forests

MDPI

Article

Genome-Wide Identification and Transcriptional Expression Profiles of the *F-box* Gene Family in Common Walnut (*Juglans regia* L.)

Feng Yan [1], Huijuan Zhou [2], Ming Yue [3], Ge Yang [4], Huaizhu Li [5], Shuoxin Zhang [2] and Peng Zhao [1,*]

1 Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an 710069, China; fengyan115@163.com
2 College of Forestry, Northwest A&F University, Yangling 712100, China; ericguanyuzhao@163.com (H.Z.); sxzhang@nwsuaf.edu.cn (S.Z.)
3 Xi'an Botanical Garden of Shaanxi Province, Xi'an 710061, China; yuming@nwu.edu.cn
4 College of Biological Sciences and Engineering, Shaanxi University of Technology, Hanzhong 723001, China; geyang0125@163.com
5 School of Chemistry and Chemical Engineering, Xianyang Normal University, Xianyang 712000, China; lihuaizhu121@126.com
* Correspondence: pengzhao@nwu.edu.cn; Tel.: +86-298-830-2411

Received: 10 February 2019; Accepted: 15 March 2019; Published: 20 March 2019

Abstract: The common walnut (or Persian walnut), *Juglans regia* L., is an economically important temperate tree species valued for both its edible nut and high-quality wood. *F-box* gene family members are involved in plant development, which includes regulating plant development, reproduction, cellular protein degradation, response to biotic and abiotic stresses, and flowering. However, in common walnut (*J. regia*), there are no reports about the *F-box* gene family. Here, we report a genome-wide identification of *J. regia F-box* genes and analyze their phylogeny, duplication, microRNA, pathway, and transcriptional expression profile. In this study, 74 *F-box* genes were identified and clustered into three groups based on phylogenetic analysis and eight subfamilies based on special domains in common walnut. These common walnut *F-box* genes are distributed on 31 different pseudo-chromosomes. The gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and microRNA profiles showed that the *F-box* gene family might play a critical role in the flowering of common walnut. The expressions were significantly higher in female flowers and male flowers compared with leaf and hull tissues at a transcriptome level. The results revealed that the expressions of the *F-box* gene in female flowers were positively correlated with male flowers, but there was no correlation between any other tissue combinations in common walnut. Our results provided insight into the general characteristics of the *F-box* genes in common walnut.

Keywords: *Juglans regia*; *F-box* gene family; transcriptome; male flower; female flower; expression profile; flowering

1. Introduction

The ubiquitin–proteasome pathway, which involves the F-box protein, is one of the most important biological regulatory systems. The F-box protein is also as a component of complexes involved in the substrate recognition subunit of Skp1-Cullin1-F-box (SCF) ubiquitin–ligase [1–4]. Recent studies indicated that the *F-box* gene is an important factor in the regulation of plant growth and development and response to stress [5–7]. The F-box protein contains a WD repeat, leucine zipper, tetratricopeptide repeats (TPRs), Kelch and ring finger structures, and zinc finger structures [2,8].

Based on the different types of conserved domains of the *F-box* gene family, F-box proteins could be involved in many biological development processes, such as leaf senescence and branching [9,10], flowering [11,12], circadian rhythms [13,14], self-incompatibility [15–17], phytochrome signaling [18], and responses to plant growth regulators [19,20] and abiotic [20,21] and biotic factors [22].

A total of 38 *F-box* genes were found in humans, whereas the common fruit fly only contains 22 *F-box* genes [2]. However, the number of *F-box* gene family members in plants is much higher than that of animals [23,24]. It was identified that model plants of *Arabidopsis* contain 694 *F-box* genes [23]. Important crops also contain *F-box* genes—for example, rice, maize, soybean, and chickpea contain 678, 509, 359, and 285 *F-box* genes, respectively [25–28]. The same applies to perennial fruit crops, such as apples and pears, which contain 517 and 226 *F-box* genes, respectively [24,29], as well as perennial woody plants, such as *Populus trichocarpa* Torr. & A.Gray ex. Hook., apple, and pear, which contain 320, 517, and 226 *F-box* genes, respectively [24,29,30]. The number of *F-box* genes of woody plants and herbaceous annual plants (*Arabidopsis thaliana* (L.) Heynh., rice, maize, and soybean) has high variation [24–27,29,30]. Under abiotic stresses, the expression profiles showed that the *F-box* gene family members play a crucial role in stress responsive pathways [25,28,31]. A total of 972 *F-box* genes were found in *Medicago* (*Medicago truncatula* Gaertn.), and most of the *F-box* genes of this species could respond to salt and heavy-metal stresses [31]. At least 43 *F-box* genes were found to be differentially expressed in rice seedlings subjected to different abiotic stress conditions. The *F-box* genes had differential expression patterns in various chickpea tissues, as well as under abiotic stress conditions [28]. Furthermore, the expression of several *F-box* genes was also influenced by light in rice [25]. In apple (*Malus domestica* Borkh.) and Japanese pear (*Pyrus pyrifolia* Nakai), since each S haplotype contains two or three related genes, the genes were named *SFBB* for S-locus *F-box* brothers. The *SFBB* genes were specifically expressed in pollen, and variable regions of the *SFBB* genes experienced positive selection [32]. Many *Prunus* species, including sweet cherry and Japanese apricot, of the Rosaceae family, display an S-RNase-based gametophytic self-incompatibility (GSI). An F-box protein is encoded by a gene located in the S locus region, named *SFB*, and it was found that there were two self-compatible (SC) haplotypes, $S^{4'}$ and S^{f} of *Prunus*, which caused a pollen-part mutant (PPM). The Sf of Japanese apricot was also considered to be a PPM. Both were examined and it was found that SFB was the pollen S gene in GSI in *Prunus* [33]; an *F-box* gene named *SLF* (*S-locus F-box*) was specifically expressed in pollen, but not in the styles or leaves [34]. Four additional alleles (*SFB1*, *SFB2*, *SFB4*, and *SFB5*) were cloned from sweet cherry (*P. avium*) and examined, and they were found to have the function of allele specificity of the GSI reaction [35]. In Japanese pear (*Pyrus pyrifolia*), the mutant haplotype S^{4sm} lacking SFBB1-S^4 caused the pollen to be rejected by pistils with an otherwise compatible S^1, while it was accepted by other non-self pistils. Moreover, the S^5 haplotype encoded a truncated SFBB1 protein, even though S^5 pollen was accepted normally by pistils with S^1 and other non-self haplotypes [36]. The various subfamilies of the *F-box* gene family were identified in different plants. The *F-box* genes of *Arabidopsis* were classified into five subfamilies [23], while there were nine in chickpea [28], 10 in rice [25], 15 in maize [26], three in *Medicago* [30], and 12 in pear [29]. The huge size of the *F-box* gene family was identified in many plants, which indicated that it evolved through duplication events [23,25,31]. In terms of the gene family expansion, this took place through whole-genome duplications, small-scale segmental duplications, local tandem duplications, or combinations of these possibilities [23,25,29]. The *F-box* genes of *Medicago* probably expanded due to localized gene duplications [31]. The dispersed and tandem duplication evolved at a high rate in pear [29]. In maize, the expansion of *F-box* genes was mainly because of tandem and segmental duplication events [27], and, in rice, the expansion might have been due to localized gene duplications [25]. The contributions made by segmental and tandem duplications in the expansion of the *F-box* gene family in chickpea were also examined [28], and whole-genome duplication and dispersed duplication made a major contribution to *F-box* family expansion [25,29].

Juglans regia L., a diploid ($2n$ = 32) walnut species, is known as Persian, English, or common walnut. It is native to the mountainous regions of central Asia [37]. Walnuts and butternuts (*Juglans*)

are known for their edible nuts and high-quality wood [38,39]. It is an ecologically important tree species valued for both its nuts and wood since ancient times [39,40]. *J. regia* is diploid monoecious and heterodichogamous. It is wind-pollinated and highly heterozygous [40]. There is no previous study regarding the *F-box* gene in common walnut; comprehensive information and functional characterization of the *F-box* gene family of common walnut (*J. regia*) is still unclear, especially the expression profiles in male and female flowers. [41] To obtain the genes belonging to the *F-box* gene family of common walnut, a genome-wide analysis was performed in this study, where the gene structure, phylogenetic tree, chromosomal location, and conserved domains were examined. Additionally, to investigate the function of the *F-box* gene family of common walnut, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, expression profiles, protein domains, and promoter predictions are presented, as they constitute useful information for the *F-box* gene family. Importantly, the expression patterns of *F-box* genes in different common walnut tissues with public transcriptome data were examined. To better understand whether the *F-box* gene may play an important role in flowering, we sampled and examined the expression level in female and male flowers, which showed that the *F-box* genes of common walnut were consistent with flowering.

2. Materials and Methods

2.1. Bioinformatics Analysis of Putative F-box Proteins from J. regia

The whole-protein sequences of common walnut were downloaded from the NCBI (National center for biotechnology information) database (https://treegenesdb.org/FTP/Genomes/Jure/v1.4/annotation/). *Arabidopsis F-box* gene family members were downloaded from the *Arabidopsis* information resource website (TAIR, https://www.arabidopsis.org/index.jsp). We used *Arabidopsis* F-box protein sequences as query sequences to search against common walnut protein sequences using a local basic local alignment search tool (BLAST), considering those with an *E*-value less than 1×10^{-10} as common walnut F-box protein sequences. We used a profile hidden Markov model (HMM) implemented with default parameters in HMMER v3.2.1 for Windows (http://hmmer.org/download.html/) to search for common walnut F-box proteins as common walnut proteins with the F-box domain in the protein family (Pfam) database (http://pfam.xfam.org/). Each of the retrieved sequences was manually checked based on the features of the F-box domain structure [29].

2.2. F-box Protein Alignment, Phylogenetic Analysis, Pfam Domain Detection, and Chromosome Location Analysis of Common Walnut F-box Genes

The complete common walnut F-box protein sequences were aligned by MEGA7.0 with default settings [42]. Subsequently, MEGA7.0 software (Molecular evolutionary genetics analysis, Pennsylvania State University: State College, PA, USA) was employed to construct an unrooted phylogenetic tree based on alignments using the neighbor-joining (NJ) method [42,43]. The neighbor-joining phylogenetic trees were constructed with pairwise deletion of 1000 bootstraps and a Poisson model [44]. All of the above 74 F-box protein sequences were detected in the search for potential domains contained in each sequence using the Pfam web server (http://pfam.xfam.org/), and those domains found by Pfam (The large collection of protein families database) were also detected by the SMART (Simple modular architecture research tool) program (http://smart.embl-heidelberg.de/) with an *E*-value cutoff of 1.0 to validate the final result. According to the special domains found in those F-box protein sequences, we classified these sequences into different subfamilies, and their distribution on the 31 common walnut pseudo-chromosomes was visualized using MapChart (Wageningen university: Wageningen, Netherlands, version 2.2) (http://mapchart.software.informer.com/2.2/) with the default parameters.

