Identification of *WD40* gene family in *Prunusmume* and its expression level under light illumination

Yue Fanfan [1] , **Chen Yougen** [1] , **Liu Bingjing** [1] , **Gan Defang** [1] , **Wang Hua** [1] , **Wang Dongliang** Corresp. [1]

[1] College of Horticulture, Anhui Agricultural University, Hefei, China

Corresponding Author: Wang Dongliang
Email address: wangdongliang@ahau.edu.cn

The *WD40* gene family in *Prunusmume* was identified by bioinformatics analysis., The evolution and function of *WD40* gene family were predicted by phylogenetic analysis, chromosome localization analysis, protein sequence analysis, exon-intron structure analysis. The results showed that the *WD40* gene family contained 77 members, which were divided into 61 groups, and only 18 groups had more than one member. When the plum plants were treated with different density of light illumination, the content of anthocyanin in plants was increased with the illumination enhancement, qRT-PCR analyses revealed that the expression of 27 *WD40* genes showed the same expression trend. The results of this study can provide useful information for further studying the function of *WD40* genes in *Prunusmume.*

Identification of *WD40* gene family in *Prunus mume* and its expression level under light illumination

Fanfan Yue, Yougen Chen, Bingjin Liu,Defang Gan, Hua Wang, Dongliang Wang*

（School of Horticulture, Anhui Agricultural University, Hefei, China）

***：wangdongliang@ahau.edu.cn**

Abstract: The *WD40* gene family in *Prunus mume* was identified using bioinformatics analysis., The evolution and function of *WD40* gene family were predicted using the methods of phylogenetic, chromosome localization, protein sequence, and exon-intron structure analyses. The results showed that the *WD40* gene family contained 77 members, which were divided into 61 groups, and only 18 groups had more than one member. When the plum plants were treated with different density of light illumination, the content of anthocyanin in plants was increased with the illumination enhancement, qRT-PCR analyses revealed that the expression of 27 *WD40* genes showed the same expression trend. The results of this study can provide useful information for further studying the function of *WD40* genes in *Prunus mume*.

Key words: *Prunus mume*; light intensity; *WD40*; expression level

1 Introduction

Prunus mume is an important ornamental plants in the landscape. The color of plum flower is one of the main embodiment of its ornamental value. At present, the color of plum flower mainly shows red, purple, pink, white, green, yellow and so on[1]. Anthocyanin is one of the main pigments that compose the color of petals and fruits. It has been found that anthocyanins can make a variety of colors, from pink to blue. Because of its influence on plant color, the breeders have paid much attention on it.

WD40 protein is an diverse super family of regulatory proteins, and is highly conserved in evolution and exists in a wide range of eukaryotes, participating in a variety of biochemical mechanisms and cellular processes [2]. WD40 protein is also known as WD-repeat protein or Trp-Asp, which contains about 40~60 amino acid residues [3]. In general, WD40 protein contains 4~10 tandem repeats of WD domains, and N-terminal has a Ganki acid histidine dipeptide (Gly-His, GH), C-terminal has a tryptophan aspartic acid dipeptide (Trp-Asp, WD) [4].

WD40 transcription factors are widely found in plants and related to anthocyanin

Funding: This research was financed by Science and Technology Research Plan of Anhui Province(1704e0302200).

biosynthetic pathway. Since the first transcription factor c1 was cloned in maize (*Zea mays*) in 1986 [5], scientists began to study the transcription regulatory factors of anthocyanin synthesis. Walker et al. [6] found the TTG1 transcription factorin WD40 family in 1999. The AN11 of *Petunia* is the first WD40 protein found in plants, and the expression of DFR is affected and decreased in *an11* mutants, which regulates the synthesis of anthocyanin[7].The mutants of Arabidopsis *ttg1* were found to reduce the expression of DFR and to reduce anthocyanin synthesis by Brueggemann et al.[8].Sompornpailin et al[9]. found that WD40 transcription factor PhAN11 in *Petunia* can be transferred into *Arabidopsis thaliana* to promote the formation of anthocyanins in *Arabidopsis thaliana*.Subsequently, the WD40 proteins of maize(PAC), perilla (Pf WD), morning glory (IpWDR1, IpWDR2), Alfalfa (MtWD40) and other plants were also found to be related to the anthocyanin biosynthetic pathway[9-15]. Xiu-Hong An identified an apple WD40 protein (MdTTG1) that promotes the accumulation of anthocyanins in 2012[16].

