forests

MDPI

Article

Mining and Expression Pattern Analysis of Genes Related to the Regulation of Flowering in Korean Pine (*Pinus koraiensis*)

Junshuai Du [1,2], Dan Hou [3], Junfei Hao [1], Junping Du [4], Hanguo Zhang [1] and Lei Zhang [1,*]

[1] State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China; 2024202040150@whu.edu.cn (J.D.); hjf@nefu.edu.cn (J.H.); 13175259125@163.com (H.Z.)
[2] National Field Station of Freshwater Ecosystem of Liangzi Lake, College of Life Sciences, Wuhan University, Wuhan 430072, China
[3] School of Big Data, Baoshan University, Baoshan 678000, China; 17667500302@163.com
[4] Haiyang Fruit Industry Development Service Center, Yantai 265100, China; djp199601@163.com
* Correspondence: zhanglei@nefu.edu.cn

Abstract: Korean pine (*Pinus koraiensis* Siebold & Zucc.) is an important timber and economic tree species in northeast China. Its seeds serve as both a primary means of propagation and a significant economic product. In this study, we identified 12 full-length MADS-box genes based on the Korean pine flower-induced transcriptome data available in our laboratory. These genes were identified through multiple sequence alignment and screening for conserved structural domains. We analyzed the genetic relationships of these genes and predicted their physicochemical properties. Additionally, we examined the expression patterns of three *SHORT VEGETATIVE PHASE* (*SVP*) genes across different tissues and developmental stages of Korean pine. The results indicate that the amino acid composition, molecular weight, isoelectric point, and other physicochemical properties of the MADS-box gene family in Pinus koraiensis are generally similar, though some individual variations are observed. A total of 12 MADS-box family genes were identified from the Korean pine transcriptome, distributed across five subfamilies. Conserved motif analysis revealed that these genes share similar conserved sequences. Structural and physicochemical analyses showed that genes with similar sequences exhibited comparable characteristics. Expression levels of the *SVP* genes varied significantly across different developmental stages and tissues, with the expression of the three *SVP* genes in leaves being markedly higher than in buds (approximately 200-fold). The expression levels of these genes in leaves were not only higher than in buds but also exceeded those in other tissues. Based on these findings, we conclude that these three SVP genes primarily play a suppressive role in the process of flower bud formation, helping Korean pine maintain a juvenile state under certain conditions, and are also involved in the growth and development of its leaves. This research provides a basis for future studies on the flowering induction mechanism in Korean pine.

Keywords: *Pinus koraiensis*; MADS-box gene; flowering mechanism; expression pattern analysis; SVP-like gene

Academic Editor: Tadeusz Malewski

Received: 14 December 2024
Revised: 3 January 2025
Accepted: 14 January 2025
Published: 17 January 2025

Citation: Du, J.; Hou, D.; Hao, J.; Du, J.; Zhang, H.; Zhang, L. Mining and Expression Pattern Analysis of Genes Related to the Regulation of Flowering in Korean Pine (*Pinus koraiensis*). *Forests* **2025**, *16*, 168. https://doi.org/10.3390/f16010168

1. Introduction

Korean pine is widely distributed in various regions worldwide (Figure 1). It is an excellent tree species that offers a balance of economic, ecological, and social benefits. In northeastern China, Korean pine is a valuable timber species and one of the most promising native trees for fruit-bearing economic forests in the region [1]. However, Korean pine

grows slowly, has a long juvenile period, and regenerates slowly under natural conditions. The fruiting cycle of Korean pine typically spans five years [2]. Due to the generally long fruiting cycle, increasing yield has become one of the urgent challenges to be addressed in most seed orchards and nut forests. Flower bud differentiation is a critical stage in the growth and development process, as it determines the number of flowers and the fruit set rate each year, directly impacting the yield. However, there are relatively few studies on the formation mechanism of flower buds in Korean pine. Analyzing the flowering mechanism is therefore crucial for improving yield.

Figure 1. Distribution map of Pinus koraiensis, where different colors of the points represent different distribution densities [3].

The term "MADS-box" is derived from the first letters of four proteins: *Saccharomyces cerevisiae MiniChromosome Maintenance 1* (*MCM1*), *Arabidopsis thaliana* (*Arabidopsis thaliana* (L.) Heynh.) floral homeotic gene *AGAMOUS* (*AG*), *Antirrhinum majus* (*Antirrhinum majus* L.) floral homeotic gene *Deficiens* (DEF), and human serum response factor [4]. Based on phylogenetic relationships and gene structure, plant MADS-box genes can be classified into two major groups: Type I and Type II. Compared to Type II, Type I MADS-box genes lack the K domain and the C segment. Due to the unique structure of Type II MADS-box genes, they are also referred to as MIKC-type MADS-box genes [5]. Meanwhile, MIKC-type genes can be further classified into two categories, MIKCC and MIKC*, based on the length of the K domain. Currently, researchers have identified the presence of the MADS-box gene family in numerous plants and have focused on its role in promoting floral organ growth [6]. It has been shown that the MADS-box gene family plays a role in regulating inflorescence differentiation [7], spike growth [8], and promoting the flowering process [9]. However, relatively few studies have been conducted on the mechanism of MADS-box-mediated flower formation in conifers.

The SVP subfamily is a crucial component of the MADS-box gene family and plays a significant role in regulating the flowering process in plants. In dicotyledonous plants, *SVP* genes can be classified into three major groups (*SVP1*, *SVP2*, and *SVP3*) [10]. Further studies in the Rosaceae family have subdivided the *SVP2* group into three subgroups: *SVP2-R1*, *SVP2-R2*, and *SVP2-R3* [11]. It has been reported that *SVP* genes function differently during flowering: they act as repressors in the early stages and later contribute to the development of specific floral organs [12]. In *Arabidopsis thaliana*, other subfamily genes of the MADS-box gene family, such as *APETALA1* (*AP1*) and *AG*, function to repress the expression of *SVP* genes, thereby ensuring the normal development of floral organs. It has also been reported that SVP-like genes in conifer species cause abnormalities in pollen, which in turn affect the germination process [13]. Another study analyzing RNA-seq data

from grape, poplar, and apple found that genes in the *SVP3* group are highly expressed in vegetative organs (such as buds, leaves, cotyledons, and dormant buds), suggesting that these genes may be involved in the dormancy process [10]. However, it has also been reported that in *Rosa multiflora* (*Rosa multiflora* Thunb.), the expression pattern of *SVP* genes differs from those observed in other Rosaceae species (such as apple (*Malus pumila* Mill.), pear (*Pyrus* L.), and peach (*Prunus persica* (L.) Batsch)). The expression pattern of *SVP* genes in *Rosa multiflora* is more similar to the expression patterns of the *FLC* gene in *Arabidopsis* and the *KSN* gene in rose (*Rosa rugosa* Thunb.), indicating that *SVP* genes may play a role in the seasonal flowering process of *Rosa multiflora*, rather than being solely involved in bud dormancy [11].

Transcriptome data on the flower formation process of Korean pine were obtained through triple sequencing technology during the pre-experiment phase, which led to the preliminary annotation of 68 genes as MADS-box genes. A total of 12 full-length MADS-box genes were then identified through multiple sequence alignment and screening for conserved structural domains. In this study, we performed bioinformatics analysis, phylogenetic analysis, and conserved motif analysis on these 12 MADS-box genes. We selected three SVP-like genes from them and analyzed their expression patterns in different tissues using real-time fluorescence quantitative PCR. Our study aims to investigate the structure and physicochemical properties of the MADS-box gene family in Korean pine, with a particular focus on its expression in different organs and tissues in order to elucidate how it regulates the flowering of Korean pine. Our study will provide a basis for further research into the biological functions of the MADS-box gene family in conifer growth and development.

2. Materials and Methods

2.1. Plant Materials

All plant materials used in this study were obtained from Qingshan Forest Farm in Linkou County (longitude 128°10′–131°08′, latitude 44°45′–45°58′, altitude 320 m). The seedling materials included 2-year-old, 4-year-old, and 6-year-old live seedlings from the nursery of Castle Peak Forest, while the 20-year-old and 40-year-old materials were collected from the primary seed orchard of Korean pine. These materials were quick-frozen in liquid nitrogen and transported back to the laboratory using dry ice, where they were stored at −80 °C. Various tissue parts were isolated in the laboratory, including roots, stems, and leaves from the 2-, 4-, and 6-year-old seedlings, as well as leaves and buds from the upper, middle, and lower branches of the 20- and 40-year-old plants. RNA was then extracted from these materials by grinding them in liquid nitrogen and using a reverse transcription kit.

