



# Effective Mechanisms for Improving Seed Oil Production in Pennycress (*Thlaspi arvense* L.) Highlighted by Integration of Comparative Metabolomics and Transcriptomics

## OPEN ACCESS

### Edited by:

Yonghua Li-Beisson,  
Commissariat à l'Energie Atomique et  
aux Energies Alternatives (CEA),  
France

### Reviewed by:

Zhenle Yang,  
Institute of Botany (CAS), China  
Chaofu Lu,  
Montana State University,  
United States

### \*Correspondence:

Ana Paula Alonso  
anapaula.alonso@unt.edu

### \*Present addresses:

Emmanuel Ortiz,  
Enviro Sciences,  
Inc., Carrollton, TX, United States  
Trevor B. Romsdahl,  
Mass Spectrometry Facility,  
Department of Biochemistry and  
Molecular Biology, University of Texas  
Medical Branch, Galveston, TX,  
United States

### Specialty section:

This article was submitted to  
Plant Metabolism and  
Chemodiversity,  
a section of the journal  
Frontiers in Plant Science

**Received:** 13 May 2022

**Accepted:** 10 June 2022

**Published:** 14 July 2022

### Citation:

Johnston C, García Navarrete LT,  
Ortiz E, Romsdahl TB, Guzha A,  
Chapman KD, Grotewold E and  
Alonso AP (2022) Effective  
Mechanisms for Improving Seed Oil  
Production in Pennycress (*Thlaspi  
arvense* L.) Highlighted by Integration  
of Comparative Metabolomics and  
Transcriptomics.  
Front. Plant Sci. 13:943585.  
doi: 10.3389/fpls.2022.943585

Christopher Johnston<sup>1</sup>, Leidy Tatiana García Navarrete<sup>2</sup>, Emmanuel Ortiz<sup>1†</sup>,  
Trevor B. Romsdahl<sup>1†</sup>, Athanas Guzha<sup>1</sup>, Kent D. Chapman<sup>1</sup>, Erich Grotewold<sup>2</sup> and  
Ana Paula Alonso<sup>1\*</sup>

<sup>1</sup>Department of Biological Sciences, BioDiscovery Institute, University of North Texas, Denton, TX, United States,

<sup>2</sup>Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, United States

Pennycress is a potentially lucrative biofuel crop due to its high content of long-chain unsaturated fatty acids, and because it uses non-conventional pathways to achieve efficient oil production. However, metabolic engineering is required to improve pennycress oilseed content and make it an economically viable source of aviation fuel. Research is warranted to determine if further upregulation of these non-conventional pathways could improve oil production within the species even more, which would indicate these processes serve as promising metabolic engineering targets and could provide the improvement necessary for economic feasibility of this crop. To test this hypothesis, we performed a comparative biomass, metabolomic, and transcriptomic analyses between a high oil accession (HO) and low oil accession (LO) of pennycress to assess potential factors required to optimize oil content. An evident reduction in glycolysis intermediates, improved oxidative pentose phosphate pathway activity, malate accumulation in the tricarboxylic acid cycle, and an anaplerotic pathway upregulation were noted in the HO genotype. Additionally, higher levels of threonine aldolase transcripts imply a pyruvate bypass mechanism for acetyl-CoA production. Nucleotide sugar and ascorbate accumulation also were evident in HO, suggesting differential fate of associated carbon between the two genotypes. An altered transcriptome related to lipid droplet (LD) biosynthesis and stability suggests a contribution to a more tightly-packed LD arrangement in HO cotyledons. In addition to the importance of central carbon metabolism augmentation, alternative routes of carbon entry into fatty acid synthesis and modification, as well as transcriptionally modified changes in LD regulation, are key aspects of metabolism and storage associated with economically favorable phenotypes of the species.

**Keywords:** ascorbate, biofuel, cell wall, central carbon metabolism, erucic acid (C22:1), lipid droplet, *Thlaspi arvense*, threonine aldolase

## INTRODUCTION

The need for renewable fuel sources has been greatly highlighted in recent years by the growing world population and increasing concern over environmental impacts of conventional fossil fuel consumption (Thelen and Ohlrogge, 2002; Banković-Ilić et al., 2012). Oilseed crops have gained prominence in the search for biofuel sources due to recent improvements in the processing of their constituent oils (Kumar et al., 2016). The global oleochemical industry was valued at USD 25.8 billion in 2020, and is expected to display an annual growth rate of 6.5% from 2021 to 2027 (Astute Analytica, 2021). By 2025, the United States' demand for renewable oils is expected to increase to a value of over USD 3 billion (Grand View Research, 2017). Additionally, energy security has been cited by both transportation and military sectors as a major incentive for producing renewable fuel sources that reduce the unpredictability associated with relying on imported fossil fuels (Drenth et al., 2014).

Native to Eurasia, pennycress (*Thlaspi arvense* L.) is a cool-season annual plant species well-adapted to North American lands, even those of marginal quality (Sedbrook et al., 2014; Zanetti et al., 2019). Natural pennycress accessions accumulate up to 36% of seed oil on a per weight basis, with production predominantly taking place in developing embryos (Moser et al., 2009b). Of particular interest is the fatty acid (FA) composition of the oil, approximately 33% of which is erucic acid, a 22-carbon monounsaturated FA (Moser et al., 2009b; Sedbrook et al., 2014). The high content of erucic acid in pennycress seed oil, along with other minor unsaturated FAs, contributes to its favorable profile for aviation fuel, including optimal combustion speed and desirable low temperature flow properties (Moser et al., 2009a). In addition to the chemical properties of its seed oil, the agronomic practices associated with successful pennycress cultivation make it interesting to growers and the agricultural sector. Government agencies have stressed that renewable fuels must be sourced from a non-food crop to minimize the impact on domestic food supplies (Blakeley, 2012). Pennycress has been identified as a lucrative candidate for introduction into crop rotations with major commercial crops like corn and soybean, because aerial broadcast of seed in the fall as a winter cover crop allows for combine harvesting the following spring (Dose et al., 2017; Markel et al., 2018; Mousavi-Avval and Shah, 2020). Not only does this benefit growers as an additional source of income, but also provides canopy cover that suppresses weed pressure, provides ecosystem services for pollinators, and maintains the soil coverage that provides erosion control and minimizes nutrient runoff (Dorn et al., 2015; Eberle et al., 2015; Johnson et al., 2015).

Despite the plethora of potential uses of pennycress seed oil, the existing germplasm must be improved *via* metabolic engineering and breeding to ensure profitability. A wide variety of oil content has been observed across natural accessions of pennycress, and relatively few, if any, of these accessions provide optimal seed oil yields as is (Cocuron et al., 2018; Cubins et al., 2019). While knowledge on FA and lipid metabolism in plants as a whole is

strong, specific metabolic characteristics of pennycress require further elucidation for the improvement of its oil-related traits.

Seeds of the Brassicaceae family are known to generate carbon precursors for plastidic *de novo* FA biosynthesis *via* glycolytic flow and reductant *via* the oxidative pentose phosphate pathway (OPPP) activity (Dennis and Miernyk, 1982; Schwender et al., 2003; Baud and Lepiniec, 2009). Metabolic models, including those in pennycress, show strong evidence that much of this occurs *via* the conversion of phosphoenolpyruvate (PEP) and/or malate into pyruvate, which is then further broken down into acetyl-CoA, the main carbon building block for *de novo* FA synthesis (Kubis et al., 2004; Cocuron et al., 2014; Tsogtbaatar et al., 2020). The *de novo* synthesis of FAs in the plastid also requires ATP and reductant (NADPH/NADH), which along with carbon precursors can be generated *via* photosynthesis in green heterotrophic seeds (Hills, 2004; Baud and Lepiniec, 2010; Tsogtbaatar et al., 2015). The first desaturation of fatty acids takes place in the plastid, namely *via* stearoyl-ACP  $\Delta 9$  desaturase, which is required for monounsaturated fatty acid production, including erucic acid (Subedi et al., 2020).

FAs produced in the plastid are released into the cytoplasm to produce an acyl-CoA pool, which contains FAs available for further modification and oil synthesis (Cagliari et al., 2011). In addition to the OPPP, the activity of the tricarboxylic acid (TCA) cycle is heavily involved in the production of carbon and/or reductant for extra-plastidic FA elongation in pennycress, which is required for the high production of erucic acid along with other long-chain and unsaturated FAs (Moser et al., 2009a; Claver et al., 2017; Tsogtbaatar et al., 2020). Elongation of FAs in the acyl-CoA pool takes place through a multi-enzyme complex, including an ER-localized condensing enzyme, 3-ketoacyl-CoA synthase (KCS; Millar and Kunst, 1997; Joubès et al., 2008). Incorporation of FAs into storage lipids then occurs *via* the Kennedy pathway, which involves incorporation into glycerol backbones to form triacylglycerols (TAGs; Kennedy, 1961; Romsdahl et al., 2021). Additional acyl-CoA-independent pathways can incorporate acyl chains directly to form TAGs as well (Dahlqvist et al., 2000; Lu et al., 2009; Cagliari et al., 2011; Hu et al., 2012; Wang et al., 2012; Kim and Chen, 2015).

The power of comparative “omics” approaches to study the intraspecific variation in oilseeds has been displayed in high and low oil maize accessions, where both increased incorporation of carbon (as glucose 6-phosphate) into the plastid and action of the plastidic malic enzyme were noted to result in higher oil accumulation and thus serve as potential targets for engineering (Cocuron et al., 2019). Furthermore, detailed metabolomic and  $^{13}\text{C}$ -labeling analyses have identified several non-conventional mechanisms (i.e., unusual and contrary to classical reports in the literature) involved in efficient oil production in pennycress (Tsogtbaatar et al., 2015, 2020). These include reversible isocitrate dehydrogenase activity, thus fixing  $\text{CO}_2$  to sustain FA elongation *via* citrate, and the incorporation of plastidic  $\text{CO}_2$  *via* Rubisco that was released by other metabolic reactions (Tsogtbaatar et al., 2020).

While these non-conventional mechanisms are associated with higher efficiency of oil production in pennycress compared to other species, it is to be determined if their upregulation

**Abbreviations:** HO, High oil; LD, Lipid droplet; LO, Low oil.

can further improve oil production in pennycress itself. The comparative approach has promise in pennycress, as representative high and low oil accessions are available with differing biomass compositions. The purpose of this research was to perform a comprehensive comparative analysis of high and low-oil pennycress accessions *via* metabolomics and transcriptomics in order to identify which components of metabolism are likely key to boosting the production of the desirable oil, from *de novo* FA synthesis all the way to oil storage. The results indicated that mechanisms powering *de novo* FA synthesis associated with higher oil production cover early central carbon metabolism all the way to organization and stability of end-product lipid droplets. These mechanisms serve as valuable targets for improving oil production in pennycress.

## MATERIALS AND METHODS

### Plant Materials

A representative accession producing a high percent weight per weight (% w/w) oil content (HO) at maturity (USDA-GRIN Acc. 'Ames 32872') and another accession containing low % w/w oil content (LO; USDA-GRIN Acc. 'Ames 31500') were chosen from existing seed stocks of USDA-GRIN pennycress accessions (**Supplementary Figure S1**). Seed stocks were stored at 4°C. Plant growth was carried out in accordance with Tsogtbaatar et al. (2015). Seeds of HO were immediately planted in soil and required no pretreatment for adequate germination. Pedicels of newly-opened flowers were hand-pollinated and tagged daily with acrylic paint to keep record of the age of developing siliques for embryo collection.

### Embryo Collection

For biomass and metabolomics analysis, siliques were selected according to how many days after pollination (DAP) had elapsed in accordance with Tsogtbaatar et al. (2015) and Cocuron and Alonso (2014). For embryos used in transcriptomics analysis, excised embryos were placed into an autoclaved microcentrifuge tube containing liquid nitrogen. Following collection of these embryos, microcentrifuge tubes were immediately transferred to a -80°C freezer prior to being capped following evaporation of all remaining liquid nitrogen in the tube. Embryos used for transcriptomics analysis were collected using autoclaved forceps and a microscope treated with RNase Zap (MilliporeSigma, Burlington, MA). For all analyses, the earliest stage selected for LO was 11 DAP and the earliest stage selected for HO was 10 DAP; these stages were in consistency with the stage in which the radicle was at a 45° angle from the central axis of the cotyledons. The subsequent developmental stages were 14, 17, and 20 DAP for both accessions. For biomass analysis, siliques at 30 DAP were senesced and therefore considered mature.

### Fatty Acid Extraction and Derivatization

Fatty acids were extracted and methylated in accordance with Cocuron et al. (2014). Combined extracts were dried under

nitrogen gas at 60°C prior to methylation with 2.5% (v/v) sulfuric acid in methanol at 80–85°C for 90 min. After methylation, the reaction was quenched with 250 µl of a 5% (w/v) sodium hydrogen sulfate solution. The organic phase containing the fatty acid methyl esters (FAMES) was diluted 5-fold with hexane in a 2 ml autosampler vial prior to GC-MS analysis.

### GC-MS Quantification of Fatty Acids

FAMES were quantified in accordance with Tsogtbaatar et al. (2015) using an Agilent 6890N gas chromatograph (Agilent Technologies, Santa Clara, CA) coupled to a Agilent 5975B mass selective detector. An Omegawax 250 capillary column (30 m × 0.25 mm × 0.25 µm; MilliporeSigma, Burlington, MA) was used to separate FAMES at a constant flow rate of 1.4 ml min<sup>-1</sup> using helium as the carrier gas. The initial temperature was 120°C and held for 30 s, followed by an increase to 245°C for the remaining 9.5 min at a rate of 100°C min<sup>-1</sup>. The injection temperature was set at 225°C, with the injection mode set to a split ratio of 10:1. Mass spectra were acquired using electron impact ionization in positive ion mode, with ion source and interface set to 240 and 150°C, respectively. Acquisition and integration of GC-MS data was carried out *via* MSD ChemStation E.02.02.1431 (Agilent Technologies, Santa Clara, CA).

