contact and ensure communication.

ACKNOWLEDGEMENTS

The authors are grateful for the support provided by Agriculture and Agri-Food Canada (S.D.S.), the Canada Research Chairs Program (G.C. and R.J.W.) and Natural Sciences and Engineering Research Council of Canada Discovery Grants awarded to G.C. (RGPIN-2016-05926) and R.J.W. (RGPIN-2014-04585). The authors also acknowledge the support provided by Alberta Innovates (S.D.S, G.C. and R.J.W.), the Canada Foundation for Innovation and the Research Capacity Program of Alberta Enterprise and Advanced Education (G.C. and R.J.W.)

ABSTRACT

Seed fatty acids (FAs) and triacylglycerol (TAG) contribute to many functions in plants, and seed lipids have broad food, feed and industrial applications. As a result, an enormous amount of attention has been dedicated towards uncovering the regulatory cascade responsible for the fine-tuning of the lipid biosynthetic pathway in seeds, which is regulated in part through the action of LEAFY COTYLEDON1, ABSCISSIC ACID INSENSITIVE 3, FUSCA3 and LEC2 (LAFL) transcription factors. Although AINTEGUMENTA-LIKE 7 (AIL7*)* is involved in meristematic function and shoot phyllotaxy, its effect in the context of lipid biosynthesis has yet to be assessed. Here we generated *AIL7* seed-specific over-expression lines and found that they exhibited significant alterations in FA composition and decreased total lipid accumulation in seeds. Seeds and seedlings from transgenic lines also exhibited morphological deviations compared to wild-type. Correspondingly, RNA-Seq analysis demonstrated that the expression of many genes related to FA biosynthesis and TAG breakdown were significantly altered in developing siliques from transgenic lines compared to wild-type plants. The seed-specific over-expression of *AIL7* also altered the expression profiles of many genes related to starch metabolism, photosynthesis and stress response, suggesting further roles for *AIL7* in plants. These findings not only advance our understanding of the lipid biosynthetic pathway in seeds, but also provide evidence for additional functions of *AIL7*, which could prove valuable in downstream breeding and/or metabolic engineering endeavours.

KEYWORDS

AINTEGUMENTA-LIKE 7, lipid, fatty acid, gene overexpression, transcriptomics, plant development

Declarations

Funding: The authors are grateful for the support provided by Agriculture and Agri-Food Canada (S.D.S.), the Canada Research Chairs Program (G.C. and R.J.W.) and Natural Sciences and Engineering Research Council of Canada Discovery Grants awarded to G.C. (RGPIN-2016-05926) and R.J.W. (RGPIN-2014-04585). The authors also acknowledge the support provided by Alberta Innovates (S.D.D, G.C. and R.J.W.), the Canada Foundation for Innovation and the Research Capacity Program of Alberta Enterprise and Advanced Education (G.C. and R.J.W.).

Conflicts of interest/Competing interests: The authors declare no conflicts of interests

Availability of data and material The RNA-Seq data sets are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (accession no. PRJNA725102).

Code availability: not applicable

Authors' contributions: S.D.S. and G.C. were responsible for project conception; R.J.W. supervised experiments; S.D.S. and K.N.J. carried out all molecular, phenotypic and lipid-related studies; C.J. carried out bioinformatic analyses; S.D.S. wrote the article with contributions from all authors. S.D.S. and G.C. agree to serve as the authors responsible for contact and ensure communication.

Ethics approval: not applicable

Consent to participate: not applicable

Consent for publication: The authors agree to publication of the manuscript

INTRODUCTION

Triacylglycerol (TAG), which constitutes three fatty acids (FAs) esterified to a glycerol backbone, is a major storage compound in eukaryotic cells. In plants, TAG is mainly stored in seeds and fruits, and is utilized for a variety of functions including seedling growth during germination, pollen development, and sexual reproduction. Plant-derived oils, and particularly seed oils, also have enormous economic importance due to their popularity for food, feed, and industrial applications, as well as growing interest in their use as a renewable feedstock for biodiesel and non-petroleum-based biomaterial production (Lu, et al. 2011; Singer and Weselake 2018). While the majority of seed oils comprise predominantly palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1Δ^{9cis}; hereafter C18:1), linoleic acid (C18:2$\Delta^{9cis,12cis}$; hereafter C18:2), and α- linolenic acid (C18:3$\Delta^{9cis,12cis,15cis}$; hereafter C18:3), the precise proportions of these FAs can differ quite substantially depending on the species, providing each oil with unique functional properties that govern their potential for particular downstream applications (Jain 2020).

Due to an escalating demand for seed oils with various functionalities, a substantial amount of research has been directed towards improving seed lipid-related traits over the last few decades by manipulating crop genetics via breeding or biotechnology (e.g., (Katche, et al. 2019; Msanne, et al. 2020; Subedi, et al. 2020)). To aid in this endeavour, considerable efforts have been undertaken to better understand the lipid biosynthetic pathway at the molecular level (e.g., (Gao, et al. 2021; Singer, et al. 2016; Woodfield, et al. 2019)). However, due to the extremely multifaceted nature of lipid biosynthetic pathways in plants, and the correspondingly large number of genes involved, gaps remain in our understanding of this process.

106 During seed development, sucrose derived from photosynthesis serves as a carbon source
107 for the synthesis of different storage compounds, including TAG. In the case of TAG
108 biosynthesis, incoming sucrose is converted to glycolytic intermediates within the seed, where
109 they are utilized for the synthesis of acetyl-coenzyme A (CoA), which serves as a precursor for
110 the *de novo* biosynthesis of various FAs in plastids. The termination of *de novo* FA biosynthesis
111 can be catalyzed by either plastidial acyl-ACP acyltransferases, which use acyl-ACP directly for
112 the production of glycerolipids in the plastid, or acyl-ACP thioesterases, which catalyze the
113 hydrolysis of acyl-ACP and result in the release of free FAs (typically C16:0, C18:0 and C18:1)
114 and ACP. These free FAs are then transferred to the outer plastid envelope and are re-esterified
115 to CoA, after which time they enter the cytosolic pool of acyl-CoA, are transported to the
116 endoplasmic reticulum (ER), and can undergo further modifications (desaturation, elongation).
117 In higher plants, C18:1 can be further desaturated to C18:2 and C18:3 through the catalytic
118 action of the ER-bound fatty acid desaturase 2 (FAD2) and FAD3, respectively (Subedi, et al.
119 2020). Elongation of C16 and C18 FAs to chain lengths of 20 carbons or greater, on the other
120 hand, is catalyzed by the ER-bound fatty acid elongase (FAE) complex, consisting of four
121 enzymatic reactions including 3-ketoacyl-CoA-synthetase (FAE1; (Huai, et al. 2015)). In plants,
122 TAG can then be synthesized *de novo* through the sequential acyl-CoA-dependent acylation of a
123 glycerol backbone on the ER (Kennedy pathway; (Weiss, et al. 1960)), or through an acyl-CoA-
124 independent pathway that uses phosphatidylcholine as the acyl donor and *sn*-1, 2-diacylglycerol
125 as the acceptor (Dahlqvist, et al. 2000).
126 Several transcription factors (TFs) have been shown to play important roles in the
127 regulation of FA biosynthesis/modification and seed lipid accumulation. For example, LEAFY
128 COTYLEDON1 (LEC1), LEC2, ABSCISIC ACID INSENSITIVE3 (ABI3) and FUSCA3

(FUS3) are known to act as master regulators of lipid biosynthesis during embryogenesis and seed maturation, and various other TFs, such as WRINKLED1 (WRI1), MYB96 and bZIP67, have also been found to play a role (Kong, et al. 2020; Kong, et al. 2020; Kong, et al. 2019; Kumar, et al. 2020). However, the functions of many TF-encoding genes that are expressed during the maturation phase of seed development when TAG is known to accumulate have yet to be elucidated in the context of lipid-related pathways during seed development/maturation (Baud, et al. 2008; Weselake, et al. 2009). One such gene is *AINTEGUMENTA-LIKE 7* (*AIL7*; also known as *PLETHORA7* [*PLT7*]; AT5G65510), which is a member of the AIL/PLT subfamily of the large APETALA2/ ETHYLENE RESPONSE FACTOR (AP2/ERF) domain TF family (Krizek 2015). The AIL/PLT subfamily comprises 8 genes in Arabidopsis, including *AINTEGUMENTA* (*ANT*), *BABY BOOM* (*BBM*), *PLT1*, *PLT2*, *AIL6*/*PLT3*, *AIL5*/*PLT5, AIL7/PLT7* and *AIL1*, which encode proteins that share a relatively high level of amino acid identity within their AP2 domains, and in certain cases between their full-length protein sequences (e.g., 74% amino acid similarity between AIL6 and AIL7; Nole-Wilson, et al. 2005). Genes within the AIL/PLT subfamily are expressed in dividing tissues where they have overlapping functions in various shoot, floral and root developmental processes (Horstman, et al. 2017; Horstman, et al. 2014; Mudunkothge and Krizek 2012; Nole-Wilson, et al. 2005). While single *ail7* Arabidopsis mutants have not been found to exhibit any obvious morphological or developmental defects due to functional redundancy with other member of the AIL/PLT subfamily (Mudunkothge and Krizek 2012; Prasad, et al. 2011), triple *ant*/*ail6*/*ail7* mutants display severe aberrations in apical meristem activity (Mudunkothge and Krizek 2012), triple *ail5*/*ail6*/*ail7* mutants exhibit altered phyllotaxis (Prasad, et al. 2011) and lateral root outgrowth

(Du and Scheres 2017), and *ant*/*ail7* double mutants exhibit alterations in flower development (Krizek 2015).

Several other TFs belonging to the AP2/ERF family are known to modulate seed oil biosynthesis indirectly via a role in sugar metabolism. For example, *ap2* mutants have been found to produce large seeds with concomitant increases in seed oil content, which has been suggested to occur, at least in part, through the alteration of source-sink relations and soluble sucrose metabolism in developing seeds (Jofuku, et al. 2005; Ohto, et al. 2005). Direct roles in seed oil biosynthesis have also been observed for members of this gene family, including WRI1, which functions as a major transcriptional regulator of fatty acid biosynthesis in seeds (Baud, et al. 2007; Kong, et al. 2020; Kong, et al. 2020). While members of the *AIL* subfamily have not been implicated in FA/TAG biosynthesis or accumulation as of yet, several (including *AIL7*) are expressed during embryo development at a time consistent with TAG deposition (Le, et al. 2010) and both BBM and PLT2 have been found to regulate the expression of the LAFL network in Arabidopsis (Horstman, et al. 2017). Furthermore, although *AIL7* has not been found to be required for the establishment of the shoot apical meristem during embryogenesis (Mudunkothge and Krizek 2012), like BBM and AIL5 its constitutive over-expression in Arabidopsis leads to the production of somatic embryos (Horstman, et al. 2017), suggesting that it may play a role in embryonic development.

Given its putative and as of yet undeciphered role in the embryo, we sought to determine the function of *AIL7* in the context of seed development and storage TAG biosynthesis/accumulation. To achieve this, we developed seed-specific *AIL7* over-expression Arabidopsis lines, which can circumvent issues related to possible functional redundancy with other AIL/PLT subfamily members, and carried out subsequent morphological characterization,

as well as TAG content and composition analyses. We also compared the gene expression profiles of the over-expression and wild-type lines using RNA-Seq in order to provide further insight into the regulatory functions of *AIL7* during seed development. Our findings provide evidence for novel storage TAG-related functions for *AIL7*, including the regulation of acyl modification and seed TAG accumulation, and also hint at additional putative roles for this TF in processes such as starch metabolism, stress response and photosynthesis.

MATERIALS AND METHODS

RNA extraction and first-strand cDNA synthesis

Siliques were harvested from Arabidopsis plants (Col-0 background) at various [14 days after flowering (DAF)], flash frozen in liquid nitrogen, and stored at -80ºC until further use. Total RNA was extracted using the Sigma Spectrum Plant Total RNA kit (Sigma-Aldrich Canada Co., Oakville, ON) and any contaminating DNA was removed using the manufacturer's on-column DNase set (Sigma-Aldrich). First-strand cDNA synthesis was carried out using the Superscript III first-strand cDNA synthesis kit according to the manufacturer's instructions (Thermo Fisher Scientific, Inc., Waltham, MA) with 250 ng RNA as template along with an oligo(dT) primer.

