



Article

Spatial Genomic Resource Reveals Molecular Insights into Key Bioactive-Metabolite Biosynthesis in Endangered *Angelica glauca* Edgew

Amna Devi ^{1,2} , Romit Seth ¹ , Mamta Masand ^{1,2} , Gopal Singh ^{1,2} , Ashlesha Holkar ^{1,2}, Shikha Sharma ^{1,2}, Ashok Singh ^{2,3} and Ram Kumar Sharma ^{1,2,*}

¹ Biotechnology Department, CSIR-Institute of Himalayan Bioresource Technology (CSIR-IHBT), Palampur 176061, Himachal Pradesh, India

² Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, Uttar Pradesh, India

³ Environmental Technology, CSIR-Institute of Himalayan Bioresource Technology (CSIR-IHBT), Palampur 176061, Himachal Pradesh, India

* Correspondence: rksharma.ihbt@gmail.com or ramsharma@ihbt.res.in

Abstract: *Angelica glauca* Edgew, which is an endangered medicinal and aromatic herb, is a rich source of numerous industrially important bioactive metabolites, including terpenoids, phenolics, and phthalides. Nevertheless, genomic interventions for the sustainable utilization and restoration of its genetic resources are greatly offset due to the scarcity of the genomic resources and key regulators of the underlying specialized metabolism. To unravel the global atlas of the specialized metabolism, the first spatial transcriptome sequencing of the leaf, stem, and root generated 109 million high-quality paired-end reads, assembled de novo into 81,162 unigenes, which exhibit a 61.53% significant homology with the six public protein databases. The organ-specific clustering grouped 1136 differentially expressed unigenes into four subclusters differentially enriched in the leaf, stem, and root tissues. The prediction of the transcriptional-interactome network by integrating enriched gene ontology (GO) and the KEGG metabolic pathways identified the key regulatory unigenes that correspond to terpenoid, flavonoid, and carotenoid biosynthesis in the leaf tissue, followed by the stem and root tissues. Furthermore, the stem and root-specific significant enrichments of phenylalanine ammonia lyase (PAL), cinnamate-4-hydroxylase (C4H), and caffeic acid 3-O-methyltransferase (COMT) indicate that phenylalanine mediated the ferulic acid biosynthesis in the stem and root. However, the root-specific expressions of NADPH-dependent alkenal/one oxidoreductase (NADPH-AOR), S-adenosyl-L-methionine-dependent methyltransferases (SDMs), polyketide cyclase (PKC), and CYP72A15 suggest the “root” as the primary site of phthalide biosynthesis. Additionally, the GC-MS and UPLC analyses corresponded to the organ-specific gene expressions, with higher contents of limonene and phthalide compounds in the roots, while there was a higher accumulation of ferulic acid in the stem, followed by in the root and leaf tissues. The first comprehensive genomic resource with an array of candidate genes of the key metabolic pathways can be potentially utilized for the targeted upscaling of aromatic and pharmaceutically important bioactive metabolites. This will also expedite genomic-assisted conservation and breeding strategies for the revival of the endangered *A. glauca*.

Keywords: *Angelica glauca*; transcriptome; specialized metabolites; phthalide; ferulic acid



Citation: Devi, A.; Seth, R.; Masand, M.; Singh, G.; Holkar, A.; Sharma, S.; Singh, A.; Sharma, R.K. Spatial Genomic Resource Reveals Molecular Insights into Key Bioactive-Metabolite Biosynthesis in Endangered *Angelica glauca* Edgew. *Int. J. Mol. Sci.* **2022**, *23*, 11064. <https://doi.org/10.3390/ijms231911064>

Academic Editor: Ning Liu

Received: 5 August 2022

Accepted: 7 September 2022

Published: 21 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Indian Himalayan Region (IHR) is an important biodiversity hotspot, with an abundance of 1748 medicinal and aromatic plant species (MAPs) that have significant utility in Indian traditional medicines [1,2]. The gradual rising demands for these MAPs have resulted in the destructive harvesting and depletion of their natural genetic resources, consequently shifting them onto the IUCN's RET list [3–6].

Angelica glauca Edgew (syn: *Angelica nuristanica*; family: Apiaceae; common names: “Chora”, “Gandhrayan”, and “Smooth angelica”) is an important medicinal and aromatic herb that is endemic to the IHR. It is an endangered perennial diploid ($2n = 22$) herb with an extensively larger genome size (11.7 Gb), and it is distributed throughout the Western Himalayan regions at an altitudinal range of 2000–3800 m above sea level (amsl) [7–10]. The plant is glabrous, tall, and erect, having 1–3 pinnately compound leaves, with a thick fistular stem and long conic roots. The clustered flowers in the form of umbel inflorescence are among the main characteristics of the plant, which attain maturity from July to August, mediated through cross-pollination [9,11,12].

The active accumulations of phthalide (butylidene phthalide and ligustilide), ferulic acid, and terpenoids (nerolidol, menthol, and carvone) have been traditionally used for the treatment of various disorders (cardiac, renal, digestive, pulmonary, rheumatism, menorrhoea, cancer, and inflammation), and they have been used widely in aromatherapy [8,12–15]. Phthalide biosynthesis involves either the condensation of malonyl CoA with acetyl-CoA or propionyl-CoA molecules, followed by the SAM-methylation-mediated production of tetraketide intermediate. Subsequently, several biochemical reactions, including cyclization, hydroxylation, carboxylation, and reduction, result in the accumulation of butylidene phthalide and ligustilide [16,17]. Similarly, the phenylpropanoid-mediated ferulic acid-biosynthesis pathway consists of the involvement of key enzymes, including phenylalanine ammonia lyase (PAL), cinnamic acid 4-hydroxylase (C4H), coumarate 3-hydroxylase (C3H), and caffeic acid 3-O-methyltransferase (COMT) [18,19]. In plants, the cytosolic MVA pathway has been suggested to contribute to sesqui- and triterpene biosynthesis, while the plastidial MEP pathway mainly synthesizes mono- and diterpenes, with limited crosstalk with the MVA pathway [20]. Nevertheless, most of these inferences were derived based on preliminary phytochemical and pharmacological studies [14,21]. Having evolved with a highly heterozygous complex genome (11.7 GB), *A. glauca* has possibly been omitted from global whole-genome-sequencing (WGS) efforts [7,10,22]. The paradigm shifts in sequencing technologies with the availability of multiple cost-effective high-throughput next-generation-sequencing (NGS) platforms has enormously contributed towards unraveling the functional and genomic insights into the specialized metabolism, irrespective of model or nonmodel plants [23,24], wherein organ-specific spatial-transcriptome-sequencing profiling has been proven to be an effective approach to the creation of comprehensive genomic resources and the elucidation of the complex biological/regulatory networks involved in the biosynthesis and transport of the specialized metabolism [25–27].

Therefore, global organ-specific transcriptome-sequencing efforts have been carried out for the creation of comprehensive genomic resources in the endangered Himalayan plant species *A. glauca*. Subsequently, the integration of spatial gene expressions with predicted protein interactome network analysis has assisted in the identification of the organ-specific key genes and regulators involved in the terpenoid, ferulic acid, and phthalide biosynthesis pathways with aromatic and pharmaceutical importance. The inferences derived in the current study will be a footstep for futuristic functional genomics and synthetic biology efforts for the nondestructive and sustainable production of the bioactive metabolites of *A. glauca*. In addition, novel genomic-resource creation can be utilized for the development of functionally relevant molecular-marker resources to strategize the conservation and domestication efforts in *A. glauca*.

2. Results

2.1. Transcriptome Sequencing, De Novo Assembly, and Functional Annotation

The organ-specific spatial transcriptome sequencing of the leaf, stem, and roots generated a total of 119 million raw reads. Quality filtering and the de novo assembly of 109 million high-quality reads resulted in 81,162 unigenes and 36,052 isoforms, with N_{50} and average transcript lengths of 1607 bp and 923.32 bp, respectively. The function assignment with six public databases annotated 72,120 (61.53%) transcripts into NCBI nr (69,731;

59.49%), TAIR (58,982; 50.31%), KOG (55,513; 47.36%), Swiss-Prot (53,957; 46.03%), Plant TFIIdb (33,040; 28.18%), and KEGG (8945; 7.63%) (Figure 1A–C, Tables 1–3 and S1).

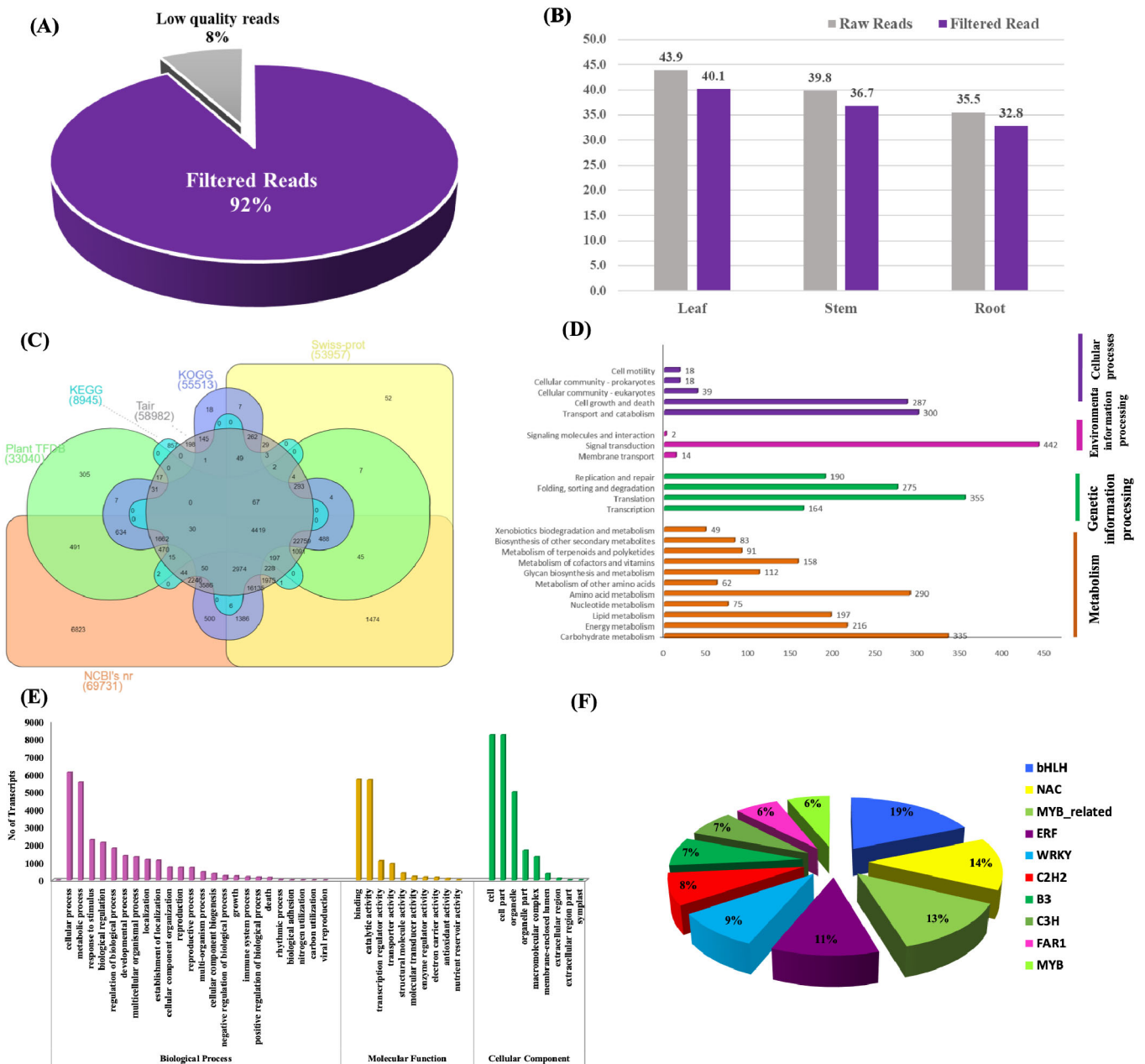


Figure 1. (A) Quality filtering of sequenced reads. (B) High-quality filtered reads in leaf, stem, and root tissues. (C) Venn diagram representing the functional annotations of transcripts with six public databases. (D) KEGG annotation grouped transcripts into four main categories: metabolism, genetic-information processing, environmental-information processing, and cellular processes. (E) Histogram representing gene ontology (GO) grouped into three functional categories: biological processes, molecular functions, and cellular components. (F) Pie chart representing classification of major transcription-factor families.

Table 1. Total raw and filtered reads generated for leaf, stem, and root tissues of *A. glauca*.

S. No.	Sample Name	Raw Reads	Filtered Reads	Retained High-Quality Reads (%)
1	Leaf	43,893,016	40,144,376	91.4
2	Stem	39,814,592	36,737,576	92.2
3	Root	35,495,806	32,757,628	92.2
	Total reads	119,203,414	109,639,580	91.9

Table 2. De novo-assembly statistics of *A. glauca* transcriptome.

De Novo-Assembly Statistics		
	Before Clustering	After Clustering
Total transcripts	1,61,836	1,17,214
Total genes	81,570	81,162
Percent GC	39.69	39.62
Contig N50	1751	1607 bp
Median contig length	715	526 bp
Average contig length	1071.45	923.32 bp

Table 3. Functional annotations of transcripts with various public databases.

