

Genome-wide identification and analysis of the WRKY gene family and low-temperature stress response in *Prunus sibirica*

Quangang Liu

Shenyang Agricultural University

Shipeng Wang

Shenyang Agricultural University

Jiaxing Wen

Shenyang Agricultural University

Jianhua Chen

Shenyang Agricultural University

Yongqiang Sun

Shenyang Agricultural University

Shengjun Dong (✉ dsj928@163.com)

Shenyang Agricultural University

Research Article

Keywords: Genome-wide identification, WRKY transcription factor, *Prunus sibirica*, expression patterns, low-temperature stress

Posted Date: May 11th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2900512/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on June 27th, 2023. See the published version at <https://doi.org/10.1186/s12864-023-09469-0>.

Abstract

Background: *WRKY* transcription factors are a prominent gene family in plants, playing a crucial role in various biological processes including development, metabolism, defense, differentiation, and stress response. Although the *WRKY* gene family has been extensively studied and analysed in numerous plant species, research on *Prunus sibirica*'s *WRKY* genes (*PsWRKY*) remains lacking.

Results: This study analysed the basic physicochemical properties, phylogeny, gene structure, cis-acting elements, and Gene ontology (GO) annotation of *PsWRKY* gene family members using bioinformatics methods based on the whole-genome data of *P. sibirica*. In total, 55 *WRKY*s were identified in *P. sibirica* and were heterogeneously distributed on eight chromosomes. Based on the phylogenetic analysis, these *WRKY*s were classified into three major groups: Group I, Group II (II-a, II-b, II-c, II-d, II-e), and Group III. Members of different subfamilies have different cis-acting elements, conserved motifs, and intron-exon structures, indicating functional heterogeneity of the *WRKY* family. Prediction of subcellular localisation indicated that *PsWRKY*s were mainly located in the nucleus. Twenty pairs of duplicated genes were identified, and segmental duplication events may play an important role in *PsWRKY* gene family expansion. Analysis of the Ka/Ks ratio showed that the *PsWRKY* family's homologous genes were primarily purified by selection. Additionally, GO annotation analysis showed that the *WRKY* gene family was mainly involved in responses to stimuli, immune system processes, and reproductive processes. Furthermore, quantitative real-time PCR (qRT-PCR) analysis showed that 23 *PsWRKY*s were highly expressed in one or more tissues (pistils and roots) and *PsWRKY*s showed specific expression patterns under different low-temperature stress conditions.

Conclusions: Our results provide a scientific basis for the further exploration and functional validation of *WRKY*s in *P. sibirica*.

1. Background

WRKY transcription factors are an essential family of transcriptional regulators in plants, with highly conserved protein structural domains [1–3]. The *WRKY* structural domain is composed of nearly 60 amino acid residues with an N-terminal DNA-binding activity-related WRKYGQK conserved structural domain (forming a β -strand). Additionally, the C-terminus contains zinc finger C_2H_2 ($Cx_4-5Cx_{22-23}HxH$) or C_2HC ($Cx_7Cx_{23}HxC$) structural motifs that are involved in zinc finger protein interactions [3–5]. *WRKY* transcription factors are classified into three groups (I, II, and III) based on the number of conserved *WRKY* structural domains and zinc finger types [5, 6]. Group I contains two *WRKY* structural domains and the $Cx_4-5Cx_{22-23}HxH$ zinc-finger motif; Group II contains one *WRKY* structural domain and the $Cx_4-5Cx_{22-23}HxH$ zinc-finger motifs, which can be further divided into five subgroups (II a–e) depending on their phylogenetic branching and assembly [7, 8]; and Group III contains one *WRKY* structural domain and the $Cx_7Cx_{23}HxC$ zinc-finger motifs [3].

The *WRKY* family has several specific biological functions owing to their distinctive structural domain. Numerous studies have demonstrated that *WRKY* transcription factors perform crucial regulatory roles in various aspects of plant biology, including growth and development, physiological and biochemical processes, as well as responses to biotic and abiotic stress [9–12]. For example, the overexpression of *MuWRKY3* enhances drought resistance in transgenic peanuts [13], whereas the overexpression of *GhWRKY34* increases the resistance of transgenic *Arabidopsis* leaves and roots to salt stress by enhancing the ability of plants to selectively absorb Na^+ and K^+ and maintain lower Na^+/K^+ levels [14]. In addition, the regulatory functions of *WRKY*s are closely related to several phytohormone-mediated signalling pathways. *VqWRKY31* enhances resistance to powdery mildew in grapevine through activation of salicylic acid defense signalling as well as promotion of the synthesis of specific disease-resistant metabolites [15]. Cold tolerance is a plant response to abiotic stress, and the gene expression levels of plants are differentially altered under low-temperature stress. Numerous studies have shown that the *WRKY* gene family plays a crucial role in regulating the response to cold stress; for example, in *Prunus mume*, overexpression of *PmWRKY57* increases cold resistance in transgenic *Arabidopsis* by reducing malondialdehyde content, increasing peroxidase and superoxide dismutase activity, proline content, and upregulating cold-responsive genes [16]. Notably, the tolerance of rice to low-temperature stress was enhanced by the upregulation of *OsWRKY71* expression under these conditions [17]. In *Acer truncatum*, 13 genes, including *AtruWRKY12*, *AtruWRKY13*, *AtruWRKY15*, and *AtruWRKY17*, were overexpressed from 0 to 12 h after low-temperature (4°C) stress [18].

Prunus sibirica is a deciduous shrub belonging to the Rosaceae family, with a natural distribution concentrated in Eastern Siberia, Eastern and Southeastern Mongolia, and Northern and Northeastern China [19]. *P. sibirica* is an excellent tree species that combines ecological and economic benefits and can grow normally in poor conditions including drought, infertile soil, cold, and sandy winds. One of the most important tree species for the afforestation of barren mountains that can contribute to improving the environment by controlling various factors such as wind, sand, and water conservation [20]. *P. sibirica* seeds have rich economic value and are used in many areas including food, medicine, and industry [21]. However, frost damage during flowering limits the development of the *P. sibirica* industry. *P. sibirica* is an early spring-flowering plant that is susceptible to severe cold and frost during flowering, resulting in flower and fruit damage, reduced fruit yield, and even crop failure, leading to huge losses in production [22]. Therefore, mining and identification of potential genes related to frost resistance in the flowering organs are important for the development of frost-resistant and productive varieties of *P. sibirica*. Recent studies on *P. sibirica* have primarily concentrated on its molecular markers, genetic diversity, and almond composition [23–28]. However, there is a scarcity of research on its gene

family. The publication of the complete genome sequence of *P. sibirica* is a significant milestone in molecular biology research and is of vital importance to further study the regulatory mechanisms of genes related to low-temperature stress, collecting cold resistance gene resources, using molecular breeding to enhance the cold resistance of *P. sibirica* and other forest tree species, and increase fruit production and quality.

In the present study, the members of the *PsWRKY* gene family were identified based on genome-wide data from *P. sibirica*, and comprehensive analyses were performed in terms of the phylogenetic tree, structural features, chromosomal localisation, conserved domains, covariance, and cis-acting elements, which can provide a theoretical foundation for further studies on the gene functions and molecular evolutionary mechanisms of the *WRKY* gene family in *P. sibirica*. Furthermore, changes in the expression levels of *PsWRKY* family members in various tissues and under low-temperature stress were investigated, revealing the mechanisms of gene expression regulation in response to low-temperature stress during the growth and development of *P. sibirica*, and providing a theoretical basis for subsequent intensive studies on *WRKYs*.

2. Results

2.1. Identification and Characterization of *PsWRKY* in *P. sibirica*

P. sibirica *WRKYs* were searched using the Hidden Markov Model (HMM) of the *WRKY* structural domain (PF03106), and 55 *WRKYs* were identified (Additional file 1: Table S1). Further analysis using Pfam and NCBI-CDD confirmed that all proteins contained the full *WRKY* domain. We named these genes *PsWRKY1–PsWRKY55* according to their respective positions on the chromosomes.

The physicochemical properties of the proteins were analysed, and it was found that the number of amino acids encoded by the 55 *PsWRKYs* varied greatly, ranging from 163 (*PsWRKY8*) to 697 (*PsWRKY40*). The relative molecular weights ranged from 18526.84 (*PsWRKY27*) to 75002.92 (*PsWRKY40*) with an average of 42431.12 Da. The theoretical isoelectric point (pI) between 4.78 (*PsWRKY39*) and 10.33 (*PsWRKY5*), and 29 and 26 *PsWRKY* members were acidic (<7) and basic (>7), respectively. The hydrophobicity results were <0, indicating that all 55 *PsWRKY* members were hydrophilic proteins with stable performance. Identification of the instability index showed that all proteins except *PsWRKY39* (<40) were unstable, with *PsWRKY8* having the strongest hydrophilicity of -1.164. The maximum and minimum aliphatic index values were 79.07 (*PsWRKY55*) and 42.19 (*PsWRKY42*), respectively. The predicted subcellular localization revealed that 51 *PsWRKYs* were found in the nucleus; *PsWRKY14*, *PsWRKY16*, and *PsWRKY45* were located in the peroxisomes; and only *PsWRKY39* was located in the chloroplast.

2.2. Multiple Sequence Alignment and Phylogenetic Analysis of *PsWRKY* Family

The amino acid sequences of the *WRKY* structural domains of the *PsWRKYs* were analysed using multiple sequence alignment, and the characteristics of the *WRKY* structural domains of individual *PsWRKYs* were determined. The results showed (Fig. 1) that most of the core sequences of *WRKY* contain complete or nearly complete domains, and individual amino acids of the conserved motifs of *PsWRKYs* have undergone specific mutations and evolution, among which *PsWRKY8* and *PsWRKY42* underwent single amino acid mutations, and the conserved motif changed from "WRKYGQK" to "WRKYGKK". Furthermore, two conserved motifs at the N-terminal and C-terminal ends of *PsWRKY39* were changed to "WRKYEQK" and "WKKYGTK", respectively.

To investigate the affinities and biological connections among members of the *P. sibirica* *WRKY* family, we selected 71 *AtWRKYs* from *Arabidopsis thaliana* as references. Phylogenetic analysis of *P. sibirica* and *A. thaliana* *WRKYs* was performed using MEGA 11.0, and the classification of *AtWRKYs* gene families was referred to further classify *PsWRKY* into three categories: Group I, Group II, and Group III, and Group II was further divided into (II-a, II-b, II-c, II-d, and II-e). Group I had 11 members with two *WRKY* domains, Group II had 37 members with one *WRKY* domain, of which Group II-d had one *WRKY* structural domain in addition to the Pant Zn Cluster transcription factor structural domain. Finally, Group III had seven members with one *WRKY* domain (Fig. 2).

2.3. Gene Structure and Motif Composition of *PsWRKYs*

Since the variety of gene structures can indicate the evolution of gene families, we performed a phylogenetic tree construction of 55 *PsWRKYs* (Fig. 3 A) and analysed the intron/exon structures and conserved motifs of each *PsWRKY* gene to gain a more in-depth understanding of the *P. sibirica* *WRKY* family. It was observed that *PsWRKY3*, *PsWRKY14*, *PsWRKY16*, *PsWRKY37*, and *PsWRKY36* contained only two exons, whereas *PsWRKY39*, *PsWRKY29*, *PsWRKY30* and *PsWRKY34* contained most exons (Fig. 3 C). The number of exons within *PsWRKYs* varied between two and six, with most gene members containing three exons. To identify conserved motifs in all 55 *PsWRKYs* and gain a better understanding of *PsWRKY* diversification, we utilized the MEME v5.1.1 program. Our analysis resulted in the annotation of 10 predicted motifs. Figure 3 D displays detailed information on the ten motifs. As shown in Fig. 3 B, the *PsWRKY* groups had highly similar conserved motifs. In Group I, 10 of the 11 proteins had structures with motifs 3, 4, 1, and 2 in tandem, whereas one protein (*PsWRKY39*) lacked motif 2. Motif 9 was unique to *PsWRKY17*, *PsWRKY47*, *PsWRKY48*, and *PsWRKY49* in group III, while motif 3 was distributed only in group I. Motifs 6, 7, and 8 were distributed in group II-b, indicating that they were highly conserved among *PsWRKYs*.

