International Journal of Life Sciences and Biotechnology
Unal, N., Genome-Wide Identification, Evolutionary Analysis and Comprehensive *In Silico* Characterization of the GPAT Gene Family in Sunflower (*Helianthus annuus* L.), 2026 9(2): p. 106-116.
https://doi.org/10.38001/ijlsb.1966973
Araştırma Makale/ Research article

Genome-Wide Identification, Evolutionary Analysis and Comprehensive In Silico Characterization of the *GPAT* Gene Family in Sunflower (*Helianthus annuus* L.)

Neslihan Ünal[1*]

ABSTRACT
Glycerol-3-phosphate acyltransferase (GPAT) are enzymes involved in glycerolipid biosynthesis and play a key role in plant growth, development, and abiotic stress responses. However, a comprehensive genome-wide and in silico analysis of the *GPAT* gene family in sunflower (*Helianthus annuus* L.) has not been performed to date. In this study, the *GPAT* gene family in the sunflower genome was characterized using bioinformatics approaches. A total of 23 *HaGPAT* genes were identified; their chromosomal distributions, phylogenetic relationships, gene structures, conserved motifs, protein properties, subcellular localizations, cis-regulatory elements, miRNA targets, protein-protein interaction networks, three-dimensional protein structures, synteny relationships, duplication events, Ka/Ks ratios, and expression profiles based on RNA-seq data were analyzed. Expression analyses based on RNA-seq data showed that *HaGPAT* genes exhibited variable expression patterns across different tissues. The findings of this study contribute to a better understanding of the structural features, evolutionary relationships, and expression profiles of the *HaGPAT* gene family, and constitute a valuable genomic resource for future functional studies.

ARTICLE HISTORY
Received
09 June 2026
Accepted
02 August 2026

KEY WORDS
Helianthus annuus,
GPAT,
genome-wide analysis,
RNA-seq

Introduction

Lipids are essential components of cell membranes and play a vital role in plant growth, development, and stress responses, as well as providing energy for cellular metabolism [1]. Glycerol-3-phosphate acyltransferase (GPAT; EC 2.3.1.15) is one of the key enzymes that catalyzes the initial reaction in the biosynthesis of phosphatidic acid, a common precursor of membrane glycolipids and triacylglycerols [2, 3]. In plant cells, GPAT enzymes are localized in plastids, cytoplasm, and mitochondria, and the activity of this enzyme was first demonstrated in the microsomal fraction obtained from the mesocarp of avocado (*Persea americana* Mill.) [4].

Members of the *GPAT* gene family have been identified in various plant species, primarily *Arabidopsis thaliana*, and other eukaryotic organisms through genome-wide analyses [1, 5, 6, 7, 8, 9]. Studies have shown that *GPAT* genes play important roles in plant development and physiology. These genes are reported to be particularly associated with reproductive success, seed development, oil accumulation, and tolerance mechanisms to environmental stresses [6]. Ten *GPAT* genes identified in the *Arabidopsis thaliana* genome have been determined to be involved in different biological processes. For example, *AtATS1* is involved in the synthesis of basic membrane phospholipids [10], while *AtGPAT1* is one of the genes necessary for pollen development and fertility [5]. Recent studies have shown that GPAT enzymes also play a role in the adaptation of plants to environmental conditions. In Arabidopsis, it has been reported that mitochondrial *GPAT1* and *GPAT2* play a role in cutin synthesis and secondary cell wall formation, thereby contributing to the maintenance of the plant's water balance [11]. *AtGPAT4, AtGPAT5*, and *AtGPAT8* genes have been reported to be involved in cutin and suberin biosynthesis [12]. Furthermore, it has been determined that *GPAT4* and *GPAT8* genes are expressed in association with suberin accumulation in the root endoderm [13]. These findings demonstrate that the *GPAT* gene family has multifaceted functions in the regulation of lipid metabolism, plant growth, and stress responses. Furthermore, it has been reported that *GPAT* genes exhibit different expression profiles under salinity, drought, low temperature, and other environmental stress conditions [6, 1]. Sunflower (*Helianthus annuus* L.) is an important oilseed crop widely cultivated worldwide due to its high-quality oil content and nutritional value. Yield and crop quality are directly affected by the plant's growth and

[1] Department of Biology, Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Türkiye
*Corresponding Autor: Neslihan Ünal, e-mail: neslihanunal.n@gmail.com

development processes [14]. However, the genome-wide organization, evolutionary relationships, and regulatory mechanisms of the *GPAT* gene family in sunflower have not yet been comprehensively investigated. This study aims to identify and comprehensively characterize the *GPAT* gene family in the sunflower genome at the genome-wide level. In this study, the phylogenetic relationships, gene structures, conserved motifs, protein characteristics, cis-regulatory elements, duplication events, synteny relationships, miRNA targets, protein-protein interaction networks, three-dimensional protein structures, and RNA-seq-based expression profiles of *HaGPAT* genes were analyzed. The aim was to contribute to a better understanding of the structural and evolutionary characteristics of the *HaGPAT* gene family and to provide a basis for future functional studies.

Material and Methods

Identification and Characterization of HaGPAT Proteins

Protein and coding DNA sequence (CDS) sequences of *Helianthus annuus* (assembly r1.2) were obtained from the Phytozome v13 database (https://phytozome-next.jgi.doe.gov/) [15]. To identify members of the *GPAT* gene family, the PF01553 Hidden Markov Model (HMM) profile corresponding to the PlsC (phosphate acyltransferase) domain was obtained from the Pfam database (http://pfam.xfam.org/) [16]. HMMER v3.4 (http://hmmer.org/) was used for HMM profile searches. In addition, Basic Local Alignment Search Tool for Proteins (BLASTp) analysis was performed using *Arabidopsis thaliana* GPAT proteins, and candidate sequences were selected based on an E-value ≤1e−5 and a minimum sequence identity of 50% [17, 18]. The obtained proteins were validated using the NCBI Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/), and sequences carrying the PF01553 domain were accepted as HaGPAT [19]. The basic physicochemical properties of the proteins were determined using ExPASy ProtParam (https://web.expasy.org/protparam/) [20].

Gene Structure and Conserved Motif Analyses

The exon-intron organization of *HaGPAT* genes was investigated using GFF (Gene Feature Format) annotation files obtained from the Phytozome v13 database [15]. Conserved motifs were determined using the Multiple Expectation Maximization for Motif Elicitation (MEME) Suite v5.5.9 (https://meme-suite.org/meme/tools/meme) with the Zero or One Occurence Per Sequence (ZOOPS) model. A maximum of 10 motifs with widths of 6–50 amino acids was used [21]. The obtained gene structure and motif information were visualized together using the TBtools program [22].

Phylogenetic Analysis of HaGPAT Proteins

In the phylogenetic analysis, GPAT protein sequences of *Helianthus annuus, Arabidopsis thaliana, Medicago truncatula* Gaertner*., and Phaseolus vulgaris* L. species were used. Multiple sequence alignment was performed using the ClustalW algorithm with the Molecular Evolutionary Genetics Analysis (MEGA) v11.0.10 program (https://www.megasoftware.net/) [23]. The phylogenetic tree was constructed using the Maximum Likelihood (ML) method with the Jones-Taylor-Thornton (JTT) substitution model and 1000 bootstrap replicates, while other parameters were left at their default settings. The resulting tree was visualized on the iTOL (Interactive Tree Of Life) platform (https://itol.embl.de/) [24].

Chromosomal Distribution, Synteny, and Duplication Analyses

Genome sequence and GFF annotation data used to determine the positions of *HaGPAT* genes on the genome were obtained from the Phytozome JGI Data Portal (https://data.jgi.doe.gov/). The chromosomal distribution of genes was visualized using the TBtools program [22]. Gene duplications and syntenic relationships were analyzed with Multiple Collinearity Scan Toolkit (MCScanX) [25], and syntenic relationships between *Helianthus annuus* and *Arabidopsis thaliana* were examined using TBtools. Ka/Ks (non-synonymous/synonymous substitution ratio) values were calculated using the Simple KaKs_Calculator 2.0 in TBtools with the Nei–Gojobori (NG) method [22]. Duplication times were estimated in millions of years ago (MYA) using the formula $T = Ks/(2\lambda)$, with $\lambda = 6.56 \times 10^{-9}$ [26].

Promoter Region and miRNA Target Analyses

To identify cis-regulatory elements in the promoter regions of *HaGPAT* genes, sequences 2000 bp upstream of the start codon (ATG) were obtained from the Phytozome v13 database. Cis-elements were identified using PlantCARE (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [27] and grouped into groups related to hormone, stress, light, and growth and development. The results were visualized using the TBtools program. miRNA sequences of *Helianthus annuus* were obtained from the PmiREN v22.1 database (http://www.pmiren.com/). miRNAs targeting *HaGPAT* genes were predicted using psRNATarget (https://www.zhaolab.org/psRNATarget/) [28]. miRNA target prediction was performed using psRNATarget with a maximum expectation score of 5.0. miRNA-gene interaction networks were generated using Cytoscape (https://cytoscape.org/download.html) [29].

