



Article

# Multi-Omics Analysis Decodes Biosynthesis of Specialized Metabolites Constituting the Therapeutic Terrains of *Magnolia obovata*

Megha Rai <sup>1,2,†</sup>, Amit Rai <sup>2,†</sup>, Towa Yokosaka <sup>1</sup>, Tetsuya Mori <sup>3</sup> , Ryo Nakabayashi <sup>3</sup>, Michimi Nakamura <sup>1</sup>, Hideyuki Suzuki <sup>4</sup>, Kazuki Saito <sup>3,5</sup> and Mami Yamazaki <sup>1,5,\*</sup>

- <sup>1</sup> Graduate School of Pharmaceutical Sciences, Chiba University, Chiba 260-8675, Japan; megha@illinois.edu (M.R.); ju.zero19990407@gmail.com (T.Y.); michimi@chiba-u.jp (M.N.)  
<sup>2</sup> Crop Sciences, University of Illinois Urbana, Champaign, IL 61801, USA; amitrai9@illinois.edu  
<sup>3</sup> RIKEN Center for Sustainable Resource Science, Yokohama 230-0045, Japan; tetsuya.mori@riken.jp (T.M.); roy.nakbayashi@gmail.com (R.N.); kazuki.saito@riken.jp (K.S.)  
<sup>4</sup> Kazusa DNA Research Institute, Chiba 292-0818, Japan; h\_suzuki@hirata.co.jp  
<sup>5</sup> Plant Molecular Science Center, Chiba University, Chiba 260-8675, Japan  
\* Correspondence: mamiy@faculty.chiba-u.jp  
† These authors contributed equally to this work.

**Abstract:** *Magnolia obovata* is renowned for its unique bioactive constituents with medicinal properties traditionally used to treat digestive disorders, anxiety, and respiratory conditions. This study aimed to establish a comprehensive omics resource through untargeted metabolome and transcriptome profiling to explore biosynthesis of pharmacologically active compounds of *M. obovata* using seven tissues: young leaf, mature leaf, stem, bark, central cylinder, floral bud, and pistil. Untargeted metabolomic analysis identified 6733 mass features across seven tissues and captured chemo-diversity and its tissue-specificity in *M. obovata*. Through a combination of cheminformatics and manual screening approach, we confirmed the identities of 105 metabolites, including neolignans, such as honokiol and magnolol, which were found to be spatially accumulated in the bark tissue. RNA sequencing generated a comprehensive transcriptome resource, and expression analysis revealed significant tissue-specific expression patterns. Omics dataset integration identified T12 transcript module from WGCNA being correlated with the biosynthesis of magnolol and honokiol in *M. obovata*. Notably, phylogenetic analysis using transcripts from T12 module identified two laccase (Mo\_LAC1 and Mo\_LAC2) and three dirigent proteins from the DIR-b/d subfamily as potential candidate genes involved in neolignan biosynthesis. This research established omics resources of *M. obovata* and laid the groundwork for future studies aimed at optimizing and further understanding the biosynthesis of metabolites of therapeutic potential.

**Keywords:** neolignans; integrative omics; metabolomics; natural products; laccases



Academic Editor: Milen I. Georgiev

Received: 5 January 2025

Revised: 21 January 2025

Accepted: 24 January 2025

Published: 26 January 2025

**Citation:** Rai, M.; Rai, A.; Yokosaka, T.; Mori, T.; Nakabayashi, R.; Nakamura, M.; Suzuki, H.; Saito, K.; Yamazaki, M. Multi-Omics Analysis Decodes Biosynthesis of Specialized Metabolites Constituting the Therapeutic Terrains of *Magnolia obovata*. *Int. J. Mol. Sci.* **2025**, *26*, 1068. <https://doi.org/10.3390/ijms26031068>

**Copyright:** © 2025 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

*Magnolia obovata*, commonly known as “Hōnoki” in Japanese, belongs to the Magnoliaceae family and is native to eastern Asia, including Japan, China, and Korea [1]. It is a deciduous tree grown widely in Europe and America for ornamental purposes and has been used as traditional medicine in Asia for centuries [2]. In addition to its medicinal uses, plants extract from genus *Magnolia* are also used as dietary supplements and cosmetic products [3,4]. From nearly 220 known *Magnolia* species, *M. obovata* and *M. officinalis* are

the most commonly used species used in traditional Chinese and Japanese herbal medicine and are incorporated in the Japanese pharmacopeia [4]. The bark of *M. officinalis* and *M. obovata*, also known as Houpu and Honoki, respectively, has been used in several medicinal formulations, including Banxia Houpu Tang and Hange-Kouboku-to, among others [4,5]. Two of the prescriptions in Japan, Sai-Boku-to and Hange-Kouboku-to, contain *Magnolia obovata* bark and are still used in modern clinical practice [4]. The bark tissue of *M. obovata* is highly regarded for its anti-cancer, anti-inflammatory, antioxidant, anti-stress, anti-anxiety, anti-depressant, anti-Alzheimer, and anti-stroke effects [3,6]. Additionally, it has also been shown to have vascular relaxation, hepatoprotective, and anti-diarrhea effects [4]. While the bark has been traditionally used for medicinal purposes, several studies have also reported the pharmacological activities of the floral buds and leaves of *Magnolia* plants, providing a basis for investigations into the use of other plant tissues for its use as drugs [4,7].

The medicinal properties of *M. obovata* are attributed to the presence of a diverse array of bioactive compounds, such as lignans, neolignans, terpenoids, alkaloids, and flavonoids [2,8,9]. To date, around 255 biologically active substances have been isolated from different tissues of the genus *Magnolia*, of which neolignans, namely magnolol and honokiol, are considered the principal active ingredients [4,10]. It is believed that magnolol and honokiol are the key metabolite constituents responsible for the therapeutic effects of *Magnolia* plants [2]. However, only 31 metabolites have been reported specifically for *M. obovata* in the KNApSack database [11]. Despite the medicinal importance of *M. obovata*, the biosynthetic pathways and regulatory mechanisms underlying the biosynthesis of therapeutic compounds in this plant species remain largely unexplored. Elucidating the biosynthetic genes and pathways associated with bioactive compounds in *M. obovata* can provide valuable insights into the molecular basis of its medicinal properties and enable potential applications in medicinal and pharmaceutical industries. The utilization of multi-omics resources has been demonstrated as an effective means to identify natural variants with desired properties, constituting an essential element for the sustainable production of bioactive metabolites from medicinal plants.

The surge in advanced omics technologies, driven by reduced cost for data generation, enhanced computational power for data integration and analysis, and high-throughput and high-resolution data generation capabilities. These advancements have emerged as powerful tools for studying the biosynthesis of bioactive compounds in plants [12]. In the absence of genomic resources, RNA-seq based de novo transcriptome assembly has been established as a valuable resource for understanding the biosynthesis of important specialized metabolites in non-model plant species such as medicinal plants [13–16]. Untargeted metabolome analysis insights into the accumulation patterns of metabolites including bioactive compounds in different plant tissues, while transcriptome analysis provides information on the expression patterns of molecular elements involved in the biosynthesis of these metabolites [12,13]. While the individual analysis of omics datasets can provide systems characteristics at the metabolite and gene expression levels, integration of multi-omics datasets assists in identifying metabolite-gene associations and to prioritize key enzymes by removing the false positives for functional characterization, thereby providing a comprehensive understanding of their biosynthetic pathway in plants [17–20].

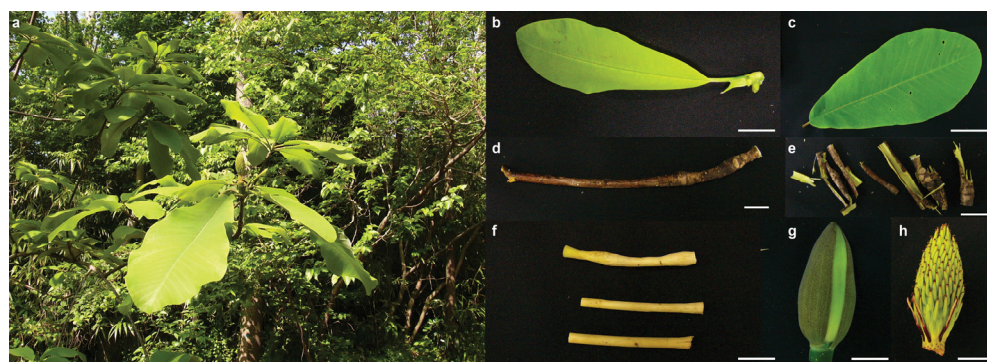
In this study, we aimed to unravel the biosynthetic genes associated with key metabolites of therapeutic importance by adopting an integrated metabolome and transcriptome analysis approach by using seven tissues of *M. obovata*, including young leaf, mature leaf, stem, bark, central cylinder, floral bud, and pistil. Using untargeted metabolomics, we investigated the metabo-constituents of different tissues of *M. obovata*. Transcriptome analysis identified the ongoing biological processes and the actively expressed transcripts.

Furthermore, module-based analysis and integration of the -omics datasets facilitated the identification of potential transcripts involved in the biosynthesis of neolignans, namely magnolol and honokiol. Our findings provide valuable insights into the molecular basis of the biosynthesis of therapeutic compounds in *M. obovata*. This study presents comprehensive -omics resources for *M. obovata* and will help identify candidate therapeutic metabolites in different tissues for future validation, structural determination, and functional characterization of associated biosynthetic enzymes.