### Protein Extraction and Quantification

Total proteins were extracted from the pellet remaining from the FA extraction with a 20 mM Tris, 150 mM NaCl, and 1% SDS buffer in accordance with Cocuron et al. (2014) to obtain a final 1.5 ml volume of extract. Protein content was then quantified colorimetrically *via* the DC Protein Assay (Bio-Rad, Hercules, CA) in a Bio-Rad SmartSpec Plus spectrophotometer at 750 nm.

### Metabolite Extraction

Extraction of soluble metabolites was carried out according to previous research (Alonso et al., 2010; Cocuron et al., 2014) using boiling deionized water. Samples were then incubated in boiling water for a total of 10 min and ultimately passed through a 0.22 µm filter into a 15 ml conical tube on ice. The remaining pellet was rinsed with ice cold water, vortexed, and centrifuged again prior to passing through the same filter as the first supernatant. Ice-cold water was then used to remove any remaining metabolites in the syringe and filter. Lyophilized filtrate was resuspended in ice-cold water and passed through a 0.22 µm Nanosep filter (Pall Corp., Port Washington, NY) to use for sugar and sugar alcohol quantification, while the remaining volume was loaded onto a 3 kDa Ultracel filter (MilliporeSigma, Burlington, MA). All extraction steps were also performed with <sup>13</sup>C-glycine, <sup>13</sup>C-glucose, and <sup>13</sup>C-fumarate internal standards without sample in order to calculate correction factors to control for errors during extraction or sample preparation.

### LC-MS/MS Quantification of Metabolites

Sugars and sugar alcohols, amino acids, organic acids, and phosphorylated compounds were quantified as previously described using programs described (Alonso et al., 2010; Koubaa

et al., 2013; Cocuron and Alonso, 2014; Cocuron et al., 2014). For sugar and sugar alcohols, extracts were diluted in a 60:40 acetonitrile:water solution prior to injection onto a Shodex Asahipak NH2P-50 2D column and a NH2P-50G 2A guard column (Showa Denko K. K., Tokyo, Japan) followed by separation *via* an acetonitrile and water gradient (Cocuron et al., 2014). Mass spectra were obtained *via* turbo-spray ionization at 4,500 V in negative ionization mode, multi-reaction monitoring (MRM) mode with a dwell time of 100 msec for each transition. For amino acids, extracts were diluted with an HCl solution and water such that the final concentration in the autosampler vial was 1 mM HCl. Samples were incubated at 15°C prior to injection onto a Hypercarb column (Thermo Fisher Scientific, Waltham, MA) and were separated *via* a gradient of acetonitrile with 0.1% formic acid and water with 0.1% formic acid. Mass spectra were obtained at 2,500 V in positive ionization mode, MRM mode with a dwell time of 35 msec for each transition. For organic acids and phosphorylated compounds, extracts were diluted with HPLC-grade water and kept at 4°C prior to injection onto an IonPac AS11 column (Dionex, Sunnyvale, CA) with an IonPac AG11 guard column. A gradient of 0.5 mM KOH to 75 mM KOH was used to separate metabolites, with mass spectra in negative ionization mode, MRM mode. All metabolite data was acquired and integrated *via* Analyst (SCIEX, Framingham, MA). All samples were run using a 1,290 Infinity II ultra-high-performance liquid chromatograph (Agilent Technologies, Santa Clara, CA) coupled with a QTRAP 6500+ mass spectrometer (SCIEX, Framingham, MA). Data acquisition and quantification was carried out using Analyst v1.7 with HotFix 3 (SCIEX, Framingham, MA).

## RNA Extraction

RNA was extracted in accordance with Horn et al. (2016), with modifications. Approximately 1,000 µg of fresh embryo tissue was manually homogenized in a 2 ml microcentrifuge tube with plastic pestles directly in liquid nitrogen. A 600 µl volume per sample of extraction buffer of 2% CTAB, 2 M NaCl, 100 mM Tris (pH=8), 25 mM EDTA (pH=8), 3% PVP, 0.5 g L<sup>-1</sup> spermidine, and 3% 2-sulfanyethan-1-ol that was preheated to 65°C was used for extraction. A 40 µl volume of 3.2 M sodium ethanoate (pH=5.5), 100 µl of RNA-grade glycogen, and 800 µl of ethanol were added to each aliquot prior to a 12 h precipitation at -20°C. Each aliquot was then centrifuged at 21,000 × g for 1 h at 4°C, immediately placed on ice, and supernatant was discarded. A total of three 70% ethanol washes were carried out on the resulting pellet prior to being dried and dissolved in RNase-free water. A lithium chloride precipitation was then performed at -20°C for 12 h. Three more ice-cold ethanol washes were then carried out as previously stated, pellets were dried, and each sample's pellet was dissolved in RNase-free water. Samples were stored at -80°C until further analysis.

## RNA Quality Confirmation and RNAseq

Analysis of the quality and quantity of RNA extracts was performed at the University of North Texas Genomics Center

using an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA) and QUBIT 3.0 fluorometer (Invitrogen, Carlsbad, CA), respectively. All samples were confirmed to have RNA integrity numbers of ≥9. Samples were shipped overnight on dry ice to the location where RNAseq was performed using DNBseq Technology: NGS 2.0 (BGI, Cambridge, MA).

## Quality Control and Mapping of RNAseq Data

A total of 2,259 million sequences of raw data were generated with an average of 37.65 million reads per sample. A raw data quality check was carried out using FastQC (Andrews, 2010) and data was deposited online (NCBI Accession: PRJNA808106). Clean reads were then mapped to the pennycress reference genome (Dorn et al., 2015) using HISAT v2.1.0 (Kim et al., 2015) with default parameters. The mapped reads were then sorted and converted to .bam files using Samtools v1.9 (Li et al., 2009). Each aligned sample had an average mapping of 97%. The R package FeatureCounts v1.5 (Liao et al., 2014) was used to generate raw count data, and StringTie v2.1.1 (Pertea et al., 2016) was used to calculate transcripts per million (TPM). All genes were paired with respective *Arabidopsis thaliana* homologs for annotation and analysis using Araport (Krishnakumar et al., 2015; Pasha et al., 2020; Navarrete et al., submitted).

## Transcriptomics Data Processing

The resulting raw count data from RNAseq were transformed using DeSeq2 v1.30.1 (Love et al., 2014) to perform PCA as previously described (Gerst and Hölzer, 2018). A pool of differentially expressed genes (DEGs) was created using a DESeq analysis, which was carried out within each developmental stage. Pre-filtering was carried out by removing genes with an average of less than 5 TPM across all samples. All fold-changes were calculated by using HO in relation to LO. Genes which returned with  $\text{padj} \leq 0.05$  and a  $\log_2\text{-fold change} (\log_2\text{FC}) \geq |1|$  were considered significant, with upregulated genes returning a  $\log_2\text{FC} \geq 1$  and downregulated genes returning a  $\log_2\text{FC} \leq -1$ . Adjusted values of  $p$  were corrected using the false discovery rate (FDR). Shrinkage of  $\log_2\text{FC}$  estimates (Zhu et al., 2019) was carried out to improve interpretability of MA plots (Supplementary Figure S2). The “rlog” function in DESeq2 was performed with a transformation not blinded by the experimental design for all downstream analyses, in accordance with the suggested workflow of the package (Love et al., 2014). Genes that maintained a statistical significance at all developmental stages were functionally profiled according to KEGG pathway membership *via* g:Profiler (Raudvere et al., 2019). Unless otherwise stated, all genes are referred to by their *A. thaliana* homolog, but also identified with their *T. arvense* number when applicable to show unique reads from our RNAseq dataset. Functional classification of DEGs based on gene ontology (GO) terms were carried out *via* PANTHER GO-Slim using the GO biological process (GO:BP) database (Mi and Thomas, 2009; Mi et al., 2021). Visualization of the DEG GO:BP profile was performed using BiNGO (Supplementary Figure S4; Maere et al., 2005).

The protein sequences of the 950 DeSeq-derived DEGs were incorporated into DeepLoc v.1.0 (Almagro Armenteros et al., 2017). This software uses recurrent neural networks (RNNs) to classify proteins in one of ten potential subcellular localizations (nucleus, cytoplasm, extracellular, mitochondrion, cell membrane, endoplasmic reticulum (ER), chloroplast, Golgi apparatus, lysosome/vacuole, and peroxisome; **Supplementary Data Sheet S1**). The cutoff value for considering a prediction was a likelihood higher than 0.7.

## Integration of Metabolomics and Transcriptomics Data

A joint pathway analysis was performed using MetaboAnalyst 5.0 to combine transcriptomics and metabolomics data in order to detect pathway enrichment and analyze pathway topology at 20 DAP (the maximum point of divergence in FA content for both accessions prior to the cessation of metabolic activity; Chong et al., 2018). A hypergeometric test was used for the enrichment analysis, relative betweenness centrality was used for the pathway topology, and “omics” integration was carried out by combining values of  $p$  via equally-weighted Z-tests (Chong et al., 2019). Pathways that returned with a  $p_{\text{adj}} \leq 0.05$  and pathway impact value  $\geq 0.2$  were considered significant. All returned pathways were from the KEGG database (Kanehisa and Goto, 2000; Kanehisa, 2019; Kanehisa et al., 2021).

## Coexpression Analysis

The “psych” package in R was used to calculate Pearson correlations of all rlog-transformed DEGs (Revelle, 2018). The correlation matrix was then filtered to include DEGs with  $>0.9$  correlation and  $p_{\text{adj}} < 0.01$ . Further network statistics including degree and betweenness centrality were calculated using the “igraph” package in R to assist with visualizations (Csardi, 2014). The coexpression network was imported into Cytoscape v3.8.0 for all network visualization. Over-representation analysis (ORA) was run using the GO:MF, GO:BP, GO:CC, and KEGG database. ORA was carried out using g:Profiler using a threshold of 0.10. General parent GO terms (response to stimulus, metabolic process, binding, etc.) and redundant terms are omitted for conciseness; full results are in Suppl. Info. CytoHubba was used on the full DEG coexpression network to identify the top 10 most important hub genes based on the maximal clique centrality algorithm (Chin et al., 2014). Separate subnetworks were made for the top upregulated and downregulated hub genes, with first neighbors of each hub gene organized by membership in over-represented GO/KEGG terms. Genes not present in any overrepresented terms were organized according to the following functional categories: 1) lipid droplet (GO:0005811), lipid metabolic process (GO:0006629), or fatty acid metabolic process (GO:0006631); 2) central carbon metabolism (KEGG:00010,00020,00030,00500); 3) Biosynthesis of amino acids (KEGG:01230) or amino acid metabolism (KEGG:00250–00400); 4) electron transport chain (GO:0022900); and 5) peroxisome (GO:0005777). Genes with unknown function were also grouped. All other known genes

were placed in the most common GO or KEGG term for visualization if not placed in any previously mentioned terms.

## MS Imaging and Confocal Microscopy

Pennycress seeds were prepared and processed for MS imaging as described by Sturtevant et al. (2021). Dry seeds were embedded into 10% gelatin, frozen at  $-80^{\circ}\text{C}$  for 24 h, and then transferred to  $-20^{\circ}\text{C}$  for 2–3 days prior to cryo-sectioning. Seeds were oriented within the gelatin so that medial-cross sections could be taken, giving both embryonic axis and cotyledonary tissues. Sections of  $30\ \mu\text{m}$  thickness were taken using a cryo-microtome (CM1950, Leica Biosystems, Germany) set at  $-16^{\circ}\text{C}$ , followed by lyophilization for 3 h to remove moisture due to condensation. Sections that were intact and showed no cosmetic defects under light microscopy were chosen for MALDI-MS imaging analysis. Tissue sections chosen for MALDI-MS imaging were coated with 2,5-dihydroxybenzoic acid (DHB) by sublimation using a protocol adapted from Hankin et al. (2007). Matrix coated sections were then analyzed by MALDI-MS imaging with a MALDI-LTQ-Orbitrap-XL mass spectrometer (ThermoScientific). The MALDI source parameters were set to:  $12\ \mu\text{J}/\text{pulse}$ , 10 laser shots per step, and raster step size of  $40\ \mu\text{m}$ . The orbitrap mass analyzer parameters were set to detect ions from 500 to  $1,200\ m/z$  at a set resolution of 60,000. Imaging data was processed using the open source software Metabolite Imager (Horn and Chapman, 2014).

The embryos were processed for Airyscan imaging, after staining of LDs with BODIPY 493/503 (Invitrogen) as previously described (Cai et al., 2015; Gidda et al., 2016). Micrographs of dry seeds were acquired with a Zeiss LSM710 microscope fitted with Airyscan confocal superresolution and the MBS 488 beam splitter using an Apochromat  $63\times/1.4$  oil DIC objective lens (Carl Zeiss Inc., Germany). Excitation and emission signals of the BODIPY stained LDs were collected as single optical section images (z-sections) of cotyledons and embryonic axis cells. Airyscan images were processed using Zeiss Zen blue software.

## RESULTS AND DISCUSSION

### Developing HO Embryos Produce Oil More Efficiently on a Dry Weight Basis and Accumulate More Erucic Acid in Cotyledons

At maturity, the total FA percentage on a per weight basis was 41% in HO embryos, whereas in LO it was 32% ( $p=0.0132$ ; **Table 1**). On a per embryo basis the total quantity of FA was not statistically significant between the two accessions; the higher dry weight in LO per embryo ( $p=0.0051$ ) was matched by a significantly greater total protein content ( $p=0.037$ ; **Supplementary Figure S1**). Increased dry weight in LO is thus at least conceivably explained by a higher production of total protein, as HO is able to produce a similar FA content while producing less biomass per embryo. A clear divergence in the biomass profile started around 17 DAP, indicating a

**TABLE 1** | Total biomass composition (% composition and % w/w) per embryo in mature pennycress embryos of HO and LO accessions.

