


RESEARCH ARTICLE

Open Access



Ribosome footprint profiling enables elucidating the systemic regulation of fatty acid accumulation in *Acer truncatum*

Qiuyue Ma¹, Yuxiao Wang², Shushun Li¹, Jing Wen¹, Lu Zhu¹, Kunyuan Yan¹, Yiming Du¹, Shuxian Li², Liping Yan³, Zhijun Xie⁴, Yunzhou Lyu^{5*}, Fei Shen^{6*}  and Qianzhong Li^{1*}

Abstract

Background The accumulation of fatty acids in plants covers a wide range of functions in plant physiology and thereby affects adaptations and characteristics of species. As the famous woody oilseed crop, *Acer truncatum* accumulates unsaturated fatty acids and could serve as the model to understand the regulation and trait formation in oil-accumulation crops. Here, we performed Ribosome footprint profiling combining with a multi-omics strategy towards vital time points during seed development, and finally constructed systematic profiling from transcription to proteomes. Additionally, we characterized the small open reading frames (ORFs) and revealed that the translational efficiencies of focused genes were highly influenced by their sequence features.

Results The comprehensive multi-omics analysis of lipid metabolism was conducted in *A. truncatum*. We applied the Ribo-seq and RNA-seq techniques, and the analyses of transcriptional and translational profiles of seeds collected at 85 and 115 DAF were compared. Key members of biosynthesis-related structural genes (*LACS*, *FAD2*, *FAD3*, and *KCS*) were characterized fully. More meaningfully, the regulators (*MYB*, *ABI*, *bZIP*, and *Dof*) were identified and revealed to affect lipid biosynthesis via post-translational regulations. The translational features results showed that translation efficiency tended to be lower for the genes with a translated uORF than for the genes with a non-translated uORF. They provide new insights into the global mechanisms underlying the developmental regulation of lipid metabolism.

Conclusions We performed Ribosome footprint profiling combining with a multi-omics strategy in *A. truncatum* seed development, which provides an example of the use of Ribosome footprint profiling in deciphering the complex regulation network and will be useful for elucidating the metabolism of *A. truncatum* seed oil and the regulatory mechanisms.

Keywords *Acer truncatum*, Ribosome profiling, Transcriptome

*Correspondence:

Yunzhou Lyu
yunzhoulv@163.com
Fei Shen
shenf1028@gmail.com
Qianzhong Li
qianzhongli@jaas.ac.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Woody plants and oil crops are important resources in China. Accelerating the development and production of woody plants and oil crops is important for ensuring food and oil security. *Acer truncatum* is a versatile oil-producing woody tree and is a valuable ornamental species [1, 2]. In 2011, seed oil was approved as a New Resource Food by the National Health and Family Planning Commission of the People's Republic of China, with important implications for various industries (e.g., food and medicine). *Acer truncatum* seed oil is composed mainly of triacylglycerols (TAGs), with approximately 90% unsaturated fatty acids, including oleic acid (C18:1), linoleic acid (C18:2), α -linolenic acid (C18:3), and nervonic acid (C24:1) [3–5]. The proportions of oleic acid and linoleic acid are higher than those in peanut (*Arachis hypogaea*), sunflower (*Helianthus annuus*), and rapeseed (*Brassica napus*) oils [6]. Additionally, nervonic acid (24:15, *cis*-15-tetracosenoic acid, n-9) is a rare fatty acid (FA) accounting for 5–6% of the *A. truncatum* seed oil, making it potentially useful for treating schizophrenia, psychosis, and attention deficit disorder [7, 8].

The complex biochemical reactions mediating lipid biosynthesis have been studied in many plants [4, 9–11], but the underlying regulatory mechanisms remain relatively uncharacterized. In plants, the general lipid biosynthesis pathway includes the following four steps: de novo FA synthesis in plastids, acyl elongation and editing, TAG assembly in the endoplasmic reticulum, and oil drop formation [9, 12]. Some studies demonstrated that ACCase and 3-ketoacyl CoA synthetase (KCS) are the key rate-determining enzymes controlling the synthesis of FAs and very long-chain monounsaturated fatty acids (VLCFAs), respectively [5, 13]. Triacylglycerols are stored in oil bodies surrounded by a lipid monolayer. Moreover, previous research indicated that TAG assembly is critical for modulating lipid accumulation in developing oil-seeds [14, 15]. The molecular basis of lipid biosynthesis and the underlying regulatory mechanism has been studied in terms of the associated genes (i.e., transcription) and proteins (e.g., enzymes) in different plants [13, 16, 17]. Changes in the abundance of the candidate enzymes involved in lipid biosynthesis are not correlated with changes in the expression of the cognate mRNAs [10]. Post-transcriptional regulation affects the production of enzymes catalyzing lipid biosynthesis-related reactions, but there is relatively little available information regarding the post-transcriptional control of lipid biosynthesis in *A. truncatum*.

Ribosome footprinting (Ribo-seq) is a recently developed high-throughput sequencing technique that can generate precise information about translation at the genome-wide level and a single-codon resolution [18, 19].

In a Ribo-seq experiment, ribosomes are immobilized and the lysate is treated with nucleases to obtain ribosome-protected mRNA fragments (ribosome footprints). Then the deep sequencing of the ribosome footprints allowed for comprehensive and precise mapping of the positions of translating ribosomes on transcripts, which could reveal the quantity and interest positions of ribosomes [20, 21]. Thus, it has been used to clarify the translational regulation of specific tissues in diverse species, including *Arabidopsis thaliana* [22], maize [23], and humans [24]. Additionally, global three-nucleotide periodicity and translational efficiency (TE) revealed by ribosome profiling also provide valuable information for identifying novel open reading frames (ORFs) and ribosomal pause sites as well as for determining elongation [25].

Results

Establishment of an experimental and data analysis pipeline for the *A. truncatum* translato

In a previous study, we found that the oil proportions rapidly increased from 85 to 115 DAF, while the content variation of the components from 115 to 180 DAF was not significant [5]. In this study, the electron microscopy images of oil bodies were also conducted in seed cells at 85 and 115 DAF, the results showed that no lipid bodies were found at 85 DAF, and the lipid bodies filled the endosperm cells in 115 DAF (Fig. 1C). Therefore, the period between 85 and 115 DAF is critical for *A. truncatum* oil production. To investigate genome-wide gene expression related to the regulation of seed development, the ribosomal profiles of the seeds collected at these two time points were analyzed based on Ribo-seq, RNA-seq, and proteomic data. The experimental strategy is presented in Fig. 1A–C. The DEGs of 115 DAF/85 DAF were identified at transcriptional, translational, and protein levels. Then we performed the KEGG analysis to identify the lipid-enriched pathways among the up-regulated key DEGs, to know more the molecular basis of *A. truncatum* seed oil production. Finally, we combined Ribo-seq and RNA-seq techniques, enabled the identification of ORFs in *A. truncatum* seeds, and revealed their translational features. For the seeds collected at 85 and 115 DAF, we obtained 569.3–667.5 million and 614.3–706.6 million clean reads from RNA-seq libraries, respectively. Meanwhile, the 5.8–7.1 million and 5.7–6.5 million clean reads were generated from the Ribo-seq libraries, respectively (Table S2). A total of 31,738 peptides were obtained and 7,039 proteins were identified during the proteomic analysis. Details (e.g., identities and quantities) regarding the transcriptome, translato

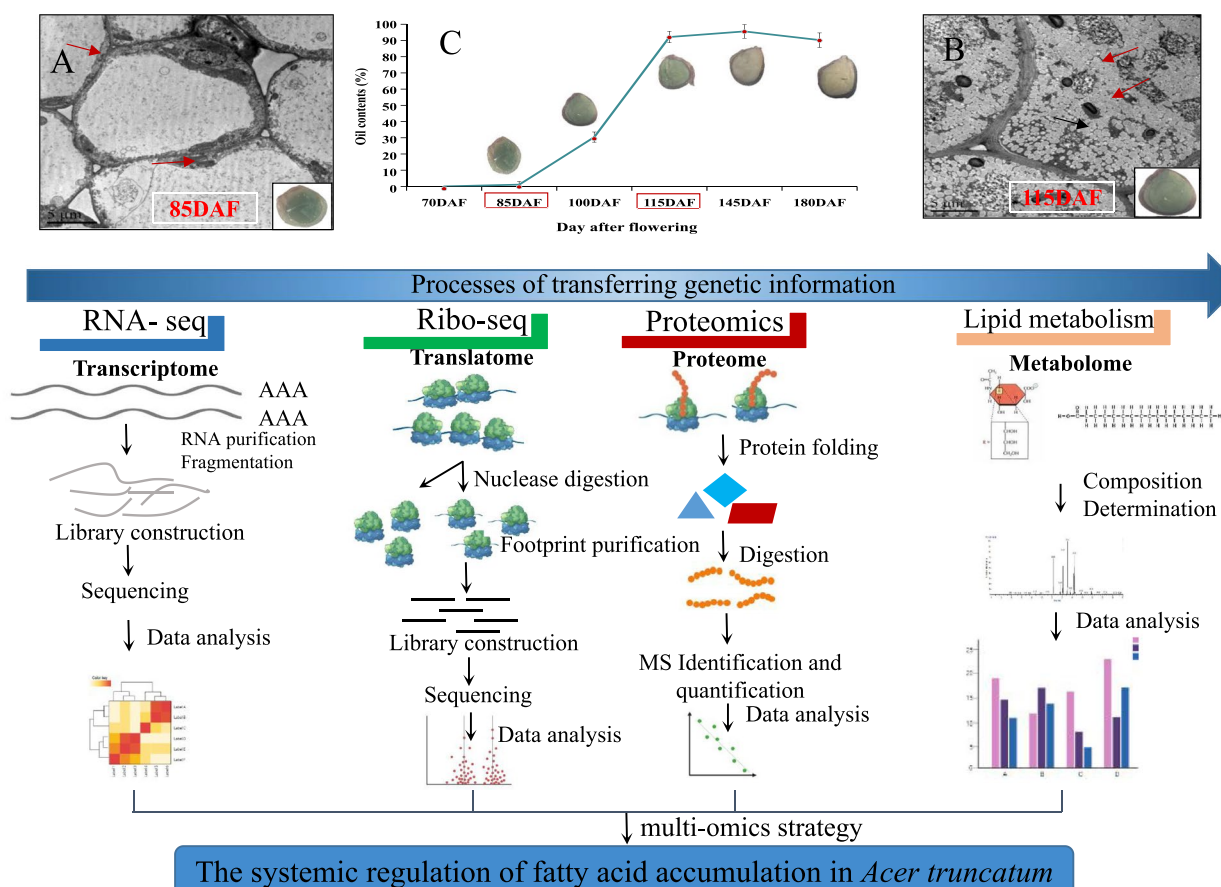


Fig. 1 Overview of the experimental design. **A, B** Electron microscopy images of oil bodies in seed cells at 85 and 115 DAF are presented. The red arrow represents the oil body. **C** The seed development of *A. truncatum*