Generation of the *AIL7* seed-specific over-expression construct

The *AIL7* seed-specific over-expression construct (AIL7-OE) was generated by first amplifying the full-length 1497-nt *AIL7* coding region (AT5G65510) from cDNA derived from Col-0 developing siliques (14 DAF) using Platinum High Fidelity polymerase (Thermo Fisher Scientific, Inc.) and primers AtAIL7F1AgeI (5' - CT**A CCG GT**A TGG CTC CTC CAA TGA CG - 3') and AtAIL7R1BamHI (5' - CCA **GGA TCC** TTA GTA AGA CTG GTT AGG - 3'),

197 which contained restriction sites near their 5' ends to facilitate cloning (shown in bold). Thermal
198 parameters for amplification were as follows: 94ºC for 2 min, 32 cycles of 94ºC for 15 s, 55ºC
199 for 30 s and 68ºC for 2 min, followed by a final extension of 68ºC for 7 min. The resulting
200 fragment was inserted between the seed-specific *Brassica napus napin* promoter (Josefsson, et
201 al. 1987) and *Pisum sativum Ribulose-1,5-bisphosphate carboxylase* transcriptional terminator
202 (*rbcS-t*) in a pGreen 0029 background ((Hellens, et al. 2000); Fig. 1A). Sequencing was carried
203 out at every step of plasmid construction to ensure the correct identity of the resulting plasmid.
204
205 **Arabidopsis transformation and plant growth conditions**
206 The AIL7-OE vector was introduced into *Agrobacterium tumefaciens* strain GV3101 via
207 electroporation along with the pSoup helper plasmid (Hellens, et al. 2000). The resulting
208 recombinant bacteria were utilized for the transformation of Arabidopsis Col-0 plants using the
209 floral dip method (Clough and Bent 1998), and first generation transformants were selected by
210 plating surface-sterilized seeds on half-strength Murashige and Skoog media (Murashige and
211 Skoog 1962) supplemented with 1% sucrose, 0.8% agar, 100 mg/L timentin, and 50 mg/L
212 kanamycin. Transgenic T_1 and T_2 seedlings were transferred to soil following approximately 10
213 days of growth on agar plates, while subsequent generations were sowed directly on soil. For
214 seedling growth assays on vertical plates, sterilized homozygous seeds were plated in a single
215 row on half-strength Murashige and Skoog media lacking sucrose and antibiotics.
216 In every case, Arabidopsis seeds were cold-treated at 4ºC for 3 days in the dark prior to
217 their placement in a growth chamber at 22ºC with a photoperiod of 18 h day/6 h night and 250
218 $\mu mol/m^2 \cdot s$ light intensity for the remainder of their life cycle. The presence of the AIL7-OE
219 construct was confirmed in transgenic plants by PCR and homozygous lines were identified

using segregation analyses. All experiments in which T_1 plants were utilized included empty-vector transformed lines as wild-type controls, while those involving subsequent generations utilized null-segregants as the wild-type control.

Quantitative real-time RT-PCR assessment of *AIL7* transcript levels

Quantitative real-time RT-PCR (qRT-PCR) assays were conducted in triplicate using 1 µl of a 1/50 dilution of each cDNA as template along with SYBR green PCR master mix (Thermo Fisher Scientific, Inc.) in a final volume of 10 µl. Assays were carried out on an ABI 7900HT Fast Real-Time PCR system (Thermo Fisher Scientific, Inc.) using the *AIL7*-specific primers shown in Table S1. Primers AtPP2AAF1 and AtPP2AAR1 (Singer, et al. 2016) were utilized to amplify a 146-nt fragment of the constitutively expressed *PROTEIN PHOSPHATASE 2A SUBUNIT 3* (*PP2AA3*) transcript (Czechowski, et al. 2005), which was used as an internal control. Thermal parameters for amplification were as follows: 95ºC for 2 min, followed by 40 cycles of 95ºC for 15s and 60ºC for 1 min. Dissociation curves were produced to confirm the presence of a single amplification product in each case. Levels of gene expression were determined using the standard curve method and SDS v2.4 software (Thermo Fisher Scientific, Inc.), with all *AIL7* expression data comprising mean values of biological replicates normalized to those of the *PP2AA3* internal control.

Seed lipid analysis

Arabidopsis seed lipid analysis was carried out as described in our previous study (Singer, et al. 2016). In brief, approximately 10 mg dry, mature Arabidopsis seeds (two technical replicates from each plant analyzed) were transmethylated in 3N methanolic HCl containing 0.1 mg

triheptadecanoin (C17:0 TAG in 100 µl chloroform) as an internal standard at 80ºC for 16 hours. The reaction was arrested through the addition of 0.9% (w/v) NaCl and the resulting FA methyl esters (FAMEs) were then extracted twice with hexanes. The FAME extracts were dried under a stream of nitrogen gas and resuspended in 1 ml iso-octane containing 0.1 mg methyl heneicosanoin (C21:0 methyl ester) as an additional internal standard. Extracted FAMEs were analyzed using an Agilent 6890 Network GC system equipped with a DB-23 capillary column (30 m x 0.25 mm x 0.25 um) and a 5975 inert XL Mass Selective Detector (Agilent Technologies Canada Inc., Mississauga, ON) with the following temperature program: 100ºC, hold for 4 min, 10ºC/min to 180ºC, hold for 5 min, and 10ºC/min to 230ºC, hold for 5 min.

Morphological analyses

Weights of T_3 seeds from homozygous AIL7-OE lines and wild-type plants were calculated by weighing small batches of seeds and counting using a FluorChem SP Imager and AlphaEase software (Alpha Innotech Corp., San Leandro, CA). Two to three technical replicates were carried out for 4-6 individual plants (biological replicates) of each line. Seed areas from the same lines were determined using the particle analysis function within ImageJ software (http://imagej.nih.gov/ij). Seedling root growth was assessed by measuring the root lengths of seedlings grown vertically on agar plates at various time points following germination. Experiments were repeated twice.

Scanning electron microscopy

Dry mature T_3 seeds from homozygous AIL7-OE-1 and AIL7-OE-19 lines, along with wild-type plants, were sputter coated with gold/palladium for 1.5 min using a Hummer 6.2 Sputter Coater

(Anatech Ltd., Battle Creek, MI). Microscopy was carried out using an FEI scanning electron microscope (model XL30; FEI Company, Hillsboro, OR) operating at 20 kV.

Illumina RNA-Seq

Total RNA from T_3 siliques (14 DAF) of three biological replicates of wild-type and three biological replicates of homozygous AIL7-OE lines (two replicates of AIL7-OE-1 and one replicate of AIL7-OE-19) was processed for RNA-Seq using the Illumina TruSeq Stranded mRNA LT Sample Prep Kit according to the manufacturer's directions (Illumina, San Diego, CA). The quality of the resulting strand-specific libraries were assessed using a BioAnalyzer (Agilent 2100 Bioanalyzer Model G2938C; Agilent Technologies Canada Inc.) and quantities were determined using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Inc.). Libraries were sequenced on an Illumina NextSeq 500 system using the single-end mode. RNA-Seq reads (75-bp in length) were initially aligned to the Arabidopsis genome using Tophat (Kim, et al. 2013), and following alignment, the count of mapped reads from each sample was derived and normalized to reads per kb of exon model per million mapped reads (RPKM). DEGs between wild-type and AIL7-OE lines were identified using the DESeq 1.8.3 package (Nadler, et al. 1997) with the raw count data. Raw P-values were adjusted for multiple testing using a false discovery rate (FDR) (Benjamini and Hochberg 1995), and genes with an FDR of less than 0.05 and fold-changes greater than 2 or less than 0.5 were regarded as DEGs. Two independent AIL7-OE lines were utilized in the analysis as a means of preventing the identification of alterations that may have arisen due to transgene insertion position rather than the transgene itself. GO functional classification of DEGs was achieved using the Plant MetGenMAP program (Joung, et al. 2009).

289 Verification of RNA-Seq results was conducted by carrying out qRT-PCR assays of
290 seven transcripts known to be involved in various lipid metabolic pathways (for primer
291 sequences see Table S1), as well as *AIL7*. Assays were carried out using total RNA derived from
292 14 DAF siliques from homozygous AIL7-OE-1 and AIL7-OE-19 lines (4-5 biological replicates
293 total), as well as wild-type controls (3 biological replicates). Reactions were carried out as
294 described above. Three technical replicates were used in each case. The correlation coefficient
295 between the $\log_2$ fold-changes in expression derived from qRT-PCR and RNA-Seq was
296 calculated using excel.
297
298 ***In silico* analyses of predicted TF binding sites**
299 The Plant Cistrome Database (http://neomorph.salk.edu/PlantCistromeDB) was utilized for the
300 identification of putative TF target genes based on their predicted DNA binding motifs
301 (O'Malley, et al. 2016).
302
303
304 **RESULTS**
305 **Confirmation of *AIL7* over-expression in AIL7-OE lines**
306 Arabidopsis seed-specific *AIL7* over-expression lines were generated through the introduction of
307 a cassette in which the seed-specific *napin* promoter was used to drive the expression of the
308 Arabidopsis *AIL7* coding sequence (Fig. 1A). This promoter is known to elicit particularly high
309 levels of expression of associated transgenes during the heart to torpedo stage of embryo
310 development, and lacks activity in vegetative tissues such as leaves and roots (Ellerström, et al.
311 1996; Stålberg, et al. 1993). To confirm the up-regulation of *AIL7* expression in our transgenic

312 lines, we analyzed developing T_2 siliques containing T_3 seeds (14 DAF) from the two
313 independent over-expression lines utilized throughout this study (AIL7-OE-1 and AIL7-OE-19),
314 as well as wild-type, using qRT-PCR. AIL7-OE-1 lines exhibited a 33-fold increase in
315 expression compared to wild-type, while AIL7-OE-19 lines exhibited a 19-fold increase in
316 expression compared to wild-type (Fig. 1B). These results confirm the over-expression of the
317 *AIL7* transcript in these lines.
318 To assess *AIL7* expression across the mid-stages of seed development when seed oil
319 begins to accumulate (Baud, et al. 2008), we also tested developing siliques at various DAF
320 during the mid-stage of seed development (Figs. S1A and S1B). Total RNA was extracted from
321 homozygous T_3 AIL7-OE-1 and wild-type siliques 10, 11, 12, 13, 14 and 15 DAF for qRT-PCR
322 assays. In wild-type plants, we observed a gradual increase in expression of *AIL7* from 11 DAF
323 until 14 DAF, at which point expression leveled off (Fig. S1A). In AIL7-OE-1 siliques,
324 expression of *AIL7* also increased between 11 and 14 DAF, after which point expression began
325 to drop (Fig. S1B). At every time point analyzed (10-15 DAF), expression of *AIL7* was
326 substantially enhanced in AIL7-OE-1 lines compared to wild-type. This enhancement due to the
327 *AIL7* overexpression remained relatively constant between 11 and 14 DAF; where we observed
328 increases of 25- and 31-fold compared to wild-type; which is reflective of the value seen in
329 AIL7-OE-1 T_2 siliques at 14 DAF. However, the difference between AIL7-OE-1 and wild-type
330 lines began to decrease at 15 DAF (11-fold increase in expression compared to wild-type).
331 Therefore, due to the peak in *AIL7* expression in wild-type lines and the very high fold-change
332 seen in the expression of this gene between AIL7-OE and wild-type lines at 14 DAF, we chose
333 this time point for subsequent RNA-Seq analysis.
334

AIL7-OE lines exhibit alterations in fatty acid composition and seed oil content

Compositional analysis of seed oil from AIL7-OE and wild-type lines demonstrated very substantial alterations in the proportions of FAs, the pattern of which was identical over both generations tested and multiple independent transgenic lines (Figs. 2A and S2A). With the exception of C16:1, the proportions of all FAs within the seed oil were affected consistently to a significant degree. In T_3 seeds derived from two independent homozygous AIL7-OE lines, average relative increases of 38.48% ± 1.23 SE (C16:0), 33.02% ± 0.35 (C18:1) and 39.93% ± 6.27 (C18:2) were observed compared to wild-type, whereas average relative decreases of 8.76% ± 0.35 (C18:0), 52.55% ± 1.72 (C18:3), 39.41% ± 1.28 (C20:0), 37.06% ± 2.59 (C20:1), 47.00% ± 1.23 (C20:2), 79.42% ± 1.06 (C20:3), 21.29% ± 0.23 (C22:0) and 52.39% ± 2.82 (C22:1) were observed compared to wild-type (Fig. 2A).

Seed oil content was also affected in AIL7-OE lines, with significant reductions apparent in the seeds of both generations tested (Figs. 2B and S2B). Indeed, in homozygous T_3 seeds, relative decreases of 11.52% (AIL7-OE-1) and 11.80% (AIL7-OE-19) were observed compared to wild-type (Fig. 2B).

Seed-specific over-expression of *AIL7* results in changes to seed morphology

Both light and scanning electron microscopy revealed that transgenic AIL7-OE seeds exhibited morphological changes compared to wild-type (Fig. 3). Light microscopy demonstrated that seeds from the transgenic lines appeared twisted or bent in shape, and were not elliptical as were wild-type seeds (Fig. 3A). SEM micrographs confirmed these findings, displaying transgenic seeds that bore ridges that twisted around the seed (Fig. 3B). While significant alterations in seed weights of AIL7-OE lines compared to wild-type plants were not evident (Fig. 3C), small

increases in seed area were observed in AIL7-OE plants, although this difference was only significant in the AIL7-OE-1, but not the AIL7-OE-19, homozygous line (Fig. 3D).

Seed-specific over-expression of *AIL7* leads to a reduction in seedling growth

In order to establish whether there were any carry-over effects in the next generation of AIL7-OE lines, we further assessed the morphology of transgenic and wild-type seedlings. As was the case for wild-type seedlings, homozygous T_3 AIL7-OE seedlings grown vertically on solid medium germinated within 24 h; however, once germinated, transgenic seedlings exhibited a significant decrease in their rates of growth compared to wild-type (Fig. 4). At 9 days post-germination, seedling root lengths were reduced on average 60.1% (AIL7-OE-1) and 47.1% (AIL7-OE-19) compared to wild-type (Fig. 4B). AIL7-OE seedlings grown on soil also appeared to exhibit reduced size of rosettes (Fig. 4C) and shorter root lengths but bolted and reached maturity at similar time points as wild-type plants (data not shown). In addition, mean overall seed yield of the two homozygous AIL7-OE T_3 lines combined (212.43 mg/plant ± 6.75 SE; n=64) was also slightly lower than that of wild-type (234.36 mg/plant ± 8.20 SE; n=34); a difference that was significant at $P \leq 0.05$.

