**IiWRKY34 positively regulates yield, lignan biosynthesis and stress tolerance in *Isatis indigotica***

Ying Xiaoa, Jingxian Fenga, Jing Li, Yangyun Zhou, Qitao Buc, Junhui Zhoud, Hexin Tan, Yingbo Yang, Lei Zhangc,*, Wansheng Chena,*,

aResearch and Development Center of Chinese Medicine Resources and Biotechnology, Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China
bDepartment of Pharmacy, Changzheng Hospital, Naval Medical University (Second Military Medical University), Shanghai 200003, China
cDepartment of Pharmaceutical Botany, School of Pharmacy, Naval Medical University (Second Military Medical University), Shanghai 200433, China
dNational Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
eKanion Research Institute, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

Received 6 November 2019; received in revised form 14 November 2019; accepted 24 December 2019

**KEY WORDS**
Polyplody vigor; WRKY transcription factor; Biomass production; Lignan biosynthesis; Stress tolerance

**Abstract**
Yield potential, pharmaceutical compounds production and stress tolerance capacity are 3 classes of traits that determine the quality of medicinal plants. The autotetraploid *Isatis indigotica* has greater yield, higher bioactive lignan accumulation and enhanced stress tolerance compared with its diploid progenitor. Here we show that the transcription factor IiWRKY34, with higher expression levels in tetraploid than in diploid *I. indigotica*, has large pleiotropic effects on an array of traits, including biomass growth rates, lignan biosynthesis, as well as salt and drought stress tolerance. Integrated analysis of transcriptome and metabolome profiling demonstrated that IiWRKY34 expression had far-reaching consequences on both primary and secondary metabolism, reprogramming carbon flux towards phenylpropanoids, such as lignans and flavonoids. Transcript—metabolite correlation analysis was applied to construct the regulatory network of IiWRKY34 for lignan biosynthesis. One candidate target Ii4CL3, a key rate-limiting enzyme of lignan biosynthesis as indicated in our previous study, has been
demonstrated to indeed be activated by IiWRKY34. Collectively, the results indicate that the differentially expressed IiWRKY34 has contributed significantly to the polyploidy vigor of *I. indigotica*, and manipulation of this gene will facilitate comprehensive improvements of *I. indigotica* herb.

© 2020 Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Polyploids often present novel phenotypes that are not found in their diploid progenitors, including enhanced organ size, biomass and stress tolerance, etc. These traits often have some adaptive significance, allowing polyploids to increase their chances of being selected by nature, which we called “polyploidy vigor” [2]. The appearance of polyploidy vigor is demonstrated under complex genetic control, involving changes in gene expression through increased variation in dosage-regulated gene expression, epigenetic regulation and regulatory interactions [3]. Therefore, study on the gene expression related to the altered phenotype is crucial to clarify the underlying molecular mechanisms of polyploidy vigor, and will prompt the discovery of rational intervention strategies towards desired phenotypes.

*Isatis indigotica* Fort., belonging to the family Cruciferae, is a prevalent Chinese medicinal herb. The root of *I. indigotica* (Radix Isatidis), with Chinese name “Ban Lan Gen”, is frequently used for the treatment of hepatitis, influenza and various kinds of inflammation [4]. Lignans, mainly including larisocresinol and its derivatives, have been identified as effective antiviral components of *I. indigotica* [5-7]. In our previous study, the tetraploid *I. indigotica* (2n = 28) with greater yield, higher lignans accumulation and enhanced stress resistance was obtained from its natural diploid progenitor (2n = 14) [8,9]. An Arabidopsis thaliana whole genome Affymetrix gene chip (ATH1) [10] was used to survey the variation of gene expression between tetraploid and diploid *I. indigotica*, and results revealed a coordinated induction and suppression of 715 and 251 ploidy-responsive genes in tetraploid *I. indigotica*, involving in various developmental, signal transduction, transcriptional regulation and metabolic pathways [11]. Some of them, such as a stomatal developmental gene IiSDD1 [12], two signal transduction genes IiCPK1 [13] and IiCPK2 [14], and a lignan biosynthetic pathway gene IiPAL [15], have been characterized to explore their contribution to the favorable physiological consequences after polyploidization. More recently, transcriptomic analysis of diploid and tetraploid *I. indigotica* indicated that the differentially expressed genes (DEGs) were mainly involved in cell growth, cell wall organization, secondary metabolite biosynthesis, stress response and photosynthetic pathways [16]. Nevertheless, further studies are required to explore the mechanisms of the autotetraploidy vigor of *I. indigotica*.

Transcription factors (TFs) play essential roles in plants by controlling the expression of genes involved in various cellular processes, and are recognized to be particularly important in the process of crop domestication and are targets of molecular breeding of crops [17-19]. The comprehensive survey of global gene expression performed by ATH1 revealed eight TFs tend to be significantly higher in tetraploid than in diploid *I. indigotica*, and among them there are 4 WRKY genes [11]. Since the physiological role of WRKY TFs are widely related to diverse developmental processes, stress responses, and specialized metabolism [20], we reason that their expression variation in diploid and tetraploid *I. indigotica* might associate with altered phenotypes.

In the present study, a total of 64 IiWRKY genes (IiWRKY1-64) were first identified in *I. indigotica* transcriptome. In particular, IiWRKY34 expression, significantly higher in tetraploids than in diploids, positively correlated with larisocresinol accumulation. Over-expression and RNAI analysis indicated that IiWRKY34 is able to regulate larisocresinol biosynthesis, meanwhile, its upregulation improves root development, and enhances salt and drought stress tolerance. This study provides new insights into the genetic bases underlying the superiority of tetraploid *I. indigotica* compared to its diploid progenitor, as well as a potential target for genetic improvement of *I. indigotica* herb.

2. Materials and methods

2.1. Identification and characterization of IiWRKY genes

The homologous of WRKYs from the assembly of diploid *I. indigotica* transcriptome sequences [21] were searched using the BLASTx algorithm [22] from *A. thaliana* WRKYs (AtWRKYs) and Chinese cabbage WRKYs (Br/WRKYs) respectively retrieved from The Arabidopsis Information Resource (http://www.arabidopsis.org) and Brassica Database (BRAD, http://brassicadb.org/brad/). The Pfam database (pfam, http://pfam.janelia.org) [23] and the Simple Modular Architecture Research Tool (SMART, http://smart.embl-heidelberg.de) [24,25] were used to identify the putative WRKY proteins. The ProtParam tool (http://web.expasy.org/protparam) was further used to analyze the chemical and physical characteristics of these IiWRKY proteins [26].

2.2. Bioinformatics analysis of IiWRKYs

The amino acid sequence alignments of IiWRKYs alone, or along with AtWRKYs were performed using CLUSTALX version 2.0.12 [27]. Phylogenetic relationships were analyzed using the Neighbor-Joining method with pairwise deletion option in MEGA 5.05 [28]. The putative polyploidy-responsive IiWRKYs were identified through comparative analysis of orthologous genes between *I. indigotica* and *Arabidopsis*. According to the multiple sequence alignment and the previously reported classification of *Ar*WRKYs [29], the IiWRKYs were assigned to different groups and subgroups. The possible conserved motifs were further detected by MEME [30]. IiWRKY protein interactions were constructed by using STRING software (http://string-db.org) [31].
2.3. Integrated analysis of IiWRKYs expression and lariciresinol accumulation

Tetraploid I. indigotica was generated followed the methods as described by Qiao8 using I. indigotica (2n = 14) as the diploid donor. The supposed diploid and tetraploid plants were sampled to analyze ploidy levels using Quanta SC Flow Cytometer (Beckman Coulter, Brea, CA, USA)16. The hairy root culture was derived after infection of diploid and tetraploid I. indigotica plantlets with a Ri T-DNA bearing Agrobacterium rhizogenes bacterium (CS8Cl)21. Methyl jasmonate (MeJA, 0.5 tor35 to generate plasmids pCAMBIA1300-ploid I. indigotica harvested at various time points (0, 1, 3, 6, 12 and 24 h). The PCR products were digested with restriction enzymes. Three independent biological samples were analyzed. The harvested hairy roots, along with roots of diploid and autotetraploid I. indigotica, were used for RNA isolation and lariciresinol content determination.