2.4. Chromosomal localization, duplication, Synteny, and Ka/Ks Analysis of *PsWRKYs*

A chromosomal location map was constructed using the genomic information of *P. sibirica* to illustrate the distribution of *PsWRKY* members on chromosomes (Fig. 4). *PsWRKYs* were randomly distributed on eight chromosomes, with most genes concentrated near the caudal middle end of Chr6 (21.8 %). Their distribution on Chr1, Chr2, Chr3, Chr4, and Chr5 was 20 %, 14.5 %, 16.36 %, 10.9 %, and 9.1 %, respectively. In addition, chromosome 8 contained 5.5 % of *PsWRKYs*, and the least *PsWRKYs* were distributed on chromosome 7 (1.8 %).

Gene replication patterns (tandem and segmental replication) were analysed using MCScanX. Among the *PsWRKY* gene families, it was found that (*PsWRKY6*, *PsWRKY7*), (*PsWRKY47*, *PsWRKY48*), and (*PsWRKY48*, *PsWRKY49*) located in Chr1 were tandem repeat replicates (Fig. 5, Table 1). Additionally, 17 pairs of segmental replication events involving 27 *PsWRKYs* were detected. These findings indicate that these gene clusters were formed due to gene duplication. The *PsWRKY* gene family underwent a few tandem duplication events during evolution, while segmental duplication events played a major role in driving its evolution. In order to assess whether these homologous *PsWRKY* underwent selection pressure, we calculated Ks, Ka, and Ka/Ks. Tandem duplication yielded Ka/Ks ratios of 0.2531–0.3812 and 0.1149–0.3493 for segmental duplication (Fig. 5 and Table 1). Except for three co-dominant *PsWRKY* gene pairs, (*PsWRKY9*, *PsWRKY43*), (*PsWRKY1*, *PsWRKY36*), and (*PsWRKY24*, *PsWRKY32*), for which Ka/Ks values were not available, the Ka/Ks ratios of all co-dominant *PsWRKY* gene pairs were less than 1, indicating that these gene pairs were purified and screened.

To gain further insight into the evolutionary relationships among *WRKY* members in different plant species, a syntenic relationships were traced between *PsWRKYs* and their homologues in other species, including *A. thaliana*, *Prunus. salicina*, *P. mume*, *Prunus. persica*, and *Prunus. avium*. The least collinear gene pairs (64 pairs) were between *P. sibirica* and *A. thaliana*, and the most collinear gene pairs (94 pairs) were between *P. sibirica* and *P. salicina*, 86 pairs between *P. mume*, 89 pairs between *P. persica*, and 87 pairs between *P. avium* (Fig. 6).

Table 1 Segmental replication and tandem repeat replication covariate gene pair Ka/Ks analysis of *PsWRKYs*

Gene pair name		Gene pair ID		Ka	Ks	Ka/Ks	Duplications type
PsWRKY2	PsWRKY27	PaF106G0100001343.01	PaF106G0300014526.01	0.2657	1.9584	0.1357	segmental duplication
PsWRKY6	PsWRKY26	PaF106G0100004627.01	PaF106G0300013492.01	0.4076	1.2112	0.3365	segmental duplication
PsWRKY4	PsWRKY21	PaF106G0100003357.01	PaF106G0300011494.01	0.2659	1.3083	0.2032	segmental duplication
PsWRKY9	PsWRKY43	PaF106G0100005037.01	PaF106G0600024291.01	0.2851	Na	Na	segmental duplication
PsWRKY15	PsWRKY19	PaF106G0200009172.01	PaF106G0200010641.01	0.5049	2.8421	0.1776	segmental duplication
PsWRKY13	PsWRKY18	PaF106G0200009086.01	PaF106G0200010574.01	0.5506	3.0714	0.1793	segmental duplication
PsWRKY15	PsWRKY37	PaF106G0200009172.01	PaF106G0500019874.01	0.4716	1.5079	0.3128	segmental duplication
PsWRKY13	PsWRKY36	PaF106G0200009086.01	PaF106G0500019756.01	0.4608	Na	Na	segmental duplication
PsWRKY18	PsWRKY36	PaF106G0200010574.01	PaF106G0500019756.01	0.4946	4.3032	0.1149	segmental duplication
PsWRKY17	PsWRKY47	PaF106G0200010134.01	PaF106G0600025042.01	0.5329	2.5901	0.2057	segmental duplication
PsWRKY22	PsWRKY34	PaF106G0300012051.01	PaF106G0400018245.01	0.3184	2.5607	0.1243	segmental duplication
PsWRKY24	PsWRKY32	PaF106G0300012509.01	PaF106G0400017434.01	0.3285	Na	Na	segmental duplication
PsWRKY22	PsWRKY30	PaF106G0300012051.01	PaF106G0400015665.01	0.3131	1.9535	0.1603	segmental duplication
PsWRKY20	PsWRKY38	PaF106G0300011395.01	PaF106G0500020691.01	0.4278	1.2246	0.3493	segmental duplication
PsWRKY34	PsWRKY30	PaF106G0400018245.01	PaF106G0400015665.01	0.4189	2.0275	0.2066	segmental duplication
PsWRKY41	PsWRKY46	PaF106G0600021917.01	PaF106G0600024936.01	0.2939	1.4943	0.1967	segmental duplication
PsWRKY50	PsWRKY54	PaF106G0600025642.01	PaF106G0800029773.01	0.2719	1.3541	0.2010	segmental duplication
PsWRKY6	PsWRKY7	PaF106G0100004627.01	PaF106G0100004628.01	0.6564	2.5930	0.2531	tandem duplication
PsWRKY47	PsWRKY48	PaF106G0600025042.01	PaF106G0600025043.01	0.2629	0.4680	0.5618	tandem duplication
PsWRKY48	PsWRKY49	PaF106G0600025043.01	PaF106G0600025044.01	0.1044	0.2739	0.3812	tandem duplication

Note: Ka, non-synonymous substitution rate; Ks, synonymous substitution rate; Ka/Ks, selection pressure ratio; Na, no result

2.5. Analysis of *PsWRKY* cis-acting elements

In order to determine the specific types and distribution of cis-acting elements present in the promoters of *PsWRKYs*, an analysis of promoter cis-acting elements was conducted using PlantCARE. From the predicted results, 24 representative cis-acting elements were selected for functional analysis (Figs. 7 and S1). Cis-acting elements are classified into three major categories according to their regulatory functions: plant growth/development, phytohormone responses, and biotic/abiotic stress responses. The first category of cis-acting elements for plant growth and development prevalent in the promoter region included the GCN4_motif, MRE, Box-4, CAT-box, O2-site, circadian rhythm, G-box, GT1-motif, TCT-motif, GATA-motif, and I-box (Fig. 7 A, B). Among these, the GCN4_motif, O2-site, Circadian rhythm, and CAT box are involved in endosperm expression, metabolic regulation of maize alcohol-soluble proteins, circadian rhythm control, and meristematic tissue expression, respectively. In addition, the promoter regions of almost all *PsWRKYs* contained at least one light-response element, with the highest percentage in G-box (32 %), followed by Box-4, GT1-motif, TCT-motif, GATA-motif, MRE, and I-box (Fig. 7 B). Among the second class of phytohormone response elements, TGA-element (5 %) and AuxRR-core (2 %), were involved in growth hormone response, TATC-box (2 %) and P-box (4 %) in gibberellin response, ABRE (37 %) in abscisic acid response, TCA-element (7 %) in salicylic acid response, and CGTCA- motif (21 %) and TGACG-motif (21 %) in methyl jasmonate (MeJA) responsive (Fig. 7 B). The third category includes biotic/abiotic stress response elements like LTR, ARE, MBS, WUN-motif, and TC-rich repeats, which are associated with cold stress, anaerobic induction, drought stress, mechanical damage, defense, and stress responses, respectively. In *P. sibirica*, ARE was present in 51 *PsWRKYs* genes, whereas the Wun motif was the least numerous, and LTR elements were found in 20 promoters of *PsWRKYs*. *PsWRKY52* and *PsWRKY6* were strongly induced by cold stress and contained five and three LTR elements, respectively. Notably, no elements associated with either biotic or abiotic stress were found in the *PsWRKY48*. Taken together, *PsWRKYs* may be associated with light, hormone, stress responses, and growth and developmental pathways in *P. sibirica*.

2.6. Gene ontology annotation analysis of *PsWRKYs*

To better inform the transcription factor functions of *PsWRKYs*, the EggNog website was used for gene ontology (GO) annotation analysis. The 47 *PsWRKYs* were classified into 21 functional groups according to their protein sequence similarity and into three classes: biological processes, molecular functions, and cellular components (Fig. 8). Within the category of biological processes, the majority of *PsWRKYs* were engaged in the regulation of cellular processes (GO:0009987), biological processes (GO:0050789), biological (GO:0065007), and metabolic (GO:0008152) processes. Furthermore, six (*PsWRKY2*, *PsWRKY24*, *PsWRKY32*, *PsWRKY47*, *PsWRKY48*, and *PsWRKY49*) (GO:0032502), 11 (*PsWRKY3*, *PsWRKY4*, *PsWRKY6*, *PsWRKY7*, *PsWRKY17*, *PsWRKY19*, *PsWRKY26*, *PsWRKY32*, *PsWRKY47*, *PsWRKY48*, and *PsWRKY49*) (GO:0050896), six (*PsWRKY2*, *PsWRKY19*, *PsWRKY32*, *PsWRKY47*, *PsWRKY48*, and *PsWRKY49*) (GO:0002376), and six *PsWRKYs* (*PsWRKY31*, *PsWRKY32*, *PsWRKY40*, *PsWRKY44*, *PsWRKY51*, and *PsWRKY53*) (GO:0022414) were projected to be engaged in the developmental, response to stimuli, immune system response, and reproductive processes, including fruit and seed development, respectively. The molecular functions of *PsWRKYs* were associated mostly with 'binding' (GO:0005488) and 'transcription regulator activity' (GO:0140110). The cellular components of *PsWRKYs* include organelles (GO:0043226), cell parts (GO:0044464), and cells (GO:0005623).

2.7. Analysis of *PsWRKYs* expression in different tissues

To characterise the potential functions of *PsWRKYs*, we screened 23 *PsWRKYs* that could potentially be associated with plant growth development using promoter cis-acting elements, GO annotation analysis, and *WRKY* homeotic genes in other plants and examined their expression levels in different tissues of *P. sibirica* using quantitative real-time PCR (qRT-PCR) (Fig. 9). The 23 *PsWRKYs* showed different expression patterns in various tissues. Firstly, ten genes were highly expressed in the pistil, including *PsWRKY4*, *PsWRKY18*, *PsWRKY21*, *PsWRKY27*, *PsWRKY42*, *PsWRKY44*, *PsWRKY51*, *PsWRKY52*, *PsWRKY54*, and *PsWRKY55*. Among them, the expression of *PsWRKY4*, *PsWRKY44*, *PsWRKY51*, *PsWRKY52*, and *PsWRKY54* in the pistil was much higher than that in other tissues. Nine *PsWRKYs* (*WRKY1*, *WRKY5*, *WRKY6*, *WRKY9*, *WRKY10*, *WRKY12*, *WRKY13*, *WRKY14*, and *WRKY15*) were dominantly expressed in the roots. Notably, *PsWRKY6*, *PsWRKY10*, and *PsWRKY15* were 75-, 70-, and 70-fold more highly expressed, respectively, in roots than in petals. Third, in the leaf, the dominantly expressed *PsWRKYs* were *PsWRKY17*, *PsWRKY36*, *PsWRKY38*, and *PsWRKY49*.