## 2. Results

### 2.1. Metabo-Constituents of *Magnolia obovata* and Its Chemo-Diversity Revealed by Untargeted Metabolomics

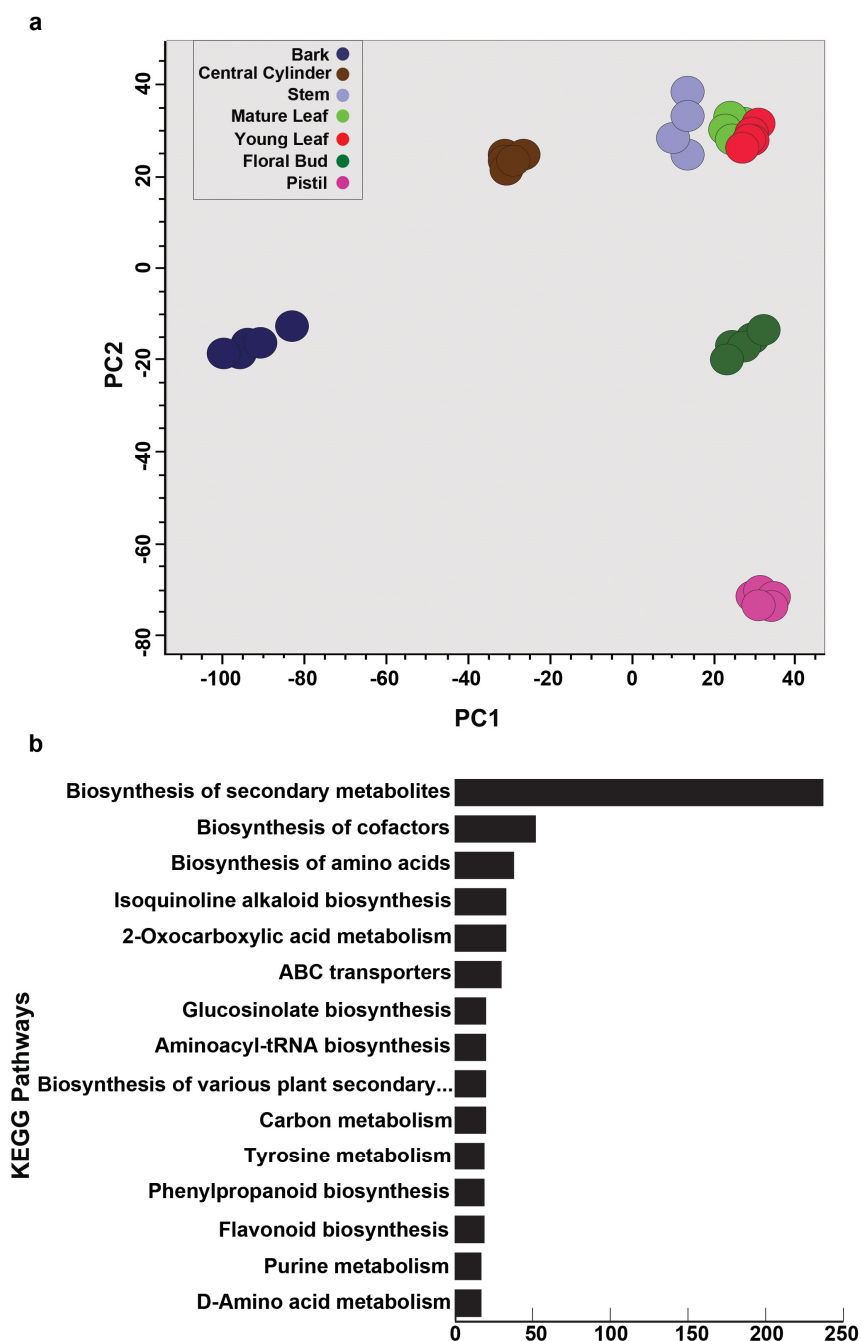
Several studies have reported a plethora of metabolites being present in *Magnolia* plants [3,4,8]. However, most of these studies have focused on investigating the components present in the bark of *Magnolia* plants and, to some extent, in the fruits, leaves, and flowers [21–29]. To establish an accurate metabolome representation for *M. obovata* covering multiple tissues, we performed untargeted metabolome profiling using seven tissues, namely young leaf, mature leaf, stem, bark, central cylinder, floral bud, and pistil (Figure 1). The acquired raw metabolome datasets were pre-processed, aligned, and metabolite features were extracted using MS-DIAL [30], resulting in 6733 mass features in the positive ionization mode (Table S1).



**Figure 1.** Tissues used for the multi-omics study on the biosynthesis of specialized metabolites in *Magnolia obovata*. (a) Plant growing under natural conditions in the herb garden, Chiba University; (b) Young leaf; (c) Mature leaf; (d) Stem; (e) Bark; (f) Central Cylinder; (g) Floral Bud; and (h) Pistil; the bars represent 1 cm.

Principal component analysis (PCA) using the normalized intensities of mass features showed metabolome datasets of seven tissues of *M. obovata* into five distinct groups (Figure 2a). Bark, the medicinally important tissue, formed a distinct group and was separated along the PC1 axis from the rest of the tissues. The young leaf, mature leaf, and stem were grouped together and separated from the floral bud and pistil along the PC2 axis, which formed a distinct tissue-group by itself. The PCA showed tissue specific metabolome signature being captured by the untargeted metabolome profiling of seven tissues of *M. obovata*, which is an ideal feature for a high-quality metabolome resource effectively capturing the tissue-specific chemo-diversity. Furthermore, we mapped metabolite features against the KEGG [31] and KNApSack [11] databases to assign chemical identities to the acquired mass features. In total, 1705 and 553 mass features were assigned with chemical identities to a corresponding KNApSack and KEGG ID, respectively (Table S2). Several of the secondary metabolite biosynthesis pathways, including “Isoquinoline alkaloid biosynthesis”, “Biosynthesis of various plant secondary metabolites”, “Phenylpropanoid biosynthesis”, and “Flavonoid biosynthesis”, were among the top 15 KEGG pathways being represented

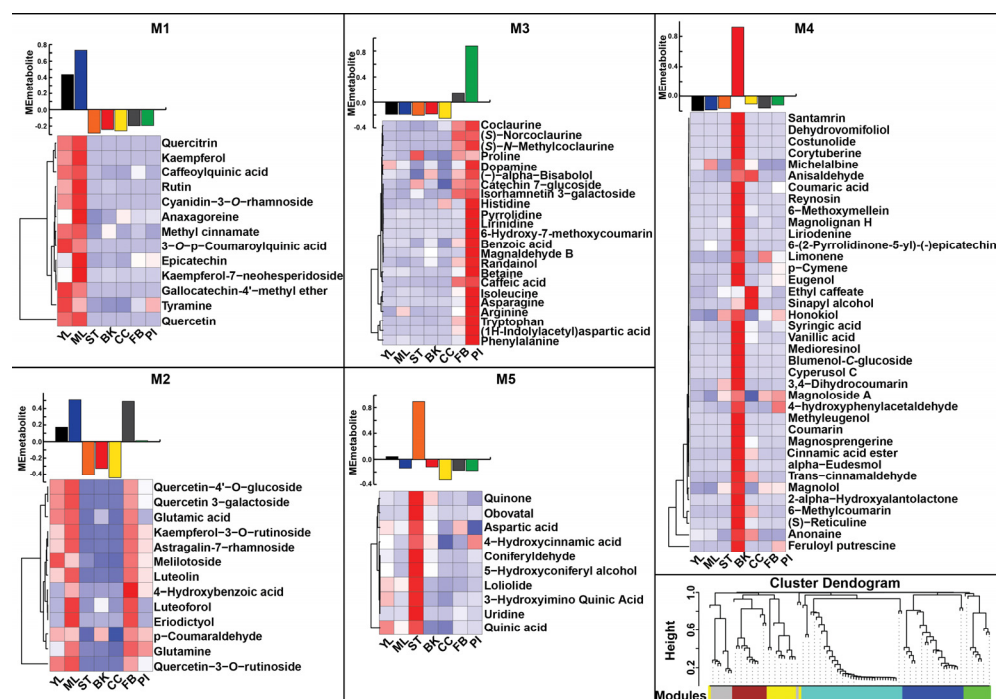
based on the number of metabolites being mapped (Figure 2b). These results suggest that *Magnolia* tissues have a much bigger diversity of metabolites compared to what is known to date as revealed by the untargeted metabolomics in this study. Therefore, our study opens new avenues for the structural validation and identification of candidate metabolites as well as exploring tissue-wise metabolite accumulation for specialized metabolites to hypothesize phytochemical's role in plant physiology and further providing a basis for investigations into the use of tissues other than the bark of *M. obovata* as a potent drug source.



**Figure 2.** Untargeted metabolome analysis for seven tissues of *Magnolia obovata*. (a) Unsupervised Principal Component Analysis (PCA) using the metabolome datasets acquired for the seven tissues of *M. obovata*. (b) Top 15 KEGG pathways based on the number of assigned metabolites. The mass-features were mapped to the KEGG database, and the chemical identities were assigned based on the  $m/z$  similarity with a mass-error window of  $\pm 10$  ppm. Furthermore, the KEGG pathways related to the assigned metabolites were extracted, the top 10 of which are shown here.

## 2.2. Distinctive Metabolic Signatures of *Magnolia obovata* Tissues

As a proof of concept to use the established metabolome resource of *M. obovata* for metabolite discovery and analysis across different tissues, we used a combination of computational chemo-informatics approach, primarily driven through MS-FINDER [32], together with manual inspection and validation using MS and MS/MS-based fragmentation data to confirm the identity of 105 metabolites across seven tissues of *M. obovata* (Table S3). The accumulation of these 105 metabolites showed a tissue-specific trend. To explore hidden signature of metabolites accumulation trend, we adopted a Weighted correlation network analysis (WCNA)-based approach using accumulation levels of 105 metabolites. WCNA is a weighted correlation-based method to define a network from high-dimensional -omics data by grouping the highly correlated variables into a signature module [33]. WCNA classified 97 of the 105 metabolites into five distinct modules (Figure 3), while eight remained unassigned. We identified 13, 13, 23, 38, and 10 metabolites being assigned to the metabolite modules, namely M1, M2, M3, M4, and M5, respectively (Table S4). Module M1 was highly accumulated in the young leaf and mature leaf of *M. obovata*, while metabolite module M2 was accumulated majorly in the mature leaf. The metabolite modules M1 and M2 were dominated by the presence of flavonoids and flavonoid glycosides, including quercetin, rutin, quercitrin, kaempferol, quercetin-4'-O-glucoside, quercetin-3-O-rutinoside. Our results are concurrent with previous studies reporting the accumulation of flavonoid-type metabolites leaves of *Magnolia* species [9]. The metabolite module M3 included several amino acids, including tryptophan, arginine, asparagine, proline, among others and was highly accumulated in the pistil tissues of *M. obovata*. The pistil, female reproductive organ, plays a crucial role in fertilization, and the accumulation of amino acids in the pistil may signify its role in facilitating successful fertilization and subsequent seed development in plants. Furthermore, magnaldehyde B, a known lignan of *Magnolia* species isolated mainly from bark tissue [9], was also highly accumulated in the pistil. Magnaldehyde B and its analogs have been reported to possess anti-tumor activity [9]. The metabolite module M4, representing the largest module with 38 metabolites, was highly accumulated in the bark tissue of *M. obovata*. Magnolol and honokiol, the principal bioactive compound of *Magnolia* species, were members of the M4 module, thus, corroborating with the tissue primarily used for medicinal purposes. The metabolite module M5, representing the smallest module, included several intermediates from the lignin biosynthetic pathway, including coniferyldehyde, 5-hydroxyconiferyl alcohol, and was highly accumulated in the stem of *M. obovata*. These lignin intermediates play crucial roles in the development and structure of plant stems. Metabolite module M5 also included loliolide, a monoterpene lactone used for treatment of central nervous system diseases. Our results clearly suggest that the WCNA of the MS<sup>2</sup>-validated metabolites successfully captured the unique metabolic profiles associated with each tissue type in *M. obovata*. Our metabolome analysis revealed that magnolol and honokiol, two principal bioactive compounds found in *Magnolia* species, exhibited highest accumulation in the bark of *M. obovata*. This underscores the significance of the bark of *M. obovata* for medicinal purposes. Nonetheless, the presence of significant flavonoids such as rutin and quercitrin, known for their antioxidant and anti-tumor properties, along with lignans like magnaldehyde B and obovatal, and monoterpene lactones such as loliolide in tissues apart from the bark, implies the potential role of different tissues for medicinal purposes.



