New Phytologist

Research

The first intron and promoter of *Arabidopsis DIACYLGLYCEROL ACYLTRANSFERASE 1* exert synergistic effects on pollen and embryo lipid accumulation

Sean T. McGuire[1]*, Jay Shockey[2]* and Philip D. Bates[1]

[1]Institute of Biological Chemistry, Washington State University, Pullman, WA 99164, USA; [2]United States Department of Agriculture, Agricultural Research Service, Southern Regional Research Service, 1100 Allen Toussaint Blvd, New Orleans, LA 70124, USA

Author for correspondence:
Philip D. Bates
Email: phil_bates@wsu.edu

Received: *24 August 2024*
Accepted: *17 October 2024*

New Phytologist (2025) **245:** 263–281
doi: 10.1111/nph.20244

Key words: *Arabidopsis thaliana*, DGAT1, embryo, first intron, pollen, regulatory elements, transcription factor, triacylglycerol.

Summary

- Accumulation of triacylglycerols (TAGs) is crucial during various stages of plant development. In *Arabidopsis*, two enzymes share overlapping functions to produce TAGs, namely acyl-CoA:diacylglycerol acyltransferase 1 (DGAT1) and phospholipid:diacylglycerol acyltransferase 1 (PDAT1). Loss of function of both genes in a *dgat1-1/pdat1-2* double mutant is gametophyte lethal. However, the key regulatory elements controlling tissue-specific expression of either gene has not yet been identified.
- We transformed a *dgat1-1/dgat1-1//PDAT1/pdat1-2* parent with transgenic constructs containing the *Arabidopsis DGAT1* promoter fused to the *AtDGAT1* open reading frame either with or without the first intron.
- Triple homozygous plants were obtained, however, in the absence of the *DGAT1* first intron anthers fail to fill with pollen, seed yield is *c*. 10% of wild-type, seed oil content remains reduced (similar to *dgat1-1/dgat1-1*), and non-Mendelian segregation of the *PDAT1/pdat1-2* locus occurs. Whereas plants expressing the *AtDGAT1pro:AtDGAT1* transgene containing the first intron mostly recover phenotypes to wild-type.
- This study establishes that a combination of the promoter and first intron of *AtDGAT1* provides the proper context for temporal and tissue-specific expression of *AtDGAT1* in pollen. Furthermore, we discuss possible mechanisms of intron mediated regulation and how regulatory elements can be used as genetic tools to functionally replace TAG biosynthetic enzymes in *Arabidopsis*.

Introduction

The economic importance of vegetable oils (primarily triacylglycerols, or TAGs) as an industrial and nutritional resource is well known. Decades of research have investigated the various aspects of how plant cells exert qualitative and quantitative control over TAG metabolism. Families of genes with known or suspected roles in lipid metabolism have been thoroughly cataloged (Li-Beisson *et al.*, 2013; Ischebeck, 2016; He *et al.*, 2020); however, the presence, production, and physiological roles of neutral lipids in non-seed tissues remains poorly understood. Considering the desire to control TAG accumulation and fatty acid composition for food and industrial uses, identification of the mechanisms of TAG accumulation in non-storage tissues represents a hurdle in lipid metabolic engineering (Vanhercke *et al.*, 2019). To improve oil quality and quantity, we require a robust understanding of the mechanisms that control the expression of genes involved in neutral lipid accumulation at essential stages of plant development. Two genes in *Arabidopsis*, acyl-CoA:diacylglycerol acyltransferase 1 (AtDGAT1, At2g19450) and phospholipid:diacylglycerol acyltransferase 1 (AtPDAT1, At5g13640), are responsible for TAG biosynthesis. At least one of these enzyme activities are necessary for pollen development and embryo viability, as evidenced by the pollen abortion and embryo lethality of a *dgat1-1/dgat1-1//pdat1-2/pdat1-2* double mutant (Zhang *et al.*, 2009). It is also known that intracellular storage lipids are a necessary source of carbon skeletons and energy for pollen grain formation and germination (Piffanelli *et al.*, 1997; Ischebeck *et al.*, 2020; Wang *et al.*, 2020b). Neutral lipid accumulation occurs in both the maternal diploid sporophyte (tapetum) and the haploid gametophyte (pollen grain), serving as specific examples of cell-specific TAG accumulation (Boavida *et al.*, 2005; Hsieh & Huang, 2007; Yang & Benning, 2018). During pollen maturation in Brassicaceae the tapetum acts as a secretory cell type, providing lipid and protein precursors to the nascent pollen grain for the formation of the pollen coat (Lévesque-Lemay *et al.*, 2016; Tidy *et al.*, 2022). Other key players in biochemical regulation of both TAG accumulation and

*These authors contributed equally to this work.

degradation during pollen tube growth have also been recently identified, such as non-specific phospholipases C2/6 (NPC2/6) (Bose *et al.*, 2021), yet many others remain to be determined. Recent discoveries such as these reveal that the process of sexual reproduction in flowering plants, though complex and poorly understood, depends on the fitness of the male gametophyte and its ability to accumulate TAG during maturation (Ischebeck, 2016; Wan *et al.*, 2020). Despite the evidence that coordinated cell- and tissue-specific expression of both *DGAT1* and *PDAT1* is critical little research has been done to determine the transcriptional elements and regulators that control TAG biosynthetic gene expression during pollen development.

Oil biosynthetic enzymes from different species can have high selectivity for various diacylglycerol (DAG) and acyl-CoA molecular species, enabling transgenic approaches to produce novel seed oil fatty acid compositions (Chen *et al.*, 2022). Most research and bioengineering approaches to alter lipid metabolism have utilized transgenic systems with constitutive or seed-specific promoters (Kay *et al.*, 1987; Bhunia *et al.*, 2014; Shockey *et al.*, 2015; Correa *et al.*, 2020). These promoters can be useful in expressing lipid biosynthetic enzymes in a manner that supplements endogenous lipid metabolism (Regmi *et al.*, 2020; Hatanaka *et al.*, 2022), with or without suppression of endogenous lipid metabolic genes to create new lipid compositions. For example, RNAi suppression of *DGAT1* or *PDAT1* in multiple plant species (Zhang *et al.*, 2009; van Erp *et al.*, 2015; Alkotami *et al.*, 2021) has demonstrated that by reducing endogenous enzyme competition the phenotypic characterization of an overexpressed transgene is simplified and the final seed oil compositions can be better controlled. However, complete replacement of endogenous TAG biosynthesis in plants, such as *Arabidopsis*, requires mutation of both *DGAT1* and *PDAT1*, which unfortunately is gametophyte lethal (Zhang *et al.*, 2009). Additionally, the proper function of TAG biosynthetic genes requires precise developmental timing and tissue-specific expression in pollen and embryos, which is not provided by all seed-specific or common constitutive promoters. For example, the widely used *Cauliflower Mosaic Virus 35S* (CaMV35S) promoter drives varying levels of tissue-dependent transgene expression and is wholly absent in pollen (Wilkinson *et al.*, 1997; Kiselev *et al.*, 2021). Thus, to functionally replace TAG biosynthesis and complement the lethality of the *Arabidopsis dgat1-1/dgat1-1//pdat1-2/pdat1-2* double mutant, it is necessary to fully mimic endogenous gene expression patterns in all relevant tissues and organs, using a transgene expressed by its native promoter and other cognate regulatory elements.

The expression profile of *AtDGAT1* has been partially resolved in previous studies. *AtDGAT1* promoter:β-glucuronidase (GUS) localization studies identified putative binding sites for transcription factors and other regulatory elements with known roles in pollen and embryo development (Lu *et al.*, 2003). However, these were reporter gene experiments performed in wild-type plants, with normal levels of endogenous *DGAT1* and *PDAT1* expression, and by their nature failed to properly contextualize the required *DGAT1* expression necessary for mutant complementation. Thus, any essential promoter elements required for proper *DGAT1* expression in embryos, endosperm, pollen, meristems or other tissues remain largely uncharacterized.

In this study we present the first known characterization of the combinatorial roles that the *Arabidopsis DGAT1* first intron and promoter play in the proper temporal and tissue-specific expression of *DGAT1*. By complementation of the pollen-lethal *dgat1-1/dgat1-1//pdat1-2/pdat1-2* double mutant our results indicate that together both promoter and intronic DNA elements mediate expression of *DGAT1* to induce TAG biosynthesis during pollen and embryo development.

Materials and Methods

Plasmid construction

Binary expression plasmids containing the AtDGAT1 open reading frame (ORF) fused to the strong, seed-specific *At2S-3* promoter (Guerche *et al.*, 1990) was described previously (Regmi *et al.*, 2020). Modified ORFs containing specific introns were designed as synthetic DNA blocks and purchased commercially (Integrated DNA Technologies, Coralville, IA, USA).

AtDGAT1 promoter cloning

BLAST searches with the *AtDGAT1* coding sequence against sequences in v.11 at The Arabidopsis Information Resource (TAIR, https://www.arabidopsis.org/) identified Bacterial Artificial Chromosome (BAC) clone stock #F3J11 as containing the entirety of the *AtDGAT1* gene, including at least several kbp of sequence both upstream of the start methionine ATG codon and downstream of the terminator stop codon. Primers AtD1prom-longAscI and AtD1term-AscI (Supporting Information Table S1) were used to PCR amplify a 6292 bp product from DNA prepared from clone F3J11. This product covered the entire AtDGAT1 coding region, including 16 exons and 15 introns, 2066 bp of 5′ flanking sequence and 1206 bp of 3′ flanking sequence. The product was digested with *AscI* and cloned into the *AscI* site of the DsRed-selectable plant binary vector pB110 (Shockey *et al.*, 2015) and sequenced to confirm sequence accuracy.

Transgenic plant production

Arabidopsis thaliana (Columbia ecotype) *dgat1-1/dgat1-1//PDAT1/pdat1-2* seed was a generous gift from John Ohlrogge. The procedure for how this background was generated can be found in Zhang *et al.* (2009). The *dgat1-1/dgat1-1//PDAT1/pdat1-2* plants were transformed with foreign binary plasmids carried by *Agrobacterium tumefaciens* strain GV3101, using the floral dip method (Clough & Bent, 1998).

Plant growth conditions

Cold-stratified seeds were sown on soil in 3.5″ square pots. Plants were grown in a 20–23°C growth chamber with a 24-h light regime (*c.* 120–130 μmol photon m^{-2}). For genotyping

experiments, a surface sterilization protocol was used: 70% (v/v) ethanol wash, followed by a 10% (v/v) sodium hypochlorite and 0.1% (w/v) SDS solution with successive sterile water rinses. Seeds were pipetted on half-strength MS agar plates containing 0.8% (w/v) Phytoagar, 0.05% (w/v) MES (pH 5.7), and 1.5% (w/v) sucrose. For pollen analysis, plants were sown on soil and randomized until many plants had several flowers at stages 11–13, from which flowers were collected, dissected, and imaged. Concurrently plants were grown for yield and lipid analyses until reaching maturity from which seeds were harvested and dried on silica. Soil grown plants were imaged after *c.* 6-wk of growth with a Canon E05 Rebel T7 with a 35 mm macro lens.

Genotyping complemented lines and RT-qPCR

Genomic DNA was isolated with a simplified DNA extraction method. Briefly, a 2 mm piece of leaf tissue was collected from *c.* 14-d old plants and crushed in buffer containing 100 mM Tris–HCl pH 9.5, 500 mM KCl and 10 mM EDTA. Samples were incubated at 96°C for 10 min. Following incubation, 0.8–1 µl of the extract was used as DNA template in a 20 µl PCR reaction using the GoTaq® G2 DNA Polymerase Mastermix (Promega Corp.) containing 400 µM dNTPs and 3 mM $MgCl_2$. An additional 1 mM $MgCl_2$ was supplied to each reaction. The presence or absence of the intron region was determined by amplifying the region corresponding to the transgene with specific primers. Following gel-purification each amplicon was sequence confirmed. To effectively test the ability of our constructs to complement the lethal *dgat1-1/dgat1-1//pdat1-2/pdat1-2* double mutant, allele-specific primers were used to determine the presence of the *pdat1-2* T-DNA in *tDGAT1/tDGAT1//dgat1-1/dgat1-1//pdat1-2/pdat1-2* plants (where *tDGAT1* represents the transgene-specific *DGAT1*). Primers used in this study are listed in Table S1. Chi-squared (χ^2) tests were performed to determine significance between observed and expected allelic ratios.

For quantitative reverse transcription polymerase chain reaction, 10 similarly staged inflorescences from three plants (30 flowers/replicate) were collected in l-N_2 and RNA was isolated with a Quick-RNA™ Plant Miniprep Kit (Zymo Research, Irvine, CA, USA). RNA quality was assured by a bioanalyzer and 2 µg of RNA was converted to cDNA using iScript™ cDNA Synthesis Kit (Bio-Rad). A 1 : 5 cDNA dilution was used as template. The $2^{-\Delta\Delta C_t}$ method was used to determine relative gene expression to the *ASAR1* gene and normalized to wild-type expression.

Pollen staining, tube growth and light microscopy

Flowers (stages 11–13) were dissected under a Leica MZ10F microscope (Leica Microsystems Inc., Wetzlar, Germany). To expose anthers and stigma a few petals were removed, and images were collected with an Axiocam 105 color camera (Zeiss Group, Oberkochen, Germany). For testing presence of pollen and pollen viability, flowers were fixed in Carnoy's solution and placed in 70% ethanol (v/v) until dissection. Anthers were prepared with detached anthers and placed into a 20 µl drop of phenol-free modified Alexander's stain (Peterson *et al.*, 2010). Slides were observed with a DM2000 (Leica) with micrograph collection with a DFC295 (Leica) camera.

For pollen tube assays we used a modified media described previously (Dickinson *et al.*, 2018). At least 12 individual prehydrated flowers (3 flowers from 4 individual plants) were brushed onto 1 cm^2 of media. Slides were prepared after 5- and 24-h incubation and imaged as above.

