

Article

Longer Duration of Active Oil Biosynthesis during Seed Development Is Crucial for High Oil Yield—Lessons from Genome-Wide In Silico Mining and RNA-Seq Validation in Sesame

Bhagwat Nawade ^{1,†} , Ajay Kumar ^{1,†}, Rasna Maurya ¹ , Rajkumar Subramani ¹ , Rashmi Yadav ² , Kuldeep Singh ^{1,‡}  and Parimalan Rangan ^{1,3,*} 

¹ Division of Genomic Resources, ICAR-National Bureau of Plant Genetic Resources, PUSA Campus, New Delhi 110012, India

² Division of Germplasm Evaluation, ICAR-National Bureau of Plant Genetic Resources, PUSA Campus, New Delhi 110012, India

³ Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane, QLD 4072, Australia

* Correspondence: r.parimalan@icar.gov.in

† These authors contributed equally to this work.

‡ Present address: Genetic Resources Division, ICRISAT, Patancheru, Hyderabad 502324, India.

Abstract: Sesame, one of the ancient oil crops, is an important oilseed due to its nutritionally rich seeds with high protein content. Genomic scale information for sesame has become available in the public databases in recent years. The genes and their families involved in oil biosynthesis in sesame are less studied than in other oilseed crops. Therefore, we retrieved a total of 69 genes and their translated amino acid sequences, associated with gene families linked to the oil biosynthetic pathway. Genome-wide in silico mining helped identify key regulatory genes for oil biosynthesis, though the findings require functional validation. Comparing sequences of the *SiSAD* (stearyl-acyl carrier protein (ACP)-desaturase) coding genes with known *SADs* helped identify two *SiSAD* family members that may be palmitoyl-ACP-specific. Based on homology with lysophosphatidic acid acyltransferase (*LPAAT*) sequences, an uncharacterized gene has been identified as *SiLPAAT1*. Identified key regulatory genes associated with high oil content were also validated using publicly available transcriptome datasets of genotypes contrasting for oil content at different developmental stages. Our study provides evidence that a longer duration of active oil biosynthesis is crucial for high oil accumulation during seed development. This underscores the importance of early onset of oil biosynthesis in developing seeds. Up-regulating, identified key regulatory genes of oil biosynthesis during early onset of seed development, should help increase oil yields.

Keywords: longer duration; early onset; sesame; oil biosynthesis; Kennedy pathway; *ACCase*; *KAS*; *SAD*; *FAD*; *LACS*; *DGAT*; *LPAAT*



Citation: Nawade, B.; Kumar, A.; Maurya, R.; Subramani, R.; Yadav, R.; Singh, K.; Rangan, P. Longer Duration of Active Oil Biosynthesis during Seed Development Is Crucial for High Oil Yield—Lessons from Genome-Wide In Silico Mining and RNA-Seq Validation in Sesame. *Plants* **2022**, *11*, 2980. <https://doi.org/10.3390/plants11212980>

Academic Editor: Valentina Pavic

Received: 13 September 2022

Accepted: 30 September 2022

Published: 4 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Background

Plants accumulate oil primarily in the form of triacylglycerols (TAG) [1]. Triacylglycerols have nutraceutical value and are the main source of edible oils for households and industrial applications. With the increasing human population, the global consumption for vegetable oils has increased by >50% over the past decade and is expected to double by 2040 [2,3]. The enzymatic reactions for TAG biosynthesis in plants are well established with the pathways for fatty acid (FA) biosynthesis in the plastid followed by TAG formation in the endoplasmic reticulum (ER) (Figure 1). For the final step in TAG biosynthesis, in addition to the reaction catalyzed by diacylglycerol O-acyl transferase (DGAT) (Figure 1),

an acyl-CoA-independent reaction is catalyzed by phospholipid:diacylglycerol acetyltransferase (PDAT) ([4,5] and citations therein). Further, a cytosolic pathway has been identified in groundnut [6].

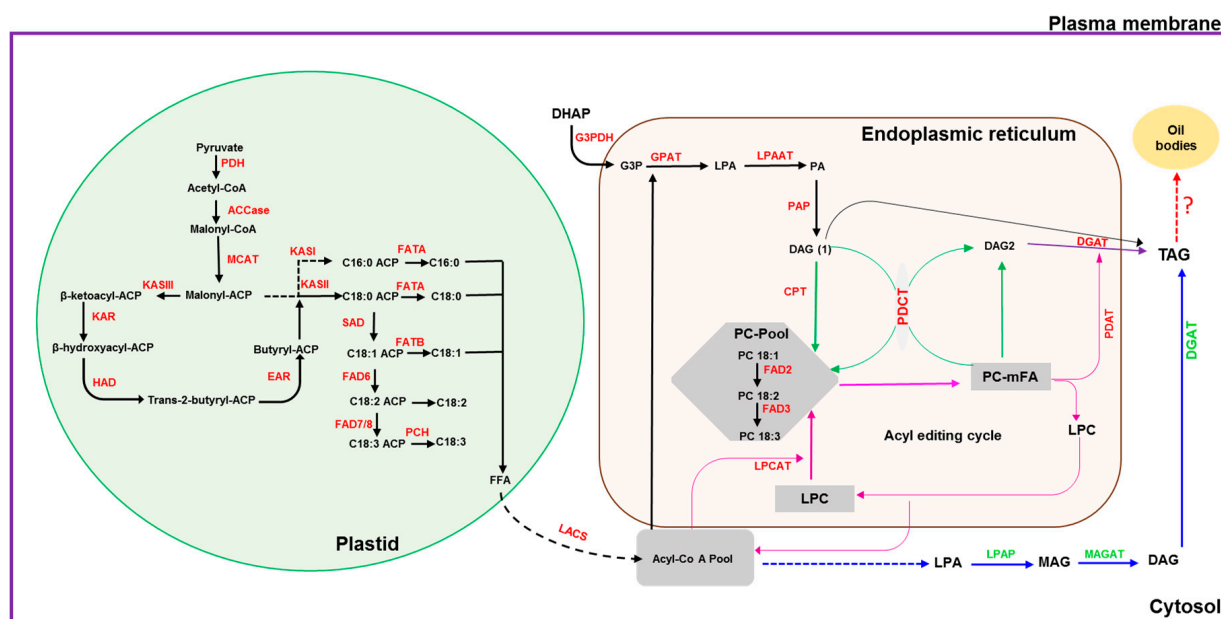


Figure 1. Schematic diagram depicting the oil biosynthesis pathway in plants (adapted from: [5–10]). The black dotted arrow indicates multiple steps in the conversion of FFA to acyl-CoA and transport; the blue dotted arrow indicates steps involved in the conversion of acyl-CoA to LPA; the red dotted arrow indicates conversion of TAG to oil bodies or lipid droplets. **ACCase**: acetyl-CoA carboxylase; **ACP**: acyl-carrier protein; **CPT**: CDP-choline:DAG cholinephosphotransferase; **DAG**: diacylglycerol; **DGAT**: diacylglycerol acyltransferase; **EAR**: enoyl-ACP reductase; **FAD2**: oleate desaturase; **FAD3**: linoleate desaturase; **FAD6**: oleate desaturase; **FAD7/8**: linoleate desaturase; **FATA**: fatty acid thioesterase A; **FATB**: fatty acid thioesterase B; **FFA**: free fatty acid; **G3PDH**: glycerol-3-phosphate dehydrogenase; **G3P**: glycerol-3-phosphate; **GPAT**: glycerol-3-phosphate acyltransferase; **HAD**: 3-hydroxyacyl-ACP dehydratase; **KASI**: 3-ketoacyl-ACP synthase I; **KASII**: 3-ketoacyl-ACP synthase II; **KASIII**: 3-ketoacyl-ACP synthase III; **KAR**: 3-ketoacyl-ACP reductase; **LACS**: long-chain acyl-CoA synthetase; **LPAAT**: lysophosphatidic acid acyltransferase; **LPCAT**: acyl-CoA:lysophosphatidylcholine acyltransferase; **LPAP**: lyso-phosphatidic acid phosphatase; **LPA**: lyso-phosphatidic acid; **LPC**: lyso-phosphatidylcholine; **MAG**: monoacylglycerol; **MAGAT**: monoacylglycerol acyltransferase; **MCAT**: malonyl Coenzyme A-ACP transacylase; **mFA**: PC-modified FA; **PAP**: phosphatic acid phosphohydrolase; **PA**: phosphatidic acid; **PC**: phosphatidylcholine; **PDH**: pyruvate dehydrogenase; **PCH**: palmitoyl-CoA hydrolase; **PDAT**: phospholipid:diacylglycerol acyltransferase; **PDCT**: PC:DAG cholinephosphotransferase; **SAD**: stearyl-ACP desaturase; **TAG**: triacylglycerol.

The first committed enzyme is acetyl-CoA carboxylase (ACCase), which acts as a control point over the carbon flux into FAs [8]. In plants, it is present in two functional forms, the heteromeric form in plastids and the homomeric form in the cytosol [11]. Heteromeric ACCase participates in FA synthesis, while the homomeric ACCase is involved in the synthesis of long-chain FAs and secondary metabolites [11]. There are four subunits present in heteromeric ACCase, encoded by *accC* for biotin carboxylase (BC), *accB* for biotin carboxyl carrier protein (BCCP), *accA* for α -carboxyl transferase, and *accD* for β -carboxyl transferase [12]. Of these four, only *accD* is chloroplast-encoded, while the others are nuclear encoded [8]. The malonyl-CoA-acyl carrier protein transacylase (MCAT) coding gene, *FabD*, is one of the key regulators in FA biosynthesis [13]. It aids in the formation of malonyl-ACP, the building block of FA synthesis, from malonyl-CoA [14]. Subsequently, β -ketoacyl-ACP

synthase III (KASIII; encoding gene: *FabH*) catalyzes the condensation and transacylation of acetyl-CoA with malonyl-ACP to form 3-ketobutyryl-ACP. β -ketoacyl-ACP synthase I (KASI; *FabB*) elongates 3-ketobutyryl-ACP to palmitoyl-ACP in six condensation cycles. In the final elongation step, β -ketoacyl-ACP synthase II (KASII, *FabF*) converts palmitoyl-ACP to stearyl-ACP in the plastid [15]. The acyl-ACP thioesterases (acyl-ACP TE; *FatA*, *FatB*) catalyze the FA chain termination through hydrolysis of the thioester bond of acyl-ACP to release the free FAs [9].

These free FAs are reactivated by a long-chain acyl-CoA synthetase (LACS) to form acyl-CoA esters and are exported to the ER [16]. FA desaturases (FAD) desaturate fatty acyl chains to form unsaturated fatty acids [17]. Plant FADs are of two types, soluble and membrane-bound desaturases. Stromal Δ^9 acyl-ACP desaturases (AADs) introduce a double bond in saturated acyl chains to form cis-monoenes. The stearyl-ACP-desaturase (SAD) is the archetype for this family and desaturates stearyl-ACP (18:0) at C-9 to form oleic acid (18:1; omega-9). An AAD isoform, palmitoyl-ACP 19 desaturase (PAD), exhibits substrate specificity to palmitoyl-ACP, thereby producing palmitoleic acid (16:1; omega-7) [18]. The substrate-specific activity of SAD enzymes determines the ratio of omega-7 and omega-9 FAs in plants [19]. Membrane-bound desaturases are localized to the ER and plastid membranes, and they are grouped into several FAD subfamilies including Omega-3-FAD (fatty acid desaturase3 FAD3, FAD7, and FAD8) and Omega-6-FAD (FAD2 and FAD6) [20]. In ER, the sequential transfer of fatty acyl moieties to a glycerol-3-phosphate backbone generates TAG in four steps, also known as the Kennedy pathway [21]. Glycerol-3-phosphate is acylated by glycerol-3-phosphate acyltransferase (GPAT) to form lyso-phosphatidic acid (LPA) and further acylation of LPA by lysophosphatidic acid acyltransferase (LPAAT) to produce phosphatidic acid (PA) [1]. Subsequently, PA is hydrolyzed by phosphatidic acid phosphatase (PAP) to generate diacylglycerol (DAG). Plant PAPs are classified into two types: soluble-type PAPs called phosphatidate phosphohydrolases (PAH) dephosphorylate PA and membrane-bound PAPs called lipid phosphate phosphatases (LPP) dephosphorylate different lipid phosphates, non-specifically. In the final step, diacylglycerol acyltransferases (DGAT) convert DAG to TAG through the transfer of an acyl group [7] (Figure 1).

Sesame (*Sesamum indicum* L., $2n = 2x = 26$) is an ancient oilseed crop of the Pedaliaceae family. Sesame seeds are nutritionally rich and possess 50–60% oil, 25% protein, and antioxidants such as sesamol and sesamin [22]. Sesame is a resilient crop with strong adaptation to environmental changes and is considered the ‘queen of oil seeds’ [23]. Sesame seed yield during 2019 was 6.5 million tons, as per FAO statistics (<https://www.fao.org/faostat/en/#data>; accessed on 28 September 2022). Understanding the genetic basis of oil biosynthesis and developing stable high oil yielding sesame varieties have been key objectives in sesame breeding in recent years [22]. To accelerate crop improvement, omics and molecular tools play key roles through the availability of high-throughput sequencing technologies. Sesame has gained increasing attention, and new genomic information is becoming available for the identification of genes and markers underlying the traits of interest, especially oil yield and productivity, with the help of omics data [24,25]. Here, we mined the existing genomic information for genes (and their translated amino acid sequences) and their families involved oil biosynthesis and subjected them to different bioinformatic tools to understand the characteristic features of the nucleotide and translated protein sequences. Our study identified the orthologs in sesame, on a homology basis, of genes that were previously unknown and also of genes associated with high oil yield with the potential to enhance oil yield in sesame.

2. Results

2.1. Physicochemical Properties of Retrieved Proteins

The translated amino acid sequences of genes involved in the oil biosynthesis pathway and their paralogs in sesame were mined from the NCBI database (Table S1). A total of seven genes that encode four subunits of heteromeric ACCase enzymes were found, including two for *SiaccB*, one for *SiaccC*, three for *SiaccA*, and one for chloroplastic *SiaccD* (Table S1).

One *MCAT*, eight *KAS*, and thirteen *FAD* genes, including seven *SAD*, two *FAD2*, and one each of *FAD3*, *FAD6*, *FAD7*, and *FAD8*, were retrieved. Six *LACS*s, seventeen *GPAT*s, six *LPAAT*s, three *PAH*s, and three *DGAT*s were retrieved in sesame (Table S1). Translated amino acid sequences (Seq_File_S1.docx) were used for the analysis. The analyzed protein sequences of sesame ranged between 271 amino acids for *SiaccB-2* (29.06 kDa) and 1004 (110.47 kDa) for *SiPAH2* (Table S2). The predicted aliphatic index of the studied sesame proteins ranged from 72.20 to 121.27. The aliphatic index indicates the thermostability and half-life of a protein [26].

