Article

Transcriptome Analysis and Identification of Lipid Genes in *Physaria lindheimeri*, a Genetic Resource for Hydroxy Fatty Acids in Seed Oil

Grace Q. Chen [1,*,†], Won Nyeong Kim [2], Kumiko Johnson [1], Mid-Eum Park [3], Kyeong-Ryeol Lee [4]
and Hyun Uk Kim [2,3,*,†]

1   Western Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture,
     Albany, CA 94710, USA; kumiko.johnson@usda.gov
2   Department of Bioindustry and Bioresource Engineering, Sejong University, Seoul 05006, Korea;
     rladnjssud@naver.com
3   Department of Molecular Biology, Graduate School, Sejong University, Seoul 05006, Korea; aledja@naver.com
4   Department of Agricultural Biotechnology, National Institute of Agricultural Sciences, Rural Development
     Administration, Jeonju 54974, Korea; realdanny@korea.kr
*   Correspondence: grace.chen@usda.gov (G.Q.C.); hukim64@sejong.ac.kr (H.U.K.)
†   Equal contribution.

Abstract: Hydroxy fatty acids (HFAs) have numerous industrial applications but are absent in most vegetable oils. *Physaria lindheimeri* accumulating 85% HFA in its seed oil makes it a valuable resource for engineering oilseed crops for HFA production. To discover lipid genes involved in HFA synthesis in *P. lindheimeri*, transcripts from developing seeds at various stages, as well as leaf and flower buds, were sequenced. Ninety-seven percent clean reads from 552,614,582 raw reads were assembled to 129,633 contigs (or transcripts) which represented 85,948 unique genes. Gene Ontology analysis indicated that 60% of the contigs matched proteins involved in biological process, cellular component or molecular function, while the remaining matched unknown proteins. We identified 42 *P. lindheimeri* genes involved in fatty acid and seed oil biosynthesis, and 39 of them shared 78–100% nucleotide identity with Arabidopsis orthologs. We manually annotated 16 key genes and 14 of them contained full-length protein sequences, indicating high coverage of clean reads to the assembled contigs. A detailed profiling of the 16 genes revealed various spatial and temporal expression patterns. The further comparison of their protein sequences uncovered amino acids conserved among HFA-producing species, but these varied among non-HFA-producing species. Our findings provide essential information for basic and applied research on HFA biosynthesis.

Keywords: hydroxy fatty acid; *Physaria lindheimeri*; seed; transcriptome; triacylglycerol; gene expression

1. Introduction

Hydroxy fatty acids (HFAs) and their derivatives are used as raw materials for numerous industrial products, such as lubricants, plastics and surfactants [1,2]. The conventional source of HFA is from castor (*Ricinus communis*), which contains 90% ricinoleic acid (12-hydroxy 9-octadecenoic acid, 18:1OH) in its seed oil. The production of castor oil, however, is hampered by the presence of the toxin ricin [3,4] and hyper-allergic 2S albumins [5–7] in its seeds. Lesquerella (*Physaria fendleri*) and *Physaria lindheimeri* seed oils contain lesquerolic acid (14-hydroxy 11-eicosenoic acid, 20:1OH), at approximately 60% [8–12] and 85% [13–15], respectively. Lesquerella possesses key domestication traits such as indehiscent seeds (or non-shattering seeds), photoperiod neutrality, upright growth habit, and extensive branching that leads to high seed yields, thus it is being developed as a safe source for HFA production [10]. Considerable breeding efforts, such as single plant or bulking recurrent selections, are made to improve lesquerella with increased HFA oil content [12]. *P. lindheimeri*

is also crossed with lesquerella, but there is no significant increase in HFA level in the hybrid off-springs of lesquerella [16]. With the success of lesquerella biotechnology [17,18], genes and regulatory elements from *P. lindheimeri* can be used as excellent targets for improving lesquerella through *Agrobacterium*-mediated genetic transformation. Besides the HFAs found in plant seeds, families of saturated hydroxy fatty acids (SHFAs) have been recently discovered in cow, goat, and human milks [19,20], and the SHFAs from human milk inhibit the growth of human cancer cells and suppress beta-cell apoptosis [19], indicating that SHFAs may play a role in the promotion and protection of human health.

The formation of seed oil (triacylglycerol, TAG) starts from *de novo* fatty acid (FA) biosynthesis in plastid and TAG assembly in the endoplasmic reticulum (ER) [21]. HFA formation and HFA-containing TAG assembly occur after the FA is exported to cytosol and activated to acyl-coenzyme A (acyl-CoA) [22]. Based on the current knowledge, a proposed pathway of 20:1OH synthesis and 20:1OH-TAG assembly in *P. lindheimeri* is depicted in Figure 1. In developing *P. lindheimeri* seeds, oleic acid (18:1) is synthesized in the plastid, exported and activated to 18:1-CoA in cytosol. The newly synthesized 18:1-CoA is acylated directly into membrane lipid phosphatidylcholine (PC) in the ER by the forward reaction of lyso-PC acyltransferase (LPCAT) [23–25], resulting in 18:1-PC (purple arrow in Figure 1). As 18:1-PC is the substrate of oleate 12-hydroxylase (FAH12) [26–28], a *P. lindheimeri* FAH12 (PlFAH12) [13] hydroxylates 18:1-PC to form 18:1OH-PC (Figure 1). Through the reverse reaction of LPCAT, the 18:1OH can be removed from PC and transferred back to cytosol to be activated as 18:1OH-CoA (purple arrow in Figure 1). The de-acylation of *sn*-2 PC can also occur with a phospholipase A (PLA$_2$)-type activity to yield a free FA, which is then activated to acyl-CoA [29] (orange arrow in Figure 1). In *Physaria* species accumulating 20:1OH, an FA-condensing enzyme KCS18 (or KCS3) elongates 18:1OH-CoA to 20:1OH-CoA [30] (Figure 1). Rapid acylation and de-acylation (or acyl editing) by LPCAT (or by PLA$_2$), and in conjunction with efficient elongation by KCS18, lead to the enrichment of 20:1OH-CoA in the cytosol during seed development. Once the 20:1OH-CoA is formed, 20-1OH-CoA can be acylated to TAG by the conventional Kennedy pathway [31–33], which consists of three sequential acylations of acyl-CoAs to a glycerol-3-phosphate (G3P) backbone (solid blue arrows in Figure 1). First, the *sn*-1 position of G3P is acylated by glycerol-3-phosphate acyltransferase (GPAT) to produce lysophosphatidic acid (LPA). Second, the *sn*-2 position of LPA is acylated by LPA acyltransferase (LPAT) to generate phosphatidic acid (PA). PA is then converted to 1,2-*sn*-diacylglycerol (DAG, or de novo DAG) by PA phosphatase (PAP). Finally, the *sn*-3 position of DAG is acylated by 1,2-*sn*-diacylglycerol acyltransferase (DGAT) to produce TAG. In addition to the Kennedy pathway, the 20:1OH-CoA can also be directly acylated to PC by LPCAT, generating 20:1OH-PC (purple arrow in Figure 1), and the 20:1OH on PC can be transferred to the *sn*-3 position of DAG by phospholipid:DAG acyltransferase (PDAT), forming 20:1OH-TAG [34–36] (solid brown arrows in Figure 1).

Metabolic labeling experiments reveal that in some oilseeds containing high amounts of poly-unsaturated FA, such as linoleic acid (18:2) and linolenic acid (18:3), a separate PC-derived DAG pool is utilized for TAG synthesis [25]. Like 18:1OH, 18:2 and 18:3 are also formed on PC through the desaturation of 18:1-PC to 18:2-PC, and 18:2-PC to 18:3-PC by FA desaturase 2 (FAD2) [37] and FA desaturase 3 (FAD3) [38,39], respectively. These modified FA-PCs (including 18:2-PC, 18:3-PC, and 18:1OH-PC) can be converted to PC-derived DAG mainly by PC:DAG cholinephosphotransferase (PDCT) through the phosphocholine headgroup exchange between PC and DAG [25,40,41] (green arrow in Figure 1). Alternatively, PC-derived DAG can be produced by the reverse action of CDP-choline:DAG cholinephosphotransferase (CPT) [42], or a lipase-based mechanism using phospholipase C, or phospholipase D plus PAP [25]. By this means, PC-derived DAG can be enriched with modified FA, such as 18:2-, 18:3-, and 18:1OH-DAG. Thus, utilizing PC-derived DAG represents a mechanism of channeling modified FA to TAG [25]. In case of *P. lindheimeri*, the modified 18:1OH has to be removed from PC and elongated to 20:1OH in cytosol, so 20:1OH-PC can be formed by the direct acylation of 20:1OH-CoA to LPC by LPCAT (purple arrows in Figure 1). The LPCAT-mediated formation of 20:1OH-PC and

PC-derived 20:1OH-DAG synthesis would be regarded as new substrates for the synthesis of TAG catalyzed by PDAT or DGAT (dotted brown and blue arrows, Figure 1).
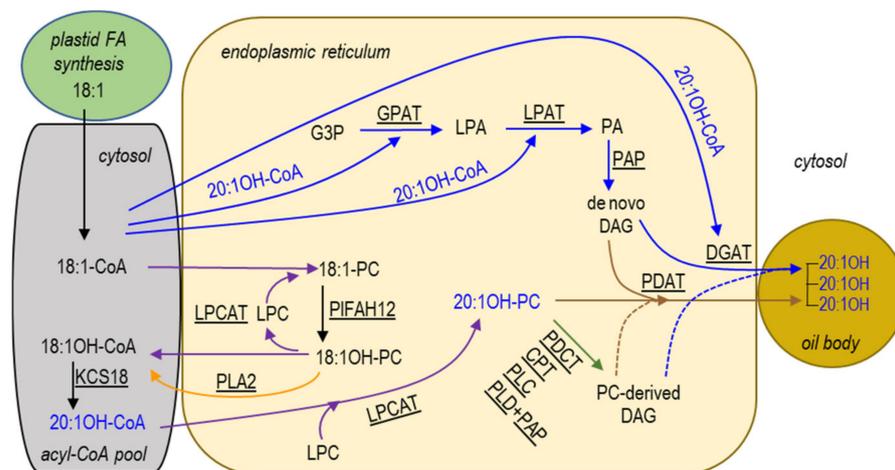

Figure 1. Proposed pathways for triacylglycerol (TAG) biosynthesis in *P. lindheimeri* seed. Blue arrows indicate reactions involved in Kennedy pathway. Purple arrows indicate reactions involved in acyl editing by LPCAT. Brown arrows indicate PDAT-mediated pathways. Green arrows indicate reactions involved in PC-derived DAG synthesis. Dotted lines indicate PC-derived DAG utilized by DGAT and PDAT. Enzymes catalyzing these reactions are underlined. FA numerical symbols: 18:1, oleic acid; 18:1OH, ricinoleic acid; 20:1OH, lesquerolic acid. Abbreviations: CoA, co-enzyme A; PC, phosphatidylcholine; LPCAT, lysophosphatidylcholine acyltransferase; LPC, lysophosphatidylcholine; PLA2, phospholipase A; PlFAH12, *P. lindheimeri* Δ12 oleic acid hydroxylase; KCS18 (or KCS3), 3-ketoacyl-CoA synthase 18; G3P, glycerol-3-phosphate; LPA, lysophosphatidic acid; PA, phosphatidic acid; DAG, diacylglycerol; GPAT, glycerol 3-phosphate acyltransferase; LPAT, lysophosphatidic acid acyltransferase; PAP, phosphatidic acid phosphatase; DGAT, diacylglycerol acyltransferase; CPT, CDP-choline:DAG cholinephosphotransferase; PDAT, phospholipid:DAG acyltransferase; PDCT, PC:DAG cholinephosphotransferase; CPT, CDP-choline:DAG cholinephosphotransferase; PLC, phospholipase C; PLD, phospholipase D; TAG, triacylglycerol.

Castor genes are introduced into non-HFA-oilseed Arabidopsis (*Arabidopsis thaliana*) or Camelina (*Camelina Sativa*) to study the HFA biosynthesis mechanism [43]. Castor *RcFAH12* is firstly isolated and demonstrated to be responsible for HFAs synthesis in transgenic seeds up to 17% [28,44,45]. Additional genes, Castor *RcDGAT2* [46], *RcPDAT1-2* (or *RcPDAT1-A*) [35,36], and *RcPDCT* [41], *RcPLCL1* [47], *RcLPAT2* [48,49], *RcLPAT3B* and *RcLPATB* [49], and *RcGPAT9* together with *RcLPAT2* and *RcPDAT1A* [50], are demonstrated to increase the HFAs content of transgenic Arabidopsis or Camelina from 17% to 28%. Recent attempts at the co-expression of certain *PLA*, including *RcsPLA$_2$α* [29] or *RcpPLAIIIβ* [51], with *RcFAH12* in Arabidopsis results in decreases in seed HFAs contents when compared with Arabidopsis expressing *RcFAH12* alone, which indicates that the cleaved HFAs at the *sn*-2 position of PC are not incorporated into seed TAG. However, when a castor lecithin:cholesterol acyltransferase-like PLA (*RcLCAT-PLA*) or a lesquerella *PfLCAT-PLA* is co-expressed with *RcFAH12* in Arabidopsis, the seed HFAs content remains constant, and the HFAs are re-distributed from the *sn*-2 position of TAG to the *sn*-1/3 positions of TAG [52].

Castor TAGs contain 90% 18:1OH, and over 70% of the TAGs have all three positions esterified with 18:1OH [53]. Lesquerella TAGs contain ~60% 20:1OH and almost all of it is esterified to the *sn*-1 and *sn*-3 positions [54] (Figure 1). The absence of HFAs at the *sn*-2 position of TAG might be caused by the preference of endogenous lesquerella LPATs (PfLPATs) for common FAs [18]. To increase the HFA content at the *sn*-2 position of lesquerella TAGs, we have overexpressed *RcLPAT2* in lesquerella seeds, and the resulting transgenic seed oil increases 3-HFA-TAGs (TAGs with all three *sn* positions acylated with HFAs) from 5% to 13–14% [18]. Regiochemical analysis reveals that RcLPAT2 increases

the 3-HFA-TAGs content by enhancing the acylation of 18:1OH at the *sn*-2 position of 20:1OH-LPA for the subsequent generation of *sn*-1/3-20:1OH-*sn*-2-18:1OH-TAG [55]. The limited accumulation of HFAs in transgenics indicates that additional genes are required to implement new strategies for increased HFA production in lesquerella. *P. lindheimeri* oil contains 85% 20:1OH, similarly high to that of castor oil, and genetically, *P. lindheimeri* is closely related to lesquerella, both belonging to the *Physaria* genus. The genes from *P. lindheimeri* may be readily adapted to the molecular machinery of gene transcription and translation in lesquerella. Genomic data from castor [56,57] and lesquerella [58,59] are freely available; however, so far, only one gene encoding oleate 12-hydroxylase from *P. lindheimeri* (*PlFAH12*) has been isolated by polymerase chain reaction (PCR) [13]. Transgenic Arabidopsis expressing *PlFAH12* accumulated 18% HFAs in seed oil [13].

In this study, we describe *P. lindheimeri* transcriptomes constructed from leaf, flower bud, and developing seeds at various stages. The lipid genes involved in FA and TAG biosynthesis were identified by the deep mining of the transcriptomes. Detailed spatial and temporal expression profiles of key genes were characterized by quantitative PCR (qPCR). *P. lindheimeri* sequences were further compared with orthologs from HFA-producing castor and lesquerella, as well as orthologs from non-HFA-producing species, including Arabidopsis and Camelina. Our findings provide essential information for future basic and applied research on HFA biosynthesis in oilseeds.

2. Results and Discussion

2.1. Analysis of Transcriptomes, Gene Ontology and Differentially Expressed Genes (DEGs)

We have previously conducted a series of seed developmental studies in *P. lindheimeri*, including changes of weight, lipid accumulation, and FA composition [15]. We identified three major periods: (1) early period, up to 21 days after pollination (DAP), when developing seeds rapidly increase in size and fresh weight; (2) mid-maturation period, from 28 to 42 DAP, when lipids and dry weights accumulate steadily, but seed size and fresh weigh cease at 42 DAP and begin to decline; and (3) late-maturation/desiccation stages, from 42 DAP to 56 DAP, when seeds enter the natural dehydrate process. The major FA 20:1OH accumulates steadily from 28 DAP to 56 DAP, eventually reach up to 77% of the total seed oil [15].

To obtain a comprehensive profile of the transcriptome of *P. lindheimeri* seeds, we collected RNA samples from seeds at 21, 35, and 42 DAP, which represent the time points of the early, middle and late stages of seed development. RNA samples were also collected from leaf and flower buds for comparative studies. Using the Illumina Hiseq2000 system (Illumina, San Diego, CA, USA), a total of 552,614,582 raw reads were generated. After trimming the adapter and those shorter than 25 bp low-quality reads, 539,709,492 clean reads were obtained (Table 1). The number of clean reads ranged from 95,336,514 to 125,026,944, or 97% to 99% of raw reads, indicating that high quality sequencing data were equally obtained from these five samples (Table 1). The de novo reconstruction of RNA-seq data resulted in a total of 129,633 contigs (or transcripts) represented by 85,948 unique genes with an average N50 contig size (50% of all bases reside in a contig) of 1373 bp (Table 2), which is comparable to the 1310 bp of the high-quality transcriptome of *P. fendleri*, a relative of *P. lindheimeri*, which also produces 20:1OH at 55–60% in seeds [58]. Gene Ontology (GO) analysis was performed through the BlastX program at NCBI [60], which compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database (go_v20150407). Among the 129,633 contigs, 40% of them did not match proteins with known functions, and the remaining 24%, 19% and 17% matched proteins in biological processes, cellular components and molecular functions, respectively (Supplementary File 1, Figure S1A). In the category of "biological process", "metabolic process" (23%), "biological regulation" (14%), "response to stimulus" (12%), and "single-organism process" (8%) were major groups (Supplementary File 1, Figure S1B). In the category of "cellular component", "cell part" had a largest percent of contigs (42%), followed by "organelle" (27%) (Supplementary File 1, Figure S1C). Within the category of

"molecular function", "binding" (39%) and "catalytic activity" (31%) counted as the two top sub-categories (Supplementary File 1, Figure S1D). The results of the GO analysis are similar to those for *P. fendleri* [58].

Table 1. Summary of sequencing data of *P. lindheimeri* seeds, flower bud and leaf transcriptomes.

	Developing Seed (DAP) [a]			Flower Bud	Leaf	Total
	21	**35**	**42**			
raw reads	129,121,576	99,873,058	127,855,980	99,729,366	96,034,602	552,614,582
clean reads [b]	125,026,944	97,400,124	123,850,518	98,065,392	95,366,514	539,709,492
ratio [c] (%)	96.8	97.5	96.9	98.3	99.3	97.7

[a] days after pollination. [b] clean reads were generated by trimming the adapter and low-quality reads. [c] ratio is calculated by clean reads/raw reads.

Table 2. Statistical summary of de novo transcriptome assembly.

	Number
Numbers of unique genes transcripts (longest isoforms)	85,948
Numbers of transcripts (contigs)	129,633
Average length of transcripts (bp)	814 (bp)
N50 [a] transcript size (bp)	1373 (bp)
Total assembled bases	105,504,114 (nt)

[a] the N50 is defined as the minimum contig length needed to cover 50% of the genome.