2.2. Bioinformatics Analysis

Homologous sequences from other species were obtained and phylogenetically analyzed using the NCBI database (https://www.ncbi.nlm.nih.gov/, accessed on 22 July 2024) and UniProt [14] following protein comparison. The phylogenetic tree was constructed using MEGA 7.0 software [15] with the neighbor-joining (NJ) method and a bootstrap value of 1000. The tree was then beautified using the online tool iTOL Version 6.0 (https://itol.embl.de/, accessed on 5 July 2024) [16]. Conserved protein motifs were predicted using the MEME online tool Version 5.5.6 (http://meme-suite.org/tools/meme, accessed on 5 August 2024) [17], with the number of motifs set to 10 and the remaining parameters set to default. The results of the conserved motif prediction were correlated with the clustering results and plotted using TBtools Version 2.130 [18]. Protein molecular weight, instability index, and other physicochemical properties were calcu-

lated using the ExPASy online tool (https://web.expasy.org/protparam/, accessed on 5 September 2023) [19]. Secondary and tertiary structures of the proteins were analyzed using the SOPMA online tool (https://npsa.lyon.inserm.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html, accessed on 18 April 2024) [20] and SWISS-MODEL Version 2024_06 (https://swissmodel.expasy.org/, accessed on 2 July 2018) [21], respectively. Signaling peptides of proteins were predicted using the iPSORT online tool (https://ipsort.hgc.jp/predict.cgi, accessed on 6 March 2024) [22]. The transmembrane structure of proteins was predicted using the TMHMM Version 2.0 online tool (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0, accessed on 24 July 2023) [23]. Subcellular localization of the proteins was predicted using the WoLF PSORT online tool (https://wolfpsort.hgc.jp/, accessed on 12 January 2005) [24].

2.3. Analysis of Gene Expression Patterns in Different Tissues

Quantitative real-time PCR was used to analyze the expression patterns of the Korean pine MADS-box genes in different tissues. In this study, total RNA was extracted from the tissues of three plants in two stages using the Universal Plant Total RNA Extraction Kit (BIOTEKE, Beijing, China) according to the manufacturer's instructions. The RNA samples were treated with DNase I (TaKaRa Bio, Shiga, Japan) to remove any contaminating genomic DNA. After extraction, the RNA quality was assessed by Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA), Qubit 2.0 (Thermo Fisher Scientific, Waltham, MA, USA), and Agilent 2100 systems (Agilent Technologies, Santa Clara, CA, USA). The Nanodrop was used to measure RNA purity, ensuring that the A260/A280 ratio was within the acceptable range, while the Qubit 2.0 quantified RNA concentration. The Agilent 2100 Bioanalyzer was used to assess RNA integrity, providing an RNA Integrity Number (RIN) to evaluate sample quality. All RNA samples had high purity, appropriate concentration, and satisfactory integrity for downstream applications. The extracted RNA was then reverse transcribed into cDNA using the ReverTra Ace® qPCR RT Master Mix (TOYOBO, Osaka, Japan) with gDNA Remover (TOYOBO, Osaka, Japan) kit. The RNA was denatured by incubating at 65 °C for 5 min and then rapidly cooled on ice. Next, 2 μL of 4× DN Master Mix, 0.5 μg of RNA template, and nuclease-free water were added to a final volume of 8 μL. The mixture was incubated at 37 °C for 5 min, then placed at 4 °C to degrade genomic DNA (gDNA). Then, 2 μL of 5× RT Master Mix II was added to initiate cDNA synthesis. The reaction was incubated at 37 °C for 15 min, followed by a 5 min incubation at 98 °C to terminate the reaction. Finally, the reaction mixture was stored at 4 °C. The cDNA library was constructed following standard protocols, and its concentration was verified by qPCR to ensure the effective concentration for sequencing. Illumina sequencing was performed on the HiSeq2500 platform with paired-end sequencing (PE125), providing 125 base-pair reads. Sample inspection, library construction, quality control, and sequencing were all completed by Biomarker Technology Co., Ltd. (Beijing, China), ensuring high-quality data for subsequent analysis. Gene-specific quantitative primers were designed using Primer Premier v5.0 software (Table 1) [25]. 18S-RNA was used as the internal reference gene [26]. Quantitative real-time PCR was performed using the qTOWER3G Real-Time Fluorescence Quantitative Gene Amplifier (Analytik Jena, Jena, Germany) and TRANSGEN's TransStart™ TipTop Green qPCR Super Mix kit (TRANSGEN, Beijing, China) as the quantification reagents, following the procedure outlined below. Initially, the temperature was raised to 94 °C and held for 30 s. The cycling process then began, with each cycle consisting of the following steps: denaturation at 94 °C for 5 s, annealing at 60 °C for 15 s, and extension at 72 °C for 10 s. This cycling process was repeated 40 to 45 times. After the cycles were completed, a melting curve analysis was performed by heating to 95 °C for 15 s, followed by a slow decrease to 60 °C, where it was held for 1 min. Finally, the reaction

was concluded by heating to 95 °C for 30 s. Dissolution curves were analyzed at the end of the reaction. The quantitative real-time PCR results were analyzed using WPS Excel 2024, and gene expression was calculated using the formula $2^{-\Delta\Delta CT}$.

Table 1. Primers used in quantitative real-time PCR.

Gene Name	Forward and Reverse Primers (5′-3′)
PkMADS1-F	ATTGGAAAATCAGGATCCTCAG
PkMADS1-R	TGCGAAGATAACCCCAACTG
PkMADS2-F	AGAAAATGCAGTGAGCAGGAAC
PkMADS2-R	CGCAAAGTATTGACAACTCCTC
PkMADS3-F	AAAGCGACTTCGGTTGTGAG
PkMADS3-R	TCAAACTGATTCCTTCAAGCTC
PkMADS4-F	AAAGCGACTTCGGTTGTGAG
PkMADS4-R	CCTTCAAGCTCATCACCTCG
18S-RNA-F	GAGGTAGCTTCGGGCGCAACT
18S-RNA-R	GCAGGTTAGCGAAATGCGATAC

3. Results

3.1. Bioinformatics Analysis Results

3.1.1. Physical and Chemical Property Analysis Results

The results of a series of physicochemical analyses of the 12 MADS-box proteins of Korean pine (Table 2) revealed that the number of amino acids in these proteins ranged from 73 to 104, with PkMADS7 having the highest number of amino acids and PkMADS10 having the lowest. Their molecular weights ranged from 8432.94 Da to 45,959.94 Da. Regarding the isoelectric point, the values of these proteins ranged from 5.29 to 10.11, with only four proteins having an isoelectric point lower than 7, while the remaining eight proteins had an isoelectric point greater than 7, suggesting a greater abundance of basic amino acids. The aliphatic index of all these proteins was greater than 70, indicating that they have good thermal stability. The average hydrophobicity coefficients of the proteins, except for PkMADS6, were negative, suggesting that the majority of these proteins are hydrophilic. The instability coefficients of all these proteins, except for PkMADS10, exceeded 40, indicating that all proteins, except for PkMADS10, are unstable. Signal peptide analysis revealed that only the PkMADS9 gene contains a signal peptide, suggesting that all other proteins are non-secretory proteins. Protein transmembrane analysis revealed that only PkMADS6 contains two transmembrane helical structures, while the others are non-transmembrane proteins. Finally, subcellular localization analysis indicated that all proteins were primarily localized in the nucleus, except for PkMADS2, which was localized in the cytoplasm, and PkMADS4, which was localized in the chloroplasts.

Table 2. Physiological and biochemical analysis of MADS-box family proteins.

Protein	Number of Amino Acid (Da)	Molecular Weight	Theoretical pI	Instability Index	Aliphatic Index	Signal Peptide	Transmembrane	Grand Average of Hydropathicity	Subcellular Localization
PkMADS1	238	27,312.8	5.94	75.06	86.05	None	None	−0.757	Nuclear
PkMADS2	183	21,820.0	9.32	66.98	90.6	None	None	−0.858	Cytoplasm
PkMADS3	163	18,981.8	9.21	51.88	80.67	None	None	−0.677	Nuclear
PkMADS4	155	18,419.3	7.64	42.9	80.45	None	None	−0.337	Chloroplast
PkMADS5	135	15,516.0	10.11	47.44	70.67	None	None	−0.416	Nuclear
PkMADS6	105	11,838.1	9.72	72.38	102.19	None	Two	0.230	Nuclear
PkMADS7	404	45,959.9	5.29	51.29	71.71	None	None	−0.567	Nuclear
PkMADS8	390	44,419.3	5.77	57.37	78.79	None	None	−0.647	Nuclear
PkMADS9	254	29,547.1	9.08	51.92	82.13	One	None	−0.614	Nuclear
PkMADS10	73	8432.9	9.93	37.44	72.19	None	None	−0.230	Nuclear
PkMADS11	162	18,845.6	9.71	42.55	89.63	None	None	−0.668	Nuclear
PkMADS12	261	30,022.7	8.91	43.78	78.39	None	None	−0.852	Nuclear

3.1.2. Protein Secondary Structure Prediction

The prediction of the secondary structure features of the 12 MADS-box family proteins in Korean pine (Figure 2, Table 3) revealed that, despite large differences in the number of amino acids, these proteins share some structural similarities. They are predominantly composed of alpha helices and random coils, with extended strand regions distributed across different amino acid chains. Specifically, the PkMADS4 protein exhibited the highest proportion of alpha helices at 80%, while PkMADS10 had the lowest proportion at 28.77%. For random coils, the PkMADS7 protein had the largest proportion at 52.97%, while PkMADS4 had the smallest at 11.61%. Finally, in the extended strand region, the PkMADS11 protein had the highest percentage at 28.77%, while PkMADS4 had the lowest at 5.81%.

