JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2024 | Volume 12 | Issue 1

Identification and expression analysis of *Oleosin* gene family in walnut (*Juglans regia* L.)

Seyyed Hamidreza Hashemipetroudi[1*], Fatemeh Ahmadi[2]

[1] Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Agricultural Sciences and Natural Resources University (SANRU), Sari, Iran
[2] Department of Horticultural Science and Engineering, Faculty of Plant Production, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

Original paper

Abstract: Oleosins as structural proteins play a crucial role in stabilization the oil-bodies, their size, number, and content. Here, exon/intron structure, conserved motifs/domains, and expression patterns of *Oleosin* gene family were identified in diploid crop *Juglans regia* L. This study found seven non-repetitive *JrOleosin* genes through a genome-wide analysis, displaying remarkable structural and physicochemical differences. Our phylogenetic analysis of the 46 oleosin genes revealed that how these genes are related, highlighting that they share conservations while also having unique divergences across different species. The *JrOleosin* intronless genes (IGs) might be examples of how evolution adapts to enhance the accuracy and effectiveness of gene expression. Transcriptome analysis showed that the *JrOleosin* genes were spatially and temporarily regulated with the highest transcript found in seeds and fruits. This suggests they are vital for storing lipids during these developmental stages. These results help us to better understand the features of the *JrOleosins* gene family, especially their structure and expression. Future research should concentrate on revealing how *oleosin* expression is controlled and how these proteins interact with others linked to lipid synthesis during seed development.

Keywords: Oil-body, oleosin, meta-analysis, walnut, genome-wide.

Edited by
Dr. Leila Ahangar,
Department of Plant Production, Faculty of Agriculture and Natural Resources,
Gonbad Kavous University, Iran

Date
Received: 25 November 2024
Accepted: 10 February 2025
Published: 12 February 2025

Correspondence
Dr. Seyyed H. Hashemipetroudi
shr.hashemi@sanru.ac.ir;
irahamidreza@yahoo.com

Citation
Hashemipetroudi, S.H., Ahmadi, F. (2024). Identification of *oleosin* gene family in *Juglans regia* L. *J Plant Mol Breed.* 12(1): 150-165. doi:10.22058/jpmb.2025.2029871.1317.

Introduction

Walnut (*Juglans regia* L.), native to central Asia, is an old fruit tree prized economically for its output of edible seeds, vegetable oil, and industrial lumber.

(Hao et al., 2024). Walnut oil seeds are rich in essential unsaturated fatty acids, including linoleic and linolenic acids, and their nutritional importance has been highlighted in the context of dried fruit (Goodarzi et al., 2023). In the majority of cultivars, walnut kernels comprise approximately 60% oil. After China and the United States, Iran ranks third in production and second in cultivated area for these trees. Consequently, Iran accounts for 7.1% of global walnut production (FAO, 2022). Studies indicate that wild populations of persian walnuts are present in the Hyrcanian woodlands of northern Iran (Shamlu et al., 2018). Given the significant genetic diversity of Persian walnuts in Iran, the core of walnut origin, breeding operations in this nation concentrate on harnessing this genetic variation. For a long time, researchers have focused on fruit characteristics like early fruit maturity, size, color, and percentage of walnut kernel oil in breeding programs, with an eye toward improving the product's quality and quantity. However, there is a lack of data on the molecular level regarding the metabolism of oil seeds and the storage processes of fatty acids.

Lipids serve as carbon stores and energy providers for plants. Seed lipids are primarily triacylglycerols, which are the primary reserves found in seeds and oil bodies (Yang and Benning, 2018). Abundant in both prokaryotic and eukaryotic cells, oil bodies—also called oleosomes—are specialized organelles that store energy in the form of neutral lipids. Oil bodies' architecture consists in an oil core and a hydrophobic exterior surrounded by a phospholipid monolayer membrane and membrane proteins, forming spherical vesicles with an approximate 0.5–2 micrometer (µm) diameter (Khor et al., 2013; Jackson, 2019). These entities are synthesized in the endoplasmic reticulum during the Initial phases of seed development and then transported to the cytoplasm via the germination process (Huang, 1996). Besides seeds, they are found in several sections of higher plants, including pollen, tapetum, leaves, fruits, and roots (Yuan et al., 2021).

Plant naturally occurring protective mechanism against lipid oxidation and severe stress circumstances is their oily bodies, which show capacity to maintain lipid stability. Along with the stability of membrane proteins and various metabolic activities, oil body shape and dimensions change constantly during seed development (Huang and Huang, 2017; Pyc et al., 2021). The size of oil bodies can fluctuate over the phase of seed development. Because of their surfactant qualities, these chemicals are environmentally beneficial and renewable. They are employed in the food, pharmaceutical, and cosmetic sectors. Recent studies demonstrate that the overexpression of some genes associated with oleosin can modulate the dimensions of oil bodies, suggesting that seed size and weight significantly influence seed oil accumulation (Chen et al., 2019). Reports indicate a negative link between the size of oily substances and the quantity of oil (Siloto et al., 2006).

Oleosin, caleosin, and streosin are the most well-known plant oil body proteins, and they all play key roles in the size, number, and stability of oil bodies, as well as lipid accumulation (Nogales-Bueno et al., 2021; Kim and Hyun, 2024). Oleosin comprises 80-90% of structural proteins among these proteins (Siloto et al., 2006; Dong et al., 2024). Relatively tiny with a molecular weight of 15–26 kilodaltons, this protein has three distinct regions and alkaline features (Tzen et al., 1997). The middle part is hydrophobic and has 72 residues. The amino terminal parts are hydrophilic and have 50 to 70 residues. The dipolar carboxyl has a length that can vary. The highly conserved non-polar middle region with beta structures in two reverse directions penetrates the lipid membrane and forms a hairpin structure to stabilize oil bodies (Roux et al., 2004; Wu et al., 2010). Two arms of this hairpin shape are protected by a proline ring containing three proline and one serine residue. Despite their dipolar alpha-helical configurations, the carboxyl and amino portions are less conserved than the hairpin area. They can bind to metabolic enzymes and regulatory proteins on the phospholipid surface and in the cytoplasm (Huang and Huang, 2017). The negative charge of these places causes steric hindrance and electrostatic repulsion, preventing the integration of stored lipids, particularly during seed drying, and thereby facilitating the movement of oil bodies during germination (Leprince et al., 1997; Baud and Lepiniec, 2010).

Research on the oleosin gene sequence began with maize and has now expanded to other plants including sunflower, rapeseed, sesame, rice, and soybean (Vance and Huang, 1987; Keddie et al., 1992; Chen et al., 1997; Sarmiento et al., 1997; Wu et al., 1998; Alexander et al., 2002). The majority of the 17 oleosin genes found in the *Arabidopsis thaliana* genome are expressed in the seeds, according to previous studies (Siloto et al., 2006). Oleosin mutant seeds are more sensitive to freezing, implying that oleosins help plants survive in cold stress condition. (Shimada et al., 2008). Fatty acid chains undergo a conformational transition from their disordered to their regular form at low temperatures, leading to the dispersion of massive, unstable oil bodies in the cytoplasm. In this case, the fluidity, physiology, and thickness of the lipid membrane are significantly reduced. Because it interferes with the plant's physiological processes, this can severely limit its growth and development potential (Parthibane et al., 2012).