To infer the relative expression level of a gene, FPKM (fragments per kilobase of transcript per million mapped reads) was used to estimate transcript abundances [61]. After filtering out of 72,932 contigs which had an FPKM value of zero in at least one of all five samples (Materials and Methods), 56,701 remaining contigs were subjected to comparative studies. Compared with the expression level in leaf, a total of 7857 contigs showed significant changes, i.e., four-fold or more (up or down) in at least one of the developing seed or flower bud samples (Supplementary File 1, Figure S2; Supplementary File 2, Table S1). Developing seeds at 21, 35 and 42 DAP had similar numbers of up-regulated contigs ranging from 2080 to 2224, and down-regulated contigs ranging from 2593 to 3082, whereas the flower bud had only 26 up- and 193 down-regulated contigs (Supplementary File 1, Figure S2; Supplementary File 2, Tables S1–S5). The results suggest that the gene expression profiles in developing seeds are distinctly different from those of the leaf and flower bud. To visualize the differential expression of genes, we constructed a heat map of hierarchically clustered expression data (Figure 2). The hierarchical clustering of differentially expressed contigs clearly showed that developing seeds vs. leaf and flower bud had different expression patterns. Approximately 45% of the total contigs were up-regulated in leaf/flower buds and down-regulated in 21/42 DAP seeds (group A of clusters, Figure 2). Oppositely, about 26% of the contigs were down-regulated in leaf/flower buds and up-regulated in 21/42 DAP seeds (group D of clusters, Figure 2). The expression patterns in 35 DAP seeds were overall similar to those of 21/42 DAP seeds, with the exception of the group B of clusters which had a pattern similar to that of leaf/flower bud (Figure 2). Besides this, 35 DAP seed had other distinct patterns: about 10% of the total contigs were highly up-regulated (blue bordered contigs from group C, Figure 2), while about 5% were extremely down-regulated (yellow bordered contigs from group A, Figure 2). Most contigs in the group D from 35 DAP seed had mixed patterns, and some of them had similar levels to that of the 21/42 seeds or leaf/flower bud, and some in between (Figure 2). As 35 DAP seeds are at the peak of seed development [15], it is not a surprise to see some genes with unique differential expression patterns. For example, *PlFAH12* and *PlKCS18*'s transcript levels were highly elevated at 35 DAP, which was verified by

qPCR (Figure 3 and Figure 6), suggesting that gene transcription played a critical role in regulating HFA synthesis during seed development.
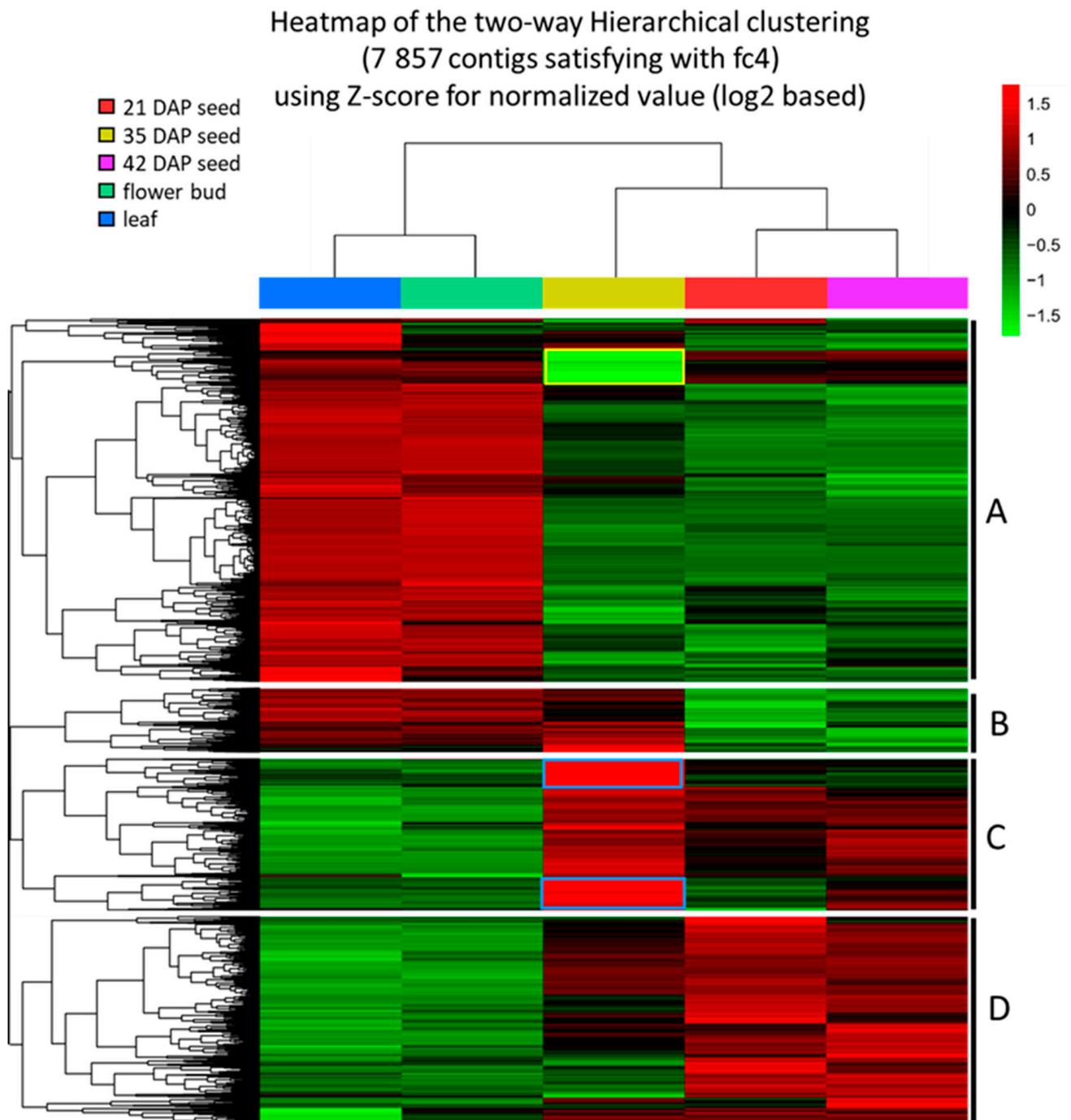

Figure 2. Heat map for hierarchical clustering. For a significant total of 7857 contigs, samples and contigs with similar expression levels were grouped by hierarchical clustering analysis (Euclidean method, complete linkage) using expression level (normalized value) for each sample and contigs. In the 35 DAP sample, yellow or blue boxes indicate that most of the contigs were highly down-regulated (green color) or up-regulated (red color). Groups A, B, C and D were visually divided. fc4, fold change ≥4.
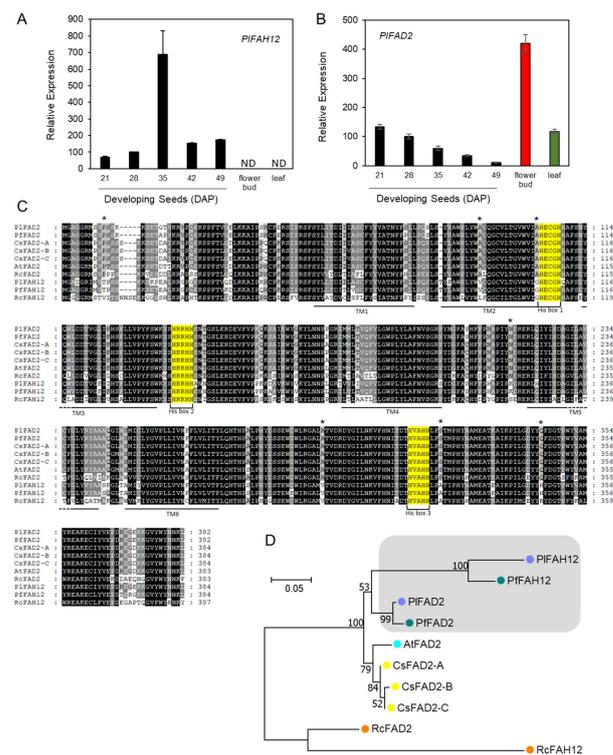
Figure 3. Characterization of PlFAH12s and PlFAD2s. Relative expression of *PlFAH12* (**A**) and *PlFAD2* (**B**) by qPCR. DAP, days after pollination. ND, not detected. (**C**) Amino acid alignment of FAH12 and FAD2 protein sequences. His-boxes were highlighted in yellow. Putative transmembrane domains (TM) were underlined. Amino acid residues identical in FAD2 but varied in FAH12 were marked with asterisks (*). (**D**) Phylogenetic tree of FAD2, FAH12 proteins from various species. Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. Genbank accession numbers are listed in Supplementary File 6, Table S7.

2.2. Mining and Characterization of Genes Involved in FA and TAG Biosynthesis during P. lindheimeri Seed Development

To understand HFA biosynthesis and metabolism in *P. lindheimeri*, we searched the contigs for genes involved in FA and TAG biosynthesis. Based on known Arabidopsis genes listed in the acyl-lipid metabolism database (https://www.arabidopsis.org/browse/genefamily/acyl_lipid.jsp), we identified 42 *P. lindheimeri* orthologous contigs involved in these processes (Supplementary File 3, Table S6). Thirty-nine of these contigs shared high sequence identity with Arabidopsis genes ranging from 85 to 98%, with two genes, *KASII* and *DGAT3*, sharing 70.8% and 83.2% identity, respectively (Supplementary File 3, Table S6). The sequence similarity between *P. lindheimeri* and Arabidopsis (Supplementary File 3, Table S6) is comparable to that between *P. fendleri* and Arabidopsis [58]. The very high sequence identity among Arabidopsis *P. fendleri* and *P. lindheimeri* genes suggests a similarly high degree of conservation of their gene functions in seeds that facilitates the translation of research findings from Arabidopsis or *P. lindheimeri* to lesquerella and other oilseeds for crop improvement.

To further characterize *P. lindheimeri* genes, we examined the expression profiles of genes in various organs, including developing seeds, flower buds and leaves. The genes involved in FA biosynthesis in plastids were expressed in all these organs, indicating their house-keeping functions. Since 20:1OH and 20:1OH-TAG synthesis occurred after 18:1 being synthesized in plastid and exported to the ER for hydroxylation and elongation, we analyzed in more detail for the ER-localized genes associated with 20:1OH formation and 20:1OH-TAG assembly.

2.3. De Novo Fatty Acid Biosynthesis in Plastids

We examined our collection of *P. lindheimeri* contigs for genes encoding the known steps of FA biosynthesis in plastids. De novo FAs are synthesized in plastids [21]. A plastidial pyruvate dehydrogenase complex (PDHC) catalyzes the conversion of pyruvate to acetyl-CoA, an initial substrate for FA synthesis (Supplementary File 4, Figure S3A). PDHC is a large multienzyme containing three components: E1 (pyruvate dehydrogenase or PDH, composed of E1α and E1ß subunits), E2 (dihydrolipoyl acyltransferase or DHLAT), and E3 (dihydrolipoamide dehydrogenase or LPD) [21]. In *P. lindheimeri*, we identified five subunit genes for *PlPDHC*: *PDH (E1α)*, *PDH (E1ß)*, *EMB3003 (E2)*, *LTA2(E2)* and *LPD1 (E3)* (Supplementary File 4, Figure S3B). Transcriptome analysis indicated that these five subunits were all expressed in developing seeds, leaves and flower buds (Supplementary File 4, Figure S3B). Once acetyl-CoA is synthesized, a heteromeric complex enzyme, acetyl-CoA carboxylase (ACC), catalyzes acetyl-CoA to form malonyl-CoA, and then a malonyl-CoA acyl-carrier protein (ACP) transferase (MCMT) further converts malonyl-CoA to malonyl-ACP (Supplementary File 4, Figure S4A). ACC is composed of four subunits; three of them, biotin carboxyl-carrier protein (BCCP), biotic carboxylase (BC) and alpha-carboxyltransferase (α-CT), are encoded from the nuclear genome, and beta-carboxyltransferase (β-CT) is encoded in the plastid genome [21,62]. *P lindheimeri* genes *PlBCCP1*, *PlBCCP2*, *Plα-CT* and *PlMCMT* were identified, and they exhibited a ubiquitous expression pattern in leaves, flower buds and developing seeds at various stages (Supplementary File 4, Figure S4B). To synthesize de novo 16- or 18-carbon FA, the initial condensation reaction of acetyl-CoA and malonyl-ACP is catalyzed by KAS isoform III (KASIII), yielding a four-carbon product (3-ketobutyrl-ACP). Subsequent condensations (up to 16:0-ACP) require KASI, whereas the final elongation of the 16-carbon palmitoyl-ACP to the 18-carbon stearoyl-ACP is catalyzed by KASII. In addition, each FA elongation cycle requires the participation of two reductases and a dehydrase. The 3-ketoacyl-ACP is first reduced by a 3-ketoacyl-ACP reductase (KAR), which is then subjected to dehydration by the enzyme hydroxyacyl-ACP dehydratase (HAD), and the enoyl-ACP thus obtained is finally reduced by the enzyme enoyl-ACP reductase (ENR) to form a saturated FA [21] (Supplementary File 4, Figure S5A). With the exception of *PlKASI*, we detected the expression of all the above genes in leaves, flower buds, and developing seeds from *P. lindheimeri* (Supplementary File 4, Figure S5B). Once FA is elongated to palmitic (16:0) or stearic acid (18:0), FA thioesterase B (FATB) cleaves acyl groups from 16:0-ACP or 18:0-ACP for the export from the plastid (Supplementary File 4, Figure S5A). In plastid, stearoyl-ACP desaturase (SAD) catalyzes the conversion of 18:0-ACP to 18:1-ACP, which is then cleaved by FATA (Supplementary File 4, Figure S5A). Arabidopsis has seven *SAD* family genes, including *FAB2* (At2g43710) and *DES6* (At4g30950) [21]. In *P. lindheimeri*, we identified *PlFATB*, *PlFATA*, *PlFAB2* and *PlDES6*, showing constitutive expression patterns in leaf, flower bud and developing seeds at various stages (Supplementary File 4, Figure S5B). Long-chain acyl-CoA synthase (LACS) esterifies free FAs to acyl-CoAs, a key activation step for FA to be utilized by many acyl-CoA-dependent acyltransferases. Arabidopsis plastid-localized LACS9 (At1g77590) and ER-localized LACS8 (At2g04350) have been reported [63,64]. In *P. lindheimeri* transcriptomes, we identified both *PlLACS8* and *PlLACS9* as showing spatial and temporal expression profiles ubiquitously expressed in all samples examined (Supplementary File 4, Figure S5B). The overall expression profiles of the genes involved in FA biosynthesis in plastids exhibited constitutive patterns, indicating their house-keeping functions in various organs.

2.4. Hydroxylase and ER-Associated FA Desaturases

Approximately 14 species from 10 families have been documented to accumulate HFAs in seed oils [28,65]. In *P. lindheimeri*, PlFAH12 is one of the key enzymes responsible for HFAs synthesis (Figure 1). FAH12 shares a high degree of sequence homology and similar reaction mechanisms with ER-associated microsomal FA desaturase 2 (FAD2), which exists in all plant species [28]. FAD2 introduces a double bond at the Δ12 position

of 18:1 on PC and converts it to 18:2 [37], whereas FAH12 introduces a hydroxy group at the same position of 18:1-PC, resulting in 18:1OH, which is structurally related to 18:2. In fact, lesquerella PfFAH12 is a bifunctional oleate 12-hydroxylase: desaturase [66]. We identified a full-length *PlFAH12* (c58040_g1_i2) and an almost full-length *PlFAD2* (1077 bp) (c58040_g1_i1) (Supplementary File 3, Table S6) from our transcriptomes. Gene expression analysis using quantitative PCR (qPCR) revealed that *PlFAH12* and *PlFAD2* had distinct expression profiles. *PlFAH12* had a bell-shaped pattern with a peak time at 35 DAP during seed development, and no expression was detected in the flower bud and leaf (Figure 3A). The increased expression of *PlFAH12* coincided with the increased synthesis and accumulation of HFA-containing TAG during *P. lindheimeri* seed development [15,30]. The overall expression profile of *PlFAH12* is similar to that of *PfFAH12* in lesquerella [58]. Our results are also consistent with castor *RcFAH12* expression, whose transcripts exist only in HFA-accumulating seed tissues, i.e., endosperm and embryo, and not in any other organs where no HFA is detected [49]. Therefore, organ-specific transcriptional activation of *FAH12* gene expression plays a key role in synthesizing HFA in seeds. During *P. lindheimeri* seed development, 18:2 is a major FA in young seeds (14–21 DAP), comprising 35.6–37.8% of total FA. When seeds develop to 28 DAP where HFA begins to accumulate, 18:2 drops to 22.9%, while during the remaining stages (35–56 DAP), 18:2 drops further to 8.6–5% [15]. The expression profile of *PlFAD2* showed a declining pattern during seed development (Figure 3B), coinciding with the decreased accumulation of 18:2 during *P. lindheimeri* seed development [15]. The *PlFAD2* transcript level in the leaf was similar to that of young seeds at 21 DAP, but in the flower bud it was three-fold higher (Figure 3B). To compare protein sequence similarities among *FAD2*-related genes, a full-length cDNA of *PlFAD2* was cloned by designing PCR primers based on the identified *PlFAD2* contig (Supplementary File 3, Table S6; Supplementary File 5, method of cDNA isolation) and *PfFAD2* sequence (Genbank ID, DQ518313), since *PlFAD2* and *PfFAD2* share a high degree of sequence homology (Supplementary File 6, Table S7). The alignment of protein sequences revealed that all FAD2s and FAH12s had three identical His-boxes and six similar transmembrane domains (Figure 3C). PlFAD2 also showed seven aa residues identical to all FAD2s, but these varied among FAH12s (Figure 3C) [13,65]. The roles of these seven aa have been assessed by using site-directed mutagenesis, and the results indicate that at least four aa substitutions are required to convert an FAD2 to an FAH12, and reciprocally, six aa substitutions are required to convert an FAH12 to an FAD2 [65]. Our isolation of full-length PlFAD2 provides a new candidate for further studying the factors that determine the activity of this group enzymes.

Arabidopsis, Camelina and *Physaria* are all from Brassicaceae, whereas castor belongs to Euphorbiaceae. Similar relationships were observed from the phylogenetic analysis of protein sequences which resulted in a tree with a big clade composed of all FAD2s and FAH12s from Brassicaceae, and a separate clade containing only castor RcFAD2 and RcFAH12 (Figure 3D). PlFAD2 shared the highest aa sequences similarity with PfFAD2 at 98.5%, followed by AtFAD2 and CsFAD2s at 93.8–94.1% (Supplementary File 6, Table S7), indicating that FAD2s were highly conserved among *Physaria*, Arabidopsis and Camelina. PlFAH12 shared 90.7% aa sequence similarity with PfFAH12, followed by RcFAH12 at 62.9% (Supplementary File 6, Table S7). It was clear that all *Physaria* sequences resided in a sub-clade (grey shaded clade, Figure 3D), showing that *Physaria* FAH12s and *Physaria* FAD2s shared a common ancestor (Figure 3D). Likewise, the castor RcFAH12 also shared a common ancestor with RcFAD2 (Figure 3D). Our results of phylogenetic analysis support the hypothesis that FAH12 has evolved from the FAD2 gene through gene duplication, and such events occur independently among species during evolution [28,65].

