

## Article

# Genome-Wide Identification of the *PgFLA* Gene Family and Association Analysis with Ginsenoside Biosynthesis in *Panax ginseng*

Aimin Wang <sup>1,2</sup>, Xiaoqian Han <sup>1</sup>, Xinyue Hu <sup>1</sup>, Lin Xu <sup>1</sup>, Mingzhu Zhao <sup>1,2</sup>, Yi Wang <sup>1,2</sup>, Kangyu Wang <sup>1,2</sup> and Meiping Zhang <sup>1,2,\*</sup>

<sup>1</sup> College of Life Science, Jilin Agricultural University, Changchun 130118, China; wangaimin@jlau.edu.cn (A.W.); hqq15568796959@163.com (X.H.); xinyuehu@jlau.edu.cn (X.H.); lin.xu@mails.jlau.edu.cn (L.X.); mingzhuzhao@jlau.edu.cn (M.Z.); yi.wang@jlau.edu.cn (Y.W.); kangyu.wang@jlau.edu.cn (K.W.)

<sup>2</sup> Jilin Engineering Research Center Ginseng Genetic Resources Development and Utilization, Jilin Agricultural University, Changchun 130118, China

\* Correspondence: meiping.zhang@jlau.edu.cn

## Abstract

*Panax ginseng* is an economically important medicinal plant, and the biosynthetic regulatory network of ginsenosides, its core bioactive constituents, is extremely complex. Fasciclin-like arabinogalactan proteins (FLAs) are a class of glycoproteins widely involved in plant growth and development that modulate secondary cell wall biosynthesis, phytohormone signal transduction, and stress responses in plants. In this study, we performed a genome-wide identification of the *FLA* gene family in *P. ginseng* and identified 21 *PgFLA* family members. Comprehensive analyses were subsequently conducted to characterize their gene structures, chromosomal distributions, collinearity relationships, phylogenetics, expression patterns, and cis-acting regulatory elements. Subsequently, combined with SNP/InDel-ginsenoside association analysis based on 344 ginseng accessions, correlation analysis between gene expression and ginsenoside content, as well as co-expression network analysis linking *PgFLA* genes with key enzyme genes of the ginsenoside biosynthetic pathway, eight candidate *PgFLA* genes associated with ginsenoside biosynthesis were identified. Further mediation effect analysis identified six *PgFLA* members that potentially affect ginsenoside biosynthesis by modulating the expression of key enzymes involved in the ginsenoside biosynthetic pathway. Quantitative real-time PCR (qRT-PCR) was used to validate the expression of these pivotal candidate genes in methyl jasmonate (MeJA)-treated ginseng adventitious roots. The results revealed that the transcript abundances of five genes (*PgFLA02*, *PgFLA05-01*, *PgFLA05-02*, *PgFLA06*, and *PgFLA09-06*) were significantly negatively correlated with the levels of protopanaxadiol-type ginsenosides (Rb1, Rb2, Rb3, Rc, and Rd). These findings provide a theoretical foundation for further dissecting the molecular regulatory network underlying ginsenoside biosynthesis and supply valuable candidate genes for marker-assisted breeding and quality improvement of cultivated *P. ginseng*.

**Keywords:** *Panax ginseng*; fasciclin-like arabinogalactan protein (FLA); SNP/InDel; ginsenoside biosynthesis; methyl jasmonate (MeJA)

Academic Editor: Jiping Gao

Received: 16 June 2026

Revised: 30 July 2026

Accepted: 1 August 2026

Published: 4 August 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

*Panax ginseng*, a perennial herb belonging to the Araliaceae family, is a well-known medicinal plant in traditional Chinese medicine, with a cultivation and medicinal application history spanning more than one thousand years [1]. The pharmaceutical efficacy of ginseng is mainly attributed to ginsenosides, triterpenoid saponins that are abundantly accumulated in ginseng roots [2]. Ginsenosides, the core bioactive compounds responsible for diverse pharmacological activities, are mainly classified into dammarane- and oleanane-types based on their aglycone structures. Dammarane-type saponins are further subdivided into protopanaxadiol (PPD) and protopanaxatriol (PPT) types according to the hydroxyl configuration at the C-20 position. To date, over 200 structurally distinct ginsenosides have been isolated and characterized [3], and although the core biosynthetic pathway of ginsenosides has been elucidated, its regulatory mechanisms remain complex. Therefore, improving ginsenoside yield via genetic manipulation of biosynthetic pathways using modern biotechnology has become a research priority.

Arabinogalactan proteins (AGPs) are hydroxyproline-rich glycosylated proteins that are localized predominantly in plant cell walls and plasma membranes. They participate in diverse biological processes, including cell proliferation, differentiation, programmed cell death, and plant-environment interactions [4–6]. Fasciclin-like arabinogalactan proteins (FLAs) constitute a unique AGP subfamily characterized by one or two conserved fasciclin (FAS) domains and canonical AGP-like glycosylation patterns. Many FLAs are anchored to the plasma membrane via glycosylphosphatidylinositol (GPI) modifications. The FAS domain was originally identified in the insect nervous system, where it mediates cell–cell adhesion [7–10]. Owing to their unique protein-glycan composite structure and cell-surface localization, FLAs are presumed to act as molecular linkers bridging the cell walls and plasma membranes, which are critical for extracellular signal perception and dynamic remodeling of the cell wall. Functional investigations have verified the indispensable roles of FLA family genes in multiple plant developmental processes. FLA genes extensively regulate plant organ morphogenesis. Analyses of *Arabidopsis fla11* and *fla12* double mutants revealed drastically reduced stem tensile strength and elastic modulus, increased cellulose microfibril angle, and decreased arabinose, galactose, and cellulose contents in the cell walls [11]. Heterologous overexpression of *OreFLA11* from the endangered *Carpinus polyneura* in *Populus alba* improved xylem development and thickened secondary cell walls (SCWs), reduced lignin content, increased cellulose abundance, and ultimately enhanced stem tensile strength [12]. In addition, FLAs participate in the regulation of plant regeneration and reproductive development. RNAi suppression of *BrFLA2*, *BrFLA28*, and *BrFLA32* in *Brassica rapa* resulted in a higher rate of precocious pollen germination in anthers under high humidity, confirming the indispensable roles of these three genes in normal pollen development [13]. In rice, the *osfla1* mutant displays abnormal Ubisch bodies and discontinuous outer exine of microspores, which eventually triggers pollen abortion [14]. The FLA family also contributes to abiotic stress tolerance in plants. Transgenic poplar plants overexpressing *OreFLA11* or *PalFLA11* exhibited retarded growth, impaired root development, decreased photosynthetic pigment levels, elevated reactive oxygen species levels, and reduced antioxidant enzyme and proline levels under salt stress, leading to weakened salt resistance [12]. Knockout of rice *OsFLA8* disturbs Casparian strip formation and increases plant sensitivity to salt and calcium stress, indicating its essential function in maintaining ionic homeostasis and stress responses [15]. Recent studies have uncovered novel functions of FLAs in secondary metabolism and crop quality. In melon, the transcript levels of multiple *CmFLA* genes were negatively correlated with the total soluble solids content in fruits. Furthermore, the transient overexpression of *CmFLA8* significantly reduced glucose accumulation in melon flesh, verifying its regulatory function in sugar metabolism [16]. In ripening tomato fruits, virus-induced

gene silencing (VIGS) of *SIFLA1*, *SIFLA5*, and *SIFLA13* accelerated fruit pigmentation and markedly increased the accumulation of lycopene,  $\alpha$ -carotene, and  $\gamma$ -carotene, suggesting that these *FLAs* are potential repressors of carotenoid biosynthesis [17].

Elicitation is an effective biotechnological approach that mimics biotic or abiotic stress cues to activate plant defense responses and boost the biosynthesis of specific secondary metabolites [18–20]. As a vital phytohormone and signaling molecule, methyl jasmonate (MeJA) plays a central role in plant growth, defense activation, and secondary metabolism regulation [21]. Exogenous MeJA application significantly upregulates the transcription of key enzyme genes, including those encoding *SS*, *SE*, *DDS*, various *CYPs*, and *UGTs*, thereby promoting ginsenoside accumulation in ginseng [22–24]. Integrated transcriptomic and metabolomic analyses have revealed differential responses to MeJA across ginseng cultivars, with protopanaxadiol (PPD)-type ginsenosides generally accumulating more than their protopanaxatriol (PPT)-type counterparts [25]. Several MeJA-responsive transcription factors, such as WNK kinases, PYL proteins, and NAC family members, have been identified to mediate MeJA signal transduction and modulate downstream ginsenoside biosynthetic genes [26–28]. These findings outline a complex MeJA-mediated regulatory network for ginsenoside synthesis, although the upstream signal perception pathway and its linkage to dynamic cell wall remodeling remain poorly characterized.

Although the biological functions of *FLA* genes in plant development and stress adaptation have been partially clarified, few studies have focused on their roles in secondary metabolism. In this study, we performed a genome-wide identification and systematic bioinformatic analysis of the *PgFLA* gene family. Candidate *PgFLA* genes tightly associated with ginsenoside biosynthesis were screened using multiple strategies, and their expression patterns were further analyzed following MeJA treatment. This study provides experimental evidence and theoretical support for the potential functions of *FLAs* in MeJA signaling and secondary metabolic regulation in *P. ginseng*.

## 2. Materials and Methods

### 2.1. Identification of *PgFLA* Gene Family Members in *P. ginseng*

Reference genomes of 20 species were retrieved from the NCBI database (<https://www.ncbi.nlm.nih.gov/>, accessed on 1 April 2026). A total of 710 amino acid sequences annotated as members of the *FLA* gene family were extracted and downloaded from the corresponding genomic annotation files. These sequences were used as query sequences, and sequence alignment was performed against the *Panax ginseng* transcriptome dataset (PRJNA302556) using BLAST software (v2.1.7). The E-value was set to  $1 \times 10^{-6}$ , and the resulting gene population was defined as gene pool I. The PF02469 file was downloaded from the InterPro database (<https://www.ebi.ac.uk/interpro/entry/pfam/#table>, accessed on 5 April 2026), and hmmsearch was performed against the *Panax ginseng* protein database using HMMER (v3.3.2) with an E-value threshold of  $\leq 0.01$ . The obtained protein sequences were subsequently used as queries to search the transcriptome, yielding gene pool II. Genes sharing 100% sequence identity in gene pools I and II were eliminated, and the remaining genes were regarded as candidate genes for characterization.

The nucleotide sequences of these preliminary candidates were extracted using customized Perl scripts and submitted to the NCBI CD-search platform (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>, accessed on 10 April 2026). Sequences harboring intact FAS1/Fascilin conserved domains were retained. The open reading frames (ORFs) of the qualified full-length sequences were extracted using TBtools (v2.468). The ORF sequences were submitted to the SMART database (<https://smart.embl.de/>, accessed on 10 April 2026) to verify the integrity of the conserved domains. Subsequently, the amino acid sequences of the AGP regions were subjected to

N-glycosylation site prediction using the NetNGlyc 1.0 server (<https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/>, accessed on 12 April 2026). Genes encoding proteins that harbored both complete FAS1/Fascidin conserved domains and two or more AGP regions were identified as *Panax ginseng* FLA family members. The finally identified genes were renamed following the format Pg + FLA + numerical serial number.

The deduced amino acid sequences of all identified family members were submitted to the SignalP 5.0 server (<https://services.healthtech.dtu.dk/services/SignalP-5.0/>, accessed on 13 April 2026) to predict N-terminal signal peptides. The big-PI Plant Predictor web server ([https://mendel.imp.ac.at/gpi/plant\\_server.html](https://mendel.imp.ac.at/gpi/plant_server.html), accessed on 13 April 2026) was used to predict the potential GPI modification sites of these proteins.

## 2.2. Conserved Domains and Motifs Analysis of PgFLA Proteins in Ginseng

The deduced amino acid sequences encoded by the PgFLA open reading frames (ORFs) were imported into MEGA (v11.0) software. A maximum-likelihood (ML) phylogenetic tree was constructed under the WAG+G substitution model with bootstrap replicates set to 2000 to classify the FLA family members' subtypes. The conserved motifs of these genes were predicted using the online MEME platform (<https://meme-suite.org/>, accessed on 15 April 2026), with the maximum number of identified motifs limited to 10, and the composition and distribution patterns of the obtained motifs were characterized. Conserved domain information for all PgFLA proteins was acquired using Batch CD-search (<https://www.ncbi.nlm.nih.gov/Structure/wrpsb-out/wrpsb.cgi>, accessed on 17 April 2026). Finally, TBtools was used to integrate the phylogenetic relationships, conserved motif distributions, and conserved domain features for integrated visualization.

## 2.3. Chromosomal Localization and Synteny Analysis of PgFLA Genes in Ginseng

BLAST alignment was performed between the gene sequences of the PgFLA family and the *P. ginseng* T2T reference genome deposited in the NGDC (China National Center for Bioinformation, BioProject: PRJCA022032). Genes with sequence identity  $\geq 95\%$  and aligned length  $\geq 1000$  bp were retained to obtain their chromosomal coordinates, and a chromosomal distribution map was plotted using the online tool MG2C (v2.0) ([http://mg2c.iask.in/mg2c\\_v2.1/index.html](http://mg2c.iask.in/mg2c_v2.1/index.html), accessed on 20 April 2026).

Subsequently, paralogous loci on chromosomes were screened using identity and alignment length thresholds of  $\geq 95\%$  and  $\geq 500$  bp, respectively. The obtained syntenic data were visualized using R software (v4.6.0), with purple lines linking collinear genomic blocks across distinct chromosomal regions.