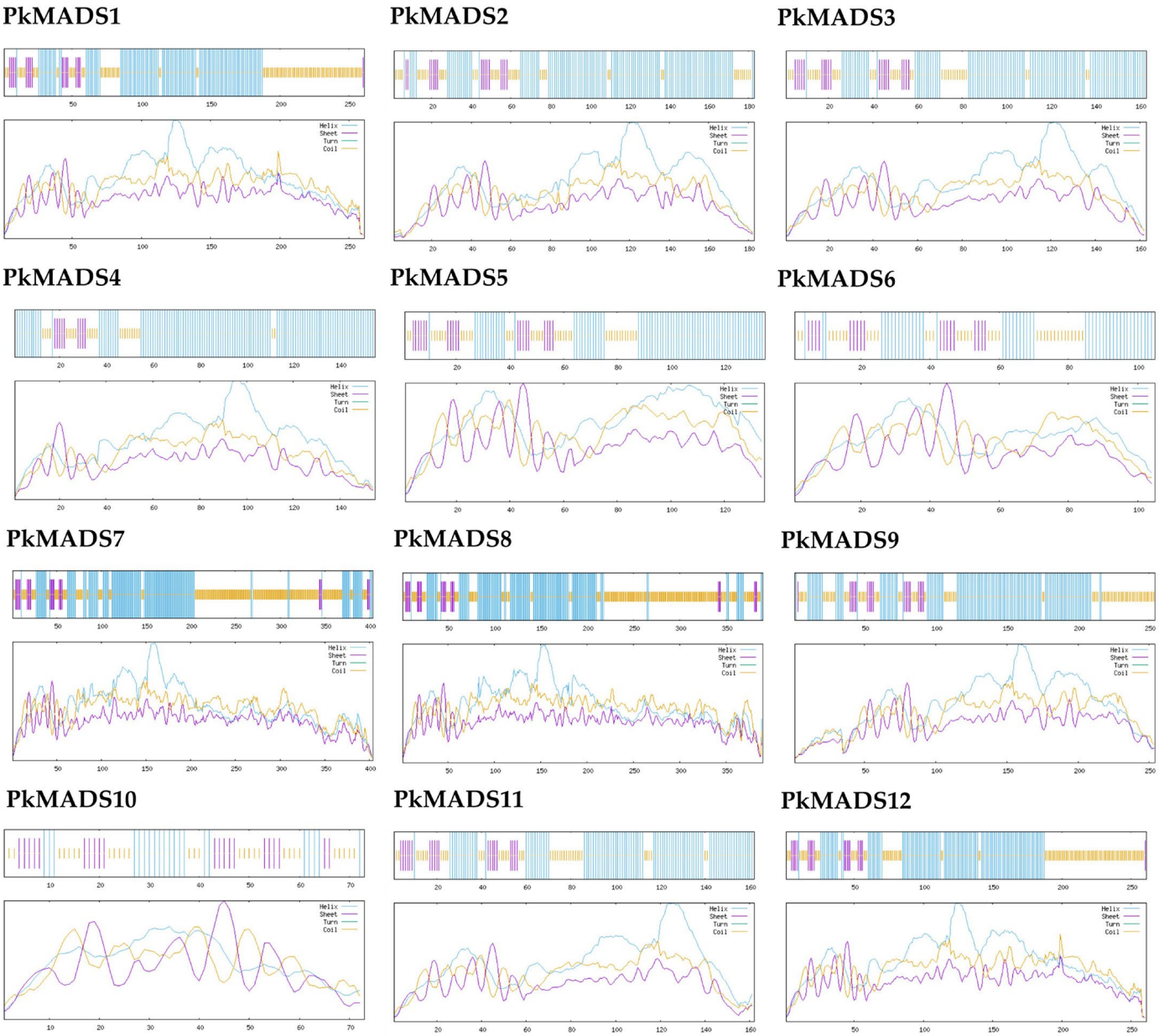

Figure 2. Predicted secondary structure of MADS-box proteins.

Table 3. Secondary structure prediction of Korean pine MADS-box protein.

Proteins	Alpha Helix	Extended Strand	Random Coil
PkMADS1	55.04	9.24	33.19
PkMADS2	66.95	9.84	18.03
PkMADS3	65.64	12.27	19.02
PkMADS4	80.00	5.81	11.61
PkMADS5	54.07	14.81	31.11
PkMADS6	44.76	17.14	38.10
PkMADS7	40.59	6.44	52.97
PkMADS8	41.03	6.67	52.31
PkMADS9	55.91	7.87	36.22
PkMADS10	28.77	28.77	42.47
PkMADS11	59.26	12.35	28.40
PkMADS12	48.28	7.66	44.06

3.1.3. Protein Tertiary Structure Prediction

In this study, the tertiary structures of the amino acid sequences of 12 MADS-box family proteins in Korean pine were predicted using the homology modeling method with online tools. The prediction results (Figure 3) show that the tertiary structures of these proteins mainly consist of random coils and alpha helices, which is consistent with the secondary structure prediction. Although there is some similarity between these amino acid sequences, their spatial structural features differ due to variations in the length and arrangement of the alpha helices and random coils.

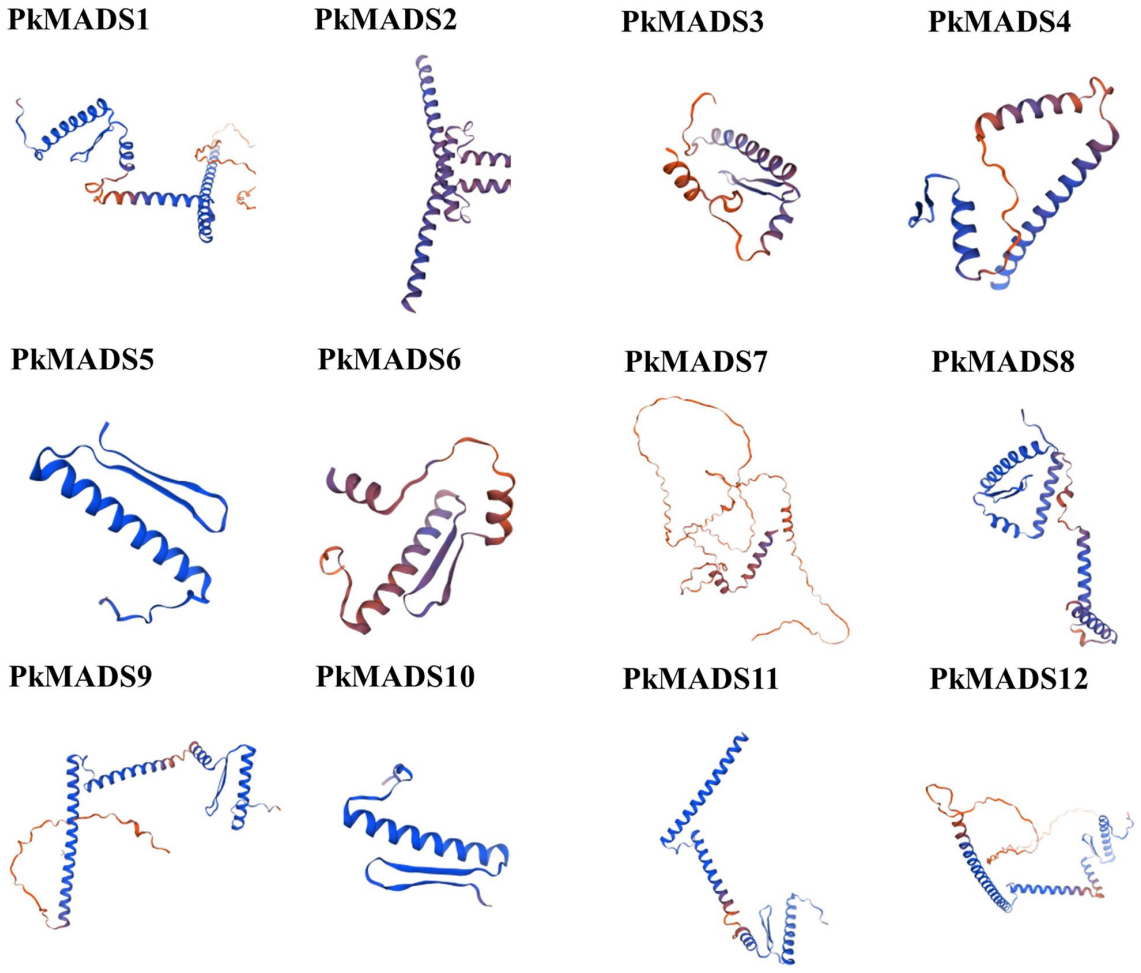

Figure 3. Tertiary structure prediction of Korean pine MADS-box proteins.

3.2. Phylogenetic and Conserved Structural Domain Analyses

Phylogenetic trees of the *Arabidopsis thaliana* MADS-box gene family and the Korean pine MADS-box gene family were constructed using the neighbor-joining method with MEGA 7.0 software (Figure 4). The twelve Korean pine MADS-box proteins were primarily classified into five branches: two proteins classified as GGM13, two as MIKC*, two as SOC1, two as AGL6, and four as SVP. Except for PkMADS2, all other proteins contain motif2 and motif7. However, it is worth noting that, despite the high consistency in the types and numbers of motifs in the protein sequences of PkMADS1, some individual MADS proteins still exhibit certain differences in the types and numbers of motifs. For example, PkMADS7 and PkMADS8 contain only four motifs. Among the ten identified motifs, motif1, motif2, and motif7 collectively constitute the core conserved domains of the SVP proteins (Figure 5).