2.3. The KEGG Pathway, Promoter, Gene Structure, and Motif Analysis of J. regia F-box Genes

The common walnut full length of *F-box* genes was submitted to online software KEGG (https://www.kegg.jp/) to acquire the KEGG (Kyoto encyclopedia of genes and genomes) (Kyoto University: Kyoto, Japan, version 1.0) terms. The 1500 bp upstream of the *F-box* genes of each species was used to perform cis-element analysis in PlantCARE (Plant cis-acting regulatory element) with default parameters [44]. To analyze the number of gene introns, we used the coding sequences (CDS) of all the candidate common walnut *F-box* genes to BLAST against the *J. regia* genome sequences with a top *E*-value less than 1×10^{-20}. The whole CDS database was downloaded (https://treegenesdb.org/FTP/Genomes/Jure/v1.1/annotation/). The exon–intron structure was illustrated using the online gene structure display server program (http://gsds.cbi.pku.edu.cn/). Related walnut gene sequences were searched on the genome browser (https://www.ncbi.nlm.nih.gov/genome/). The motifs were identified using the multiple motif alignment for motif elicitation (MEME) program with default parameters [45]. The parameters were as follows: the maximum number of motifs was set to 20, and the optimum motif width was set to 30–50.

2.4. Synteny Analysis, and Calculating Ka, Ks, and Ka/Ks Values of Duplicated Gene Pairs

BLASTP was carried out to identify potential homologous gene pairs ($E < 1 \times 10^{-5}$, top 5 matches) across multiple genomes. Then, these homologous pairs were used as the input for MCScanX to identify syntenic chains [46]. MCScanX was further used to identify the duplicate gene pairs derived from whole-genome duplication (WGD)/segmental and tandem duplication, and other types of gene pairs. The Ka/Ks value was the ratio between the number of nonsynonymous substitutions per nonsynonymous site (Ka) and the number of synonymous substitutions per synonymous site (Ks). To estimate the type of selection *F-box* genes are under, the ratio of the rates of nonsynonymous to synonymous substitutions (Ka/Ks) of all sister pairs was calculated for each terminal branch of the phylogenetic trees of *J. regia* using DnaSP software (University of Barcelona, Barcelona, Spain, version 5.0) with default parameters [47]. To confirm the selection pressure, a Ka/Ks ratio greater than 1, less than 1, and equal to 1 represented positive selection, negative selection, and neutral selection, respectively [48]. For each gene pair, the Ks value was used to estimate the divergence time in millions of years based on a rate of 6.1×10^{-9} substitutions per site per year, and the divergence time (T) was calculated as $T = Ks/(2 \times 6.1 \times 10^{-9}) \times 10^{-6}$ million years ago (Mya) [49]. We did a maximum test of Ks values for the *F-box* gene family to discover the T belonging to common walnut.

2.5. Microarray Expression Profiles of F-box Genes

To investigate the expression pattern in common walnut, we downloaded the public transcriptional raw reads data from (https://treegenesdb.org/FTP/Genomes/Jure/v1.0/transcriptome/raw_reads/) of common walnut in different tissues [41] (Table S1, Supplementary Materials). We used the bowtie software (Johns Hopkins university, Baltimore, MD, USA, version 2-2.2.6 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) with default parameters to align sequencing reads to long reference sequences. The reference sequences were downloaded (https://treegenesdb.org/FTP/Genomes/Jure/v1.4/genome/), using Cufflinks software (University of Washington: Seattle, WA, USA, version 2.2.1) with default parameters (http://cufflinks.cbcb.umd.edu/) to assemble transcripts, estimate their abundances, and test for differential expression and regulation in RNA sequencing (RNA-Seq) samples, and we represented these results using HemI software (Huazhong University of Science and Technology, Wuhan, China, version1.0) with default parameters [50]. A hierarchical map was constructed and viewed with the software Mev (Multiple Experiment Viewer) (George Washington University, Washington, DC, USA, v4.9.0) [51] based on the normalization data.

2.6. Plant Materials, Treatments, and Collections

To assess the expression of common walnut *F-box* genes, we sampled a total of 12 fresh flowers, specifically nine fresh female flowers and three male flowers from common walnut individual trees which were grown on the Qinling Mountains, collected at different development times (Table S1, Supplementary Materials). Each tissue was collected from equivalent positions on the same tree. For each female flower development time, we collected female inflorescence as three, three, two, and one biological replicates from an adult tree growing on the Qinling Mountains on 10 April, 15 April, 22 April, and 1 May, respectively (Table S1, Supplementary Materials). For each male flower development time, we collected one male inflorescence from an adult tree (around 15 years old) growing on the Qinling Mountains on 10 April, 11 April, and 2 May (for details on flower tissue collection, see Table S1, Supplementary Materials). In this study, for female flower development stages, the first opening of female flowers occurred on 10 April, full female flowers occurred on 15 April and 22 April (specifically, the stigma was not fully developed on 15 April, but was on 22 April), and the end date of female flowers occurred on 1 May. For male flowers, the three respective stages occurred on 10 April, 11 April, and 2 May (Table S1, Supplementary Materials). We put the fresh female flowers and male flowers into liquid nitrogen prior to storage at −80 °C until use [38] (Table S1, Supplementary Materials). Total RNA was isolated using an RNA-prep Pure Plant Kit (Tiangen, Beijing, China) [32,51,52]. Libraries for RNA-seq were produced using NEBNext Ultra RNA Library Prep Kit (NEB, Beverly, MA, USA). Paired-end sequencing was performed on the Illumina HiSeq2500 platform to generate 100-bp reads with default parameters with Novogene Bioinformatics Technology Co., Ltd., Beijing, China (www.novogene.cn). The de novo transcriptome was assembled using default settings in Trinity [50] based on the well genome reference of *J. regia* [36]. To further characterize the different temporal and spatial gene expression patterns of the *F-box* gene family, we analyzed RNA sequencing (RNA-seq) data. The transcriptome sequencing datasets were deposited in the BioProject identifier (ID) PRJNA358784, which was used to perform RNA-seq of different *J. regia* wild female flowers and male flowers. We analyzed the total RNA-seq data of the female flowers and male flowers at the germination initial flowering stages (Table S1, Supplementary Materials). We quantified these gene expression levels based on their fragments per kilobase of exon per million reads mapped (FPKM) values using Cufflinks with default parameters [53] and represented these results using HemI 1.0 software with default parameters [45,54]. For the expression level of several *F-box* genes with error bars, the ggplot2 R package was used [55]. Differential gene expression (DESeq) analysis was performed using the DESeq R package (1.10.1). Genes with an adjusted p-value <0.05 found by DESeq were assigned as differentially expressed. We normalized the number of reads for the differential gene expression from the RNAseq data [54]. Unigenes were annotated using data from the NCBI gene ontology (GO) and protein family (Pfam) databases. GO annotations were performed in Blast2GO v2.5 with a cutoff E-value of 1×10^{-6} [56].

3. Results

3.1. Identification, Classification, and Genomic Distribution of F-box Genes in J. regia

To identify the *F-box* genes from common walnut, an HMMER (biosequence analysis using profile hidden Markov models) search was performed against a local genome database using the seed files for the F-box (PF00646), F-box-like (PF12937), F-box-like 2 (PF13013), FBA (PF04300), FBA_1 (PF07734), FBA_2 (PF07735), FBA_3 (PF08268), and FBD (PF08387) domains. We acquired 111 candidate *F-box* genes from the common walnut whole genome. Those 111 candidate protein sequences were submitted to the SMART online software and Pfam online software to check the presence of the F-box domain. In total, in this study, 74 *F-box* genes were detected in common walnut (Table S2, Supplementary Materials). Based on the similar domain in which common walnut *F-box* genes were identified, the *F-box* genes were classified into eight subfamilies (Figure 1a). The most abundant *F-box* genes (34, 45.9%) were those with a similar known functional domain FBA_1, classified as the FBA_1

subfamily. The other 40 genes displayed the presence of one or more known functional domains and were classified as *F-box* subfamily (15, 20.3%) which contained the F-box domain, the F-box-like subfamily (14, 19.1%) which had F-box-like repeats, the FBD subfamily (3, 4.2%) which contained FBD domains, the FBOX subfamily (3, 4.2%) which had FBOX domains, the Kelch subfamily (3, 4.2%) which contained Kelch domains, and the Amn1 subfamily (2, 2.2%) which had AMN1 (Antagonist of mitotic exit network protein 1) domains (Figure 1b).

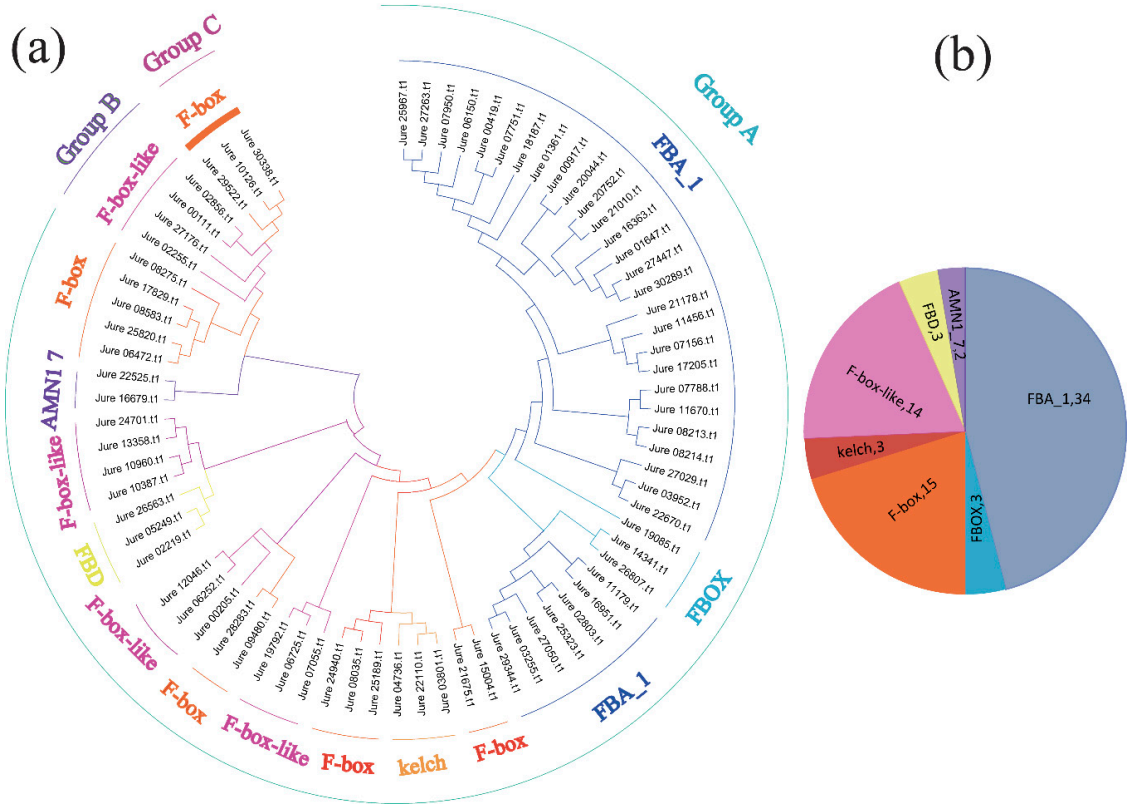

Figure 1. (**a**) Phylogenetic analysis of F-box proteins in common walnut (74). These 74 sequences were used to construct a neighbor-joining (NJ) tree. The tree was divided into eight subfamilies; the names of different groups are displayed. (**b**) The number of each group of the *F-box* gene family of common walnut.