Total RNAs were extracted using TRIzol Reagent (Thermo, Waltham, MA, USA), and the mRNAs were reversely transcribed by oligo dT to generate cDNA as a template. Real-time quantitative PCR (RT-qPCR) was used to analyze the transcripts of IiWRKY33, IiWRKY34, IiWRKY48, IiWRKY49 and IiWRKY50. Gene-specific RNA primers for these IiWRKYs and the I. indigotica actin gene reported by Li et al.32 were listed in Supporting Information Table S1. The RT-qPCR was performed according to manufacturer’s instruction (Waltham, MA, USA) as described by Li et al.32. The content of lariciresinol was determined by using LC–MS as described above. The selected transitions of m/z were 179 → 146 for conifer alcohol, 357 → 151 for pinoresinol, 359 → 329 for lariciresinol, 361 → 164 for secoisolariciresinol, 519 → 357 for pinoresinol 4-O-glucopuranoisde, and 685 → 523 for secoiso-lariciresinol diglucoside, respectively. All the standards were purchased from Sigma–Aldrich. Microscopic analysis of hairy roots was done essentially as described by De & Aronne37. Phloroglucinol-HCl staining was conducted to detect lignans, lignins, or wall-bound phenolics and derivatives, based on our previously published protocols35.

2.4. Plasmid vector construction and transgenic hairy roots generation

The coding sequence of IiWRKY34 was amplified by PCR using gene-specific primers IiWRKY34-F and IiWRKY34-R (Table S1). The PCR products were digested with Bcl I and Spe I, and ligated into plasmid PHB-flag to generate PHB-IiWRKY34-flag. For construction of the RNAi vector, an appropriate 351 bp fragment of IiWRKY34 was amplified by PCR using primers IiWRKY34-sNcoI-aSalI and IiWRKY34-sKpnI-aXbaI (Table S1). The PCR products were then subcloned in opposite orientations on either side of the Pdk intron of the pCAMBIA1300-pHANNIBAL vector35 to generate plasmids pCAMBIA1300-IiWRKY34. After sequencing confirmation, the above two plasmids, together with PHB-flag and pCAMBIA1300-pHANNIBAL as vector controls (control check, CK), were introduced separately into leaf explants of diploid I. indigotica by using Agrobacterium tumefaciens C58Cl strain and the generated hairy roots were screened using hygromycin. Hairy root lines generated through transformation with the blank C58Cl strain were used as wild-type (WT) control. The hairy roots were cultured as described by Chen et al.21 The fresh weight (determined as the difference between the whole flask with and without the harvested root tissues) was recorded at Day 9, 18, 27, 36, and 45 post-inoculation. The hairy roots harvested at the Day 45 were used for DNA extraction, RNA extraction, metabolite determination, microscopic analysis and phloroglucinol-HCl staining. Genomic DNA was subjected for PCR analysis to detect exogenous IiWRKY34 transformations using primers IiWRKY34-ovx-F and IiWRKY34-ovx-R (designed specifically to cover the gene sequence and the vector sequence, Table S1). For RNAi transgenic hairy roots, primers JPDK-1F and JPDKR (Table S1) were used to detect the inserted IiWRKY34 fragment. The transformed status of hairy roots was also verified for the presence of genes hpt and rolB or rolC35. PCR-positive hairy roots were analyzed for IiWRKY34 expression by using RT-qPCR analysis as described above.

The content of lignans was determined by LC–MS as described above. The selected transitions of m/z were 179 → 146 for conifer alcohol, 357 → 151 for pinoresinol, 359 → 329 for lariciresinol, 361 → 164 for secoisolariciresinol, 519 → 357 for pinoresinol 4-O-glucopuranoisde, and 685 → 523 for secoiso-lariciresinol diglucoside, respectively. All the standards were purchased from Sigma–Aldrich. Microscopic analysis of hairy roots was done essentially as described by De & Aronne37. Phloroglucinol-HCl staining was conducted to detect lignans, lignins, or wall-bound phenolics and derivatives, based on our previously published protocols35.

2.5. Expression profile of IiWRKY34 in different tissues and under various treatments

Leaves of 2-month-old diploid I. indigotica seedlings were sprayed with salicylic acid (SA, 100 μmol/L), MeJA (100 μmol/L) or NaCl (200 mmol/L) and sampled at 0, 1, 3, 6, and 12 h after treatment. For drought treatment, the seedlings were subjected to 2.5% polyethylene glycol (PEG) for the indicated times. For UV-B treatment, the seedlings were exposed to 1500 J/m2 UV-B light for 30 min, and then sampled at 0, 20, 40, 60 and 80 min during treatment, and at 30, 60, and 120 min post-treatment. The expression level of IiWRKY34 in different tissues (roots, stems, leaves and flowers) and under various treatments was examined using RT-qPCR analysis as described above.

2.6. Determination of ROS level, proline content and total antioxidant capacity of transgenic hairy roots under stress treatments

The IiWRKY34 overexpression and depression, along WT type hairy roots (20 g) were respectively subjected to salt and drought treatments by adding NaCl (75 mmol/L) and PEG (2.5%) into the liquid culture medium. After treatment for 5 days, reactive oxygen species (ROS) level was detected using the red fluorescence probe dihydroethidium Vigorous (http://www.vigorosbiol.com/) following the methods as described by Wang et al.38 Free proline content was determined as described by Bates et al.39 Total antioxidant activities were evaluated for trolox equivalent antioxidant capacity (TEAC) using the methods as described by Zhang et al.40.
2.7. Transcript profiling

The Illumina HiSeq2000 platform (San Diego, CA, USA) was used to investigate gene expression profile of the transgenic and WT hairy roots (3 groups × 5 biological replicates) harvested at the day 45 after inoculation. The raw reads were first generated using Solexa GA pipeline 1.6. After the removal of low-quality reads, the retained high-quality reads were mapped to previous annotation of *I. indigotica* transcriptome21. Tags were assigned to have significantly differential expression if they had a P-value of <0.05, a false discovery rate (FDR) of <0.05, and an estimated absolute fold-change >2 in sequence counts across libraries. These DEGs were further applied for Gene Ontology (GO) enrichment analysis16 and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis16 to find out their biological implications or the involved pathway.

2.8. Untargeted metabolite profiling

Metabolites from 15 different samples (100 mg of fresh weight, 3 groups × 5 biological replicates) were extracted according to Lise et al.33 and determined by GC–TOF/MS (Agilent 7890, Agilent). The programs of temperature-rise was set as followed: 70 °C for 2 min, 10 °C/min rate up to 140 °C, 4 °C/min rate up to 240 °C, 10 °C/min rate up to 300 °C and staying at 300 °C for 8 min. Full-scan method with range from 50 to 600 (m/z) was used. The total mass of signal integration area was normalized for each sample, and the normalized data were imported into Simca-P software (version 11.5, http://www.umetrics.com/simca), employing PLS-DA model using the first principal component of VIP (variable importance in the projection) values (VIP>1) combined with Student’s t test (t-test, P < 0.05) to find differentially expressed metabolites. The volcano plot was used to reveal significantly altered metabolite features via delineating a log transformation plot of the fold-change difference (log2 fold change value as x-axis) and the level of statistical significance (−log10P value as the y-axis) of each metabolite. Metabolite alterations between *IiWRKY34-OVX* and *IiWRKY34-RNAi* lines were depicted in a map drawn according to KEGG pathway database.

2.9. Data mining

The assembly of transcriptome sequences was searched for *IiWRKY*s and the lariacerinol biosynthetic genes using our previously published protocols33, and these gene expression patterns in different *IiWRKY34* transgenic lines (15 samples) were visualized with a heat map using the log2-transformed data in Multi-Experiment Viewer (version 4.9.0)33. Genes with different expression patterns were grouped through hierarchical clustering. Accordingly, a heat map showing lignans accumulation in different *IiWRKY34* transgenic lines was also constructed.

Correlation among *IiWRKY34*, lariacerinol biosynthetic genes and lignans were constructed using the Pearson correlation coefficient according to the co-occurrence principle33. The correlation network was generated using Cytoscape (version 3.6.0)35.