Component	HO		LO		Value of p	
	Mean	SD	Mean	SD		
Fatty acid composition	% C16:0	3.184	0.076	3.272	0.117	0.250
	% C18:0	<b>0.598</b>	0.033	<b>0.831</b>	0.114	<b>0.00775</b>
	% C18:1	<b>13.691</b>	0.376	<b>18.470</b>	0.316	<b>1.19E-06</b>
	% C18:2	<b>19.480</b>	0.150	<b>17.145</b>	0.239	<b>3.11E-06</b>
	% C18:3	<b>8.938</b>	0.171	<b>9.590</b>	0.171	<b>0.00673</b>
	% C20:1	<b>9.522</b>	0.227	<b>11.067</b>	0.329	<b>0.00025</b>
	% C20:2	<b>1.332</b>	0.084	<b>1.003</b>	0.080	<b>0.00126</b>
	% C22:1	<b>40.189</b>	0.309	<b>36.184</b>	0.450	<b>6.26E-06</b>
	% C24:1	<b>3.067</b>	0.267	<b>2.438</b>	0.098	<b>0.00441</b>
Fatty acid	% w/w	<b>41.084</b>	1.726	<b>31.789</b>	2.683	<b>0.0132</b>
Protein	% w/w	31.196	1.556	32.185	4.746	0.706
Starch	% w/w	0.024	0.009	0.030	0.008	0.339
Hemicellulose composition	% Arabinose	<b>30.285</b>	1.432	<b>37.148</b>	1.908	<b>0.00120</b>
	% Galactose	13.304	1.505	13.897	1.132	0.552
	% Glucose	12.996	1.103	11.934	1.492	0.296
	% Mannose	<b>22.791</b>	1.413	<b>17.619</b>	3.255	<b>0.0268</b>
	% Xylose	20.625	0.818	19.402	1.692	0.240
Hemicellulose	% w/w	1.693	0.362	1.267	0.281	0.112
Cellulose	% w/w	0.526	0.126	0.343	0.169	0.134
Soluble sugar and sugar alcohol composition	% Arabinol	<b>0.010</b>	0.002	<b>0.023</b>	0.005	<b>0.00180</b>
	% Fructose	<b>0.033</b>	0.007	<b>0.018</b>	0.004	<b>0.00546</b>
	% Galactinol	0.241	0.046	0.203	0.043	0.271
	% Hexitols	0.016	0.001	0.022	0.005	0.0528
	% Glucose	4.433	0.240	4.283	0.219	0.389
	% Inositol	<b>0.056</b>	0.006	<b>0.041</b>	0.004	<b>0.00577</b>
	% Raffinose	<b>1.160</b>	0.204	<b>0.625</b>	0.020	<b>0.00196</b>
	% Stachyose	<b>3.762</b>	0.434	<b>1.789</b>	0.115	<b>0.000120</b>
	% Sucrose	<b>90.251</b>	0.767	<b>92.976</b>	0.343	<b>0.00064</b>
	% Tetrols	<b>0.037</b>	0.006	<b>0.021</b>	0.002	<b>0.00269</b>
Soluble sugars and sugar alcohols	% w/w	<b>10.651</b>	1.457	<b>14.455</b>	1.607	<b>0.0127</b>
	Total recovery	86.928		77.354		

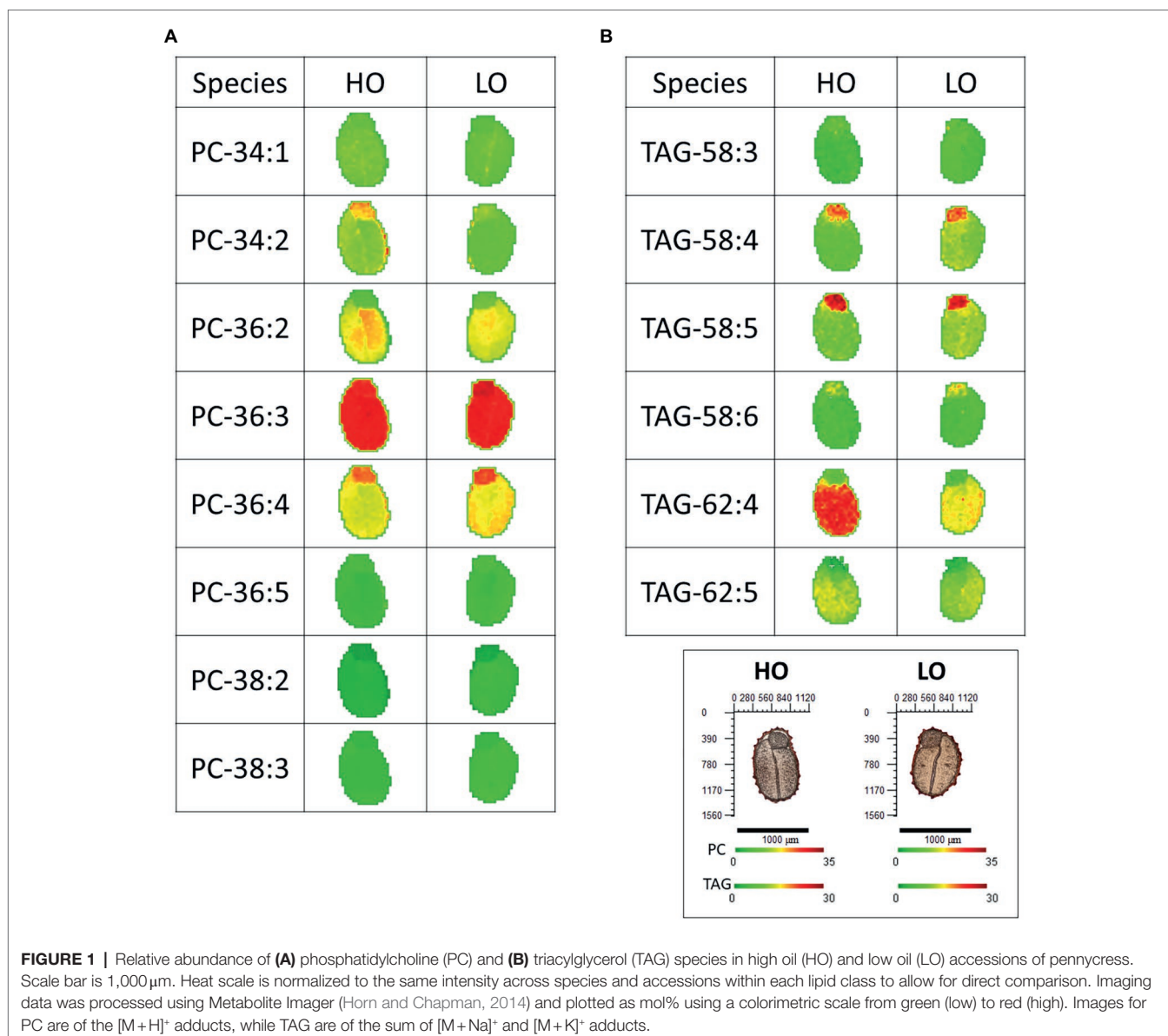
SD, standard deviation of the mean. Value of p corresponds to t-test results assuming equal variance with 4 biological replicates. Fatty acid (FA), hemicellulose, and soluble sugar and sugar alcohol composition indicate the percentage of total content for each individual class of compound. FA data was quantified using C17:0 internal standard. Biomass components in mature embryos were subjected to t-tests assuming equal variance to compare each accession. Bold text indicates significance at  $p < 0.05$ .

potentially critical point in assessing differences between the two accessions. All FA species except for palmitic acid (C16:0) exhibited a significant difference between the two accessions, however the magnitude of these differences did not exceed 4% of the total composition in any case. Notably, long chain fatty acid content was higher in HO embryos (54% vs. 51% in LO). Erucic acid comprised 40% of the fatty acid composition in HO embryos compared to 36% in LO embryos. This is a modest difference compared to the range reported in previous research across wild populations, displaying a range of 30 to 39% erucic acid (Phippen and Phippen, 2013). Soluble sugars and sugar alcohols made up a significantly larger portion of LO embryos (14%) in comparison with HO embryos (11%;  $p = 0.0127$ ). Starch content was negligible in both accessions ( $< 0.1\%$ ).

The distributions and relative amounts of phosphatidylcholine (PC) and TAG in embryos of both genotypes were visualized by MALDI-MS imaging, and representative images of selected molecular species are shown in **Figure 1**. PC is an intermediate in the desaturation and modification of acyl chains for TAG assembly, so these lipid classes often display precursor-product

relationships. HO and LO genotypes displayed similar distributions and relative proportions of most PC molecular species, although heterogeneous distributions were evident within the embryos (between the cotyledons and the embryonic axis) for some PC species, such as PC 36:2 and PC 36:4 (**Figure 1A**).

The spatial distribution of TAG molecular species in embryos were similar between the two genotypes (**Figure 1B**). There were species that were notably concentrated in the embryonic axis (e.g., TAG- 58:4 and TAG 58:5). In contrast, erucic-acid-containing TAG molecular species appeared to be mostly restricted to cotyledons in both genotypes (TAG 62:4 and TAG 62:5), indicating a spatial distinction for FA elongation in the embryos. The HO genotype appeared to have relatively higher amounts of TAG 62:4 compared to the LO genotype (**Figure 1B**), consistent with the higher amounts of erucic acid in these seeds (**Table 1**). Taken together it appears that the HO genotype more readily produces very long chain ( $> 18^\circ\text{C}$ ) monounsaturated FAs and incorporates them into the storage form of TAGs, and that this elongation occurs most predominantly in the cotyledons of the embryo. This is strongly supported by previous research in Brassicaceae, where in addition to pennycress, both



*Brassica napus* and *Camelina sativa* are shown to be enriched for longer-chain ( $\geq 20^\circ\text{C}$ ) fatty acid species in the cotyledonary tissue compared to that of the embryonic axis (Horn et al., 2013; Lu et al., 2020; Romsdahl et al., 2021).

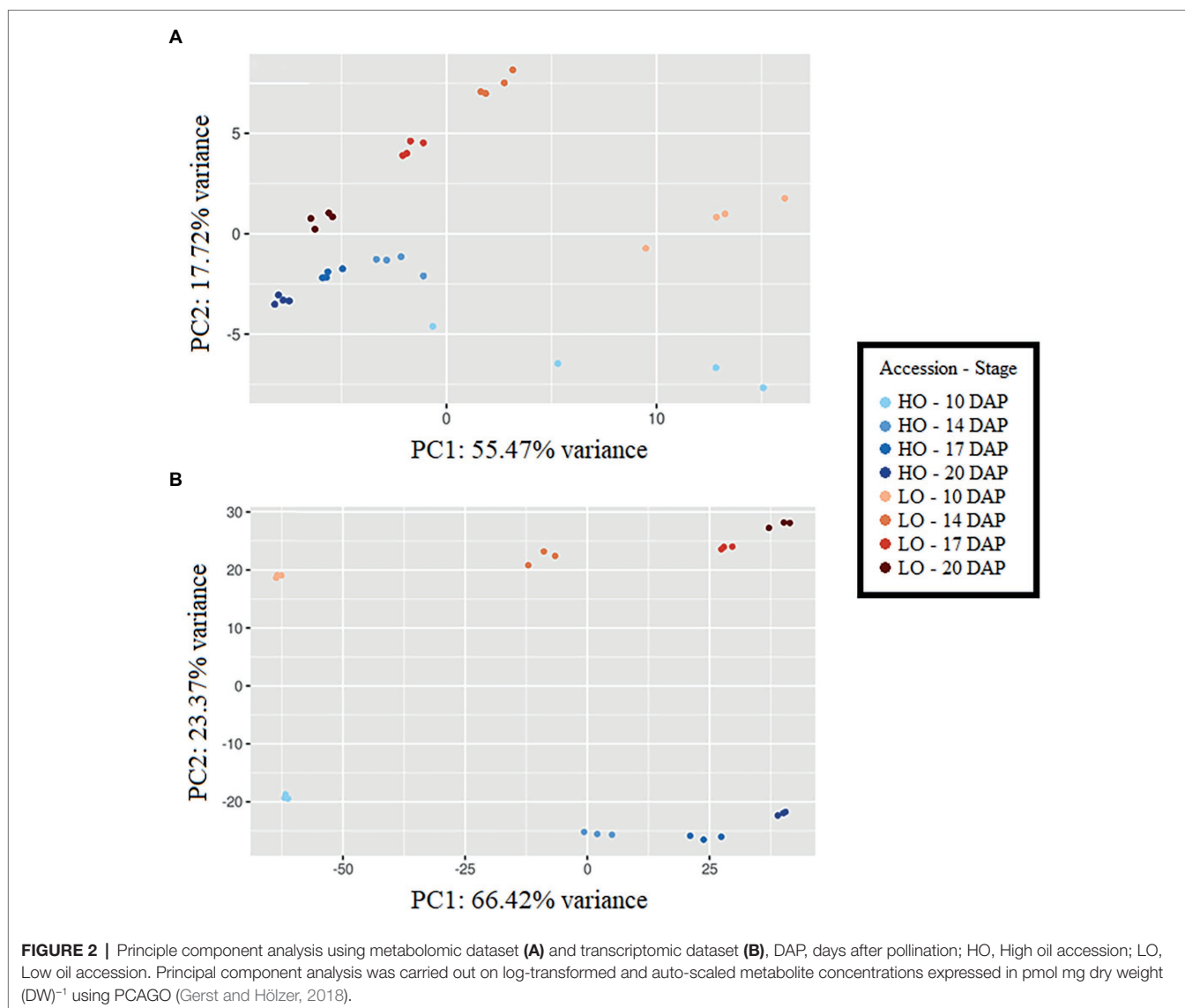
### Metabolomic and Transcriptomic Profiles Show Divergence Between Lines

Principal component analyses (PCAs) showed clear separation by accession and developmental stage using metabolomic and transcriptomic datasets (Figure 2). In metabolomics and transcriptomics datasets, principal component 1 (PC1) generally separated developmental stages, accounting for 55.47 and 66.42% variance, respectively. PC2, explaining 17.72 and 23.37% of the variance in metabolomics and transcriptomics datasets, respectively, was the primary separator of accessions. Time-series metabolite quantities between HO and LO genotypes were placed

into four clusters (Supplementary Figure S3). Clusters 1 and 2 notably consisted of the majority of sugar phosphates, all basic and most polar amino acids, 3-carbon glycolysis intermediates (phosphoglycerates, phosphoenolpyruvate), and TCA intermediates upstream of isocitrate dehydrogenase (Figure 3). These metabolites were consistently higher in LO embryos throughout development. Cluster 3, encompassing metabolites that were constantly higher in HO, consisted of the main carbon sources glucose and sucrose, as well as key cell wall precursors. Specialization of metabolite investment thus appears to be a phenomenon associated with the differences between accessions.