We found a total of five gene pairs of *F-box* gene family members (*Jure_21010.t1* and *Jure_20752.t1*; *Jure_08213.t1* and *Jure_08214.t1*; *Jure_20752.t1* and *Jure_21010.t1*; *Jure_27447.t1* and *Jure_30289.t1*; *Jure_30289.t1* and *Jure_27447.t1*) (Table S3, Supplementary Materials). Interestingly, the Ka/Ks ratios of these five *F-box* sister pairs were less than 1, and divergence was estimated to be 100 Mya (Table S3, Figure S1, Supplementary Materials). The maximum number of *F-box* genes detected on pseudo-chromosome 34 (6) (Figure 2; Table S4, Supplementary Materials). The results showed that *F-box* genes were randomly distributed on the pseudo-chromosomes of common walnut. In all 74 *F-box* genes, one gene pair of *F-box* gene family members (*Jure_08214.t1* and *Jure_08213.t1*) exhibited tandem duplication and three gene pairs (*Jure_06252.t1* and *Jure_12046.t1*; *Jure_25820.t1* and *Jure_06472.t1*; *Jure_11670.t1* and *Jure_07788.t1*) exhibited segmental duplication (Figure 2).

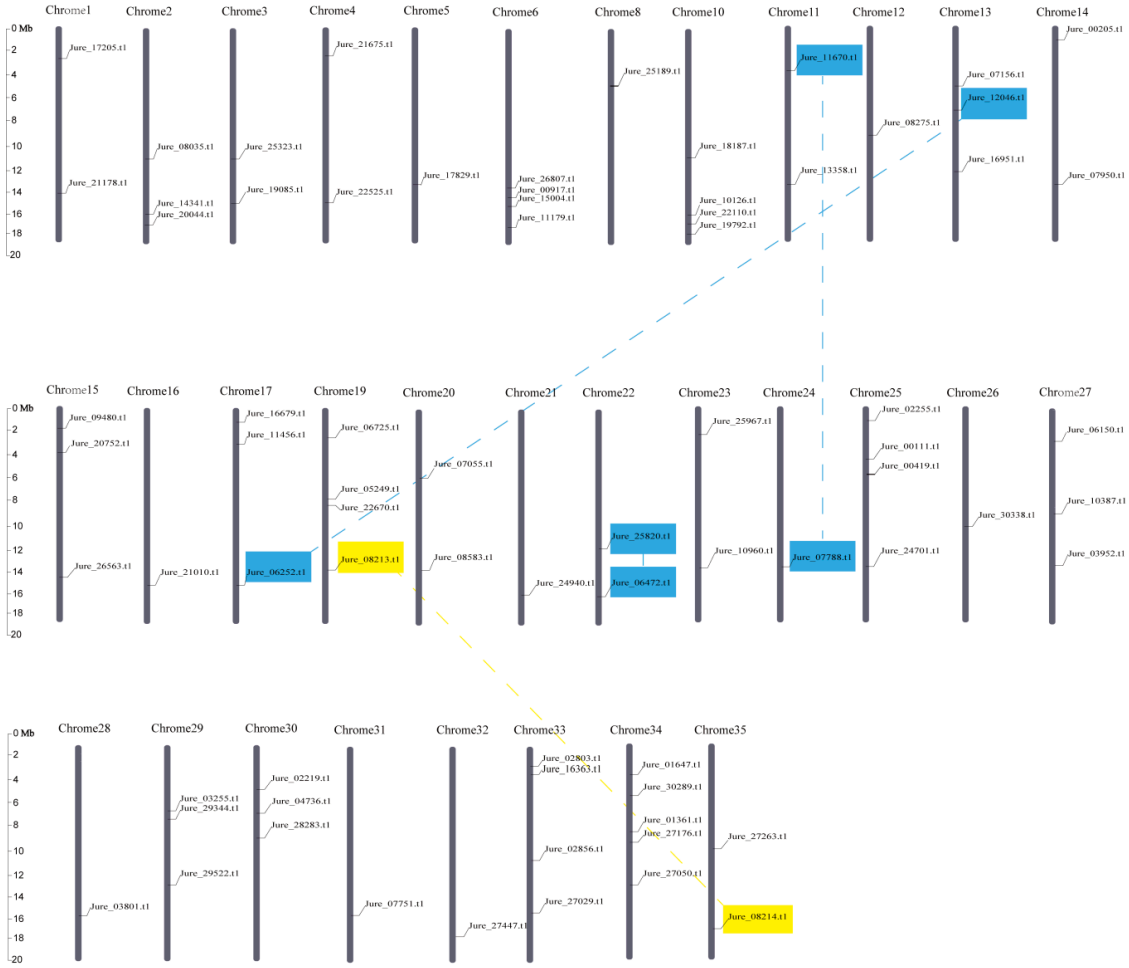

Figure 2. Chromosomal locations in common walnut of *F-box* genes on 31 pseudo-chromosomes. The pseudo-chromosome name is at the top of each bar. The scale of the pseudo-chromosome is millions of base pairs (Mb). The blue box represents the gene pairs of the *F-box* gene family of segmental duplication. The orange box represents the gene pairs of the *F-box* gene family of tandem duplication.

3.2. Phylogenetic Relationship, Gene Structure, and Protein Domains of the F-box Gene Family of Common Walnut

Based on the completed alignment of the sequences of *F-box* genes, they were clustered into three groups (Figure 1). To compare the structural components of 74 *F-box* genes, their structures including exons and introns were mapped (Figure 3c). The analysis of *F-box* gene structure revealed great variation among the genes. It was observed that the exon number varied from 1 to 23 in common walnut *F-box* transcripts. The maximum number of exons was found in *Jure 09480.t1* (23 exons), and the minimum number of one exon was found in 20 *F-box* genes. Regarding the remainder of genes, two exons were found in seven *F-box* members, three exons were found in four *F-box* members, four exons were found in two *F-box* members, and five exons in two *F-box* members (Figure 3). Using the available information of the Pfam database, the exact number and position of domains was identified for each common walnut protein (Figure 3b). Conserved domain and motif analysis showed that Motif 2 existed in all *F-box* genes in common walnut (Figure 3b). Motifs 2 and 1 were typical domains represented in the FBA_1 subfamily; Motif 17 was a typical motif for the FBOX subfamily; Motifs 2, 1, and 3 were typical for the F-box and the Kelch subfamily; Motif 12 was typical for the F-box-like subfamily; Motif 2 was typical for the FBD subfamily; and Motif 5 was typical for the AMN1 subfamily (Figure 3b). These results suggested that all *F-box* genes contained at least one typical

domain. The typical F-box domains, including FBOX and FBA_1, as well as the conserved motifs, are shown in Figure S1 (Supplementary Materials).

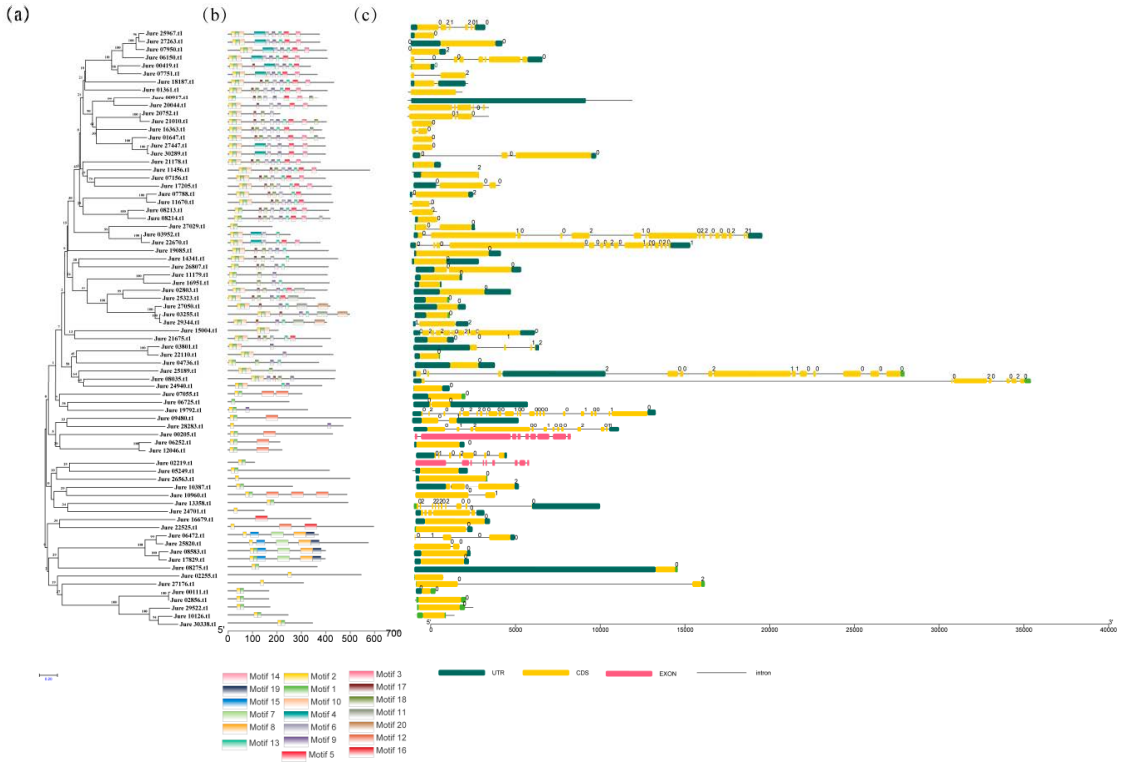

Figure 3. (**a**) Phylogenetic relationships, (**b**) motif compositions, and (**c**) gene structures of the 74 *F-box* genes identified in common walnut. (**a**) Phylogenetic relationships used the NJ method, and different colors represent different groups. (**b**) Colored boxes indicate conserved motifs, and gray lines represent non-conserved sequences. The length of motifs in each protein is shown proportionally. (**c**) UTR represents upstream, CDS represents coding sequence, and EXON represents exon; 0, 1, and 2 represent different types of phase. Phase 0: located between two consecutive codons; Phase 1: splitting codon between the first and second nucleotides; Phase 2: between the second and third nucleotides of a codon.

3.3. KEGG Analysis, and Cis-Acting Regulatory Elements in the Promoter of the F-box Gene Family of Common Walnut

KEGG enrichment for common walnut *F-box* genes was determined; a total of six terms were found, which included plant hormone signal transduction, genetic information processing, signal transduction, folding, sorting and degradation, environmental information processing, and 04120 ubiquitin-mediated proteolysis. It was obvious that plant hormone signal transduction plays a critical role in *F-box* genes in common walnut (Figure S4, Supplementary Materials). Otherwise, the *F-box* genes had a specific KEGG enrichment term, which was ubiquitin-mediated proteolysis (Figure S4, Supplementary Materials). To further clarify the gene function and transcriptional regulation mechanism of *F-box* genes, we amplified the 1500 bp upstream of the *F-box* genes and analyzed their cis-elements. We detected 23 types of cis-elements of the *F-box* gene family (Figure S3, Supplementary Materials). The most abundant cis-elements of *F-box* genes were part of a conserved DNA module array (CMA3), and the cis-elements could be classified into two main types, which contained light response elements and binding elements (Figure S3, Supplementary Materials).