Lipid and fatty acid analysis

Dried seed from homozygous *tDGAT1/tDGAT1//dgat1/dgat1//pdat1-2/pdat1-2* lines was weighed (*c.* 2–3 mg). For transmethylation, seeds were incubated in 1 ml of 5% (v/v) methanolic-H_2SO_4 with 300 µl of toluene at 85°C for 1.5 h. Ten microgram of pentadecanoic acid (15:0) TAG was included as an internal standard. Fatty acid methyl esters (FAMEs) were collected and extracted with 500 µl hexanes and 1 ml 0.88% (w/v) KCl. FAMEs were analyzed by an Agilent 7890B Gas Chromatograph (GC) (Agilent Technologies Inc., Santa Clara, CA, USA) on a DB-HeavyWAX column (30 m, 0.250 mm internal diameter and 0.25 µm film thickness) equipped with a flame ionization detector (FID). The GC-FID conditions were as follows: split mode injection (1 : 10), 5 µl injection volume, injector at 250°C and FID at 255°C, with oven temperature programmed at 140°C for 1.5 min, with a ramping rate of 15°C min^{-1} until reaching 200°C, this was followed by a 6°C min^{-1} increase until reaching 260°C and holding for 4 min.

Silique fertility analysis

Siliques collected from the entire length of the primary shoot were attached with double-sided tape to a slide for analysis. Selected siliques at *c.* 12 DAF were dissected to reveal ovule remnants and aborted embryos. Greening siliques were collected and cleared in 70% (v/v) ethanol for 24 h. Next, the liquid was decanted and replaced with 95% (v/v) ethanol, 5% (v/v) acetic acid for an additional 24 h. Siliques were placed on a slide and imaged using a MZ10F (Leica) microscope.

Statistical analysis

Raw data for GC analysis was collected and analyzed with Microsoft Excel v.2311 (Microsoft, Redmond, WA, USA). Statistical tests and graphs were generated with GraphPad Prism v.10 (GraphPad Software, Boston, MA, USA).

In silico assessment

Orthologs of the *A. thaliana DGAT1* genomic sequence were deduced by Blast search using Phytozome v.13 (https://phytozome-next.jgi.doe.gov/) (Goodstein *et al.*, 2011). Sequences with the highest identity, alignment length similarity and a putative PANTHER (https://www.pantherdb.org/) definition of diacylglycerol acyltransferase 1, were chosen for alignment. Accessions unavailable in Phytozome were found on National Centre for Biotechnology Information (NCBI) via nucleotide Blast search

(https://blast.ncbi.nlm.nih.gov/Blast.cgi). Species ID or Accession: *Ricinus communis*, 29912.t000099 (Chan *et al.*, 2010); *Jatropha curcas*, JQ319812.1; *Glycine max*, Glyma.13G106100 (Valliyodan *et al.*, 2019); *Oryza sativa*, LOC_Os06g36800 (Ouyang *et al.*, 2006); *Zea mays*, ZmPHB47.06G136000 (*Z. mays* PHB47 v.1.0, DOE-JGI, http://phytozome.jgi.doe.gov/); *Helianthus annuus*, HanXRQChr10g0319461 (Badouin *et al.*, 2017); *Thlaspi arvense*, Thlar.0020s0160 (Brassicales Map Alignment Project, DOE-JGI, http://bmap.jgi.doe.gov/); *Brassica napus*, JN224476.1 (Greer *et al.*, 2015); *A. thaliana*, AT2G19450; *Arabidopsis lyrata*, AL3G48800; *Arabidopsis helleri*, Ah3G41810; *Brassica rapa*, Brara.I01120; *Physaria fendleri*, OVBV01015132.1; *Camelina sativa* Svalof v.1.1 (DOE-JGI, https://phytozome.jgi.doe.gov/bap).

Multiple-sequence alignments of genomic DNA were generated by the ClustalW (v.2.0) algorithm (Larkin *et al.*, 2007) using default parameters. The Maximum-likelihood phylogenetic tree was generated with Molecular Evolutionary Genetics Analysis (Mega11) (Tamura *et al.*, 2021) with default settings and a partial deletion cut-off of 95% with 1500 bootstrap replications. The first intron of each *DGAT1* was identified by conserved 3′ GT and 5′ AG splice sites and transcription factor binding sites were found using PlantRegMap (http://plantregmap.gao-lab.org/index.php) with default constraints. A protein multiple-sequence alignment of DGAT1s and predicted or known genomic architectures was mapped with GenePainter (https://genepainter.motorprotein.de/) using default constraints.

Results

Development of a *dgat1-1/dgat1-1//PDAT1/pdat1-2* complementation-segregation strategy to define essential regulatory elements

Actualization of the goal of engineering *Arabidopsis* and other oilseed crops to produce unique TAG molecules via transgenic overexpression of alternative DGAT1s with various substrate selectivities – while maintaining both oil yields and tissue-specific expression patterns of endogenous DGAT1 activity – will require a better understanding of the transcriptional regulation of *AtDGAT1*. Therefore, we investigated a synthetic approach to rescue the *dgat1-1/dgat1-1//pdat1-2pdat1-2* double mutant through the transgenic complementation of a heterozygous *dgat1-1/dgat1-1//PDAT1/pdat1-2* parental line (Zhang *et al.*, 2009). Due to the requirement of TAG biosynthesis in pollen, viable double mutant pollen cannot be obtained from a heterozygous parent, leading to non-Mendelian segregation of the *PDAT1/pdat1-2* alleles, with an anticipated 1 : 1 : 0 segregation ratio. We considered the possibility that expression of a transgenic *AtDGAT1* (*tDGAT1*), if paired with the necessary regulatory DNA elements, should complement the *dgat1-1/dgat1-1* lesion and its accompanying phenotypes, allowing for Mendelian segregation of the heterozygous *PDAT1/pdat1-2* locus. Since gametophytes are haploid and the *dgat1-1/pdat1-2* genotype is pollen lethal, there are three hypothetical genotypes for the resulting male and female gametes produced from primary T_1 transgenic plants, either *tdgat1/dgat1-1/PDAT1*, *tDGAT1/dgat1-1/PDAT1* or *tDGAT1/dgat1-1/pdat1-2*. Ultimately, after fertilization, if the transgenic constructs complemented the lethal phenotype, it would be possible to obtain T_2 plants homozygous for *tDGAT1* and *dgat1-1/dgat1-1* with either *PDAT1/PDAT1*, *PDAT1/pdat1-2*, or *pdat1-2/pdat1-2* genotypes.

The first attempt to identify the proper *AtDGAT1* promoter via the *dgat1-1/dgat1-1//PDAT1/pdat1-2* complementation strategy used a transgenic construct containing 1822 bp of the 5′ UTR and promoter, fused in-frame with the *AtDGAT1* open reading frame (ORF) and flanked on its 3′ end with a 444 bp soybean *glycinin G1* subunit transcriptional terminator (Sims & Goldberg, 1989) (plasmid E546) (Fig. S1). Transgenic T_1 seeds containing this complementation construct were selected by the fluorescence of the DsRed-selectable marker driven by the cassava vein mosaic virus (CVMV) promoter (Verdaguer *et al.*, 1996; Stuitje *et al.*, 2003). Here, we used a PCR-based genotyping assay specific to both wild-type *PDAT1* (723 bp) and mutant *pdat1-2* (908 bp) alleles. For selection of the transgenic lines a rigorous screening procedure was necessary because the lethality exhibited by the double homozygous genotype (*dgat1-1/dgat1-1//pdat1-2/pdat1-2*) was partially indicated when screening for *dgat1-1/dgat1-1//PDAT1/pdat1-2* plants as well. Three separate sets of progeny plants produced from parents with this genotype yielded 27%, 22% and 27% *PDAT1/pdat1-2* alleles, respectively. This transmission frequency was significantly less than the 50% expected from a normal, unbiased 1 : 2 : 1 segregation pattern. Additionally, it also varied significantly from the 66.6% transmission rate of heterozygous alleles expected from the 2 : 1 *PDAT1/pdat1-2 : PDAT1/PDAT1* ratio anticipated by a homozygous lethal/heterozygous normal trait, and from the 47% heterozygous allelic transmission frequency initially reported by Zhang *et al.* (2009).

The *tDGAT1*-transformed T-DNA insertion lines chosen contained single transgenic loci, as determined by 3 : 1 red : brown seed segregation ratios. These were propagated to homozygosity as indicated by the presence of uniformly red seeds. As indicated above, due to the abnormal segregation of the *PDAT1* locus < 24% of the individual transgenic lines would potentially be *PDAT1/pdat1-2* heterozygotes, indicating that the *dgat1-1/pdat1-2* eggs may be less amenable to transformation. Therefore, to unambiguously determine the *PDAT1* locus zygosity in each plant line – even from a single transgenic event – a substantially higher than normal number of individual plants needed to be screened in each generation to identify both single transgenic loci and *PDAT1/pdat1-2* heterozygotes for further analysis. Once identified, T_3 transgenic lines homozygous for *tDGAT1/tDGAT1* and the *dgat1-1/dgat1-1* mutation, but heterozygous for the *PDAT1/pdat1-2* mutation (*tDGAT1/tDGAT1//dgat1-1/dgat1-1//PDAT1/pdat1-2*) were utilized for segregation analysis of the *PDAT1/pdat1-2* locus. Of four independent transgenic lines evaluated, only two produced homozygous *dgat1-1/dgat1-1//pdat1-2-/pdat1-2* double mutants at *c.* 6% and *c.* 15% of the progeny. However, the complemented *dgat1-1/dgat1-1//pdat1-2/pdat1-2* double mutants from both E546 *tDGAT1/tDGAT1//dgat1-1/-dgat1-1//pdat1-2/pdat1-2* lines were sterile and failed to produce

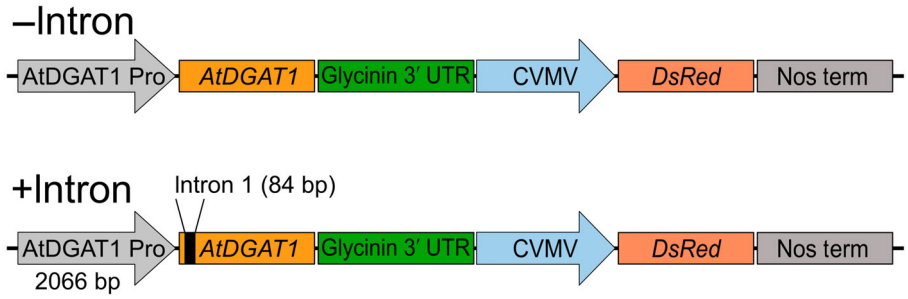

Fig. 1 Transgenic construct design. The *AtDGAT1* gene in the absence (−intron) or presence (+intron), constructs E826 and E827, respectively. Both constructs utilize the 2066 bp 5′ UTR/promoter region of *AtDGAT1* and terminated by the soybean *glycinin G1* subunit transcriptional terminator. For transgenic screening and seed selection the selectable *DsRed* marker was driven by the Cassava vein mosaic virus (CVMV) promoter and terminated with the Nos terminator.

siliques or fertile progeny, indicating that the E546 construct did not fully complement the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* lethality (Fig. S2).

In the second iteration, two new complementation constructs (plasmids E700 and E702) were produced and tested for their ability to rescue the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* lethality and produce viable seeds. Both constructs contained 2066 bp of *AtDGAT1* 5′ UTR/promoter region and the full genomic DNA complement of 16 exons and 15 introns, and the terminator regions of either *AtDGAT1* or soybean *glycinin G1* terminators, respectively (Fig. S1). The different terminators were included to compare the contributions of native vs heterologous downstream elements to gene expression. Similar selection and genotyping procedures confirmed the presence of *tDGAT1/tDGAT1//dgat1-1/dgat1-1//pdat1-2/pdat1-2* triple homozygotes which were grown to maturity to verify fertility. Unlike those transgenic lines expressing the *AtDGAT1* ORF (E546), each line containing the full *AtDGAT1* protein-coding region, the intervening introns and the longer promoter produced siliques and normal seed (Fig. S2). These data indicated that a complementation-segregation strategy could be useful in revealing the essential regulatory elements controlling *DGAT1* expression necessary for viable pollen and embryo development in a DGAT1/PDAT1 dysfunctional mutant.

The successful rescue of triple homozygous progeny using the full-length *AtDGAT1* genomic DNA constructs (E700, E702) clearly demonstrated the presence of a key regulatory element that was absent in the E546 construct, but it was unclear whether this element was present in the promoter or introns of the gene. We investigated whether the extended promoter or the first intron was responsible for this attenuation of *AtDGAT1* expression. A previous *AtDGAT1promoter*:GUS analysis (Lu *et al.*, 2003) did not identify any well-known transcription factor binding sites in the region between −1822 and −2066. However, it is known that first introns can play important transcriptional roles for many plant (Callis *et al.*, 1987; Rose *et al.*, 2008; Back & Walther, 2021; Cao *et al.*, 2023) and nonplant genes (Watanabe *et al.*, 2002; Fu *et al.*, 2006). Hence, we developed two constructs to focus on the effects of the first intron on the transgenic expression of a new *tDGAT1* to complement the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* background. We transformed the *dgat1-1/dgat1-1/PDAT1/pdat1-2* parental line with *tDGAT1* constructs containing the 2066 bp promoter fused to the *AtDGAT1* ORF, in the absence (−intron, E826) or presence (+intron, E827) of the first intron of *Arabidopsis DGAT1* (Fig. 1) and utilized the extensive screening/propagation approach outlined above to obtain multiple single insertion *tDGAT1/tDGAT1//dgat1-1/dgat1-1//PDAT1/pdat1-2* lines for segregation analysis.