2.2. Subcellular Localization

The subcellular localization for the translated AA sequences was predicted in silico using the DeepLoc tool. For the nuclear-encoded subunits of ACCase, *SiaccB1*, *SiaccB2*, and *SiaccC* were localized to the plastid while *SiaccA* localized in the plastid membrane (Table S2). All the members of *SiKASI* and *SiKASII* were plastid localized (Table S2).

The TMHMM tool predicted the existence of transmembrane helix domains in the *FAD*, *GPAT*, *LPAAT*, and *DGAT* families. In the case of membrane-bound *FAD*s (*SiFAD2-1*, *SiFAD2-2*, *SiFAD3*, *SiFAD6*, *SiFAD7*, and *SiFAD8*), the number of transmembrane domains among proteins exhibited variability between a total of five transmembrane domains in *SiFAD2-1* and only one in *SiFAD7* (Figure S1). *SiGPAT* proteins were predicted to have two to four transmembrane domains (Figure S2). The TMHMM tool had predicted nine and two transmembrane domains for *SiDGAT1-1* and *SiDGAT-2*, respectively (Figures S1–S3).

2.3. Homology Relationship

A tree was constructed based on an amino acid sequence alignment to understand the homology relationships between gene families of oil biosynthesis of sesame and other related oilseed crops (Seq_File_S1.docx). The gene products involved in oil biosynthesis were grouped into 11 distinct clades (Figure 2). All desaturase proteins formed a single clade with four subclades, in which one subclade consist of two groups, *FAD2* in one and *FAD3* and *FAD7/8* in the second group. The second subclade possesses most of the *SAD* proteins excluding six *SAD*s (*SiSAD5*, *SiSAD6*, *SiSAD7*, *AtSAD7*, *AhSAD2*, and *AhSAD3*, that are grouped in third subclade). Lastly, the fourth subclade grouped all *FAD6* proteins (Figure 2). *KAS* and *LACS* proteins formed a separate clade (Figure 2). All *GPAT* proteins were grouped into a single clade, except *SiGPAT1* and *AtATS1*. In a clade containing *FatA* and *FatB* proteins, α -CT and β -CT subunits of ACCase were grouped in two distinct subclades (Figure 2). The other subunit proteins of ACCase proteins, BCCP and BC, were grouped in a different clade. Among the Kennedy pathway gene products, *DGAT3* and *PAH* were more closely related and fall into one clade. Similarly, *DGAT1* and *DGAT2* forms a distinct clade (Figure 2). The protein identified by us as *SiLPAAT-B* on the basis of homology relationship with class-B *LPAAT* of *Ricinus communis* 410 (*RcLPAAT-B*, Figure 2) is annotated in the NCBI protein database as 1-acyl-sn-glycerol-3-phosphate acyltransferase, (XP_020554192.1).

2.4. Multiple Sequence Alignment

Multiple sequence alignments of BCCP proteins indicated that the C-terminal region was conserved with a typical biotinyl domain (CIIEAMKLMNEIE) (Figure S4A), while sequence alignments of BC (accC) proteins indicated their highly conserved nature (Figure S4B). The N-terminal domain of α -CT proteins showed more conserved sites than the C-terminal one, whereas in the β -CT proteins, the C-terminal domain was more conserved (Figure S5).

The specific amino acid residues that are conserved in the active sites of *FabD* proteins are present in *SiFabD* as Gln112, Ser197, Arg222, His310, and Gln359 were highly conserved in *AhFabD* as Gln89, Ser174, Arg199, His287, and Gln336 [13,28] (Figure S6). An active site triad Cys-His-Asn and the motif GNTSAAS were found to be conserved in *SiKASIII* proteins (Figure S7). The seven catalytic sites' residues (Cys, His, Thr, Thr,

Lys, His, and Phe) and cation site residues [29] are highly conserved in the KASI and KASII proteins (Figure S8). The presence of two histidine boxes in the *SiSAD* genes is consistent with plastidial stearyl-ACP desaturases containing EENRHG and DEKRHE boxes (Figure S9; Table 1).

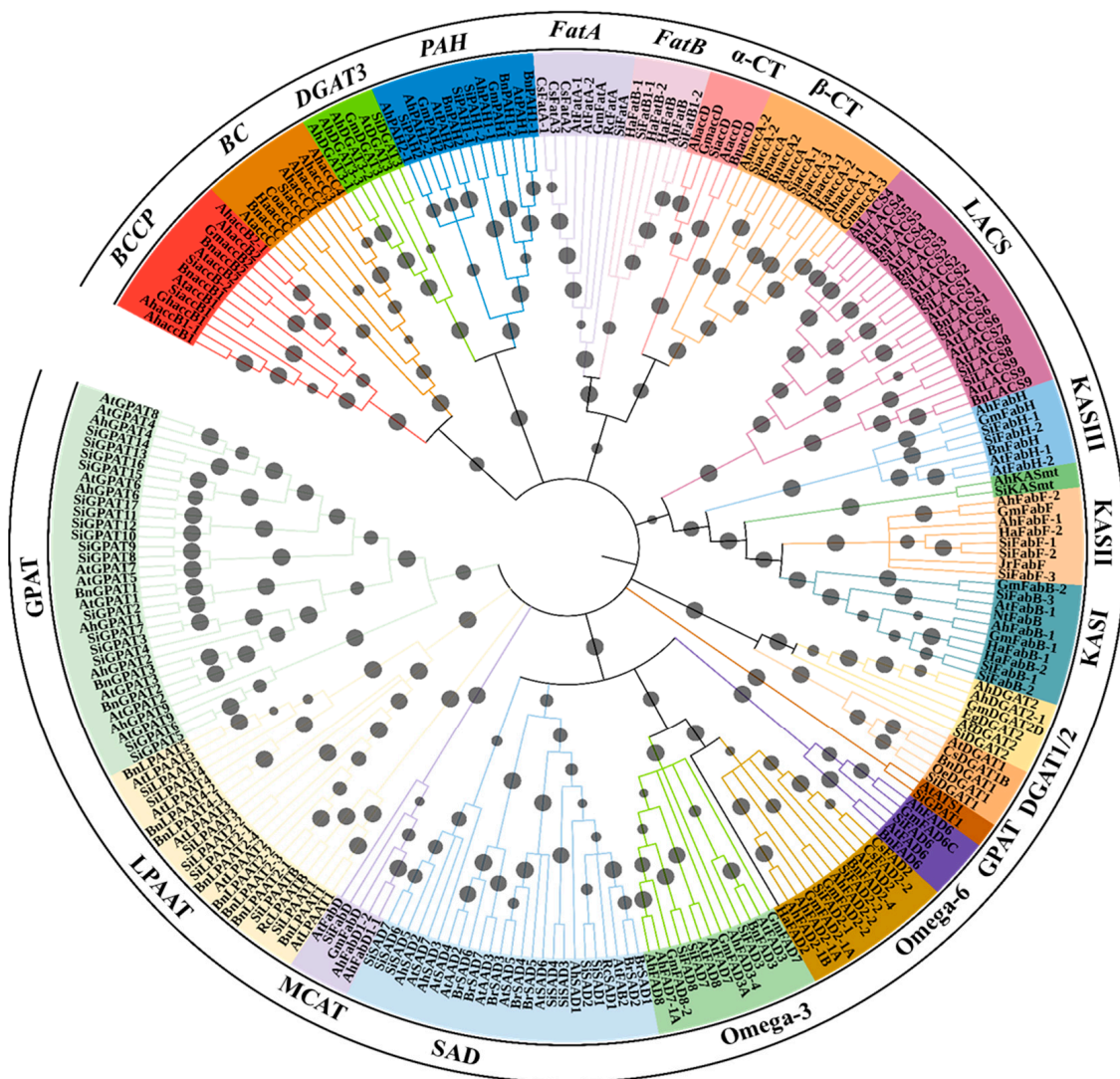


Figure 2. Circular dendrogram for the gene products involved in oil biosynthesis constructed using the N-J method with 10,000 bootstraps using MEGA v10 [27]. For accession number details of each gene product, please refer to Supplementary Table S1. Bootstrap values are represented by grey circles on the branches. The subfamily names of gene products are denoted on the outermost circle of the tree. Ah—*Arachis hypogaea*, At—*Arabidopsis thaliana*, Bn—*Brassica napus*, Br—*Brassica rapa*, Co—*Camellia oleifera*, Cs—*Camelina sativa*, Eg—*Elaeis guineensis*, Gh—*Gossypium hirsutum*, Gm—*Glycine max*, Ha—*Helianthus annuus*, Hi—*Handroanthus impetiginosus*, Ji—*Jatropha curcas*, Jr—*Juglans regia*, Muc—*Macfadyena unguis-cati*, Nt—*Nicotiana tabacum*, Oe—*Olea europaea*, Pf—*Perilla frutescens*, Rc—*Ricinus communis*, Sa—*Striga asiatica*, Sh—*Salvia hispanica*, Si—*Sesamum indicum*, Tm—*Tropaeolum majus*.

Table 1. Conserved histidine-rich boxes for the FAD family members identified in sesame.

| Protein Name | Histidine Box-1 | | Histidine Box-2 | | Histidine Box-3 | | Membrane Nature |
|-----------------|-----------------|-------------|-----------------|-------------|-----------------|-------------|-----------------|
| | Conserved Motif | AA Position | Conserved Motif | AA Position | Conserved Motif | AA Position | |
| <i>SiSAD-1</i> | EENRHG | 175–180 | DEKRHE | 261–266 | - | - | soluble |
| <i>SiSAD-2</i> | EENRHG | 175–180 | DEKRHE | 261–266 | - | - | soluble |
| <i>SiSAD-3</i> | EENRHG | 169–174 | DEKRHE | 255–260 | - | - | soluble |
| <i>SiSAD-4</i> | EENRHG | 169–174 | DEKRHE | 255–260 | - | - | soluble |
| <i>SiSAD-5</i> | EENRHG | 160–165 | DEKRHE | 246–251 | - | - | soluble |
| <i>SiSAD-6</i> | EENRHG | 159–164 | DEKRHE | 245–250 | - | - | soluble |
| <i>SiSAD-7</i> | EENRHG | 166–171 | DEKRHE | 252–257 | - | - | soluble |
| <i>SiFAD2-1</i> | HECGH | 105–109 | HRRHH | 141–145 | HVTHH | 315–319 | membrane-bound |
| <i>SiFAD2-2</i> | HECGH | 105–109 | HRRHH | 141–145 | HVAHH | 315–319 | membrane-bound |
| <i>SiFAD3</i> | HDCGH | 123–127 | HKTHH | 159–163 | HVIHH | 326–330 | membrane-bound |
| <i>SiFAD6</i> | HDCAH | 167–171 | HDRHH | 203–207 | HIPHH | 363–367 | membrane-bound |
| <i>SiFAD7</i> | HDCGH | 167–171 | HRTTH | 203–207 | HVIHH | 370–374 | membrane-bound |
| <i>SiFAD8</i> | HDCGH | 161–165 | HRTTH | 197–201 | HVIHH | 364–368 | membrane-bound |

To identify specific *SiSAD* forms potentially possessing palmitoyl-ACP activity, sesame SAD sequences were aligned with well-characterized SADs, including SADs characterized as 18:0-ACP specific, *AtFAB2* [18], and *RcSAD1* [30]. The palmitoyl-ACP (16:0-ACP)-specific enzymes were characterized using sequence information from *AtAAD2* and *AtAAD3* in *Arabidopsis* [31] and *MucPAD* in *Macfadyena unguis-cati* [32]. These studies reported that the substrate specificity of SAD is determined by eight amino acid residues from the region of the catalytic sites (Figure S9; Table 2). The comparison showed that 5 *SiSAD* members (*SiSAD1*–5) have the same amino acid residues at catalytic sites as *AtFAB2* and *RcSAD1* (except that *SiSAD3* and *SiSAD5* have isoleucine instead of leucine at the 142nd position), while *SiSAD6* and *SiSAD7* showed divergent residues in the aforesaid catalytic sites, suggesting that they might possess 16:0-ACP substrate specificity (Table 2). Sequence alignment of *SiFatA* and *SiFatB* with representative proteins identified four conserved residues necessary for thioesterase activity (Figure S10). All membrane-bound FAD proteins contain three conserved histidine motifs, H (X)3–4H, H (X)2–3HH, and H/Q (X)2–3HH. Eight conserved histidine residues from these boxes were reported to be important for di-iron center formation (Figure S11).

Table 2. Comparison of the amino acid residues of the catalytic sites of *SiSADs* with cell highlights showing a different AA when compared to *RcSAD1* and *AtFAB2* sequences.

| <i>Rc</i> | | <i>At</i> | | <i>Mu</i> | <i>Si</i> | | | | | | |
|-----------|------|-----------|------|-----------|-----------|------|------|------|------|------|------|
| SAD1 | FAB2 | AAD2 | AAD3 | PAD | SAD1 | SAD2 | SAD3 | SAD4 | SAD5 | SAD6 | SAD7 |
| M147 | M152 | T157 | T149 | M147 | M147 | M147 | M141 | M141 | M132 | T131 | M138 |
| L148 | L153 | L158 | L150 | L148 | L148 | L148 | I142 | L142 | I133 | I132 | I139 |
| T150 | T155 | T160 | T152 | T150 | T150 | T150 | T144 | T144 | T135 | A134 | T141 |
| L151 | L156 | L161 | L153 | W151 | L151 | L151 | L145 | L145 | L136 | C135 | C142 |
| P212 | P217 | S222 | S214 | P212 | P212 | P212 | P206 | P206 | P197 | V196 | I203 |
| T214 | T219 | F224 | F216 | T214 | T214 | T214 | T208 | T208 | T199 | W198 | L205 |
| G221 | G226 | G231 | G223 | G221 | G221 | G221 | G215 | G215 | G206 | T205 | T212 |
| F222 | F227 | F232 | F224 | F222 | F222 | F222 | F216 | F216 | F207 | F206 | F213 |

Rc: *Ricinus communis*; **At:** *Arabidopsis thaliana*; **Mu:** *Macfadyena unguis-cati*; **Si:** *Sesamum indicum*.

Two conserved domains are typically present in the LACS family (Figure S12), an AMP-binding domain (involved in adenylate formation) and the ACS signature domain for catalytic activity [33]. Four conserved acyltransferase domains necessary for acyltransferase activity [34] were found in all SiGPAT proteins (Figure S13). Two protein motifs (NHX₄D and EGTK) essential for LPAAT activity [35] were present in all LPAAT proteins (Figure S14). The catalytic motif DI/VDGT, characteristic of the Mg²⁺-dependent PA phosphatases, was identified in all PAH proteins (Figure S15). The sequences of DGAT subfamilies were distinct with no common features (Figures S16–S18).