### Alteration of Metabolic Pathways Spans Central Metabolism

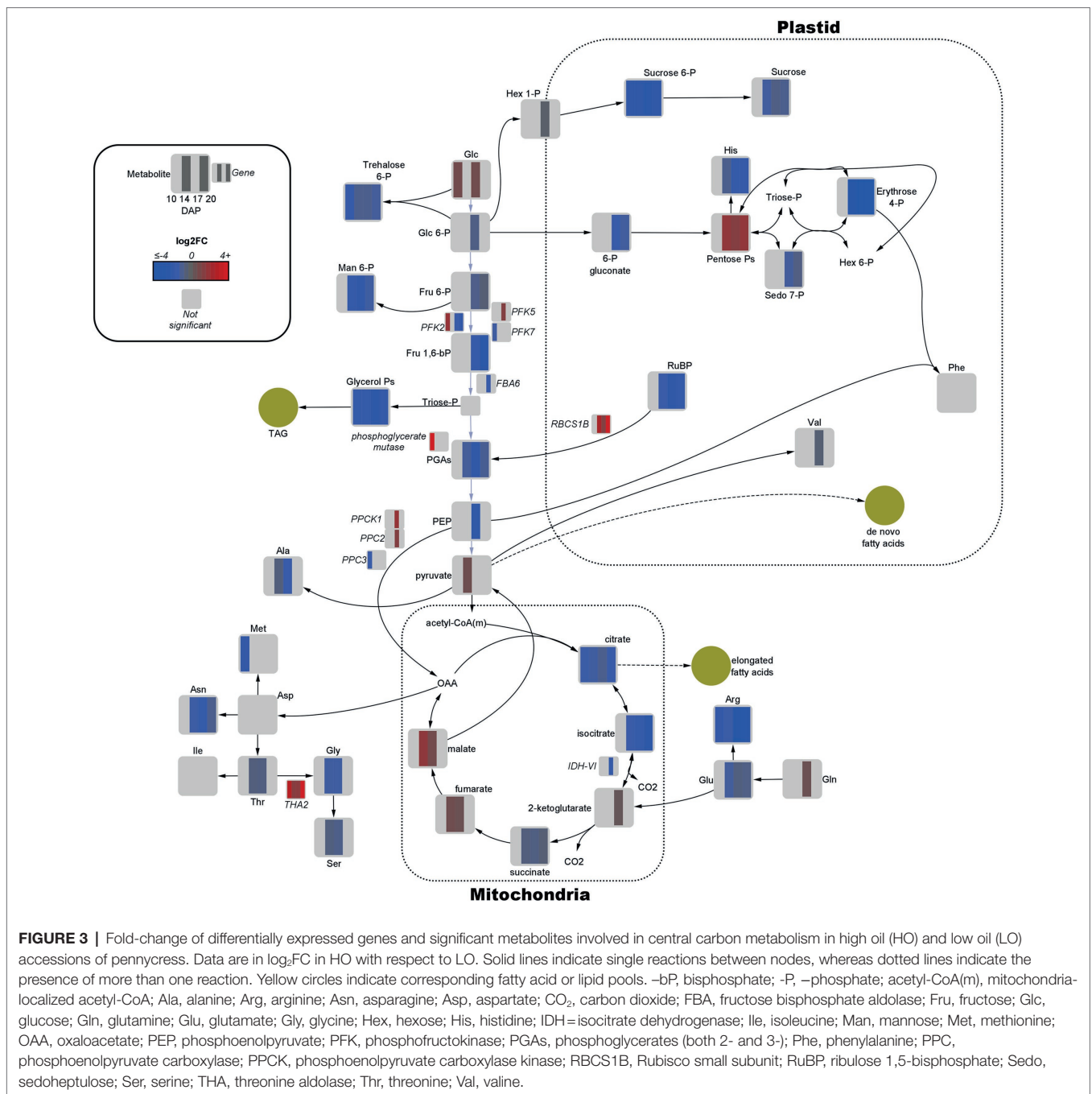
Quantitative joint pathway analysis identified all key parts of central carbon metabolism as being significantly enriched



( $p_{\text{adj}} \leq 0.05$ ) and impacted (impact factor  $\geq 0.2$ ), consisting of glycolysis, carbon fixation, the TCA cycle, and the OPPP (Figure 4). Of all significant pathways involved in central carbon metabolism, glycolysis was the most heavily impacted (Impact = 2.0067,  $p_{\text{adj}} = 0.0312$ ). Key mapped metabolites and genes were downregulated in HO around the divergence point in the rate of oil accumulation. Specifically, pyruvate displayed a 1.6-fold increase in abundance in HO at 14 DAP, matched by a 1.7-fold increase in glucose at 17 DAP (Figure 3). The gene *phosphofructokinase 5* (*PFK5*), predicted by our modeling to be localized to the plastid, was upregulated 2.7-fold in HO at 17 DAP and may serve as a chief target in central carbon metabolism. The PFK step is known to be the rate-limiting step in glycolytic flux in previous research on *B. napus*, and its upregulation has been shown to explain the higher content of oil in oil palm vs. the sugar-accumulating date palm, presumably *via* increased pyruvate production (Junker et al., 2007; Bourgis et al., 2011).

Similar to glycolysis, there were only a few key metabolites in both the PPP and TCA cycle that were upregulated in HO at any developmental stage (Figure 3). Of note was the 2.8- to 3.7-fold increase in pentose phosphates from 10 to 17 DAP; coupled with a  $\leq 2.4$ -fold decrease in 6-phosphogluconate content at 17 and 20 DAP, this strongly suggests a higher flow of carbon through the OPPP in HO embryos. As the OPPP is a major source of reductant (i.e., NADPH) and was shown to be mainly active in the cytosol in pennycress, increased total activity may contribute to FA elongation, and therefore to differential oil accumulation (Tsogtbaatar et al., 2020). Furthermore, malate, a carbon source for *de novo* FA biosynthesis, was 1.8- to 3.5-fold higher in HO embryos at 14 and 17 DAP, respectively. This directly correlates with the same increase in another 4-carbon TCA intermediate, fumarate. A significant 3.2-fold decrease in transcript levels of *isocitrate dehydrogenase catalytic subunit 6* (*IDH-VI*) expression at 17 DAP was observed in HO in the TCA cycle, which according to previous research

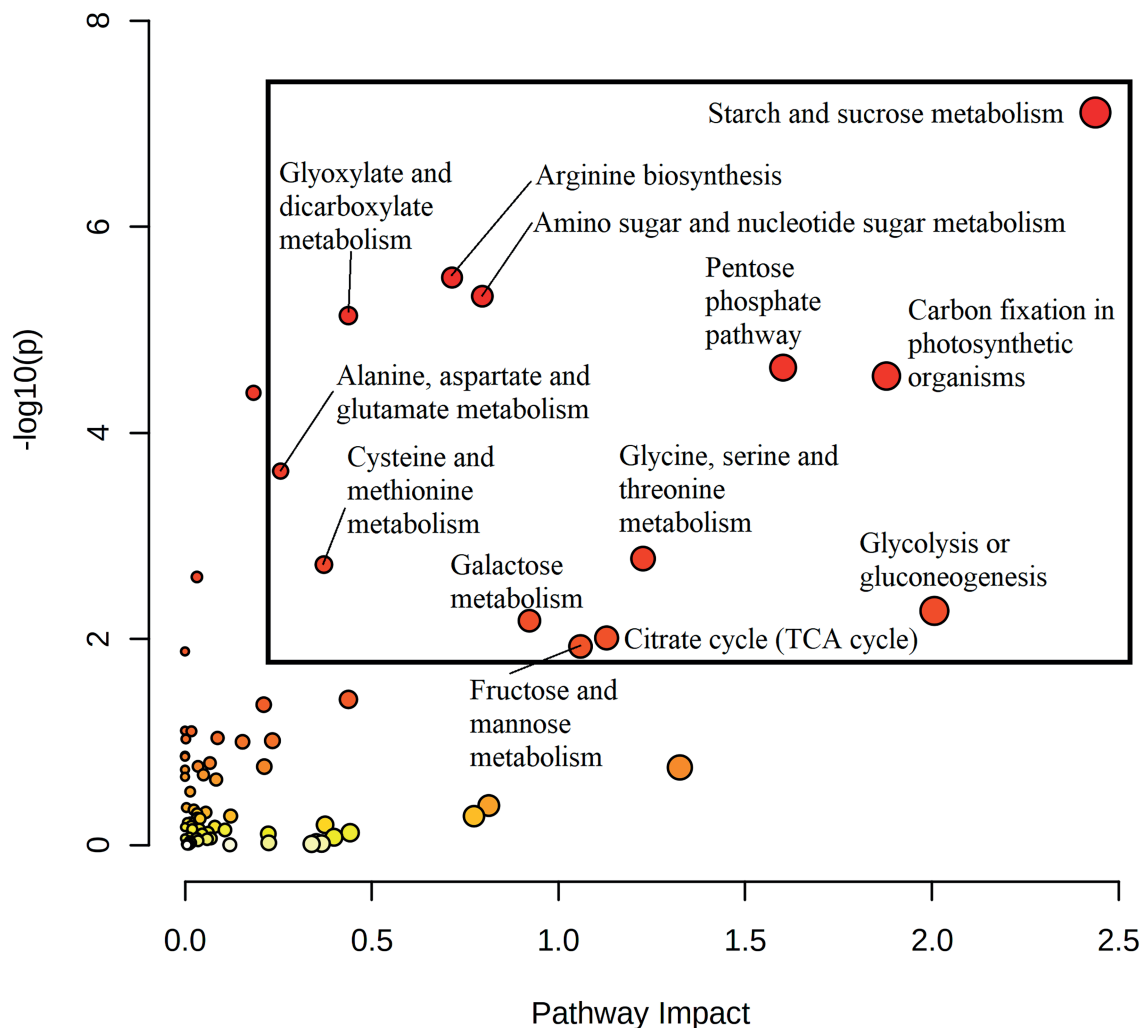




on pennycress suggests a reduction of reverse flow through isocitrate dehydrogenase (Tsogtbaatar et al., 2020). This, along with a 1.9- to 2.8-fold reduction in citrate and a 2.3- to 4.9-fold reduction in isocitrate content, may suggest a reduced flow of carbon from fumarate/malate towards the 6-carbon metabolites of the TCA cycle. As such, this would preserve malate for potential carbon incorporation into *de novo* FA metabolism in the plastid. Interestingly, at 17 DAP there was a 2.8-fold increase in *phosphoenolpyruvate carboxylase 2* (*PPC2*) expression in HO embryos, which was simultaneous with a 4.1-fold increase in *phosphoenolpyruvate carboxylase kinase* (*PPCK1*) expression.

This provides good evidence of both transcriptional (*via PPC2* expression) and post-translational (*via* increased abundance of *PPCK1*) upregulation of the flow of carbon through the anaplerotic pathway from phosphoenolpyruvate (PEP) to oxaloacetate (OAA), as *PPCKs* phosphorylate *PPC* which is generally associated with increased anaplerotic pathway flow (Nimmo, 2003).

The gene *RBCS1B*, encoding a Rubisco subunit, displayed 3-fold upregulation at 17 DAP, increasing all the way to an 1,176-fold increase at 20 DAP in HO. This in addition to up to a 5-fold decrease in ribulose 1,5-bisphosphate (RuBP)



**FIGURE 4** | A joint pathway analysis was performed in order to detect pathway enrichment and analyze pathway topology at 20 DAP (the maximum point of divergence in FA content for both accessions prior to the cessation of metabolic activity). Node size indicates magnitude of pathway impact, whereas node color indicates enrichment analysis significance (with red indicating higher significance). Pathways that returned with a  $p_{\text{adj}} \leq 0.05$  and pathway impact value  $\geq 0.2$  were considered significant. All returned pathways were from the KEGG database (Kanehisa and Goto, 2000; Kanehisa, 2019; Kanehisa et al., 2021).

abundance at 17 DAP suggests a potentially higher degree of carbon fixation. An additional function of Rubisco of re-fixing plastidic  $\text{CO}_2$  lost *via* other metabolic processes has been demonstrated in pennycress embryos and other Brassicaceae, which provides additional C source for *de novo* FA synthesis (Schwender et al., 2004; Tsogtbaatar et al., 2020). The upregulation of Rubisco activity suggested by the increase of *RBCS1B* may allow HO to more efficiently reincorporate lost  $\text{CO}_2$ .

In light of the current knowledge of pennycress, results were interpreted in order to determine if the unusual pathways previously found in pennycress were more active in HO vs. LO embryos. These pathways included: (1) high activity of the OPPP in the cytosol, (2) the reversibility of the isocitrate dehydrogenase that produces isocitrate from  $\alpha$ -ketoglutarate;

and (3) re-fixation of plastidic  $\text{CO}_2$  released by the NADP-dependent malic enzyme and pyruvate dehydrogenase *via* Rubisco (Tsogtbaatar et al., 2020). Our results demonstrate conflicted findings concerning which of these pathways are positively associated with oil production in HO. Specifically, isocitrate and citrate content as well as *TaIDH-VI* expression were lower in HO, suggesting that reversibility of isocitrate dehydrogenase is actually downregulated in HO (Tsogtbaatar et al., 2015, 2020). However, despite increased malate content in HO, no upregulation of genes encoding an NADP-dependent malic enzyme or pyruvate dehydrogenase were noted, which suggests that achieving improved oil content in pennycress is not necessarily limited by carbon incorporation into plastidic *de novo* FA *via* these canonical pathways.