Database	No. of Transcripts	Percent (%)
NCBI nr	69,731	59.49
TAIR	58,982	50.31
KOG	55,513	47.36
Swiss-Prot	53,957	46.03
Plant TFDB	33,040	28.18
KEGG	8945	7.63
Total	72,120	61.53

The KEGG-pathway curation assigned 397 enriched metabolic pathways with top representations of metabolism (1668), genetic-information processing (984), the cellular processes (662), and environmental-information processing (458), including the transcripts that correspond to terpenoids and polyketides (91) and secondary-metabolite biosynthesis (83). Among the secondary-metabolite biosynthesis, terpenoid-backbone biosynthesis (ko00900; 30) represented the largest group, followed by phenylpropanoid biosynthesis (ko00940; 20), carotenoid biosynthesis (ko00906; 18), and flavonoid biosynthesis (ko00941; 11) (Figure 1D).

In addition, KOG annotated 55,513 transcripts that represent 25 different major groups, exhibiting the maximum enrichment of the 'general function prediction only' (8117), followed by the 'signal transduction mechanisms' (4821), 'transcription' (4057), 'posttranslational modification, protein turnover, chaperones' (3302), and 'secondary metabolites biosynthesis, transport and catabolism' (1875).

Furthermore, the GO annotation classified 17,293 transcripts into 66,091 GO terms, including biological processes (26,794; 24 groups), molecular functions (14,418; 10 groups), and cellular components (24,879; 9 groups). Furthermore, the majority of the transcripts were grouped into cellular processes (6105), binding (5691), and cell including cell part (8206) in the biological processes, molecular functions, and cellular components, respectively (Figure 1E).

2.2. Differential Transcript Expression and Enrichment Analysis

Pairwise combinations identified 4184 transcripts exhibiting significant differential expressions in leaf vs. stem (1894: Ups: 962; DOWNS: 902), leaf vs. root (1637: Ups: 1139; DOWNS: 498), and stem vs. root (683: Ups: 514; DOWNS: 169) (Figure 2A, Table S2). Furthermore, the clustering of 1136 key DE transcripts was categorized into four subclusters representing 373 (Subcluster 1), 49 (Subcluster 2), 614 (Subcluster 3), and 100 (Subcluster 4)

transcripts. Subclusters 2 and 3 represent the transcripts significantly enriched in the leaf, while Subclusters 1 and 4 were enriched solitarily in the stem and root tissues, respectively (Figure 2B, Table S3).

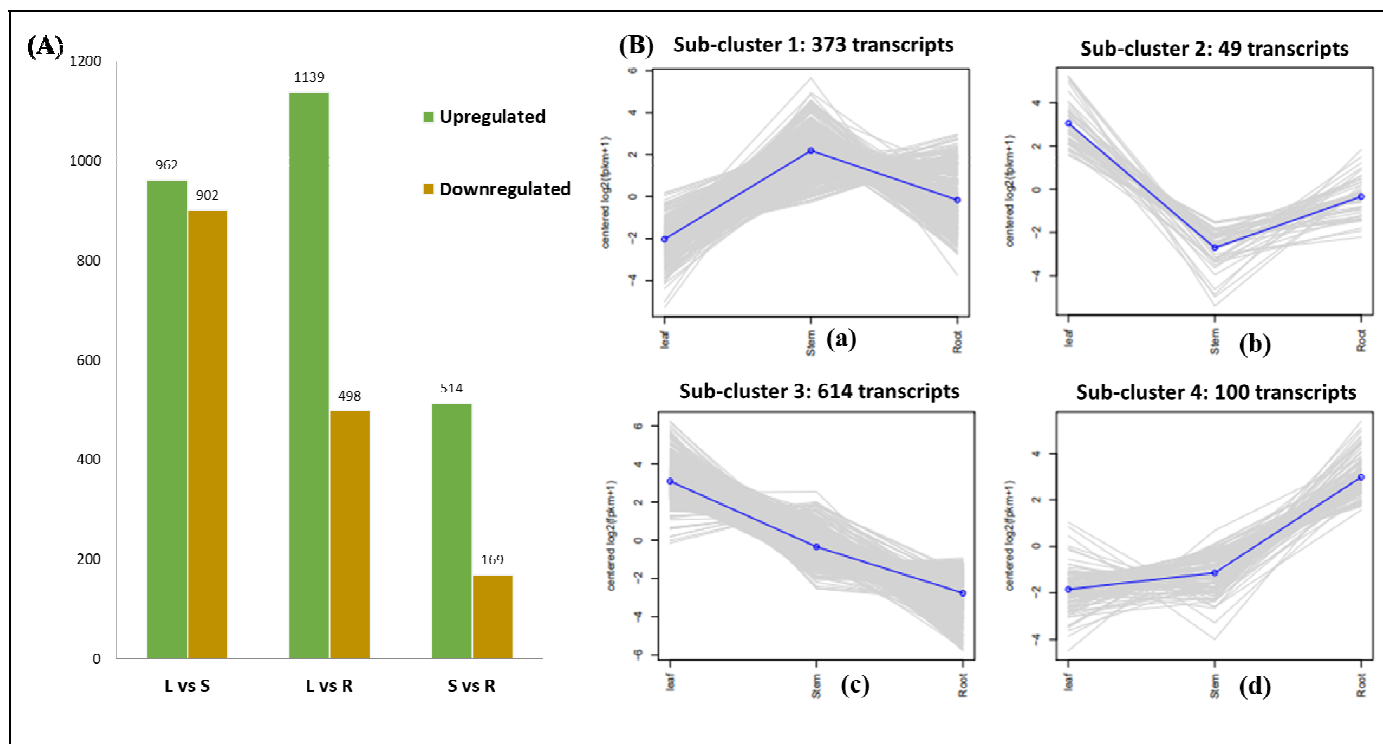


Figure 2. (A) Histogram representing significant DEGs in three combinations: leaf vs. stem (L vs. S), leaf vs. root (L vs. R), and stem vs. root (S vs. R). (B) Clustering of 1136 significant transcripts in leaf, stem, and root: (a) Subcluster 1 represents higher expressions of transcripts in the stem; (b) Subcluster 2 in the leaf; (c) Subcluster 3 in the leaf; (d) Subcluster 4 in the root. The color line in subfigure B represents the median values.

GO-Enrichment Analysis

The organ-specific GO-enrichment analysis provided insights into the transcriptional dynamics of the significantly enriched transcripts in the leaf, stem, and root tissues of *A. glauca*. Among the biological processes, the response to gibberellin (GO:0009739), the abscisic acid-activated signaling pathway (GO:0009738), and the regulation of the cellular macromolecule biosynthetic process (GO:2000112) were recorded with significantly higher enrichments in the root tissues. However, the photosynthesis-related GO terms (GO:0009769, GO:0009768, GO:0009773, GO:0015979, GO:0009657), along with stomatal regulation (GO:0090333, GO:0010119), the cellular protein metabolic process (GO:0044267), and the organonitrogen compound biosynthetic process (GO:1901566) were significantly enriched in the leaf. Meanwhile, the stem exhibited a significantly higher enrichment of cell-wall biogenesis (GO:0042546), the cellular polysaccharide biosynthetic process (GO:0033692), and ion transport (GO:0006811). Interestingly, the secondary metabolic process (GO:0019748) showed significant enrichment in both the stem and root tissues (Figure 3A, Table S4).

Among the molecular functions, the GO terms corresponding to the transcription-factor (GO:0003700, GO:0043565) and enzyme-inhibitor (GO:0004857) activities were enriched in the root tissue. Moreover, the hydrolase (GO:0016798, GO:0004553, GO:0016787), transporter (GO:0022857, GO:0005215), and transferase (GO:0016740) activities were significantly enriched in both the stem and root tissues. The GO terms related to binding, including pigment binding (GO:0031409), chlorophyll binding (GO:0016168), and cofactor binding (GO:0048037), were enriched in the leaf tissue (Figure 3B, Table S4).

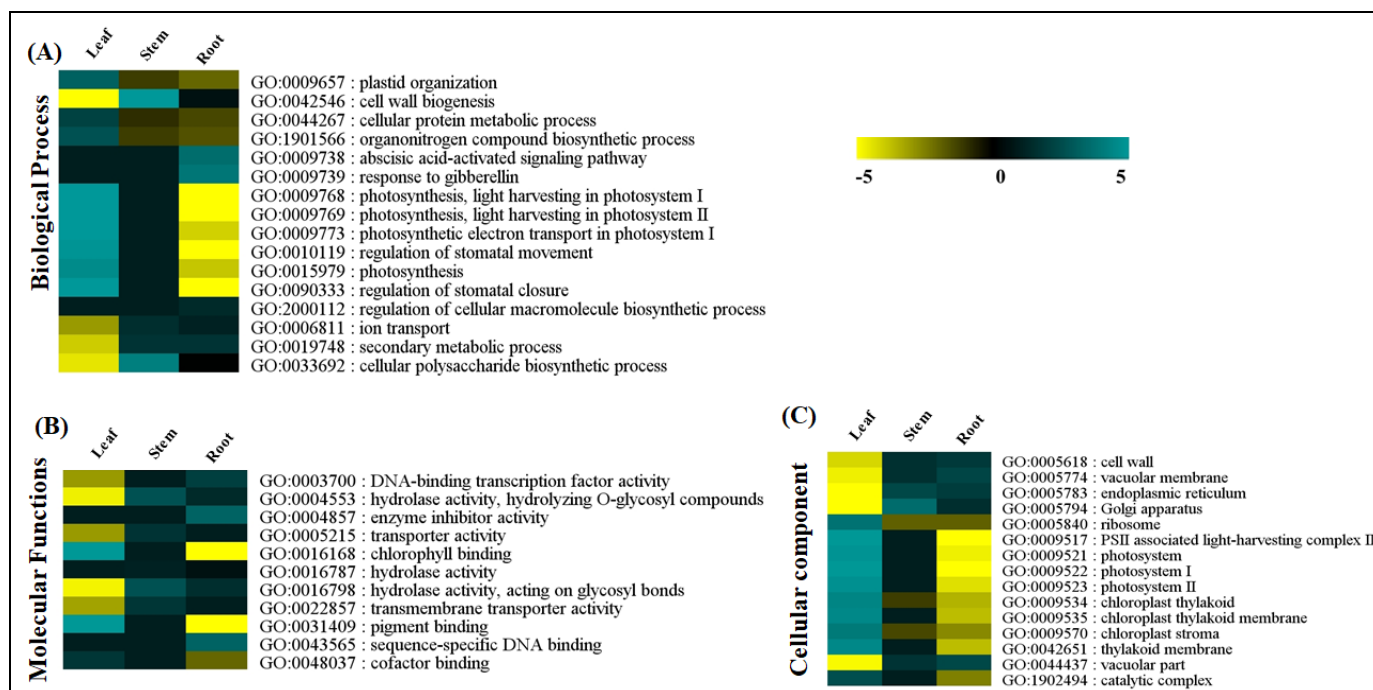


Figure 3. Heatmap representing GO enrichment classified into three categories: (A) biological processes; (B) molecular functions; (C) cellular components, in leaf, stem, and root tissues of *A. glauca*.

Among the cellular components, the vacuolar part (GO:0044437, GO:0005774), endoplasmic reticulum (GO:0005783), cell wall (GO:0005618), and Golgi apparatus (GO:0005794) were significantly enriched in the stem and root. In concordance with the biological processes, the leaf tissue recorded significantly higher enrichments of chloroplast thylakoid related (GO:0009535, GO:0042651, GO:0009534, GO:0009570), along with photosystems (GO:0009517, GO:0009521, GO:0009522, GO:0009523), catalytic complexes (GO:1902494), and ribosomes (GO:0005840) (Figure 3C, Table S4).

2.3. Protein–Protein-Interactome-Network Analysis

The protein–protein-network prediction recorded significant interactions of 651 nodes (unigenes), having 24,017 edges (average number of neighbors: 73.785; coefficient: 0.475). The predicted network revealed a significantly higher enrichment in the leaf tissue, followed by the root and stem tissues (1×10^{-16} ; average degree: 67.27; Figure 4A,B; Table S5). Interestingly, the key candidates, including caffeoyl CoA methyl transferase, having significant direct interactions with CYPs (CYP84 and CYP98), phenylalanine ammonia lyase (PAL), cinnamoyl CoA reductase (IRX4), and 4-coumarate: CoA ligase (4CL2), with significantly high enrichment in the root tissues, suggests phenylpropanoid-pathway-mediated ferulic acid biosynthesis (Figure 4C). Furthermore, the terpenoid biosynthesis recorded a significantly higher enrichment in the leaf and root, whereas terpene synthase (TPS), exhibiting a direct interaction with CYP82, Tubulin (TUB5), and Aspartic protease (APS1), exhibited higher enrichment in the roots, while the photosynthesis- and chlorophyll-biosynthesis pathways were recorded with significant enrichment in the leaf (Figure 4D,E).

