Article

Coexpression Network Analysis Based Characterisation of the R2R3-MYB Family Genes in Tolerant Poplar Infected with *Melampsora larici-populina*

Qiaoli Chen [1,2], **Feng Wang** [1,2] **and Danlei Li** [1,2,*]

1   Key Laboratory of Alien Forest Pests Monitoring and Control-Heilongjiang Province, School of Forestry, Northeast Forestry University, Harbin 150040, China
2   Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, Northeast Forestry University, Harbin 150040, China
*   Correspondence: danleili@nefu.edu.cn; Tel.: +86-0451-82190384

Abstract: R2R3-MYB protein is the most abundant class of MYB transcription factor family in plants. The transcript profiles of two E4 races of *Melampsora larici-populina*-tolerant poplars and an intolerant poplar were investigated to characterise the role of the R2R3-MYB family genes in the poplar–E4 interaction. In this study, 217 R2R3-MYBs were identified, and 83 *R2R3-MYB* genes were assigned to 22 different coexpression modules by weighted gene coexpression network analysis. Most *R2R3-MYB* genes were unchanged in the early period of E4 infection in both tolerant and intolerant poplars. However, there were obvious increases in differentially expressed *R2R3-MYB* genes in tolerant poplars at 2 and 4 dpi when defence responses occurred, suggesting that differently expressed *R2R3-MYB* genes at these time points may play an important role in poplar resistance to E4 infection. In total, 34 *R2R3-MYB* genes showed differential expression at 2 and 4 dpi between tolerant and intolerant poplars. Among them, 16 differentially expressed *R2R3-MYB* genes were related to 43 defence-related genes that had significant differences between tolerant and intolerant poplars. There might be coregulatory relationships between R2R3-MYBs and other TFs during poplar–E4 interaction. Some differentially expressed *R2R3-MYB* genes were related to genes involved in flavonoid biosynthesis and IAA or free SA signal transduction and might help activate defence response during poplar–E4 interaction. MYB194 could be an important node in the convergence of IAA and SA signalling.

Keywords: R2R3-MYB; *Populus*; rust; *Melampsora larici-populina*; transcription factor

1. Introduction

Melampsora larici-populina causes serious foliar rust disease of poplar worldwide [1,2]. Hybrids between *Populus deltoides*, *P. nigra*, or *P. trichocarpa* were selected for their immunity to rust in the mid-20th century in Europe. However, outbreaks of rust on clones of these hybrids were caused by E4, a new rust race that occurred in the late 20th century. Nevertheless, some hybrid poplars were found to be tolerant to infection by E4 and developed hypersensitive response (HR) cell death at the infection site [3]. Further study to identify the genetic basis of host susceptibility or tolerance will provide new insight into how rust overcomes poplar resistance.

To ensure the optimal intensity and duration of immune responses, plant innate immunity is regulated at different levels. In plant immune signalling pathways, mitogen-activated protein kinases (MAPKs) [4–6] and calcium-dependent protein kinases (CDPKs) [7] can regulate the expression of plant immune-related genes by phosphorylating downstream transcription factors (TFs). TFs are proteins that control target gene expression levels and modulate rates of transcription. Many TFs can regulate plant immunity by regulating the expression of downstream defence-related genes [8] and participate in regulating the crosstalk between different defensive signalling pathways [9]. Several small messenger

molecules, such as salicylic acid (SA), jasmonic acid (JA), and ethylene (ET), are involved in translating pathogen-induced early signalling events into the activation of effective defence responses [10,11]. Some TFs are important nodes for the convergence of phytohormone signalling and play an important role in the regulation of phytohormone-responsive genes [12]. Therefore, many TFs are key participants in plant immune responses and are considered key targets for genetic engineering to enhance adaptation to not only abiotic but also biotic stresses in valuable plants.

The transcript levels of TFs, such as MYBs, AP2/ERFs, and WRKYs, were previously found to be altered in poplars infected with E4 [3,13,14]. The MYB family is one of the largest groups of TFs in plants. The MYB domain, which is a highly conserved DNA-binding domain, is characteristic of MYB proteins. Structurally, the MYB domain consists of 1–4 MYB repeat units, and the MYB family can be divided into four types according to the number of replicates, namely 4R-MYB (containing 4 MYB replicates), 3R-MYB (containing 3 MYB replicates), R2R3-MYB (containing 2 MYB replicates), and MYB-related (containing 1 MYB). In plants, the R2R3-MYB protein is the most abundant class of MYB family, containing two MYB domains at the N-terminus. The plant MYB family has selectively expanded, particularly through the large R2R3-MYB family, and many (if not all) R2R3-MYBs play central roles in plant-specific processes [15,16]. As the MYB family is expected to play an important role in plant defence against biotic and abiotic stresses [17,18], placing each member in an organised nomenclature system and providing maps of MYB family gene expression should contribute to unravelling the complexity of the transcriptional regulation of defence-related genes in poplar–rust interactions.

Our previous study indicated that the hybrid poplar *P. nigra* × *P. deltoides* ('Intolerant') is susceptible to the virulent E4 race of *M. larici-populina*, whereas *P. deltoides* × *P. trichocarpa* ('Tolerant 2') and *P. trichocarpa* × *P. deltoides* ('Tolerant 1') are tolerant to E4 and timely activation or inhibition of the SA or JA pathways is the key difference between tolerant and intolerant poplars [3]. However, the molecular characteristics of TFs, especially defence-related TFs such as R2R3-MYBs, need to be further analysed and explored [19,20]. Therefore, in this study, weighted gene coexpression network analysis (WGCNA) [21] was further performed to obtain a better understanding of the expression patterns of *R2R3-MYB* genes and their related genes in E4-infected poplars to provide a better understanding of the interaction between poplar and rust on the basis of the previous study [3]. The changes in transcriptome profiling and contents of indole-3-acetic acid (IAA) [22] and free SA [20] after E4 infection were also investigated to study whether and how R2R3-MYBs might interfere with phytohormone signalling pathways.

2. Materials and Methods

2.1. E4 Isolates, Plant Materials, and Inoculation Procedure

E4-infected poplar leaves were collected from *P. trichocarpa* cv. Trichobel at Markington (northern England) [2], and E4 rust isolates were obtained from single uredinial pustules as previously reported [2,3,13,14]. One-year-old hybrid poplars, including an intolerant poplar, *P. nigra* × *P. deltoides* ('Intolerant'), and two tolerant poplars, *P. deltoides* × *P. trichocarpa* ('Tolerant 2') and *P. trichocarpa* × *P. deltoides* ('Tolerant 1'), were used as plant tissue sources. E4 completed its vegetative cycle in 7 days on 'Intolerant'. The growth of E4 was inhibited in 'Tolerant 2' and 'Tolerant 1'. By 7 days post-inoculation (dpi), only a few new or barely mature urediniospores were found, and visible scattered lesions and confluent necrosis appeared on 'Tolerant 2' and 'Tolerant 1', respectively [3].

These hybrid poplars were grown as described previously [3,13,14,23]. Leaves of hybrid poplars were inoculated with E4 as described by Pei et al. [2] and Chen et al. [3,13,14]. In brief, fully expanded leaves from leaf plastochrony index 5–9 were detached from plants and spray-inoculated on their abaxial surface with a rust spore suspension in deionised water containing 0.004% Tween 20 adjusted to 100,000 spores ml^{-1}, or with deionised water containing 0.004% Tween 20 as a control. After inoculation, the E4-inoculated leaves were incubated in a phytotron with 16 h day^{-1} illumination (80 µE m^{-2} s^{-1}) for different periods

as described previously (2, 6, and 12 h and 1, 2, 4, and 7 d) [2,3,13,14,23]. The control groups contained E4-free leaves (leaves were treated with deionised water only) incubated under the same conditions. Each leaf sample was frozen and ground using liquid nitrogen for RNA extraction.

2.2. Transcriptome Library Preparation and Sequencing

The CTAB (cetyltrimethylammonium bromide) method was used to extract total RNA. The RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA) was used to assess RNA quality. A NanoDrop^{TM} spectrophotometer (Thermo Scientific, Wilmington, DE, USA) was used to determine the RNA purity. Genomic DNA was removed using DNase I, Amplification Grade (Invitrogen, Foster, CA, USA, cat. no. 18068-015). The construction of the libraries and sequencing of the three poplars were performed on a BGISEQ-500 RNA-seq platform (BGI, Shenzhen, China). The average insert size for the paired-end libraries was 300 bp ($\pm$50 bp).

Low-quality (more than 20% of bases in the total read have quality score lower than 15), adaptor-polluted and high content (5%) of unknown base (N) reads were trimmed to acquire clean reads using SOAPnuke (v1.5.2, https://github.com/BGI-flexlab/SOAPnuke, accessed on 24 November 2016). After clean reads were obtained, HISAT2 (Hierarchical Indexing for Spliced Alignment of Transcripts, v2.0.4, http://www.ccb.jhu.edu/software/hisat, accessed on 18 May 2016) was used to align clean reads to the genome sequence of *P. trichocarpa* (version 3.0, http://www.phytozome.net/poplar.php, accessed on 26 November 2018) [24,25]. The uniformity of the mapping result for each sample suggested that the samples were comparable. After comparison with the reference genome, StringTie (v1.0.4, http://ccb.jhu.edu/software/stringtie, accessed on 19 May 2016) [26] was used to reconstruct the transcriptome of each sample.

2.3. Identification of R2R3-MYBs

Extensive BLAST (Basic Local Alignment Search Tool) searches (https://www.ncbi.nlm.nih.gov/, accessed on 15 May 2020) were conducted to select R2R3-MYB family members based on transcriptome sequencing results. The sequences of 192 R2R3-MYBs identified by Wilkins et al. [27] were used as the query sequences to perform BLASTP local search with the E value set to 1×10^{-10}. The selected candidate sequences were then assessed by the hidden Markov model of Pfam (PF00249, http://xfam.org/, accessed on 15 May 2020) with the E value set to 1×10^{-3}. After that, the sequences were submitted to the website InterPro (http://www.ebi.ac.uk/interpro/, accessed on 15 May 2020) and SMART (http://smart.embl-heidelberg.de/, accessed on 15 May 2020) for DNA-binding domain test, and proteins containing two repeated sequences (R2 and R3) in the DNA-binding domain were taken as R2R3-MYB family members in the three hybrid poplars.

2.4. WGCNA

WGCNA (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/, accessed on 19 May 2020) [21] was performed to obtain a better understanding of the expression patterns of *R2R3-MYB* genes and their related genes in E4-infected poplars to provide a better understanding of the interaction between poplar and rust on the basis of our previous study [3]. WGCNA was implemented in the R software package (http://www.r-project.org/, accessed on 19 May 2020). The input data and parameter settings were consistent with our previous study [3]. After filtering out genes with median FPKM levels that did not exceed 1 [3], the expression of 97 of the *R2R3-MYB* genes were included in the construction of the coexpression module (grey module was reserved for unassigned genes) with WGCNA package tools.

According to the tutorials of the WGCNA package, the module eigengene (ME) is defined as the first principal component of a given module, which can be considered representative of the gene expression profiles in a module. Intramodular connectivity (IC) was defined only for the genes inside a given module and was calculated for each gene

by summing the connection strengths with those of other module genes and dividing this number by the maximum intramodular connectivity. IC measures how connected, or coexpressed, a given gene is with respect to the genes of a particular module. Correlation analyses between MEs and external traits (SA and IAA levels) were performed to look for the most significant associations. For each expression profile, gene significance (GS) was calculated as the absolute value of the Pearson correlation between the expression profile and each trait. Module membership (MM) was defined based on the Pearson correlation of the expression profile and each ME. Genes with higher MM were defined as the more important (central) elements of the modules.