Differential expression in AIL7-OE vs. wild-type silique transcriptomes

To determine how the over-expression of *AIL7* affects the transcriptome in developing Arabidopsis seeds, we carried out RNA-Seq analysis using RNA derived from developing siliques (14 DAF) of homozygous AIL7-OE lines and wild-type plants (Table S2). We identified a total of 27,202 expressed sequences in our transgenic and wild-type lines with 1,682 of these genes displaying significantly increased transcript abundance up-regulated Differentially

Expressed Genes (DEGs) and 378 exhibiting significantly decreased transcript abundance (down-regulated DEGs) in transgenic lines compared to wild-type (Supplemental Data S1).

Gene ontology classification revealed differential expression of genes related to many biological processes, molecular functions and cellular compartments

Gene ontology (GO) classification of the resulting DEGs indicated that they fell within several biological process categories (Fig. S3). For instance, many DEGs were involved in photosynthesis (GO:0015979; 28 genes up-regulated), response to stress (GO:0006950; 137 genes up-regulated and 27 genes down-regulated), embryonic development (GO:0009790; 19 genes up-regulated and 5 genes down-regulated) and flower development (GO:0009908; 17 genes up-regulated and 8 genes down-regulated), for example. Furthermore, 45 genes were up-regulated and 20 genes were down-regulated within the lipid metabolic process (GO:0006629; Table S3).

DEGs were also classified based on their molecular function (Fig. S4), as well as their cellular component (Fig. S5). Within the molecular function category, groups with a high abundance of DEGs included those relating to protein binding (GO:0005515; 191 up-regulated and 22 down-regulated), hydrolase activity (GO:0016786; 180 up-regulated and 40 down-regulated), kinase activity (GO:0016301; 129 up-regulated and 10 down-regulated), catalytic activity (GO:0003824; 143 up-regulated and 54 down-regulated), nucleotide-binding (GO:0000166; 129 up-regulated and 11 down-regulated) and DNA bindng (GO0003677; 104 up-regulated and 22 down-regulated). Other interesting groups included TF activity (GO:0030528; 91 up-regulated and 45 down-regulated) and lipid binding (GO:0008289; 13 up-regulated and 3 down-regulated). Within the cellular component category, a number of DEGs were classified

within the nucleus (GO:0005634; 131 up-regulated and 28 down-regulated), plastid (GO:0009536; 196 up-regulated and 19 down-regulated) and thylakoid (GO:0009579; 55 up-regulated and 1 down-regulated). Thirteen and four genes encoding proteins localized to the ER were also differentially up-regulated or down-regulated, respectively, between AIL7-OE and wild-type developing siliques (GO:0005783).

AIL7 over-expression affects the expression of genes involved in fatty acid biosynthesis and triacylglycerol accumulation

At least 17 DEGs with possible involvement in seed FA or TAG biosynthesis were identified in developing siliques from our AIL7-OE lines compared to wild-type plants (Table 1; Fig. 5). Up-regulated DEGs included an *ACYL CARRIER PROTEIN 5* (*ACP5*; AT5G27200), *LONG CHAIN ACYL-COA SYNTHETASE 9* (*LACS9*; AT1G77590), 3-*KETOACYL-ACYL CARRIER PROTEIN SYNTHASE I* (*KASI*; AT5G46290), and *LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE 5* (*LPAT5*; AT3G18850) for example. Down-regulated DEGs included *3-KETOACYL-COA SYNTHASE 5* (AT1G25450), *KCS7* (AT1G71160), *FATTY ACID ELONGATION 1* (*FAE1*; AT4G34520) and *B-KETOACYL-COA REDUCTASE 2* (*KCR2*; AT1G24470). In addition, although it did not meet our chosen threshold ratio of 0.5, *FATTY ACID DESATURASE 3* (*FAD3;* AT2G29980), which is essential for the desaturation of C18:2 to C18:3, was also significantly down-regulated by a ratio of 0.52 in AIL7-OE lines compared to wild-type. Similarly, *PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE 2* (*PDAT2*; AT3G44830) and *OLEOSIN 1* (*OLEO1*; AT4G25140) were also significantly down-regulated with a ratio of 0.59 and 0.76 in AIL7-OE lines. Conversely, *VIVIPAROUS1/ABI3-LIKE1* (*VAL1*; AT2G30470) was significantly up-regulated in AIL7-OE lines with a ratio of 1.67. A number of

genes with possible functions related to FA or TAG breakdown were also differentially expressed in AIL7-OE lines compared to wild-type, including *SUGAR-DEPENDENT1-LIKE* (*SDP1L*; AT3G57140; Table 2), which plays a known role in catalyzing TAG hydrolysis in Arabidopsis seeds (Kelly et al. 2011). The expression levels of many other genes encoding TFs, enzymes and biologically active proteins known to function in FA and TAG biosynthesis in developing seeds were not significantly affected by the over-expression of *AIL7* (Table S4).

AIL7 *over-expression leads to the up-regulation of genes involved in starch metabolism*

Fourteen genes involved in starch biosynthesis and degradation were also differentially expressed in developing siliques from AIL7-OE lines compared to wild-type, with the genes being significantly up-regulated in every case (Table S5). These included genes involved in starch biosynthesis such as *ADP-GLUCOSE PYROPHOSPHORYLASE SMALL SUBUNIT 2* (*APS2*; AT1G05610) and *ADP-GLUCOSE PYROPHOSPHORYLASE LARGE SUBUNIT 2* and *4* (*APL2* and *APL4*; AT1G27680 and AT2G21590). While up-regulation of *STARCH SYNTHASE 3* (*SS3*; AT1G11720) did not meet our threshold ratio of 2, its 1.8-fold up-regulation in AIL7-OE lines compared to wild-type was significant. Up-regulation was also observed in genes involved in starch branching, including *STARCH BRANCHING ENZYME 2.1* and *2.2* (*SBE2.1* and *SBE2.2*; AT2G36390 and AT5G03650), and de-branching (*ISOAMYLASE 1* [*ISA1*]; AT2G39930). Finally, up-regulation was also apparent in various genes involved in starch degradation, including *α-GLUCAN PHOSPHORYLASE 2* (*PHS2*; AT3G46970), *β-AMYLASE 2* (*BAM2*; AT4G00490), *DISPROPORTIONATING ENZYME 1* and *2* (*DPE1* and *DPE2*; AT5G64860 and AT2G40840), *STARCH EXCESS 1* and *4* (*SEX1* and *SEX4*; AT1G10760 and AT3G52180), and *LIKE SEX4 2* (*LSF2*; AT3G10940).

Effect of* AIL7 *over-expression on the expression of genes involved in processes other than storage compound accumulation

In addition to its effect on the expression of genes involved in TAG and starch accumulation, the over-expression of *AIL7* also significantly affected the expression of five of the seven other *AIL/PLT* subfamily genes, including *BBM* and *PLT2*, which were up-regulated, and *AIL1*, *AIL5* and *AIL6*, which were down-regulated (Table S6). A number of other genes falling within meristem-related GO biological process categories were also differentially expressed in developing siliques from AIL7-OE lines (Table S7), with the majority being up-regulated. However, genes encoding major meristem regulators such as *CLAVATA 1*, *2* and *3* (*CLV1*, *2* and *3*; AT1G75820, AT1G65380, AT2G27250), *WUSCHEL* (*WUS*; AT2G17950) *SHOOTMERISTEMLESS* (*STM*; AT1G62360) and *CUP SHAPED COTYLEDON 1* and *2* (*CUC1* and *2*; AT3G15170 and AT5G53950) were not altered at the transcriptional level in AIL7-OE lines compared to wild-type.

Interestingly, a large proportion of genes with functions within the photosynthetic process, and especially the light reactions, were also up-regulated in developing siliques from AIL7-OE lines compared to wild-type, while no genes with roles within this pathway were down-regulated (Table S8). In terms of genes functioning within the response to stress GO cellular process category, 164 genes were differentially expressed (Table S9), including 27 that were significantly down-regulated and 137 that were up-regulated. Genes within anatomical structure morphogenesis (Table S10) and embryonic development (Table S11) GO biological process categories were also differentially regulated in AIL7-OE lines compared to wild-type.

Verification of RNA-Seq data

To validate our RNA-Seq expression profile data, we performed quantitative real-time RT-PCR on seven genes encoding enzymes involved in lipid metabolic pathways, including *AIL7*, *FAE1*, *FAD3*, *ABCG12, ACP5*, *GPAT1*, *LACS9* and *KASI*, which were found to exhibit significant alterations in their expression levels in developing siliques from transgenic AIL7-OE lines compared to wild-type in the RNA-Seq experiment. In every case, $\log_2$ fold-changes in expression between transgenic lines and wild-type plants corresponded well with our RNA-Seq results (Fig. 6), yielding a correlation coefficient of 0.981173.

Predicted direct targets of AIL7 and other TFs in the seed lipid regulatory network

Comparison of predicted AIL7 target genes derived from the Plant Cistrome Database, which were established using DNA affinity purification sequencing (DAP-Seq; O'Malley et al. 2016), with DEGs revealed through our RNA-Seq analysis led to the identification of 51 genes that are possible direct targets of AIL7 in the seeds of our over-expression lines (Table S12). All but three of these DEGs were up-regulated in AIL7-OE lines compared to wild-type plants. Interestingly, *AIL7* was also listed as a putative target gene for the AGL15 (AT5G13790) and ABI5 (AT2G36270) TFs in the Plant Cistrome Database.

DISCUSSION

The biosynthesis and accumulation of FAs and TAG in plants are extremely complex processes involving an incredibly large number of genes, which are regulated at the epigenetic, transcriptional and post-transcriptional levels (e.g., (Kong, et al. 2019; Kumar, et al. 2020)). Although enormous efforts have been undertaken to unravel these regulatory pathways in full,

several gaps remain in our overall understanding of them. The LAFL TFs have been identified as master regulators of seed maturation and seed oil accumulation, affecting the expression of multiple other genes encoding TFs, enzymes and biologically active proteins involved in lipid biosynthesis (Fatihi, et al. 2016). However, this list of TFs is by no means exhaustive, and other, as of yet unidentified TFs likely also contribute to this process. As such, we sought to examine the function of Arabidopsis *AIL7*, which is expressed at relatively high levels during the maturation phase of seed development when storage oil begins to accumulate (Baud, et al. 2008) and encodes a TF that has yet to be characterized in this context (Le, et al. 2010), by investigating the effects of its over-expression in Arabidopsis seeds and assessing global transcriptomic changes in these lines compared to wild-type.

Quantitative real-time RT-PCR of wild-type plants indicated that *AIL7* expression increased during the mid-stages of Arabidopsis seed development (at which time storage oil reserves are accumulating) until 14 DAF, after which time expression began to level off (Fig. S1A), which corresponds well with a previous expression analysis of this gene in developing Arabidopsis seeds (Le, et al. 2010). Over-expression of *AIL7* in Arabidopsis seeds also resulted in high levels of up-regulation during the maturation phase of seed development, peaking in siliques harvested 14 DAF (Figs. 1B and S1B), which derives from the spatiotemporal activity of the *napin* promoter used to drive *AIL7* expression.

Intriguingly, in AIL7-OE lines, both T_2 and T_3 seeds exhibited consistent and significant alterations in the FA profile of their seed oil (Figs. 2A and S2A). Specifically, the proportions of C16:1, C18:1 and C18:2, increased significantly in transgenic lines. Conversely, the proportion of C18:0 and C18:3, as well as all very-long chain FAs (VLCFAs) with carbon chain lengths above 20, decreased significantly in AIL7-OE lines compared to wild-type. These alterations in

519 FA composition correlated well with changes in the expression of particular genes that were
520 observed in our RNA-Seq analysis (Supplemental Data S1; Fig. 5). For instance, an increased
521 amount of C16:0 in seed oil could have resulted, at least in part, from the up-regulation of *KASI*
522 (Table 1, Fig. 6), which utilizes C4:0- to C14:0-ACPs as substrate in a condensation reaction
523 with malonyl-ACP, resulting in the production of C6:0- to C16:0-ACPs (Wu and Xue 2010).
524 Similarly, increased levels of C18:2 and decreased proportions of C18:3 could be explained by
525 down-regulation of *FAD3*, which functions to catalyze the desturation of C18:2 to produce C18:3
526 (Table 1; Fig. 6; (Lemieux, et al. 1990)). Furthermore, the reduction in the proportion of FAs
527 with chain lengths longer than C18 likely stemmed from a significant decrease in the expression
528 of genes encoding components of the FAE complex, including the β-ketoacyl-CoA synthases
529 *FAE1*, which encodes a condensing enzyme that is responsible for the production of C20 and
530 C22 FAs in seeds (James, et al. 1995), and *KCS7*, which has yet to be functionally characterized
531 in seeds (Table 1). Given that *fae1* mutants display increased accumulation of C16 and C18 FAs
532 concomitant to reductions in VLCFAs, reduction in the expression of this gene could also
533 contribute to the overall increase in C16:0, C18:1 and C18:2 observed in AIL7-OE lines (James,
534 et al. 1995). Similarly, a significant increase in the expression of *LACS9*, which encodes a
535 plastidial long-chain acyl-CoA synthetase that is involved in the activation of free FA to acyl-
536 CoA and displays substrate preferences for C16:0, C18:0, C18:1 and C18:2 (Shockey, et al.
537 2002), could also have been a factor in the alterations in fatty acid composition observed in the
538 AIL7-OE lines (Table 1; Fig. 6).