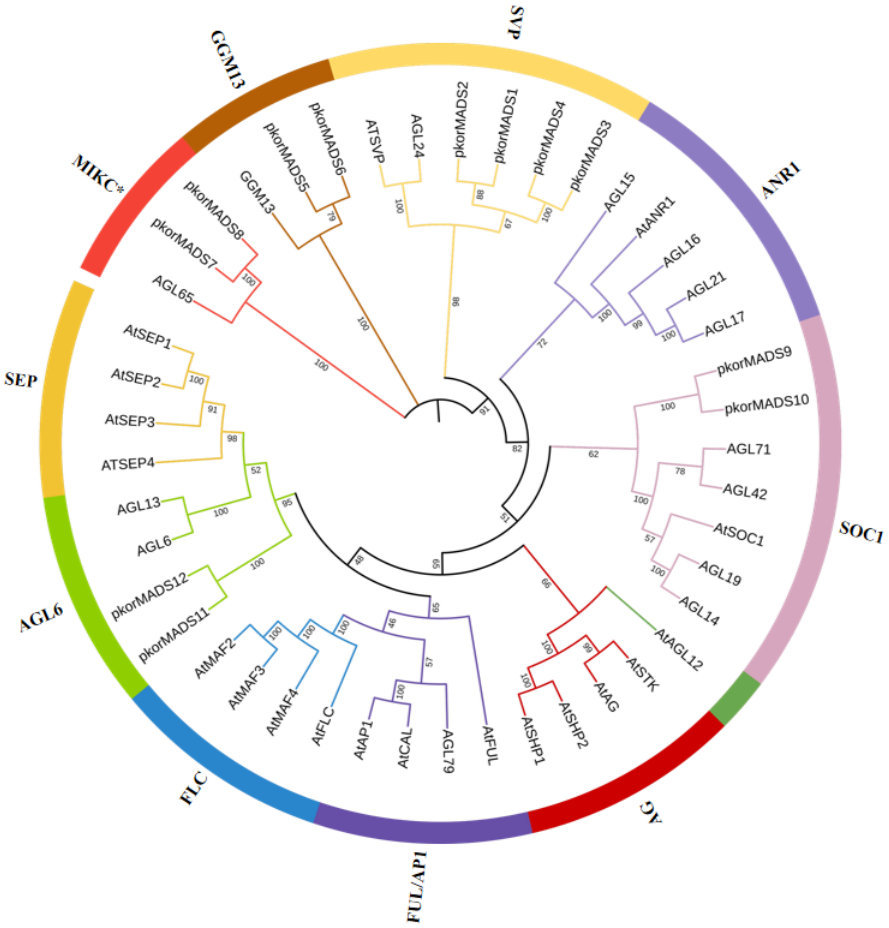

Figure 4. Phylogenetic tree of MADS-box proteins in Korean pine and *Arabidopsis thaliana*.

3.3. Analysis of the Expression Pattern of MADS Gene in Different Tissues and Organs of Korean Pine

The expression levels of Pinus koraiensis MADS-box genes *PkMADS1*, *PkMADS3*, and *PkMADS4* were analyzed in the needles, upper buds, middle buds, and lower buds of twenty-year-old and forty-year-old plants; the roots, stems, and leaves of two-year, four-year, and six-year-old seedlings; and the female flowers, male flowers, and fruits, to explore their tissue-specific expression patterns.

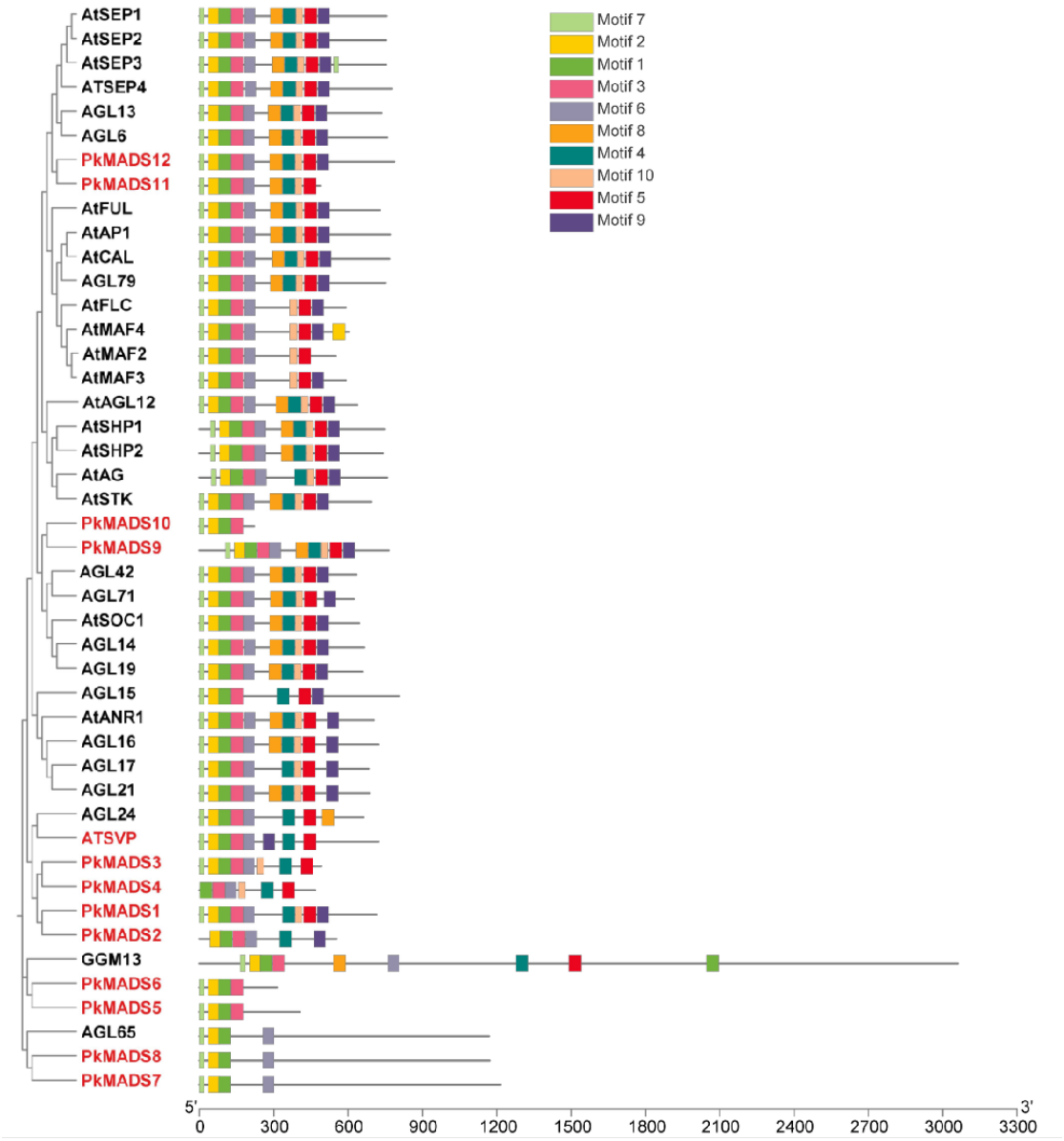

Figure 5. Phylogenetic tree and conserved motif analysis of MADS-box proteins.

3.3.1. Results of the Analysis of PkMADS1 Expression Pattern

The results of real-time quantitative PCR analysis (Figure 6) show that the expression of *PkMADS1* in leaves first increases and then decreases with age, peaking in the 20-year-old leaf samples. In roots, the expression of *PkMADS1* initially increases and then decreases, with significantly higher expression in the 4-year-old roots compared to the other age. In stems, *PkMADS1* expression gradually declines with age, reaching its lowest level in the 6-year-old samples. In buds, *PkMADS1* expression does not exhibit a clear age-dependent trend, with the highest expression in the central buds of the 20-year-old plants and the lowest in those of the 40-year-old plants. In reproductive organs, the expression of *PkMADS1* is significantly higher in the cones than in the flowers.