The eigengene dendrogram and heatmap were used to identify groups of correlated eigengenes. Gene coexpression network depictions were constructed using Cytoscape software [28]. In gene coexpression networks and *R2R3-MYB*-gene-focused gene networks, genes with weight values were connected by lines and the colour of the line represented the weight value between two genes. Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analysis was performed on selected genes [29] using the hypergeometric test to determine which pathway was significantly enriched in selected genes compared with the whole genome background and p value ≤ 0.05 after the correction was used as the threshold. PossionDis algorithms were used to detect the DEG (differentially expression genes). In this study, DEGs were defined by default as those with a false discovery rate (FDR) ≤ 0.05 and differences of more than twofold ($\log_2$ value of E4-inoculated expression to E4-free expression ≥ 1 or ≤ -1). The depictions of heatmaps of gene expression pattern were constructed using the OmicShare tools, a free online platform for data analysis (http://www.omicshare.com/tools, accessed on 21 April 2022).

2.5. Quantification of IAA and Free SA Levels

SA and IAA were quantified by high-performance liquid chromatography (HPLC) mass spectrometry from crude plant extracts based on the method of Pan et al. [30] and Chen et al. [3]. According to Pan et al., poplar leaves were frozen in liquid nitrogen and then ground followed by adding working solution and extraction solvent. Then, SA and IAA were extracted and quantified by an ultra-HPLC-Q-Exactive™ system (Thermo Scientific, San José, CA, USA) using an ODS column (μ-Bondasphere C_{18}, 5 μm, 3.9 × 150 mm; Waters, Milford, MA, USA). Authentic SA and IAA (Sigma Aldrich, Burlington, MA, USA, cat. no. S5922 and I3750) were used as external standards. The amounts of SA and IAA were calculated by comparing with the corresponding internal standard. Three separate biological replicates of each treatment were performed, and each replicate was assessed three times.

2.6. RT-qPCR

Quantitative real-time reverse transcription PCR (RT–qPCR) was performed with the GoTaq 2-Step RT–qPCR System Kit (Promega, Madison, WI, USA, cat. no. A6010) and the Stratagene Mx3000P qPCR system (Agilent Technologies, Santa Clara, CA, USA) to validate the transcript levels of selected genes at 2 hpi, 6 hpi, 12 hpi, 1 dpi, 2 dpi, 4 dpi, and 7 dpi. The 18S ribosomal RNA was used as the internal control. All primers used in this study are listed in Table S1. The PCR program was 95 °C for 10 s and 40 cycles of 95 °C for 5 s, 58 °C for 30 s, and 72 °C for 30 s. Each treatment was performed on three separate biological replicates, and each replicate was measured three times. Quantification of the RT–qPCR results was conducted as previously described [3]. The normalisation of the data followed the instructions of the GoTaq 2-Step RT–qPCR System Kit and the $2^{-\Delta\Delta C_T}$ method [31]. Significance was determined by Student's *t*-test.

3. Results

3.1. Identification of R2R3-MYBs

The transcriptomes of the leaves of the three poplars were measured, producing an average of 6.54 GB of data per sample (Table S2). The average genome mapping

rate is 79.58% and the average gene mapping rate is 77.77%. A total of 34,279 genes were detected, including 33,896 known genes and 406 predicted new genes. A total of 12,994 new transcripts were detected, of which 11,604 were new alternative splicing subtypes of known protein-coding genes, 406 were transcripts of new protein-coding genes, and the remaining 984 were long noncoding RNAs. Based on the results of Wilkins et al. [27], BLASTP was used to search for translated nucleotides of R2R3-MYBs in two tolerant poplars, *P. trichocarpa* × *P. deltoides* ('Tolerant 1') and *P. deltoides* × *P. trichocarpa* ('Tolerant 2'), and an intolerant poplar, *P. nigra* × *P. deltoides* ('Intolerant'). Our extensive search for R2R3 MYB DNA binding-domain-containing proteins identified 217 putative distinct R2R3-MYBs in the three poplars and named them MYB001 to MYB217 (Table S3).

3.2. Expression Characteristics of R2R3-MYB Genes Based on WGCNA

Of the 97 *R2R3-MYB* genes analysed by WGCNA, 83 were assigned to 22 different coexpression modules (Table S4), and 14 were assigned to the grey module. As the grey module was reserved for unassigned genes, the modules that contained more than 5 of the *R2R3-MYB* genes were blue, brown, black, green, green–yellow, yellow, and turquoise. Correlation of MEs was analysed and groups of correlated MEs of modules that these *R2R3-MYB* genes assigned to were identified, indicating that the correlation of expressions of most *R2R3-MYB* genes was low (Figure S1A,B). This also suggested that different *R2R3-MYB* genes might play different roles in response of poplar to rust infestation. To further analyse the gene expression differences in *R2R3-MYB* genes in the three poplars at different time points, the expression patterns of those 83 *R2R3-MYB* genes in the 22 modules were compared (Figure 1A). The results indicated that obvious differences in the expression levels between tolerant and intolerant poplars were observable mainly at 4 dpi.
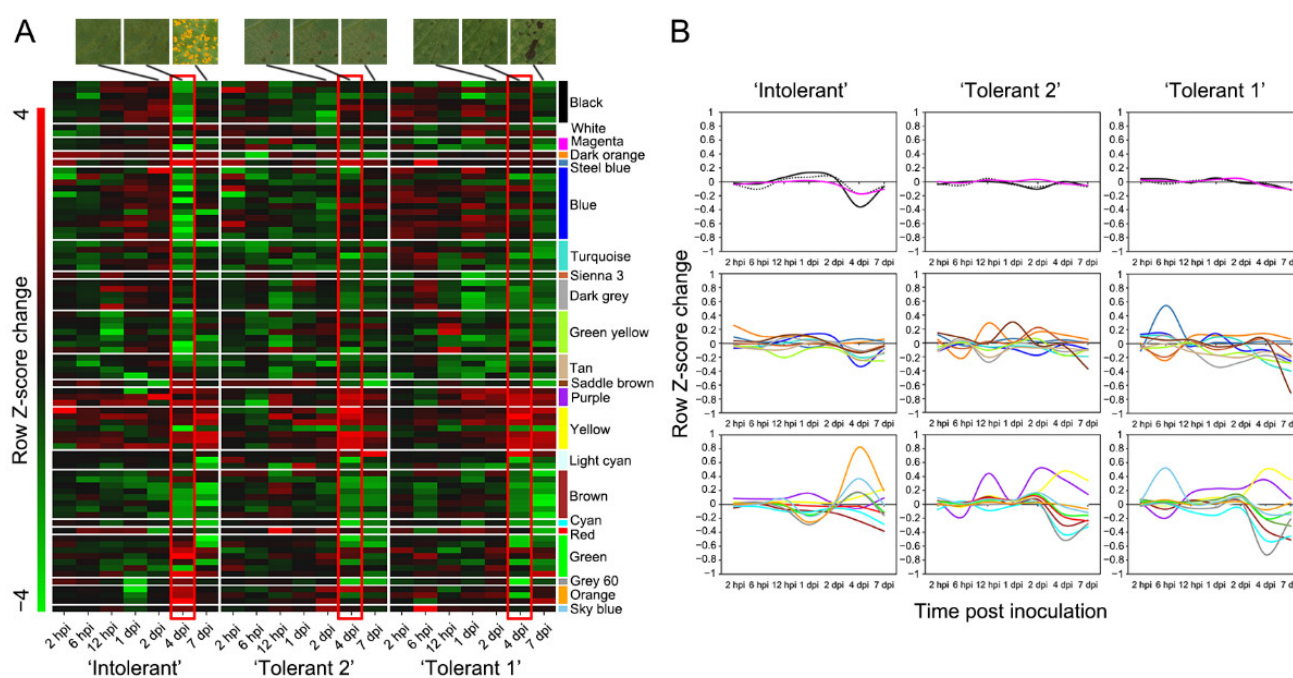

Figure 1. WGCNA revealed the expression characteristics of *R2R3-MYB* genes. (**A**) Heatmap of *R2R3-MYB* genes (rows, the Z-score change was calculated for the expression of each gene based on FPKM, and the results are presented in E4-inoculated samples minus E4-free samples) across the samples (columns). The red rectangle highlights the difference in expression. (**B**) Comparison of eigengene expressions (*y*-axis, row Z-score change in eigengene expression, E4-inoculated sample minus E4-free sample for each time point) in different modules across the samples (*x*-axis). Coexpression modules were shown in different colours.

The MEs of the 22 modules to which those 83 *R2R3-MYB* genes were assigned were compared across the samples (Figure 1B). The most obvious difference among the three poplars was that changes in MEs for different modules were more similar in 'Tolerant 1' and 'Tolerant 2'. In many modules, the MEs did not change obviously in 'Intolerant' at many time points, but the MEs in 'Tolerant 1' and 'Tolerant 2' at the same time points showed obvious increasing or decreasing trends. In contrast, when the MEs did not change obviously in 'Tolerant 1' and 'Tolerant 2', the MEs in 'Intolerant' showed obvious increasing or decreasing trends. These situations occurred at 4 dpi. The results showed that there were significant differences in the MEs between tolerant and intolerant poplars mainly at 4 dpi, which agreed with the expression characteristics of *R2R3-MYB* genes, indicating that the expression of these *R2R3-MYB* genes was consistent with the basic expression characteristics of the genes in the specific modules.

The MM and IC for each *R2R3-MYB* gene in the 22 modules to which the 83 *R2R3-MYB* genes were assigned were calculated (Table S5). Among them, 4 *R2R3-MYB* genes had relatively high MMs (>0.90), and 21 *R2R3-MYB* genes had high MMs (>0.80), indicating that these *R2R3-MYB* genes were central elements in their respective modules. The results also indicated that highly connected intramodular *R2R3-MYB* genes tend to have high MMs in the respective modules. Therefore, *R2R3-MYB* genes with high MMs and ICs of a particular module should occupy important regulatory positions within the module. This result was also found in other TF genes associated with plant–pathogen interactions, such as *AP2*, *ERF*, *NAC*, and *WRKY*.

To identify the functions and pathways of *R2R3-MYB* gene-related genes, KEGG enrichment analyses were performed on genes in the 22 modules with which those 83 *R2R3-MYB* genes had weight values (Table S6). Genes with weight values of *R2R3-MYB* genes were enriched in plant hormone signal transduction (ko04075) in 15 modules, plant–pathogen interaction (ko04626) in 13 modules, and MAPK signalling-pathway–plant (ko04016) in 13 modules. Genes enriched in these three pathways accounted for a large proportion of all genes in approximately half of all these modules, especially the yellow and grey 60 modules. These results suggested that R2R3-MYBs played an important role in the interaction between poplar and E4.