In our initial segregant screens we analyzed the bulk T_2 seed and compared fatty acid composition between nontransgenic (brown) and transgenic seed (red) to evaluate the ability of the *tDGAT1* to rescue the seed fatty acid phenotype incurred by the *dgat1-1/dgat1-1* mutation. Both constructs indicated the ability to rescue the *dgat1-1/dgat1-1* seed fatty acid phenotype (Fig. S3). We further screened multiple independent insertion lines for complementation of the non-Mendelian segregation of the *PDAT1/pdat1-2* locus. At the T_2 generation the *tDGAT1* insertion locus may still be segregating, but of seedlings that contained DsRed fluorescence we observed transmission differences of the *PDAT1/pdat1-2* locus between the −intron and +intron complements with the +intron lines resembling Mendelian inheritance (Table S2). During subsequent screening of true homozygotes (T_3 generation) for the *tDGAT1*, we successfully obtained double mutants of *dgat1-1/dgat1-1//pdat1-2/pdat1-2* in both −intron and +intron complemented lines (Fig. 2a). However, the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* genotype of the −intron lines displayed severe developmental defects, whereby 90% of siliques were shorter than 10 mm. From the siliques ≥ 10 mm the average total number of developing embryos was significantly lower than control lines and *c.* 12% of the embryos therein were aborted. Conversely, +intron transgene expression restored morphological characteristics of silique length, seed number and reduced embryo abortion count to or near wild-type (Figs 2b–h, S4). Interestingly, mature embryos excised from seeds and whole mature seed from both −intron or +intron lines appeared to have normal morphologies and the mature seed could readily germinate on sucrose-supplemented media (Fig. S5a–l). Here, we demonstrate that *dgat1-1/pdat1-2* seed can be obtained by complementation with each construct, indicating that the extended

promoter in E826 *tDGAT1* lines (Fig. 1) provides sufficient expression to at least partially rescue gametophytic lethality. Nevertheless, due to the differences in fecundity between −intron and +intron lines we posited whether the first intron is necessary for sufficient wild-type levels of *AtDGAT1* gene expression in pollen.

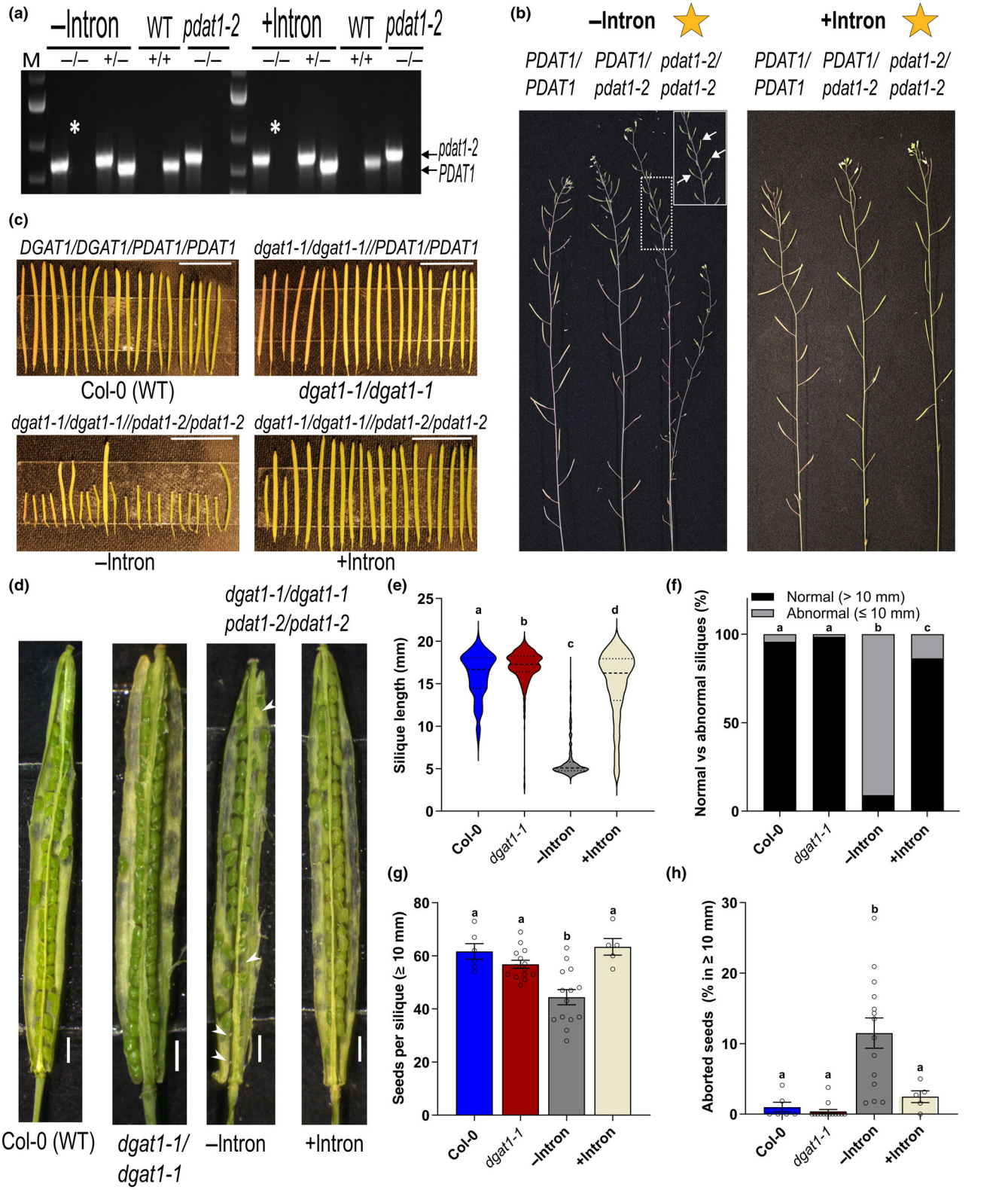

Fig. 2 Absence of the first intron of *AtDGAT1* results in silique abortion in a TAG biosynthetic deficient mutant of Arabidopsis. (a) Genotyping zygosity of the *PDAT1/pdat1-2* locus of seedlings from self-pollinated *dgat1-1/dgat1-1//PDAT1/pdat1-2* plants in the absence (−intron, line #2) or presence (+intron, line #2) of the *Arabidopsis DGAT1* first intron. Representative gel image showing allele-specific PCR products (*pdat1-2* = 908 bp; *PDAT1* = 723 bp). *PDAT1/PDAT1* (+/+), *PDAT1/pdat1-2* (+/−), and *pdat1-2/pdat1-2* (−/−). Asterisks indicate recovery of *pdat1-2/pdat1-2* homozygous mutants. For clarity each genotype was resolved alongside wild-type and *pdat1-2/pdat1-2* controls. (b) Main shoots of self-pollinated *tDGAT1/tDGAT1//dgat1-1/dgat1-1//PDAT1/pdat1-2* imaged after 6-wk of development. Each genotype for the *PDAT1/pdat1-2* locus is indicated. Inset is a close-up of siliques on the main shoot and white arrows highlight a few representative aborted siliques. Stars (orange) indicate the *pdat1-2/pdat1-2* genotype that was used for analysis in panels c–h. (c–h) Silique fertility analysis. For simplicity only line #2 was assessed for the −/+ intron genotype. (c) Siliques from the primary shoot of each indicated genotype; Bars, 10 mm. (d) Representative images of dissected siliques longer than 10 mm, to reveal embryo abortion and empty slots in the −intron line (white arrowheads); Bars, 1 mm. (e) Silique length measurements from the primary shoots of Col-0, *dgat1-1/dgat1-1*, −intron and +intron lines (n = 393, 339, 452, 300 siliques per genotype, respectively), violin plot shows the mean (dashed line) and the third and first quartiles (dotted lines). (f) Percentage of abnormal (< 10 mm) and normal (≥ 10 mm) siliques. (g) Number of seeds per silique from randomly selected siliques ≥ 10 mm (n = 6, 14, 14, 5 siliques per genotype, respectively). (h) Percentage of embryos aborted in dissected siliques. Data represent the mean values ± SEM. Different letters indicate significant differences ($P \leq 0.05$, one-way ANOVA, Dunnett's multiple comparison). Additional silique and seed photographs are in Supporting Information Figs S4 and S5. Primer sequences used in this study are listed in Table S1.

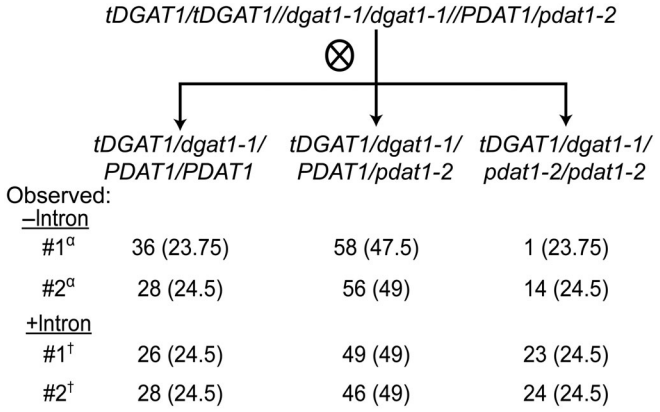

Fig. 3 Genetic screening of complemented plants for *PDAT1/pdat1-2* locus segregation in Arabidopsis. Analysis of progeny from multiple independent transgenic events of self-pollinated *dgat1-1/dgat1-1// PDAT1/pdat1-2* homozygous for each transgenic DGAT1 construct (*tDGAT1*) without (−intron) or with the (+intron). Observed and expected (in parenthesis) segregants are indicated (−intron line #1, for n = 95 plants; all other lines, n = 98 plants). χ^2 tests were performed for a 1 : 2 : 1 segregation. α indicates statistically significant differences compared to the null hypothesis with a $P \leq 0.05$. † indicates no statistical difference compared with expected segregation ratios, therefore, following Mendelian laws of inheritance.

A fully effective complementation strategy would yield progeny from a *tDGAT1/tDGAT1//dgat1-1/dgat1-1//PDAT1/pdat1-2* plant with a 1 : 2 : 1 ratio of the *PDAT1* alleles, following Mendelian laws of inheritance. However, incomplete complementation may not restore Mendelian principles and atypical segregation ratios would be anticipated. As a proxy to determine the effectiveness of each genetic construct in rescuing pollen lethality, self-pollinated plants from multiple transgenic events of *tDGAT1/tDGAT1//dgat1-1/dgat1-1//PDAT1/pdat1-2* transformants with −intron or +intron constructs were analyzed for segregation ratios at the T_3 and T_4 generations (Fig. 3; Table S2). Triple homozygous genotypes could be obtained from the −intron segregants, but at abnormal non-Mendelian ratios indicating only partial rescue of the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* pollen lethality. These results indicate that the complementation construct used in the −intron lines still lacks a key regulatory element. On the other hand, the genotyped segregants from the +intron lines displayed a 1 : 2 : 1 *PDAT1/pdat1-2* locus segregation ratio and apparent complete restoration of the pollen lethality phenotype.

Seed yield and oil amounts are impaired in *tDGAT1/dgat1-1/pdat1-2* lines lacking the first intron of *Arabidopsis* DGAT1

To better understand the role of the first intron in *AtDGAT1* expression we assessed seed yield, oil composition and total oil amounts from the top-performing transgenic event (Line #2) for the −/+intron constructs in comparison to the *dgat1-1/dgat1-1* mutant and wild-type controls. In the *Arabidopsis dgat1-1/dgat1-1* mutant, only PDAT1 functionally produces TAG (Banaś *et al.*, 2000; Zhang *et al.*, 2009; Xu *et al.*, 2012). The *dgat1-1/dgat1-1* mutant has slightly reduced seed oil yield (*c.* 70–80% of WT), and a unique fatty acid composition, with reduced monounsaturated fatty acid (18:1, 20:1) and increased polyunsaturated fatty acid (PUFA, 18:2, 18:3) content (Katavic *et al.*, 1995). We omitted comparison to the *pdat1-2/pdat1-2* single mutant since it shows no major changes in seed yield, oil amounts or composition when compared to wild-type plants (Zhang *et al.*, 2009; Xu *et al.*, 2012). During the growth of these plants, we could not detect changes during vegetative growth until bolting, at which point the −intron line began to produce many secondary branches and more inflorescences (Fig. S5), likely in an attempt to increase outcrossing capabilities or produce enough viable pollen. Seed yield analysis from triple homozygous −intron *tDGAT1/tDGAT1//dgat1-1/dgat1-1//pdat1-2/pdat1-2* plants indicated reduced fertilization rates (Fig. 4a). These seeds were also not a result of outcrossing, as the sterility and limited seed yield was maintained in their progeny. Inversely, the +intron plants restored seed yields to near wild-type levels. The reduced seed set in −intron *tDGAT1* plants is consistent with the segregation results indicating an impairment in fertilization.