In these results, only a suggested upregulation of the Calvin cycle and higher OPPP activity are consistent with previously published reasons that pennycress has improved carbon conversion efficiency compared to other Brassicaceae (Tsogtbaatar et al., 2015, 2020). Furthermore, improved glycolytic flux *via* increased *PFK5* expression, and enhanced anaplerotic pathway flux *via* increased *PPC2* and *PPCK1* expression may play a key role in HO to achieve higher % w/w oil (or less protein) content. The lack of evidence for increased malate flow into the plastid *via* NADP-dependent malic enzyme, and evidence that higher oil content is not due to reversible ICDH activity in HO therefore suggests that additional pathways or reactions are key to HO producing more oil than its LO counterpart.

### Threonine Aldolase Upregulation May Provide Carbon Sources for Oil Production

The reduced protein accumulation in HO prompted investigation into amino acid metabolism. One of the most upregulated genes involved in central carbon metabolism was *THA2*, encoding a threonine aldolase with no specific predicted localization. Expression was significantly higher in HO at all developmental stages, and by 20 DAP a 17-fold increase in expression was observed (Figure 5A). In *Arabidopsis* and pennycress, significant research has established that *THA2* is generally unidirectional, resulting in the catabolism of threonine to glycine and acetaldehyde, under non-photorespiratory conditions (Jander et al., 2004; Joshi et al., 2006; Tsogtbaatar et al., 2020). As expected, a transient ~1.5-fold decrease in threonine content was observed in HO at 14 and 17 DAP. However, the same trend was observed with glycine (2.5-fold decrease). This provided evidence that accumulation of glycine was not the final result of upregulated *TaTHA2*, and that further flow of carbon through glycine and acetaldehyde was likely.

While multiple roles of threonine aldolases have been suggested in plants, a particular role of interest with regard to FA production is the conversion of its breakdown product acetaldehyde to acetyl-CoA *via* aldehyde dehydrogenases (ALDH) followed by acetyl-CoA synthase (Jander et al., 2004). This strategy serves as a pyruvate dehydrogenase bypass, allowing for carbon from aldehydes to be directly converted to acetate for acetyl-CoA production (Wei et al., 2009). One specific ALDH transcript, *ALDH2C4*, was upregulated 3-fold at 20 DAP, which is notable considering FA production continues in HO embryos after LO has generally ceased accumulation (Supplementary Figure S1A). In *A. thaliana*, *ALDH2C4* is cytosolic and generally associated with phenylpropanoid metabolism (Wei et al., 2009). However, research has suggested it may also play at least a partial role in the pyruvate dehydrogenase bypass to funnel carbon towards acetyl-CoA, in addition to the activity of mitochondrial ALDH enzymes (Nair et al., 2004; Wei et al., 2009). Another route of aldehyde formation is *via* ethanolic fermentation by alcohol dehydrogenase (Wei et al., 2009). A cytosolic *ADH-like 1* gene was upregulated 4.5 to 12.4-fold in HO embryos from 10 to 17 DAP, suggesting a further source of acetaldehyde for acetyl-CoA production is occurring.

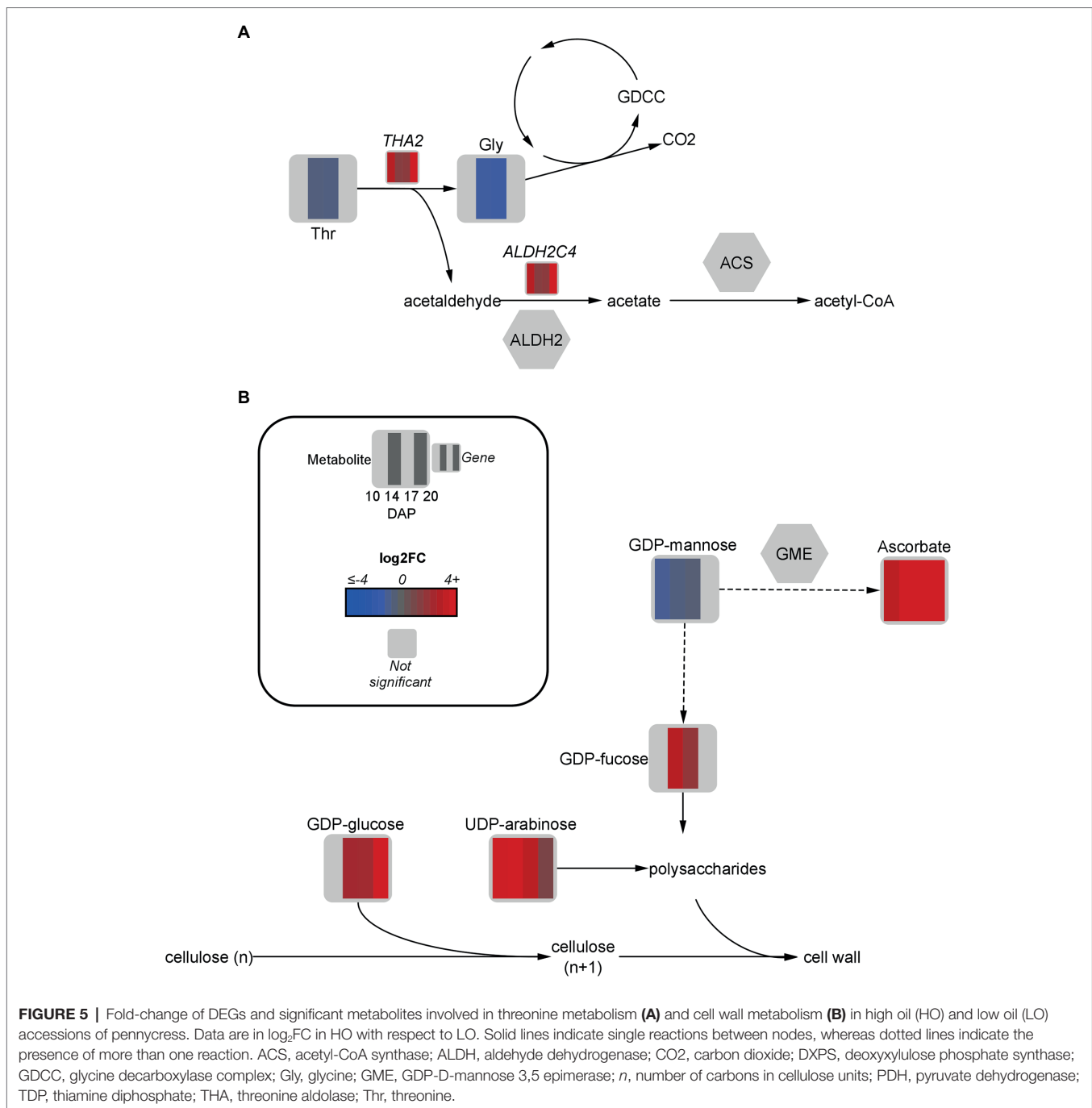
The glycine decarboxylase complex is another potential downstream use of threonine aldolase products, which upon metabolizing glycine releases CO<sub>2</sub> (Igamberdiev et al., 2004). As previously mentioned, this CO<sub>2</sub> can theoretically then be reincorporated *via* the reversible ICDH activity, or into OAA *via* phosphoenolpyruvate carboxykinase (Tsogtbaatar et al., 2020). Further research should be done to determine the flux of threonine-derived glycine and acetaldehyde through these pathways to incorporate carbon towards FA synthesis and elongation.

### Ascorbate and Nucleotide Sugar Metabolism Suggest Cell Wall Contributions

Among the most upregulated metabolites in HO from 17 to 20 DAP were the key cell-wall precursors UDP-arabinose (9.4- and 2.3-fold), GDP-glucose (5.2- and 278.2-fold, respectively), and GDP-fucose (3.8-fold at 17 DAP). Consistent with this upregulation was a slightly altered hemicellulose composition between each accession. Indeed, the accumulation of UDP-arabinose in HO was associated with a 19% reduction in arabinose content in the hemicellulose fraction of the cell wall (Figure 5B). Furthermore, a nearly 30% increase in mannose content in the hemicellulose fraction of HO was observed. The only developmental stages in which GDP-mannose, the precursor for hemicellulosic mannose, was significantly different between accessions was 10 and 14 DAP, where a 1.5 to 2-fold reduction was observed in HO (Brar and Elbein, 1972).

GDP-D-mannose 3,5 epimerase is an enzyme responsible for diverting GDP-mannose away from cell wall production and toward the production of ascorbate (Wolucka and Van Montagu, 2003; Gilbert et al., 2009). Interestingly, ascorbate was confirmed in HO at 8,135 and 11,656 pmol/mg dry weight (DW)<sup>-1</sup> from 17 to 20 DAP, respectively. However, ascorbate was below the limit of quantification in LO and in general has not been previously mentioned as present in pennycress. Ascorbate reduces membrane lipid peroxidation during stress which may preserve total lipid content in HO (Ebrahimian and Bybordi, 2012). Furthermore, ascorbate can act as an alternate electron donor in various areas of plant metabolism including that of embryos, and as such may contribute extra reductant for FA synthesis and/or modification (Mano et al., 2004; Tóth et al., 2009; Hoang et al., 2021). Ascorbate has also been shown to regulate amino acid synthesis, which may relate to the different findings on amino acid and protein content in HO (Gulyás et al., 2017). The novel finding of ascorbate accumulation, along with the altered hemicellulose composition, suggests a difference in cell wall metabolism or redox regulation exists between accessions.

A slight upregulation of *BGLC1* (3.3- and 2.9-fold at 17 and 20 DAP, respectively) was observed in HO. This gene encodes a glycosyl hydrolase that is thought to be able to cleave glucose from xyloglucan, which may have implications on cellular morphology and development (Saffer, 2018; Rubianes et al., 2019). Consistently, a difference in the cellular morphology in the embryonic axis of HO can be seen when compared to LO,

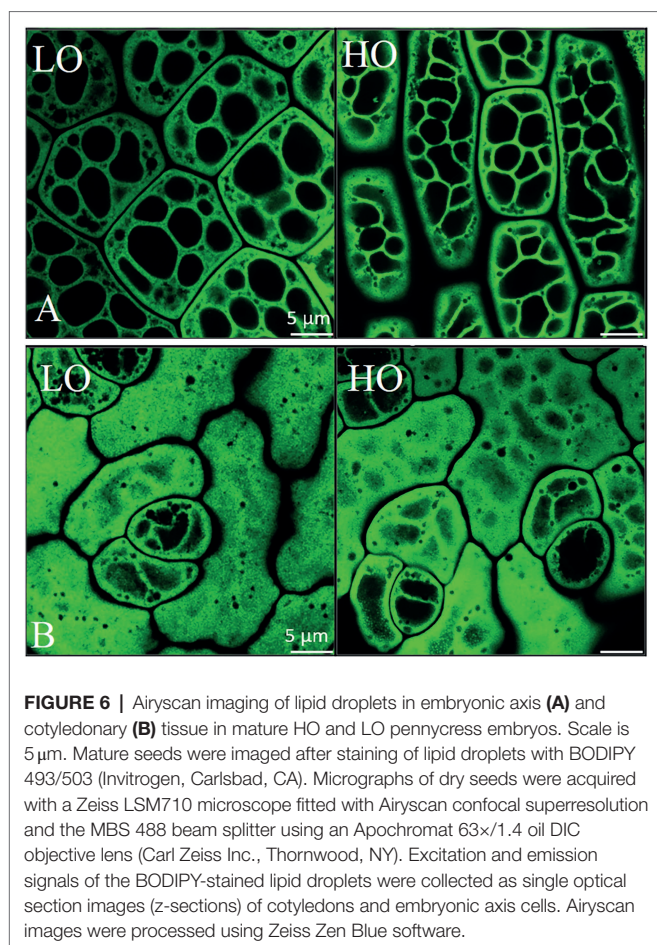


consequently affecting lipid arrangement in this organ (**Figure 6A**). A potential reduction in xyloglucan content at any point of embryo development may allow cleaved glucose to be redirected elsewhere. More research is necessary to determine if the glucose removed from xyloglucan has any fate in oil metabolism or energy production. The increased glycolytic flow suggested in HO could theoretically funnel this glucose to produce more pyruvate. Ultimately, total free sugar content (**Table 1**) was 33% lower in HO on a % w/w basis compared to LO, highlighting that sugar allocation in high oil pennycress is likely altered. The substantial difference in cell wall features between accessions, as well as

their association with ascorbate production, suggests a potential avenue for future research on contributions to lipid metabolism.

### Downregulation of the Ubiquitin Proteasome System and Increased Lipid Droplet Stability Are Associated With Higher Oil Production

Interconnected transcriptional control of protein dynamics, amino acid metabolism, RNA/DNA modification, redox regulation, and lipid and FA metabolism are illustrated by



**FIGURE 6 |** Airyscan imaging of lipid droplets in embryonic axis (A) and cotyledonary (B) tissue in mature HO and LO pennycress embryos. Scale is 5  $\mu$ m. Mature seeds were imaged after staining of lipid droplets with BODIPY 493/503 (Invitrogen, Carlsbad, CA). Micrographs of dry seeds were acquired with a Zeiss LSM710 microscope fitted with Airyscan confocal superresolution and the MBS 488 beam splitter using an ApoChromat 63 $\times$ /1.4 oil DIC objective lens (Carl Zeiss Inc., Thornwood, NY). Excitation and emission signals of the BODIPY-stained lipid droplets were collected as single optical section images (z-sections) of cotyledons and embryonic axis cells. Airyscan images were processed using Zeiss Zen Blue software.