3.3. Differential Expression of R2R3-MYB Genes

To further explore the effect of the differential expression of *R2R3-MYB* genes on the interaction between poplar and E4, the expression of *R2R3-MYB* genes at different time points in different poplars after E4 infection was analysed (Figure 2). We focused on genes with different expression patterns between the tolerant and intolerant poplars. It was found that most *R2R3-MYB* genes were unchanged during the interaction between poplar and E4, but there were many more unchanged *R2R3-MYB* genes at 2 and 4 dpi in 'Intolerant' than in both 'Tolerant 2' and 'Tolerant 1'. Additionally, many *R2R3-MYB* genes were downregulated at 4 and 7 dpi, which were the late stages of E4 infection, in both tolerant and intolerant poplars. To find the main difference between tolerant and intolerant poplars, the *R2R3-MYB* genes with different expression patterns between tolerant and intolerant poplars at 2 and 4 dpi were screened. In total, 15 of the 22 modules had 34 *R2R3-MYB* genes with differential expression at 2 and 4 dpi between tolerant and intolerant poplars (Table S7).
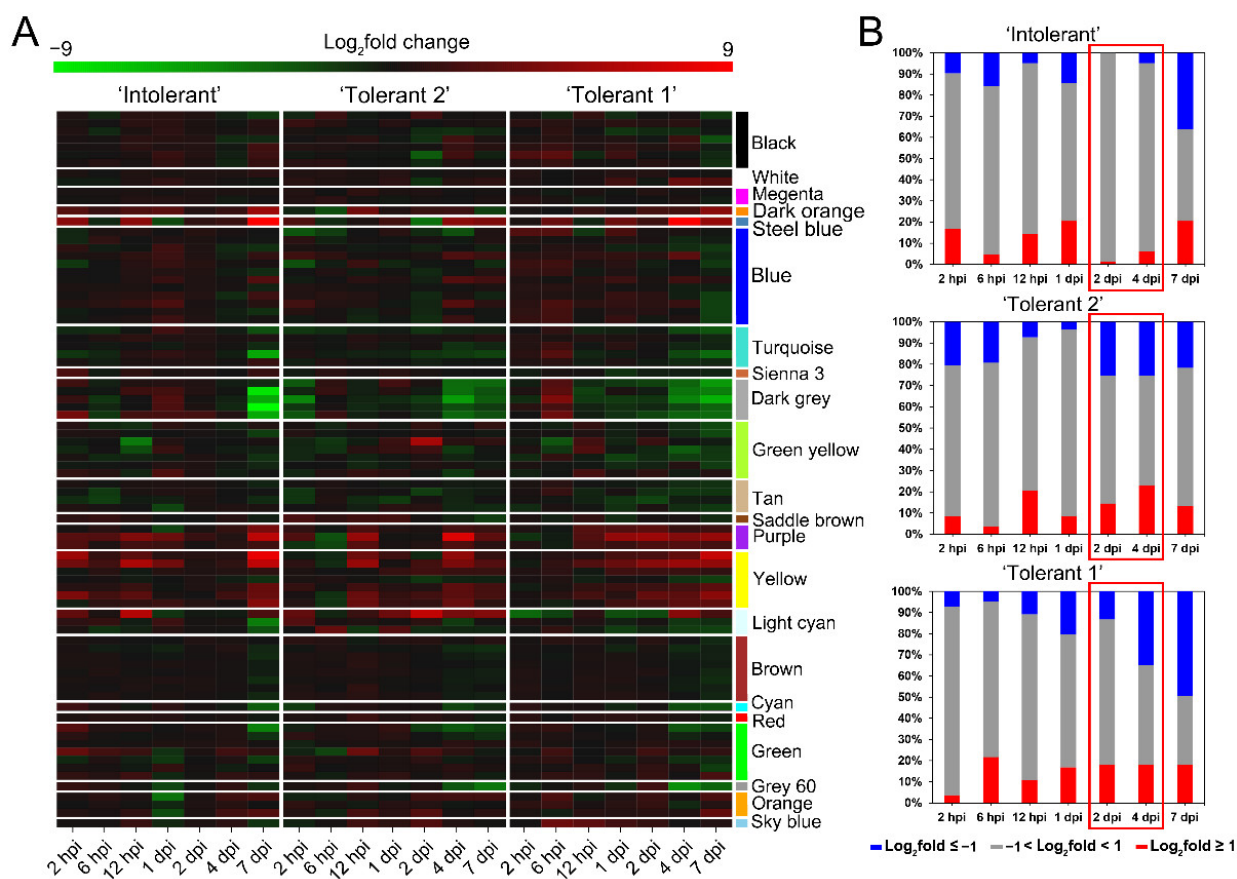
Figure 2. Expression analysis of *R2R3-MYB* genes at different time points in different poplars after E4 infection. (**A**) Heatmap of *R2R3-MYB* genes. (**B**) Proportion of differentially expressed *R2R3-MYB* genes. Coexpression modules were shown in different colours. The red rectangle highlights the difference in proportion.

The expression of 16 *R2R3-MYB* genes in 'Intolerant' was unchanged or significantly lower than that in 'Tolerant 2' and 'Tolerant 1' at 2 or 4 dpi (Figure S2A–C and Table S7), and 6 of these 16 *R2R3-MYB* genes were assigned to the yellow module (Figure S3A–C and Table S7). These 6 *R2R3-MYB* genes from the yellow module showed a gradual increase in expression with increasing inoculation time in 'Tolerant 1'. However, in 'Tolerant 2' and 'Intolerant', the expression of these genes fluctuated greatly, with no obvious change patterns. On the other hand, the expression levels of 18 *R2R3-MYB* genes in 'Intolerant' were unchanged or significantly higher than those in 'Tolerant 2' and 'Tolerant 1' at 2 or 4 dpi (Figure S2D–F and Table S7), and 5 of these *R2R3-MYB* genes were assigned to the dark grey module (Figure S3D–F and Table S7). These 5 *R2R3-MYB* genes from the dark grey module showed a gradual decrease in expression after 6 hpi with the increase in inoculation time in 'Tolerant 1'. However, in 'Tolerant 2' and 'Intolerant', there was no noticeable change in the expression of such genes.

To further explore the relationship between the expression of the *R2R3-MYB* genes and their related genes in the interactions between different poplars and E4, KEGG enrichment results on genes with weighted values of all these differently expressed *R2R3-MYB* genes were further analysed. The modules with more genes enriched in plant–pathogen interaction (ko04626) and signalling-related pathways (ko04016 and ko04075) were in the black, purple, yellow, light cyan, and grey 60 modules (Figure 3 and Table S8). Additionally, many genes were enriched in pathways related to the biosynthesis of other secondary metabolites that were important in plant stress resistance and development, such as phenylpropanoid biosynthesis (ko00940), flavonoid biosynthesis (ko00941), isoflavonoid biosynthesis (ko00943), anthocyanin biosynthesis (ko00942), and flavone and flavonol biosynthesis

(ko00944). The modules with more genes enriched in these biosynthesis of other secondary metabolite pathways were black, dark grey, yellow, and light cyan (Figure 3 and Table S8). Thus, *R2R3-MYB* genes and their related genes in the six modules, black, dark grey, purple, yellow, light cyan, and grey 60 modules, were more likely to be involved in the interaction between poplars and E4.
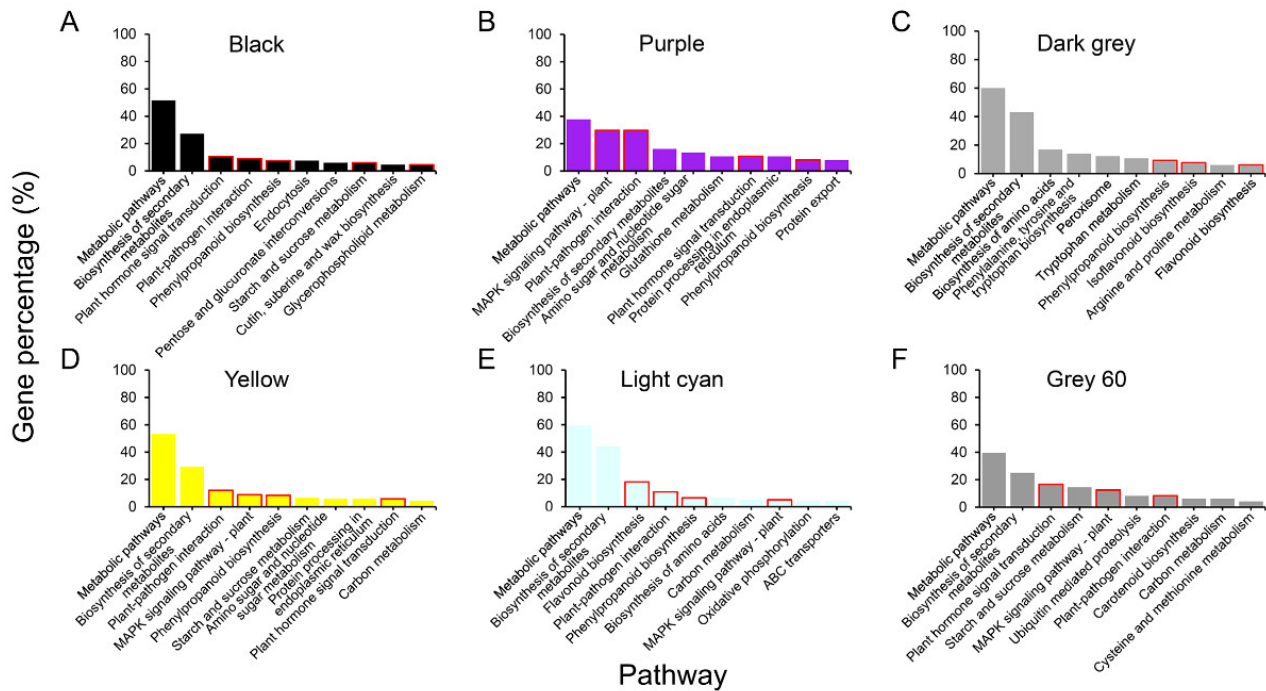

Figure 3. KEGG enrichment analysis for genes that differentially expressed *R2R3-MYB* genes had weight values within selected modules (top 10 pathways, detailed in Table S8). (**A**) The black module. (**B**) The purple module. (**C**) The dark grey module. (**D**) The yellow module. (**E**) The light cyan module. (**F**) The grey 60 module. The red rectangle highlights the pathways involved in the interaction between poplars and E4.

To further explore the regulatory mechanisms of *R2R3-MYB* genes, the functions of the proteins encoded by genes from the six modules, black, dark grey, purple, yellow, light cyan, and grey 60, were annotated with BLASTP based on the NR (nonredundant proteins) database. We focused on the functions of the genes that were enriched in plant–pathogen interaction pathways (ko04626), plant hormone signal transduction (ko04075), MAPK signalling pathway—plant (ko04016), phenylpropanoid biosynthesis (ko00940), flavonoid biosynthesis (ko00941), isoflavonoid biosynthesis (ko00943), anthocyanin biosynthesis (ko00942), flavone and flavonol biosynthesis (ko00944), and cutin, suberine and wax biosynthesis (ko00073), which were all related to the interaction between poplars and E4.