Protein-Protein Interaction (PPI) and 3D Structure Analysis

Possible protein-protein interactions between HaGPAT proteins were investigated using the STRING v12.0 database (https://string-db.org/) [30]. A confidence score of 0.900 was used in the analyses, and the number of interactions displayed in the network was limited to 10. The three-dimensional (3D) structures of HaGPAT proteins were modeled using the Phyre2 server (https://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index), and predictions with a confidence value above 90% were considered [31].

Tissue-Specific RNA-seq Expression Profiles of *HaGPAT* Genes

To examine the expression profiles of *HaGPAT* genes, RNA-Seq expression as Reads Per Kilobase of transcript per Million mapped reads (RPKM) values data from *Helianthus annuus* were obtained from the published NCBI Sequence Read Archive (SRA) database (Study accession: SRP092742) (https://www.ncbi.nlm.nih.gov/sra). Tissue-specific expression analyses were performed using samples from stamen (SRR4996809), ray floret ovary (SRR4996808), pistil (SRR4996814), corolla (SRR4996833), bract (SRR4996800), disc floret ovary (SRR4996831), ligule (SRR4996822), and stem (SRR4996799) [32]. According to the original RNA-seq study, each tissue sample consisted of three biological replicates. RPKM values for *HaGPAT* genes were subjected to $\log_2$(RPKM+1) transformation [33], and the results were visualized as a heatmap using the TBtools program.

Results

Identification and Characterization of HaGPAT Proteins

Scanning of the *Helianthus annuus* genome revealed 23 members of the *GPAT* gene family, numbered from *HaGPAT1* to *HaGPAT23*. Physicochemical analyses showed that HaGPAT proteins have lengths ranging from 271 to 572 amino acids and molecular weights between 31.02 and 64.84 kDa (Table 2). Theoretical isoelectric point (pI) values of the proteins ranged from 6.34 to 9.88. Instability index results revealed that the vast majority of proteins had values below 40, thus exhibiting stable characteristics. Furthermore, Grand Average of Hydropathicity (GRAVY) values showed that some HaGPAT proteins exhibited hydrophobic properties, while others were hydrophilic (Table 2). Subcellular localization prediction suggested that most HaGPAT proteins are localized in plasma membrane and endoplasmic reticulum (ER). Furthermore, it was noted that HaGPAT7 was mostly localized in the chloroplast (12), while HaGPAT23 had the highest score (5) equally for the cytoplasm and ER (Table 1).

Table 1. Predicted subcellular localization of HaGPAT proteins

Protein ID	Subcellular localization
HaGPAT1	plas: 4, E.R.: 4, cyto: 3, chlo: 1, nucl: 1, vacu: 1
HaGPAT2	E.R.: 4, cyto: 3.5, cyto_nucl: 3, chlo: 2, nucl: 1.5, mito: 1, plas: 1, pero: 1
HaGPAT3	E.R.: 4, chlo: 3, cyto_nucl: 3, nucl: 2.5, cyto: 2.5, mito: 1, plas: 1
HaGPAT4	plas: 4, vacu: 4, E.R.: 3, cyto: 1, extr: 1, golg: 1
HaGPAT5	plas: 7, vacu: 4, extr: 2, E.R.: 1
HaGPAT6	plas: 7, E.R.: 4, vacu: 2, extr: 1
HaGPAT7	chlo: 12, cyto: 1, extr: 1
HaGPAT8	plas: 8, E.R.: 4, golg: 2
HaGPAT9	plas: 9, E.R.: 2, mito: 1, vacu: 1, pero: 1
HaGPAT10	plas: 5, E.R.: 4, golg: 2, chlo: 1, vacu: 1, pero: 1
HaGPAT11	plas: 7, cyto: 3, E.R.: 3, nucl: 1
HaGPAT12	plas: 9, E.R.: 3, cyto: 1, vacu: 1
HaGPAT13	plas: 7, vacu: 3, E.R.: 3, extr: 1
HaGPAT14	plas: 7, E.R.: 2, golg: 2, cyto: 1, extr: 1, vacu: 1
HaGPAT15	chlo: 4, vacu: 4, mito: 2, plas: 2, extr: 2
HaGPAT16	plas: 8, E.R.: 4, extr: 1, vacu: 1
HaGPAT17	plas: 9, E.R.: 3, cyto: 1, vacu: 1
HaGPAT18	plas: 7, E.R.: 3, vacu: 2, cyto: 1, golg: 1
HaGPAT19	plas: 6, vacu: 3, chlo: 2, mito: 2, extr: 1
HaGPAT20	plas: 4, E.R.: 3, chlo: 2, cyto: 2, nucl: 1, mito: 1, vacu: 1
HaGPAT21	plas: 7, vacu: 3, E.R.: 2, cyto: 1, extr: 1
HaGPAT22	plas: 5, mito: 3, vacu: 2, E.R.: 2, chlo: 1, extr: 1
HaGPAT23	cyto: 5, E.R.: 5, chlo: 1, nucl: 1, mito: 1, vacu: 1

*Abbreviations: plas, plasma membrane; E.R., endoplasmic reticulum; cyto, cytoplasm; cyto_nucl, cytoplasm/nucleus; chlo, chloroplast; nucl, nucleus; vacu, vacuole; mito, mitochondrion; pero, peroxisome; extr, extracellular; golg, golgi apparatus)

Gene Structure and Conserved Motifs

The phylogenetic relationships, conserved motif distributions, and exon-intron organizations of the *HaGPAT* gene family were evaluated together (Figure 1). Genes belonging to the same phylogenetic group exhibited similar motif compositions (Figure 1a). In particular, 10 conserved motifs were largely conserved in the HaGPAT4, HaGPAT5, HaGPAT6, HaGPAT9, HaGPAT11, HaGPAT13, HaGPAT14, HaGPAT16, HaGPAT18, HaGPAT19, HaGPAT20, and HaGPAT21 proteins. In contrast, HaGPAT1, HaGPAT2, HaGPAT3, HaGPAT7, HaGPAT8, HaGPAT12, HaGPAT15, HaGPAT17, HaGPAT22, and HaGPAT23 proteins contained fewer motifs (Figure 1b). According to the gene structure analysis results, there were differences in the number of exons and introns among the *HaGPAT* genes, and it was determined that the *HaGPAT11* and *HaGPAT18* genes covered longer genomic regions compared to other members of the family (Figure 1c). In general, it has been determined that phylogenetically closely related genes exhibit similar motif structures and exon-intron organizations.