2.8. Analysis of *PsWRKYs* expression under low-temperature stress

WRKY transcription factors play an important regulatory role in the response of plants to low-temperature stress. To investigate the expression pattern of *PsWRKYs* that may be associated with the low-temperature response, qRT-PCR was used to observe the expression levels of 23 *PsWRKYs* at different stages (0 h, 15 min, 30 min, 1 h, and 2 h) of low-temperature stress (-4 °C). The results showed that 23 *PsWRKYs* showed significant differences, with most of them showing high expression after 15 min and 1 h of low-temperature stress (Fig. 10). Among them, 10 genes were highly expressed after 15 min of low-temperature stress, including *PsWRKY1*, *PsWRKY9*, *PsWRKY10*, *PsWRKY13*, *PsWRKY14*, *PsWRKY21*, *PsWRKY27*, *PsWRKY42*, *PsWRKY51*, and *PsWRKY54*; 12 *PsWRKYs* (*PsWRKY4*, *PsWRKY5*, *PsWRKY6*, *PsWRKY12*, *PsWRKY15*, *PsWRKY17*, *PsWRKY18*, *PsWRKY38*, *PsWRKY44*, *PsWRKY49*, *PsWRKY52*, and *PsWRKY55*) were expressed highest after 1 h of low-temperature stress. Notably, the largest increase in expression was detected for *PsWRKY14* (approximately 30-fold) after 15 min of low-temperature stress, whereas *PsWRKY12* expression was upregulated 10-fold after 1 h of low-temperature stress. In contrast, the *PsWRKY51* gene was observed significantly upregulated at 0 h after low-temperature stress.

3. Discussion

WRKYs are one of the largest transcription regulatory factor families in plants, and play a crucial role in plant growth, development, and defense mechanisms [4]. In the present study, 55 *WRKYs* were identified throughout the genome of *P. sibirica* using bioinformatic analysis. The number of *PsWRKYs* was similar to the identified *WRKYs* in *P. mume* (58) [29], *P. persica* (58), and *Prunus armeniaca* (56) [30]. This indicates that *WRKY* members have similar evolutionary patterns among closely related species; however, it was significantly less than that of model plants, like *Arabidopsis* (72) [31] and tomato (88) [32], and the difference in number was most pronounced in Groups II-c and III, suggesting that *WRKY* members in these groups had more severe gene loss during evolution, similar to that observed in *P. mume* [29] and *P. armeniaca* [30] *WRKY* gene family

The predicted subcellular localisation of the 55 *PsWRKYs* indicated that most *WRKYs* (51) were located in the nucleus, indicating that their functions may be closely linked to the regulation of target gene expression. Notably, *PsWRKY14*, *PsWRKY16*, and *PsWRKY45* are localised in the peroxisome, and the peroxisome may play a crucial role in the cold stress response [33], so the role of these genes in cold stress in *P. sibirica* is of interest to us. Furthermore, *PsWRKY39* is localised in the chloroplast and is speculated to be involved in chloroplast formation, thereby affecting photosynthesis. Phylogenetic analysis based on *AtWRKYs* divided the 55 *PsWRKYs* into three groups, with the largest proportion of *PsWRKYs* in Group II (67%), suggesting that this taxon may have undergone gene duplication events during evolution in *Hypericum perforatum* [34], *Siraitia siamensis* [35] and *P. armeniaca* [30]. Group II was subdivided into five subgroups in which members II-e and II-d were concentrated in two adjacent branches of the same subgroup, as were members of II-a and II-b. This indicates a closer evolutionary relationship between them, consistent with the evolutionary findings of other studies on eggplant [36] and plum [37].

The *WRKY* transcription factor is named after its highly conserved sequence (WRKYGQK) which is crucial for its ability to recognize and bind to the W-box element in target gene promoters [3, 38]. Previous studies have identified variants of the conserved WRKYGQK motif in multiple species. Two variants (WKKYGKK and WKKYGEK) have been reported for *Taraxacum kok-saghyz* [39], while *Ginkgo biloba* has four variants: WRKYAQK, WRKYGEK, WRKYGRK, and WRKYGKK [40]. Moreover, three variants have been observed in *Kandelia obovata* (WRKYGEK, WRKYGKK, and WRKYGQK) [41] and one in *Cannabis sativa* (WRKYGKK) [42]. Variants of WRKYGKK were noted in *PsWRKY8* and *PsWRKY42* protein sequences of Group II-c. Variations in this conserved structural domain may affect the binding specificity of *WRKYs* to downstream target genes, resulting in altered gene functions [43]. For example, the *NtWRKY12* protein with a WRKYGKK motif in *Nicotiana tabacum* can bind to WK-box elements but not to W-box motifs. Further verification is required to determine whether *PsWRKY8* and *PsWRKY42* bind to the W-box [44]. Notably, variations in WKKYGTK and WRKYEQK were observed in the *PsWRKY39* protein sequence of Group I. WKKY variation is common in Asteraceae and has been observed in different legumes, but at a lower frequency. However, no WKKY taxa have been reported in other plant species [45]. Therefore, the functional and DNA-binding properties of *PsWRKY39* require further analysis.

Upon analysing the conserved structural domains showed that the number and type of conserved motifs differed between *PsWRKYs*. However, gene members of the same subgroup exhibited similar conserved motif types, indicating the presence of analogous structures and functions within the same *WRKY* group [46]. All *PsWRKYs* contained motifs 1 and 2, which could be key elements retained in evolution. Motif 9 is exclusive to Group III, while motifs 6, 7, and 10 are only found in Group II-b. The similarity between conserved motifs in II-a and II-b suggests a close evolutionary relationship. A similar motif composition within the same group or subgroup indicates that the proteins are structurally conserved and may have similar functions [47].

Gene duplication events are the primary drivers for acquiring new genes and enabling the expansion of an organism's gene family, which consists of two main forms: tandem and segmental duplication [48]. In this study, we found that among the 55 *PsWRKYs*, only three pairs of genes, (*PsWRKY6*; *PsWRKY7*), (*PsWRKY47*; *PsWRKY48*), and (*PsWRKY48*; *PsWRKY49*), showed tandem duplications. A total of 17 *PsWRKYs* had segmental duplications between them, which far outnumbered tandem duplications; these gene pairs were distributed on distinct chromosomes, indicating that segmental duplication at these positions may contribute to the extension of the *PsWRKY* family. Ka/Ks analysis of the identified homologous gene pairs showed that all *PsWRKY* gene family members underwent purifying selection, indicating that the *PsWRKY* gene family evolved to be highly conserved. Based on this, we constructed a collinearity map of the *PsWRKY* family in *A. thaliana*, *P. salicina*, *P. mume*, *P. persica*, and *P. avium*. *P. sibirica* and *A. thaliana*, and *P. salicina*, *P. mume*, *P. persica*, and *P. avium* there were 64, 92, 86, 89, and 87 collinear gene pairs, respectively. These results indicate that *P. sibirica* shares more collinear connections with other Rosaceae genomes compared to *A. thaliana*, implying a possible evolutionary relationship between them. The large number of collinear pairs shared between *P. sibirica* and *P. salicina* suggests that these two species have a comparatively close evolutionary relationship.

Transcription factors perform crucial functions in signal transduction and the induction of downstream functional gene expression by specifically binding to gene promoter regions and activating or repressing the transcription of downstream genes [49]. The promoter regions of *PsWRKYs* were found to contain numerous cis-elements with multiple potential regulatory functions. Among these cis-elements, some were overrepresented in the *PsWRKYs*, including G-box and Box 4, which regulate plant growth and development, and biotic/abiotic response-associated elements, such as ARE, MBS, and LTR, followed by involvement in phytohormone response elements, among which signalling molecule elements, such as ABRE, TGACG-motif, and CGTCA-motif, were the most prominent, sensing and transmitting signals in response to

external environmental changes, and maintaining plant homeostasis through hormonal networks [50]. These results suggest the functional diversity of *PsWRKYs*. An LTR cis-acting element was found in 20 *WRKYs*, among which *PsWRKY52* contained five LTR cis-acting elements, which simultaneously verified the adaptation of the gene to a freezing point environment.

Gene expression is closely related to gene function [51]. Research has demonstrated that the expression of *WRKYs* in specific tissues has a significant impact on the growth and development of plants by controlling the expression of genes engaged in growth and differentiation [52]. In this study, many *PsWRKYs* were constitutively expressed in the pistils and roots of *P. sibirica*, and numerous of these, for instance, *PsWRKY6*, *PsWRKY13*, *PsWRKY14*, *PsWRKY27*, *PsWRKY44*, *PsWRKY51*, *PsWRKY52*, and *PsWRKY54*, showing tissue-specific expression patterns, meaning that *PsWRKYs* play crucial roles in various aspects of plant development and exhibit distinct functions in different tissue types. *P. sibirica* is an early spring flowering species, and frost damage to its floral organs is a significant constraint to almond production. Among apricot floral organs, the pistil is the least resistant to frost and directly affects fruit yield [53]. In our study, *PsWRKYs* were predominantly expressed in the pistils, with 10 being highly regulated. *PsWRKY4*, *PsWRKY18*, *PsWRKY21*, *PsWRKY27*, *PsWRKY42*, *PsWRKY44*, *PsWRKY51*, *PsWRKY52*, *PsWRKY54*, and *PsWRKY55* were the predominant transcripts that accumulated in pistils, indicating that these *PsWRKYs* are essential in pistil growth and development. The results of GO annotation analysis confirmed the possible involvement of *PsWRKY44* and *PsWRKY51* in pistil development and fruit formation in *P. sibirica*. This was consistent with the results of previous studies on *Juglans regia* [54]. The root system plays a crucial role in the uptake of water and nutrients in plants. Consequently, it is often the first to react to stressful conditions such as drought, high salt levels, and other external factors. In this study, nine *PsWRKY* genes, including *PsWRKY1*, *PsWRKY5*, *PsWRKY6*, *PsWRKY9*, *PsWRKY10*, *PsWRKY12*, *PsWRKY13*, *PsWRKY14*, and *PsWRKY15*, were highly expressed in roots, suggesting that *P. sibirica WRKY* may play a specific role in root growth and development. Studies have suggested that *AtWRKY75* is implicated in regulating the nutrient starvation response and root development [55], and *AtWRKY46* inhibits ammonium efflux from the root elongation zone to promote ammonium tolerance in Arabidopsis [56]. The results of the phylogenetic analysis indicated that *PsWRKY14* exhibited a closer genetic relationship with *AtWRKY75*, while *PsWRKY15* was found to be more closely related to *AtWRKY46*, indicating that these *PsWRKYs* may be involved in regulating root growth and material uptake in *P. sibirica*; however, the related mechanisms need to be further studied.

The *WRKY* family has been linked to both biotic and abiotic stress responses [38]. Research has demonstrated the significant regulatory role of the *WRKY* family in plant responses to low-temperature stress. For example, in rice, *WRKY53* negatively affects cold tolerance during gestation by repressing GA transcription in anthers [57], and overexpression of *KoWRKY40* and *VvWRKY28* in Arabidopsis enhances cold tolerance in transgenic Arabidopsis [58, 59]. In our study, we examined the expression patterns of *PsWRKYs* in *P. sibirica* pistils under low-temperature stress. Our findings revealed that out of the 23 *PsWRKYs* analysed, most genes were upregulated in response to the stress, while only *PsWRKY36* was downregulated. Similar results have been found in *K. obovata* [41], *Coffea canephora* [47], and *P. mume* [37], suggesting that *PsWRKY* transcription factors may have positive or negative regulatory effects. Significantly, seven *PsWRKYs* (*PsWRKY4*, *PsWRKY6*, *PsWRKY12*, *PsWRKY14*, *PsWRKY18*, *PsWRKY49*, and *PsWRKY52*) were upregulated more than 2-fold after low-temperature stress, suggesting that these *PsWRKYs* are likely to have a positive function in the low-temperature stress response. *PsWRKY14* was upregulated by approximately 30-fold after 15 min of low-temperature stress, and its relative phenotype was higher in pistils, indicating that it might have a significant effect on pistil freezing resistance. Notably, the cis-acting element prediction results of *PsWRKY6* and *PsWRKY52* contained more LTR cis-acting elements, while GO annotation results showed that *PsWRKY4*, *PsWRKY6*, and *PsWRKY49* were engaged in the regulation of the stimulus response, demonstrating that these genes might have major roles in the freezing resistance of *P. sibirica* organs and could be used in future gene function validation. In addition, 10 *PsWRKYs* were highly expressed after 15 min of low-temperature treatment, and 12 *PsWRKYs* were highly expressed after 1 h of low-temperature stress, suggesting that these *PsWRKY* transcription factors may play different roles at different times during the stress response. Recent studies have shown that overexpression of *PmWRKY57* augments tolerance to cold stress in Arabidopsis [16], and sequence comparison has revealed that *PsWRKY5* and *PmWRKY57* are homologous genes, implying that *PsWRKY5* may have an essential function in the response to cold stress and may be associated with cold resistance in *P. sibirica*.