Among the six modules, the yellow module had the largest number of genes enriched in these pathways, at 147. Sixty-nine of these genes were annotated with specific functions (Table S9). Many of the genes encoded pathogenesis-related family proteins, peroxidase family proteins, WRKY transcription factors, and disease-resistance proteins (Figure 4A). The light cyan module had the second highest number of genes enriched in these pathways at 49. Twenty-three of these genes were annotated as having specific functions (Table S9). Many of the genes encoded naregenin-chalcone synthase family proteins, receptor-like proteins, and leucoanthocyanidin dioxygenase family proteins (Figure 4B).
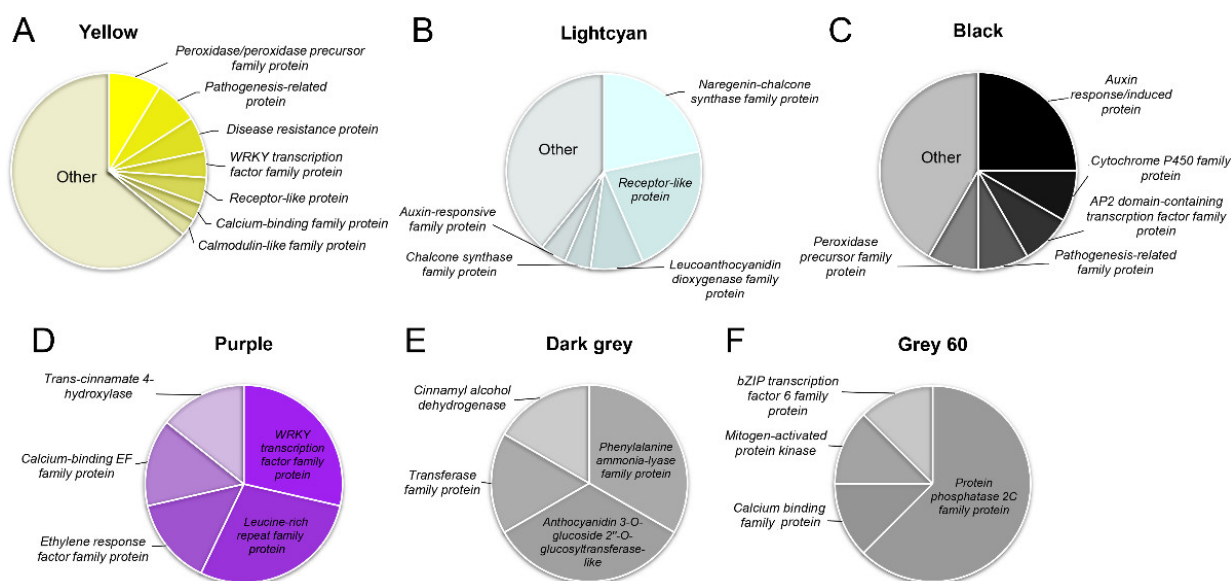
Figure 4. Proportion of differently expressed *R2R3-MYB* gene-related genes involved in poplar–E4 interactions in the selected modules. (**A**) The yellow module. (**B**) The light cyan module. (**C**) The black module. (**D**) The purple module. (**E**) The dark grey module. (**F**) The grey 60 module.

The numbers of genes enriched in these pathways in the black, dark grey, purple, and grey 60 modules were 24, 19, 16, and 13, respectively. Of these, 12, 6, 7, and 8 genes were annotated with specific functions, respectively (Table S9). Many of the genes encoded auxin responsive/induced proteins in the black module; WRKY transcription factors, ethylene response factors, leucine-rich repeat family proteins, and calcium-binding EF hand family proteins in the purple module; phenylalanine ammonia-lyase family proteins and anthocyanidin 3-O-glucoside 2″-O-glucosyltransferase-like proteins in the dark grey module; and protein phosphatase 2C family proteins in the grey 60 module (Figure 4C–F). Thus, different R2R3-MYBs and their related genes should regulate different genes in the pathways.

3.4. Gene Networks of Differently Expressed R2R3-MYB Genes

To further comprehensively analyse the regulatory interactions between differentially expressed *R2R3-MYB* genes and their related genes (genes have weight value with *R2R3-MYB* genes) enriched in the pathways involved in the interaction between poplars and E4 in the black, dark grey, purple, yellow, light cyan, and grey 60 modules, gene networks were predicted based on the weight values between genes for each selected module (Figure 5). In the yellow module, the weight value between *linamarase family protein* (No. 261), *beta-glucosidase 12-like* genes (No. 182 and 205), and *peroxidase* genes (No. 393 and 408) was higher (top 1%, weight value > 0.4, Figure S4). Therefore, there were high correlations among these genes. *K$^+$ rectifying channel family protein* (No. 32) had the highest IC in the network, indicating that this gene occupied a central position among these genes and that the regulation between *MYB169* (No. 204), *MYB194* (No. 371), *MYB024* (No. 653), *MYB129* (No. 1040), *MYB046* (No. 1588), and *MYB011* (No. 1654) and the other genes was affected by *K$^+$ rectifying channel family protein* (No. 32). Additionally, many genes were related to both MYBs (Nos. 204, 371, 653, 1040, 1588, and 1654) and WRKYs (Nos. 200, 305, and 509).
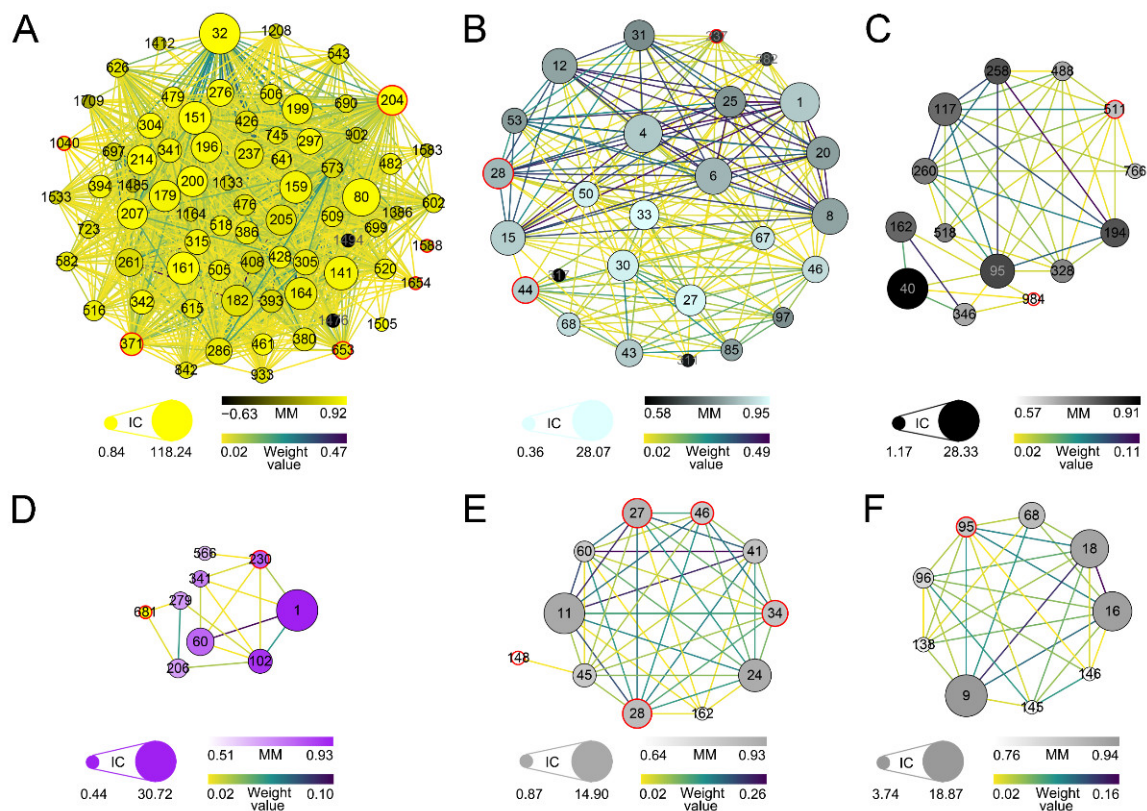
Figure 5. WGCNA revealed gene networks for differentially expressed *R2R3-MYB* genes and their related genes (detailed in Table S10) involved in poplar–E4 interactions in the selected modules. (**A**) The yellow module. (**B**) The light cyan module. (**C**) The black module. (**D**) The purple module. (**E**) The dark grey module. (**F**) The grey 60 module. The size of the dots represents IC. The colour of the dot represents MM. The colour of the line represents the weight value between two genes. The label of the dot is listed based on IC of the specific gene in the specific module. The *R2R3-MYB* genes are highlighted with red borders. IC, intramodular connectivity; MM, module membership.

In the light cyan module, based on weight values between genes, the correlations between *MYB162* (No. 28) and the other genes were greater than the correlations between *MYB086* (No. 44) or *MYB202* (No. 237) and the other genes (Figure 5B). *MYB162* (No. 28) was highly related to *dihydroflavonol reductase family protein* (No. 1), which was the gene with the highest IC in the light cyan module. Additionally, *dihydroflavonol reductase family protein* (No. 1), *chalcone synthase family protein* (No. 4), *naregenin-chalcone synthase family protein* genes (No. 6, 8, 12, 15, 20, and 53), *leucoanthocyanidin dioxygenase family protein* (No. 25), *MYB162* (No. 28), and *leucoanthocyanidin reductase family protein* (No. 31) had high weight values with each other, suggesting that they were highly related (Figure S5).

In the black module, *MYB146* (No. 511) and *MYB101* (No. 984) were involved in regulating the genes of different groups. *Auxin-responsive family protein* (No. 40) was the gene with the highest IC in the network, and this gene was only related to *MYB101* (No. 984). However, *glucose-methanol-choline oxidoreductase family protein* (No. 95) played a role in the connection between *MYB146* (No. 511) and *MYB101* (No. 984). Except for *glucose-methanol-choline oxidoreductase family protein* (No. 95), *Fe(III)-Zn(II) purple acid phosphatase family protein* (No. 117), *AP2 domain-containing transcription factor family protein* (No. 194) *cytochrome P450 family protein* (No. 258), and *beta-ketoacyl-CoA synthase family protein* (No. 260), which had higher weight values with each other, *quinone oxidoreductase family protein* (No. 162) and *transferase family protein* (No. 346) also had high weight values with each other, indicating that they were highly related (Figure 5C). Additionally, *auxin response factor 2 family protein* (No. 488) was related to both *MYB146* (No. 511) and *AP2 domain-containing transcription factor family protein* (No. 194), indicating that it was regulated by both MYB and AP2.

In the purple module, *MYB137* (No. 230) was predicted to be related to *leucine-rich repeat family protein* (No. 1), which was the gene with the highest IC in the module. *The calcium-binding EF hand family protein* (No. 279) played a role in the connection between *MYB137* (No. 230) and *MYB104* (No. 681). *WRKY transcription factor 51 family protein* (No. 60) and *leucine-rich repeat family protein* (No. 1) had a high weight value, indicating they were highly related (Figure 5D). Additionally, *leucine-rich repeat family protein* genes (Nos. 1 and 102) were related to both *MYB137* (No. 230) and *WRKY transcription factor 51 family protein* (No. 60), and *calcium-binding EF hand family protein* (No. 279) was related to both *MYB104* (No. 681) and *WRKY transcription factor 42 family protein* (No. 206), indicating that these genes were regulated by both MYBs and WRKYs.

In the dark grey module, *cinnamyl alcohol dehydrogenase 6* (No. 11) was the gene with the highest IC in the network and was related to *MYB078* (No. 27), *MYB214* (No. 28), *MYB195* (No. 34), and *MYB156* (No. 46). *Phenylalanine ammonia-lyase family protein* (No. 45) played a role in the connection between *MYB082* (No. 148) and the other four *R2R3-MYB* genes. Additionally, *cinnamyl alcohol dehydrogenase 6* (No. 11) and *anthocyanidin 3-O-glucoside 2''-O-glucosyltransferase-like* genes (No. 41 and 60) had high weight values with each other, indicating that they were highly related (Figure 5E).