There are many related domestic studies in the past time. For example, Xu studied the related genes in the anthocyanin synthesis process of turnip in 2008[17].Luo studied the cloning and expression of anthocyanin transcription-activating gene *StWD40* in 2008[18]. Han studies on the key genes of anthocyanin biosynthesis in chrysanthemum[19]. Zhang further studied the homology between PsWD40-1 and PsWD40-2 protein sequences of peony and WD40 protein related to anthocyanin biosynthesis in other species[20]. Min found that when the WD40 transcription factor TTG1 was over expressed, the expression of related genes in anthocyanin biosynthetic pathway was upregulated, and the anthocyanin accumulation was increased[21]. Hong summarized the transcription factors of anthocyanin synthesis in *Arabidopsis thaliana*, and elaborated that the pathways of anthocyanin synthesis in *Arabidopsis* were mainly regulated by four transcription factors, MYB, bHLH, WD40 and bZIP[22].The mechanism of anthocyanin biosynthesis in the leaves of the red-claw Maple Leaf was expounded by An et al. [23], and the transcription factor WD40 gene, which regulates anthocyanin expression, was obtained successfully.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

At present, *WD40* genes in *Prunus mume* has not been reported in the literature. In this paper, the *WD40* genes in *Prunus mume* were studied to provide further understanding of the nature and function of WD40 protein family. Finding the genes related to anthocyanin biosynthesis in plum flower can provide a basis for future research on color improvement.

2 Materials and methods

2.1 Plant materials and treatments

2.1.1 Plant materials

One-year-old plum "Pink cinnabar" in the School of Horticulture, Anhui Agricultural University was used in this study."Pink cinnabar" is a species of red plum.

2.1.2 Different light intensity treatments

Plants with similar growth potential and similar leaf area were selected and placed in an artificial growth box (16-h light from 06:00 to 22:00, 23-27 °C). Different light intensity treatments were: Normal-light: 7,000-8,000 lx; Low-light: 3,000-4,000 lx; High-light: 11,000-12,000 lx. After7 days, all leave samples were clooected and immediately frozen in liquid nitrogen, and stored at -80 °C for determination of anthocyanin content and RNA extraction.

2.2 Extraction of anthocyanin from plum leaves and determination of its content

Anthocyanin was isolated from plum leaves according to Xiong (2003), anthocyanin content was determined using the formula:（nmol/cm^2）=OD_λ/ε_λ × V/S × 〖10〗^6。OD_λ: the optical density of anthocyanin; ε_λ: anthocyanin molar extinction coefficient of 4.62×10^4; V: total volume of extract (mL); S: sampling area (cm^2); 〖10〗^6: the result of the calculation is converted into a multiple of nmol.

2.3 Database search and identification of *WD40* family genes in *Prunus mume*

To identify WD40 proteins in plum, the plum genome data base (http://prunus mume

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

genome.bjfu.edu.cn/) [24] was searched using BLASTP (Basic Local Alignment Search Tool algorithms), with the published *Arabidopsis* WD40 protein sequences as queries. The *Arabidopsis* WD40 protein sequences are obtained from the Tair website (https://www.arabidopsis.org/). All acquired protein sequences were analyzed manually for the presence of the WD40 domain using the Hidden Markov Model of Pfam (http://pfam.xfam.org/)[25] and SMART (Simple Modular Architecture Research Tool) tools (http://smart.embl-heidelberg.de/smart/)[26].

2.4 Analysis of physicochemical properties

Physical parameters of the predicted WD40 proteins, including the amino acid length (aa), molecular mass (Da), isoelectric point (pI) and instability index for each gene product were calculated using the online ExPASy programs (http://www.expasy.org/protparam/)[27].

2.5 Phylogenetic analysis

Phylogenetic trees were constructed with the NJ method [28]. Bootstrap analysis was conducted using 1000 replicates. Multiple sequences of all WD40 protein sequences from Arabidopsis and plum were aligned using ClustalX 2.0 [29]. An unrooted NJ tree was constructed using MEGA 6 [30].

2.6 Identification of conserved motifs

We used the online MEME (Multiple Expectation Maximization for Motif Elicitation) (http://meme-suite.org/tools/meme) program to analyze the conserved motif structures of the proteins encoded by the *WD40*genes [31]. The program was set to detect up to 20 motif. Multiple sequence alignmen was performed using DNAMAN 6.0 software (LynnonBiosoft, Qc, Canada).

2.7 Gene structure analysis

Exon and intron structures of individual plum *WD40* genes were analyzed using the Gene structure display server (GSDS; http://gsds.cbi.pku.edu.cn/) by alignment of the cDNAs with their corresponding genomic DNA sequences[32].