The diploid walnut seed (2n=2x=32) contains around 108 genes that are involved in lipid biosynthesis. Among these genes are 33 genes that are engaged in triacylglycerol biosynthesis, seven genes that are involved in oil bodies, and eight genes that are transcription factors (Yan et al., 2021). The function of oleosins in walnut oil seeds remains ambiguous. Thus, the present study utilized bioinformatics approaches to clarify the *JrOleosin* gene structure, protein domains, chromosomal distribution, expression patterns, and evolutionary relationships of oleosin with other plant species.

Materials and Methods

Identification of the JrOleosin gene family members

The protein sequences of the *Oleosin* gene family members from *Arabidopsis thaliana* were initially identified and downloaded from the TAIR (http://www.arabidopsis.org/) database. Genomic sequences of *A. thaliana* (TAIR10.1), *Juglans regia* (Walnut 2.0), *Elaeis guineensis* (African oil palm) and *Olea europaea* (Olive) were downloaded from public databases NCBI (https://www.ncbi.nlm.nih.gov/) and Phytozome (phytozome.jgi.doe.gov). To identify oleosin homologs, the oleosin domain profile (PF01277) was used for HMMER (v3.3, http://hmmer.org/) searches (E-value < 1E-5). Finally, the coding sequence (CDS), protein, and genomic sequences of *JrOleosin* family were identified. The presence of the conserved oleosin domain in deduced peptides was confirmed using InterProScan(https://www.ebi.ac.uk/interpro/search/sequence-search) (Jones et al., 2014), Pfam (http://pfam.xfam.org/) (Finn et al., 2016) and SMART (http://smart.embl-heidelberg.de/) (Letunic et al., 2015) protein database.

Physiochemical analysis of JrOleosin proteins

The subcellular localization of JrOleosin proteins was predicted using the WoLF PSORT algorithm (https://wolfpsort.hgc.jp/) (Horton et al., 2007). The sequences of JrOleosin were analyzed with ExPASy ProtParam (http://www.expasy.org/tools/protparam.html) to obtain the number of amino acids, theoretical isoelectric point (pI), molecular weight, GRAVY, instability index, and aliphatic index.

Exon/intron structures and motif organization

The *JrOleosin* gene'' exon/intron structures were visualized via Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn) (Hu et al., 2015). The following parameters were used to find the conserved protein motifs of all JrOleosin using the MEME program (http://meme-suite.org/tools/meme/): Ten is the maximum number of motifs, and the lengths of the motifs range from six to fifty amino acid residues (Bailey et al., 2009).

Phylogenetic tree construction

Oleosin protein sequences in walnut, Arabidobsis, olive and oil palm were obtained from Phytozome database (Goodstein et al., 2012). Multiple sequence alignment of oleosin proteins in *J. regia*, Arabidopsis, olive and oil palm were performed using MUSCLE and finally the maximum likelihood (ML) phylogenetic tree were constructed using MEGA 11 (Tamura et al., 2013).

Expression profiling of JrOleosin genes based on RNA-seq data

To examine the function and expression of the *JrOLE* gene family, 210 RNA-seq samples from 13 projects were gathered from publicly accessible RNA-seq data associated with *J. regia*.

Transcriptome datasets were selected from various treatments or genetic backgrounds, their accession number were listed as follow: PRJNA232394, PRJNA235890, PRJNA237044, PRJNA609369, PRJNA622910, PRJNA642991, PRJNA643637, PRJNA673559, PRJNA688391, PRJNA734671, PRJNA776681, PRJNA794344, PRJNA806342, PRJNA862472. TPM (transcripts per million) had been utilized to measure transcript expression levels.

Results

JrOleosin gene family identification

Seven *Oleosin* non-repetitive genes were detected in *J. regia* genome based on Oleosin specific domain with the Pfam number of PF01277 (Table 1). Twelve, nine, and eighteen *Oleosin* genes were found in the genomes of Arabidopsis, African oil palm, and olive, respectively, by a hidden Markov model (HMM) tools based on the Oleosin domain (Pfam accession number PF01277). The comparative genomics analysis revealed the existence of varied numbers of *Oleosin* genes among various species, therefore offering understanding of the evolutionary divergences and functional conservation of these genes in respect to lipid storage and metabolism. In a previous report, 17 genes belonging to the *Oleosin* family were identified in the Arabidopsis (Siloto et al., 2006; Yuan et al., 2021). In the current investigation, one new *AtOleosin* gene within arabidopsis genome was identified (18 loci, 28 isoforms), which was made possible due to the well-annotated genome of *A. thaliana* (Araport11) and the high-quality genome sequencing data available. The *Oleosin* genes were named based on their positions on phylogenetic tree, and by using genus and species abbreviations for clarity. This systematic nomencalture may support upcoming research into gene functions across plant species (Table 1, Figure 1).

Table 1. Physiochemical properties of the *Oleosin* genes in *Juglans regia.*

Gene name	Gene ID	mRNA RefSeqs ID	Homolog to Arabidopsis genes		PL	MW (kDa)	pI	GRAVY	II	AI
JrOleosin01	LOC108998787	XM_018975486.2	AT4G25140.1	*AtOleosin02*	140	14.838	9.65	0.316	46.25	103.86
JrOleosin02	LOC108983309	XM_019000720.2	AT4G25140.1		139	14.691	10.14	0.308	36.55	105.97
JrOleosin03	LOC109009462	XM_018989925.2	AT4G25140.1		147	15.606	9.56	0.42	29.42	108.84
JrOleosin04	LOC109010520	XM_018991380.2	AT5G51210.1	*AtOleosin01*	156	16.776	9.03	0.435	39.78	98.27
JrOleosin05	LOC109007243	XM_018986840.2	AT3G01570.1	*AtOleosin05*	154	16.278	10.1	0.230	52.53	112.14
JrOleosin06	LOC108982740	XM_018954195.2	AT3G01570.1		155	15.896	9.65	0.290	31.56	95.23
JrOleosin07	LOC109003841	XM_018982154.2	AT3G01570.1		159	16.620	9.91	0.181	34.65	95.16

A.C.: Accession number; PL: Protein length (aa); MW: Molecular weight; pI: Isoelectric point; II: Instability index; AI: Aliphatic index.

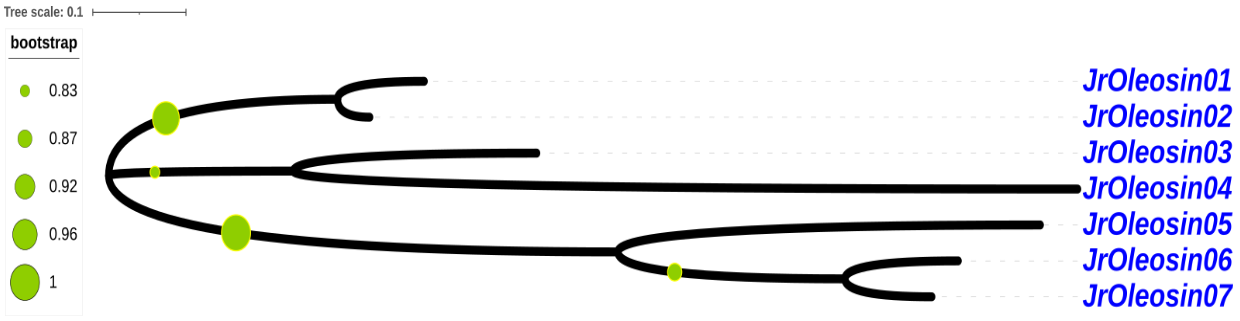

Figure 1. Phylogenetic analysis of Oleosin in *J. regia*.