clearly distinguished the seeds collected at 85 DAF from the seeds collected at 115 DAF (Fig. S1). In the present study, the seeds collected at 115 DAF contained 2.65 g/100 g oleic acid, 3.25 g/100 g linoleic acid, and 0.69 g/100 g nervonic acid, which were respectively 3.04-fold, 4.58-fold higher, and 13.8-fold than the corresponding contents in the seeds collected at 85 DAF (Table S6 and Fig. S2). These results reflected the reproducibility and the clear differences between the seeds collected at 85 and 115 DAF at the transcription, translation, and protein levels.

Features of ribosomal footprints

High-quality ribosomal footprints are critical for detecting ORFs. To explore whether the ribosomal footprint features change as *A. truncatum* seeds develop, the basic ribosomal profiles of the RPFs were compared between the seeds collected at 85 and 115 DAF. Our findings were consistent with the reported RPF length distributions in *A. thaliana* and maize [22, 26]. The predominant RPF length was 32–33 nt for the *A. truncatum* seeds

(Fig. 2). In our study, the three-nucleotide periodicity was detected around the start and stop codon regions of the RPFs with differing read lengths (Fig. 3). Moreover, 86.94% and 88.55% of the RPFs for the seeds collected at 85 and 115 DAF, respectively, were located in the coding sequence (CDS) regions, whereas less than 6% of the RPFs were detected in the other regions (Fig. 3). Similar results were reported for other plant species [26–28]. Previous studies confirmed the importance of 5' and 3' UTRs for the post-transcriptional regulation of gene expression [29, 30]. Additionally, the overall ribosome binding profiles were similar; the robustness of the three-nucleotide periodicity was quantified based on the percentage of reads in the expected reading frame. Considered together, these results indicated high-quality Ribo-seq data were obtained in this study.

Multi-omics-based functional comparison and analysis of the seeds at critical time points

To further characterize the transcription, translation, and protein level differences between *A. truncatum*

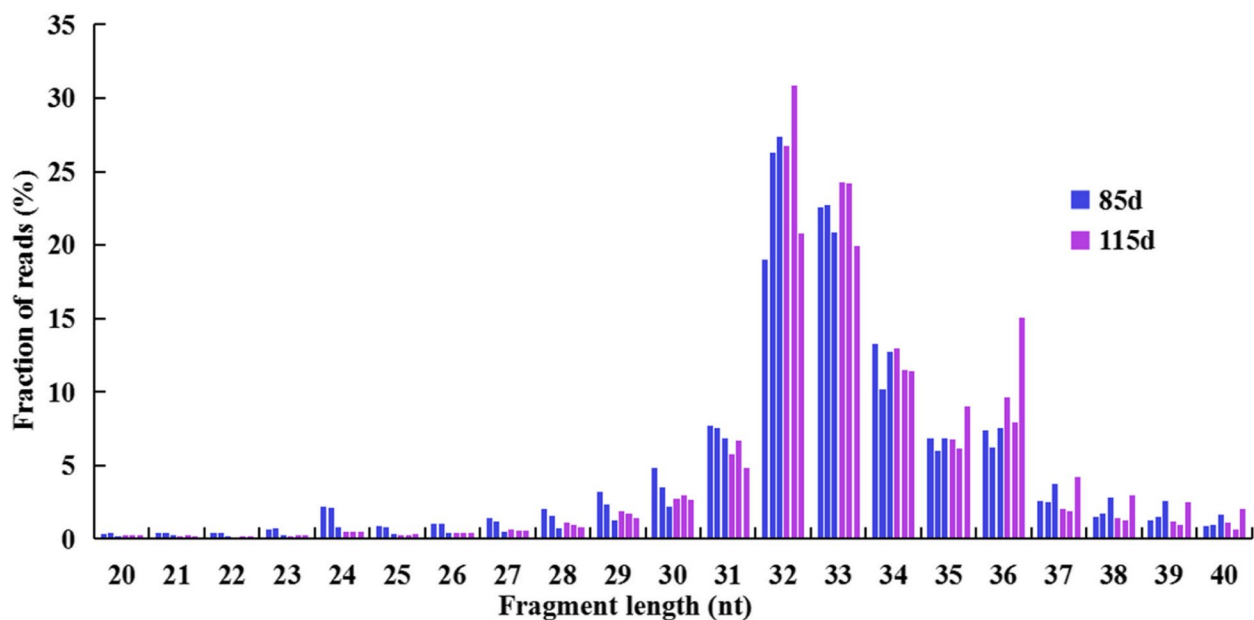


Fig. 2 Characteristics of ribosome profiling data in 85 and 115 DAF seeds. Length distribution of RPFs in 85 and 115 DAF seeds. The same color bars refer to three biological replicates

seeds at different stages of development, we applied a multi-omics approach. A differential expression analysis was conducted to characterize the changes in global gene expression in the transcription, translation, and protein levels, respectively. A total of 1543 up-regulated and 3013 down-regulated DEGs (transcription level), 320 up-regulated and 411 down-regulated DEGs (translation level), and 953 up-regulated and 1720 down-regulated DRPs (protein level) in 115 DAF vs. 85 DAF. (Fig. 4A). To clarify the overall tendencies in the gene expression changes, the scatter plots of the fold-changes in the transcription and translation levels (Fig. 4B) and the transcription and protein levels (Fig. 4C) were constructed by the comparison of the seeds collected at 115 DAF vs. 85 DAF. We subdivided all of the DEGs into nine categories ($|\log_2[\text{fold-change in FPKM}]| \geq 1$ and $\text{FDR} < 0.05$). Quadrant E included the genes with non-significant changes in expression at the transcription and translation levels (79.07%) as well as at the transcript and protein levels (49.76%). Quadrant C comprised the genes that were up-regulated at the transcription and translation levels (2.97%) as well as at the transcript and protein levels (3.34%). Details regarding the genes in different quadrants are listed in Tables S7–8. To elucidate the biological functions of the DEGs between the seeds collected at the two time points, we also performed a KEGG analysis to identify the enriched pathways among the genes. The genes that were translated at higher levels in the seeds collected at 115 DAF than in the seeds collected at 85

DAF were in quadrant C (Fig. 4D). The examination of the transcription and translation levels revealed 20 DEGs associated with FA metabolism and three DEGs associated with suberine and wax biosynthesis. The analysis of the transcript and protein levels detected seven and three DEGs associated with FA metabolism and suberine and wax biosynthesis, respectively. These DEGs included *LACS* and *KCS* genes (Tables S7–8). This was consistent with the observation that the FA proportions rapidly increased from 85 to 115 DAF. In addition, we also found 151 up-regulated and 362 down-regulated DEGs about translational efficiency in 115 DAF vs. 85 DAF ($\text{FDR} < 0.05$ and absolute fold-change ≥ 1.5) (Fig. S3A). The KEGG analysis results showed that six differential translational efficiency genes enriched in lipid metabolism (Fig. S3B and Table S9). We speculated that they might play an important role in the transcription and translation of lipid biosynthesis. Our findings regarding the FA metabolism at these two-time points may provide researchers with important insights into the accumulation of oil and the mechanism regulating FA synthesis in *A. truncatum* seeds.

Insights into lipid biosynthesis derived from a multi-omics analysis

The morphological characteristics of developing seeds at five time points are presented in Fig. 1C. From 85 to 115 DAF, the lipid bodies gradually increased in size and eventually filled the endosperm cells (Fig. 1A and B).