Additional seed lipid analysis of both types of complemented *tDGAT1* lines showed that the −intron seeds still contained *c.* 22% less total oil than wild-type, similar to the *dgat1-1/dgat1-1* mutant, while the +intron seeds had significantly higher oil levels

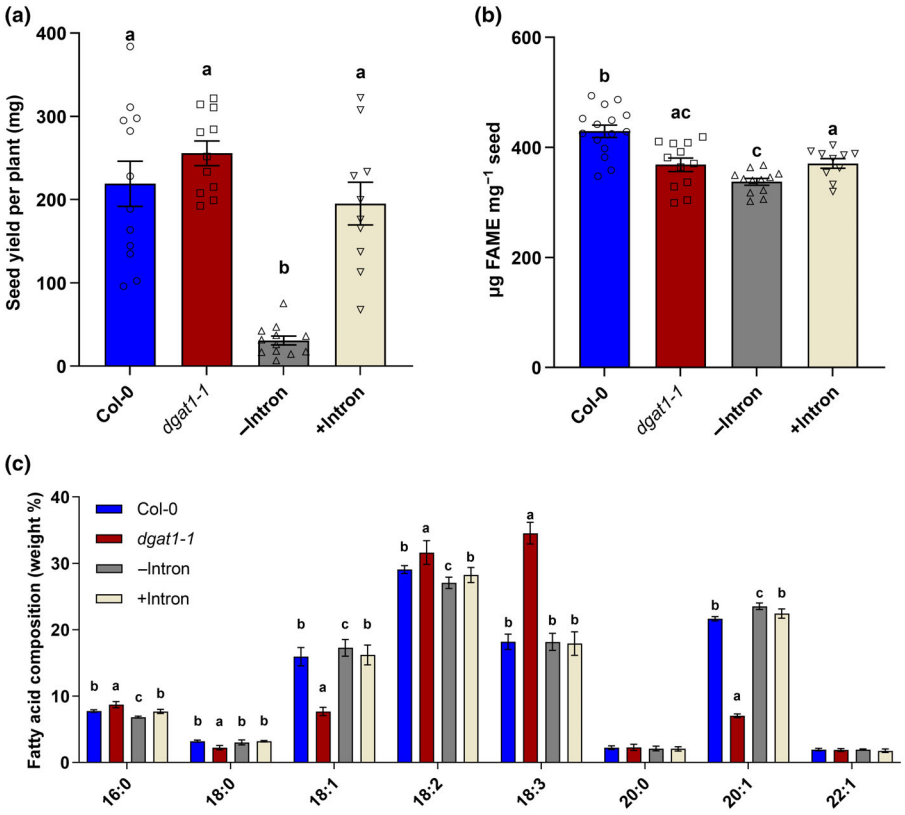

Fig. 4 Absence of the first intron hinders seed yield while oil composition remains unaffected in Arabidopsis. (a) Mean seed yield of individual lines $n \geq 10$ plants. (b) Total seed fatty acid content quantified as fatty acid methyl esters (FAMEs). Different letters in (a, b) indicate significant differences ($P \leq 0.05$, one-way ANOVA, Dunnett's multiple comparison). (c) Weight % fatty acid composition of total seed FAMEs. Values are the mean ± SD harvested from $n \geq 10$ plants. Different letters indicate statistical significance ($P \leq 0.05$, two-way ANOVA analysis with Dunnett's multiple comparison). The genotypes of the plants analyzed in all subfigures are wild-type (*DGAT1/DGAT1/PDAT1/PDAT1*), *dgat1-1/dgat1-1//PDAT1/PDAT1* and −/+ intron *dgat1-1//dgat1-1//pdat1-2/pdat1-2* (both line #2).

when compared to the −intron line (*c.* 9% more) (Fig. 4b). Although the +intron line did not fully restore total oil amount to wild-type levels, we can rule out the possibility that the *pdat1-2* mutation is the causative agent, since the *pdat1-2* knockout shows no change in TAG content or fatty acid composition (Mhaske *et al.*, 2005). However, it is possible +intron transgene expression, even when expressed behind the native promoter, still does not fully replicate the wild-type *AtDGAT1* expression profile. Furthermore, seed lipids in both types of complemented lines returned to wild-type fatty acid composition, with only minor differences between the −intron or +intron lines demonstrating functional DGAT1 in both lines (Figs 4c, S3). Taken together, these results provide additional evidence that the first intron of *AtDGAT1* is critical for normal, properly nuanced gene expression patterns in pollen and developing embryos, as evidenced by increased total seed set and seed oil levels between +intron and −intron lines.

Proper pollen development is dependent on the presence of the first *AtDGAT1* intron

To further confirm and investigate the necessity of the first intron on *DGAT1* expression during pollen development, we assessed flower and anther morphology alongside pollen viability. Stage 11–13 flowers (Alvarez-Buylla *et al.*, 2010) from wild-type, *dgat1-1/dgat1-1*, −intron (Line #2), and +intron (Line #2) plants were dissected for morphological inspection to assess early anther/-flower development. We found that the anthers in the −intron line were either incapable of dehiscing or carried no observable pollen as early as stage 12 and the stigmas had long outreaching stigmata papillae (Figs 5g, S6). Both phenotypes have been surveyed previously, including in lipid biosynthetic dysfunctional mutants, where smooth anthers are often incapable of dehiscence and fail to produce mature pollen grains (Sanders *et al.*, 1999; Fan *et al.*, 2013). Likewise, increased papillae surface area is a morphological effect due to limited viable pollen or decreased availability and reception (Katano *et al.*, 2020). Contrastingly, the +intron line was similar to both wild-type and the *dgat1-1/dgat1-1* mutant clearly demonstrating transfer of pollen from anther to stigma (Fig. 5a,d,j).

The presence of indehiscent anthers poised us to examine anthers for mature pollen grains. Fixed and cleared anthers were subjected to a modified Alexander's staining (Peterson *et al.*, 2010). Healthy viable pollen appears a deep purple while nonviable pollen fails to retain stain and either appears clear purple or green. Anthers from the −intron line appeared to have an undeveloped or pollenless phenotype as pollen grains failed to retain stain and crushed anthers yielded no stainable pollen (Fig. 5h). Instead, debris and other cell structures were found that may be attributed to deformed or immature pollen grains (Fig. S7). Conversely, the +intron line restored pollen viability to near wild-type levels (Fig. 5b,e,k). It should be noted that during multiple independent sample preparations of anthers, only a few anthers in the −intron line had definitive pollen staining with 67% of anthers containing no discernable pollen (Fig. S8). Furthermore, we tested the capacity of pollen grains to form pollen tubes. Based on the prior results it was unsurprising that any randomized imaging field contained only a few or no pollen grains/tubes in the −intron line, whereas in the +intron line pollen and pollen tubes were abundant (Figs 5c,f,i,l, S7, S9).

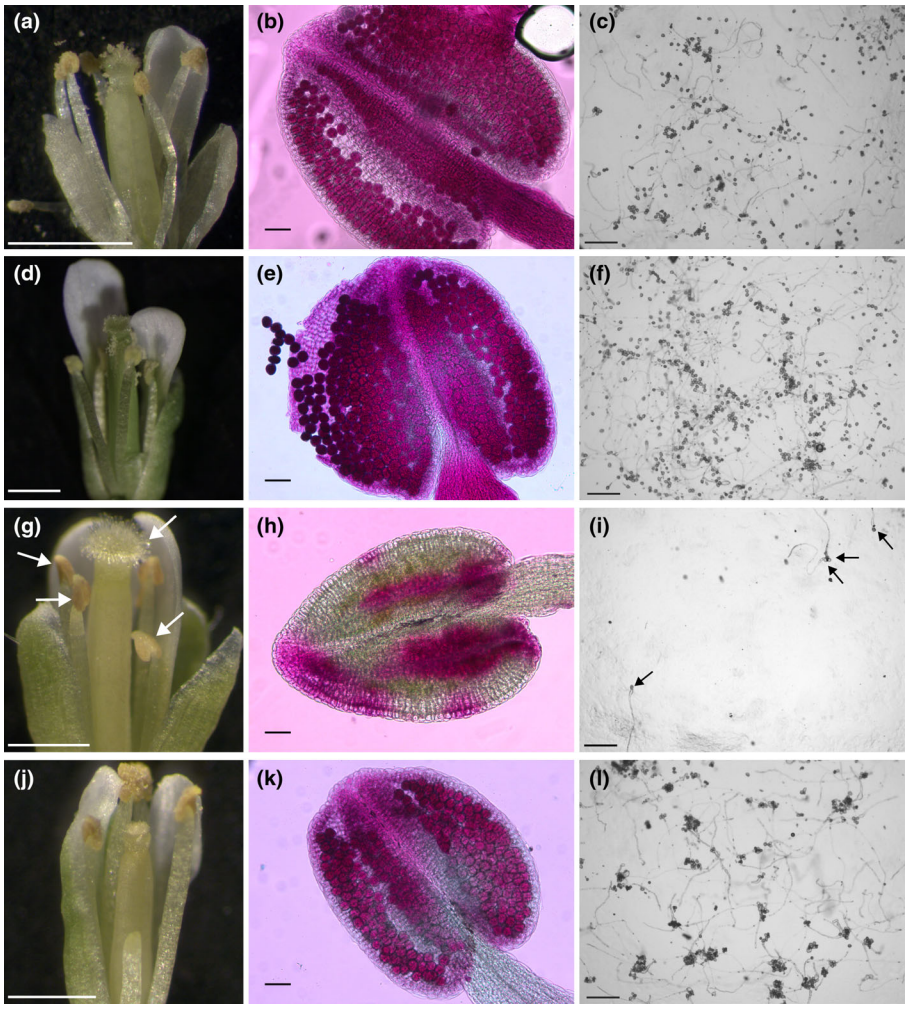

Fig. 5 Absence of the first intron in *DGAT1* impairs pollen development and ceases pollen production in Arabidopsis. The plant lines identified in Fig. 3, −intron (#2) and +intron (#2) from a *dgat1-1/dgat1-1//pdat1-2/pdat1-2* genotype were grown for analysis. (a, d, g, j) Flowers from Col-0, *dgat1-1/dgat1-1*, −intron, and +intron plants, respectively. White arrows indicate the presence of smooth anthers and outreaching stigmata papillae (g). Bars, 1 mm. (b, e, h, k) Light microscopy imaging of anthers from Col-0, *dgat1-1/dgat1-1*, −intron, and +intron flowers, respectively, pollen was stained with Alexander's solution. Viable pollen stain purple; Bars, 50 μm. (c, f, i, l) Pollen tube growth of Col-0, *dgat1-1/dgat1-1*, −intron, and +intron flowers, respectively, after 24-h incubation. Bars, 250 μm. Black arrows in (i) indicate the presence of a few pollen grains capable of proper germination and tube growth. Additional images are in Supporting Information Figs S6 (flowers), S7 (pollen), and S8 (pollen tubes).

These results support the hypothesis that the reduced fertility rates in the −intron lines are due to a failure to produce many viable pollen. This lack of pollen contributes to decreased pollen tube growth and minimal seed set; as pollen density is a determinant for pollen tube proliferation (Boavida & McCormick, 2007). Furthermore, this is consistent with the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* pollen lethality previously described and correlates positively with our observations of reduced seed yield, silique abortion, and anther morphology phenotypes (Figs 4a, 2b–h, S2, S4–S6). Although *DGAT1* expression was detected in developing whole flowers in all lines tested (Fig. S10b), the other phenotypic characteristics observed reinforce our conclusion that without the *AtDGAT1* first intron, the proper expression of *AtDGAT1* for pollen development is greatly impaired.

In silico assessment of the promoter and first intron of AtDGAT1 identified putative binding motifs for transcription factors involved in pollen development and oil biosynthesis

Unsurprisingly, *AtDGAT1* expression is likely tightly regulated, with the promoter containing interaction sequences for abscisic acid-mediated regulation by *ABI3/4/5* transcription factors (Yang *et al.*, 2011), several napin-like motifs and binding sites for ethylene response factors (Lu *et al.*, 2003). However, comprehensive dissection of the regulatory elements in the *AtDGAT1* promoter does not yet exist, and our segregation and microscopy results strongly supported that the distal 5′ region of the promoter (243 bp, from position −1822 to −2066) and the first intron (84 bp) may contain uncharacterized regulatory elements necessary for correct gene expression during pollen development. Upon closer inspection of the first intron sequence, we found multiple -NGATY- core sequences that are putative intron mediated enhancer (IME) motifs (Gallegos & Rose, 2019; Fig. S11). Next, to interrogate the DNA elements (promoter and intron) for putative transcription factor (TF) binding sites, a web-based plant TF database (http://plantregmap.gao-lab.org/index.php) (Tian *et al.*, 2020) was used. Several candidate TFs with recognized roles in lipid biosynthesis, neutral lipid accumulation and pollen development were identified based on our sequence queries in both the *DGAT1* first intron (Table 1) and the *DGAT1* Promoter from −1822 to −2066 bp (Table 2). Of note, there were no WRI1 binding motifs present in either the promoter or first intron, this coincides with the consensus

Table 1 Predicted transcription factor binding sites in the first intron of *Arabidopsis thaliana DGAT1*.

Locus ID	Gene	Hypothetical or known biological role	*P*-value	Reference
AT3G23250	MYB15	Lignin biosynthesis	1.13E-05	Kim *et al*. (2020)
AT1G79180	MYB63	Lignin biosynthesis, vascular development	1.49E-05	Zhou *et al*. (2009)
AT3G08500	MYB83	Lignin biosynthesis	2.00E-05	McCarthy *et al*. (2009)
AT4G38620	MYB4	Lignin biosynthesis, vascular development	2.16E-05	Panda *et al*. (2020)
AT5G62470	MYB96†	Oil accumulation	3.86E-05	Lee *et al*. (2018)
AT1G16490	MYB58	Lignin biosynthesis	3.96E-05	Zhou *et al*. (2009)
AT3G12820	MYB10	Iron sensing and signaling	4.52E-05	Palmer *et al*. (2013)
AT1G18570	MYB51	Glucosinolate biosynthesis	4.75E-05	Gigolashvili *et al*. (2007)
AT3G61910	NAC66	Secondary cell wall synthesis in anthers	4.78E-05	Mitsuda *et al*. (2005)
AT1G06180	MYB13	Lignin biosynthesis	4.83E-05	Taylor-Teeples *et al*. (2015)
AT5G62320	MYB99†	Stamen/pollen development	4.91E-05	Alves-Ferreira *et al*. (2007)
AT3G12720	MYB67	–	5.51E-05	–
AT5G12870	MYB46	Secondary cell wall synthesis	7.25E-05	Ko *et al*. (2009)
AT3G06490	MYB108†	Stamen/pollen development, jasmonate signaling	8.45E-05	Mandaokar & Browse (2009)
AT5G04410	NAC2/78	Flavonoid biosynthesis	9.05E-05	Morishita *et al*. (2009)

Biological roles are not exhaustive and are based on TAIR10 annotation data. – indicates no known attribution to the gene or its function. See Fig. 6 phylogenetic tree for further analysis.
†Genes selected for further analysis with known roles during TAG accumulation and stamen or gametophyte development.

findings that, although WRI1 is a master regulator of fatty acid synthesis (Lee *et al*., 2018; Kuczynski *et al*., 2022), it indirectly modulates *AtDGAT1*.