2.5. Conserved Motif Analysis

Proteins having related functions contain highly conserved short (<20 amino acids) amino acid sequences recurring in fixed-length patterns called motifs. A motif may represent important biological features, such as protein-binding or targeting a particular subcellular location. Conserved motifs in all studied genes (Figures S19–S36) were identified using the MEME web tool.

2.6. Promoter Analysis

Based on the homology relationships for genes associated with oil biosynthesis in model plants and genotypes exhibiting higher oil yield, we have identified 12 genes in sesame that could be potential targets for enhancement in the sesame oil yield (Table 3). The upstream sequences (3 kb from the start codon) for each of the 12 genes were subjected to promoter analysis using the PlantCARE database to predict the potential regulatory features [36]. A total of 79 CAREs (cis-acting regulatory elements) were detected from the upstream sequences of the aforementioned 12 genes potentially targeted to improve oil yield (Figure 3; Table S3). Different TFs associated with FA biosynthesis regulation and TAG accumulation are well documented, including DNA binding with one finger (Dof), LEAFY COTYLEDON1 (LEC1), LEC2, sequences over-represented in light-induced promoters 5 (SORLIP5), AGAMOUS-LIKE15 (AGL15), BASIC LEUCINE ZIPPER 67 (bZIP67), SPATULA (SPT), WRINKLED1 (WRI1), MYB96, ABSCISIC ACID INSENSITIVE3 (ABI3), and FUSCA3 (FUS3) [37,38]. Of these 11 TFs, binding sites for TFs DOF, SORLIP5, AGL15, and LEC1 were identified in the promoter regions of all the 12 genes studied (Figure 4). Although the number of DOF binding sites is over-represented in all these promoters studied, the literature suggests that not all the AAG motifs in plant promoters are targets for the DOF domain containing proteins and are variably regulated [39]. DOF4 and DOF11 positively regulate ACCase and LACS and, thereby, increase seed oil content in Arabidopsis, while this negatively regulates CRA1, associated with seed storage protein [40]. LEC1 and LEC2 positively regulate WRI1 and AGL15, respectively [41,42]. WRI1 in turn positively regulates TAG biosynthesis through DGAT1 [43]. Our transcriptome analysis also revealed the higher expression levels of WRI1 (SIN_1023649), specifically in high oil yielding genotypes. Binding sites for MYB96, a positive regulator of seed oil accumulation and activator of the DGAT1 gene in Arabidopsis seeds [44], was detected in the upstream sequences of the SiDGAT2 and SiDGAT3 genes in sesame (Figure 4), but not in the SiDGAT1 gene. It is reported that MYB96 binds to the promoter of ABI4 and activates its expression [45], and ABI4 has been shown to directly regulate DGAT1 expression in Arabidopsis [46].

Table 3. List of orthologs in sesame identified from genes reported in model plants as associated with improving oil content.

| S. No | Gene Name | Genes from Model Plants | Effect on Oil Accumulation | Reference | Orthologs in <i>S. indicum</i> |
|-------|-----------|--|---|-----------|--------------------------------|
| 1 | ACCase | GhaccB1 (<i>Gossypium hirsutum</i>) | Overexpression driven by the seed specific AGP promoter improved cotton oil content by 9.19 to 21.92% | [47,48] | SiaccB1 |
| | | NtaccD (<i>Nicotiana tabacum</i>) | Chloroplast transformation with modified accD operon increased seed and oil yield | [49] | SiaccD |
| 2 | MCAT | AtFabD (<i>Arabidopsis</i>) | Overexpression of AtFabD driven by the promoter of the senescence-associated 1 (SEN1) gene increased by 15–20% storage oil content of <i>Arabidopsis</i> | [50] | SiFabD |
| 3 | KASI | NtFabB (<i>Nicotiana tabacum</i>) | NtKASI-1 overexpression enhanced oil accumulation | [15] | SiFabB-2 |
| 4 | KASII | JcFabF (<i>Jatropha curcas</i>) | Virus-induced gene silencing of JcKASII significantly altered TAG biosynthesis | [51] | SiFabF-3 |
| 5 | LACS | AtLACS4 and AtLACS9 (<i>Arabidopsis</i>) | TAG content of the lacs4 lacs9 double mutant seeds reduced by 27% compared to wild type seeds | [16] | SiLACS4 and SiLACS9 |
| 6 | GPAT | AtGPAT9 (<i>Arabidopsis</i>) | Knockdown of AtGPAT9 resulted in 26% to 44% reduced seed oil content | [52] | SiGPAT5 and SiGPAT6 |
| 7 | LPAAT | AhLPAAT2 (<i>Arachis hypogea</i>) | Seed-specific overexpression of AhLPAAT2 in <i>Arabidopsis</i> increased oil content by 7.4% in transgenic plants | [53] | SiLPAAT2-1 and SiLPAAT2-2 |
| | | BnLPAAT2-2 and BnLPAAT2-4 (<i>Brassica napus</i>) | Overexpression of rapeseed LPAAT genes in <i>Arabidopsis</i> resulted in transgenic seeds with 13% increased TAG content | [54] | |
| 8 | DGAT1 | AtDGAT1 | Overexpression of AtDGAT1 under the control of a seed-specific promoter in <i>Arabidopsis</i> and canola increased seed oil content by 28 and 16%, respectively | [55] | SiDGAT1 |
| | | CsDGAT1B (<i>Camelina sativa</i>) | ~24% increase in seed oil content in transgenic <i>C. sativa</i> overexpressing CsDGAT1B | [56] | |
| | | TmDGAT1 (<i>Tropaeolum majus</i>) | Overexpression of embryo-specific TmDGAT1 increased the storage oil content in transgenic <i>Arabidopsis</i> and rapeseed by ~8 and ~15%, respectively | [57] | |
| | | BnDGAT1 (<i>Brassica napus</i>) | Seed-specific napin promoter driven overexpression in <i>Arabidopsis</i> and rapeseed increased oil content by ~5 and 14%, respectively | [58] | |
| 9 | DGAT2 | EgDGAT2 (<i>Elaeis guineensis</i>) | Heterologously expressed EgDAGT2 in mutant yeast H1246 restored TAG biosynthesis and overexpression of EgDAGT2 in <i>Arabidopsis</i> increased the content of polyunsaturated FAs | [59] | SiDGAT2 |
| | | GmDGAT2D (<i>Glycine max</i>) | Overexpression of GmDGAT2D in <i>Arabidopsis</i> increased TAG biosynthesis by 5.7% | [60] | |

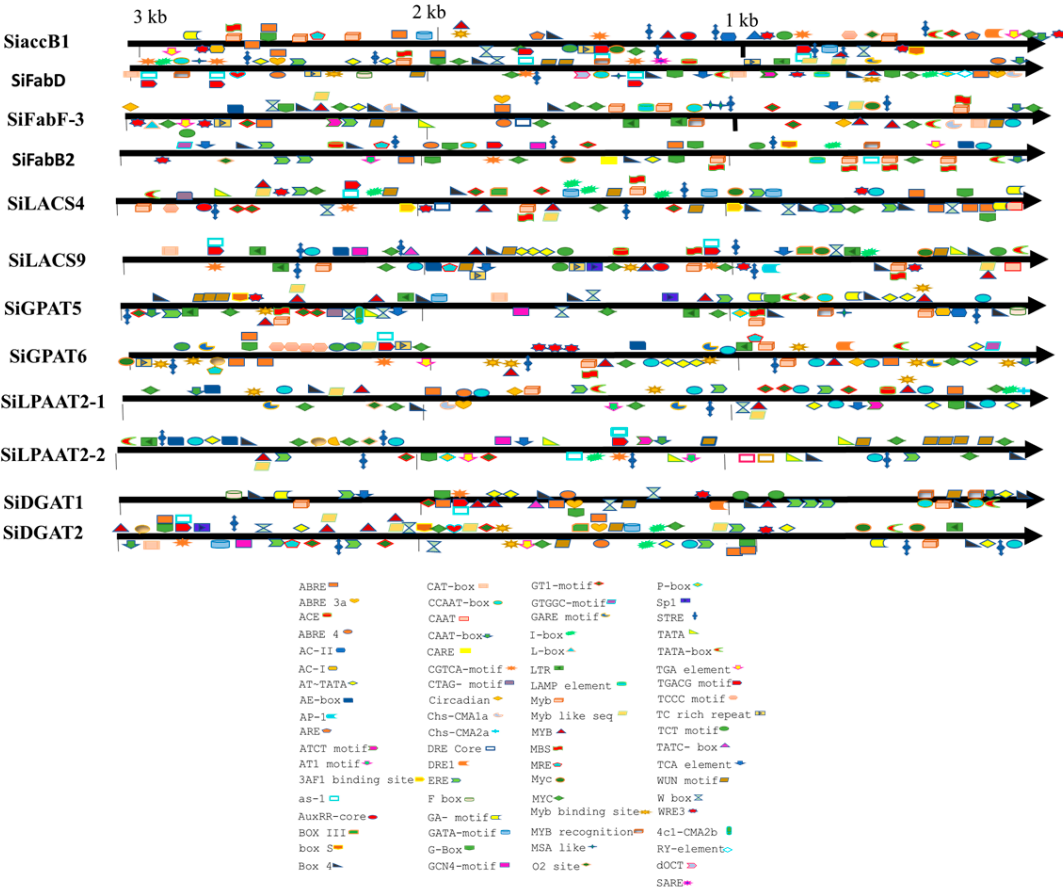


Figure 3. Distribution of *cis*-acting regulatory elements (CAREs) in the promoters of selected genes detected in the PlantCare tool.

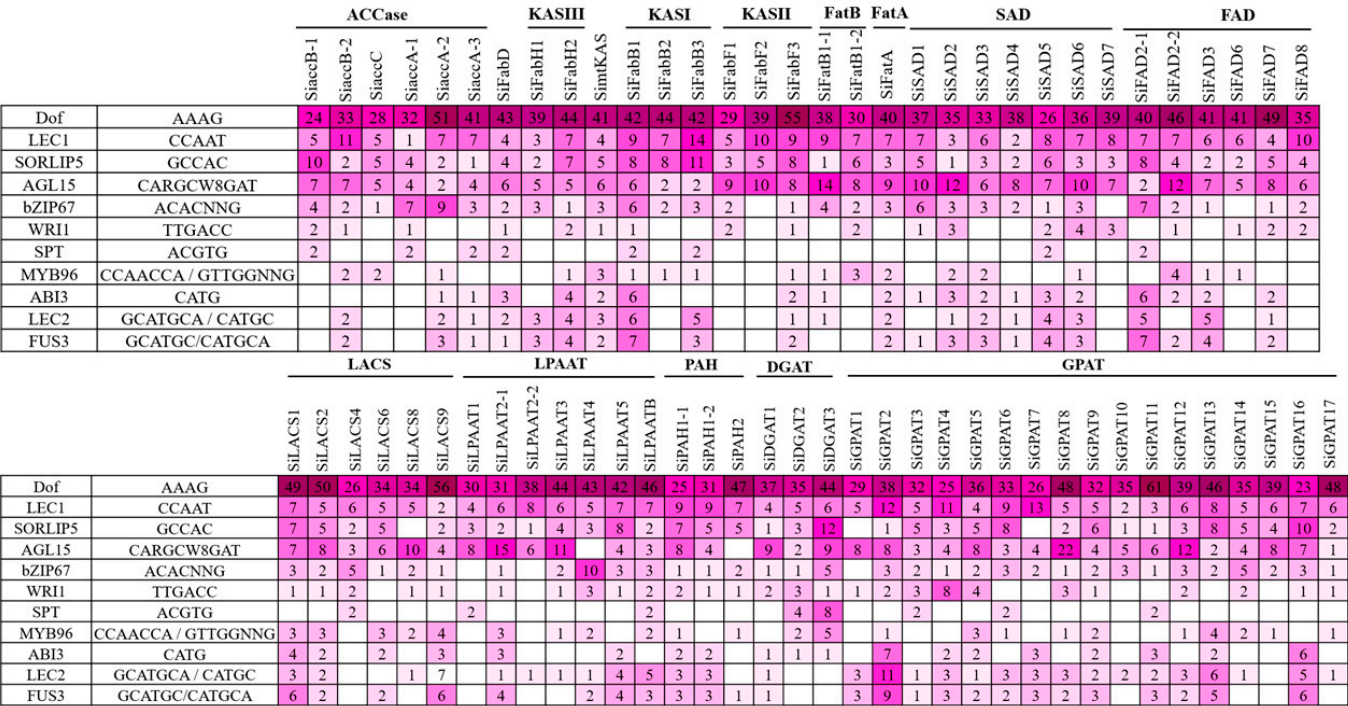


Figure 4. The number of binding sites identified using PLACE and PlantPAN 3.0 tools for each of the CARE (cis-acting regulatory elements) associated with oil accumulation (represented in a heat map).

2.7. Validation of In Silico Predicted Oil Biosynthesis Genes Using RNA-Seq

The key regulatory genes associated with oil biosynthesis (Table 3) were identified through genome-wide in silico mining approaches. Expression of these genes was validated through RNA-seq based differential expression studies using the publicly available transcriptome data for three genotypes (high oil yielding-ZZM4728, low oil yielding-ZZM3495, ZZM2161) at four developmental stages (10, 20, 25, and 30 days-post-anthesis: DPA) [24]. In total, 501 genes were found to be significantly differentially expressed in the high oil yielding sesame genotype (ZZM4728) when compared to the low oil yielding genotypes (ZZM3495 and ZZM2161) (Table S4). Further, of the 501 genes, 27 significantly differentially expressed genes were associated with the lipid biosynthesis (including lipid droplet biogenesis) pathway. From these, *DGAT*, *LPAAT*, *FAD*, *SAD*, *KCS* (β -ketoacyl-CoA synthase), *Oleosin*, *nsLTP* (non-specific lipid transfer protein), and *DIR1* (defective in induced resistance1) genes were selected manually.