4. Conclusions

In summary, 55 *WRKYs* were identified in the *P. sibirica* genome, all of which are located on eight chromosomes. By sequence alignment and phylogenetic analysis, *PsWRKYs* were classified into six subclasses, and phylogenetic analysis; these subclasses shared similarities in their gene structures. Segmental duplication significantly contributed to the expansion of the *P. sibirica WRKY* gene family. Evolutionary difference analysis (Ka/Ks) indicated that *PsWRKYs* were strongly subjected to purifying selection during evolution. GO annotation results indicated that *WRKYs* were primarily involved in the regulation of molecular functions and biological processes, such as transcriptional regulatory activity, response to stimuli, immune system processes, reproductive processes, regulation of metabolic processes, and regulation of cellular processes. The qRT-PCR analysis revealed that 23 *PsWRKYs* were highly expressed in one or more tissues. The *PsWRKYs* were significantly overexpressed after 15 min and 1 h of low-temperature stress. The results presented in this study serve as a foundation for understanding the function of the *WRKY* gene in enhancing plant cold tolerance and the associated molecular mechanisms.

5. Materials and methods

5.1 Plant material and treatment

P. sibirica clone No. 453 (Appraisal of improved varieties of forest tree, S-SV-PS-002-2021, Liaoning Provincial Department of Forest and Grassland) from the National Forest Germplasm Resource Preservation Repository for *Prunus* species at Shenyang Agricultural University (Kazuo, Liaoning, China) was used as the experimental material. The leaf, stem, root, petal, and pistil tissues of *P. sibirica* clone No. 453 were sampled in the absence of mechanical damage or pests. Additionally, whole plants of *P. sibirica* clone No. 453 were placed in an artificial frost chamber at -4°C, and pistils were taken as samples after 0 h, 15 min, 30 min, 1 h, and 2 h of stress, with 0 h as control, placed in liquid nitrogen, and stored at -80°C. The sampling and stress treatment experiments were performed in triplicate.

5.2 Identification and Characterization of *PsWRKYs*

Psibirica genome data were downloaded from the Rosaceae genome database (<https://www.rosaceae.org/Analysis/10254124>), and *A. thaliana* WRKY protein sequences were downloaded from the Arabidopsis (*A. thaliana*) database (<https://www.arabidopsis.org/>). A Hidden Markov Model file (PF03106) for *WRKY* was downloaded from the Pfam database (<https://www.ebi.ac.uk/interpro/entry/pfam/PF03106/>). HMMER 3.0 (<http://hmmer.janelia.org/>) was used to screen for candidate *PsWRKYs*^[60]. Two online software programs, the Pfam database (<https://www.ebi.ac.uk/interpro/entry/pfam>) and NCBI-CDD (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>), were used to determine whether all candidate *PsWRKYs* contained the entire *WRKY* structural domain. ExPASy (<https://web.expasy.org/protparam/>) was used to calculate the amino acid length, molecular weight (MW), isoelectric point (pI), and instability index of all candidate *PsWRKYs*, whereas WoLF PSORT (<https://wolfpsort.hgc.jp>) was used to predict subcellular localisation [61, 62].

5.3 Phylogenetic analysis and multiple sequence alignment of *PsWRKYs*

MEGA 11.0 software with the Cluster W tool was used for multi-protein sequence alignment analysis of *WRKYs* from *P. sibirica* and *A. thaliana*, and a phylogenetic evolutionary tree was constructed using default parameter values and the neighbour-joining method with 1000 bootstrap iterations [63]. Subsequently, the ITOLS v6.7.2 program (<http://itol.embl.de>) was used for phylogenetic tree visualization [64]. Classification of *PsWRKY* into relevant taxa based on a previously reported classification of the *A. thaliana* *WRKY* gene family.

5.4 Conserved Motifs and Gene structure analysis of *PsWRKYs*

The TBtools v1.089 software was used to analyse and visualise the exon-intron structure of *PsWRKY*. The MEME v5.1.1 online program (<https://meme-suite.org/meme/tools/meme>) was used to predict the conserved motifs of *WRKY* proteins with the number of motifs set to 10 [65].

5.5 Chromosomal localization, gene duplication, and covariance analysis of *PsWRKYs*

The *P. sibirica* genomic data annotation file was downloaded to determine chromosome length and gene location based on *PsWRKY* member IDs. TB tools were used to map the distribution of chromosomal localisation. MCScanX [66] was used to check for gene duplication events using default parameters. Dual systemic plots using TBtools for *P. sibirica* with *A. thaliana*, *P. salicina*, *P. mume*, and *P. persica* *WRKYs* interspecies covariance were analysed, and intraspecific gene duplicates of *P. sibirica* were analysed using Advanced Circos from TBtools. The Ka Ks_Calculator tool was used to calculate the frequency of synonymous mutation rate (Ks), frequency of non-synonymous mutation rate (Ka), and the ratio of nonsynonymous mutation rate to synonymous mutation rate (Ka/Ks) for *WRKY* duplicated genes in *P. sibirica*.

5.6 Promoter *Cis*-Acting Element and GO annotation analysis of *PsWRKYs*

The PlantCARE database (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) was used to predict cis-acting elements of the 2000 bp promoter sequence of *PsWRKYs* [67]. Egglog 5.0 (<http://egglog5.embl.de/>) was used for GO annotation and the results were visualized using WeGo 2.0 (<https://wego.genomics.cn/>).

5.7 RNA isolation, and qRT-PCR validation

The primers were done using Primer Premier 5.0; their synthesis was entrusted to GENEWIZ (Suzhou, China), and 18S rRNA served as a reference gene [68] (Table S2). Total RNA was extracted from the samples using an RNAprep Pure Plant Kit (Tiangen, Beijing, China), and cDNA was obtained using a FastKing RT Kit (Tiangen), according to the manufacturer's instructions. The qRT-PCR was performed on a StepOne Real-Time PCR System (Applied Biosystems) using SuperReal PreMix Plus (SYBR Green, Tiangen). The PCR reaction program consisted of 95°C for 30 s, followed by 40 cycles of 95°C for 10 s, 60°C for 30 s, 95°C for 15 s, and 60°C for 1 min, finishing with 95°C for 15 s. The experiment was conducted three times and the relative expression level was determined using the $2^{-\Delta\Delta CT}$ method. [69]. Statistical analysis was performed using SPSS software version 26.0. Bar graphs were created using GraphPad Prism 8.4.0 software.

Abbreviations

Ka: Non-synonymous substitution rate

Ks: Synonymous substitution rate

Ka/Ks: Ratio of the non-synonymous substitution rate to the synonymous substitution rate

qRT-PCR: Quantitative real-time PCR

GO: Gene ontology

Declarations

Ethics approval and consent to participate.

Plant material used in the study complies with relevant institutional, national, and international guidelines and legislation. Plant material used is not wild species and also permission for the plant material is not required.

Consent for publication

Not applicable.

Availability of data and materials

The WRKY domain HMM (Hidden Markov Model) profile numbered PF03106 was extracted from the Pfam protein family database (<https://www.ebi.ac.uk/interpro/entry/pfam/PF03106/>). *Psibirica* genome data were downloaded from the Rosaceae genome database (<https://www.rosaceae.org/Analysis/10254124>). The AtWRKY sequences were obtained from the Arabidopsis (*A. thaliana*) database (<https://www.arabidopsis.org/>). The datasets analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the Science and Technology Project Program of Liaoning Provincial Science and Technology Department (2021-BS-139) and China Postdoctoral Science Foundation (2022MD723804).

Authors' contributions

QGL and SPW conceived and designed the experiments, performed the experiments, analysed the data, prepared figures and/or tables, wrote the manuscript. JHC and JXW conducted the experiments and analysed the data. YQS participated in the research work and partially designed the study. SJD contributed to proofreading and critical review of this manuscript. All authors read and approved the final manuscript.

Acknowledgements

We appreciate the Key Laboratory for Silviculture of Liaoning Province in order to provide the necessary facilities.

Author's information

¹College of Forestry, Shenyang Agricultural University, Shenyang, China.

References

1. Ding ZJ, Yan JY, Li CX, Li GX, Wu YR, Zheng SJ. Transcription factor WRKY46 modulates the development of Arabidopsis lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. *Plant J.* 2015;84:56–69.
2. Ding ZJ, Yan JY, Li GX, Wu ZC, Zhang SQ, Zheng SJ. WRKY41 controls Arabidopsis seed dormancy via direct regulation of *ABI3* transcript levels not downstream of ABA. *Plant J.* 2014;79:810–23.
3. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends Plant Sci.* 2010;15:247–58.
4. Jiang J, Ma S, Ye N, Jiang M, Cao J, Zhang J. WRKY transcription factors in plant responses to stresses: WRKY in plant responses to stresses. *J Integr Plant Biol.* 2017;59:86–101.

5. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 2000;5:199–206.
6. Viana VE, Carlos Da Maia L, Busanello C, Pegoraro C, Costa De Oliveira A. When rice gets the chills: comparative transcriptome profiling at germination shows WRKY transcription factor responses. *Plant Biol J.* 2021;23:100–12.
7. Bao W, Wang X, Chen M, Chai T, Wang H. A WRKY transcription factor, PcWRKY33, from *Polygonum cuspidatum* reduces salt tolerance in transgenic *Arabidopsis thaliana*. *Plant Cell Rep.* 2018;37:1033–48.
8. Qin Y, Tian Y, Liu X. A wheat salinity-induced WRKY transcription factor TaWRKY93 confers multiple abiotic stress tolerance in *Arabidopsis thaliana*. *Biochem Biophys Res Commun.* 2015;464:428–33.
9. Xie L, Yan T, Li L, Chen M, Ma Y, Hao X, et al. The WRKY transcription factor AaGSW2 promotes glandular trichome initiation in *Artemisia annua*. *J Exp Bot.* 2021;72:1691–701.
10. Zhang L, Wan X, Xu Y, Niyitanga S, Qi J, Zhang L. De novo assembly of transcriptome and genome-wide identification reveal GA₃ stress-responsive WRKY transcription factors involved in fiber formation in jute (*Corchorus capsularis*). *BMC Plant Biol.* 2020;20:403.
11. Zhu H, Jiang Y, Guo Y, Huang J, Zhou M, Tang Y, et al. A novel salt inducible WRKY transcription factor gene, AhWRKY75, confers salt tolerance in transgenic peanut. *Plant Physiol Biochem.* 2021;160:175–83.
12. Wang D, Wang L, Su W, Ren Y, You C, Zhang C, et al. A class III WRKY transcription factor in sugarcane was involved in biotic and abiotic stress responses. *Sci Rep.* 2020;10:20964.
13. Kiranmai K, Lokanadha Rao G, Pandurangaiah M, Nareshkumar A, Amaranatha Reddy V, Lokesh U, et al. A Novel WRKY Transcription Factor, MuWRKY3 (*Macrotyloma uniflorum* Lam. Verdc.) Enhances Drought Stress Tolerance in Transgenic Groundnut (*Arachis hypogaea* L.) Plants. *Front Plant Sci.* 2018;9:346.
14. Zhou L, Wang N-N, Gong S-Y, Lu R, Li Y, Li X-B. Overexpression of a cotton (*Gossypium hirsutum*) WRKY gene, GhWRKY34, in *Arabidopsis* enhances salt-tolerance of the transgenic plants. *Plant Physiol Biochem.* 2015;96:311–20.
15. Yin W, Wang X, Liu H, Wang Y, Van Nocker S, Tu M, et al. Overexpression of *VqWRKY31* enhances powdery mildew resistance in grapevine by promoting salicylic acid signaling and specific metabolite synthesis. *Hortic Res.* 2022;9:uhab064.
16. Wang Y, Dong B, Wang N, Zheng Z, Yang L, Zhong S, et al. A WRKY Transcription Factor PmWRKY57 from *Prunus mume* Improves Cold Tolerance in *Arabidopsis thaliana*. *Mol Biotechnol.* 2022. <https://doi.org/10.1007/s12033-022-00645-3>.
17. Kim C-Y, Vo KTX, Nguyen CD, Jeong D-H, Lee S-K, Kumar M, et al. Functional analysis of a cold-responsive rice WRKY gene, OsWRKY71. *Plant Biotechnol Rep.* 2016;10:13–23.
18. Li Y, Li X, Wei J, Cai K, Zhang H, Ge L, et al. Genome-Wide Identification and Analysis of the WRKY Gene Family and Cold Stress Response in *Acer truncatum*. *Genes.* 2021;12:1867.
19. Niu J, Zhu B, Cai J, Li P, Wang L, Dai H, et al. Selection of Reference Genes for Gene Expression Studies in Siberian Apricot (*Prunus sibirica* L.) Germplasm Using Quantitative Real-Time PCR. *PLoS ONE.* 2014;9:e103900.
20. Ma Y, Wang S, Liu X, Yu H, Yu D, Li G, et al. Oil content, fatty acid composition and biodiesel properties among natural provenances of Siberian apricot (*Prunus sibirica* L.) from China. *GCB Bioenergy.* 2021;13:112–32.
21. Chen J, Zhang J, Liu Q, Wang X, Wen J, Sun Y, et al. Mining for genes related to pistil abortion in *Prunus sibirica* L. *PeerJ.* 2022;10:e14366.
22. Xu W, Bao W, Liu H, Chen C, Bai H, Huang M, et al. Insights Into the Molecular Mechanisms of Late Flowering in *Prunus sibirica* by Whole-Genome and Transcriptome Analyses. *Front Plant Sci.* 2022;12:802827.
23. Wen J, Chen J, Sun Y, Liu Q, Jin L, Dong S. Association mapping of major economic traits and exploration of elite alleles in *Prunus sibirica*. *Euphytica.* 2023;219:39.
24. Wang Z, Liu H, Liu J, Li Y, Wu R, Pang X. Mining new microsatellite markers for Siberian apricot (*Prunus sibirica* L.) from SSR-enriched genomic library. *Sci Hort.* 2014;166:65–9.
25. Liu H-B, Liu J, Wang Z, Ma L-Y, Wang S-Q, Lin X-G, et al. Development and Characterization of Microsatellite Markers in *Prunus sibirica* (Rosaceae). *Appl Plant Sci.* 2013;1:1200074.
26. Fan S, Liang T, Yu H, Bi Q, Li G, Wang L. Kernel characteristics, oil contents, fatty acid compositions and biodiesel properties in developing Siberian apricot (*Prunus sibirica* L.) seeds. *Ind Crops Prod.* 2016;89:195–9.
27. Yin M, Wuyun T, Jiang Z, Zeng J. Amino acid profiles and protein quality of Siberian apricot (*Prunus sibirica* L.) kernels from Inner Mongolia. *J For Res.* 2020;31:1391–7.
28. Buer H, Rula S, Wang ZY, Fang S, Bai Y. Analysis of genetic diversity in *Prunus sibirica* L. in inner Mongolia using SCoT molecular markers. *Genet Resour Crop Evol.* 2022;69:1057–68.
29. Chen M, Tan Q, Sun M, Li D, Fu X, Chen X, et al. Genome-wide identification of WRKY family genes in peach and analysis of WRKY expression during bud dormancy. *Mol Genet Genomics.* 2016;291:1319–32.