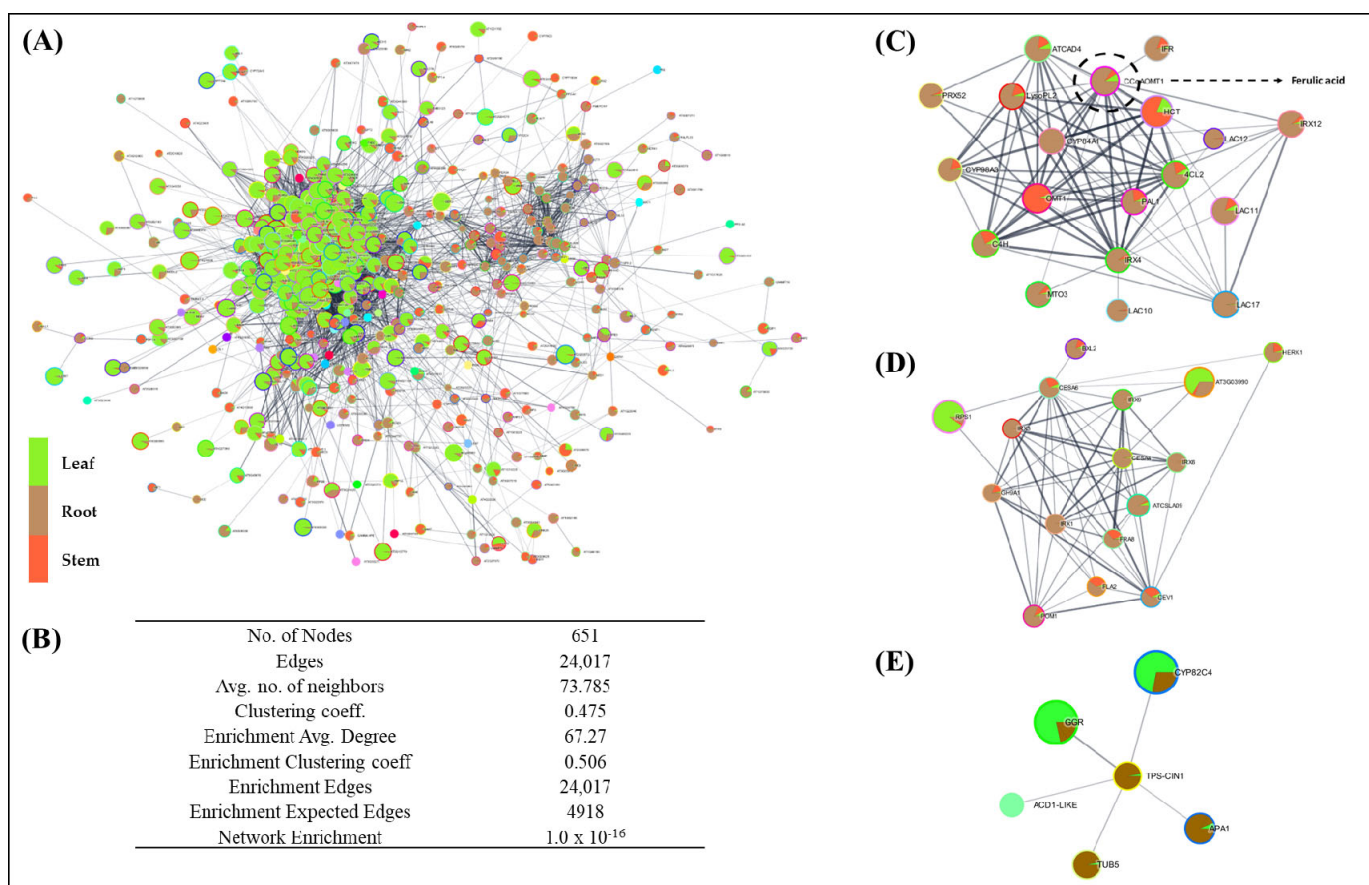


Figure 4. (A) Predicted spatial transcriptional interactome of *A. glauca*. (B) Network statistics of the interactome build. (C) Enriched network of ferulic acid pathway. (D,E) Terpenoid biosynthesis. The green color represents significant enrichment in the leaf tissue, orange in the stem tissue, and brown in the root tissue.

2.4. Organ-Specific Gene-Expression Elucidation of Key Metabolite Biosynthesis

The organ-specific KEGG-pathway-enrichment analysis recorded a significantly higher enrichment of the photosynthesis-related metabolic pathways, including carbon fixation and porphyrin and chlorophyll metabolism in the leaf tissue. However, diterpenoid and phenylpropanoid biosynthesis, including ferulic acid synthesis, were found to be enriched in both the stem and root (Figure 5). Moreover, the transcripts involved in the biosynthesis of terpenoid backbone, flavonoids, and carotenoids confined their enrichment in the leaf, followed by the stem and root (Figures 5 and 6).

The root-specific higher expressions of the key candidates (NADPH-AOR, SDM, PKC, and CYP72A15) suggest the ‘root’ as the site of the synthesis and accumulation of phthalide biosynthesis. Likewise, the confined expressions of DHQ-SDH, EPSP, CM, ADT, PAL, C4H, and COMT in the stem and root complement the PPIN inference of the phenylpropanoid-pathway-mediated ferulic acid biosynthesis. Furthermore, the leaf-specific expressions of DXS, MCT, and HDR suggest the MEP-mediated biosynthesis of aromatic mono- and sesquiterpenoids. Notably, the committed step (catalyzed by GPPS) was specific to the leaf tissue, while the downstream monoterpene-diversification-related genes, including LS, NMD, and CD, exhibited higher expressions in the stem and root. Nevertheless, STS and NES, involved in sesquiterpene biosynthesis, confined their expressions in the leaf (Figure 7). This transcriptional-level study was further complemented by the metabolite profiling of the leaf, stem, and root tissues.

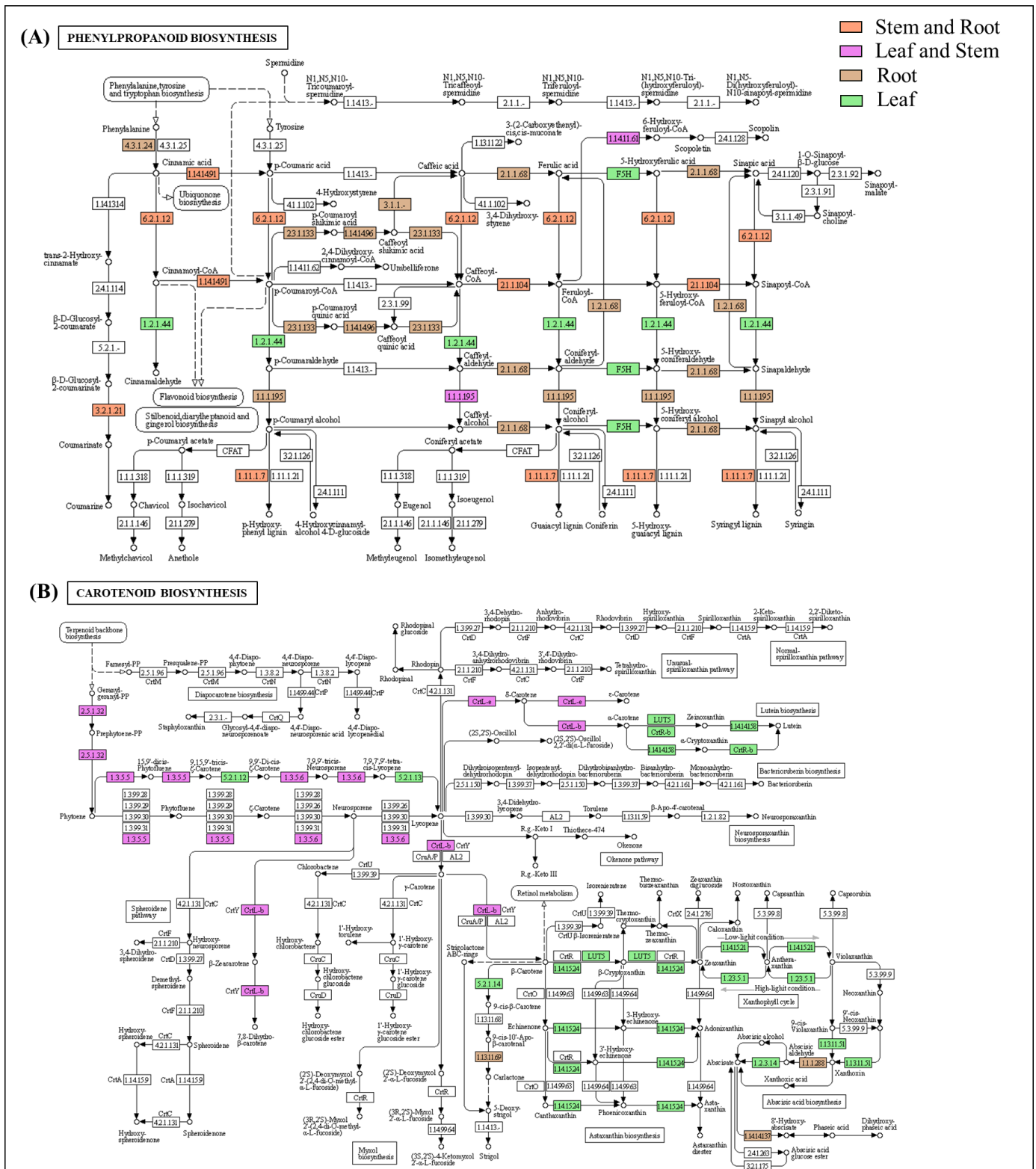


Figure 5. Representation of significant (A) phenylpropanoid- and (B) carotenoid-pathway enrichment in leaf, stem, and root tissues of *A. glauca*.

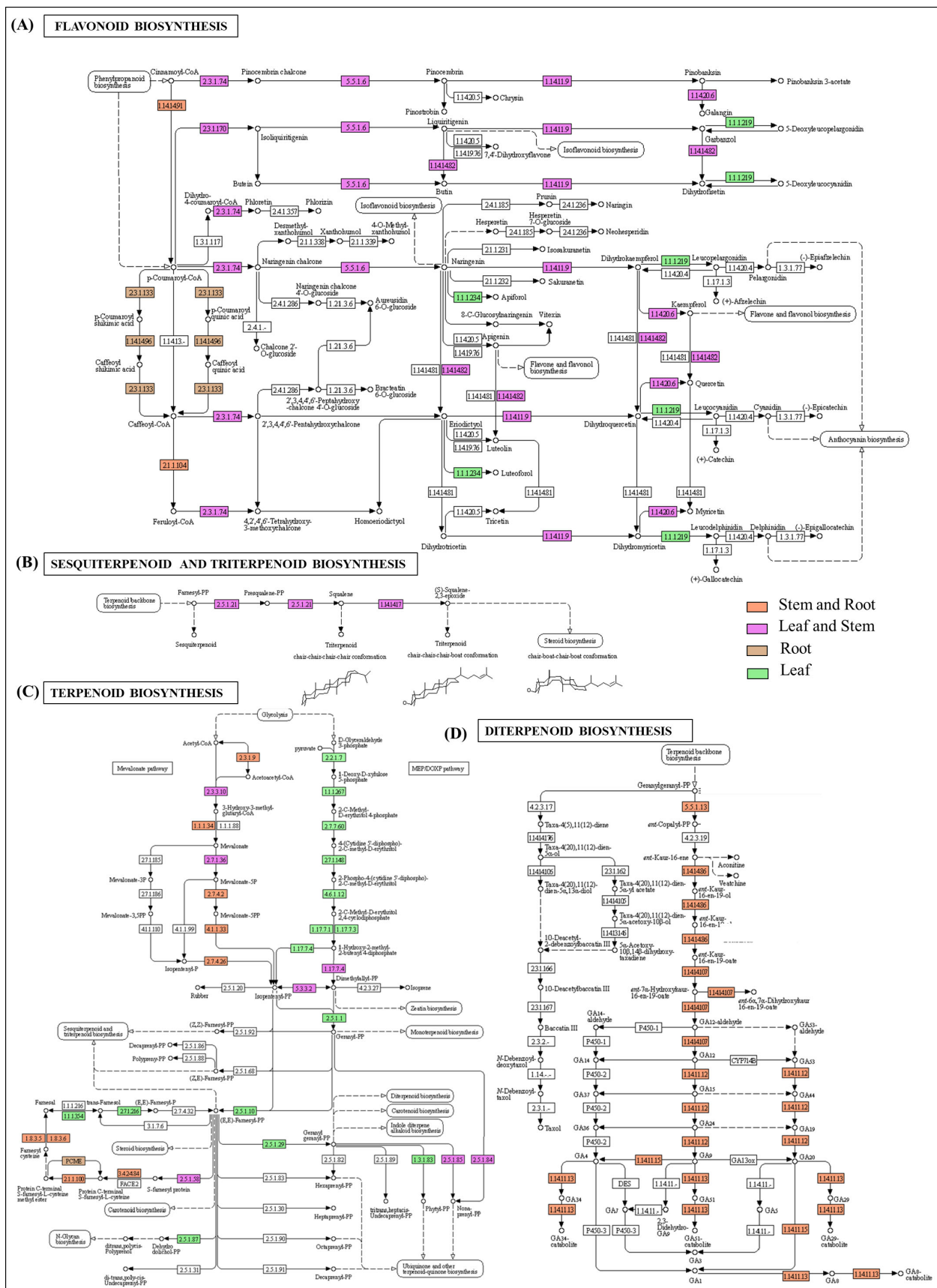


Figure 6. Representation of significant (A) flavonoids, (B) sesquiterpenoid and triterpenoid biosynthesis; (C) terpene backbone; (D) diterpenoids, in leaf, stem, and root tissues of *A. glauca*.