The *SAD* gene showed higher expression at the early stage (10 DPA) in the high oil yielding genotype while its higher expression was found during middle stages (20 and 25 DPA) in the low oil yielding genotypes (Figure 5). Similar to *SAD* expression, *FAD2* gene expression also increases consistently from the early seed developmental stage (10 DPA) and is decreased at 30 DPA in the high oil yielding genotype, while, in case of low oil yielding genotypes, its expression level increased only during the middle developmental stage (20 DPA) (Figure 5). We also found genes associated with lipid droplet biogenesis (oleosin gene family) to be significantly more highly expressed in the high oil yielding genotypes, at 10 DPA. Interestingly, these oleosin genes exhibited higher expression levels in both high and low oil yielding genotypes during later developmental stages (25 and 30 DPA). The genes *WSD1/DGAT*, *LPAAT*, and *FAD4L2* exhibited higher expression levels during the early and middle developmental stages (10 and 20 DPA) in both high and low oil yielding genotypes. However, the expression level of these genes decreased towards later developmental stages (25 and 30 DPA) in low oil yielding genotypes while the expression pattern in high oil yielding genotypes is sustained at comparatively higher levels than the low oil yielding genotypes. Genes for lipid transfer proteins include *nsLTP*-encoding genes and *DIR1* [61,62]. The gene product of *KCS* catalyzes the first step of the very long chain FA biosynthesis [63,64]. These three genes, *nsLTP*, *DIR1*, and *KCS*, tended to show higher expression in the high oil content sesames genotype when compared to the low oil yielding genotypes (Figure 5J–L). Of these, expression pattern for *nsLTP1* and *DIR1* transcripts were lower at 10 DPA in low oil yielding genotypes when compared to high oil yielding genotypes, whereas *KCS* expression was found to be lower at 30 DPA in the low oil yielding genotypes than the high oil yielding genotype (Figure 5M–O). These results suggest that early onset of oil biosynthesis during seed development is important for higher oil yield, and this is regulated transcriptionally by the key regulatory genes studied here. In low oil yielding genotypes, the capacity for oil biosynthesis evidently initiates during the mid to late seed developmental stage, which may explain the low oil yield. Since the time duration between anthesis and capsule maturity is uniformly 30–45 days [24], early onset of oil biosynthesis during seed development results in the longer duration of active oil formation and, thereby, higher oil yield.

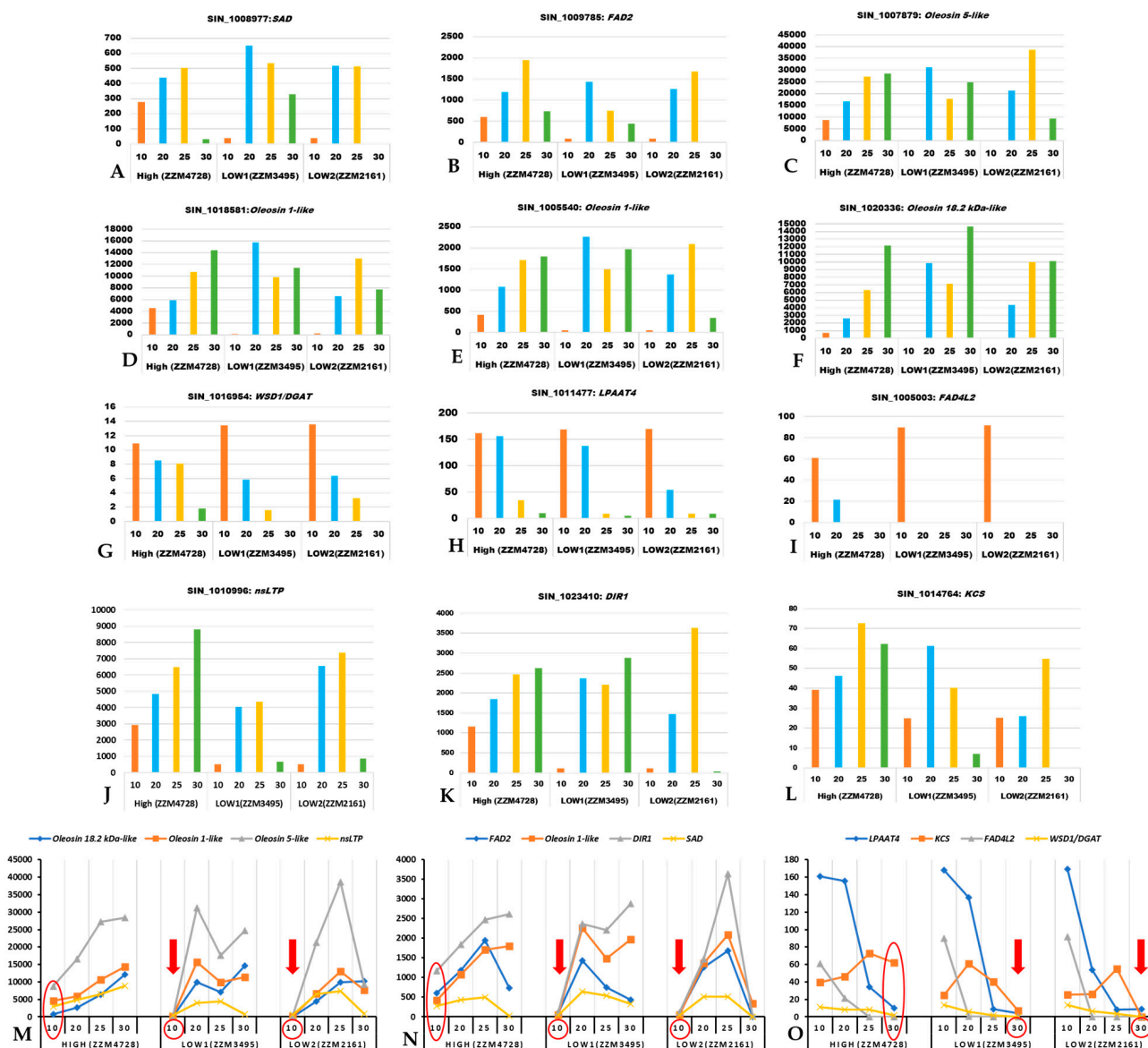


Figure 5. Expression profiles (in FPKM values, Y-axis) for key regulatory genes of oil biosynthesis in sesame using the transcriptome dataset of Wang et al. (2019). The X-axis represents the developmental stages in days post anthesis (DPA). (A): *Sad*; (B): *Fad2*; (C): *Oleosin 5-like*; (D): *Oleosin 1-like*; (E): *Oleosin 1-like*; (F): *Oleosin 18.2 kDa-like*; (G): *Dgat*; (H): *Lpaat4*; (I): *Fad4L2*; (J): *nsLTP*; (K): *DIR1*; (L): *KCS*; (M,N): Set of genes that are expressed significantly less at 10 DPA in both of the low oil yielding genotypes (indicated with a red colored down arrow at 10 DPA); (O): Set of genes that are significantly less expressed at 30 DPA in both of the low oil yielding genotypes (indicated with a red colored down arrow at 30 DPA).

3. Discussion

With the availability of high-throughput technologies and the sesame genome datasets, various genes involved in oil biosynthesis have been mined from the publicly available sesame genome and RNA-seq datasets and were studied using various bioinformatics tools. The present report focused on studies pertaining to conserved sites, catalytic sites, domains and motifs, physicochemical properties, subcellular location, and their homology with genes of other oilseed crop and model plants. TAG biosynthesis in oilseeds can be summarized in three major steps (Figure 1): (i) biosynthesis of FA in plastids, (ii) desaturation of

FA in plastid and ER, and (iii) TAG assembly in ER, where the free fatty acids are exported from the plastid to undergo stepwise acylation onto the glycerol backbone, forming TAG.

Plant de novo FA synthesis occurs in the stroma of the plastid in a series of reactions: initiation, elongation, and termination [5]. The heteromeric ACCase subunit genes have been characterized in different crop plants, including peanut [13], soybean [12], *Gossypium* species [48], and pea [65]. Overexpression of the *BCCP* gene modulates the oil content in seeds of *Arabidopsis* and cotton [47]. Additionally, overexpression of the *accD* gene through chloroplast transformation resulted in increased ACCase levels that led to higher oil production in transgenic tobacco [49].

3.1. Characterization of the In Silico Mined Genes of Oil Biosynthesis

An evaluation of the physicochemical properties of the amino acids of FA synthesis genes highlighted that, on an isoelectric point (pI) basis, most of the proteins were alkaline in nature except *SiaccB-2*, *SiaccC*, *SiaccD*, *SiKASIII*, *SiSAD*, and *FATB* (Table S2). Similarly, the isoelectric points (pIs) of FAD genes from sunflower (pIs ranged from 6.24 to 9.61) and *Brassica napus* (7.8 to 9.5) showed alkaline nature [66,67]. The instability index is a measure of correlation with the in vivo half-life of a protein [68]. The aliphatic index is defined as a relative volume occupied by aliphatic side chains (alanine, valine, leucine, and isoleucine). Aliphatic amino acids are hydrophobic; therefore, a high aliphatic index indicates that a protein is thermo-stable over a wide temperature range and is regarded as a positive factor for the increase in thermostability of globular proteins [69]. The GRAVY value indicates the solubility of a protein, and a low GRAVY value indicates the hydrophilicity of the protein [69]. Most studied proteins have negative values revealing their hydrophilic nature (Table S2).

The catalytic residues are well conserved within KASI, KASII, and KASIII proteins (Figures S7 and S8). These residues are reported to each have a key function: cysteine (C) is a substrate-binding residue, two histidines (H) are involved in the decarboxylation, and two threonines (T) are involved in forming hydrogen bonds with the ACP phosphopantetheine moiety [15] (Figure S8). In the C-terminal region, a conserved Gly-rich GNTSAAS motif was also found, that has been reported to be involved in forming oxide anion free radical [15]. Hence, formed oxide anion free radicals participate in redox reactions leading to oxidative modifications. The plant KASIII enzyme-specific catalytic triad composed of Cys-His-Asn was conserved in all *SiKASIII* proteins [70] (Figure S7). Although the N-terminal domains of *FatA* and *FatB* showed high divergence, the specific catalytic residues were conserved (D268, N270, H272, and E306 in *SiFatA*; D282, N284, H286, and E320 in *SiFatB-1*, and D319, N321, H323, and E357 in *SiFATB1-2*) (Figure S10).

The FAD gene family is well characterized in oilseed crops, including 31 in peanut [71], 68 in rapeseed [66], 29 in soybean [72], 40 in sunflower [67], 12 in *B. juncea*, and 8 in black mustard [73]. SAD enzymes catalyze the first desaturation in the plant FA biosynthesis pathway. The members of the SAD family have been identified in oilseed crops, including peanut, *B. juncea*, *B. rapa*, *B. nigra*, olive, and camelina, and possess 3, 12, 7, 8, 3, and 3 SAD genes, respectively [74]. Among seven *AtSAD* members, *SSI2/FAB* was characterized as a typical 18:0-ACP-specific acyl-ACP Δ^9 desaturase [18], while two members from *AtSAD*, *AtAAD2*, and *AtAAD3* catalyze 16:0-ACP Δ^9 desaturation [19]. The key amino acid residues located in the catalytic domain are predicted to shape a deep substrate-binding pocket for 18:0-ACP, while in 16:0-ACP specific enzymes, a shorter channel substrate-binding pocket was observed [75]. Based on the alignment of eight AA residues at the substrate-binding pocket, five *SiSADs* (*SiSAD1-5*) might be of 18:0-ACP specificity and two *SiSADs* (*SiSAD6* and 7) possibly exhibit 16:0-ACP specificity (Table 2; Figure S9). Of the two *SiSADs*, *SiSAD6* and 7, structural and functional association analysis of the amino acid residues (for amino acid positions refer to Table 2) reveals the presence of a bulky amino acid in *SiSAD6* (W198), which is similar to *AtAAD2* (F224), *AtAAD3* (F216), and *MuPAD* (W151), but not in *SiSAD7*. This observation favors *SiSAD6* as a better candidate for PAD than *SiSAD7*. Since the orthologous genes from different species tend to group together rather than with the

paralogous genes within a species, homology relationships (Figure 2) indicate *SiSAD6* and *AtAAD2* and 3 would have evolved their catalytic specificity independently. Hence, the *SiSAD6* is not as similar (with respect to orthologous relationship) as *BrSAD6* with *AtAAD2* or *BrSAD7* with *AtAAD3* (Figure 2). As reported in cotton (*GhA-SAD6* and *GhD-SAD8*) and Arabidopsis (*AtAAD2* and 3) for the preferential expression (of the respective gene copies) in the endosperm and aleurone tissues [19,31,75–77], the gene copies do exhibit tissue-specific selective expression patterns in addition to their specificity for 16:0-ACP or 18:0-ACP.

Membrane-bound FADs include omega-6-FADs (FAD2 and FAD6), omega-3-FADs (FAD3, FAD7 and FAD8), and palmitate desaturase (FAD4) [67]. FAD2 and FAD6 synthesize linoleic acid from oleic acid in plastids and ER, respectively [9]. In this study, three omega-6-FADs were retrieved from the sesame genome (Table S2). The presence of multiple copies of FAD2 is common in oilseed crops. For example, six FAD2 have been reported in peanut, five in soybean, six in safflower, three in sunflower, four in cotton, and two in flax [20]. Previously, FAD2 from sesame was isolated and characterized [78]. Furthermore, screening of the variants for *SiFAD2-1* from 705 accessions detected a mutation causing an amino acid change (R142H), which probably affects the desaturase activity, with mutant accessions accumulating extremely high levels of oleic acid (48%) in seed oil [25]. In the present study, a FAD2 member, designated as *SiFAD2-2*, was retrieved from the sesame genome (Table 1). The omega-3 desaturase genes (FAD3, FAD7, and FAD8) were found as single copies in sesame (Table S1).

The physicochemical properties of the FA desaturase proteins revealed that both *SiSAD* and *SiFAD* families were hydrophilic in nature. *SiFAD* proteins were predicted as alkaline proteins while *SiSADs* were slightly acidic (Table S2). The motifs identified using MEME software for the *SiFAD* genes were highly conserved within each subfamily (Figure S28), in accordance with a previous report [71]. Generally, soluble and membrane-bound desaturases were found to possess two and three histidine boxes, respectively (Table 1). The histidine boxes are essential for FAD catalytic activity [79]. The presence of conserved transmembrane domains is a typical characteristic of membrane-bound FADs [20]. The altered expression pattern of FAD family members is known to alter fatty acid profiles in various oil crops such as *Brassica* [80], cotton [81], and soybean [82]. The mechanism through which the FAD family members' regulation alters the protein and oil content and yield is not known yet.

3.2. Identification of Key Regulatory Genes

The *MCAT* gene has been isolated and characterized in rapeseed, soybean, and peanut [13,83]. *MCAT* is a key enzyme that catalyzes the first committed step in the FA pathway, and its expression level is closely associated with increased storage oil content [50]. The total seed yield and oil content were increased in transgenic *Arabidopsis* plants overexpressing *MCAT* [50].

β -ketoacyl-acyl carrier synthases (*KAS*) catalyze chain-initiation, -elongation, and -condensation steps and are classified as *KASIII*, *KASI*, and *KASII* [14]. A total of eight *KAS* gene family members, including two *KASIII*, three *KASI*, and three *KASII*, were found in sesame (Table S2). In silico genome-wide analysis of the *KAS* gene family in flax (*Linum usitatissimum* L.) led to the identification of twelve genes consisting of four *KASIII*, six *KASI*, and two *KASII* [84]. *KAS* genes are reported to affect oil production. *KASI*-deficient mutant *Arabidopsis* seeds have been reported to produce significantly lower oil content [85]. Overexpression of *NtKASI-1* significantly enhanced oil accumulation in tobacco [15]. Virus-induced gene silencing in *Jatropha curcas* demonstrated that silencing of *JcKASII* significantly altered FA composition and TAG biosynthesis [51]. Statistically significant associations for the two major genes, *KASI* (*SiFabB-2*) and *SiDGAT2*, were reported to determine the unsaturated to saturated fat ratio [25]. No variability for *SiFabB-2* at sequence level was found among the 705 sesame accessions studied [25].