3.4. Expression Profile Analysis of Walnut F-box Genes

To further reveal the expression patterns of *F-box* family genes in common walnut, public transcript data, transcript sequencing data, and digital expression profiles were investigated in common walnut

tissues including male and female flowers (Figure 4; Table S5, Supplementary Materials). For female and male flowers, we collected nine periods of female flowers and three periods of male flowers; the period from April to May was a key period of flower development for the flowering of common walnut, so we collected these tissues. Of course, we could notice that the members of the walnut *F-box* gene family were not much different within the period. The difference was that there was no obvious increase in the expression level with the developmental period. Based on phylogenetic analysis and expression level, we found a clear pattern that Group A had a higher expression in female and male flowers than other tissues; the genes which were classified into Group B were both lowly expressed in all tissues of common walnut; in addition, for the genes classified into Group C, the *F-box* gene family members (*Jure_29522.t1*, *Jure_10126.t1*, and *Jure_30338.t1*) were highly expressed in other tissues, including those of vegetative bud, embryo, somatic embryo, leaves, young leaf, immature hull, hull cortex, hull peel, hull dehiscing, and root (Figure 4). For the *F-box* gene family, we were most interested in those which played an important role during common walnut male and female flower development and ripening (Figure 4). Expression profile analysis showed that some *F-box* genes had a higher expression in multiple tissues. For instance, *Jure_07950.t1*, *Jure_20044.t1*, *Jure_07156.t1*, *Jure_11456.t1*, *Jure_06725.t1*, *Jure_21675.t1*, and *Jure_26807.t1* were highly expressed in male and female flowers, especially *Jure_05249.t1* and *Jure_02219.t1* which had a quite different expression in male and female flowers, *Jure_05249.t1* which had a high expression in female flowers, and *Jure_02219.t1* which had a low expression in male flowers, and we found that *Jure_05249.t1* had a special domain which was FBOX and *Jure_02219.t1* had a special domain which was F-box-like (Figure 4). During the female flower ripening, it was shown that some genes had an increased expression level with female growing, for example, *Jure_21178.t1*, *Jure_11456.t1*, *Jure_07156.t1*, *Jure_07788.t1*, *Jure_14341.t1*, *Jure_21675.t1*, *Jure_06725.t1*, *Jure_19792.t1*, *Jure_05249.t1*, *Jure_26563.t1*, *Jure_13358.t1*, and *Jure_16679.t1*. In male flowers, *Jure_10126.t1*, *Jure_30338.t1*, *Jure_26563.t1*, and *Jure_07156.t1* had a higher expression level on 10 April than in the other two periods. Overall, in the reproductive tissues (male and female flowers) and vegetative tissues (vegetative bud, embryo, somatic embryo, leaves, young leaf, immature hull, hull cortex, hull peel, hull dehiscing, and root), a total of 13 members (*Jure_07950.t1*, *Jure_21178.t1*, *Jure_07156.t1*, *Jure_07788.t1*, *Jure_14341.t1*, *Jure_21675.t1*, *Jure_06725.t1*, *Jure_19792.t1*, *Jure_06252.t1*, *Jure_02219.t1*, *Jure_05249.t1*, *Jure_26563.t1*) of the *F-box* gene family were highly expressed in reproductive tissues, while those in other vegetative tissues were low in expression, and three *F-box* genes (*Jure_29522.t1*, *Jure_10126.t1*, *Jure_30338.t1*) were higher in reproductive tissues than in nutritive tissues (Figure 4).

Based on the different domains of *F-box* subfamilies, including the FBA-1 subfamily, F-box subfamily-box-like subfamily, FBD subfamily, FBOX subfamily, Kelch subfamily, and AMN1 subfamily, in female stages, all *F-box* genes of common walnut had a high expression level compared to other stages, followed by those in male flower tissues, where the *F-box* genes in leaf and hull had a low expression level. Notably, different subfamilies had quite different expression levels; for example, the *F-box* genes which contained the FBA-1 domain had a high expression in the reproductive tissues compared to the nutritive tissues, indicating that the FBA-1 domain might play an important role in flowering. In the FBA-1 subfamily, we could notice that three genes (*Jure 21178.t1*, *Jure 11456.t1*, and *Jure 07156.t1*) had a high expression level in female and male flowers, with values even reaching 100 (especially high values), followed by six genes (*Jure 07950.t1*, *Jure 06150.t1*, *Jure 00419.t1*, *Jure 07751.t1*, *Jure 18187.t1*, and *Jure 07788.t1*) which had a high expression level contributing in the FBA-1 domains. We found that higher expression levels of three genes (*Jure 29522.t1*, *Jure 10126.t1*, and *Jure 30338.t1*) were detected in the nutritive tissues (leaf and hull) than in the reproductive tissues of common walnut. In the FBD, FBOX, Kelch, and AMN1 subfamilies, higher expression levels were detected in the reproductive tissues (male and female flowers) than in the nutritive tissues of common walnut.

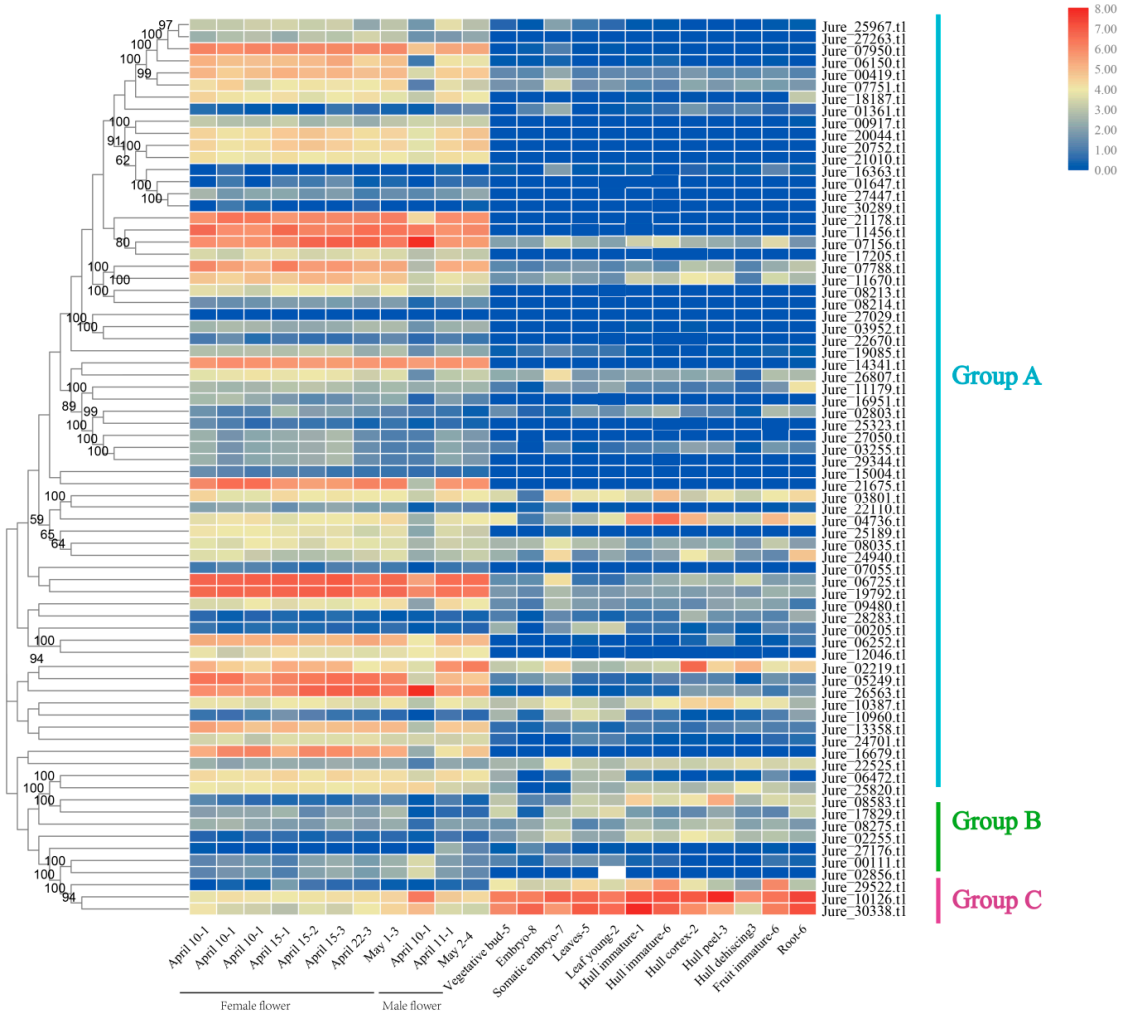

Figure 4. Expression patterns in multiple tissues of common walnut. RNA sequencing (RNA-Seq) data of all tissues in common walnut were used to analyze the expression pattern. The heat map was drawn in $\log_{10}$-transformed expression values. Red and green represent relatively high and low expression compared to the control, respectively.

In order to compare the differences between the *F-box* gene expression level in reproductive tissues and vegetative tissues, we selected male and female flowers, young leaves, and hull with biological replicates to create a clear expression pattern, as shown in Figure 5. Transcript expression analysis of different tissues of common walnut (male flower, female flower, young leaf, and hull) showed that the *F-box* gene family expression level was high in female flowers, and male flowers had a similar pattern, while the expression level was low in leaf and hull (Figure 5). However, the linear relative regression of the *F-box* gene family of common walnut was different in male flowers, female flowers, leaf, and hull (Figure 6). The results revealed that the expression of *F-box* genes in male flowers was positively correlated with female flowers in common walnut (Figure 6a). However, there was no correlation of expression levels between leaf and female flowers, hull and female flowers, leaf and male flowers, hull and male flowers, and vegetative tissues and germinal tissues (Figure 6). There was no significant positive or negative correlation between the expression level of the *F-box* gene in leaves and the expression level in female flowers; this was also the case in hull and female flowers, leaves and male flowers, and hull and male flowers. We averaged the expression values of vegetative and reproductive tissues at various stages and performed regression analysis, which did not show an obvious positive or negative correlation.