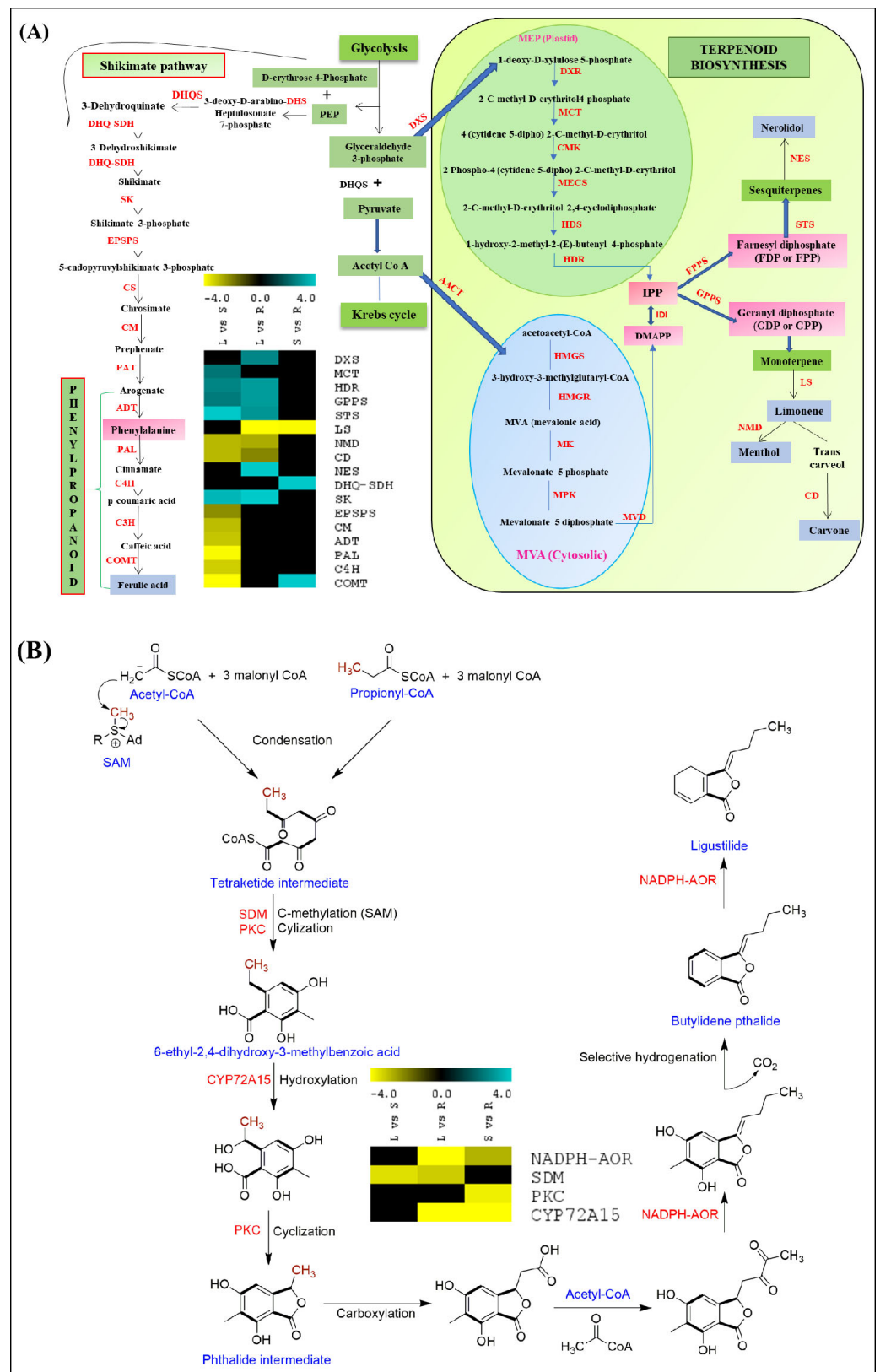


Figure 7. Representations of: **(A)** terpenoid and ferulic acid biosynthesis; **(B)** putative pathway of phthalide biosynthesis in *A. glauca*. Heatmap representing expression profiles of candidates across three combinations (leaf vs. stem; leaf vs. root; stem vs. root).

2.5. Expression Dynamics of Transcription Factors

Transcription factors (TFs) are key elements that regulate the spatiotemporal expression of bioactive-metabolite biosynthesis [28]. Overall, 33,040 transcripts encoding 58 TF families, with significant abundances of bHLH (3386), NAC (2439), MYB related (2398), ERF (1998), WRKY (1655), C2H2 (1456), B3 (1328), C3H (1243), FAR1 (1121), and MYB (1113), were identified in *A. glauca* (Figure 2F). Among these, 2764 TFs exhibited significant spatial expressions in the leaf (L), stem (S), and root (R), wherein bHLH (L: 65, S: 47, R: 19), NAC (L: 33, S: 54, R: 16), MYB related (L: 42, S: 46, R: 24), ERF (L: 77, S: 71, R: 22), and WRKY (L: 18, S: 20, R: 4) were the most abundant (Tables 4 and S6). Additionally, the Trihelix, TCP, YABBY, WOX, SBP, and SRS TF families also recorded tissue-specific expressions, with higher enrichments of WRKY, TCP, and YABBY in the leaf tissue, while Trihelix, WOX, SBP, and SRS showed significant enrichments in the root tissue in *A. glauca* (Figure 8A).

Table 4. Tissue-specific expressions of abundant TFs encoding transcripts in *A. glauca*.

S. No.	TF Family	Leaf (L)	Stem (S)	Root (R)
1	bHLH	65	47	19
2	NAC	33	54	16
3	MYB related	42	46	24
4	ERF	77	71	22
5	WRKY	18	20	4
6	C2H2	20	21	9
7	B3	33	14	9
8	C3H	16	23	2
9	FAR1	21	16	14
10	MYB	17	21	5

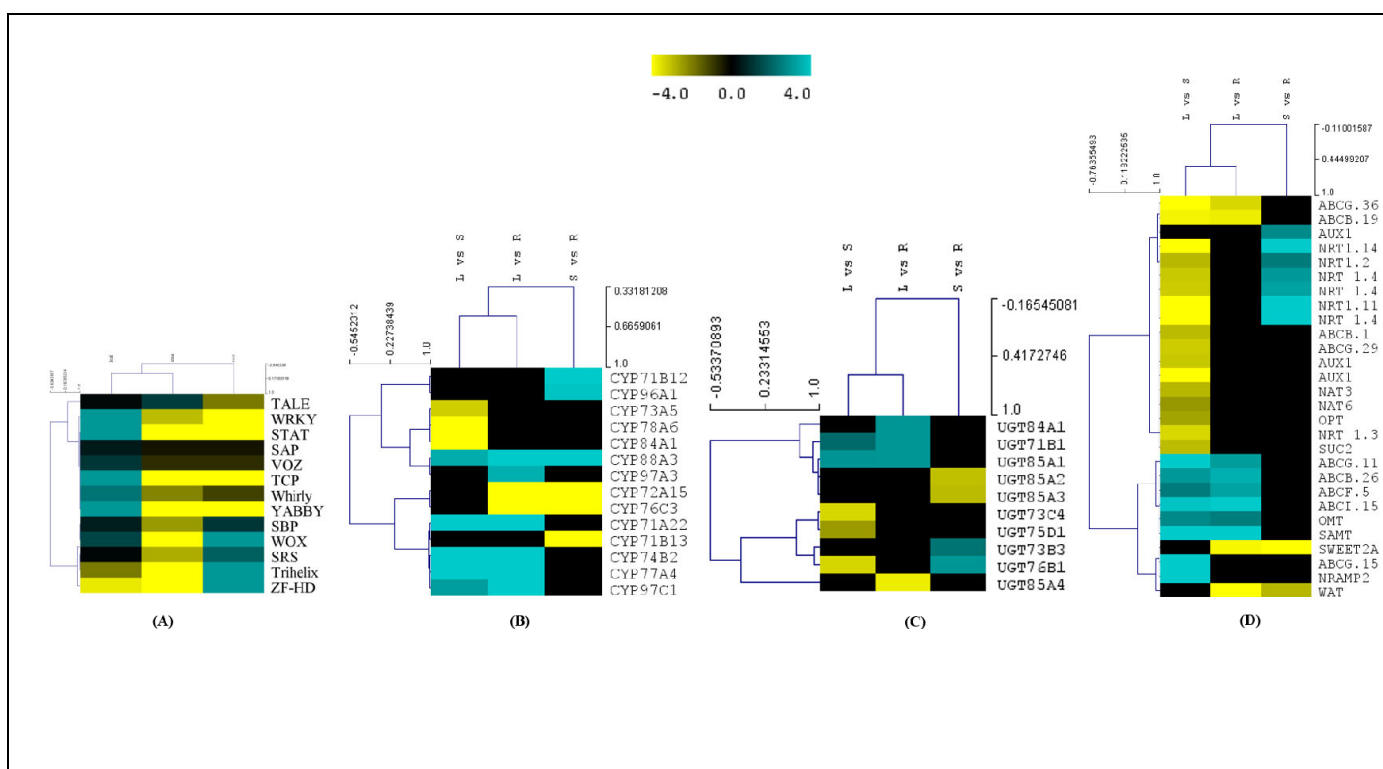


Figure 8. Heatmap visualization of differentially expressed transcripts of: (A) transcription factors (in leaf, stem, and root); (B) CYP; (C) UGT; (D) transporters in three combinations (L vs. S; L vs. R; S vs. R), of *A. glauca*.

2.6. Differential Expressions of CYPs, UGTs, and Transporters

The CYPs and UGTs, which are a specific class of oxidoreductases and transferases, have been proven to be multifunctional candidates that regulate the biosynthesis and diversification of secondary metabolites [29,30]. In total, 602 transcripts encoding 44 CYP families were identified in *A. glauca*, with abundances of CYP71, CYP72, CYP76, CYP704, and CYP81. Of these, 30 CYPs recorded significant spatial expressions, with CYP71A22, CYP74B2, CYP77A4, CYP82C4, CYP84A1, CYP97A3, and CYP97C1 having higher expressions in the leaf. Likewise, CYP71B12, CYP73A5, CYP78A6, and CYP96A1 exhibited higher expressions in the stem, while CYP71B13, CYP72A (10, 13, 15), CYP72C, CYP76C3, CYP716A1, CYP81D8, and CYP98A3 exhibited higher expressions in the root (Figure 8B, Table S7).

Similarly, 264 transcripts were assigned to 21 UGT families, with the majority of them belonging to three UGT families (UGT85: 65 transcripts; UGT73: 47 transcripts; UGT76: 29 transcripts). Of these, UGT71B1, UGT84A1, and UGT85A1 recorded higher expressions in the leaf, while the stem exhibited higher expressions of UGT73B3, UGT73C4, UGT75D1, and UGT76B1. Likewise, UGT85A2, UGT85A3, and UGT85A4 showed higher expressions in the root (Figure 8C, Table S7). Considering the key role of transporters in membrane transport and the accumulation of various endogenous secondary metabolites in plants [31], interestingly, 78 differentially expressed transporters identified with the ABCB.26, ABCF.5, ABCG.11, ABCG.15, dicarboxylate transporter 1 (OMT), and SAMT1 families, exhibiting higher expressions in the leaf, while the ABCB.1, ABCG.29, sucrose-transporter (SUC2), NAT, and NRT families were significantly expressed in the stem. Transcripts encoding CAT, SWEET 2A, MATE, and WAT1-related protein recorded higher expressions in the root tissues, while ABCB.19 and ABCG.36 exhibited higher expressions in the stem compared with the root (Figure 8D, Table S7).

2.7. RNA-Seq Validation Using qRT-PCR

The RNA-Seq expression patterns of 21 genes involved in the key metabolic biosynthesis pathways, including terpenoid (DXS, MCT, HDR, GPPS, STS, LS, NMD, CD, NES), ferulic acid (DHQ-SDH, SK, EPSPS, CM, ADT, PAL, C4H, COMT), and phthalide (NADH-AOR, SDM, PKC, CYP72A15), were well complemented by the quantitative real-time expression data, and were recorded with a positive correlation ($r^2 = 0.779$) (Figure 9A–C, Table S8). However, the expression differences (0.23) between the RNA-Seq and qRT-PCR are owing to intrinsic properties, such as the detection ranges and sensitivities of these methods [32,33] (Figure 9A–C, Table S8).

Furthermore, eight candidate genes (viz., ADT, PAL, C4H, COMT, NADH-AOR, SDM, PKC, and CYP72A15) were validated with four diverse genotypes of *A. glauca*, and the qRT-PCR expression profiles follow a similar trend of expression as that recorded in the RNA-Seq expression data (Figure 10, Table 5). However, the expression variations recorded across the genotypes may be due to the genotypic diversity and unique features of the methods.

Table 5. Geographical description of four diverse genotypes of *A. glauca* from Himachal Pradesh.