Acyl-ACP thioesterases play an essential role and exhibit two distinct classes (FatA and FatB) [86]. FatA has higher specificity for unsaturated acyl groups, while FatB is more active in saturated acyl-ACPs. FatA and FatB grouped with α -CT and β -CT subunits of ACCase (Figure 2), which might be due to the presence of a conserved motif in these proteins (Figure S36). Genetic engineering of acyl-ACP thioesterases has been demonstrated to be effective for oil improvement [87]. Overexpression of *FatB1* from *Umbellularia californica* (California Bay laurel) and *MFatB* from *Madhuca longifolia* in *Brassica juncea* increased the laurate levels by over 50% and stearate levels by 16-fold, respectively [87,88]. A genome-wide association study of 705 sesame accessions reported that the candidate genes underlying the variation in FA composition and oil content include *SiFatB*, *SiFatA*, *KASII*, *KASI*, and *SAD* [25]. Based on homology relationships, we identified them as *SiFatB1-1*, *SiFatA*, *SiFabF-1*, *SiFabB-2*, and *SiSAD-1*, respectively. Availability of the sequence information is much helpful in identifying, on a homology basis, the actual gene copy (among the gene copies in a genome) being transcriptionally regulated with tissue (space) and developmental stage (time) specificity. Hence, the evolutionary or comparative genomics tools are helpful in establishing a homology relationship and, thereby, predict the actual gene copy being specifically expressed with reference to space and time in an organism. Unraveling such biological information on a homology-basis leads to novel insights due to the exceptional conservation of synteny among ortholog blocks [89–94]. Adding such homology-based information through bioinformatic tools enriches our understanding on the expression patterns for the biological pathways, with reference to space and time, in the organism of study.

TAGs are the major form of energy storage in plants and contribute to many specific developmental stages of the plant [1]. Long-chain acyl-coenzyme A synthetase (LACS) catalyzes the formation of acyl-CoAs from free fatty acids, which is pivotal for TAG biosynthesis [33]. LACS enzymes esterify free fatty acids to fatty acyl-CoA thioesters [95]. These fatty acyl-CoA thioesters are utilized in many metabolic pathways for FA elongation, triacylglycerols, membrane lipids, wax, cutin, and suberin as well as in FA catabolism [33]. *Arabidopsis* contains one of the largest known LACS families, having nine LACS genes [95]. In sesame, we found six LACS genes. In comparison to *Arabidopsis*, three LACS types, *LACS3*, *LACS5*, and *LACS7*, were absent in sesame (Table S2). Likewise, substantial variation in 629 LACS homologs from 122 species was reported [96].

The functions of most of the LACS members in *Arabidopsis* are well characterized. *AtLACS9* functions redundantly with either *AtLACS1* or *AtLACS4* in seed TAG biosynthesis [16]. *AtLACS-6* and *-7* localize in the peroxisome and are involved in β -oxidation. *AtLACS-1*, *-2*, and *-4* are involved in surface wax and cutin biosynthesis [16]. The overall functions of *AtLACS3* and *AtLACS5* are unknown yet [96]. In oilseed crops, LACS from sunflower, *HaLACS1* and *HaLACS2*, showed high sequence homology with *AtLACS9* and *AtLACS8* and are essential in oil production [97]. *BnLACS2* from rapeseed is predominantly expressed during seed development and is involved in seed oil synthesis [98]. The specific gene homologs of LACS and DGAT in flax, *LuLACS8A* and *LuDGAT2*, respectively, have been reported to contribute in the enrichment of flaxseed oil with α -linolenic acid [99].

Plant GPATs are classified into three types, based on location, as membrane-bound, chloroplastic, and cytosolic [100]. In *Arabidopsis*, ten GPATs were extensively characterized, revealing their multifaceted functions [101]. Based on acyl-transfer specificity of *AtGPATs*, they are categorized into sn-1- (*AtATS1* and *AtGPAT9*), GPATs that catalyze esterification of fatty acyl moiety from acyl-CoA (or -ACP) to the Sn-1 position of the G3P, and sn-2-GPAT (*AtGPAT1-8*), which esterifies the Sn-2 position of the G3P [102]. Based on their functions, *AtGPATs* are classified into three different groups, cutin (*AtGPAT4*, 6, and 8), suberin (*AtGPAT5* and 7), and mitochondria-localized (*AtGPAT1-3*) [103]. *AtGPAT1* was reported to be involved in tapetum differentiation and male fertility [104], while the functions of *AtGPAT2* and *GPAT3* are yet unknown [100]. Plastidial *AtATS1* and ER-localized *AtGPAT9* contribute to glycerolipid biosynthesis but not extracellular lipids [52]. GPAT family members in *Brassica* are well studied [105]. Similarly, in *Gossypium* species, 28,

27, 16, and 14 GPAT homologs were identified from *G. hirsutum*, *G. barbadense*, *G. arboreum*, and *G. raimondii*, respectively [106]. In sesame, SiGPAT10–SiGPAT17 represent the cutin group (Figure 2), while the suberin group contains SiGPAT8 and SiGPAT9. Among three mitochondrially targeted GPATs, AtGPAT1 is found to have homology with SiGPAT2, while AtGPAT-2 and -3 are grouped along with three SiGPATs (SiGPAT-3, -4, and -7) (Figure 2). SiGPAT1 groups with AtATS while SiGPAT5 and SiGPAT6 group with AtGPAT9 (Figure 2). Therefore, it could be inferred that sesame has two GPATs (SiGPAT5 and SiGPAT6) that may be involved in seed TAG biosynthesis.

LPAAT incorporates FA at the sn-2 position of PA and is a crucial enzyme in membrane phospholipid and storage lipids biosynthesis [107]. In plants, LPAATs are classified into two classes, plastid LPAATs and ER LPAATs, with the ER LPAATs further classified into Class-A and Class-B [108]. In *Arabidopsis*, five LPAATs have been reported, one plastidial (AtLPAAT1) and four (AtLPAAT2–5) microsomal targeted [109]. Class-A LPAATs are ubiquitously present in all parts of most of the plants and show substrate specificity with C18 unsaturated FAs [107], whereas class-B LPAATs are expressed in seeds and are associated with acylation of unusual acyl-CoAs [110]. Among eight sesame LPAATs, one (SiLPAAT1) was identified to be plastid localized in this study and exhibits similarity with *Arabidopsis* AtLPAAT1 (Figure 2). We have identified the protein annotated in the NCBI protein database as 1-acyl-sn-glycerol-3-phosphate acyltransferase (XP_020554192.1) as SiLPAATB because it was found to be homologous with class-B LPAAT of *Ricinus communis* (RcLPAATB) (Figure 2). The microsomal LPAAT2/3 group includes SiLPAAT2-1 and SiLPAAT2-2, along with AhLPAAT2 [53], BnLPAAT2 of *B. napus* [54], and AtLPAAT2 of *Arabidopsis* [110] (Figure 2). We retrieved SiLPAAT3 from the sesame genome, a probable ortholog of the male gametophyte isoenzyme AtLPAAT3 characterized in *Arabidopsis* [110] (Figure 2). Overexpression of LPAAT genes has proven to increase seed oil content. Specifically, LPAAT2 from rapeseed and peanut enhanced oil content by 13% and 7.4%, respectively, when transformed in *Arabidopsis* seeds [53,54]. The overexpression of yeast SLC1 and SLC1-1 genes (homologs of ER AtLPAATs) in *Arabidopsis*, soybean, and rapeseed resulted in an 8–48% increase in seed oil content [111].

The dephosphorylation of PA to produce DAG catalyzed by PAPs is a committed step in TAG biosynthesis [112]. Transient expression of N-terminal green fluorescent protein (GFP) fused AtPAH1 in tobacco leaves showed signals in cytosol and is predicted to become localized to the ER membrane to bind with PA [113], which is consistent with previous reports in yeast Pah1 and *Tetrahymena thermophila* TtPah1 [114]. A double mutant of the two AtPAH genes (AtPAH1 and 2) is affected in phosphatidylcholine production at the ER [113]. These two proteins have been reported as being involved in phospholipid homeostasis in the ER and in TAG synthesis, especially under N starvation [115].

DGATs are studied for enhancing oil production as they are one of the rate-limiting factors in plant storage lipid accumulation. Three structurally unrelated classes of DGATs were identified in plants: DGAT1, DGAT2, and DGAT3. DGAT1 and DGAT2 are associated predominantly with membranes, while DGAT3 is cytosolic [6,116]. Three members of the DGAT1 and four members of the DGAT2 family were found in soybean [60], and seven DGAT1, eight DGAT2 and three DGAT3 were identified in peanut [117].

The putative C-terminal ER retrieval motif is detectable in SiDGAT1 (–YYHDL) and in other plant DGAT1s [118]. These putative ER retrieval motifs possess hydrophobic amino acid residues [119]. The SiDGAT2 was found to have five conserved domains (Figure S17) as signature motifs within the DGAT2 subfamily. The Pfam analysis revealed that SiDGAT1 is a membrane-bound O-transferase (MBOAT) protein with nine transmembrane domains, consistent with the DGAT1 proteins of *Arabidopsis*, *B. napus*, castor, peanut, and soybean, as reported earlier [116]. In SiDGAT2, two transmembrane domains at the N-terminal of the protein were predicted (Figure S1), as in the other characterized plant DGAT2 proteins [120]. Overexpression of DGAT1 genes from *Arabidopsis* and *Camelina sativa* in *Arabidopsis*, rapeseed, and *Camelina* led to a 5–25% increase in storage oil content [55,56,121]. Similar results were documented when DGAT1 from *Tropaeolum majus* was transformed

into *B. napus* [57] and *DGAT1* from *B. napus* and *Sesamum indicum* was transformed into *Arabidopsis* [58,122]. Overexpression of soybean and oil palm *DGAT2* in *Arabidopsis* also increased TAG biosynthesis [59,60].

3.3. Validation of the Identified Regulatory Genes through RNA-Seq Studies

Oil biosynthesis in plants is performed through de novo fatty acid biosynthesis, TAG assembly, and lipid droplet biogenesis ([5] and citations therein). *LPAAT* and *WSD1/DGAT* are known to regulate key steps of TAG or lipid biosynthesis in oil seed plants [123,124]. Higher expression levels of the genes involved in TAG biosynthesis such as *LPAAT*, *GPAT*, and *DGAT* were known to be associated with enhanced oil accumulation in the seeds of the oilseed crops [24,123,125]. *SAD* and *FAD2* desaturate fatty acyl chains of unsaturated fatty acids. *SAD* desaturates stearyl-ACP (18:0) to form oleic acid (18:1), and *FAD2* catalyzes synthesis of linoleic acid (18:2) from oleic acid (18:1) [24]. The expression pattern of *SAD* and *FAD2* corresponds with oleic acid (18:1) and linoleic acid (18:2) at the early and middle stages of developing seeds of sesame [126]. The oleosins have been linked to lipids' biosynthesis/metabolism and the size of seed oil bodies [127,128]. The expression patterns observed for oleosins are consistent with their critical roles in oil body biogenesis.

Validation of the identified key regulatory genes for oil biosynthesis was completed using RNA-seq data. The transcriptome dataset used for this study was downloaded from publicly available datasets for the high and low oil yielding genotypes at four developmental stages [24]. Our results indicate that there are 27 genes (of 501 genes that are significantly differentially expressed between the high- and low-oil yielding genotypes) associated with oil biosynthesis and lipid droplet biogenesis (Table S4; Figure 5). These include those identified through our genome-wide in silico studies. In general, oil accumulation is said to increase rapidly during seed maturation [129]. Transcriptomic studies performed by Wang et al. (2019) underscore the importance of higher expression levels of oil biosynthetic genes at 30 DPA for increased oil accumulation in high oil yielding genotypes, compared to the low oil yielding genotypes [24]. In addition to this, our analysis using the same dataset suggests that early onset of higher expression levels for oil biosynthetic genes is also crucial for higher oil yield. Combining these two, early onset by 10 DPA and sustained expression until 30 DPA, the high oil yielding genotype accommodates a broader window of time for active oil accumulation when compared to the low oil yielding ones. Hence, for higher oil yield, in addition to expression patterns at 30 DPA, expression for genes of the oil biosynthesis pathway to be triggered during the early seed developmental stage (10 DPA) itself also seems to be a determining factor, as identified through the study. This helps in oil accumulating for a longer duration within the anthesis-capsule maturity window, thereby yielding higher oil content in the seeds at maturity.

4. Material and Methods

4.1. Sequence Retrieval

Using the available sesame genome sequence information [25,130], the gene family members involved in oil biosynthesis were retrieved. The sequence information for genes (and their translated amino acid sequences, Table S1 and Seq_File_S1) in oil biosynthesis in *Arabidopsis*, peanut, soybean, and other plant species was also retrieved. Transcriptome raw reads (90 bp paired-end sequencing using the Illumina HiSeq 2000 platform) of 12 samples representing different stage of seed development (10, 20, 25, and 30 DPA) for 1 high (ZMZ4728) oil- and 2 low (ZMZ3495 and ZMZ2161) oil-yielding genotypes were downloaded from the GenBank SRA database [24] for validation studies.

4.2. Physicochemical Analysis

Retrieved translated amino acid sequences were analyzed using the ExPASy ProtParam tool to estimate the molecular weight (MW), isoelectric point (pI), grand average of hydropathy (GRAVY), aliphatic index, and instability index [131]. The subcellular localization was predicted using DeepLoc [132]. TargetP-2.0 and ChloroP 1.1 were used to

predict signal peptides. The transmembrane structures were predicted using TMHMM-2.0 (<http://www.cbs.dtu.dk/services/>; accessed on 28 September 2022).

4.3. Phylogenetic Analysis and Conserved Motifs Screening

The conserved domains for the retrieved proteins were also identified from the data reported in the NCBI conserved domain database [133]. Domain analysis of the retrieved proteins was performed using the Pfam database [134]. Retrieved sequences were also verified using a simple modular architecture research tool (SMART, <http://smart.embl.de>; accessed on 28 September 2022), an annotation resource [135]. Multiple sequence alignments (MSA) of the retrieved protein sequences were performed using the Clustal [136] with default parameters. Sequence analyses were conducted in MEGA v10 [27] with a bootstrap value set to 10,000. The MEME v4.12.0 [137] server was utilized to identify the conserved amino acid motifs (Table S1) with the following setting: the maximum number of motifs—ten and minimum motif width—six.