Physiochemical properties of JrOleosin

Our analysis on the *JrOleosin* gene family showed significant differences in the biochemical properties of gene family members, which are listed in Table 1, and include the following critical parameters: molecular weight, theoretical isoelectric point (pI), aliphatic index, instability index, and grand average of hydropathicity (GRAVY). Each gene product was assessed for critical parameters, including amino acid (aa) length, molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and grand average of hydropathicity (GRAVY). The average length of the JrOleosin proteins ranged from 139aa (JrOleosin02) to 159aa (JrOleosin07), with an average of roughly 151aa. The pI values also ranged from 9.03 (JrOleosin04) to 10.14 (JrOleosin02). rOleosin07 having the highest molecular weight of 16.62 kDa and the lowest JrOleosin02 having the lowest at 14.691 kDa indicate a varied functional capacity within the gene family, which may influence their molecular function and biological processes.

The increased pI values are likely connected with JrOleosins' specific features and functions, which impact their capacity to interact with other biomolecules and their overall performance. The instability indices ran from 29.42 (JrOleosin03) to 52.53 (JunrOleosin05). The higher instability index of the latter implies it could be less stable and more prone to degradation, which would affect its physiological functions under stress conditions. From 95.16 (JrOleosin07) to 112.14 (JrOleosin05), the aliphatic index—which shows protein stability at high temperatures— were ranged. Higher values show thermal stability and capacity for operation at several temperatures. Essential for their interaction with water-soluble components inside plant cells, the GRAVY values for these proteins ranged from 0.181 (JrOleosin07) to 0.435 (JrOleosin04), indicating a generally hydrophilic character for them (Table 1).

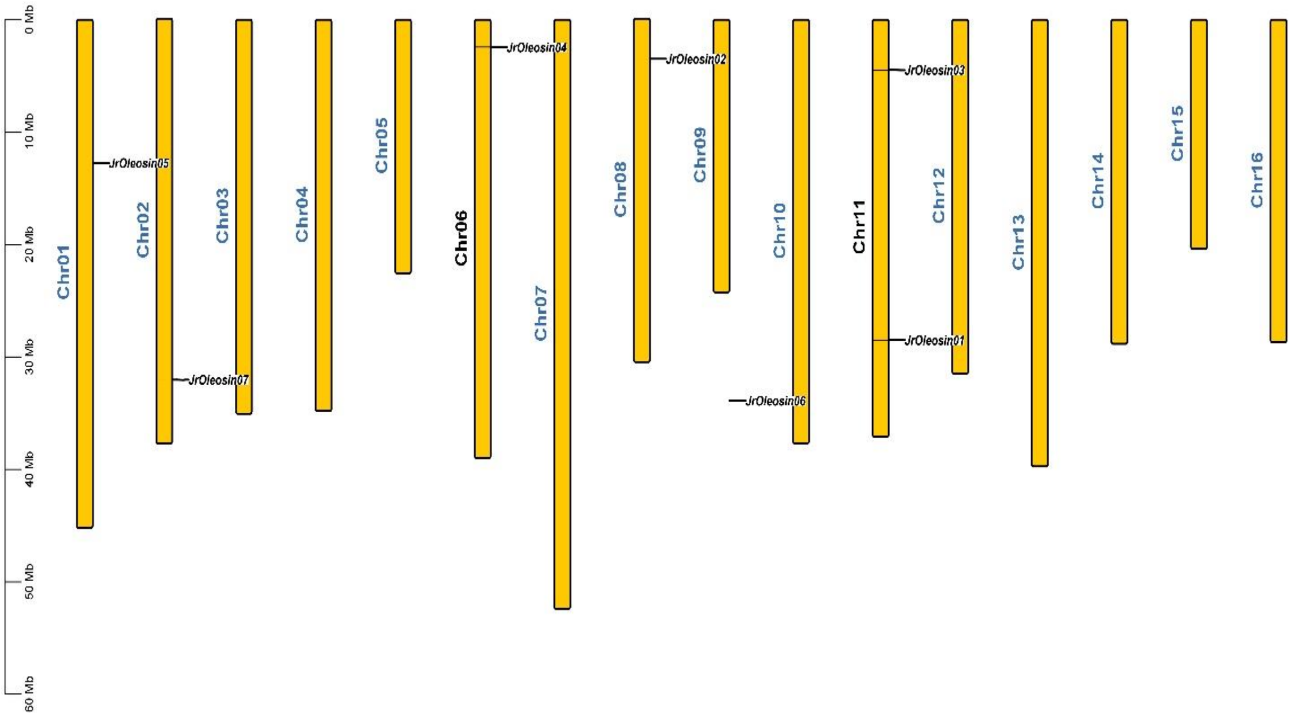

Figure 2. The chromosomal distribution of the *Oleosin* gene family in *Juglans regia*. The left side of each chromosome displays the chromosome numbers and their respective approximate sizes.

Phylogenetic analysis of JrOleosin

Figure 1 depicts the phylogenetic tree that elucidates the evolutionary relationships among the Oleosin gene family members in *J. regia*. The analysis employed bootstrap values to evaluate the confidence of the branching patterns, with circles indicating the level of support for each node. The bootstrap values span from 0.83 to 1.0, showing a significant level of confidence in the evolutionary relationships among the *JrOleosin* genes. The larger circles (bootstrap values of 0.96 and 1.0) indicate strong support for the clades established by *JrOleosin* genes, hence enhancing the credibility of the inferred evolutionary relationships.

Figure 3. Conserved structural domains (a) distribution of identified motifs (b), and motif logo sequences in the JrOleosin family. In the above forms, motifs and domains were shown with different color boxes. Scale indicates amino acid length.

The JrOleosin genes are categorized into distinct groups, with JrOleosin01, JrOleosin02, and JrOleosin03 forming a closely related cluster. JrOleosin04, exhibiting unique characteristics, is also included in this cluster, suggesting potential shared functional roles or similar evolutionary pressures. In contrast, JrOleosin05, JrOleosin06, and JrOleosin07 appear more distantly related, indicating possible divergence in function or adaptation. The phylogenetic relationships illustrated in the tree offer insights into the evolutionary background of the JrOleosin gene family. The close clustering of certain JrOleosins may indicate conserved functions associated with lipid metabolism and stress responses, which are essential for the survival of *J. regia* in diverse environmental conditions.

Chromosomal distribution of JrOleosin gene

The results of the distribution pattern of *JrOleosin* genes on 16 chromosomes showed that seven were located on six chromosomes (Chr 01, 02, 06, 08, 10 and 11). The highest numbers of genes on Chr11 were *JrOleosin01* and *JrOleosin03* genes. Other chromosomes (Ch 01, 02 , 06, 08 and 10) had one *JrOleosin* gene. Therefore, Chr01 had *JrOleosin05* gene, Chr02 had *JrOleosin07* gene, Chr06 had *JrOleosin04* gene, Chr08 had *JrOleosin* gene, and Chr10 had *JrOleosin06* gene. In addition, most of the *JrOleosin* genes are distributed near the end of each chromosome (Figure 2).