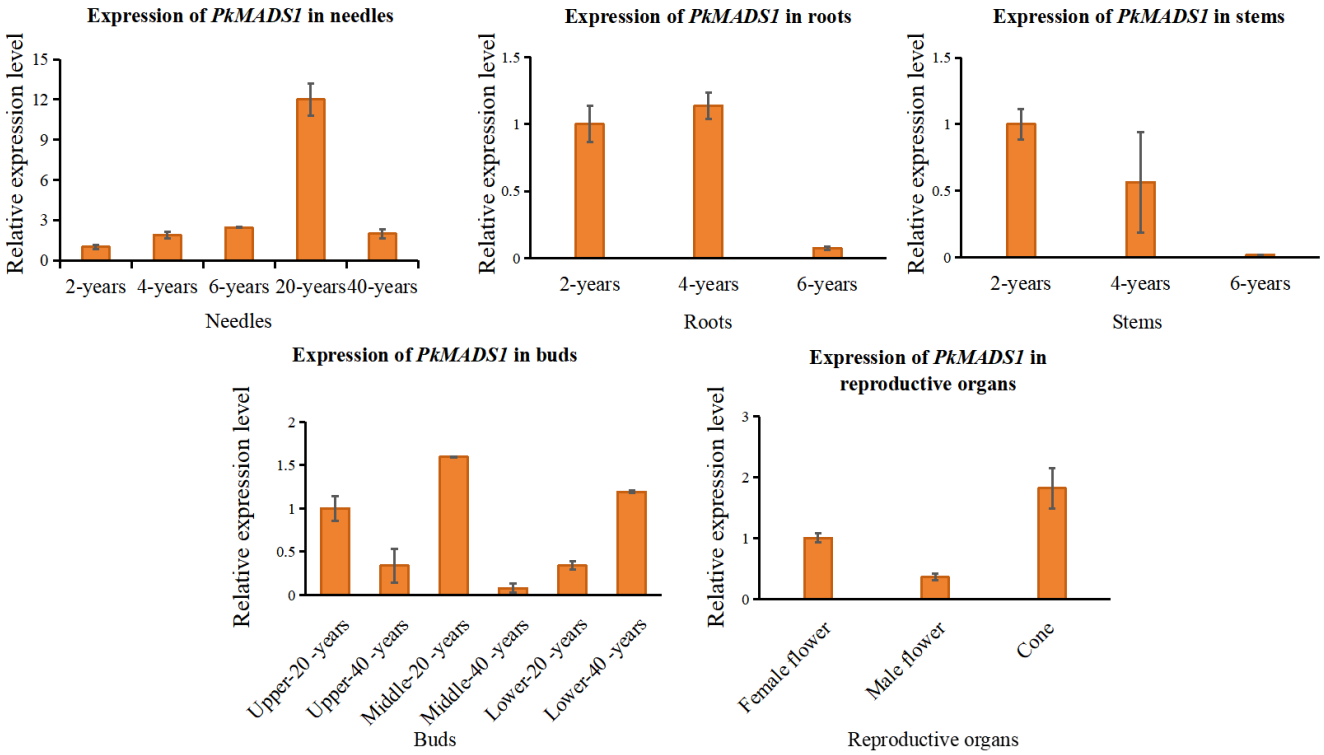

Figure 6. Expression of *PkMADS1* in different organs.

3.3.2. Results of the Analysis of *PkMADS3* Expression Pattern

Real-time quantitative PCR analysis (Figure 7) revealed that the expression pattern of *PkMADS3* in leaves from different ages was similar to that of *PkMADS1*. Its expression initially increased and then decreased with age, peaking in the 20-year-old leaves. In roots, *PkMADS3* expression exhibited an age-dependent trend, decreasing initially and then increasing, with much lower expression in the 4-year-old roots compared to the 2-year-old and 6-year-old roots. In stems, *PkMADS3* expression first increases and then decreases with age, with the 4-year-old stems showing significantly higher expression levels than those of the other age. In buds, the highest expression of *PkMADS3* was observed in the middle buds of the 40-year-old samples, followed by the lower buds of the 20-year-old samples. In reproductive organs, *PkMADS3* expression followed a pattern similar to *PkMADS1*, with the highest expression in cones and significantly lower expression in flowers.

3.3.3. Results of the Analysis of PkMADS4 Expression Pattern

Real-time quantitative PCR analysis (Figure 8) revealed that the expression pattern of *PkMADS4* in leaves from different ages was similar to that of *PkMADS1* and *PkMADS3*, with expression levels increasing initially and then decreasing with age. In roots, *PkMADS4* expression initially decreases and then increases, peaking in 6-year-old roots. *PkMADS4* expression in stems decreased with age. In buds, *PkMADS4* expression was significantly higher in middle buds than in other buds. *PkMADS4* expression showed slight differences between reproductive organs, with the highest expression in cones, followed by male flowers, and the lowest in female flowers.

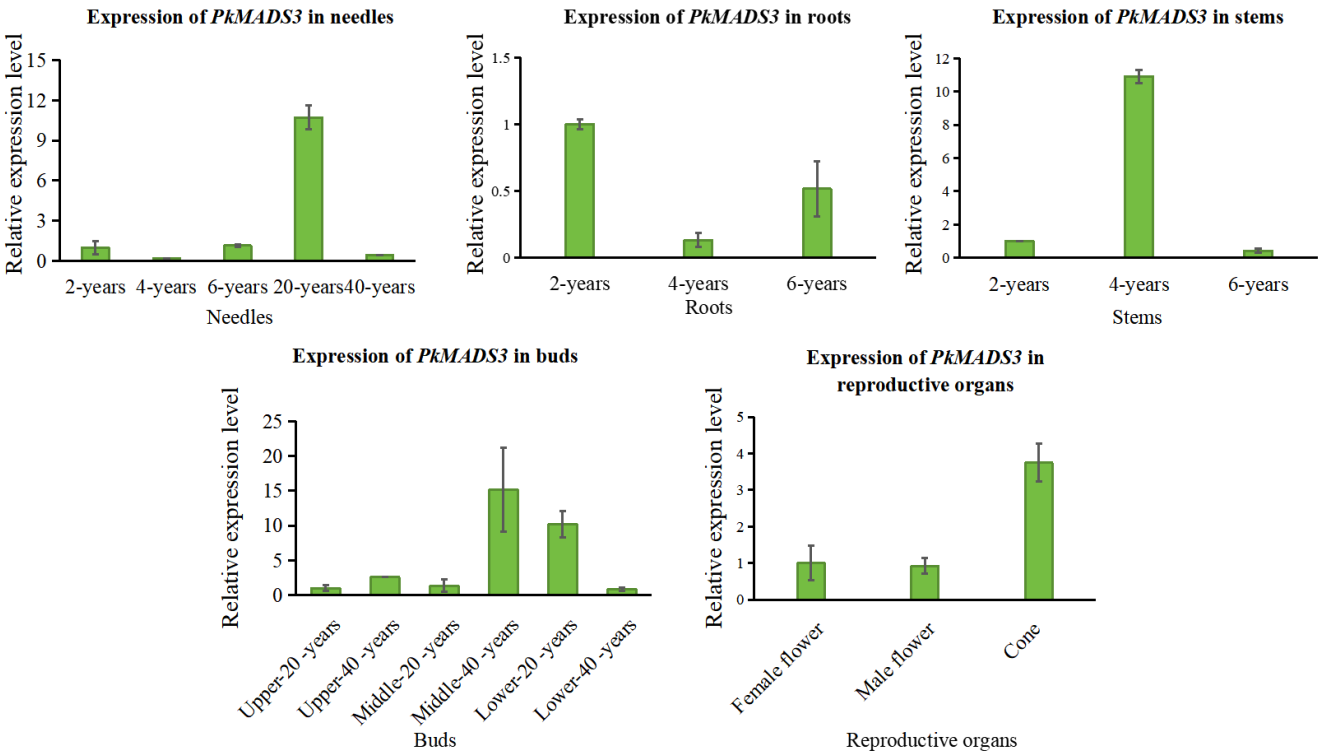

Figure 7. Expression of *PkMADS3* in different organs.

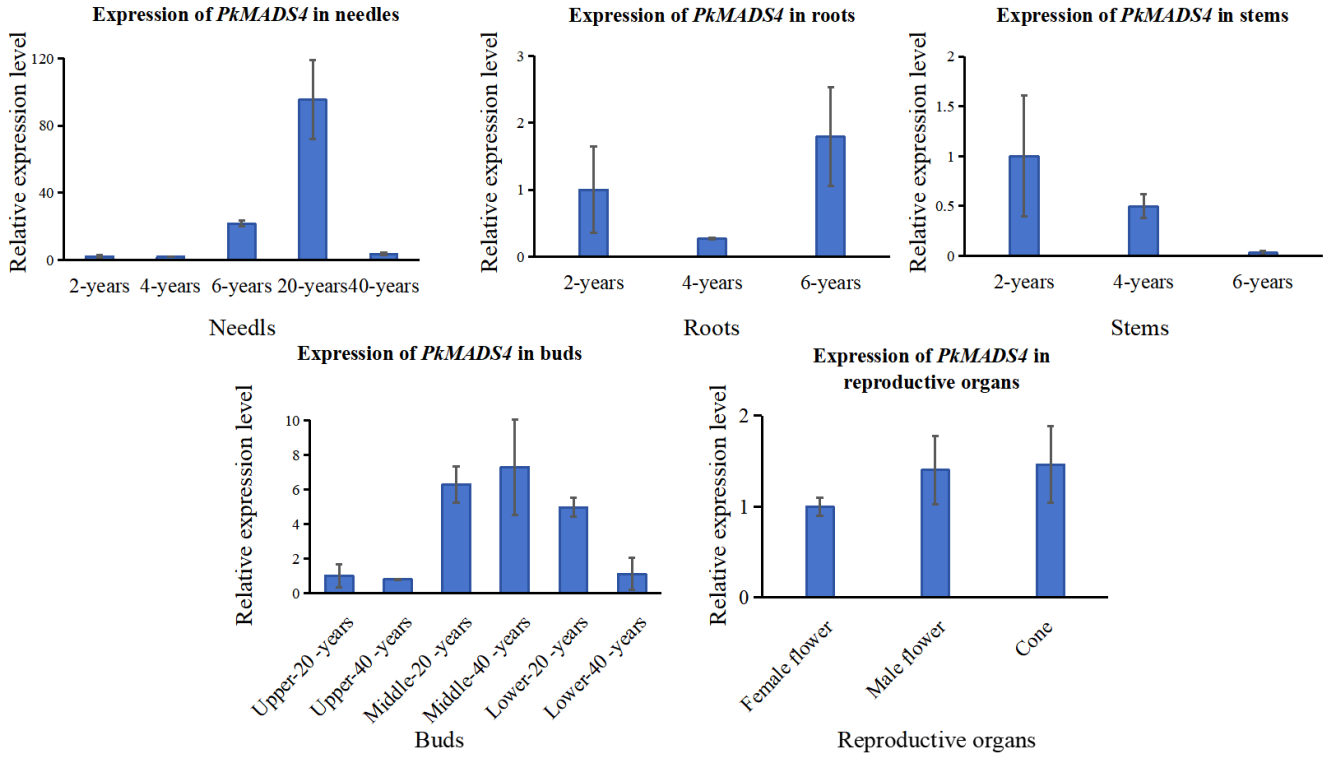

Figure 8. Expression of *PkMADS4* in different organs.