## 2.4. Phylogenetic Analysis of PgFLA Genes

The ORFs of PgFLA family members were subjected to phylogenetic analysis, together with 24 genes derived from eight exogenous species (two model plants, three monocotyledons, and three dicotyledons). Multiple sequence alignment of amino acid sequences was performed using the ClustalW algorithm in MEGA (v11.0) software. The Find Best DNA/Protein Models (ML) module was applied to screen the optimal evolutionary model, and the WAG+G model was identified. A maximum likelihood (ML) phylogenetic tree was constructed with a bootstrap replicate value of 2000. The resulting phylogenetic tree was visualized using the ChiPlot online platform (<https://www.chip-lot.online/>, accessed on 25 April 2026) [29].

### 2.5. Expression Pattern Analysis of PgFLA Genes in Ginseng

TPM (Transcripts Per Million) values of PgFLA gene family members were extracted from published transcriptome datasets (PRJNA302556), including main roots of ginseng at four growth ages, 14 different tissues from 5-year-old plants, and main roots derived from 42 landraces, using R software (v4.6.0). Subsequently, TBtools (v2.468) was used for expression normalization, hierarchical clustering, and heatmap visualization.

### 2.6. Cis-Acting Element Analysis of PgFLA Genes in Ginseng

The 2000 bp upstream promoter sequences of each PgFLA gene were retrieved based on their chromosomal location. All acquired sequences were submitted to the PlantCARE database (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>, accessed on 27 April 2026) to identify the distribution of cis-acting regulatory elements in the promoter regions. Statistical visualization of the category and quantity of cis-elements was performed using the ggplot2 package (v4.0) in R language, and TBtools was used to visualize the positional distribution of cis-elements along promoter sequences.

### 2.7. Association Analysis Between SNP/InDel Variations in PgFLA Genes and Ginsenoside Contents

SNP and InDel variation information for all PgFLA gene family members was extracted from an in-house SNP/InDel dataset consisting of 344 ginseng accessions (NCBI, <https://www.ncbi.nlm.nih.gov/>, BioProject accession: PRJNA302556). The SNP/InDel dataset used in this study was derived from 344 cultivated varieties and local landraces collected and screened from 16 major ginseng-producing counties in Jilin Province. Transcriptome sequencing, SNP/InDel variant calling, population structure analysis, and phylogenetic reconstruction among individuals had been performed on these samples, which improved the reliability of the loci retrieved from the database [30]. Genotypes were numerically coded as follows: homozygous reference allele = 0, heterozygous genotype = 1, and homozygous variant allele = 2. Association analysis was conducted between genotypic variants and nine types of ginsenosides. For loci with only two genotypes and more than three biological replicates per genotype, the car package in R was used to test the homogeneity of the variances. An independent two-sample t-test was used to compare ginsenoside content when variance homogeneity was satisfied; otherwise, Welch's corrected t-test was performed. For loci containing three genotypes with qualified sample sizes, one-way ANOVA based on the Type-III sum of squares was performed, and post hoc multiple comparisons were implemented using Tukey's test. To eliminate false positive biases induced by multiple testing, the Benjamini–Hochberg (BH) procedure was utilized to conduct FDR correction on the association analysis outputs. Grouped independent correction was implemented for the *p*-values of all loci corresponding to each ginsenoside, and a corrected FDR value of less than 0.05 was defined as the screening cutoff for significant associations. Genes harboring variant loci that exerted significant effects on ginsenoside content were screened out, and these genes were grouped into candidate gene set A.

### 2.8. Correlation Analysis Between PgFLA Expression Levels and the Contents of Nine Ginsenosides

Spearman's correlation analysis between transcript abundances of PgFLA genes and the concentrations of nine ginsenosides in main roots from 344 ginseng landraces was performed using SPSS software (v26.0). Correlation was regarded as statistically significant at  $p \leq 0.05$ , and genes with significant correlations were presumed to participate in the regulation of ginsenoside biosynthesis. The correlation coefficients ( $\rho$ ) and corresponding *p*-values were summarized, and a correlation heatmap was generated for data

visualization. Genes with significant correlations were collected and designated as candidate gene set B.

### 2.9. Co-Expression Network Analysis Between PgFLA Genes and Key Enzyme Genes Involved in Ginsenoside Biosynthesis

Spearman's correlation analysis was performed using the expression data of PgFLA genes and 17 functionally verified key enzyme genes from the ginsenoside biosynthetic pathway. The resulting data matrix was imported into BioLayout Express 3D (v3.4) for network visualization, in which nodes represented individual genes and edges indicated correlated gene pairs. Co-expression networks were constructed using eight different screening thresholds (Table S6). Genes that maintained tight co-expression links with key enzyme genes at a stringent threshold of  $p = 1 \times 10^{-8}$  were assigned to candidate gene set C. Finally, overlapping genes shared among sets A, B, and C were extracted and defined as core candidate gene set D.

### 2.10. Mediation Effect Analysis Among PgFLA Genes, Key Enzymes, and Ginsenosides in *P. ginseng*

Highly correlated genes were identified using the three aforementioned approaches. However, both the association analysis between SNPs/InDels and ginsenosides and the correlation analysis were performed to screen key genes from a single perspective. In contrast, mediation effect analysis identified an integrated regulatory pathway linking genetic variation, transcription, and metabolism to ginsenoside accumulation. Thus, the expression levels of genes in gene pool D derived from 344 main root samples of *P. ginseng* were adopted as independent variables, the expression data of 17 key enzymes were set as mediator variables, and the contents of nine monomeric ginsenosides were treated as dependent variables. Mediation effect analysis was performed among independent, mediator, and dependent variables using the mediation package implemented in R software (v4.6.0), with the bootstrap resampling number set to 5000. FDR multiple correction was applied to the final p-values to improve the reliability of the analytical results.

Genes satisfying the criteria of  $FDR \leq 0.05$ ,  $PropMediated \geq 0.3$ , and identical symbols of ACME and ADE were screened out, indicating that these genes indirectly regulated ginsenoside accumulation by modulating the transcription of ginsenoside-synthetic key enzyme-encoding genes. Significant mediating relationships among genes, key enzymes, and ginsenosides were visualized using the OmicStudio cloud platform [31] (<https://www.omicstudio.cn/tool>, accessed on 1 May 2026). TBtools was used to plot the Average Causal Mediation Effect (ACME) values of screened genes with significant mediation effects, and GraphPad Prism (v9.5) was applied for the visualization of PropMediated values.

### 2.11. Expression Profiling of Candidate PgFLA Genes Under MeJA Treatment

Ginseng adventitious roots were induced from *P. ginseng* cv. Damaya, a local cultivar collected from Jilin Province, China. The induced adventitious roots were cultured on solid B5 medium supplemented with 30 g/L sucrose and 4 mg/L IBA, with the pH adjusted to 5.95. Subculturing was performed at 28-day intervals, and three successive subcultures were completed before the root materials were used for subsequent elicitation treatments. For suspension proliferation, 1 g of pre-cultured ginseng adventitious roots was inoculated into 125 mL of liquid B5 medium in 250 mL Erlenmeyer flasks. Suspension cultures were incubated at 22 °C in continuous darkness on an orbital shaker at 110 rpm for 28 days. After proliferation, 10 g of ginseng adventitious roots were transferred into separate 250 mL Erlenmeyer flasks containing 125 mL liquid B5 medium supplemented with 200  $\mu$ M methyl jasmonate (MeJA, dissolved in ethanol) for elicitor treatment. A total of 27

independent root samples were prepared for analysis. A solvent vehicle control treated with an equal volume of pure ethanol was established in parallel. Three independent root lines were harvested as biological replicates at each sampling time point, and sample collection was performed at 0, 6, 12, 24, 36, 48, 60, 72, and 84 h post-treatment.

Total RNA was extracted from samples harvested at different time points using the TRIzol reagent method. RNA integrity was assessed using agarose gel electrophoresis. Only RNA samples displaying three distinct clear bands, with the intensity of the 28S rRNA band approximately twice that of the 18S rRNA band, were subjected to reverse transcription to synthesize complementary DNA (cDNA) for further analysis. The resulting cDNA was diluted five-fold and used as the template for subsequent quantitative real-time PCR (qRT-PCR) assays. The *CYP* gene was used as the internal reference gene [32]. Quantitative real-time PCR was performed on an ABI 7500 system to quantify the transcript abundance of candidate *PgFLA* genes, and relative gene expression values were calculated using the  $2^{-\Delta\Delta Ct}$  method (Table S8). One-way analysis of variance (ANOVA) was used to compare transcriptional differences across sampling time points, with the 0 h group serving as the control.

The residual adventitious root materials were freeze-dried and ground into fine powder. Total ginsenosides were extracted with water and further purified using ODS column chromatography for subsequent high-performance liquid chromatography (HPLC) detection. The standards used for these ginsenosides were purchased from the National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China). The mobile phase consisted of acetonitrile (A) and purified water (B). The separation time was 120 min, with gradient elution. The injection volume was 10  $\mu$ L, the column temperature was 30  $^{\circ}$ C, and the flow rate of the mobile phase was 1.0 mL/min. Three technical replicates were also performed for each biological replicate. Individual ginsenosides were separated using a Waters e2695 HPLC system coupled with an e2498 UV detector (Waters, Milford, MA, USA) set at an absorbance wavelength of 203 nm. The contents of eight monomeric ginsenosides, including Rb1, Rg2, Rh1, Rc, Rb2, Rb3, Rd, and Rg3, were quantified based on their corresponding chromatographic peak areas [32–34].

#### 2.12. Correlation Analysis Between *PgFLA* Expression and Ginsenoside Contents Under MeJA Induction

Intersecting genes derived from candidate sets A, B, and C were screened to obtain core *PgFLA* genes that exhibited significant regulatory correlations with ginsenoside biosynthesis across all three analytical strategies. Spearman's correlation analysis between the relative expression levels of these core *PgFLA* genes and the protopanaxadiol-type ginsenoside content in matched samples was performed using SPSS (v26.0). The correlation coefficients, significance values, and gene-ginsenoside association relationships were visualized using the OmicStudio platform.

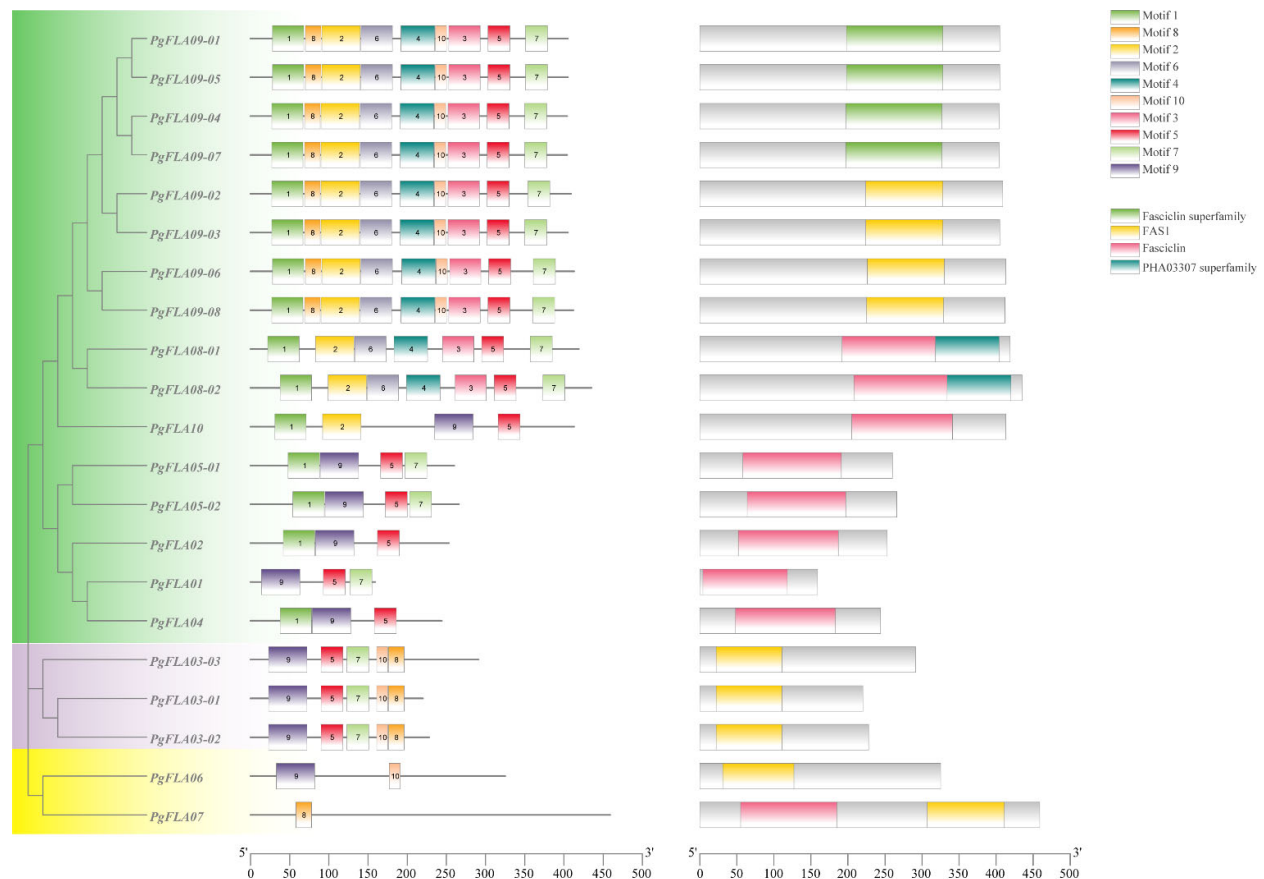
### 3. Results

#### 3.1. Analysis of Conserved Domains and Motifs of *PgFLA* Proteins

Based on the aforementioned procedures, 101 genes were retrieved from gene pool I, and 310 genes were obtained from gene pool II. After redundant sequences with 100% identity were removed, a union set containing 344 genes was acquired. A total of 21 *PgFLA* family members were identified via homologous alignment and conserved domain prediction of deduced proteins. All 21 family members harbored intact FAS1/Fasciclin conserved domains and contained no less than two AGP regions. Signal peptides were absent in six members, namely *PgFLA01*, *PgFLA03-01*, *PgFLA03-02*, *PgFLA03-03*, *PgFLA06*, and *PgFLA08-02*, whereas the remaining 66.6% of the family proteins were predicted to

possess N-terminal signal peptides. In addition, 12 members of this gene family were identified to carry GPI-anchor motifs. (Table S1). Phylogenetic classification separated these members into three major clades, and all predicted FLA proteins harbored conserved canonical FAS1 (fasciclin) domains (Figure 1).