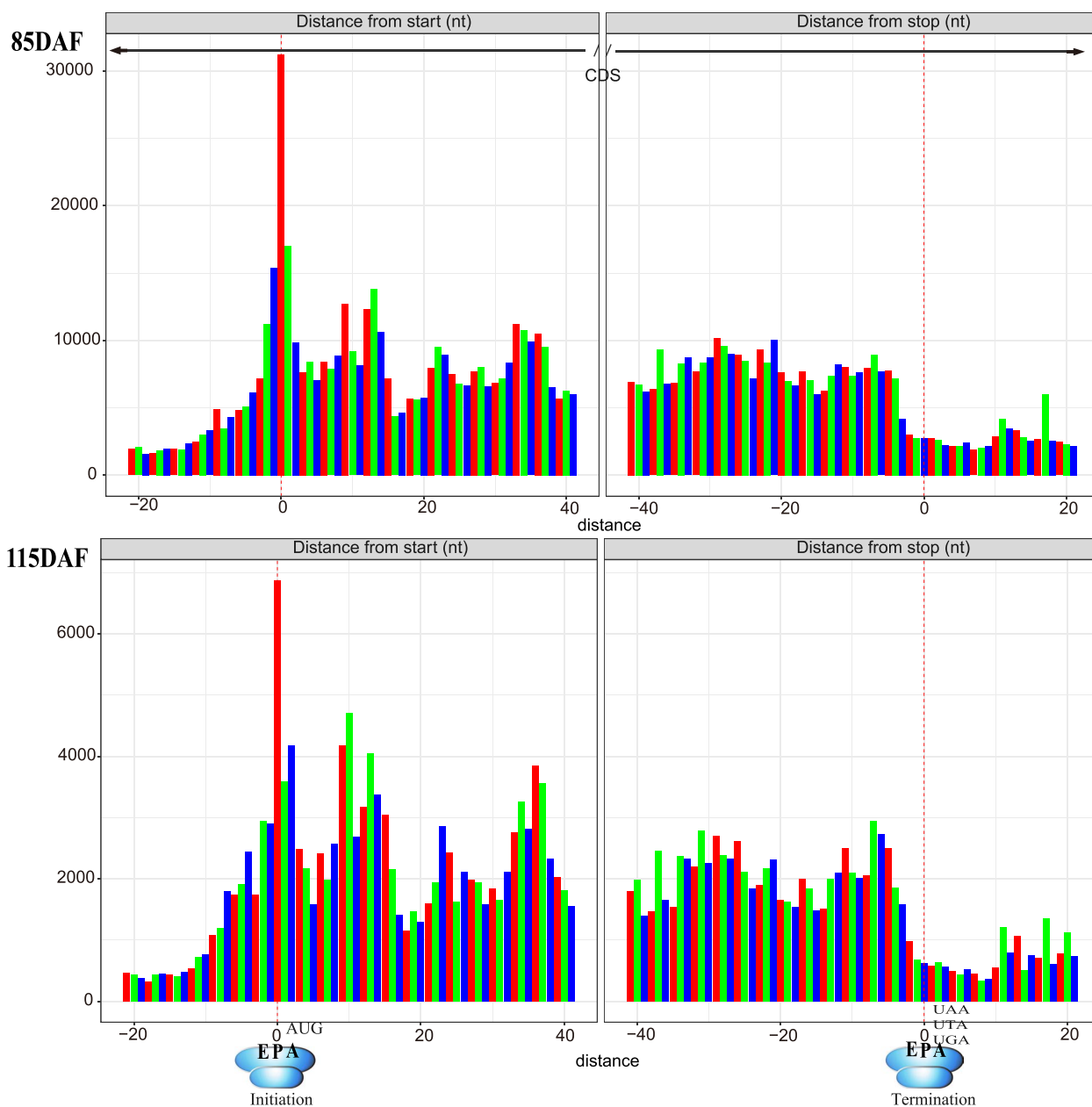


Fig. 3 The total number of RPFs along the start and stop codon regions in 85 and 115 DAF of *A. truncatum* seeds. The red, green, and blue bars represent the first, second, and third reading frames, respectively

Moreover, the analysis of the oil content and composition detected the following eight dominant components: palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), eicosenoic acid (C20:1), erucic acid (C22:1), and nervonic acid (C24:1) (Table S6 and Fig. 1). The oil content was much higher in the seeds collected at 115 DAF than in the seeds collected at 85 DAF. Furthermore, individual FAs

were biosynthesized and accumulated at different rates during the development of *A. truncatum* seeds.

To further investigate the major lipid biosynthesis pathway in the critical period for oil accumulation in *A. truncatum* seeds, we combined lipid biosynthesis-related pathways with heatmaps for the mRNA, RPF, and protein levels and constructed the oil biosynthesis pathway (Fig. 5A). The results of the KEGG pathway analysis indicated 25 genes identified in our multi-omics

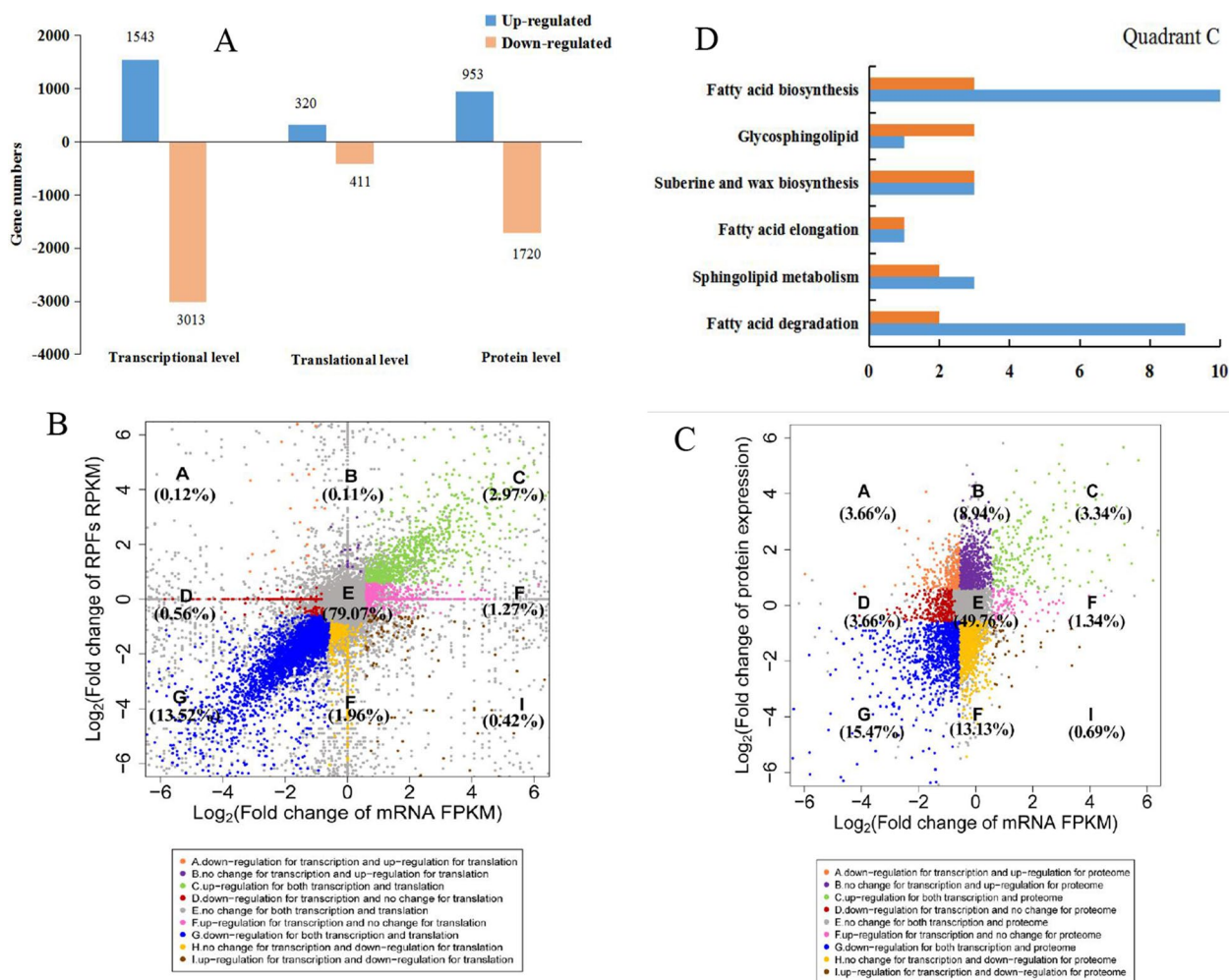


Fig. 4 Transcriptional, translation, and protein regulation of 85 and 115 DAF seeds in *A. truncatum*. **A** The number of DEGs at transcriptional or translational levels, the blue and orange bars refer to the number of up-regulation and down-regulated genes, respectively. **B** Scatter plot of the fold change of 115 DAF/85 DAF group at transcriptional and translational levels. **C** Scatter plot of the fold change of 115 DAF/85 DAF group at transcriptional and protein levels. **D** KEGG enrichment analysis of lipid genes in quadrant C, up-regulation for transcriptional and translational (orange column), up-regulation for both transcriptional and protein (blue column)

analysis encode key enzymes involved in FA biosynthesis; their expression levels were determined for the seeds collected at 85 and 115 DAF. Ten genes were selected for qRT-PCR analysis to validate the RNA-seq results. The

qRT-PCR data for the analyzed genes were generally consistent with the corresponding RNA-seq data (Fig. 5B). In plants, FA production involves the de novo synthesis and elongation of FAs in plastids and the desaturation of FAs

(See figure on next page.)