2.8 Chromosomal location

The chromosomal location image of *WD40* genes was generated by Map Draw V2.1

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

software [33] according to chromosomal position information provided by the plum genome database.

2.9 RNA extraction and qRT-PCR analysis

Total RNA was extracted from the collected plum leaf samples using the UNlQ-10 Column Trizol Total RNA Isolation Kit (Sangon Biotech) according to the manufacturer's instruction. The integrity of the RNA was assessed by1.5% agarose gel electrophoresis. First-stand cDNAs were synthesized from RNA using the Trans Script One-Step gDNA Removal and cDNA Synthesis Super Mix (Trans Gen Biotech) according to the manufacturer's instructions.

The qRT-PCR analysis was performed on a BIO-RAD CFX96 Real-Time system. Each reaction contained 10μl of SYBR Green Master Mix reagent (Applied Biosystems), 400 nM gene-specific primers, and 2.0 μl of diluted cDNA sample in a total volume of 20μl. The qPCR thermal cycle reaction conditions as follows: 95 °C for 30 s; followed by 40 cycles of 5 s at 95 °C and then 60 °C for 30 s. A melting curve was generated at the end of the 40 cycles and used to analyze the specificity for each gene, three biological replicates were conducted for each sample. The actin gene was used as internal reference.The relative expression level was calculated as $2^{-\Delta\Delta CT}$ [$\Delta CT = CT_{Target} — CT_{Actin}$， $\Delta\Delta CT = \Delta CT_{Test} — \Delta CT_{Control}$].

3 Results

3.1 Identification of WD40 family genes in *Prunus mume*

The *A. thaliana* full-length WD40 protein sequences (AEC10837.1) were used as queries to search against the plum genome database. All potential candidates were subjected to SMART and Pfam analyses to ensure that the predicted proteins contained WD40 motifs. Through an extensive search for *WD40* genes, 77 non-redundant WD40 genes (named *PmWD4001* to *PmWD4077*) were identified (Table1). The identified plum WD40 gene family members encode predicted proteins ranging from 269 (*PmWD4038*) to 3612 (*PmWD4077*) amino acid residues (aa) in length, with an average length of 712 aa, molecular masses of 30746.79 Da (*PmWD4038*) to 402874.69 Da (*PmWD4077*) (average of 79041.88Da)in theory, and its oelectric point of 4.58 (*PmWD4057*) to 9.53

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

(*PmWD4007*) (average of 6.76). Moreover, when the instability index was more than 40, the gene was in an unstable state [35]. Only 42.9% of the predicted WD40 proteins were below 40, indicating that the WD40 proteins are in an unstable state and prone to degradation. The details of the *WD40* gene family, such as locus names, were provided in Table 1.

3.2 Phylogenetic analyses of *WD40* gene family members

To explore the phylogenetic relationships of WD40 proteins from different species, we constructed a phylogenetic tree of the WD40 protein families from *Arabidopsis* and plum (Figure 1). Two hundred and thirty seven WD40 proteins in *A.thaliana* were divided into 143 different groups, and only 49 of the groups have more than one member. The Figure 1 showed that all the plum WD40 proteins were distributed into 61 groups, 18 groups have more than one member. From the NJ tree, we noted that most PmWD40 proteins clustered with Arabidopsis had a high bootstrap support. The distribution of PmWD40 proteins are similar to that of AtWD40 protains, although there were less groups in plum than that in *Arabidopsis*.

3.3 Chromosomal locations of *PmWD40* genes

According to the chromosomal location map, 72 of the 77 plum *WD40* genes were widely, but unevenly,distributed on the 8 poplar chromosomes, while five genes (*PmWD4073-77*) were mapped to as yet unattributed scaffolds (Figure 2). The number of *WD40* genes on each chromosome varied. Chromosome 5 and 8 contained only five plum *WD40* genes, while chromosome 2 had the largest number of *WD40* genes (20). The genes on the same chromosome were also unevenly distributed, genes in the chromosomes 1 and 7 were mainly on the long arm, and the genes in chromosome 2, 3, 4 and 6 were mainly on the short arm, A gene cluster (*PmWD4009* and *PmWD4010*) was found on chromosome 2. The distribution of genes on other chromosomes was randomly.

3.4 Domain structure sequence analysis of *PmWD40* genes

We used the MEME software of online tools to search for conserved motifs that are shared among related proteins within each subfamily for 77 WD40 proteins. Twenty

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

distinct motifs were identified, and the details of the conserved amino acid sequences and their lengths were shown in Figure3. The results showed that the same subfamily members contained very similar conserved structural motifs. We speculated that these genes with the same or similar domain might have the same or similar functions.