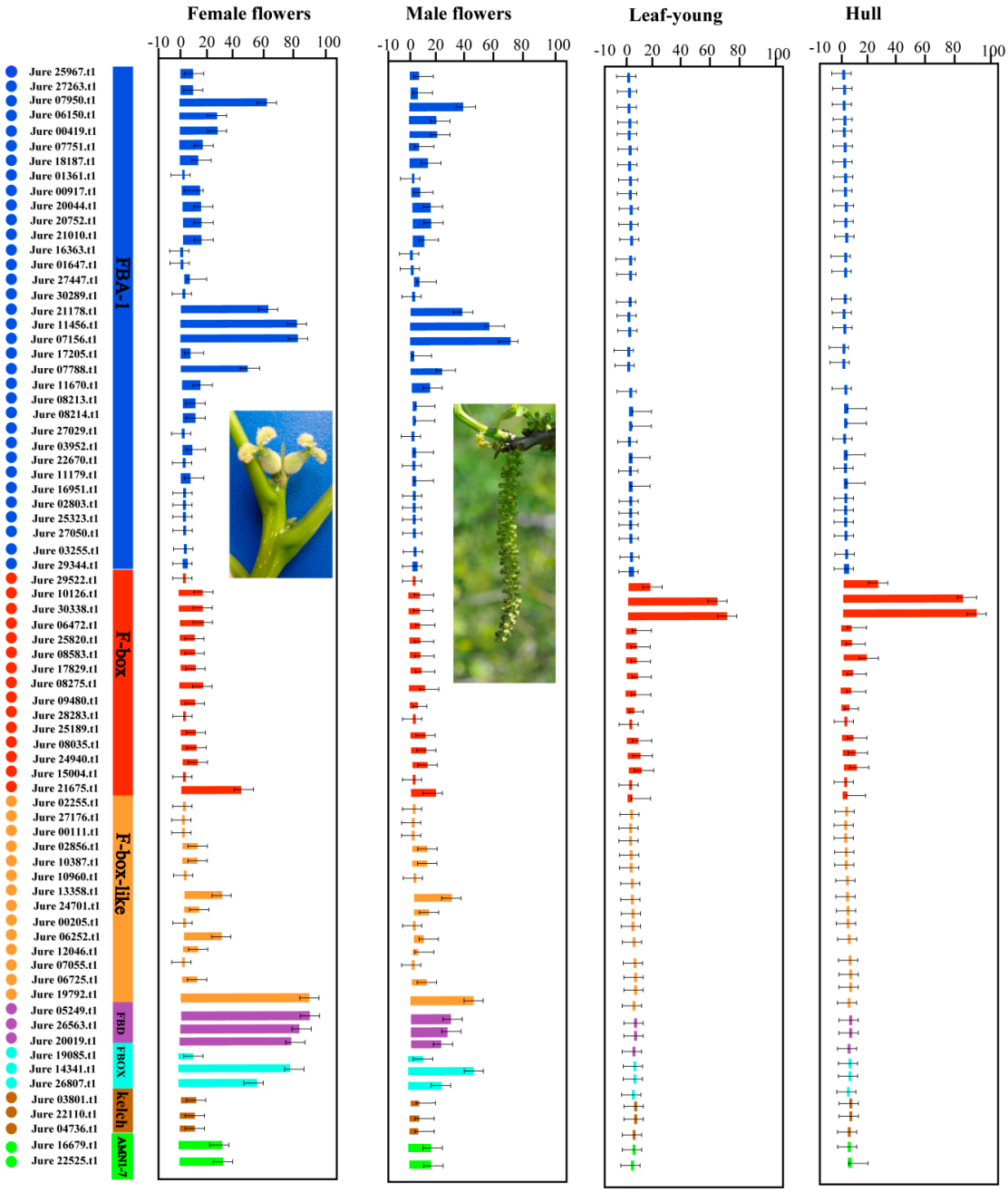

Figure 5. The values of *F-box* gene family member expression for different subfamilies of common walnut.

Jure_06725.t1, *Jure_07950.t1*, *Jure_19792.t1*, and *Jure_11456.t1* were highly expressed in female and male flowers during all stages, and *Jure_07788.t1*, *Jure_21178.t1*, *Jure_02219.t1*, *Jure_06252.t1*, *Jure_13358.t1*, *Jure_21675.t1*, *Jure_16679.t1*, and *Jure_05249.t1* were lowly expressed on 10 April in male flowers; *Jure_07156.t1* and *Jure_26563.t1* were especially high on 10 April in male flowers. We could notice that the *F-box* genes of common walnut did not have a similar trend to that with the plant growth; the *F-box* genes of common walnut had a rising expression (Figures 6 and 7). For male flowers, we found that two genes (*Jure_07156.t1*, *Jure_26563.t1*) had higher expressed levels on 10 April than other stages. For female flowers, some genes (*Jure_06725.t1*, *Jure 07950.t1*, *Jure_19792.t1*, *Jure_11456.t1*, *Jure_07788.t1*, *Jure_21178.t1*, *Jure_02219.t1*, *Jure_06252.t1*, *Jure_13358.t1*, *Jure_21675.t1*, *Jure_16679.t1*, and *Jure_05249.t1*) were expressed highly on 10 April (Figure 7).

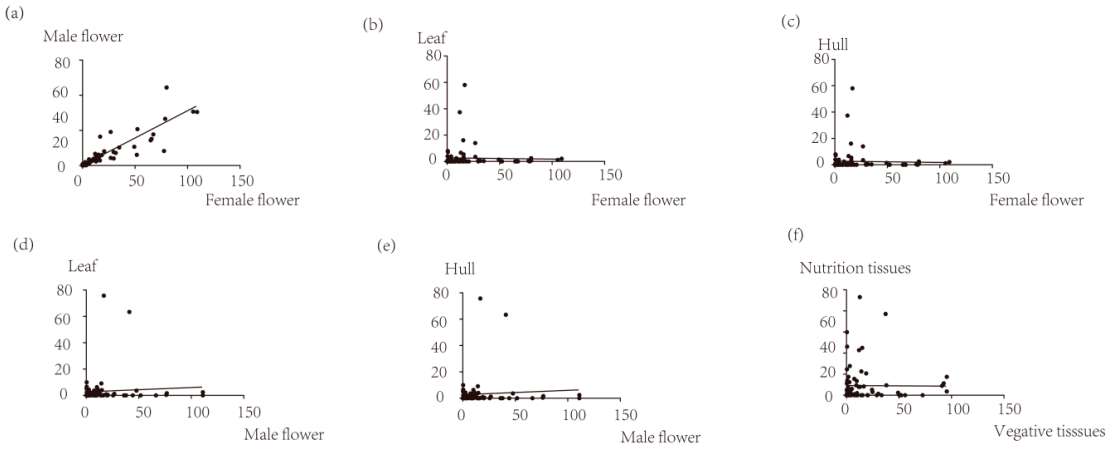

Figure 6. The regression analysis of the *F-box* gene family of common walnut.(**a**) the regression analysis of expression level between male and female flower; (**b**) the regression analysis of expression level between leaf and female flower; (**c**) the regression analysis of expression level between hull and female flower; (**d**) the regression analysis of expression level between leaf and male flower; (**e**) the regression analysis of expression level between hull and male flower; (**f**) the regression analysis of expression level between nutrition tissues and vegative tissues.

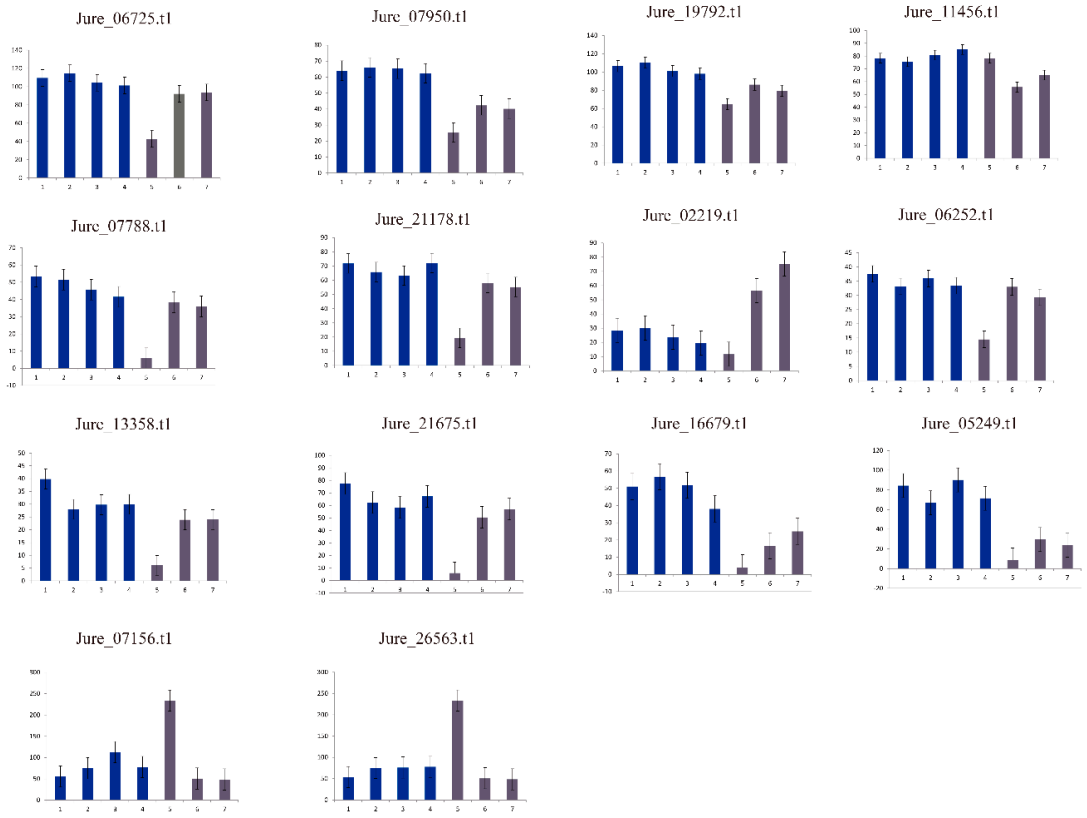

Figure 7. Expression of 14 *F-box* genes in various stages of female and male flowers. The *X*-axis represents female and male samples from different growth periods; 1, 2, 3, 4, 5, 6, and 7 represent 10 April, 15 April, 22 April, 1 May, 10 April, 11 April, and 2 May. The blue represents the expression level in female flowers, and the gray represents the expression level in male flowers. The *Y*-axis represents the relative expression level of *F-box* gene of common walnut. Error bars represent the standard deviations of three replicates. The expression level data are as shown in Table S5 (Supplementary Materials). We used two or three biological replicates from the same organization in the same period in Table S5 (Supplementary Materials).

4. Discussion

4.1. The Characteristics of the F-box Gene Family of Common Walnut

The present study identified 74 *F-box* genes from common walnut (Figure 1; Table S2, Supplementary Materials), compared to the higher number in plants such as *Arabidopsis* (672), *M. truncatula* (972), maize (359), and rice (687) [23,25,26,31]. A previous study claimed that the number of *F-box* genes was significantly less than that of herbaceous plants in perennial woody plants [30], due to the selection pressure of herbaceous plants in a single growth cycle, resulting in many *F-box* genes in herbaceous plants [30]. In this study, the number of *F-box* genes in common walnut was less than other studies identified from herbaceous annual plants, which was consistent with another previous study [30] (Figure 1).

The genomes of plants are different; for example, the maize genome is about 2300 Mb, the *Arabidopsis* genome is about 125 Mb, and the rice genome is about 389 Mb. However, the common walnut genome is 667 Mb [36,57–61]. These plants contain *F-box* genes—672 in Arabidopsis, 359 in maize, 687 in rice, and 74 in common walnut. Comparative analysis suggested that maize, *Arabidopsis*, and rice had similar numbers of gene families [60,61]. In terms of the percentage of each plant, the common walnut was 1.91%, *Arabidopsis* was 2.71%, maize was 2.03%, and rice was 2.21%. We noticed that the percentage was not consistent with the genome size; thus, we could speculate that the genome size was not the main reason determining the number of gene families [62] (Figure 1).