539 The increases in the proportions of C18:1 and C18:2 combined with the decrease in the
540 proportion of C18:3 for the two AIL7-OE lines suggests that the resulting oils would be more
541 stable to oxidation than oil from wild-type plants. Previous studies have clearly demonstrated

542 that vegetable oils containing lower proportions of C18:3 are more stable under frying conditions
543 (Eskin, et al. 1989). Indeed, reducing 18:3 content to increase oxidative stability is a major goal
544 of many Brassica oilseed breeding programs (e.g., (Rahman, et al. 2013; Singer, et al. 2014).
545 Given that the penalty on seed oil content is relatively low (Fig. 2B), *AIL7* over-expression might
546 be a useful molecular tool for decreasing the C18:3 content of seed oil from Brassica oilseed
547 species to generate an oil with increased oxidative stability.
548 In addition to the clear shift in FA composition in AIL7-OE seed oil, we also observed a
549 slight but significant reduction in total seed oil content compared to wild-type plants in both the
550 T_2 and T_3 generations (Figs. 2B and S2B). While such an effect may have been elicited in a
551 variety of distinct manners, the differential expression of genes with functions in TAG
552 hydrolysis, the generation of oil bodies and starch biosynthesis were particularly conspicuous as
553 possible contributors. In the context of oilseeds, the breakdown of storage TAG occurs primarily
554 during seedling establishment to provide a source of energy prior to the establishment of
555 photosynthetic competence. Oil breakdown is initiated through the enzymatic action of lipases
556 that catalyze the hydrolysis of TAG to release glycerol and free FAs, which are then transferred
557 to the glyoxysome where they are activated to acyl-CoAs and subsequently catabolized through
558 β-oxidation (Eastmond 2006). Our RNA-Seq results demonstrated that genes encoding several
559 enzymes with possible functions in TAG breakdown were differentially expressed in developing
560 siliques from AIL7-OE lines compared to wild-type (Table 2). Perhaps the most noteworthy is
561 the significant up-regulation of *SDP1L*, which encodes a TAG lipase that has been shown
562 previously to play a major role in oil breakdown following seed germination in Arabidopsis
563 (Kelly, et al. 2011) and could plausibly have a negative impact on seed oil accumulation. While a
564 small number of the other genes listed have been reported to be involved in fatty acid

565 degradation and/or lipid metabolism (Chen, et al. 2012), the biological functions of GDSL
566 proteins in particular remain poorly understood. As such, future studies involving the functional
567 characterization of such genes in lipid catabolism would provide additional knowledge in this
568 area.

569 In seeds, TAG is stored in specialized oil bodies that are associated with various proteins,
570 with oleosins making up the majority of these (Huang 2018). It has been shown previously that
571 down-regulation/mutation of the major oleosin, *OLEO1*, in Arabidopsis led to a substantial
572 reduction in seed TAG, along with a simultaneous increase in storage protein content (Siloto, et
573 al. 2006). Therefore, it is feasible that the significant down-regulation of *OLEO1* in AIL7-OE
574 lines compared to wild-type may have also played a role in the observed decrease in seed oil
575 content (Table 1).

576 A tradeoff in the accumulation of distinct seed storage compounds (e.g., oil, protein and
577 starch), whereby an increase in the accumulation of one type of compound is associated with a
578 decrease in another, is a fairly common occurrence in plants due to the fact that carbon supplies
579 tend to be limited during seed maturation and the distinct biosynthetic pathways are in
580 competition for resources (e.g., (Lin, et al. 2006; Wang, et al. 2019; Zhang, et al. 2010). While
581 starch is an important seed storage product in many cereal grains, it only accumulates transiently
582 during oilseed development, with oil and storage proteins predominantly stored in mature seeds
583 (Mansfield and Briarty 1992). Interestingly, 14 genes involved in starch metabolism were
584 differentially expressed in developing siliques from AIL7-OE lines, with every one of them
585 being significantly up-regulated (Table S5). This included three genes encoding various subunits
586 of ADP-glucose pyrophosphorylase, which catalyzes the first and rate-limiting step of starch
587 biosynthesis (Cross, et al. 2004). These findings hint at the possibility that reduced seed oil

588 content in AIL7-OE lines could be linked to an increase in transitory starch. Such a phenomenon
589 has been observed in Arabidopsis *wri1* mutants, where a maximal augmentation in seed starch
590 content was noted 9 days after flowering, with levels dissipating to those seen in wild-type plants
591 by maturity (Focks and Benning 1998). Although starch content was not assessed in AIL7-OE
592 transgenic lines, the fact that we also detected the up-regulation of several genes involved in
593 starch degradation (Table S5) suggests that a similar decline in starch enhancement could be
594 evident. However, further characterization will be required to gain insight into the precise
595 interactions between starch and lipid metabolic pathways in AIL7-OE transgenic plants in the
596 future.

597 It is not known whether the effects of *AIL7* over-expression on FA biosynthesis and lipid
598 accumulation in seeds result from direct or indirect transcriptional regulation of downstream
599 genes involved in this process. However, a search of putative AIL7 gene targets in the Plant
600 Cistrome Database indicated that although it may directly regulate a number of genes encoding
601 enzymes that were differentially expressed in developing siliques from AIL7-OE lines, the only
602 potential direct target identified in the context of seed storage compounds was the starch
603 biosynthetic gene *SBE2.1* (Table S12). While the list of target genes in this database was by no
604 means exhaustive, it may indicate that AIL7 elicits the transcriptional regulation of FA
605 biosynthetic and TAG accumulation-related genes by acting upstream of other TFs in the lipid
606 biosynthetic network.

607 Interestingly, many of the enzyme-encoding fatty acid biosynthetic and TAG
608 accumulation-related genes that were differentially expressed in AIL7-OE lines are also known
609 to be transcriptionally regulated by LAFL TFs and their TF targets (Kumar, et al. 2020). These
610 include, for example, *KAS1*, *ACP1*, *FAD3*, *FAE1* and *PDAT2* (Table 1). This insinuates that

611 AIL7 may possibly function within the LAFL transcriptional cascade. While additional research
612 will be required to demonstrate this with certitude, if this were indeed the case, AIL7 would not
613 be the only AIL protein to function in this manner. For example, BBM has been shown to bind
614 and/or transcriptionally activate LEC1, LEC2, ABI3 and AGL15 (directly), as well as FUS3
615 (indirectly), while PLT2 directly activates LEC1, LEC2 and FUS3 (Horstman, et al. 2017). This
616 suggests that these two AIL proteins function directly upstream of LAFL/AGL15 genes.
617 While no shifts in the expression of most known regulators of seed maturation and fatty
618 acid biosynthesis (e.g., LEC1, LEC2, FUS3, ABI3, ABI4, ABI5, L1L, AGL15, MUM4, GL2,
619 TT2, TT8, ASIL1, TTG1, MYB89, MYB96, MYB118, WRKY6, bZIP67, VAL2, PKL and
620 WRI1) were noted in developing siliques from our AIL7-OE lines (Table S4), significant up-
621 regulation of *VAL1* was observed (Table 1). This gene encodes a B3 domain protein that
622 functions to silence the MADS-box gene *AGL15* (Chen, et al. 2018), which is known to control
623 *LEC2* expression through a positive feedback loop (Fatihi, et al. 2016). However, the up-
624 regulation of *VAL1* in AIL7-OE lines did not elicit any alteration in the expression of *AGL15* or
625 *LEC2*. Similarly, the up-regulation of *BBM* and *PLT2* in AIL7-OE lines (Table S6) did not
626 provoke any associated increases in *LAFL/AGL15* expression, suggesting that other, as of yet
627 unidentified, TFs within the cascade may be associated with the lipid-related phenotype.
628 There is also evidence that at least certain AIL proteins may function downstream of
629 LAFL proteins. For example, *BBM* expression is reduced in *lafl* mutant seeds (Horstman, et al.
630 2017), and it has been found previously that an ABI3-like factor from *Phaseolus vulgaris* can
631 bind and activate Arabidopsis *AIL5*, the down-regulation of which led to changes in the
632 expression of several seed maturation-related genes (Sundaram, et al. 2013). In addition,
633 Arabidopsis FUS3 has been shown to bind *BBM*, *PLT2*, *AIL6* and *AIL7* promoters (Wang and

634 Perry 2013), and up-regulation of *AIL5*, *AIL6* and *AIL7* has been observed in Arabidopsis
635 seedlings following the induced over-expression of *FUS3* (Zhang, et al. 2016). A search for
636 putative direct gene targets of a small number of available TFs with known functions in TAG
637 accumulation in the Plant Cistrome Database in this study indicated that *AIL7* may also be a
638 direct target of AGL15 and ABI5, the latter of which functions synergistically with ABI4 in the
639 activation of *DGAT1* in Arabidopsis seedlings under stress conditions (Kong, et al. 2013).
640 However, further investigation will be required to elucidate the precise positioning and role of
641 AIL7 in the lipid biosynthetic regulatory cascade.
642 Seeds from AIL7-OE lines also displayed abnormalities in terms of morphology and
643 seedling growth. While AIL7-OE seeds did not exhibit consistent discrepancies in weight or area
644 compared to wild-type seeds (Figs. 3C and 3D), and both AIL7-OE and wild-type plants
645 possessed seed coats with epidermal cells exhibiting a characteristic reticulated pattern with
646 central columellae, unlike their wild-type counterparts, *AIL7* over-expressing seeds were twisted
647 in shape rather than elliptical, which suggests a defect in morphological differentiation (Fig. 3).
648 A similar twisted seed phenotype has been observed in Arabidopsis plants bearing defects in the
649 embryonic cuticle or extracuticular sheath at the embryo surface, which leads to the improper
650 separation of the embryo and endosperm during embryo elongation (Kazaz, et al. 2020; Moussu,
651 et al. 2017). For example, mutations in several genes involved in embryonic cuticle deposition,
652 including the simultaneous disruption/reduction in activity of three Δ9 stearoyl-ACP desaturases
653 (*FATTY ACID BIOSYNTHESIS 2* [*FAB2*], *ACYL-ACYL CARRIER PROTEIN 1* and *5* [*AAD1* and
654 *AAD5*]), as well as *zhoupi* (*zou*), *abnormal leaf shape1* (*ale1*), *twisted seed1* (*tws1*) and
655 *gassho1/2* (*gso1/2*) mutants, have been found to display an embryo/endosperm adhesion
656 phenotype, which presents as twisted seeds (Creff, et al. 2019; Fiume, et al. 2016; Kazaz, et al.

2020; Xing, et al. 2013). Similarly, mutations in the *KERBEROS* (*KRS*) gene lead to a very similar phenotype, which has been suggested to occur due to a deficiency in the production of the embryonic sheath rather than the cuticle itself (Moussu, et al. 2017). While none of these genes were differentially expressed in developing siliques from AIL7-OE lines (Supplemental Data S1), the substantial alteration in FA composition seen in transgenic seeds compared to wild-type (Figs. 2A and S2A) could feasibly have led to cuticle defects in the seed. However, further analyses will be required to establish this definitively.

Morphological defects were also observed in early vegetative and root development in AIL7-OE seedlings compared to wild-type controls, with AIL7-OE seedlings exhibiting reductions in the growth of aboveground vegetative organs, along with shorter primary roots. This decrease in initial seedling growth may be attributable to the decrease in seed oil content seen in these lines, which is akin to deficits in seedling establishment observed in *wri1* mutants that presumably result from a sparsity of seed storage oil (Cernac, et al. 2006). However, it is also possible that the defects in seedling growth that occur with *AIL7* over-expression stem from its known functions in meristematic-related processes ((Mudunkothge and Krizek 2012); Table S7) or embryo development ((Horstman, et al. 2017); Table S11). In line with this, *plt1plt2ail6* triple mutants have been reported to exhibit an aberrant organization of the embryonic root pole, resulting in rootless seedlings (Aida, et al. 2004). Although the *napin* promoter used to drive the expression of the *AIL7* coding sequence in the transgenic cassette in this study has been shown previously to be active specifically in seeds (Ellerström et al. 1996; Stålberg et al. 1993), we cannot rule out at this point that some level of leaky expression in seedlings may have also contributed to these morphological abnormalities.

679 In addition to its roles in FA biosynthesis, seed oil accumulation, and seed/seedling
680 growth, our RNA-Seq data suggests that *AIL7* may also have other roles in processes with
681 potential agronomic importance, such as photosynthesis and stress response (Tables S8 and S9).
682 Interestingly, all DEGs related to photosynthesis were up-regulated in AIL7-OE lines. This is
683 reminiscent of other genes functioning within the LAFL regulatory network such as *LEC1* and
684 *FUS3*, which have also been found to positively regulate genes involved in photosynthesis either
685 directly or indirectly (Pelletier, et al. 2017; Yamamoto, et al. 2010). Similarly, various genes
686 with roles in photosynthesis and defense response have also been observed to be up-regulated in
687 *ant ail6* inflorescences (Krizek, et al. 2016), which may correlate with the fact that *AIL6* was
688 found to be significantly down-regulated in the developing siliques of AIL7-OE lines (Table S6)
689 However, it remains to be determined whether the up-regulation of such genes translates into
690 increased photosynthetic efficiency/capacity and/or improvements in stress tolerance in
691 developing seeds and/or siliques in the AIL7-OE lines.
692 Taken together, our data demonstrate that in addition to its known functions in meristem
693 development and maintenance, as well as cell proliferation and shoot phyllotaxy, *AIL7* also plays
694 an important role in seed FA biosynthesis and oil accumulation. While the precise mechanisms
695 through which *AIL7* regulates these processes remain to be unraveled, our findings suggest that
696 like BBM and PLT2, AIL7 may function within the LAFL regulatory cascade. Moreover, our
697 RNA-Seq data suggests additional possible functions for *AIL7* in starch metabolism, embryonic
698 cuticle deposition, photosynthesis and stress response. However, it remains to be determined
699 whether such transcriptional changes translate into alterations in the traits themselves in AIL7-
700 OE lines, and therefore further verification of phenotypic and physiological assessments based
701 on our transcriptomic data will be essential for confirming such roles. These findings not only

enhance our understanding of the regulatory network controlling lipid biosynthesis in seeds, but also hint at additional functions of *AIL7*, which could benefit downstream breeding and/or metabolic engineering endeavours.