3.3.4. Comparison of Gene Expression in Upper Buds and Leaves

Korean pine plantations generally reach fruiting age at around 20 years, with female flowers and cones typically borne at the tips of the upper branches. To analyze the potential effects of the three genes on flowering and fruiting in Korean pine, the results of real-time fluorescence quantitative PCR in the leaves and upper buds of the three genes were

compared and analyzed (Figure 9). The results indicated that all three genes exhibited significantly higher expression in the leaves compared to the upper buds.

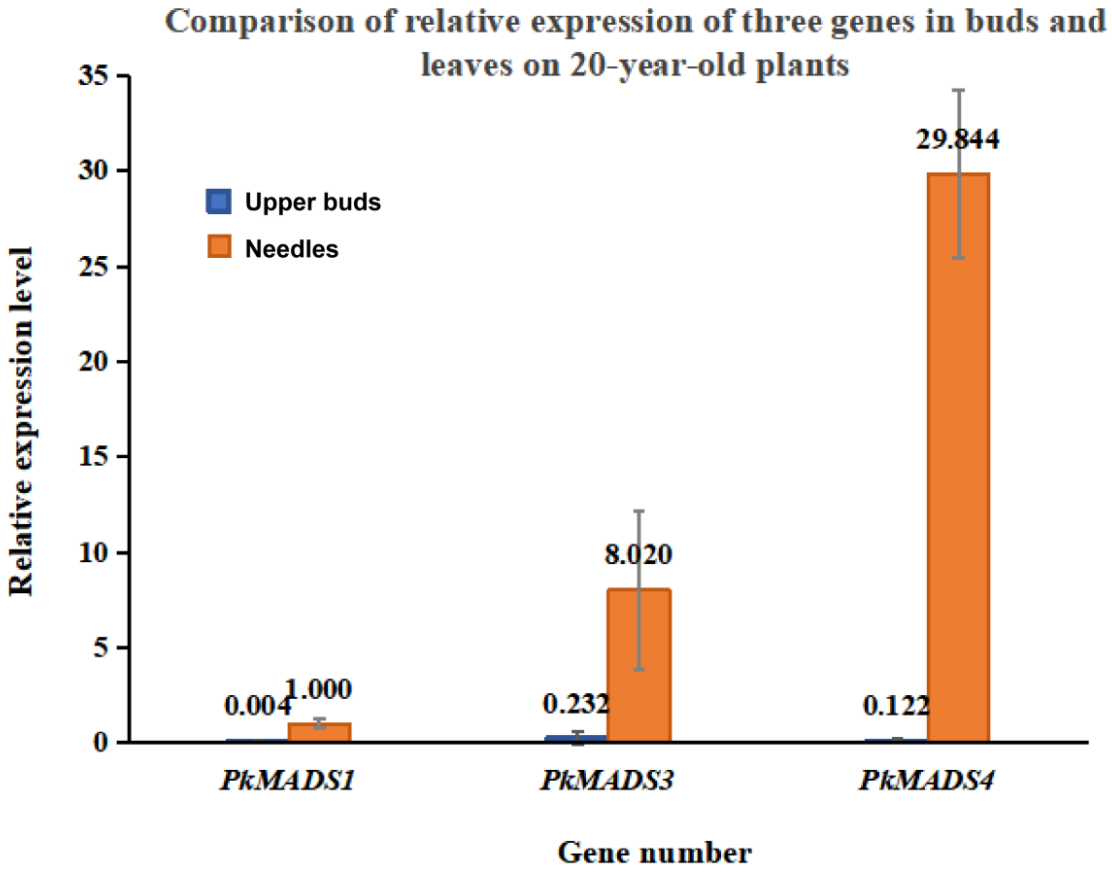

Figure 9. Comparison of the expression of the three genes in leaves and buds.

4. Discussion

Over the years, MADS-box genes have been recognized as one of the key gene families regulating reproductive growth and the development of vegetative organs. Currently, *Arabidopsis thaliana* is the most extensively studied plant in MADS-box gene research, followed by tomato (*Solanum lycopersicum* L.), apple, and rice (*Oryza sativa* L.) [27]. With the advances in genetics and genomics, the expression and function of MADS-box genes in different species are being investigated in greater depth. Genes like *FUL* and *AGL24* are reported to be directly or indirectly involved in stem growth and development in plants. Genes such as *AGL13* are associated with pollen development. *SOC1* and *AGL24* inhibit root growth, with *SOC1* acting as a negative regulator of root stem cell differentiation [28]. It has been reported that genes such as *SVP* and *AP1* are directly or indirectly associated with leaf growth and development [27]. In this study, the expression level of SVP-like genes in leaves was significantly higher than in other tissues, suggesting that *SVP* genes in Korean pine may also play a role in regulating leaf growth and development. A gene may show associations in different tissues. For example, recent reports indicate that *AGL24* is transported from the leaves to the distal growing point, promoting floral organ differentiation [29]. MADS-box genes are widely expressed in both reproductive and vegetative organs of ferns, while their expression pattern in seed plants is the opposite. MADS-box genes show evolutionary conservation in ferns and land plants, particularly in regulating floral organ development [30]. In this study, compared to *Arabidopsis*, the MADS-box gene family in Korean pine lacks the FLC subfamily. It has been reported that FLC plays a significant regulatory role in the vernalization pathway of barley and wheat [31]. *FLC* gene

silencing is regulated by the coordinated action of two pathways: COOLAIR and PRC2. Under low-temperature conditions, plant growth slows, leading to the accumulation of the transcription factor *NTL8*, which promotes *FLC* silencing [32], such that the longer the duration of low temperatures, the weaker the inhibition of flowering by *FLC* [33]. It is noteworthy that, despite inhabiting mid- to high-latitude regions, Korean pine, as a non-winter annual plant, does not require vernalization for its flowering process [34]. Consequently, the absence of the *FLC* gene in Korean pine can be interpreted as an adaptation. The results of the conserved motif analysis indicate that, while there is a high degree of similarity among the conserved motifs of different MADS-box genes, structural differences exist, with some genes lacking certain motifs. This is consistent with the notion that during the process of evolution, not only may the functions of genes undergo changes, but their structures may also be subject to variation [35]. However, the specific relationship between the structural diversity and functional diversity of Korean pine MADS-box genes remains to be further investigated. From the results of the physicochemical property analysis of Korean pine MADS-box proteins, it is observed that the majority of these proteins exhibit similar properties, with only a minority displaying differences. For instance, all proteins except PkMADS7 are hydrophilic, all proteins except PkMADS9 are non-secretory, and all proteins except PkMADS10 are unstable. Some studies have shown that the MADS-box family of proteins, a widely distributed transcription factor in plants, is normally expressed in the nucleus of cells [36]. From the results of subcellular localization prediction, it can be seen that most of the MADS proteins in Korean pine are expressed in the nucleus, which suggests that the MADS-box proteins of Korean pine identified in this study share characteristics with those of MADS-box proteins in other plant species. According to the results of protein secondary and tertiary structure prediction (Figure 3, Table 3), it can be seen that the secondary structure of MADS-box genes is mainly composed of alpha helices and random coils, which is similar among all MADS-box proteins. However, in the tertiary structure, the spatial distribution characteristics of alpha helices, random coils, and extended strands are very different among different proteins, which may be reflected in their biological functions. The specific relationship may need further research [37]. According to existing reports on *SVP* genes, the relationship between *Pinus tabuliformis* (*Pinus tabuliformis* Carrière) and the *Arabidopsis thaliana* SVP subfamily is relatively close [13]. The results of the conserved motif analysis of Korean pine and *Arabidopsis thaliana* MADS-box genes in this study also showed that the relationship between Korean pine and *Arabidopsis thaliana* SVP family is relatively close (Figure 4). Therefore, we hypothesize that there may be a certain degree of similarity in the *SVP* gene functions between these two species. It has been reported that the SVP subfamily genes of *Arabidopsis thaliana* are central regulators in the flowering regulatory network [38]. While *Arabidopsis thaliana* has two major SVP subfamily genes, four SVP subfamily genes were found in Korean pine in this study. This suggests that the SVP subfamily genes in Korean pine may be involved in more complex regulation of flowering time and flower development.