The number and combination of motifs in the WD40 proteins were difference. We observed that all the amino acid sequences encoded by the *PmWD40* genes contain multiple identical conserved amino acid sequences. The amino acid sequences shown in Table 2 are the core sequence of PmWD40 proteins. In addition, each of the putative motifs was annotated by searching Pfam and SMART tool

To further understand the similarity between plum WD40 domains, we compared the sequences of all 77 WD40 domains (Figure 4), the result showed that 77 WD40 proteins were only 4.09% identity. The 77 sequences ranged from 43 ~ 60 amino acids, the colored part is the homologous part of their sequence.As shown in Fig. 4, the sequences have very limited homologous portions, which made it is difficult to determine the conservative sequence of the WD40 domain.According to the literature, the WD40 structure domain usually contains GH and WD, and the universal sequence is $\{X^{6\text{-}94}\text{-}[GH\text{-}X^{23\text{-}41}\text{-}WD]\}^{4\text{-}16[37]}$. This sequence is similar to the putative WD40 domain sequences from Table 2. Therefore, we speculated that these sequences might be the conserved sequences of the WD40 domains of plum.

3.5 Structural analyses of *PmWD40* gene family members

To further understand the structural diversity of *WD40* genes, we analyzed the exon/intron organization from the coding sequences of *PmWD40* genes (Figure 5). The intron number of the 77 *PmWD40* genes ranged from 0 to 35, of which 12 genes have no intron, *PmWD4002*,*-07*,*-21*,*-22*,*-23*,*-26*,*-48*,*-73*,*-74*,*-75*,*-76 and -77*, and accounted for 15.6% of the total. The most closely related gene members in the same subfamilies shared similar exon/intron structures, although exception to this observation was also found. And the *PmWD40* genes appears to be more compact than normal in the cluster of the same structure, for instance, *PmWD4034* and *PmWD4063*. This is similar to the pattern shown in the Figure 3.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

3.6 Anthocyanin content in leaves of plum

The content of anthocyanin in plum leaves was analyzed and shown in Figure 6. The content of Anthocyanin was the highest in the high-light treatment, the content of the pigment was the lowest in the low-light treatment, indicating that the content of anthocyanin increased with the increase of light illumination.

3.7 Expression analysis of *PmWD40* gene family under different light illumination

The level of gene expression in the high-light treatment was higher than that in normal-light treatment, the lowest level of gene expression was detected in the low-light treatment. Data of gene expression (Figure 7) showed that 27 genes that conform to the trend (*PmWD4001*、*-06*、*-07*、*-10*、*-12*、*-15*、*-19*、*-20*、*-22*、*-23*、*-24*、*-25*、*-27*、*-29*、*-31*、*-36*、*-37*、*-42*、*-48*、*-53*、*-56*、*-57*、*-62*、*-64*、*-65*、*-70* and *-72*). However, 2 genes (*PmWD4047* and *PmWD4050*) were completely opposite to what was expected, with the lowest expression in high-light and the highest expression in low-light treatments.

The remaining 48 genes showed no specific trend, occupying 53.2% of the total number of *PmWD40* genes. The expression of 24 genes (*PmWD4002*, *-03*, *-04*, *-05*, *-09*, *-11*, *-13*, *-14*, *-16*, *-21*, *-26*, *-32*, *-33*, *-39*, *-40*, *-41*, *-44*, *-51*, *-54*, *-61*, *-63*, *-68*, *-69* and *-75*) was reduced under high-light and low-light, and the expression was highest under normal-light treatment. The expression of 17 genes (*PmWD4008*, *-17*, *-18*, *-28*, *-34*, *-35*, *-38*, *-45*, *-46*, *-52*, *-58*, *-59*, *-60*, *-66*, *-74*, *-76* and *-77*) was the lowest under normal-light, and was up regulated under either high-light or low-light.

4 Discussion

WD40 is an astonishing variety super family of regulatory proteins that are highly conserved in evolution and present in a wide range of eukaryotes, involving in various biochemical mechanisms and cellular processes[37].According to the previous report, anthocyanin biosynthesis pathway is regulated by two groups of genes: one is the structural gene encoding the enzymes required in the biosynthetic pathway, and the other

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

is the regulator gene that encode the temporal and spatial expression of structure genes[38]. In general, the structural gene determines the type of anthocyanin, and the regulatory gene determines the expression of the structure gene. As a member of the regulatory gene, the*WD40* gene family plays an important role in anthocyanin biosynthetic pathway. We hope to identify the related genes to improve the flower color of plum.