Fig. 5 **A** Regulation on lipid biosynthesis-related pathways. The icons next to the key enzymes represent the relative expression levels of the transcripts at 85 and 115 DAF. The heat map is displayed according to $\log_{10}[\text{fragments per kilobase per million (FPKM)} + 1]$ values. PDHC, pyruvate dehydrogenase complex; ACCase, acetyl-CoA carboxylase; MCMT, malonyl-CoA-ACP transacylase; KASI/II/III, ketoacyl-ACP synthase I/II/III; KAR, ketoacyl-ACP reductase; HAD, hydroxyacyl-ACP dehydrase; EAR, enoyl-ACP reductase; SAD, stearyl-ACP desaturase; FAD2/6, oleate desaturase; FAD3/7/8, linoleate desaturase; FATA/B, acyl-ACP thioesterase A/B; PCH, palmitoyl-CoA hydrolase; LACS, long-chain acyl-CoA synthetase; LPCAT, lysophosphatidylcholine acyltransferase; GK, glycerol kinase; GPAT, glycerol-3-phosphate acyltransferase; LPAAT, 1-acylglycerol-3-phosphate acyltransferase; PAP, phosphatidic acid phosphatase; DGAT = diacylglycerol acyltransferase; PDAT, phospholipid:diacylglycerol acyltransferase. **B** The qRT-PCR validation of oil biosynthesis genes at two-time points of *A. truncatum*. The relative expression levels of genes were normalized with the internal reference gene Actin (Wang et al., 2018). Three biological replicates (nested with three technical replicates) are represented by the error bars ($P < 0.05$). **C** TF-oil biosynthesis network. Co-expressed network of the correlation among transcription factors and oil biosynthesis genes (*FAD2*, *FAD3*, and *KCS*). The gray lines indicate the correlations between the two genes

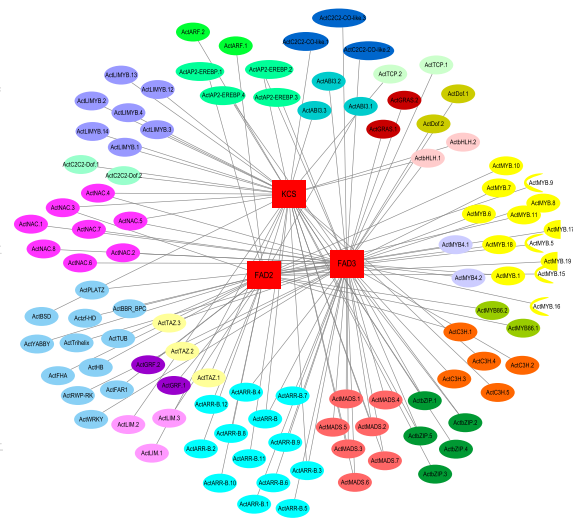
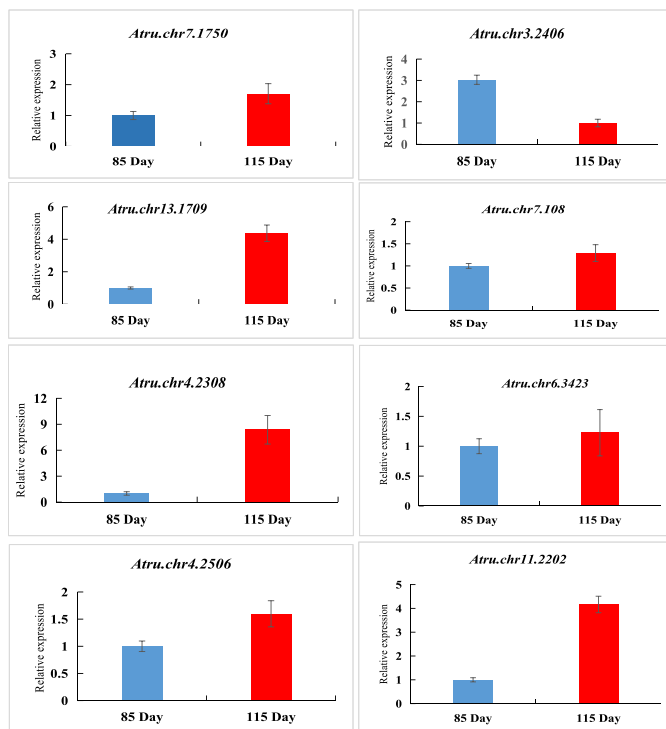
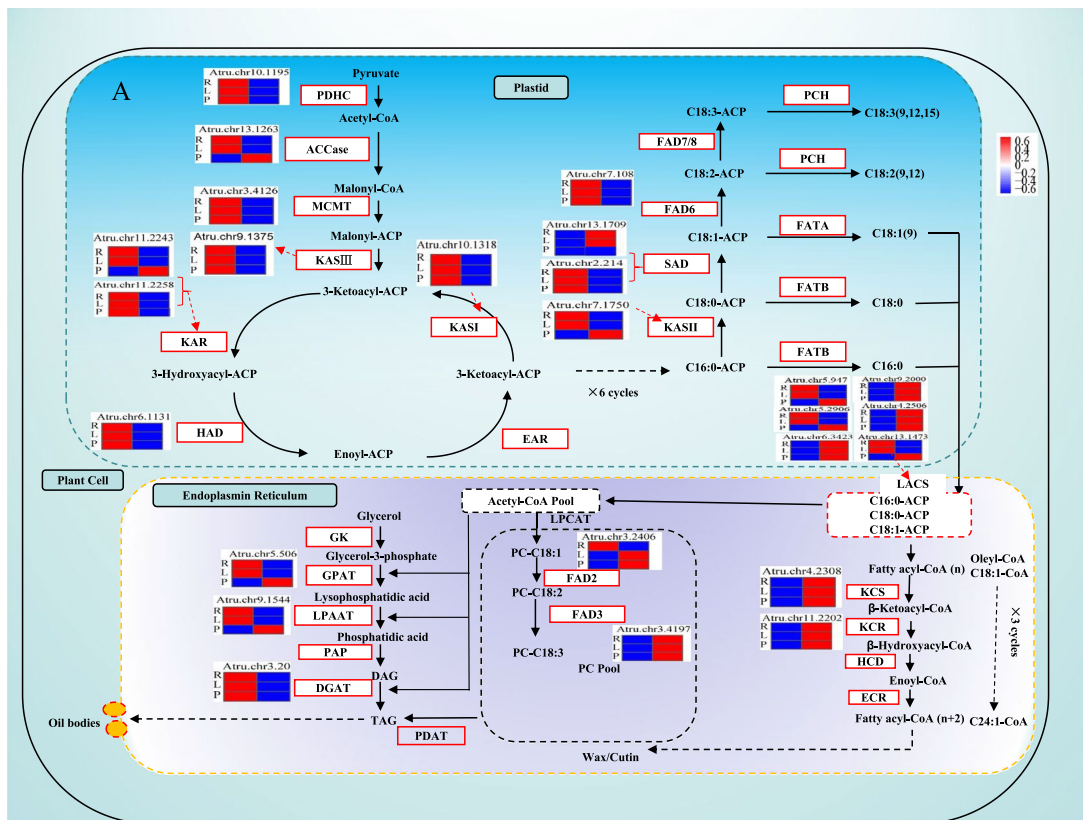


Fig. 5 (See legend on previous page.)

and TAG assembly in the endoplasmic reticulum [5, 31, 32]. However, there have been relatively few studies that examined FA synthesis based on a multi-omics approach (i.e., transcriptomic, translational, and proteomic). Fatty acids are synthesized by a series of reactions, after which LACS on the outer membrane of the plastid catalyzes the esterification of FAs to generate the acyl-CoA pool. In the current study, six homologous LACS genes were identified, among which *Atru.chr.9.2000*, *Atru.chr1.2506*, and *Atru.chr6.342* were highly expressed at the transcription, translation, and protein levels. The C18:1 and C18:2 FAs accounted for the largest proportion of the FAs in *A. truncatum* seeds. Previous studies demonstrated that the accumulation of C18:1 and C18:2 FAs is correlated with the expression of *FAD2* and *FAD3*, respectively [5, 33]. In addition, *FAD2* (*Atru.chr3.2406*) and *FAD3* (*Atru.chr3.4197*) were more highly expressed in the seeds at 115 DAF than at 85 DAF. Similarly, *KCS* is a rate-limiting enzyme mediating FA elongation, which is important for nervonic acid biosynthesis [2]. The data generated in the present study indicated the nervonic acid content in the seeds was 13.8-fold higher at 115 DAF than at 85 DAF. We previously identified 28 genes in the *A. truncatum* *KCS* gene family [2]. However, only *Atru.chr4.2308* was differentially expressed between the seeds at 85 and 115 DAF at the transcription, translation, and protein levels in this study. Accordingly, this gene may be important for nervonic acid biosynthesis. To know more the overall transcription and translation features of *FAD2* (*Atru.chr3.2406*), *FAD3* (*Atru.chr3.4197*), and *KCS* (*Atru.chr4.2308*). Figure S4 showed the three genes of RNA-seq and Ribo-seq coverage tracks by using Integrative Genomic Viewer (IGV). The results showed that the change tendency of mRNA and RPFs was similar in the seeds at 85 DAF and at 115 DAF, respectively. While the abundance of the seeds at 115 DAF was higher than that of 85 DAF. This may indicate that they are efficient in transcription and translation levels.

Transcription factors (TFs) have important roles related to post-transcriptional regulation. The regulatory effects of several TFs on lipid biosynthesis have been reported, including WRINKLED1 (WRI1) [34], leafy cotyledon1 (LEC1), fusca3 (FUS3), abscisic acid insensitive 3 (ABI3), v-myb avian myeloblastosis viral oncogene homolog (MYB), and DNA-binding with one finger (Dof) [5, 35–37]. However, the transcriptional regulation of lipid biosynthesis in *A. truncatum* seeds has not been comprehensively elucidated. In this study, 1,510 TFs were identified during the examination of the seeds collected at two time points (Table S10). *A. truncatum* is a versatile oil-producing woody tree and a rich source of oleic acid (C18:1), linoleic acid (C18:2), and nervonic acid (C24:1) [3–5]. To identify the TFs that

directly regulate the expression of the oil biosynthesis pathway genes *FAD2* (*Atru.chr3.2406*), *FAD3* (*Atru.chr3.4197*), and *KCS* (*Atru.chr4.2308*), we completed a co-expression analysis to identify TF genes with expression patterns that were significantly and highly correlated ($r > 0.7$, $P < 0.05$) with the expression patterns of the three genes in the oil biosynthesis pathway (Fig. 5C). We detected 31, 51, and 37 TF genes that were co-expressed with *FAD2* (*Atru.chr3.2406*), *FAD3* (*Atru.chr3.4197*), and *KCS* (*Atru.chr4.2308*), respectively (Table S11). Additionally, the known MYB and ABI TFs affected the expression of *FAD2* (*Atru.chr3.2406*), whereas the bZIP, MYB, and Dof TFs modulated the *FAD3* (*Atru.chr3.4197*) and *KCS* (*Atru.chr4.2308*) expression levels. These findings may provide important clues regarding the molecular basis of lipid biosynthesis in *A. truncatum* seeds.