4.4. Analysis of Cis-Acting Regulatory Elements

The retrieved candidate promoter regions (3 kb upstream) were used to predict the potential *cis*-acting regulatory elements (CAREs) using the PlantCare tool (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>; accessed on 28 September 2022). Function annotation of identified CAREs was retrieved from the PlantCare database (Lescot et al. 2002).

4.5. Analysis of CAREs for Oil Related Transcription Factors Binding Sites (TFbs)

A total of 11 TFs were identified to be involved in regulating seed oil accumulation [37,38,40,138]. Potential TF binding sites in all retrieved promoters sequences were identified using PLACE [36] (<http://www.dna.affrc.go.jp/PLACE/>; accessed on 28 September 2022) and PlantPAN 3.0 [139] (<http://plantpan.itps.ncku.edu.tw/index.html>; accessed on 28 September 2022).

4.6. RNA-Seq Analysis between High and Low Oil Content Yielding Sesame Genotypes

The fastq files of raw reads [24] were processed to check the quality parameters using FastQC [140]. The adapter, low-quality sequences, were trimmed using sickle (v1.33) [141]. The reads were subjected to a quality check using FastQC post-trimming. The filtered high-quality reads were mapped against the reference sesame genome [130] using TopHat (v2.1.1) [142] with default parameters. A reference-guided assembly of the transcriptome data from all 12 samples was performed using Cufflinks (v2.1.1) [143], and a consensus assembly was generated using Cuffmerge. The differentially expressed genes were identified using cuffdiff [144] with default parameters. The transcripts exhibiting differences of at least two-fold with FDR value ≤ 0.05 were considered to be significantly differentially expressed. The transcriptome database (cDNA sequences) for sesame genotype *Zhongzhi-13* [130] was downloaded and functionally annotated using omicsbox (v1.4.11) [145]. All significantly differentially expressed gene IDs were searched in the annotated file to assign the functionality of the unknown differentially expressed genes.

5. Conclusions

We report here on the predicted physicochemical properties, subcellular locations, conserved sites, and homology relationships for the gene products involved in oil biosynthesis in sesame, using the available genome information. Of eight SAD members, which are known as stearyl-(acyl-carrier-protein) 9-desaturases, two members were predicted to be possibly 16:0-ACP specific (Table 2). Similarly, *SiLPAAT1*, previously uncharacterized, has been functionally characterized by its homology to LPAAT1 from *Arabidopsis* (Figure 2). Moreover, *SiLPAATB*, previously annotated as 1-acyl-sn-glycerol-3-phosphate acyltransferase, is a class-B LPAAT, identified through a homology relationship with *Ricinus communis* *RcLPAATB* (Figure 2). The genome-wide in silico mining revealed key regulatory genes

associated with the oil biosynthesis pathway (Table 3). Validation of these genes was performed through RNA-seq approaches using the publicly available transcriptome dataset [24]. Our validation studies underscored the requirement to trigger the early onset of oil biosynthesis during seed development to have a longer duration of oil accumulation, resulting in higher oil yield. This is especially required when the anthesis-capsule maturity window period does not vary much. Therefore, identifying genotypes with early onset (prior 10 DPA) and sustained expression patterns until 30 DPA for the genes involving oil biosynthesis would help enhance oil yield in the oilseed crop sesame.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants11212980/s1>, Figures S1–S36: ‘Fig_S1-S36_revAug2022.pptx’; Sequence file S1: ‘Seq_File_S1.docx’; Table S1: ‘Table_S1.docx’; Table S2: ‘Table_S2.docx’; Table S3: ‘Table_S3.xlsx’; Table S4: ‘Table_S4.xlsx’;

Author Contributions: Conceptualization and supervision, P.R.; formal investigation and analysis, B.N., A.K. and R.M.; methodology, B.N., A.K., R.M. and P.R.; Resources, P.R., R.Y. and R.S.; writing—original draft, B.N., A.K. and P.R.; Writing—review and editing, P.R., B.N., A.K., R.S., R.Y. and K.S.; project administration, P.R.; funding acquisition, K.S., R.Y. and P.R. All authors have read and agreed to the published version of the manuscript.

Funding: Department of Biotechnology, Government of India (grant id: 16113200037-1012166).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data pertaining to the study reported here are provided in the supplementary information.

Acknowledgments: The authors KS, RY, PR, and RS are thankful to the Department of Biotechnology, Government of India, for funding through research grant number-code 16113200037-1012166. The authors are thankful to the anonymous reviewers for their critical suggestions and input.

Conflicts of Interest: All authors declare that there are no direct or indirect, including financial and non-financial, competing interests, with reference to this article.

References

1. Bates, P.D.; Stymne, S.; Ohlrogge, J. Biochemical pathways in seed oil synthesis. *Curr. Opin. Plant Biol.* **2013**, *16*, 358–364. [CrossRef] [PubMed]
2. Lu, C.; Napier, J.A.; Clemente, T.E.; Cahoon, E.B. New frontiers in oilseed biotechnology: Meeting the global demand for vegetable oils for food, feed, biofuel, and industrial applications. *Curr. Opin. Biotechnol.* **2011**, *22*, 252–259. [CrossRef] [PubMed]
3. Bruinsma, J. *World Agriculture: Towards 2015/2030—An FAO Perspective*; Taylor and Francis: London, UK, 2017; p. 444.
4. Banaś, W.; Sanchez Garcia, A.; 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] [PubMed]
5. Rangan, P.; Maurya, R.; Singh, S. Can omic tools help generate alternative newer sources of edible seed oil? *Plant Direct* **2022**, *6*, e399. [CrossRef] [PubMed]
6. Saha, S.; Enugutti, B.; Rajakumari, S.; Rajasekharan, R. Cytosolic triacylglycerol biosynthetic pathway in oilseeds. Molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *Plant Physiol.* **2006**, *141*, 1533–1543. [CrossRef] [PubMed]
7. Bhatt-Wessel, B.; Jordan, T.W.; Miller, J.H.; Peng, L. Role of DGAT enzymes in triacylglycerol metabolism. *Arch. Biochem. Biophys.* **2018**, *655*, 1–11. [CrossRef] [PubMed]
8. Nikolau, B.J.; Ohlrogge, J.B.; Wurtele, E.S. Plant biotin-containing carboxylases. *Arch. Biochem. Biophys.* **2003**, *414*, 211–222. [CrossRef]
9. Ohlrogge, J.; Browse, J. Lipid biosynthesis. *Plant Cell* **1995**, *7*, 957–970.
10. Bates, P.D.; Browse, J. The significance of different diacylglycerol synthesis pathways on plant oil composition and bioengineering. *Front. Plant Sci.* **2012**, *3*, 147. [CrossRef]
11. Sasaki, Y.; Nagano, Y. Plant acetyl-CoA carboxylase: Structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 1175–1184. [CrossRef]
12. Reverdatto, S.; Beilinson, V.; Nielsen, N.C. A multisubunit acetyl coenzyme A carboxylase from soybean. *Plant Physiol.* **1999**, *119*, 961–978. [CrossRef] [PubMed]

13. Li, M.-J.; Li, A.-Q.; Xia, H.; Zhao, C.-Z.; Li, C.-S.; Wan, S.-B.; Bi, Y.-P.; Wang, X.-J. Cloning and sequence analysis of putative type II fatty acid synthase genes from *Arachis hypogaea* L. *J. Biosci.* **2009**, *34*, 227–238. [\[CrossRef\]](#)
14. White, S.W.; Zheng, J.; Zhang, Y.-M.; Rock, C.O. The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* **2005**, *74*, 791–831. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Yang, T.; Xu, R.; Chen, J.; Liu, A. β -Ketoacyl-acyl carrier protein synthase I (KASI) plays crucial roles in the plant growth and fatty acids synthesis in tobacco. *Int. J. Mol. Sci.* **2016**, *17*, 1287. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jessen, D.; Roth, C.; Wiermer, M.; Fulda, M. Two activities of long-chain acyl-coenzyme A synthetase are involved in lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis. *Plant Physiol.* **2015**, *167*, 351–366. [\[CrossRef\]](#)
17. Dong, C.-J.; Cao, N.; Zhang, Z.-G.; Shang, Q.-M. Characterization of the fatty acid desaturase genes in cucumber: Structure, phylogeny, and expression patterns. *PLoS ONE* **2016**, *11*, e0149917. [\[CrossRef\]](#)
18. Lightner, J.; Wu, J.; Browse, J. A mutant of Arabidopsis with increased levels of stearic acid. *Plant Physiol.* **1994**, *106*, 1443–1451. [\[CrossRef\]](#)
19. Bryant, F.M.; Munoz-Azcarate, O.; Kelly, A.A.; Beaudoin, F.; Kurup, S.; Eastmond, P.J. ACYL-ACYL CARRIER PROTEIN DESATURASE2 and 3 are responsible for making omega-7 fatty acids in the Arabidopsis aleurone. *Plant Physiol.* **2016**, *172*, 154–162. [\[CrossRef\]](#)
20. Dar, A.A.; Choudhury, A.R.; Kancharla, P.K.; Arumugam, N. The FAD2 gene in plants: Occurrence, regulation, and role. *Front. Plant Sci.* **2017**, *8*, 1789. [\[CrossRef\]](#)
21. Kennedy, E.P. Biosynthesis of complex lipids. *Fed. Proc.* **1961**, *20*, 934–940.
22. Dossa, K.; Diouf, D.; Wang, L.; Wei, X.; Zhang, Y.; Niang, M.; Fonceka, D.; Yu, J.; Mmadi, M.A.; Yehouessi, L.W. The emerging oilseed crop *Sesamum indicum* enters the “Omics” era. *Front. Plant Sci.* **2017**, *8*, 1154. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Bedigian, D. History and lore of sesame in Southwest Asia. *Econ. Bot.* **2004**, *58*, 329–353. [\[CrossRef\]](#)
24. Wang, L.; Zhang, Y.; Li, D.; Dossa, K.; Wang, M.L.; Zhou, R.; Yu, J.; Zhang, X. Gene expression profiles that shape high and low oil content sesames. *BMC Genet.* **2019**, *20*, 45. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Wei, X.; Liu, K.; Zhang, Y.; Feng, Q.; Wang, L.; Zhao, Y.; Li, D.; Zhao, Q.; Zhu, X.; Zhu, X. Genetic discovery for oil production and quality in sesame. *Nat. Commun.* **2015**, *6*, 8609. [\[CrossRef\]](#)
26. Ikai, A. Thermostability and aliphatic index of globular proteins. *J. Biochem.* **1980**, *88*, 1895–1898.
27. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [\[CrossRef\]](#)
28. Keatinge-Clay, A.T.; Shelat, A.A.; Savage, D.F.; Tsai, S.-C.; Miercke, L.J.; O’Connell, J.D., III; Khosla, C.; Stroud, R.M. Catalysis, specificity, and ACP docking site of *Streptomyces coelicolor* malonyl-CoA: ACP transacylase. *Structure* **2003**, *11*, 147–154. [\[CrossRef\]](#)
29. Von Wettstein-Knowles, P.; Olsen, J.G.; McGuire, K.A.; Henriksen, A. Fatty acid synthesis: Role of active site histidines and lysine in Cys-His-His-type β -ketoacyl-acyl carrier protein synthases. *FEBS J.* **2006**, *273*, 695–710. [\[CrossRef\]](#)
30. Lindqvist, Y.; Huang, W.; Schneider, G.; Shanklin, J. Crystal structure of delta9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. *EMBO J.* **1996**, *15*, 4081–4092. [\[CrossRef\]](#)
31. Troncoso-Ponce, M.A.; Barthole, G.; Tremblais, G.; To, A.; Miquel, M.; Lepiniec, L.; Baud, S. Transcriptional activation of two delta-9 palmitoyl-ACP desaturase genes by MYB115 and MYB118 is critical for biosynthesis of omega-7 monounsaturated fatty acids in the endosperm of Arabidopsis seeds. *Plant Cell* **2016**, *28*, 2666–2682. [\[CrossRef\]](#)
32. Cahoon, E.B.; Shah, S.; Shanklin, J.; Browse, J. A determinant of substrate specificity predicted from the acyl-acyl carrier protein desaturase of developing cat’s claw seed. *Plant Physiol.* **1998**, *117*, 593–598. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Wu, T.; Fu, Y.; Shi, Y.; Li, Y.; Kou, Y.; Mao, X.; Liu, J. Functional characterization of long-chain acyl-CoA synthetase gene family from the oleaginous alga *Chromochloris zofingiensis*. *J. Agric. Food Chem.* **2020**, *68*, 4473–4484. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Mañas-Fernández, A.; Li-Beisson, Y.; Alonso, D.L.; García-Maroto, F. Cloning and molecular characterization of a glycerol-3-phosphate O-acyltransferase (GPAT) gene from *Echium* (Boraginaceae) involved in the biosynthesis of cutin polyesters. *Planta* **2010**, *232*, 987–997. [\[CrossRef\]](#)
35. Koerbes, A.P.; Kulcheski, F.R.; Margis, R.; Margis-Pinheiro, M.; Turchetto-Zolet, A.C. Molecular evolution of the lysophosphatidic acid acyltransferase (LPAAT) gene family. *Mol. Phylogenetics Evol.* **2016**, *96*, 55–69. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Higo, K.; Ugawa, Y.; Iwamoto, M.; Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **1999**, *27*, 297–300. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Kumar, N.; Chaudhary, A.; Singh, D.; Teotia, S. Transcriptional regulation of seed oil accumulation in *Arabidopsis thaliana*: Role of transcription factors and chromatin remodelers. *J. Plant Biochem. Biotechnol.* **2020**, *29*, 754–768. [\[CrossRef\]](#)
38. Niu, Y.; Zhang, G.; Wan, F.; Zhang, Y.-M. Integration of RNA-Seq profiling with genome-wide association study predicts candidate genes for oil accumulation in soybean. *Crop Pasture Sci.* **2020**, *71*, 996–1009. [\[CrossRef\]](#)
39. Mehrotra, R.; Jain, V.; Shekhar, C.; Mehrotra, S. Genome wide analysis of *Arabidopsis thaliana* reveals high frequency of AAAGN7CTTT motif. *Meta Gene* **2014**, *2*, 606–615. [\[CrossRef\]](#)
40. Wang, H.W.; Zhang, B.; Hao, Y.J.; Huang, J.; Tian, A.G.; Liao, Y.; Zhang, J.S.; Chen, S.Y. The soybean Dof-type transcription factor genes, *GmDof4* and *GmDof11*, enhance lipid content in the seeds of transgenic Arabidopsis plants. *Plant J.* **2007**, *52*, 716–729. [\[CrossRef\]](#)