JrOleosin domain and motif organization

Seven different JrOleosin proteins were identified through the investigation of the Oleosin superfamily, and each one showed a notable degree of alignment with the Oleosin superfamily domain (Figure 3). For every protein, the data is summarized in Table 1. While the others genes had bigger domain sizes of 113 to 114 amino acids, JrOleosin01 had the shortest domain size of 87 amino acids, hence the lengths of the protein domains varied. Different activities or genetic modifications occurring in the JrOleosin proteins could be the reason of this variation in domain size. Several Oleosin superfamily members have constant domain sizes, suggesting a conserved structural characteristic that may be essential for lipid storage and oil body stabilization.

Motifs 1, 2, 3, and 4 were found in all seven JrOleosin proteins out of a total of ten, while several of them shared some motifs. For example, JrOleosin 06 and 07 had motifs 5 and 7, JrOleosin 03 and 05 had motifs 6, JrOleosin 07 and 08 had motifs 8, JrOleosin 01 and 02 had motifs 9, and JrOleosin 04 and 07 had motifs 10. Also, the highest number of motifs (eight numbers) belonged to JrOlesin07 protein, and the width of 6 and 50 residues for each conserved motif (Figure 3).

JrOleosin exon/intron structure

The study indicates that the *Oleosin* gene family possesses comparable coding sequences (CDS) and untranslated region (UTRs) patterns, implying structural conservation, functional similarity, and regulatory functions, while also featuring a distinct UTR region (Figure 4).

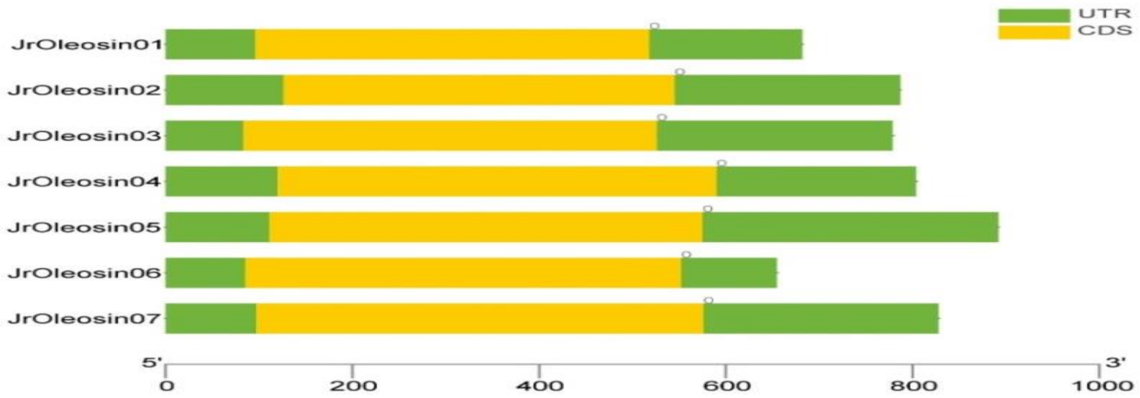

Figure 4. Analysis of *JrOleosin* gene family structures. The green and yellow boxes denote the *JrOleosin* UTR regions and CDS, respectively.

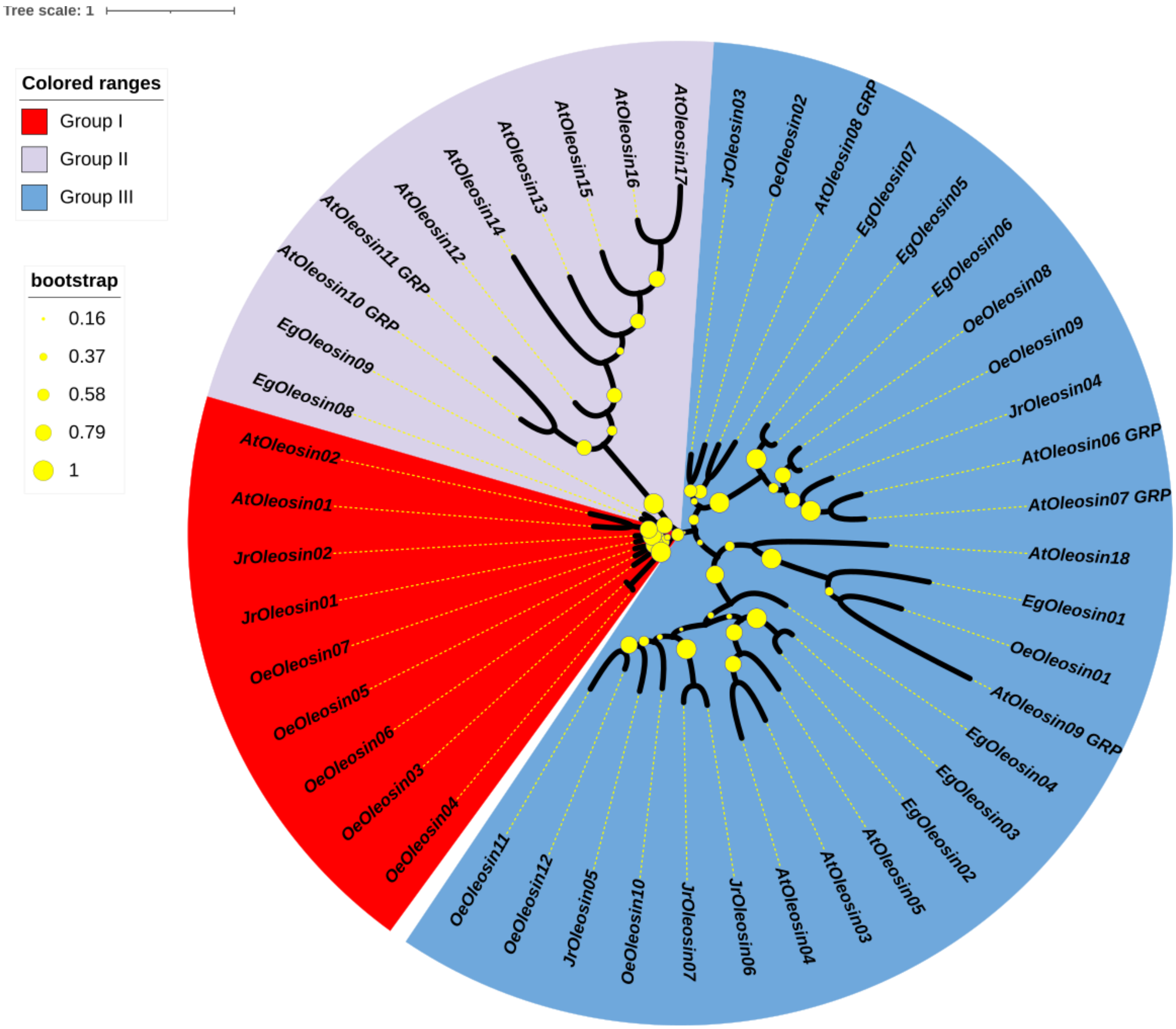

Figure 5. Phylogenetic tree of the Oleosin family based on an alignment of the proteins found in *Arabidopsis thaliana*, *Olea europaea*, *Elaeis guineensis* and *Juglans regia*. Different Oleosin groups are presented by different-colored.