**Figure 3.** WCNA and accumulation pattern of MS<sup>2</sup> confirmed metabolites in *Magnolia obovata*. The bar plots show the module eigenmetabolite of MS<sup>2</sup>-confirmed metabolites. The relative accumulation levels of metabolites included in each module are also shown by the heat maps above each bar plot. The dendrogram represents MS<sup>2</sup>-confirmed metabolites used for WCNA and clustering of module eigenmetabolites. Abbreviations: YL: Young leaf, ML: Mature Leaf, ST: Stem, BK: Bark, CC: Central cylinder, FB: Floral Bud, and PI: Pistil.

### 2.3. Genomic Features Characteristics Through De Novo Transcriptome Assembly of *Magnolia obovata*

The availability of comprehensive genomic datasets is fundamental for precise species selection, trait delineation, and the formulation of biotechnology-driven approaches to enhance desired characteristics. This includes the strategic identification of specific species variants exhibiting elevated levels of bioactive metabolites, considering factors such as genetic diversity and environmental influences. While high-quality genome assemblies of non-model medicinal plants are still scarce, the applications of high-quality RNA-seq-based de novo transcriptome assembly in identifying the potential biosynthetic genes have been well established [34–38]. Therefore, to establish a comprehensive transcriptome resource, capturing the diversity of active transcripts across multiple tissues, we performed RNA-sequencing for the same set of seven tissues of *M. obovata* as was used for establishing metabolome resource. The RNA samples from the seven tissues of *M. obovata* were used to synthesize individual cDNA libraries, and they were subsequently sequenced using Illumina HiSeq™ 2000 platform, generating over 15.76 Gbp of total raw sequence data.

The raw data were preprocessed using Trimmomatic [39] to exclude adapters and low-quality reads in the de novo assembly process, which dropped less than 1% of the sequencing reads. It is widely known that different de novo transcriptome assembly algorithms, even though are based on similar working principles, result in different assembled transcripts with varying lengths, transcriptome completeness and varying overlaps of transcriptome coverage. Merging individual RNA-seq assemblies from different assemblers leverages the strengths of each tool, resulting in a more comprehensive and accurate transcriptome representation. Previous studies have shown an advantage of using a hybrid de novo transcriptome assembly approach to capture diversity of transcripts with near to complete full-length transcripts for further functional characterizations [40,41]. Therefore, in or-

der to achieve a high-quality transcriptome assembly for *M. obovata*, we adopted a hybrid assembly approach using the individual transcriptome assembly derived from three popular assemblers, namely CLC genomics workbench (<https://www.qiagenbioinformatics.com/>, accessed on 8 May 2022), Trinity v3.0 [42], and SOAPdenovo-Trans [43]. The resulting three de novo transcriptome assemblies were merged, followed by removing the redundancy using CD-HIT-EST [44] (Table 1). The hybrid assembly for *M. obovata*, thus obtained, included 284,165 transcripts with an average length and N50 value of 762 bp and 1280 bp, respectively. The length distribution of the assembled transcripts varied between 196 and 16,986 bp (Figure S1a). To evaluate the completeness of our hybrid RNA-seq assembly, we employed BUSCO (Benchmarking Universal Single-Copy Orthologs), which assesses assembly completeness by mapping transcript sequences against a set of conserved single-copy orthologs [45]. The results demonstrated a high level of completeness in our hybrid RNA-seq assembly, with 98.82% of the BUSCO groups having complete gene representation (single-copy or duplicated), while 1.18% are fragmented, and 0% are missing (Table S5). Consistent with previous studies, our results showed improved assembly stats for the hybrid de novo transcriptome assembly approach compared to the individual de novo transcriptome assemblies [14,18,19].

**Table 1.** Assembly statistics for *Magnolia obovata* de novo transcriptome assembly based on three popular assemblers and their combination.

Assembler	Kmer	No. of Contigs	N50	Average Length	Median Length	Max Length
CLC	20	126,234	1126	719	408	15,296
Trinity	25	404,394	1230	734	403	15,776
	31	147,639	1239	720	384	15,241
	41	148,423	1257	715	373	15,334
SOAPdenovo	51	146,829	1212	687	355	15,635
	63	135,683	1171	661	334	15,335
	71	109,186	1214	695	353	15,321
	91	28,652	1415	936	650	12,222
CLC_Trinity_SOAPdenovo (kmer31) CD-HIT-EST	N.A.	284,165	1280	762	424	17,676

To characterize the de novo transcriptome assembly of *M. obovata*, we performed a Blastx-based homology search of the assembled transcripts against protein sequences deposited at the NCBI non-redundant (nr) database. Blastx-based searching resulted in the annotation of 105,980 transcripts, with 82,367 transcripts and 17,275 transcripts being associated with Gene Ontology (GO) annotation and GO mapping, respectively, (Figure S1b, Table S6). Notably, despite being phylogenetically distant, our analysis revealed significant sequence similarity with several plant species, with *Vitis vinifera*, *Populus trichocarpa*, and *Ricinus communis* emerging as the top three species based on the number of transcripts with high-sequence similarity with the assembled transcripts of *M. obovata* (Figure S1c). This can be attributed to the conservation of core plant genes essential for fundamental biological processes, as well as genes involved in shared biosynthetic pathways. Evolutionary homology arising from a common ancestor may also play a role, with ancient gene families retaining high sequence similarity across angiosperms. This observation also highlights the limited representation of sequences from the genus *Magnolia* in existing NCBI databases, underscoring the significance of our study as a valuable resource of transcripts for future research and annotation efforts. The transcripts were assigned with annotations using the top hit from the Blastx search, providing valuable insights into the genetic underpinnings of its active biological processes including assignment of transcripts to different metabolic

pathways, shedding light on potential biochemical pathways active within *M. obovata*. A total of 26,400 transcripts were assigned to 146 KEGG pathways, with the top 15 pathways based on the number of transcripts being assigned summarized in Figure S1d. Among the top 15 pathways, starch and sucrose metabolism, glycolysis/gluconeogenesis, and drug metabolism-other enzymes were among the top three pathways, with 1158, 977, and 875 transcripts being mapped to these metabolic pathways, respectively. *M. obovata* is widely reported to contain phenylpropanoids and flavonoids as their major secondary metabolites [46,47]. KEGG pathway mapping assigned 478 and 306 transcripts to phenylpropanoid biosynthesis and flavonoid biosynthesis, respectively, with homologs of major known enzymes identified in our transcriptome datasets. Collectively, these findings establish a comprehensive framework for elucidating the biological significance of *M. obovata* transcripts and pave the way for future investigations into its molecular intricacies.

#### 2.4. Transcripts Expression Analysis Reveals the Tissue-Specific Molecular Signatures in *Magnolia obovata*

To unravel the complex genetic makeup for a plant species, it is vital to understand how the different genes expression and associations to different tissues orchestrates the plant's growth, development, and responses to different biotic–abiotic challenges. Transcriptome expression analysis provides a comprehensive overview of the active biological processes, where the spatial expression and relative expression abundance of a given gene provides valuable hints for its potential functions. To calculate the qualitative expression level of each transcript across different tissues, the processed paired-end RNA-seq reads were mapped to the de novo transcriptome assembly of *M. obovata*, and normalized transcript expression in FPKM (Fragments Per Kilobase of transcript per Million mapped reads) was determined. Among the seven tissues of *M. obovata*, stem and bark exhibited the highest number of transcriptionally active transcripts, with 186,216 and 185,959 transcripts, respectively, displaying expression levels above FPKM > 0. On the other hand, floral bud and pistil displayed the lowest number of transcriptionally active transcripts, with 127,323 and 135,322 transcripts, respectively, (Table S6). The differential distribution of transcriptionally active genes across tissues underscores the functional diversity and the requirement of specific sets of transcripts activity to achieve the specific functionality. A total of 69,697 transcripts were shared among all seven tissues, suggesting their potential role in the essential biological functions such as primary metabolic and core metabolic processes. Six of the seven tissues of *M. obovata* showed expression specific to the tissue for transcripts in the range of 2500–5250. Bark, on the other hand, showed specific expression for 24,242 transcripts (Table S7), which co-incidentally is also the tissue that produces a range of specialized metabolites with bioactivity. The huge parity among the unique active transcripts between bark and the rest of the tissues suggest that bark includes a unique collection of transcripts for fine-tuning *M. obovata*'s response to environmental cues and to produces a diverse array of secondary metabolites.

To uncover the underlying patterns and relationships within transcriptome datasets across seven tissues of *M. obovata*, we conducted unsupervised principal component analysis (PCA). The PCA results exhibited the presence of four distinct groups (Figure S2a), with leaf tissues, including both young and mature leaves, clustered in quadrant III and segregated from other tissues along the PC1 axis. Along the PC2 axis, five tissues were partitioned into two quadrants, with bark and pistil grouping in quadrant I, exhibiting clear separation along the PC2 axis. Conversely, the floral bud, stem, and central cylinder were clustered in quadrant IV. Consistent with the PCA-based clustering, expression-based correlation analysis also identified all seven tissues being clustered into four clusters (Figure S2b), wherein young leaf and mature leaf were clustered together, and bark and pistil tissues formed individual clusters. The floral bud, stem, and central cylinder were also



clustered together, with the stem and central cylinder sharing a closer association. These analyses underscore the presence of tissue-specific transcriptomic signatures in *M. obovata*, highlighting their specialized functions. Further, the overlap of transcript expression across tissues seems to stem from their association at different stages of development.