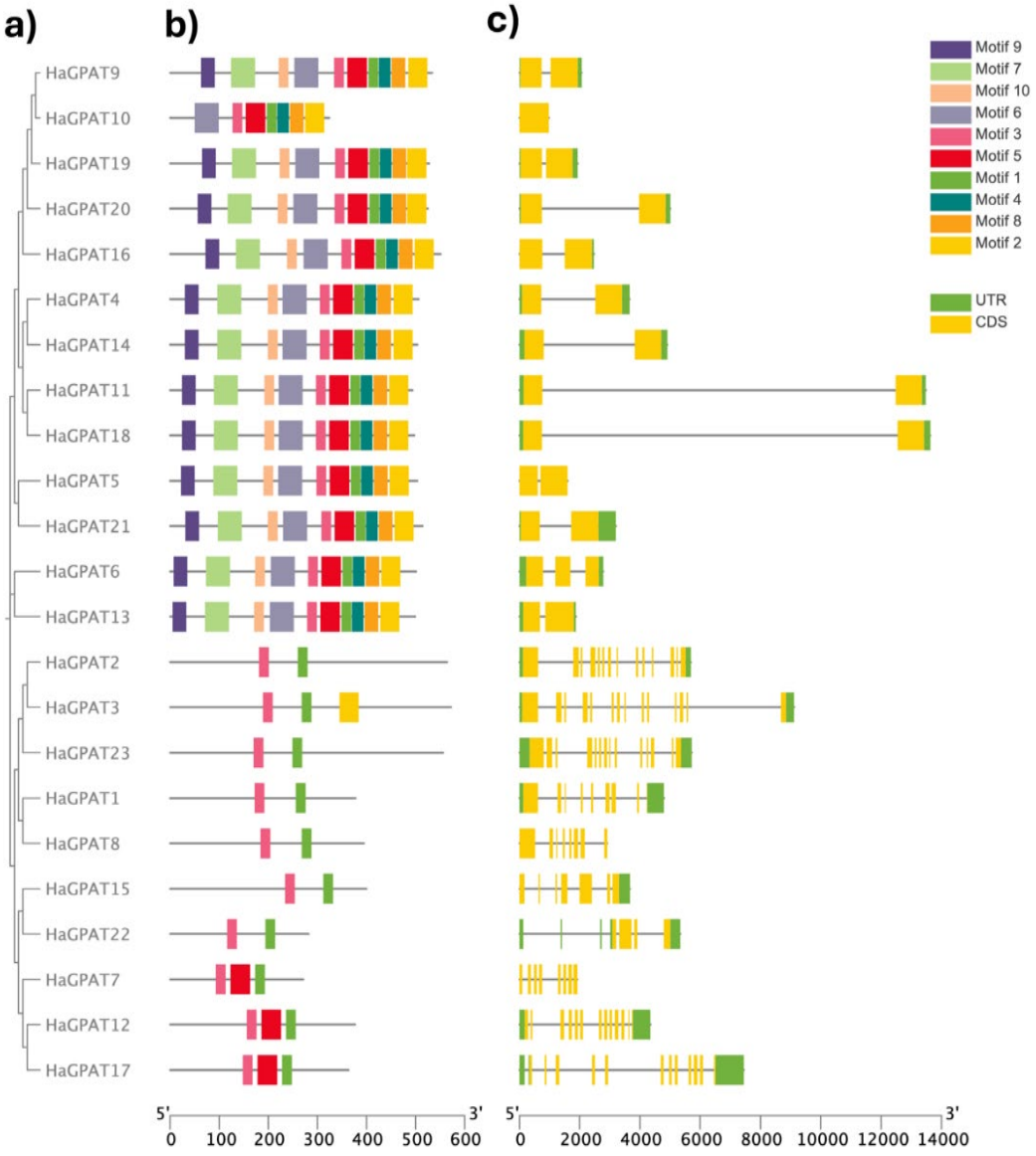

Fig 1. Phylogenetic relationships, conserved motifs, and gene structures of *HaGPAT* genes. (a) Phylogenetic tree of HaGPAT proteins. (b) Conserved motif distribution in HaGPAT proteins. (c) Exon–intron organisation of *HaGPAT* genes, showing coding sequences (CDS), untranslated regions (UTRs), and introns

Table 2. Structural features and physicochemical properties of *HaGPAT* proteins, Abbreviations: Chr, chromosome number; kDa, kilodalton; P_I, isoelectric point; GRAVY, grand average of hydropathicity

Gene ID	Gene name	Chr	Gene start position	Gene end Position	Protein length	Molecular weight (kDa)	pI	Instability index	GRAVY	Aliphatic index	Stability
HanXRQChr02g0040241	*HaGPAT1*	02	54831836	54836630	378	42.53	6.42	55.50	-0.103	92.75	unstable
HanXRQChr04g0122151	*HaGPAT2*	04	161373233	161378923	564	63.78	6.70	43.29	-0.024	88.87	unstable
HanXRQChr06g0169291	*HaGPAT3*	06	13900608	13909721	572	64.84	6.75	42.60	-0.115	86.80	unstable
HanXRQChr06g0169981	*HaGPAT4*	06	15971169	15974827	506	56.17	9.29	37.44	0.112	100.14	stable
HanXRQChr08g0209131	*HaGPAT5*	08	3241901	3243499	504	55.27	8.89	37.55	0.160	94.40	stable
HanXRQChr08g0220991	*HaGPAT6*	08	47422250	47425033	501	55.58	9.24	35.93	0.293	101.86	stable
HanXRQChr08g0236381	*HaGPAT7*	08	146725721	146727643	271	31.02	8.84	37.70	0.155	99.19	stable
HanXRQChr09g0268531	*HaGPAT8*	09	182970551	182973464	395	44.85	8.43	55.44	-0.287	86.84	unstable
HanXRQChr10g0284681	*HaGPAT9*	10	36354795	36356864	534	61.28	9.02	40.27	-0.034	96.33	unstable
HanXRQChr10g0284711	*HaGPAT10*	10	36740644	36741618	324	37.20	9.03	34.45	-0.010	94.75	stable
HanXRQChr10g0301051	*HaGPAT11*	10	160740769	160754251	494	55.16	9.20	39.12	0.128	98.42	stable
HanXRQChr11g0335221	*HaGPAT12*	11	72130573	72134916	377	43.12	9.01	45.89	-0.013	98.22	unstable
HanXRQChr12g0369551	*HaGPAT13*	12	61269302	61271175	499	55.20	9.11	41.05	0.292	101.32	unstable
HanXRQChr13g0426601	*HaGPAT14*	13	197124131	197129040	504	56.12	9.26	35.70	0.086	99.58	stable
HanXRQChr14g0448981	*HaGPAT15*	14	136858916	136862585	400	45.17	9.88	39.47	-0.022	88.15	stable
HanXRQChr14g0455271	*HaGPAT16*	14	156565923	156568393	551	62.13	8.84	33.59	0.096	96.53	stable
HanXRQChr14g0458991	*HaGPAT17*	14	165578860	165586308	363	41.72	8.76	38.73	-0.012	97.96	stable
HanXRQChr15g0464011	*HaGPAT18*	15	2690070	2703704	497	55.63	9.31	39.79	0.064	100.20	stable
HanXRQChr15g0465081	*HaGPAT19*	15	5186683	5188614	528	60.14	9.31	36.50	0.008	96.50	stable
HanXRQChr15g0482551	*HaGPAT20*	15	71979459	71984469	525	60.18	9.20	38.94	0.112	96.88	stable
HanXRQChr15g0488511	*HaGPAT21*	15	106654508	106657706	514	56.31	8.99	38.39	0.154	94.14	stable
HanXRQChr15g0490751	*HaGPAT22*	15	121343052	121348388	282	31.25	9.82	40.45	0.286	99.86	unstable
HanXRQChr16g0523541	*HaGPAT23*	16	150108280	150113999	556	62.82	6.34	39.18	-0.090	84.98	stable

Phylogenetic Tree

Phylogenetic tree constructed using GPAT proteins from *Helianthus annuus, Arabidopsis thaliana, Phaseolus vulgaris*, and *Medicago truncatula* species showed that the GPAT proteins were divided into three main groups (Figure 2). HaGPAT proteins are represented in all of these groups, with 2 members in Group I, 8 in Group II, and 13 in Group III. Group III has the most HaGPAT proteins, while Group I has the fewest members (Figure 2). Furthermore, in all three groups, the proteins of *Helianthus annuus*, *Arabidopsis thaliana*, *Medicago truncatula*, and *Phaseolus vulgaris* were observed to cluster together.

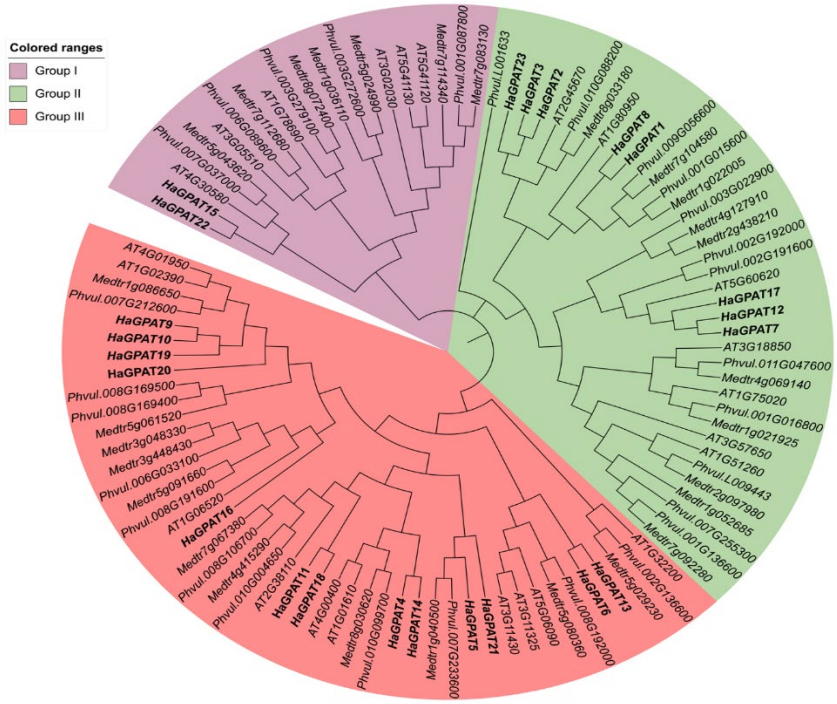

Fig 2. Phylogenetic relationships of GPAT proteins from *Helianthus annuus, Arabidopsis thaliana, Medicago truncatula,* and *Phaseolus vulgaris.* Multiple sequence alignment was performed using ClustalW, and the phylogenetic tree was constructed using the Maximum Likelihood (ML) method with 1000 bootstrap replicates in MEGA. The tree was visualized using iTOL. Different colors indicate the three major phylogenetic groups

Chromosomal Distribution, Duplication and Synteny Analyses

According to chromosomal localization analysis, the 23 *HaGPAT* genes are irregularly located on 12 chromosomes in the *Helianthus annuus* genome. The highest gene density was observed on chromosome Chr15 with five genes. In contrast, only one *HaGPAT* gene was found on chromosomes Chr02, Chr04, Chr09, Chr11, Chr12, Chr13, and Chr16 (Figure 3). These results suggests that the *HaGPAT* gene family exhibits an uneven distribution across chromosomes.