41. Mu, J.; Tan, H.; Zheng, Q.; Fu, F.; Liang, Y.; Zhang, J.; Yang, X.; Wang, T.; Chong, K.; Wang, X.-J. *LEAFY COTYLEDON1* is a key regulator of fatty acid biosynthesis in Arabidopsis. *Plant Physiol.* **2008**, *148*, 1042–1054. [\[CrossRef\]](#)
42. Baud, S.; Mendoza, M.S.; To, A.; Harscoët, E.; Lepiniec, L.; Dubreucq, B. *WRINKLED1* specifies the regulatory action of *LEAFY COTYLEDON2* towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant J.* **2007**, *50*, 825–838. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Behera, J.R.; Rahman, M.M.; Bhatia, S.; Shockey, J.; Kilaru, A. Functional and predictive structural characterization of *WRINKLED2*, a unique oil biosynthesis regulator in avocado. *Front. Plant Sci.* **2021**, *12*, 648494. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Lee, H.G.; Kim, H.; Suh, M.C.; Kim, H.U.; Seo, P.J. The MYB96 transcription factor regulates triacylglycerol accumulation by activating *DGAT1* and *PDAT1* expression in Arabidopsis seeds. *Plant Cell Physiol.* **2018**, *59*, 1432–1442. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Lee, K.; Lee, H.G.; Yoon, S.; Kim, H.U.; Seo, P.J. The Arabidopsis MYB96 transcription factor is a positive regulator of *ABSCISIC ACID-INSENSITIVE4* in the control of seed germination. *Plant Physiol.* **2015**, *168*, 677–689. [\[CrossRef\]](#)
46. Kong, Y.; Chen, S.; Yang, Y.; An, C. ABA-insensitive (*ABI*) 4 and *ABI5* synergistically regulate *DGAT1* expression in Arabidopsis seedlings under stress. *FEBS Lett.* **2013**, *587*, 3076–3082. [\[CrossRef\]](#)
47. Cui, Y.; Liu, Z.; Zhao, Y.; Wang, Y.; Huang, Y.; Li, L.; Wu, H.; Xu, S.; Hua, J. Overexpression of heteromeric GhACCase subunits enhanced oil accumulation in upland cotton. *Plant Mol. Biol. Report.* **2017**, *35*, 287–297. [\[CrossRef\]](#)
48. Cui, Y.; Zhao, Y.; Wang, Y.; Liu, Z.; Ijaz, B.; Huang, Y.; Hua, J. Genome-wide identification and expression analysis of the biotin carboxyl carrier subunits of heteromeric acetyl-CoA carboxylase in *Gossypium*. *Front. Plant Sci.* **2017**, *8*, 624. [\[CrossRef\]](#)
49. Madoka, Y.; Tomizawa, K.-I.; Mizoi, J.; Nishida, I.; Nagano, Y.; Sasaki, Y. Chloroplast transformation with modified *accD* operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco. *Plant Cell Physiol.* **2002**, *43*, 1518–1525. [\[CrossRef\]](#)
50. Jung, S.H.; Kim, R.J.; Kim, K.J.; Lee, D.H.; Suh, M.C. Plastidial and mitochondrial malonyl CoA-ACP malonyltransferase is essential for cell division and its overexpression increases storage oil content. *Plant Cell Physiol.* **2019**, *60*, 1239–1249. [\[CrossRef\]](#)
51. Ye, J.; Qu, J.; Bui, H.T.N.; Chua, N.H. Rapid analysis of *Jatropha curcas* gene functions by virus-induced gene silencing. *Plant Biotechnol. J.* **2009**, *7*, 964–976. [\[CrossRef\]](#)
52. Shockey, J.; Regmi, A.; Cotton, K.; Adhikari, N.; Browse, J.; Bates, P.D. Identification of Arabidopsis *GPAT9* (At5g60620) as an essential gene involved in triacylglycerol biosynthesis. *Plant Physiol.* **2016**, *170*, 163–179. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Chen, S.; Lei, Y.; Xu, X.; Huang, J.; Jiang, H.; Wang, J.; Cheng, Z.; Zhang, J.; Song, Y.; Liao, B. The peanut (*Arachis hypogaea* L.) gene *AhLPAT2* increases the lipid content of transgenic Arabidopsis seeds. *PLoS ONE* **2015**, *10*, e0136170. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Maisonneuve, S.; Bessoule, J.-J.; Lessire, R.; Delseny, M.; Roscoe, T.J. Expression of rapeseed microsomal lysophosphatidic acid acyltransferase isozymes enhances seed oil content in Arabidopsis. *Plant Physiol.* **2010**, *152*, 670–684. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Jako, C.; Kumar, A.; Wei, Y.; Zou, J.; Barton, D.L.; Giblin, E.M.; Covello, P.S.; Taylor, D.C. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol.* **2001**, *126*, 861–874. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Kim, H.; Park, J.H.; Kim, D.J.; Kim, A.Y.; Suh, M.C. Functional analysis of *diacylglycerol acyltransferase1* genes from *Camelina sativa* and effects of *CsDGAT1B* overexpression on seed mass and storage oil content in *C. sativa*. *Plant Biotechnol. Rep.* **2016**, *10*, 141–153. [\[CrossRef\]](#)
57. Xu, J.; Francis, T.; Mietkiewska, E.; Giblin, E.M.; Barton, D.L.; Zhang, Y.; Zhang, M.; Taylor, D.C. Cloning and characterization of an acyl-CoA-dependent diacylglycerol acyltransferase 1 (*DGAT1*) gene from *Tropaeolum majus*, and a study of the functional motifs of the *DGAT* protein using site-directed mutagenesis to modify enzyme activity and oil content. *Plant Biotechnol. J.* **2008**, *6*, 799–818. [\[CrossRef\]](#)
58. Weselake, R.J.; Shah, S.; Tang, M.; Quant, P.A.; Snyder, C.L.; Furukawa-Stoffer, T.L.; Zhu, W.; Taylor, D.C.; Zou, J.; Kumar, A. Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. *J. Exp. Bot.* **2008**, *59*, 3543–3549. [\[CrossRef\]](#)
59. Jin, Y.; Yuan, Y.; Gao, L.; Sun, R.; Chen, L.; Li, D.; Zheng, Y. Characterization and functional analysis of a type 2 diacylglycerol acyltransferase (*DGAT2*) gene from oil palm (*Elaeis guineensis* Jacq.) mesocarp in *Saccharomyces cerevisiae* and transgenic *Arabidopsis thaliana*. *Front. Plant Sci.* **2017**, *8*, 1791. [\[CrossRef\]](#)
60. Chen, B.; Wang, J.; Zhang, G.; Liu, J.; Manan, S.; Hu, H.; Zhao, J. Two types of soybean diacylglycerol acyltransferases are differentially involved in triacylglycerol biosynthesis and response to environmental stresses and hormones. *Sci. Rep.* **2016**, *6*, 28541. [\[CrossRef\]](#)
61. Liang, Y.; Huang, Y.; Chen, K.; Kong, X.; Li, M. Characterization of non-specific lipid transfer protein (*nsLtp*) gene families in the *Brassica napus* pangenome reveals abundance variation. *BMC Plant Biol.* **2022**, *22*, 21. [\[CrossRef\]](#)
62. Alotaibi, S.S.; Elseehy, M.M.; Aljuaid, B.S.; El-Shehawi, A.M. Transcriptome analysis of jojoba (*Simmondsia chinensis*) during seed development and liquid wax ester biosynthesis. *Plants* **2020**, *9*, 588. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Javelle, M.; Vernoud, V.; Rogowsky, P.M.; Ingram, G.C. Epidermis: The formation and functions of a fundamental plant tissue. *New Phytol.* **2011**, *189*, 17–39. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Samuels, L.; Kunst, L.; Jetter, R. Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* **2008**, *59*, 683–707. [\[CrossRef\]](#)

65. Nováková, E.; Zablatzká, L.; Brus, J.; Nesrstová, V.; Hanáček, P.; Kalendar, R.; Cvrčková, F.; Majeský, L.; Smýkal, P. Allelic diversity of acetyl coenzyme A carboxylase accD/bccp genes implicated in nuclear-cytoplasmic conflict in the wild and domesticated pea (*Pisum* sp.). *Int. J. Mol. Sci.* **2019**, *20*, 1773. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Xu, L.; Zeng, W.; Li, J.; Liu, H.; Yan, G.; Si, P.; Yang, C.; Shi, Y.; He, Q.; Zhou, W. Characteristics of membrane-bound fatty acid desaturase (FAD) genes in *Brassica napus* L. and their expressions under different cadmium and salinity stresses. *Environ. Exp. Bot.* **2019**, *162*, 144–156. [\[CrossRef\]](#)
67. Li, J.; Liu, A.; Najeeb, U.; Zhou, W.; Liu, H.; Yan, G.; Gill, R.A.; Yun, X.; Bai, Q.; Xu, L. Genome-wide investigation and expression analysis of membrane-bound fatty acid desaturase genes under different biotic and abiotic stresses in sunflower (*Helianthus annuus* L.). *Int. J. Biol. Macromol.* **2021**, *175*, 188–198. [\[CrossRef\]](#)
68. Heidari, P.; Ahmadizadeh, M.; Izanlo, F.; Nussbaumer, T. In silico study of the CESA and CSL gene family in *Arabidopsis thaliana* and *Oryza sativa*: Focus on post-translation modifications. *Plant Gene* **2019**, *19*, 100189. [\[CrossRef\]](#)
69. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [\[CrossRef\]](#)
70. González-Mellado, D.; von Wettstein-Knowles, P.; Garcés, R.; Martínez-Force, E. The role of β -ketoacyl-acyl carrier protein synthase III in the condensation steps of fatty acid biosynthesis in sunflower. *Planta* **2010**, *231*, 1277–1289. [\[CrossRef\]](#)
71. Chi, X.; Zhang, Z.; Chen, N.; Zhang, X.; Wang, M.; Chen, M.; Wang, T.; Pan, L.; Chen, J.; Yang, Z. Isolation and functional analysis of fatty acid desaturase genes from peanut (*Arachis hypogaea* L.). *PLoS ONE* **2017**, *12*, e0189759. [\[CrossRef\]](#)
72. Chi, X.; Yang, Q.; Lu, Y.; Wang, J.; Zhang, Q.; Pan, L.; Chen, M.; He, Y.; Yu, S. Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Mol. Biol. Report.* **2011**, *29*, 769–783. [\[CrossRef\]](#)
73. Xue, Y.; Jiang, J.; Yang, X.; Jiang, H.; Du, Y.; Liu, X.; Xie, R.; Chai, Y. Genome-wide mining and comparative analysis of fatty acid elongase gene family in *Brassica napus* and its progenitors. *Gene* **2020**, *747*, 144674. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Ahmadizadeh, M.; Rezaee, S.; Heidari, P. Genome-wide characterization and expression analysis of fatty acid desaturase gene family in *Camelina sativa*. *Gene Rep.* **2020**, *21*, 100894. [\[CrossRef\]](#)
75. Liu, B.; Sun, Y.; Xue, J.; Mao, X.; Jia, X.; Li, R. Stearoyl-ACP $\Delta 9$ desaturase 6 and 8 (GhA-SAD6 and GhD-SAD8) are responsible for biosynthesis of palmitoleic acid specifically in developing endosperm of upland cotton seeds. *Front. Plant Sci.* **2019**, *10*, 703. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Kryuchkova-Mostacci, N.; Robinson-Rechavi, M. Tissue-specificity of gene expression diverges slowly between orthologs, and rapidly between paralogs. *PLoS Comput. Biol.* **2016**, *12*, e1005274. [\[CrossRef\]](#)
77. Kryuchkova-Mostacci, N.; Robinson-Rechavi, M. A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **2017**, *18*, 205–214. [\[CrossRef\]](#)
78. Jin, U.-H.; Lee, J.-W.; Chung, Y.-S.; Lee, J.-H.; Yi, Y.-B.; Kim, Y.-K.; Hyung, N.-I.; Pyee, J.-H.; Chung, C.-H. Characterization and temporal expression of a ω -6 fatty acid desaturase cDNA from sesame (*Sesamum indicum* L.) seeds. *Plant Sci.* **2001**, *161*, 935–941. [\[CrossRef\]](#)
79. Shanklin, J.; Whittle, E.; Fox, B.G. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* **1994**, *33*, 12787–12794. [\[CrossRef\]](#)
80. Scarth, R.; Tang, J. Modification of Brassica oil using conventional and transgenic approaches. *Crop Sci.* **2006**, *46*, 1225–1236. [\[CrossRef\]](#)
81. Liu, F.; Zhao, Y.-P.; Zhu, H.-g.; Zhu, Q.-H.; Sun, J. Simultaneous silencing of GhFAD2-1 and GhFATB enhances the quality of cottonseed oil with high oleic acid. *J. Plant Physiol.* **2017**, *215*, 132–139. [\[CrossRef\]](#)
82. Clemente, T.E.; Cahoon, E.B. Soybean oil: Genetic approaches for modification of functionality and total content. *Plant Physiol.* **2009**, *151*, 1030–1040. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Qu, C.; Jiang, H.; Zhang, L. Cloning and enzymatic activity analysis of the malonyl-CoA: Acyl carrier protein transacylase in *Brassica napus* cultivars with different oil content. *Am. J. Plant Sci.* **2014**, *5*, 2599–2610. [\[CrossRef\]](#)
84. You, F.M.; Li, P.; Kumar, S.; Ragupathy, R.; Li, Z.; Fu, Y.-B.; Cloutier, S. Genome-wide identification and characterization of the gene families controlling fatty acid biosynthesis in flax (*Linum usitatissimum* L.). *J. Proteom. Bioinform* **2014**, *7*, 310–326.
85. Wu, G.-Z.; Xue, H.-W. Arabidopsis β -ketoacyl-[acyl carrier protein] synthase I is crucial for fatty acid synthesis and plays a role in chloroplast division and embryo development. *Plant Cell* **2010**, *22*, 3726–3744. [\[CrossRef\]](#)
86. Salas, J.N.J.; Ohlrogge, J.B. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Arch. Biochem. Biophys.* **2002**, *403*, 25–34. [\[CrossRef\]](#)
87. Bhattacharya, S.; Sinha, S.; Das, N.; Maiti, M.K. Increasing the stearate content in seed oil of *Brassica juncea* by heterologous expression of MIFatB affects lipid content and germination frequency of transgenic seeds. *Plant Physiol. Biochem.* **2015**, *96*, 345–355. [\[CrossRef\]](#)
88. Eccleston, V.S.; Cranmer, A.M.; Voelker, T.A.; Ohlrogge, J.B. Medium-chain fatty acid biosynthesis and utilization in *Brassica napus* plants expressing lauroyl-acyl carrier protein thioesterase. *Planta* **1996**, *198*, 46–53. [\[CrossRef\]](#)
89. Koonin, E.V. Orthologs, paralogs, and evolutionary genomics. *Annu. Rev. Genet.* **2005**, *39*, 309–338. [\[CrossRef\]](#)
90. Fujimoto, M.S.; Suvorov, A.; Jensen, N.O.; Clement, M.J.; Bybee, S.M. Detecting false positive sequence homology: A machine learning approach. *BMC Bioinform.* **2016**, *17*, 101. [\[CrossRef\]](#)
91. McGimpsey, S. The Twilight Zone of Nucleotide Homology. Ph.D. Thesis, University of Otago, Dunedin, New Zealand, 2019.