#### 2.5. Construction of Gene-Metabolite Co-Expression Network Uncovers Tissue-Specific Transcriptome Signature and Specialized Metabolites Biosynthesis Pathways in *Magnolia obovata*

The ability of weighted gene co-expression network analysis (WGCNA) to transform a comprehensive transcriptome expression dataset into a representation of interconnectedness, thus revealing the intricate relationships within, has been widely acknowledged, facilitating the exploration of interconnected biomolecules [18,48–50]. Rooted in the guilt-by-association concept, WGCNA operated on the premise that genes with correlated expression are often functionally related, participating in common pathways, regulatory mechanisms, or responses to environmental stimuli [33]. WGCNA helps in understanding the gene interactions and underlying relationships within a biological system by grouping these co-expressed genes into modules, with genes within the same module tending to have similar biological functions. Leveraging this approach, it becomes feasible to predict novel genes involved in specialized metabolite biosynthesis pathways by exploring modules containing well-characterized ‘bait genes’ of interest. This approach has been proven effective in identifying functional genes associated with the biosynthesis of plants’ specialized metabolites in several model and medicinal plants [18,51–54]. Using a similar approach for *M. obovata*, we performed WGCNA using the highly expressed (FPKM > 5 in at least one of the seven tissues) and annotated transcripts and identified highly co-expressed transcript modules (Table S8). WGCNA identified 12 transcript modules, of which T12 (5429 member transcripts) and T6(70 member transcripts) represented the largest and the smallest modules, respectively, (Table S9). Notably, transcript expression patterns mirrored the observed accumulation profiles of metabolites, displaying tissue-specific trends. The transcript modules, namely T1, T2, T3, T8, T9, T11, and T12 showed tissue-specific expression trends for young leaf, pistil, floral bud, central cylinder, stem, mature leaf, and bark, respectively (Figure 4).

The integration of multi-omics datasets enables a broader exploration beyond gene expression patterns and metabolite accumulation by establishing gene-to-metabolite associations to reveal complex relationships among biomolecules. The integration allows for the examination of how metabolite accumulation patterns correlate with gene expression patterns within the same pathway. By combining gene co-expression data with metabolic profiling data, it becomes possible to identify genes whose expression patterns correlate with metabolite accumulation patterns in specific pathways, thereby suggesting their involvement in orchestrating observed metabolic shifts [12,13]. This integrated omics approach has facilitated the achievement of a refined and comprehensive prioritization of genes involved in a specific biosynthetic pathway of interest [12,18,19,28]. To implement the integrated omics strategy in *M. obovata*, we performed a Pearson-based correlation analysis using the average accumulation or expression value of the metabolites and transcripts assigned to the identified metabolome, and transcriptome modules, respectively. A notably high correlation ( $R^2 > 0.9$ ) was observed for four module pairs, M1–T11, M3–T2, M4–T12, and M5–T9 (Figure 5a, Table S10). Among the metabolite–gene module pairs, M4–T12 was specific to the bark tissues and included magnolol and honokiol, the major constituents that give the characteristic medicinal properties to the *M. obovata*. Therefore, we further explored the transcript module T12. The top 15 KEGG pathways based on the number of transcripts from T12 module being mapped to the pathway database are summarized in Figure 5b. While the top KEGG pathways were mainly associated with primary metabolic pathways, transcripts within the T12 module also represented homologs of all genes involved in the

biosynthesis of phenylpropanoids and flavonoids in *M. obovata* (Figures S3 and S4). Gene ontology enrichment analysis using annotation of transcripts from the T12 module showed enrichment of several metabolic processes associated with stress responses, including some response type processes, namely response to water deprivation and defense response to bacterium (Figure 5c).

## 2.6. Merging Omics Landscapes to Decode Neolignans Biosynthesis in *Magnolia obovata*

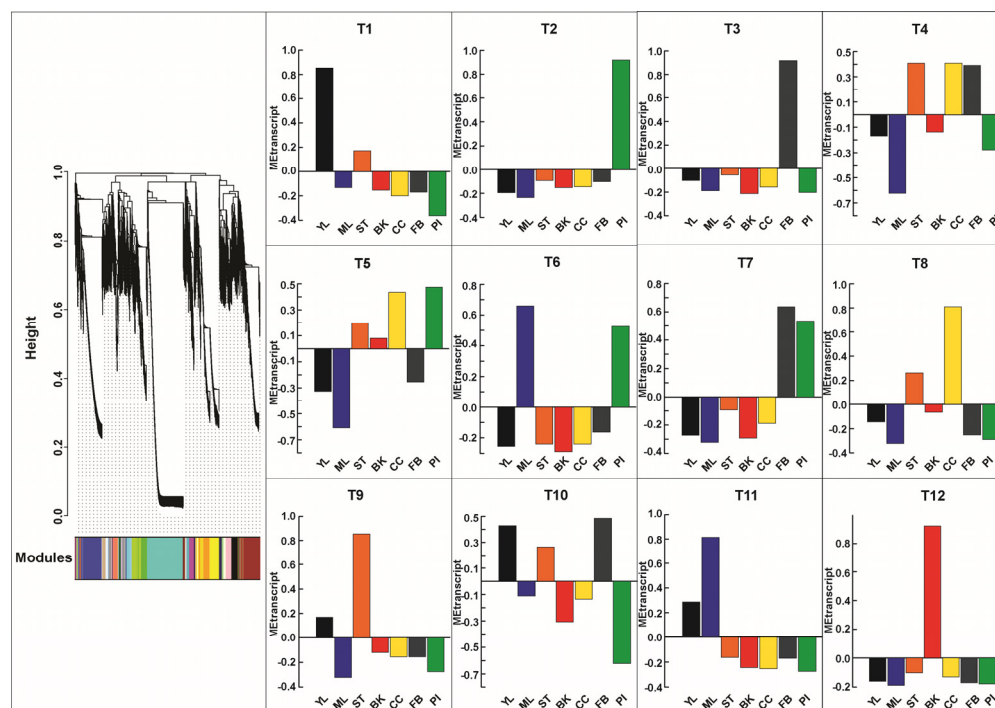
Neolignans, namely honokiol and magnolol, are the key bioactive constituents attributing to the characteristic medicinal properties of *M. obovata* [2]. Monolignols from the cinnamate/monolignol biosynthetic pathway are reported to be the intermediates in the biosynthesis of neolignans [55]. The biosynthesis of lignans is well characterized and is known to be initiated by oxidation of coniferyl alcohol to form corresponding radical through the enzymatic activity of oxidases such as laccase followed by the stereoselective dimerization catalyzed by dirigent proteins [56–58]. In contrast, our understanding of the biosynthesis pathway of neolignans is currently limited. However, several crude enzyme preparation-based studies have highlighted that similar to the lignans biosynthesis, polymerization of phenylpropanoid monomers also results in the formation of neolignans [55,59–62]. However, the oxidases and dirigent proteins involved in the biosynthesis of neolignans are still largely unknown.

Laccases (LACs), a class of multi-copper oxidases, are widely distributed across different kingdoms, including bacteria, fungi, plants, and animals [63,64]. Characterized by broad substrate specificity, laccases catalyze independently of any cofactors. Laccases mainly oxidize aromatic phenolics but also participate in the oxidizing reaction of diamines and benzenethiols through the reduction in molecular oxygen into water [63]. The radicals generated by laccase enzyme require the action of dirigent proteins (DIRs) that orient the two radicals correctly for them to undergo a regio- and stereo-selective coupling reaction [63]. These mechanisms of action have been majorly reported for the biosynthesis of lignans, ellagitannins, and stilbenoid dimers [63,65–67]. A recent study on *Arabidopsis* highlighted the role of laccase and dirigent proteins in the biosynthesis of seed-protective neolignans [68]. Considering the diverse substrate specificity observed in laccases, and reports of its novel functionalities observed in the biosynthesis of lignans, neolignans, and related compounds [63], we hypothesize that neolignans in *M. obovata* originate from the precursor coniferyl alcohol through a two-step process (Figure 6a). As the first step, laccases catalyze the generation of monolignol radicals. These radicals, stabilized through mesomeric forms, render multiple positions on the carbon skeleton reactive. Subsequently, in the second step facilitated by dirigent proteins, these radicals undergo dimerization, forming new C-C or C-O bonds.

The transcript module T12, which shared a high association with the metabolite module M4, which included the neolignans magnolol and honokiol, included homologs of all the enzymes involved in the biosynthesis of coniferyl alcohol. Therefore, LACs and DIRs included in the T12 module and sharing high co-expression with the other homologs of phenylpropanoid biosynthetic pathway are strong candidates with potential role in catalyzing the reaction to produce neolignans in *M. obovata*. Using homology-based annotation and functional characterization, we identified three LACs and seven DIRs, respectively, in the transcript module T12. To understand the type of reaction catalyzed by these LACs and DIRs in *M. obovata*, we performed a phylogenetic analysis of them with the functionally characterized laccases and dirigent proteins from other plant species.

Phylogenetic analysis of three of the laccases included in the T12 transcript module of *M. obovata* was performed with the 29 of the known laccases across different plant species, including *Arabidopsis thaliana*, *Populus trichocarpa*, *Zea mays*, among others