In the grey 60 module, *MYB049* (No. 95) had a high weight value with *protein phosphatase 2C family protein* genes (No. 16 and 18). *MYB049* (No. 95) also had a high weight value with *bZIP transcription factor 6 family protein* (No. 9), which was the gene with the highest IC in the network and had a high weight value with *protein phosphatase 2C family protein* genes (No. 16 and No. 18, Figure 5F). Additionally, *protein phosphatase 2C family protein* genes (Nos. 16, 18, 96, and 145), *mitogen-activated protein kinase homologues* (No. 68), and *calcium binding family protein* (No. 138) were related to both *MYB049* (No. 95) and *bZIP transcription factor 6 family protein* (No. 9), indicating that they were regulated by both MYB and bZIP.

3.5. Analysis of Interactions between R2R3-MYB Genes, IAA, and Free SA

Many genes in these modules involved in the interaction between poplars and E4 were related to auxin. In the yellow module, *aux/IAA family protein* (No. 1494) and *auxin-responsive GH3 family protein* (No. 1583) were related to 31 genes (Figure S6A). In the light cyan module, *auxin-responsive family protein* (No. 33) was related to 24 genes (Figure S6B). In the black module, *auxin-responsive family protein* genes (Nos. 40 and 488) and *auxin-induced protein IAA4* were related to 10 genes (Figure S6C). This result suggested that R2R3-MYBs could be involved in the interaction between poplars and E4 by regulating the expression of genes associated with IAA.

In the light cyan module, compared to other genes, two *receptor-like protein 12 isoform X9* genes (Nos. 85 and 97) had higher GSs with IAA (>0.20). In the black module, GSs for all auxin-related genes with IAA were negative or very low. However, in the yellow module, six genes (Nos. 304, 371, 505, 543, 602, and 1583) had higher GSs with IAA, and there was a certain correlation between them (Figure 6A,B). A *MYB194* (No. 371) was included here. There was a high correlation between genes in the yellow module and free SA (Figure 6B). These results suggested that *MYB194* (No. 371) was involved in regulating IAA and free SA changes in the poplar interaction with E4 and affecting the defence response of poplar.
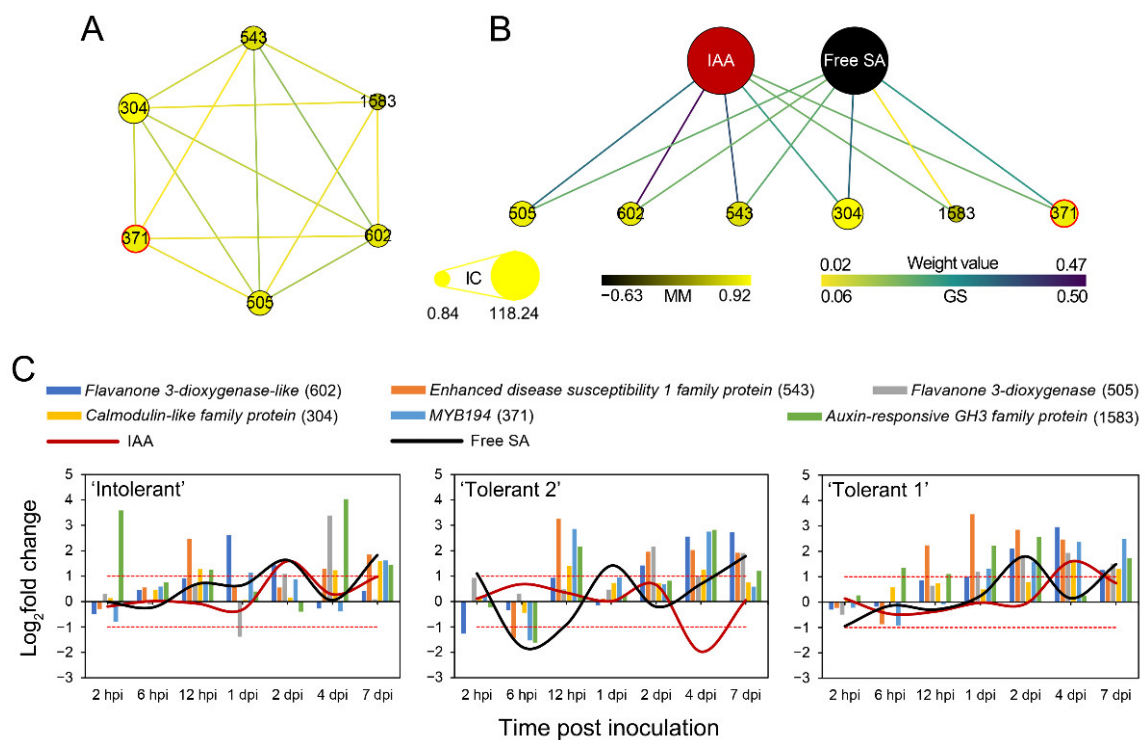
Figure 6. Coexpression characteristics of auxin-related genes. (**A**) Coexpression relationship between auxin-related genes. (**B**) Correlation between auxin-related genes and auxin (indole-3-acetic acid, IAA) and free salicylic acid (free SA). The size of the dots represents IC. The colour of the dot represents MM. The colour of the line represents the weight value between two genes in (**A**) and the GS for IAA or free SA in (**B**). The label of the dot is listed based on IC for the specific gene in the specific module. The R2R3-MYB genes are highlighted with a red border. GS, gene significance; IC, intramodular connectivity. (**C**) Comparison of the changes in the expression ($\log_2$E4-inoculated/E4-free) of auxin-related genes and levels of IAA and free SA.

Comparison of the expression changes in these auxin-related genes and levels of IAA and free SA indicated that changes in gene expression should play a role in affecting the levels of IAA and free SA (Figure 6B,C). In 'Intolerant', the changes in the levels of IAA and free SA showed the same trend, while in 'Tolerant 2' and 'Tolerant 1', the changes in the levels of IAA and free SA showed the opposite trend, which indicated completely different strategies between tolerant and intolerant poplars. On the other hand, 'Tolerant 2' and 'Tolerant 1' also likely adopt different strategies in terms of IAA and free SA-related tolerance as they changed differently at different time points. These results suggested that interactions between IAA and free SA were significantly different in the tolerant and intolerant poplars and that IAA and free SA acted in different ways in poplars with different tolerances.

3.6. Expression Characteristics of R2R3-MYB Gene-Related Genes Involved in the Poplar–E4 Interaction

To further explore the differences between tolerant and intolerant poplars in defence-related genes after E4 infection, the expression of *R2R3-MYB* gene-related genes involved in poplar–E4 interactions was investigated. RT-qPCR was performed to validate the changes in transcript levels of the genes. In the yellow module, we found that the expression of many genes was downregulated in the middle and late stages of E4 infection in 'Intolerant'. For example, *enhanced disease susceptibility 1 family protein* (*EDS1*; No. 543) was continuously upregulated after 2 dpi in 'Tolerant 2' and 'Tolerant 1'. However, although the expression of *EDS1* was upregulated at 4 and 7 dpi, there was no significant change in the expression of the gene at 2 dpi in 'Intolerant'. Similar situations occurred for the *pathogenesis-related protein*

(No. 179), *calcium-binding family protein* (No. 196), and *NBS-LRR type disease resistance protein* (No. 516). Some *pathogenesis-related family protein* genes (*PR*s; Nos. 164, 141, 237, and 626) were continuously upregulated after 12 hpi in 'Tolerant 1' and continuously upregulated after 2 dpi in 'Tolerant 2'. However, there were no significant changes in the expression of these *PR*s at 4 dpi in 'Intolerant'. The upregulations of *WRKY transcription factor 72* (No. 305), *calcium-dependent protein kinase* (No. 341), *peroxidase family protein* (No. 408), calmodulin-like protein (No. 699), and *NBS-LRR type disease resistance protein* (No. 1208) did not happen at 4 dpi in 'Intolerant'. The upregulations of *Leucine-rich repeat receptor-like protein kinase* genes (Nos. 380 and 386) and *disease resistance protein* genes (Nos. 428 and 506) did not happen at 2 and 4 dpi in 'Intolerant'. In addition, the upregulations of *Cytochrome P450 family protein* (No. 207), *flavonol synthase/flavanone 3-hydroxylase-like* (No. 276), *peroxidase family protein* (No. 315), *WRKY transcription factor 47 family protein* (No. 509), and *flavanone 3-dioxygenase-like* (No. 602) did not happen at 4 and 7 dpi in 'Intolerant'. However, the upregulation of *auxin-responsive GH3 family protein* (No. 1583) did not happen at 4 dpi, and the downregulation of *aux/IAA family protein* (No. 1494) happened at 4 and 7 dpi in 'Tolerant 1' (Figures S7 and S13 and Table S10).

In the light cyan module, *LRR receptor-like serine/threonine-protein kinase* genes (Nos. 27 and 30), *receptor-like protein* genes (Nos. 43, 85, and 97), and *chitinase* (No. 317) were only downregulated at 2 dpi and/or 4 dpi in 'Intolerant' (Figures S8 and S14 and Table S10). In the black module, *cytochrome P450 family protein* (No. 258) and *PR* (No. 766) were only downregulated at 4 dpi in 'Intolerant' (Figures S9 and S15 and Table S10). In the purple module, *leucine-rich repeat family protein* (No. 1) and *WRKY transcription factor 42 family protein* (No. 206) were only not upregulated at 2 and 4 dpi in 'Intolerant' (Figures S10 and S16 and Table S10). In the dark grey module, *phenylalanine ammonia-lyase family protein* (No. 24) and *cinnamyl alcohol dehydrogenase 6* (No. 11) were only upregulated at 12 hpi and 1 dpi in 'Intolerant' (Figures S11 and S17 and Table S10). In the grey 60 module, most genes were only upregulated at 4 dpi in 'Intolerant', including *protein phosphatase* genes (Nos. 16, 18, 96, 145, and 146), *bZIP transcription factor 6 family protein* (No. 9), and *calcium binding family protein* (No. 138, Figures S12 and S18 and Table S10).

These results suggested that the continuous expression of some defence-related genes in the middle and late stages of E4 infection could play an important role in the timely development of the defence response in tolerant poplars. Additionally, genes involved in flavonoid biosynthesis, phenylpropanoid biosynthesis, and anthocyanin biosynthesis were likely to play an important role in poplar defence, but their responses might be delayed or suppressed in 'Intolerant' in the late stages of E4 infection (Figure 7, Figures S13–S18 and Table S10).
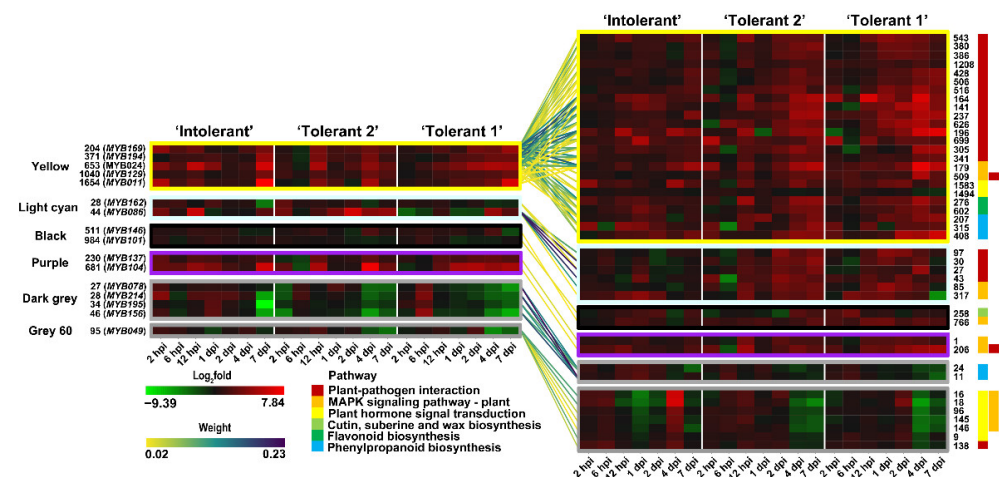

Figure 7. Expression characteristics of *R2R3-MYB* gene-related genes (detailed in Table S10) involved in poplar–E4 interactions.