Furthermore, all these genes demonstrate tissue-specific expression in seed tissue, underscoring their specialized roles in lipid metabolism during seed development. The analysis of the JrOleosin gene architecture also indicates the lack of introns in the examined genes.

Phylogenetic relationships of Oleosin gene family

The phylogenetic analysis of the *Oleosin* subfamily highlights the evolutionary dynamics within the oil body genes. The high bootstrapping score for the observed clades suggests that these relationships are well-established. Understanding evolutionary relationships, especially identification *AtOLE* ortholog in *J. regia*, will enhance our understanding of *JrOleosins'* roles in plant biology and their potential applications in stress tolerance. To analyze the evolutionary relationships of *JrOleosin* gene family members, 18 *A. thaliana*, 12 *O. europaea*, 9 *E. guineensis* oleosin proteins and 7 *J. regia* oleosin proteins were used to construct a phylogenetic tree. So JrOleosin proteins were divided into three main groups. The number of JrOleosins in groups I and III were 2 and 5, respectively. Moreover, the group II contained no JrOleosins mamber. Group III

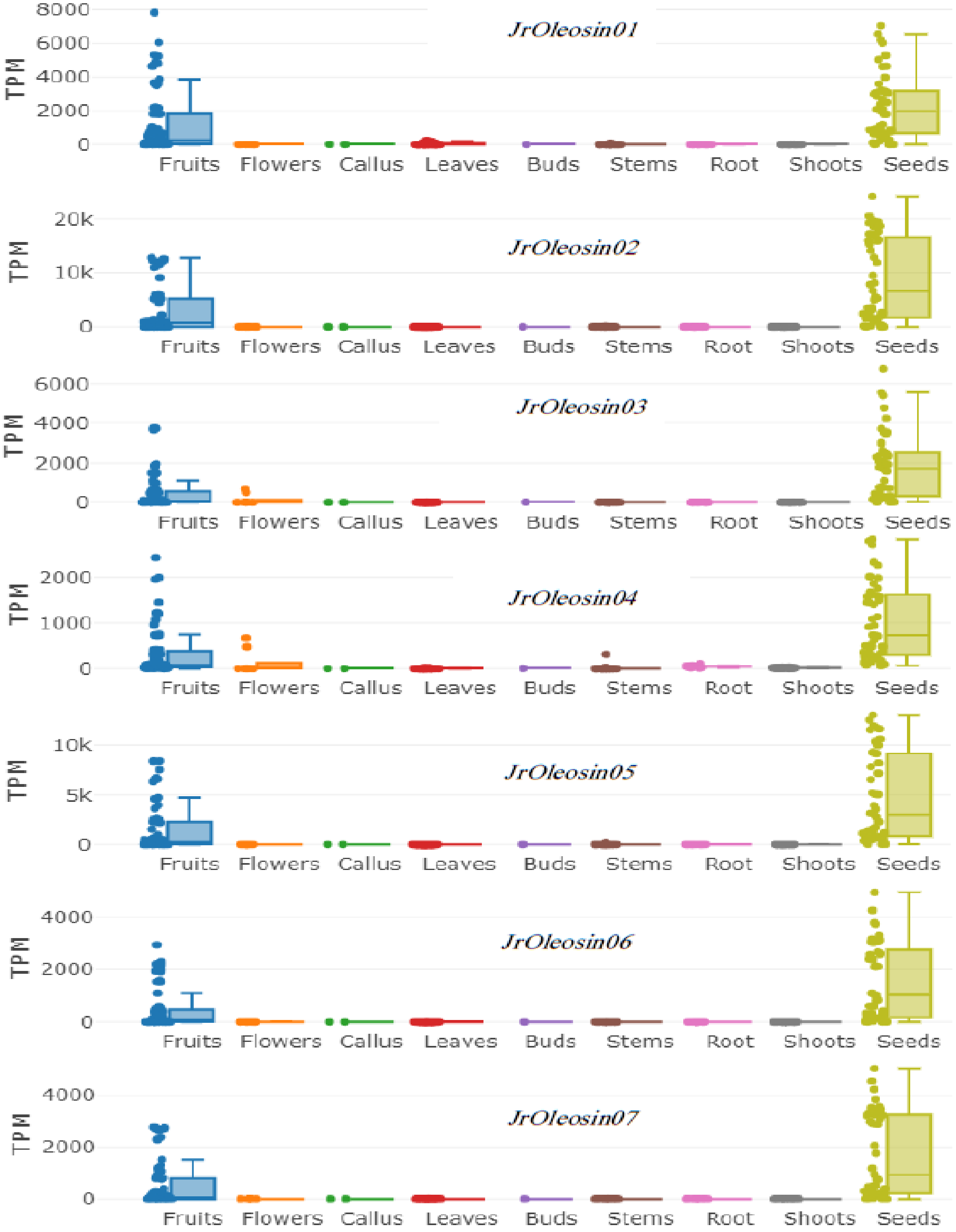

Figure 6. Meta-analysis of the *Oleosin* 's expression pattern in nine distinct tissues of *Juglans regia*. The box-plot illustrates the relative expression levels of the *JrOleosin* genes across tissues, including roots, stems, leaves, flowers, seeds, and other organs in TPM (transcripts per million) format. This analysis highlights tissue-specific expression patterns of *JrOleosin* genes in the sues of seeds and fruits.

encompass the largest number of JrOleosin genes family mambers (Figure 5). In addition, by comparing the distribution of oleosin proteins in different groups and species, it can be found that the distribution of oleosin proteins in various groups was similar, indicating that their functions may be similar and strongly conserved during evolution.

Expression profiles of JrOleosin

As presented in Figure6, the transcript levels of *JrOleosin* gene family mamber in different walnut organs including fruit, flower, callus, leaf, bud, stem, root, shoot and seed were different, and the highest level of gene expression found in seed and fruit organs (Yang et al., 2024). The lowest and highest levels of gene expression (TPM) were observed in *JrOleosin01* and *JrOleosin02* genes, respectively. Transcriptome analyses indicate that the *JrOleosin* genes were expressed spatially and temporally during seed and fruit development.

Discussion

Oleosins are well-known to play a role in lipid metabolism and oil body stabilization, both of which are necessary for seed germination and development (Shao et al., 2019). Oleosins are the most abundant protein constituents that play essential roles in the formation and stabilization of lipid droplets in seeds of oil crops (Zhi et al., 2017; Kim and Hyun, 2024). According to Huang and Huang (2017)Huang and Huang (2017)Huang and Huang (2017)Huang and Huang (2017)Huang and Huang (2017), the oleosin gene is primarily found in higher plants and green algae. Oleosins in plants can be classified into six types: primary, general, tapetum, mesocarp, and seed, with low and high molecular weights (Huang and Huang, 2015). Nine forms of this protein were found in tapetum branches, three in general branches, two in seeds with low molecular weight, and three in seeds with high molecular weight, according to research on *Arabidopsis* (Chen et al., 2019). The identified characteristics of JrOleosins convey significant insights into their potential biological functions. Due to their differing molecular weights and stability indices, these proteins may serve unique functions in lipid metabolism, seed development, or stress response pathways in *J. regia*. The preference for higher aliphatic indices and GRAVY values may indicate a specific function in preserving cellular integrity and regulating osmotic balance under environmental stresses (Hashemipetroudi et al., 2023; Mohammadi et al., 2023). Members of the JrOleosin family have comparable physicochemical characteristics to those reported in prior research, including low molecular weight, alkaline isoelectric point, and a positive GRAVY value (Zou et al., 2022; Zou et al., 2024).