During seed development, *P. lindheimeri* accumulates moderate levels of 18:3 in a declining pattern, showing 7–11.6% in 14–21 DAP young seeds, 7.3–2.9% in 28–42 DAP mid-mature seeds, and 1.9–2.1% in 49–56 DAP dry seeds [15]. 18:3 is synthesized through the desaturation of 18:2 by FAD3 [38]. In the *P. lindheimeri* transcriptome, we identified two contigs, c29986_g5_i1 and c37413_g1_i1, encoding PlFAD3s (Supplementary File 3, Table S6). Contig c29986_g5_i1 was a partial coding sequence and shared sequence homology with

a previously identified lesquerella *PfFAD3-1* (MF611845), thus designated as *PlFAD3-1* (Supplementary File 3, Table S6). Using sequence information from *PfFAD3-1*, we cloned a full-length *PlFAD3-1* (Supplementary File 5). Contig c37413_g1_i1 shared 93.5% nucleotide identity with MF611846, a lesquerella ortholog *PfFAD3-2* [67], therefore designated as *PlFAD3-2* (Supplementary File 3, Table S6). Using qPCR, *PlFAD3-1* and *PlFAD3-2* showed distinct expression profiles (Figure 4A,B). During seed development, *PlFAD3-1* had a bell-shaped pattern peaking at 35 DAP, while *PlFAD3-2* had a declining pattern from 21–49 DAP (Figure 4A,B). The expression level of *PlFAD3-1* was 11-fold higher in the flower bud and 1.7-fold higher in the leaf than that of 35 DAP seeds; however, *PlFAD3-2* transcripts were not detected in flower buds or in leaves (Figure 4A,B). The protein sequences of PlFAD3-1 and PlFAD3-2 shared 98.3% and 96.8% similarity with PfFAD3-1 and PfFAD3-2, respectively (Supplementary File 6, Table S8). Lesquerella PfFAD3-1 and PfFAD3-2 have been extensively characterized in the Arabidopsis FAD3-deficient mutant (*fad3-2*), which had a reduced 18:3 (from 20.0% in the wild type to 1.6% in *fad3-2*) [67]. The over-expression of PfFAD3-1 in *fad3-2* recovers 18:3 up to 29.9%; however, overexpressing PfFAD3-2 fails to recover the loss of 18:3 in the *fad3-2* mutant [67]. The PfFAD3-2 aa sequence has been examined to explain the loss of its function [67]. Interestingly, those altered aa found in PfFAD3-2 were also exhibited in PlFAD3-2 (Figure 4C). Among them, two clusters of alternations occurred in the highly conserved regions of His-Boxes (Red asterisks in Figure 4C). The first is in the second His box, where PlFAD3-2 and PfFAD3-2 altered the conserved amino acid sequences from HRTHH [68] to HKIHH and HKLHH, respectively (Figure 4C). Thr contains an uncharged polar R group, which is a critical position in the second His box for FAD3 activity. A single aa change from HRTHH to HRIHH diminishes the desaturase activity or changes the desaturase to an enzyme with a different function [69,70]. The second cluster change was in the third His-box HYCGTHVIHH in *Physaria* FAD3-2s, which was different from the known conserved region of the third His box, HDIGTHVIHH, in other FAD3s (Figure 4C) [67]. Although it is possible that these major alterations in *Physaria* FAD3-2s result in changes of enzyme activity or substrate specificity, additional in vivo and in vitro experiments are required for testing such possibilities. Our previous study also shows that PfFAD3-1 belongs to common FAD3s, whereas PfFAD3-2 has similarities with a divergent FAD3 group from the Brassicaceae family [67]. Phylogenetic analysis grouped PlFAD3-1 and PlFAD3-2 with PfFAD3-1 and PfFAD3-2, respectively (Figure 4D). The identification and isolation of PlFAD3s from this study provides new candidates for investigating the biological function and molecular evolution process among plant FAD3s.

2.5. Lyso-PC Acyltransferase, FA-Condensing Enzyme, and Phospholipases A

The major HFA in *P. lindheimeri* is 20:1OH, comprising 85% of seed oil. As depicted in Figure 1, the 18:1 in cytosol can be acylated to PC in the ER by the forward reaction of LP-CAT, followed by hydroxylation through FAH12 to form 18:1OH-PC. The reverse reaction of LPCAT removes the 18:1OH from PC to the cytosol to be elongated to 20:1OH by KCS18. We did identify contig c56783_g1_i1 encoding a full-length *PlLPCAT* (Supplementary File 3, Table S6). The expression of *PlLPCAT* exhibited a constitutive pattern among the leaf, flower bud, and developing seeds at various stages (Figure 5A), indicating its important house-keeping function in the lipid metabolism of various organs. PlLPCAT shared 83.5% to 96.8% sequence similarities with Arabidopsis, Camelina and lesquerella, and 76.4% with RcLPCAT (Supplementary File 6, Table S9). The similarity among these LPCATs was also shown in a phylogenetic tree (Figure 5B) and protein alignment (Figure 5C). Studies of seven LPCATs from five different species, including Arabidopsis, lesquerella and castor, reveal that all seven LPCATs remove 18:1OH from PC at a three- to six-fold faster rate than acylating 18:1 to PC [24]. It is highly likely that PlLPCAT plays a critical role in removing the 18:1-OH from PC to the cytosol acyl-CoA pool.
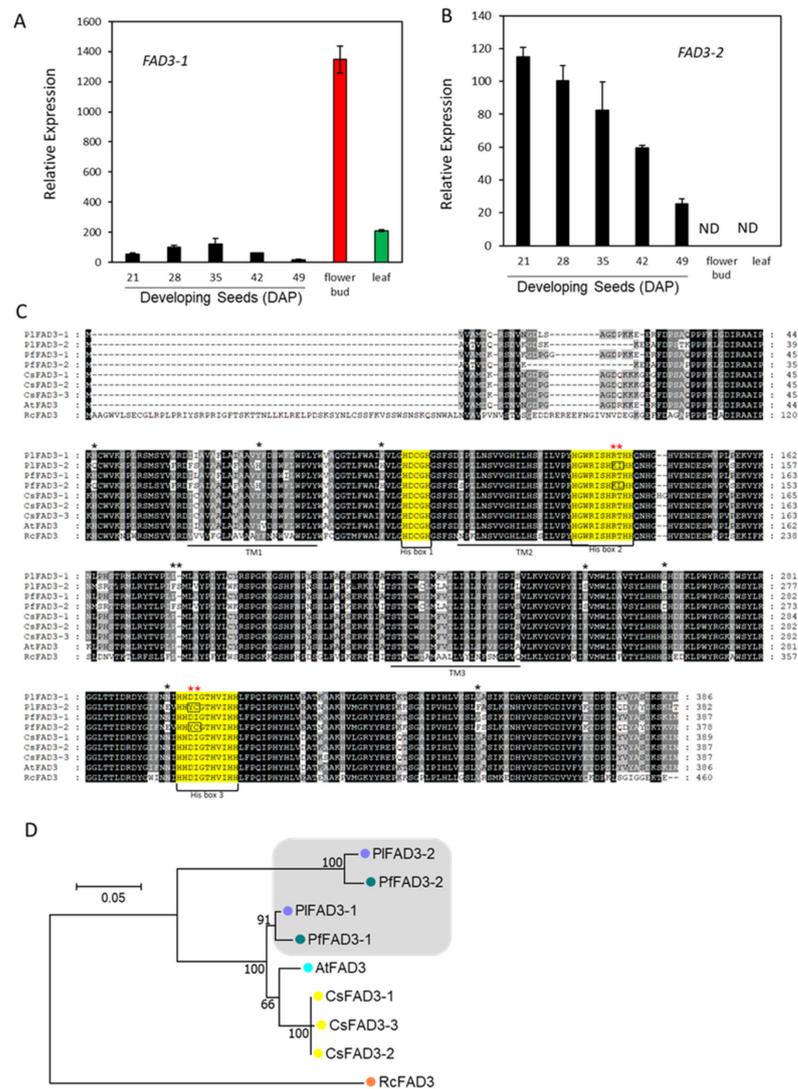
Figure 4. Characterization of FAD3s. Relative expression of *PlFAD3-1* (**A**) and *PlFAD3-2* (**B**) by qPCR. DAP, days after pollination. ND, not detected. (**C**) Amino acid alignment of FAD3 protein sequences. His-boxes were highlighted in yellow. Putative transmembrane domains (TM) were underlined. Asterisks indicate the aa that were conserved between PlFAD3-2 and PfFAD3-2, but that were different from the rest of the plant FAD3s. (**D**) Phylogenetic tree of FAD3 from various species. Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. Genbank accession numbers were listed in Supplementary File 6, Table S8.
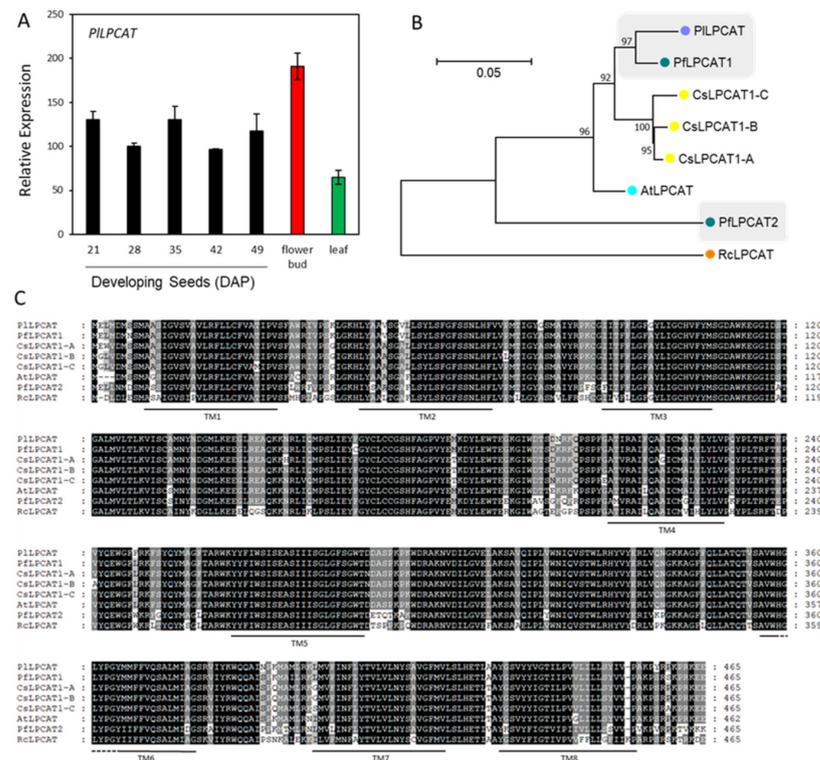
Figure 5. Characterization of LPCAT proteins. (**A**) Relative expression level of *PlLPCAT* in developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of LPCAT proteins in various species. *Physaria lindheimeri*, *Physaria fendleri* were highlighted with grey background. (**C**) Amino acid alignment of LPCAT protein sequence in *Physaria lindheimeri* (Pl), *Physaria fendleri* (Pf), *Arabidopsis thaliana* (At), *Camelina sativa* (Cs), and *Ricinus communis* (Rc). Putative transmembrane domains were underlined. Genbank accession numbers were listed in Supplementary File 6, Table S9.

PlKCS18 is directly responsible for elongating 18:1OH-CoA to 20:1-OH-CoA in *P. lindheimeri* (Figure 1). In our transcriptome, we identified a full-length *PlKSC18* (c61684_g3_i1) showing 82% nucleotide sequence identity with *PfKCS18* (Supplementary File 3, Table S6). Like *PlFAH12*, *PlKCS18* had a bell-shaped pattern peaked at 35 DAP during seed development, and no expression in the flower bud and leaf (Figure 6A). The nearly identical temporal and spatial expression patterns between *PlFAH12* and *PlKCS18* suggested that HFA synthesis and accumulation in *P. lindheimeri* were coordinately regulated in terms of gene expression, similarly to *PfFAH12* and *PfKCS18* in lesquerella [58]. A protein sequence comparison indicated that PlKCS18 shared 94.6% aa similarity with PfKCS18, 82.1–82.8% with AtKCS18 or CsKCS18, and 43.7% with RcKCS18 (Supplementary File 6, Table S10). The sequence similarities among KCS18 proteins were also shown in the phylogenetic tree (Figure 6B). Alignments of the aa of PlKCS18 exhibited two transmembrane domains and a region containing an active Cys site conserved among all KCS enzymes (Figure 6C) [71]. Although AtKCS18 (or AtFAE1 in some publication) exhibits activity upon elongating 18:1OH to 20:1OH [28,44,72], PfKCS18 (or LfKCS3 in some publication) is highly efficient in elongating 18:1OH in transgenic Arabidopsis [30] and Camelina [73]. The identification of PlKCS18 in this study provides another candidate for the further investigation of the mechanism of substrate recognition and specificity of the KCS family enzymes.
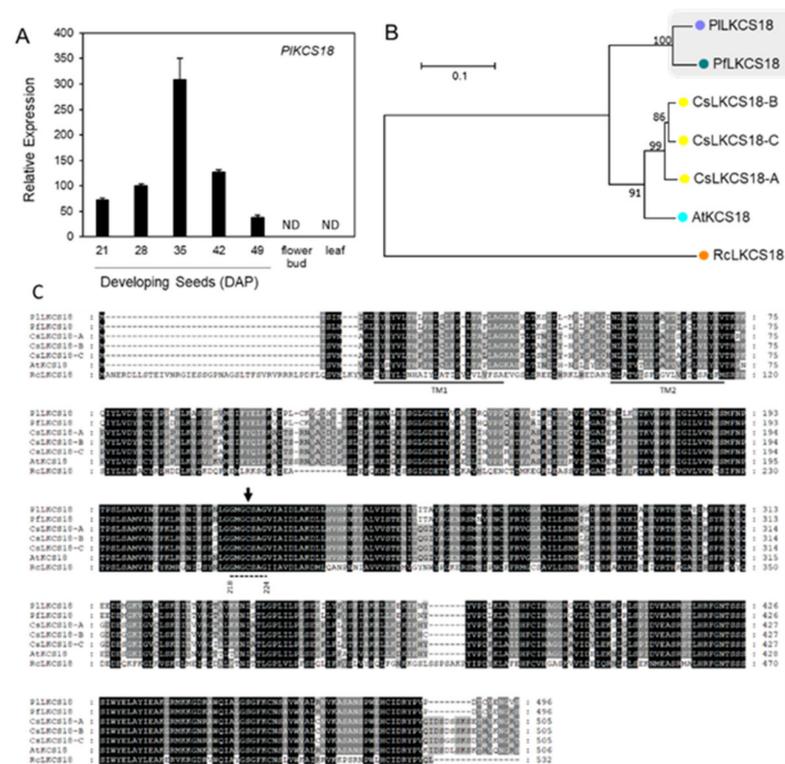
Figure 6. Characterization of KCS18 proteins. (**A**) Relative expression of *PlKCS18* by qRT-PCR. DAP, days after pollination. (**B**) Phylogenetic tree of KCS18 proteins from various species. Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. (**C**) Amino acid alignment of KCS18 proteins. Putative transmembrane domains were underlined. Conserved region (dash-underlined between Gly218 and Gly224) containing an active Cys site among all KCS enzymes [71] was indicated by the arrow. Genbank accession numbers were listed in Supplementary File 6, Table S10.

We depicted a PLA$_2$-tpye enzyme that could release HFA at the *sn*-2 position of PC to the cytosol acyl-CoA pool (Figure 1). Based on the published PLAs known to cleave 18:1OH at the *sn*-2 of PC, including *RcsPLA$_2$α* [29], *RcpPLAIIIβ* [51], *PfLCAT-PLA* or *RcLCAT-PLA* [52], we searched our transcriptomes for the orthologs in *P. lindheimeri*. We did not find *PlsPLA2α*, but we identified contigs c48215_g2_i1 and c55398_g1_i1 encoding full-length *PlpPLAIIIβ* and *PlLCAT-PLA*, respectively (Supplementary File 3, Table S6). Both transcripts were present ubiquitously in the leaf, flower bud, and developing seeds at 21, 35 and 42 DAP, indicating their house-keeping functions in *P. lindheimeri* (Supplementary File 4. Figure S6). The mechanisms of PLA-mediated HFA cleavage and the incorporation of the HFA into seed TAG remain largely unknown. The co-expression of *RcpPLAIIIβ* and *RcFHA12* in Arabidopsis results in reductions in HFA contents in both PC and TAG, and it is suggested that the HFAs released from PC may be directed to peroxisome for degradation via β-oxidation, and are thus unavailable for TAG synthesis [51]. However, the co-expression of *PfLCAT-PLA* or *RcLCAT-PLA* with *RcFAH12* in Arabidopsis does not change the total HFA content in TAG, but instead, it increases the proportion of HFA on the *sn*-1/3 positions of TAG at the expense of HFA at the *sn*-2 of TAG [52]. The efficient utilization of HFA cleaved by PfLCAT-PLA or RcLCAT-PLA is explained by the coordinated function of a LACS which preferentially activates HFA to HFA-CoAs, and the HFA-CoAs can be acylated to *sn*-1/3 TAG by GPAT9 or DGATs through the Kennedy pathway [52]. We identified *PlLACS8* and *PlLACS9* as constitutively expressed among different organs and during seed development (Supplementary File 3, Table S6, Supplementary File 4, Figure S5B). The role of PlLACS8 and PlLACS9 in HFA-CoA formation could be investi-

gated. It is also possible that there is an acyl-CoA-binding protein that specifically binds the HFA-CoA and transports the HFA-CoA to a correct subcellular compartment for TAG biosynthesis [52]. In *P. lindheimeri* seed cytosol, PlKCS18 is expected to efficiently elongate 18:1OH-CoA to 20:1OH-CoA; therefore, PlKCS18 should also be a key factor in this PLA-mediated generation of 20:1OH-CoA. The *PlpPLAIIIβ* and *PlLCAT-PLA* identified in this study provide candidates for the future investigation of the role of PLA-type enzymes in HFA-TAG synthesis in *P. lindheimeri*.