S. No.	District	Population	Latitude (°N)	Longitude (°E)	Altitude (m)
1	Kinnaur	Asrang (G1)	31.667	78.313	3376
2	Kullu	Kandhaghai (G2)	31.408	77.444	2250
3	Mandi	Thunag (G3)	31.577	77.152	2177
4	Kullu	Gulaba (G4)	32.317	77.189	2521

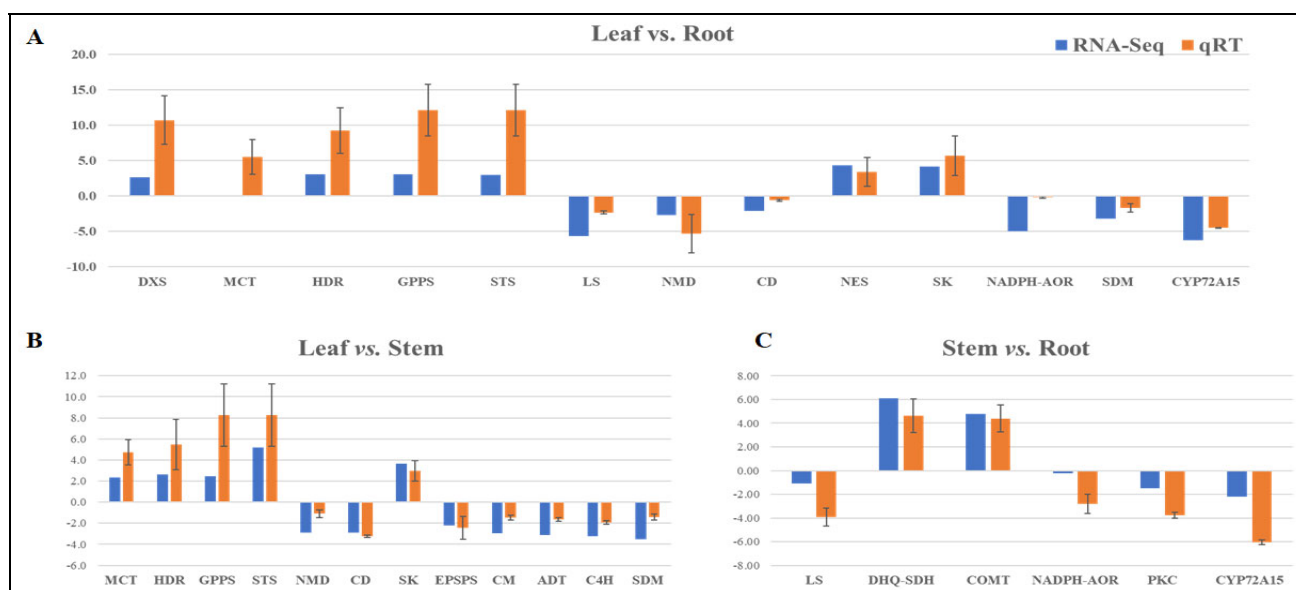


Figure 9. Representation of comparative expression profiles of genes (RNA-Seq and qRT) involved in specialized-metabolite-biosynthesis pathways across three combinations: (A) leaf vs. root; (B) leaf vs. stem; (C) stem vs. root.

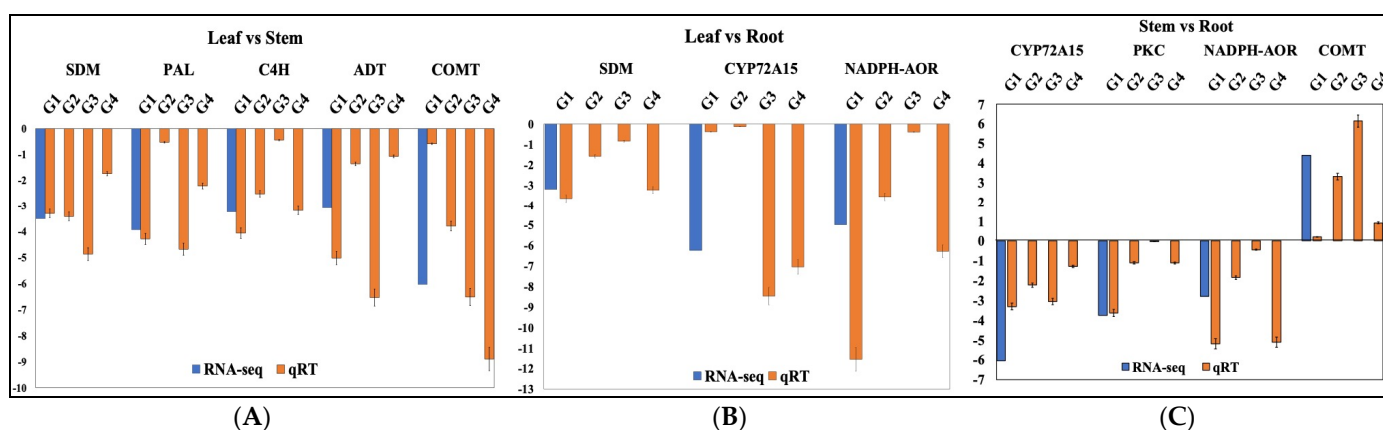


Figure 10. Comparative organ-specific RNA-Seq and qRT-PCR expression profiles of bioactive metabolite pathway-related genes in diverse genotypes from geographically separated populations (G1: Kinnaur_Asrang; G2: Kullu_Kandhaghai; G3: Mandi_Thunag; G4: Kullu_Gulaba) of *A. glauca* across three combinations: (A) leaf vs. stem; (B) leaf vs. root; (C) stem vs. root.

2.8. GC-MS and UPLC Analysis

GC-MS analysis was used to determine the volatile compounds in the leaf, stem, and root tissues of *A. glauca*. The root tissue reported higher contents of limonene, with an area percent of 7.213%, as well as ligustilide (40.48%) and butylidenephthalide (0.66%), while germacrene D (5.01%), alpha farnesene (1.96%), and alpha bergamotene, with 2.39% area, exhibited dominance presence in the leaf tissue. Moreover, the results of the UPLC analysis recorded a higher accumulation of ferulic acid in the stem (206.66 $\mu\text{g}/100\text{ mg}$), followed by the root (22.82 $\mu\text{g}/100\text{ mg}$) and leaf (15.61 $\mu\text{g}/100\text{ mg}$) tissues (Figures 11, S1 and S2).

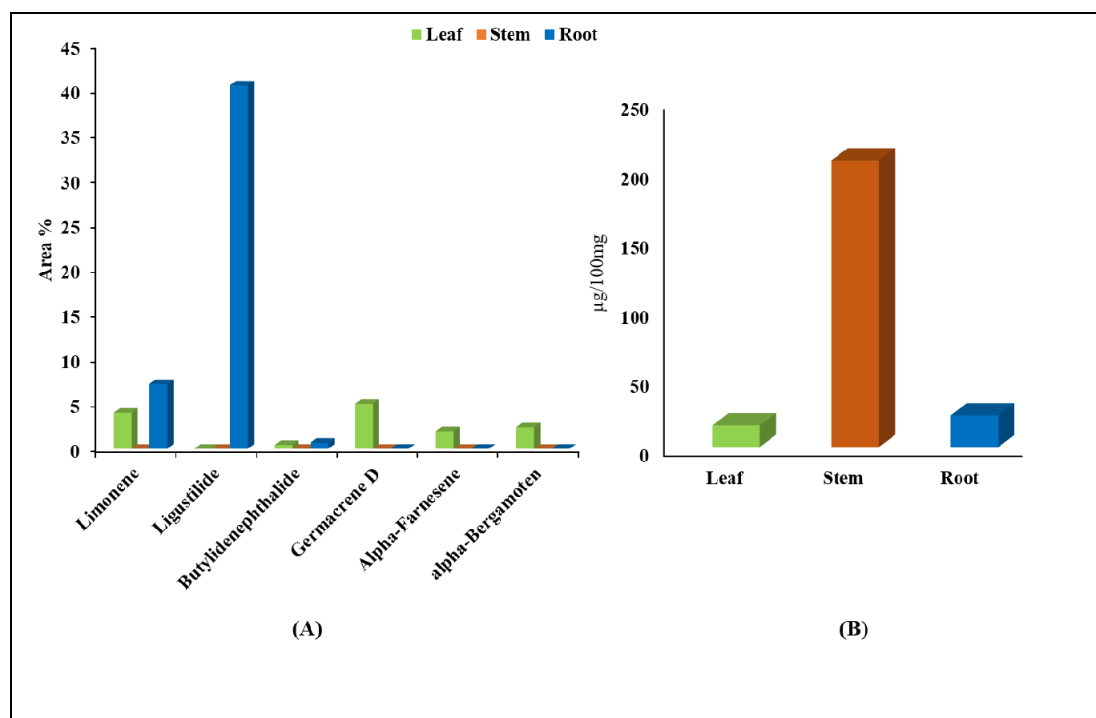


Figure 11. (A) GC-MS analysis of bioactive metabolites and (B) UPLC analysis of ferulic acid in leaf, stem, and root tissues of *A. glauca*.

3. Discussion

A. glauca is a rich repository of secondary metabolites, and especially terpenoids, phenylpropanoids, and phthalides having cardioactive, carminative, stimulant, digestive, expectorant, anti-inflammatory, and anticancerous properties are its key constituents [15,34]. Consequently, the cumulative demand of the phytopharmaceutical and aroma industries has recklessly harvested the genetic resources of *A. glauca* from its natural habitat. Additionally, soil erosion and excessive animal grazing have caused *A. glauca* to be listed among the IUCN's endangered plants [35,36]. More recently, high-throughput cost-effective next-generation sequencing (NGS) has significantly assisted in the conservation efforts through the rapid creation of genomic resources and gene-expression patterns for the identification of the key biosynthetic pathways, irrespective of model or nonmodel medicinal plant species [23,27].

Therefore, in the current study, global organ-specific transcriptome sequencing was carried out, with the successful generation of 109 million paired-end reads, assembled into 81,162 unigenes in organ-specific NGS-assisted transcriptome sequencing, which were in consensus with the recent genomic efforts in related *Angelica* species [37,38]. Overall, a 61.53% functional annotation rate of the de novo assembled unigenes with public protein databases is comparable with *Ferula asafoetida* (58%) and *Thapsia laciniata* (61.9%) of the family Apiaceae [39,40]. Moreover, the integration of the predicted protein-interactome network with leaf-, stem-, and root-specific gene expressions has assisted in the identification of the key regulators for the successful elucidation of the specialized-metabolite-biosynthesis pathway.

3.1. Organ-Specific Transcriptional Dynamics in *A. glauca*

The global understanding of organ-specific gene expression has been successfully utilized to reveal the intricate key biosynthetic pathways and biological mechanisms, including different gene families, at the genome and transcriptional levels [41–43]. Significant pair-wise DEGs (leaf vs. stem; leaf vs. root; stem vs. root) and KEGG-enrichment analysis, bolstered in deciphering the complex secondary-metabolite-biosynthesis pathways

in *A. glauca*, have been complemented with various previous studies, including those on *Podophyllum hexandrum* [44], *Trillium govanianum* [45], and *Fritillaria roylei* [27]. Subsequently, the prediction of the transcriptional-interactome network complementing the organ-specific expression-profiling-assisted elucidation of the complex specialized-metabolite pathways has been performed successfully in earlier studies [27,46–48].

The cumulative enrichment of the terpenoids, flavonoids, and carotenoids in the leaf, and the phenylpropanoid and diterpenoid biosynthesis in the root and stem tissues, suggest the intra-/intercellular trafficking of key bioactive metabolites via membrane transporters [49,50]. This was further supported by the successful prediction of the transcriptional-interactome network, which suggests phenylpropanoid-mediated ferulic acid biosynthesis in the root and stem tissues with the active involvement of caffeoyl CoA methyl transferase, phenylalanine ammonia lyase, cinnamoyl CoA reductase, and 4-coumarate: CoA ligase. Nevertheless, the root-specific enrichment of regulators having significant interactions with the key candidates that correspond to ferulic acid synthesis further support the ‘stem and root’ as the sites of ferulic acid biosynthesis. However, the enrichment of the photosynthesis and the stomatal regulation in the predicted network in the leaf tissue indicates enough primary-metabolite flux to synthesize and accumulate secondary metabolites [51,52].

Furthermore, the abundance of TFs (viz., bHLH, NAC, ERF, WRKY, and MYB) indicates their role in mono- and sesquiterpenoid and phenylpropanoid biosynthesis, as reported in previous studies [53,54]. Likewise, the leaf-specific higher enrichments of WRKY, TCP, YABBY, ERF, and MYB suggest their key involvement in regulating terpenoid and flavonoid biosynthesis [28,55–57], while the root-specific enrichment of TFs (viz., Trihelix, WOX, and SBP) probably regulates secondary-metabolite biosynthesis [58].

3.2. Elucidation of Key Biosynthetic Pathways

Phenylpropanoid biosynthesis is one of the most important pathways in secondary metabolism, and it contributes to the biosynthesis of a wide range of bioactive metabolites in plants [59]. The significant stem- and root-specific enrichments of phenylalanine ammonia lyase, cinnamate-4-hydroxylase, and caffeic acid 3-O methyl transferase suggest the tissue-specific biosynthesis and accumulation of ferulic acid. Considering that isoprenoids (C-5 isopentenyl diphosphate (IPP), and its isomer dimethylallyl diphosphate (DMAPP) act as the precursors for the terpenoid pool in plant systems, the significant KEGG enrichment of the cytosolic MVA pathway in the stem and root tissues suggests MVA-mediated diterpenoid biosynthesis [20,60]. Nevertheless, the higher enrichments of the transcripts involved in the MEP pathway in the leaf suggests MEP-mediated mono-, sesqui-, and triterpenoid and carotenoid biosynthesis in the leaf tissues, as reported earlier for *Salvia miltiorrhiza* [61] and *Cymbopogon winterianus* [62]. Furthermore, inferences in biosynthetic studies on the phthalide-like structure (marilone and mycophenolic acid) in fungi were used to elucidate the condensation-, methylation-, cyclization-, and hydroxylation-mediated putative biosynthesis pathway of phthalide in *A. glauca* [17,63–65]. The higher accumulation of ferulic acid in the stem, followed by the root and leaf tissues, observed in the UPLC analysis, well complemented both the RNA-Seq and qRT expression analyses. Moreover, the root-specific higher accumulation of limonene, ligustilide, and butylidene, well corroborated by the transcriptional dynamics reported in the current investigation, indicates the root as a storage site for the accumulation of these bioactive metabolites, as reported earlier [8,9].