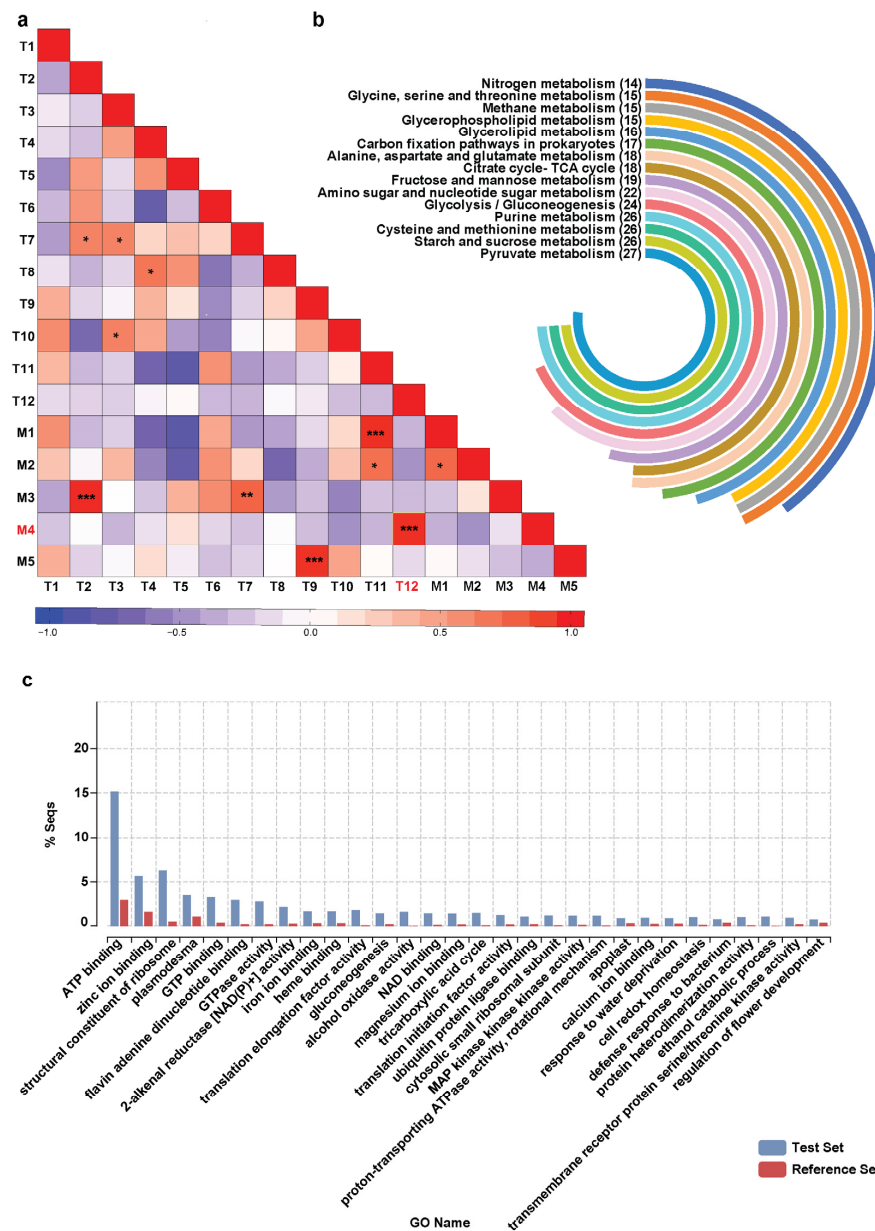
(Table S11). While one of the laccases of *M. obovata* formed a separate clade, two of the laccases clustered together with the laccase 5, laccase 12, laccase 3, and laccase 13 of *A. thaliana* and laccase 90 of *P. trichocarpa* (Figure 6b). While most of these Arabidopsis laccase remains uncharacterized, a recent study showed *AtLAC5* to oxidize coniferyl alcohol [68]. *PtLAC90* has also been shown to oxidize a wide range of substrates, including coniferyl alcohol [69]. Moreover, both laccase encoding transcripts in *M. obovata* contain the known functional multicopper oxidase domain. Therefore, inferring from our phylogenetic analysis, we propose these two laccases, namely *Mo\_LAC1* and *Mo\_LAC2*, as strong candidate genes potentially involved in the oxidizing of coniferyl alcohol, the first reaction towards the biosynthesis of neolignans in *M. obovata*.



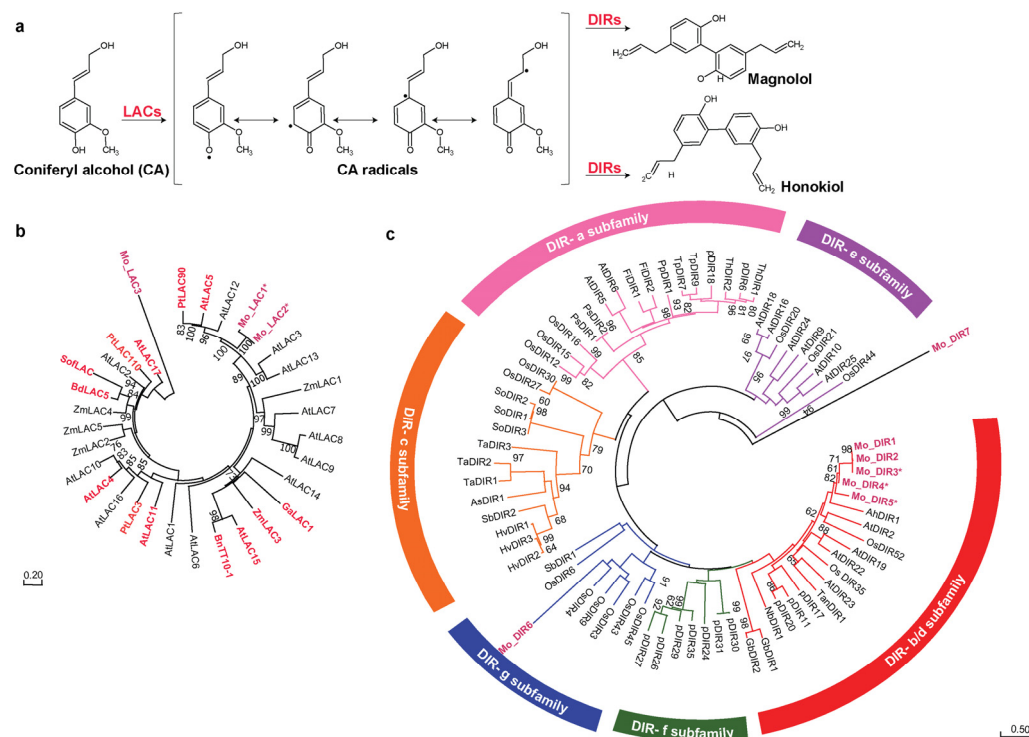
**Figure 4.** WCNA of the assembled transcripts in *Magnolia obovata*. The bar plots show the module eigentranscript of assembled transcripts. The relative accumulation levels of metabolites included in each module are also shown by the heat maps above each bar plot. The dendrogram represents expression level of transcripts used for WCNA and clustering of module eigentranscripts. Abbreviations: YL: Young leaf, ML: Mature Leaf, ST: Stem, BK: Bark, CC: Central cylinder, FB: Floral Bud, and PI: Pistil.

To identify the potential DIRs involved in the second step of neolignans biosynthesis, we performed a phylogenetic analysis of seven of the DIRs of *M. obovata* with the known DIRs spanning across various plant species (Table S11). Our comprehensive phylogenetic analysis provides invaluable insights into the relationships between various subfamilies of dirigent proteins with dirigent proteins of *M. obovata* from T12 module, therefore, shedding light on the complexity of neolignans biosynthesis. While one of the transcripts formed a separate clade (Figure 6c), notably, the six transcripts span two subfamilies, five to DIR-b/d subfamily and one with DIR-g subfamily, signifying multifaceted repertoire of dirigent proteins at play in *M. obovata*'s biochemical landscape. While less characterized, the DIR-g is a species-specific subfamily, and putatively participates in the biosynthesis of lignins [70,71], and the *M. obovata* transcript clustered together within the DIR-g subfamily may suggest its role in the biosynthesis of lignins. Most transcripts included in the T12 transcript module of *M. obovata* clustered within the DIR-b/d subfamily, making this subfamily a focal point of our investigation. The majority of the genes from the DIR-a subfamily of dirigent proteins

is primarily associated with synthesizing pinoresinol lignans [71,72]. On the other hand, DIR-b/d proteins are more versatile and renowned for their remarkable substrate specificity, guiding the synthesis of a broader range of compounds [73]. It has also been reported that the DIR-b/d subfamily has expanded in many plant families, suggesting their role in the plant adaptive evolution under different biotic and abiotic stresses [71,73–75]. Therefore, five transcripts being grouped with the DIR-b/d subfamily may suggest their potential role in the biosynthesis of neolignans in *M. obovata*. Moreover, three of the five transcripts from the DIR-b/d subfamily had functional dirigent-like protein domain, thus, prioritizing their role in the biosynthesis of neolignans in *M. obovata*. Results from the phylogenetic analysis open doors to further exploration. Future research to delve deeper into the functional characterization of transcripts affiliated with different DIR subfamilies may help uncover how they influence neolignans biosynthesis in *M. obovata*.



**Figure 5.** Module based integration of the metabolome and transcriptome datasets. (a) Pearson correlation coefficient between each MetM and TransM module pair. The red font highlights the transcript-metabolite module pair of interest. \*\*\* pairs with PCC > 0.9, \*\* pairs with PCC > 0.8, \* PCC > 0.7. (b) The top 15 KEGG pathways based represented by transcripts included in T2 module based on the number of transcripts being mapped. (c) Gene Ontology (GO) enrichment analysis of the transcripts included in the T12 module.



**Figure 6.** Proposed biosynthesis pathway for neolignans, honokiol and magnolol, in *Magnolia obovata*. (a) Proposed biosynthetic pathway. Abbreviations: LACs (laccases), and DIRs (dirigent proteins); Phylogenetic analysis of transcripts included in T12 module and annotated as LACs (b); and DIRs (c). The nucleotide sequences for candidate LACs and DIRs were translated and the corresponding protein sequences were aligned with LACs and DIRs spanning across different plant species using MUSCLE program. \* DIRs annotated sequences with functional dirigent-like protein domain. The evolutionary history was inferred by using the Neighbor-Joining method with bootstrap values obtained after applying 1000 replications using the MEGA X program. Bootstrap values above 60% are shown here. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The accession number of all the LACs and DIRs used for phylogenetic analysis are included as Table S11.

### 3. Materials and Methods

#### 3.1. Plant Materials

The *Magnolia obovata* plant was maintained at the medicinal plant gardens, Chiba University, under natural conditions (Figure 1). The seven tissues of *M. obovata*, including young leaf, mature leaf, stem, bark, central cylinder, floral bud, and pistil, were harvested and immediately frozen in liquid nitrogen. The tissues were stored at  $-80^{\circ}\text{C}$  until further use. For RNA extraction, tissues were homogenized to fine powder using mortar and pestle. For metabolites extraction, tissues were homogenized to fine powder using a mixer mill, MM300 (Retsch, Haan, Germany). Here, we adopted the split experimental design, where same tissues were split into two parts, each used for generating transcriptome and metabolite datasets.

#### 3.2. Untargeted Metabolite Analysis for Seven Tissues of *Magnolia obovata*

Each of the homogenized tissues of *M. obovata* were freeze-dried and metabolites were extracted using 50  $\mu\text{L}$  of 80% (*v/v*) LC-MS-grade methanol (Wako chemicals, Osaka, Japan) and 20% (*v/v*) LC-MS-grade water (Wako chemicals, Osaka, Japan) containing 2.5  $\mu\text{M}$  of 10-camphoursulfonic acid (TCI, Tokyo, Japan), and 2.5  $\mu\text{M}$  of lidocaine (TCI, Tokyo, Japan) per milligram of dry weight. After centrifugation using an Eppendorf 5417R centrifuge (Eppendorf SE, Hamburg, Germany) at  $17,800 \times g$  for 10 min, the supernatant was

subsequently processed using Oasis HLB  $\mu$ Elution plate (Waters Inc., Milford, MA, USA) for removal of lipids and impurities, and the MS and MS<sup>2</sup> datasets were acquired in the positive and negative ionization mode as described previously [18]. Peaks were aligned using MS-DIAL v4.80 with its default parameters. Peak normalization was conducted by comparison with an internal standard (lidocaine). Metabolites were annotated using MS-FINDER v3.52 and using  $m/z$  values of diterpene alkaloids being reported previously with tolerance 10 ppm. Putative chemical identities were assigned to the acquired metabolite peaks by mapping onto the KNApSAcK database and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database. The MS<sup>2</sup>-based validation of metabolites was conducted as reported previously [18].