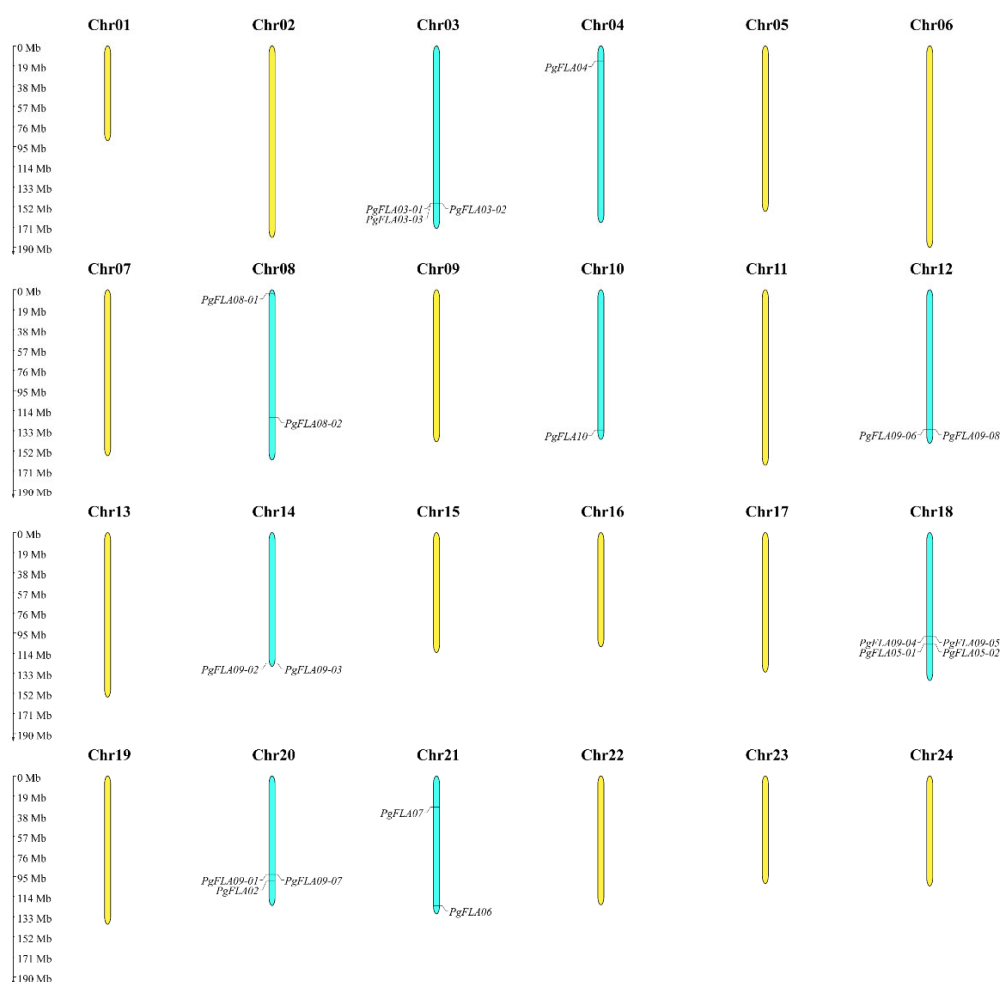
Conserved motif distribution patterns revealed that FLA proteins from distinct clades exhibit characteristic motif compositions and arrangements. All members of the subclade comprising *PgFLA09*, *PgFLA08*, and *PgFLA10* within Clade I possessed a tandem array of motifs 1-2-5, which may represent a unique signature specific to this subclade. Compared with *PgFLA08*, two additional conserved motifs (Motif 8 and Motif 10) were detected in *PgFLA09*. All proteins belonging to Clade II (*PgFLA03*) contained an intact tandem arrangement of Motif 9-5-7-10-8. However, *PgFLA03-01* and *PgFLA03-02* transcripts clustered separately to form an independent minor branch, whereas *PgFLA03-03* transcripts may not have shared sufficient sequence similarity to cluster alongside them. Further divergence in motif composition was observed in Clade III (*PgFLA06/07*): *PgFLA06* carried motifs 9 and 10, whereas only motif 8 was detected in *PgFLA07*. Collectively, the motif composition and domain architecture of the FLA gene family may exhibit pronounced subtype-specific characteristics.



**Figure 1.** Phylogenetic analysis, motif distribution, and conserved domain analysis of *PgFLA* gene family members. In the phylogenetic tree, *PgFLA* members shaded in green belong to the FLA\_A subtype, those shaded in purple belong to the FLA\_B subtype, and those shaded in yellow correspond to the FLA\_C subtype. The phylogenetic tree was constructed using the Maximum Likelihood (ML) method with the WAG+G model.

### 3.2. Chromosomal Localization and Synteny Analysis of PgFLA Genes#

Chromosomal localization results revealed that *PgFLA* genes were unevenly distributed across nine ginseng chromosomes, with multiple genes clustered on individual chromosomes. Four *PgFLA* members were mapped onto Chr 18, representing the highest copy number among all chromosomes. Notably, these four genes were derived from two physically adjacent loci, indicative of tandem duplication events during the evolution of ginseng. A similar pattern was observed on Chr 20, suggesting that *PgFLA* genes on Chr 18 and Chr 20 were retained after whole-genome duplication. Two transcript isoforms originating from identical genes were separately anchored on Chr 8, Chr 12, and Chr 14, which arose from alternative splicing and post-transcriptional modifications. Multiple transcripts derived from a single gene can produce structurally distinct proteins that fulfill diverse biological functions. Although *PgFLA09* was originally assembled as an alternative splice variant from a single gene based on transcriptomic data, genome mapping assigned these transcripts to Chr 12, Chr 14, Chr 18, and Chr 20. Consequently, only the position corresponding to the transcript with the highest alignment score was reserved for the final localization dataset (Figure 2A). This phenomenon may be attributed to two factors. On one hand, *P. ginseng* is an allotetraploid species, and the co-occurrence of selfing and outcrossing during sexual reproduction may cause substantial genetic divergence among sequenced samples. However, this discrepancy may also stem from artifacts generated by de novo transcriptome assembly without a reference genome.



(A)

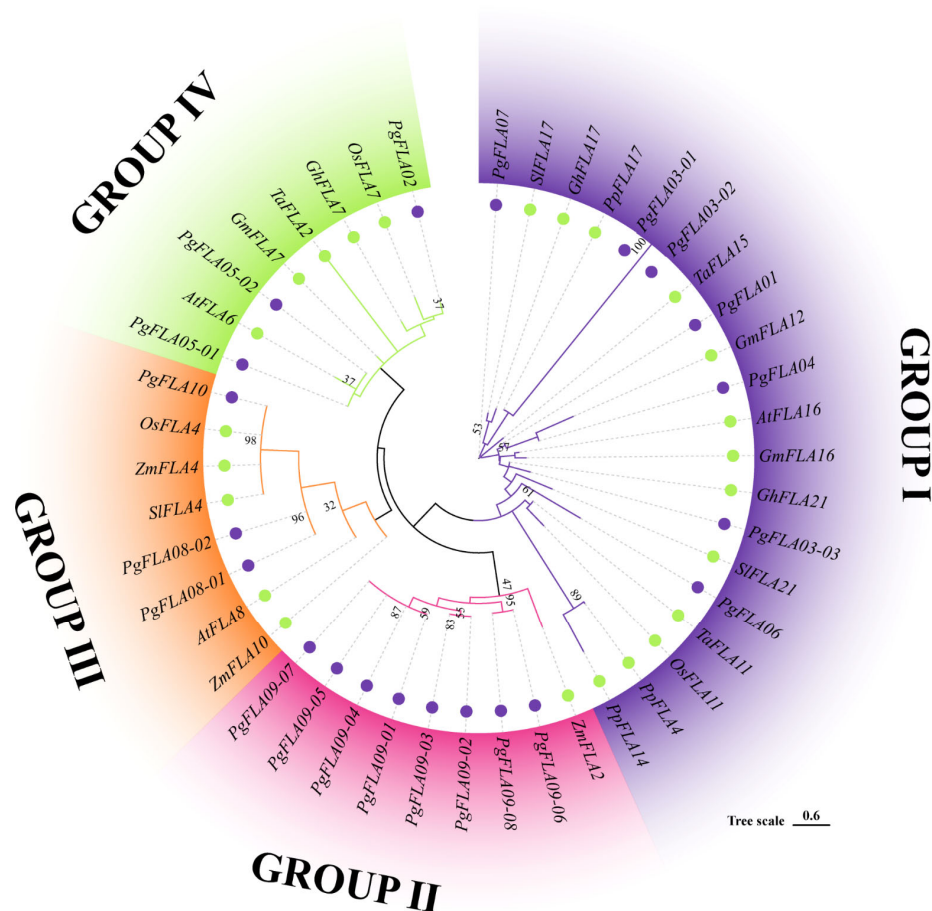


**Figure 2.** Chromosomal localization and synteny analyses of *PgFLA* genes. (A) Chromosomal localization and distribution of *PgFLA* genes. The scale on the left represents the physical position of the chromosomes. Chromosomes mapped with *PgFLA* genes are shown in green, while the remaining chromosomes are shown in yellow. (B) Intrasppecific collinearity and duplication event analysis of the *PgFLA* gene family. The outer color blocks indicate chromosome number and physical position, and purple curves link the chromosomal positions of collinear *PgFLA* gene pairs.

Synteny analysis further revealed the evolutionary characteristics of the *PgFLA* gene family in the *Panax ginseng* genome. The results showed that prominent syntenic blocks were detected at both ends of chromosomes Chr 04 and Chr 10, suggesting that these two chromosomes may have undergone identical whole-genome duplication or large-segmental duplication events during evolution, which could have driven the expansion of the *PgFLA* gene family. Multiple homologous genes were identified on Chr 18, which may indicate that the gene family members on this chromosome might have collectively originated from a gene segment on Chr 20; moreover, the two chromosomes may have experienced gene duplication events such as chromosomal crossover, transposition, or retro transposition, thereby resulting in a syntenic relationship between them. In addition, *PgFLA* genes with high sequence similarity were also detected at both ends of the same chromosome (Chr 21), which may have been caused by intrachromosomal structural variations such as segmental duplication or inversion. Notably, the terminal sequences of Chr 23 exhibited significant homology with genes located at the telomeric region of Chr 21. However, no matching *PgFLA* genes were detected at the corresponding loci on Chr 23 during chromosomal localization analysis. This finding indicates that Chr 23 harbors highly homologous sequences derived from FLA family genes; however, these sequences may lack integral gene architectures and consequently fail to encode functional proteins.

### 3.3. Phylogenetic Analysis of PgFLA Gene Family Members

A phylogenetic tree constructed using FLA protein sequences from various plant species grouped *PgFLA* members in ginseng into four distinct clades (Table S2). All *PgFLA09* transcripts were tightly clustered into a single subclade within Group II and formed an independent lineage with *ZmFLA12*. This finding indicates that, despite the nucleotide sequence divergence among different *PgFLA09* transcripts, their deduced amino acid sequences shared closer phylogenetic relationships relative to other paralogs (Figure 3).



**Figure 3.** Phylogenetic tree analysis of the FLA gene family among *Panax ginseng*, *Arabidopsis thaliana*, *Physcomitrium patens*, *Glycine max*, *Solanum lycopersicum*, *Gossypium hirsutum*, *Triticum aestivum*, *Zea mays*, and *Oryza sativa* Japonica Group. Orange solid circles indicate FLA members of *P. ginseng*, and green solid circles represent FLA members from other exogenous species. All members were classified into distinct subclades according to evolutionary distances, which clarified the evolutionary status and subfamily classification of *PgFLA* genes and their phylogenetic relationships with homologous genes from other plant species.

Within Group I, *PgFLA07* possessed two canonical conserved domains (FAS1 and Fasciclin) and clustered with *SIFLA17*, *GhFLA17*, and *PpFLA17*, all of which contained identical dual-domain architectures, further confirming the structural and evolutionary conservation of genes in this clade. Only *PgFLA02* and *PgFLA05* from ginseng fell into Group IV. Combined with previous chromosomal localization and synteny data, this clustering result supports their high structural similarity and shared origin from an ancestral gene duplication event, implying functional coordination in the same biological pathways.

Notably, the alternative transcript *PgFLA03-03* did not cluster with *PgFLA03-01* and *PgFLA03-02*, although all three transcripts originated from the same genomic locus. Amino acid sequence comparison revealed that *PgFLA03-03* had a remarkably longer protein sequence, presumably stemming from alternative splicing-induced alterations in the composition or arrangement of the conserved domains. Such structural variation led to divergent clustering patterns on the phylogenetic tree and implied functional divergence among different transcript isoforms.

### 3.4. Expression Pattern Analysis of *PgFLA* Genes

The expression profiles of the *PgFLA* gene family were characterized across the main roots of ginseng with different cultivation ages, diverse tissues, and various landraces, and most *PgFLA* genes exhibited obvious spatiotemporal expression specificity (Figure 4; Table S3). For the main roots sampled from plants of four different ages, the transcription level of *PgFLA08-01* decreased gradually with increasing growth years, whereas *PgFLA01*, *PgFLA03-02*, and *PgFLA09-08* remained undetectable in all age groups (Figure 4A).



**Figure 4.** Heatmap spatiotemporal expression pattern analysis of *PgFLA* genes in *P. ginseng*. (A) Expression patterns of *PgFLA* genes in the main roots of ginseng during four different growth years.