ACCESSION NUMBERS

The RNA-Seq data sets are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (accession no. PRJNA725102).

SUPPLEMENTAL MATERIAL

Supplemental Data S1. RNA-Seq data for all genes identified in AIL7-OE and wild-type siliques.

Supplemental Figure 1. Quantitative real-time RT-PCR analysis of *AIL7* expression in siliques from wild-type (A) and homozygous AIL7-OE-1 lines (T_3; with wild-type results included as a reference) (B) at various time points after flowering. Each point represents the mean of two biological replicates relative to the internal control, *PP2AA3*. Grey denotes AIL7-OE lines while black indicates wild-type. Three technical replicates were carried out for every qRT-PCR assay, and in all cases, bars indicate standard errors. DAF, days after flowering; wt, wild-type.

Supplemental Figure 2. Fatty acid composition and lipid content of T_2 AIL7-OE and wild-type seeds. **(A)** Mean fatty acid composition of oil from T_2 AIL7-OE seeds. Blocks represent mean values of wild-type (n=18) and AIL7-OE (n=21) independent lines. **(B)** Mean seed oil content of

T_2 seeds. Blocks represent mean values from wild-type (*n*=18) and AIL7-OE (*n*=24) independent lines. Two technical replicates were carried out for each line analyzed. Bars denote standard errors. Significant increases compared to wild-type (as measured by 2-tailed Student's t-tests assuming unequal variance) are indicated by ++ ($P \leq 0.01$) while significant decreases are denoted by – ($P \leq 0.05$) and - - ($P \leq 0.01$). wt, wild type.

Supplemental Figure 3. GO term functional classification for biological process. **(A)** Up-regulated DEGs. **(B)** Down-regulated DEGs.

Supplemental Figure 4. GO term functional classification for molecular function. **(A)** Up-regulated DEGs. **(B)** Down-regulated DEGs.

Supplemental Figure 5. GO term functional classification for cellular component. **(A)** Up-regulated DEGs. **(B)** Down-regulated DEGs.

Supplemental Table S1. Primers used for qRT-PCR validation of RNA-Seq results.

Supplemental Table S2. RNA-Seq read and alignment data.

Supplemental Table S3. Effect of *AIL7* over-expression on the expression of genes within the lipid metabolic GO biological process category.

Supplemental Table S4. Genes with roles in FA and TAG biosynthetic pathways that did not exhibit significant alterations in expression levels in AIL7-OE lines compared to wild-type plants.

Supplemental Table S5. Effect of *AIL7* over-expression on the expression of genes with known or putative functions in starch biosynthesis and degradation.

Supplemental Table S6. Effect of *AIL7* over-expression on the expression of *AIL* family members.

Supplemental Table S7. Effect of *AIL7* over-expression on the expression of genes within meristem-related GO biological process categories.

Supplemental Table S8. Effect of *AIL7* over-expression on the expression of genes with functions in the photosynthetic process.

Supplemental Table S9. Effect of *AIL7* over-expression on the expression of genes within the response to stress GO biological process category.

Supplemental Table S10. Effect of *AIL7* over-expression on the expression of genes within the anatomical structure and morphogenesis GO biological process category.

Supplemental Table S11. Effect of *AIL7* over-expression on the expression of genes within the embryo development GO biological process category.

Supplemental Table S12. Differentially expressed genes derived from comparative RNA-Seq analysis that are putative direct targets of AIL7.

774 **LITERATURE CITED**

775 Aida M, Beis D, Heidstra R, Willemsen V, Blilou I, Galinha C, Nussaume L, Noh YS, Amasino
776 R, Scheres B (2004) The *PLETHORA* genes mediate patterning of the Arabidopsis root stem cell
777 niche. Cell 119:109-120

778 Alamery S, Tirnaz S, Bayer P, Tollenaere R, Chaloub B, Edwards D, Batley J (2017) Genome-
779 wide identification and comparative analysis of NBS-LRR resistance genes in *Brassica napus*.
780 Crop and Pasture Science 69:72-93

781 Baud S, Dubreucq B, Miquel M, Rochat C, Lepiniec L (2008) Storage reserve accumulation in
782 Arabidopsis: metabolic and developmental control of seed filling. The Arabidopsis book /
783 American Society of Plant Biologists 6:e0113

784 Baud S, Mendoza MS, To A, Harscoet E, Lepiniec L, Dubreucq B (2007) WRINKLED1
785 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during
786 seed maturation in Arabidopsis. Plant J 50:

787 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful
788 approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological)
789 57:289-300

790 Cernac A, Andre C, Hoffmann-Benning S, Benning C (2006) WRI1 is required for seed
791 germination and seedling establishment. Plant Physiol 141:745-757

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

794 Chen M, Du X, Zhu Y, Wang Z, Hua S, Li Z, Guo W, Zhang G, Peng J, Jiang L (2012) Seed
795 *Fatty Acid Reducer* acts downstream of gibberellin signalling pathway to lower seed fatty acid
796 storage in Arabidopsis. Plant, Cell & Environment 35:2155-2169

797 Chen N, Veerappan V, Abdelmageed H, Kang M, Allen RD (2018) HSI2/VAL1 silences AGL15
798 to regulate the developmental transition from seed maturation to vegetative growth in
799 Arabidopsis. The Plant Cell 30:600-619

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

802 Creff A, Brocard L, Joubes J, Taconnat L, Doll NM, Marsollier AC, Pascal S, Galletti R, Boeuf
803 S, Moussu S, Widiez T, Domergue F, Ingram G (2019) A stress-response-related inter-
804 compartmental signalling pathway regulates embryonic cuticle integrity in Arabidopsis. PLoS
805 genetics 15:e1007847

806 Cross JM, Clancy M, Shaw JR, Greene TW, Schmidt RR, Okita TW, Hannah LC (2004) Both
807 subunits of ADP-glucose pyrophosphorylase are regulatory. Plant Physiol 135:137-144

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

811 Dahlqvist A, Stahl U, Lenman M, Banas A, Lee M, Sandager L, Ronne H, Stymne H (2000)
812 Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-
813 independent formation of triacylglycerol in yeast and plants. Proc Natl Acad Sci USA 97:6487-
814 6492

Du Y, Scheres B (2017) PLETHORA transcription factors orchestrate de novo organ patterning during Arabidopsis lateral root outgrowth. Proc Natl Acad Sci USA 114:11709-11714

Eastmond PJ (2006) SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating Arabidopsis seeds. The Plant Cell 18:665-675

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

Eskin NAM, Vaisey-Genser M, Durance-Todd S, Przybylski R (1989) Stability of low linolenic acid canola oil to frying temperatures. Journal of the American Oil Chemists' Society 66:1081-1084

Fatihi A, Boulard C, Bouyer D, Baud S, Dubreucq B, Lepiniec L (2016) Deciphering and modifying LAFL transcriptional regulatory network in seed for improving yield and quality of storage compounds. Plant Sci 250:198-204

Fiume E, Guyon V, Remoué C, Magnani E, Miquel M, Grain D, Lepiniec L (2016) TWS1, a novel small protein, regulates various aspects of seed and plant development. Plant Physiol 172:1732-1745

Focks N, Benning C (1998) Wrinkled1: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. Plant Physiol 118:91-101

Gao H, Gao Y, Zhang F, Liu B, Ji C, Xue J, Yuan L, Li R (2021) Functional characterization of an novel acyl-CoA:diacylglycerol acyltransferase 3-3 (CsDGAT3-3) gene from *Camelina sativa*. Plant Sci 303:110752

836 Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: a versatile and
837 flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol Biol
838 42:819-832

839 Horstman A, Li M, Heidmann I, Weemen M, Chen B, Muino JM, Angenent GC, Boutilier K
840 (2017) The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to
841 induce somatic embryogenesis. Plant Physiol 175:848-857

842 Horstman A, Willemsen V, Boutilier K, Heidstra R (2014) AINTEGUMENTA-LIKE proteins:
843 hubs in a plethora of networks. Trends Plant Sci 19:146-157

844 Huai D, Zhang Y, Zhang C, Cahoon EB, Zhou Y (2015) Combinatorial effects of fatty acid
845 elongase enzymes on nervonic acid production in *Camelina sativa*. Plos One 10:e0131755

846 Huang AHC (2018) Plant lipid droplets and their associated proteins: potential for rapid
847 advances. Plant Physiol 176:1894-1918

848 Jain T (2020) Fatty acid composition of oilseed crops: A review. In: Thakur M, Modi VK (eds)
849 Emerging Technologies in Food Science: Focus on the Developing World. Springer Singapore,
850 Singapore, pp 147-153

851 James DW, Jr., Lim E, Keller J, Plooy I, Ralston E, Dooner HK (1995) Directed tagging of the
852 Arabidopsis *FATTY ACID ELONGATION1* (*FAE1*) gene with the maize transposon activator.
853 The Plant Cell 7:309-319

854 Jofuku KD, Omidyar PK, Gee Z, Okamuro JK (2005) Control of seed mass and seed yield by the
855 floral homeotic gene *APETALA2*. Proc Natl Acad Sci USA 102:3117-3122

856 Josefsson LG, Lenman M, Ericson ML, Rask L (1987) Structure of a gene encoding the 1.7 S
857 storage protein, napin, from *Brassica napus*. J Biol Chem 262:12196-12201

858 Joung JG, Corbett AM, Fellman SM, Tieman DM, Klee HJ, Giovannoni JJ, Fei Z (2009) Plant
859 MetGenMAP: an integrative analysis system for plant systems biology. Plant Physiol 151:1758-
860 1768

861 Katche E, Quezada-Martinez D, Katche EI, Vasquez-Teuber P, Mason AS (2019) Interspecific
862 hybridization for Brassica crop improvement. Crop Breeding, Genetics and Genomics 1:e190007

863 Kazaz S, Barthole G, Domergue F, Ettaki H, To A, Vasselon D, De Vos D, Belcram K, Lepiniec
864 L, Baud S (2020) Differential activation of partially redundant Δ9 stearoyl-ACP desaturase genes
865 is critical for omega-9 monounsaturated fatty acid biosynthesis during seed development in
866 Arabidopsis. The Plant Cell 32:3613-3637

867 Kelly AA, Quettier A-L, Shaw E, Eastmond PJ (2011) Seed storage oil mobilization is important
868 but not essential for germination or seedling establishment in Arabidopsis. Plant Physiol
869 157:866-875

870 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate
871 alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome
872 Biol 14:R36

873 Kong Q, Singh SK, Mantyla JJ, Pattanaik S, Guo L, Yuan L, Benning C, Ma W (2020) Teosinte
874 branched1/cycloidea/proliferating cell factor4 interacts with wrinkled1 to mediate seed oil
875 biosynthesis. Plant Physiol 184:658-665

876 Kong Q, Yang Y, Low PM, Guo L, Yuan L, Ma W (2020) The function of the WRI1-TCP4
877 regulatory module in lipid biosynthesis. Plant Signal Behav 15:1812878

878 Kong Q, Yuan L, Ma W (2019) WRINKLED1, a “master regulator” in transcriptional control of
879 plant oil biosynthesis. Plants 8:238

880 Kong Y, Chen S, Yang Y, An C (2013) ABA-insensitive (ABI) 4 and ABI5 synergistically
881 regulate DGAT1 expression in Arabidopsis seedlings under stress. FEBS Lett 587:3076-3082

882 Krizek BA (2015) *AINTEGUMENTA-LIKE* genes have partly overlapping functions with
883 *AINTEGUMENTA* but make distinct contributions to *Arabidopsis thaliana* flower development.
884 J Exp Bot 66:4537-4549

885 Krizek BA, Bequette CJ, Xu K, Blakley IC, Fu ZQ, Stratmann JW, Loraine AE (2016) RNA-Seq
886 links the transcription factors *AINTEGUMENTA* and *AINTEGUMENTA-LIKE6* to cell wall
887 remodeling and plant defense pathways. Plant Physiol 171:2069-2084

888 Kumar N, Chaudhary A, Singh D, Teotia S (2020) Transcriptional regulation of seed oil
889 accumulation in *Arabidopsis thaliana*: role of transcription factors and chromatin remodelers.
890 Journal of Plant Biochemistry and Biotechnology 29:754-768

891 Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M,
892 Kirkbride R, Horvath S, Drews GN, Fischer RL, Okamuro JK, Harada JJ, Goldberg RB (2010)
893 Global analysis of gene activity during Arabidopsis seed development and identification of seed-
894 specific transcription factors. Proc Natl Acad Sci USA 107:8063-8070