4. Discussion

In this study, 217 R2R3-MYB-encoding genes were identified by transcriptome sequencing based on the genome sequence of *P. trichocarpa* [24,27]. The transcript profiles of two tolerant poplars ('Tolerant 1' and 'Tolerant 2') and an intolerant poplar ('Intolerant') at different time points after E4 inoculation were investigated to study the roles of *R2R3-MYB* genes and their related genes in the interaction between poplar and E4. Our previous study indicated that the expression of most TF genes of 'Intolerant' did not react to E4 infection, particularly many defence-related TF genes, including *MYB* genes [3,13,14]. The susceptibility of 'Intolerant' may be related to the lack of response of most TFs at the infection phase of E4 [13]. Here, we found that the expression of *R2R3-MYB* genes and their related genes varied greatly at different time points in different poplars, revealing the special role of R2R3-MYBs in poplar defence against E4 infection.

By weighted gene coexpression network analysis, we found that 83 *R2R3-MYB* genes were assigned to 22 different coexpression modules, which had different expression patterns. Among them, 25 *R2R3-MYB* genes were central elements in their respective modules. However, it was found that the expression of most *R2R3-MYB* genes did not change significantly in both tolerant and intolerant poplars in the early period of E4 infection (2 hpi to 1 dpi), suggesting that the defence response was not actively mobilised or that only a few R2R3-MYBs were needed in the early interaction. Then, at 2 and 4 dpi, there were almost no differentially expressed *R2R3-MYB* genes in 'Intolerant', but there was an obvious increase in differentially expressed genes in 'Tolerant 1' and 'Tolerant 2', indicating that the response of some *R2R3-MYB* genes at this period played an important role in the defence response in tolerant poplars.

The major stages of *M. larici-populina* infection include germination and penetration, early colonisation of plant tissue, colonisation of plant mesophyll, and uredinia formation [14,32,33]. In the compatible interaction, stomatal penetration normally occurs from 2 to 6 hpi after the germination of urediniospores, followed by substomatal vesicle formation from 6 to 12 hpi and the development of infectious hyphae from 12 hpi to 1 dpi. Then, the formation of haustoria occurs from 1 to 2 dpi, and dense infection hyphae and haustorial networks grow from 2 to 4 dpi. By 4 dpi, the whole plant mesophyll is colonised by infection structures, and differentiation of the first sporogenous hyphae is observed at this time point [14,32,34]. Therefore, 2–4 dpi can be considered the important biotrophic growth period of E4 during its infection in 'Intolerant'. As an obligate biotrophic fungus, *M. larici-populina* must produce haustoria to derive nutrients from the host to achieve spore production and sporulation and to suppress host defences, enabling its proliferation between 1 and 4 dpi [34,35]. 'Intolerant' should have responded to rust invasion, as several *R2R3-MYB* genes, some other TF genes, and defence-related genes were differentially expressed during the infection phase [13,14,36]. Therefore, E4 must overcome plant surveillance systems in 'Intolerant' [37], and the manipulation of plant defence mechanisms in poplar is most likely to begin at 2 dpi. Differences in gene expression between poplars at 2–4 dpi should be related to differences in their susceptibility.

Many *R2R3-MYB* gene-related genes are involved in plant–pathogen interaction- and signalling-related pathways. We concentrated on genes with different expression patterns between tolerant and intolerant poplars at 2 and 4 dpi. In total, 15 of the 22 modules had 34 *R2R3-MYB* genes with differential expression at 2 and 4 dpi between tolerant and intolerant poplars. Of the 34 differentially expressed *R2R3-MYB* genes, 11 had distinct expression characteristics in 'Tolerant 1', which had confluent necrosis at 7 dpi [3]. Six of these *R2R3-MYB* genes were assigned to the yellow module and showed a gradual increase in expression with increasing inoculation time in 'Tolerant 1'. The other 5 *R2R3-MYB* genes were assigned to the dark grey module and showed a gradual decrease in expression after 6 hpi with the increase in inoculation time in 'Tolerant 1'. The largest difference between 'Tolerant 1' and the other two poplars was that 'Tolerant 1' gradually showed programmed cell death (PCD) in the late stage of infection [3]. Therefore, these results suggested that

these *R2R3-MYB* genes might be associated with the regulation of PCD in 'Tolerant 1' after E4 infection [38–42]. However, this correlation needs to be further verified.

KEGG enrichment analysis of genes related to differentially expressed *R2R3-MYB* genes showed that different *R2R3-MYB* genes were involved in regulating different levels of defence-related responses (Figure 8). We focused on modules with more genes enriched in plant–pathogen interaction- and signalling-related pathways, as well as pathways related to the biosynthesis of secondary metabolites that are important in plant stress resistance. The groups of genes associated with defence in different modules were quite different, indicating that different R2R3-MYBs participated in different defence-related processes. It is generally accepted that eukaryotic genes are regulated by more than one TF and that their target genes are also dependent on several TFs [43]. The regulation of plant tolerance to disease is complex, with a number of TF families playing important roles [44]. Here, we found coregulatory relationships between R2R3-MYBs. Additionally, other TFs, such as the AP2 domain-containing transcription factors bZIP and WRKY, were coregulated with R2R3-MYBs.
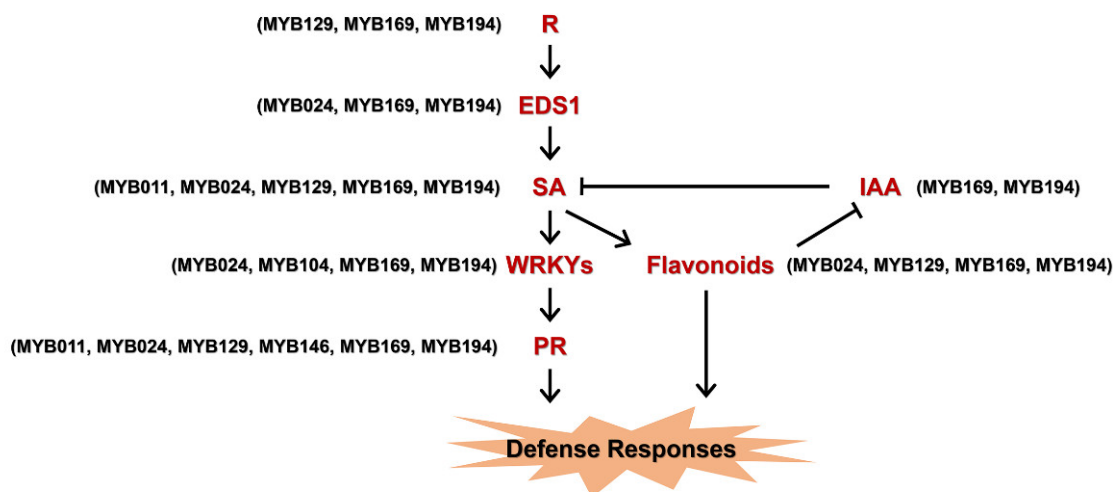
Figure 8. Model of how R2R3-MYBs are involved in regulating defence-related responses. R, disease resistance protein; EDS1, enhanced disease susceptibility 1 family protein; SA, salicylic acid; IAA, indole-3-acetic acid; WRKY, WRKY transcription factor; PR, pathogenesis-related family protein.

We analysed the expression patterns of all differentially expressed *R2R3-MYB* gene-related genes and identified 43 defence-related genes with significant differences between tolerant and intolerant poplars. These genes related to 16 different *R2R3-MYB* genes, most of which were assigned to the yellow module, which was related to free SA [3], followed by the dark grey module, suggesting that many differentially expressed *R2R3-MYB* genes and *R2R3-MYB* gene-related genes might be associated with free SA. Among them, *MYB169* had the highest MM (0.92) and IC (72.81), indicating that it had regulatory relationships with more genes within the module and therefore played an essential regulatory role. Most of the differentially expressed genes were enriched in plant–pathogen interactions and were expressed at low levels or downregulated in 'Intolerant' at 2 and/or 4 dpi. These genes included *enhanced disease susceptibility 1 family protein*, *leucine-rich repeat receptor-like protein kinase*, *disease resistance protein*, *pathogenesis-related family protein*, and *calcium-binding family protein*. These results suggested that most of the differentially expressed *R2R3-MYB* genes and their associated genes played a positive role in the interaction between poplars and E4 to improve defence.

However, genes related to *MYB049* were mainly enriched in plant hormone signal transduction and the MAPK signalling pathway, and their significant downregulation at 1 dpi and upregulation at 4 dpi in 'Intolerant' might have been manipulated by E4, as these two time points were exactly the time points when E4 began and ended bioaccumulation [34]. This suggested that E4 might no longer inhibit the transmission of defensive

signals after colonisation, possibly to maintain the physiological state of the leaf and enable it to successfully complete its life cycle [3]. On the other hand, the downregulated expression of these genes in tolerant poplars might weaken signal transmission and reduce the involvement of the defence response in the late stage of E4 infection [3].

Some *flavanone 3-hydroxylase* (also named *flavanone 3-dioxygenase*) genes were downregulated in 'Intolerant' at 2 and/or 4 dpi. Flavanone 3-hydroxylase is involved in the flavonoid biosynthesis pathway, which is part of secondary metabolite biosynthesis. Flavonoids are widely distributed in plants and play important roles. Plants promote the accumulation of flavonoids under stress conditions, resulting in the production of compounds that have, for example, antimicrobial activity (phytoalexins), thereby protecting themselves [45–47]. Overexpression of the *flavanone 3-hydroxylase* gene confers tolerance to both biotic and abiotic stresses, and flavanone 3-hydroxylase induction significantly promoted SA and inhibited JA accumulation [48,49]. The R2R3-MYB family has been demonstrated to act as the main flavonoid biosynthesis regulator in many plant species [50–55]. The ME of the yellow module to which these *flavanone 3-hydroxylase* genes were assigned was associated with free SA [3], so their related R2R3-MYBs (MYB011, MYB024, MYB129, MYB169, and MYB194) might play an important regulatory role in flavonoid biosynthesis and free SA signal transduction, and the accumulation of flavonoids and free SA should be highly correlated with the defence response of poplar to E4 infection.

However, *phenylalanine ammonia-lyase family protein* and *cinnamyl alcohol dehydrogenase 6* were only upregulated at 12 hpi and 1 dpi in 'Intolerant'. A pathogen contacting the plant cell wall is the first signal that triggers the phenylpropanoid pathway for plant defence [56]. Recently, some key players, such as phenylalanine ammonia-lyases and cinnamyl alcohol dehydrogenase from the phenylpropanoid pathway, have been proposed to have broad-spectrum disease resistance [56–59]. Their gene expression is regulated by R2R3-MYBs [60–62]. The upregulation of *phenylalanine ammonia-lyase family protein* and *cinnamyl alcohol dehydrogenase 6* at the infective hyphae development period suggested that in the early developmental stages of the E4 hyphae, some defence-related metabolites were produced in the intolerant poplar, but these substances might not be key defensive substances and were not enough to inhibit further E4 infection.