There have been many reports describing the expression of SVP subfamily genes in different tissues of various species. For instance, it has been reported that SVP subfamily genes are expressed in several tissues and organs of *Paeonia lactiflora* (*Paeonia lactiflora* Pall.), but the expression levels vary across different tissues, with particularly low expression in the petals and relatively higher expression in the roots [39]. In *Actinidia chinensis* (*Actinidia chinensis* Planch.), SVP subfamily genes are highly expressed in vegetative organs such as leaves, apical shoots, and axillary buds, with lower expression in flowers [40]. In *Pinus tabuliformis*, four SVP-like genes exhibit elevated expression in both vegetative buds and female cones, with *PISVL2* and *PISVL3* also showing higher expression in needles, suggesting that SVP-like genes may play a crucial role in the reproductive transition and cone

development of *Pinus tabuliformis* [13]. The expression levels of the *SVP* genes in mango vary with age within the same tissue. These genes show high expression levels in buds, leaves, and stems, with *MiSVP3* exhibiting the highest expression in stems, *MiSVP4* in leaves, and the lowest expression in flowers [36]. The expression patterns of the reported *SVP* genes in various organs and tissues across different species share many similarities with the expression patterns of *SVP* genes in Korean pine observed in this study, suggesting that *SVP* genes may play similar roles or be regulated by similar mechanisms in multiple species.

PkMADS3 expression was the lowest in various tissues compared to the other two genes (Figure 10). Overall, the expression trends of these three genes across different tissues were similar. The expression of *PkMADS1*, *PkMADS3*, and *PkMADS4* genes was relatively high in leaves and lowest in reproductive organs (Figure 9), which aligns with other studies on the expression patterns of SVP subfamily genes in different tissues across various species [37]. The expression of these three genes in the leaves of plants of different ages exhibited a regular pattern, showing an initial increase followed by a decrease, with the highest expression observed in the leaves of 20-year-old plants before starting to decline (Figure 10). In Korean pine plantations, under full sunlight conditions, flowers typically begin to bloom around 10 years of age, with mature seeds being harvested at approximately 15 years [41]. Furthermore, some studies have demonstrated that the expression of SVP subfamily genes in the leaves of species like mango, saffron, spring orchid, and hedgehog tree varies according to the developmental stage of the flowers [42]. It is thus hypothesized that the expression of Korean pine SVP subfamily genes in leaves is also correlated with the stage of flower development. In terms of stems, the expression patterns of *PkMADS1* and *PkMADS4* were similar, both decreasing over time. However, *PkMADS3* exhibited a distinct pattern, with the highest expression in 4-year-old stems, following an increase and subsequent decrease, though the overall expression trend was still downward, similar to the expression pattern of SVP subfamily genes in mango [43].

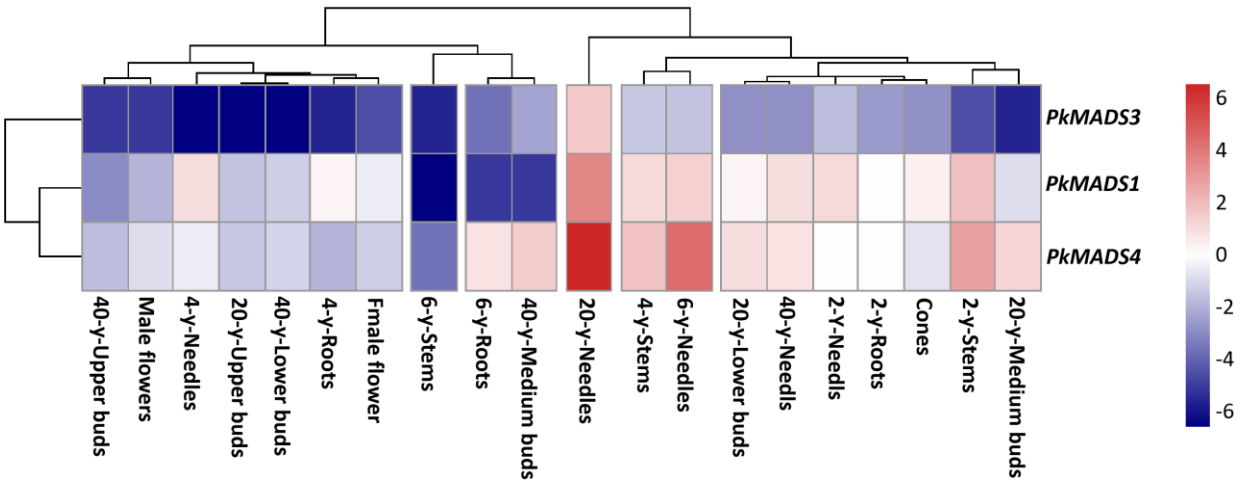

Figure 10. Differences in the expression of Korean pine MADS-box gene in different tissues.

The SVP subfamily genes, as pivotal members of the MADS-box gene family, have been reported to regulate the mechanisms of flowering, floral organ development, and dormancy in numerous plant species [13]. For instance, scientific research has identified the presence of SVP subfamily genes in various ornamental plants, such as wintersweet (*Chimonanthus praecox* (L.) Link), Japanese morning glory (*Ipomoea nil* (L.) Roth), German iris (*Iris germanica* L.), cotton (*Gossypium* Linn.), and roses, and has investigated their roles in the flowering mechanisms of these plants [44]. Upon *PtSVL1* gene of *Pinus tabuliformis*

transformed into *Arabidopsis thaliana*, it was observed that the pollen tube development was affected [13]. Additionally, research has found that the heterologous transformation of the sweet cherry (*Prunus avium* L.) *PavSVP* gene into *Arabidopsis thaliana* results in a significant reduction in the length of the *Arabidopsis siliques* [45]. However, there are no reports on how SVP subfamily genes regulate the flowering process in Korean pine.

5. Conclusions

In conclusion, 12 MADS-box family genes were identified from Korean pine transcriptome data, and these genes were mainly found in five subfamilies. Conserved motif analysis revealed that these genes share a common conserved motif. By analyzing the physicochemical properties and structures of the genes, many similarities were found in these genes, indicating that genes with similar sequences have similar physicochemical properties. The analysis of gene expression patterns in different tissues and organs revealed that the expression levels of the three *SVP* genes in leaves were significantly higher than in buds (approximately 200-fold higher). Additionally, the heatmap (Figure 10) further demonstrated that the expression levels of these genes in leaves were not only higher than in buds but also exceeded those in other tissues. Based on these results, we speculate that the roles of these three SVP genes in Korean pine are similar to those in other species, primarily playing a suppressive role in the process of flower bud formation, which helps Korean pine maintain a juvenile state under certain conditions. They may also be involved in the growth and development of leaves. However, the specific mechanisms of action require further investigation. This study provides valuable insights for further research into the flowering mechanisms of Korean pine.

Author Contributions: Conceptualization, H.Z. and L.Z.; Validation, J.D. (Junshuai Du); Formal analysis, J.D. (Junshuai Du); Investigation, J.D. (Junshuai Du), D.H., J.H. and J.D. (Junping Du); Data curation, J.D. (Junshuai Du); Writing—original draft, J.D. (Junshuai Du); Writing—review & editing, L.Z.; Visualization, J.D. (Junshuai Du) and J.H.; Project administration, L.Z.; Funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number No. 2023YFD2200605.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: This work was supported by the staff of the Qingshan Larch National Improved Seed Base in Linkou County.

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

References

1. Ren, Q.X.; Ren, S.X.; Li, W.Y. Problems and countermeasures in the development of *Pinus koraiensis* nut grove industry in Xing'an league. *North. Fruits* **2023**, *6*, 47–49.
2. Wu, H.B. Photosynthetic Physiological Characteristics and Distribution of Photosynthate During Cone Growth of Pinus koraiensis. Doctor's Thesis, Northeast Forestry University, Harbin, China, 2023.
3. Farjon, A. *Handbook of the World's Conifers*; Brill: Leiden, The Netherlands; Boston, MA, USA, 2010; Volume 1 & 2, p. 1112.
4. Schwarz-Sommer, Z.; Huijser, P.; Nacken, W.; Saedler, H.; Sommer, H. Genetic Control of Flower Development by Homeotic Genes in *Antirrhinum majus*. *Science* **1990**, *250*, 931–936. [CrossRef] [PubMed]
5. Mou, Y.; Yuan, C.; Sun, Q.; Yan, C.; Zhao, X.; Wang, J.; Wang, Q.; Shan, S.; Li, C. MIKC-type MADS-box transcription factor gene family in peanut: Genome-wide characterization and expression analysis under abiotic stress. *Front. Plant Sci.* **2022**, *13*, 980933. [CrossRef] [PubMed]
6. Wang, S.L.; Viswanath, K.K.; Tong, C.G.; An, H.R.; Jang, S.; Chen, F.C. Floral Induction and Flower Development of Orchids. *Front. Plant Sci.* **2019**, *10*, 1258. [CrossRef]