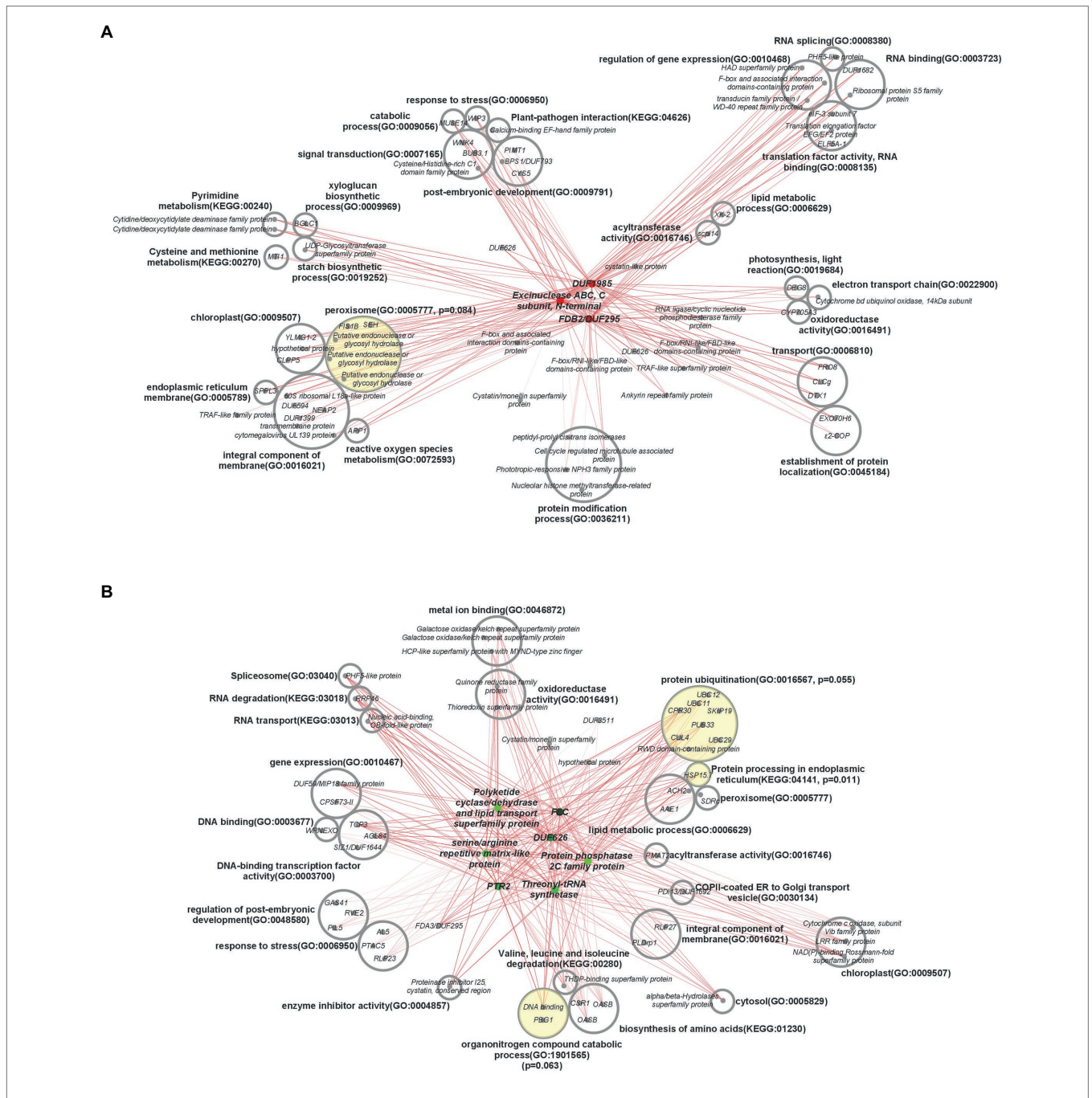
coexpression analysis (Figure 7). Specifically, these results indicated a perturbation in genes associated with the stability of proteins, lipids, and lipid droplets (LD). Protein ubiquitination ( $p_{\text{adj}}=0.055$ ) and protein processing in the endoplasmic reticulum ( $p_{\text{adj}}=0.011$ ) were significantly over-represented connections to downregulated hub genes. In plants, the ubiquitin-proteasome system is associated with a diverse array of functions, particularly proteolytic action related to plant development, immunity, stress, and hormone response (Liu et al., 2020). Transcription of *UBC11*, associated in previous research with protein degradation in aging seeds (Figure 8A), was downregulated 150.1–739.3 fold throughout development in HO (Wang et al., 2018). This same trend was exhibited by both *UBC12* (17.4–278.2 fold decrease) and *UBC29* (16.4–634.7 fold decrease). Recent research has also suggested a link between cullin-RING ligases, enzymes critical to targeting proteins for ubiquitination, and lipid production in photosynthetic organisms. *Chlamydomonas reinhardtii* was shown to display a lower biomass per cell but higher total lipid content per cell when one of these ligases, *CULLIN4* (*CUL4*), was silenced using RNAi (Luo et al., 2021). This is highly consistent with our biomass results in pennycress, as *CUL4* was constitutively downregulated 4.3–5.7-fold throughout development. Taken together, the overall knockdown of a specialized component

of the protein ubiquitination system appears to be related to oil production in pennycress.

HO embryos displayed a 5.4-fold increase in expression (17 DAP) of *FIS1B*, homologous with an *A. thaliana* gene which is noted to be directly responsible for peroxisome fission following elongation (Lingard et al., 2008; Zhang and Hu, 2008). Peroxisomes encapsulate a large portion of the system required for FA catabolism in plants (Wright and Bartel, 2020). Knockout of *FIS1B* activity has been shown to result in elongated, enlarged peroxisomes, while the inverse is true following insertion of *FIS1A* (Zhang and Hu, 2009; Kao et al., 2018). An overabundance of this protein in pennycress may therefore result in differential effects on FA catabolism (Figure 8B). Although no DEGs were detected that were directly involved TAG hydrolysis or  $\beta$ -oxidation (Supplementary Data Sheet S1), a reduction in the cell volume available for FA catabolism could theoretically shift the balance between oil synthesis and catabolism toward synthesis.

We examined the organization of lipid storage in cells of both genotypes by enhanced-resolution, confocal fluorescence microscopy. BODIPY 495/503 selectively stains neutral lipids in cells and here the arrangement of LDs, packed with neutral lipids, is visualized in cotyledons of both HO and LO embryos (Figure 6B). The cellular distribution of LDs in both genotypes was similar although the LDs may be more tightly packed together in HO. Transcriptomic contributions to the cellular organization of lipid storage are suggested by DEG results. The gene encoding the Coat Protein I (COPI) epsilon subunit,  $\epsilon 2$ -COP, was upregulated 3.3–6.2 fold throughout development (Figure 8C). The COPI complex is has been demonstrated to associate the LD with the ER, allowing for targeting of proteins to the LD membrane (Wilfling et al., 2014; Huang et al., 2019). Furthermore, a  $\geq 81$ -fold increase throughout development in HO of *EXO70H6*, part of the largely uncharacterized exocyst complex, was found to affect the phospholipid composition in membranes, cell wall thickening, and overall recognition of lipids at the membrane surface (He et al., 2007). In *A. thaliana*, the loss of *serine-carboxypeptidase-like 41* (*SCPL41*), involved in protein turnover, has resulted in increased levels of membrane lipids (Chen et al., 2020). The observed 74–286 fold increase in *SCPL14* in HO may suggest that this isoform could interact with LDs or membrane lipids to facilitate the efficient subcellular organization of LDs (Figure 8C).

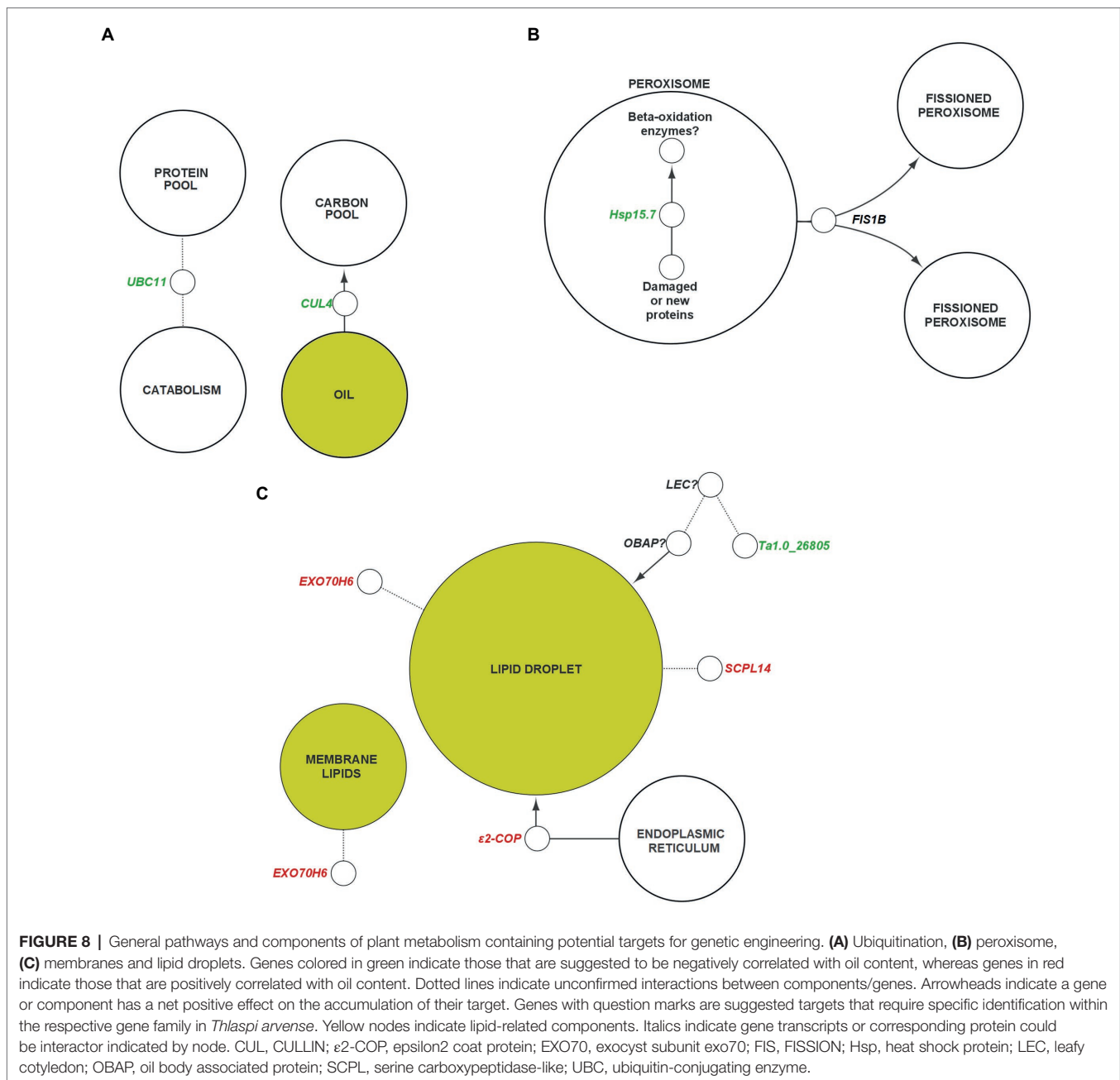
A polyketide cyclase/dehydrase and lipid transport superfamily protein, *Ta1.0\_26805*, displayed high expression in LO throughout development, but was undetectable in HO. Previous research in *A. thaliana* showed it to be involved in a network that contains a mutual regulator/interactor (*LEC1*) of *OBAP1a*, a gene involved in LD stability (López-Ribera et al., 2014; Depuydt and Vandepoel, 2021). While there was no differential expression of the pennycress *LEC1* homolog, the lack of apparent *Ta1.0\_26805* transcription in HO warrants investigation to determine if any interactions are present between it and another transcriptional regulator in pennycress. The discovery of a regulator similar in function to *LEC1* in pennycress might provide potential as a regulatory target for improving LD assembly or stability, however the possibility that there are



**FIGURE 7 |** Central coexpression network of DEGs containing top 10 major hub genes and their respective first neighbors. **(A)** Three main hub genes that were upregulated in high oil accession, **(B)** seven main hub genes that were downregulated in high oil accession. Hub genes are in the center of each network, in color, and first neighbor genes are categorized by gene ontology terms. Circles with yellow shading indicate over-represented gene ontology terms, with *p*-values based on the false discovery rate, and their associated genes. DUF = gene with a domain of unknown function. Darkness of lines indicates correlation, with darker indicating higher correlation. ORA was carried out using a threshold of 0.10.

sequence dissimilarities in *Ta1.0\_26805* between accessions must also be considered. LEC1 is part of a regulatory network of transcription factors known as LAFL responsible for controlling seed maturation and accumulation of storage components (Santos-Mendoza et al., 2008; Verdier and Thompson, 2008). Interestingly none of the other genes of LAFL nor several

known targets of these genes were differentially expressed in our data (Fatihi et al., 2016; **Supplementary Data Sheet S1**). Together these results suggest an enhanced system in HO for maintaining the access of biosynthetic enzymes to membranes and LDs, together with a reduction in a component of the ubiquitin-proteasome system.



## Comparative “Omics” Analysis Suggests Future Targets for Genetic Engineering

Oil production, from *de novo* FA synthesis to lipid packaging, is an integrated process and requires multiple approaches to assess its associated metabolism. This research illustrates evidence of several specific regions of plant metabolism that are perturbed during high oil production. Additionally, it displays evidence that previously-identified mechanisms that allow pennycress to produce oil more efficiently than fellow Brassicaceae are not necessarily targets for improving oil production within pennycress itself.

This research highlights four metabolic processes that appear to be positively associated with pennycress oil content:

(1) carbon partitioning in central carbon metabolism that accumulates malate, pentose phosphates, increased flow through the glycolysis and anaplerotic pathway; (2) production of carbon precursors *via* threonine aldolase activity, either by pyruvate bypass and/or the glycine decarboxylase complex; (3) alteration of sugar incorporation into cell wall components, potentially *via* ascorbate metabolism; and (4) expression-level alteration of ubiquitination and LD organization. Augmentation of these cellular processes may provide lucrative genetic engineering targets for pennycress and other oilseeds as well as critical insights into the cross-talk between lipid metabolism and other pathways in plants in general.

## DATA AVAILABILITY STATEMENT

The raw data presented in the study are deposited as **Supplementary Material** and in the [https://www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/repository) repository, accession number PRJNA808106.