To expand our TF analysis, we selected related species as well as other distant lineages that accumulate seed oil through unique TAG biosynthetic mechanisms. These ER-localized mechanisms define the flux of acyl units into TAG and control the final TAG molecular species profile. There exist at least three defined pathways that contribute DAG for subsequent TAG biosynthesis via DGAT activity (Bates, 2016): (1) *de novo* DAG (Kennedy pathway); (2) phosphatidylcholine (PC)-derived DAG, (utilized by the *Arabidopsis* DGAT1); and (3) TAG remodeling, here TAG is first generated from either the *de novo* or PC-derived DAG pathways, and subsequent lipase activity generates an *sn*-1,2 or *sn*-2,3 TAG-derived DAG to be converted back to TAG by a DGAT (Bhandari & Bates, 2021; Parchuri *et al*., 2024). As such our query also included the *DGAT1* first intron from several species (Table S3), including, but not limited to *R. communis* (Kennedy pathway) (Bafor *et al*., 1991), *C. sativa* (PC-derived DAG pathway) (Yang *et al*., 2017), and *P. fendleri* (PC-derived DAG/TAG remodeling) (Bhandari & Bates, 2021; Parchuri *et al*., 2024). To visualize the relationship between *DGAT1*s a maximum-likelihood phylogenetic tree was generated from an alignment of 14 full-length genomic DNA sequences (Fig. 6a). Using GenePainter (https://genepainter.motorprotein.de/, Hammesfahr *et al*., 2013) we also evaluated a positional protein multiple-sequence alignment mapped to the genomic architecture of each gene (Fig. 6b) demonstrating that many *DGAT1s* have a conserved exon-intron structure with a highly variable *N*-terminal domain. While somewhat promiscuous, DGAT1s are known to have preferred substrates (Shockey *et al*., 2006; Winichayakul *et al*., 2022) and this variability in the *N*-terminus likely assists in determining acyl-CoA substrate specificities.

The first intron of *DGAT1*s from each species analyzed contain conserved binding sites for at least one TF (Fig. 6a; Table S3). Several of the identified TFs have confirmed roles in regulating TAG accumulation and plant reproduction. For example, MYB96 (At5g62470), is attributed to gatekeeping the ABA response (Lee *et al*., 2015, 2016, 2019) and directly regulates TAG accumulation during seed maturation by influencing expression of *DGAT1* and *PDAT1* (Lee *et al*., 2018). Intriguingly, the first introns contain many other putative binding motifs for TFs involved in secondary cell wall development (MYB99, At5g62320), (Alves-Ferreira *et al*., 2007) and jasmonate signaling (MYB108, At3g06490), (Mandaokar & Browse, 2009) which contributes to stamen and anther development. Additionally, existing DNA Affinity Purification Sequencing (DAP-seq) data provides *in vitro* evidence that at least a few of the TF discussed have binding sites in the *AtDGAT1* first intron sequence (O'Malley *et al*., 2016; Fig. S11). It is interesting to note that in the Brassicaceae family, only *Arabidopsis* and *Camelina* contain these binding sites (Fig. 6a). This may be due to evolution of the specific sequences between varying tribes to which each genus belongs, as only *Arabidopsis* and *Camelina* reside in the *Camelineae* tribe (Koch, 2018). Additionally, the distal promoter region (−1822 to −2066 bp, Table 2) has a putative binding site for MYB80 (At5g56110), which is known to regulate tapetal development; *myb80* mutant plants are male sterile (Phan *et al*., 2012). This region also has a site for REM34 (At4g31610) which is known to control female and male gametophyte development (Caselli *et al*., 2019). Thus, the distal promoter region of *DGAT1* may be controlled, at least in part, by MYB80 and REM34 and the lack of DGAT1 activity may contribute to the defective developmental phenotypes of their respective mutants. Together, the identification of potential binding sites for MYB96, MYB99, MYB108 and other TFs in the first intron of many *DGAT1*s, and various TFs in the distal promoter region of *Arabidopsis DGAT1* provides exciting avenues of research to explore when attempting to understand the intricate transcriptional regulatory landscape controlling

Table 2 Predicted transcription factor binding sites in the distal −1822 to −2066 bp region of the *Arabidopsis thaliana DGAT1* promoter.

Locus ID	Gene	Hypothetical or known biological role	*P*-value	Reference
AT1G69570	CDF5	Circadian clock regulation	3.20E-05	Martín *et al*. (2020)
AT4G38620	MYB4	Lignin biosynthesis, vascular development	3.54E-05	Panda *et al*. (2020)
AT3G23690	BHLH77	Secondary cell wall synthesis	3.71E-05	Taylor-Teeples *et al*. (2015)
AT5G56110	MYB80†	Anther development/tapetum regulation	3.74E-05	Verma & Burma (2017)
AT1G25340	MYB116	–	4.20E-05	–
AT3G55370	OBP3	Phytochrome regulation	4.24E-05	Ward *et al*. (2005)
AT3G46640	LUX	Circadian rhythm	4.74E-05	Zhang *et al*. (2019a)
AT4G17785	MYB39	Suberin biosynthesis	5.03E-05	Wang *et al*. (2020a)
AT1G72050	C2H2	TFIIIA for 5S rRNA production	5.34E-05	Layat *et al*. (2012)
AT5G16600	MYB43	Lignin biosynthesis	5.85E-05	Geng *et al*. (2019)
AT3G20310	ERF7	Represses GCC-box mediated transcription	6.19E-05	Nakano *et al*. (2006)
AT1G73360	HDG11	Trichome development	6.49E-05	Nakamura *et al*. (2006)
AT5G14340	MYB40	Arsenic resistance	6.64E-05	Chen *et al*. (2021)
AT3G50700	IDD2	–	6.67E-05	–
AT3G53200	MYB27	Anthocyanin repressor	6.97E-05	Albert *et al*. (2014)
AT2G43010	PIF4	Red light sensing by interacting with PhyB	7.76E-05	Sakuraba *et al*. (2014)
AT5G54230	MYB49	Interaction with ABI5/cuticle formation	8.14E-05	Zhang *et al*. (2019b)
AT4G31610	REM34†	Control of male/female gametophyte development	8.44E-05	Caselli *et al*. (2019)
AT4G00730	ANL2	Accumulation of anthocyanin	9.17E-05	Mabuchi *et al*. (2016)
AT2G20570	G2-like	Expression of the photosynthetic apparatus	9.20E-05	Susila *et al*. (2023)
AT5G08130	BIM1	Involved in brassinosteroid signaling	9.42E-05	Chandler *et al*. (2009)
AT1G09530	PIF3	Modulates photoreceptors phyA/B	9.47E-05	Leivar *et al*. (2020)
AT5G14960	E2F/DP	Cell proliferation	9.51E-05	Sozzani *et al*. (2010)
AT3G60580	C2H2†	Pollen germination and tube growth	9.57E-05	Wang *et al*. (2008)
AT5G17430	AP2	Expression in embryos	9.87E-05	Horstman *et al*. (2017)

– indicates no known attribution to the gene or its function.
†Genes with known or implicated functions in gametophytic development or processes.

TAG accumulation, particularly during tapetal and pollen development.

Discussion

The combined roles of the promoter and first intron in tissue-specific expression patterns of *AtDGAT1*

Our analyses support the hypothesis that the first intron is critical for proper timing of expression and modulation of DGAT1 enzyme abundance, specifically during pollen and embryo development. Though clearly important for normal fertility, the *AtDGAT1* intron does not act alone. Had the absence of the intron in our −intron *tDGAT1/tDGAT1//dgat1/dgat1-1//pdat1-2/pdat1-2* line resulted in cessation of seed development or massively impaired mature seed TAG levels, we would have anticipated a similar phenomenon to Zhang *et al*. (2009), where a *dgat1-1/dgat1-1 PDAT1* RNAi line had strongly reduced oil content, variable fatty acid composition and altered seed morphology. Additionally, the *dgat1-1/dgat1-1 PDAT1* RNAi plants could not produce cotyledons or readily germinate in the presence of sucrose. However, the opposite was true for our −intron triple homozygotes, that had near wild-type fatty acid compositions and similar embryo/seed morphologies, albeit with less total oil in the limited seeds that were produced (Figs 4, S5a–h). Of the few seeds produced in the −intron complemented lines they readily germinated on sucrose-containing media (Fig. S5i–l). Combined with the observation that a phenotype did not manifest until late in the plant life cycle in −intron plants (Figs 2b–h, 5, S4–S6) this may indicate that the first intron is most important for pollen development, but less so for embryo development, as it appears the main regulatory elements necessary for *DGAT1* expression during embryo maturation is found within the promoter and not the first intron.

The pollen-specific importance of the first intronic region of *AtDGAT1* was clarified by analyzing the contrasting phenotypes of the −intron and the +intron lines that were restored to wild-type, these results are consistent with prior expression data (Fig. S10a) and observations made by Zhang *et al*. (2009). Studies across a multitude of species demonstrate that complex regulatory functions are modulated by TFs binding to various important *cis*-regulatory motifs distributed throughout key metabolic genes (Peng & Weselake, 2011; Blayney *et al*., 2023). Mutation and selection of *cis*-regulatory region alleles have been critical to species development, during both naturally occurring evolutionary timeframes (Hendelman *et al*., 2021), and in targeted breeding for crop domestication and improvement (Wang *et al*., 2021). Though often masked due to gene family complexity or functional redundancy, such alleles paradoxically may be valuable because of their generally weaker phenotypic effects relative to mutations in protein-coding regions, which often result in drastic pleiotropic consequences (Rodríguez-Leal *et al*., 2017). While promoters have been more widely studied, the roles that introns play have gained additional importance in recent years. In particular, the first intron is implicated in containing many regulatory motifs that contain binding sites for TFs that can

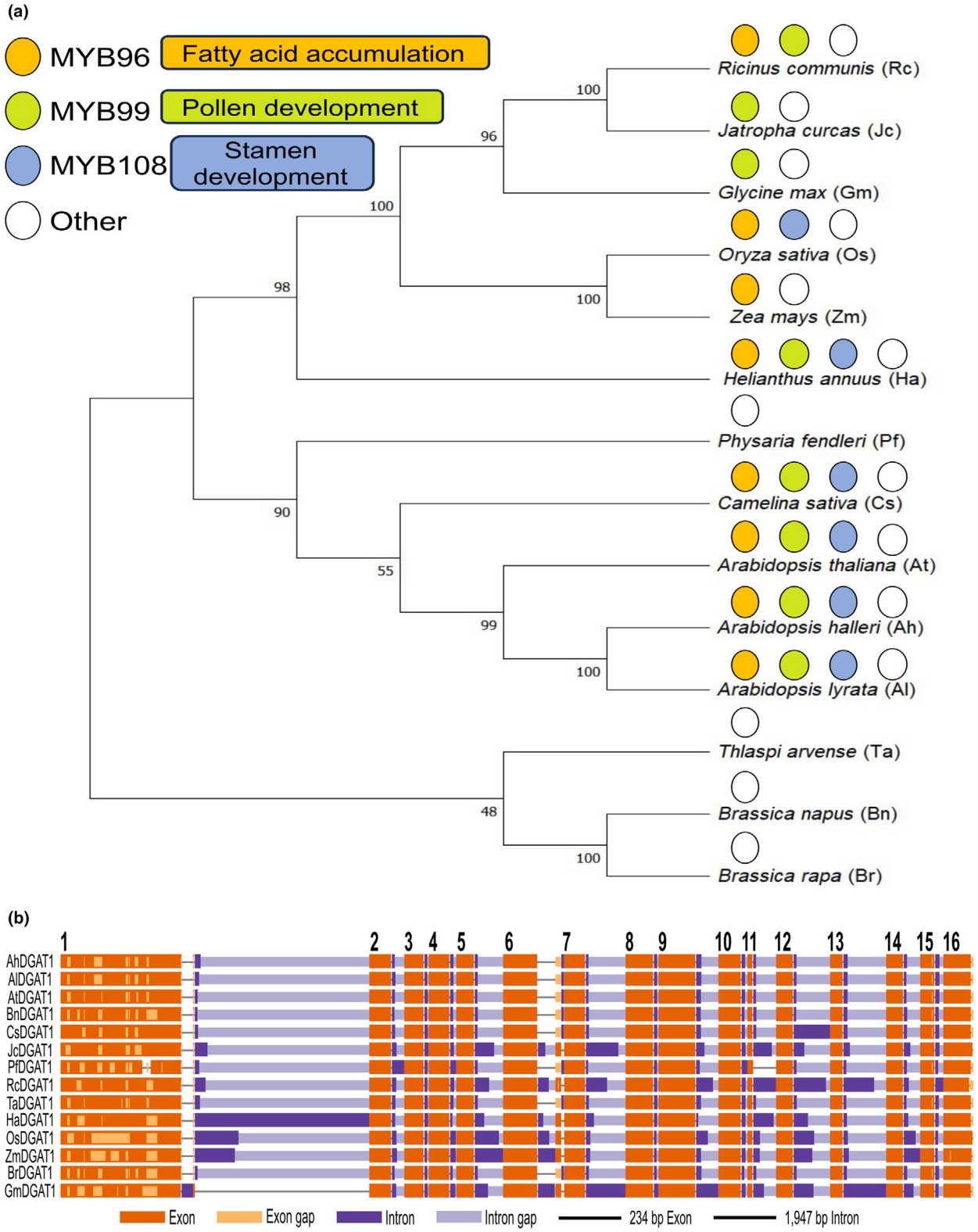

Fig. 6 *In silico* analysis reveals conservation of putative transcription factor binding motifs and genomic DNA architectures in plant *DGAT1*. (a) Maximum-likelihood phylogenetic relationship of *DGAT1* genomic sequences and the presence of MYB96, MYB99, MYB108 or other transcription factor binding motifs in the first intron of each (different colored circles). Numbers at branch points are bootstrap values from 1500 replicates. For a full list of transcription factor binding sites identified in the first intron of *DGAT1* from each species see Supporting Information Table S3. (b) GenePainter analysis mapping gene structures onto protein multiple-sequence alignments amongst selected *DGAT1* genes. Scales for exon and intron lengths are provided.