2.6. Kennedy Pathway for TAG Synthesis in the ER

The Kennedy pathway for TAG synthesis utilizes three acyl-CoA-dependent acyltransferases, GPAT, LPAT and DGAT, that sequentially acylate the *sn*-1- and *sn*-2- and then the *sn*-3-position of G3P with acyl-CoA (Figure 1, blue arrows). Since the synthesis of membrane glycerolipids also begins with the sequential acylation of the *sn*-1- and *sn*-2-positions of G3P, GPAT and LPAT are common to the synthesis of TAG and membrane glycerolipids. Arabidopsis GPAT9 (At5g60620) has been demonstrated to participate in seed oil and membrane lipids biosynthesis [74,75]. Using At5g60620 as a query, we identified a contig, c58129_g2_i1, encoding a full-length *PlGPAT9* (Supplementary File 3, Table S6). Gene expression profiling indicated that *PlGPAT9* was expressed in various organs, showing a bell-shaped temporal pattern peaking at 35 DAP (Figure 7A). The *PlGPAT9* transcript level in the leaf was similar to that of the 35 DAP seeds, and the level in the flower bud was two-fold higher than that in the leaf (Figure 7A). Phylogenetic analysis showed a close relationship between PlGPAT9 and PfGPAT9 (Figure 7B). The PlGPAT9 protein sequence shared 85.4% aa similarity with PfGPAT9, which has 1–51 aa deletion at the N-terminal (Figure 7C, Supplementary File 6, Table S11). When the first 51 aa were removed from PlGPAT9, the remaining aa similarity between PlGAPT9 and PfGPAT9 reached 98.8% (Supplementary File 6, Table S11). Compared with GPAT9s from Arabidopsis, Camelina, and castor, PlDGAT9 showed similarities from 88.1 to 95.8% (Supplementary File 6, Table S11). High protein sequence similarity among GPAT9 proteins indicates that GPAT9s have evolved conservatively. Most GPAT members have a broad acyl-CoA substrate specificity [25,76], but limited studies have been conducted to understand the role of GPAT9 in HFA-TAG biosynthesis. Over-expressing castor *RcGPAT9* alone does not change the HFAs content in transgenic Arabidopsis CL37 [48,50], which accumulates a background of 17% HFAs in seed TAG due to the expression of *RcFAH12* [28]. This is in contrast with over-expressing single castor gene, *RcLPAT2* [48,49], *RcDGAT2* [46], and *RcPDAT1A* or *RcPDAT1-2* [35,36], where increased HFAs production has been observed in CL37. However, when *RcGPAT9* is co-expressed with *RcLPAT2* and *RcDGAT2* in CL37, seed TAG contains more HFAs, and besides this, most HFAs are acylated in the *sn*-1/3 position compared with the line only co-expressing of *RcLPAT2* and *RcDGAT2* [50]. In addition, co-expressing *RcGPAT9* with *RcLPAT2* and *RcPDAT1A* increases not only total HFAs but also substantial amounts of tri-HFA-TAG (TAG species with all three *sn* positions occupied with HFAs) in CL37 seeds, whereas in the line co-expressing only *RcLPAT2* and *RcPDAT1A*, there is no tri-HFA-TAG detected [50]. These results suggest that RcGPAT9 plays an important role in acylating HFAs at the *sn*-1 position of G3P, resulting in *sn*-1-HFA-LPA, which facilitates the subsequent incorporation of *sn*-2 and *sn*-3 HFA into seed TAG by *sn*-2- and *sn*-3-specific acyltransferases. It would be interesting to investigate if PlGPAT9 plays similar role to RcGPAT9 in HFA-TAG synthesis. Sequence alignment showed that in the second transmembrane domain, one aa at the 105th position of PlDGAT9 was Thr (Figure 7C). The corresponding aa in lesquerella and castor GPAT9s was also Thr, but for Arabidopsis and Camelina it was Ala (Figure 7C). The further examination of 39 plant GPAT9s from 30 species and 11 families revealed that the Thr[105] was predominantly present in GPAT9s from Fabaceae, Euphorbiaceae, Ranunculaceae, Phrymaceae, Malvaceae, Solanaceae and Rutaceae, while the Ala was predominantly present in GPAT9s from Brassicaceae and Poaceae (Supplementary File 7, Figure S7A,B). It was intriguing that among 11 Brassicaceae GPAT9s, only *Physaria* GPAT9s had the Thr corresponding to the 105th position in PlDGAT9

(Supplementary File 7, Figure S7A,B). Thr and Ala are neutral aa, but Thr is polar, whereas Ala is nonpolar. Whether the difference in hydrophobicity between Thr and Ala affects GPAT9 substrate selection remains to be investigated.
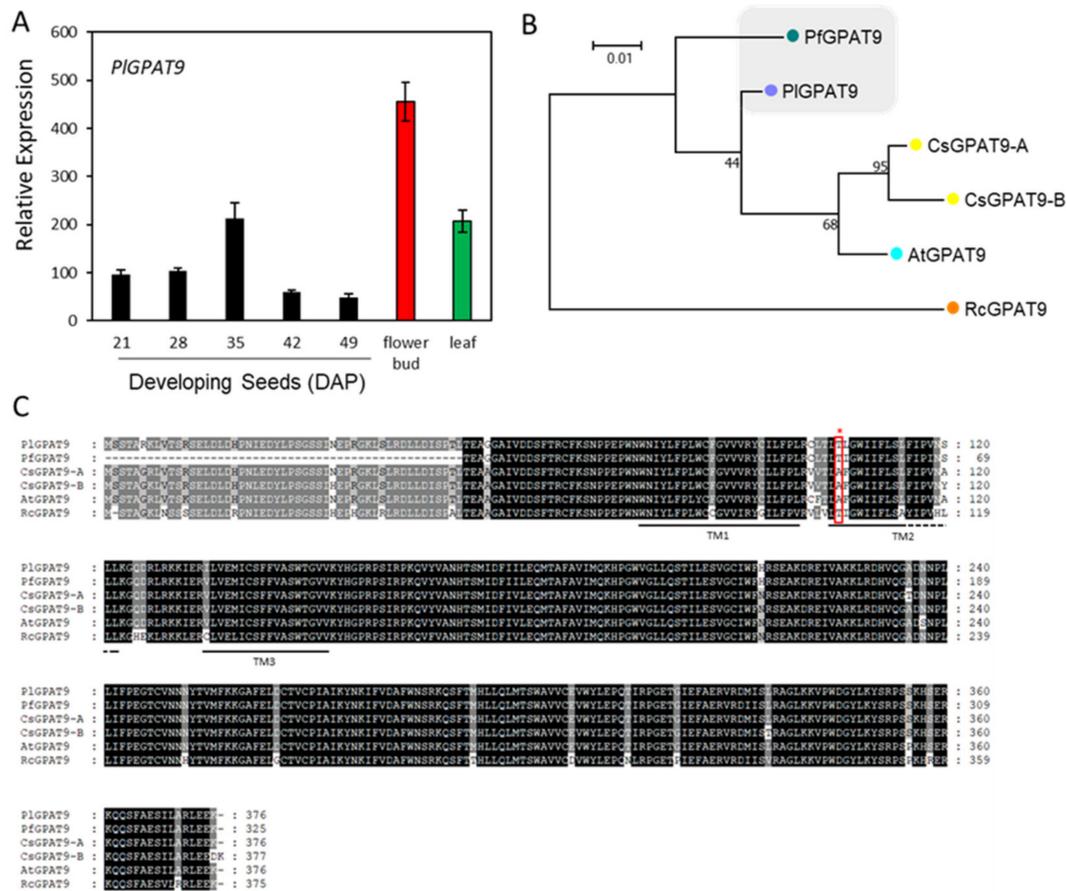

Figure 7. Characterization of GPAT9 proteins. (**A**) Relative expression level of *PlGPAT9* in developing seed, flower, and leaf tissues by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of GPAT9 proteins in various species. Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. PfDGAT9 and PlDGAT9 are highlighted in a grey background. (**C**) Amino acid alignment of GPAT9 protein sequences. Putative transmembrane domains are underlined. The red asterisk and box indicate the Thr[105] of PlGPAT9 conserved among Pl, Pf, and Rc, but the corresponding position was Ala in At and Cs. Genbank accession numbers are listed in Supplementary File 6, Table S11.

LPAT acylates the *sn*-2 position on LPA with an acyl-CoA to produce the PA precursor of membrane lipids and TAG. In most seed TAG synthesis pathways, this reaction is typically catalyzed by an LPAT2 that has high substrate specificity for C_{18} unsaturated FAs (i.e., 18:1, 18:2, and 18:3) [77]. Using Arabidopsis *LPAT2* (At3g57650) as query, we identified *PlLPAT2* (c53827_g2_i7, Supplementary File 3, Table S6). *PlLPAT2* expressed constitutively during seed development, and the level was 1.2-fold higher in the leaf and 5-fold higher in the flower bud than that of 21 DAP seeds (Figure 8A). Our spatial and temporal expression profiles of *PlLPAT2* indicate that *PlLPAT2* was likely involved in membrane lipids and storage TAG biosynthesis. The results of protein sequence analysis indicate that PlLPAT2 and PfLPAT2 share 95.2% aa similarity (Supplementary File 6, Table S12). Phylogenetic analysis grouped PlLPAT2 with PfPLPAT2 (Figure 8B). Despite sequence similarity, PlLPAT2 may differ from PfLPAT2 in acyl-CoA substrate specificity. Lesquerella TAGs contain ~60% 20:1OH and almost all of it is esterified to the *sn*-1 and *sn*-3 positions, while the *sn*-2 positions are exclusively occupied with C_{18} unsaturated FAs [54]. The *sn*-2 FA profile of lesquerella TAG implies that PfLPAT2 may act like a typical plant LPAT2 with a high substrate specificity to C_{18} unsaturated FAs [77]. In *P. lindheimeri*, seed TAGs contain 85%

20:1OH, with most of them having 20:1OH at all three *sn*-positions, similar to castor oil, where over 70% are tri-18:1OH-TAGs [53]. PlLPAT2 probably plays a key role in acylating 20:1-OH-LPA during seed development in *P. lindheimeri*. The alignment of protein sequences revealed that PlLPAT2 shared two aa, Asn[230] (neutral, polar) and Leu[255] (neutral, non-polar) with RcLPAT2 (indicated by red asterisks and boxes in Figure 8C), whereas the corresponding aa among PfLPAT2, AtLPAT2 and CsLPAT2 were Thr and Ser (both neutral, polar) (Figure 8C). To further examine the differences, 32 available LPAT2s from plants were aligned (Supplementary File 8, Figure S8A). Although the Leu[255] was present in 15 sequences representing 11 families, the combination of Asn[230] and Leu[255] was conserved in *P. lindheimeri* and the species from Euphorbiaceae, including castor, Tung tree (*Vernicia fordii*) and Jatropha (*Jatropha curcas*) (Supplementary File 8, Figure S8A,B). Like *P. lindheimeri*, castor and Tung tree seed oils contain over 80% unusual FAs, ricinoleic acid and eleostearic acid (conjugated FA, 18:3Δ9cis,11cis,13trans, or ESA), respectively [78]. The ESA in Tung oil is synthesized by an FADX, a conjugase X also derived from FAD2 [79]. Although the major FAs in Jatropha oil are 18:1 and 18:2, there are 1–2% of ESAs in seed TAG [80], suggesting that Jatropha seeds have a weak conjugase activity. The possibility exists that Jatropha acyltransferases, including JcLPAT2, acylate ESA into TAG. More experiments are required to investigate the substrate specificity of PlLPAT2, RcLPAT2, VfLPAT2 and JcLPAT2, and to examine the role of Asn[230] and Leu[255].
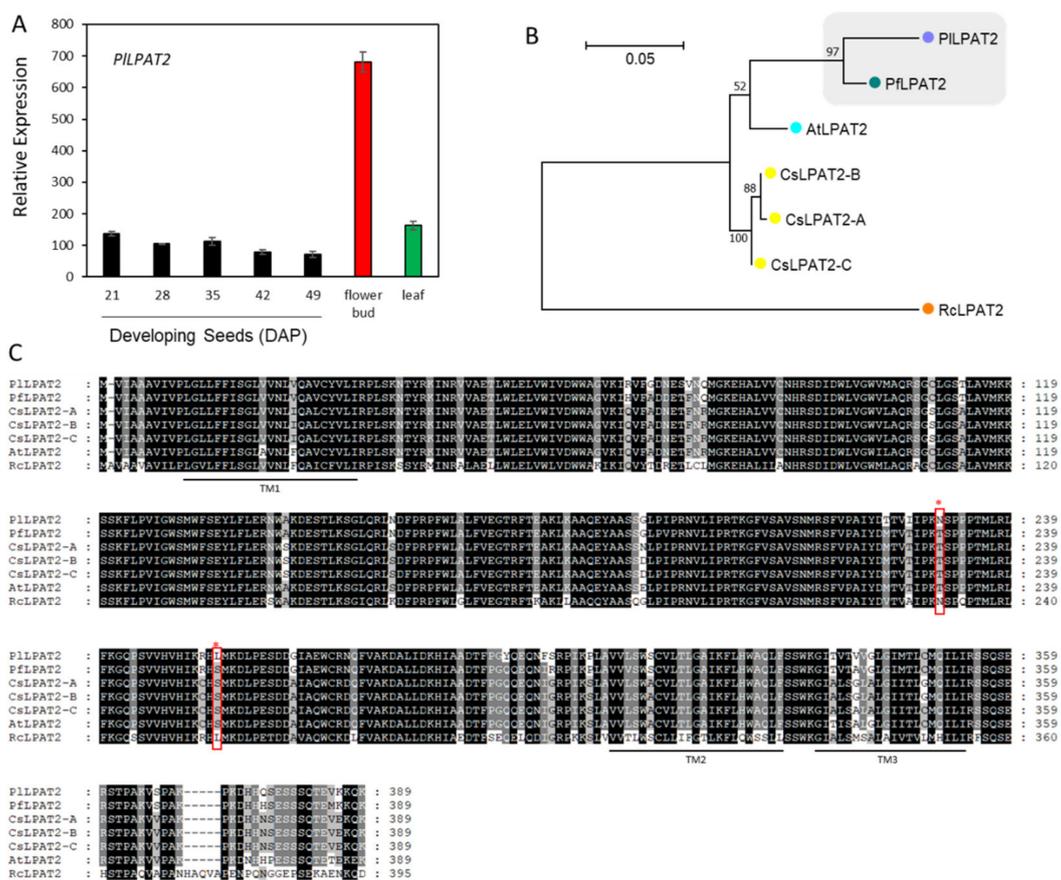

Figure 8. Characterization of LPAT2 proteins. (**A**) Relative expression level of *PlLPAT2* in the developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of LPAT2s in various species. Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. PlLPAT2 and PfLPAT2 are highlighted with a grey background. (**C**) Amino acid alignment of LPAT2 protein sequences. Putative transmembrane domains are underlined. The 230th and 255th amino acids of PlLPAT2 are indicated by red asterisks and boxes. Genbank accession numbers are listed in Supplementary File 6, Table S12.

Most plant species contain five Class-A LPATs, LPAT1 to LPAT5 [77,81,82]. LPAT1 is the only member localized in plastids showing high substrate specificity for palmitic acid (16:0)-CoA and 18:0-CoA [81,83]. The rest of the LPATs are ER-localized [82]. In some seed oils accumulating unusual FAs, there are additional Class-B LPATs. Class-B LPATs are often expressed in seeds and are associated with the acylation of unusual acyl-CoAs for TAGs, such as erucic acid (22:1) in Limnanthes [84], lauric acid (12:0) in Cocos [85], and caprylic (8:0) and capric acid (10:0) in Cuphea [86]. *LPAT-B* genes have been used to improve the unusual FA content in other oilseeds by genetic engineering [86–89]. Seven castor LPATs (RcLPATs) have been identified [49,90], and all of them encode functional enzymes [49]. In vitro enzyme assays have indicated that RcLPAT2 has a strong preference for 18:1OH-CoA when *sn*-1-18:1OH-LPA was used as an acyl acceptor; however, when *sn*-1-18:1-LPA was used as an acyl acceptor, RcLPAT2 selected 18:1-CoA [90]. The phenomenon of RcLPAT2 having different specificities toward acyl-donor and acyl-acceptor is supported by in vivo analysis of TAG species in transgenic lesquerella lines expressing *RcLPAT2* [55]. Interestingly, *RcLPAT-B* is expressed in diverse organs [49,90], and exhibits broad substrate specificity, with saturated CoAs (12:0–16:0) being the preferred substrates [90]. When all seven *RcLPATs* are expressed in an HFA-producing Arabidopsis CL37, *RcLPAT2, RcLPAT3A*, and *RcLPATB* increase the HFAs content in Arabidopsis [49]. Increasing evidence indicates that unusual FAs are acylated to TAG by variant LPATs [48,86,91,92]. In our transcriptomes, we identified *PlLPAT1* to *PlLPAT5* (Supplementary File 9, Figure S9). *PlLPAT2* transcript levels were much higher than those of *PlLPAT3, PlLPAT4* and *PlLPAT5* in the developing seed, leaf and flower bud (Supplementary File 9, Figure S9). We did not find orthologs of *RcLPAT*-B and *RcLPAT3A* in *P. linheimeri* transcriptomes. *PlLPAT2* likely is the isoform for acylating 20:1OH at the *sn*-2 position of 20:1OH-LPA. Further studies, such as in vitro enzyme assays and the in vivo transgenic expression of *PlLPAT2* in lesquerella or other oilseeds, would determine its role in HFA-TAG synthesis.

DGAT catalyzes the final step of TAG biosynthesis in the Kennedy pathway (Figure 1). There are three sequence-unrelated classes of DGATs reported in plants: membrane-bound DGAT1 and DGAT2, and cytosolic DGAT3 [93]. Using Arabidopsis *DGAT1* (At2g19450), *DGAT2* (At3g51520), and *DGAT3* (At1g48300) as queries, we identified contigs, c56288_g1_i1, c47089_g1_i3, and c56599_g1_i1, encoding PlDGAT1, PlDGAT2, and PlDGAT3, respectively (Supplementary File 3, Table S6). All three *PlDGATs* were ubiquity expressed in the developing seed, flower buds and leaf (Figure 9A,C,E), indicating that PlDGATs acted at least in part in the general TAG metabolism in *P. lindheimeri*. Phylogenetic analyses indicated that *P. lindheimeri* DGATs were closely related to their lesquerella orthologs (Figure 9B,D,F) (Figure 10A–C), showing aa similarity at 84.3% with PfDGAT1-1, 84.9% with PfDGAT2, and 89.4% with PfDGAT3, respectively (Supplementary File 6, Tables S13–S15). Among these three DGAT families, the role of DGAT3 in seed TAG synthesis has not yet been established [94], but it has been shown to be involved in recycling FAs in general lipid metabolism [94–96]. DGAT1 and DGAT2 have been intensively studied. Large amounts of evidence indicate that DGAT1 is a major determinant in seed oil accumulation in most species studied [93,97–99]. Recent studies using enzyme assays suggest that different DGAT families have different specificities toward acyl-donor and acyl-acceptor. Camelina CsDGAT1 excludes, whereas CsDGAT2 selects, acyl-acceptors containing only polyunsaturated FAs [100]. In *Brassica napus*, DGAT1 variants prefer acyl-donors with saturated (e.g., 16:0-CoA) and monounsaturated (e.g., 18:1-CoA) acyl-donors, whereas DGAT2 variants strongly select 18:3-CoA [101]. In plant seeds accumulating unusual FAs, DGAT2s are essential enzymes in acylating unusual *FAs* to the *sn*-3 position of DAG. For example, castor RcDGAT2 prefers 18:1OH to common FAs [46,102]. The *RcDGAT2* transcript is expressed 18-fold higher than that of *RcDGAT1* in developing seeds, supporting the role of RcDGAT2 in acylating 18:1OH-CoA to DAG [102]. Tung tree VfDGAT2 selects ESA [48,103]. It seems that tung tree DGAT1 and DGAT2 are localized in different regions of the ER, and consequently they may select different substrates [103]. *Brassica napus* DAGT2 variants, BnaA.DGAT2.d and BnaA.DGAT2.e, select not only 18:3 but also unusual erucic