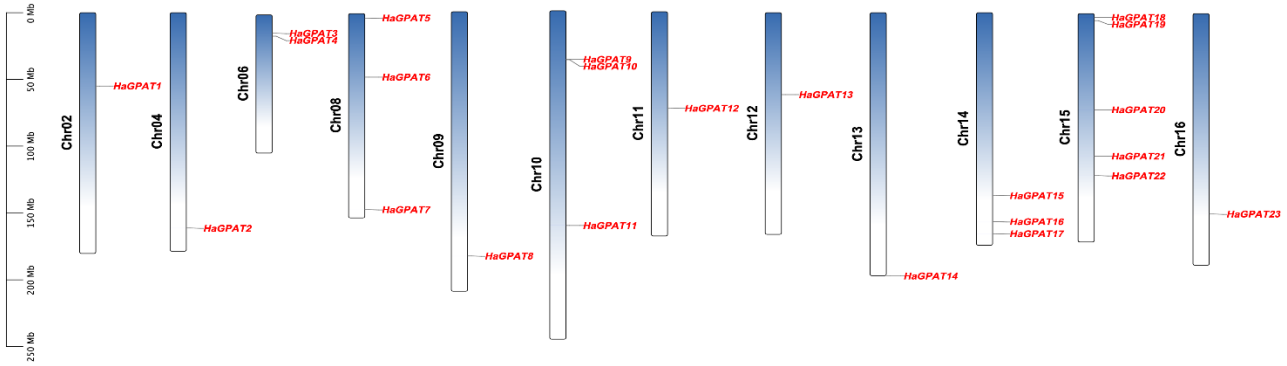

Fig 3. Chromosomal distribution of *HaGPAT* genes in the sunflower genome

Duplication and synteny relationships of the *HaGPAT* gene family were analyzed. Analysis within the same genome revealed segmental duplication gene pairs between *HaGPAT2-HaGPAT3*, *HaGPAT3-HaGPAT23*,

and *HaGPAT6-HaGPAT13*. These duplications were found to occur between chromosomes Chr04–Chr06, Chr06–Chr16, and Chr08–Chr12 (Figure 4).

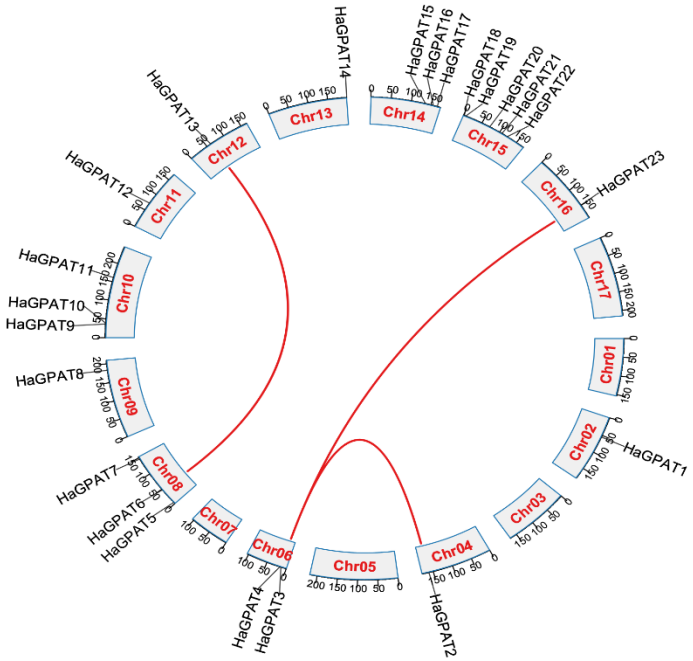

Fig 4. Chromosomal distribution and segmental duplication relationships of *HaGPAT* genes in the sunflower genome. The positions of *HaGPAT* genes on sunflower chromosomes are shown around the circle. Red connecting lines indicate segmentally duplicated gene pairs identified within the *HaGPAT* gene family. The Figure was visualized using TBtools

Comparative synteny analysis performed between the genomes of *Helianthus annuus* and *Arabidopsis thaliana* revealed that four *HaGPAT* genes exhibit collinear relationships with Arabidopsis orthologs (Figure 5). *HaGPAT2-AT2G45670, HaGPAT3-AT2G45670, HaGPAT4-AT4G00400, HaGPAT14-AT4G00400,* and *HaGPAT4-AT1G01610* showed syntenic relationships. These results suggest that *GPAT* genes may have been conserved between the two species throughout evolution.

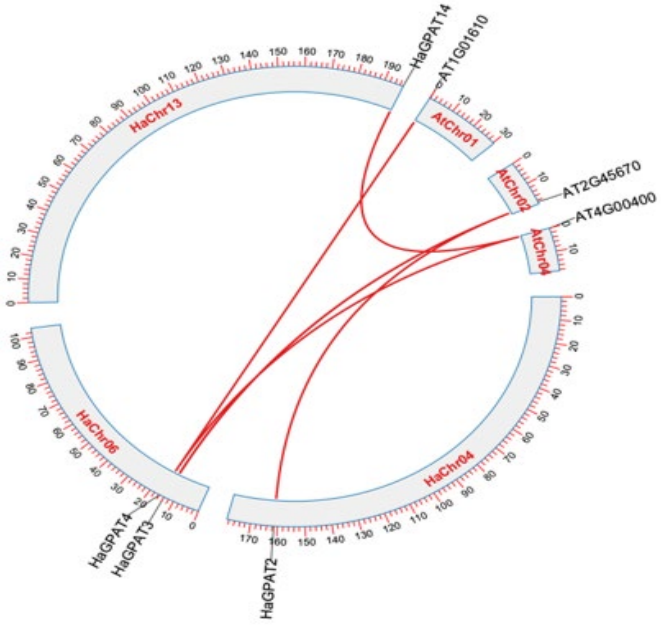

Fig 5. Comparative synteny analysis of *GPAT* genes between *Helianthus annuus* and *Arabidopsis thaliana*. Chromosomes of sunflower and Arabidopsis are displayed on the circular map. Red connecting lines represent syntenic *GPAT* gene pairs identified between the two species. The synteny map was generated using TBtools

In the orthologous gene pairs identified between *Helianthus annuus* and *Arabidopsis thaliana*, Ka/Ks values ranged from 0.0634 to 0.1747, and the estimated separation times of these gene pairs were calculated as 141.22–224.47 MYA. In the segmental duplication pairs between *HaGPAT* genes, Ka/Ks values were found to be between 0.1165 and 0.8816, and it was estimated that the duplication events occurred approximately between 39.85 and 67.20 MYA (Table 3). The fact that Ka/Ks < 1 in all gene pairs suggests that *HaGPAT* genes may have been conserved under negative (purifying) selection pressure throughout the evolutionary process.

Table 3. Ka/Ks based evolutionary analysis and divergence times MYA of *HaGPAT* gene pairs

Gene Pair	Ka	Ks	Ka/Ks	Selection Type	MYA
HaGPAT2 – AT2G45670	0.3238	1.8528	0.1747	Purifying selection	141.22
HaGPAT3 – AT2G45670	0.3302	2.1606	0.1528	Purifying selection	164.68
HaGPAT4 – AT4G00400	0.1951	2.5462	0.0766	Purifying selection	194.07
HaGPAT4 – AT1G01610	0.1866	2.9451	0.0634	Purifying selection	224.47
HaGPAT14 –AT4G00400	0.1839	NaN	NaN	-	-
HaGPAT2 – HaGPAT3	0.1335	0.5228	0.2553	Purifying selection	39.85
HaGPAT3 – HaGPAT23	0.1882	0.8816	0.2135	Purifying selection	67.20
HaGPAT6 – HaGPAT13	0.0633	0.5432	0.1165	Purifying selection	41.40

Promoter Regulatory Elements and miRNA Prediction

Promoter analyses revealed a wide variety of cis-regulatory elements in *HaGPAT* genes associated with growth and development, hormone response, light response, and environmental stresses. The identified elements include regulatory regions such as MYC, MYB, Anaerobic Responsive Element (ARE), Light-sensitive element (G-box), Abscisic Acid Responsive Element (ABRE), Light-sensitive element (Box 4) and CAT-box (Figure 6, Supplementary Table 1). These results indicate that *HaGPAT* genes may be regulated by light signals, hormone-mediated regulatory mechanisms, growth and development processes, and various abiotic stress conditions.