30. Jing C, Wang D, Liu Z, Chen X, Dong H, Wu X. Identification of the WRKY gene family in apricot and its response to drought stress. *Hortic Environ Biotechnol.* 2023;1–14.
31. Berri S, Abbruscato P, Faivre-Rampant O, Brasileiro AC, Fumasoni I, Satoh K, et al. Characterization of WRKYco-regulatory networks in rice and Arabidopsis. *BMC Plant Biol.* 2009;9:120.
32. Liu G, Zhang D, Zhao T, Yang H, Jiang J, Li J, et al. Genome-wide analysis of the WRKY gene family unveil evolutionary history and expression characteristics in tomato and its wild relatives. *Front Genet.* 2022;13:962975.
33. Zhou P, Li X, Liu X, Wen X, Zhang Y, Zhang D. Transcriptome profiling of *Malus sieversii* under freezing stress after being cold-acclimated. *BMC Genomics.* 2021;22:681.
34. Zhou W, Yang S, Yang L, Xiao R, Chen S, Wang D, et al. Genome-Wide Identification of the *Hypericum perforatum* WRKY Gene Family Implicates HpWRKY85 in Drought Resistance. *IJMS.* 2022;24:352.
35. Mu D, Chen W, Shao Y, Wilson IW, Zhao H, Luo Z, et al. Genome-Wide Identification and Expression Analysis of WRKY Transcription Factors in *Siraitia siamensis*. *Plants.* 2023;12:288.
36. Yang Y, Liu J, Zhou X, Liu S, Zhuang Y. Identification of *WRKY* gene family and characterization of cold stress-responsive *WRKY* genes in eggplant. *PeerJ.* 2020;8:e8777.
37. Bao D, Cheng W. Genome-Wide Analysis of Members of the WRKY Gene Family and Their Cold Stress Response in *Prunus mume*. *Genes.* 2019;10:911.
38. Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol.* 2005;5:1.
39. Cheng Y, Luo J, Li H, Wei F, Zhang Y, Jiang H, et al. Identification of the WRKY Gene Family and Characterization of Stress-Responsive Genes in *Taraxacum kok-saghyz* Rodin. *IJMS.* 2022;23:10270.
40. Li W, Xiao N, Wang Y, Liu X, Chen Z, Gu X, et al. Genome-Wide Identification, Evolutionary and Functional Analyses of WRKY Family Members in *Ginkgo biloba*. *Genes.* 2023;14:343.
41. Du Z, You S, Zhao X, Xiong L, Li J. Genome-Wide Identification of WRKY Genes and Their Responses to Chilling Stress in *Kandelia obovata*. *Front Genet.* 2022;13:875316.
42. Wei H, Chen S, Niyitanga S, Liu T, Qi J, Zhang L. Genome-wide identification and expression analysis response to GA₃ stresses of WRKY gene family in seed hemp (*Cannabis sativa* L). *Gene.* 2022;822:146290.
43. Xie T, Chen C, Li C, Liu J, Liu C, He Y. Genome-wide investigation of WRKY gene family in pineapple: evolution and expression profiles during development and stress. *BMC Genomics.* 2018;19:490.
44. Van Verk MC, Pappaioannou D, Neeleman L, Bol JF, Linthorst HJM. A Novel WRKY Transcription Factor Is Required for Induction of *PR-1a* Gene Expression by Salicylic Acid and Bacterial Elicitors. *Plant Physiol.* 2008;146:1983–95.
45. Li J, Islam F, Huang Q, Wang J, Zhou W, Xu L, et al. Genome-wide characterization of WRKY gene family in *Helianthus annuus* L. and their expression profiles under biotic and abiotic stresses. *PLoS ONE.* 2020;15:e0241965.
46. Stegmaier P, Kel AE, Wingender E, Systematic. DNA-Binding Domain Classification of Transcription Factors. 2004.
47. Dong X, Yang Y, Zhang Z, Xiao Z, Bai X, Gao J, et al. Genome-Wide Identification of WRKY Genes and Their Response to Cold Stress in *Coffea canephora*. *Forests.* 2019;10:335.
48. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. *J Genet.* 2013;92:155–61.
49. Schwechheimer C, Zourelidou M, Bevan MW. PLANT TRANSCRIPTION FACTOR STUDIES. *Annu Rev Plant Physiol Plant Mol Biol.* 1998;49:127–50.
50. Dong T, Park Y, Hwang I. Absciscic acid: biosynthesis, inactivation, homeostasis and signalling. *Essays Biochem.* 2015;58:29–48.
51. Xu Z, Sun L, Zhou Y, Yang W, Cheng T, Wang J, et al. Identification and expression analysis of the SQUAMOSA promoter-binding protein (SBP)-box gene family in *Prunus mume*. *Mol Genet Genomics.* 2015;290:1701–15.
52. Li H-L, Guo D, Yang Z-P, Tang X, Peng S-Q. Genome-wide identification and characterization of WRKY gene family in *Hevea brasiliensis*. *Genomics.* 2014;104:14–23.
53. Liu X, Yang Y, Xu H, Yu D, Bi Q, Wang L. Transcriptome Analysis of Apricot Kernel Pistils Reveals the Mechanisms Underlying ROS-Mediated Freezing Resistance. *Forests.* 2022;13:1655.
54. Yang G, Zhang W, Liu Z, Yi-Maer A-Y, Zhai M, Xu Z. Both JrWRKY2 and JrWRKY7 of *Juglans regia* mediate responses to abiotic stresses and abscisic acid through formation of homodimers and interaction. *Plant Biol J.* 2017;19:268–78.
55. Devaiah BN, Karthikeyan AS, Raghothama KG. WRKY75 Transcription Factor Is a Modulator of Phosphate Acquisition and Root Development in Arabidopsis. *Plant Physiol.* 2007;143:1789–801.
56. Di D, Sun L, Wang M, Wu J, Kronzucker HJ, Fang S, et al. WRKY46 promotes ammonium tolerance in Arabidopsis by repressing NUDX9 and indole-3-acetic acid-conjugating genes and by inhibiting ammonium efflux in the root elongation zone. *New Phytol.* 2021;232:190–207.

57. Tang J, Tian X, Mei E, He M, Gao J, Yu J, et al. WRKY53 negatively regulates rice cold tolerance at the booting stage by fine-tuning anther gibberellin levels. *Plant Cell*. 2022;34:4495–515.
58. Liu W, Liang X, Cai W, Wang H, Liu X, Cheng L, et al. Isolation and Functional Analysis of VvWRKY28, a Vitis vinifera WRKY Transcription Factor Gene, with Functions in Tolerance to Cold and Salt Stress in Transgenic *Arabidopsis thaliana*. *IJMS*. 2022;23:13418.
59. Fei J, Wang Y-S, Cheng H, Su Y-B, Zhong Y-J, Zheng L. The Kandelia obovata transcription factor KoWRKY40 enhances cold tolerance in transgenic Arabidopsis. *BMC Plant Biol*. 2022;22:274.
60. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 2019;47:D427–32.
61. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ et al. WoLF PSORT: protein localization predictor. *Nucleic Acids Res*. 2007;35 Web Server:W585–7.
62. Wilkins MR, Gasteiger E, Bairoch A, Sanchez J-C, Williams KL, Appel RD et al. Protein Identification and Analysis Tools in the ExPASy Server. 2-D Proteome Analysis Protocols. New Jersey: Humana Press; 1998. 531–52.
63. Wu G-Q, Li Z-Q, Cao H, Wang J-L. Genome-wide identification and expression analysis of the *WRKY* genes in sugar beet (*Beta vulgaris* L.) under alkaline stress. *PeerJ*. 2019;7:e7817.
64. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019;47:W256–9.
65. 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;37:202–8. Web Server:W.
66. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res*. 2012;40:e49–9.
67. Lescot M. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. 2002;30:325–7.
68. Wu Y, Zhang C, Yang H, Lyu L, Li W, Wu W. Selection and Validation of Candidate Reference Genes for Gene Expression Analysis by RT-qPCR in Rubus. *IJMS*. 2021;22:10533.
69. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;25:402–8.