2.10. EMSA

The *IiWRKY34* coding sequence was amplified using primers *IiWRKY34*-pET-F and *IiWRKY34*-pET-R (Table S1), and inserted into pET32a vector (Novagen, Darmstadt, Germany) between the sites NcoI and SacI to generate pET-*IiWRKY34* plasmid. This plasmid was then transformed into the *Escherichia coli* BL21 strain, and the recombinant protein was purified by using a His Spin Trap column (GE Healthcare, Buckinghamshire, UK). EMSA was performed using biotin-labelled probes and a Light-Shift chemiluminescent EMSA kit (Thermo, Chicago, IL, USA). To design biotin-labelled probes, a 1500-bp upstream region of *Ii4CL3* was amplified following the instructions of the Genome Walker Kit (Clontech, Mountain View, CA, USA), and analysed for the presence of W-boxes (C/T)TGAC (T/C). The biotin-labelled probes (Table S1) were synthesized by Sangon Biotech Company (Shanghai, China).

2.11. Dual luciferase assay

The coding sequence of *IiWRKY34* was subcloned into the PHB vector (Biovector, Beijing, China) to generate the effector, and the promoter of *Ii4CL3* was fused into the vector pGreenII 0800 (Biovector) to generate a reporter. The reporter and effector constructs were then separately transformed into *A. tumefaciens* strain GV3101. The bacterial cells were resuspended in MS medium with 10 mmol/L methylester sulfonate and 150 μmol/L acetosyringone to OD600 0.6 and then incubated at room temperature for 3 h. The bacteria-harboring constructs were infiltrated into tobacco leaves according to Zhang et al.46 The leaves were collected after 48 h for dual-LUC assays using a Dual-Luciferase Reporter Assay System according to the manufacturer’s instructions (Promega, Madison, WI, USA). Three independent biological replicates were measured for each sample.

2.12. Statistics

Experiments were performed in triplicate, and statistical analysis was performed using SPSS 22.0 software. Paired, two-tailed Student’s t-test was used to compare group differences. P-values<0.05 were regarded as statistically significant.

2.13. Data availability

The nucleotides and amino acid sequences of *IiWRKY*s (*IiWRKY1* to *IiWRKY64*) are deposited in the GenBank databases under the accession numbers (MN480587 to MN480650). The raw RNA-seq read data are accessible through accession number PRJNA491805 (http://www.ncbi.nlm.nih.gov/sra/).

3. Results

3.1. Identification and characterization of WRKY genes in *I. indigotica*

A total of 64 putative *IiWRKY* genes (*IiWRKY1* to *IiWRKY64*) were identified (Supporting Information Table S2). The ORFs were extracted from the putative *IiWRKY* sequences, and then converted into amino acid sequences (Supporting Information Table S3). Detailed information about each *IiWRKY* is given in Table 1.

Sequence alignment of the unique DNA-binding domain, spanning approximately 60 amino acids of all 64 *IiWRKY*s revealed that all *IiWRKY*s contain the highly conserved DNA binding domain composed of the conserved WRKYGQK sequence followed by a C2H2- or C2HC-type zinc finger motif
| Group | Subgroup | Gene ID | Gene locus ID | CDS (bp) | ORF (aa) | Mass (kDa) | pI | Atorholog |
|-------|----------|---------|--------------|---------|----------|------------|----|-----------|
| I     | IiWRKY3  | comp12944_c0_seq1 | 1443 | 480 | 52.36 | 5.47 | AtIiWRKY32 |
|       | IiWRKY5  | comp14011_c0_seq1 | 594 | 197 | 22.08 | 8.86 | AtIiWRKY20 |
|       | IiWRKY47 | comp22052_c1_seq1 | 717 | 238 | 26.16 | 9.30 | AtIiWRKY44 |
|       | IiWRKY37 | comp22961_c0_seq1 | 1392 | 463 | 50.46 | 8.73 | AtIiWRKY58 |
|       | IiWRKY33 | comp27067_c0_seq1 | 1548 | 515 | 56.76 | 8.37 | AtIiWRKY32 |
|       | IiWRKY36 | comp27813_c0_seq2 | 972 | 323 | 36.24 | 9.45 | AtIiWRKY26 |
|       | IiWRKY37 | comp27813_c0_seq4 | 375 | 124 | 14.46 | 9.56 | AtIiWRKY26 |
|       | IiWRKY48 | comp32055_c0_seq1 | 1065 | 354 | 39.73 | 7.70 | AtIiWRKY25 |
|       | IiWRKY49 | comp32055_c0_seq2 | 1224 | 407 | 45.56 | 7.19 | AtIiWRKY25 |
|       | IiWRKY53 | comp32270_c3_seq7 | 1521 | 506 | 54.83 | 8.67 | AtIiWRKY94 |
|       | IiWRKY54 | comp32270_c3_seq8 | 1548 | 515 | 56.23 | 6.98 | AtIiWRKY32 |
|       | IiWRKY60 | comp34055_c0_seq1 | 1431 | 476 | 52.80 | 8.70 | AtIiWRKY11 |
|       | IiWRKY61 | comp36248_c0_seq1 | 1728 | 575 | 62.32 | 8.81 | AtIiWRKY72 |
|       | IiWRKY2  | comp12464_c0_seq1 | 972 | 323 | 36.26 | 5.94 | AtIiWRKY49 |
|       | IiWRKY5  | comp14006_c0_seq1 | 1236 | 411 | 46.24 | 6.37 | AtIiWRKY48 |
|       | IiWRKY31 | comp14018_c0_seq1 | 585 | 194 | 22.38 | 6.30 | AtIiWRKY59 |
|       | IiWRKY15 | comp22685_c0_seq2 | 465 | 154 | 18.02 | 8.48 | AtIiWRKY48 |
|       | IiWRKY39 | comp28388_c0_seq1 | 885 | 294 | 33.43 | 8.12 | AtIiWRKY94 |
|       | IiWRKY50 | comp32255_c0_seq2 | 729 | 242 | 26.81 | 8.61 | AtIiWRKY60 |
|       | IiWRKY48 | comp32055_c0_seq1 | 909 | 302 | 33.43 | 7.55 | AtIiWRKY40 |
|       | IiWRKY2  | comp12464_c0_seq1 | 972 | 323 | 36.26 | 5.94 | AtIiWRKY49 |
|       | IiWRKY61 | comp36248_c0_seq1 | 1728 | 575 | 62.32 | 8.81 | AtIiWRKY72 |
|       | IiWRKY25 | comp26532_c0_seq2 | 852 | 283 | 31.95 | 7.74 | AtIiWRKY71 |
|       | IiWRKY26 | comp26532_c0_seq3 | 852 | 283 | 31.95 | 7.74 | AtIiWRKY71 |
|       | IiWRKY28 | comp26776_c0_seq1 | 528 | 175 | 20.20 | 9.51 | AtIiWRKY57 |
|       | IiWRKY29 | comp26776_c0_seq5 | 999 | 332 | 36.95 | 6.98 | AtIiWRKY57 |
|       | IiWRKY30 | comp26776_c0_seq9 | 540 | 179 | 20.73 | 9.63 | AtIiWRKY57 |
|       | IiWRKY32 | comp26972_c0_seq2 | 441 | 146 | 16.77 | 9.37 | AtIiWRKY75 |
|       | IiWRKY35 | comp27574_c0_seq1 | 1113 | 370 | 41.52 | 7.32 | AtIiWRKY23 |
|       | IiWRKY38 | comp27813_c0_seq7 | 576 | 191 | 21.52 | 7.00 | AtIiWRKY56 |
|       | IiWRKY40 | comp28803_c0_seq1 | 519 | 172 | 19.82 | 8.97 | AtIiWRKY12 |
|       | IiWRKY41 | comp28803_c0_seq2 | 399 | 133 | 15.51 | 9.30 | AtIiWRKY12 |
|       | IiWRKY56 | comp32704_c0_seq5 | 432 | 143 | 16.51 | 9.01 | AtIiWRKY45 |
|       | IiWRKY28 | comp16130_c0_seq1 | 1038 | 345 | 37.72 | 9.62 | AtIiWRKY51 |
|       | IiWRKY45 | comp31187_c4_seq1 | 1041 | 346 | 38.24 | 9.75 | AtIiWRKY74 |
|       | IiWRKY46 | comp31187_c4_seq3 | 996 | 331 | 36.74 | 9.51 | AtIiWRKY39 |
|       | IiWRKY47 | comp31456_c3_seq4 | 588 | 195 | 21.83 | 9.44 | AtIiWRKY15 |
|       | IiWRKY50 | comp32075_c0_seq1 | 687 | 228 | 25.35 | 9.64 | AtIiWRKY21 |
|       | IiWRKY62 | comp37405_c0_seq1 | 1038 | 345 | 37.81 | 9.77 | AtIiWRKY77 |
|       | IiWRKY7  | comp15334_c0_seq1 | 1269 | 422 | 45.77 | 5.17 | AtIiWRKY14 |
|       | IiWRKY9  | comp17322_c0_seq1 | 900 | 299 | 32.35 | 6.39 | AtIiWRKY22 |
|       | IiWRKY12 | comp19645_c0_seq1 | 870 | 289 | 31.64 | 5.60 | AtIiWRKY35 |
|       | IiWRKY16 | comp22855_c0_seq1 | 828 | 275 | 30.59 | 5.09 | AtIiWRKY69 |
|       | IiWRKY78 | comp24498_c0_seq1 | 777 | 258 | 28.91 | 5.52 | AtIiWRKY65 |
|       | IiWRKY43 | comp30090_c0_seq1 | 552 | 183 | 20.37 | 4.57 | AtIiWRKY27 |
|       | IiWRKY44 | comp30090_c0_seq3 | 1047 | 348 | 38.68 | 4.93 | AtIiWRKY27 |
|       | IiWRKY64 | comp66853_c0_seq1 | 921 | 306 | 34.08 | 6.27 | AtIiWRKY29 |
|       | IiWRKY11 | comp7344_c0_seq2 | 786 | 261 | 29.77 | 5.66 | AtIiWRKY11 |
|       | IiWRKY4  | comp36399_c0_seq2 | 894 | 297 | 33.21 | 6.01 | AtIiWRKY70 |
|       | IiWRKY20 | comp25578_c2_seq1 | 870 | 289 | 32.90 | 5.71 | AtIiWRKY46 |
|       | IiWRKY27 | comp26620_c0_seq1 | 1020 | 339 | 37.96 | 5.63 | AtIiWRKY54 |