895 Lemieux B, Miquel M, Somerville C, Browse J (1990) Mutants of Arabidopsis with alterations
896 in seed lipid fatty acid composition. Theor Appl Genet 80:234-240

897 Lin Y, Ulanov AV, Lozovaya V, Widholm J, Zhang G, Guo J, Goodman HM (2006) Genetic and
898 transgenic perturbations of carbon reserve production in Arabidopsis seeds reveal metabolic
899 interactions of biochemical pathways. Planta 225:153-164

900 Lu CF, Napier JA, Clemente TE, Cahoon EB (2011) New frontiers in oilseed biotechnology:
901 meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications.
902 Curr Opin Biotechnol 22:252-259

903 Mansfield SG, Briarty LG (1992) Cotyledon cell development in *Arabidopsis thaliana* during
904 reserve deposition. Canadian Journal of Botany 70:151-164

905 Moussu S, Doll NM, Chamot S, Brocard L, Creff A, Fourquin C, Widiez T, Nimchuk ZL,
906 Ingram G (2017) ZHOUPI and KERBEROS mediate embryo/endosperm separation by
907 promoting the formation of an extracuticular sheath at the embryo surface. The Plant Cell
908 29:1642-1656

909 Msanne J, Kim H, Cahoon EB (2020) Biotechnology tools and applications for development of
910 oilseed crops with healthy vegetable oils. Biochimie 178:4-14

911 Mudunkothge JS, Krizek BA (2012) Three Arabidopsis AIL / PLT genes act in combination to
912 regulate shoot apical meristem function. Plant J 71:108-121

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

915 Nadler SG, Tritschler D, Haffar OK, Blake J, Bruce AG, Cleaveland JS (1997) Differential
916 expression and sequence-specific interaction of karyopherin alpha with nuclear localization
917 sequences. J Biol Chem 272:4310-4315

918 Nole-Wilson S, Tranby TL, Krizek BA (2005) AINTEGUMENTA-like (AIL) genes are
919 expressed in young tissues and may specify meristematic or division-competent states. Plant Mol
920 Biol 57:613-628

O'Malley RC, Carol S-s, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR (2016) Cistrome and epicistrome features shape the regulatory DNA landscape resource cistrome and epicistrome features shape the regulatory DNA landscape. Cell 165:1280-1292

Ohto M-a, Fischer RL, Goldberg RB, Nakamura K, Harada JJ (2005) Control of seed mass by APETALA2. Proc Natl Acad Sci USA 102:3123-3128

Pelletier JM, Kwong RW, Park S, Le BH, Baden R, Cagliari A, Hashimoto M, Munoz MD, Fischer RL, Goldberg RB, Harada JJ (2017) LEC1 sequentially regulates the transcription of genes involved in diverse developmental processes during seed development. Proc Natl Acad Sci USA 114:E6710-E6719

Prasad K, Grigg SP, Barkoulas M, Yadav RK, Sanchez-Perez GF, Pinon V, Blilou I, Hofhuis H, Dhonukshe P, Galinha C, Mahonen AP, Muller WH, Raman S, Verkleij AJ, Snel B, Reddy GV, Tsiantis M, Scheres B (2011) Arabidopsis PLETHORA transcription factors control phyllotaxis. Curr Biol 21:1123-1128

Rahman H, Singer SD, Weselake RJ (2013) Development of low-linolenic acid *Brassica oleracea* lines through seed mutagenesis and molecular characterization of mutants. Theor Appl Genet 126:1587-1598

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

Siloto RMP, Findlay K, Lopez-Villalobos A, Yeung EC, Nykiforuk CL, Moloney MM (2006) The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. The Plant Cell 18:1961-1974

943 Singer SD, Chen G, Mietkiewska E, Tomasi P, Jayawardhane K, Dyer JM, Weselake RJ (2016)
944 Arabidopsis GPAT9 contributes to synthesis of intracellular glycerolipids but not surface lipids.
945 J Exp Bot 67:4627-4638

946 Singer SD, Weselake RJ (2018) Production of biodiesel from plant oils. In: Chen G, Weselake
947 RJ, Singer SD (eds) Plant Bioproducts. Springer New York, New York, NY, pp 41-58

948 Singer SD, Weselake RJ, Rahman H (2014) Development and characterization of low alpha-
949 linolenic acid *Brassica oleracea* lines bearing a novel mutation in a 'class a' *FATTY ACID*
950 *DESATURASE 3* gene. BMC Genetics 15:94

951 Stålberg K, Ellerstrom M, Josefsson LG, Rask L (1993) Deletion analysis of a 2S seed storage
952 protein promoter of *Brassica napus* in transgenic tobacco. Plant Mol Biol 23:671-683

953 Subedi U, Jayawardhane KN, Pan X, Ozga J, Chen G, Foroud NA, Singer SD (2020) The
954 potential of genome editing for improving seed oil content and fatty acid composition in oilseed
955 crops. Lipids 55:495-512

956 Sundaram S, Kertbundit S, Shakirov EV, Iyer LM, Juricek M, Hall TC (2013) Gene networks
957 and chromatin and transcriptional regulation of the phaseolin promoter in Arabidopsis. The Plant
958 Cell 25:2601-2617

959 Tian B, Lu T, Xu Y, Wang R, Chen C (2019) Identification of genes associated with ricinoleic
960 acid accumulation in *Hiptage benghalensis* via transcriptome analysis. Biotechnol Biofuels 12:
961 1-16

962 Wang F, Perry SE (2013) Identification of direct targets of FUSCA3, a key regulator of
963 Arabidopsis seed development. Plant Physiol 161:1251-1264

964 Wang Z, Yang M, Sun Y, Yang Q, Wei L, Shao Y, Bao G, Li W (2019) Overexpressing
965 *Sesamum indicum* L.'s DGAT1 increases the seed oil content of transgenic soybean. Mol
966 Breeding 39:101

967 Weiss SB, Kennedy EP, Kiyasu JY (1960) The enzymatic synthesis of triglycerides. J Biol Chem
968 235:40-44

969 Weselake RJ, Taylor DC, Rahman MH, Shah S, Laroche A, McVetty PBE, Harwood JL (2009)
970 Increasing the flow of carbon into seed oil. Biotechnol Adv 27:866-878

971 Woodfield HK, Fenyk S, Wallington E, Bates RE, Brown A, Guschina IA, Marillia EF, Taylor
972 DC, Fell D, Harwood JL, Fawcett T (2019) Increase in lysophosphatidate acyltransferase activity
973 in oilseed rape (Brassica napus) increases seed triacylglycerol content despite its low intrinsic
974 flux control coefficient. New Phytol 224:700-711

975 Wu G-Z, Xue H-W (2010) Arabidopsis β-Ketoacyl-[Acyl Carrier Protein] Synthase I is crucial
976 for fatty acid synthesis and plays a role in chloroplast division and embryo development. The
977 Plant Cell 22:3726-3744

978 Xing Q, Creff A, Waters A, Tanaka H, Goodrich J, Ingram GC (2013) ZHOUPI controls
979 embryonic cuticle formation via a signalling pathway involving the subtilisin protease
980 ABNORMAL LEAF-SHAPE1 and the receptor kinases GASSHO1 and GASSHO2.
981 Development 140:770-779

982 Yamamoto A, Kagaya Y, Usui H, Hobo T, Takeda S, Hattori T (2010) Diverse roles and
983 mechanisms of gene regulation by the Arabidopsis seed maturation master regulator FUS3
984 revealed by microarray analysis. Plant Cell Physiology 51:2031-2046

985 Zhang LZ, Tan QM, Lee R, Trethewy A, Lee YH, Tegeder M (2010) Altered xylem-phloem
986 transfer of amino acids affects metabolism and leads to increased seed yield and oil content in
987 Arabidopsis. The Plant Cell 22:3603-3620

988 Zhang M, Cao X, Jia Q, Ohlrogge J (2016) FUSCA3 activates triacylglycerol accumulation in
989 Arabidopsis seedlings and tobacco BY2 cells. Plant J 88:95-107

990 **Table 1.** Effect of *AIL7* over-expression on the expression of a selection of genes with known or

991 putative functions in seed fatty acid and triacylglycerol biosynthesis

Gene ID	Ratio	p-value	Description
Transcriptional regulators			
AT2G30470	1.67	1.87e-2	VAL1 (VIVIPAROUS1/ABI3-LIKE1)
FA biosynthesis/modification			
AT1G25450	0.09	5.08e-6	KCS5 (3-ketoacyl-CoA synthase 5)
AT1G71160	0.09	3.08e-2	KCS7 (3-ketoacyl-CoA synthase 7)
AT4G34520	0.16	8.91e-6	FAE1 (FATTY ACID ELONGASE 1)
AT1G24470	0.32	1.70e-3	KCR2 (B-KETOACYL-COA REDUCTASE 2)
AT4G16210	0.48	1.07e-7	Enoyl-Coa hydratase/isomerase family protein
AT2G29980	0.52	4.44e-4	FAD3 (FATTY ACID DESATURASE 3)
AT5G46290	2.42	5.59e-4	KAS I (3-ketoacyl-acyl carrier protein synthase I)
AT1G36180	2.43	7.74e-7	ACC2 (ACETYL-COA CARBOXYLASE 2)
AT4G00520	2.51	7.64e-5	Acyl-CoA thioesterase family protein
AT5G27200	5.54	1.65e-5	ACP5 (ACYL CARRIER PROTEIN 5)
AT1G77590	5.63	1.41e-7	LACS9 (LONG CHAIN ACYL-COA SYNTHETASE 9)
TAG biosynthesis and storage			
AT3G44830	0.59	4.3e-2	PDAT2 (PHOSPHOLIPID:DIACYLGLYCEROL ACYLTRANSFERASE 2)
AT4G25140	0.76	4.2e-2	OLEO1 (OLEOSIN 1)
AT4G37740	2.38	3.07e-4	GRF2 (GROWTH-REGULATING FACTOR 2-LIKE)
AT3G18850	6.23	5.64e-11	LPAT5 (LYSOPHOSPHATIDYL ACYLTRANSFERASE 5)
AT3G11325	12.3	2.79e-2	Putative acyltransferase

992

993 Raw P-values derived from comparative RNA-Seq analysis were adjusted for multiple testing

994 using a false discovery rate (FDR), and genes that are up-regulated or down-regulated in AIL7-

995 OE lines compared to wild-type plants with an FDR of less than 0.05 are listed here. Ratios

996 indicate relative expression levels in AIL7-OE lines compared to wild-type plants.

997

998

Table 2. Effect of *AIL7* over-expression on the expression of genes with known or putative functions in fatty acid or triacylglycerol breakdown

Gene ID	Ratio	p-value	Description
AT1G54000	0.11	5.94e-3	GDSL-like lipase/acylhydrolase superfamily protein
AT2G42990	0.17	3.85e-2	GDSL-like lipase/acylhydrolase superfamily protein
AT1G54010	0.24	2.91e-3	GDSL-like lipase/acylhydrolase superfamily protein
AT3G05180	0.26	2.00e-2	GDSL-like lipase/acylhydrolase superfamily protein
AT5G03610	0.34	2.05e-11	GDSL-like lipase/acylhydrolase superfamily protein
AT1G51440	0.39	1.52e-2	DAD1-like lipase 2
AT5G22810	0.41	4.13e-3	GDSL-like lipase/acylhydrolase superfamily protein
AT1G28610	0.45	1.03e-2	GDSL-like lipase/acylhydrolase superfamily protein
AT4G16210	0.48	1.07e-7	enoyl-CoA hydratase/isomerase family protein
AT3G07400	2.19	2.18e-5	lipase class 3 family protein
AT5G14450	2.24	1.86e-4	GDSL-motif lipase/acylhydrolase superfamily protein
AT1G09390	2.26	1.97e-2	GDSL-motif lipase/acylhydrolase superfamily protein
AT4G01130	2.44	3.29e-2	GDSL-like lipase/acylhydrolase superfamily protein
AT1G71250	2.83	3.94e-3	GDSL-motif lipase/acylhydrolase superfamily protein
AT1G06800	2.87	4.51e-6	DAD1-like lipase 4
AT3G57140	3.1	1.24e-5	SDP1L (SUGAR-DEPENDENT 1-LIKE)
AT1G75900	3.45	3.63e-4	GDSL-motif lipase/acylhydrolase superfamily protein
AT5G45670	3.59	1.62e-3	GDSL-motif lipase/acylhydrolase superfamily protein
AT5G03820	4.33	5.98e-3	GDSL-motif lipase/acylhydrolase superfamily protein
AT5G24210	4.69	1.48e-2	lipase class 3 family protein
AT1G71120	6.13	1.39e-10	GLIP6 (GDSL-motif lipase/hydrolase 6)
AT4G10950	80.6	2.62e-8	GDSL-type esterase/lipase 77

Raw P-values derived from comparative RNA-Seq analysis were adjusted for multiple testing using a false discovery rate (FDR), and genes with an FDR of less than 0.05 are listed here. Ratios indicate relative expression levels in AIL7-OE lines compared to wild-type plants.