Figures

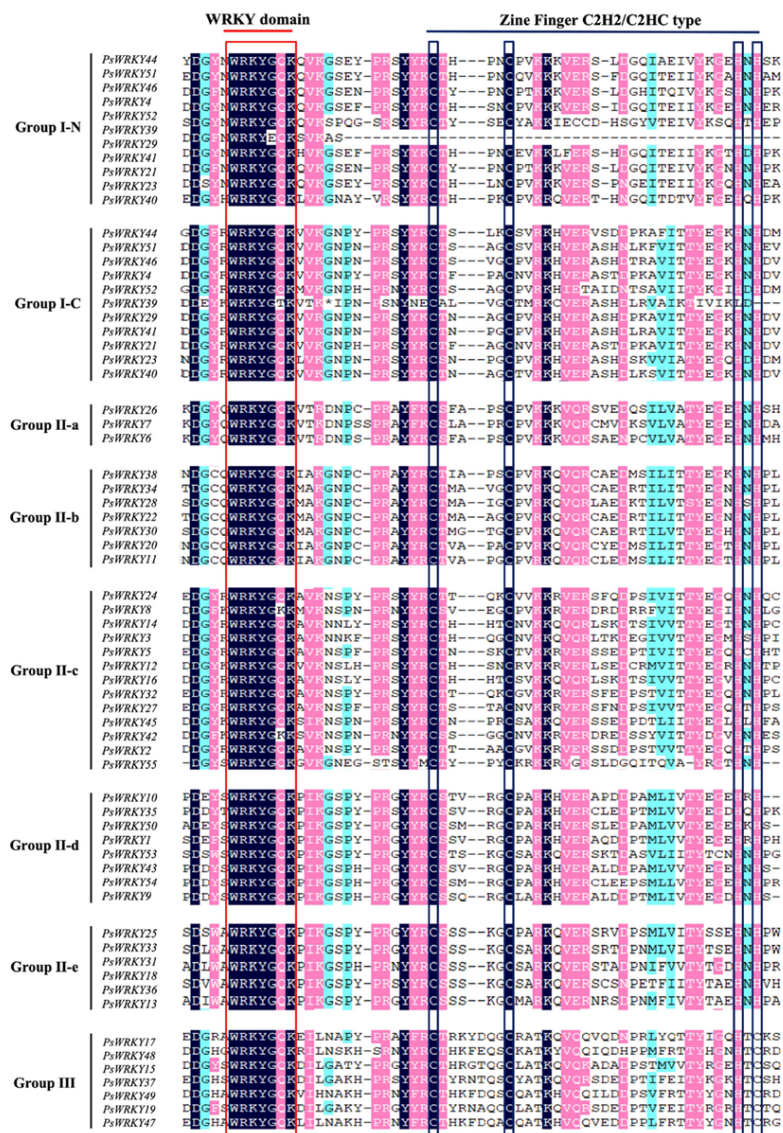


Figure 1

Multiple sequence alignments of conserved domains in PsWRKY transcription factors. Red box represent the conserved WRKY domains, and black box represent the zinc finger motif.

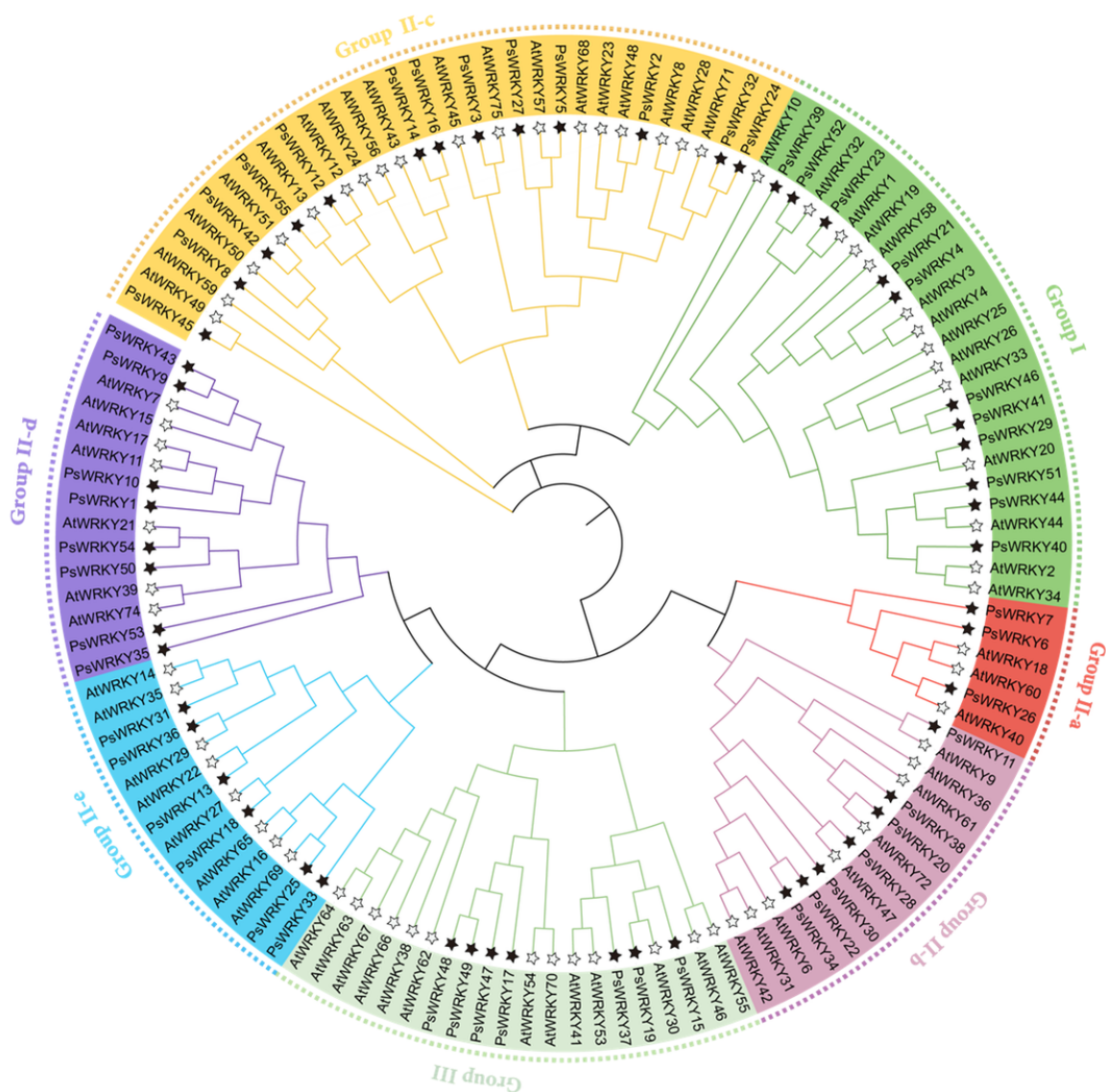


Figure 2

Phylogenetic analysis of *P. sibirica* and *A. thaliana* WRKY domain. Different coloured branches represent different subfamilies. *P. sibirica* and *A. thaliana* are marked as black and white pentagrams respectively.

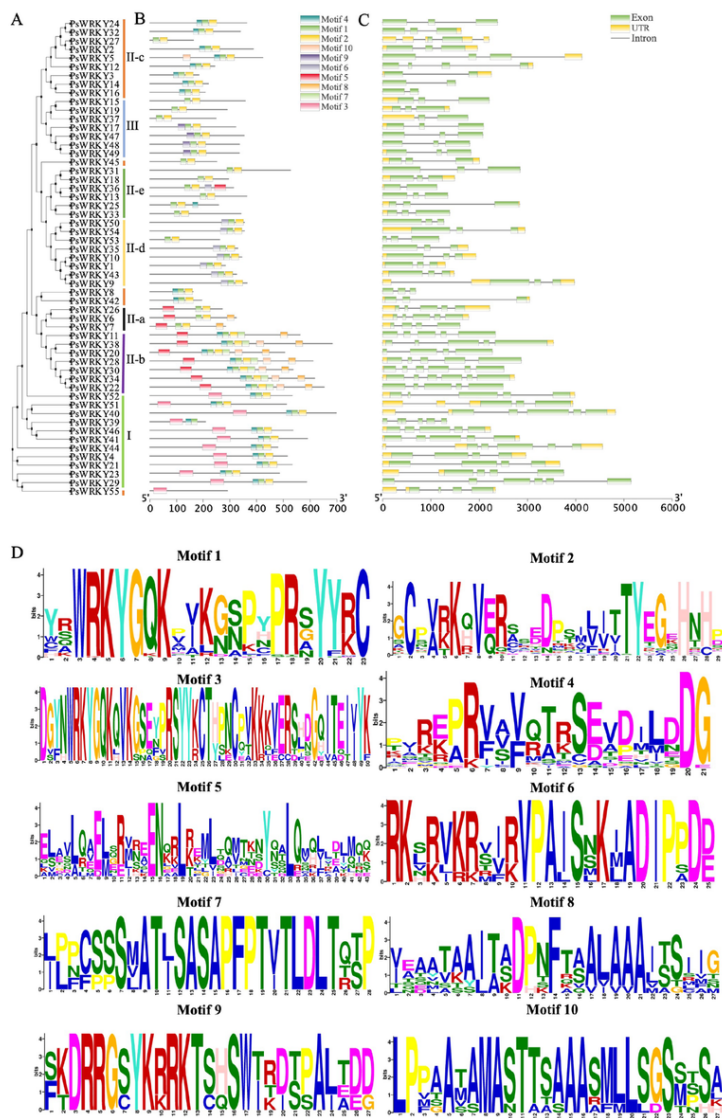


Figure 3

Phylogenetic relationships, conserved motifs, and gene structural analyses of the *PsWRKYs*.

(A) Neighbour-joining phylogenetic relationships of *PsWRKYs*. (B) Locations of conserved motifs in *PsWRKYs*. The ten predicted motifs are indicated by boxes of various colours. (C) Gene structure of *PsWRKYs*. Black lines and green and yellow boxes indicate introns, exons, and UTR, respectively. (D) Detailed sequence information for motifs 1–10.

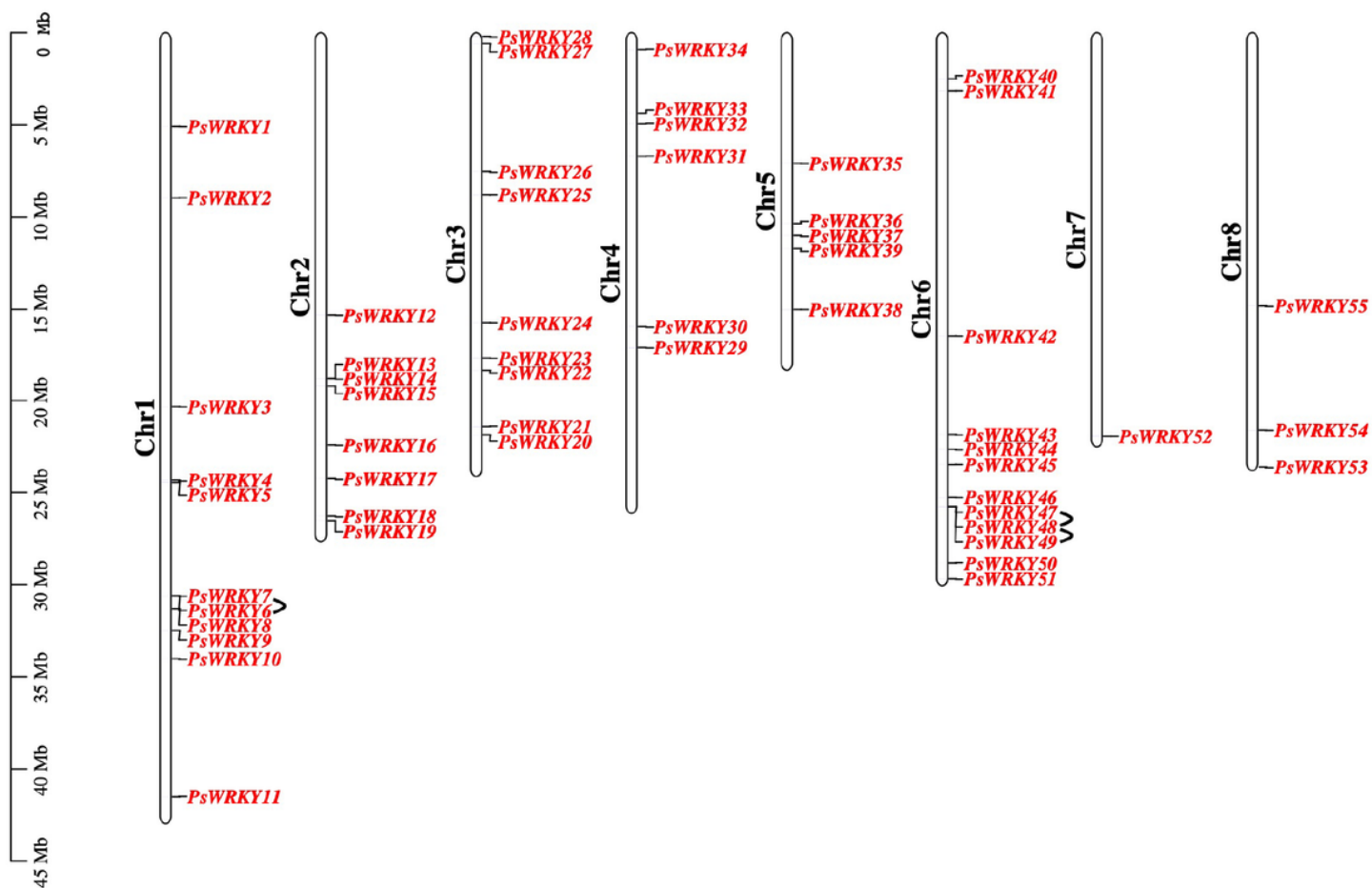


Figure 4

Chromosome distributions of *PsWRKY*s. The black arcs connected were tandem duplication genes.