(continued on next page)
Factors affecting plant stress responses, including STZ (salt tolerance) largely participate in plant defense regulatory pathways, as most Arabidopsis and 84.54% of identity to IiWRKY48 (orthologous to AtWRKY21). Phylogenetic and structural analysis of IiWRKYs were shown in Supporting Information Fig. S3. Results indicated that IiWRKYs evolved from group I to group II and finally to group III, paralleled with a WRKY evolutionary process and grouping of their WRKY domains, and the Orthologous WRKYs functions in the present study.

3.3. Integrated analysis of polyploidy-responsive IiWRKYs and lariciresinol

The ploidy levels of diploid and tetraploid I. indigotica were confirmed via flow cytometric analysis as shown in Supporting Information Fig. S4. RT-qPCR analysis indicated that the transcript levels of IiWRKY33, IiWRKY34, IiWRKY48, IiWRKY49 and IiWRKY50 in roots of tetraploid I. indigotica were significantly higher than that of diploid ones (P < 0.05), and the fold changes were comparable with our microarray findings. The role of IiWRKY34, IiWRKY33, IiWRKY48, IiWRKY49 and IiWRKY50 were responsive to MeJA treatment in both diploid and tetraploid I. indigotica hairy roots, but with different patterns. It was obvious to note that IiWRKY34 was more responsive to MeJA than other members, and its expression was dramatically up-regulated at 1 h post-treatment then lasted to the end of the experiment in both samples. It was interesting to note that the expression pattern of IiWRKY33, IiWRKY34, IiWRKY48, IiWRKY49 and IiWRKY50 in the induced hairy roots of tetraploid I. indigotica greatly differed from that in its original roots, suggesting the expression of these IiWRKY genes is under strict developmental and tissue-specific control. LC–MS analysis showed roots of tetraploid I. indigotica accumulated more lariciresinol than diploid progenitor (P < 0.05), consistent with our earlier finding that tetraploid I. indigotica exhibited higher antiviral effect compared with its diploid counterpart. In addition, MeJA treatment greatly triggered lariciresinol production in both diploid and tetraploid I. indigotica hairy roots, but with different patterns that in diploids lariciresinol accumulation increased gradually and peaked at 12 h, whereas in tetraploids it decreased gradually until 3 h and then continuously increased from 3 to 24 h post treatment (Fig. 2A).

A correlation analysis between the above IiWRKYs and lariciresinol presented as a heat map (Fig. 2B) indicated that IiWRKY34 was most highly correlated with lariciresinol with a correlation coefficient of 0.812, suggesting IiWRKY34 probably positively regulated lariciresinol production.

3.4. IiWRKY34 positively regulates lignan biosynthesis in I. indigotica hairy roots

The role of IiWRKY34 in lignan biosynthesis was investigated using a transgenic hairy root assay (Fig. 3). Two constructs (PHB-IiWRKY34-flag and pCAMBIA1300-IiWRKY34).
Information Fig. S5) were generated for over-expression and RNA interference analysis (IiWRKY34-OVX and IiWRKY34-RNAi), respectively. The transformants were identified by PCR analysis: all the hairy roots contained the rolb or rolc gene, which indicated successful transformation of pRiA436. The hygromycin resistance gene hpt was detected in both transgenic and CK lines. In addition, transgenic lines also contained IiWRKY34-specific fragments (Supporting Information Fig. S6).

RT-qPCR analysis indicated that IiWRKY34 expression level was successfully regulated through genetic manipulation that transgenic roots overexpressing IiWRKY34 showed a dramatic increase in IiWRKY34 expression, whereas IiWRKY34-RNAi roots showed a significant reduction compared with WT and CK (P < 0.05, Fig. 3G). It was interesting to note IiWRKY34-OVX roots grew fast and vigorously with thick branches whereas IiWRKY34-RNAi roots grew slowly with slender branches (Fig. 3A). At the Day 45 after inoculation, the biomass of IiWRKY34-OVX and IiWRKY34-RNAi roots were approximately 3.7- and 0.3-fold of WT respectively, and no significant difference was detected between WT and CK (PHB, P1300) lines during the whole hairy root culture period (Fig. 3F). The morphology of different transgenic roots at the Day 45 was shown in Fig. 3B, and these roots were used to microscopic analysis and phloroglucinol-HCl staining. Microscopic analysis showed that when compared

Figure 1  The interaction network of 64 IiWRKY proteins identified in I. indigotica and related proteins in Arabidopsis. Stronger associations are represented by thicker lines. The polyploidy-responsive IiWRKYs are denoted with red color.
with WT and CK controls, the interfascicular fibers and xylem cells of *I. indigotica*-OVX roots were relatively compacted with a higher lignification level, on the contrary those of *I. indigotica*-RNAi counterparts were dispersed and had a lower lignification level (Fig. 3C). Phloroglucinol-HCl staining showed *I. indigotica*-OVX roots presented a violet-red colour, whereas *I. indigotica*-RNAi counterparts presented a weaker browning compared with WT and CK (Fig. 3D). Similar color was also found in their corresponding eluents (Fig. 3E). These results indicate that *IiWRKY34* positively improves the accumulation of lignans, lignins, and/or wall-bound phenolics and derivatives.

To test whether *IiWRKY34* positively improves the pharmacologically important lignans, 6 compounds involved in lariciresinol biosynthetic pathway (Fig. 3H) were determined by LC–MS. Results showed that overexpression of *IiWRKY34* dramatically enhanced the production of the 6 lignans, and line OVX-4, with the highest *IiWRKY34* expression (20-fold of WT, Fig. 3G), produced the most abundant conifer alcohol (82.8 μg/g DW), pinoresinol (14.9 μg/g DW), lariciresinol (400.4 μg/g DW), secoisolariciresinol (184.5 μg/g DW), and secoisolariciresinol diglucoside (31.1 μg/g DW), which were ~6.7-, 2.5-, 7.6-, 14.1- and 16.4-fold more than in WT, respectively. In contrast, RNAi suppression of *IiWRKY34* decreased the production of conifer alcohol, lariciresinol, secoisolariciresinol, and pinoresinol 4-O-glucopuranoside with different degrees. There was no significant difference in lignan content between CK and WT lines (Fig. 3I).