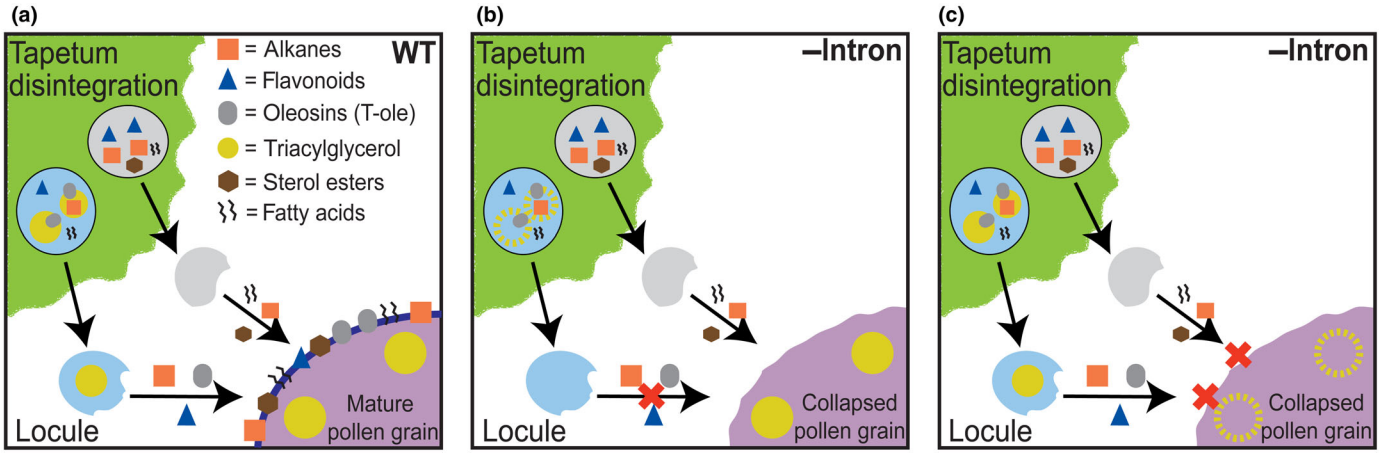

Fig. 7 Schematic hypothesizing two roles for the first intron of *DGAT1* during pollen development in Brassicaceae. (a) Inset view of a wild-type anther locule where the tapetum has accumulated elaioplasts (grey circles) and tapetosomes (blue circles). Tapetum disintegration releases materials that are deposited onto a maturing pollen grain. (b) Hypothesis 1: the absence of the first intron (−intron) of *DGAT1* impairs transcript abundance, in turn altering TAG accumulation (yellow-dashed circle) in the tapetosome and subsequent release of material during tapetal disintegration to the pollen coat (red cross) causing the pollen grain to collapse. (c) Hypothesis 2: the tapetum preserves TAG production but the pollen grain fails to accumulate TAG (yellow-dashed circle) and collapses. Thus, the deposition of lipids and proteins to the pollen coat does not occur (red crosses). These hypotheses are not mutually exclusive and may occur synchronously.

positively or negatively modulate gene expression (Back & Walther, 2021; Kuczynski *et al.*, 2022). Our results indicating the role of the *AtDGAT1* first intron in proper pollen development, the identification of various likely TF binding sites in the first intron of *AtDGAT1* and many other plant *DGAT1*s (Fig. 6a; Tables 1, S3), and the existing *Arabidopsis* DAP-seq data indicating TF binding to the *DGAT1* first intron (Fig. S11), together support that tissue-specific plant lipid metabolism is controlled in part by modulation of *DGAT1* gene expression through its first intron.

The evidence laid out above indicates that the first intron may recruit TFs for regulation of *DGAT1*; however, it is possible that other features of the first intron are responsible for fine-tuning gene expression. For example, putative enhancer motifs were identified in the first intron (Fig. S11). Although we performed in-depth *in silico* analysis of the first intron sequence itself, it is possible that a more cryptic regulatory sequence may be partially comprised of the intron sequence and its flanking regions, for example the 5′ exon 1 or 3′ exon 2. Further, it may be considered that structural alterations of the *DGAT1* pre-mRNA cause the observed differences between the −intron and +intron lines. Transcriptional efficiency is linked to proper folding of the RNA molecule with respect to processing by the polymerase, spliceosome and a cohort of other RNA-binding proteins (Zemora & Waldsich, 2010). Thus, it is conceivable that the pre-mRNA is differentially folded and the first intron is critical in the conformation for proper RNA processing. From the data presented here describing the role of the *DGAT1* first intron on pollen development, and the *in silico* analysis, it is clear that carefully controlled gene expression required for life is not limited to the promoter region. It should be noted that the same TF can regulate transcription of different genes by binding to distinct locations, as an example, MYB96 is known to bind to the promoter of *PDAT1* and it was suggested to indirectly regulate *DGAT1*. Additionally, a *myb96-2* knockout is known to simultaneously suppress *DGAT1* and *PDAT1* causing a fatty acid composition change and a 20% reduction in seed oil content (Lee *et al.*, 2018). When combined with our data this indicates that for MYB96-based regulation of *DGAT1* the binding motif is not found within the promoter region but in the first intron of *DGAT1* (Figs 6, S11). It is highly likely that other genes involved in essential central carbon metabolism (like lipid biosynthesis) are tightly regulated by not only their promoters but also the regulatory elements that exist within or proximal to the gene. In plants, this stringency of gene regulation likely evolved to either restrict or promote the use of particular metabolites for cell or tissue needs during development. As such researchers should carefully consider how introns may participate either co- or posttranscriptionally to modify gene abundance through any of the aforementioned mechanisms.

DGAT1-mediated TAG production may be equally important in gametophytic and sporophytic tissues

During pollen development in Brassicaceae the sporophytic tapetal cells accumulate lipids, proteins and sterol esters (Ting *et al.*, 1998; Hsieh & Huang, 2007). Specifically, two organelles, the tapetosome and the elaioplast house these metabolites. The tapetosome is wholly unique and can only be found within the tapetum. In wild-type cells the tapetosome accumulates flavonoids, alkanes and TAG-based oil bodies lined with tapetal-specific oleosin proteins (T-ole) (Fig. 7a; Hsieh & Huang, 2004, 2007; Lévesque-Lemay *et al.*, 2016). During tapetal disintegration, TAGs degrade, raising questions as to the fate of this lipid class. Several possibilities have been proposed, including their active breakdown for use as an energy source for tapetal cells during programmed cell death. Some of the linolenic acid (18:3)

present could be converted into jasmonic acid. Fatty acids could also be decarboxylated to form alkanes for deposition onto primordial pollen surfaces (Hsieh & Huang, 2004, 2007).

Although the precise role of TAG in tapetosomes remains elusive, a knockout mutant of phosphoserine phosphatase (PSP) is male sterile, a phenotype attributed to the absence of TAG oil bodies in the tapetosome (Flores-Tornero *et al.*, 2015). Furthermore, our data from the original complementation of *dgat1-1/dgat1-1//PDAT1/pdat1-2* with construct E546 (containing 1822 bp of the 5′ UTR and promoter) indicated that gametophytes with *tDGAT1/dgat1-1/pdat1-2* genotypes could be recovered and were functional; however, the failure of the resultant *tDGAT1/tDGAT1//dgat1-1/dgat1-1//pdat1-2/pdat1-2* parent to ever set seed (Fig. S2) indicates a maternal (perhaps tapetal) inability to accumulate TAG – likely due to insufficient *DGAT1* expression as driven by the 1822 bp promoter alone. In comparison, the −intron triple homozygous lines contained only an extra 243 bp (from position −1822 to −2066) of the promoter but could set a limited number of seed. These differences could be due to presence of a MYB80 TF binding motif in the distal promoter region in the −intron line (Table 2); coinciding with the sterility phenotype of plants expressing construct E546, a *myb80* knockout is male sterile due to tapetal and pollen deficiencies (Phan *et al.*, 2012). Perhaps revealing that *DGAT1* is in fact transcriptionally controlled by MYB80 for TAG production in the tapetum and pollen. Our −intron/+intron triple homozygous lines revealed an additional layer of regulation as both constructs likely contain varying levels of DGAT1 activity or *DGAT1* transcript abundance, thus providing further insights into how perturbations of TAG production may influence tapetal and pollen development. Whereby insufficient levels of DGAT1 activity – such as in the −intron line – cause a concomitant decrease in oil body size and shape, ultimately leading to an absence of resident oil bodies in the tapetosome or pollen grain. Thus, we propose two hypotheses as to the role of the *DGAT1* first intron of *Arabidopsis* in pollen development. Hypothesis 1: impaired TAG accumulation in the tapetosome alters the distribution of pollen coat precursors, including proteins such as T-ole and other lipidic material within the tapetum that are *required* to be donated to a maturing pollen grain (Fig. 7b). This hypothesis is not unfounded as in the *dgat1-1/dgat1-1* mutant, both tapetal and seed lipid body abundance and morphologies were altered (Zhang *et al.*, 2009; Bai *et al.*, 2023). Due to the dependency of pollen maturation on tapetal development (Tidy *et al.*, 2022; Wei & Ma, 2023), impairment of the latter often leads to precocious pollen abortion along with subsequent silique and embryo malformation, which was observed in our −intron lines (Figs 2b–h, 5h,i, S3–S8). Hypothesis 2: immature pollen grains in the −intron lines fail to accumulate sufficient TAG and collapse before receiving material from the sporophyte, contributing to the pollen abortion phenotype (Fig. 7c). Furthermore, the absence of TAG in the pollen grain may have adverse effects on pollen germination and polarization, as pollen oil bodies play roles in directing growth into a receptive stigma and the TAGs therein are also broken down, likely providing reductant for cytoskeletal remodeling during pollen tube growth (Rodríguez-García *et al.*, 2003; Ischebeck, 2016; Müller & Ischebeck, 2018; Bose *et al.*, 2021). It is foreseeable that many tissues have a minimal necessary threshold of *DGAT1* expression and rely on a certain rate of TAG synthesis that must occur within a specific timeframe for proper biological function.

Implications for plant bioengineering

As the primary gatekeeper enzymes of TAG accumulation, DGAT1 and PDAT1 are ideal candidates for genetic engineering studies focused on enhancing overall seed oil content, or accumulation of unique TAG molecular species containing novel fatty acids (Busta *et al.*, 2022; Chen *et al.*, 2022). To date, much research has been done on characterizing oil biosynthesis across the plant kingdom, with a largely singular focus on the overexpression of DGAT1s or PDAT1s. Typically, this expression has been 'laid on top' of the endogenous TAG biosynthetic pathways or used to complement a *dgat1-1/dgat1-1* knockout (Jako *et al.*, 2001; van Erp *et al.*, 2011; Regmi *et al.*, 2020; Hatanaka *et al.*, 2022; Klińska-Bąchor *et al.*, 2023). This past research has allowed us to ascertain some substrate preferences, metabolic pathways, and partial enzymatic functions of select DGAT1 isozymes, with the caveat that some endogenous TAG biosynthesis remains, complicating selectivity interpretations. Thus far, however, the ability to complement a *dgat1-1/dgat1-1//pdat1-2/pdat1-2* mutant has been unobtainable, namely due to limited TAG accumulation in specific nonseed tissues (Zhang *et al.*, 2009; Shakya & Bhatla, 2010) and the experimental focus on non-native expression systems. Given the desire to logically control lipid metabolism to produce TAGs of known fatty acid composition and other valuable lipophilic molecules, while minimizing deleterious effects such as impaired germination and reduced seed yield which can occur in overexpression studies (Jarvis *et al.*, 2021; Lunn *et al.*, 2021) our research demonstrates that identifying the functional regulatory elements within the promoter and introns requires more scrutiny. Paired with our assessment of the highly conserved genomic structure of *DGAT1* the most variable region appears to be the *N*-terminal domain (Fig. 6b). Therefore, both noncoding/flanking sequences (promoters, introns, terminators) and coding regions should be adequately considered in future engineering designs.

The finding that the first intron and promoter of *AtDGAT1*, when paired together, provides proper temporal expression of an exogenous *DGAT1* in multiple essential plant cell and tissue types, is a step forward in plant bioengineering. While the current study focuses on DGAT1, it is likely that DGAT2 may have also evolved to fill a similar role during gametophyte development in other species, this is supported by the finding that DGAT2, not DGAT1, is the active DGAT in olive (*Olea europaea*) anthers and ovules (Banilas *et al.*, 2011; Table S3). In this vein, our research describes a minimal unit (promoter + first intron) that can be used to replace endogenous TAG biosynthetic enzymes with various DGAT1 or DGAT2s from across the plant kingdom to produce the designer plant oils of the future. Similarly, one recent study concluded that inclusion of a first intron enhanced *DGAT1* expression in transgenic TAG-producing sugarcane

(Cao *et al.*, 2023). Additionally, regulatory introns in *Chlamydomonas reinhardtii* are crucial, with reports of > 5-fold increases in transgene expression when first intron sequences are included (Jaeger *et al.*, 2019). When combined with our study the importance of introns in regulating transcription is gaining interest for not only fundamental biology studies but also biotechnology applications. Moreover, identifying regulatory elements for tissue-specific expression is the next natural step in the plant science research community to accurately elucidate a genes' true function across both time and space.

Acknowledgements

We would like to thank Alexandra Richards for assistance with microscopy and Anu Regmi for technical assistance propagating plants and genotyping. This material is based upon work supported by the National Science Foundation under grant no. 2242822; the United States Department of Agriculture National Institute of Food and Agriculture Hatch Project no. 1015621 and Multi-State Project no. 1013013; and the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under award no. DE-SC0023142. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement from the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

Competing interests

None declared.

Author contributions

PDB and JS conceived the project and procured the funding. PDB agrees to serve as the author responsible for contact and ensures communication. STM, JS and PDB designed experiments, analyzed the data and prepared figures. STM and JS performed the experiments. All authors contributed to writing the article and approve the submitted version. STM and JS contributed equally to this work.

ORCID

Philip D. Bates https://orcid.org/0000-0002-1291-3363
Sean T. McGuire https://orcid.org/0000-0002-5629-7873
Jay Shockey https://orcid.org/0000-0002-5057-5457

Data availability

The data that supports the findings of this study are available in the main text and the Figures S1–S11 and Tables S1–S3 of the Supporting Information of this article. The transgenic lines produced for this manuscript can be made available after publication and upon request to the corresponding author.