### 3.3. RNA Extraction, and cDNA Library Preparation

Frozen tissues of *M. obovata*, the same tissues used for untargeted metabolome analysis, were used for RNA extraction and cDNA library preparation. RNA extraction, RNA integrity analysis, mRNA sample preparation, fragmentation of isolated mRNA, and cDNA library preparation for Illumina sequencing were performed as described previously [16].

### 3.4. Illumina Sequencing

The cDNA library thus prepared for each tissue of *M. obovata* was sequenced using an Illumina HiSeq™ 2000 sequencer (Illumina, Inc., San Diego, CA, USA), and paired-end reads were obtained with an average length of 101 bps. Preparation and shearing of mRNA, cDNA library preparation, and sequencing were performed at Kazusa DNA Research Institute, Chiba, Japan. The raw sequence reads for all seven tissues of *M. obovata*, their expression value and the de novo transcriptome assembly used in this study have been deposited in NCBI's Gene Expression Omnibus (GEO) and are available at GEO series accession number GSE275868.

### 3.5. RNA-seq Raw Reads Pre-Processing, De Novo Transcriptome Assembly, and Functional Classification of Transcripts

Raw sequencing reads thus generated were preprocessed through the Trimmomatic program v0.39 [39] to remove adaptor sequences, short reads, reads with ambiguous 'N' base > 5%, and low-quality reads (Phred score < 30). Paired-end processed reads, as well as unpaired high-quality reads that lost their corresponding sequence partner due to Trimmomatic filtering, for all seven tissue types were combined to build a de novo transcriptome assembly of *M. obovata*. The de novo transcriptome assembly resulting from the Trinity program v3.0 [42] was further processed through the CD-HIT-EST program v4.8.1 [44] for sequence redundancy removal, and it was and subsequently used for annotation and characterization. The transcripts obtained were, thus, used for the read alignment and abundance estimation in individual tissues of *M. obovata* using Bowtie 2.0 v2.3.5.1 [76] and RSEM v1.3 [77], respectively. Transcript expression was calculated in terms of Fragments per Kilobase exon per million mapped fragments (FPKM). Correlation analysis was performed for all seven tissues of *M. obovata* through the DeSeq2 program v3.17 [78] of the R-package. The completeness and quality of the *M. obovata* de novo transcriptome assembly were assessed using Benchmarking Universal Single-Copy Orthologs (BUSCO) [45] analysis against the viridiplantae\_odb10 database, conducted through OmicsBox v3.1.2 (<https://www.biobam.com/omicsbox>; access date 25 January 2025).

The annotation of de novo transcriptome assembly of *M. obovata* was performed using a Blastx-based homology search against the NCBI-nr database with an E-value cutoff < 10<sup>-5</sup>, and the top Blastx hit was used to assign sequence description and putative functionality to the transcripts. Furthermore, OmicsBox v3.1.2 was used to obtain GO terms, EC number, and KEGG pathway-based annotation for the transcripts of *M. obovata*.

For the identification of transcripts involved in the biosynthesis of diterpene alkaloid, the annotated transcriptome assembly was screened for the homologs of the associated enzymes, and candidate transcripts were selected that had a length greater than 500 bps and sequence similarity with top Blastx hit of over 70%.

### 3.6. Module Construction Using Omics Dataset of *Magnolia obovata*

For metabolome data, 105 MS<sup>2</sup>-validated metabolites were used for the module construction. For the identification of modules in the transcriptome datasets, we first filtered the assembled transcripts using the following parameters: (i) length > 500 bps and similarity > 70%, and (ii) FPKM > 5 in at least one of the seven tissues, resulting in narrowing down of 25,728 transcripts, which were subsequently used for WGCNA in R. Hierarchical clustering based on the topological overlap (TO) similarity was used to obtain network modules, and the highly similar modules (dissimilarity < 0.25) were merged to prevent over-splitting of the modules. Module eigenvalues, representing the overall profile of the module, were calculated for each of the identified modules.

### 3.7. Integrative Omics Analysis Through Correlation of Metabolite and Transcript Modules of *Magnolia obovata*

The average metabolite levels and the FPKM values of transcripts clustered in the individual metabolite and transcript modules, respectively, were used for correlation analysis. A Pearson correlation analysis was performed to calculate correlation coefficient between each metabolite and transcript module and generate a pairwise Pearson correlation matrix. The correlation relationships between metabolites and transcripts were visualized using Hatmap2.0 in the R package.

### 3.8. Phylogenetic Analysis of Candidate Laccase and Dirigent Proteins in *Magnolia obovata*

The transcripts of *M. obovata* annotated as laccase and dirigent proteins, included in transcript module T12 and having a sequence length of over 500 bps as well as having sequence similarity with the top Blastx hit of 70% were selected for phylogenetic analysis. These selected transcripts were translated to their corresponding protein sequences by selecting the translation frame that resulted in the longest amino acid sequence while starting with methionine using OmicsBox v3.1.2. Protein sequences of candidate transcripts and previously known laccase and dirigent proteins involved in the biosynthesis of specialized metabolites in other plant species were aligned using the MUSCLE program v3.8.31 [79], and evolutionary distances were inferred using the Neighbor-Joining method [80] with bootstrap values obtained after applying 1000 replications using the MEGA X program v10.2.6 [81].

## 4. Conclusions

This study significantly enhances our understanding on the biosynthetic pathways of neolignans in *Magnolia obovata*, underscoring the complex interplay between metabolite accumulation and gene expression across various tissues. In this study, we established high-quality metabolome and transcriptome resources for *M. obovata*, which holds a unique position within the diverse *Magnolia* genus for its extensive pharmacological relevance in various traditional medicinal practices. The multi-omics approach focused on metabolomics and transcriptomics integration has provided valuable insights into the genetic and molecular underpinnings of this medicinal plant. By focusing on seven tissues, namely young leaf, mature leaf, stem, bark, central cylinder, floral bud, and pistil, we aimed to gain insight into the biosynthesis of specialized metabolites with therapeutic properties in *M. obovata*. The establishment of a robust omics resource revealed distinct tissue-specific accumulation patterns of metabolites, particularly in bark tissue, highlighting the adaptive

strategies employed by this species. Through comprehensive untargeted metabolome and transcriptome profiling, we identified candidate genes associated with the biosynthesis of key bioactive compounds pivotal to the medicinal properties of *M. obovata*, including honokiol and magnolol. The integration of gene expression data with metabolite profiles facilitated the identification of candidate genes associated with neolignan biosynthesis, particularly enzymes that code for laccases and dirigent proteins. Phylogenetic analyses of these enzymes provided insights into their evolutionary relationships and potential roles in catalyzing specific biosynthetic reactions. The identification of Mo\_LAC1 and Mo\_LAC2 as strong candidates for oxidizing coniferyl alcohol lays a foundation for future functional characterization studies. Overall, this research not only clarifies the molecular mechanisms underlying neolignan biosynthesis in *M. obovata* but also offers valuable resources for the exploration of secondary metabolism in medicinal plants. The findings pave the way for further investigations into the regulatory networks governing these pathways, contributing to the broader understanding of plant biology and the potential development of natural products with therapeutic applications. Our findings also lay the groundwork for future endeavors in harnessing the therapeutic potential of *M. obovata*.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms26031068/s1>.

**Author Contributions:** A.R., K.S. and M.Y.; conceived and designed the experiments. M.N.; prepared the tissues and RNA extraction for transcriptome sequencing. H.S.; acquired the RNA-seq datasets. R.N. and T.M.; performed the metabolome profiling. T.M., R.N., M.R. and A.R.; analyzed and annotated the metabolome datasets. M.R. and T.Y.; performed the transcriptome analysis and provided the biological interpretations for the integrative omics analysis. A.R., M.R., K.S. and M.Y.; contributed to the manuscript preparation. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by Research and Development Grant of Japan Agency for Medical Research and Development (AMED), Grant-in Aid for Transformative Research Area (A) (22H05124), and Grant-in-Aid for Scientific Research-KAKENHI (S), Japan Society for the Promotion of Science (JSPS; grant number-19H05652).

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The Illumina raw sequence reads, the de novo transcriptome assembly, annotations, and expression value for the unigenes have been deposited in NCBI's Gene Expression Omnibus (GEO) and are available at GEO series accession number GSE275868.

**Acknowledgments:** We thank Sayaka Shinpo, from the Kazusa DNA Research Institute, for technical support in Illumina sequencing. The super-computing resources were provided by the National Institute of Genetics (NIG), Research Organization of Information and Systems, Japan. The computing resources were provided by the Medical Mycology Research Center, Chiba University, Japan.

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

## References

1. Kwon, O.J.; Oh, C.H. Naturalization of landscaping woody plant, *Magnolia obovata* potentially invasive species. *J. Mt. Sci.* **2015**, *12*, 30–38. [[CrossRef](#)]
2. Lovecka, P.; Svobodova, A.; Macurkova, A.; Vrchotova, B.; Demnerova, K.; Wimmer, Z. Decorative *Magnolia* Plants: A Comparison of the Content of Their Biologically Active Components Showing Antimicrobial Effects. *Plants* **2020**, *9*, 879. [[CrossRef](#)]
3. Lee, Y.J.; Lee, Y.M.; Lee, C.K.; Jung, J.K.; Han, S.B.; Hong, J.T. Therapeutic applications of compounds in the *Magnolia* family. *Pharmacol. Ther.* **2011**, *130*, 157–176. [[CrossRef](#)]
4. Poivre, M.; Duez, P. Biological activity and toxicity of the Chinese herb *Magnolia officinalis* Rehder & E. Wilson (Houpo) and its constituents. *J. Zhejiang Univ. Sci. B* **2017**, *18*, 194–214. [[CrossRef](#)] [[PubMed](#)]