In this paper, we identified 77 nonredundant *WD40* genes in the plum by bioinformatics method. More than half of the proteins were found to be in an unstable state and prone to degradation. According to phylogenetic analysis, the 77 *PmWD40* genes were divided into 61 groups, and only 18 groups had more than one member, this might be related to the diversity of WD40 protein functions. Based on the phylogenetic tree of the WD40 protein families from *Arabidopsis* and plum, similar to what has been reported for *Arabidopsis thaliana*[39], and the WD40 protein family in plum also showed low conservative of sequence. We found that some or homologous proteins exhibited a closer relationship, implying that those or tholog pairs may have derived from a common ancestor. In general, genes clustered into the same group of a phylogenetic tree often share similar functional features.

By analyzing the distribution of *PmWD40* genes in chromosome position, it was found that the number of genes distributed on each chromosome was different, and the distribution of genes on the same chromosome was not uniform. The composition and number of WD40 proteins in the same group were similar. Futher more, the composition and number of exon and intron were similar, which indicated that the genes of the same group might have the same function. Previous studies revealed that WD40 proteins have defined as $\{X^{6-94}\text{-}[GH\text{-}X^{23-41}\text{-}WD]\}^{4-16}$[37]. As shown in this study, after analysis of the WD40 gene family in plum, motif 1, 2, 3, 5 and 9 were identified as core sequences of PmWD40 proteins, which are widely distributed in the PmWD40 protein family. In general, the similarities in gene structures and motifs of the majority of *WD40* genes within each group supported the phylogenetic analysis. Similarly, the differences in the characteristics among the different groups may imply that the plum *WD40* genes are functionally diverse.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

The previous studies showed that illumination was one of the most important regulatory factors affecting the synthesis of plant anthocyanins, and illumination can not only regulate the synthesis of anthocyanin in plants, but also play an important role in flower development. The optical signal conduction factor and the photoreceptors play an important role in the synthesis and accumulation of anthocyanin.

In this study，the anthocyanin content in plum leaves was determined with different light illumination treatments, and the content of anthocyanin in leaves increased with the increase of light illumination. By qRT-PCR, the expression of PmWD40 genes was obtained under different light illumination. The results showed that expression of 27 in high-light treatment was higher than that in normal-light treatment. The expression of two genes was completely opposite to previous trend, with the lowest level of expression in high-light and the highest level of expression in low-light. These genes with obvious regularity might be closely related to anthocyanin regulation in plum leaves. The remaining 48 genes did not show any special expression regularity. If we want to understand the rationale behind these different manifestations, we need to continue to work more intensively in the future.

5 Conclusions

In conclusion, *PmWD40* gene family might be involved in anthocyanin synthesis and regulation process. The regulatory mechanisms of WD40 proteins in plum remain poorly understood. Therefore, systematic analysis of the *WD40* gene family may provide useful information for further studies on the family structure and biological functions of plum WD40 proteins. Furthermore, in order to better study the changes of flower color, plum blossom petals can be used as samples in later experiments.

References

[1] Hodges SA, Derieg NJ, Avise JC, et al. Adaptive radiations: from field to genomic studies [J].Proceedings of the National Academy of Sciences of the United States of America, 2009,106(Suppl 1): 9947-9954.

[2] Stirnimann CU, Petsalaki E, Russell RB, et al. WD40 proteins propel cellular networks

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

[J].Trends in biochemical sciences, 2010, 35(10): 565-574.

[3] Van Nocker S, Ludwig P. The WD-repeat protein superfamily in Arabidopsis: conservation anddivergence in structure and function [J]. BMC Genomics, 2003, 4: 50.

[4] Smith TF, Gaitatzes C, Saxena K, et al. The WD repeat: a common architecture for diversefunctions [J]. Trends in Biochemical Sciences, 1999, 24(5): 181-185.

[5] Cone KC, Burr FA,Benjamin B. Molecular analysis of the maize anthocyanin regulatory locus C1[J]. Proc Natl Acad Sci USA, 1986, 83: 9631-9635.

[6] Walker AR, Davison PA, Bolognesi-Winfield AC, et al The TRANSPARENT TESTA GLABRA1 locus,which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein[J].Plant Cell, 1999, 11: 1337-1350.

[7] de Vetten N, Quattrocchio F, Mol J, et al. The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast， plants,and animals[J]. Genes & Development, 1997, 11:1422-1434.