7. Liu, C.; Teo, Z.W.; Bi, Y.; Song, S.; Xi, W.; Yang, X.; Yin, Z.; Yu, H. A conserved genetic pathway determines inflorescence architecture in *Arabidopsis* and rice. *Dev. Cell* **2013**, *24*, 612–622. [CrossRef]
8. Yoshida, A.; Sasao, M.; Yasuno, N.; Takagi, K.; Daimon, Y.; Chen, R.; Yamazaki, R.; Tokunaga, H.; Kitaguchi, Y.; Sato, Y.; et al. *TAWAWA1*, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 767–772. [CrossRef]
9. Zhang, X.H.; Shen, H.Y.; Wen, B.B. Molecular mechanism of PpCMB1 gene in peach MADS-box family regulating flower development. *Plant Physiol. J.* **2021**, *57*, 1211–1217.
10. Liu, X.; Sun, Z.; Dong, W.; Wang, Z.; Zhang, L. Expansion and Functional Divergence of the SHORT VEGETATIVE PHASE (SVP) Genes in Eudicots. *Genome Biol. Evol.* **2018**, *10*, 3026–3037. [CrossRef]
11. Liu, J.; Ren, M.; Chen, H.; Wu, S.; Yan, H.; Jalal, A.; Wang, C. Evolution of SHORT VEGETATIVE PHASE (SVP) genes in Rosaceae: Implications of lineage-specific gene duplication events and function diversifications with respect to their roles in processes other than bud dormancy. *Plant Genome* **2020**, *13*, e20053. [CrossRef]
12. Lee, J.H.; Yoo, S.J.; Park, S.H.; Hwang, I.; Lee, J.S.; Ahn, J.H. Role of *SVP* in the control of flowering time by ambient temperature in *Arabidopsis*. *Genes Dev.* **2007**, *21*, 397–402. [CrossRef]
13. Zhou, C.; Liu, H.; Wang, H.; Niu, S.; El-Kassaby, Y.A.; Li, W. Deciphering the Role of *SVP*-Like Genes and Their Key Regulation Networks During Reproductive Cone Development in *Pinus tabuliformis*. *Plant Cell Environ* **2025**, *48*, 365–386. [CrossRef] [PubMed]
14. Coudert, E.; Gehant, S.; de Castro, E.; Pozzato, M.; Baratin, D.; Neto, T.; Sigrist, C.J.A.; Redaschi, N.; Bridge, A.; UniProt Consortium. Annotation of biologically relevant ligands in UniProtKB using ChEBI. *Bioinformatics* **2023**, *39*, btac793. [CrossRef] [PubMed]
15. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]
16. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **2024**, *52*, W78–W82. [CrossRef]
17. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49. [CrossRef]
18. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y.; et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* **2023**, *16*, 1733–1742. [CrossRef]
19. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999**, *112*, 531–552.
20. King, R.D.; Sternberg, M.J. Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Sci.* **1996**, *5*, 2298–2310. [CrossRef]
21. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef]
22. Bannai, H.; Tamada, Y.; Maruyama, O.; Nakai, K.; Miyano, S. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* **2002**, *18*, 298–305. [CrossRef]
23. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [CrossRef] [PubMed]
24. Horton, P.; Park, K.J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.J.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **2007**, *35*, W585–W587. [CrossRef] [PubMed]
25. Juan, W. Primer Design with Primer Premier 5.0. *Northwest Med. Educ.* **2008**, *4*, 695–698.
26. Li, Y.X. Physiological and Molecular Mechanisms of Pinuskoraiensis Seedlings in Response to Different Light Conditions Research. Doctor's Thesis, Northeast Forestry University, Harbin, China, 2023.
27. Castañón-Suárez, C.A.; Arrizubieta, M.; Castelán-Muñoz, N.; Sánchez-Rodríguez, D.B.; Caballero-Cordero, C.; Zluhan-Martínez, E.; Patiño-Olvera, S.C.; Arciniega-González, J.A.; García-Ponce, B.; Sánchez, M.P.; et al. The MADS-box genes SOC1 and AGL24 antagonize XAL2 functions in *Arabidopsis thaliana* root development. *Front. Plant Sci.* **2024**, *15*, 1331269. [CrossRef]
28. Adhikari, P.B.; Kasahara, R.D. An overview on MADS box members in plants: A meta-review. *Int. J. Mol. Sci.* **2024**, *25*, 8233. [CrossRef]
29. Huang, N.C.; Tien, H.C.; Yu, T.S. Arabidopsis leaf-expressed AGAMOUS-LIKE 24 mRNA systemically specifies floral meristem differentiation. *New Phytol.* **2024**, *241*, 504–515. [CrossRef]
30. Zhang, R.; Zhang, J.; Xu, Y.X.; Sun, J.M.; Dai, S.J.; Shen, H.; Yan, Y.H. Dynamic evolution of MADS-box genes in extant ferns via large-scale phylogenomic analysis. *Front. Plant Sci.* **2024**, *15*, 1410554. [CrossRef]
31. Alexandre, C.M.; Hennig, L. *FLC* or not *FLC*: The other side of vernalization. *J. Exp. Bot.* **2008**, *59*, 1127–1135. [CrossRef]
32. Nielsen, M.; Menon, G.; Zhao, Y.; Mateo-Bonmati, E.; Wolff, P.; Zhou, S.; Howard, M.; Dean, C. COOLAIR and PRC2 function in parallel to silence FLC during vernalization. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e2311474121. [CrossRef]

33. Hong, W.; Cao, J. The Function of *FLC* in Vernalization Process. *Chin. Bull. Bot.* **2002**, *19*, 406–411.
34. Xu, S.; Chong, K. Remembering winter through vernalisation. *Nat. Plants* **2018**, *4*, 997–1009. [CrossRef] [PubMed]
35. Babenko, V.N.; Rogozin, I.B.; Mekhedov, S.L.; Koonin, E.V. Prevalence of intron gain over intron loss in the evolution of paralogous gene families. *Nucleic Acids Res.* **2004**, *32*, 3724–3733. [CrossRef] [PubMed]
36. Mo, X.; Luo, C.; Xia, L.; Mo, W.; Zhu, J.; Zhang, Y.; Hu, W.; Liu, Y.; Xie, F.; He, X. Overexpression of mango *MiSVP3* and *MiSVP4* delays flowering time in transgenic *Arabidopsis*. *Sci. Hortic.* **2023**, *317*, 112021. [CrossRef]
37. Xu, D.; Hao, J.; Wang, C.; Zhang, L.; Zhang, H. Analysis of the Expression Patterns of 13 DREB Family GenesRelated to Cone-Setting Genes in Hybrid Larch (*Larix kaempferi* × *Larix olgensis*). *Forests* **2023**, *14*, 2300. [CrossRef]
38. Li, D.; Liu, C.; Shen, L.; Wu, Y.; Chen, H.; Robertson, M.; Helliwell, C.A.; Ito, T.; Meyerowitz, E.; Yu, H. A repressor complex governs the integration of flowering signals in *Arabidopsis*. *Dev. Cell* **2008**, *15*, 110–120. [CrossRef]
39. Ji, F.; Ma, Y.; Qi, S.; Guo, X.; Chen, J. Cloning and functional analysis of peony *PlSVP* gene in regulating flowering. *Acta Hortic. Sin.* **2022**, *49*, 2367–2376.
40. Wu, R.M.; Walton, E.F.; Richardson, A.C.; Wood, M.; Hellens, R.P.; Varkonyi-Gasic, E. Conservation and divergence of four kiwifruit *SVP*-like MADS-box genes suggest distinct roles in kiwifruit bud dormancy and flowering. *J. Exp. Bot.* **2012**, *63*, 797–807. [CrossRef]
41. Yang, K.; Gu, H.Y. Dynamic changes of hormone in the plants from teneral stage to blossom phase of *Pinus koraiensis* Fruit Forests. *Sci. Silvae Sin.* **2005**, *41*, 33–37.
42. Sun, L.; Xu, Z.; Huang, W.; Wu, S.; Lin, X.; Zhu, F.; Liu, N.; Huang, M.; Chen, R.; Zeng, H. Preliminary study of differentiating smears from cancerous and non-cancerous nasopharyngeal tissue using confocal Raman spectroscopy. *J. Cancer Res. Clin. Oncol.* **2016**, *142*, 823–831. [CrossRef]
43. Mo, X.; Luo, C.; Yu, H.; Chen, J.; Liu, Y.; Xie, X.; Fan, Z.; He, X. Isolation and Functional Characterization of Two *SHORT VEGETATIVE PHASE* Homologous Genes from Mango. *Int. J. Mol. Sci.* **2021**, *22*, 9802. [CrossRef]
44. Chen, J.W.; He, X.H.; Luo, C.; Fan, Y.; Zhang, X.J.; Yu, H.X. Research Progress of Plants Flowering Suppressor Homologous Genes of SVPs. *Mol. Plant Breed.* **2017**, *15*, 4888–4898.
45. Wang, J.; Jiu, S.; Xu, Y.; Sabir, I.A.; Wang, L.; Ma, C.; Xu, W.; Wang, S.; Zhang, C. *SVP*-like gene *PavSVP* potentially suppressing flowering with *PavSEP*, *PavAP1*, and *PavJONITLESS* in sweet cherries (*Prunus avium* L.). *Plant Physiol. Biochem.* **2021**, *159*, 277–284. [CrossRef]