5. Maruyama, Y.; Kuribara, H. Overview of the Pharmacological Features of Honokiol. *CNS Drug Rev.* **2000**, *6*, 35–44. [[CrossRef](#)]
6. Zhou, L.; Hou, F.; Wang, L.; Zhang, L.; Wang, Y.; Yin, Y.; Pei, J.; Peng, C.; Qin, X.; Gao, J. The genome of *Magnolia hypoleuca* provides a new insight into cold tolerance and the evolutionary position of magnoliids. *Front. Plant Sci.* **2023**, *14*, 1108701. [[CrossRef](#)]
7. Oshima, N.; Kume, H.; Umeda, T.; Takito, H.; Tsukimoto, M.; Hada, N. Structures and Inhibitory Activities for Interleukin-2 Production of Seasonally Variable Constituents in Flower Parts of *Magnolia kobus* at Different Growth Stages. *Chem. Pharm. Bull.* **2020**, *68*, 91–95. [[CrossRef](#)]
8. Sarker, S.D.; Maruyama, Y.J. *Magnolia: The Genus Magnolia*; Taylor & Francis: New York, NY, USA, 2002; p. xi. 187p.
9. Luo, H.; Wu, H.; Yu, X.; Zhang, X.; Lu, Y.; Fan, J.; Tang, L.; Wang, Z. A review of the phytochemistry and pharmacological activities of *Magnoliae officinalis* cortex. *J. Ethnopharmacol.* **2019**, *236*, 412–442. [[CrossRef](#)]
10. Dominiak, K.; Gostynska, A.; Szulc, M.; Stawny, M. The Anticancer Application of Delivery Systems for Honokiol and Magnolol. *Cancers* **2024**, *16*, 2257. [[CrossRef](#)] [[PubMed](#)]
11. Shinbo, Y.; Nakamura, Y.; Altaf-Ul-Amin, M.; Asahi, H.; Kurokawa, K.; Arita, M.; Saito, K.; Ohta, D.; Shibata, D.; Kanaya, S. KNAPSAcK: A Comprehensive Species-Metabolite Relationship Database. In *Plant Metabolomics*; Saito, K., Dixon, R.A., Willmitzer, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 165–181.
12. Rai, A.; Saito, K.; Yamazaki, M. Integrated omics analysis of specialized metabolism in medicinal plants. *Plant J.* **2017**, *90*, 764–787. [[CrossRef](#)]
13. Rai, A.; Yamazaki, M.; Saito, K. A new era in plant functional genomics. *Curr. Opin. Syst. Biol.* **2019**, *15*, 58–67. [[CrossRef](#)]
14. Rai, A.; Kamochi, H.; Suzuki, H.; Nakamura, M.; Takahashi, H.; Hatada, T.; Saito, K.; Yamazaki, M. De novo transcriptome assembly and characterization of nine tissues of *Lonicera japonica* to identify potential candidate genes involved in chlorogenic acid, luteolosides, and secoiridoid biosynthesis pathways. *J. Nat. Med.* **2017**, *71*, 1–15. [[CrossRef](#)]
15. Rai, A.; Nakamura, M.; Takahashi, H.; Suzuki, H.; Saito, K.; Yamazaki, M. High-throughput sequencing and de novo transcriptome assembly of *Swertia japonica* to identify genes involved in the biosynthesis of therapeutic metabolites. *Plant Cell Rep.* **2016**, *35*, 2091–2111. [[CrossRef](#)]
16. Rai, A.; Yamazaki, M.; Takahashi, H.; Nakamura, M.; Kojoma, M.; Suzuki, H.; Saito, K. RNA-seq Transcriptome Analysis of *Panax japonicus*, and Its Comparison with Other *Panax* Species to Identify Potential Genes Involved in the Saponins Biosynthesis. *Front. Plant Sci.* **2016**, *7*, 481. [[CrossRef](#)] [[PubMed](#)]
17. Tsugawa, H.; Rai, A.; Saito, K.; Nakabayashi, R. Metabolomics and complementary techniques to investigate the plant phytochemical cosmos. *Nat. Prod. Rep.* **2021**, *38*, 1729–1759. [[CrossRef](#)]
18. Rai, M.; Rai, A.; Mori, T.; Nakabayashi, R.; Yamamoto, M.; Nakamura, M.; Suzuki, H.; Saito, K.; Yamazaki, M. Gene-Metabolite Network Analysis Revealed Tissue-Specific Accumulation of Therapeutic Metabolites in *Mallotus japonicus*. *Int. J. Mol. Sci.* **2021**, *22*, 8835. [[CrossRef](#)] [[PubMed](#)]
19. Rai, A.; Rai, M.; Kamochi, H.; Mori, T.; Nakabayashi, R.; Nakamura, M.; Suzuki, H.; Saito, K.; Yamazaki, M. Multiomics-based characterization of specialized metabolites biosynthesis in *Cornus officinalis*. *DNA Res.* **2020**, *27*, dsaa009. [[CrossRef](#)]
20. Rai, A.; Nakaya, T.; Shimizu, Y.; Rai, M.; Nakamura, M.; Suzuki, H.; Saito, K.; Yamazaki, M. De Novo Transcriptome Assembly and Characterization of *Lithospermum officinale* to Discover Putative Genes Involved in Specialized Metabolites Biosynthesis. *Planta Med.* **2018**, *84*, 920–934. [[CrossRef](#)]
21. Dharmananda, S. *Safety Issues Affecting Chinese Herbs: The Case of Kava*; ITM: Portland, OR, USA, 2002.
22. Yan, R.; Wang, W.; Guo, J.; Liu, H.; Zhang, J.; Yang, B. Studies on the alkaloids of the bark of *Magnolia officinalis*: Isolation and on-line analysis by HPLC-ESI-MS(n). *Molecules* **2013**, *18*, 7739. [[CrossRef](#)] [[PubMed](#)]
23. Yu, S.X.; Yan, R.Y.; Liang, R.X.; Wang, W.; Yang, B. Bioactive polar compounds from stem bark of *Magnolia officinalis*. *Fitoterapia* **2012**, *83*, 356–361. [[CrossRef](#)] [[PubMed](#)]
24. Seo, K.H.; Lee, D.Y.; Lee, D.S.; Park, J.H.; Jeong, R.H.; Jung, Y.J.; Shrestha, S.; Chung, I.S.; Kim, G.S.; Kim, Y.C.; et al. Neolignans from the fruits of *Magnolia obovata* and their inhibition effect on NO production in LPS-induced RAW 264.7 cells. *Planta Med.* **2013**, *79*, 1335–1340. [[CrossRef](#)]
25. Barros, L.F.; Barison, A.; Salvador, M.J.; de Mello-Silva, R.; Cabral, E.C.; Eberlin, M.N.; Stefanello, M.E. Constituents of the leaves of *Magnolia ovata*. *J. Nat. Prod.* **2009**, *72*, 1529–1532. [[CrossRef](#)] [[PubMed](#)]
26. Lee, J.; Kim, N.H.; Nam, J.W.; Lee, Y.M.; Jang, D.S.; Kim, Y.S.; Nam, S.H.; Seo, E.K.; Yang, M.S.; Kim, J.S. Scopoletin from the flower buds of *Magnolia fargesii* inhibits protein glycation, aldose reductase, and cataractogenesis ex vivo. *Arch. Pharm. Res.* **2010**, *33*, 1317–1323. [[CrossRef](#)] [[PubMed](#)]
27. Park, C.H.; Park, S.Y.; Lee, S.Y.; Kim, J.K.; Park, S.U. Analysis of Metabolites in White Flowers of *Magnolia denudata* Desr. and Violet Flowers of *Magnolia liliiflora* Desr. *Molecules* **2018**, *23*, 1558. [[CrossRef](#)] [[PubMed](#)]
28. Dhandapani, S.; Jin, J.; Sridhar, V.; Sarojam, R.; Chua, N.H.; Jang, I.C. Integrated metabolome and transcriptome analysis of *Magnolia champaca* identifies biosynthetic pathways for floral volatile organic compounds. *BMC Genom.* **2017**, *18*, 463. [[CrossRef](#)] [[PubMed](#)]

29. Hyeon, H.; Hyun, H.B.; Go, B.; Kim, S.C.; Jung, Y.H.; Ham, Y.M. Profiles of Essential Oils and Correlations with Phenolic Acids and Primary Metabolites in Flower Buds of *Magnolia heptapeta* and *Magnolia denudata* var. *purpurascens*. *Molecules* **2021**, *27*, 221. [[CrossRef](#)]
30. Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **2015**, *12*, 523–526. [[CrossRef](#)] [[PubMed](#)]
31. Kanehisa, M.; Sato, Y. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci.* **2020**, *29*, 28–35. [[CrossRef](#)] [[PubMed](#)]
32. Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal. Chem.* **2016**, *88*, 7946–7958. [[CrossRef](#)]
33. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
34. Su, X.; Yang, L.; Wang, D.; Shu, Z.; Yang, Y.; Chen, S.; Song, C. 1 K Medicinal Plant Genome Database: An integrated database combining genomes and metabolites of medicinal plants. *Hortic. Res.* **2022**, *9*, uhac075. [[CrossRef](#)] [[PubMed](#)]
35. Rai, A.; Hirakawa, H.; Nakabayashi, R.; Kikuchi, S.; Hayashi, K.; Rai, M.; Tsugawa, H.; Nakaya, T.; Mori, T.; Nagasaki, H.; et al. Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis. *Nat. Commun.* **2021**, *12*, 405. [[CrossRef](#)] [[PubMed](#)]
36. Sun, L.; Rai, A.; Rai, M.; Nakamura, M.; Kawano, N.; Yoshimatsu, K.; Suzuki, H.; Kawahara, N.; Saito, K.; Yamazaki, M. Comparative transcriptome analyses of three medicinal *Forsythia* species and prediction of candidate genes involved in secondary metabolisms. *J. Nat. Med.* **2018**, *72*, 867–881. [[CrossRef](#)] [[PubMed](#)]
37. Rai, M.; Rai, A.; Kawano, N.; Yoshimatsu, K.; Takahashi, H.; Suzuki, H.; Kawahara, N.; Saito, K.; Yamazaki, M. De Novo RNA Sequencing and Expression Analysis of *Aconitum carmichaelii* to Analyze Key Genes Involved in the Biosynthesis of Diterpene Alkaloids. *Molecules* **2017**, *22*, 2155. [[CrossRef](#)] [[PubMed](#)]
38. Udomsom, N.; Rai, A.; Suzuki, H.; Okuyama, J.; Imai, R.; Mori, T.; Nakabayashi, R.; Saito, K.; Yamazaki, M. Function of AP2/ERF Transcription Factors Involved in the Regulation of Specialized Metabolism in *Ophiorrhiza pumila* Revealed by Transcriptomics and Metabolomics. *Front. Plant Sci.* **2016**, *7*, 1861. [[CrossRef](#)]
39. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
40. Kumar, S.; Blaxter, M.L. Comparing de novo assemblers for 454 transcriptome data. *BMC Genom.* **2010**, *11*, 571. [[CrossRef](#)] [[PubMed](#)]
41. Nakasugi, K.; Crowhurst, R.; Bally, J.; Waterhouse, P. Combining transcriptome assemblies from multiple *de novo* assemblers in the allo-tetraploid plant *Nicotiana benthamiana*. *PLoS ONE* **2014**, *9*, e91776. [[CrossRef](#)] [[PubMed](#)]
42. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
43. Xie, Y.; Wu, G.; Tang, J.; Luo, R.; Patterson, J.; Liu, S.; Huang, W.; He, G.; Gu, S.; Li, S.; et al. SOAPdenovo-Trans: De novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **2014**, *30*, 1660–1666. [[CrossRef](#)]
44. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [[CrossRef](#)]
45. Seppey, M.; Manni, M.; Zdobnov, E.M. BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol. Biol.* **2019**, *1962*, 227–245. [[CrossRef](#)]
46. Cristea, R.M.; Sava, C.; Căpățână, C.; Kanellou, A. Phytochemical Analysis and Specific Activities of Bark and Flower Extracts from Four *Magnolia* Plant Species. *Horticulturae* **2024**, *10*, 141. [[CrossRef](#)]
47. Kelm, M.A.; Nair, M.G. A Brief Summary of Biologically Active Compounds from *Magnolia* spp. In *Studies in Natural Products Chemistry*; Attaur, R., Ed.; Elsevier: Amsterdam, The Netherlands, 2000; Volume 24, pp. 845–873.
48. Li, D.; Liu, K.; Zhao, C.; Liang, S.; Yang, J.; Peng, Z.; Xia, A.; Yang, M.; Luo, L.; Huang, C.; et al. GWAS Combined with WGCNA of Transcriptome and Metabolome to Excavate Key Candidate Genes for Rice Anaerobic Germination. *Rice* **2023**, *16*, 49. [[CrossRef](#)] [[PubMed](#)]
49. Abid, M.; Gu, S.; Zhang, Y.J.; Sun, S.; Li, Z.; Bai, D.F.; Sun, L.; Qi, X.J.; Zhong, Y.P.; Fang, J.B. Comparative transcriptome and metabolome analysis reveal key regulatory defense networks and genes involved in enhanced salt tolerance of *Actinidia* (kiwifruit). *Hortic. Res.* **2022**, *9*, uhac189. [[CrossRef](#)] [[PubMed](#)]
50. Dowle, E.J.; Powell, T.H.Q.; Doellman, M.M.; Meyers, P.J.; Calvert, M.B.; Walden, K.K.O.; Robertson, H.M.; Berlocher, S.H.; Feder, J.L.; Hahn, D.A.; et al. Genome-wide variation and transcriptional changes in diverse developmental processes underlie the rapid evolution of seasonal adaptation. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23960–23969. [[CrossRef](#)]