acid (22:1) [101]. In some species, isoforms of DGAT1 and DGAT2 are both capable of acylating unusual FA, such as vernolic acid (cis-12,13-epoxy–cis-9-octadecenoic, 18:1>O) by VgDGAT1 and VgDGAT2 from ironweed (*Vernonia galamensis*) [104], and 22:1 by ChDGAT1 and ChDGAT2 from Crambe (*Crambe hispanica*) [105]. It is worth noting that *Euonymus alatus* diacylglycerol acetyltransferase (EaDAcT) acylates acetyl-CoA to the *sn*-3 position of DAG, and it belongs to the membrane-bound O-acyltransferase (MBOAT) super-family, which also includes DGAT1 [106,107]. Nevertheless, it is clear that castor RcDGAT2, but not RcDGAT1, acylates 18:1OH both in vitro [102] and in vivo [46].
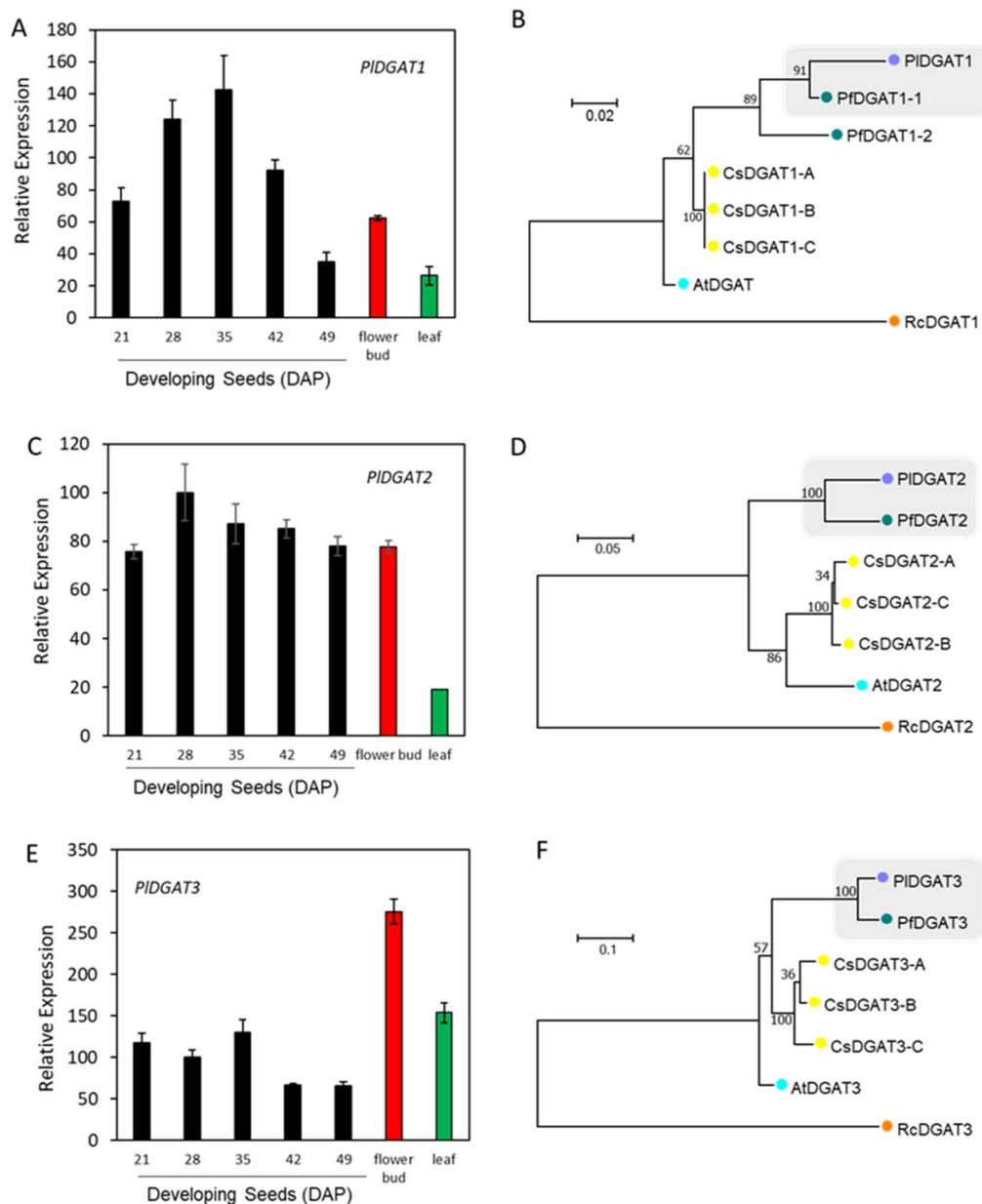

Figure 9. DGATs expression profiles and phylogenic trees. Relative expression level of *PlDGAT1* (**A**), *PlDGAT2* (**C**), and *PlDGA3* (**E**) in the developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. Phylogenetic tree of proteins for DGAT1 (**B**), DGAT2 (**D**) and DGAT3 (**F**) from *Physaria lindheimeri* (Pl), *Physaria fendleri* (Pf), *Arabidopsis thaliana* (At), *Camelina sativa* (Cs) and *Ricinus communis* (Rc). Pl and Pf genes are highlighted with a grey background. Genbank accession numbers are listed in Supplementary File 6, Tables S13–S15.
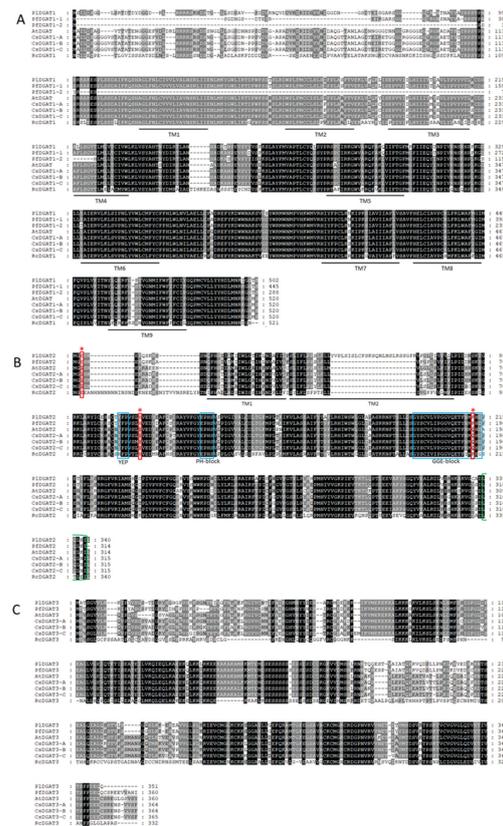
Figure 10. DGAT protein sequences alignment. Amino acid alignment of DGAT1 (**A**), DGAT2 (**B**), and DGAT3 (**C**). Putative transmembrane domains are underlined. The red asterisks and boxes indicate the amino acid corresponding to the 3rd, 115th, and 212th of PlDGAT2. The conserved YFP motif [108], PH-block and GGE-block [98] in plant DGAT2s are indicated in blue boxes. The modestly conserved ER retrieval motif in plant DGAT2s is in green brackets [103]. Abbreviation: Pl, *Physaria lindheimeri*; Pf, *Physaria fendleri*; At, *Arabidopsis thaliana*; Cs, *Camelina sativa*; Rc, *Ricinus communis*. Genbank accession numbers were listed in Supplementary File 6, Tables S13–S15.

Considering the similarities between 18:1OH and 20:1OH, it is likely that PlDGAT2 plays a key role in assembling seed TAG in *P. lindheimeri*. No detailed structural analysis of DGAT2 has been reported, although various motifs have been proposed to serve as important binding or active sites [93,98,99]. Protein sequence alignment of DGAT2s from non-HFA-producing plants, Arabidopsis and Camelina, as well as HFA-producing plants, castor, lesquerella and *P. lindheimeri* (Figure 10B), showed the following features: (1) castor RcDGAT2 had an extra 25 aa at the N-terminal of the protein, and nine Asn residues were located from the 6th to the 18th positions; (2) *P. lindheimeri* also had 25 extra aa (position 41th to 75th), but these were located in the second transmembrane domain; (3) three aa, Glu (polar, negative), His (polar, positive) and Lys (polar, positive), at the 3rd, 115th, and 212th positions of PlDGAT2, were conserved among HFA-producing plants, while the corresponding positions are Gly (nonpolar, neutral), Tyr (polar, neutral), and Gln (polar, neutral), respectively, in Arabidopsis and Camelina. Various conserved motifs have been proposed to serve as important binding or active sites for DGAT2s [93,98,99]. In vitro analysis reveals that the N-terminal region (the first 30–50 aa) of mouse DGAT2 or yeast DGAT2 is not essential for DGAT activity in vitro [108,109]. The alignment of 32 DGAT2s from different plant species showed that the N-terminals of DGAT2s were highly variable in length (Supplementary File 10, Figure S10A), thus it was unlikely that the extra 25 aa at the N-terminal of RcDGAT2 was critical for HFA-TAG assembly. The 3rd aa also varied among different DGAT2s capable of acylating unusual FAs (Supplementary File 10, Figure S10B). Although the Glu[3] is found in ironweed VgDGAT2 and Crambe ChDGAT2s,

which acylate unusual epoxy vernolic acid FA (18:1 > O) [104] and 22:1 [105], respectively, different aa Met[3] (non-polar, neutral) and Lys[3] are located in tung tree VfDGAT2 and rapeseed BnaA.DGAT2.d and BnaA.DGAT2.e (Supplementary File 10, Figure S10B), which also select unusual FAs. VfDGAT2 has been demonstrated to acylates ESA [48,103], whereas both BnaA.DGAT2.d and BnaA.DGAT2.e select 18:3 and 22:1 in high-erucic acid MAPLUS cultivar [101]. Thus the 3rd aa of DGAT2s is not critical for DGAT2s in selecting unusual FAs. Among the 38 plant DGAT2s, PlDGAT2 was the only one having the additional 25 aa in the second transmembrane domain (Supplementary File 10, Figure S10A). It is doubtful that these extra aa are essential for PlDGAT2 activity, but further studies are required to determine the relationship between the structure and function of PlDGAT2. Two motifs, YEP and PH-block (indicated by blue boxes in Figure 10B), are found to be essential for DGAT2 function in mouse and yeast [108,109]. The His[115] (red asterisk in Figure 10B) located between the YEP and PH-block is conserved among 23 DGAT2s from a broad range of families, including Poaceae, Euphorbiaceae, Celastraceae, Brassicaceae, and Fabaceae (Supplementary File 10, Figures S11A and S10B), thus it is unlikely that the His[115] is a specific aa for DGAT2s in selecting unusual FAs. The Lys[212] is located inside a conserved GGE-block (indicated by a blue box in Figure 10B) whose function has not yet been identified [98]. It is intriguing that 11 out of 35 of the DGAT2s have corresponding Lys (Supplementary File 10, Figure S10B), and eight of them are from species accumulating various unusual FAs, such as 18:1OH in castor, 20:1OH in lesquerella and *P. lindheimeri*, ESA in tung, and 22:1 in crambe. The remaining three have substrate preferences for polyunsaturated FA, including rapeseed BnaA.DGAT2.2b and BnaA.DGAT2.2c for selecting 18:3 [101], and Jatropha JcDGAT2 for selecting 18:2 [110]. Future experiments are required to determine the importance of the Lys[212] in PlDGAT2 and other DGAT2s in conducting enzyme activity.

2.7. Acyl-CoA-Independent and PC-Mediated Pathways for TAG Synthesis

Unlike DGAT, which utilizes the acyl-CoA as acyl donor, PDAT transfers the acyl moiety at the *sn*-2 position of PC to the *sn*-3 position of DAG, yielding TAG [34,111] (Figure 1). Using Arabidopsis *AtPDAT1* (AT5G13640.1) and *AtPDAT2* (AT3G44830.1) as queries, we identified c61686_g2_i4 and c62019_g1_i1 encoding *PlPDAT1* and *PlPDAT2*, respectively (Supplementary File 3, Table S6). Gene expression analysis showed that the *PlPDAT1* transcript level was high in the 35 DAP seed and flower bud, moderate in 21, 28 and 42 DAP seeds and leaf, but not detectable in 49 DAP desiccating seed (Figure 11A). *PlPDAT2* expressed ubiquitously during seed development as well as in the flower bud and leaf (Figure 12A). Almost all plant species contain at least PDAT1, thus PDAT1s have been characterized more extensively than PDAT2s. In Arabidopsis, AtPDAT1 plays an essential role in TAG biosynthesis in developing seeds and pollen [112–114] for maintaining membrane lipid homeostasis through directing the FA toward β-oxidation [115], and in stress responses [116]. Similar to AtPDAT1, Camelina CsPDAT1 also participates in stress responses [117]. The role of CsPDAT1 in controlling seed oil content remains controversial. A mutation in *CsPDAT1* reduces oil content [118], and either up-regulating or down-regulating *CsPDAT1* does not change seed oil content [118,119]. Studies indicate that CsPDAT1 has a substrate preference for polyunsaturated FAs, 18:2 and 18:3 [100,117–119]. Similar results are also reported for *AtPDAT1* [111], sunflower *PDAT1* [120] and flax PDATs (*LuPDAT1* and *LuPDAT5*) [121], with the exception of sesame (*Sesamum indicum*), where the over-expression of *SiPDAT1* in yeast H1246 results in increased 18:1 and 18:2 levels, but not 18:3 [122]. The role of *PlPDAT1* and *PlPDAT2* remains to be determined. Based on the spatial and temporal expression patterns of *PlPDAT1* (Figure 11A) and *PlPDAT2* (Figure 12A), *PlPDAT1* might play an important role in TAG synthesis, as it was expressed at higher levels in developing seeds than *PlPDAT2*. The ubiquitous expression patterns of *PlPDAT2* suggest its house-keeping function for maintaining normal plant growth and development. The results of protein sequence comparison showed that PlPDAT1 shared higher aa similarity with AtPDAT1 (92.1%) and CsPDAT1s (91.5–92.1%) than with RcPDAT1 (74.7%)

(Figure 11B,C) (Supplementary File 6, Table S16). PlPDAT2 displayed 86.2–86.2% aa similarities with AtPDAT2 and CsPDAT2, and 60.5% with RcPDAT2 (Figure 12B,C) (Supplementary File 6, Table S17). We excluded PfPDATs from protein sequence analysis, because PfPDATs were not full-length. Castor has two PDAT1s, RcPDAT1-1 (or RcPDAT1B) and RcPDAT1-2 (or RcPDAT1A) [35,36]. RcPDAT1-2 (RcPDAT1A), not RcPDAT1-1 (RcPDAT1B), is demonstrated to be the isoform having substrate preference for 18:1OH-PC, and it participates in TAG synthesis [35,36]. Castor *RcPDAT2*, Arabidopsis *AtPDAT2*, and flax *PDAT2s* (*LuPDAT3* and *LuPDAT6*) do not show an apparent function in TAG biosynthesis [35,36,111,121]. Considering the similarly high HFA contents in castor and *P. lindheimeri*, PlPDAT1 may play a similar role to RcLPAT1-2 (RcLPAT1A) in trans-acylating 18:1OH from *sn*-2 PC to *sn*-3 DAG for TAG formation. More detailed examinations of PDAT1 sequences revealed that PlPDAT1 had the typical conserved features of lecithin:cholesterol acyltransferase LCAT enzymes [123–125], including the 'lid' region between two cysteines with a tryptophan in the middle, a salt bridge, and a catalytic triad with a set of three coordinated amino acids S-D-H (or Ser-Asp-His) in the active site of acyltransferase [126] (Figure 11C). The Thr and Asp corresponding to the aa at the 390th and 415th positions of PlPDAT1 were conserved between PlPDAT1 and RcPDAT1-2, but varied among AtPDAT1, CsPDAT1s, and RcPDAT1-1 (Figure 11C). The further alignment of 32 plant PDAT1s showed that Thr^{390} was present in 10 PDATs from four different families—Brassicaceae, Malvaceae, Ranunculaceae, and Euphorbiaceae (Supplementary File 11, Figure S11A,B). Asp^{415} was conserved among PlPDAT1, RcPDAT1-2, and two *Citrus* PDAT1s (Supplementary File 11, Figure S11A,B). It seems unlikely that either Thu^{390} or Asp^{415} alone would be indispensable to determine the selectivity of HFAs by PlPDAT1 or RcPDAT1-2. However, Thu^{390} and Asp^{415} were close to each other (25 aa apart), located in the catalytic trial domain, and the combination of Thu^{390} and Asp^{415} was found only in PlPDAT1 and RcPDAT1-2, so it is possible that the combination of Thu^{390} and Asp^{415} in PlPDAT1 and RcPDAT1-2 could play a role in catalytic site geometry for determining the enzyme specificity of PlPDAT1 and RcPDAT1-2 in trans-acylating HFA-PC. Future experiments are required to verify this possibility. The isolation of *PlPDAT1* will be of value in studying the factors that determine the HFA-preferring substrate of plant PDATs.

PDCT catalyzes the interconversion between DAG and PC through the transfer of the phosphocholine group from PC to DAG. In *P. lindheimeri*, acyl editing and PC-DAG interconversion through PlLPCAT and PlPDCT, respectively, may co-contribute to the formation of TAGs with enriched 20:1OH (Figure 1). In *P. lindheimeri* transcriptomes, we identified one contig, c58276_g1_i1, showing 85% nucleotide identity with Arabidopsis *PDCT*, and thus designated it as *PlPDCT* (Supplementary File 3, Table S6). Gene expression analysis indicated that *PlPDCT* was expressed ubiquitously in developing seeds at various stages, in flower bud and in leaf tissues (Figure 13A). The PlPDCT protein sequence shared 73.1–86.4% with AtPDCT and CsPDCTs, and 61.8% with RcPDCT (Supplementary File 6, Table S18, Figure 13B). AtPDCT [41] and CsPDCT [100] are not able to use HFAs substrates, whereas in species where the seed oils contain high levels of unusual FAs, such as 18:1OH in castor and cyclopropane FAs (CPAs) in lychee (*Litchi chinensis*), their PDCTs have been shown to facilitate the efficient accumulation of HFAs in Arabidopsis [41] and CPAs in Camelina [127]. The alignment of the PlPDCT with AtPDCT, CsPDCT, and RcPDCT revealed a central core of high aa similarity (Figure 13C), showing a catalytic triad with a set of three coordinated amino acids, H-H-D (or His-His-Asp), in the conserved C2 and C3 domains of the lipid phosphatase/phosphotransferase family [128]. In addition, PlPDCT had Leu, Lys, and Leu at the 40th, 71th and 295th positions that were identical to the aa aligned with RcPDCT (Figure 13C), and the corresponding aa of AtPDCT and CsPDCTs were Arg, Thr, and Val, respectively (Figure 13C). The PfPDCT sequence was partial, so it was not included in the comparison. To further examine if the Leu^{40}, Lys^{71}, and Leu^{295} were unique to PlPDCT and RcPDCT, 28 plant PDCTs were aligned (Supplementary File 12, Figure S12A). The Leu^{40} and Leu^{295} were located in the N- and C-terminal regions that were highly variable in length and aa identity (Supplementary File 12, Figure S12A), thus

the Leu[40] and Leu[295] were considered as conserved aa for neither PlPDCT nor for RcPDCT. The Lys[71] was present in 9 out of 28 PDATs representing six different families (Supplementary File 12, Figure S12B). Therefore, it was also unlikely that the Lys[71] played an essential role for RcPDCT or for PlPDCT in HFA-PC-mediated TAG synthesis. Using a transgenic approach [41] and flux labeling analysis [129] should help to determine the role of PlPDCT in channeling HFAs into seed TAG. In addition to PDCT, plant phospholipase C (PLC) hydrolyzes membrane phospholipids, including PC, yielding DAG and phosphorylated head-groups. HFAs esterified at the *sn*-2 position of PCs are then converted to *sn*-2-HFA-DAGs to be utilized by PDAT and DGAT (Figure 1). A phospholipase C-like protein (*RcPLCL1*) from castor increases the amount of HFAs when co-expressed with *RcFAH12* in transgenic Camelina seed [47]. In our transcriptomes, we did not find an ortholog sequence encoding a putative *PlPLCL1*. Besides, two *Arabidopsis phospholipase D$_\zeta$* genes (*AtPLD$_{\zeta1}$* and *AtPLD$_{\zeta2}$*) have been recently isolated and found to increase seed TAG content in Camelina through PC to DAG conversion [130] (Figure 1). We found two contigs, c53747_g1_i1 and c55834_g1_i1, representing *PlPLD$_{\zeta1}$* and *PlPLD$_{\zeta2}$*, respectively (Supplementary File 3, Table S6). Both *PlPLD$_{\zeta1}$* and *PlPLD$_{\zeta2}$*, had constitutive spatial and temporal expression profiles, indicating their house-keeping function in *P. lindheimeri* (Supplementary File 13, Figure S13). The finding of *PlPLD$_{\zeta1}$* and *PlPLD$_{\zeta2}$*, provides targets for the future investigation of their roles in HFA accumulation in seed TAGs.
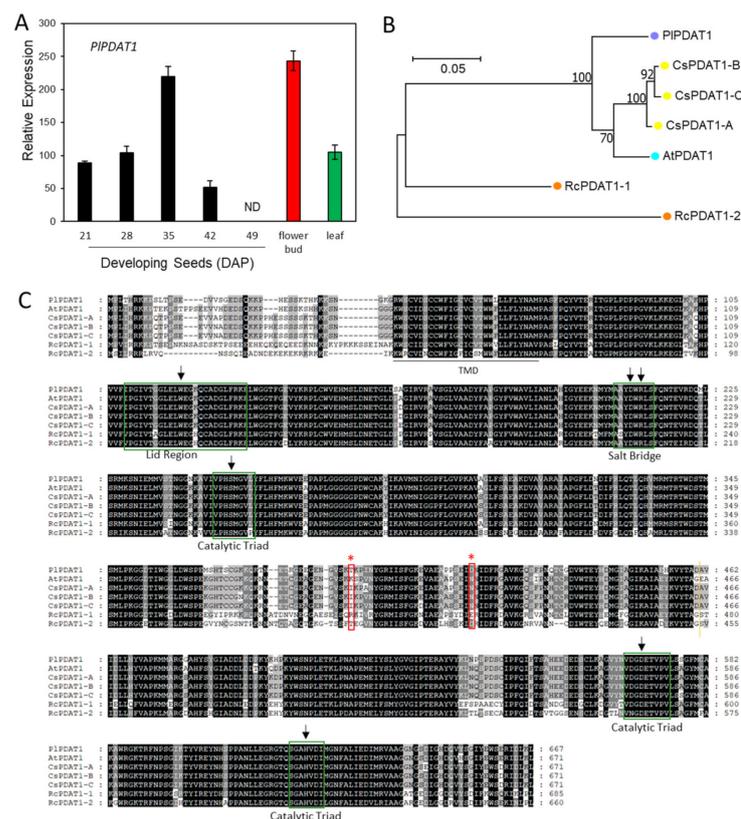