## AUTHOR CONTRIBUTIONS

AA and EG designed the experiments. CJ, EO, TR, and AG performed the research. CJ, LTG, EO, TR, and AG performed data analysis and collection. CJ, LTG, EO, TR, AG, KC, EG, and AA performed data interpretation. CJ wrote the manuscript. LTG, EO, TR, AG, KC, EG, and AA assisted with manuscript preparation. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431
- Alonso, A. P., Piasecki, R. J., Wang, Y., LaClair, R. W., and Shachar-Hill, Y. (2010). Quantifying the labeling and the levels of plant cell wall precursors using ion chromatography tandem mass spectrometry. *Plant Physiol.* 153, 915–924. doi: 10.1104/pp.110.155713
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed January 24, 2022).
- Astute Analytica (2021). Global Oleochemicals Market, By Product (Alkoxyates, Fatty Acid MethylEster, Glycerin, Fatty Alcohol, Others), By Application (F&B, Textiles, Industrial, Paints & Inks, Others), By Sales Channel (Direct, Indirect), Estimation & Forecast, 2017–2027. Available at: [https://www.researchandmarkets.com/reports/5411590/global-oleochemicals-market-by-product?utm\\_source=CI&utm\\_medium=PressRelease&utm\\_code=xz5vkw&utm\\_campaign=1584894+-+Global+Oleochemicals+Market+Report+2021-2027+Featuring+Major+Players+from+Indonesia%2C+M](https://www.researchandmarkets.com/reports/5411590/global-oleochemicals-market-by-product?utm_source=CI&utm_medium=PressRelease&utm_code=xz5vkw&utm_campaign=1584894+-+Global+Oleochemicals+Market+Report+2021-2027+Featuring+Major+Players+from+Indonesia%2C+M) (Accessed November 16, 2021).
- Banković-Ilić, I. B., Stamenković, O. S., and Veljković, V. B. (2012). Biodiesel production from non-edible plant oils. *Renew. Sust. Energ. Rev.* 16, 3621–3647. doi: 10.1016/j.rser.2012.03.002
- Baud, S., and Lepiniec, L. (2009). Regulation of de novo fatty acid synthesis in maturing oilseeds of *Arabidopsis*. *Plant Physiol. Biochem.* 47, 448–455. doi: 10.1016/j.plaphy.2008.12.006
- Baud, S., and Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Prog. Lipid Res.* 49, 235–249. doi: 10.1016/j.plipres.2010.01.001
- Blakeley, K. (2012). CRS Report for Congress DOD Alternative Fuels: Policy, Initiatives and Legislative Activity. Available at: [www.crs.gov](http://www.crs.gov) (Accessed March 15, 2021).
- Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G.-F., Drira, N., Ohlrogge, J. B., et al. (2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc. Natl. Acad. Sci.* 108, 12527–12532. doi: 10.1073/PNAS.1106502108
- Brar, S. S., and Elbein, A. D. (1972). Incorporation of mannose into hemicellulose a in mung bean seedlings. *Phytochemistry* 11, 651–656. doi: 10.1016/0031-9422(72)80028-1
- Cagliari, A., Margis, R., Dos Santos Maraschin, F., Turchetto-Zolet, A. C., Loss, G., and Margis-Pinheiro, M. (2011). Biosynthesis of Triacylglycerols (TAGs) in plants and algae. *Int. J. Plant Biol.* 2, 40–52. doi: 10.4081/PB.2011.E10
- Cai, Y., Goodman, J. M., Pyc, M., Mullen, R. T., Dyer, J. M., and Chapman, K. D. (2015). *Arabidopsis* SEIPIN proteins modulate triacylglycerol accumulation

## FUNDING

This work was supported by the Department of Energy Office of Science, Office of Biological and Environmental Research (BER), grant # DE-SC0019233 and # DE-SC0020325.

## ACKNOWLEDGMENTS

We thank Jordan LaChance for technical assistance with this project.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.943585/full#supplementary-material>

- and influence lipid droplet proliferation. *Plant Cell* 27, 2616–2636. doi: 10.1105/TPC.15.00588
- Chen, J., Li, W., and Jia, Y. (2020). The serine carboxypeptidase-like gene SCPL41 negatively regulates membrane lipid metabolism in *Arabidopsis thaliana*. *Plants* 9:696. doi: 10.3390/PLANTS9060696
- Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., and Lin, C.-Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8:S11. doi: 10.1186/1752-0509-8-S4-S11
- Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 46, W486–W494. doi: 10.1093/nar/gky310
- Chong, J., Wishart, D. S., and Xia, J. (2019). Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Curr. Protoc. Bioinformatics* 68:e86. doi: 10.1002/cpbi.86
- Claver, A., Rey, R., López, M. V., Picorel, R., and Alfonso, M. (2017). Identification of target genes and processes involved in erucic acid accumulation during seed development in the biodiesel feedstock pennycress (*Thlaspi arvense* L.). *J. Plant Physiol.* 208, 7–16. doi: 10.1016/j.jplph.2016.10.011
- Cocuron, J. C., and Alonso, A. P. (2014). Liquid chromatography tandem mass spectrometry for measuring <sup>13</sup>C-labeling in intermediates of the glycolysis and pentose phosphate pathway. *Methods Mol. Biol.* 1090, 131–142. doi: 10.1007/978-1-62703-688-7\_9
- Cocuron, J. C., Anderson, B., Boyd, A., and Alonso, A. P. (2014). Targeted metabolomics of *Physaria fendleri*, an industrial crop producing hydroxy fatty acids. *Plant Cell Physiol.* 55, 620–633. doi: 10.1093/pcp/pcu011
- Cocuron, J.-C., Koubaa, M., Kimmelfield, R., Ross, Z., and Alonso, A. P. (2019). A combined metabolomics and fluxomics analysis identifies steps limiting oil synthesis in maize embryos. *Plant Physiol.* 181, 961–975. doi: 10.1104/PP.19.00920
- Cocuron, J. C., Tsogtbaatar, E., and Alonso, A. P. (2018). Tracking the flow of carbon toward oil synthesis in an emerging crop lesquerella (*Physaria fendleri*). *FASEB J.* 31:629.16. doi: 10.1096/FASEB.31.1\_SUPPLEMENT.629.16
- Csardi, G. (2014). The igraph software package for complex network research. Available at: <https://www.researchgate.net/publication/221995787> (Accessed July 12, 2021).
- Cubins, J. A., Wells, M. S., Frels, K., Ott, M. A., Forcella, F., Johnson, G. A., et al. (2019). Management of pennycress as a winter annual cash cover crop. A review. *Agron. Sustain. Dev.* 39, 1–11. doi: 10.1007/s13593-019-0592-0
- Dahlqvist, A., Ståhl, U., Lenman, M., Banas, A., Lee, M., Sandager, L., et al. (2000). Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6487–6492. doi: 10.1073/PNAS.120067297
- Dennis, D. T., and Miernyk, J. A. (1982). Compartmentation of nonphotosynthetic carbohydrate metabolism. *Annu. Rev. Plant Physiol.* 33, 27–50. doi: 10.1146/annurev.pp.33.060182.000331