Although the genes associated with *R2R3-MYB* genes varied greatly from one module to another, we found that some of the genes associated with *R2R3-MYB* genes in the different modules were associated with IAA, namely, *auxin-responsive Gretchen Hagen3 (GH3) family protein*, *aux/IAA family protein*, *auxin-responsive family protein*, and *auxin-induced protein IAA4*. Auxin has long been recognised as a regulator of plant defence [63]. Auxin biosynthesis, transport, and signalling antagonise SA biosynthesis and signalling that is required for resistance to biotrophic pathogens [63]. The changes in free SA and IAA in 'Intolerant' showed the same trend all the time, suggesting that IAA might have been inhibiting the accumulation of free SA. In 'Tolerant 1' and 'Tolerant 2', the free SA and IAA contents showed a more complementary pattern, suggesting that IAA would decrease at some time point to promote SA-dominated defence responses. This difference might eventually lead to different susceptibilities of poplar to E4 infection.

Plants can quickly sense and respond to changes in auxin levels, and these responses involve several major classes of auxin-responsive genes, including the *auxin/indole-3-acetic acid* family, the *auxin response factor* family, *small auxin upregulated RNA*, and the *auxin-responsive GH3 family* [64]. This suggested that there might be multiple R2R3-MYBs involved in IAA-related pathways, and the differential expression of IAA-related genes might be associated with the occurrence of defence responses in poplar–E4 interactions. Among them, the *R2R3-MYB* gene with the highest correlation with IAA levels was *MYB194*, which was assigned to the yellow module. Genes that were related to *MYB194* in the yellow module and had high relationships with IAA levels included *flavanone 3-dioxygenase*, *enhanced disease susceptibility 1 family protein*, *calmodulin-like family protein*, and *auxin-responsive GH3 family protein*. As one of the three major auxin-responsive families, the auxin-responsive GH3 family maintains hormonal homeostasis by conjugating excess IAA, SA, and JA to amino

acids during hormone and stress-related signalling [65–68]. Our previous study found that these genes were also related to free SA [3], which is essential in poplar–E4 interactions and defence responses against E4, indicating that MYB194 might be an important node for the convergence of IAA and SA signalling. Flavonoids act as endogenous negative regulators of auxin transport [69] and are possibly involved in SA-related stress signalling [70]. Therefore, R2R3-MYBs were involved in multiple processes from IAA or free SA signal transduction to flavonoid biosynthesis during poplar–E4 interactions (Figure 8).

Considerable interest exists in identifying and utilising key TFs in plant defence to engineer increased resistance to plant pathogens. A comprehensive analysis of the physiological functions and biological roles of the R2R3-MYB family and their related genes in tolerant and intolerant poplars when under attack by E4 is required to fully describe the R2R3-MYB family, and such an analysis will provide rich resources and opportunities to understand rust tolerance in poplar and to screen effective *R2R3-MYB* genes for utilisation of transgenic technology to improve poplar resistance to *M. larici-populina*.

5. Conclusions

In this study, 217 R2R3-MYBs were identified and 83 *R2R3-MYB* genes were assigned to 22 different coexpression modules. Most *R2R3-MYB* genes were unchanged in the early period of E4 infection (2 hpi to 1 dpi) in both tolerant and intolerant poplars. However, there were obvious increases in differentially expressed *R2R3-MYB* genes in 'Tolerant 1' and 'Tolerant 2' later at 2 and 4 dpi, which was an important biotrophic growth period of E4 during its infection of 'Intolerant' and the period when 'Tolerant 1' and 'Tolerant 2' developed hypersensitive cell death responses at the infection sites. These results suggested that the expression of *R2R3-MYB* genes is associated with the occurrence of defence responses, and differently expressed *R2R3-MYB* genes at 2 and 4 dpi between tolerant and intolerant poplars may play an important role in poplar resistance to E4 infection. In total, 34 *R2R3-MYB* genes showed differential expression at 2 and 4 dpi between tolerant and intolerant poplars, and they may participate in different defence-related processes. Among them, 16 differentially expressed *R2R3-MYB* genes were related to 43 defence-related genes that had significant differences between tolerant and intolerant poplars. There might be coregulatory relationships between R2R3-MYBs and other TFs during poplar–E4 interaction. Some differentially expressed *R2R3-MYB* genes were related to genes involved in flavonoid biosynthesis and IAA or free SA signal transduction and might help activate defence response during poplar–E4 interaction. MYB194 could be an important node in the convergence of IAA and SA signalling.

RT–qPCR; Figure S16: Expression validation of genes in purple module by RT–qPCR; Figure S17: Expression validation of genes in dark grey module by RT–qPCR; Figure S18: Expression validation of genes in grey 60 module by RT–qPCR; Table S1: Primers used in this study; Table S2: Quality statistics of filtered reads; Table S3: Basic information of *R2R3-MYB* genes; Table S4: Number of *R2R3-MYB* genes in each module; Table S5. MMs and ICs of *R2R3-MYB* genes in each module; Table S6: Top 10 pathways for genes that *R2R3-MYB* genes had weight values within each module; Table S7: Differentially expressed *R2R3-MYB* genes; Table S8: Top 10 pathways for genes that differently expressed *R2R3-MYB* genes had weight values with in each module; Table S9: Annotation of differently expressed *R2R3-MYB* genes-related genes that involved in poplar–E4 interaction in the selected modules; Table S10: Differently expressed *R2R3-MYB* genes and their related defence-related genes.

References

1. Pinon, J.; Van Dam, B.C.; Genetet, I.; De Kam, M. Two pathogenic races of *Melampsora larici-populina* in Northwestern Europe. *Eur. J. Forest Pathol.* **1987**, *17*, 47–53. [CrossRef]
2. Pei, M.; Ruiz, C.; Bayon, C.; Hunter, T.; Lonsdale, D. Pathogenic variation in poplar rust *Melampsora larici-populina* from England. *Eur. J. Plant Pathol.* **2005**, *111*, 147–155. [CrossRef]
3. Chen, Q.; Zhang, R.; Li, D.; Wang, F. Integrating transcriptome and coexpression network analyses to characterize salicylic acid- and jasmonic acid-related genes in tolerant poplars infected with rust. *Int. J. Mol. Sci.* **2021**, *22*, 5001. [CrossRef]
4. Zhang, S.; Klessig, D.F. MAPK cascades in plant defense signaling. *Trends Plant Sci.* **2001**, *6*, 520–527. [CrossRef]
5. Meng, X.; Zhang, S. MAPK cascades in plant disease resistance signaling. *Annu. Rev. Phytopathol.* **2013**, *51*, 245–266. [CrossRef]
6. Thulasi Devendrakumar, K.; Li, X.; Zhang, Y. MAP kinase signalling: Interplays between plant PAMP- and effector-triggered immunity. *Cell. Mol. Life Sci.* **2018**, *75*, 2981–2989. [CrossRef]
7. Boudsocq, M.; Sheen, J. CDPKs in immune and stress signaling. *Trends Plant Sci.* **2013**, *18*, 30–40. [CrossRef]
8. Ng, D.W.; Abeysinghe, J.K.; Kamali, M. Regulating the regulators: The control of transcription factors in plant defense signaling. *Int. J. Mol. Sci.* **2018**, *19*, 3737. [CrossRef]
9. Chen, W.; Provart, N.J.; Glazebrook, J.; Katagiri, F.; Chang, H.S.; Eulgem, T.; Mauch, F.; Luan, S.; Zou, G.; Whitham, S.A.; et al. Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* **2002**, *14*, 559–574. [CrossRef]
10. Nimchuk, Z.; Eulgem, T.; Holt III, B.F.; Dangl, J.L. Recognition and response in the plant immune system. *Annu. Rev. Genet.* **2003**, *37*, 579–609. [CrossRef]
11. Glazebrook, J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.* **2005**, *43*, 205–227. [CrossRef] [PubMed]
12. Pieterse, C.M.J.; Leon-Reyes, A.; Van der Ent, S.; Van Wees, S.C.M. Networking by small-molecule hormones in plant immunity. *Nat. Chem. Biol.* **2009**, *5*, 308–316. [CrossRef]
13. Chen, Q.; Wang, J.; Li, D.; Wang, Z.; Wang, F.; Zhang, R. Molecular characterization of the transcription factors in susceptible poplar infected with virulent *Melampsora larici-populina*. *Int. J. Mol. Sci.* **2019**, *20*, 4806. [CrossRef] [PubMed]
14. Chen, Q.; Wang, Z.; Li, D.; Wang, F.; Zhang, R.; Wang, J. Molecular characterization of the ERF family in susceptible poplar infected by virulent *Melampsora larici-populina*. *Physiol. Mol. Plant Pathol.* **2019**, *108*, 101437. [CrossRef]
15. Martin, C.; Paz-Ares, J. MYB transcription factors in plants. *Trends Genet.* **1997**, *13*, 67–73. [CrossRef]
16. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **2010**, *15*, 573–581. [CrossRef]
17. Liu, J.; Osbourn, A.; Ma, P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol. Plant* **2015**, *8*, 689–708. [CrossRef]
18. Chezem, W.R.; Memon, A.; Li, F.S.; Weng, J.K.; Clay, N.K. SG2-type R2R3-MYB transcription factor MYB15 controls defense-induced lignification and basal immunity in *Arabidopsis*. *Plant Cell* **2017**, *29*, 1907–1926. [CrossRef]

19. Lee, M.W.; Qi, M.; Yang, Y. A novel jasmonic acid-inducible rice myb gene associates with fungal infection and host cell death. *Mol. Plant Microbe Interact.* **2001**, *14*, 527–535. [CrossRef]

20. Zhang, Z.; Liu, X.; Wang, X.; Zhou, M.; Zhou, X.; Ye, X.; Wei, X. An R2R3 MYB transcription factor in wheat, TaPIMP1, mediates host resistance to *Bipolaris sorokiniana* and drought stresses through regulation of defense- and stress-related genes. *New Phytol.* **2012**, *196*, 1155–1170. [CrossRef]

21. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef] [PubMed]

22. Wang, R.K.; Cao, Z.H.; Hao, Y.J. Overexpression of a R2R3 MYB gene *MdSIMYB1* increases tolerance to multiple stresses in transgenic tobacco and apples. *Physiol. Plant* **2014**, *150*, 76–87. [CrossRef] [PubMed]

23. Li, D.; Wang, F.; Wang, C.; Zou, L.; Wang, Z.; Chen, Q.; Niu, C.; Zhang, R.; Ling, Y.; Wang, B. MicroRNA-mediated susceptible poplar gene expression regulation associated with the infection of virulent *Melampsora larici-populina*. *BMC Genom.* **2016**, *17*, 59. [CrossRef] [PubMed]

24. Tuskan, G.A.; Difazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **2006**, *313*, 1596–1604. [CrossRef]

25. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef]

26. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]

27. Wilkins, O.; Nahal, H.; Foong, J.; Provart, N.J.; Campbell, M.M. Expansion and diversification of the Populus R2R3-MYB family of transcription factors. *Plant Physiol.* **2009**, *149*, 981–993. [CrossRef]

28. 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]

29. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]

30. Pan, X.; Welti, R.; Wang, X. Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatography-mass spectrometry. *Nat. Protoc.* **2010**, *5*, 986–992. [CrossRef]

31. 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] [PubMed]