Figure 11. Characterization of PDAT1s in various species. (**A**) Relative expression level of *PlPDAT1* in the developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of PDAT1 proteins from *Physaria lindheimeri* (Pl), *Arabidopsis thaliana* (At), *Camelina sativa* (Cs), *Ricinus communis* (Rc). (**C**) Protein alignment of PDAT1s. Putative transmembrane domains are underlined. The Thr and Asp corresponding to the aa at the 390th and 415th positions of PlPDAT1 are indicated by red asterisks and boxes. Green boxes indicate lid region, salt bridge, and catalytic triad [126], where conserved Trp (W), Asp (D), Arg (R), and three coordinated amino acids Ser-Asp-His (or S-D-H) are indicated by arrows. Genbank accession numbers are listed in Supplementary File 6, Table S16.
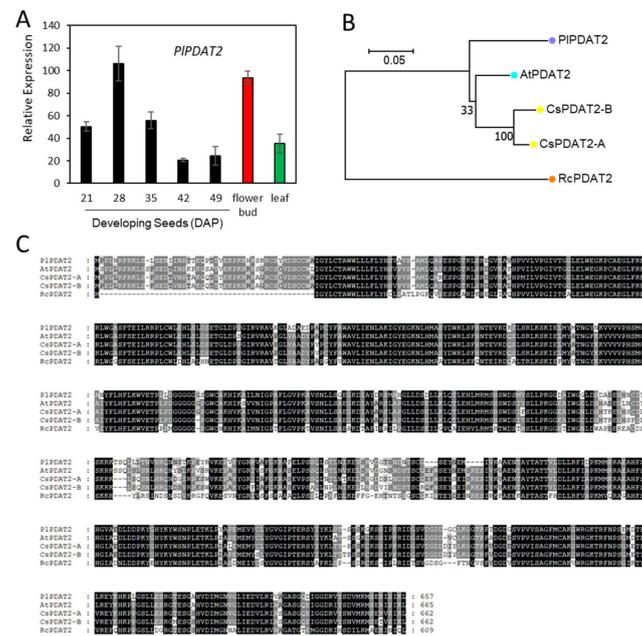
Figure 12. Characterization of PDAT2s in various species. (**A**) Relative expression level of *PlPDAT2* in the developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of PDAT2 proteins from *Physaria lindheimeri* (Pl), *Arabidopsis thaliana* (At), *Camelina sativa* (Cs), *Ricinus communis* (Rc). (**C**) Protein sequences alignment of PDAT2s. Genbank accession numbers are listed in Supplementary File 6, Table S17.
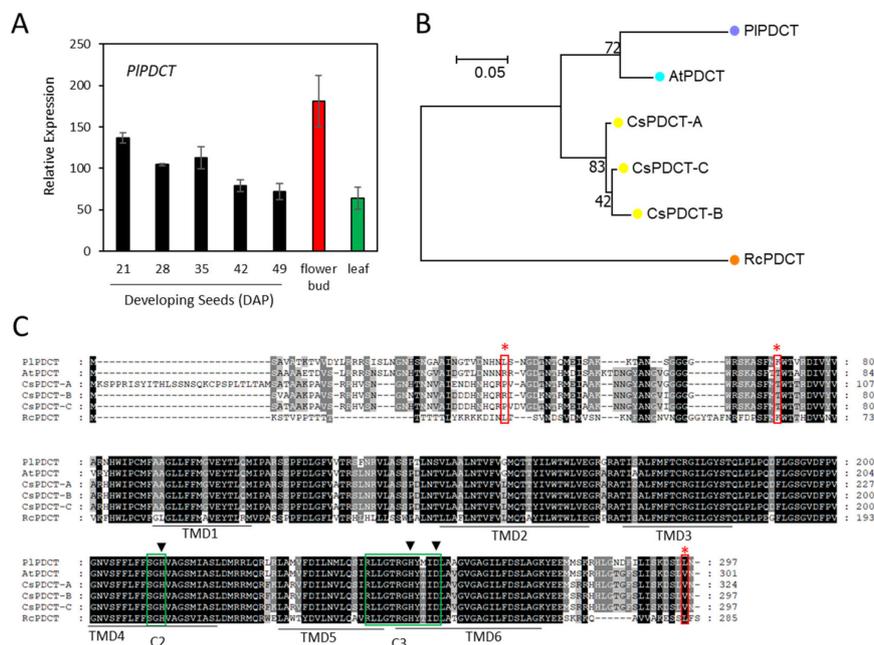

Figure 13. Characterization of PDCTs in various species. (**A**) Relative expression level of *PlPDCT* in the developing seed, flower, and leaf tissue by qRT-PCR. DAP; day after pollination. (**B**) Phylogenetic tree of PDCT proteins from *Physaria lindheimeri*, (Pl), *Arabidopsis thaliana* (At), *Camelina sativa* (Cs), *Ricinus communis* (Rc). (**C**) Protein alignment of PDCTs. Putative transmembrane domains were underlined. The red asterisks and boxes indicate the amino acids corresponding to the Leu, Lys, and Leu at the 40th, 71th and 295th positions of PlPDCT. Green boxes are the C2 and C3 domains, where a set of three coordinated amino acids, H-H-D (or His-His-Asp), of a catalytic triad [128] are indicated by arrows. Genbank accession numbers are listed in Supplementary File 6, Table S18.

3. Materials and Methods

3.1. Plant Materials and RNA Extraction

The *P. lindheimeri* seeds (accession number PI 643174) were obtained from United States Department of Agriculture, National Plant Germplasm System (http://www.ars-grin.gov/npgs/). Ten plants were grown in a greenhouse under a condition as described [131]. Developing seeds from 21, 35 and 42 days after pollination (DAP) were generated using the method as described [131]. Unopened mature flower buds, fully expended young leaves, and developing seeds were pooled from ten plants for each sample and frozen immediately in liquid nitrogen after harvest and stored at −80 °C. Total RNA was prepared using RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA).

3.2. cDNA Library Construction

To generate an mRNA-focused sequencing library, the mRNA from a total RNA sample was converted into a cDNA library using the reagents and instruction was provided by the Illumina® TruSeq™ RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). The first step in the workflow involves purifying the poly-A-containing mRNA molecules using poly-T oligo-attached magnetic beads. Following the purification, the mRNA is fragmented into small pieces using divalent cations under elevated temperatures. The cleaved RNA fragments are copied into first strand cDNA using reverse transcriptase and random primers. This is followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments then go through an end repair process, the addition of a single 'A' base, and then ligation of the adapters. The products are then purified and enriched with PCR to create the final cDNA library.

3.3. Sequencing and Sequence Quality Check

Illumina utilizes a unique "bridged" amplification reaction that occurs on the surface of the flow cell. A flow cell containing millions of unique clusters is loaded into the HiSeq 2000 (Illumina) for automated cycles of extension and imaging. Solexa's Sequencing-by-Synthesis (Illumina) utilizes four proprietary nucleotides possessing reversible fluorophore and termination properties. Each sequencing cycle occurs in the presence of all four nucleotides leading to higher accuracy than methods where only one nucleotide is present in the reaction mix at a time. This cycle is repeated, one base at a time, generating a series of images each representing a single base extension at a specific cluster.

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was applied to check the quality of raw sequence data obtained from high-throughput sequencing pipelines. Trimmomatic (0.32) was used to remove adapters which are undesirable for downstream application [132].

The sequence raw data from this study have been submitted to the NCBI Sequence Read Archive (SRA) (http://www.ncbi.nlm.nih.gov/), under the BioProject: PRJNA672105.

3.4. Transcriptome Analysis

RNA-seq data were subjected to de novo transcriptome assembly using Trinity (r20140717) [133], which also incorporates RSEM (RNA-Seq by Expectation-Maximization) algorithm (1.2.15) for abundance estimation. RSEM is particularly useful for the quantification of de novo transcriptome assemblies when a reference genome is not available, such as the case of *P. lindhemeri* in this study. The gene expression levels (transcripts or contigs levels) were calculated as fragments per kilobase of transcript per million mapped reads (FPKM) [61]. Contigs with an FPKM value of zero in at least one sample of all five samples were excluded to eliminate transcripts (contigs) with extremely low coverage and to minimize bias of statistical analysis. Logarithm (based 2) (fold-change) was calculated for filtered contigs to define significantly differentially expressed genes (DEGs). If the FPKM value was between 0 and 1, it was converted to a negative value during log transformation. Therefore, log2-based transformation was performed on Raw Signal (FPKM) +1 value. Furthermore, quantile normalization [134] was performed to reduce

the systematic/technique bias that may affect data presentation. After data scaling with logarithm and quantile normalization, a heatmap plot was generated using the "pheatmap" package in the R program [135].

3.5. Protein Sequence Analysis

NCBI ORF-finder (http://www.ncbi.nlm.nih.gov/orffinder/) was applied to find the open reading frame (ORF) of a nucleotide sequence from transcriptome analysis, and the resulting ORF or deduced amino acid sequence was confirmed by NCBI BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/) for its protein function. Putative transmembrane domains (TM) were predicted by the TMHMM web tool (http://www.cbs.dtu.dk/services/TMHMM/). Protein sequences were aligned using the ClustalW method of MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets [136]. Aligned sequences were converted in to fasta format for generating a phylogenetic tree. The maximum likelihood statistical method was used and replicated 100 times with bootstrapping. The numbers shown next to the branches indicate repeated times of replicated trees clustered together during the bootstrap test.

3.6. Quantitative PCR

Total RNA was reverse transcribed using the QuantiTect Reverse Transcription Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's directions. The resulting cDNA samples were used in PCR reactions. Standard PCR amplification reactions were carried out as described [58]. PCR product specificity was confirmed by melting-curve analysis and by electrophoresis on 4% agarose gel to ensure that PCR reactions were free of primer dimers and non-specific amplicons. Information of primer pairs and their PCR efficiencies is listed in Supplementary File 14 Table S19. The method of Pfaffl [137] was applied to calculate comparative expression levels between samples. The *P. lindheimeri* 18S gene was used as internal reference to normalize the relative amounts of RNAs for all samples.

4. Conclusions

To develop alternative oilseed crops for safe sources of HFAs production, we have built *P. lindheimeri* transcriptomes from developing seeds, leaf and flower bud samples that provide a valuable reference for discovering genes involved in the synthesis of TAGs enriched with HFAs. Clean reads were assembled into 129,633 contigs which represented 85,948 unique genes. GO analysis indicated that 60% of contigs matched genes with known functions. The comparison of DEGs revealed that the gene expression profiles in developing seeds were distinctly different from those of the leaf and flower bud. Using well-defined Arabidopsis lipid genes as queries, we identified 40 contigs, with 37 of them showing 78–100% high nucleotide sequence identity with Arabidopsis genes, which permits the translating of gene function from Arabidopsis to *P. lindheimeri*. Based on current knowledge, we proposed a pathways and characterized 16 essential *P. lindheimeri* lipid genes involved in 20:1OH-TAG biosynthesis in the seed. Manual annotation revealed that 14 out of these 16 essential genes encode full-length proteins, suggesting that the clean reads had good coverage of the assembled contigs. Detailed spatial and temporal expression profiling of these transcripts revealed various patterns. Two genes, *PlFAH12* and *PlKCS18,* directly responsible for 20:1OH synthesis were expressed only in developing seeds, and showed high peaks at 35 DAP, coinciding with the maximum 20:1OH accumulation rate during seed development. As such, the transcriptional up-regulation of *PlFAH12* and *PlKCS18* should play a key role in synthesizing 20:1OH. To understand how *P. lindheimeri* evolves to channel 20:1OH into seed TAG at a high level, the *P lindheimeri* protein sequences of enzymes in the Kennedy pathway (PlGPAT9, PlLAPT2, and PlDGAT2) and PC-mediated pathway (PlPDAT1 and PlPDCT) were compared with 28–39 correspondent orthologs from a published database. We identified the single or multiple aa conserved in each enzyme among species accumulating HFA-TAG. Our transcriptome data itself are insufficient to

deduce the regulatory mechanism of HFA synthesis and accumulation in seeds. Additional studies such as genome-wide proteotypes analyses [138] and genome-wide metabolic flux analyses [139] are also essential to enhance our understanding of HFA metabolism. Nevertheless, the information obtained from this study provides a valuable resource not only for studying the mechanisms of how HFAs are accumulated, but also for the biotechnological production of HFAs, including health beneficial SHFAs in existing oilseed crops.

Supplementary Materials: The following are available online at https://www.mdpi.com/1422-006 7/22/2/514/s1. Supplementary File 1: Figure S1. Gene ontology (GO) annotation of all assembled contigs (or transcripts). Figure S2. Changes in gene expression in developing seeds and flower buds using leaf sample as a control. Supplementary File 2: Table S1. Contig list with 4-fold or more expression changes in at least one of the five sample comparisons. Table S2. Contig list with more than 4-fold expression change in developing seed of 21 days based on leaf expression gene. Table S3. Contig list with more than 4-fold expression change in developing seed of 35 days based on leaf expression gene. Table S4. Contig list with more than 4-fold expression change in developing seed of 42 days based on leaf expression gene. Table S5. Contig list with more than 4-fold expression change in flower buds based on leaf expression gene. Supplementary File 3: Table S6. List of expressed genes involved in fatty acid and TAG biosynthesis in *P. lindheimeri*. Supplementary File 4: Figure S3. Synthesis of acetyl-CoA from pyruvate in plastids. Figure S4. Synthesis of malonyl-CoA from acetyl-CoA. Figure S5. Fatty acid biosynthesis in plastid. Figure S6. Transcript level of LPA-type A contigs in *P. lindhemeri* transcriptomes. Supplementary File 5: Method of cloning full length cDNAs of *PlFAD2* and *PlFAD3-1*. Supplementary File 6: Table S7. Protein sequence similarities among FAD2s and FAH12s. Table S8. Protein sequence similarities among FAD3s. Table S9. Protein sequence similarities among LPCATs. Table S10. Protein sequence similarities among KCS18s. Table S11. Protein sequence similarities among GPAT9s. Table S12. Protein sequence similarities among LPAT2s. Table S13. Protein sequence similarities among DGAT1s. Table S14. Protein sequence similarities among DGAT2. Table S15. Protein sequence similarities among DGAT3s. Table S16. Protein sequence similarities among PDAT1s. Table S17. Protein sequence similarities among PDAT2s Table S18. Protein sequence similarities among PDCTs. Supplementary File 7: Figure S7. GPAT9 protein analysis. Supplementary File 8: Figure S8. LPAT2 protein analysis. Supplementary File 9: Figure S9. Characterization of *PlLPAT* gene family. Supplementary File 10: Figure S10. DGAT2 protein analysis. Supplementary File 11: Figure S11. PDAT1 protein analysis. Supplementary File 12: Figure S12. PDCT protein analysis. Supplementary File 13: Figure S13. Transcript level *PlPLDζ1* and *PlPLDζ2* contigs in *P. lindhemeri* transcriptomes. Supplementary File 14: Table S19. Lipid genes, primer sequences, and PCR efficiencies.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Sequence raw data of Transcriptomes was registered in the NCBI Sequence Read Archive (SRA) (http://www.ncbi.nlm.nih.gov/), under the BioProject: PRJNA672105. All data was shared in supplementary file.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Caupin, H.J. Products from castor oil: Past, present and future. In *Lipid Technologies and Applications*; Gunstone, F.D., Padley, F.B., Eds.; Marcel Dekker: New York, NY, USA, 1997.
2. Roetheli, J.C. *Lesquerella as a Source of Hydroxy Fatty Acids for Industrial Products*; Growing Industrial Materials Series; USDA CSRS: Washington, DC, USA, 1991. Available online: https://www.ars.usda.gov/ARSUserFiles/53442010/NewCrops/Lesquerella/PDF/LesquerellaBooklet.pdf (accessed on 5 January 2021).
3. Chen, G.Q.; He, X.; McKeon, T.A. A Simple and Sensitive Assay for Distinguishing the Expression of Ricin andRicinus communisAgglutinin Genes in Developing Castor Seed (*R. communis* L.). *J. Agric. Food Chem.* **2005**, *53*, 2358–2361. [CrossRef] [PubMed]
4. Lord, M.J.; He, X.; Liao, L.P.; McKeon, T.A. Ricin: Mechanisms of cytotoxicity. *Toxicol. Rev.* **2003**, *22*, 53–64. [CrossRef] [PubMed]
5. Chen, G.Q.; He, X.; Liao, L.P.; McKeon, T.A. 2S albumin gene expression in castor plant (*Ricinus communis* L.). *J. Am. Oil Chem. Soc.* **2004**, *81*, 867–872. [CrossRef]
6. Machado, O.L.T.; Silva, J.G., Jr. An allergenic 2S storage protein from *Ricinus communis* seeds which is a part of the 2S albumin precursor predicted by C-DNA data. *Braz. J. Med. Biol. Res.* **1992**, *25*, 567–582.
7. Youle, R.J.; Huang, A.H.C. Albumin Storage Proteins in the Protein Bodies of Castor Bean. *Plant Physiol.* **1978**, *61*, 13–16. [CrossRef]
8. Chen, G.Q.; Lin, J.-T.; Lu, C. Hydroxy fatty acid synthesis and lipid gene expression during seed development in Lesquerella fendleri. *Ind. Crop Prod.* **2011**, *34*, 1286–1292. [CrossRef]
9. Hayes, D.G.; Carlson, K.D.; Kleiman, R. The isolation of hydroxy acids from lesquerella oil lipolysate by a saponification/extraction technique. *J. Am. Oil Chem. Soc.* **1996**, *73*, 1113–1119. [CrossRef]
10. Dierig, D.; Wang, G.; McCloskey, W.; Thorp, K.; Isbell, T.; Ray, D.; Foster, M. Lesquerella: New crop development and commercialization in the U.S. *Ind. Crop Prod.* **2011**, *34*, 1381–1385. [CrossRef]
11. Smith, C.R., Jr.; Wilson, T.L.; Miwa, T.K.; Zobel, H.; Lohmar, R.L.; Wolff, I.A. Lesquerolic acid. A new hydroxy acid from Lesquerella seed oil. *J. Org. Chem.* **1961**, *26*, 2903–2905. [CrossRef]
12. Von Cruz, M.V.; Dierig, D.A. Germplasm improvement to develop commercially viable lines of the new oilseed crop lesquerella. In *Industrial Crops: Breeding for Bioenergy and Bioproducts*; Springer: New York, NY, USA, 2015; pp. 315–334.
13. Dauk, M.; Lam, P.; Kunst, L.; Smith, M.A. A FAD2 homologue from Lesquerella lindheimeri has predominantly fatty acid hydroxylase activity. *Plant Sci.* **2007**, *173*, 43–49. [CrossRef]
14. Jenderek, M.; Dierig, D.; Isbell, T. Fatty-acid profile of Lesquerella germplasm in the National Plant Germplasm System collection. *Ind. Crop Prod.* **2009**, *29*, 154–164. [CrossRef]
15. Chen, G.Q.; Riiff, T.J.; Johnson, K.; Morales, E.; Kim, H.U.; Lee, K.-R.; Lin, J.-T. Seed development and hydroxy fatty acid biosynthesis in Physaria lindheimeri. *Ind. Crop Prod.* **2017**, *108*, 410–415. [CrossRef]
16. Dierig, D.A.; Tomasi, P.M.; Salywon, A.M.; Ray, D.T. Improvement in hydroxy fatty acid seed oil content and other traits from interspecific hybrids of three Lesquerella species: Lesquerella fendleri, L. pallida, and L. lindheimeri. *Euphytica* **2004**, *139*, 199–206. [CrossRef]
17. Chen, G.Q. Effective Reduction of Chimeric Tissue in Transgenics for the Stable Genetic Transformation of Lesquerella fendleri. *HortScience* **2011**, *46*, 86–90. [CrossRef]
18. Chen, G.Q.; Van Erp, H.; Martin-Moreno, J.; Johnson, K.; Morales, E.; Browse, J.; Eastmond, P.J.; Lin, J.-T. Expression of Castor LPAT2 Enhances Ricinoleic Acid Content at the *sn*-2 Position of Triacylglycerols in Lesquerella Seed. *Int. J. Mol. Sci.* **2016**, *17*, 507. [CrossRef] [PubMed]
19. Kokotou, M.G.; Kokotos, A.C.; Gkikas, D.; Mountanea, O.G.; Mantzourani, C.; Almutairi, A.; Lei, X.; Ramanadham, S.; Politis, P.K.; Kokotos, G. Saturated Hydroxy Fatty Acids Exhibit a Cell Growth Inhibitory Activity and Suppress the Cytokine-Induced β-Cell Apoptosis. *J. Med. Chem.* **2020**, *63*, 12666–12681. [CrossRef] [PubMed]
20. Kokotou, M.G.; Mantzourani, C.; Bourboula, A.; Mountanea, O.G.; Kokotos, G. A Liquid Chromatography-High Resolution Mass Spectrometry (LC-HRMS) Method for the Determination of Free Hydroxy Fatty Acids in Cow and Goat Milk. *Molecules* **2020**, *25*, 3947. [CrossRef]
21. Li-Beisson, Y.; Shorrosh, B.; Beisson, F.; Andersson, M.X.; Arondel, V.; Bates, P.D.; Baud, S.; Bird, D.; Debono, A.; Durrett, T.P.; et al. Acyl-Lipid Metabolism. *Arab. Book* **2013**, *11*, e0161. [CrossRef]
22. Bates, P.D.; Browse, J. The pathway of triacylglycerol synthesis through phosphatidylcholine in Arabidopsis produces a bottleneck for the accumulation of unusual fatty acids in transgenic seeds. *Plant J.* **2011**, *68*, 387–399. [CrossRef]
23. Lands, W.E. Lipid Metabolism. *Annu. Rev. Biochem.* **1965**, *34*, 313–346. [CrossRef]
24. Lager, I.; Lindberg Yilmaz, J.; Zhou, X.-R.; Jasieniecka, K.; Kazachkov, M.; Wang, P.; Zou, J.; Weselake, R.; Smith, M.A.; Bayon, S.; et al. Plant acyl-CoA:lysophosphatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. *J. Biol. Chem.* **2013**, *288*, 36902–36914. [CrossRef] [PubMed]
25. Bates, P.D. Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochim. et Biophys. Acta (BBA) Mol. Cell Biol. Lipids* **2016**, *1861*, 1214–1225. [CrossRef]