References

Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* **26**: 962–980.

Alkotami L, Kornacki C, Campbell S, McIntosh G, Wilson C, Tran TNT, Durrett TP. 2021. Expression of a high-activity diacylglycerol acetyltransferase results in enhanced synthesis of acetyl-TAG in camelina seed oil. *The Plant Journal* **106**: 953–964.

Alvarez-Buylla ER, Benítez M, Corvera-Poiré A, Chaos Cador A, de Folter S, Gamboa de Buen A, Garay-Arroyo A, García-Ponce B, Jaimes-Miranda F, Pérez-Ruiz RV *et al.* 2010. Flower development. *Arabidopsis Book* **8**: e0127.

Alves-Ferreira M, Wellmer F, Banhara A, Kumar V, Riechmann JL, Meyerowitz EM. 2007. Global expression profiling applied to the analysis of Arabidopsis stamen development. *Plant Physiology* **145**: 747–762.

Back G, Walther D. 2021. Identification of *cis*-regulatory motifs in first introns and the prediction of intron-mediated enhancement of gene expression in *Arabidopsis thaliana*. *BMC Genomics* **22**: 390.

Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, Cottret L, Lelandais-Brière C, Owens GL, Carrère S, Mayjonade B *et al.* 2017. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**: 148–152.

Bafor M, Smith MA, Jonsson L, Stobart K, Stymne S. 1991. Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor-bean (*Ricinus communis*) endosperm. *The Biochemical Journal* **280**: 507–514.

Bai M, Gao H, Yang Y, Wu H. 2023. Changes in the content of pollen total lipid and TAG in *Arabidopsis thaliana* DGAT1 mutant as11. *AoB Plants* **15**: plad012.

Banaś A, Dahlqvist A, Ståhl U, Lenman M, Stymne S. 2000. The involvement of phospholipid:diacylglycerol acyltransferases in triacylglycerol production. *Biochemical Society Transactions* **28**: 703–705.

Banilas G, Karampelias M, Makariti I, Kourti A, Hatzopoulos P. 2011. The olive DGAT2 gene is developmentally regulated and shares overlapping but distinct expression patterns with DGAT1. *Journal of Experimental Botany* **62**: 521–532.

Bates PD. 2016. Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta* **1861**: 1214–1225.

Bhandari S, Bates PD. 2021. Triacylglycerol remodeling in *Physaria fendleri* indicates oil accumulation is dynamic and not a metabolic endpoint. *Plant Physiology* **187**: 799–815.

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

Blayney JW, Francis H, Rampasekova A, Camellato B, Mitchell L, Stolper R, Cornell L, Babbs C, Boeke JD, Higgs DR *et al.* 2023. Super-enhancers include classical enhancers and facilitators to fully activate gene expression. *Cell* **186**: 5826–5839.

Boavida LC, Becker JD, Feijó JA. 2005. The making of gametes in higher plants. *International Journal of Developmental Biology* **49**: 595–614.

Boavida LC, McCormick S. 2007. Temperature as a determinant factor for increased and reproducible *in vitro* pollen germination in *Arabidopsis thaliana*. *The Plant Journal* **52**: 570–582.

Bose D, Ngo AH, Nguyen VC, Nakamura Y. 2021. Non-specific phospholipases C2 and C6 redundantly function in pollen tube growth via triacylglycerol production in Arabidopsis. *The Plant Journal* **106**: 409–418.

Busta L, Chapman KD, Cahoon EB. 2022. Better together: protein partnerships for lineage-specific oil accumulation. *Current Opinion in Plant Biology* **66**: 102191.

Callis J, Fromm M, Walbot V. 1987. Introns increase gene expression in cultured maize cells. *Genes & Development* **1**: 1183–1200.

Cao VD, Luo G, Korynta S, Liu H, Liang Y, Shanklin J, Altpeter F. 2023. Intron-mediated enhancement of DIACYLGLYCEROL ACYLTRANSFERASE1 expression in energycane promotes a step change for lipid accumulation in vegetative tissues. *Biotechnology for Biofuels and Bioproducts* **16**: 153.

Caselli F, Beretta VM, Mantegazza O, Petrella R, Leo G, Guazzotti A, Herrera-Ubaldo H, de Folter S, Mendes MA, Kater MM *et al.* 2019. REM34 and REM35 control female and male gametophyte development in *Arabidopsis thaliana*. *Frontiers in Plant Science* **10**: 351.

Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G *et al.* 2010. Draft genome sequence of the oilseed species *Ricinus communis*. *Nature Biotechnology* **28**: 951–956.

Chandler JW, Cole M, Flier A, Werr W. 2009. BIM1, a bHLH protein involved in brassinosteroid signalling, controls Arabidopsis embryonic patterning via interaction with DORNROSCHEN and DORNROSCHEN-LIKE. *Plant Molecular Biology* **69**: 57–68.

Chen G, Harwood JL, Lemieux MJ, Stone SJ, Weselake RJ. 2022. Acyl-CoA: diacylglycerol acyltransferase: properties, physiological roles, metabolic engineering and intentional control. *Progress in Lipid Research* **88**: 101181.

Chen Y, Wang HY, Chen YF. 2021. The transcription factor MYB40 is a central regulator in arsenic resistance in Arabidopsis. *Plant Communications* **2**: 100234.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**: 735–743.

Correa SM, Fernie AR, Nikoloski Z, Brotman Y. 2020. Towards model-driven characterization and manipulation of plant lipid metabolism. *Progress in Lipid Research* **80**: 101051.

Dickinson H, Rodriguez-Enriquez J, Grant-Downton R. 2018. Pollen germination and pollen tube growth of *Arabidopsis thaliana*: *in vitro* and semi *in vivo* methods. *Bio-Protocol* **8**: e2977.

van Erp H, Bates PD, Burgal J, Shockey J, Browse J. 2011. Castor phospholipid: diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. *Plant Physiology* **155**: 683–693.

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

Fan J, Yan C, Xu C. 2013. Phospholipid:diacylglycerol acyltransferase-mediated triacylglycerol biosynthesis is crucial for protection against fatty acid-induced cell death in growing tissues of Arabidopsis. *The Plant Journal* **76**: 930–942.

Flores-Tornero M, Anoman AD, Rosa-Téllez S, Ros R. 2015. Lack of phosphoserine phosphatase activity alters pollen and tapetum development in *Arabidopsis thaliana*. *Plant Science* **235**: 81–88.

Fu L, Tomita A, Wang H, Buchholz DR, Shi YB. 2006. Transcriptional regulation of the *Xenopus laevis* stromelysin-3 gene by thyroid hormone is mediated by a DNA element in the first intron. *Journal of Biological Chemistry* **281**: 16870–16878.

Gallegos JE, Rose AB. 2019. An intron-derived motif strongly increases gene expression from transcribed sequences through a splicing independent mechanism in *Arabidopsis thaliana*. *Scientific Reports* **9**: 13777.

Geng P, Zhang S, Liu J, Zhao C, Wu J, Cao Y, Fu C, Han X, He H, Zhao Q. 2019. MYB20, MYB42, MYB43, and MYB85 regulate phenylalanine and lignin biosynthesis during secondary cell wall formation. *Plant Physiology* **182**: 1272–1283.

Gigolashvili T, Berger B, Mock H-P, Müller C, Weisshaar B, Flügge U-I. 2007. The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *The Plant Journal* **50**: 886–901.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N *et al.* 2011. PHYTOZOME: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**: D1178–D1186.

Greer MS, Truksa M, Deng W, Lung S-C, Chen G, Weselake RJ. 2015. Engineering increased triacylglycerol accumulation in *Saccharomyces cerevisiae* using a modified type 1 plant diacylglycerol acyltransferase. *Applied Microbiology and Biotechnology* **99**: 2243–2253.

Guerche P, Tire C, Grossi-de-Sá M, Clercq A, Van Montagu M, Krebbers E. 1990. Differential expression of the Arabidopsis 2S albumin genes and the effect of increasing gene family size. *Plant Cell* **2**: 469–478.

Hammesfahr B, Odronitz F, Mühlhausen S, Waack S, Kollmar M. 2013. GENEPAINTER: a fast tool for aligning gene structures of eukaryotic protein families, visualizing the alignments and mapping gene structures onto protein structures. *BMC Bioinformatics* **14**: 77.

Hatanaka T, Tomita Y, Matsuoka D, Sasayama D, Fukayama H, Azuma T, Soltani Gishini MF, Hildebrand D. 2022. Different acyl-CoA:diacylglycerol acyltransferases vary widely in function, and a targeted amino acid substitution enhances oil accumulation. *Journal of Experimental Botany* **73**: 3030–3043.

He M, Qin CX, Wang X, Ding NZ. 2020. Plant unsaturated fatty acids: biosynthesis and regulation. *Frontiers in Plant Science* **11**: 390.

Hendelman A, Zebell S, Rodriguez-Leal D, Dukler N, Robitaille G, Wu X, Kostyun J, Tal L, Wang P, Bartlett ME *et al.* 2021. Conserved pleiotropy of an ancient plant homeobox gene uncovered by *cis*-regulatory dissection. *Cell* **184**: 1724–1739.

Horstman A, Bemer M, Boutilier K. 2017. A transcriptional view on somatic embryogenesis. *Regeneration* **4**: 201–216.

Hsieh K, Huang AH. 2004. Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiology* **136**: 3427–3434.

Hsieh K, Huang AHC. 2007. Tapetosomes in *Brassica* tapetum accumulate endoplasmic reticulum–derived flavonoids and alkanes for delivery to the pollen surface. *Plant Cell* **19**: 582–596.

Ischebeck T. 2016. Lipids in pollen – they are different. *Biochimica et Biophysica Acta (BBA) – Molecular and Cell Biology of Lipids* **1861**: 1315–1328.

Ischebeck T, Krawczyk HE, Mullen RT, Dyer JM, Chapman KD. 2020. Lipid droplets in plants and algae: distribution, formation, turnover and function. *Seminars in Cell & Developmental Biology* **108**: 82–93.

Jaeger D, Baier T, Lauersen KJ. 2019. INTRONSERTER, an advanced online tool for design of intron containing transgenes. *Algal Research* **42**: 101588.

Jako C, Kumar A, Wei Y, Zou J, Barton DL, Giblin EM, Covello PS, Taylor DC. 2001. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiology* **126**: 861–874.

Jarvis BA, Romsdahl TB, McGinn MG, Nazarenus TJ, Cahoon EB, Chapman KD, Sedbrook JC. 2021. CRISPR/Cas9-induced fad2 and rod1 mutations stacked with fae1 confer high oleic acid seed oil in pennycress (*Thlaspi arvense* L.). *Frontiers in Plant Science* **12**: 652319.

Katano K, Oi T, Suzuki N. 2020. Failure of pollen attachment to the stigma triggers elongation of stigmatic papillae in *Arabidopsis thaliana*. *Frontiers in Plant Science* **11**: 989.

Katavic V, Reed DW, Taylor DC, Giblin EM, Barton DL, Zou J, Mackenzie SL, Covello PS, Kunst L. 1995. Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. *Plant Physiology* **108**: 399–409.

Kay R, Chan A, Daly M, McPherson J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* **236**: 1299–1302.

Kim SH, Lam PY, Lee M-H, Jeon HS, Tobimatsu Y, Park OK. 2020. The Arabidopsis R2R3 MYB transcription factor MYB15 is a key regulator of lignin biosynthesis in effector-triggered immunity. *Frontiers in Plant Science* **11**: 583153.

Kiselev KV, Aleynova OA, Ogneva ZV, Suprun AR, Dubrovina AS. 2021. 35S promoter-driven transgenes are variably expressed in different organs of *Arabidopsis thaliana* and in response to abiotic stress. *Molecular Biology Reports* **48**: 2235–2241.

Klińska-Bąchor S, Kędzierska S, Demski K, Banaś A. 2023. Phospholipid: diacylglycerol acyltransferase1-overexpression stimulates lipid turnover, oil production and fitness in cold-grown plants. *BMC Plant Biology* **23**: 370.

Ko JH, Kim WC, Han KH. 2009. Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *The Plant Journal* **60**: 649–665.

Koch MA. 2018. The plant model system Arabidopsis set in an evolutionary, systematic, and spatio-temporal context. *Journal of Experimental Botany* **70**: 55–67.

Kuczynski C, McCorkle S, Keereetaweep J, Shanklin J, Schwender J. 2022. An expanded role for the transcription factor WRINKLED1 in the biosynthesis of triacylglycerols during seed development. *Frontiers in Plant Science* **13**: 955589.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* 2007. ClustalW and ClustalX v.2.0. *Bioinformatics* **23**: 2947–2948.

Layat E, Cotterell S, Vaillant I, Yukawa Y, Tutois S, Tourmente S. 2012. Transcript levels, alternative splicing and proteolytic cleavage of TFIIIA control 5S rRNA accumulation during *Arabidopsis thaliana* development. *The Plant Journal* **71**: 35–44.

Lee HG, Kim H, Suh MC, Kim HU, Seo PJ. 2018. The MYB96 transcription factor regulates triacylglycerol accumulation by activating DGAT1 and PDAT1 expression in Arabidopsis seeds. *Plant & Cell Physiology* **59**: 1432–1442.

Lee HG, Mas P, Seo PJ. 2016. MYB96 shapes the circadian gating of ABA signaling in Arabidopsis. *Scientific Reports* **6**: 17754.

Lee HG, Park M-E, Park BY, Kim HU, Seo PJ. 2019. The Arabidopsis MYB96 transcription factor mediates ABA-dependent triacylglycerol accumulation in vegetative tissues under drought stress conditions. *Plants* **8**: 296.

Lee K, Lee HG, Yoon S, Kim HU, Seo PJ. 2015. The Arabidopsis MYB96 transcription factor is a positive regulator of ABSCISIC ACID-INSENSITIVE4 in the control of seed germination. *Plant Physiology* **168**: 677–689.