Identification and analysis of small ORFs in UTRs

Untranslated regions are key mediators of post-transcriptional regulation. Moreover, CDS translation is modulated by the combined effects of uORFs, dORFs, the secondary structure around uORFs, and the distance between the uORF and CDS [36]. Earlier studies suggested that uORFs affect TE [38, 39]. However, it remains unknown whether the annotated uORFs in *A. truncatum* can influence gene transcription or translation. The ribosomal profile data generated in the current study will enable the identification and analysis of uORFs in *A. truncatum*. Information regarding the uORFs, dORFs, and mORFs identified by our analysis of *A. truncatum* seeds is provided in Table S12. The uORFs and dORFs were distinguished from the known protein-coding genes according to RRS and ORFscore data (Fig. S5A and B). A total of 2,151 uORFs were identified, which accounted for 9.67% of the transcripts.

Figure 6 presents the features of the uORFs detected in the seeds collected at 85 DAF. Most of the identified uORFs consisted of 60–460 nt, with 80% of the uORFs shorter than 160 nt (Fig. 6A). This observation was following the uORF length distribution in *A. thaliana*, tea, rice, and other higher plants [38, 40, 41]. Furthermore, we observed that the relative distance from uORF to the start codon of the mORFs (P -value = $8.22e - 6$) in translated uORFs was shorter than non-translated uORFs (Fig. 6B). Of the genes with a relatively high TE (> 0.5), the cumulative curve showed that TE tended to be lower for the genes with a translated uORF than for the genes with a non-translated uORF. In contrast, there were no significant differences between the genes with a single translated uORF and the genes with multiple translated uORFs (Fig. 6C). Moreover, Kozak consensus sequence with “GCCA/GCCAUGG,” especially the position of $-3(A/G)$ and $4(G)$ around AUG start codon

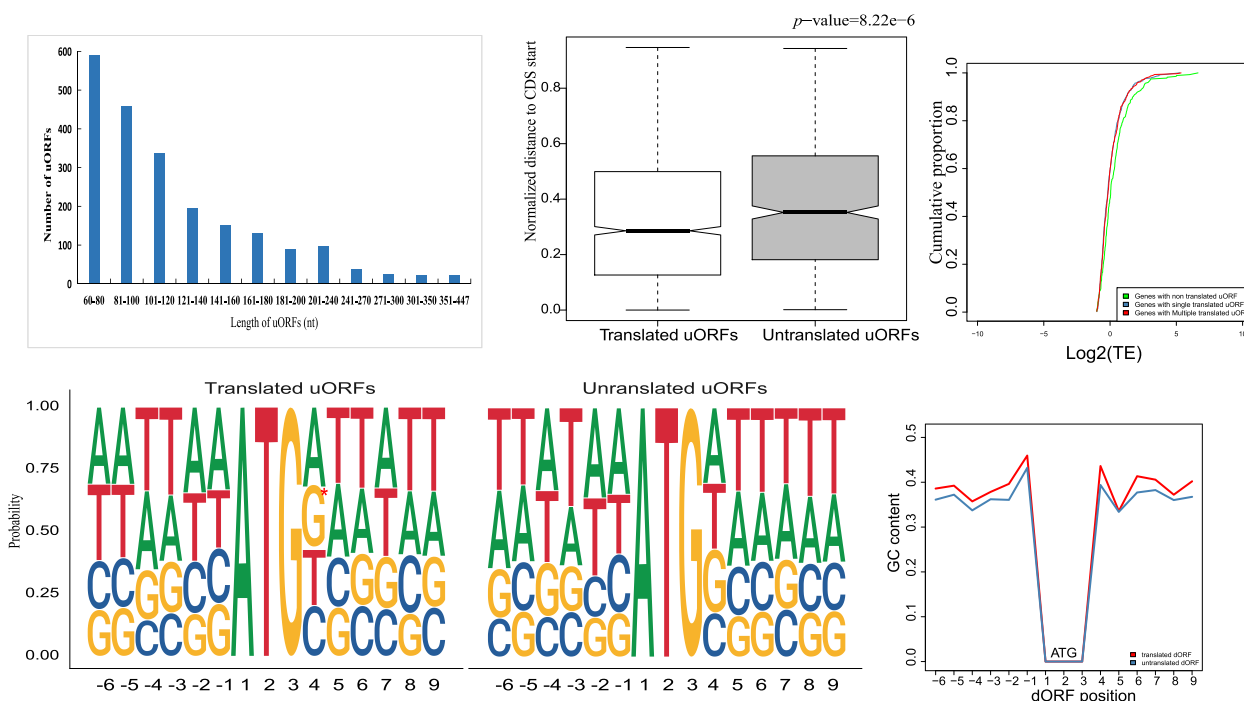


Fig. 6 Features of uORFs in 85DAF *A. truncatum* seeds. **A** Length distribution of uORFs. **B** Normalized distance from uORFs to CDS ATG start codon (85DAF). **C** The relationship between the length of uORF and TE (85DAF). **D** Sequence composition between translated and untranslated uORFs around the ATG start codon. Chi-squared test was used for significance analysis. The asterisks indicate P -value < 0.05 (85DAF). **E** GC content of the flanking sequence of translated (red line) and untranslated (blue line) uORFs' ATG start codons, respectively

is important for start codon recognition as well as translation initiation [42]. We first performed motif analysis around the start codon of mORFs, the frequency of each base of the flanking sequence around the uORF start codon between translated and untranslated was analyzed, and a high GC content of the flanking sequence in translated uORFs was observed (Fig. 6D and E). Furthermore, the probabilities of guanine at the position of +4 (P -value=0.02) in translated uORFs were found to be higher than in untranslated ones (Fig. 6D). Similar results were obtained for the seeds collected at 115 DAF (Fig. S6).

Discussion

A. truncatum is an important woody tree species that produce seeds with large amounts of unsaturated fatty acids [2]. Thus, the molecular mechanism underlying lipid synthesis in the key seed development stages must be characterized. Previous research determined that 85 and 115 DAF are crucial time points for FA synthesis during the *A. truncatum* seed development stage. Our comparative analyses of transcriptional and translational profiles of seeds collected at 85 and 115 DAF provide new insights into the global mechanisms underlying the developmental regulation of lipid metabolism.

Based on our analyses, we identified the expressed uORFs in *A. truncatum*. Translational efficiency is influenced by the presence of uORFs that repress translation, which is consistent with our finding that genes with translated uORFs have a lower TE than genes with non-translated uORFs [43]. Some researchers have speculated that long uORFs may compete for translation initiation factors, and the distance between expressed uORFs and the CDS affects TE in mammalian and yeast cells [43]. Wu et al. [11] reported that the uORF length is not highly correlated with TE. Our results suggested that the sequence flanking ATG was perfectly accorded with the Kozak sequence in mORFs, the repressive effects of uORFs on translation may be associated with their unique features in different species.

In oilseed plants, oil synthesis mainly involves de novo FA biosynthesis, TAG assembly, and oil body formation [4]. Although the genes in the main FA biosynthesis pathways have been identified [2, 5], the number of key genes and their expression patterns in developing seeds vary among species. In the current study, the oil synthesis pathway was constructed by combining the results of a multi-omics investigation. We also examined the key genes at the translational and proteomic levels. At the transcription, translation, and protein levels, *LACS*

(*Atru.chr.9.2000*, *Atru.chr.1.2506*, *Atru.chr.6.342*), *FAD2* (*Atru.chr.3.2406*), *FAD3* (*Atru.chr.3.4197*), and *KCS* (*Atru.chr.4.2308*) expression levels were highly correlated with FA contents. Hence, the expression of these genes, which encode key enzymes in the lipid biosynthesis pathway, is potentially regulated by post-translational modifications. Additionally, our findings are suggestive of the post-translational regulation of lipid biosynthesis by the identified TFs (e.g., MYB, ABI, bZIP, and Dof). Notably, lipid biosynthesis in plants is influenced by various factors, including the cultivar, tissue, growth stage, and environmental stimuli. In this study, we combined our analyses of phenotypic changes, differential expression, and co-expression with an examination of TE to reveal the molecular basis of *A. truncatum* seed oil production. However, future studies will need to precisely decipher the associated post-translational regulation in *A. truncatum*.

Conclusions

We combined Ribo-seq and RNA-seq techniques, enabled the identification of ORFs in *A. truncatum* seeds and revealed their translational features. To the best of our knowledge, this is the first comprehensive multi-omics analysis of lipid metabolism.

We compared the analyses of transcriptional and translational profiles of seeds collected at 85 and 115 DAF. And the key genes and TFs were examined at the translational and proteomic levels. The features of uORFs in *A. truncatum* seeds also provide new insights into the global mechanisms underlying the developmental regulation of lipid metabolism. The results will serve as an important foundation for future explorations of the mechanisms regulating lipid accumulation in *A. truncatum* and oil-seed crops.