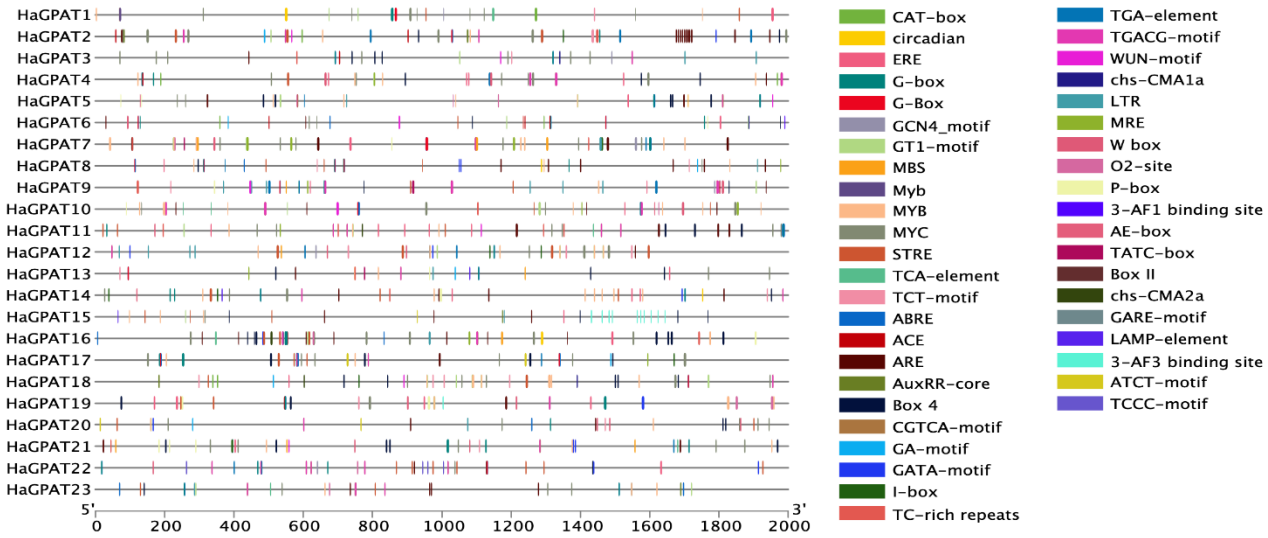

Fig 6. Cis-acting regulatory elements in the promoter regions of *HaGPAT* genes. The 2000 bp upstream promoter sequences of *HaGPAT* genes were analyzed using PlantCARE. Cis-acting elements associated with light responsiveness, growth and development, hormone response, and stress response were identified and visualized using TBtools

miRNA target analysis revealed that 15 *HaGPAT* genes were targeted by a total of 27 different miRNAs (Figure 7, Supplementary Table 2). In this study, several miRNAs targeting *HaGPAT* genes and associated with plant growth and development, hormone signaling, and abiotic stress responses were identified. Furthermore, HaGPAT16 and *HaGPAT1* genes had the most miRNA interactions, being targeted by five and

four different miRNAs, respectively. It was found that Han-miR160 targeted both *HaGPAT6* and *HaGPAT13* genes, while Han-miRN5691 targeted *HaGPAT7* and *HaGPAT12* genes (Figure 7). These results suggest that *HaGPAT* genes are regulated by numerous miRNAs.

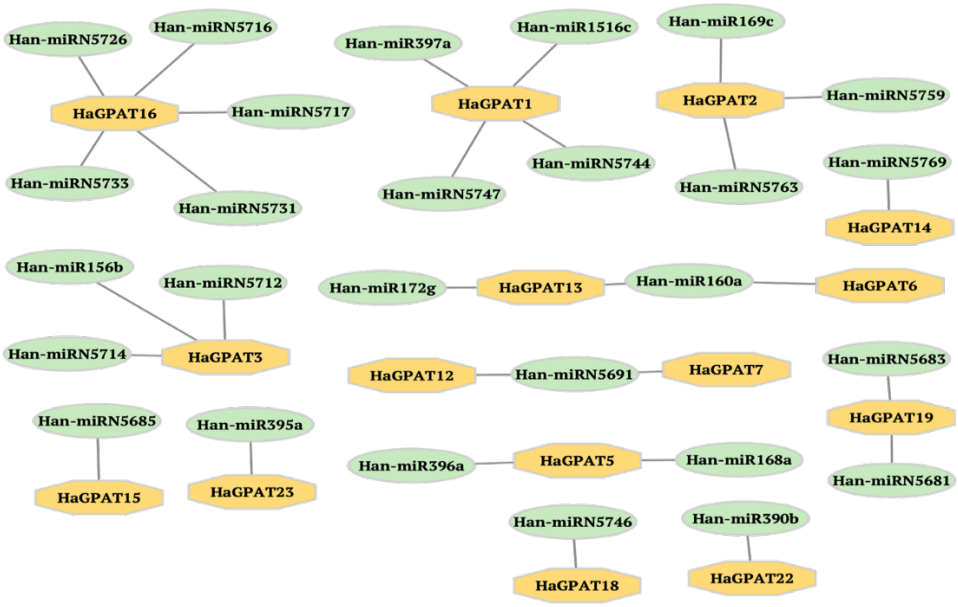

Fig 7. miRNA interaction network of *HaGPAT* genes. miRNAs targeting *HaGPAT* genes were predicted using psRNATarget, and the interaction network was visualized in Cytoscape. Yellow hexagons represent *HaGPAT* genes, and green ellipses represent the corresponding miRNAs. Lines represent the predicted interactions between the miRNA and the target gene

Protein-Protein Interaction (PPI) and 3D Structure Analysis

PPI analysis suggest that HaGPAT proteins form different interaction clusters (Figure 8, Table 4). HaGPAT2, HaGPAT3, HaGPAT8, and HaGPAT23 proteins were predicted to interact extensively with Lyso-PhosphatidylCholine (PC) AcylTransferase 1 (LPCAT1), Lysocardiolipin Acyltransferase 1 (LCLT1), Phosphatidyl-N-methylethanolamine N-methyltransferase (ATPLMT), 1-acylglycerophosphocholine O-acyltransferase **(**A0A251VHS5), phospholipase A2 (A0A251FVE9), phospholipase A2 **(**A0A251SFU6), and 1-acylglycerophosphocholine O-acyltransferase (A0A251RKJ8) proteins. In additional, HaGPAT6, HaGPAT9, HaGPAT11, HaGPAT12, HaGPAT13, HaGPAT16, HaGPAT4, HaGPAT14, HaGPAT17, HaGPAT18, HaGPAT10, HaGPAT19, and HaGPAT20 proteins formed a separate interaction network with glycerol-1-phosphatase **(**GS1), glycerol kinase (NHO1), and Manganese-dependent ADP-ribose/CDP-alcohol diphosphatase (A0A251U166) proteins. However, it was observed that HaGPAT1, HaGPAT5, HaGPAT7, HaGPAT15, HaGPAT21, and HaGPAT22 proteins did not show any interaction with other proteins in the network and were found as isolated nodes (Figure 8). Furthermore, according to the STRING analysis results, the majority of proteins showing strong interaction with HaGPAT proteins were determined to be associated with acyltransferase, phospholipase, and methyltransferase activities. Specifically, LPCAT1, LCLT1, A0A251VHS5, and A0A251RKJ8 proteins were identified as belonging to the acyltransferase family, while A0A251V6E9 and A0A251SFU6 proteins belonged to the phospholipase A2 family. Additionally, the ATPLMT protein was determined to be involved in the phosphatidylcholine biosynthesis pathway (Table 4).

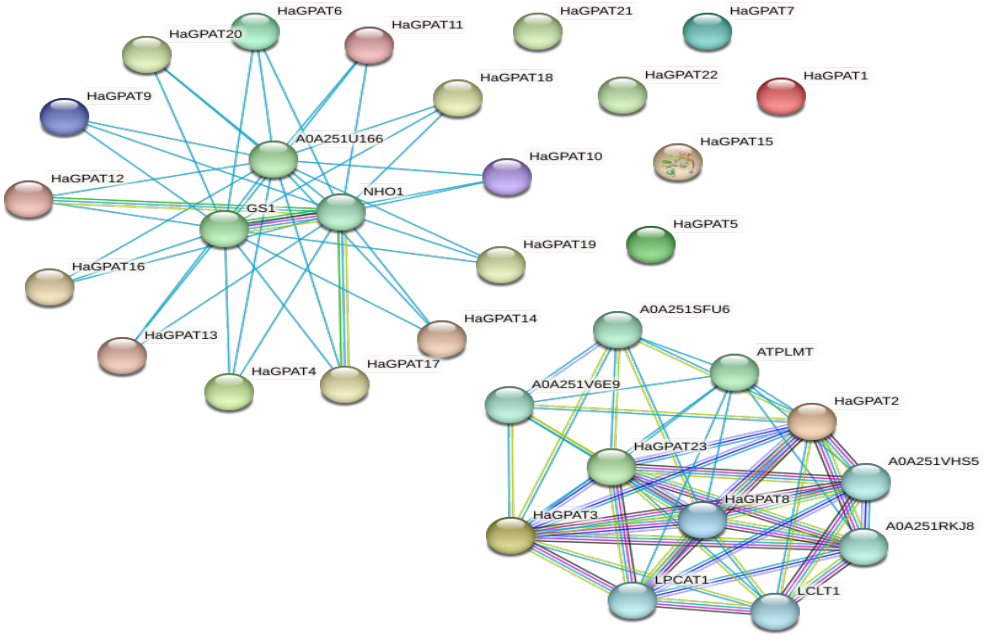

Fig 8. Protein-protein interaction (PPI) network of HaGPAT proteins. In the network, nodes represent proteins, and connections represent predicted PPIs between HaGPAT proteins