51. Zheng, J.; He, C.; Qin, Y.; Lin, G.; Park, W.D.; Sun, M.; Li, J.; Lu, X.; Zhang, C.; Yeh, C.T.; et al. Co-expression analysis aids in the identification of genes in the cuticular wax pathway in maize. *Plant J.* **2019**, *97*, 530–542. [[CrossRef](#)] [[PubMed](#)]
52. Shahan, R.; Zawora, C.; Wight, H.; Sittmann, J.; Wang, W.; Mount, S.M.; Liu, Z. Consensus Coexpression Network Analysis Identifies Key Regulators of Flower and Fruit Development in Wild Strawberry. *Plant Physiol.* **2018**, *178*, 202–216. [[CrossRef](#)]
53. Xu, Y.; Magwanga, R.O.; Yang, X.; Jin, D.; Cai, X.; Hou, Y.; Wei, Y.; Zhou, Z.; Wang, K.; Liu, F. Genetic regulatory networks for salt-alkali stress in *Gossypium hirsutum* with differing morphological characteristics. *BMC Genom.* **2020**, *21*, 15. [[CrossRef](#)] [[PubMed](#)]
54. Sun, W.; Wang, B.; Yang, J.; Wang, W.; Liu, A.; Leng, L.; Xiang, L.; Song, C.; Chen, S. Weighted Gene Co-expression Network Analysis of the Dioscin Rich Medicinal Plant *Dioscorea nipponica*. *Front. Plant Sci.* **2017**, *8*, 789. [[CrossRef](#)]
55. Umezawa, T.; Yamamura, M.; Nakatsubo, T.; Suzuki, S.; Hattori, T. Stereoselectivity of the Biosynthesis of Norlignans and Related Compounds. In *The Biological Activity of Phytochemicals*; Gang, D.R., Ed.; Springer: New York, NY, USA, 2011; pp. 179–197.
56. Suzuki, S.; Umezawa, T. Biosynthesis of lignans and norlignans. *J. Wood Sci.* **2007**, *53*, 273–284. [[CrossRef](#)]
57. Davin, L.B.; Lewis, N.G. An historical perspective on lignan biosynthesis: Monolignol, allylphenol and hydroxycinnamic acid coupling and downstream metabolism. *Phytochem. Rev.* **2003**, *2*, 257–288. [[CrossRef](#)]
58. Umezawa, T. Diversity in lignan biosynthesis. *Phytochem. Rev.* **2003**, *2*, 371–390. [[CrossRef](#)]
59. Katayama, T.; Kado, Y. Formation of optically active neolignans from achiral coniferyl alcohol by cell-free extracts of *Eucommia ulmoides*. *J. Wood Sci.* **1998**, *44*, 244–246. [[CrossRef](#)]
60. Lourith, N.; Katayama, T.; Ishikawa, K.; Suzuki, T. Biosynthesis of a syringyl 8-O-4' neolignan in *Eucommia ulmoides*: Formation of syringylglycerol-8-O-4'-(sinapyl alcohol) ether from sinapyl alcohol. *J. Wood Sci.* **2005**, *51*, 379–386. [[CrossRef](#)]
61. Orr, J.D.; Lynn, D.G. Biosynthesis of dehydrodiconiferyl alcohol glucosides: Implications for the control of tobacco cell growth. *Plant Physiol.* **1992**, *98*, 343–352. [[CrossRef](#)] [[PubMed](#)]
62. Sartorelli, P.c.; Benevides, P.J.C.; Ellensohn, R.M.; Rocha, M.V.A.F.; Moreno, P.R.H.; Kato, M.J. Enantioselective conversion of p-hydroxypropenylbenzene to (+)-conocarpan in *Piper regnellii*. *Plant Sci.* **2001**, *161*, 1083–1088. [[CrossRef](#)]
63. Cardullo, N.; Muccilli, V.; Tringali, C. Laccase-mediated synthesis of bioactive natural products and their analogues. *RSC Chem. Biol.* **2022**, *3*, 614–647. [[CrossRef](#)] [[PubMed](#)]
64. Janusz, G.; Pawlik, A.; Swiderska-Burek, U.; Polak, J.; Sulej, J.; Jarosz-Wilkolazka, A.; Paszczyński, A. Laccase Properties, Physiological Functions, and Evolution. *Int. J. Mol. Sci.* **2020**, *21*, 966. [[CrossRef](#)] [[PubMed](#)]
65. Niemetz, R.; Gross, G.G. Oxidation of pentagalloylglucose to the ellagitannin, tellimagrandin II, by a phenol oxidase from *Tellima grandiflora* leaves. *Phytochemistry* **2003**, *62*, 301–306. [[CrossRef](#)]
66. Niemetz, R.; Gross, G.G. Ellagitannin biosynthesis: Laccase-catalyzed dimerization of tellimagrandin II to cornusiin E in *Tellima grandiflora*. *Phytochemistry* **2003**, *64*, 1197–1201. [[CrossRef](#)] [[PubMed](#)]
67. Davin, L.B.; Wang, H.B.; Crowell, A.L.; Bedgar, D.L.; Martin, D.M.; Sarkanen, S.; Lewis, N.G. Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science* **1997**, *275*, 362–366. [[CrossRef](#)] [[PubMed](#)]
68. Yonekura-Sakakibara, K.; Yamamura, M.; Matsuda, F.; Ono, E.; Nakabayashi, R.; Sugawara, S.; Mori, T.; Tobimatsu, Y.; Umezawa, T.; Saito, K. Seed-coat protective neolignans are produced by the dirigent protein AtDP1 and the laccase AtLAC5 in *Arabidopsis*. *Plant Cell* **2021**, *33*, 129–152. [[CrossRef](#)] [[PubMed](#)]
69. Ranocha, P.; McDougall, G.; Hawkins, S.; Sterjiades, R.; Borderies, G.; Stewart, D.; Cabanes-Macheteau, M.; Boudet, A.M.; Goffner, D. Biochemical characterization, molecular cloning and expression of laccases—A divergent gene family—in poplar. *Eur. J. Biochem.* **1999**, *259*, 485–495. [[CrossRef](#)]
70. Razna, K.; Harencar, L.; Kucka, M. The Involvement of microRNAs in Plant Lignan Biosynthesis-Current View. *Cells* **2022**, *11*, 2151. [[CrossRef](#)]
71. Corbin, C.; Drouet, S.; Markulin, L.; Auguin, D.; Laine, E.; Davin, L.B.; Cort, J.R.; Lewis, N.G.; Hano, C. A genome-wide analysis of the flax (*Linum usitatissimum* L.) dirigent protein family: From gene identification and evolution to differential regulation. *Plant Mol. Biol.* **2018**, *97*, 73–101. [[CrossRef](#)]
72. Yadav, V.; Wang, Z.; Yang, X.; Wei, C.; Changqing, X.; Zhang, X. Comparative Analysis, Characterization and Evolutionary Study of Dirigent Gene Family in Cucurbitaceae and Expression of Novel Dirigent Peptide against Powdery Mildew Stress. *Genes* **2021**, *12*, 326. [[CrossRef](#)]
73. Meng, Q.; Kim, S.J.; Costa, M.A.; Moinuddin, S.G.A.; Celay, R.M.; Smith, C.A.; Cort, J.R.; Davin, L.B.; Lewis, N.G. Dirigent protein subfamily function and structure in terrestrial plant phenol metabolism. *Methods Enzym.* **2023**, *683*, 101–150. [[CrossRef](#)]
74. Luo, R.; Pan, W.; Liu, W.; Tian, Y.; Zeng, Y.; Li, Y.; Li, Z.; Cui, L. The barley DIR gene family: An expanded gene family that is involved in stress responses. *Front. Genet.* **2022**, *13*, 1042772. [[CrossRef](#)] [[PubMed](#)]
75. Gong, L.; Li, B.; Zhu, T.; Xue, B. Genome-wide identification and expression profiling analysis of DIR gene family in *Setaria italica*. *Front. Plant Sci.* **2023**, *14*, 1243806. [[CrossRef](#)] [[PubMed](#)]
76. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009**, *10*, R25. [[CrossRef](#)]

77. 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](#)]
78. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)] [[PubMed](#)]
79. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
80. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [[CrossRef](#)]
81. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]

**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.