4.2. The Expansion and Evolution of the F-box Gene Family of Common Walnut

Gene families were most likely to choose tandem duplication and segmental duplication as forms of expansion [62]. The duplication events were also a prospective which could explain the expansion of the *F-box* gene family of common walnut (Figure 2). *M. truncatula*'s expansion was probably due to localized gene duplications [30], tandem duplications mainly existed in chickpea [28], whole-genome duplication and dispersed duplication occurred in pear [29], and tandem and segmental duplication events were responsible for the rise in the number of *F-box* genes of maize [26]. In this study, the results showed that tandem duplications and segmental duplications were low-rate events in common walnut (Figure 2). Additionally, the Ka/Ks ratios of these five *F-box* sister pairs were less than 1 and divergence was estimated to be 100 Mya (Table S1, Figure S5, Supplementary Materials), but the *Juglans*' divergence time was 40 Mya [62], indicating that the *F-box* gene family was an original gene family which played a vital role in flowering. We found that, among the *F-box* gene family members, only one gene pair experienced tandem duplication events, and three gene pairs of F-box gene family members went through segmental events. These results indicated that the expansion of the *F-box* gene family of common walnut might be due to whole-genome duplication (WGD) events (Figure 2; Table S3, Supplementary Materials).

The *F-box* genes of common walnut were clustered into three main groups based on phylogenetic analysis (Figure 1). By contrast, based on the specific and special domains, a total of five subfamilies were identified in *Arabidopsis* [23], nine clusters were found in chickpea [28], three subfamilies were found in *M. truncatula* [30], 12 subfamilies were found in pear [29], 10 subfamilies were found in rice [25], and 15 subfamilies were found in maize [26]. In this study, all F-box proteins in common walnut were classified into eight subfamilies (Figure 1). F-box proteins with the same or a similar domain organization were usually grouped together in the same branches (Figures 1 and 3b), indicating that these genes might have the same function, Motif 2 existed in the entire *F-box* gene family of common walnut, and it was proven that the *F-box* gene family was conservative in evolution (Figure 3a).

Gene structure is an important factor for gene family evolution [29,43]. Intron positions always determine the gene coding sequence position and protein structure, and increased protein diversity through exon shuffling and alternative splicing causes an evolutionary advantage [63–65]. The association of the intron phase with conservation at splice site sequences is related to the evolution

of spliceosomal introns [66,67]. Intron phase 0 indicates the highest conservation, intron Phase 1 also shows a high conservation, and intron Phase 2 shows the lowest conservation [68,69]. In the case of the *F-box* genes in common walnut, almost all gene structures exhibited intron phase 0; thus, the *F-box* genes were conservative in evolution (Figure 3c). In conclusion, both analyses including motif and gene structure showed that *F-box* genes were conservative in evolution.

4.3. Function of the F-box Genes

Many F-box proteins were found to contain various protein–protein interaction domains at their C terminus [23]. In *Arabidopsis*, there are 16 types (including one unknown type) of domains [23], and at least 15 types of specific domains were examined in rice [25]. Fifteen types of domains were identified from *M. truncatula* [31] (Figure 1b). In *Arabidopsis* and rice, the number of some domains, such as FBA, LRR (Leucine-rich repeat protein), DUF (Protein of unknown function, DUF), and Kelch, is larger than that of other domains, such as WD40 (WD domain, G-beta repeat), JmjC, and LysM. In *M. truncatula*, four domains that were predicted in the F-box protein seem to not be species-specific in their conservation, including PP2 (Phloem protein 2), PAS (Signal transduction mechanisms)-Kelch, GSH-synth (C-terminal, alpha helical domain of the Glutathione S-transferase family) ATP, and ARM (armadillo repeat). In this study, 25 types of F-box proteins of common walnut were detected. We could speculate that it was precisely because the F-box protein family possessed such a large number of conserved domains that the *F-box* genes were relatively conservative in evolution (Figure S2, Supplementary Materials). In *M. truncatula*, the FBA domain (165) was the most abundant [31]. The most abundant domains in the presence of ZmFBX proteins (37) were Kelch repeats [26]. Kelch repeats exist in a large fraction of rice and *Arabidopsis* F-box proteins and appear to be the unique feature of plant F-box proteins [31]. In this study, the FBA-1 domain was the most abundant compared to other domains, which is consistent with previous studies [23,30].

A large number Kelch domains exist in *Arabidopsis* and rice, and they appear more likely to be the unique feature of plant F-box proteins [31]. In this study, three F-box proteins (*Jure 04736.t1*, *Jure 22110.t1*, and *Jure 03801.t1*) were identified with the Kelch domain at their C-terminal regions. In the phloem sap, there was a large number of phloem protein 2 (PP2), indicating that PP2 plays a crucial role in the phloem [68] (Figure 1). In *Arabidopsis*, there were 18 F-box proteins that had C-terminal homology to the squash lectin PP2, while *Jure 24701.t1* may detect glycosylated substrates in common walnut. Following a comparison of walnuts and *Arabidopsis*, it was found that *Arabidopsis* (18 F-box proteins) contained PP2, but, in common walnut, Jure 24701.t1 might detect glycosylated substrates (Figure S1, Supplementary Materials). Regarding FBOX, a receptor for ubiquitination targets [69], 12 F-box proteins of common walnut (*Jure_14341.t1*, *Jure_06150.t1*, *Jure_20752.t1*, *Jure_08035.t1*, *Jure_26807.t1*, *Jure_29344.t1*, *Jure_07950.t1*, *Jure_30289.t1*, *Jure_27447.t1*, *Jure_25967.t1*, *Jure_21010.t1*, and *Jure_27263.t1*) had the FBOX domain in this study, indicating that these genes exhibited a ubiquitination function (Figure 1). Importantly, a total of three genes contained the F-box domains (*Jure 29522.t1*, *Jure 10126.t1*, and *Jure 30338.t1*) and had a higher expression level compared to the genes contained in other domains (Figure 6). AMN1 (antagonist of mitotic exit network protein 1) is a leucine-rich repeat (LRR) protein and a negative regulator of the signal transduction pathway MEN (mitotic exit network), and overexpression of AMN1 slows the growth of wild-type cells [70]. A total of five genes (*Jure_10960.t1*, *Jure_12046.t1*, *Jure_22525.t1*, *Jure_07055.t1*, and *Jure_06252.t1*) had this domain and they might have a function connected to the right-handed beta helix region, which contains a parallel beta helix region that shares some similarity with pectate lyases (Figure 6).

Based on KEGG analysis, we could notice that the major pathway for the *F-box* genes of common walnut was plant hormone signal transduction, which was useful information (Figure S4, Supplementary Materials). These F-box proteins containing similar cis-elements might perform similar functions in rice and *Arabidopsis*. Taken together, these results suggested that the expression of *F-box* protein-coding genes was regulated by light, indicating light control of protein

turnover in rice, which might be critical for several light-dependent cellular processes (Figure S3, Supplementary Materials). In this study, a comprehensive analysis was performed to investigate the expression patterns of *F-box* gene family genes in different common walnut tissues, including vegetative bud, embryo, somatic embryo, leaves, young leaf, immature hull, hull cortex, hull peel, hull dehiscing, and root (Figure 4). Interestingly, some specific gene clusters presenting high expression were found in different tissues, suggesting their close relationship with the development of the specific tissue. The expression pattern of *F-box* genes of Fabaceae was complex; the *F-box* genes were expressed differently in different tissues: in flowers, some genes were expressed highly, but the other genes exhibited low expression [27]. *F-box* genes showed an increasing trend with pollen tube development, implying that these genes of *F-box* genes of pear played potential roles in pollen tube elongation [29]. A clear pattern in rice was that only a few genes were highly expressed in root, leaf, young leaf, and seeding. The same situation is present in common walnut (Figure 4). Apparently, Group A had more *F-box* genes in the whole *F-box* gene family of common walnut (Figure 1), while a clear expression pattern was that the Group A genes had a higher expression level than other groups; this showed that more *F-box* genes of common walnut had a higher expression level in male and female flowers, implying that the F-box genes might play a critical role in flowering (Figure 4) [11,12]. During the critical development period of walnuts, from April to May, there was no increase in the expression of the *F-box* gene family between walnuts, regardless of the period of female flower or male flower. In the three periods, only different genes had different expression levels, but they were not different at different times. Therefore, we speculate that this is because the walnut *F-box* gene family played an important role in the flowering and development process, so the key genes were expressed throughout the flowering period, and there was no difference in the developmental stage, consistent with previously reported research results (Figure 4). In chickpea, there was a clear expression pattern, which was that different group genes of the *F-box* gene family members had a special expression level in different tissues, with a group of genes showing a higher expression level in flower buds than other tissues, and several common walnut *F-box* genes showing tissue-specific expression profiles, exhibiting a high similarity with well-documented *F-box* genes in *Arabidopsis* and chickpea [23,28]. In the FBA-1 subfamily, we could notice that three genes (*Jure 21178.t1*, *Jure 11456.t1*, and *Jure 07156.t1*) had a high expression level in female and male flowers, with values even reaching 100 (especially high values), followed by six genes (*Jure 07950.t1*, *Jure 06150.t1*, *Jure 00419.t1*, *Jure 07751.t1*, *Jure 18187.t1*, and *Jure 07788.t1*) having a high expression level, contributing in the FBA-1 domains; thus, we could speculate that the FBA-1 domain might be an important factor in the formation stages. These proteins contain pfam00646 at the N terminus, suggesting that they are effectors linked with ubiquitination (Figure 5). However, in the F-box subfamily, we found that higher expression levels of three genes (*Jure 29522.t1*, *Jure 10126.t1*, and *Jure 30338.t1*) were detected in the vegetative tissues (leaf and hull) than in the germinal tissues of common walnut (Figure 5). These genes (*Jure_07950.t1*, *Jure_21178.t1*, *Jure_11456.t1*, *Jure_07156.t1*, *Jure_06725.t1*, *Jure_19792.t1*, *Jure_05249.t1*, and *Jure_26563.t1*) were highly expressed in female and male flowers, showing a clear expression level in male and female flowers (Figure 6). These genes (*Jure_06725.t1*, *Jure_07950.t1*, *Jure_19792.t1*, and *Jure_11456.t1*) were highly expressed in the female and male flower stages at all times, indicating that these genes might have a vital role in flowering [11,12] (Figure 7). Based on the expression analysis, regression analysis, and the values for the *F-box* gene family of common walnut, it was found that the *F-box* genes might play a vital role in flowering (Figures 5–7). Some previous studies showed that *F-box* genes and mutants play an important role in the development of pollen tubes in woody plants. In apples, the *F-box* gene was named *SFBB*, which is specifically expressed in pollen tubes. In Rosaceae plants, such as sweet cherry and Japanese, the *F-box* gene was named *SFB* and plays an important role in the process leading to self-incompatibility. *SLF* in *Prunus mume* was also confirmed to be specifically expressed in pollen tubes by cloning the *F-box* gene. Alleles *SFB1*, *SFB2*, *SFB4*, and *SFB5* in sweet cherry proved to play a self-incompatible effect by affecting the development of pollen tubes. Allele S4sm, lacking the SFBB1-S4 protein, caused no significant effect on pollen tubes, while the S5 haplotype encoding the

SFBB1 protein could cause pistils to accept other genes [29,35–40]; thus, there is strong evidence that the *F-box* gene is relevant in pollen development. In this study, the results showed that some genes (*Jure 07950.t1*, *Jure 21178.t1*, *Jure 11456.t1*, *Jure 07156.t1*, *Jure 07788.t1*, *Jure 21675.t1*, *Jure 19792.t1*, *Jure 05249.t1*, *Jure 26563.t1*, *Jure 20019.t1*, *Jure 14341.t1*, and *Jure 26807.t1*) were expressed highly in female flowers, suggesting that the *F-box* gene family of common walnut was vital for regulating the flowering (Figures 4–6). We found that a total of 12 genes were highly expressed in the first opening of female flowers (10 April), while the two genes were highly expressed in the first opening of male flowers, suggesting that the *F-box* genes were relevant in female and male flowering in common walnut (Figure 7).