(B) Expression patterns of *PgFLA* genes in 14 tissues of five-year-old ginseng plants. (C) Expression patterns of *PgFLA* genes in the main roots of 42 four-year-old ginseng landraces. All expression levels were normalized for analysis, and clustering was performed using the Markov clustering algorithm. #

Among the 14 tissues, *PgFLA02*, *PgFLA06*, *PgFLA08-01*, *PgFLA09-02*, *PgFLA09-03*, *PgFLA09-04*, *PgFLA09-06*, and *PgFLA09-07* displayed significantly higher transcript abundance in underground organs than in aerial tissues, suggesting their potential involvement in the growth, development, and physiological metabolism of underground structures of plants. Three transcript isoforms derived from *PgFLA03* showed divergent expression patterns: *PgFLA03-01* and *PgFLA03-02* were barely detectable in most tissues, whereas *PgFLA03-03* was expressed in all 14 tissues (Figure 4B). This expression feature was consistent with its independent clustering in a previous phylogenetic analysis, further supporting the substantial structural and functional divergence of this transcript. *PgFLA04* was highly expressed in seeds and fruit flesh, implying its putative roles in seed and fruit development. Two transcript variants of *PgFLA05* were preferentially expressed in underground tissues compared to the aerial parts. Although they shared similar expression trends, their expression levels differed markedly, with *PgFLA05-01* accumulating far more transcripts than *PgFLA05-02*, indicating a distinct transcriptional preference between isoforms. *PgFLA07* maintained abundant transcription in all examined tissues, suggesting constitutive functions associated with fundamental cellular metabolism.

Significant expression divergence was also observed in 42 ginseng landraces (Figure 4C). In accordance with the above findings, *PgFLA03-01*, *PgFLA03-02*, *PgFLA08-08*, *PgFLA09-05*, *PgFLA09-08*, *PgFLA09-01*, and *PgFLA01* were silenced in more than half of the landraces, whereas *PgFLA07* maintained a stable high expression across all accessions. An intriguing observation was made: paralogous gene pairs residing on the identical chromosome, including *PgFLA06/PgFLA07* and *PgFLA08-01/PgFLA08-02*, displayed divergent expression patterns, in which one gene exhibited high expression abundance, whereas its counterpart showed low expression or nearly undetectable transcription levels.

### 3.5. Cis-Acting Element Analysis of *PgFLA* Genes

Cis-acting element prediction revealed that the expression of *PgFLA* gene family members is modulated by diverse environmental cues (including light, temperature, water availability, and oxygen concentration) and multiple phytohormones (methyl jasmonate, abscisic acid, salicylic acid, auxin, and gibberellin). Summarizing the cis-element profiles across all family members, light-responsive elements were the most abundant (182 copies), followed by methyl jasmonate (MeJA)-responsive elements (60 copies), with anaerobic induction elements ranking third (41 copies in total) (Figure 5).

All members of the *FLA* gene family in *P. ginseng* harbored light-responsive cis-elements, and 80% possessed MeJA-responsive motifs. Only four genes (*PgFLA02*, *PgFLA05-01*, *PgFLA05-02*, and *PgFLA08-02*) lacked these MeJA-inducible elements. Nevertheless, subsequent expression assays demonstrated that *PgFLA02*, *PgFLA05-01*, and *PgFLA05-02* exhibited significant transcriptional responses to MeJA treatment. This observation indicates that the MeJA-dependent induction of these three genes does not rely on the canonical SCF<sup>COI1</sup>-JAZ-MYC2 jasmonate signaling cascade; instead, their transcription may be governed by alternative transcription factors that perceive jasmonate signals.

Notably, 90% of the *PgFLA* family members contained anaerobic induction elements, with copy numbers ranging from one to nine. Only *PgFLA04* and *PgFLA08-01* lacked such motifs. The *PgFLA07* promoter carried nine anaerobic induction elements, implying that this gene likely plays a vital role in plant tolerance to waterlogging stress.

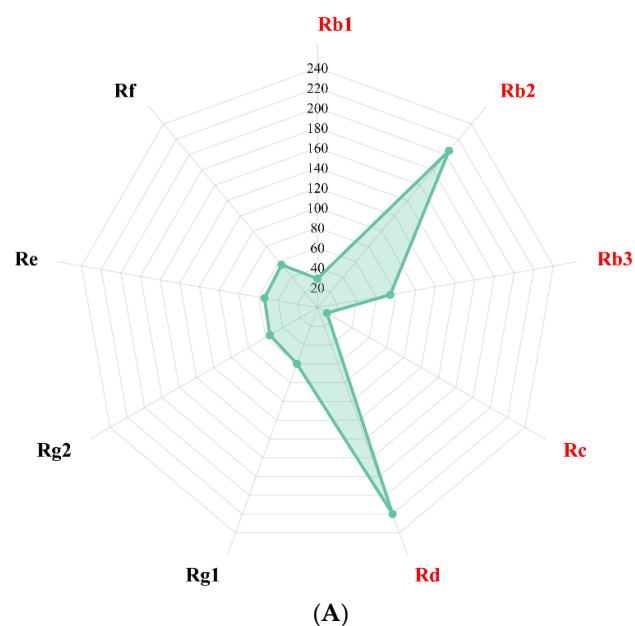
As a crucial gene family involved in signal perception, 45% of the *PgFLA* members contained drought-responsive elements, 65% carried defense and stress responsiveness motifs, and 40% possessed low temperature-responsive elements. These data further demonstrate that subsets of *PgFLA* proteins broadly participate in plant processes, including osmotic homeostasis, mechanical defense, and temperature acclimation.

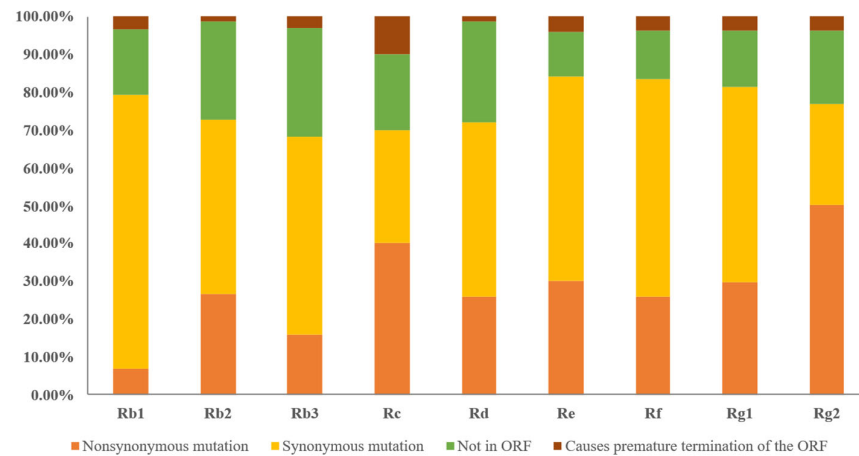
<https://doi.org/10.3390/agriculture16151679>

associated with protopanaxadiol-type ginsenosides (Rb1, Rb2, Rb3, Rc, and Rd), accounting for 97.94% of all significant variants (Figure 6A). In contrast, protopanaxatriol-type ginsenosides corresponded to only 82 significant loci (33.74%). These findings indicate that SNP/InDel mutations in *PgFLA* genes preferentially participate in the genetic regulation of protopanaxadiol-type ginsenoside biosynthesis.

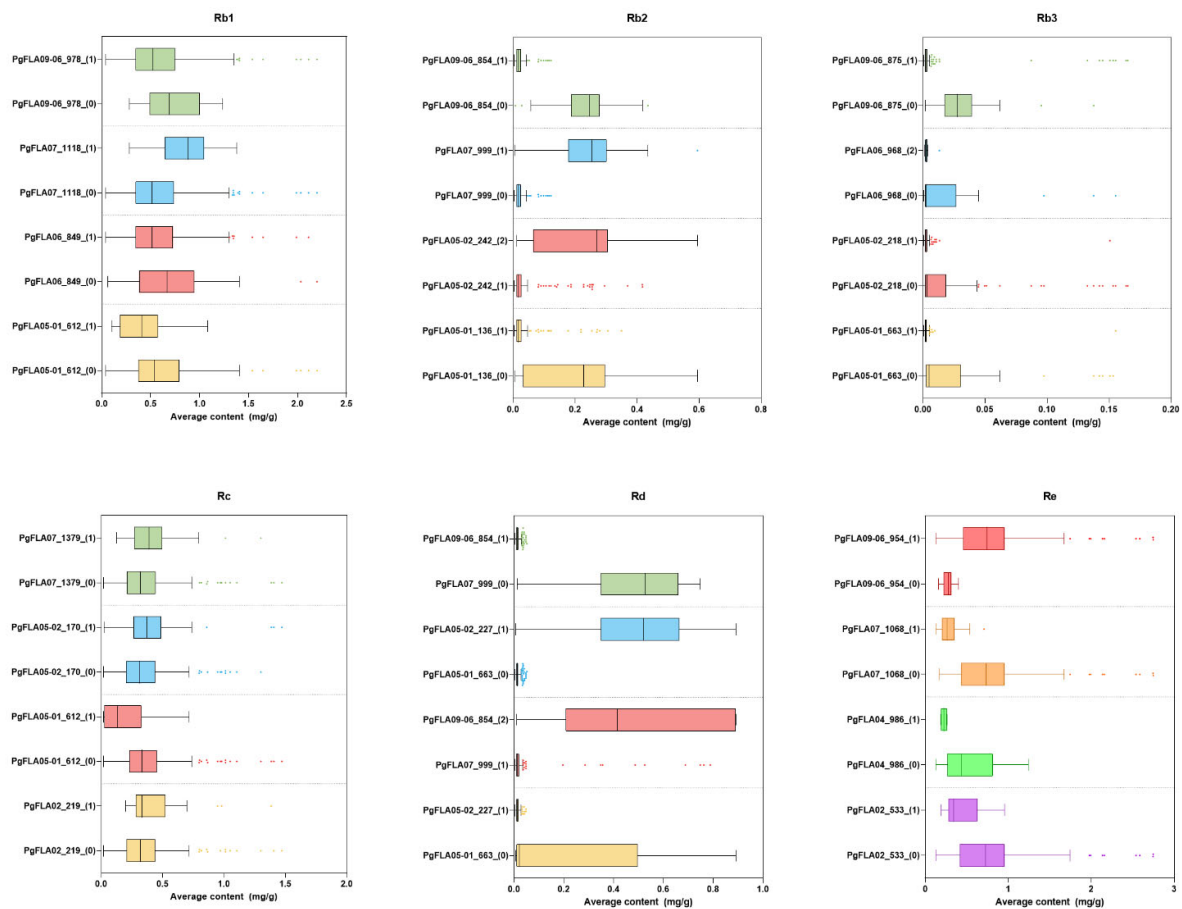
Of the 248 significant variant sites, 47 loci from five genes (*PgFLA02*, *PgFLA05-01*, *PgFLA05-02*, *PgFLA07*, and *PgFLA09-06*) significantly altered the accumulation of at least five individual ginsenosides, and variants in *PgFLA05-02* and *PgFLA07* had significant effects on eight or more ginsenoside monomers. An additional 41 loci exhibited significant associations with only one type of ginsenoside monomer.

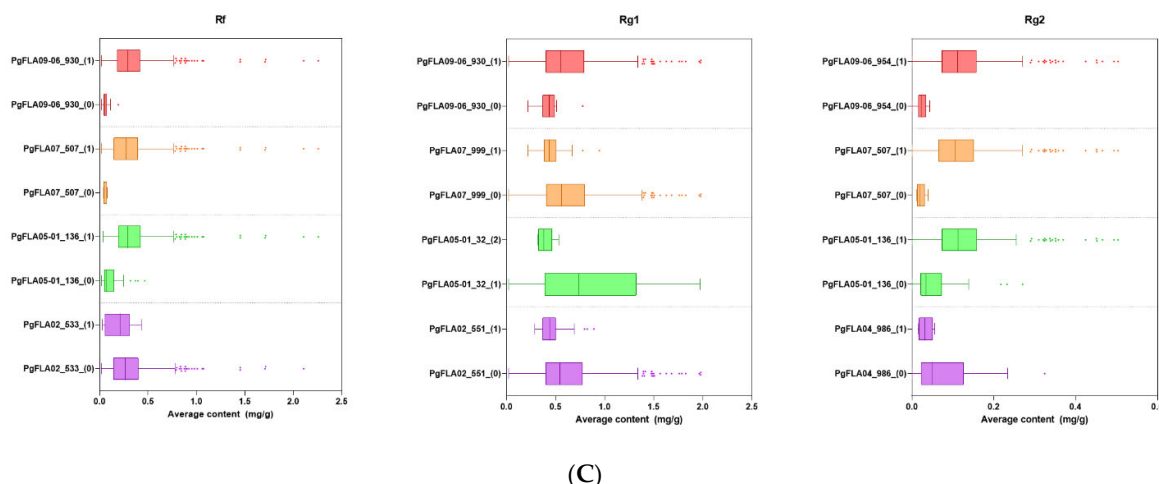
The classification of mutation types yielded 118 synonymous mutations (47.58%) and 58 non-synonymous mutations (23.39%). Seventy variants (28.23%) were mapped to non-ORF genomic regions, and two variants (0.81%) caused premature termination of the reading frames (Figure 6B,C). Distinct mutation categories modulate ginsenoside accumulation via different mechanisms. Non-ORF variants alter gene transcription by interfering with mRNA stability, alternative splicing, and epigenetic modifications. Non-synonymous substitutions change amino acid sequences and potentially disrupt protein conformation and biological function. Synonymous mutations modulate translational efficiency by modifying the mRNA secondary structure, codon usage bias, or m6A methylation, ultimately influencing downstream ginsenoside production.