Methods

Plant materials and sampling

A. truncatum plants were grown in the Jiangsu Province Aceraceae Germplasm Repository (Lishui district, Nanjing, China; latitude: 31°65'N, longitude: 119°02'E) under natural conditions (each has three biological replicates). Seeds were collected at 85 and 115 DAF from three individuals for the RNA-seq and Ribo-seq analyses, oil content measurements, and protein extraction and digestion. The collected samples were immediately frozen in liquid nitrogen and stored at −80 °C until analyzed.

Lipid analysis and oil body examination

Fifty seeds collected from three plants (each has three biological replicates) at 85 and 115 DAF were used for extracting oil and FA methyl esters as described by Gao

et al. [44] and Ma et al. [5]. The FAs were analyzed using the TRACE DSQ gas chromatography–mass spectrometry system (Thermo Fisher Scientific Inc., Waltham, MA, USA) as described by Zhang et al. [45]. The embryos were cut in the middle and then immediately fixed in a solution comprising 4% (v/v) glutaraldehyde and sodium phosphate buffer (pH 7.2). They were subsequently post-fixed in osmium tetroxide, dehydrated, and embedded in Spurr's resin. The sections were examined using the JEM-1400 transmission electron microscope (60 kV) (JEOL, Tokyo, Japan).

RNA extraction, library construction, and sequencing

TRIzol reagent (Invitrogen, NJ, USA) was used to extract total RNA from the developing seeds collected from three plants at 85 and 115 DAF. The RNA-seq libraries were constructed using the NEBNext Ultra Directional RNA Library Prep Kit (New England BioLabs, Ltd., USA). Six cDNA libraries were sequenced on the NovaSeq 6000 platform (Illumina, San Diego, CA). After removing duplicate reads, adapters, and ambiguous sequences (i.e., reads with an N rate exceeding 10%), the remaining clean reads were mapped to our *A. truncatum* reference genome (the Sequence Read Archive (SRA) under study accession number SUB6287730) [2]. using HISAT2 (version 2.0.4) (parameters: -rna-strandness RF) [46]. Differentially expressed genes (DEGs) were analyzed using the DESeq2 software (FDR < 0.05 and absolute fold-change ≥ 1.5) [47]. Significant DEGs were identified based on the following criteria: false discovery rate (FDR) < 0.05 and |log₂fold-change| ≥ 1.5. The Gene Ontology (GO) (<http://www.geneontology.org/>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>) databases were used to functionally annotate genes and identify enriched metabolic pathways (R project package Rscript.3.6.0), respectively. To search for transcription factors (TFs) among DEGs, the data were screened against TF database PlantTFDB (<http://planttfdb.gao-lab.org/>).

Ribosome footprints (RFs) extraction and Ribo-seq library construction

The RFs were extracted as below: the seeds frozen in liquid nitrogen were treated with 400 μL lysis buffer for 10 min. The lysate was centrifuged at 20,000 × g for 10 min at 4 °C, after which the supernatant was collected. To prepare the RFs, 10 μL RNase I (NEB, Ipswich, MA, USA) and 6 μL DNase I (New England BioLabs, Ltd., USA) were added to 400 μL lysate supernatant, which was incubated for 45 min at room temperature with gentle mixing on a Nutator mixer. The RNases in the samples were digested with 10 μL SUPERase in RNase Inhibitor (Ambion, Austin, TX, USA). The digested

Ribosome Protected Fragments (RPFs) were loaded onto the Illustra MicroSpin S-400 HR column and centrifuged at $600 \times g$ for 3 min. Next, 10 μ L 10% (w/v) SDS was added to the eluted samples, then RPFs longer than 17 nt were isolated using the RNA Clean and Concentrator-25 kit (Zymo Research). To eliminate rRNA, antisense DNA probes (50–80 bases) complementary to rRNA sequences were added to the RPFs, after which RNase H (New England BioLabs, Ltd., USA) and DNase I (New England BioLabs, Ltd., USA) were added to digest the rRNA and residual DNA probes. Finally, RPFs were further purified using magnetic beads (Vazyme, Nanjing, Jiangsu, China) [48]. Additionally, Ribo-seq libraries were constructed using the NEBNext[®] Multiple Small RNA Library Prep Set for Illumina[®] [49] and then sequenced on the Illumina HiSeq[™] X10 system by Gene Denovo Biotechnology Co. (Guangzhou, China).

Ribosome profiling raw data processing and analysis

The raw data were filtered to remove reads containing adapters, poly-N sequences, and more than 50% low-quality bases using an in-house Perl script. After mapping the retained reads to the sequences in the ribosomal RNA (rRNA), GenBank, and Rfam databases using Bowtie2 (version 2.2.8) (parameters: -p 4 -k 10 -no-unal -phred64), the reads mapped to rRNAs, transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and miRNAs were eliminated [50]. Next, the processed reads were mapped to our *A. truncatum* reference genome (SUB6287730) [2] using STAR (version 3.1.0) (ThreadN 4 -runMode alignReads -readFilesCommand zcat -twopassMode Basic), with no mismatches allowed. Read counts in the ORFs of protein-coding genes were calculated using the RSEM software (version 1.2.19) (est_method RSEM -aln_method bowtie2 -seqType fq -thread_count 4 -seed-length 10) and the gene expression level was normalized by using FPKM (fragments per kilobase of transcript per million mapped reads) values. The DEGs (FDR < 0.05 and absolute fold-change ≥ 1.5) were functionally annotated and enriched metabolic pathways were identified according to the GO and KEGG databases, respectively. The three-nucleotide periodicity was plotted using the riboWaltz R package [50], whereas the potential ribosomal pause sites (single-codon resolution) were investigated using PausePred (<https://pausepred.ucc.ie/>).

Identification and quantitative analysis of ORFs

A custom ORFfinder search was conducted using transcript sequences annotated as non-coding regions, including 5' and 3' UTRs. Noncanonical ORFs, including uORFs and downstream ORFs (dORFs), were extracted using start codons (e.g., ATG). The raw reads containing

adapters as well as reads containing poly-N sequences and over 50% low-quality bases were removed using an in-house Perl script. The reads mapped to rRNA and tRNA database sequences were eliminated using Bowtie2. The training reads were mapped to our *A. truncatum* reference genome (SUB6287730) [2] using Bowtie2, with no mismatches allowed. The reads mapped to snRNAs, snoRNAs, and microRNA precursor regions were not considered further. Finally, the main ORFs (mORFs) of the protein-coding genes as well as the uORFs and dORFs were used to analyze the translation level. The DESeq2 package (<http://www.r-project.org/>) was used for identifying differentially translated ORFs (FDR < 0.05 and absolute fold-change ≥ 1.5). The ribosome release score (RRS) and ORFscore were used to evaluate the coding potential of non-canonical ORFs [51]. The 95th percentiles of the RRS and ORFscore data as well as the Fickett and hexamer scores calculated using the CPAT tool [52] were used to assess the coding potential of different ORFs.

DIA-based protein quantification

Proteins were extracted from seeds and digested as described by Wu et al. [11]. Briefly, 2 mL lysis buffer (Roche Ltd., Basel, Switzerland) was added to the seeds frozen in liquid nitrogen. The solution was sonicated (on ice for 30 min) and then centrifuged (13,000 rpm for 30 min at 4 °C). The proteins in the supernatant were precipitated with ice-cold acetone during an overnight incubation at -20 °C. Protein quality was evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The extracted protein (50 μ g) was suspended and incubated. The obtained proteins were digested with modified trypsin (Promega Co, Ltd., Madison, USA) at a substrate: enzyme ratio of 50:1 (w/w). The peptide mixture was dissolved and fractionated using the Ultimate 3000 system (Thermo Fisher Scientific, MA, USA). The peptides were redissolved and analyzed by LC-MS/MS. Specifically, the Orbitrap Fusion Lumos mass spectrometer coupled to the EASY-nLC 1200 system (Thermo Fisher Scientific) was used. The raw DIA data were processed and quantified using the default parameters of Spectronaut Pulsar 11.0 (Biognosys AG, Zurich, Switzerland). The differentially regulated proteins (DRPs) were filtered (fold-change ≥ 1.5 or < 0.67; $P < 0.05$) [53]. The identified proteins were annotated using the GO and KEGG databases.

Quantitative real-time PCR (qRT-PCR)

The expression levels of the DEGs related to FA biosynthesis were validated by qRT-PCR analysis. Specific

primers were designed using Primer5 (Table S1). Total RNA was extracted as previously described. Next, 1 µg RNA served as the template for the synthesis of first-strand cDNA using the PrimeScript Reagent Kit (TaKaRa, Dalian, China). The cDNA templates were diluted 20-fold before use. The qRT-PCR analysis was performed using SYBR Premix ExTaq (TaKaRa Bio, Kusatsu, Japan) and the StepOne Plus Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific). *Actin* was selected as the internal reference gene [54]. Gene expression levels were analyzed according to the $2^{-\Delta\Delta Ct}$ method [55].