The diverse biochemical properties of the *Oleosin* gene family underline their potential functional diversification within *J. regia*. The findings show that, while domain sizes vary, a significant proportion of JrOleosin proteins have similar lengths, emphasizing their possible functional commonalities within the Oleosin subfamily. Further research into these protein activities may reveal their significance in plant lipid metabolism and storage. Future functional studies, along with expression analyses under diverse environmental conditions, are crucial for clarifying their specific roles in plant stress genomics. Grasping the distinct characteristics of each JrOleosin can improve our understanding of their function in the overall adaptability of *J. regia*. The distribution of UTRs and CDS among the *JrOleosin* genes reveals heterogeneity in gene structure, potentially indicating variances in regulatory mechanisms and functional roles. Genes with 5′ and 3′-UTR may have regulatory functions that influence translation efficiency, mRNA localization, and stability (Heidari et al., 2023). Additional functional investigations are necessary to clarify the specific roles of these UTRs and coding areas in the overall functionality of JrOleosin proteins in plant biology. The examining of the *JrOleosin* gene architecture reveals the absence of introns in the analyzed genes. A fascinating group of genes that are also present in eukaryotes are intronless genes (IGs), which are a common feature of prokaryotes (Chen et al., 2023). The absence of introns can be ascribed to multiple factors: gene structure, evolutionary adaptation, functional consequences, and genomic context (Aviña-Padilla et al., 2021). The lack of intron of *JrOleosin* gene family may suggest evolutionary adaptations that promote gene expression efficiency. Some reports believe that amount of introns provide advantages to plant genomes, such as regulating gene expression and enhancing protein diversity (Xiao et al., 2022). New funding on the *Poaceae* genome comparison also revealed that IGs were expressed at low levels (Chen et al., 2023). However, IGs that do not undergo splicing have a more degree of transcriptional fidelity, which increases their imapct in regulatory roles (Aviña-Padilla et al., 2021). IGs, likely due to selective forces, may exhibit lower post-transcriptional gene expression variability and prefer simpler gene architectures for specific functions (Aviña-Padilla et al., 2022).

The results show that the oleosin proteins in walnut were divided into three groups with a total of 46 members, which were similar to those in olive and oil palm and much higher than in *Arabidopsis*. Phylogenetic analysis indicates that JrOleosin01, JrOleosin02, JrOleosin03, and JrOleosin04 are oil body proteins that control oil body dynamics; JrOleosin05, JrOleosin06, and JrOleosin07 are proteins that delay germination and affect the seeds' ability to withstand freezing. Therefore, the expansion and diversification of the *oleosin* gene family may be closely related to the evolutionary process of plants from lower to higher (Shao et al., 2019).

In this study, all seven *JrOleosin* genes show a strong seed-specific expression pattern. Fruits show the second-highest expression, though considerably lower than seeds. Other tissues have negligible expression. The up-regulation of JrOleosin genes during seed maturation was reported in *Camelina sativa* (Abdullah et al., 2016), soybean (Yang et al., 2019), *Paeonia section* (Li et al., 2015), *Perilla frutescens* (Zhang et al., 2021), and *Prunus pedunculata* (Bao et al., 2021) showing specific expression patterns in oil-rich tissue. In *Cyperus esculentus*, the expression levels of different *oleosin* genes had low or high expression across different developing stages (Dong et al., 2024). The expression profile of *CpOLE* genes was detected in at least twelve of the studied tissues, though gene abundances are highly diverse. Total transcripts of the whole gene family were most abundant in shoot tissue (Zou et al., 2024). Analysis of *Arabidopsis* oleosin homolog proteins indicated that *OLE1* and *OLE4* negatively regulated oil body size, whereas *OLE2* positively influenced oil body size (Miquel et al., 2014). These comparisons highlight the evolutionary conservation of oleosin gene functions while also pointing to species-specific adaptations.

Conclusion

The results elucidate the structural properties of the *JrOleosin* genes, establishing a foundation for subsequent inquiries into their biological functions and regulatory mechanisms. The lack of introns in the *JrOleosin* genes underscores their potential for efficient expression and functional specialization, which may be essential for their involvement in plant lipid metabolism. The evolutionary and functional effects of this intronless gene structure could be studied. These JrOleosin proteins' expression patterns under different environmental circumstances and interactions with other lipid-associated proteins could be studied in the future. These dynamics may reveal oil body formation and stability mechanisms, advancing agricultural biotechnology and crop development.

Supplementary Materials

There is no supplementary material for this article.

Author Contributions

Conceptualization, S.H.H.; bioinformatic analysis, S.H.H.; investigation, F.A.; writing—original draft preparation, S.H.H. and F.A.; writing—review and editing, S.H.H.; All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

This research is supported by the Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT) and Sari Agricultural Sciences and Natural Resources University (SANRU). The authors also gratefully acknowledge the use of the services and facilities of the GABIT during this research.

Conflicts of Interest

The authors declare no conflict of interest.

References

Abdullah, H.M., Akbari, P., Paulose, B., Schnell, D., Qi, W., Park, Y., Pareek, A., and Dhankher, O.P. (2016). Transcriptome profiling of *Camelina sativa* to identify genes involved in triacylglycerol biosynthesis and accumulation in the developing seeds. *Biotechnol. Biofuels* 9: 1-19.

Alexander, L.G., Sessions, R.B., Clarke, A.R., Tatham, A.S., Shewry, P.R., and Napier, J.A. (2002). Characterization and modelling of the hydrophobic domain of a sunflower oleosin. *Planta* 214: 546-551.

Aviña-Padilla, K., Ramírez-Rafael, J.A., Herrera-Oropeza, G.E., Muley, V.Y., Valdivia, D.I., Díaz-Valenzuela, E., García-García, A., Varela-Echavarría, A., and Hernández-Rosales, M. (2021). Evolutionary perspective and expression analysis of intronless genes highlight the conservation of their regulatory role. *Front Genet.* 12: 654256.

Aviña-Padilla, K., Ramírez-Rafael, J.A., Zambada-Moreno, O., Herrera-Oropeza, G.E., Romero, G., Gupta, I., and Hernández-Rosales, M. (Year). "Deciphering the tissue-specific regulatory role of intronless genes across cancers", in: *RECOMB International Workshop on Comparative Genomics*: Springer), 311-339.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server issue): W202-208. doi: 10.1093/nar/gkp335.

Bao, W., Ao, D., Wang, L., Ling, Z., Chen, M., Bai, Y., Wuyun, T.-N., Chen, J., Zhang, S., and Li, F. (2021). Dynamic transcriptome analysis identifies genes related to fatty acid biosynthesis in the seeds of *Prunus pedunculata* Pall. *BMC Plant Biol.* 21: 1-16.

Baud, S., and Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Prog. Lipid Res.* 49(3): 235-249.

Chen, J.C., Lin, R.-H., Huang, H.-C., and Tzen, J.T. (1997). Cloning, expression and isoform classification of a minor oleosin in sesame oil bodies. *J Biochem.* 122(4): 819-824.