5. Conclusions

In this study, we identified 74 *F-box* genes in common walnut (*Juglans regia*). Phylogenetic analysis showed that the *F-box* genes can be grouped into three groups. Based on specific and special domains, the common walnut *F-box* genes were classified into eight subfamilies. In addition, expression profile analysis revealed that the *F-box* genes display diverse expression patterns in different common walnut tissues. Most of the common walnut *F-box* genes grouped into Group A were highly expressed in female and male flowers, while Group C members were lowly expressed in leaf and hull tissues. A total of 14 genes (*Jure_06725.t1*, *Jure_07950.t1*, *Jure_19792.t1*, *Jure_11456.t1*, *Jure_07788.t1*, *Jure_21178.t1*, *Jure_02219.t1*, *Jure_06252.t1*, *Jure_13358.t1*, *Jure_21675.t1*, *Jure_16679.t1*, *Jure_05249.t1*, *Jure_07156.t1*, and *Jure_26563.t1*) were highly expressed in female and male flowers at all stages, which might play an important role in the flowering of common walnut. During female flower ripening, it was shown that some genes exhibited an increased expression level with the female growing. Overall, in the reproductive tissues and vegetative tissues, a total of 13 members of the *F-box* gene family were highly expressed in reproductive tissues, while those in other vegetative tissues were low in expression, and three *F-box* genes were higher in reproductive tissues than in nutritive tissues. The expressions were significantly higher in female flowers and male flowers compared with leaf and hull tissues at the transcription level. The results revealed that the expressions of *F-box* genes in male flowers were positively correlated with female flowers, but there was no correlation between any other tissue combinations in common walnut.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4907/10/3/275/s1: Figure S1. The whole conserved domains of the *F-box* gene family of common walnut; Figure S2. The maximum estimation for KS (synonymous replacement) values; Figure S3. The cis-elements of the *F-box* gene family of common walnut; Figure S4. The KEGG pathway of the *F-box* gene family of common walnut; Table S1. A total of 12 samples of common walnut used for expression profiling in this study; Table S2. The information and accession number of the *F-box* gene family of common walnut; Table S3. Pairwise identities and divergence period estimation of common tobacco *F-box* genes; Table S4. Information for a total of 40 scaffolds (pseudo-chromosomes) for common walnut (*J. regia*); Table S5. Expression data of *F-box* genes in flowers at different developmental stages in common walnut (*J. regia*).

Author Contributions: Conceptualization, P.Z., F.Y., and H.L.; experiment design, P.Z. and F.Y.; data curation, P.Z., H.Z., M.Y., G.Y., H.L., and F.Y.; analysis, P.Z., H.L., G.Y., and F.Y.; funding acquisition, P.Z.; project administration, P.Z.; contribution of materials/analysis tools, P.Z., H.Z., M.Y., H.L., and F.Y.; writing—original draft, P.Z., H.Z., G.Y., M.Y., and F.Y.; writing—review and editing, P.Z. and F.Y.; review and editing, S.Z.

Funding: This work was funded by the National Natural Science Foundation of China (No. 41471038; No. 31200500), the Program for Excellent Young Academic Backbones funded by the Northwest University, Shaanxi Academy of Science Research Funding Project, and the Natural Science Basic Foundation of Shaanxi Province (2018JM3021).

Acknowledgments: The authors wish to thank Yiheng Hu, Xiaojia Feng, Yiwei Sun and Dang Meng for sample collection.

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

References

1. Kumar, A.; Paietta, J.V. The sulfur controller-2 negative regulatory gene of Neurospora crassa encodes a protein with beta-transducin repeats. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3343–3347. [CrossRef] [PubMed]
2. Kipreos, E.T.; Pagano, M. The F-box protein family. *Genome Biol.* **2000**, *1*, 1–7. [CrossRef] [PubMed]
3. Jin, J.; Ang, X.L.; Shirogane, T.; Harper, J.W. Identification of substrates for F-box proteins. *Method Enzym.* **2005**, *399*, 287–309.
4. Xu, G.X.; Hong, M.; Masatoshi, N.; Hongzhi, K. Evolution of *F-box* genes in plants: Different modes of sequence divergence and their relationships with functional diversification. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 835–840. [CrossRef] [PubMed]
5. Stefanowicz, K.; Lannoo, N.; Zhao, Y.; Eggermont, L.; Hove, J.V.; Atalah, B.A.; Damme, E.J.M.V. Glycan-binding F-box protein from *Arabidopsis thaliana* protects plants from pseudomonas syringae infection. *BMC Plant Biol.* **2016**, *16*, 213. [CrossRef] [PubMed]
6. Baute, J.; Polyn, S.; De, B.J.; Blomme, J.; Van, L.M.; Inzé, D. F-box protein FBX92 affects leaf size in Arabidopsis thaliana. *Plant Cell Physiol.* **2017**, *58*, 962–975. [CrossRef] [PubMed]
7. Zhao, Z.; Zhang, G.; Zhou, S.; Ren, Y.; Wang, W. The improvement of salt tolerance in transgenic tobacco by overexpression of wheat *F-box* gene *TaFBA*1. *Plant Sci.* **2017**, *259*, 71–85. [CrossRef] [PubMed]
8. Xue, F.; Cooley, L. Kelch encodes a component of intercellular bridges in drosophila egg chambers. *Cell* **1993**, *72*, 681–693. [CrossRef]
9. Woo, H.R.; Chung, K.M.; Park, J.H.; Oh, S.A.; Ahn, T.; Hong, S.H.; Jang, S.K.; Nam, H.G. ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. *Plant Cell* **2001**, *13*, 1779–1790. [CrossRef]
10. Petra, S.; Furner, I.J.; Ottoline Leyser, H.M. Max2 participates in an SCF complex which acts locally at the node to suppress shoot branching. *Plant J.* **2007**, *50*, 80–94.
11. Durfee, T.; Roe, J.L.; Sessions, R.A.; Inouye, C.; Serikawa, K.; Feldmann, K.A.; Weigel, D.; Zambryski, P.C. The F-box-containing protein ufo and agamous participate in antagonistic pathways governing early petal development in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8571–8576. [CrossRef] [PubMed]
12. Takato, I.; Schultz, T.F.; Harmon, F.G.; Ho, L.A.; Kay, S.A. FKF1 F-box protein mediates cyclic degradation of a repressor of constans in Arabidopsis. *Science* **2005**, *309*, 293–297.
13. Eva, K.; Péter, G.; Anthony, H.; László, K.B.; Woe-Yeon, K.; Eriksson, M.E.; Réka, T.; Shigeru, H.; Balázs, F.; Southern, M.M. Forward genetic analysis of the circadian clock separates the multiple functions of zeitlupe. *Plant Physiol.* **2006**, *140*, 933–945.
14. Han, L.Q.; Mason, M.E.; Risseeuw, E.P.; Crosby, W.L.; Somers, D.E. Formation of an SCF(ZTL) complex is required for proper regulation of circadian timing. *Plant J.* **2010**, *40*, 291–301. [CrossRef]
15. Paja, S.; Xi, W.; Skirpan, A.L.; Yan, W.; Dowd, P.E.; Mccubbin, A.G.; Shihshieh, H.; Teh-Hui, K. Identification of the pollen determinant of s-RNase-mediated self-incompatibility. *Nature* **2004**, *429*, 302–305.
16. Hong, Q.; Fei, W.; Lan, Z.; Zhou, J.; Zhao, L.; Zhang, Y.; Robbins, T.P.; Xue, Y. The F-box protein AhSLF-S2 controls the pollen function of S-RNase-based self-incompatibility. *Plant Cell* **2004**, *16*, 2307–2322.
17. Wang, L.; Dong, L.; Zhang, Y.E.; Zhang, Y.; Wu, W.; Deng, X.; Xue, Y. Genome-wide analysis of s-locus *F-box-like* genes in *Arabidopsis thaliana*. *Plant Mol. Biol.* **2004**, *56*, 929–945. [CrossRef]
18. Dieterle, M.; Zhou, Y.C.; SchäFer, E.; Funk, M.; Kretsch, T. Eid1, an F-box protein involved in phytochrome a-specific light signaling. *Genes Dev.* **2001**, *15*, 939–944. [CrossRef]
19. Alyssa, D.; Thomas, S.G.; Jianhong, H.; Steber, C.M.; Tai-Ping, S. The Arabidopsis F-box protein sleepy1 targets gibberellin signaling repressors for gibberellin-induced degradation. *Plant Cell* **2004**, *16*, 1392–1405.
20. Binder, B.M.; Walker, J.M.; Gagne, J.M.; Emborg, T.J.; Georg, H.; Bleecker, A.B.; Vierstra, R.D. The arabidopsis EIN3 binding F-box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. *Plant Cell* **2007**, *19*, 509–523. [CrossRef]
21. Calderón-Villalobos, L.I.A.; Nill, C.; Marrocco, K.; Kretsch, T.; Schwechheimer, C. The evolutionarily conserved arabidopsis thaliana F-box protein *AtFBP7* is required for efficient translation during temperature stress. *Gene* **2007**, *392*, 106–116. [CrossRef] [PubMed]
22. Suk, K.H.; Delaney, T.P. Arabidopsis SON1 is an F-box protein that regulates a novel induced defense response independent of both salicylic acid and systemic acquired resistance. *Plant Cell* **2002**, *14*, 1469–1482.