26. Bafor, M.; A Smith, M.; Jonsson, L.; Stobart, K.; Stymne, S. Ricinoleic acid biosynthesis and triacylglycerol assembly in microsomal preparations from developing castor-bean (Ricinus communis) endosperm. *Biochem. J.* **1991**, *280*, 507–514. [CrossRef] [PubMed]

27. Moreau, R.A.; Stumpf, P.K. Recent Studies of the Enzymic Synthesis of Ricinoleic Acid by Developing Castor Beans. *Plant Physiol.* **1981**, *67*, 672–676. [CrossRef] [PubMed]

28. Van De Loo, F.J.; Broun, P.; Turner, S.; Somerville, C. An oleate 12-hydroxylase from *Ricinus communis* L. is a fatty acyl desaturase homolog. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6743–6747. [CrossRef]

29. Bayon, S.; Chen, G.; Weselake, R.J.; Browse, J. A Small Phospholipase A2-α from Castor Catalyzes the Removal of Hydroxy Fatty Acids from Phosphatidylcholine in Transgenic Arabidopsis Seeds. *Plant Physiol.* **2015**, *167*, 1259–1270. [CrossRef]

30. Moon, H.; Smith, M.A.; Kunst, L. A condensing enzyme from the seeds of Lesquerella fendleri that specifically elongates hydroxy fatty acids. *Plant Physiol.* **2001**, *127*, 1635–1643. [CrossRef]

31. Weiss, S.B.; Kennedy, E.P. The enzymatic synthesis of triglycerides. *J. Am. Chem. Soc.* **1956**, *78*, 3550. [CrossRef]

32. Weiss, S.B.; Kennedy, E.P.; Kiyasu, J.Y. The enzymatic synthesis of triglycerides. *J. Biol. Chem.* **1960**, *235*, 40–44. [CrossRef]

33. Kennedy, E.P. Biosynthesis of complex lipids. *Fed. Proc.* **1961**, *20*, 934–940.

34. Dahlqvist, A.; Ståhl, U.; Lenman, M.; Banas, A.; Lee, M.; Sandager, L.; Ronne, H.; Stymne, S. Phospholipid:diacylglycerol acyltransferase: An enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6487–6492. [CrossRef] [PubMed]

35. Kim, H.U.; Lee, K.-R.; Go, Y.S.; Jung, J.H.; Suh, M.-C.; Kim, J.B. Endoplasmic Reticulum-Located PDAT1-2 from Castor Bean Enhances Hydroxy Fatty Acid Accumulation in Transgenic Plants. *Plant Cell Physiol.* **2011**, *52*, 983–993. [CrossRef] [PubMed]

36. Van Erp, H.; Bates, P.D.; Burgal, J.; Shockey, J. Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. *Plant Physiol.* **2011**, *155*, 683–693. [CrossRef] [PubMed]

37. Okuley, J.; Lightner, J.; Feldmann, K.; Yadav, N.; Lark, E. Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* **1994**, *6*, 147. [PubMed]

38. Arondel, V.; Lemieux, B.; Hwang, I.; Gibson, S.; Goodman, H.M.; Somerville, C.R. Map-based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis. *Science* **1992**, *258*, 1353. [CrossRef]

39. Yadav, N.S.; Wierzbicki, A.; Aegerter, M.; Caster, C.S.; Perez-Grau, L.; Kinney, A.J.; Hitz, W.D.; Booth, J.R., Jr.; Schweiger, B.; Stecca, K.L.; et al. Cloning of higher plant omega-3 fatty acid desaturases. *Plant Physiol.* **1993**, *103*, 467–476. [CrossRef]

40. Lu, C.; Xin, Z.; Ren, Z.; Miquel, M.; Browse, J. An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18837–18842. [CrossRef]

41. Hu, Z.; Ren, Z.; Lu, C. The Phosphatidylcholine Diacylglycerol Cholinephosphotransferase Is Required for Efficient Hydroxy Fatty Acid Accumulation in Transgenic Arabidopsis. *Plant Physiol.* **2012**, *158*, 1944–1954. [CrossRef]

42. Slack, C.R.; Campbell, L.C.; Browse, J.A.; Roughan, P.G. Some evidence for the reversibility of the cholinephosphotransferase-catalysed reaction in developing linseed cotyledons in vivo. *Biochim. Biophys. Acta (BBA) Lipids Lipid Metab.* **1983**, *754*, 10–20. [CrossRef]

43. Lee, K.-R.; Chen, G.Q.; Kim, H.U. Current progress towards the metabolic engineering of plant seed oil for hydroxy fatty acids production. *Plant Cell Rep.* **2015**, *34*, 603–615. [CrossRef]

44. Broun, P.; Somerville, C.; Morell, M.K.; Blennow, A.; Kosar-Hashemi, B.; Samuel, M.S. Accumulation of Ricinoleic, Lesquerolic, and Densipolic Acids in Seeds of Transgenic Arabidopsis Plants That Express a Fatty Acyl Hydroxylase cDNA from Castor Bean. *Plant Physiol.* **1997**, *113*, 933–942. [CrossRef] [PubMed]

45. Lu, C.; Fulda, M.; Wallis, J.G.; Browse, J. A high-throughput screen for genes from castor that boost hydroxy fatty acid accumulation in seed oils of transgenic Arabidopsis. *Plant J.* **2006**, *45*, 847–856. [CrossRef] [PubMed]

46. Burgal, J.; Shockey, J.; Lu, C.; Dyer, J.; Larson, T.; Graham, I.; Browse, J. Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil. *Plant Biotechnol. J.* **2008**, *6*, 819–831. [CrossRef] [PubMed]

47. Aryal, N.; Lu, C. A Phospholipase C-Like Protein From *Ricinus communis* Increases Hydroxy Fatty Acids Accumulation in Transgenic Seeds of Camelina sativa. *Front. Plant Sci.* **2018**, *9*, 1576. [CrossRef] [PubMed]

48. Shockey, J.; Lager, I.; Stymne, S.; Kotapati, H.K.; Sheffield, J.; Mason, C.; Bates, P.D. Specialized lysophosphatidic acid acyl-transferases contribute to unusual fatty acid accumulation in exotic Euphorbiaceae seed oils. *Planta* **2019**, *249*, 1285–1299. [CrossRef]

49. Kim, H.U.; Park, M.-E.; Lee, K.-R.; Suh, M.C.; Chen, G.Q. Variant castor lysophosphatidic acid acyltransferases acylate ricinoleic acid in seed oil. *Ind. Crop Prod.* **2020**, *150*, 112245. [CrossRef]

50. Lunn, D.; Wallis, J.G.; Browse, J. Tri-Hydroxy-Triacylglycerol Is Efficiently Produced by Position-Specific Castor Acyltransferases. *Plant Physiol.* **2019**, *179*, 1050–1063. [CrossRef]

51. Lin, Y.; Chen, G.; Mietkiewska, E.; Song, Z.; Caldo, K.M.P.; Singer, S.D.; Dyer, J.; Smith, M.; McKeon, T.; Weselake, R.J. Castor patatin-like phospholipase A IIIβ facilitates removal of hydroxy fatty acids from phosphatidylcholine in transgenic Arabidopsis seeds. *Plant Mol. Biol.* **2019**, *101*, 521–536. [CrossRef]

52. Xu, Y.; Caldo, K.M.P.; Singer, S.D.; Mietkiewska, E.; Greer, M.S.; Tian, B.; Dyer, J.M.; Smith, M.; Zhou, X.R.; Qiu, X.; et al. *Physaria fendleri* and *Ricinus communis* LCAT-like phospholipases selectively cleave hydroxy acyl chains from phosphatidylcholine. *Plant J.* **2020**. [CrossRef]

53. Lin, J.-T.; Turner, C.; Liao, L.P.; McKeon, T.A. Identification and Quantification of the Molecular Species of Acylglycerols in Castor Oil by HPLC Using ELSD. *J. Liq. Chromatogr. Relat. Technol.* **2003**, *26*, 773–780. [CrossRef]

54. Hayes, D.G.; Kleiman, R. A detailed triglyceride analysis ofLesquerella fendleri oil: Column chromatographic fractionation followed by supercritical fluid chromatography. *J. Am. Oil Chem. Soc.* **1996**, *73*, 267–269. [CrossRef]

55. Chen, G.Q.; Lin, J.-T.; Van Erp, H.; Johnson, K.; Lu, C. Regiobiochemical analysis reveals the role of castor LPAT2 in the accumulation of hydroxy fatty acids in transgenic lesquerella seeds. *Biocatal. Agric. Biotechnol.* **2020**, *25*, 101617. [CrossRef]

56. Chan, A.P.; Crabtree, J.; Zhao, Q.; Lorenzi, H.; Orvis, J.; Puiu, D.; Melake-Berhan, A.; Jones, K.M.; Redman, J.C.; Chen, G.; et al. Draft genome sequence of the oilseed species Ricinus communis. *Nat. Biotechnol.* **2010**, *28*, 951–956. [CrossRef] [PubMed]

57. Troncoso-Ponce, M.A.; Kilaru, A.; Cao, X.; Durrett, T.P.; Fan, J.; Jensen, J.K.; Thrower, N.A.; Pauly, M.; Wilkerson, C.; Ohlrogge, J.B. Comparative deep transcriptional profiling of four developing oilseeds. *Plant J.* **2011**, *68*, 1014–1027. [CrossRef] [PubMed]

58. Kim, H.U.; Chen, G.Q. Identification of hydroxy fatty acid and triacylglycerol metabolism-related genes in lesquerella through seed transcriptome analysis. *BMC Genom.* **2015**, *16*, 230. [CrossRef] [PubMed]

59. Horn, P.J.; Liu, J.; Cocuron, J.-C.; McGlew, K.; Thrower, N.A.; Larson, M.; Lu, C.; Alonso, A.P.; Ohlrogge, J. Identification of multiple lipid genes with modifications in expression and sequence associated with the evolution of hydroxy fatty acid accumulation in Physaria fendleri. *Plant J.* **2016**, *86*, 322–348. [CrossRef]

60. McGinnis, S.; Madden, T.L. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **2004**, *32*, W20–W25. [CrossRef]

61. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; Van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [CrossRef]

62. Li, X.; Ilarslan, H.; Brachova, L.; Qian, H.-R.; Li, L.; Che, P.; Wurtele, E.S.; Nikolau, B.J. Reverse-Genetic Analysis of the Two Biotin-Containing Subunit Genes of the Heteromeric Acetyl-Coenzyme A Carboxylase in Arabidopsis Indicates a Unidirectional Functional Redundancy. *Plant Physiol.* **2010**, *155*, 293–314. [CrossRef]

63. Schnurr, J.A.; Shockey, J.M.; De Boer, G.-J.; Browse, J.A. Fatty Acid Export from the Chloroplast. Molecular Characterization of a Major Plastidial Acyl-Coenzyme A Synthetase from Arabidopsis. *Plant Physiol.* **2002**, *129*, 1700–1709. [CrossRef]

64. Zhao, L.; Katavic, V.; Li, F.; Haughn, G.W.; Kunst, L. Insertional mutant analysis reveals that long-chain acyl-CoA synthetase 1 (LACS1), but not LACS8, functionally overlaps with LACS9 in Arabidopsis seed oil biosynthesis. *Plant J.* **2010**, *64*, 1048–1058. [CrossRef] [PubMed]

65. Broun, P.; Shanklin, J.; Whittle, E.; Somerville, C. Catalytic Plasticity of Fatty Acid Modification Enzymes Underlying Chemical Diversity of Plant Lipids. *Science* **1998**, *282*, 1315. [CrossRef] [PubMed]

66. Broun, P.; Boddupalli, S.; Somerville, C. A bifunctional oleate 12-hydroxylase: Desaturase fromLesquerella fendleri. *Plant J.* **1998**, *13*, 201–210. [CrossRef] [PubMed]

67. Lee, K.-R.; Kim, E.-H.; Jeon, I.; Lee, Y.; Chen, G.Q.; Kim, H.U. Lesquerella FAD3-1 gene is responsible for the biosynthesis of trienoic acid and dienoic hydroxy fatty acids in seed oil. *Ind. Crop Prod.* **2019**, *134*, 257–264. [CrossRef]

68. Los, D.A.; Murata, N. Structure and expression of fatty acid desaturases. *Biochim. Biophys. Acta (BBA) Lipids Lipid Metab.* **1998**, *1394*, 3–15. [CrossRef]

69. Broadwater, J.A.; Whittle, E.; Shanklin, J. Desaturation and hydroxylation. Residues 148 and 324 of Arabidopsis FAD2, in addition to substrate chain length, exert a major influence in partitioning of catalytic specificity. *J. Biol. Chem.* **2002**, *277*, 15613–15620. [CrossRef]

70. Yurchenko, O.P.; Park, S.; Ilut, D.C.; Inmon, J.J.; Millhollon, J.C.; Liechty, Z.; Page, J.T.; Jenks, M.A.; Chapman, K.D.; Udall, J.A.; et al. Genome-wide analysis of the omega-3 fatty acid desaturase gene family in Gossypium. *BMC Plant Biol.* **2014**, *14*, 312. [CrossRef]

71. Ghanevati, M.; Jaworski, J.G. Active-site residues of a plant membrane-bound fatty acid elongase beta-ketoacyl-CoA synthase, FAE1 KCS. *Biochim. Biophys. Acta* **2001**, *1530*, 77–85. [CrossRef]

72. Smith, M.A.; Moon, H.; Chowrira, G.; Kunst, L. Heterologous expression of a fatty acid hydroxylase gene in developing seeds of Arabidopsis thaliana. *Planta* **2003**, *217*, 507–516. [CrossRef]

73. Snapp, A.R.; Kang, J.; Qi, X.; Lu, C. A fatty acid condensing enzyme from Physaria fendleri increases hydroxy fatty acid accumulation in transgenic oilseeds of Camelina sativa. *Planta* **2014**, *240*, 599–610. [CrossRef]

74. Shockey, J.; Regmi, A.; Cotton, K.; Adhikari, N.D.; Browse, J.; Bates, P.D. Identification of Arabidopsis GPAT9 (At5g60620) as an Essential Gene Involved in Triacylglycerol Biosynthesis. *Plant Physiol.* **2016**, *170*, 163–179. [CrossRef] [PubMed]

75. Singer, S.D.; Chen, G.; Mietkiewska, E.; Tomasi, P.; Jayawardhane, K.; Dyer, J.M.; Weselake, R.J. Arabidopsis GPAT9 contributes to synthesis of intracellular glycerolipids but not surface lipids. *J. Exp. Bot.* **2016**, *67*, 4627–4638. [CrossRef] [PubMed]