- Depuydt, T., and Vandepoele, K. (2021). Multi-omics network-based functional annotation of unknown *Arabidopsis* genes. *Plant J.* 108, 1193–1212. doi: 10.1101/2021.06.17.448819
- Dorn, K. M., Fankhauser, J. D., Wyse, D. L., and Marks, M. D. (2015). A draft genome of field pennycress (*Thlaspi arvense*) provides tools for the domestication of a new winter biofuel crop. *DNA Res.* 22, 121–131. doi: 10.1093/dnares/dsu045
- Dose, H. L., Eberle, C. A., Forcella, F., and Gesch, R. W. (2017). Early planting dates maximize winter annual field pennycress (*Thlaspi arvense* L.) yield and oil content. *Ind. Crop. Prod.* 97, 477–483. doi: 10.1016/j.indcrop.2016.12.039
- Drenth, A. C., Olsen, D. B., Cabot, P. E., and Johnson, J. J. (2014). Compression ignition engine performance and emission evaluation of industrial oilseed biofuel feedstocks camelina, carinata, and pennycress across three fuel pathways. *Fuel* 136, 143–155. doi: 10.1016/j.fuel.2014.07.048
- Eberle, C. A., Thom, M. D., Nemeček, K. T., Forcella, F., Lundgren, J. G., Gesch, R. W., et al. (2015). Using pennycress, camelina, and canola cash cover crops to provision pollinators. *Ind. Crop. Prod.* 75, 20–25. doi: 10.1016/j.indcrop.2015.06.026
- Ebrahimian, E., and Bybordi, A. (2012). Effect of salinity, salicylic acid, silicium and ascorbic acid on lipid peroxidation, antioxidant enzyme activity and fatty acid content of sunflower. *Afr. J. Agric. Res.* 7, 3685–3694. doi: 10.5897/AJAR12.492
- Fatihhi, A., Boulard, C., Bouyer, D., Baud, S., Dubreucq, B., and Lepiniec, L. (2016). Deciphering and modifying LAFL transcriptional regulatory network in seed for improving yield and quality of storage compounds. *Plant Sci.* 250, 198–204. doi: 10.1016/J.PLANTSCI.2016.06.013
- Gerst, R., and Hölzer, M. (2018). PCAGO: an interactive tool to analyze RNA-Seq data with principal component analysis. bioRxiv [Preprint]. 433078. doi: 10.1101/433078
- Gidda, S. K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P., et al. (2016). Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant Physiol.* 170, 2052–2071. doi: 10.1104/PP.15.01977
- Gilbert, L., Alhaghdow, M., Nunes-Nesi, A., Quemener, B., Guillon, F., Bouchet, B., et al. (2009). GDP-d-mannose 3,5-epimerase (GME) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *Plant J.* 60, 499–508. doi: 10.1111/J.1365-313X.2009.03972.X
- Grand View Research (2017). *U.S. Specialty Oleochemicals Market Analysis by Product (Specialty Esters, Fatty Acid Methyl Esters, Glycerol Esters, Alkoxyates, Fatty Amines), by Application, Competitive Landscape, and Segment Forecasts, 2014–2025*. San Francisco, CA, Grand View Research.
- Gulyás, Z., Simon-Sarkadi, L., Badics, E., Novák, A., Mednyánszky, Z., Szalai, G., et al. (2017). Redox regulation of free amino acid levels in *Arabidopsis thaliana*. *Physiol. Plant.* 159, 264–276. doi: 10.1111/PPL.12510
- Hankin, J. A., Barkley, R. M., and Murphy, R. C. (2007). Sublimation as a method of matrix application for mass spectrometric imaging. *J. Am. Soc. Mass Spectrom.* 18, 1646–1652. doi: 10.1016/J.JASMS.2007.06.010
- He, B., Xi, F., Zhang, X., Zhang, J., and Guo, W. (2007). Exo70 interacts with phospholipids and mediates the targeting of the exocyst to the plasma membrane. *EMBO J.* 26, 4053–4065. doi: 10.1038/SJ.EMBOJ.7601834
- Hills, M. J. (2004). Control of storage-product synthesis in seeds. *Curr. Opin. Plant Biol.* 7, 302–308. doi: 10.1016/j.pbi.2004.03.003
- Hoang, M. T. T., Almeida, D., Chay, S., Alcon, C., Corratge-Faillie, C., Curie, C., et al. (2021). AtDTX25, a member of the multidrug and toxic compound extrusion family, is a vacuolar ascorbate transporter that controls intracellular iron cycling in *Arabidopsis*. *New Phytol.* 231, 1956–1967. doi: 10.1111/NPH.17526
- Horn, P. J., and Chapman, K. D. (2014). Metabolite imager: customized spatial analysis of metabolite distributions in mass spectrometry imaging. *Metabolomics* 10, 337–348. doi: 10.1007/S11306-013-0575-0/FIGURES/6
- Horn, P. J., Liu, J., Cocuron, J. C., McGlew, K., Thrower, N. A., Larson, M., et al. (2016). Identification of multiple lipid genes with modifications in expression and sequence associated with the evolution of hydroxy fatty acid accumulation in *Physaria fendleri*. *Plant J.* 86, 322–348. doi: 10.1111/tpj.13163
- Horn, P. J., Silva, J. E., Anderson, D., Fuchs, J., Borisjuk, L., Nazarenus, T. J., et al. (2013). Imaging heterogeneity of membrane and storage lipids in transgenic *Camelina sativa* seeds with altered fatty acid profiles. *Plant J.* 76, 138–150. doi: 10.1111/TPJ.12278
- Hu, Z., Ren, Z., and Lu, C. (2012). The phosphatidylcholine diacylglycerol cholinephosphotransferase is required for efficient hydroxy fatty acid accumulation in transgenic *Arabidopsis*. *Plant Physiol.* 158, 1944–1954. doi: 10.1104/PP.111.192153
- Huang, S., Jiang, L., and Zhuang, X. (2019). Possible roles of membrane trafficking components for lipid droplet dynamics in higher plants and green algae. *Front. Plant Sci.* 10:207. doi: 10.3389/FPLS.2019.00207
- Igamberdiev, A. U., Mikkelsen, T. N., Ambus, P., Bauwe, H., Lea, P. J., and Gardestrom, P. (2004). Photorespiration contributes to stomatal regulation and carbon isotope fractionation: a study with barley, potato and *Arabidopsis* plants deficient in glycine decarboxylase. *Photosynth. Res.* 81, 139–152. doi: 10.1023/B:PRES.0000035026.05237.EC
- Jander, G., Norris, S. R., Joshi, V., Fraga, M., Rugg, A., Yu, S., et al. (2004). Application of a high-throughput HPLC-MS/MS assay to *Arabidopsis* mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. *Plant J.* 39, 465–475. doi: 10.1111/J.1365-313X.2004.02140.X
- Johnson, G. A., Kantar, M. B., Betts, K. J., and Wyse, D. L. (2015). Field pennycress production and weed control in a double crop system with soybean in Minnesota. *Agron. J.* 107, 532–540. doi: 10.2134/agronj14.0292
- Joshi, V., Laubengayer, K. M., Schauer, N., Fernie, A. R., and Jander, G. (2006). Two *Arabidopsis* threonine aldolases are nonredundant and compete with threonine deaminase for a common substrate pool. *Plant Cell* 18, 3564–3575. doi: 10.1105/TPC.106.044958
- Joubès, J., Raffaele, S., Bourdenx, B., Garcia, C., Laroche-Traineau, J., Moreau, P., et al. (2008). The VLCFA elongase gene family in *Arabidopsis thaliana*: phylogenetic analysis, 3D modelling and expression profiling. *Plant Mol. Biol.* 67, 547–566. doi: 10.1007/S11103-008-9339-Z
- Junker, B. H., Lonien, J., Heady, L. E., Rogers, A., and Schwender, J. (2007). Parallel determination of enzyme activities and in vivo fluxes in *Brassica napus* embryos grown on organic or inorganic nitrogen source. *Phytochemistry* 68, 2232–2242. doi: 10.1016/J.PHYTOCHEM.2007.03.032
- Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* 28, 1947–1951. doi: 10.1002/pro.3715
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 49, D545–D551. doi: 10.1093/nar/gkaa970
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kao, Y. T., Gonzalez, K. L., and Bartel, B. (2018). Peroxisome function, biogenesis, and dynamics in plants. *Plant Physiol.* 176, 162–177. doi: 10.1104/PP.17.01050
- Kennedy, P. E. (1961). Biosynthesis of complex lipids. *Fed. Proc.* 20, 934–940.
- Kim, H. U., and Chen, G. Q. (2015). Identification of hydroxy fatty acid and triacylglycerol metabolism-related genes in *lesquerella* through seed transcriptome analysis. *BMC Genomics* 16, 1–22. doi: 10.1186/S12864-015-1413-8
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Koubaa, M., Cocuron, J. C., Thomasset, B., and Alonso, A. P. (2013). Highlighting the tricarboxylic acid cycle: liquid and gas chromatography-mass spectrometry analyses of <sup>13</sup>C-labeled organic acids. *Anal. Biochem.* 436, 151–159. doi: 10.1016/j.ab.2013.01.027
- Krishnakumar, V., Hanlon, M. R., Contrino, S., Ferlanti, E. S., Karamycheva, S., Kim, M., et al. (2015). Araport: the *Arabidopsis* information portal. *Nucleic Acids Res.* 43, D1003–D1009. doi: 10.1093/NAR/GKU1200
- Kubis, S. E., Pike, M. J., Everett, C. J., Hill, L. M., and Rawsthorne, S. (2004). The import of phosphoenolpyruvate by plastids from developing embryos of oilseed rape, *Brassica napus* (L.), and its potential as a substrate for fatty acid synthesis. *J. Exp. Bot.* 55, 1455–1462. doi: 10.1093/jxb/erh157
- Kumar, R., Sundari, I. S., Sen, S., Dasgupta, N., and Chidambaram, R. (2016). “Animal fat- and vegetable oil-based platform chemical biorefinery” in *Platform Chemical Biorefinery: Future Green Industry*. eds. S. Kaur Brar, S. Jyoti Sarma and K. B. T. Pakshirajan (Amsterdam: Elsevier), 361–377.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656
- Lingard, M. J., Gidda, S. K., Bingham, S., Rothstein, S. J., Mullen, R. T., and Trelease, R. N. (2008). Arabidopsis PEROXIN11c-e, FISSION1b, and DYNAMIN-RELATED PROTEIN3A cooperate in cell cycle-associated replication of peroxisomes. *Plant Cell* 20, 1567–1585. doi: 10.1105/TPC.107.057679
- Liu, W., Tang, X., Qi, X., Fu, X., Ghimire, S., Ma, R., et al. (2020). The ubiquitin conjugating enzyme: an important ubiquitin transfer platform in ubiquitin-proteasome system. *Int. J. Mol. Sci.* 21, 2894. doi: 10.3390/IJMS21082894
- López-Ribera, I., La Paz, J. L., Repiso, C., García, N., Miquel, M., Hernández, M. L., et al. (2014). The evolutionary conserved oil body associated protein OBAP1 participates in the regulation of oil body size. *Plant Physiol.* 164, 1237–1249. doi: 10.1104/PP.113.233221
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Lu, S., Aziz, M., Sturtevant, D., Chapman, K. D., and Guo, L. (2020). Heterogeneous distribution of erucic acid in *Brassica napus* seeds. *Front. Plant Sci.* 10:1744. doi: 10.3389/FPLS.2019.01744/FULL
- Lu, C., Xin, Z., Ren, Z., Miquel, M., and Browse, J. (2009). An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 106, 18837–18842. doi: 10.1073/PNAS.0908848106
- Luo, Q., Zou, X., Wang, C., Li, Y., and Hu, Z. (2021). The roles of Cullins E3 ubiquitin ligases in the lipid biosynthesis of the green microalgae *Chlamydomonas reinhardtii*. *Int. J. Mol. Sci.* 22:4695. doi: 10.3390/IJMS22094695
- Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/BIOINFORMATICS/BTI551
- Mano, J., Hídeg, É., and Asada, K. (2004). Ascorbate in thylakoid lumen functions as an alternative electron donor to photosystem II and photosystem I. *Arch. Biochem. Biophys.* 429, 71–80. doi: 10.1016/J.ABB.2004.05.022
- Markel, E., English, B. C., Hellwinckel, C., and Menard, R. J. (2018). Potential for pennycress to support a renewable jet fuel industry. *Ecol. Pollut. Environ. Sci.* 1, 95–102.
- Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albu, L. P., Mushayama, T., et al. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* 49, D394–D403. doi: 10.1093/nar/gkaa1106
- Mi, H., and Thomas, P. (2009). PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol.* 563, 123–140. doi: 10.1007/978-1-60761-175-2\_7
- Millar, A. A., and Kunst, L. (1997). Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant J.* 12, 121–131. doi: 10.1046/J.1365-313X.1997.12010121.X
- Moser, B. R., Knothe, G., Vaughn, S. F., and Isbell, T. A. (2009a). Production and evaluation of biodiesel from field pennycress (*Thlaspi arvense* L.) oil. *Energy Fuel* 23, 4149–4155. doi: 10.1021/ef900337g
- Moser, B. R., Shah, S. N., Winkler-Moser, J. K., Vaughn, S. F., and Evangelista, R. L. (2009b). Composition and physical properties of cress (*Lepidium sativum* L.) and field pennycress (*Thlaspi arvense* L.) oils. *Ind. Crop. Prod.* 30, 199–205. doi: 10.1016/j.indcrop.2009.03.007
- Mousavi-Avval, S. H., and Shah, A. (2020). Techno-economic analysis of pennycress production, harvest and post-harvest logistics for renewable jet fuel. *Renew. Sust. Energ. Rev.* 123:109764. doi: 10.1016/j.rser.2020.109764
- Nair, R. B., Bastress, K. L., Ruegger, M. O., Denault, J. W., and Chapple, C. (2004). The Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. *Plant Cell* 16, 544–554. doi: 10.1105/tpc.017509
- Nimmo, H. G. (2003). Control of the phosphorylation of phosphoenolpyruvate carboxylase in higher plants. *Arch. Biochem. Biophys.* 414, 189–196. doi: 10.1016/S0003-9861(03)00115-2
- Pasha, A., Subramaniam, S., Cleary, A., Chen, X., Berardini, T., Farmer, A., et al. (2020). Araport lives: an updated framework for Arabidopsis bioinformatics. *Plant Cell* 32, 2683–2686. doi: 10.1105/TPC.20.00358
- Perteau, M., Kim, D., Perteau, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095
- Phippen, W. B., and Phippen, M. E. (2013). Seed oil characteristics of wild field pennycress (*Thlaspi arvense* L.) populations and USDA accessions. Available at: <http://www.wiu.edu/pennycress/current-experiments/Pennycressseed-oil2013.pdf> (Accessed September 28, 2018).
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47, W191–W198. doi: 10.1093/nar/gkz369
- Revelle, W. (2018). psych: Procedures for Personality and Psychological Research — Northwestern Scholars. Available at: <https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research> (Accessed July 12, 2021).
- Romsdahl, T. B., Kambhampati, S., Koley, S., Yadav, U. P., Alonso, A. P., Allen, D. K., et al. (2021). Analyzing mass spectrometry imaging data of 13C-labeled phospholipids in *Camelina sativa* and *Thlaspi arvense* (pennycress) embryos. *Metabolites* 11, 148. doi: 10.3390/METABO11030148
- Rubianes, D., Valdivia, E. R., Revilla, G., Zarra, I., and Sampedro, J. (2019). Xyloglucan exoglycosidases in the monocot model *Brachypodium distachyon* and the conservation of xyloglucan disassembly in angiosperms. *Plant Mol. Biol.* 100, 495–509. doi: 10.1007/s11103-019-00875-1
- Saffer, A. M. (2018). Expanding roles for pectins in plant development. *J. Integr. Plant Biol.* 60, 910–923. doi: 10.1111/JIPB.12662
- Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., and Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *Plant J.* 54, 608–620. doi: 10.1111/J.1365-313X.2008.03461.X
- Schwender, J., Goffman, F., Ohlrogge, J. B., and Shachar-Hill, Y. (2004). Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432, 779–782. doi: 10.1038/nature03145
- Schwender, J., Ohlrogge, J. B., and Shachar-Hill, Y. (2003). A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. *J. Biol. Chem.* 278, 29442–29453. doi: 10.1074/jbc.M303432200
- Sedbrook, J. C., Phippen, W. B., and Marks, M. D. (2014). New approaches to facilitate rapid domestication of a wild plant to an oilseed crop: example pennycress (*Thlaspi arvense* L.). *Plant Sci.* 227, 122–132. doi: 10.1016/j.plantsci.2014.07.008
- Sturtevant, D., Aziz, M., Romsdahl, T. B., Corley, C. D., and Chapman, K. D. (2021). In situ localization of plant lipid metabolites by matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI). *Methods Mol. Biol.* 2295, 417–438. doi: 10.1007/978-1-0716-1362-7\_24
- Subedi, U., Jayawardhane, K. N., Pan, X., Ozga, J., Chen, G., Foroud, N. A., et al. (2020). The potential of genome editing for improving seed oil content and fatty acid composition in oilseed crops. *Lipids* 55, 495–512. doi: 10.1002/LIPD.12249
- Thelen, J. J., and Ohlrogge, J. B. (2002). Metabolic engineering of fatty acid biosynthesis in plants. *Metab. Eng.* 4, 12–21. doi: 10.1006/mben.2001.0204
- Tóth, S. Z., Puthur, J. T., Nagy, V., and Garab, G. (2009). Experimental evidence for ascorbate-dependent electron transport in leaves with inactive oxygen-evolving complexes. *Plant Physiol.* 149, 1568–1578. doi: 10.1104/PP.108.132621
- Tsogtbaatar, E., Cocuron, J.-C., Alonso, A. P., and Lunn, J. (2020). Non-conventional pathways enable pennycress (*Thlaspi arvense* L.) embryos to achieve high efficiency of oil biosynthesis. *J. Exp. Bot.* 71, 3037–3051. doi: 10.1093/jxb/eraa060
- Tsogtbaatar, E., Cocuron, J.-C., Sonera, M. C., and Alonso, A. P. (2015). Metabolite fingerprinting of pennycress (*Thlaspi arvense* L.) embryos to assess active pathways during oil synthesis. *J. Exp. Bot.* 66, 4267–4277. doi: 10.1093/jxb/erv020
- Verdier, J., and Thompson, R. D. (2008). Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol.* 49, 1263–1271. doi: 10.1093/PCP/PCN116
- Wang, T., Hou, L., Jian, H., Di, F., Li, J., and Liu, L. (2018). Combined QTL mapping, physiological and transcriptomic analyses to identify candidate genes involved in *Brassica napus* seed aging. *Mol. Gen. Genomics* 293, 1421–1435. doi: 10.1007/s00438-018-1468-8

- Wang, L., Shen, W., Kazachkov, M., Chen, G., Chen, Q., Carlsson, A. S., et al. (2012). Metabolic interactions between the lands cycle and the Kennedy pathway of glycerolipid synthesis in Arabidopsis developing seeds. *Plant Cell* 24, 4652–4669. doi: 10.1105/TPC.112.104604
- Wei, Y., Lin, M., Oliver, D. J., and Schnable, P. S. (2009). The roles of aldehyde dehydrogenases (ALDHs) in the PDH bypass of Arabidopsis. *BMC Biochem.* 10:7. doi: 10.1186/1471-2091-10-7
- Wilfling, F., Thiam, A. R., Olarte, M. J., Wang, J., Beck, R., Gould, T. J., et al. (2014). Arf1/COPI machinery acts directly on lipid droplets and enables their connection to the ER for protein targeting. *elife* 3:e01607. doi: 10.7554/ELIFE.01607
- Wolucka, B. A., and Van Montagu, M. (2003). GDP-mannose 3',5'-epimerase forms GDP-L-gulose, a putative intermediate for the de novo biosynthesis of vitamin C in plants. *J. Biol. Chem.* 278, 47483–47490. doi: 10.1074/JBC.M309135200
- Wright, Z. J., and Bartel, B. (2020). Peroxisomes form intralumenal vesicles with roles in fatty acid catabolism and protein compartmentalization in Arabidopsis. *Nat. Commun.* 11, 1–13. doi: 10.1038/s41467-020-20099-y
- Zanetti, F., Isbell, T. A., Gesch, R. W., Evangelista, R. L., Alexopoulou, E., Moser, B., et al. (2019). Turning a burden into an opportunity: pennycress (*Thlaspi arvense* L.) a new oilseed crop for biofuel production. *Biomass Bioenergy* 130:105354. doi: 10.1016/j.biombioe.2019.105354
- Zhang, X. C., and Hu, J. P. (2008). FISSION1A and FISSION1B proteins mediate the FISSION of peroxisomes and mitochondria in Arabidopsis. *Mol. Plant* 1, 1036–1047. doi: 10.1093/MP/SSN056
- Zhang, X., and Hu, J. (2009). Two small protein families, DYNAMIN-RELATED PROTEIN3 and FISSION1, are required for peroxisome fission in Arabidopsis. *Plant J.* 57, 146–159. doi: 10.1111/J.1365-313X.2008.03677.X
- Zhu, A., Ibrahim, J. G., and Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 35, 2084–2092. doi: 10.1093/bioinformatics/bty895

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Johnston, García Navarrete, Ortiz, Romsdahl, Guzha, Chapman, Grotewold and Alonso. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.