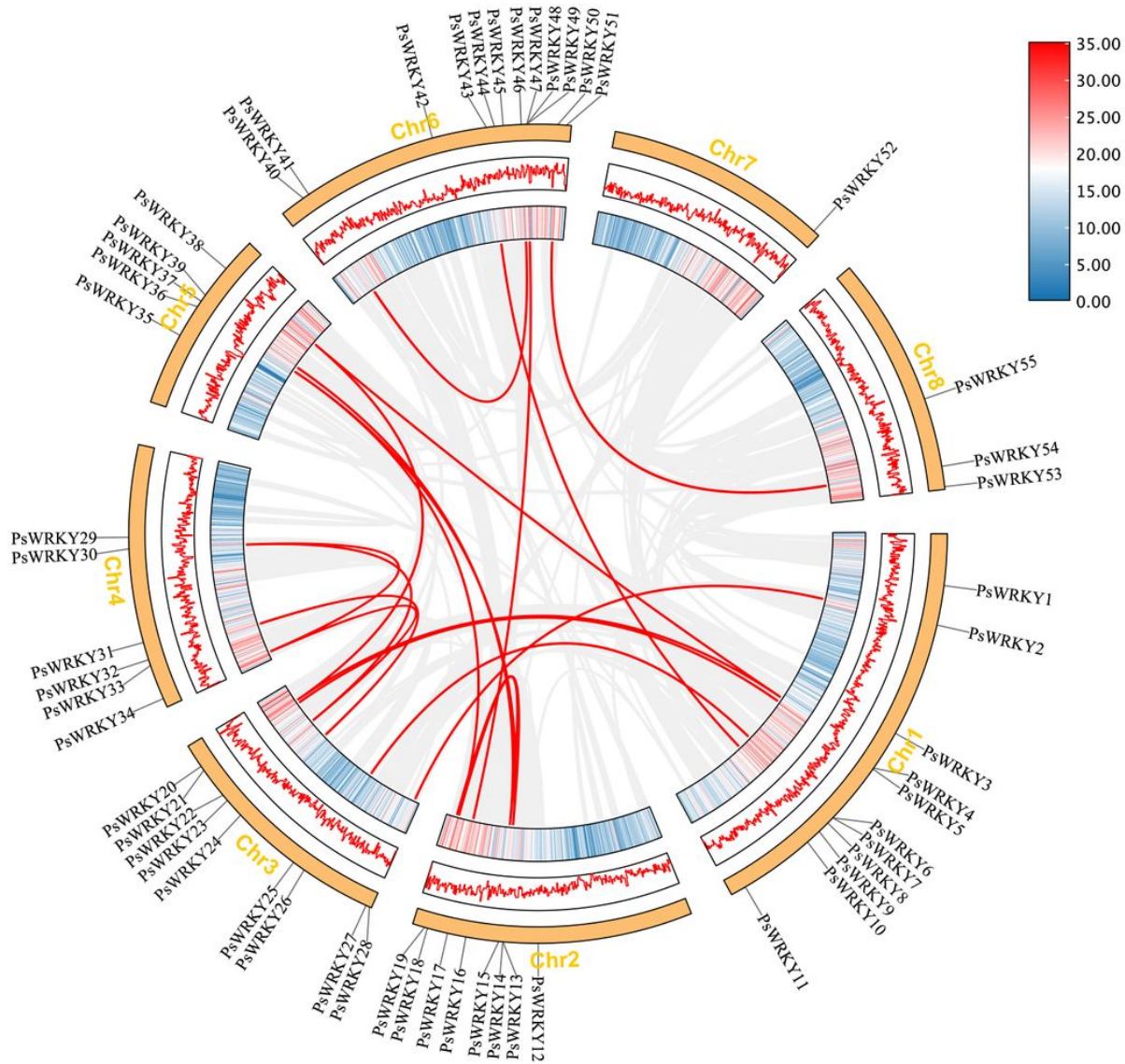


Figure 5

Syntenic analysis of *WRKYs* in *P. sibirica* genome.

Gray lines represent all syntenic blocks in the *P. sibirica* genome, while red lines indicate segmental duplication genes between *PsWRKYs*. Chromosome gene density is represented by a heatmap (inner circle) and column map (middle circle), with the outer circle showing the length of chromosomes.

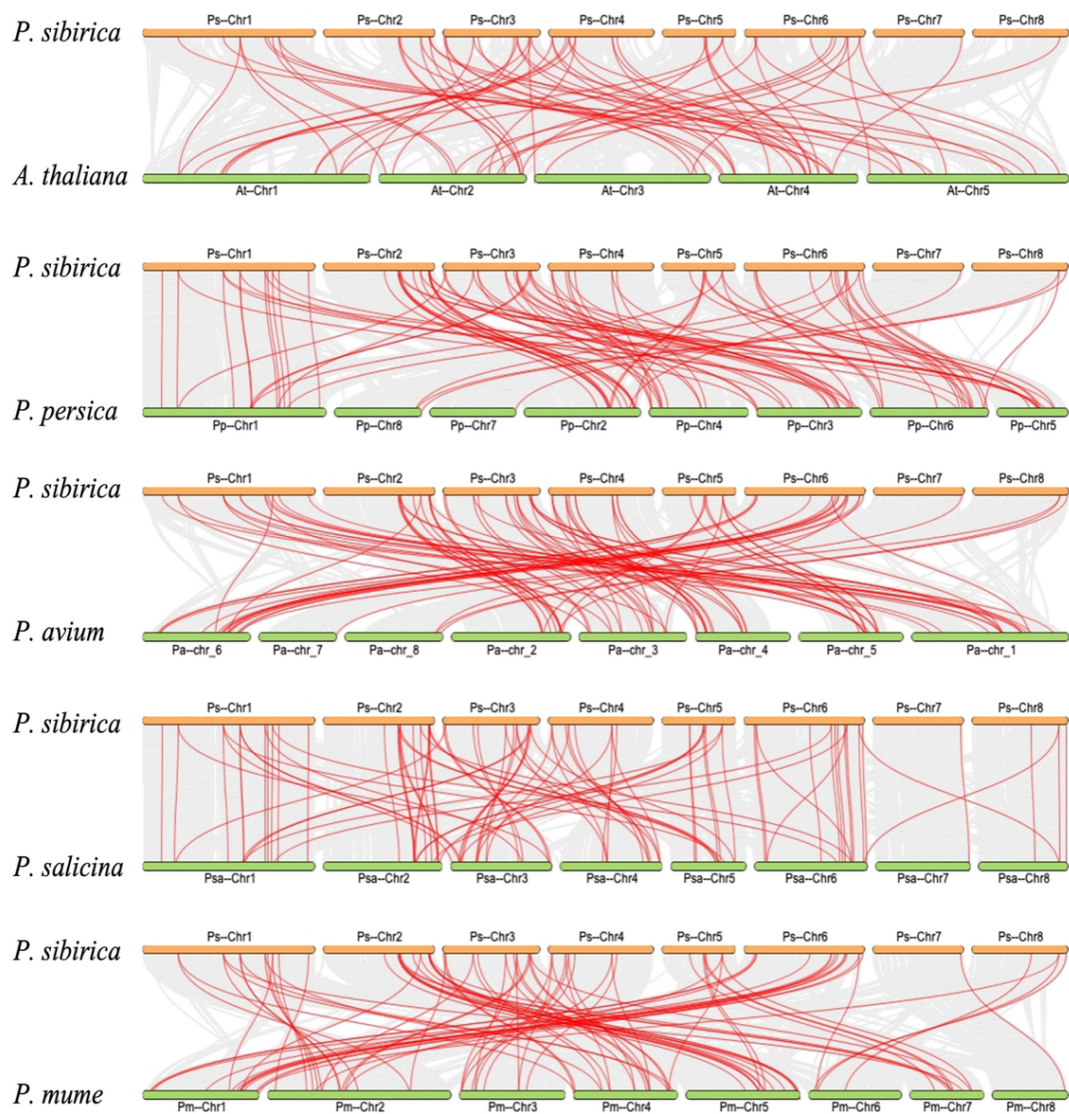


Figure 6

Synteny analysis of *WRKYs* in *P. sibirica* with, *A. thaliana*, *P. persica*, *P. avium*, *P. salicina*, and *P. mume*. Gray lines represent the collinear blocks within *P. sibirica* and other plant genomes, while red lines emphasize the pairs of syntenic *PsWRKYs*.

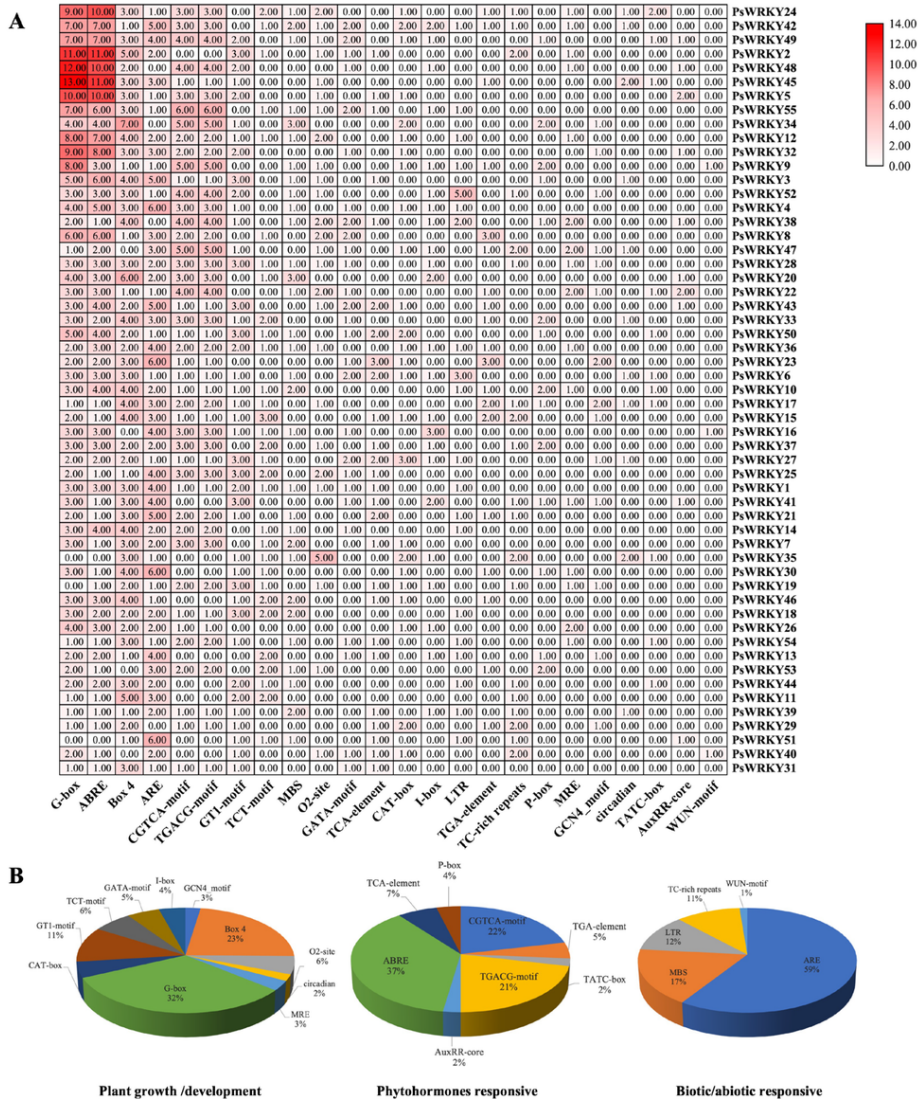


Figure 7

Analysis of the cis-acting elements of the *PsWRKYs*.

(A) Different colours represent the number of cis-acting elements. (B) The Pie chart size represents the percentage of promoter elements in each category, such as phytohormone-responsive, biotic/abiotic stress, and plant growth and development.