92. Pearson, W.R. An introduction to sequence similarity (“homology”) searching. *Curr. Protoc. Bioinform.* **2013**, *42*, 3.1.1–3.1.8. [\[CrossRef\]](#)
93. Movahedi, S.; Van de Peer, Y.; Vandepoele, K. Comparative network analysis reveals that tissue specificity and gene function are important factors influencing the mode of expression evolution in Arabidopsis and rice. *Plant Physiol.* **2011**, *156*, 1316–1330. [\[CrossRef\]](#) [\[PubMed\]](#)
94. Kille, B.; Balaji, A.; Sedlazeck, F.J.; Nute, M.; Treangen, T.J. Multiple genome alignment in the telomere-to-telomere assembly era. *Genome Biol.* **2022**, *23*, 182. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Shockey, J.M.; Fulda, M.S.; Browse, J.A. Arabidopsis contains nine long-chain acyl-coenzyme A synthetase genes that participate in fatty acid and glycerolipid metabolism. *Plant Physiol.* **2002**, *129*, 1710–1722. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Ayaz, A.; Saqib, S.; Huang, H.; Zaman, W.; Lü, S.; Zhao, H. Genome-wide comparative analysis of long-chain acyl-CoA synthetases (LACSs) gene family: A focus on identification, evolution and expression profiling related to lipid synthesis. *Plant Physiol. Biochem.* **2021**, *161*, 1–11. [\[CrossRef\]](#)
97. Aznar-Moreno, J.A.; Venegas Caleron, M.; Martínez-Force, E.; Garcés, R.; Mullen, R.; Gidda, S.K.; Salas, J.J. Sunflower (*Helianthus annuus*) long-chain acyl-coenzyme A synthetases expressed at high levels in developing seeds. *Physiol. Plant.* **2014**, *150*, 363–373. [\[CrossRef\]](#)
98. Ding, L.-N.; Gu, S.-L.; Zhu, F.-G.; Ma, Z.-Y.; Li, J.; Li, M.; Wang, Z.; Tan, X.-L. Long-chain acyl-CoA synthetase 2 is involved in seed oil production in *Brassica napus*. *BMC Plant Biol.* **2020**, *20*, 21. [\[CrossRef\]](#)
99. Xu, Y.; Holic, R.; Li, D.; Pan, X.; Mietkiewska, E.; Chen, G.; Ozga, J.; Weselake, R.J. Substrate preferences of long-chain acyl-CoA synthetase and diacylglycerol acyltransferase contribute to enrichment of flax seed oil with α -linolenic acid. *Biochem. J.* **2018**, *475*, 1473–1489. [\[CrossRef\]](#)
100. Bai, Y.; Shen, Y.; Zhang, Z.; Jia, Q.; Xu, M.; Zhang, T.; Fang, H.; Yu, X.; Li, L.; Liu, D. A GPAT1 Mutation in Arabidopsis Enhances Plant Height but Impairs Seed Oil Biosynthesis. *Int. J. Mol. Sci.* **2021**, *22*, 785. [\[CrossRef\]](#)
101. Chen, X.; Snyder, C.L.; Truksa, M.; Shah, S.; Weselake, R.J. sn-Glycerol-3-phosphate acyltransferases in plants. *Plant Signal. Behav.* **2011**, *6*, 1695–1699. [\[CrossRef\]](#)
102. Jayawardhane, K.N.; Singer, S.D.; Weselake, R.J.; Chen, G. Plant sn-glycerol-3-phosphate acyltransferases: Biocatalysts involved in the biosynthesis of intracellular and extracellular lipids. *Lipids* **2018**, *53*, 469–480. [\[CrossRef\]](#)
103. Yang, W.; Simpson, J.P.; Li-Beisson, Y.; Beisson, F.; Pollard, M.; Ohlrogge, J.B. A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: Substrate specificity, sn-2 preference, and evolution. *Plant Physiol.* **2012**, *160*, 638–652. [\[CrossRef\]](#) [\[PubMed\]](#)
104. Zheng, Z.; Xia, Q.; Dauk, M.; Shen, W.; Selvaraj, G.; Zou, J. Arabidopsis AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* **2003**, *15*, 1872–1887. [\[CrossRef\]](#) [\[PubMed\]](#)
105. Wang, J.; Singh, S.K.; Geng, S.; Zhang, S.; Yuan, L. Genome-wide analysis of glycerol-3-phosphate O-acyltransferase gene family and functional characterization of two cutin group GPATs in *Brassica napus*. *Planta* **2020**, *251*, 93. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Cui, Y.; Ma, J.; Liu, G.; Wang, N.; Pei, W.; Wu, M.; Li, X.; Zhang, J.; Yu, J. Genome-wide identification, sequence variation, and expression of the glycerol-3-phosphate acyltransferase (GPAT) gene family in *Gossypium*. *Front. Genet.* **2019**, *10*, 116. [\[CrossRef\]](#)
107. 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. Crops Prod.* **2020**, *150*, 112245. [\[CrossRef\]](#)
108. Frentzen, M. Acyltransferases from basic science to modified seed oils. *Lipid/Fett* **1998**, *100*, 161–166.
109. Kim, H.U.; Huang, A.H. Plastid lysophosphatidyl acyltransferase is essential for embryo development in Arabidopsis. *Plant Physiol.* **2004**, *134*, 1206–1216. [\[CrossRef\]](#)
110. Kim, H.J.; Silva, J.E.; Iskandarov, U.; Andersson, M.; Cahoon, R.E.; 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. [\[CrossRef\]](#)
111. Rao, S.S.; Hildebrand, D. Changes in oil content of transgenic soybeans expressing the yeast SLC1 gene. *Lipids* **2009**, *44*, 945–951. [\[CrossRef\]](#)
112. Nakamura, Y. Plant phospholipid diversity: Emerging functions in metabolism and protein–lipid interactions. *Trends Plant Sci.* **2017**, *22*, 1027–1040. [\[CrossRef\]](#)
113. Eastmond, P.J.; Quettier, A.-L.; Kroon, J.T.; Craddock, C.; Adams, N.; Slabas, A.R. Phosphatidic Acid Phosphohydrolase 1 and 2 regulate phospholipid synthesis at the endoplasmic reticulum in Arabidopsis. *Plant Cell* **2010**, *22*, 2796–2811. [\[CrossRef\]](#)
114. Pillai, A.N.; Shukla, S.; Rahaman, A. An evolutionarily conserved phosphatidate phosphatase maintains lipid droplet number and endoplasmic reticulum morphology but not nuclear morphology. *Biol. Open* **2017**, *6*, 1629–1643. [\[PubMed\]](#)
115. Yoshitake, Y.; Sato, R.; Madoka, Y.; Ikeda, K.; Murakawa, M.; Suruga, K.; Sugiura, D.; Noguchi, K.; Ohta, H.; Shimojima, M. Arabidopsis phosphatidic acid phosphohydrolases are essential for growth under nitrogen-depleted conditions. *Front. Plant Sci.* **2017**, *8*, 1847. [\[CrossRef\]](#) [\[PubMed\]](#)
116. Gao, H.; Gao, Y.; Zhang, F.; Liu, B.; Ji, C.; Xue, J.; Yuan, L.; Li, R. Functional characterization of a novel acyl-CoA: Diacylglycerol acyltransferase 3-3 (CsDGAT3-3) gene from *Camelina sativa*. *Plant Sci.* **2021**, *303*, 110752. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Zheng, L.; Shockey, J.; Bian, F.; Chen, G.; Shan, L.; Li, X.; Wan, S.; Peng, Z. Variant amino acid residues alter the enzyme activity of peanut type 2 diacylglycerol acyltransferases. *Front. Plant Sci.* **2017**, *8*, 1751. [\[CrossRef\]](#)

118. Chi, X.; Hu, R.; Zhang, X.; Chen, M.; Chen, N.; Pan, L.; Wang, T.; Wang, M.; Yang, Z.; Wang, Q. Cloning and functional analysis of three diacylglycerol acyltransferase genes from peanut (*Arachis hypogaea* L.). *PLoS ONE* **2014**, *9*, e105834. [\[CrossRef\]](#)
119. McCartney, A.W.; Dyer, J.M.; Dhanoa, P.K.; Kim, P.K.; Andrews, D.W.; McNew, J.A.; Mullen, R.T. Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *Plant J.* **2004**, *37*, 156–173. [\[CrossRef\]](#)
120. 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. [\[CrossRef\]](#)
121. Lu, C.L.; de Noyer, S.B.; Hobbs, D.H.; Kang, J.; Wen, Y.; Krachtus, D.; Hills, M.J. Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in triacylglycerol biosynthesis, in *Arabidopsis thaliana*. *Plant Mol. Biol.* **2003**, *52*, 31–41. [\[CrossRef\]](#)
122. Wang, Z.; Huang, W.; Chang, J.; Sebastian, A.; Li, Y.; Li, H.; Wu, X.; Zhang, B.; Meng, F.; Li, W. Overexpression of SiDGAT1, a gene encoding acyl-CoA: Diacylglycerol acyltransferase from *Sesamum indicum* L. increases oil content in transgenic *Arabidopsis* and soybean. *Plant Cell Tissue Organ Cult.* **2014**, *119*, 399–410. [\[CrossRef\]](#)
123. Liu, F.; Xia, Y.; Wu, L.; Fu, D.; Hayward, A.; Luo, J.; Yan, X.; Xiong, X.; Fu, P.; Wu, G. Enhanced seed oil content by overexpressing genes related to triacylglyceride synthesis. *Gene* **2015**, *557*, 163–171. [\[CrossRef\]](#)
124. Abdullah, H.M.; Akbari, P.; Paulose, B.; Schnell, D.; Qi, W.; Park, Y.; Pareek, A.; Dhankher, O.P. Transcriptome profiling of *Camelina sativa* to identify genes involved in triacylglycerol biosynthesis and accumulation in the developing seeds. *Biotechnol. Biofuels* **2016**, *9*, 136. [\[CrossRef\]](#) [\[PubMed\]](#)
125. Zhou, W.; Song, S.; Dossou, S.S.K.; Zhou, R.; Wei, X.; Wang, Z.; Sheng, C.; Zhang, Y.; You, J.; Wang, L. Genome-wide association analysis and transcriptome reveal novel loci and a candidate regulatory gene of fatty acid biosynthesis in sesame (*Sesamum indicum* L.). *Plant Physiol. Biochem.* **2022**, *186*, 220–231. [\[CrossRef\]](#) [\[PubMed\]](#)
126. Zhang, Y.-P.; Zhang, Y.-Y.; Thakur, K.; Zhang, F.; Hu, F.; Zhang, J.-G.; Wei, P.-C.; Wei, Z.-J. Integration of miRNAs, Degradome, and Transcriptome Omics Uncovers a Complex Regulatory Network and Provides Insights Into Lipid and Fatty Acid Synthesis During Sesame Seed Development. *Front. Plant Sci.* **2021**, *12*, 709197. [\[CrossRef\]](#) [\[PubMed\]](#)
127. Quettier, A.-L.; Eastmond, P.J. Storage oil hydrolysis during early seedling growth. *Plant Physiol. Biochem.* **2009**, *47*, 485–490. [\[CrossRef\]](#) [\[PubMed\]](#)
128. Parthibane, V.; Rajakumari, S.; Venkateshwari, V.; Iyappan, R.; Rajasekharan, R. Oleosin is bifunctional enzyme that has both monoacylglycerol acyltransferase and phospholipase activities. *J. Biol. Chem.* **2012**, *287*, 1946–1954. [\[CrossRef\]](#)
129. Li, X.; Xiao, L.; Wu, G. Accumulation pattern of fatty acids during the seed development of sesame (*Sesamum indicum* L.). *Chin. J. Oil Crop Sci.* **2008**, *30*, 84.
130. Wang, L.; Xia, Q.; Zhang, Y.; Zhu, X.; Zhu, X.; Li, D.; Ni, X.; Gao, Y.; Xiang, H.; Wei, X. Updated sesame genome assembly and fine mapping of plant height and seed coat color QTLs using a new high-density genetic map. *BMC Genom.* **2016**, *17*, 31. [\[CrossRef\]](#)
131. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Walker, J., Ed.; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607.
132. Almagro Armenteros, J.J.; Sønderby, C.K.; Sønderby, S.K.; Nielsen, H.; Winther, O. DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **2017**, *33*, 3387–3395. [\[CrossRef\]](#)
133. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [\[CrossRef\]](#)
134. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.; Tosatto, S.C.; Paladin, L.; Raj, S.; Richardson, L.J. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [\[CrossRef\]](#)
135. Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **2018**, *46*, D493–D496. [\[CrossRef\]](#) [\[PubMed\]](#)
136. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.; Potter, S.C.; Finn, R.D. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37* (Suppl. 2), W202–W208. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Liu, B.; Hua, C.; Song, G.; Wu, M.; Cui, R.; Zhang, A.; Liu, Y.; Huang, L.; Yan, A.; Ali, I. The SPATULA transcription factor regulates seed oil content by controlling seed specific genes in *Arabidopsis thaliana*. *Plant Growth Regul.* **2017**, *82*, 111–121. [\[CrossRef\]](#)
139. Chow, C.-N.; Lee, T.-Y.; Hung, Y.-C.; Li, G.-Z.; Tseng, K.-C.; Liu, Y.-H.; Kuo, P.-L.; Zheng, H.-Q.; Chang, W.-C. PlantPAN3. 0: A new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res.* **2019**, *47*, D1155–D1163. [\[CrossRef\]](#)
140. Andrews, S. FastQC: A quality control tool for high throughput sequence data. In *Babraham Bioinformatics*; Babraham Institute: Cambridge, UK, 2010.
141. Joshi, N.; Fass, J. Sickel: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files, Software Version 1.33. 2011. Available online: <https://github.com/najoshi/sickle> (accessed on 28 September 2022).
142. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*, 1105–1111. [\[CrossRef\]](#)

-
143. 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](#)]
 144. Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D.R.; Pimentel, H.; Salzberg, S.L.; Rinn, J.L.; Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **2012**, *7*, 562–578. [[CrossRef](#)]
 145. Bioinformatics, B.; Valencia, S. OmicsBox-Bioinformatics made easy. March 2019. Available online: <https://www.biobam.com/omicsbox/> (accessed on 28 September 2022).