76. Waschburger, E.; Kulcheski, F.R.; Veto, N.M.; Margis, R.; Margis-Pinheiro, M.; Turchetto-Zolet, A.C. Genome-wide analysis of the Glycerol-3-Phosphate Acyltransferase (GPAT) gene family reveals the evolution and diversification of plant GPATs. *Genet. Mol. Biol.* **2018**, *41* (Suppl. 1), 355–370. [CrossRef] [PubMed]

77. Frentzen, M. Acyltransferases from basic science to modified seed oils. *Lipid/Fett* **1998**, *100*, 161–166. [CrossRef]

78. Ohlrogge, J.B.; Thrower, N.; Mhaske, V.; Stymne, S.; Baxter, M.; Yang, W.; Liu, J.; Shaw, K.; Shorrosh, B.; Zhang, M.; et al. PlantFAdb: A resource for exploring hundreds of plant fatty acid structures synthesized by thousands of plants and their phylogenetic relationships. *Plant J.* **2018**, *96*, 1299–1308. [CrossRef]

79. Dyer, J.M.; Chapital, D.C.; Kuan, J.-C.W.; Mullen, R.T.; Turner, C.; McKeon, T.A.; Pepperman, A.B. Molecular Analysis of a Bifunctional Fatty Acid Conjugase/Desaturase from Tung. Implications for the Evolution of Plant Fatty Acid Diversity1. *Plant Physiol.* **2002**, *130*, 2027–2038. [CrossRef]

80. Jonas, M.; Ketlogetswe, C.; Gandure, J. Variation of Jatropha curcas seed oil content and fatty acid composition with fruit maturity stage. *Heliyon* **2020**, *6*, e03285. [CrossRef]

81. Kim, H.U.; Huang, A.H. Plastid Lysophosphatidyl Acyltransferase Is Essential for Embryo Development in Arabidopsis. *Plant Physiol.* **2004**, *134*, 1206–1216. [CrossRef]

82. Kim, H.U.; Li, Y.; Huang, A.H. Ubiquitous and Endoplasmic Reticulum–Located Lysophosphatidyl Acyltransferase, LPAT2, Is Essential for Female but Not Male Gametophyte Development in Arabidopsis. *Plant Cell* **2005**, *17*, 1073–1089. [CrossRef]

83. Bourgis, F.; Kader, J.-C.; Barret, P.; Renard, M.; Robinson, D.; Robinson, C.; Delseny, M.; Roscoe, T.J. A Plastidial Lysophosphatidic Acid Acyltransferase from Oilseed Rape. *Plant Physiol.* **1999**, *120*, 913–922. [CrossRef]

84. Phillips, B.E.; Smith, C.R.; Tallent, W.H. Glycerides ofLimnanthes douglasii seed oil. *Lipids* **1971**, *6*, 93–99. [CrossRef]

85. Balachandran, C.; Arumughan, C.; Mathew, A.G. Distribution of major chemical constituents and fatty acids in different regions of coconut endosperm. *J. Am. Oil Chem. Soc.* **1985**, *62*, 1583–1586. [CrossRef]

86. Kim, H.J.; Silva, J.E.; Iskandarov, U.; Andersson, M.; Cahoon, E.B.; Mockaitis, K.; Cahoon, E.B. Structurally divergent lysophosphatidic acid acyltransferases with high selectivity for saturated medium chain fatty acids from Cuphea seeds. *Plant J.* **2015**, *84*, 1021–1033. [CrossRef] [PubMed]

87. Brough, C.L.; Coventry, J.M.; Christie, W.W.; Kroon, J.T.M.; Brown, A.P.; Barsby, T.L.; Slabas, A.R. Towards the genetic engineering of triacylglycerols of defined fatty acid composition: Major changes in erucic acid content at the *sn*-2 position affected by the introduction of a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase from Limnanthes douglasii into oil seed rape. *Mol. Breed.* **1996**, *2*, 133–142.

88. Knutzon, D.S.; Hayes, T.R.; Wyrick, A.; Xiong, H.; Davies, H.M.; Voelker, T.A. Lysophosphatidic Acid Acyltransferase from Coconut Endosperm Mediates the Insertion of Laurate at the *sn*-2 Position of Triacylglycerols in Lauric Rapeseed Oil and Can Increase Total Laurate Levels. *Plant Physiol.* **1999**, *120*, 739–746. [CrossRef]

89. Lassner, M.W.; Levering, C.K.; Davies, H.M.; Knutzon, D.S. Lysophosphatidic Acid Acyltransferase from Meadowfoam Mediates Insertion of Erucic Acid at the *sn*-2 Position of Triacylglycerol in Transgenic Rapeseed Oil. *Plant Physiol.* **1995**, *109*, 1389–1394. [CrossRef]

90. Arroyo-Caro, J.M.; Chileh, T.; Kazachkov, M.; Zou, J.; Alonso, D.L.; García-Maroto, F. The multigene family of lysophosphatidate acyltransferase (LPAT)-related enzymes in Ricinus communis. Cloning and molecular characterization of two LPAT genes that are expressed in castor seeds. *Plant Sci.* **2013**, *199–200*, 29–40. [CrossRef]

91. Yu, X.-H.; Prakash, R.R.; Sweet, M.; Shanklin, J. Coexpressing Escherichia coli cyclopropane synthase with Sterculia foetida Lysophosphatidic acid acyltransferase enhances cyclopropane fatty acid accumulation. *Plant Physiol.* **2013**, *164*, 455–465. [CrossRef]

92. Yu, X.-H.; Cahoon, R.E.; Horn, P.J.; Shi, H.; Prakash, R.R.; Cai, Y.; Hearney, M.; Chapman, K.D.; Cahoon, E.B.; Schwender, J.; et al. Identification of bottlenecks in the accumulation of cyclic fatty acids in camelina seed oil. *Plant Biotechnol. J.* **2018**, *16*, 926–938. [CrossRef]

93. Xu, Y.; Caldo, K.M.P.; Pal-Nath, D.; Ozga, J.; Lemieux, M.J.; Weselake, R.J.; Chen, G. Properties and Biotechnological Applications of Acyl-CoA:diacylglycerol Acyltransferase and Phospholipid:diacylglycerol Acyltransferase from Terrestrial Plants and Microalgae. *Lipids* **2018**, *53*, 663–688. [CrossRef]

94. Aymé, L.; Arragain, S.; Canonge, M.; Baud, S.; Touati, N.; Bimai, O.; Jagic, F.; Louis-Mondésir, C.; Briozzo, P.; Fontecave, M.; et al. Arabidopsis thaliana DGAT3 is a [2Fe-2S] protein involved in TAG biosynthesis. *Sci. Rep.* **2018**, *8*, 17254. [CrossRef] [PubMed]

95. Hernández, M.L.; Whitehead, L.; He, Z.; Gazda, V.; Gilday, A.D.; Kozhevnikova, E.; Vaistij, F.E.; Larson, T.R.; Graham, I. A Cytosolic Acyltransferase Contributes to Triacylglycerol Synthesis in Sucrose-Rescued Arabidopsis Seed Oil Catabolism Mutants. *Plant Physiol.* **2012**, *160*, 215–225. [CrossRef] [PubMed]

96. Chi, X.; Hu, R.; Zhang, X.; Chen, M.; Chen, N.; Pan, L.; Wang, T.; Wang, M.; Yang, Z.; Wang, Q.; et al. Cloning and Functional Analysis of Three Diacylglycerol Acyltransferase Genes from Peanut (*Arachis hypogaea* L.). *PLoS ONE* **2014**, *9*, e105834. [CrossRef] [PubMed]

97. Liu, Q.; Siloto, R.M.; Lehner, R.; Stone, S.J.; Weselake, R.J. Acyl-CoA:diacylglycerol acyltransferase: Molecular biology, biochemistry and biotechnology. *Prog. Lipid Res.* **2012**, *51*, 350–377. [CrossRef] [PubMed]

98. Cao, H. Structure-Function Analysis of Diacylglycerol Acyltransferase Sequences from 70 Organisms. *BMC Res. Notes* **2011**, *4*, 249. [CrossRef]

99. Aznar-Moreno, J.A.; Denolf, P.; Van Audenhove, K.; De Bodt, S.; Engelen, S.; Fahy, D.; Wallis, J.G.; Browse, J. Type 1 diacylglycerol acyltransferases of Brassica napus preferentially incorporate oleic acid into triacylglycerol. *J. Exp. Bot.* **2015**, *66*, 6497–6506. [CrossRef]

100. Lager, I.; Jeppson, S.; Gippert, A.-L.; Feussner, I.; Stymne, S.; Marmon, S. Acyltransferases Regulate Oil Quality in Camelina sativa Through Both Acyl Donor and Acyl Acceptor Specificities. *Front. Plant Sci.* **2020**, *11*, 1144. [CrossRef]

101. Demski, K.; Jeppson, S.; Lager, I.; Misztak, A.; Jasieniecka-Gazarkiewicz, K.; Waleron, M.; Stymne, S.; Banaś, A. Isoforms of Acyl-CoA:Diacylglycerol Acyltransferase2 Differ Substantially in Their Specificities toward Erucic Acid1. *Plant Physiol.* **2019**, *181*, 1468–1479. [CrossRef]

102. Kroon, J.T.M.; Wei, W.; Simon, W.J.; Slabas, A.R. Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animals. *Phytochemistry* **2006**, *67*, 2541–2549. [CrossRef]

103. Shockey, J.M.; Gidda, S.K.; Chapital, D.C.; Kuan, J.-C.; Dhanoa, P.K.; Bland, J.M.; Rothstein, S.J.; Mullen, R.T.; Dyer, J.M. Tung Tree DGAT1 and DGAT2 Have Nonredundant Functions in Triacylglycerol Biosynthesis and Are Localized to Different Subdomains of the Endoplasmic Reticulum. *Plant Cell* **2006**, *18*, 2294–2313. [CrossRef]

104. Li, R.; Yu, K.; Jamboonsri, W. DGAT1, DGAT2 and PDAT Expression in Seeds and Other Tissues of Epoxy and Hydroxy Fatty Acid Accumulating Plants. *Lipids* **2010**, *45*, 145–157. [CrossRef] [PubMed]

105. Jeppson, S.; Demski, K.; Carlsson, A.S.; Zhu, L.H.; Banaś, A.; Stymne, S.; Lager, I. Crambe hispanica Subsp. abyssinica Diacylglycerol Acyltransferase Specificities Towards Diacylglycerols and Acyl-CoA Reveal Combinatorial Effects That Greatly Affect Enzymatic Activity and Specificity. *Front. Plant Sci.* **2019**, *10*, 1442. [CrossRef] [PubMed]

106. Durrett, T.P.; McClosky, D.D.; Tumaney, A.W.; Elzinga, D.A.; Ohlrogge, J.; Pollard, M. A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in Euonymus and transgenic seeds. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9464–9469. [CrossRef] [PubMed]

107. Liu, J.; Rice, A.; McGlew, K.; Shaw, V.; Park, H.; Clemente, T.E.; Pollard, M.; Ohlrogge, J.; Durrett, T.P. Metabolic engineering of oilseed crops to produce high levels of novel acetyl glyceride oils with reduced viscosity, freezing point and calorific value. *Plant Biotechnol. J.* **2015**, *13*, 858–865. [CrossRef] [PubMed]

108. Liu, Q.; Siloto, R.M.; Snyder, C.L.; Weselake, R.J. Functional and topological analysis of yeast acyl-CoA:diacylglycerol acyltransferase 2, an endoplasmic reticulum enzyme essential for triacylglycerol biosynthesis. *J. Biol. Chem.* **2011**, *286*, 13115–13126. [CrossRef]

109. Stone, S.; Levin, M.; Farese, R.V. Membrane Topology and Identification of Key Functional Amino Acid Residues of Murine Acyl-CoA:Diacylglycerol Acyltransferase-2*. *J. Biol. Chem.* **2006**, *281*, 40273–40282. [CrossRef] [PubMed]

110. Xu, R.; Yang, T.; Wang, R.; Liu, A. Characterisation of DGAT1 and DGAT2 from Jatropha curcas and their functions in storage lipid biosynthesis. *Funct. Plant Biol.* **2014**, *41*, 321–329. [CrossRef]

111. Stahl, U.; Carlsson, A.S.; Lenman, M.; Dahlqvist, A.; Huang, B.; Banaś, W.; Banaś, A.; Stymne, S. Cloning and Functional Characterization of a Phospholipid:Diacylglycerol Acyltransferase from Arabidopsis. *Plant Physiol.* **2004**, *135*, 1324–1335. [CrossRef]

112. Zhang, M.; Fan, J.; Taylor, D.C.; Ohlrogge, J.B. DGAT1 and PDAT1 acyltransferases have overlapping functions in Arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *Plant Cell* **2009**, *21*, 3885–3901. [CrossRef]

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

114. Fan, J.; Yan, C.; Zhang, X.; Xu, C. Dual role for phospholipid: Diacylglycerol acyltransferase: Enhancing fatty acid synthesis and diverting fatty acids from membrane lipids to triacylglycerol in Arabidopsis leaves. *Plant Cell* **2013**, *25*, 3506–3518. [CrossRef] [PubMed]

115. Fan, J.; Yan, C.; Roston, R.; Shanklin, J.; Xu, C. Arabidopsis lipins, PDAT1 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward β-oxidation, thereby maintaining membrane lipid homeostasis. *Plant Cell* **2014**, *26*, 4119–4134. [CrossRef] [PubMed]

116. Mueller, S.P.; Unger, M.; Guender, L.; Fekete, A.; Mueller, M.J. Phospholipid:Diacylglycerol Acyltransferase-Mediated Triacylglyerol Synthesis Augments Basal Thermotolerance. *Plant Physiol.* **2017**, *175*, 486–497. [CrossRef] [PubMed]

117. Yuan, L.; Mao, X.; Zhao, K.; Ji, X.; Ji, C.; Xue, J.; Li, R. Characterisation of phospholipid: Diacylglycerol acyltransferases (PDATs) from Camelina sativa and their roles in stress responses. *Biol. Open* **2017**, *6*, 1024–1034. [CrossRef] [PubMed]

118. Aznar-Moreno, J.A.; Durrett, T.P. Simultaneous Targeting of Multiple Gene Homeologs to Alter Seed Oil Production in Camelina sativa. *Plant Cell Physiol.* **2017**, *58*, 1260–1267. [CrossRef] [PubMed]

119. Marmon, S.; Sturtevant, D.; Herrfurth, C.; Chapman, K.D.; Stymne, S.; Feussner, I. Two Acyltransferases Contribute Differently to Linolenic Acid Levels in Seed Oil. *Plant Physiol.* **2017**, *173*, 2081–2095. [CrossRef]

120. Banaś, W.; Garcia, A.S.; Banaś, A.; Stymne, S. Activities of acyl-CoA:diacylglycerol acyltransferase (DGAT) and phospholipid:diacylglycerol acyltransferase (PDAT) in microsomal preparations of developing sunflower and safflower seeds. *Planta* **2013**, *237*, 1627–1636. [CrossRef]

121. Pan, X.; Siloto, R.M.P.; Wickramarathna, A.D.; Mietkiewska, E.; Weselake, R.J. Identification of a Pair of Phospholipid:Diacylglycerol Acyltransferases from Developing Flax (*Linum usitatissimum* L.) Seed Catalyzing the Selective Production of Trilinolenin. *J. Biol. Chem.* **2013**, *288*, 24173–24188. [CrossRef]

122. Chellamuthu, M.; Kumaresan, K.; Subramanian, S.; Muthumanickam, H. Functional Analysis of Sesame Diacylglycerol Acyltransferase and Phospholipid: Diacylglycerol Acyltransferase Genes Using In Silico and In Vitro Approaches. *Plant Mol. Biol. Rep.* **2019**, *37*, 146–156. [CrossRef]

123. Peelman, F.; Vinaimont, N.; Vanloo, B.; Labeur, C.; Rosseneu, M.; Verhee, A.; Verschelde, J.L.; Vandekerckhove, J.; Tavernier, J.; Seguret-Mace, S.; et al. A proposed architecture for lecithin cholesterol acyl transferase (LCAT): Identification of the catalytic triad and molecular modeling. *Protein Sci.* **1998**, *7*, 587–599. [CrossRef]

124. Glukhova, A.; Hinkovska-Galcheva, V.; Kelly, R.; Abe, A.; Shayman, J.A.; Tesmer, J.J. Structure and function of lysosomal phospholipase A2 and lecithin:cholesterol acyltransferase. *Nat. Commun.* **2015**, *6*, 6250. [CrossRef] [PubMed]

125. Piper, D.E.; Romanow, W.G.; Gunawardane, R.N.; Fordstrom, P.; Masterman, S.; Pan, O.; Thibault, S.T.; Zhang, R.; Meininger, D.; Schwarz, M.; et al. The high-resolution crystal structure of human LCAT1. *J. Lipid Res.* **2015**, *56*, 1711–1719. [CrossRef] [PubMed]

126. Falarz, L.J.; Xu, Y.; Caldo, K.M.P.; Garroway, C.J.; Singer, S.D.; Chen, G. Characterization of the diversification of phospholipid:diacylglycerol acyltransferases in the green lineage. *Plant J.* **2020**, *103*, 2025–2038. [CrossRef] [PubMed]

127. Yu, X.-H.; Cai, Y.; Chai, J.; Schwender, J.; Shanklin, J. Expression of a Lychee PHOSPHATIDYLCHOLINE: DIACYLGLYCEROL CHOLINEPHOSPHOTRANSFERASE with an Escherichia coli CYCLOPROPANE SYNTHASE enhances cyclopropane fatty acid accumulation in Camelina Seeds. *Plant Physiol.* **2019**, *180*, 1351–1361. [CrossRef]

128. Sigal, Y.J.; McDermott, M.I.; Morris, A.J. Integral membrane lipid phosphatases/phosphotransferases: Common structure and diverse functions. *Biochem. J.* **2005**, *387*, 281–293. [CrossRef]

129. Bates, P.D.; Fatihi, A.; Snapp, A.R.; Carlsson, A.S.; Browse, J.; Lu, C. Acyl Editing and Headgroup Exchange Are the Major Mechanisms That Direct Polyunsaturated Fatty Acid Flux into Triacylglycerols. *Plant Physiol.* **2012**, *160*, 1530–1539. [CrossRef]

130. Yang, W.; Wang, G.; Li, J.; Bates, P.D.; Wang, X.; Allen, D.K. Phospholipase Dζ Enhances Diacylglycerol Flux into Triacylglycerol. *Plant Physiol.* **2017**, *174*, 110–123. [CrossRef]

131. Chen, G.Q.; Vang, L.; Lin, J.-T. Seed Development in Lesquerella fendleri (L.). *HortScience* **2009**, *44*, 1415–1418. [CrossRef]

132. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

133. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef]

134. Bolstad, B.; Irizarry, R.; Astrand, M.; Speed, T. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **2003**, *19*, 185–193. [CrossRef] [PubMed]

135. Kolde, R.; Laur, S.; Adler, P.; Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **2012**, *28*, 573–580. [CrossRef] [PubMed]

136. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]

137. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [CrossRef] [PubMed]

138. Walley, J.W.; Shen, Z.; Sartor, R.; Wu, K.J.; Osborn, J.; Smith, L.G.; Briggs, S.P. Reconstruction of protein networks from an atlas of maize seed proteotypes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4808–E4817. [CrossRef]

139. Schwender, J.; König, C.; Klapperstück, M.; Heinzel, N.; Munz, E.; Hebbelmann, I.; Hay, J.O.; Denolf, P.; De Bodt, S.; Redestig, H.; et al. Transcript abundance on its own cannot be used to infer fluxes in central metabolism. *Front. Plant Sci.* **2014**, *5*, 668. [CrossRef]