3.5. *IiWRKY34* positively improves salt and drought stress tolerance in *I. indigotica* hairy roots

Expression of *IiWRKY34* was examined in various organs of 2-month-old *I. indigotica* seedlings, result showed that it was abundantly expressed in roots, stems and leaves, but only slightly expressed in flowers. *IiWRKY34* responded to drought, salt, SA, MeJA and UV-B treatments, but with different patterns of variation. After drought and salt treatment, the expression level of *IiWRKY34* increased gradually and reached a maximum at 6 h after treatment, which was approximately 7.3- and 24.2-fold, respectively, higher than that before treatment. When *I. indigotica* was treated with SA, *IiWRKY34* expression increased sharply at 1 h and peaked at 6 h (29.1-fold of that before treatment). Parallelled with MeJA-treated *I. indigotica* hairy roots (Fig. 2A), *IiWRKY34* expression in *I. indigotica* seedlings was induced by MeJA at 1 h (28.0-fold of that before treatment) post-treatment but decreased at 3 h, and then gradually increased afterward. For UV-B treatment, *IiWRKY34* expression gradually increased until 60 min and decreased at 80 min under UV-B, but its expression then increased after UV-B was turned off, and the expression level at 120 min achieved approximate 31.6-fold of that before treatment (Supporting Information Fig. S7). These results indicate that *IiWRKY34* can be significantly induced when subjected to environment stresses, which is in agreement with that the expression level of *IiWRKY34* is involved in plant defence regulatory pathways as indicated in Fig. 1.

Both bioinformatics analysis (Fig. 1) and stress induction (Fig. S7) suggested a role of *IiWRKY34* in stress response. Thus, its capacity for stress tolerance was further investigated using transgenic hairy roots. Results showed *IiWRKY34* expression indeed could positively improve salt and drought stress tolerance. As shown in Fig. 4A, after 5 days of salt or drought treatment, both WT and *IiWRKY34*-OVX roots grew well as normal, but *IiWRKY34*-RNAi counterparts showed an early senescence phenotype with severe growth retardation compared with that without treatment (CK).

The intracellular ROS level in the WT and transgenic hairy roots was tested by fluorescence staining. Under normal conditions (CK), *IiWRKY34*-OVX root tips displayed a lower level of ROS whereas *IiWRKY34*-RNAi displayed a relatively higher level
compared to WT. After salt or drought treatment, ROS accumulation enhanced in both WT and RNAi roots (especially in RNAi ones), while its accumulation in IiWRKY34-OVX roots still stayed at a low level as that in CK (Fig. 4B). These results imply that IiWRKY34 may reduce the ROS level to confer salinity and drought stress tolerance.

Since proline accumulation is widely recognized as a sign of stress tolerance in plants, we examined whether the proline content in transgenic hairy roots was altered. Under salt or drought stress condition, both WT and IiWRKY34 lines accumulated more proline than CK, and the proline content of IiWRKY34-OVX lines was much higher than that of WT. In contrast, the proline accumulation in IiWRKY34-RNAi lines was significantly decreased after salt treatment ($P < 0.05$), and remained approximately constant after drought condition (Fig. 4C). These results indicate that IiWRKY34 may enhance salt and drought stress tolerance by promoting proline production.

Moreover, we measured TEAC to investigate the physiological effects of transgenic hairy roots. Compared with WT, IiWRKY34-OVX lines showed an approximate 1.5-fold increase in TEAC level, whereas IiWRKY34-RNAi showed a reduction by 1-s. After salt and drought treatment, the TEAC level of both WT and IiWRKY34-RNAi lines significantly decreased ($P < 0.05$), but that of IiWRKY34-OVX was barely changed (Fig. 4D). This result indicates IiWRKY34 may maintain total antioxidant capacity to confer stress tolerance.

3.6. Gene expression profiles of transgenic I. indigotica hairy roots

Totally, 144,731 isogenes were identified by assembly. Differences in gene expression of the 3 groups (IiWRKY34-OVX, IiWRKY34-RNAi and WT hairy roots, 5 lines in each group) were shown in Fig. 5. Gene expression from individual groups showed a distinct sample separation (Fig. 5A), and a larger variation was found between IiWRKY34-OVX and IiWRKY34-RNAi lines significantly decreased ($P < 0.05$), and that of IiWRKY34-OVX was barely changed (Fig. 4D). This result indicates IiWRKY34 may maintain total antioxidant capacity to confer stress tolerance.

Figure 3 Phenotype analysis of IiWRKY34 transgenic hairy roots. Phenotype of developed root lines on solid medium for 20 days (A), and their corresponding root culture in liquid medium for 45 days (B). (C) Cross sections of hairy roots stained with safranin O/fast green FCF. Bar = 100 µm. Phenotype (D) and eluents (E) of hairy roots after phloroglucinol-HCl staining. (F) Biomass accumulation during the culture period. IiWRKY34 transcript expression (G) and lignan content (I) in different lines. Quantitative PCR analysis showing IiWRKY34 expression relative to the wild-type lines (WT-2) set at 1. (H) lignan biosynthetic pathway. Data were expressed as means±SD ($n = 3$). Asterisks represent significant difference at 0.05 level by Student’s t-test.

IiWRKY34 has pleiotropic effects
indicated that these DEGs were involved in biosynthesis of amino acids, carbon metabolism and phenylpropanoid, etc. The top 10 enriched pathways via pairwise contrasts presented in Fig. 5C distinguished a prominent variation between IiWRKY34-OVX and IiWRKY34-RNAi samples, and there were a total of 60 DEGs (17%) involved in phenylpropanoid biosynthesis.

The transcript levels of 64 IiWRKY in WT, IiWRKY34-OVX and IiWRKY34-RNAi hairy roots were presented as a heat map in Fig. 5D. As expected, IiWRKY34 expression in IiWRKY34-OVX lines was higher than WT, whereas lower in IiWRKY34-RNAi counterparts, and the fold changes were paralleled well with that examined by RT-qPCR (Fig. 3G), indicating that the RNA-Seq expression profile is robust and gene expression level obtained from this database is reliable. It was interesting to note that IiWRKY33, IiWRKY48 and IiWRKY49, which were found highly responsive to autopolyploidy (Fig. 2A), displayed a similar expression pattern with IiWRKY34 and they grouped together, indicating they may associate with each other in some manners.

3.7. Metabolite profiling of transgenic I. indigotica hairy roots

To assess the impact of IiWRKY34 on the metabolic shifts, nontargeted metabolic profiling was performed using GC–TOF/MS. Totally, 662 independent analytes were obtained from 15 samples. Differentially expressed metabolites (VIP > 1, P < 0.05) in the 3 groups (IiWRKY34-OVX, IiWRKY34-RNAi and WT hairy roots, 5 lines in each group) were shown in Fig. 6. Similar with gene expression profiles (Fig. 5A), metabolite accumulation from individual groups also showed a distinct sample separation (Fig. 6A). Variation between IiWRKY34-OVX and IiWRKY34-RNAi lines was more significant than that between other pairwise samples, there were 113 and 52 metabolites found decreased and enhanced in abundances for IiWRKY34-OVX versus IiWRKY34-RNAi samples, respectively (Fig. 6B and Supporting Information Table S4). Generation of volcano plots further visualized the significantly altered metabolite features in IiWRKY34-OVX, IiWRKY34-RNAi and WT hairy roots (Fig. 6C).
To integrate both primary and secondary metabolism that had been modified by *IiWRKY34* expression, we used a pathway scheme to summarize the metabolic changes (VIP > 1, *P* < 0.05) in *IiWRKY34-OVX* compared with *IiWRKY34-RNAi* roots. As shown in Fig. 6D, there were significantly higher amounts of phenylpropanoids such as flavonoids and lignans, whereas the basic sugar and the products of the TCA cycle, were reduced significantly, indicating that *IiWRKY34* appeared to reprogram primary metabolism, driving carbon flux towards specific secondary metabolism.