Abbreviations

| | |
|-------------|---|
| RPFs | Ribosome-protected fragments |
| uORFs | Upstream open reading frames |
| dORFs | Downstream open reading frames |
| 5'UTR | 5' upstream untranslated region |
| CDS | Protein-coding sequences |
| DEG | Differentially expressed genes |
| TE | Translational efficiency |
| PDHC | Pyruvate dehydrogenase complex |
| ACCase | Acetyl-CoA carboxylase |
| MCMT | Malonyl-CoA-ACP transacylase |
| KASI/II/III | Ketoacyl-ACP synthase I/II/III |
| KAR | Ketoacyl-ACP reductase |
| HAD | Hydroxyacyl-ACP dehydrase |
| EAR | Enoyl-ACP reductase |
| SAD | Stearoyl-ACP desaturase |
| FAD2/6 | Oleate desaturase |
| FAD3/7/8 | Linoleate desaturase |
| FATA/B | Acyl-ACP thioesterase A/B |
| PCH | Palmitoyl-CoA hydrolase |
| LACS | Long-chain acyl-CoA synthetase |
| LPCAT | Lysophosphatidylcholine acyltransferase |
| GK | Glycerol kinase |
| GPAT | Glycerol-3-phosphate acyltransferase |
| LPAAT | 1-Acylglycerol-3-phosphate acyltransferase |
| PAP | Phosphatidic acid phosphatase |
| DGAT | Diacylglycerol acyltransferase |
| PDAT | Phospholipid:diacylglycerol acyltransferase |
| WRI1 | WRINKLED1 |
| LEC1 | Leafy cotyledon1 |
| FUS3 | fusca3 |
| ABI3 | Abscisic acid insensitive 3 |
| MYB | v-myb avian myeloblastosis viral oncogene homolog |
| Dof | DNA-binding with one finger |

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-023-01564-8>.

Additional file 1: Fig. S1. Principal component analysis on the RNA-seq (A) and Ribo-seq (B) and proteomic profiles (C) of three replicates of 85 and 115 DAF seeds in *A. truncatum*. **Fig. S2.** The GC-MS analysis of 85 DAF and 115 DAF seeds in *A. truncatum*. **Fig. S3.** (A): DEGs of translational efficiency in 115 DAF vs. 85 DAF. (B): KEGG pathway enrichment analysis of DEGs of translational efficiency in 115 DAF vs. 85 DAF. **Fig. S4.** The RNA-seq and Ribo-seq coverage tracks of *FAD2* (*Atru.chr3.2406*), *FAD3* (*Atru.chr3.4197*) and *KCS* (*Atru.chr4.2308*) by using Integrative Genomic Viewer (IGV). **Fig. S5.** The scatterplots of \log_{10} (Ribosome release score) against ORF score of ORFs (A) and Fickett score against Hexamer score. (B) The dashed lines represent the 95th percentiles set as threshold values. **Fig. S6.** Features of uORFs in 115DAF *A. truncatum* seeds.

Additional file 2: Supplementary Table S1. Primers (5'-3') used in the quantitative real-time polymerase chain reaction (qRT-PCR) to validate the reliability of the RNA-seq results.

Additional file 3: Supplementary Table S2. Statistic on the RNA-seq data and ribosomal profiling data.

Additional file 4: Supplementary Table S3. The identification and quantification information of transcriptome.

Additional file 5: Supplementary Table S4. The identification and quantification information of translatoome.

Additional file 6: Supplementary Table S5. The identification and quantification information of proteome.

Additional file 7: Supplementary Table S6. The gene information in different quadrants at transcription and translation levels.

Additional file 8: Supplementary Table S7. The gene information in different quadrants at the transcript and protein levels.

Additional file 9: Supplementary Table S8. The contents of fatty acids in 85 DAF and 115 DAF *A. truncatum* seeds.

Additional file 10: Supplementary Table S9. The KEGG analysis of differential translational efficiency genes in 85 DAF and 115 DAF *A. truncatum* seeds.

Additional file 11: Supplementary Table S10. The information of identified TFs in 85 DAF and 115 DAF *A. truncatum* seeds.

Additional file 12: Supplementary Table S11. The information of identified TF genes were co-expressed with *FAD2* (*Atru.chr3.2406*), *FAD3* (*Atru.chr3.4197*), and *KCS* (*Atru.chr4.2308*).

Additional file 13: Supplementary Table S12. The information of uORFs and dORFs identified as potential coding sequences based on RRS and ORF score.

Acknowledgements

Thanks to all the members of the Institute of Leisure Agriculture team for their contributions. We thank Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac) for editing the English text of a draft of this manuscript.

Authors' contributions

QM, QL, FS, and Y Lyu designed the project and the strategy; SL1, JW, LZ, and KY were responsible for plant sample collection, DNA, RNA protein preparation, library construction, and sequencing; QM, YW, SL2, and ZX were responsible for software comparative analyses; LY, ZX, and YL performed investigation and software; QM, FS, QL, and YLyu wrote and revised the manuscript. All authors have read and approved the manuscript for publication.

Funding

This work was funded by the Natural Science Foundation of China (32001357), Natural Science Foundation of Jiangsu Province (BK20211139), Supported by the National Key Research and Development Program of China (2022YFD2200401), the Independent Innovation Fund Project of Agricultural Science and Technology in Jiangsu Province (CX (21)2032).

Availability of data and materials

The RNA-seq raw data (accession number: SUB12016538) and Ribo-seq raw data (accession number: SUB12025717) were uploaded to National Center for Biotechnology Information under Project No. PRJNA877423 [56]. The mass spectrometry proteomics data (Project No. IPX0005014000) was uploaded to the iProX integrated proteome resources.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing financial interests.

Author details

¹Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences, Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement Nanjing, Nanjing 210014, China. ²Nanjing Forestry University, Nanjing 210037, China. ³Shandong Academy of Forestry Sciences, Jinan 250014, China. ⁴Xiangyang Forestry Science and Technology Extension Station, Xiangyang 441000, China. ⁵Jiangsu Academy of Forestry, Nanjing 211153, China. ⁶Institute of Biology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100197, China.