Table 4. Predicted interaction partners of HaGPAT proteins and their putative functions

String ID	UniProt ID	Protein Function
GS1	A0A251UNS9	Putative haloacid dehalogenase-like hydrolase (HAD) superfamily protein
NHO1	A0A251VCU0	Putative actin-like ATPase superfamily protein; Belongs to the FGGY kinase family
LPCAT1	V9H0K6	Putative MBOAT (Membrane bound O-acyl transferase) family protein; Belongs to the membrane-bound acyltransferase family
LCLT1	A0A251TMS4	Putative lysocardiolipin acyltransferase 1
A0A251U166	A0A251U166	Putative metallo-dependent phosphatase-like protein
A0A251FVE9	A0A251V6E9	Putative phospholipase A2, Phospholipase A2 domain protein
A0A251SFU6	A0A251SFU6	Putative phospholipase A2 family protein
ATPLMT	A0A251USI6	Phosphatidyl-N-methylethanolamine N-methyltransferase; Catalyzes the second two steps of the methylation pathway of phosphatidylcholine biosynthesis, the SAM-dependent methylation of phosphatidylmonomethylethanolamine (PMME) to phosphatidyldimethylethanolamine (PDME) and of PDME to phosphatidylcholine (PC)
A0A251VHS5	A0A251VHS5	Putative membrane bound O-acyl transferase, MBOAT; Belongs to the membrane-bound acyltransferase family
A0A251RKJ8	A0A251RKJ8	Putative membrane bound O-acyl transferase, MBOAT; Belongs to the membrane-bound acyltransferase family

Three-dimensional structural analyses of HaGPAT proteins showed that the proteins generally exhibit similar folding properties. Notably, HaGPAT2/HaGPAT23, HaGPAT1/HaGPAT21, HaGPAT11/HaGPAT18, and HaGPAT6/HaGPAT13/HaGPAT17 exhibited highly similar three-dimensional structural organizations (Figure 9). Three-dimensional structural analyses showed that HaGPAT proteins generally exhibit α-helix-rich structures and contain fewer β-sheets and loop regions. Overall, the similar structural features of the proteins suggest functional conservatism among members of the *HaGPAT* gene family.

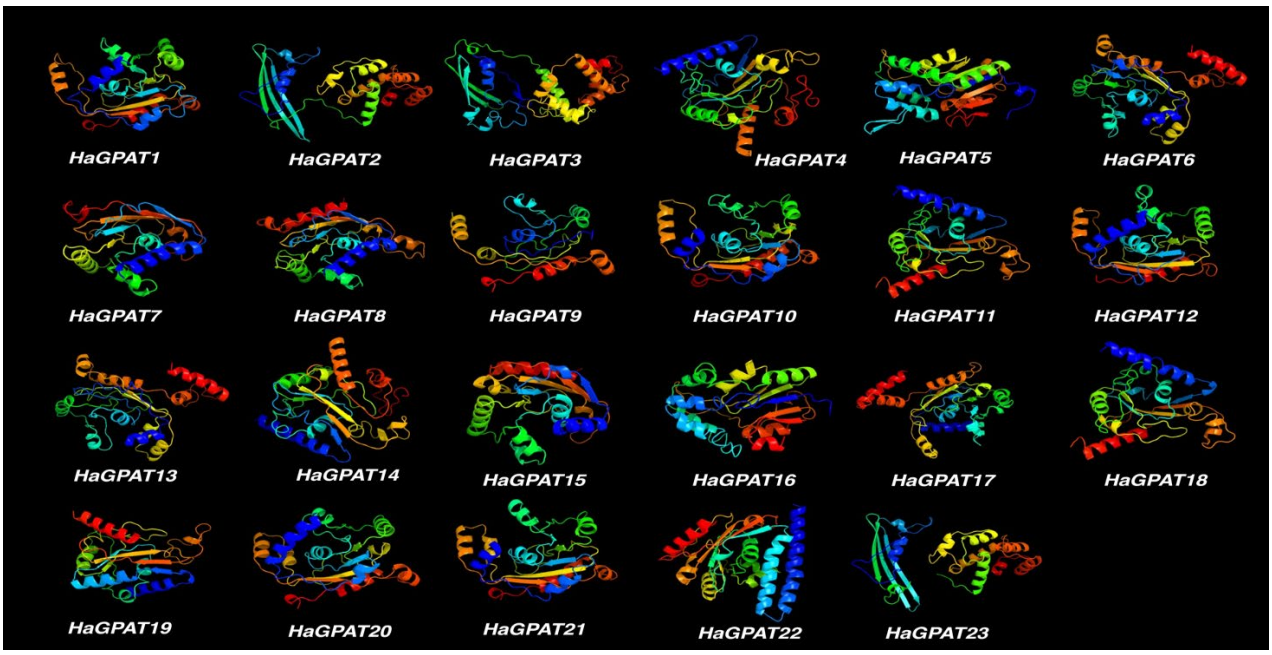

Fig 9. Three-dimensional (3D) structures of HaGPAT proteins**.** The three-dimensional structures of HaGPAT proteins were predicted using the Phyre2 server. The models illustrate the spatial folding patterns and secondary structural elements (α-helices and β-sheets) of proteins

Tissue-Specific RNA-seq Expression Profiles

According to tissue-specific expression analysis, *HaGPAT* genes showed different expression profiles in different organs (Figure 10). *HaGPAT4* showed the most pronounced high expression level in most tissues. High expression levels were also detected in *HaGPAT18* and *HaGPAT21* genes, particularly in ligula tissue. In contrast, *HaGPAT2, HaGPAT5, HaGPAT7, HaGPAT8*, and *HaGPAT10* genes showed low expression in all tissues. Hierarchical clustering analysis shows that genes with similar expression profiles cluster in the same groups.

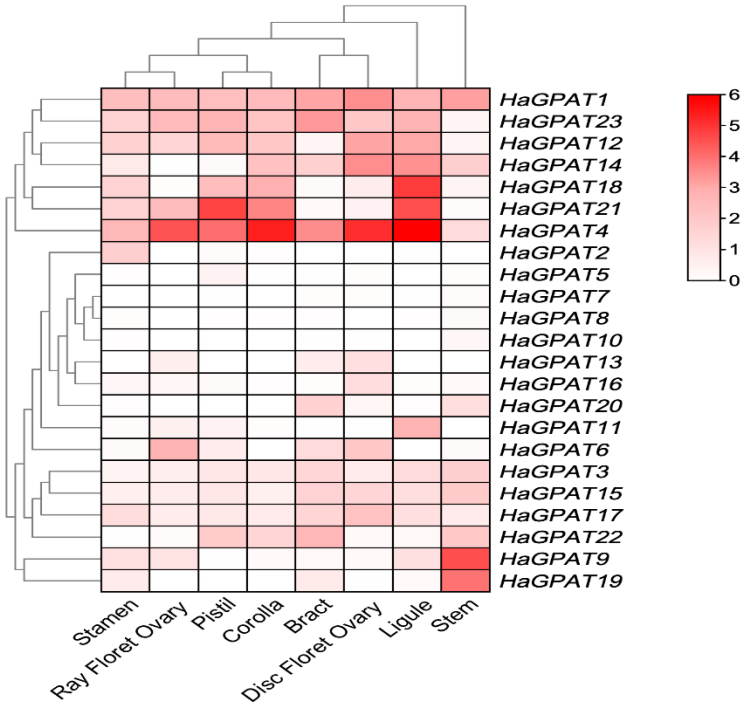

Fig 10. Expression profiles of *HaGPAT* genes in different tissues. The heatmap, created using $\log_2$(RPKM + 1) values obtained from RNA-seq data, was visualized in TBtools. Red and white colors indicate high and low expression levels, respectively

Conclusion and Discussion

GPAT is an enzyme involved in glycerolipid biosynthesis and plays an important role in various physiological processes [34]. According to the results of this study, a total of 23 *HaGPAT* genes were identified in the *Helianthus annuus* genome. The *GPAT* gene family has been identified in different plant species. This family is found in *Arabidopsis thaliana* (n =8) [35], *Phaseolus vulgaris* (n =13) [1], *Perilla frutescens* L. (Britton) (n =14) [36], *Medicago sativa* (n =15) [8], *Zea mays* L. (n =20) [37], *Oryza sativa* (n =26) [9], *Glycine max* (n = 27) [38], and *Brassica napus* L. (n =32) [39].

Phylogenetic trees constructed using GPAT proteins from *Helianthus annuus, Arabidopsis thaliana, Medicago truncatula, and Phaseolus vulgaris* showed that HaGPAT proteins clustered into three main groups. Similar phylogenetic classifications of the GPAT proteins have been previously reported in different plant species. Studies in *Zea mays* [37], *Phaseolus vulgaris* [1], *Gossypium hirsutum* L. [6], and *Medicago sativa* [8], revealed that GPAT proteins clustered under three main phylogenetic groups. Furthermore, it was noted that genes within the same group exhibited similar pattern structures. These results suggest that the phylogenetic classification of *GPAT* gene family may be evolutionarily conserved across different plant species.