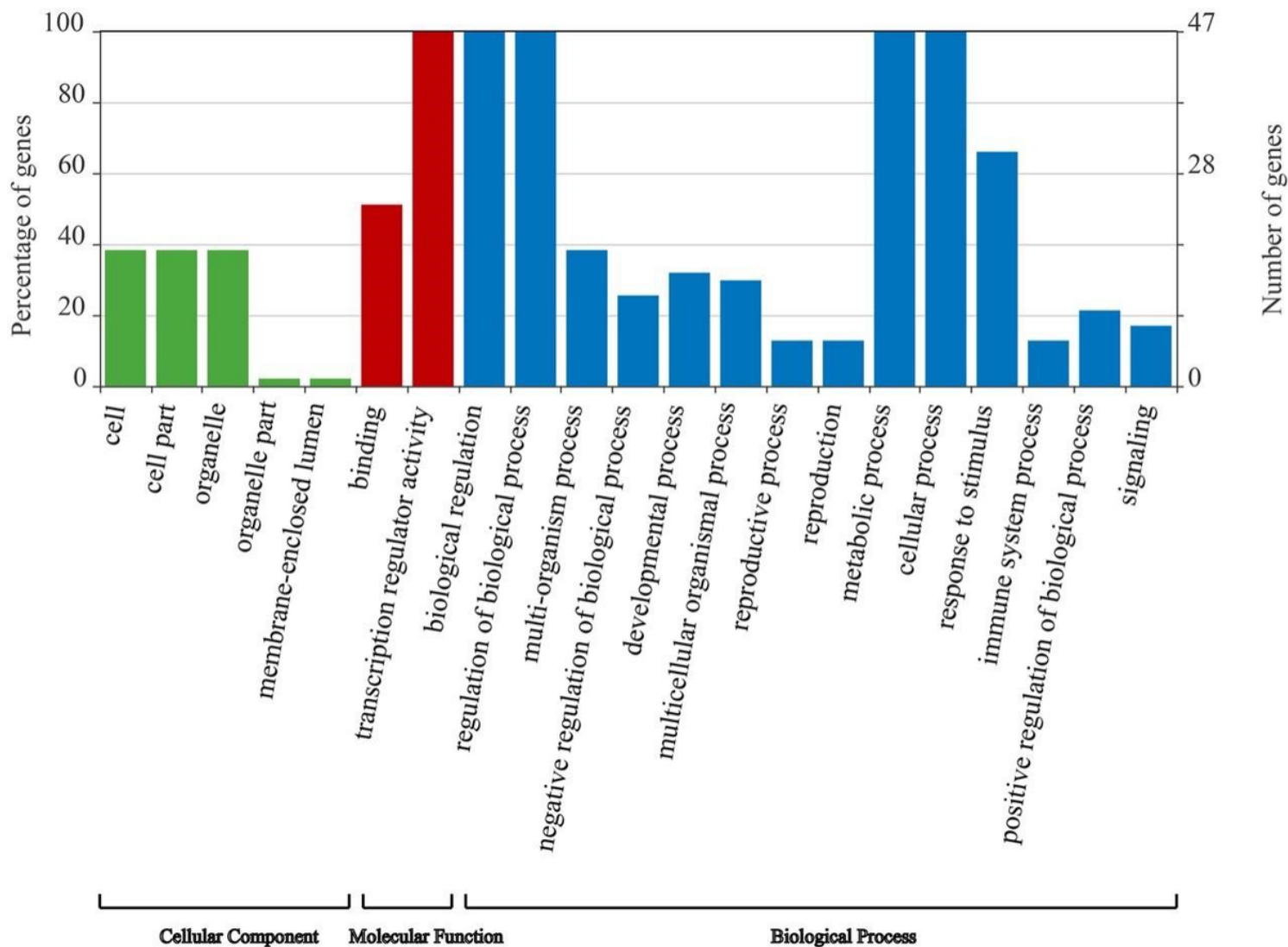


Figure 8

Gene ontology (GO) analysis of *PsWRKYs*.

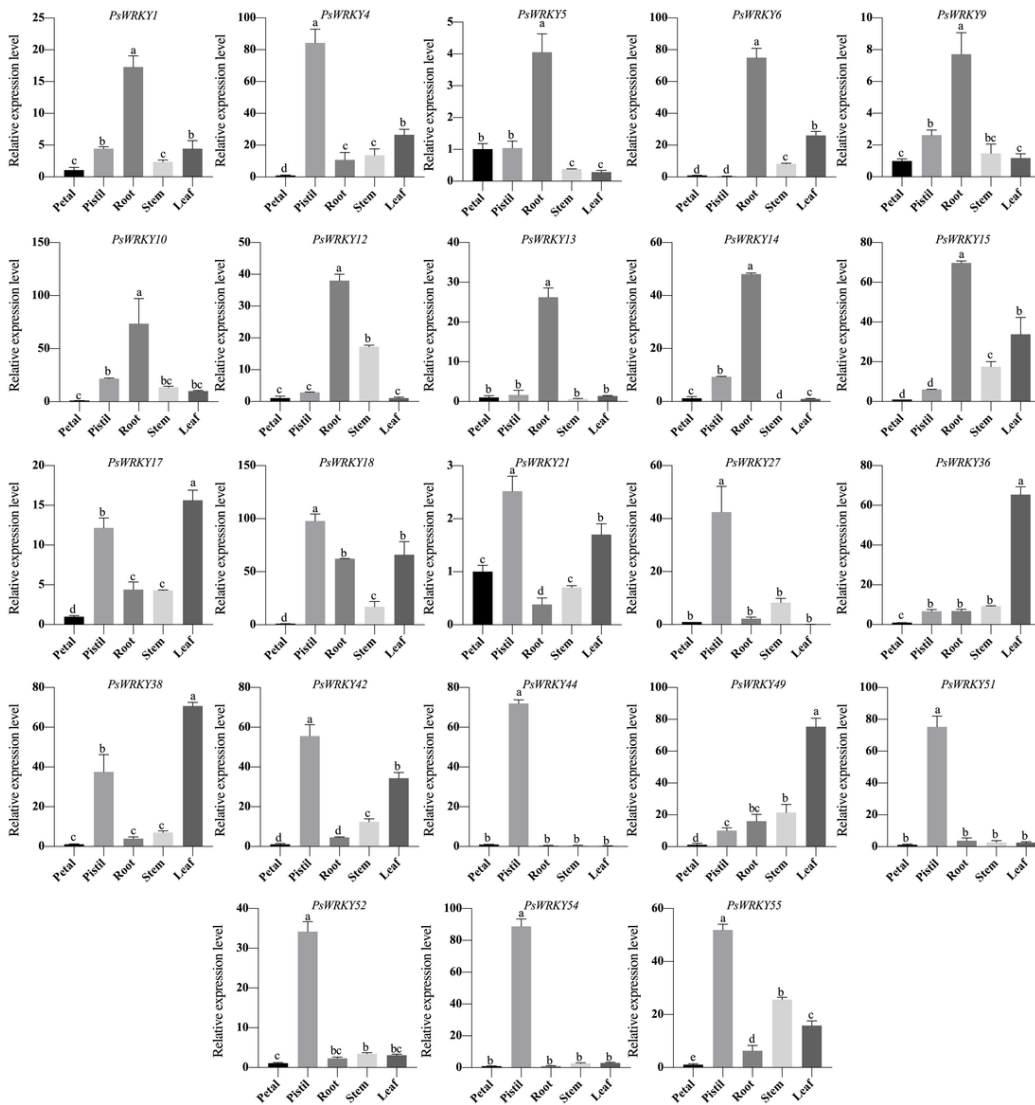


Figure 9

Expression analysis of the 23 *PsWRKYs* using quantitative real-time PCR (qRT-PCR) in different tissues of *P. sibirica*.

Total RNA was extracted from the petal, pistil, root, stem, and leaf. Petal was used as an internal control and normalized to 18S rRNA levels. The standard error was represented by error bars and the significant differences were identified using ANOVA followed by Duncan's multiple range test ($p < 0.05$, $n = 3$), where different letters indicate significant differences.

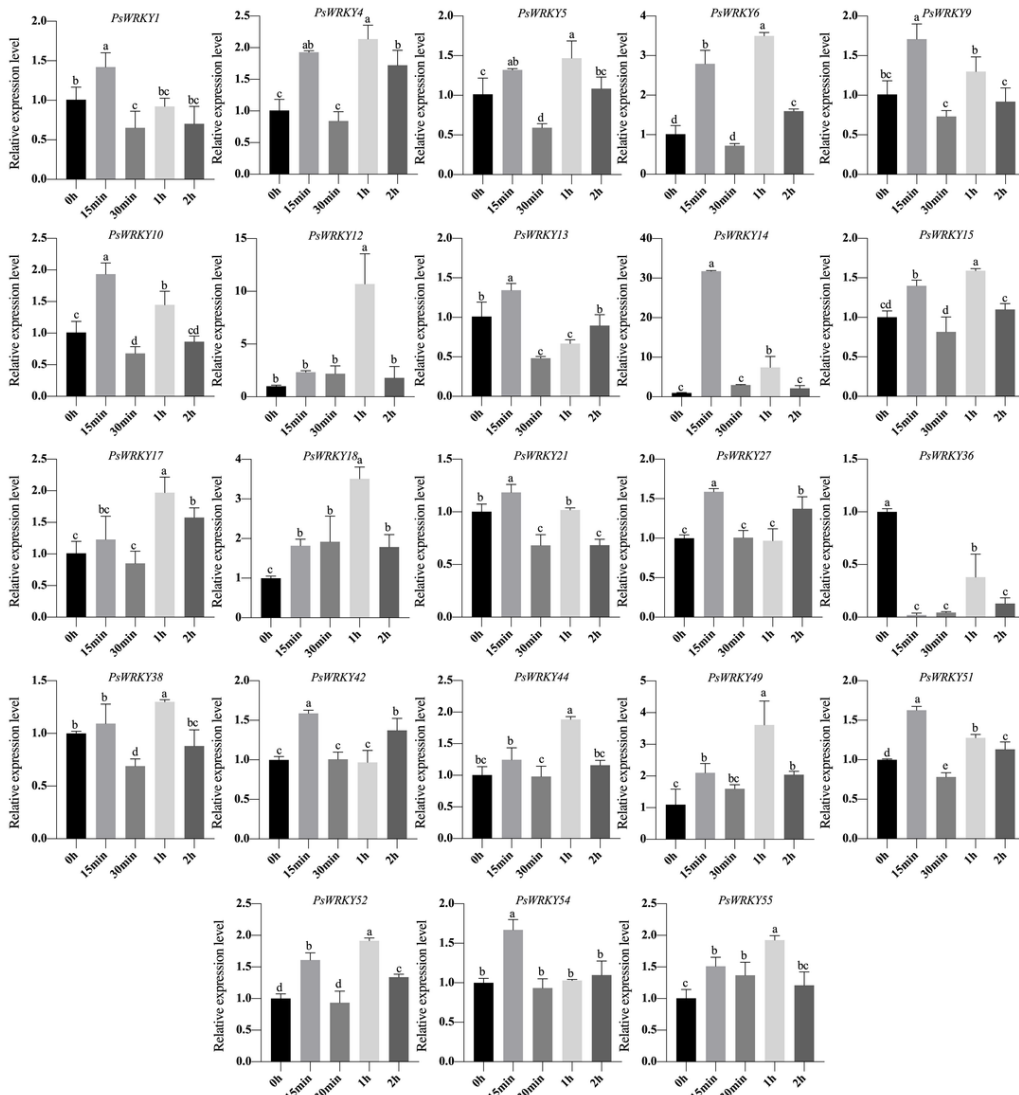


Figure 10

Expression analysis of the 23 *PsWRKYs* in *P. sibirica* pistils under low-temperature stress.

The relative expression levels of 23 *PsWRKYs* were examined by qRT-PCR at -4 °C treatment for 0 h, 5 min, 30 min, 1 h, and 2 h. The 0 h was used as a control and 18S rRNA as an internal reference gene. The standard error was represented by error bars and the significant differences were identified using ANOVA followed by Duncan's multiple range test ($p < 0.05$, $n = 3$), where different letters indicate significant differences.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)