32. Duplessis, S.; Hacquard, S.; Delaruelle, C.; Tisserant, E.; Frey, P.; Martin, F.; Kohler, A. *Melampsora larici-populina* transcript profiling during germination and timecourse infection of poplar leaves reveals dynamic expression patterns associated with virulence and biotrophy. *Mol. Plant Microbe Interact.* **2011**, *24*, 808–818. [CrossRef] [PubMed]

33. Hacquard, S.; Petre, B.; Frey, P.; Hecker, A.; Rouhier, N.; Duplessis, S. The poplar-poplar rust interaction: Insights from genomics and transcriptomics. *J. Pathog.* **2011**, *2011*, 716041. [CrossRef] [PubMed]

34. Hacquard, S.; Delaruelle, C.; Legué, V.; Tisserant, E.; Kohler, A.; Frey, P.; Martin, F.; Duplessis, S. Laser capture microdissection of uredinia formed by *Melampsora larici-populina* revealed a transcriptional switch between biotrophy and sporulation. *Mol. Plant Microbe Interact.* **2010**, *23*, 1275–1286. [CrossRef]

35. Rinaldi, C.; Kohler, A.; Frey, P.; Duchaussoy, F.; Ningre, N.; Couloux, A.; Wincker, P.; Thiec, D.L.; Fluch, S.; Martin, F.; et al. Transcript profiling of poplar leaves upon infection with compatible and incompatible strains of the foliar rust *Melampsora larici-populina*. *Plant Physiol.* **2007**, *144*, 347–366. [CrossRef] [PubMed]

36. Hamel, L.P.; Benchabane, M.; Major, I.T.; Nicole, M.C.; Sheen, J.; Séguin, A. The genomics of poplar-rust interactions to improve tree resistance against fungal disease. *BMC Proc.* **2011**, *5*, 12. [CrossRef]

37. Dodds, P.N.; Rathjen, J.P. Plant immunity: Towards an integrated view of plant–pathogen interactions. *Nat. Rev. Genet.* **2010**, *11*, 539–548. [CrossRef]

38. Vailleau, F.; Daniel, X.; Tronchet, M.; Montillet, J.L.; Triantaphylidès, C.; Roby, D. A R2R3-MYB gene, AtMYB30, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10179–10184. [CrossRef]

39. Peng, S.Q.; Wu, K.X.; Huang, G.X.; Chen, S.C. HbMyb1, a Myb transcription factor from *Hevea brasiliensis*, suppresses stress induced cell death in transgenic tobacco. *Plant Physiol. Biochem.* **2011**, *49*, 1429–1435. [CrossRef]

40. Kaurilind, E.; Xu, E.; Brosché, M. A genetic framework for H_2O_2 induced cell death in *Arabidopsis thaliana*. *BMC Genom.* **2015**, *16*, 837. [CrossRef]

41. Chen, B.; Niu, F.; Liu, W.Z.; Yang, B.; Zhang, J.; Ma, J.; Cheng, H.; Han, F.; Jiang, Y.Q. Identification, cloning and characterization of R2R3-MYB gene family in canola (*Brassica napus* L.) identify a novel member modulating ROS accumulation and hypersensitive-like cell death. *DNA Res.* **2016**, *23*, 101–114. [CrossRef] [PubMed]

42. Yao, L.; Yang, B.; Xian, B.; Chen, B.; Yan, J.; Chen, Q.; Gao, S.; Zhao, P.; Han, F.; Xu, J.; et al. The R2R3-MYB transcription factor BnaMYB111L from rapeseed modulates reactive oxygen species accumulation and hypersensitive-like cell death. *Plant Physiol. Biochem.* **2020**, *147*, 280–288. [CrossRef] [PubMed]

43. Sorrells, T.R.; Johnson, A.D. Making sense of transcription networks. *Cell* **2015**, *161*, 714–723. [CrossRef] [PubMed]

44. Singh, K.; Foley, R.C.; Oñate-Sánchez, L. Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* **2002**, *5*, 430–436. [CrossRef]

45. Treutter, D. Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol.* **2005**, *7*, 581–591. [CrossRef]

46. Pourcel, L.; Routaboul, J.M.; Cheynier, V.; Lepiniec, L.; Debeaujon, I. Flavonoid oxidation in plants: From biochemical properties to physiological functions. *Trends Plant Sci.* **2007**, *12*, 29–36. [CrossRef]

47. Hernández, I.; Alegre, L.; Van Breusegem, F.; Munné-Bosch, S. How relevant are flavonoids as antioxidants in plants? *Trends Plant Sci.* **2009**, *14*, 125–132. [CrossRef]

48. Mahajan, M.; Yadav, S.K. Overexpression of a tea flavanone 3-hydroxylase gene confers tolerance to salt stress and *Alternaria solani* in transgenic tobacco. *Plant Mol. Biol.* **2014**, *85*, 551–573. [CrossRef]

49. Jan, R.; Aaqil Khan, M.; Asaf, S.; Lubna; Park, J.R.; Lee, I.J.; Kim, K.M. Flavonone 3-hydroxylase relieves bacterial leaf blight stress in rice via overaccumulation of antioxidant flavonoids and induction of defense genes and hormones. *Int. J. Mol. Sci.* **2021**, *22*, 6152. [CrossRef]

50. Espley, R.V.; Hellens, R.P.; Putterill, J.; Stevenson, D.E.; Kutty-Amma, S.; Allan, A.C. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* **2007**, *49*, 414–427. [CrossRef]

51. Stracke, R.; Ishihara, H.; Huep, G.; Barsch, A.; Mehrtens, F.; Niehaus, K.; Weisshaar, B. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* **2007**, *50*, 660–677. [CrossRef] [PubMed]

52. Zhou, H.; Lin-Wang, K.; Wang, H.; Gu, C.; Dare, A.P.; Espley, R.V.; He, H.; Allan, A.C.; Han, Y. Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. *Plant J.* **2015**, *82*, 105–121. [CrossRef] [PubMed]

53. Liu, C.; Long, J.; Zhu, K.; Liu, L.; Yang, W.; Zhang, H.; Li, L.; Xu, Q.; Deng, X. Characterization of a citrus R2R3-MYB transcription factor that regulates the flavonol and hydroxycinnamic acid biosynthesis. *Sci. Rep.* **2016**, *6*, 25352. [CrossRef] [PubMed]

54. Wang, N.; Liu, W.; Zhang, T.; Jiang, S.; Xu, H.; Wang, Y.; Zhang, Z.; Wang, C.; Chen, X. Transcriptomic analysis of red-fleshed apples reveals the novel role of MdWRKY11 in flavonoid and anthocyanin biosynthesis. *J. Agric. Food Chem.* **2018**, *66*, 7076–7086. [CrossRef]

55. Zhai, R.; Zhao, Y.; Wu, M.; Yang, J.; Li, X.; Liu, H.; Wu, T.; Liang, F.; Yang, C.; Wang, Z.; et al. The MYB transcription factor PbMYB12b positively regulates flavonol biosynthesis in pear fruit. *BMC Plant Biol.* **2019**, *19*, 85. [CrossRef]

56. Yadav, V.; Wang, Z.; Wei, C.; Amo, A.; Ahmed, B.; Yang, X.; Zhang, X. Phenylpropanoid pathway engineering: An emerging approach towards plant defense. *Pathogens* **2020**, *9*, 312. [CrossRef]

57. Tronchet, M.; Balagué, C.; Kroj, T.; Jouanin, L.; Roby, D. Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in *Arabidopsis*. *Mol. Plant Pathol.* **2010**, *11*, 83–92. [CrossRef]

58. Rong, W.; Luo, M.; Shan, T.; Wei, X.; Du, L.; Xu, H.; Zhang, Z. A wheat cinnamyl alcohol dehydrogenase TaCAD12 contributes to host resistance to the sharp eyespot disease. *Front. Plant Sci.* **2016**, *7*, 1723. [CrossRef]

59. Liu, X.; Cui, X.; Ji, D.; Zhang, Z.; Li, B.; Xu, Y.; Chen, T.; Tian, S. Luteolin-induced activation of the phenylpropanoid metabolic pathway contributes to quality maintenance and disease resistance of sweet cherry. *Food Chem.* **2021**, *342*, 128309. [CrossRef]

60. Ma, Q.H.; Wang, C.; Zhu, H.H. TaMYB4 cloned from wheat regulates lignin biosynthesis through negatively controlling the transcripts of both cinnamyl alcohol dehydrogenase and cinnamoyl-CoA reductase genes. *Biochimie* **2011**, *93*, 1179–1186. [CrossRef]

61. Craven-Bartle, B.; Pascual, M.B.; Cánovas, F.M.; Avila, C. A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *Plant J.* **2013**, *74*, 755–766. [CrossRef] [PubMed]

62. He, J.; Liu, Y.; Yuan, D.; Duan, M.; Liu, Y.; Shen, Z.; Yang, C.; Qiu, Z.; Liu, D.; Wen, P.; et al. An R2R3 MYB transcription factor confers brown planthopper resistance by regulating the phenylalanine ammonia-lyase pathway in rice. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 271–277. [CrossRef] [PubMed]

63. Kazan, K.; Manners, J.M. Linking development to defense: Auxin in plant-pathogen interactions. *Trends Plant Sci.* **2009**, *14*, 373–382. [CrossRef] [PubMed]

64. Luo, J.; Zhou, J.J.; Zhang, J.Z. Aux/IAA gene family in plants: Molecular structure, regulation, and function. *Int. J. Mol. Sci.* **2018**, *19*, 259. [CrossRef]

65. Staswick, P.E.; Tiryaki, I.; Rowe, M.L. Jasmonate response locus JAR1 and several related *Arabidopsis* genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. *Plant Cell* **2002**, *14*, 1405–1415. [CrossRef]

66. Nobuta, K.; Okrent, R.A.; Stoutemyer, M.; Rodibaugh, N.; Kempema, L.; Wildermuth, M.C.; Innes, R.W. The GH3 acyl adenylase family member PBS3 regulates salicylic acid-dependent defense responses in *Arabidopsis*. *Plant Physiol.* **2007**, *144*, 1144–1156. [CrossRef]

67. Zhang, Z.; Li, Q.; Li, Z.; Staswick, P.E.; Wang, M.; Zhu, Y.; He, Z. Dual regulation role of GH3.5 in salicylic acid and auxin signaling during *Arabidopsis-Pseudomonas syringae* interaction. *Plant Physiol.* **2007**, *145*, 450–464. [CrossRef]

68. Jiang, W.; Yin, J.; Zhang, H.; He, Y.; Shuai, S.; Chen, S.; Cao, S.; Li, W.; Ma, D.; Chen, H. Genome-wide identification, characterization analysis and expression profiling of auxin-responsive GH3 family genes in wheat (*Triticum aestivum* L.). *Mol. Biol. Rep.* **2020**, *47*, 3885–3907. [CrossRef]

69. Brown, D.E.; Rashotte, A.M.; Murphy, A.S.; Normanly, J.; Tague, B.W.; Peer, W.A.; Taiz, L.; Muday, G.K. Flavonoids act as negative regulators of auxin transport in vivo in *Arabidopsis*. *Plant Physiol.* **2001**, *126*, 524–535. [CrossRef]

70. Gondor, O.K.; Janda, T.; Soós, V.; Pál, M.; Majláth, I.; Adak, M.K.; Balázs, E.; Szalai, G. Salicylic acid induction of flavonoid biosynthesis pathways in wheat varies by treatment. *Front. Plant Sci.* **2016**, *7*, 1447. [CrossRef]