The cytochrome P450 monooxygenases (CYPs) and UDP-glycosyltransferases (UGTs) are known to play an imperative role in the structural and chemical diversity of plant secondary metabolites (phenylpropanes, terpenoids, alkaloids, and carotenoids) by catalyzing numerous hydroxylation reactions [66–69]. The leaf-specific expressions of CYP71, 74, 77, 82, and 97, along with the UGTs (71, 84, and 85), suggest their key involvement in the modification of shikimate products, the 3-hydroxylation of carotenoid and flavonoid biosynthesis [70–73]. Nonetheless, the root-specific expressions of the CYP families (72,

76, 81, and 98) and UGT85 (A2, A3, and A4) signify their potential role in the biosynthesis, modification, and diversification of phthalide and phenylpropanoid biosynthesis in *A. glauca* [74,75].

The inter-/intracellular transport of plant secondary metabolites is a critical process to attain metabolic homeostasis for the maintenance of the spatiotemporal biosynthesis and storage of plant secondary metabolites [47]. The leaf-specific expressions of membrane transporters, such as ABCB.26, ABCF.5, and ABCG, indicate the involvement in the cellular transport of mono-sesquiterpenoids and volatile compounds [48,76]. However, the stem-specific enrichments of the ABCB.1, ABCG.29, and NRT families, while SWEET14 and MATE were expressed in the root, indicate their role in the metabolite-flux transportation from the biosynthetic sites to the storage site [31,48].

4. Methods and Materials

4.1. Plant Materials and RNA Isolation

For the tissue-specific transcriptomic sequencing, leaf, stem, and root tissues of *A. glauca* were collected from Ranikot, Chamba district, Himachal Pradesh, India (32.5601° N, 76.1049° E), at an elevation of 2708 amsl. The plant samples were snap-frozen in liquid nitrogen and stored at -80°C till RNA isolation. Total high-quality RNA was extracted using the iRIS protocol [77]. The concentration of RNA was determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA), and the quality was checked in 1% formaldehyde MOPS gel electrophoresis. The integrity of the RNA was measured in terms of the RNA-integrity-number (RIN) value using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and samples having RIN values ≥ 7.0 were subsequently utilized for library preparation. The methodology adopted in this study is represented in Figure 12.

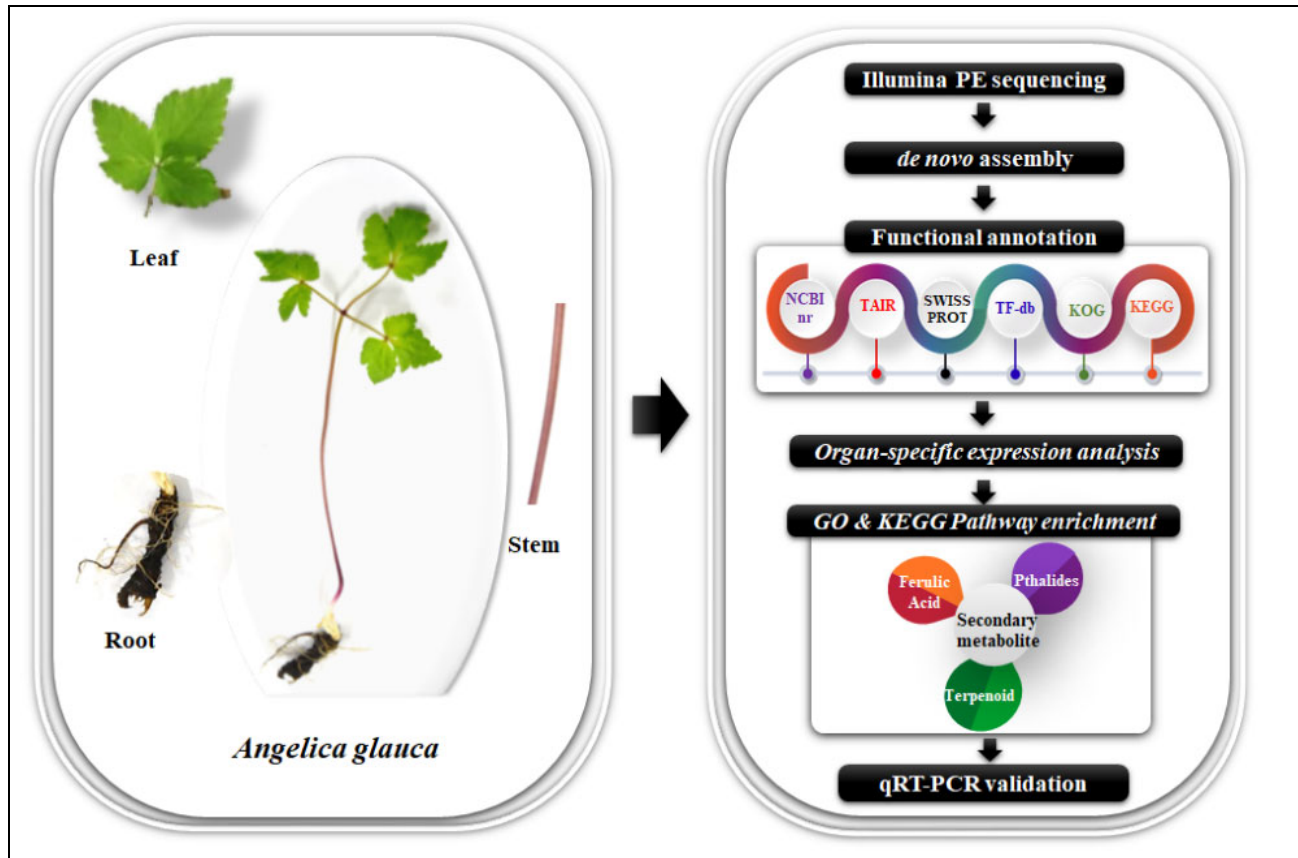


Figure 12. Schematic representation of methodology used for tissue-specific (leaf, stem, and root) transcriptomics study on *A. glauca*.

4.2. Library Construction, Transcriptome Sequencing, and De Novo Assembly

Three RNA-Seq libraries corresponding to the leaf, stem, and root tissues were prepared using the Illumina TruSeq RNA sample prep kit v2 (Illumina Inc., San Diego, CA, USA), following the manufacturer's instructions, targeting 200–300 bp insert-sized libraries. The prepared libraries were quantified using a Qubit 2.0 Fluorimeter (Invitrogen, Waltham, MA, USA), followed by a quality assessment using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The paired-end sequencing (2×100 bp) was performed using the NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA). The raw sequencing data obtained from all three tissues were deposited to the SRA (Sequence Read Archive) in the NCBI database, with the accession number SRP359017, under Bio Project PRJNA 804685. The high-quality clean reads having Phred scores (Q value) > 30 were extracted using the NGS Toolkit [78], and they were subsequently assembled (de novo) using Trinity RNA-Seq version 2.4.0 with default parameters [79]. Furthermore, the assembled sequences were clustered based on a 90% sequence homology using CD-HIT-EST version 4.6 [80].

4.3. Functional Annotation of De Novo-Assembled Transcripts

The functional annotations of the nonredundant (NR) assembled transcripts were performed with the six publicly available protein databases, including NCBI nr, Arabidopsis proteome (TAIR), Swiss-Prot, KEGG, KOG, and Plant TFII-db, using the BLASTx algorithm, with an e-value cutoff $\leq 1 \times 10^{-5}$ [81]. AgriGO version 1.2 was used for the gene-ontology (GO) annotation, classifying transcripts into three major categories (viz., biological processes, cellular components, and molecular functions [82]). The pathway-curation and enrichment analyses were performed using the KEGG pathway-annotation database accessed on 2 July 2020 (<https://www.genome.jp/kegg>) [83,84]. The CYP and UGT families were identified using the *Arabidopsis thaliana* database (<http://www.p450.kvl.dk/index.shtml> (accessed on 2 July 2020)).

4.4. Differential Expression Analysis

To determine the tissue specificity of the differential-gene-expression patterns, the high-quality reads were mapped onto the NR transcripts using the Bowtie 2 tool [85]. The reads mapped onto the assembled transcripts were quantified in terms of FPKM using the RSEM tool [86]. Furthermore, the EdgeR package was used to assess the differential transcript expressions in the following pair-wise comparisons: leaf vs. stem; leaf vs. root; stem vs. root, with the false-discovery rate (FDR < 0.05) and \log_2 fold change (≥ 2) as the differentially expressed genes [87]. The differential-pathway-enrichment analysis was performed by gene-enrichment analysis using the R Bioconductor package, as previously used [46]. The Fischer exact test (Hochberg FDR adjustment cutoff < 0.05) was used for the identification of the organ-specific significant differentially enriched pathways.

4.5. Protein–Protein-Interactome-Network Prediction

The protein–protein-interactome network was predicted by using the predetermined PPI network of *Arabidopsis thaliana* [88]. The orthologs of the significant organ-specific enriched transcripts were obtained with the TAIR database using the BLASTx tool. The interactome network was built and analyzed using Cytoscape ver. 3.4.0 an open-source software maintained by National Institute of General Medical Sciences (NIGMS), U.S [89]. The nodes of the network represent transcripts, with the respective orthologs having significant correlation edges (FDR ≤ 0.05) of the predicted network. The R Bioconductor package was used to perform the pathway curation and gene-set enrichment of the predicted network [90].

4.6. qRT-PCR Validation of RNA-Seq Data

The validation of the RNA-Seq data was performed with the key genes of the pathways using qRT-PCR analysis (ST 8). Gene-specific primers were designed using Batch

Primer 3 ver1.0 (Albany, NY, US) (<http://probes.pw.usda.gov/batchprimer3/> (accessed on 2 July 2020)).

The cDNA was prepared from 2 µg of high-quality total RNA using a cDNA synthesis kit (Thermo Scientific, Waltham, MA, USA, Revert H Minus). The reaction mixture was composed of 1 µL of template cDNA, 0.5 µL of gene-specific primers, 5 µL Power SYBR Green mix (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA), and 3 µL nuclease-free water, and the amplifications were performed in the real-time PCR system (QuantStudio 5, Applied Biosystems, Waltham, MA, USA). The relative expression levels of the genes were calculated from the cycle-threshold (ct) values using the $2^{-\Delta\Delta ct}$ method, considering EF-1 α as a reference gene [91]. Additionally, the expression stabilities of the key candidate genes were validated in four diverse genotypes of *A. glauca*, collected from the natural habitats of the geographically separated populations of Western Himalaya, India (viz., Kinnaur_Asrang (G1: 3376 m, 31.667° N, 78.313° E); Kullu_Kandhaghai (G2: 2250 m, 31.408° N, 77.444° E); Mandi_Thunag (G3: 2177 m, 31.577° N, 77.152° E); Kullu_Gulaba (G4: 2521 m, 32.317° N, 77.189° E) (Table 5)).

4.7. Metabolite Profiling

The phytochemical analysis of the specialized metabolites in the leaf, stem, and root tissues was performed using gas chromatography–mass spectrometry (GC-MS) and ultra-performance liquid chromatography (UPLC-MS). Volatile compounds were extracted and analyzed by a mass spectrometer (QP2010 Shimadzu, Tokyo, Japan), attached with an AOC-5000 auto injector and a DB-5 (SGE International, Ringwood, Australia) fused-silica capillary column of a 30 m length, 0.25 mm i.d., and 0.25 µm film thickness. Helium was utilized as a carrier gas, with a constant flow rate of 1.1 mL/min, and the program was set as in a prior study [92].

Ultra-performance liquid chromatography (UPLC) was used to identify the ferulic acid, and the analysis was carried out using the Water Acquity UPLC-H class system, which includes a binary solvent manager, an autosampler, a column heater, and an e λ photodiode array detector (PDA). All the parameters used in the analysis followed a previous study [93], and the chromatographic data were recorded for each sample and were quantified based on internal standard curves.

5. Conclusions

The specialized metabolites of the high-altitude Himalayan endangered species *A. glauca* are a vital resource for the pharmaceutical and aromatic industries. The increasing demand has led to the indeterminate extraction of natural plant populations. Therefore, the creation of the first NGS-assisted comprehensive spatial genomic resource has assisted in the identification of the tissue-specific key regulators that influence the specialized-metabolite biosynthetic pathways in *A. glauca*, which will enable the identification of the targeted secondary-metabolite-biosynthesis pathways in other medicinal plants. Furthermore, the tissue-specific molecular insights into the terpenoid, ferulic acid, and phthalide biosynthesis will assist leaf-, stem-, and root-specific genetic interventions to upscale the targeted metabolite synthesis with the active involvement of TFs, CYPs, UGTs, and transporters that correspond to the biosynthesis, transportation, and diversification of secondary metabolites. The present study paves the way to understand the multifarious metabolic genes for the synthesis of diverse biomolecules in plant species. Creating genomic resources from the current study will formulate future steps to enhance metabolite production by genetic engineering and molecular functionally relevant marker resources for the population diversity and conservation planning of this important Himalayan medicinal plant.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231911064/s1>.

Author Contributions: R.K.S. conceived and designed the study; A.D., A.H. and S.S. performed the experiments; A.S. provided the plant materials; A.D., R.S., M.M. and G.S. analyzed the data; A.D. and R.S. wrote the draft manuscript; R.K.S. edited and approved the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Council of Scientific and Industrial Research grant number MLP0201.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw reads were deposited to the NCBI SRA database with the accession number SRP359017 under Bio Project PRJNA 804685.