[8] Brueggemann J, Weisshaar B, Sagasser M. A WD40-repeat gene from Malus x domestica is a functional homologue of Arabidopsis thaliana TRANSPARENT TESTA GLABRA1[J]. Plant cell roport, 2010, 29: 285-294.

[9] Sompornpailin K, Makita Y, Yamazaki M, et al. A WD-repeat-containing putative regulatory protein in anthocyanin biosynthesis in Perilla frutescens[J]. Plant molecular biology, 2002, 50: 485-495.

[10] Morita Y, Saitoh M, Hoshino A, et al. Isolation of cDNAs for R2R3-MYB, bHLH and WDR transcriptional regulators and identification of c and ca mutations conferring white flowers in the Japanese morning glory[J]. Plant Cell Physiol, 2006, 47:457-470.

[11] Baudry A, Heim MA, Dubreucq B, et al. TT2, TT8 and TTGl synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana[J]. The Plant Journal, 2004, 39: 366-380.

[12] Carey CC, Strahle JT, Selinger DA, et al. Mutations in the pale aleurone color 1 regulatory gene of the Zea mays anthocyanin pathway have distinct phenotypes relative to the functionally similar TRANSPARENT TESTA GLABRA1 gene in Arabidopsis thalian[J]. Plant Cell, 2004, 16(2):450-464.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

[13] Ramsay NA, Glover BJ. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity[J]. Trends in Plant Science, 2005, 10(2):63-70.

[14] Yamazaki M, Makita Y, Springob K, et al. Regulatory mechanisms for anthocyanin biosynthesis in chemotypes of Perilla frutescensvar Crispa[J]. Bioehemical Engineering Joumal, 2003, 14 (3):191-197.

[15] Pang Y, Wenger JP, Saathoff K, et al. A WD40 repeat protein from Medicago truncatula is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development[J]. Plant Physiolgy, 2009, 151(3):1114-1129.

[16] Xiu-Hong An,Yi Tian,Ke-Qin Chen,Xiao-Fei Wang,Yu-Jin Hao. The apple WD40 protein MdTTG1 interacts with bHLH but not MYB proteins to regulate anthocyanin accumulation[J]. Journal of Plant Physiology,2012,169(7).

[17] Xu ZR, Li CL, Sun Y, et al. Screening the Genes Associated with Anthocyanin Biosynthesisi in 'Tsuda' and 'Yurugi Akamaru' Turnip (Chinese) [J]. Letters in Biotechnology,2008, 19(5):693-696.

[18] Luo ZX, Liu SY, Zhang SZ, et al. Cloning and Expression Analysis of the Anthocyanin Transcriptional ActivatorGene stwd40 of Solanum tuberosum (Chinese) [J]. Acta Horticulturae Sinica, 2008, 35(9): 1317-1322.

[19] Han KT, Zhao L, Tang XJ, et al. The Relationship Between the Expression of Key Genes in AnthocyaninBiosynthesis and the Color of Chrysanthemum (Chinese) [J]. Acta Horticulturae Sinica, 2012, 39(3): 516-524.

[20] Zhang C, Gao SL, Du DN, et al. Isolation and Sequence Analysis of the Paeonia suffruticosa WD40Transcription Factor Genes PsWD40-1 and PsWD40-2 (Chinese) [J]. Biotechnology Bulletin, 2014,(2):85-90.

[21] Min YQ, Yan HF, Li YH. WD40 Proteins of Anthocyanin Biosynthesis in Plant (Chinese) [J].Plant Physiology Journal, 2010, 46: 863-870.

[22] Hong X, Zhang SB, Xu J, et al. Research Progress on important transcription factors of anthocyanin biosynthesis pathway in Arabidopsis thaliana (Chinese) [J].Chinese Horticulture Abstracts, 2014,(4):1-3.

[23] An LJ, Feng AL, Shi BS. Cloning and sequence analysis of WD40 transcription factor from leaves of Acer negundo (Chinese) [J].Northern Horticulture,2012,(12):148-152.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

[24] Zhang Q, Chen W, Sun L, et al. The genome of*Prunusmume*[J]. Nature Communications, 2012, 3(4):1318.

[25] Finn RD, Penelope C, Eberhardt RY, et al. The Pfam protein families database: towards a more sustainable future[J]. Nucleic Acids Research, 2016, 44(Database issue):D279-D285.

[26] Letunic I, Bork P. 20 years of the SMART protein domain annotation resource[J]. Nucleic Acids Research, 2017.