Received: 18 November 2022 Accepted: 14 March 2023

Published online: 03 April 2023

References

- Ma X, Tian W, Wu L, Cao X, Ito Y. Isolation of quercetin-3-O-L-rhamnoside from *Acer truncatum* Bunge by high-speed counter-current chromatography. *J Chromatogr A*. 2005;1070(1–2):211–4. <https://doi.org/10.1016/j.chroma.2005.02.052>.
- Ma Q, Sun T, Li S, Wen J, Zhu L, Yin T, et al. The *Acer truncatum* genome provides insights into the nervonic acid biosynthesis. *Plant J*. 2020;104(3):662–78. <https://doi.org/10.1111/tpj.14954>.
- Wang X, Fan J, Wang S, Sun R. A new resource of nervonic acid from purpleblow maple (*Acer truncatum*) seed oil. *Forest Products J*. 2006;56(11–1):147–50. <https://doi.org/10.1007/s00226-006-0085-6>.
- Liu X, Fu H, Chen Y. Study on the physico-chemical properties and fatty acid composition of *Acer truncatum* Buge oil. *China Oils Fats*. 2013;28:66–7. <https://doi.org/10.1142/S0252959903000104>.
- Ma Q, Li S, Wen J, Zhu L, Zhang B, Li S, et al. Transcriptome analysis of *Acer truncatum* seeds reveals candidate genes related to oil biosynthesis and fatty acid metabolism. *J Am Soc Hortic Sci*. 2020;146(2):1–10. <https://doi.org/10.21273/JASHS05013-20>.
- Zhao W, Zhang J, Wang Z, Zhang Y, Tian W. The extract of leaves of *Acer truncatum* Bunge: a natural inhibitor of fatty acid synthase with antitumor activity. *J Enzyme Inhib Med Chem*. 2006;21(5):589–96. <https://doi.org/10.1080/14756360600774579>.
- Tanaka K, Shimizu T, Ohtsuka Y, Yamashiro Y, Oshida K. Early dietary treatments with Lorenzo's oil and docosahexaenoic acid for neurological development in a case with Zellweger syndrome. *Brain Dev-Jpn*. 2007;29(9):586–9. <https://doi.org/10.1016/j.braindev.2007.02.005>.
- Amminger G, Schäfer M, Klier C, Slavik J, Holzer I, Holub M. Decreased nervonic acid levels in erythrocyte membranes predict psychosis in help-seeking ultra-high-risk individuals. *Mol Psychiatry*. 2012;17(12):1150–2. <https://doi.org/10.1038/mp.2011.167>.
- Bates P, Stymne S, Ohlrogge J. Biochemical pathways in seed oil synthesis. *Curr Opin Plant Biol*. 2013;16(3):358–64. <https://doi.org/10.1016/j.pbi.2012.12.003>.
- Wang X, Liang H, Guo D, Duan X, Jia Q, Hou X. Integrated analysis of transcriptomic and proteomic data from tree peony (*P. ostii*) seeds reveals key developmental stages and candidate genes related to oil biosynthesis and fatty acid metabolism. *Hort Res*. 2019;6(1):19. <https://doi.org/10.1038/s41438-019-0194-7>.
- Li J, Liu A, Najeeb U, Zhou W, Liu H, Yan G, et al. 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;175:188–98. <https://doi.org/10.1016/j.ijbiomac.2021.02.013>.
- Chapman K, Ohlrogge J. Compartmentation of triacylglycerol accumulation in plants. *J Biol Chem*. 2012;287(4):2288–94. <https://doi.org/10.1074/jbc.R111.290072>.
- Lin P, Wang K, Zhou C, Xie Y, Yao X, Yin H. Seed transcriptomics analysis in *Camellia oleifera* uncovers genes associated with oil content and fatty acid composition. *Int J Mol Sci*. 2018;19(1):1–17. <https://doi.org/10.3390/ijms19010118>.
- Lardizabal K, Effertz R, Levering C, Mai J, Pedrosa M, Jury T, et al. Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiol*. 2008;148(1):89–96. <https://doi.org/10.1104/pp.108.123042>.
- Tang M, Gu I, O'Hara P, Slabas A, Quant P, Fawcett T, et al. Metabolic control analysis of developing oilseed rape (*Brassica napus* cv Westar) embryos shows that lipid assembly exerts significant control over oil accumulation. *New Phytol*. 2012;196(2):414–26. <https://doi.org/10.1111/j.1469-8137.2012.04262.x>.
- Nguyen Q, Talou T, Cerny M, Evon P, Merah O. Oil and fatty acid accumulation during coriander (*Coriandrum sativum* L.) fruit ripening under organic cultivation. *Crop J*. 2015;4(3):366–9. <https://doi.org/10.1016/j.cj.2015.05.002>.
- Wu P, Xu X, Li J, Zhang J, Chang S, Yang X, et al. Seed-specific overexpression of cotton GhDGAT1 gene leads to increased oil accumulation in cottonseed. *Crop J*. 2021;9(2):487–90. <https://doi.org/10.1016/j.cj.2020.10.003>.
- Ying P, Calviello L, Larry H, Fay L, Rothfels C, Ohler U, et al. Super-resolution ribosome profiling reveals unannotated translation events in *Arabidopsis*. *Natl Acad Sci*. 2016;113(45):E7126–35. <https://doi.org/10.1073/pnas.1614788113>.
- Yoo H, Greene G, Yuan M, Xu G, Burton D, Liu L, et al. Translational regulation of metabolic dynamics during effector-triggered immunity. *Mol Plant*. 2020;13(1):88–98. <https://doi.org/10.1016/j.molp.2019.09.009>.
- Wu H, Song G, Walley J, Ying H. The tomato translational landscape revealed by transcriptome assembly and ribosome profiling. *Plant Physiol*. 2019;181(1):367–80. <https://doi.org/10.1104/pp.19.00541>.
- Ingolia N, Ghaemmaghami S, Newman J, Weissman J. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*. 2009;324(5924):218–23. <https://doi.org/10.1126/science.1168978>.
- Juntawong P, Girke T, Bazin J, Bailey-Serres J. Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. *Proc Natl Acad Sci USA*. 2014;111(1):E203–12. <https://doi.org/10.1073/pnas.1313249111>.
- Williams CR, Brewster C, Belcher S, Rojas M, Chotewutmontri P, Ljungdahl S, et al. The *Arabidopsis* pentatricopeptide repeat protein LPE1 and its maize ortholog are required for translation of the chloroplast psbJ RNA. *Plant J*. 2019;99(1):56–66. <https://doi.org/10.1111/tpj.14308>.
- Michel A, Choudhury K, Firth A, Ingolia N, Atkins J, Baranov P. Observation of dually decoded regions of the human genome using ribosome profiling data. *Genome Res*. 2012;22(11):2219–29. <https://doi.org/10.1101/gr.133249.111>.
- Chung B, Hardcastle T, Jones J, Irigoyen N, Firth A, Baulcombe D, et al. The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. *RNA*. 2015;21(10):1731–45. <https://doi.org/10.1261/rna.052548.115>.
- Lei L, Shi J, Chen J, Zhang M, Sun S, Xie S, et al. Ribosome profiling reveals dynamic translational landscape in maize seedlings under drought stress. *Plant J*. 2015;84(6):1206–18. <https://doi.org/10.1111/tpj.13073>.
- Liu M, Wu S, Wu J, Lin W, Wu Y, Tsai T, et al. Translational landscape of photomorphogenic *Arabidopsis*. *Plant Cell*. 2013;25(10):3699–710. <https://doi.org/10.1105/tpc.113.114769>.
- Calviello L, Ohler U. Beyondread-counts: Ribo-seq data analysis to understand the functions of the transcriptome. *Trends Genet*. 2017;33(10):728–44. <https://doi.org/10.1016/j.tig.2017.08.003>.
- Araujo P, Yoon K, Ko D, Smith A, Qiao M, Suresh U, et al. Before it gets started: regulating translation at the 5'UTR. *Comp Funct Genomics*. 2012;2012:475731. <https://doi.org/10.1155/2012/475731>.
- Cabrera QL, Herberg S, Pauli A. Decoding sORF translation - from small proteins to gene regulation. *RNA Biol*. 2016;13(1):105159. <https://doi.org/10.1080/15476286.2016.1218589>.
- Pan L, Woodfield H, Harwood J, Chye M, Scofield S. Comparative transcriptomics analysis of *Brassica napus* L. during seed maturation reveals dynamic changes in gene expression between embryos and seed coats and distinct expression profiles of acyl-CoA-binding proteins for lipid accumulation. *Plant Cell Physiol*. 2019;60(12):2812–25. <https://doi.org/10.1093/pcp/pcz169>.
- Alotaibi S, Elseehy M, Aljuaid B, El-Shehawi A. Transcriptome analysis of *Joboba* (*Simmondsia chinensis*) during Seed development and liquid wax ester biosynthesis. *Plants (Basel)*. 2020;9(5):588. <https://doi.org/10.3390/plants9050588>.
- Liao B, Hao Y, Lu J, Bai H, Guan L, Zhang T. Transcriptomic analysis of *Perilla frutescens* seed to insight into the biosynthesis and metabolic of unsaturated fatty acids. *BMC Genomics*. 2018;19:213. <https://doi.org/10.1186/s12864-018-4595-z>.

34. Ji X, Mao X, Hao Q, Liu B, Xue J, Li R. Splice variants of the castor WRI1 gene upregulate fatty acid and oil biosynthesis when expressed in tobacco leaves. *Intl J Mol Sci*. 2018;19:146. <https://doi.org/10.3390/ijms19010146>.
35. Prystenski J. Studies on transcription factors involved in seed oil biosynthesis. 2011.
36. Dong L, Jin C, Duan S, Zhu Y, Qi S, Liu K, et al. MYB89 transcription factor represses seed oil accumulation. *Plant Physiol*. 2016;173:1211–5. <https://doi.org/10.1104/pp.16.01634>.
37. Chew G, Pauli A, Schier A. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat Commun*. 2016;7:11663.
38. Hayden C, Jorgensen AR. Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol*. 2007;5(32):1–30. <https://doi.org/10.1186/1741-7007-5-32>.
39. Kim B, Cai X, Vaughn J, Arnim A. On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. *Genome Biol*. 2007;8:R60. <https://doi.org/10.1186/gb-2007-8-4-r60>.
40. Kochetov A. Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *BioEssays*. 2008;30(7):683–91. <https://doi.org/10.1002/bies.20771>.
41. Wu Q, Cao Y, Chen C, Gao Z, Guy R. Transcriptome analysis of metabolic pathways associated with oil accumulation in developing seed kernels of *Styrax tonkinensis*, a woody biodiesel species. *BMC Plant Biol*. 2020;20:121. <https://doi.org/10.1186/s12870-020-2327-4>.
42. Kozak M. Pushing the limits of the scanning mechanism for initiation of translation. *Gene*. 2002;299(1):1–34. [https://doi.org/10.1016/S0378-1119\(02\)01056-9](https://doi.org/10.1016/S0378-1119(02)01056-9).
43. Rajkowsch L, Vilela C, Berthelot K, Ramirez C, McCarthy J. Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *J Mol Biol*. 2004;335(1):71–85. <https://doi.org/10.1016/j.jmb.2003.10.049>.
44. Gao J. Experimental guidance for plant physiology. Beijing, Chinese: China Higher Education Press; 2006.
45. Zhang Z, Luo Y, Wang X, Yu F. Fruit spray of 24-epibrassinolide and fruit shade alter pericarp photosynthesis activity and seed lipid accumulation in *Styrax tonkinensis*. *J Plant Growth Regul*. 2018;37:1066–84. <https://doi.org/10.1007/s00344-017-9769-4>.
46. Kim D, Ben L, Steven LS. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60. <https://doi.org/10.1038/nmeth.3317>.
47. Love M, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.
48. Morlan J, Qu K, Sinicropi D. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS one*. 2012;7(8):e42882. <https://doi.org/10.1371/journal.pone.0042882>.
49. Ingolia N, Brar G, Rouskin S, McGeachy A, Weissman J. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc*. 2012;7(8):1534–50. <https://doi.org/10.1038/nprot.2012.086>.
50. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
51. Bazzini A, Johnstone T, Christiano R, Mackowiak S, Obermayer B, Fleming E, et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J*. 2014;33(9):981–93. <https://doi.org/10.1002/embj.201488411>.
52. Wang L, Park H, Dasari S, Wang S, Kocher J, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res*. 2013;41(6):e74. <https://doi.org/10.1093/nar/gkt006>.
53. Wu L, Lv Y, Ye Y, Liang Y, Ye J. Transcriptomic and translational analyses reveal insights into the developmental regulation of secondary metabolism in the young shoots of tea plants (*Camellia sinensis* L.). *J Agric Food Chem*. 2020;68(39):10750–62. <https://doi.org/10.1021/acs.jafc.0c03341>.
54. Wang R, Liu P, Fan J, Li L. Comparative transcriptome analysis two genotypes of *Acer truncatum* Bunge seeds reveals candidate genes that influence seed VLCFAs accumulation. *Sci Rep*. 2018;8(1):15504. <https://doi.org/10.1038/s41598-018-33999-3>.
55. Jeyaraj A, Zhang X, Hou Y, Shangguan M, Gajjeraman P, Li Y, et al. Genome-wide identification of conserved and novel microRNAs in one bud and two tender leaves of tea plant (*Camellia sinensis*) by small RNA sequencing, microarray-based hybridization and genome survey scaffold sequences. *BMC Plant Biol*. 2017;17(1):212. <https://doi.org/10.1186/s12870-017-1169-1>.
56. Ribosome footprint profiling enables elucidating the systemic regulation of fatty acid accumulation in *Acer truncatum*. NCBI BioProject accession: PRJNA877423. 2022. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA877423>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