Chen, K., Yin, Y., Liu, S., Guo, Z., Zhang, K., Liang, Y., Zhang, L., Zhao, W., Chao, H., and Li, M. (2019). Genome-wide identification and functional analysis of oleosin genes in *Brassica napus* L. *BMC Plant Biol.* 19: 1-20.

Chen, Y., Ma, T., Zhang, T., and Ma, L. (2023). Trends in the evolution of intronless genes in Poaceae. *Front Plant Sci.* 14: 1065631.

Dong, Y., Cui, Y., Wang, Y., Luan, S., Liu, X., Yang, Q., Liu, W., Li, X., Wang, N., and Wang, F. (2024). Identification of the *Oleosin* genes and functional analysis of *CeOle4* gene in *Cyperus esculentus* L. *Horticulturae* 10(9): 945.

FAO (2022). "Word food and agriculture. Statistical Yearbook".).

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., and Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44(D1): D279-285. doi: 10.1093/nar/gkv1344.

Goodarzi, H., Hassani, D., Pourhosseini, L., Kalantari, S., and Lashgari, A. (2023). Total lipid and fatty acids components of some Persian walnut (*Juglans regia*) cultivars. *Sci Hort.* 321: 112252.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40(Database issue): D1178-1186. doi: 10.1093/nar/gkr944.

Hao, Y., Ge, X., Xu, R., Zhao, X., and Zhai, M. (2024). Transcriptome analysis of lipid biosynthesis during kernel development in two walnut (*Juglans regia* L.) varieties of 'Xilin 3'and 'Xiangling'. *BMC Plant Biol.* 24(1): 828.

Hashemipetroudi, S.H., Mohammadi, S., and Fatemi, F. (2023). Identification and expression analysis of *HSP100* gene family in *Aeluropus littoralis*. *J Plant Mol Breed.* 11(2): 66-77. doi: 10.22058/jpmb.2024.2023541.1294.

Heidari, P., Sabari, B., and Seifi, A. (2023). Magnesium transporter family: sequence, evolution and expression analysis in soybean (*Glycine max* L.). *J Plant Mol Breed.* 11(1): 62-73. doi: 10.22058/jpmb.2024.2020070.1287.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35(Web Server issue): W585-587. doi: 10.1093/nar/gkm259.

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31(8): 1296-1297. doi: 10.1093/bioinformatics/btu817.

Huang, A. (1996). Oleosins and oil bodies in seeds and other organs. *Plant Physiol.* 110(4): 1055.

Huang, C.-Y., and Huang, A.H. (2017). Unique motifs and length of hairpin in oleosin target the cytosolic side of endoplasmic reticulum and budding lipid droplet. *Plant Physiol.* 174(4): 2248-2260.

Huang, M.-D., and Huang, A.H. (2015). Bioinformatics reveal five lineages of oleosins and the mechanism of lineage evolution related to structure/function from green algae to seed plants. *Plant Physiol.* 169(1): 453-470.

Jackson, C.L. (2019). Lipid droplet biogenesis. *Curr. Opin. Cell Biol.* 59: 88-96.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., and Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9): 1236-1240. doi: 10.1093/bioinformatics/btu031.

Keddie, J.S., Hübner, G., Slocombe, S.P., Jarvis, R.P., Cummins, I., Edwards, E.-w., Shaw, C.H., and Murphy, D.J. (1992). Cloning and characterisation of an oleosin gene from *Brassica napus*. *Plant Mol Biol.* 19: 443-453.

Khor, V.K., Shen, W.-J., and Kraemer, F.B. (2013). Lipid droplet metabolism. *Curr. Opin. Clin. Nutr. Metab. Care.* 16(6): 632-637.

Kim, E., and Hyun, T.K. (2024). Genome-wide identification of oleosin family and expression analysis in response to abiotic stresses in balloon flower (*Platycodon grandiflorus*). *J Plant Biotechnol.* 51(1): 71-76.

Leprince, O., Van Aelst, A., Pritchard, H., and Murphy, D. (1997). Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and-sensitive oilseeds. *Planta* 204: 109-119.

Letunic, I., Doerks, T., and Bork, P. (2015). SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43(Database issue): D257-260. doi: 10.1093/nar/gku949.

Li, S.-S., Wang, L.-S., Shu, Q.-Y., Wu, J., Chen, L.-G., Shao, S., and Yin, D.-D. (2015). Fatty acid composition of developing tree peony (*Paeonia section* Moutan DC.) seeds and transcriptome analysis during seed development. *BMC Genomics* 16: 1-14.

Miquel, M., Trigui, G., d'Andréa, S., Kelemen, Z., Baud, S., Berger, A., Deruyffelaere, C., Trubuil, A., Lepiniec, L., and Dubreucq, B. (2014). Specialization of oleosins in oil body dynamics during seed development in *Arabidopsis*. *Plant Physiol.* 164(4): 1866-1878.

Mohammadi, S., Sohrevardi, F., and Nematzade, G. (2023). Genome-wide analysis of the *HSP90* gene family and their roles in soybean growth and development. *J Plant Mol Breed.* 11(2): 119-132. doi: 10.22058/jpmb.2024.555256.1256.

Nogales-Bueno, J., Baca-Bocanegra, B., Hernández-Hierro, J.M., Garcia, R., Barroso, J.M., Heredia, F.J., and Rato, A.E. (2021). Assessment of total fat and fatty acids in walnuts using near-infrared hyperspectral imaging. *Front. Plant Sci.* 12: 729880.

Parthibane, V., Rajakumari, S., Venkateshwari, V., Iyappan, R., and Rajasekharan, R. (2012). Oleosin is bifunctional enzyme that has both monoacylglycerol acyltransferase and phospholipase activities. *J. Biol. Chem.* 287(3): 1946-1954.

Pyc, M., Gidda, S.K., Seay, D., Esnay, N., Kretzschmar, F.K., Cai, Y., Doner, N.M., Greer, M.S., Hull, J.J., and Coulon, D. (2021). LDIP cooperates with SEIPIN and LDAP to facilitate lipid droplet biogenesis in Arabidopsis. *Plant Cell* 33(9): 3076-3103.

Roux, É., Baumberger, S., Axelos, M.A., and Chardot, T. (2004). Oleosins of *Arabidopsis thaliana*: expression in *Escherichia coli*, purification, and functional properties. *J. Agric. Food Chem.* 52(16): 5245-5249.

Sarmiento, C., Ross, J.H., Herman, E., and Murphy, D.J. (1997). Expression and subcellular targeting of a soybean oleosin in transgenic rapeseed. Implications for the mechanism of oil‐body formation in seeds. *Seeds Plant J.* 11(4): 783-796.

Shamlu, F., Rezaei, M., Lawson, S., Ebrahimi, A., Biabani, A., and Khan-Ahmadi, A. (2018). Genetic diversity of superior Persian walnut genotypes in Azadshahr, Iran. *Physiol. Mol. Biol. Plants.* 24: 939-949.

Shao, Q., Liu, X., Su, T., Ma, C., and Wang, P. (2019). New insights into the role of seed oil body proteins in metabolism and plant development. *Front. Plant Sci.* 10: 1568.

Shimada, T.L., Shimada, T., Takahashi, H., Fukao, Y., and Hara‐Nishimura, I. (2008). A novel role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *Plant J.* 55(5): 798-809.

Siloto, R.M., Findlay, K., Lopez-Villalobos, A., Yeung, E.C., Nykiforuk, C.L., and Moloney, M.M. (2006). The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. *Plant Cell* 18(8): 1961-1974.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12): 2725-2729. doi: 10.1093/molbev/mst197.