Acknowledgments: The Director, CSIR-IHBT, is gratefully acknowledged for providing the necessary facilities for conducting the research. This is CSIR-IHBT publication no. 4920.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

IHR	Indian Himalayan Region
MAPs	Medicinal and Aromatic Plant species
UGT	UDP-Dependent Glucuronosyltransferase
ABC	ATP-binding cassette
CYP450	Cytochrome P450
DXS	1-deoxy-D-xylulose-5-phosphate synthase
DXR	1-deoxy-D-xylulose 5-phosphate reductoisomerase
MCT	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase
CMK	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
MECS	2-C-methyl-D-erythritol 2,4-cyclodiphosphatesynthase
HDS	4-hydroxy-3-methylbut-2-enyl diphosphate synthase
HDR	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
AACT	Acetyl Co-A acetyltransferase
HMGS	HMG-CoA synthase
HMGR	HMG-CoA reductase
MK	Mevalonate kinase
MPK	Phosphomevalonate kinase
MVD	Di phosphomevalonate decarboxylase
IDI	Isopentenyl-diphosphate delta-isomerase
GPPS	Geranyl pyrophosphate synthase
FPPS	Farnyl pyrophosphate synthase
STS	Sesquiterpene synthase
LS	Limonene synthase
NMD	Neomenthol dehydrogenase
CD	Carveol dehydrogenase
NES	Nerolidol synthase
DHS	3-deoxy-D-arabino-hept-2-ulosonate 7-phosphate synthase
DHQS	3-dehydroquininate synthase
DHQ-SDH	3-dehydroquininate dehydratase/shikimate 5-dehydrogenase
SK	Shikimate kinase
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase
CS	Chorismate synthase
CM	Chorismate mutase
PAT	Prephenate aminotransferase
ADT	Arogenate dehydratase
PAL	Phenylalanine ammonia lyase
C4H	Cinnamate-4-hydroxylase
C3H	Coumarate 3-hydroxylase
COMT	Caffeic acid 3-O-methyltransferase
NADPH-AOR	NADPH-dependent alkenal/one oxidoreductase

SDM S-adenosyl-L-methionine-dependent Methyltransferases
 PKC Polyketide cyclase

References

- Rodgers, W.A.; Panwar, H.S. *Planning a Wildlife Protected Area Network in India*; A Report; Wildlife Institute of India: Dehradun, India, 1988.
- Pal, M.; Samant, S.S. Diversity, distribution and conservation of economically important medicinal and aromatic plants of the Indian Himalayan region. *Blood Purif.* **2005**, *1500*, 2500.
- Kumar, A.; Jnanasha, A.C. Conservation of Rare and Endangered Plant Species for Medicinal. *Use. Int. J. Sci. Res.* **2016**, *5*, 1370–1372.
- Samarth, R.M.; Samarth, M.; Matsumoto, Y. Medicinally important aromatic plants with radioprotective activity. *Future Sci.* **2017**, *3*, 247. [[CrossRef](#)] [[PubMed](#)]
- Gopi, D.K.; Mattummal, R.; Narayana, S.K.K.; Parameswaran, S. IUCN Red Listed Medicinal Plants of Siddha. *J. Res. Sid. Med.* **2018**, *1*, 15–22.
- Thakur, S.D. Phytochemical constituents of some important medicinal plants from Dachigam National Park, Srinagar, Jammu and Kashmir. *Pharma Innov.* **2019**, *8*, 68–71.
- Nag, A.; Chanda, S.; Rajkumar, S. Estimation of nuclear genome size of important medicinal plant species from western Himalaya using flow cytometry. *J. Cell Plant Sci.* **2011**, *2*, 19–23.
- Purohit, V.K.; Andola, H.C.; Haider, S.Z.; Tiwari, D.; Bahuguna, Y.M.; Gairola, K.C.; Arunachalam, K. Essential oil constituents of *Angelica glauca* Edgew. roots: An endangered species from Uttarakhand Himalaya (India). *Natl. Acad. Sci. Lett.* **2015**, *38*, 445–447. [[CrossRef](#)]
- Joshi, R.K. *Angelica (Angelica glauca and A. archangelica) Oils*. In *Essential Oils in Food Preservation, Flavor and Safety*; Academic Press: Cambridge, MA, USA, 2016; pp. 203–208. [[CrossRef](#)]
- Gautam, K.; Raina, R. Floral architecture, breeding system, seed biology and chromosomal studies in endangered Himalayan *Angelica glauca* Edgew.(Apiaceae). *Caryologia Int. J. Cytol. Cytosystematics Cytogenet.* **2019**, *72*, 23–34. [[CrossRef](#)]
- Bisht, A.K.; Bhatt, A.; Rawal, R.S.; Dhar, U. Assessment of reproductive potential of different populations of *Angelica glauca* Edgew., a critically endangered Himalayan medicinal herb. *J. Mt. Sci.* **2008**, *5*, 84–90. [[CrossRef](#)]
- Rawat, J.M.; Bhandari, A.; Mishra, S.; Rawat, B.; Dhakad, A.K.; Thakur, A.; Chandra, A. Genetic stability and phytochemical profiling of the in vitro regenerated plants of *Angelica glauca* Edgew: An endangered medicinal plant of Himalaya. *Plant Cell Tissue Org. Cult.* **2018**, *135*, 111–118. [[CrossRef](#)]
- Agnihotri, V.K.; Thappa, R.K.; Meena, B.; Kapahi, B.K.; Saxena, R.K.; Qazi, G.N.; Agarwal, S.G. Essential oil composition of aerial parts of *Angelica glauca* growing wild in North-West Himalaya(India). *Phytochemistry* **2004**, *65*, 2411–2413. [[CrossRef](#)] [[PubMed](#)]
- Irshad, M.; Shahid, M.; Aziz, S.; Ghous, T. Antioxidant, antimicrobial and phytotoxic activities of essential oil of *Angelica glauca*. *Asian J. Chem.* **2011**, *23*, 1947.
- Chung, J.W.; Choi, R.J.; Seo, E.K.; Nam, J.W.; Dong, M.S.; Shin, E.M.; Guo, L.Y.; Kim, Y.S. Anti-inflammatory effects of (Z)-ligustilide through suppression of mitogen-activated protein kinases and nuclear factor- κ B activation pathways. *Arch. Pharm. Res.* **2012**, *35*, 723–732. [[CrossRef](#)] [[PubMed](#)]
- Butola, J.S.; Vashistha, R.K. An overview on conservation and utilization of *Angelica glauca* Edgew. in three Himalayan states of India. *Med. Plants-Int. J. Phytomed. Relat. Ind.* **2013**, *5*, 171–178. [[CrossRef](#)]
- El Maddah, F.; Eguereva, E.; Kehraus, S.; König, G.M. Biosynthetic studies of novel polyketides from the marine sponge-derived fungus *Stachylidium* sp. 293K04Org. *Biomol. Chem.* **2019**, *17*, 2747–2752. [[CrossRef](#)]
- Kong, J.Q. Phenylalanine ammonia-lyase, a key component used for phenylpropanoids production by metabolic engineering. *RSC Adv.* **2015**, *5*, 62587–62603. [[CrossRef](#)]
- Rastogi, S.; Shah, S.; Kumar, R.; Kumar, A.; Shasany, A.K. Comparative temporal metabolomics studies to investigate interspecies variation in three *Ocimum* species. *Sci. Rep.* **2020**, *10*, 5234. [[CrossRef](#)]
- Henry, L.K.; Thomas, S.T.; Widhalm, J.R.; Lynch, J.H.; Davis, T.C.; Kessler, S.A.; Bohlmann, J.; Noel, J.P.; Dudareva, N. Contribution of isopentenyl phosphate to plant terpenoid metabolism. *Nat. Plants* **2018**, *4*, 721–729. [[CrossRef](#)]
- Sharma, S.; Rasal, V.P.; Patil, P.A.; Joshi, R.K. Effect of *Angelica glauca* essential oil on allergic airway changes induced by histamine and ovalbumin in experimental animals. *Indian J. Pharmacol.* **2017**, *49*, 55. [[CrossRef](#)]
- Lewin, H.A.; Richards, S.; Aiden, E.L.; Allende, M.L.; Archibald, J.M.; Bálint, M.; Barker, K.B.; Baumgartner, B.; Belov, K.; Bertorelle, G.; et al. The earth BioGenome project 2020: Starting the clock. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2115635118. [[CrossRef](#)]
- Unamba, C.I.; Nag, A.; Sharma, R.K. Next generation sequencing technologies: The doorway to the unexplored genomics of non-model plants. *Front. Plant Sci.* **2015**, *6*, 1074. [[CrossRef](#)] [[PubMed](#)]
- Chen, X.; Li, J.; Wang, X.; Zhong, L.; Tang, Y.; Zhou, X.; Liu, Y.; Zhan, R.; Zheng, H.; Chen, W.; et al. Full-length transcriptome sequencing and methyl jasmonate-induced expression profile analysis of genes related to patchoulol biosynthesis and regulation in *Pogostemon cablin*. *BMC Plant Biol.* **2019**, *19*, 266. [[CrossRef](#)] [[PubMed](#)]

25. Pattison, R.J.; Csukasi, F.; Zheng, Y.; Fei, Z.; van der Knaap, E.; Catalá, C. Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. *Plant Physiol.* **2015**, *168*, 1684–1701. [[CrossRef](#)] [[PubMed](#)]
26. Feng, G.; Wu, J.; Yi, H. Global tissue-specific transcriptome analysis of *Citrus sinensis* fruit across six developmental stages. *Sci. Data* **2019**, *6*, 153. [[CrossRef](#)] [[PubMed](#)]
27. Sharma, B.; Seth, R.; Thakur, S.; Parmar, R.; Masand, M.; Devi, A.; Singh, G.; Dhyani, P.; Choudhary, S.; Sharma, R.K. Genome-wide transcriptional analysis unveils the molecular basis of organ-specific expression of isosteroidal alkaloids biosynthesis in critically endangered *Fritillaria roylei* Hook. *Phytochemistry* **2021**, *187*, 112772. [[CrossRef](#)] [[PubMed](#)]
28. Yang, C.Q.; Fang, X.; Wu, X.M.; Mao, Y.B.; Wang, L.J.; Chen, X.Y. Transcriptional regulation of plant secondary metabolism. *J. Integr. Plant Biol.* **2012**, *54*, 703–712. [[CrossRef](#)] [[PubMed](#)]
29. Jung, S.C.; Kim, W.; Park, S.C.; Jeong, J.; Park, M.K.; Lim, S.; Lee, Y.; Im, W.T.; Lee, J.H.; Choi, G.; et al. Two ginseng UDP-glycosyltransferases synthesize ginsenoside Rg3 and Rd. *Plant Cell Physiol.* **2014**, *55*, 2177–2188. [[CrossRef](#)]
30. Naveed, A.; Li, H.; Liu, X. Cytochrome P450s: Blueprints for Potential Applications in Plants. *J. Plant Biochem. Physiol.* **2018**, *6*, 204. [[CrossRef](#)]
31. Yazaki, K.; Sugiyama, A.; Morita, M.; Shitan, N. Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites. *Phytochem. Rev.* **2008**, *7*, 513–524. [[CrossRef](#)]
32. Younesi-Melerdi, E.; Nematzadeh, G.A.; Pakdin-Parizi, A.; Bakhtiarizadeh, M.R.; Motahari, S.A. De novo RNA sequencing analysis of *Aeluropus litoralis* halophyte plant under salinity stress. *Sci. Res.* **2020**, *10*, 9148. [[CrossRef](#)]
33. Lei, B.; Chang, W.; Zhao, H.; Zhang, K.; Yu, J.; Yu, S.; Cai, K.; Zhang, J.; Lu, K. Nitrogen application and differences in leaf number retained after topping affect the tobacco (*Nicotiana tabacum*) transcriptome and metabolome. *BMC Plant Biol.* **2022**, *22*, 38. [[CrossRef](#)]
34. Pillai, S.G.; Yadav, Y.; Sharma, K.C. A review on an endangered himalayan medicinal aromatic plant-choraka (*Angelica glauca* edgew). *IAMJ* **2020**, *8*, 4509–4515. [[CrossRef](#)]
35. Singh, G.; Chandra, N.; Lingwal, S.; Bisht, M.P.S.; Tiwari, L.M. Distribution and threat assessment of an endemic and endangered species *Angelica glauca* in high ranges of western Himalaya. *J. Herbs Spices Med. Plants* **2020**, *26*, 394–404. [[CrossRef](#)]
36. IUCN. The IUCN Red List of Threatened Species. Version 2021. Available online: <http://www.iucnredlist.org> (accessed on 7 November 2021).
37. Chen, C.; Chen, Y.; Huang, W.; Jiang, Y.; Zhang, H.; Wu, W. Mining of simple sequence repeats (SSRs) loci and development of novel transferability-across EST-SSR markers from de novo transcriptome assembly of *Angelica dahurica*. *PLoS ONE* **2019**, *14*, e0221040. [[CrossRef](#)]
38. Xu, R.; Xu, J.; Li, Y.C.; Dai, Y.T.; Zhang, S.P.; Wang, G.; Liu, Z.G.; Dong, L.L.; Chen, S.L. Integrated chemical and transcriptomic analyses unveil synthetic characteristics of different medicinal root parts of *Angelica sinensis*. *Chin. Herb. Med.* **2020**, *12*, 19–28. [[CrossRef](#)] [[PubMed](#)]
39. Drew, D.P.; Dueholm, B.; Weitzel, C.; Zhang, Y.; Sensen, C.W.; Simonsen, H.T. Transcriptome analysis of *Thapsia laciniata* Rouy provides insights into terpenoid biosynthesis and diversity in Apiaceae. *Int. J. Mol. Sci.* **2013**, *14*, 9080–9098. [[CrossRef](#)]
40. Amini, H.; Naghavi, M.R.; Shen, T.; Wang, Y.; Nasiri, J.; Khan, I.A.; Fiehn, O.; Zerbe, P.; Maloof, J.N. Tissue-specific transcriptome analysis reveals candidate genes for terpenoid and phenylpropanoid metabolism in the medicinal plant *Ferula assafoetida*. *G3 Genes Genomes Genet.* **2019**, *9*, 807–816. [[CrossRef](#)]
41. Lovén, J.; Orlando, D.A.; Sigova, A.A.; Lin, C.Y.; Rahl, P.B.; Burge, C.B.; Levens, D.L.; Lee, T.I.; Young, R.A. Revisiting global gene expression analysis. *Cell* **2012**, *151*, 476–482. [[CrossRef](#)]
42. Huang, Q.; Huang, X.; Deng, J.; Liu, H.; Liu, Y.; Yu, K.; Huang, B. Differential gene expression between leaf and rhizome in *Atractylodes lancea*: A comparative transcriptome analysis. *Front. Plant Sci.* **2016**, *7*, 348. [[CrossRef](#)]
43. Wang, T.; Li, B.; Nelson, C.E.; Nabavi, S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinform.* **2019**, *20*, 40. [[CrossRef](#)]
44. Nag, A.; Choudhary, S.; Masand, M.; Parmar, R.; Bhandawat, A.; Seth, R.; Singh, G.; Dhyani, P.; Sharma, R.K. Spatial transcriptional dynamics of geographically separated genotypes revealed key regulators of podophyllotoxin biosynthesis in *Podophyllum hexandrum*. *Ind. Crop. Prod.* **2020**, *147*, 112247. [[CrossRef](#)]
45. Singh, P.; Singh, G.; Bhandawat, A.; Singh, G.; Parmar, R.; Seth, R.; Sharma, R.K. Spatial transcriptome analysis provides insights of key gene (s) involved in steroidal saponin biosynthesis in medicinally important herb *Trillium govanianum*. *Sci. Res.* **2017**, *7*, 45295. [[CrossRef](#)] [[PubMed](#)]
46. Seth, R.; Bhandawat, A.; Parmar, R.; Singh, P.; Kumar, S.; Sharma, R.K. Global transcriptional insights of pollen-pistil interactions commencing self-incompatibility and fertilization in tea [*Camellia sinensis* (L.) O. Kuntze]. *Int. J. Mol. Sci.* **2019**, *20*, 539. [[CrossRef](#)] [[PubMed](#)]
47. Maritim, T.K.; Seth, R.; Parmar, R.; Sharma, R.K. Multiple-genotypes transcriptional analysis revealed candidate genes and nucleotide variants for improvement of quality characteristics in tea (*Camellia sinensis* (L.) O. Kuntze). *Genomics* **2021**, *113*, 305–316. [[CrossRef](#)]
48. Seth, R.; Maritim, T.K.; Parmar, R.; Sharma, R.K. Underpinning the molecular programming attributing heat stress associated thermotolerance in tea (*Camellia sinensis* (L.) O. Kuntze). *Hortic. Res.* **2021**, *8*, 99. [[CrossRef](#)] [[PubMed](#)]
49. Smetanska, I. Production of secondary metabolites using plant cell cultures. *Food Biotechnol.* **2008**, *111*, 187–228. [[CrossRef](#)]