(B)





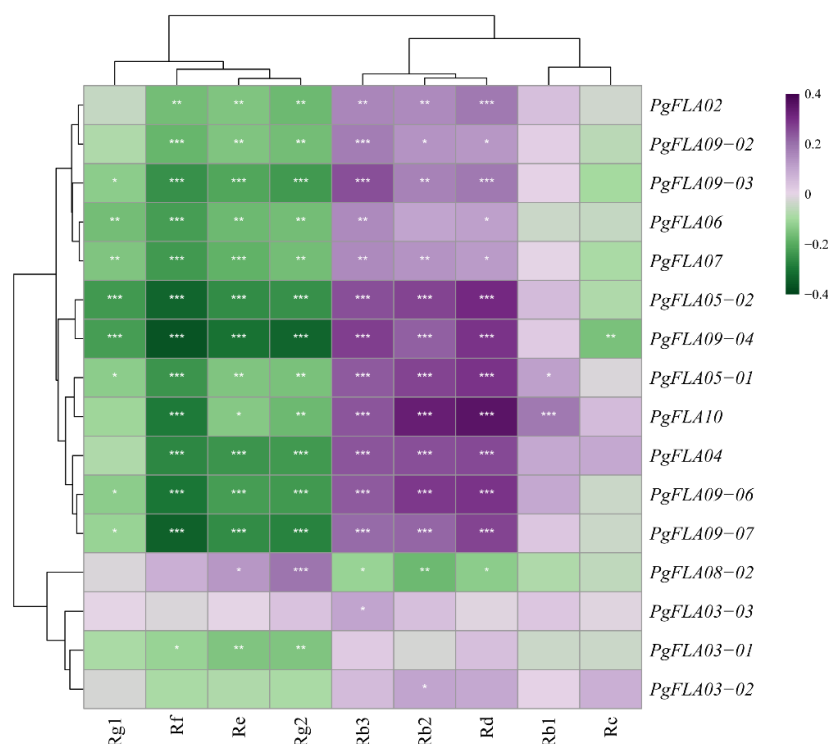
**Figure 6.** Analysis of the regulatory effects of SNP/InDels in *PgFLA* genes on ginsenosides. (A) Significantly affected the distribution of the number of loci for each ginsenoside monomer. (B) Significantly affected the effect and proportion of each ginsenoside locus. (C) Biological effects of significantly influential SNPs/InDels on the content of each ginsenoside monomer.

### 3.7. Correlation Analysis Between *PgFLA* Transcript Abundance and the Contents of Nine Ginsenosides

Correlation analysis between *PgFLA* gene expression levels and ginsenoside content revealed that 16 of the 21 *PgFLA* genes exhibited statistically significant correlations with ginsenoside accumulation, accounting for 76.2% of the total number of genes analyzed. This suggests that most *PgFLA* family members play essential regulatory roles in ginsenoside biosynthesis and its accumulation.

*PgFLA09-04* and *PgFLA05-01* were significantly correlated with eight individual ginsenoside subtypes simultaneously, and another six genes showed significant correlations with seven ginsenosides, representing 37.5% of all significantly associated genes. These observations suggest that these genes are likely involved in the synergistic regulation of multiple ginsenoside biosynthetic pathways. Among the 16 correlated genes, 93.75% displayed a consistent expression trend, with significantly negative correlations with protopanaxatriol-type (PPT) ginsenosides and significantly positive correlations with protopanaxadiol-type (PPD) ginsenosides. In contrast, only *PgFLA08-02* followed an opposite regulatory pattern, with positive correlations with PPT-type ginsenosides and negative correlations with PPD-type ginsenosides, implying its unique physiological function in ginsenoside metabolism (Figure 7).

Moreover, the three transcript isoforms of *PgFLA03* were divided into two distinct correlation patterns. *PgFLA03-01* and *PgFLA03-02* exhibited identical correlation trends, whereas *PgFLA03-03* exhibited completely reversed correlation trends. This finding was consistent with previous results from phylogenetic tree construction, gene structure characterization, and tissue-specific expression profiling, further verifying the substantial functional divergence among different transcripts derived from identical genomic loci.



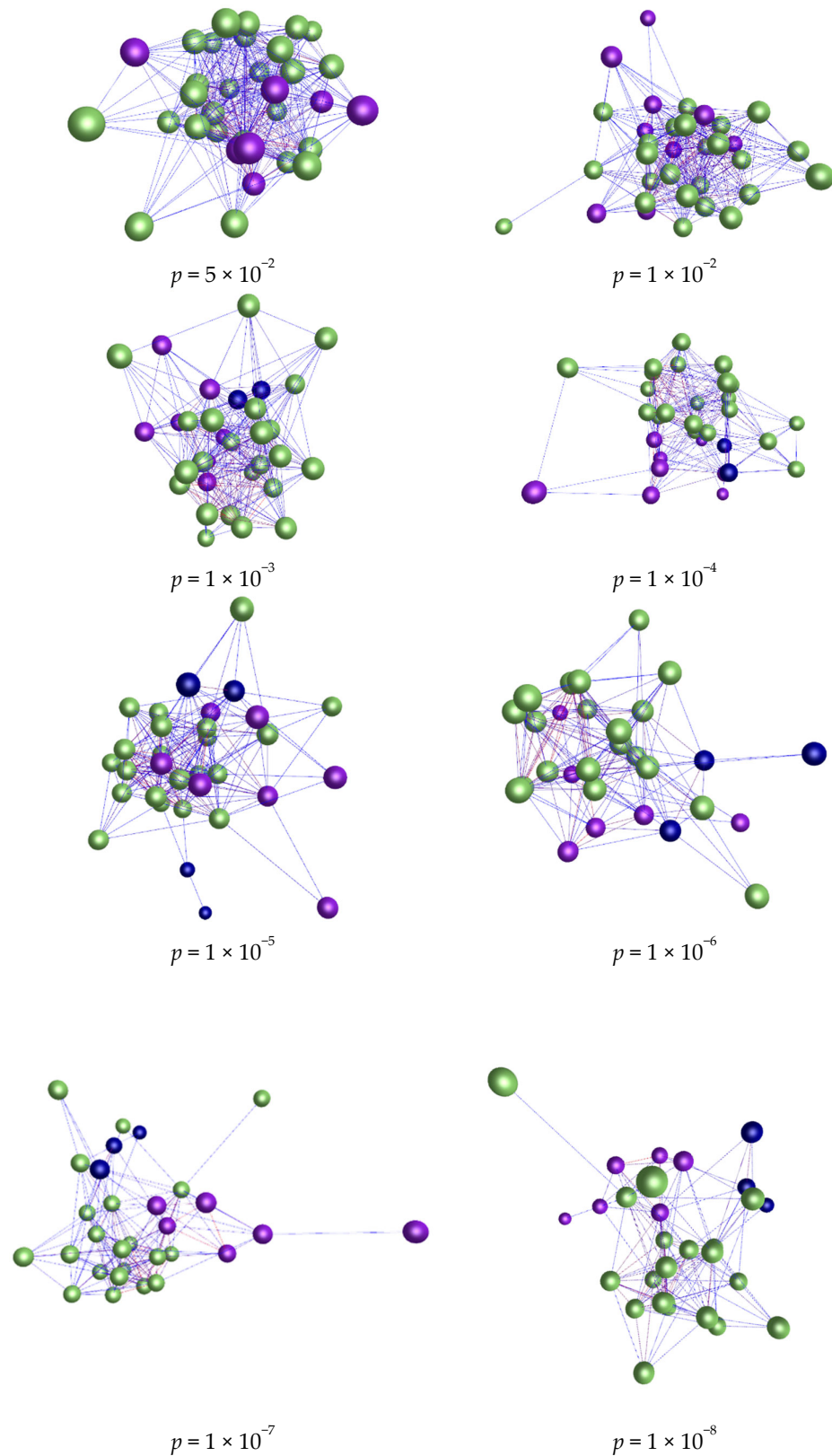
**Figure 7.** Correlation analysis between *PgFLA* gene expression and the contents of nine ginsenosides in 344 ginseng landraces. Only genes and ginsenosides with significant correlations are shown in the figure. \* indicates significant correlation at  $p \leq 0.05$ ; \*\* indicates extremely significant correlation at  $p \leq 0.01$ ; \*\*\* indicates extremely significant correlation at  $p \leq 0.001$ .

### 3.8. Co-Expression Network Analysis Between *PgFLA* Genes and Key Enzyme Genes for Ginsenoside Biosynthesis

Correlation analysis of transcript abundances between *PgFLA* genes and ginsenoside-biosynthetic key enzyme genes showed that at a threshold of  $p = 0.05$ , *PgFLA01*, *PgFLA09-01*, *PgFLA09-05*, and *PgFLA09-08* did not establish significant co-expression interactions with enzyme genes. The remaining 17 *PgFLA* genes formed eight co-expression networks with 17 key enzyme genes (Table S9), which were grouped into two distinct clusters. When the screening threshold was tightened to  $p = 1 \times 10^{-8}$ , 15 *PgFLA* genes still retained significant interactions with 11 enzyme genes, and the network topology was maintained as two separate clusters (Figure 8).

Notably, the three transcript isoforms of *PgFLA03* exhibited obvious modular divergence in the network. All three isoforms clustered into the same module ( $p = 0.05$ ). After the threshold was raised to  $p = 0.01$ , *PgFLA03-03* formed an independent cluster, whereas *PgFLA03-01* and *PgFLA03-02* were assigned to an alternative cluster, indicating that the transcriptional behavior of *PgFLA03-03* diverged substantially from that of the two other transcripts. Collectively, these results demonstrate the extensive interactions and prominent functional divergence among *PgFLA* genes.

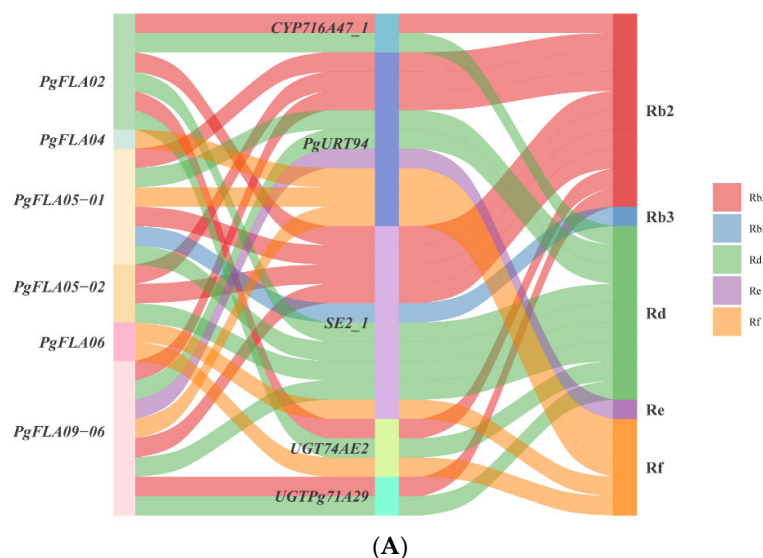
Genes from the three datasets were subsequently intersected: Set A (genes screened by SNP/InDel-ginsenoside association analysis), Set B (genes with significant expression-ginsenoside correlations), and Set C (genes with remarkable co-expression with key biosynthetic enzymes). Eight overlapping core genes were identified that displayed significant associations with ginsenoside accumulation at the genomic variation, transcriptional expression, and co-regulatory interaction levels.

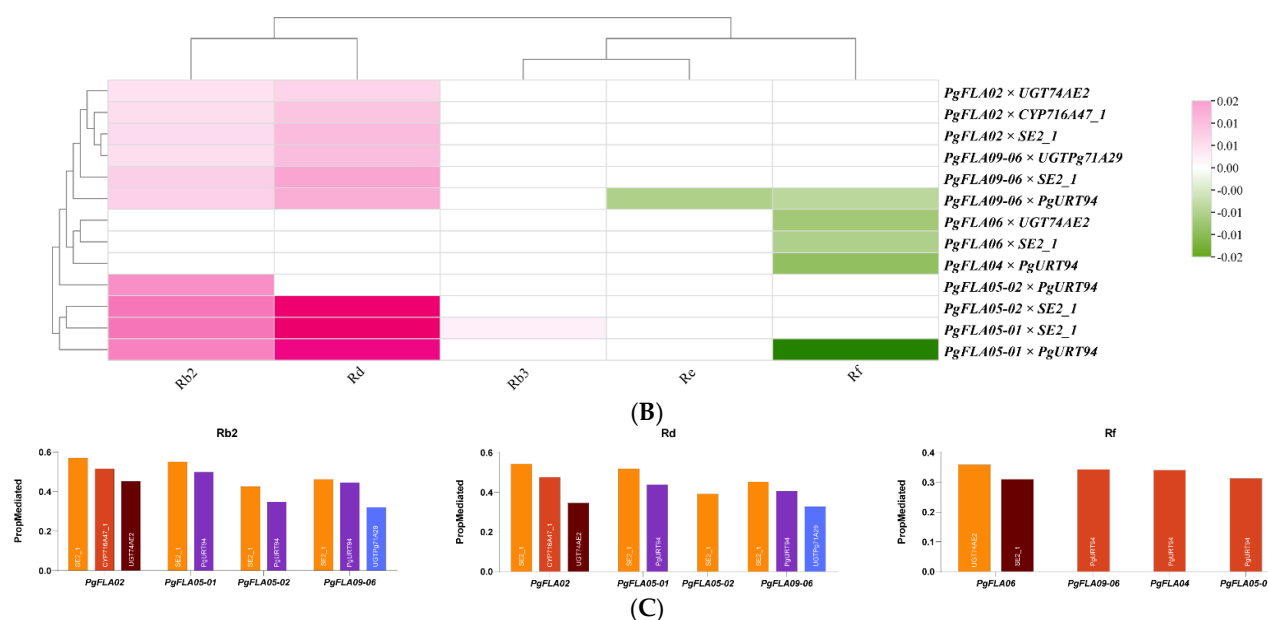


**Figure 8.** Analysis of the co-expression network between *PgFLA* genes and 17 key enzyme genes under different  $p$  values; nodes in distinct colors denote genes with analogous expression patterns.