[27] Artimo P, Jonnalagedda M, Arnold K, et al. ExPASy: SIB bioinformatics resource portal[J]. Nucleic Acids Research, 2012, 40(Web Server issue):597-603.

[28] Retief JD. Phylogenetic analysis using PHYLIP[J]. Methods in Molecular Biology, 2000, 132(132):243-258.

[29] Tang Z, Blacquiere G, Leus G. Clustal W and Clustal X version 2.0[J]. Bioinformatics, 2007, 23(21):2947-2948.

[30] Tamura K, Peterson D, Peterson N, et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0 [J].Molecular Biology & Evolution, 2013,30(12): 2725-2729.

[31] Bailey TL, Mikael B, Buske FA, et al. MEME Suite: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37(Web Server issue):W202-W208.

[32] Guo AY, Zhu QH, Chen X, et al. GSDS: a gene structure display server [J]. Hereditas, 2007, 29(8):1023-1026.

[33] Liu RH, Meng JL. MapDraw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data [J]. Yi Chuan, 2003, 25(3):317-321.

[34] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-timequantitative PCR and the 2(-Delta Delta C(T)) Method [J]. Methods, 2001, 25(4): 402-408.

[35] Guruprasad K, Reddy BV, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence[J]. Protein Engineering, 1990, 4(2):155-161.

[36] Van Nocker S, Ludwig P. The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function [J]. BMC Genomics, 2003, 4: 50.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

374 [37] Stirnimann CU, Petsalaki E, Russell RB, et al. WD40 proteins propel cellularnetworks [J].
375 Trends in biochemical sciences, 2010, 35(10): 565-574.
376 [38] Jaakoa L, Hohtola A. ExPression of genes involved in anthocyan in biosynthesis in relation
377 toanthocyanin, Proanthocyanidin, and fiavonol levels during bilberry fruit development [J].
378 PlantPhysiology, 2002, 130(2): 729-739.
379 [39] Evaluating the Phylogenetic Position of Chinese Tree Shrew(Tupaia belangeri chinensis) Based
380 on Complete Mitochondrial Genome:Implication for Using Tree Shrew as an Alternative
381 Experimental Animal to Primates in Biomedical Research[J].遗传学报,2012,39(03):131-137.

Figure 1(on next page)

Phylogenetic relationships in WD40 protein family of *Prunus mume* and *Arabidopsis thaliana*

To explore the phylogenetic relationships of WD40 proteins from different species, we constructed a phylogenetic tree of the WD40 protein families from *Arabidopsis* and plum (Figure 1).