50. Gani, U.; Vishwakarma, R.A.; Misra, P. Membrane transporters: The key drivers of transport of secondary metabolites in plants. *Plant Cell Rep.* **2020**, *40*, 1–18. [[CrossRef](#)]
51. Singh, G.; Singh, G.; Singh, P.; Parmar, R.; Paul, N.; Vashist, R.; Swarnkar, M.K.; Kumar, A. Molecular dissection of transcriptional reprogramming of steviol glycosides synthesis in leaf tissue during developmental phase transitions in *Stevia rebaudiana* Bert. *Sci. Rep.* **2017**, *7*, 11835. [[CrossRef](#)]
52. Shih, M.L.; Morgan, J.A. Metabolic flux analysis of secondary metabolism in plants. *Metab. Eng. Commun.* **2020**, *10*, 00123. [[CrossRef](#)]
53. Patra, B.; Schluttenhofer, C.; Wu, Y.; Pattanaik, S.; Yuan, L. Transcriptional regulation of secondary metabolite biosynthesis in plants. *Biochim. Biophys. Acta* **2013**, *1829*, 12361247. [[CrossRef](#)]
54. Van Moerkercke, A.; Steensma, P.; Schweizer, F.; Pollier, J.; Gariboldi, I.; Payne, R.; Bossche, R.V.; Miettinen, K.; Espoz, J.; Purnama, P.C.; et al. The bHLH transcription factor BIS1 controls the iridoid branch of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8130–8135. [[CrossRef](#)] [[PubMed](#)]
55. Schluttenhofer, C.; Yuan, L. Regulation of specialized metabolism by WRKY transcription factors. *Plant Physiol.* **2015**, *167*, 295–306. [[CrossRef](#)] [[PubMed](#)]
56. Chahel, A.A.; Zeng, S.; Yousaf, Z.; Liao, Y.; Yang, Z.; Wei, X.; Ying, W. Plant-specific transcription factor LrTCP4 enhances secondary metabolite biosynthesis in *Lycium ruthenicum* hairy roots. *Plant Cell Tissue Org. Cult.* **2019**, *136*, 323–337. [[CrossRef](#)]
57. Kayani, S.I.; Shen, Q.; Rahman, S.U.; Fu, X.; Li, Y.; Wang, C.; Hassani, D.; Tang, K. Transcriptional regulation of flavonoid biosynthesis in *Artemisia annua* by AaYABBY5. *Hortic. Res.* **2021**, *8*, 257. [[CrossRef](#)]
58. Lv, Z.; Wang, Y.; Liu, Y.; Peng, B.; Zhang, L.; Tang, K.; Chen, W. The SPB-box transcription factor AaSPL2 positively regulates artemisinin biosynthesis in *Artemisia annua* L. *Front. Plant Sci.* **2019**, *10*, 409. [[CrossRef](#)] [[PubMed](#)]
59. Turgut-Kara, N.; Cakir, O. Comparative phylogenetic analysis of phenylpropanoid metabolism genes of legume plants. *Plant Omics* **2015**, *8*, 55. [[CrossRef](#)]
60. Vranová, E.; Coman, D.; Gruissem, W. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* **2013**, *64*, 665–700. [[CrossRef](#)]
61. Yang, L.; Ding, G.; Lin, H.; Cheng, H.; Kong, Y.; Wei, Y.; Fang, X.; Liu, R.; Wang, L.; Chen, X.; et al. Transcriptome analysis of medicinal plant *Salvia miltiorrhiza* and identification of genes related to tanshinone biosynthesis. *PLoS ONE* **2013**, *8*, e80464. [[CrossRef](#)]
62. Devi, K.; Mishra, S.K.; Sahu, J.; Panda, D.; Modi, M.K.; Sen, P. Genome wide transcriptome profiling reveals differential gene expression in secondary metabolite pathway of *Cymbopogon winterianus*. *Sci. Rep.* **2016**, *6*, 21026. [[CrossRef](#)]
63. Regueira, T.B.; Kildegaard, K.R.; Hansen, B.G.; Mortensen, U.H.; Hertweck, C.; Nielsen, J. Molecular basis for mycophenolic acid biosynthesis in *Penicillium brevicompactum*. *Appl. Environ. Microbiol.* **2011**, *77*, 3035–3043. [[CrossRef](#)]
64. Sanchez, J.F.; Entwistle, R.; Corcoran, D.; Oakley, B.R.; Wang, C.C. Identification and molecular genetic analysis of the cichorine gene cluster in *Aspergillus nidulans*. *MedChemComm* **2012**, *3*, 997–1002. [[CrossRef](#)]
65. Yaegashi, J.; Praseuth, M.B.; Tyan, S.W.; Sanchez, J.F.; Entwistle, R.; Chiang, Y.M.; Oakley, B.R.; Wang, C.C. Molecular genetic characterization of the biosynthesis cluster of a prenylated isoindolinone alkaloid aspermidine A in *Aspergillus nidulans*. *Org. Lett.* **2013**, *15*, 2862–2865. [[CrossRef](#)] [[PubMed](#)]
66. Yonekura-Sakakibara, K.; Hanada, K. An evolutionary view of functional diversity in family 1 glycosyltransferases. *Plant J.* **2011**, *66*, 182–193. [[CrossRef](#)] [[PubMed](#)]
67. Seki, H.; Tamura, K.; Muranaka, T. P450s and UGTs: Key players in the structural diversity of triterpenoid saponins. *Plant Cell Physiol.* **2015**, *56*, 1463–1471. [[CrossRef](#)] [[PubMed](#)]
68. Hori, K.; Yamada, Y.; Purwanto, R.; Minakuchi, Y.; Toyoda, A.; Hirakawa, H.; Sato, F. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant Cell Physiol.* **2018**, *59*, 222–233. [[CrossRef](#)] [[PubMed](#)]
69. Yuan, Y.; Zhang, J.; Liu, X.; Meng, M.; Wang, J.; Lin, J. Tissue-specific transcriptome for *Dendrobium officinale* reveals genes involved in flavonoid biosynthesis. *Genomics* **2020**, *112*, 1781–1794. [[CrossRef](#)] [[PubMed](#)]
70. Morant, M.; Schoch, G.A.; Ullmann, P.; Ertunç, T.; Little, D.; Olsen, C.E.; Petersen, M.; Negrel, J.; Werck-Reichhart, D. Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. *Plant Mol. Biol.* **2007**, *63*, 1–19. [[CrossRef](#)]
71. Cui, H.; Yu, X.; Wang, Y.; Cui, Y.; Li, X.; Liu, Z.; Qin, S. Evolutionary origins, molecular cloning and expression of carotenoid hydroxylases in eukaryotic photosynthetic algae. *BMC Genom.* **2013**, *14*, 457. [[CrossRef](#)]
72. Magome, H.; Nomura, T.; Hanada, A.; Takeda-Kamiya, N.; Ohnishi, T.; Shinma, Y.; Katsumata, T.; Kawaide, H.; Kamiya, Y.; Yamaguchi, S. 2013. CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1947–1952. [[CrossRef](#)]
73. Lin, J.S.; Huang, X.X.; Li, Q.; Cao, Y.; Bao, Y.; Meng, X.F.; Li, Y.J.; Fu, C.; Hou, B.K. UDP-glycosyltransferase 72B1 catalyzes the glucose conjugation of monolignols and is essential for the normal cell wall lignification in *Arabidopsis thaliana*. *Plant J.* **2016**, *88*, 26–42. [[CrossRef](#)]
74. Irmeler, S.; Schröder, G.; St-Pierre, B.; Crouch, N.P.; Hotze, M.; Schmidt, J.; Strack, D.; Matern, U.; Schröder, J. Indole alkaloid biosynthesis in *Catharanthus roseus*: New enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. *Plant J.* **2000**, *24*, 797–804. [[CrossRef](#)] [[PubMed](#)]

75. Jun, X.U.; Wang, X.Y.; Guo, W.Z. The cytochrome P450 superfamily: Key players in plant development and defense. *J. Integr. Agric.* **2015**, *14*, 1673–1686. [[CrossRef](#)]
76. Fu, X.; Shi, P.; He, Q.; Shen, Q.; Tang, Y.; Pan, Q.; Ma, Y.; Yan, T.; Chen, M.; Hao, X.; et al. AaPDR3, a PDR transporter 3, is involved in sesquiterpene β -caryophyllene transport in *Artemisia annua*. *Front. Plant Sci.* **2017**, *8*, 723. [[CrossRef](#)]
77. Ghawana, S.; Paul, A.; Kumar, H.; Kumar, A.; Singh, H.; Bhardwaj, P.K.; Rani, A.; Singh, R.S.; Raizada, J.; Singh, K.; et al. An RNA isolation system for plant tissues rich in secondary metabolites. *BMC Res. Notes* **2011**, *4*, 85. [[CrossRef](#)] [[PubMed](#)]
78. Patel, R.K.; Jain, M. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS ONE* **2012**, *7*, e30619. [[CrossRef](#)] [[PubMed](#)]
79. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644. [[CrossRef](#)]
80. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)]
81. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
82. Du, Z.; Zhou, X.; Ling, Y.; Zhang, Z.; Su, Z. agriGO: A GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* **2010**, *38*, 64–70. [[CrossRef](#)]
83. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)]
84. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, 457–462. [[CrossRef](#)] [[PubMed](#)]
85. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357. [[CrossRef](#)] [[PubMed](#)]
86. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)] [[PubMed](#)]
87. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. EdgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)] [[PubMed](#)]
88. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P.; et al. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2015**, *43*, 447–452. [[CrossRef](#)]
89. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
90. Yu, G.; He, Q.Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol. BioSyst.* **2016**, *12*, 477–479. [[CrossRef](#)]
91. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2[−] $\Delta\Delta$ CT method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)]
92. Walia, S.; Kumar, R. Elucidating the yield and quality response of *Tagetes minuta* L. intercropped with *Zea mays* L. under different spacing in the western Himalayas. *Ind. Crop. Prod.* **2021**, *171*, 113850. [[CrossRef](#)]
93. Dadwal, V.; Agrawal, H.; Sonkhla, K.; Joshi, R.; Gupta, M. Characterization of phenolics, amino acids, fatty acids and antioxidant activity in pulp and seeds of high altitude Himalayan crab apple fruits (*Malus baccata*). *Food Sci. Technol.* **2018**, *55*, 2160–2169. [[CrossRef](#)]