### 3.9. Mediation Effect Analysis Among PgFLA Genes, Key Enzymes, and Ginsenosides

Statistical calculations based on mediation effect analysis revealed that the six screened candidate genes (*PgFLA02*, *PgFLA04*, *PgFLA05-01*, *PgFLA05-02*, *PgFLA06*, and *PgFLA09-06*) significantly altered ginsenoside contents by regulating the expression of key enzyme genes. (Figure 9A). Five key enzymes served as statistically significant mediating variables, along with five dependent variables corresponding to individual monomeric ginsenosides: Rb2, Rb3, Rd, Re, and Rf. Among all the genes tested, *PgFLA05-01* and *PgFLA09-06* regulated the greatest number of monomeric ginsenoside species, with each gene significantly affecting the accumulation of four monomeric ginsenosides by modulating the expression of the key enzymes. Both *PgFLA02* and *PgFLA05-02* indirectly promoted the accumulation of protopanaxadiol-type ginsenosides (Rb2 and Rd) by transcriptionally regulating the relevant key enzymes (Figure 9B). The results revealed that *PgFLA05-01* and *PgFLA09-06* exerted opposite regulatory effects on the expression of *PgURT94*, and such divergent positive/negative modulation specifically differentiated between protopanaxadiol-type and protopanaxatriol-type ginsenosides, respectively. This finding suggests that these two genes maintain a dynamic balance between these two categories of ginsenosides by regulating the transcriptional abundance of *PgURT94*. Notably, all five genes displayed detectable mediating effects in the mediation effect analysis, whereby they modulated the expression of the key enzyme gene *SE2\_1* and thereby altered the accumulation of multiple ginsenosides. The enzyme encoded by *SE2\_1* catalyzes the conversion of squalene to 2,3-oxidosqualene, a critical intermediate step in dammarenediol biosynthesis and subsequent diversification into diverse ginsenoside derivatives (Table S7). Accordingly, we inferred that these *FLA* genes predominantly regulate ginsenoside biosynthesis via this metabolic pathway (Figure 9C).





**Figure 9.** Mediation effect analysis among *PgFLA* genes, key enzyme genes, and ginsenosides. (A) Regulatory pathways among *PgFLA* genes, key enzyme genes, and ginsenosides with significant mediation effects. (B) Average causal mediation effect of the significant regulatory pathways. Gene combinations corresponding to the white blocks showed no significant mediation effects. (C) Proportion of the mediated effect (PropMediated) for pathways significantly regulating the three monomeric ginsenosides.

### 3.10. Expression Profiling Analysis of Core *PgFLA* Genes Under MeJA Treatment

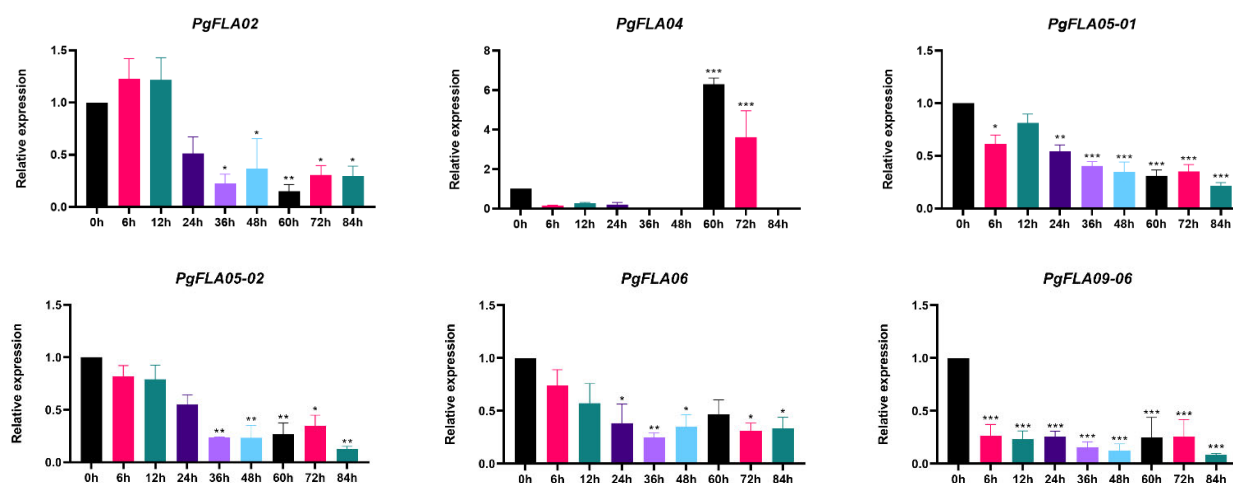
Quantitative real-time PCR (qRT-PCR) was used to detect the expression dynamics of *PgFLA* genes in ginseng adventitious roots at multiple time points after MeJA elicitation. The quantification results revealed that the gene abundance of five of the six *PgFLA* genes was significantly downregulated after MeJA treatment (Figure 10), whereas only *PgFLA04* exhibited remarkable upregulation at 60 h post-treatment.

Notably, *PgFLA05-01* and *PgFLA09-06* were significantly induced as early as 6 h after treatment, classifying them as early rapid-response genes. Significant transcriptional alterations were detected for *PgFLA06* at 24 h, whereas clear responses were only observed for *PgFLA02* and *PgFLA05-02* at 36 h. This observation may indicate that these genes exert their biological functions downstream of the jasmonate signal transduction cascade. Combined analyses of promoter cis-acting elements further demonstrated that eight MeJA-responsive cis-elements were distributed within the promoter region of *PgFLA09-06*, whereas only four MeJA cis-elements were identified in the promoter of *PgFLA06*. The discrepancy in the number of hormone-responsive motifs may be a critical factor that enables *PgFLA09-06* to respond to MeJA signals at an earlier time point. Although no canonical MeJA cis-acting elements were detected in the promoter of *PgFLA05-01*, this gene was rapidly induced at 6 h upon elicitation. It is speculated that the transcriptional activation of this gene may occur independent of MeJA-specific cis-elements and may be rapidly triggered via binding sites for core jasmonate signaling transcription factors, such as MYC2, harbored in its promoter sequence.

For *PgFLA04*, eight MeJA-responsive cis-elements were also identified in its promoter; however, its transcriptional response was markedly delayed, with significant induction detected only at 36 h post-treatment. Combined with tissue expression profiling data, *PgFLA04* maintained low transcript levels in most ginseng tissues. This finding suggests that this gene may not participate in early rapid stress responses and may instead

function as a late-responsive gene located far downstream of the MeJA signaling pathway. Given its opposite expression trend relative to other *PgFLA* members, it was hypothesized that this gene may participate in the termination of jasmonate signal transduction and act as a negative regulator of this signaling cascade.

Collectively, these results further confirmed that pronounced functional differentiation and distinct temporal response patterns may exist among members of the *PgFLA* gene family within the MeJA signaling pathway.

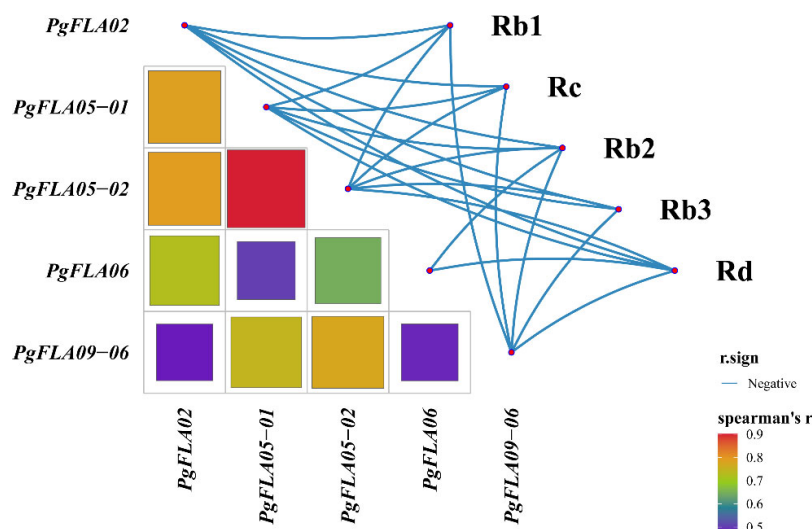


**Figure 10.** Expression pattern analysis of six *PgFLA* candidate genes in ginseng adventitious roots under MeJA treatment. \* indicates a significant difference at  $p \leq 0.05$ ; \*\* indicates an extremely significant difference at  $p \leq 0.01$ ; \*\*\* indicates an extremely significant difference at  $p \leq 0.001$ .

### 3.11. Correlation Analysis Between *PgFLA* Gene Expression and Ginsenoside Contents Under MeJA Induction

Correlation analysis revealed that the transcript abundances of *PgFLA02*, *PgFLA05-01*, *PgFLA05-02*, *PgFLA06*, and *PgFLA09-06* were significantly negatively correlated with the levels of protopanaxadiol-type ginsenosides, including Rb1, Rc, Rb2, Rb3, and Rd, following MeJA treatment (Figure 11). Pairwise expression correlation analysis revealed significant positive correlations among these five *PgFLA* genes, with correlation coefficients ranging from 0.47 to 0.90. The correlation coefficient between *PgFLA05-01* and *PgFLA05-02* was as high as 0.90, indicating an extremely strong positive correlation between them. Consistent expression patterns of the two genes were previously observed across 42 ginseng landraces, which supported their coordinated participation in identical biological processes and agreed with their synchronous transcriptional responses to MeJA. The correlation coefficients between *PgFLA02* and *PgFLA05* and between *PgFLA05* and *PgFLA09-06* were greater than 0.70. Combined with phylogenetic clustering data, these genes fell into the same phylogenetic clade, suggesting high sequence homology, close evolutionary relatedness, and conserved functions.

Notably, *PgFLA02* and *PgFLA06* also shared a correlation coefficient above 0.70 and exhibited highly coordinated co-expression, despite their separation into distinct phylogenetic branches, implying functional convergence with divergent evolutionary origins. Promoter cis-element assays revealed no canonical MeJA-responsive motifs in *PgFLA02*, whereas multiple MeJA-related cis-elements were enriched in the promoter of *PgFLA06*. Their nearly identical expression profiles under MeJA induction indicate that both genes are likely modulated by common upstream transcription factors in the MeJA signaling pathway to achieve synchronized transcription.



**Figure 11.** Correlation analysis between five *PgFLA* gene expressions and five protopanaxadiol-type monomeric ginsenoside contents under MeJA induction. Blue connecting lines between genes and ginsenosides indicate negative correlations, and all displayed lines represent statistically significant correlations between gene expression and ginsenoside content at  $p \leq 0.05$ .

#### 4. Discussion

In the present study, 21 members of the ginseng *FLA* gene family were identified using homologous sequence alignment, protein structural characterization, and conserved domain annotation. These *PgFLA* genes were unevenly distributed across nine ginseng chromosomes. Chromosomal localization and synteny analyses revealed that tandem and segmental duplication were the major driving forces underlying *PgFLA* family expansion. The gene clusters and syntenic blocks between Chr 18 and Chr 20, as well as the conserved syntenic segments spanning Chr 04 and Chr 10, corroborated the essential contribution of whole-genome or large-scale genomic duplication to *PgFLA* evolutionary expansion, and such duplication-driven expansion patterns have been widely documented for *FLA* families across numerous plant species [35,36]. Phylogenetic clustering grouped *PgFLA02* and *PgFLA05* into a single clade, indicative of high sequence similarity and close orthology; hence, these two paralogs presumably execute analogous biological functions [11,37]. Tissue expression profiling revealed highly consistent transcription signatures for *PgFLA02* and *PgFLA05* across 14 distinct ginseng tissues, with both genes exhibiting significantly higher transcript abundances in underground organs than in aerial tissues. Cis-acting element prediction further demonstrated remarkable conservation in promoter architecture between the two genes: neither possessed canonical methyl jasmonate (MeJA)-responsive cis-elements, whereas both were enriched with abscisic acid- and gibberellin-responsive regulatory motifs. Moreover, co-expression network analysis placed *PgFLA02* and *PgFLA05* within identical regulatory modules, providing additional evidence of functional redundancy.