23. Gagne, J.M.; Downes, B.P.; Shin-Han, S.; Durski, A.M.; Vierstra, R.D. The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11519–11524. [CrossRef]
24. Cui, H.R.; Zhang, Z.R.; Lv, W.; Xu, J.N.; Wang, X.Y. Genome-wide characterization and analysis of F-box protein-encoding genes in the *Malus domestica* genome. *Mol. Genet. Genom.* **2015**, *290*, 1435–1446. [CrossRef] [PubMed]
25. Mukesh, J.; Aashima, N.; Rita, A.; Pinky, A.; Swatismita, R.; Pooja, S.; Sanjay, K.; Tyagi, A.K.; Khurana, J.P. F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* **2007**, *143*, 1467–1483.
26. Jia, F.J.; Wu, B.J.; Li, H.; Huang, J.G.; Zheng, C.C. Genome-wide identification and characterisation of *F-box* family in maize. *Mol. Genet. Genom.* **2013**, *288*, 559–577. [CrossRef] [PubMed]
27. Jia, Q.; Xiao, Z.X.; Wong, F.L.; Sun, S.; Liang, K.J.; Lam, H.M. Genome-wide analyses of the soybean *F-box* gene family in response to salt stress. *Int. J. Mol. Sci.* **2017**, *18*, 818. [CrossRef]
28. Gupta, S.; Garg, V.; Kant, C.; Bhatia, S. Genome-wide survey and expression analysis of *F-box* genes in chickpea. *BMC Genom.* **2015**, *16*, 67. [CrossRef]
29. Wang, G.M.; Yin, H.; Qiao, X.; Tan, X.; Gu, C.; Wang, B.H.; Cheng, R.; Wang, Y.Z.; Zhang, S.L. F-box genes: Genome-wide expansion, evolution and their contribution to pollen growth in pear (*pyrus bretschneideri*). *Plant Sci.* **2016**, *253*, 164–175. [CrossRef]
30. Yang, X.H.; Kalluri, U.C.; Sara, J.; Gunter, L.E.; Tongming, Y.; Tschaplinski, T.J.; Weston, D.J.; Priya, R.; Tuskan, G.A. The *F-box* gene family is expanded in herbaceous annual plants relative to woody perennial plants. *Plant Physiol.* **2008**, *148*, 1189–1200. [CrossRef]
31. Jian, B.S.; Yan, X.W.; Hai, B.L.; Bo, W.L.; Zhao, S.Z.; Shuai, G.; Zhi, M.Y. The *F-box* family genes as key elements in response to salt, heavy mental, and drought stresses in *Medicago truncatula*. *Funct. Integr. Genom.* **2015**, *15*, 495–507.
32. Hidenori, S.; Hiroyuki, K.; Mayu, M.; Yusuke, S.; Toshio, H.; Koichiro, U.; Makoto, K.; Hisashi, H.; Takato, K.J.G. S locus F-box brothers: Multiple and pollen-specific *F-box* genes with S haplotype-specific polymorphisms in apple and Japanese pear. *Genetics* **2007**, *175*, 1869–1881.
33. Ushijima, K.; Yamane, H.; Watari, A.; Kakehi, E.; Ikeda, K.; Hauck, N.R.; Iezzoni, A.F.; Tao, R.J.P.J. The S haplotype-specific F-box protein gene, *SFB*, is defective in self-compatible haplotypes of *Prunus avium* and *P. mume*. *Plant J.* **2010**, *39*, 573–586. [CrossRef] [PubMed]
34. Entani, T.; Iwano, M.; Shiba, H.; Che, F.S.; Isogai, A.; Takayama, S.J.G.t.C. Comparative analysis of the self-incompatibility (S-) locus region of Prunus mume: Identification of a pollen-expressed *F-box* gene with allelic diversity. *Genes Cells* **2010**, *8*, 203–213. [CrossRef]
35. Ikeda, K.; Igic, B.; Ushijima, K.; Yamane, H.; Hauck, N.R.; Nakano, R.; Sassa, H.; Iezzoni, A.F.; Kohn, J.R.; Tao, R.J.S.P.R. Primary structural features of the S haplotype-specific F-box protein, *SFB*, in Prunus. *Sex. Plant Reprod.* **2004**, *16*, 235–243. [CrossRef]
36. Hiroyuki, K.; Masaki, K.; Koichiro, U.; Miyoko, K.; Shu, K.; Hidenori, S.J.P.J. Sequence divergence and loss-of-function phenotypes of S locus F-box brothers genes are consistent with non-self recognition by multiple pollen determinants in self-incompatibility of Japanese pear (*Pyrus pyrifolia*). *Plant J.* **2011**, *68*, 1028–1038.
37. Han, H.; Woeste, K.E.; Hu, Y.; Meng, D.; Tian, Z.; Gao, X.X.; Zhou, H.; Feng, X.; Zhao, G.; Peng, Z. Genetic diversity and population structure of common walnut (*Juglans regia*) in china based on est-ssrs and the nuclear gene phenylalanine ammonia-lyase (*PAL*). *Tree Genet. Genomes* **2016**, *12*, 111. [CrossRef]
38. Hu, Y.; Woeste, K.E.; Zhao, P. Completion of the chloroplast genomes of five chinese juglans and their contribution to chloroplast phylogeny. *Front. Plant Sci.* **2016**, *7*, 1955. [CrossRef]
39. Zhao, P.; Zhou, H.J.; Potter, D.; Hu, Y.H.; Feng, X.J.; Dang, M.; Feng, L.; Zulfiqar, S.; Liu, W.Z.; Zhao, G. Population genetics, phylogenomics and hybrid speciation of juglans in china determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). *Mol. Phylogenet. Evol.* **2018**, *126*, 250–265. [CrossRef]
40. Feng, X.; Yuan, X.; Sun, Y.; Hu, Y.; Zulfiqar, S.; Ouyang, X.; Meng, D.; Zhou, H.; Woeste, K.; Peng, Z. Resources for studies of iron walnut (*Juglans sigillata*) gene expression, genetic diversity, and evolution. *Tree Genet. Genomes* **2018**, *14*, 51. [CrossRef]

41. Martínez-García, P.J.; Crepeau, M.W.; Puiu, D.; Gonzalez-Ibeas, D.; Whalen, J.; Stevens, K.A.; Paul, R.; Butterfield, T.S.; Britton, M.T.; Reagan, R.L. The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of nonstructural polyphenols. *Plant J.* **2016**, *87*, 507–532. [CrossRef]
42. Kumar, S.; Nei, M.; Dudley, J.; Tamura, K. Mega: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* **2008**, *9*, 299–306. [CrossRef]
43. Yan, F.; Li, H.Z.; Zhao, P. Genome-Wide Identification and transcriptional expression of the *PAL* Gene family in common Walnut (*Juglans regia* L.). *Genes* **2019**, *10*, 46. [CrossRef]
44. Magali, L.; Patrice, D.; Gert, T.; Kathleen, M.; Yves, M.; Yves, V.D.P.; Pierre, R.; Stephane, R. 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*, 325–327.
45. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The meme suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49. [CrossRef]
46. Wang, Y.; Tang, H.; DeBarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.-h.; Jin, H.; Marler, B.; Guo, H.; et al. Mcscanx: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef]
47. Rozas, J.; Librado, P. Dnasp v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452.
48. Doerks, T.; Copley, R.R.; Schultz, J.; Ponting, C.P.; Bork, P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res.* **2002**, *12*, 47–56. [CrossRef]
49. Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science* **2000**, *290*, 1151–1155. [CrossRef]
50. Deng, W.K.; Wang, Y.B.; Liu, Z.X.; Chen, H.; Xue, Y. Hemi: A toolkit for illustrating heatmaps. *PLoS ONE* **2014**, *9*, e111988. [CrossRef]
51. Saeed, A.I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M. TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* **2003**, *34*, 374–378. [CrossRef]
52. Zhu, H.; Han, X.; Lv, J.; Zhao, L.; Xu, X.; Zhang, T.; Guo, W. Structure, expression differentiation and evolution of duplicated fiber developmental genes in gossypium barbadense and *G. hirsutum*. *BMC Plant Biol.* **2011**, *11*, 40. [CrossRef]
53. Grabherr, M.G.; Haas, B.J.; Moran, Y.; Levin, J.Z.; Thompson, D.A.; Ido, A.; Xian, A.; Lin, F.; Raktima, R.; Qiandong, Z. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644. [CrossRef]
54. Li, P.; Piao, Y.; Shon, H.S.; Ryu, K.H. Comparing the normalization methods for the differential analysis of illumina high-throughput RNA-seq data. *BMC Bioinform.* **2015**, *16*, 347. [CrossRef]
55. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*, 2nd ed.; Springer: New York, NY, USA, 2009.
56. Ana, C.; Stefan, G.T.; Juan Miguel, G.G.; Javier, T.; Manuel, T.; Montserrat, R. Blast2go: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676.
57. Schnable, P.S.; Ware, D.; Fulton, R.S.; Stein, J.C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T.A.; et al. The B73 maize genome: Complexity, diversity, and dynamics. *Science* **2009**, *326*, 1112–1115. [CrossRef]
58. Marienfeld, J.; Unseld, M.; Brandt, P.; Brennicke, A. Genomic recombination of the mitochondrial Atp6 gene in *Arabidopsis thaliana* at the protein processing site creates two different presequences. *DNA Res.* **1996**, *3*, 287–290. [CrossRef]
59. Goff, S.A.; Darrell, R.; Tien-Hung, L.; Gernot, P.; Ronglin, W.; Molly, D.; Jane, G.; Allen, S.; Paul, O.; Hemant, V. A draft sequence of the rice genome (*Oryza sativa* L; Ssp. Japonica). *Science* **2002**, *296*, 92–100. [CrossRef]
60. Hua, Z.H.; Cheng, Z.; Shin-Han, S.; Vierstra, R.D. Phylogenetic comparison of *F-box* (FBX) gene superfamily within the plant kingdom reveals divergent evolutionary histories indicative of genomic drift. *PLoS ONE* **2011**, *6*, e16219. [CrossRef]
61. Initiative, A.G. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **2000**, *408*, 796–815. [CrossRef]
62. Dong, W.; Xu, C.; Li, W.; Xie, X.; Lu, Y.; Liu, Y.; Jin, X.; Suo, Z. Phylogenetic resolution in Juglans based on complete chloroplast genomes and nuclear DNA sequences. *Front. Plant Sci.* **2017**, *8*, 1148. [CrossRef]

63. Gorlova, O.; Fedorov, A.; Logothetis, C.; Amos, C.; Gorlov, I. Genes with a large intronic burden show greater evolutionary conservation on the protein level. *BMC Evol. Biol.* **2014**, *14*, 50. [CrossRef]
64. Long, M.; Deutsch, M. Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. *Mol. Biol. Evol.* **1999**, *16*, 1528–1534. [CrossRef]
65. Zheng, T.Z.; Zhuo, X.K.; Li, L.L.; Wang, J.; Cheng, T.R.; Zhang, Q.X. Genome-Wide analysis of the *D-type cyclin* gene family reveals differential expression patterns and stem development in the woody plant *Prunus mume*. *Forests* **2019**, *10*, 147. [CrossRef]
66. Islam, M.S.; Choudhury, M.; Majlish, A.N.K.; Islam, T.; Ghosh, A. Comprehensive genome-wide analysis of glutathione s-transferase gene family in potato (*Solanum tuberosum* L.) and their expression profiling in various anatomical tissues and perturbation conditions. *Gene* **2017**, *639*, 149–162. [CrossRef]
67. Dong, C.-J.; Shang, Q.-M. Genome-wide characterization of phenylalanine ammonia-lyase gene family in watermelon (*Citrullus lanatus*). *Planta* **2013**, *238*, 35–49. [CrossRef]
68. Dinant, S.; Clark, A.M.; Zhu, Y.; Vilaine, F.; Palauqui, J.C.; Kusiak, C.; Thompson, G.A.; Thompson, G.A. Diversity of the superfamily of phloem lectins (phloem protein 2) in angiosperms. *Plant Physiol.* **2003**, *131*, 114–128. [CrossRef]
69. Hofmann, K. Skp1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the *F-box*. *Cell* **1996**, *86*, 263–274.
70. Wang, N.Y.; Shirogane, T.; Liu, D.; Harper, J.W.; Elledge, S.J. Exit from exit: Resetting the cell cycle through am inhibition of g protein signaling. *Cell* **2003**, *112*, 697–709. [CrossRef]