### 3.8. Regulatory network of *IiWRKY34* for lignan biosynthesis

Metabolic analysis revealed that *IiWRKY34* positively regulated lignans production. To have a systematic view on the variation of lignan biosynthesis pathway, we examined abundances of 37 transcripts coding 9 catalytic genes (Supporting Information Tables S5) and 6 metabolites involved in lariciresinol biosynthesis in WT, *IiWRKY34-OVX* and *IiWRKY34-RNAi* hairy roots (5 lines in each group). The RNA-Seq expression profile indicated that lariciresinol biosynthetic genes responded to *IiWRKY34* transgene with various patterns (Fig. 7A). The accumulation levels of six lignans (Fig. 3I) were normalized and presented as a heat map in Fig. 7B. Correlation coefficient cut-off values were applied to construct *IiWRKY34*-pathway genes—lignans correlation networks. Fig. 7C presents one example with a cut-off *R* > 0.5: a total of 11 pathway genes are correlated with *IiWRKY34* and at least one lignan, indicating *IiWRKY34* may improve lignan biosynthesis by modulating these lignan biosynthetic genes. Among these genes, one *Ii4CL* family member *Ii4CL3* has been demonstrated as a key rate-limiting enzyme of lariciresinol production, severing as a hub gene of lignan regulatory network in our earlier study56. In order to test whether *IiWRKY34* regulate lignan by direct interacting with *Ii4CL3*, EMSA was performed using *IiWRKY34* recombinant protein and the promoter of *Ii4CL3*. Result showed *IiWRKY34* indeed specifically bound to the W-box in the promoter region of *Ii4CL3*. There are two W-boxes in the 1500-bp *Ii4CL3* promoter region (Supporting Information Fig. S9), and the biotin-modified probe representing the two TGAC core elements formed a DNA–protein complex with *IiWRKY34*. Mutation of the two elements disrupted protein binding, and no retarded band representing complex formation was observed in the binding assay (Fig. 7D). Dual luciferase assay was further used to investigate how *IiWRKY34* regulate *Ii4CL3* by direct binding to its promoter sequence (Fig. 7E). Result showed *IiWRKY34* activated the promoter of *Ii4CL3* in vivo, as evidenced by a higher value of LUC/REN than the control (Fig. 7F), supporting the hypothesis that *IiWRKY34* interacts with the promoter of *Ii4CL3* and thus activates its transcription.

**Figure 5** Gene expression profiles of *IiWRKY34* transgenic hairy roots. (A) Heat map showing DEGs in *IiWRKY34-OVX*, *IiWRKY34-RNAi* and WT lines. (B) Number of DEGs via pairwise contrasts of *IiWRKY34-OVX*, *IiWRKY34-RNAi* and WT roots with fold-change > 2 and FDR <0.05. (C) Numbers of DEGs in the top ten enriched pathways. In parentheses: percentage of the total number of genes in the respective pathway. (D) Expression profiles of 64 *IiWRKYs* in different lines. *IiWRKY33, IiWRKY34, IiWRKY48, IiWRKY48* and *IiWRKY50* are highlighted with red markers.
Figure 6  Metabolite profiling of *IfWRKY34* transgenic hairy roots. (A) Heat map showing differentially expressed metabolites of *IfWRKY34-OVX, IfWRKY34-RNAi* and WT lines. (B) Number of metabolites with significant changes in concentration (VIP > 1, *P* < 0.05) via pairwise contrasts. (C) Volcano plot of differentially expressed metabolites via pairwise contrasts. (D) Pathway scheme summarizing the metabolic changes in *IfWRKY34-OVX* compared with *IfWRKY34-RNAi* transgenic roots. Metabolites which changed significantly (VIP > 1, *P* < 0.05) are highlighted in red (for increased) and blue (for decreased), metabolites without significant changes are highlighted in black.
4. Discussion

Yield potential, medicinal compounds concentration and stress
tolerance capacity are 3 classes of traits determining the quality of
herbs. Tetraploid *I. indigotica* has been appealing to people
because of its greater yield, higher bioactive compounds accu-
mulation and enhanced stress tolerance compared to its diploid
counterpart. Elucidation of the underlying molecular basis of the
significantly qualitative difference is of great importance for the
improvement of *I. indigotica*.

In the present study, *IiWRKY33*, *IiWRKY34*, *IiWRKY48*,
*IiWRKY49* and *IiWRKY50*, which express especially higher in
tetraploid *I. indigotica* than diploids, were proposed to be
particularly important in the trait development seen in the tetra-
ploids. An interaction network constructed using *Arabidopsis*
database revealed *IiWRKY33*, *IiWRKY34*, *IiWRKY48* and
*IiWRKY49* located as hub genes and associated with various
defence regulatory pathways (Fig. 1), suggesting they might have
contributed to the higher stress resistance of tetraploid *I. indigo-
tica*. Previous reports of their *Arabidopsis* homologues
*AtWRKY33* (orthologous to *IiWRKY33*), *AtWRKY40* (ortholo-
gous to *IiWRKY34*) and *AtWRKY25* (orthologous to *IiWRKY48*
and *IiWRKY49*) confirmed the reliability of this functional
protein association network. For instance, overexpression of
*AtWRKY25* and *AtWRKY33* increased salt tolerance and ABA
sensitivity57, *AtWRKY40* was induced in response to microbial
pathogen infection58 as well as MeJA treatment59, *AtWRKY25*
was involved in plant defense against *Pseudomonas syringae*,
and also acted as a cold resistance gene60. Another polyploid-
responsible member *IiWRKY50*, not found integrated in the
network, was designated as a calmodulin binding protein ac-
cording to the protein annotation of its *Arabidopsis* ortholog
*AtWRKY21*. Since calcium serves as an important second
messenger in plants, and changes in calcium concentration are
closely related to plant responses to various stimuli61, the
expression of calcium-related proteins such as *IiWRKY50* might
indirectly influence plant performance. To sum up, it can be
conceived that, during plant development, the up-regulation of
*WRKY* caused by the environmental stresses (e.g., pathogen,
salinity, coldness, drought, etc.) would result in higher stress
tolerance, thus prompting higher growth performance (e.g., higher
yield and enhanced metabolites biosynthetic efficiency). There-
fore, we can expect that the up-regulation of *WRKY* TFs
(*IiWRKY33*, *IiWRKY34*, *IiWRKY48*, *IiWRKY49* and *IiWRKY50*)
induced by autopolyploidy suggested tetraploid *I. indigotica* had a
much stronger adaptation capacity than diploid progenitor.

---

Figure 7  Regulatory network of *IiWRKY34* for lignan biosynthesis. Heat maps showing expression profiles of lignan biosynthetic genes (A) and lignan accumulations (B) in *IiWRKY34-OVX*, *IiWRKY34-RNAi* and WT lines. (C) *IiWRKY34*-lignan pathway genes—lignans correlation network with a cut-off $R>0.5$. *IiWRKY34*, pathway genes and lignans are drawn in red, yellow and green, respectively. The thickness of lines represents the level of correlation. *IiWRKY34*—lignans correlation was highlighted by black lines. (D) *IiWRKY34* specifically binds to the promoter of *Ii4CL3*. (E) Schematic diagram of the reporter and effector constructs used in the transient dual luciferase assay. (F) Transient dual luciferase analysis showing *IiWRKY34* activation of the transcription of *Ii4CL3* in *N. benthamiana* leaves. LUC/REN represents the luciferase/renilla ratio of $n=3$ independent experiments; Data were expressed as means±SD. Asterisk represents significant difference at 0.05 level by Student’s $t$-test.
showed a distinct sample separation (Figs. 5A and 6A), indicating IiWRKY34 expression made a marked effect on reshaping metabolic profiling from individual groups (metabolism. Generally speaking, both transcriptome and metabolome profiling of diploid and tetraploid I. indigotica showed a distinct sample separation (Figs. 5A and 6A), indicating IiWRKY34 expression made a marked effect on reshaping metabolic profiling from individual groups (metabolism. Generally speaking, both transcriptome and metabolome profiling of diploid and tetraploid I. indigotica.