This study showed that *HaGPAT* genes exhibit an uneven distribution in the *Helianthus annuus* genome. A similar distribution has been reported for the *GPAT* gene family in *Phaseolus vulgaris* [1], *Medicago sativa* [8], and *Zea mays* [37].

Ka (non-synonymous substitution rate) and Ks (synonymous substitution rate) values are commonly used to assess the selection pressures experienced by genes throughout evolution [40]. Ka/Ks > 1 indicates positive selection, Ka/Ks < 1 indicates negative (purifying) selection, and Ka/Ks ≈ 1 indicates neutral evolution [41]. In this study, Ka/Ks was determined to be < 1 for all orthologous and paralogous gene pairs, suggesting that *HaGPAT* genes may have been subjected to purifying selection during evolution.

Homologous genes resulting from speciation events are defined as orthologous, while homologous genes resulting from gene duplications are defined as paralogous. Determining collinear and syntenic relationships between genomes provides important information about the evolutionary history of gene families [25, 42]. This study suggests that *HaGPAT* genes may have conserved orthologous relationships with *Arabidopsis thaliana*. Furthermore, segmental duplication events may contribute to the expansion of the *HaGPAT* gene family.

Promoter region analyses have shown that *HaGPAT* genes carry numerous cis-regulatory elements associated with stress response, hormone signaling, light response, and growth and development processes. These results suggest that *HaGPAT* genes can be regulated by environmental factors and plant hormones. Similar cis-element profiles have also been reported in studies of the *GPAT* gene family in *Oryza sativa* [9], *Medicago sativa* [8], and *Phaseolus vulgaris* [1].

MicroRNAs (miRNAs) are small, non-coding RNA molecules that regulate the expression of target genes at the post-transcriptional level and are involved in numerous biological processes [43]. miRNAs also play a role in maintaining nutrient balance and adapting to adverse environmental conditions in plants [44]. A significant portion of the miRNAs identified in this study belong to miRNA families that are associated in the literature with plant growth and development, hormone signaling, and abiotic stress responses [45, 46, 47]. In this study, Han-miR160 and Han-miR5691 are miRNAs that target multiple *HaGPAT* genes. miR160 regulates Auxin Response Factor (ARF) genes involved in the auxin signaling pathway in plants and plays an important role in controlling growth and development processes under stress conditions [47]. According to these results, miRNAs may play a role in the post-transcriptional regulation of *HaGPAT* genes.

Protein-protein interactions (PPI), play a role in regulating biological processes, and many proteins perform their functions through interactions with other proteins [48]. LPCAT1 has been reported to be associated with phospholipid metabolism, phosphorus homeostasis, and zinc deficiency responses [49]. PPI analysis has shown that HaGPAT proteins can interact with proteins associated with acyltransferase, phospholipase, and methyltransferase activities. Similar interaction patterns have also been reported for the *Phaseolus vulgaris GPAT* gene family [1].

Investigating the three-dimensional structures of proteins provides important information for understanding their functional properties and biological roles [50]. This study showed that HaGPAT proteins consist largely of α-helices and contain fewer β-sheets. Similar structural features have also been reported for *Phaseolus vulgaris* GPAT proteins [1].

HaGPAT genes have shown variable expression profiles in different tissues. Previous studies have reported that *GPAT6* plays an important role in cutin biosynthesis in sepals and petals in *Arabidopsis thaliana* [51], while *GPAT2* contributes to root tip cuticle formation in young seedlings [52]. In addition, *GPAT* genes have been shown to be involved in processes such as cutin and suberin biosynthesis, cell wall formation, and water

balance maintenance in various plant species [11, 53]. In conclusion, some *HaGPAT* members can be expressed at high levels, especially in floral organs, and these genes may be involved in processes related to flower development. While the tissue-specific expression profiles obtained provide preliminary information about the possible functions of *HaGPAT* genes, further research is needed to confirm the biological functions of these genes.

In conclusion, this study provides important insights into a better understanding of the functions of *GPAT* genes in sunflower. The findings contribute to a better understanding of the structural features, evolutionary relationships, and expression profiles of *HaGPAT* genes. Future studies can correlate this data with results from gene expression and functional analyses. Therefore, this study represents a significant step in the genome-wide identification and in silico characterization of the *HaGPAT* gene family.

Abbreviations

ABRE, abscisic acid-responsive element; ARE, anaerobic responsive element; bp, base pair; CDD, Conserved Domain Database; CDS, coding DNA sequence; Chr, chromosome; ExPASy, Expert Protein Analysis System; GPAT, glycerol-3-phosphate acyltransferase; GRAVY, grand average of hydropathicity; HMM, hidden Markov model; iTOL, Interactive Tree Of Life; Ka, non-synonymous substitution rate; Ks, synonymous substitution rate; kDa, kilodalton; MeJA, methyl jasmonate; MEME, Multiple Expectation Maximization for Motif Elicitation; miRNA, microRNA; ML, maximum likelihood; MYA, million years ago; NCBI, National Center for Biotechnology Information; Pfam, Protein families database; pI, theoretical isoelectric point; PPI, protein–protein interaction; psRNATarget, plant small RNA target analysis server; RPKM, reads per kilobase of transcript per million mapped reads; SRA, Sequence Read Archive; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; UTR, untranslated region

Acknowledgments

The author gratefully acknowledges the Scientific and Technological Research Council of Türkiye (TÜBİTAK) for the financial support provided through the 2211-A Scholarship Program. The author also thanks Prof. Dr. İlker Büyük (Department of Biology, Ankara University) for his assistance in accessing the public NCBI SRA Helianthus annuus RNA-seq datasets.

Funding

The author did not receive support from any organization for the submitted work.

Data Availability statement

The author confirms that the data supporting this study are cited in the article.

Compliance with ethical standards

Conflict of interest / Çıkar çatışması

The author declare no conflict of interest.

Ethical standards

The study is proper with ethical standards.

Authors' contributions

All aspects of this study were conducted by Author.

References

1. Kasapoğlu, Ayşe Gül, et al. Genome-wide characterization of the GPAT gene family in bean (Phaseolus vulgaris L.) and expression analysis under abiotic stress and melatonin. *Genetic Resources and Crop Evolution*, 2024, 71.8: 45494569. https://doi.org/10.1007/s10722-024-01899-3
2. Niu, Ying-Fang, et al. Molecular characterization of a glycerol-3-phosphate acyltransferase reveals key features essential for triacylglycerol production in Phaeodactylum tricornutum. *Biotechnology for Biofuels*, 2016, 9.1: 60 . https://doi.org/10.1186/s13068-016-0478-1
3. Li, Yanli, et al. Phylogenomics and functional analysis of Glycerol-3-Phosphate Acyltransferase (GPAT) Genes: A critical role in lipid biosynthesis. *Physiologia Plantarum*, 2024, 176.5: e14509.. https://doi.org/10.1111/ppl.14509
4. Murata, Norio and Tasaka, Yasushi. Glycerol-3-phosphate acyltransferase in plants. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, 1997, 1348.1-2: 10-16. https://doi.org/10.1016/s0005-2760(97)00115-x
5. Zheng, Zhifu, et al. Arabidopsis AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *The Plant Cell*, 2003, 15.8: 1872-1887. https://doi.org/10.1105/tpc.012427
6. Cui, Yupeng, et al. Genome-wide identification, sequence variation, and expression of the glycerol-3-phosphate acyltransferase (GPAT) gene family in Gossypium. *Frontiers in Genetics*, 2019, 10: 116. https://doi.org/10.3389/fgene.2019.00116
7. Li, Xueting, et al. Characterization of the glycerol-3-phosphate acyltransferase gene and its real-time expression under cold stress in Paeonia lactiflora Pall. *PLoS One*, 2018, 13.8: e0202168. https://doi.org/10.1371/journal.pone.0202168
8. Ma, Jianzhi, et al. Genome-wide identification of the GPAT family in medicago sativa L. and expression profiling under abiotic stress. *Plants*, 2024, 13.23: 3392. https://doi.org/10.3390/plants13233392