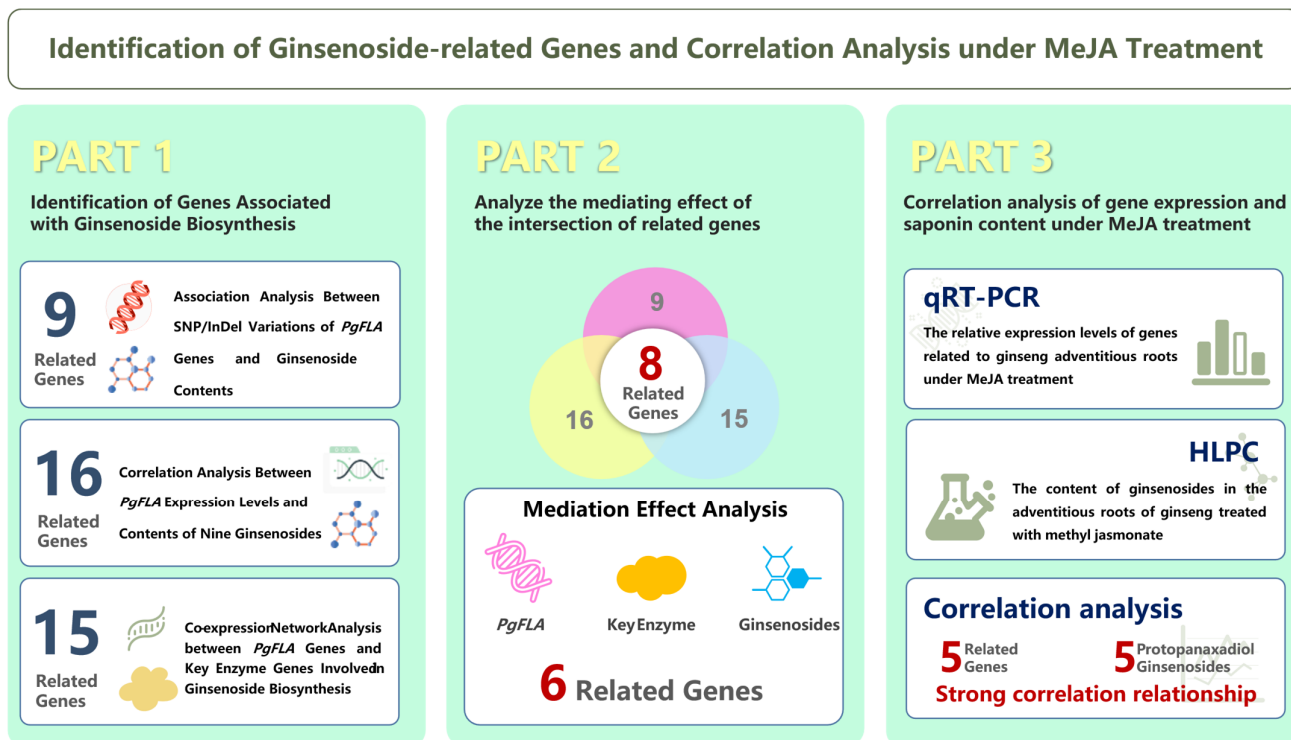
*PgFLA04* was preferentially transcribed in seeds, implying its involvement in ginseng sexual reproductive development. Consistent findings have been reported in cotton: CRISPR/Cas9-mediated knockout of *GhFLA19* in upland cotton resulted in male sterility [38], whereas heterologous overexpression of *Arabidopsis FLA14* caused shrunken, morphologically aberrant pollen grains in up to 39.2% of pollen populations and drastically reduced seed-setting rates [39]. Collectively, these reports suggest an evolutionarily conserved role for the *FLA* family in modulating anther, pollen, and seed development across

plant lineages, rendering them indispensable regulators of plant reproductive growth. *PgFLA07* was constitutively highly expressed in all 14 tested tissues and root samples from 42 ginseng landraces, a phenotype strongly associated with its encoded protein harboring two tandem FAS1 domains. Its Arabidopsis ortholog *AtFLA16* also contains dual FAS1 domains and is specifically enriched in cells undergoing secondary cell wall biosynthesis [40]. The striking overlap in tissue-specific expression between *PgFLA07* and *AtFLA16* suggests that the dual-FAS1 structural feature confers ubiquitous expression in secondary-wall-forming tissues, and *PgFLA07* most likely retains a conserved function governing secondary cell wall formation in ginseng. SNP/InDel-ginsenoside association analysis indicated that natural sequence polymorphisms within *PgFLA* genes significantly altered ginsenoside accumulation, with variants preferentially regulating the biosynthesis of protopanaxadiol-type (PPD) ginsenosides.

In the present study, a multidimensional screening strategy was adopted to improve the reliability of candidate genes. Conventional gene function prediction relies on a single association test or expression profiling, which is prone to false-positive outcomes. Three lines of evidence from distinct layers were integrated in this study to achieve this goal. First, an association analysis between SNP/InDel polymorphisms and the contents of nine predominant ginsenosides across 344 ginseng accessions identified trait-linked genes at the genomic variation level. Second, correlation analysis linking gene transcription abundance to ginsenoside accumulation among diverse landraces validated the coordinated variation between gene expression and saponin biosynthesis at the transcriptional level. Third, the construction of co-expression networks between *PgFLA* and key ginsenoside biosynthetic enzyme genes uncovered potential regulatory interactions between candidate genes and functionally characterized enzymes from a systems biology perspective. Cross-validation of the three approaches yielded eight core candidate genes (*PgFLA02*, *PgFLA04*, *PgFLA05-01*, *PgFLA05-02*, *PgFLA06*, *PgFLA07*, *PgFLA08-01*, and *PgFLA09-06*) that were robustly associated with ginsenoside biosynthesis across variation, transcription, and co-interaction dimensions. This screening largely eliminated random screening bias and laid a solid foundation for subsequent in vivo functional studies. Omics-integrated screening pipelines of this kind have proven to be efficient for gene family investigations in numerous plant species [41–43]. Nevertheless, several limitations are inherent to this analytical approach. The screened SNPs and InDels that exerted significant effects on ginsenoside accumulation may not function independently to drive phenotypic variation. Instead, multiple SNPs may form haplotypes via linkage disequilibrium to jointly regulate phenotypic traits, which represents a major aspect requiring further improvement in this study. Conventional association approaches have inherent limitations: they only detect statistical correlations between genetic loci and phenotypic traits but fail to resolve causal relationships or the directional regulatory cascade, restricting mechanistic dissection of how candidate genes modulate the formation of metabolites. To address this bottleneck, mediation effect analysis [44], a rigorous statistical framework, was introduced beyond routine correlation analyses to resolve the hierarchical regulatory pathways linking target genes, intermediate regulatory enzymes, and the final ginsenoside phenotypes.

Promoter cis-element profiling provided fundamental clues for deciphering the transcriptional regulatory network governing *PgFLA* expression. The present results demonstrated that *PgFLA* promoters were enriched with cis-motifs related to light responsiveness, phytohormone signaling (abscisic acid, methyl jasmonate, auxin, etc.), and abiotic stress adaptation; such conserved features have also been documented in *FLA* families of soybean, maize, and sorghum [45–47]. Accumulating functional evidence has illustrated the universal stress response of *FLA* genes. Under salt stress, the Arabidopsis *SOS5* mutant displays swollen root tips and exhibits arrested root elongation [8]. After cold acclimation, transgenic banana lines overexpressing *MaFLA27* exhibited reduced pectin methyl

esterification and lower reactive oxygen species (ROS) accumulation than wild-type controls [48]. To date, most published *FLA* research has focused on cold and salt stress responses, whereas investigations into phytohormone-triggered transcriptional changes are limited. The exogenous MeJA application used in this study addresses this research gap. As a pivotal endogenous phytohormone, MeJA strongly triggers ginsenoside biosynthesis in ginseng plants. In this experiment, five of the six screened candidate genes were significantly transcriptionally repressed upon MeJA elicitation, whereas only *PgFLA04* underwent late-stage upregulation after treatment. Combined with the variable copy numbers of MeJA-responsive cis-elements in the respective promoter sequences, these divergent induction patterns indicate the disparate biological roles and temporally separated functions of *PgFLA* paralogs within the MeJA signaling cascade. For instance, *PgFLA09-06*, which harbors eight MeJA-responsive motifs, was rapidly induced at an early treatment stage, whereas *PgFLA06*, which has fewer relevant cis-elements, responded much later after treatment. Notably, *PgFLA05-01*, which lacks canonical MeJA cis-elements, exhibited an early and rapid transcriptional response. These results suggest that *PgFLA05-01* is modulated via an indirect, cis-element-independent regulatory pathway, potentially through physical interactions with master transcription factors, such as MYC2. However, these results only revealed that the expression of *PgFLA* genes was significantly altered in response to MeJA, and significant correlations were detected between ginsenoside content and the screened *PgFLA* genes under MeJA treatment (Figure 12). However, such correlations were established only at the biostatistical level, and the authentic functional linkage between *PgFLA* genes and ginsenoside biosynthesis has not been verified via biochemical or genetic experiments. This aspect may serve as a vital research direction for further in-depth exploration in future studies.



**Figure 12.** Schematic representation of *PgFLA* candidate genes and ginsenoside biosynthesis analysis under MeJA treatment.

Collectively, these observations lay a solid foundation for future mechanistic dissection of the link between *PgFLA* genes, phytohormone signaling, and ginsenoside biosynthesis.

## 5. Conclusions

In summary, the present study systematically dissected the evolutionary characteristics, structural diversity, and expression patterns of the *PgFLA* gene family in *P. ginseng*. Candidate genes markedly associated with ginsenoside biosynthesis were identified from multiple dimensions, and their relevance to ginsenoside synthesis under MeJA regulation was further verified experimentally. These findings not only deepen the understanding of the biological functions of plant *FLA* family genes but also provide valuable candidate gene resources and theoretical foundations for improving ginseng quality and elevating ginsenoside content via molecular breeding strategies. Further in-depth functional validation assays will be performed on the screened candidate genes to clarify the precise roles of individual *PgFLA* members within the regulatory network of ginseng secondary metabolism. This study provides a new theoretical framework and offers innovative perspectives for the development and utilization of high-quality ginseng germplasm resources.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/agriculture16151679/s1>, Table S1. Basic information of the *PgFLA* gene family in ginseng. Table S2. The identified genes used as evolutionary controls for the *PgFLA* gene phylogenetic analysis. Table S3. The TPM expressions of the *PgFLA* genes in 42 agricultural cultivars of ginseng roots, 14 tissues of 4-year-old ginseng, and 4 different 1-year-old ginseng roots. Table S4. The *PgFLA* genes contain SNP/InDels mutations that significantly affect the content of nine monomeric saponins (Rb1, Rb2, Rb3, Rc, Rd, Re, Rg1, Rg2, and Rf) and the effect of their gene SNPs/InDels on the content of the nine ginsenosides. Table S5. Significant loci from association analysis between *PgFLA* genes and nine types of ginsenosides, significantly associated ginsenoside subtypes, genotype comparison groups, raw significant *p*-values, and FDR values corrected by the BH (Benjamini–Hochberg) method. Table S6. The correlation coefficients calculated for the co-expression network analysis were derived from a total of 344 samples. Table S7. Mediation effect analysis results of genes related to ginsenoside biosynthesis. Table S8. Primer information used for quantitative real-time PCR (qRT-PCR) in this study. Table S9. Gene IDs and GO functional annotations of 17 key enzyme genes involved in the ginsenoside biosynthetic pathway of *Panax ginseng*.#

**Author Contributions:** K.W., Y.W., and M.Z. (Meiping Zhang) designed the experiments of the study. A.W. and K.W. wrote and revised the main manuscript. A.W., X.H. (Xiaoqian Han), X.H. (Xinyue Hu), L.X., and M.Z. (Mingzhu Zhao) performed the experiments and contributed to data analysis. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by an award from the Department of Science and Technology of Jilin Province (20250205032GH) and the Jilin Agricultural University Undergraduate Students Innovation and Entrepreneurship Project (2025, 2026).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Mancuso, C.; Santangelo, R. *Panax ginseng* and *panax quinquefolius*: From pharmacology to toxicology. *Food. Chem. Toxicol.* **2017**, *107*, 362–372. <https://doi.org/10.1016/j.fct.2017.07.019>.
- Lee, M.H.; Jeong, J.H.; Seo, J.W.; Shin, C.G.; Kim, Y.S.; In, J.G.; Yang, D.C.; Yi, J.S.; Choi, Y.E. Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* overexpressing squalene synthase gene. *Plant Cell Physiol.* **2004**, *45*, 976–984. <https://doi.org/10.1093/pcp/pch126>.
- Geng, X.; Wang, J.; Liu, Y.; Liu, L.; Liu, X.; Zhao, Y.; Wang, C.; Liu, J. Research progress on chemical diversity of saponins in panax ginseng. *Chin. Herb. Med.* **2024**, *16*, 529–547. <https://doi.org/10.1016/j.chmed.2024.08.005>.
- Tan, L.; Showalter, A.M.; Egelund, J.; Hernandez-Sanchez, A.; Doblin, M.S.; Bacic, A. Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* **2012**, *3*, 140. <https://doi.org/10.3389/fpls.2012.00140>.
- Ayorinde, D.A.; Olanrewaju, G.O.; Showalter, A.M. Transcriptomic insights into arabinogalactan protein mechanism of action in galactosyltransferase octuple mutants. *Front. Plant Sci.* **2026**, *16*, 1706954. <https://doi.org/10.3389/fpls.2025.1706954>.
- Kapusta, M.; Narajczyk, M.; Plachno, B.J. Arabinogalactan proteins mark the generative cell–vegetative cell interface in monocotyledonous pollen grains. *Cells* **2025**, *14*, 1549. <https://doi.org/10.3390/cells14191549>.
- Johnson, K.L.; Jones, B.J.; Bacic, A.; Schultz, C.J. The fasciclin-like arabinogalactan proteins of Arabidopsis. A multigene family of putative cell adhesion molecules. *Plant Physiol.* **2003**, *133*, 1911–1925. <https://doi.org/10.1104/pp.103.031237>.
- Shi, H.; Kim, Y.; Guo, Y.; Stevenson, B.; Zhu, J.K. The arabidopsis *SOS5* locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell* **2002**, *15*, 19–32. <https://doi.org/10.1105/tpc.007872>.
- Borner, G.H.H.; Lilley, K.S.; Stevens, T.J.; Dupree, P. Identification of glycosylphosphatidylinositol-anchored proteins in arabidopsis. A proteomic and genomic analysis. *Plant Physiol.* **2003**, *132*, 568–577. <https://doi.org/10.1104/pp.103.021170>.
- Ma, Y.; Shafee, T.; Mudiyansele, A.M.; Ratcliffe, J.; MacMillan, C.P.; Mansfield, S.D.; Bacic, A.; Johnson, K.L. Distinct functions of FASCICLIN-LIKE ARABINOGLACTAN PROTEINS relate to domain structure. *Plant Physiol.* **2023**, *192*, 119–132. <https://doi.org/10.1093/plphys/kiad097>.
- Macmillan, C.P.; Mansfield, S.D.; Stachurski, Z.H.; Evans, R.; Southerton, S.G. Fasciclin-like arabinogalactan proteins: Specialization for stem biomechanics and cell wall architecture in arabidopsis and eucalyptus. *Plant J. Cell Mol. Biol.* **2010**, *62*, 689–703. <https://doi.org/10.1111/j.1365-313X.2010.04181.x>.
- Niu, Z.; Bai, Q.; Lv, J.; Tian, W.; Mao, K.; Wei, Q.; Zheng, Y.; Yang, H.; Gao, C.; Wan, D. The fasciclin-like arabinogalactan protein FLA11 of *ostrea reherdiana* impacts wood formation and salt stress in populus. *Environ. Exp. Bot.* **2024**, *219*, 105651. <https://doi.org/10.1016/j.envexpbot.2024.105651>.
- Huang, H.; Miao, Y.; Zhang, Y.; Huang, L.; Cao, J.; Lin, S. Comprehensive analysis of arabinogalactan protein-encoding genes reveals the involvement of three BrFLA genes in pollen germination in brassica rapa. *Int. J. Mol. Sci.* **2021**, *22*, 13142. <https://doi.org/10.3390/ijms222313142>.
- Deng, Y.; Wan, Y.; Liu, W.; Zhang, L.; Zhou, K.; Feng, P.; He, G.; Wang, N. OsFLA1 encodes a fasciclin-like arabinogalactan protein and affects pollen exine development in rice. *TAG Theor. Appl. Genet. Theor. Angew. Genet.* **2022**, *135*, 1247–1262. <https://doi.org/10.1007/s00122-021-04028-1>.
- Chao, D.; Sun, X.; Chen, X.; Meng, R.; Zhang, B.; Liu, C.; Liu, J.; Liu, R.; Chen, Z.; Xia, J. A fasciclin-like arabinogalactan protein OsFLA8 is involved in casparian strip formation and salt tolerance in rice. *Plant Physiol. Biochem. PPB* **2025**, *227*, 110169. <https://doi.org/10.1016/j.plaphy.2025.110169>.
- Wen, S.; Huang, R.; Gao, M.; Li, J.; Zhou, Y.; Sun, R.; Bie, Z.; Cheng, J. Genome-wide identification of fasciclin-like arabinogalactan proteins (FLAs) in melon and the role of CmFLA8 in glucose accumulation in fruit. *Veg. Res.* **2025**, *5*, e009. <https://doi.org/10.48130/vegres-0025-0004>.
- Hu, J.; Wang, J.; Muhammad, T.; Tuerdiyusufu, D.; Yang, T.; Li, N.; Yang, H.; Wang, B.; Yu, Q. Functional analysis of fasciclin-like arabinogalactan in carotenoid synthesis during tomato fruit ripening. *Plant Physiol. Biochem. PPB* **2024**, *210*, 108589. <https://doi.org/10.1016/j.plaphy.2024.108589>.
- Hembram, D.B.; Panda, S.P.; Das, B.K.; Soren, D.; Singh, N.R. Microbial elicitors to metabolic reprogramming: An integrative model of plant-microbe interactions. *Front. Microbiol.* **2026**, *17*, 1816468. <https://doi.org/10.3389/fmicb.2026.1816468>.
- Kersch-Becker, M.F.; Pinheiro, G.P.; Schweller, T.S.; Pareja, M.; Pocius, V.M.; Trigo, J.R.; Romero, G.Q. Complex multitrophic species interactions and fitness costs: Intricate consequences of jasmonate and salicylate induced plant defences. *J. Anim. Ecol.* **2026**, 1–17. <https://doi.org/10.1111/1365-2656.70243>.