For medicinal plants, the concentration of pharmaceutical compounds is the most important factor affecting the practice of medicine. Compared with diploid I. indigotica, the tetraploids accumulate more lignans including lariciresinol and its derivatives, which present effective antiviral ingredients of I. indigotica. The potency of plant-specific signaling molecules such as MeJA, to elicit secondary metabolism in cell cultures has made them powerful tools to cause the genetic diversity and help to unravel the complex cellular process. Here, MeJA-elicited diploid and tetraploid I. indigotica hairy roots harvested at different time points, along with the original roots of diploid and tetraploid I. indigotica, were employed as a resource of genetic variation to explore the potential correlation between polyploidy-responsive IiWRKYS (IiWRKY33, IiWRKY34, IiWRKY48, IiWRKY49 and IiWRKY50) and lariciresinol. Result showed IiWRKY34 was positively correlated with lariciresinol with a high correlation coefficient value ($R = 0.812$), suggesting IiWRKY34 probably participated in lariciresinol biosynthesis. Further evidence for a role of IiWRKY34 in the regulation of secondary metabolites has been found in its orthologs from other plants. For instance, GaWRKY1, also an ortholog of AtWRKY40, regulates the production of gossypol in cotton. CrWRKY13, another ortholog of AtWRKY40, is involved in biosynthesis of terpenoidindole alkaloids in Catharanthus. Therefore, we further investigated the effect of IiWRKY34 expression in lariciresinol biosynthesis using transgenic hairy root assays. Overexpression and RNAi analysis demonstrate that IiWRKY34 is an activator of lignans including lariciresinol, and it also plays a positive role in biomass accumulation (Fig. 3), as well as salt and drought tolerance as indicated by the changes of ROS level, proline content and total antioxidant capacity of transgenic hairy roots under stress conditions (Fig. 4).

Since large alterations were observed in the developmental phenotype for IiWRKY34-transgenic hairy roots, their molecular phenotype was characterized by changes to both transcript and metabolism. Generally speaking, both transcriptome and metabolome profiling from individual groups (IiWRKY34-OVX, IiWRKY34-RNAi and WT hairy roots, 5 lines in each group) showed a distinct sample separation (Figs. 5A and 6A), indicating IiWRKY34 expression made a marked effect on reshaping molecular phenotype of I. indigotica. Pathway classification of the DEGs revealed that IiWRKY34 appeared to affect both primary and secondary metabolism, including carbon metabolism, starch and sucrose metabolism, amino acids biosynthesis and phenylpropanoid biosynthesis, etc (Fig. 3C). Measurement of metabolic shifts supported this conclusion, and IiWRKY34 expression can drive carbon flux to specific overaccumulations of phenylpropanoids including flavonoids and lignans (Fig. 6). However, the content of the downstream compounds in the biosynthetic pathway of macrolcyclopropanediol did not change significantly, suggesting IiWRKY34 modulated the flux not through genetic regulation of the enzymatic steps involved in this pathway. Therefore, IiWRKY34 modifications on lignan accumulations are the combined outcome of a much more complex interplay of various metabolic pathways, not merely due to the activated phenylpropanoid biosynthetic steps.

To get insight into the specific molecular mechanism of IiWRKY34 for lignan biosynthesis, a IiWRKY34–lignan pathway genes–lignans network (Fig. 7C) was constructed based on transcript–metabolite correlation. Ii4CL3, which has been demonstrated as a hub gene for lignan biosynthesis, was found to be a potential target gene of IiWRKY34. EMSA and dual luciferase assays demonstrated that IiWRKY34 indeed activated the transcription of Ii4CL3 by binding to the promoter (Fig. 7D and F). These results indicate that IiWRKY34 modulates lignan biosynthesis, at least in part, due to regulate Ii4CL3, revealing the regulatory network of IiWRKY34 for lignan biosynthesis is robust and the identified target genes are worthy to be intensively investigated.

5. Conclusions

Numerous genes that individually control plant growth, secondary metabolism and stress response have been identified. Compared with these genes, IiWRKY34 has large pleiotropic effects on an array of traits, including yield, lignan biosynthesis and stress tolerance, which are inferred to has contributed significantly to the high level of polyploidy vigor of I. indigotica. Strong expression of IiWRKY34 in tetraploid I. indigotica corresponded well with greater yield, higher lignan accumulation and enhanced stress tolerance of the tetraploids. The major effects of IiWRKY34 will prompt the possibility of this gene based molecular marker-assisted selection and transformation for the improvement of herbs instead of individually manipulating the component traits using multiple genes of small effects.

Acknowledgments

This work was sponsored by National Natural Science Foundation of China (Grant Nos. 31872665, 81874335 and 31670292) and Shanghai Rising-Star Program (18QB1402700, China).

Author contributions

Ying Xiao, Wansheng Chen and Lei Zhang planned and designed the research. Ying Xiao, Jingxian Feng, Qing Li, Yangyun Zhou, Qitao Bu, Junhui Zhou and Hexin Tan performed experiments. Ying Xiao, Jingxian Feng, YingboYang, Wansheng Chen and Lei Zhang analysed data. Ying Xiao and Jingxian Feng wrote the manuscript. All authors read and approved the final manuscript.

Conflicts of interest

The authors have no conflicts of interest to declare.

Appendix A. Supporting information

Supporting data to this article can be found online at https://doi.org/10.1016/j.apsb.2019.12.020.
References

1. Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZZ, Lee HS, et al. Understanding mechanisms of novel gene expression in polyploids. Trends Genet 2003;19:141–7.
2. Alix K, Gérard PR, Schwarzerche T, Heslop-Harrison J. Polyploidy and interspecific hybridization: partners for adaptation, speciation and evolution in plants. Ann Bot-London 2017;120:183–94.
3. Cheng F, Wu J, Cai X, Liang J, Freeling M, Wang X. Gene retention, fractionation and subgenome differences in polyploid plants. Nat Plants 2018;4:258–68.
4. Council NP. China pharmacopoeia. 2013 ed. Beijing: Beijing Chemical Industry; 2015.
5. Li B. Studies on active constituents and quality evaluation of Banlange. Shanghai: Second Military University; 2003.
6. Li J, Zhou B, Li C, Chen Q, Wang Y, Li Z, et al. Lariciresinol-4-O-β-D-glucopyranoside from the root of Isatis indigotica inhibits influenza virus-induced pro-inflammatory response. J Ethnopharmacol 2015;174:379–86.
7. Yang Z, Wang Y, Zheng Z, Zhao S, Zhao J, Lin Q, et al. Antiviral activity of Isatis indigotica root-derived clemastin B against human and avian influenza A and B viruses in vitro. Int J Mol Med 2013;31:867–73.
8. Qiao CZ. Studies on polyploid breeding of Isatis indigotica Fort. J Integr 1989;31:678–83.
9. Qiao C, Li H. Cultivation and popularization for tetraploidy strain of Isatis indigotica. J Chin Med Mater 1994;17:3–6.
10. Initiative AG. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 2000;408:796–815.
11. Lu B, Pan X, Zhang L, Huang B, Sun L, Li B, et al. A genome-wide comparison of genes responsive to autoploidy in Isatis indigotica using Arabidopsis thaliana Affymetrix genechips. Plant Mol Biol Report 2006;24:197–204.
12. Xiao Y, Yu X, Chen J, Di P, Chen W, Zhang L. ISSDD1, a gene responsive to autoploidy and environmental factors in Isatis indigotica. Mol Biol Rep 2010;37:987–94.
13. Pan X, Xiao Y, Wang Z, Zhang L, Tang K. Tetraploids Isatis indigotica are more responsive and adaptable to stresses than the diploid progenitor based on changes in expression patterns of a cold inducible ICPK1. Biologia 2008;63:535–41.
14. Lu B, Ding R, Zhang L, Yu X, Huang B, Chen W. Molecular cloning and characterization of a novel calcium-dependent protein kinase gene ICPK2 responsive to polyploidy from tetraploid Isatis indigotica. J Biochem Mol Biol 2006;39:667–71.
15. Lu B, Du Z, Ding R, Zhang L, Yu X, Liu C, et al. Cloning and characterization of a differentially expressed phenylalanine ammonia-lyase gene (UPAL) after genome duplication from tetraploid Isatis indigotica Fort. J Integr 2006;48:1439–49.
16. Zhou Y, Kang L, Liao S, Pan Q, Ge X, Li Z. Transcriptomic analysis reveals differential gene expressions for cell growth and functional secondary metabolites in induced autotetraploid of Chinese Woad (Isatis indigotica Fort.). PLoS One 2015;10:e16392.
17. Colinas M, Goossens A. Combinatorial transcriptional control of plant specialized metabolism. Trends Plant Sci 2018;17:S1360–85.
18. Chu Y, Xiao S, Su H, Liao B, Zhang J, Xu J, et al. Genome-wide characterization and analysis of BHLH transcription factors in Panax ginseng. Acta Pharm Sin B 2018;8:666–77.
19. Zhang Y, Xu Z, Ji A, Luo H, Song J. Genomic survey of bZIP transcription factor genes related to tanshinone biosynthesis in Salvia miltiorrhiza. Acta Pharm Sin B 2018;8:295–305.
20. Chen F, Hu Y, Vannozzi A, Wu K, Cai H, Qin Y, et al. The WRKY transcription factor family in model plants and crops. Crit Rev Plant Sci 2018;36:1–25.
21. Chen J, Xin D, Li Q, Xun Z, Gao S, Chen R, et al. Biosynthesis of the active compounds of Isatis indigotica based on transcriptomic sequencing and metabolites profiling. BMC Genomics 2013;14:857.
22. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402.
23. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res 2004;32:263–6.
24. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 1998;95:5857–64.
25. Letunic I, Doerks T, Bork P, Smart S. recent updates to the protein domain annotation resource. Nucleic Acids Res 2012;40:302–5.
26. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. Methods Mol Biol 1999;112:531.
27. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mewhild H, et al. Clustal W and clustal X version 2.0. Bioinformatics 2005;21:2947–8.
28. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGAS: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.
29. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factor. Trends Plant Sci 2000;5:199–206.
30. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. Meme suite: tools for motif discovery and searching. Nucleic Acids Res 2009;37:W202–8.
31. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 2010;39:D561–8.
32. Li Q, Chen J, Xiao Y, Di P, Zhang L, Chen W. The dirigent multigene family in Isatis indigotica: gene discovery and differential transcript abundance. BMC Genomics 2014;15:388.
33. Xiao Y, Ji Q, Gao S, Tan H, Chen R, Li Q, et al. Combined transcriptome and metabolite profiling reveals that iPLR1 plays an important role in lariciresinol accumulation in Isatis indigotica. J Exp Bot 2015;66:6259.
34. Saito K, Matsuda F. Metabolomics for functional genomics, systems biology, and biotechnology. Ann Rev Plant Biol 2010;61:463–89.
35. Wesley SV, Hellolulu CA, Smith NA, Wang MB, Rouse DT, Liu Q, et al. Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J 2001;27:581–90.
36. Chilton MD, Tepper DA, Petit A, David C, Cassedelbart F, Tempé J. Agrobacterium rhizogenes inserts T-DNA into the genomes of the host plant root cells. Nature 1982;295:432–4.
37. de Micco V, Aronne G. Combined histochemistry and autofluorescence for identifying lignin distribution in cell walls. Biotech Histochem 2007;82:209–16.
38. Wang F, Chen HW, Li QT, Wei W, Li W, Zhang WK, et al. GmWRKY27 interacts with GmMYB174 to reduce expression of GmNAC29 for stress tolerance in soybean plants. Plant J 2015;83:224–36.
39. Bates LS, Waldren RP, Teare ID. Rapid determination of free proline for water-stress studies. Plant Soil 1973;39:205–7.
40. Zhang Y, Yan YP, Wang ZZ. The Arabidopsis PAPI transcription factor plays an important role in the enrichment of phenolic acids in Salvia miltiorrhiza. J Agric Food Chem 2010;58:12188–75.
41. Consortium TGO, Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. Gene ontology: tool for the unification of biology. Nat Genet 2000;25:25–9.
42. Draguchi S, Khati P, Tarca AL, Amin K, Done A, Voichita C, et al. A systems biology approach for pathway level analysis. Genome Res 2007;17:1537.
43. Lise J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc 2006;1:387–96.
44. Saeed AI, Sharow V, White J, Li J, Liang W, Bhagabati N, et al. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 2003;34:374–8.
45. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using cytoscape. *Nat Protoc* 2007;2:2366–82.
46. Zhang F, Fu X, Lv Z, Lu X, Shen Q, Zhang L, et al. A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol Plant* 2015;8:163–75.
47. Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, et al. Solution structure of an *Arabidopsis* WRKY DNA binding domain. *Plant Cell* 2005;17:944–56.
48. Guo C, Guo R, Xu X, Gao M, Li X, Song J, et al. Evolution and expression analysis of the grape (*Vitis vinifera*) WRKY gene family. *J Exp Bot* 2014;65:1513–28.
49. He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, et al. Genome-wide survey and characterization of the gene family in *Populus trichocarpa*. *Plant Cell Rep* 2012;31:1199–217.
50. Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, et al. Genome-wide analysis of WRKY gene family in *Cucumis sativus*. *BMC Genomics* 2011;12:471.
51. Sakamoto H, Maruyama K, Sakuma Y, Meshi T, Iwabuchi M, Shinozaki K, et al. *Arabidopsis* Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiol* 2004;136:2734–46.
52. Zhou C, Zhang L, Duan J, Miki B, Wu K. Histone deacetylase19 is involved in jasmonic acid and ethylene signaling of pathogen response in *Arabidopsis*. *Plant Cell* 2005;17:1196–204.
53. Xie Y, Li W, Guo D, Dong J, Zhang Q, Fu Y, et al. The *Arabidopsis* gene sigma factor-binding protein 1 plays a role in the salicylate- and jasmonate-mediated defence responses. *Plant Cell Environ* 2010;33:828–39.
54. Hano C, Addi M, Bensaddek L, Crè Nier D, Baltora-Rosset S, Doussot J, et al. Differential accumulation of monolignol-derived compounds in elicited flux (*Linum usitatissimum*) cell suspension cultures. *Planta* 2006;223:975–89.
55. Ashraf M, Foolad MR. Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environ Exp Bot* 2007;59:206–16.
56. Zhang L, Chen J, Zhou X, Chen X, Li Q, Tan H, et al. Dynamic metabolic and transcriptomic profiling of methyl jasmonate-treated hairy roots reveals synthetic characters and regulators of lignan biosynthesis in *Isatis indigotica* Fort. *Plant Biotechnol J* 2016;14:2217–27.
57. Jiang Y, Deyholos MK. Functional characterization of *Arabidopsis* NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant Mol Biol* 2009;69:91–105.
58. Rushton DL, Tripathi P, Babara RC, Lin J, Ringler P, Boken AK, et al. WRKY transcription factors: key components in abscisic acid signaling. *Plant Biotechnol J* 2011;10:2–11.
59. Pauwels L, Morreel K, Witte ED, Lammertyn F, Montagu MV, Boerjan W, et al. Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *P Natl Acad Sci USA* 2008;105:1380–5.
60. Zheng Z, Mosher SL, Fan B, Klessig DF, Chen Z. Functional analysis of *Arabidopsis* WRKY25 transcription factor in plant defense against *Pseudomonas syringae*. *BMC Plant Biol* 2007;7:2.
61. Bush DS. Calcium regulation in plant cells and its role in signaling. *Annu Rev Plant Physiol Plant Mol Biol* 1995;46:95–122.
62. Gechev TS, Hille J, Woerdemah BJ, Benina M, Melterov N, Toneya V, et al. Natural products from resurrection plants: potential for medical applications. *Biotechnol Adv* 2014;32:1091–101.
63. Xu Y, Wang J, Wang S, Wang J, Chen X. Characterization of GoWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-δ-cadinene synthase-A. *Plant Physiol* 2004;135:507–15.
64. Schluttenhofer C, Pattanaik S, Patra B, Yuan L. Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. *BMC Genomics* 2014;15:502.