Figure captions

Figure 1. Generation of seed-specific *AIL7* over-expression Arabidopsis lines. (A) Schematic representation (not to scale) of the AIL7-OE construct used in this study. Arrows indicate the direction of transcription. **(B)** Quantitative real-time RT-PCR analysis of *AIL7* expression in T_2 siliques (14 DAF) from AIL7-OE-1, AIL7-OE-19, and wild-type lines. Blocks denote mean *AIL7* transcript levels from three biological replicates relative to the internal control, *PP2AA3*. Three technical replicates were carried out for qRT-PCR assays and bars indicate standard errors. AIL7, Arabidopsis *AINTEGUMENTA 7* coding sequence; LB, left T-DNA border; napin-p, *Brassica napus napin* promoter; nosp, *NOPALINE SYNTHASE* promoter; nost, *NOPALINE SYNTHASE* transcriptional terminator; NPTII, *NEOMYCIN PHOSPHOTRANSFERASE II*; RB, right T-DNA border; rbcSt, *Pisum sativum RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE* transcriptional terminator; wt, wild-type.

Figure 2. Fatty acid composition and oil content of homozygous T_3 AIL7-OE and wild-type seeds. (A) Mean fatty acid composition of oil from homozygous T_3 AIL7-OE seeds. Blocks indicate mean values of wild-type (n=6) and two independent AIL7-OE lines (AIL7-OE-1 [n=6] and AIL7-OE-19 [n=5]). **(B)** Mean seed oil content of T_3 seeds. Blocks represent mean values from wild-type (n=6) and two independent homozygous AIL7-OE lines (AIL7-OE-1 [n=6] and AIL7-OE-19 [n=5]). Two technical replicates were carried out for each line analyzed. Bars denote standard errors. Very significant increases and decreases compared to wild-type (as measured by 2-tailed Student's t-test assuming unequal variance) are indicated by + and - ($P \leq 0.01$). wt, wild-type.

Figure 3. Seed morphology of AIL7-OE lines. (A) Representative light microscopic images of wild-type and homozygous T_3 AIL7-OE mature seeds. Scale bars = 1 mm. **(B)** Representative SEM images of wild-type and homozygous T_3 AIL7-OE seeds. Scale bars = 100 μm. **(C)** Mean mature seed weights from wild-type (from 6 individual plants), as well as homozygous T_3 AIL7-OE-1 (from 6 individual plants) and AIL7-OE-19 (from 4 individual plants) lines. Two to three replicate seed batches were used for each measurement. (D) Mean mature seed areas of wild-type (*n*=61 from 3 individual plants), as well as homozygous T_3 AIL7-OE-1 (*n*=88 from 3 individual plants) and AIL7-OE-19 (*n*=53 from 3 individual plants) lines. Asterisks denote significant differences from wild-type as determined by 2-tailed Student's t-tests assuming unequal variance, $P \leq 0.01$.

Figure 4. Growth of homozygous T_3 AIL7-OE and wild-type seedlings. (A) Seedlings grown vertically on solid medium were photographed 9 days post-germination and are representative of two independent experiments. AIL7-OE seedlings are shown to the left of the black vertical line, while wild-type seedlings are present to the right**. (B)** Root lengths of wild-type (*n*=35), AIL7-OE-1 (*n*=20), and AIL7-OE-19 (*n*=16) seedlings grown vertically on solid medium 9 days post-germination. Blocks denote mean values and bars indicate standard errors. Asterisks denote means that are significantly different from wild-type as determined by 2-sided Student's t-tests assuming unequal variance ($P \leq 0.01$). **(C and D)** Representative soil-grown AIL7-OE **(C)** and wild-type **(D)** lines 24 days post-germination. DPG, days post-germination; wt, wild type.

Figure 5. Simplified diagrammatic representation of fatty acid and triacylglycerol biosynthesis in Arabidopsis seeds. A selection of genes with known or putative functions in

seed fatty acid and triacylglycerol biosynthesis are displayed, with those transcriptionally affected by *AIL7* over-expression labeled with red or blue arrows to represent up- or down-regulation, respectively. Enzymes/genes are displayed within green rectangles. Abbreviations: ACC, acetyl-CoA carboxylase; ACP, acyl carrier protein; ACT, acyl-CoA thioesterase; DAG, diacylglycerol; DGAT, diacylglycerol acyltransferase; ECR, enoyl-CoA reductase; ER, endoplasmic reticulum; FAD, fatty acid desaturase; FAE, fatty acid elongase; FAT, acyl-ACP thioesterase; FFA, free fatty acid; G3P, *sn*-glycerol-3-phosphate; GPAT, *sn*-glycerol-3-phosphate acyltransferase; GRF2, growth-regulating factor 2-like; KAS, ketoacyl-ACP synthase; KCR, ketoacyl-CoA reductase; KCS, ketoacyl-CoA synthase; LACS, long-chain acyl-CoA synthase; LPA, lysophosphatidic acid; LPAAT, acyl-CoA:lysophosphatidic acid acyltransferase; MCAT, malonyl-CoA: ACP acyltransferase; OLEO, Oleosin; PA, phosphatidic acid; PAP, phosphatidic acid phosphatase; PC, phosphatidylcholine; PDAT, phospholipid:diacyglycerol acyltransferase; SAD, stearoyl-ACP desaturase; TAG, triacylglycerol; VAL1, VIVIPAROUS1/ABI3-LIKE1. This model was developed based on the transcriptome data garnered in the current study, as well as information from Chen et al. (2015) and Tian et al. (2019).

Figure 6. **qRT-PCR validation of RNA-Seq results.** qRT-PCR was carried out on RNA derived from the developing siliques (14 DAF) of AIL7-OE (4-5 biological replicates) and wild-type (3 biological replicates) lines, using the internal control, *PP2AA3,* as a reference. Three technical replicates were carried out in each case. Blocks represent $\log_2$ fold-changes in expression derived from qRT-PCR and RNA-Seq data, respectively. ABCG12, ABC-2 type transporter family protein; ACP5, acyl-carrier protein 5; AIL7, AINTEGUMENTA-LIKE 7;

1076 FAD3, fatty acid desaturase 3; FAE1, fatty acid elongase 1; GPAT1, glycerol-3-phosphate
1077 acyltransferase 1; KASI, 3-ketoyacyl-acyl carrier protein synthase I; LACS9, long-chain acyl-
1078 CoA synthetase 9.

Figure1

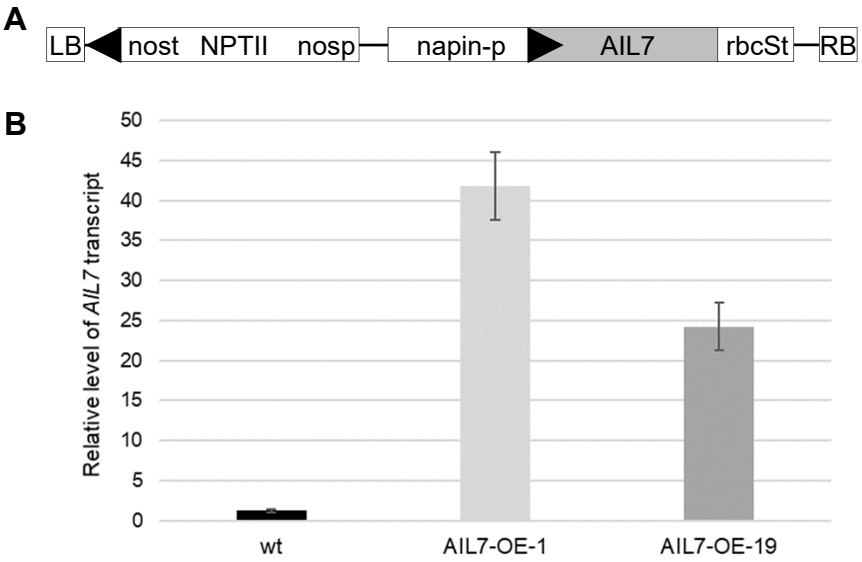

Figure 1. Generation of seed-specific *AIL7* over-expression Arabidopsis lines. (A) Schematic representation (not to scale) of the AIL7-OE construct used in this study. Arrows indicate the direction of transcription. **(B)** Quantitative real-time RT-PCR analysis of *AIL7* expression in T_2 siliques (14 DAF) from AIL7-OE-1, AIL7-OE-19, and wild-type lines. Blocks denote mean *AIL7* transcript levels from three biological replicates relative to the internal control, *PP2AA3.* Three technical replicates were carried out for qRT-PCR assays and bars indicate standard errors. AIL7, Arabidopsis *AINTEGUMENTA 7* coding sequence; LB, left T-DNA border; napin-p, *Brassica napus NAPIN* promoter; nosp, *NOPALINE SYNTHASE* promoter; nost, *NOPALINE SYNTHASE* transcriptional terminator; NPTII, *NEOMYCIN PHOSPHOTRANSFERASE II*; RB, right T-DNA border; rbcSt, *Pisum sativum RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE* transcriptional terminator; wt, wild-type.

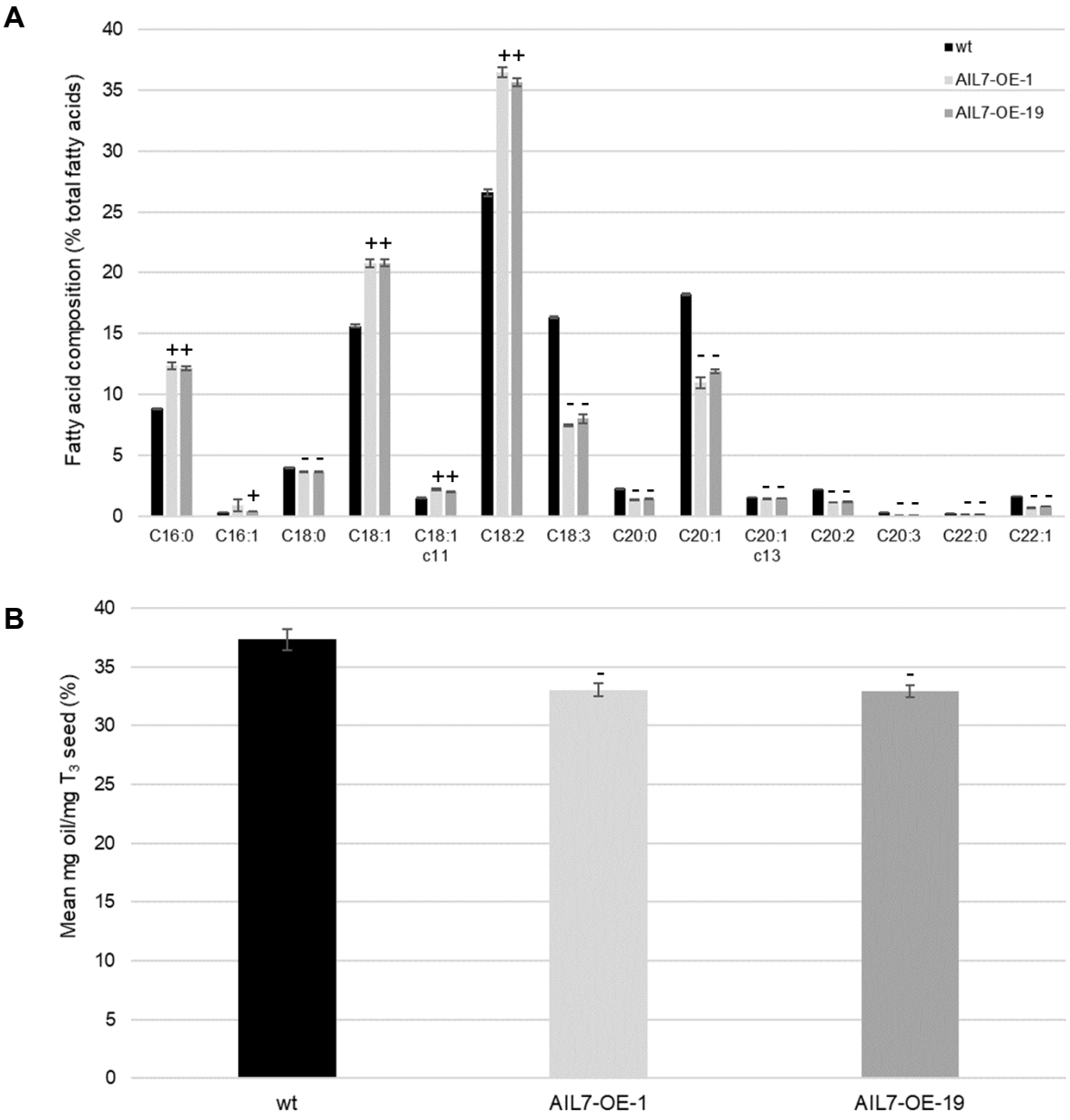

Figure 2. Fatty acid composition and oil content of homozygous AIL7-OE and wild-type T_3 seeds. (A) Mean fatty acid composition of oil from homozygous T_3 AIL7-OE seeds. Blocks indicate mean values of wild-type (*n*=6) and two independent AIL7-OE lines (AIL7-OE-1 [*n*=6] and AIL7-OE-19 [*n*=5]). **(B)** Mean seed oil content of T_3 seeds. Blocks represent mean values from wild-type (*n*=6) and two independent homozygous AIL7-OE lines (AIL7-OE-1 [*n*=6] and AIL7-OE-19 [*n*=5]). Two technical replicates were carried out for each line analyzed. Bars denote standard errors. Very significant increases and decreases compared to wild-type (as measured by 2-tailed Student's t-test assuming unequal variance) are indicated by + and - ($P \leq 0.01$). wt, wild-type.