9. Safder, Imran, et al. Identification and analysis of the structure, expression and nucleotide polymorphism of the GPAT gene family in rice. *Plant Gene*, 2021, 26: 100290. https://doi.org/10.1016/j.plgene.2021.100290
10. Singer, Stacy D., et al. Arabidopsis GPAT9 contributes to synthesis of intracellular glycerolipids but not surface lipids. *Journal of experimental botany*, 2016, 67.15: 4627-4638. https://doi.org/10.1093/jxb/erw242
11. Jia, Qianru, et al. Mitochondrial GPATs protect Arabidopsis from water stress by impacting cuticle and secondary cell wall development. *Plant Cell Reports*, 2025, 44.10: 225. https://doi.org/10.1007/s00299-025-03608-5
12. Beisson, Fred, et al. The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *The Plant Cell*, 2007, 19.1: 351-368. https://doi.org/10.1105/tpc.106.048033
13. Zhang, Xuanhao, et al. Function identification of Arabidopsis GPAT4 and GPAT8 in the biosynthesis of suberin and cuticular wax. *Plant Science*, 2024, 339: 111933. https://doi.org/10.1016/j.plantsci.2023.111933
14. Riccucci, Ettore, et al. Genome-wide analysis of WOX multigene family in sunflower (Helianthus annuus L.). *International Journal of Molecular Sciences*, 2023, 24.4: 3352. https://doi.org/10.3390/ijms24043352
15. Goodstein, David M., et al. Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, 2012, 40.D1: D1178-D1186. https://doi.org/10.1093/nar/gkr944
16. Potter, Simon C., et al. HMMER web server: 2018 update. *Nucleic acids research*, 2018, 46.W1: W200-W204. https://doi.org/10.1093/nar/gky448
17. Mount, David W. Using the basic local alignment search tool (BLAST). *Cold spring harbor Protocols*, 2007, 2007.7: pdb. top17.https://doi.org/10.1101/pdb.top17
18. Liu, Yue, et al. Genome-wide analysis of wheat DNA-binding with one finger (Dof) transcription factor genes: evolutionary characteristics and diverse abiotic stress responses. *BMC genomics*, 2020, 21.1: 276. https://doi.org/10.1186/s12864-020-6691-0
19. Marchler-Bauer, Aron, et al. CDD: NCBI's conserved domain database. *Nucleic acids research*, 2015, 43.D1: D222-D226. https://doi.org/10.1093/nar/gku1221
20. Gasteiger, Elisabeth, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 2003, 31.13: 3784-3788. https://doi.org/10.1093/nar/gkg563
21. Bailey, Timothy L., et al. The MEME suite. *Nucleic acids research*, 2015, 43.W1: W39-W49. https://doi.org/10.1093/nar/gkv416
22. Chen, Chengjie, et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Molecular plant*, 2023, 16.11: 1733-1742. https://doi.org/10.1016/j.molp.2023.09.010
23. Tamura, Koichiro; Stecher, Glen; Kumar, Sudhir. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, 2021, 38.7: 3022-3027. https://doi.org/10.1093/molbev/msab120
24. Letunic, Ivica and Bork, Peer. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic acids research*, 2024, 52.W1: W78-W82. https://doi.org/10.1093/nar/gkae268
25. Wang, Yupeng, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids research*, 2012, 40.7: e49-e49. https://doi.org/10.1093/nar/gkr1293
26. Gaut, Brandon S., et al. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. *Proceedings of the National Academy of Sciences*, 1996, 93.19: 10274-10279. https://doi.org/10.1073/pnas.93.19.10274
27. Lescot, Magali, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, 2002, 30.1: 325-327.. https://doi.org/10.1093/nar/30.1.325
28. Dai, Xinbin; Zhuang, Zhaohong; Zhao, Patrick Xuechun. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic acids research*, 2018, 46.W1: W49-W54. https://doi.org/10.1093/nar/gky316
29. Shannon, Paul, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 2003, 13.11: 2498. https://doi.org/10.1101/gr.1239303
30. Szklarczyk, Damian, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*, 2019, 47.D1: D607-D613. https://doi.org/10.1093/nar/gky1131
31. Kelley, Lawrence A., et al. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 2015, 10.6: 845-858. https://doi.org/10.1038/nprot.2015.053
32. Badouin, H., et al. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. Nature. 2017;546(7656):148–152. https://doi.org/10.1038/nature22380
33. Xu, Yan-Hong, et al. Genome-wide analysis of WRKY transcription factors in Aquilaria sinensis (Lour.) Gilg. *Scientific Reports*, 2020, 10.1: 3018. https://doi.org/10.1038/s41598-020-59597-w
34. Waschburger, Edgar, et al. Genome-wide analysis of the Glycerol-3-Phosphate Acyltransferase (GPAT) gene family reveals the evolution and diversification of plant GPATs. *Genetics and molecular biology*, 2018, 41.1 suppl 1: 355-370. https://doi.org/10.1590/1678-4685-gmb-2017-0076
35. Yang, Weili, et al. A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, sn-2 preference, and evolution. *Plant Physiology*, 2012, 160.2: 638-652. https://doi.org/10.1104/pp.112.201996
36. Zhou, Yali, et al. Genome-Wide analysis of Glycerol-3-Phosphate Acyltransferase (GPAT) family in Perilla frutescens and functional characterization of PfGPAT9 crucial for biosynthesis of storage oils rich in high-value lipids. *International Journal of Molecular Sciences*, 2023, 24.20: 15106. https://doi.org/10.3390/ijms242015106
37. Zhu, Taotao, et al. Genome-wide analysis of maize GPAT gene family and cytological characterization and breeding application of ZmMs33/ZmGPAT6 gene. *Theoretical and Applied Genetics*, 2019, 132.7: 2137-2154. https://doi.org/10.1007/s00122-019-03343-y
38. Li, Xin, et al. Characterization of the Soybean GPAT Gene Family Identifies GmGPAT1 as a Key Protein in Salt Stress Tolerance. *Plants*, 2025, 14.18: 2862. https://doi.org/10.3390/plants14182862
39. Wang, Jingxue, et al. Genome-wide analysis of glycerol-3-phosphate O-acyltransferase gene family and functional characterization of two cutin group GPATs in Brassica napus. *Planta*, 2020, 251.4: 93. https://doi.org/10.1007/s00425-020-03384-4

40. Zhang, Zhang, et al. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics, proteomics & bioinformatics*, 2006, 4.4: 259-263. https://doi.org/10.1016/s1672-0229(07)60007-2
41. Li, Jun, et al. Correlation between Ka/Ks and Ks is related to substitution model and evolutionary lineage. *Journal of molecular evolution*, 2009, 68.4: 414-423. https://doi.org/10.1007/s00239-009-9222-9
42. Wang, Yupeng, et al. Detection of colinear blocks and synteny and evolutionary analyses based on utilization of MCScanX. *Nature Protocols*, 2024, 19.7: 2206-2229. https://doi.org/10.1038/s41596-024-00968-2
43. Pu, Mengfan, et al. Regulatory network of miRNA on its target: coordination between transcriptional and post-transcriptional regulation of gene expression. *Cellular and molecular life sciences: CMLS*, 2018, 76.3: 441. https://doi.org/10.1007/s00018-018-2940-7
44. Zeng, Houqing, et al. Role of microRNAs in plant responses to nutrient stress. *Plant and Soil*, 2014, 374.1: 1005-1021. https://doi.org/10.1007/s11104-013-1907-6
45. Shriram, Varsha, et al. MicroRNAs as potential targets for abiotic stress tolerance in plants. *Frontiers in Plant Science*, 2016, 7: 817. https://doi.org/10.3389/fpls.2016.00817
46. Dong, Xiangxiang, et al. miR390-tasiRNA3-ARF4 pathway is involved in regulating flowering time in woodland strawberry. *Plant Cell Reports*, 2022, 41.4: 921-934. https://doi.org/10.1007/s00299-021-02828-9
47. Mukherjee, Amrita, et al. Enhancement of ABA sensitivity through conditional expression of the ARF10 gene in Brassica juncea reveals fertile plants with tolerance against Alternaria brassicicola. *Molecular Plant-Microbe Interactions*, 2019, 32.10: 1429-1447. https://doi.org/10.1094/mpmi-05-19-0132-r
48. Athanasios, Alexiou, et al. Protein-protein interaction (PPI) network: recent advances in drug discovery. *Current drug metabolism*, 2017, 18.1: 5-10. https://doi.org/10.2174/1389200218011701192048 32
49. Kisko, Mushtak, et al. LPCAT1 controls phosphate homeostasis in a zinc-dependent manner. *Elife*, 2018, 7: e32077. https://doi.org/10.7554/elife.32077
50. Pan, Archana; Pranavathiyani, G.; Chakraborty, Sibani Sen. Computational modeling of protein three-dimensional structure: Methods and resources. In: *Molecular docking for computer-aided drug design*. Academic Press, 2021. p. 155-178.https://doi.org/10.1016/b978-0-12-822312-3.00023-0
51. Li, Xiao-Chuan, et al. Glycerol-3-phosphate acyltransferase 6 (GPAT6) is important for tapetum development in Arabidopsis and plays multiple roles in plant fertility. *Molecular Plant*, 2012, 5.1: 131-142. https://doi.org/10.1093/mp/ssr057
52. Berhin, Alice, et al. GPAT2 is required for root cap cuticle formation in young Arabidopsis seedlings. *Cell reports*, 2026, 45.1. https://doi.org/10.1016/j.celrep.2025.116686
53. Wang, Wenlong, et al. TaGPAT6 enhances salt tolerance in wheat by synthesizing cutin and suberin monomers to form a diffusion barrier. *Journal of Integrative Plant Biology*, 2025, 67.2: 208-225. https://doi.org/10.1111/jipb.13808