Tzen, J.T., Peng, C.-C., Cheng, D.-J., Chen, E.C., and Chiu, J.M. (1997). A new method for seed oil body purification and examination of oil body integrity following germination. *J Exp Bot.* 121(4): 762-768.

Vance, V.B., and Huang, A. (1987). The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. *J Biol Chem.* 262(23): 11275-11279.

Wu, L.S., Wang, L.-D., Chen, P.-W., Chen, L.-J., and Tzen, J.T. (1998). Genomic cloning of 18 kDa oleosin and detection of triacylglycerols and oleosin isoforms in maturing rice and postgerminative seedlings. *J Biochem.* 123(3): 386-391.

Wu, Y.-Y., Chou, Y.-R., Wang, C.-S., Tseng, T.-H., Chen, L.-J., and Tzen, J.T. (2010). Different effects on triacylglycerol packaging to oil bodies in transgenic rice seeds by specifically eliminating one of their two oleosin isoforms. *Plant Physiol Biochem.* 48(2-3): 81-89.

Xiao, Y., Li, M., and Wang, J. (2022). The impacts of allopolyploidization on Methyl-CpG-Binding Domain (MBD) gene family in *Brassica napus*. *BMC Plant Biol.* 22(1): 103.

Yan, S., Wang, X., Yang, C., Wang, J., Wang, Y., Wu, B., Qiao, L., Zhao, J., Mohammad, P., and Zheng, X. (2021). Insights into walnut lipid metabolism from metabolome and transcriptome analysis. *Front Genet.* 12: 715731.

Yang, S., Miao, L., He, J., Zhang, K., Li, Y., and Gai, J. (2019). Dynamic transcriptome changes related to oil accumulation in developing soybean seeds. *Int. J. Mol. Sci.* 20(9): 2202.

Yang, S., Zong, W., Shi, L., Li, R., Ma, Z., Ma, S., Si, J., Wu, Z., Zhai, J., and Ma, Y. (2024). PPGR: a comprehensive perennial plant genomes and regulation database. *Nucl Acids Res.* 52(D1): D1588-D1596.

Yang, Y., and Benning, C. (2018). Functions of triacylglycerols during plant development and stress. *Curr Opin Biotechnol.* 49: 191-198.

Yuan, Y., Cao, X., Zhang, H., Liu, C., Zhang, Y., Song, X.-L., and Gai, S. (2021). Genome-wide identification and analysis of *Oleosin* gene family in four cotton species and its involvement in oil accumulation and germination. *BMC Plant Biol.* 21: 1-18.

Zhang, Y., Shen, Q., Leng, L., Zhang, D., Chen, S., Shi, Y., Ning, Z., and Chen, S. (2021). Incipient diploidization of the medicinal plant *Perilla* within 10,000 years. *Nature Comm.* 12(1): 5508.

Zhi, Y., Taylor, M.C., Campbell, P.M., Warden, A.C., Shrestha, P., El Tahchy, A., Rolland, V., Vanhercke, T., Petrie, J.R., and White, R.G. (2017). Comparative lipidomics and proteomics of lipid droplets in the mesocarp and seed tissues of Chinese tallow (*Triadica sebifera*). *Front. Plant Sci.* 8: 1339.

Zou, Z., Zhang, L., and Zhao, Y. (2024). Integrative analysis of *oleosin* genes provides insights into lineage-specific family evolution in Brassicales. *Plants* 13(2): 280.

Zou, Z., Zhao, Y., and Zhang, L. (2022). Genomic insights into lineage-specific evolution of the *oleosin* family in Euphorbiaceae. *BMC Genomics* 23(1): 178.

JPMB
Journal of Plant Molecular Breeding
Published by Genetics and Agricultural Biotechnology Institute of Tabarestan
2024 | Volume 12 | Issue 1

مقاله پژوهشی

شناسایی و آنالیز بیان خانواده ژنی اولئوسین در گردو (*Juglans regia* L.)

سیدحمیدرضا هاشمی‌پطرودی[*1]**، فاطمه احمدی**[2]

[1] پژوهشکده ژنتیک و زیست فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
[2] گروه مهندسی باغبانی و فضای سبز، دانشکده تولید گیاهی، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران

چکیده: اولئوسین‌ها به عنوان پروتئین‌های ساختاری نقش مهمی در تثبیت اجسام روغنی، اندازه، تعداد و محتوای آنها دارند. در مطالعه حاضر، شناسایی ساختار اگزون/اینترون، موتیف‌ها/دامنه‌های حفاظت‌شده و الگوهای بیان خانواده ژن اولئوسین در گیاه دیپلوئید *Juglans regia* L. مورد توجه قرار گرفت. در این تحقیق هفت ژن غیر تکراری *JrOleosin* در تجزیه و تحلیل گسترده ژنومی شناسایی گردید که تفاوت‌های ساختاری و فیزیکوشیمیایی قابل توجهی را نشان دادند. تجزیه و تحلیل فیلوژنتیکی ۴۶ ژن اولئوسین نشان داد که این ژن‌ها از قرابت ژنتیکی برخوردار بوده، در عین دارا بودن توالی حفاظت‌شده مشترک، از واگرایی‌های منحصربه‌فردی در گونه‌های مختلف برخورداند. ژن‌های بدون اینترون (IGs) *JrOleosin* ممکن است نمونه‌هایی از چگونگی سازگاری تکامل برای افزایش دقت و اثربخشی بیان ژن باشند. تجزیه و تحلیل رونوشت نشان داد که ژن‌های *JrOleosin* به صورت بافت/زمان- اختصاصی با بالاترین رونوشت یافت شده در دانه‌ها و میوه‌ها بیان می‌شوند. این امر می‌تواند بیانگر نقش کلیدی این ژن‌ها در ذخیره‌سازی لیپیدها در این مراحل رشدی هستند. یافته‌های این تحقیق درک ما از خصوصیات خانواده ژن *JrOleosins* به ویژه ساختار و بیان آنها را عمیق‌تر می‌نماید. پیشنهاد می‌شود در مطالعات آتی چگونگی تنظیم بیان اولئوسین و نحوه تعامل این پروتئین‌ها با سایر پروتئین‌های مرتبط با تولید لیپید در مرحله نمو بذر مدنظر قرار گیرد.

کلمات کلیدی: اجسام روغنی، اولئوسین، تجزیه تحلیل فراداده، گردو، گستره ژنومی.

ویراستار علمی
دکتر لیلا آهنگر،
گروه تولیدات گیاهی، دانشکده کشاورزی و منابع طبیعی، دانشگاه گنبدکاووس، ایران

تاریخ
دریافت: ۵ آذر ۱۴۰۳
پذیرش: ۲۲ بهمن ۱۴۰۳
چاپ: ۲۴ بهمن ۱۴۰۳

نویسنده مسئول
دکتر سیدحمیدرضا هاشمی‌پطرودی
shr.hashemi@sanru.ac.ir;
irahamidreza@yahoo.com

ارجاع به این مقاله
Hashemipetroudi, S.H., Ahmadi, F. (2024). Identification of *oleosin* gene family in *Juglans regia* L. *J Plant Mol Breed.* 12(1): 150-165. doi:10.22058/jpmb.2025.2029871.1317.