Leivar P, Martín G, Soy J, Dalton-Roesler J, Quail PH, Monte E. 2020. Phytochrome-imposed inhibition of PIF7 activity shapes photoperiodic growth in Arabidopsis together with PIF1, 3, 4 and 5. *Physiologia Plantarum* **169**: 452–466.

Lévesque-Lemay M, Chabot D, Hubbard K, Chan JK, Miller S, Robert LS. 2016. Tapetal oleosins play an essential role in tapetosome formation and protein relocation to the pollen coat. *New Phytologist* **209**: 691–704.

Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, Debono A, Durrett TP *et al.* 2013. Acyl-lipid metabolism. *Arabidopsis Book* **11**: e0161.

Lu CL, de Noyer SB, Hobbs DH, Kang J, Wen Y, Krachtus D, Hills MJ. 2003. Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in triacylglycerol biosynthesis, in *Arabidopsis thaliana*. *Plant Molecular Biology* **52**: 31–41.

Lunn D, Wallis JG, Browse J. 2021. A multigene approach secures hydroxy fatty acid production in Arabidopsis. *Journal of Experimental Botany* **73**: 2875–2888.

Mabuchi A, Soga K, Wakabayashi K, Hoson T. 2016. Phenotypic screening of Arabidopsis T-DNA insertion lines for cell wall mechanical properties revealed ANTHOCYANINLESS2, a cell wall-related gene. *Journal of Plant Physiology* **191**: 29–35.

Mandaokar A, Browse J. 2009. MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in Arabidopsis. *Plant Physiology* **149**: 851–862.

Martín G, Veciana N, Boix M, Rovira A, Henriques R, Monte E. 2020. The photoperiodic response of hypocotyl elongation involves regulation of CDF1 and CDF5 activity. *Physiologia Plantarum* **169**: 480–490.

McCarthy RL, Zhong R, Ye ZH. 2009. MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant & Cell Physiology* **50**: 1950–1964.

Mhaske V, Beldjilali K, Ohlrogge J, Pollard M. 2005. Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid: diacylglycerol transacylase gene (At5g13640). *Plant Physiology and Biochemistry* **43**: 413–417.

Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M. 2005. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* **17**: 2993–3006.

Morishita T, Kojima Y, Maruta T, Nishizawa-Yokoi A, Yabuta Y, Shigeoka S. 2009. Arabidopsis NAC transcription factor, ANAC078, regulates flavonoid biosynthesis under high-light. *Plant & Cell Physiology* **50**: 2210–2222.

Müller AO, Ischebeck T. 2018. Characterization of the enzymatic activity and physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New Phytologist* **217**: 1062–1076.

Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto KT, Takahashi T. 2006. Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant Physiology* **141**: 1363–1375.

Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology* **140**: 411–432.

O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. 2016. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**: 1280–1292.

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L *et al.* 2006. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Research* **35**: D883–D887.

Palmer CM, Hindt MN, Schmidt H, Clemens S, Guerinot ML. 2013. MYB10 and MYB72 are required for growth under iron-limiting conditions. *PLoS Genetics* **9**: e1003953.

Panda C, Li X, Wager A, Chen H-Y, Li X. 2020. An importin-beta-like protein mediates lignin-modification-induced dwarfism in Arabidopsis. *The Plant Journal* **102**: 1281–1293.

Parchuri P, Bhandari S, Azeez A, Chen G, Johnson K, Shockey J, Smertenko A, Bates PD. 2024. Identification of triacylglycerol remodeling mechanism to synthesize unusual fatty acid containing oils. *Nature Communications* **15**: 3547.

Peng FY, Weselake RJ. 2011. Gene coexpression clusters and putative regulatory elements underlying seed storage reserve accumulation in Arabidopsis. *BMC Genomics* **12**: 286.

Peterson R, Slovin JP, Chen C. 2010. A simplified method for differential staining of aborted and non-aborted pollen grains. *International Journal of Plant Biology* **1**: e13.

Phan HA, Li SF, Parish RW. 2012. MYB80, a regulator of tapetal and pollen development, is functionally conserved in crops. *Plant Molecular Biology* **78**: 171–183.

Piffanelli P, Ross JHE, Murphy DJ. 1997. Intra- and extracellular lipid composition and associated gene expression patterns during pollen development in *Brassica napus*. *The Plant Journal* **11**: 549–562.

Regmi A, Shockey J, Kotapati HK, Bates PD. 2020. Oil-producing metabolons containing DGAT1 use separate substrate pools from those containing DGAT2 or PDAT. *Plant Physiology* **184**: 720–737.

Rodríguez-García MI, M'Rani-Alaoui M, Fernández MC. 2003. Behavior of storage lipids during development and germination of olive (*Olea europaea* L.) pollen. *Protoplasma* **221**: 237–244.

Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB. 2017. Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**: 470–480.

Rose AB, Elfersi T, Parra G, Korf I. 2008. Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression. *Plant Cell* **20**: 543–551.

Sakuraba Y, Jeong J, Kang MY, Kim J, Paek NC, Choi G. 2014. Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in Arabidopsis. *Nature Communications* **5**: 4636.

Sanders PM, Bui AQ, Weterings K, McIntire KN, Hsu Y-C, Lee PY, Truong MT, Beals TP, Goldberg RB. 1999. Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sexual Plant Reproduction* **11**: 297–322.

Shakya R, Bhatla SC. 2010. A comparative analysis of the distribution and composition of lipidic constituents and associated enzymes in pollen and stigma of sunflower. *Sexual Plant Reproduction* **23**: 163–172.

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

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM. 2006. Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* **18**: 2294–2313.

Sims TL, Goldberg RB. 1989. The glycinin Gy1 gene from soybean. *Nucleic Acids Research* **17**: 4386.

Sozzani R, Maggio C, Giordo R, Umana E, Ascencio-Ibañez JT, Hanley-Bowdoin L, Bergounioux C, Cella R, Albani D. 2010. The E2FD/DEL2 factor is a component of a regulatory network controlling cell proliferation and development in Arabidopsis. *Plant Molecular Biology* **72**: 381–395.

Stuitje AR, Verbree EC, van der Linden KH, Mietkiewska EM, Nap JP, Kneppers TJ. 2003. Seed-expressed fluorescent proteins as versatile tools for

easy (co)transformation and high-throughput functional genomics in Arabidopsis. *Plant Biotechnology Journal* **1**: 301–309.

Susila H, Nasim Z, Gawarecka K, Jung J-Y, Jin S, Youn G, Ahn JH. 2023. Chloroplasts prevent precocious flowering through a GOLDEN2-LIKE–B-BOX DOMAIN PROTEIN module. *Plant Communications* **4**: 100515.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis v.11. *Molecular Biology and Evolution* **38**: 3022–3027.

Taylor-Teeples M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, Young NF, Trabucco GM, Veling MT, Lamothe R *et al.* 2015. An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* **517**: 571–575.

Tian F, Yang DC, Meng YQ, Jin J, Gao G. 2020. PLANTREGMAP: charting functional regulatory maps in plants. *Nucleic Acids Research* **48**: D1104–D1113.

Tidy AC, Ferjentsikova I, Vizcay-Barrena G, Liu B, Yin W, Higgins JD, Xu J, Zhang D, Geelen D, Wilson ZA. 2022. Sporophytic control of pollen meiotic progression is mediated by tapetum expression of ABORTED MICROSPORES. *Journal of Experimental Botany* **73**: 5543–5558.

Ting JT, Wu SS, Ratnayake C, Huang AH. 1998. Constituents of the tapetosomes and elaioplasts in *Brassica campestris* tapetum and their degradation and retention during microsporogenesis. *The Plant Journal* **16**: 541–551.

Valliyodan B, Cannon SB, Bayer PE, Shu S, Brown AV, Ren L, Jenkins J, Chung CY-L, Chan T-F, Daum CG *et al.* 2019. Construction and comparison of three reference-quality genome assemblies for soybean. *The Plant Journal* **100**: 1066–1082.

Vanhercke T, Dyer JM, Mullen RT, Kilaru A, Rahman MM, Petrie JR, Green AG, Yurchenko O, Singh SP. 2019. Metabolic engineering for enhanced oil in biomass. *Progress in Lipid Research* **74**: 103–129.

Verdaguer B, de Kochko A, Beachy RN, Fauquet C. 1996. Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. *Plant Molecular Biology* **31**: 1129–1139.

Verma N, Burma PK. 2017. Regulation of tapetum-specific A9 promoter by transcription factors AtMYB80, AtMYB1 and AtMYB4 in *Arabidopsis thaliana* and *Nicotiana tabacum*. *The Plant Journal* **92**: 481–494.

Wan X, Wu S, Li Z, An X, Tian Y. 2020. Lipid metabolism: critical roles in male fertility and other aspects of reproductive development in plants. *Molecular Plant* **13**: 955–983.

Wang C, Wang H, Li P, Li H, Xu C, Cohen H, Aharoni A, Wu S. 2020a. Developmental programs interact with abscisic acid to coordinate root suberization in Arabidopsis. *The Plant Journal* **104**: 241–251.

Wang L, Li Q, Xia Q, Shen W, Selvaraj G, Zou J. 2020b. On the role of DGAT1 in seed glycerolipid metabolic network and critical stages of plant development in Arabidopsis. *Lipids* **55**: 457–467.

Wang X, Aguirre L, Rodríguez-Leal D, Hendelman A, Benoit M, Lippman ZB. 2021. Dissecting *cis*-regulatory control of quantitative trait variation in a plant stem cell circuit. *Nature Plants* **7**: 419–427.

Wang Y, Zhang W-Z, Song L-F, Zou J-J, Su Z, Wu W-H. 2008. Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in Arabidopsis. *Plant Physiology* **148**: 1201–1211.

Ward JM, Cufr CA, Denzel MA, Neff MM. 2005. The Dof transcription factor OBP3 modulates phytochrome and cryptochrome signaling in Arabidopsis. *Plant Cell* **17**: 475–485.

Watanabe M, Rebbert ML, Andreazzoli M, Takahashi N, Toyama R, Zimmerman S, Whitman M, Dawid IB. 2002. Regulation of the Lim-1 gene is mediated through conserved FAST-1/FoxH1 sites in the first intron. *Developmental Dynamics* **225**: 448–456.

Wei S, Ma L. 2023. Comprehensive insight into tapetum-mediated pollen development in *Arabidopsis thaliana*. *Cells* **12**: 247.

Wilkinson JE, Twell D, Lindsey K. 1997. Activities of CaMV 35S and nos promoters in pollen: implications for field release of transgenic plants. *Journal of Experimental Botany* **48**: 265–275.

Winichayakul S, Curran A, Moraga R, Cookson R, Xue H, Crowther T, Roldan M, Bryan G, Roberts N. 2022. An alternative angiosperm DGAT1 topology and potential motifs in the N-terminus. *Frontiers in Plant Science* **13**: 951389.

Xu J, Carlsson AS, Francis T, Zhang M, Hoffman T, Giblin ME, Taylor DC. 2012. Triacylglycerol synthesis by PDAT1 in the absence of DGAT1 activity is dependent on re-acylation of LPC by LPCAT2. *BMC Plant Biology* **12**: 4.

Yang W, Wang G, Li J, Bates PD, Wang X, Allen DK. 2017. Phospholipase Dζ enhances diacylglycerol flux into triacylglycerol. *Plant Physiology* **174**: 110–123.

Yang Y, Benning C. 2018. Functions of triacylglycerols during plant development and stress. *Current Opinion in Biotechnology* **49**: 191–198.

Yang Y, Yu X, Song L, An C. 2011. ABI4 activates DGAT1 expression in arabidopsis seedlings during nitrogen deficiency. *Plant Physiology* **156**: 873–883.

Zemora G, Waldsich C. 2010. RNA folding in living cells. *RNA Biology* **7**: 634–641.

Zhang C, Gao M, Seitz NC, Angel W, Hallworth A, Wiratan L, Darwish O, Alkharouf N, Dawit T, Lin D *et al.* 2019a. LUX ARRHYTHMO mediates crosstalk between the circadian clock and defense in Arabidopsis. *Nature Communications* **10**: 2543.

Zhang M, Fan J, Taylor DC, Ohlrogge JB. 2009. DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *Plant Cell* **21**: 3885–3901.

Zhang P, Wang R, Ju Q, Li W, Tran LP, Xu J. 2019b. The R2R3-MYB transcription factor MYB49 regulates cadmium accumulation. *Plant Physiology* **180**: 529–542.

Zhou J, Lee C, Zhong R, Ye ZH. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell* **21**: 248–266.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Graphical representation of full-length genomic *AtDGAT1* binary plasmids.

Fig. S2 Initial identification of transgenic construct elements to rescue the lethality of the *dgat1-1/dgat1-1//pdat1-2/pdat1-2* mutant.

Fig. S3 T_2 screening of brown vs red independent transformants of *tDGAT1* (constructs E826, −intron; and E827 +intron) in *dgat1-1/dgat1-1//PDAT1/pdat1-2*.

Fig. S4 Silique and embryo abortion phenotypes in developing siliques of *tDGAT1/tDGAT1//dgat1-1/dgat1-1//pdat1-2/pdat1-2*.

Fig. S5 Representative images of plant development of *dgat1-1/dgat1//pdat1-2/pdat1-2* complements.

Fig. S6 Floral development is unaffected in the absence of the first intron of AtDGAT1.

Fig. S7 Alexander's staining of mature pollen grains.

Fig. S8 The absence of the first intron (−intron, line #2) ceases pollen development.

Fig. S9 Pollen tube growth after 5-h of incubation.

Fig. S10 TAIR10 data of relative expression patterns of *DGAT1*, with quantitative polymerase chain reaction determined *DGAT1* and *PDAT1* expression in flowers (stage 10–11).

Fig. S11 *In silico* analysis of the *AtDGAT1* first intron.

Table S1 Primers used during this study. *AscI* site is underlined.

Table S2 Entire genetic screening of complemented plants for *PDAT1/pdat1-2* locus segregation.

Table S3 Entire TF binding motif analysis of analyzed *DGAT1* first introns.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.