Figure3

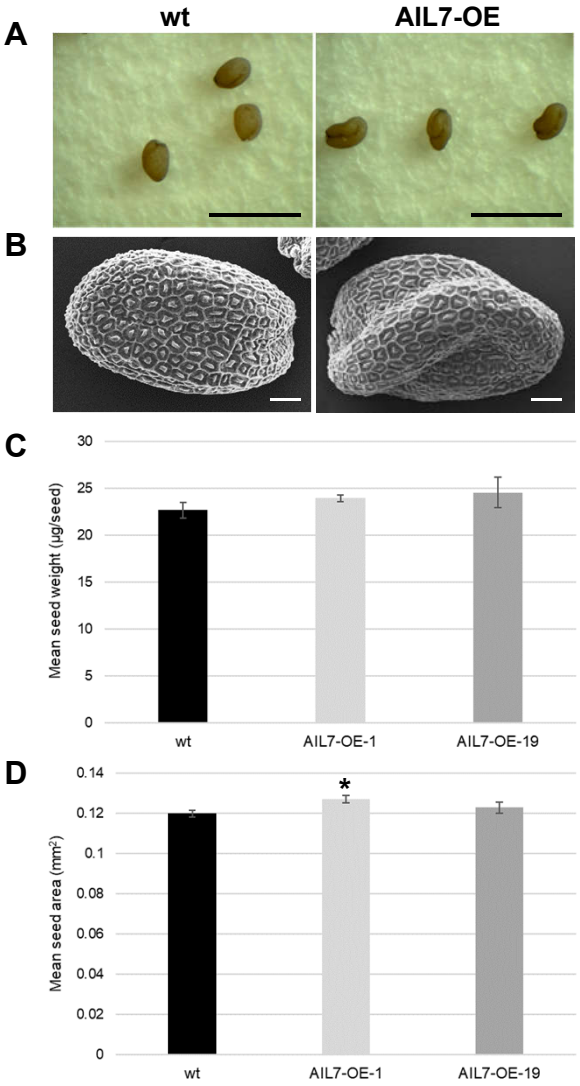

Figure 3. Seed morphology of AIL7-OE lines. (A) Representative light microscopic images of wild-type and homozygous T_3 AIL7-OE mature seeds. Scale bars = 1 mm. **(B)** Representative SEM images of wild-type and homozygous T_3 AIL7-OE seeds. Scale bars = 100 µm. **(C)** Mean mature seed weights from wild-type (from 6 individual plants), as well as homozygous T_3 AIL7-OE-1 (from 6 individual plants) and AIL7-OE-19 (from 4 individual plants) lines. Two to three replicate seed batches were used for each measurement. (D) Mean mature seed areas of wild-type (n=61 from 3 individual plants), as well as homozygous T_3 AIL7-OE-1 (n=88 from 3 individual plants) and AIL7-OE-19 (n=53 from 3 individual plants) lines. Asterisks denote significant differences from wild-type as determined by 2-tailed Student's t-tests assuming unequal variance, $P \leq 0.01$.

Figure4

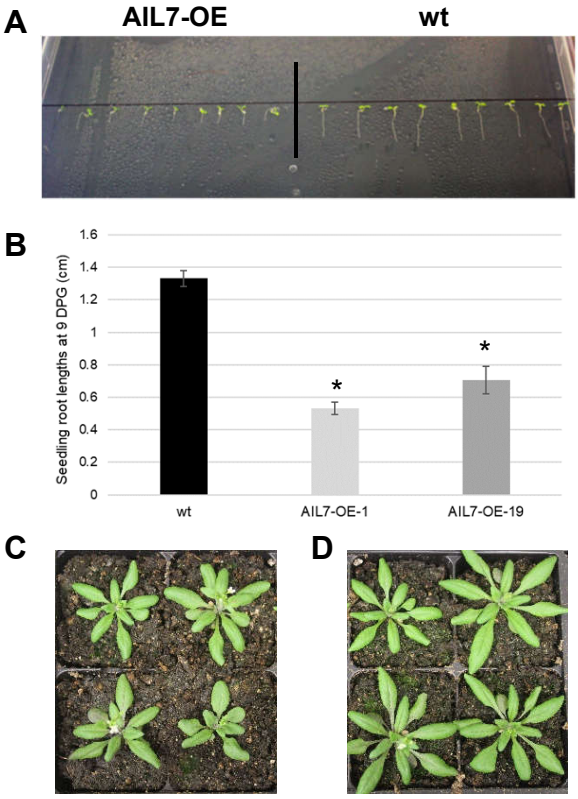

Figure 4. Growth of homozygous T_3 AIL7-OE and wild-type seedlings. (A) Seedlings grown vertically on solid medium were photographed 9 days post-germination and are representative of two independent experiments. AIL7-OE seedlings are shown to the left of the black vertical line, while wild-type seedlings are present to the right**. (B)** Root lengths of wild-type (*n*=35), AIL7-OE-1 (*n*=20), and AIL7-OE-19 (*n*=16) seedlings grown vertically on solid medium 9 days post-germination. Blocks denote mean values and bars indicate standard errors. Asterisks denote means that are significantly different from wild-type as determined by 2-sided Student's t-tests assuming unequal variance (P≤0.01). **(C and D)** Representative soil-grown AIL7-OE **(C)** and wild-type **(D)** lines 24 days post-germination. DPG, days post-germination; wt, wild type.

Figure 5

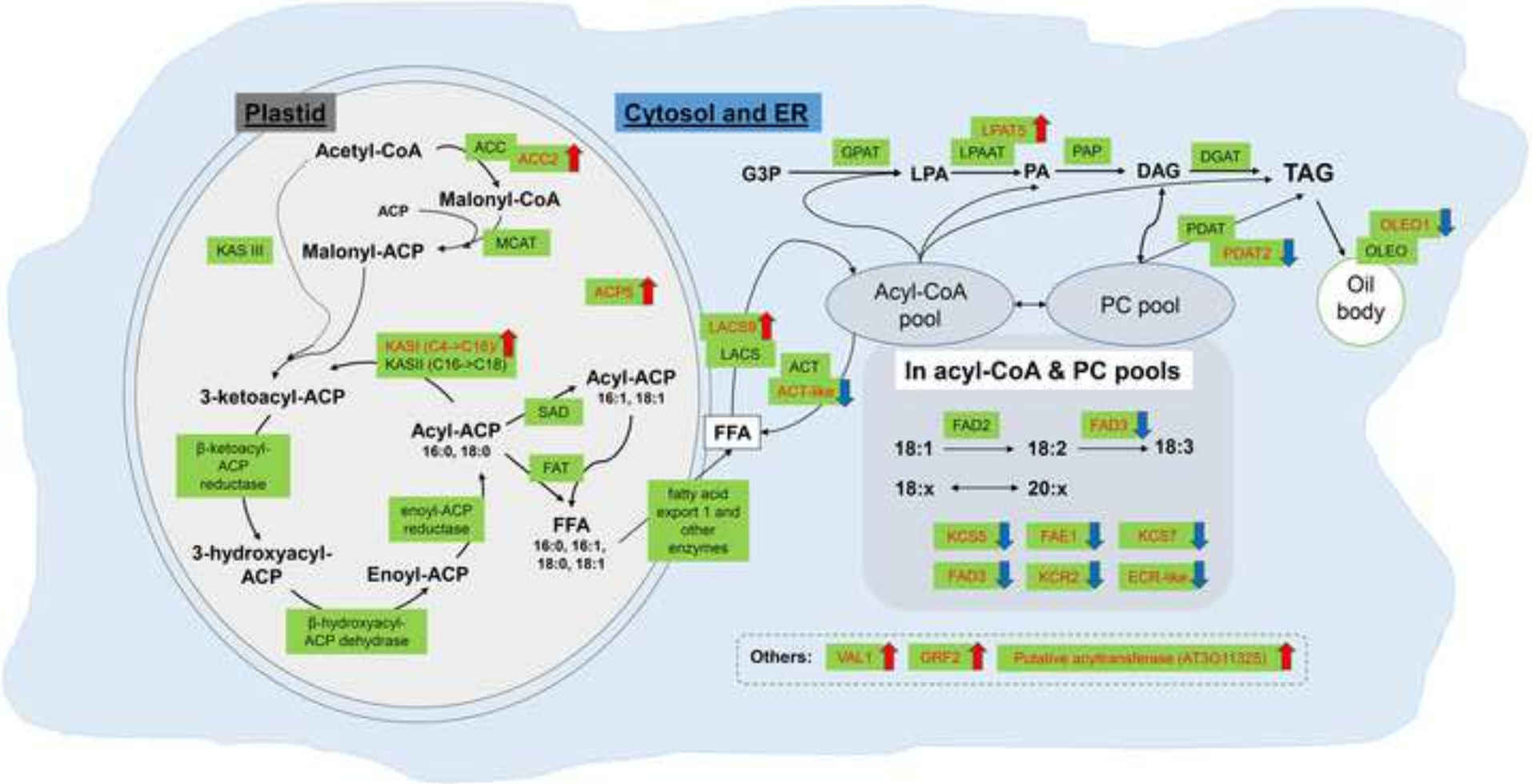
Plastid
Cytosol and ER
Acetyl-CoA
ACC
ACC2
Malonyl-CoA
ACP
KAS III
Malonyl-ACP
MCAT
ACP5
KASI (C4->C16)
KASII (C16->C18)
3-ketoacyl-ACP
β-ketoacyl-ACP reductase
3-hydroxyacyl-ACP
β-hydroxyacyl-ACP dehydrase
Enoyl-ACP
enoyl-ACP reductase
Acyl-ACP
16:0, 18:0
SAD
Acyl-ACP
16:1, 18:1
FAT
FFA
16:0, 16:1, 18:0, 18:1
fatty acid export 1 and other enzymes
FFA
LACS9
LACS
ACT
ACT-like
G3P
GPAT
LPA
LPAAT
LPAT5
PA
PAP
DAG
DGAT
TAG
PDAT
PDAT2
OLEO1
OLEO
Oil body
Acyl-CoA pool
PC pool
In acyl-CoA & PC pools
FAD2
FAD3
18:1
18:2
18:3
18:x
20:x
KCS5
FAE1
KCS7
FAD3
KCR2
ECR-like
Others:
VAL1
GRF2
Putative acyltransferase (AT3G11325)

Figure 6 qPCR validation

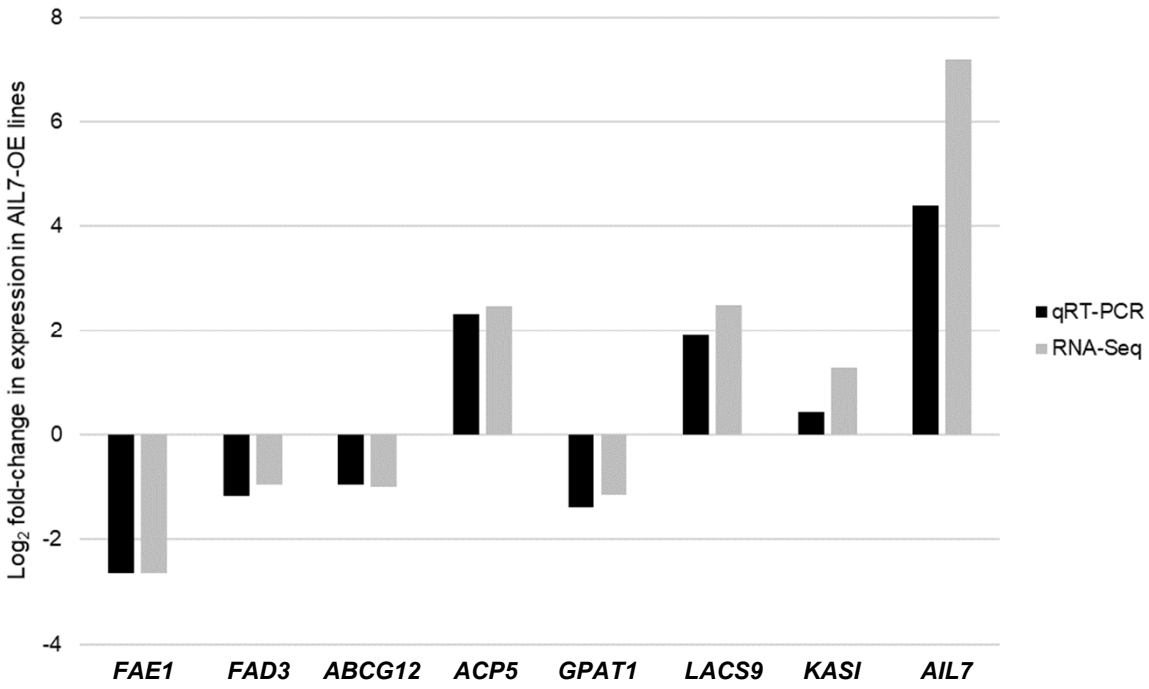

Figure 6. **qRT-PCR validation of RNA-Seq results.** qRT-PCR was carried out on RNA derived from the developing siliques (14 DAF) of AIL7-OE (4-5 biological replicates) and wild-type (3 biological replicates) lines, using the internal control, *PP2AA3,* as a reference. Three technical replicates were carried out in each case. Blocks represent $\log_2$ fold-changes in expression derived from qRT-PCR and RNA-Seq data, respectively. ABCG12, ABC-2 type transporter family protein; ACP5, acyl-carrier protein 5; AIL7, AINTEGUMENTA-LIKE 7; FAD3, fatty acid desaturase 3; FAE1, fatty acid elongase 1; GPAT1, glycerol-3-phosphate acyltransferase 1; KASI, 3-ketoyacyl-acyl carrier protein synthase I; LACS9, long-chain acyl-CoA synthetase 9.