20. Mubeen, B.; Murtaza, S.; Yaqoob, S.; Aregbe, A.Y.; Boasiako, T.A.; Xiong, Y.; Ma, Y. Harnessing plant defense: Elicitors, hormones and immunity-driven production of medicinally valuable secondary metabolites. *S. Afr. J. Bot.* **2025**, *179*, 280–292. <https://doi.org/10.1016/j.sajb.2025.02.018>.
21. Farmer, E.E.; Ryan, C.A. Interplant communication: Airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 7713–7716. <https://doi.org/10.1073/pnas.87.19.7713>.
22. Han, J.Y.; Kim, H.J.; Kwon, Y.S.; Choi, Y.E. The cyt p450 enzyme CYP716a47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* **2011**, *52*, 2062–2073. <https://doi.org/10.1093/pcp/pcr150>.
23. Sun, C.; Li, Y.; Wu, Q.; Luo, H.; Sun, Y.; Song, J.; Lui, E.M.; Chen, S. De novo sequencing and analysis of the american ginseng root transcriptome using a GS FLX titanium platform to discover putative genes involved in ginsenoside biosynthesis. *BMC Genom.* **2010**, *11*, 262. <https://doi.org/10.1186/1471-2164-11-262>.
24. Han, J.Y.; In, J.G.; Kwon, Y.S.; Choi, Y.E. Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of squalene epoxidase gene in panax ginseng. *Phytochemistry* **2010**, *71*, 36–46. <https://doi.org/10.1016/j.phytochem.2009.09.031>.
25. Lee, Y.S.; Park, H.S.; Lee, D.K.; Jayakodi, M.; Kim, N.H.; Koo, H.J.; Lee, S.C.; Kim, Y.J.; Kwon, S.W.; Yang, T.E.J. Integrated transcriptomic and metabolomic analysis of five *Panax ginseng* cultivars reveals the dynamics of ginsenoside biosynthesis. *Front. Plant Sci.* **2017**, *8*, 1048. <https://doi.org/10.3389/fpls.2017.01048>.
26. Zhang, K.; Wang, A.; Shi, L.; Liu, J.; Zhang, Y.; Jiang, Y.; Zhang, M.; Wang, Y.; Zhao, M.; Wang, K. Genome-wide identification and integrated analysis of the WNK kinase gene family and expression analysis of PgWNK genes under methyl jasmonate treatment in *Panax ginseng*. *BMC Plant Biol.* **2025**, *25*, 791. <https://doi.org/10.1186/s12870-025-06818-x>.
27. Zhang, K.; Wang, A.; Fan, M.; Liu, J.; Ye, K.; Zhang, Y.; Zhao, M.; Zhang, M.; Wang, Y.; Zhu, L.; et al. Genome-wide identification and integrated analysis of the PYL gene family in *Panax ginseng* and analysis of the expression of PgPYL genes involved in protopanaxadiol-type ginsenoside biosynthesis under MeJA treatment. *Horticulturae* **2026**, *12*, 572. <https://doi.org/10.3390/horticulturae12050572>.
28. Jiang, T.; Zhang, Y.; Zuo, G.; Luo, T.; Wang, H.; Zhang, R.; Luo, Z. Transcription factor PgNAC72 activates DAMMARENEDIOL SYNTHASE expression to promote ginseng saponin biosynthesis. *Plant Physiol.* **2024**, *195*, 2952–2969. <https://doi.org/10.1093/plphys/kiae202>.
29. Xie, J.; Chen, Y.; Cai, G.; Cai, R.; Hu, Z.; Wang, H. Tree visualization by one table (tvBOT): A web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids. Res.* **2023**, *51*, W587–W592. <https://doi.org/10.1093/nar/gkad359>.
30. Liu, S.; Jiang, Y.; Wang, Y.; Huo, H.; Cilkiz, M.; Chen, P.; Han, Y.; Li, L.; Wang, K.; Zhao, M.; et al. Genetic and molecular dissection of ginseng (*Panax ginseng* Mey.) Germplasm using high-density genic SNP markers, secondary metabolites, and gene expressions. *Front. Plant Sci.* **2023**, *14*, 1165349. <https://doi.org/10.3389/fpls.2023.1165349>.
31. Lyu, F.; Han, F.; Ge, C.; Mao, W.; Chen, L.; Hu, H.; Chen, G.; Lang, Q.; Fang, C. OmicStudio: A composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. *iMeta* **2023**, *2*, e85. <https://doi.org/10.1002/imt2.85>.
32. Li, L.; Wang, K.; Zhao, M.; Li, S.; Jiang, Y.; Zhu, L.; Chen, J.; Wang, Y.; Sun, C.; Chen, P.; et al. Selection and validation of reference genes desirable for gene expression analysis by qRT-PCR in MeJA-treated ginseng hairy roots. *PLoS ONE* **2019**, *14*, e0226168. <https://doi.org/10.1371/journal.pone.0226168>.
33. Li, L.; Wang, Y.; Zhao, M.; Wang, K.; Sun, C.; Zhu, L.; Han, Y.; Chen, P.; Lei, J.; Wang, Y.; et al. Integrative transcriptome analysis identifies new oxidosqualene cyclase genes involved in ginsenoside biosynthesis in Jilin ginseng. *Genomics* **2021**, *113*, 2304–2316. <https://doi.org/10.1016/j.ygeno.2021.05.023>.
34. Zhao, M.; Lin, Y.; Wang, Y.; Li, X.; Han, Y.; Wang, K.; Sun, C.; Wang, Y.; Zhang, M. Transcriptome analysis identifies strong candidate genes for ginsenoside biosynthesis and reveals its underlying molecular mechanism in *Panax ginseng* C.A. Meyer. *Sci. Rep.* **2019**, *9*, 615. <https://doi.org/10.1038/s41598-018-36349-5>.
35. Turupcu, A.; Almohamed, W.; Oostenbrink, C.; Seifert, G.J. A speculation on the tandem fasciclin 1 repeat of FLA4 proteins in angiosperms. *Plant Signal. Behav.* **2018**, *13*, e1507403. <https://doi.org/10.1080/15592324.2018.1507403>.
36. Zhang, Y.; Zhou, F.; Wang, H.; Chen, Y.; Yin, T.; Wu, H. Genome-wide comparative analysis of the fasciclin-like arabinogalactan proteins (FLAs) in salicacea and identification of secondary tissue development-related genes. *Int. J. Mol. Sci.* **2023**, *24*, 1481. <https://doi.org/10.3390/ijms24021481>.
37. Costa, M.; Pereira, A.M.; Pinto, S.C.; Silva, J.; Pereira, L.G.; Coimbra, S. In silico and expression analyses of fasciclin-like arabinogalactan proteins reveal functional conservation during embryo and seed development. *Plant Reprod.* **2019**, *32*, 353–370. <https://doi.org/10.1007/s00497-019-00376-7>.

38. Zhang, M.; Wei, H.; Liu, J.; Bian, Y.; Ma, Q.; Mao, G.; Wang, H.; Wu, A.; Zhang, J.; Chen, P.; et al. Non-functional GoFLA19s are responsible for the male sterility caused by hybrid breakdown in cotton (*Gossypium* spp.). *Plant J. Cell Mol. Biol.* **2021**, *107*, 1198–1212. <https://doi.org/10.1111/tpj.15378>.
39. Miao, Y.; Cao, J.; Huang, L.; Yu, Y.; Lin, S. FLA14 is required for pollen development and preventing premature pollen germination under high humidity in arabidopsis. *BMC Plant Biol.* **2021**, *21*, 254. <https://doi.org/10.1186/s12870-021-03038-x>.
40. Liu, E.; Macmillan, C.P.; Shafee, T.; Ma, Y.; Ratcliffe, J.; van de Meene, A.; Bacic, A.; Humphries, J.; Johnson, K.L. Fasciclin-like arabinogalactan-protein 16 (FLA16) is required for stem development in arabidopsis. *Front. Plant Sci.* **2020**, *11*, 615392. <https://doi.org/10.3389/fpls.2020.615392>.
41. Zhang, Y.; Wang, A.; Wu, J.; Gui, Y.; Liu, J.; Zhang, K.; Zhao, M.; Zhang, M.; Wang, Y.; Tian, Z.; et al. Comprehensive genome identification analysis of WOX gene family and its expression in protopanaxatriol ginsenosides under MeJA treatment in ginseng. *Physiol. Plant.* **2026**, *178*, e70920. <https://doi.org/10.1111/pp1.70920>.
42. Su, J.; Song, S.; Wang, Y.; Zeng, Y.; Dong, T.; Ge, X.; Duan, H. Genome-wide identification and expression analysis of DREB family genes in cotton. *BMC Plant Biol.* **2023**, *23*, 169. <https://doi.org/10.1186/s12870-023-04180-4>.
43. Sun, Y.; Jia, X.; Yang, Z.; Fu, Q.; Yang, H.; Xu, X. Genome-wide identification of PEBP gene family in solanum lycopersicum. *Int. J. Mol. Sci.* **2023**, *24*, 9185. <https://doi.org/10.3390/ijms24119185>.
44. Mackinnon, D. *Introduction to Statistical Mediation Analysis*; Routledge: New York, NY, USA, 2008. <https://doi.org/10.4324/9780203809556>.
45. Cao, Y.; Chen, X.; Wei, B.; Zeng, T.; Wang, H.; Zhu, B.; Du, X.; Gu, L. Genome-wide identification and analysis of the FLA gene family in maize (*Zea mays* L.) And its expression in response to abiotic stresses. *BMC Plant Biol.* **2025**, *25*, 982. <https://doi.org/10.1186/s12870-025-06981-1>.
46. Zhang, J.; Yang, L.; Dou, J.; Wang, C.; Yao, Z. Identification of the FLA gene family in soybean and preliminary functional analysis of its drought-responsive candidate genes. *Genes* **2025**, *16*, 1425. <https://doi.org/10.3390/genes16121425>.
47. Zhuang, X.; Kong, L.; Lv, W.; Zhou, Y.; Guan, S.; Yang, H.; Gao, J.; Kang, Y.; Chi, J.; Chai, S.; et al. Genome-wide identification and characterization of the FLA gene family in sorghum under salt-alkali stress. *3 Biotech* **2025**, *15*, 117. <https://doi.org/10.1007/s13205-025-04283-9>.
48. Liu, J.; Chen, C.; Chen, L.; Sharif, R.; Meng, J.; Gulzar, S.; Yi, Z.; Chen, S.; Zhan, H.; Liu, H.; et al. The banana MaFLA27 confers cold tolerance partially through modulating cell wall remodeling. *Int. J. Biol. Macromol.* **2025**, *290*, 138748. <https://doi.org/10.1016/j.ijbiomac.2024.138748>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.