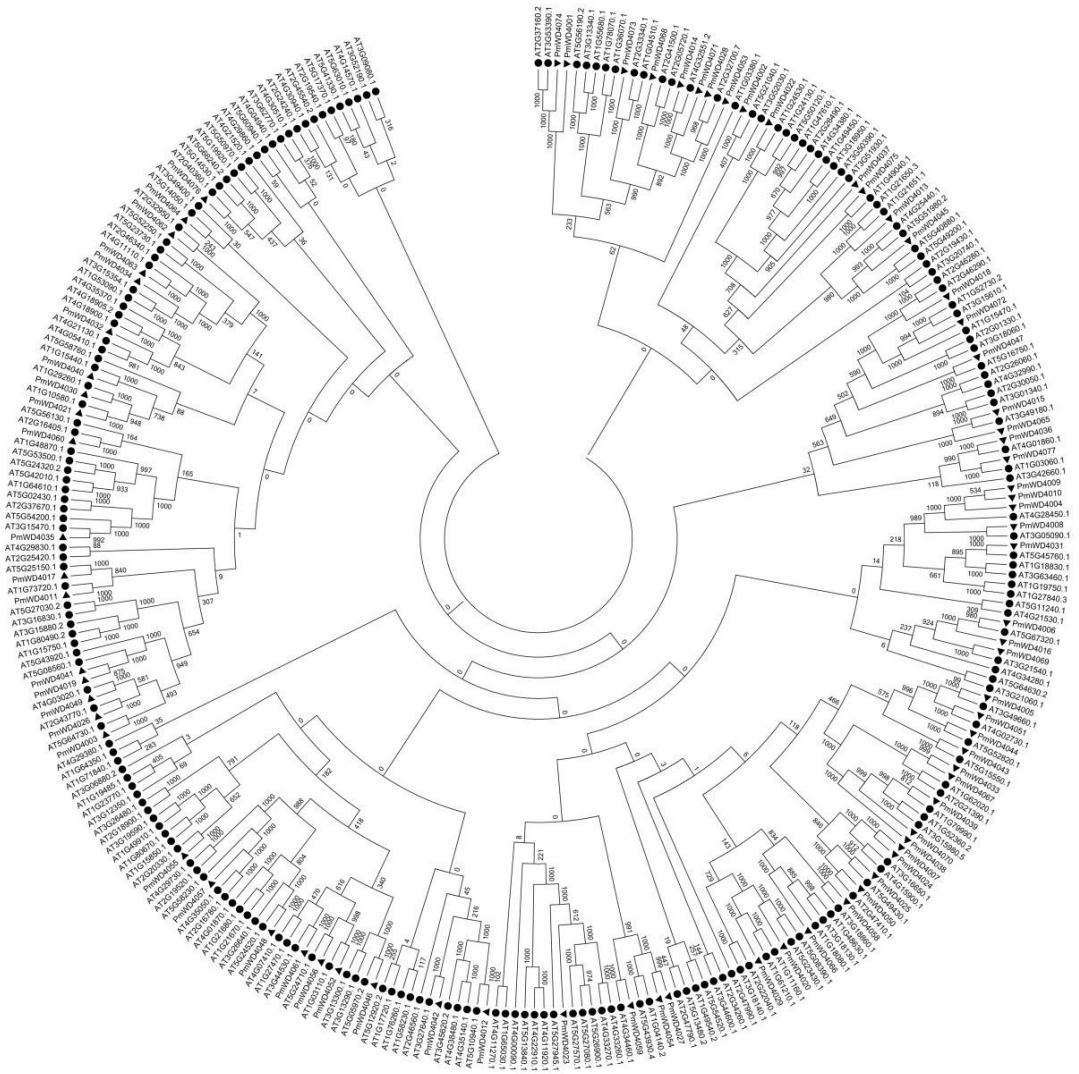


Figure 1 Phylogenetic relationships in WD40 protein family of *Prunus mume* and *Arabidopsis thaliana*

Figure 2(on next page)

Distributions of *WD40* gene family members on the *Prunus mume* chromosome s

According to the chromosomal location map, 72 of the 77 plum *WD40* genes were widely, but unevenly,distributed on the 8 poplar chromosomes, while five genes (*PmWD4073-77*) were mapped to as yet unattributed scaffolds (Figure 2).

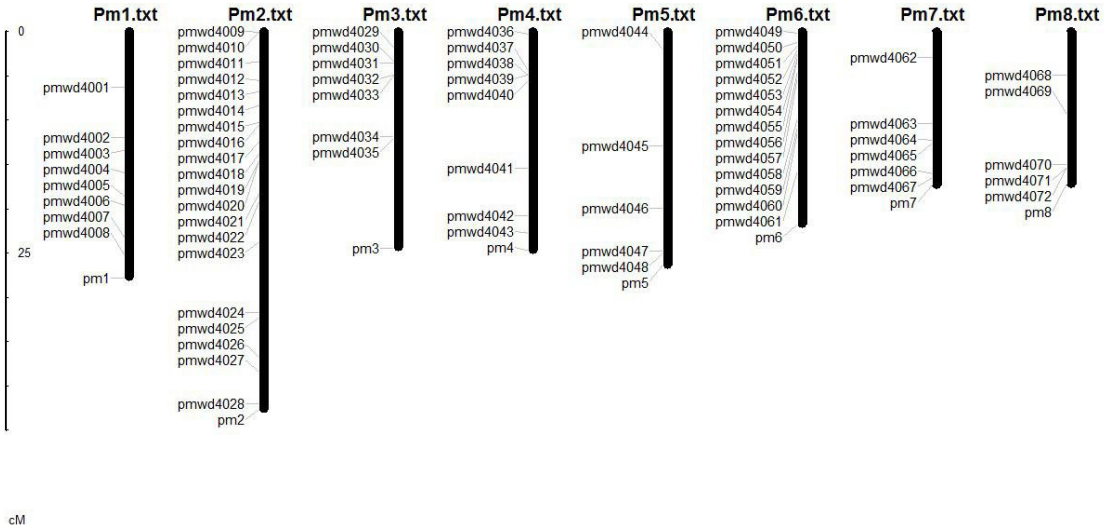


Figure 2 Distributions of *WD40* gene family members on the *Prunus mume* chromosomes

Figure 3(on next page)

Classification of PmWD40 protein family and conversed motifs

Twenty distinct motifs were identified, and the details of the conserved amino acid sequences and their lengths were shown in Figure3.

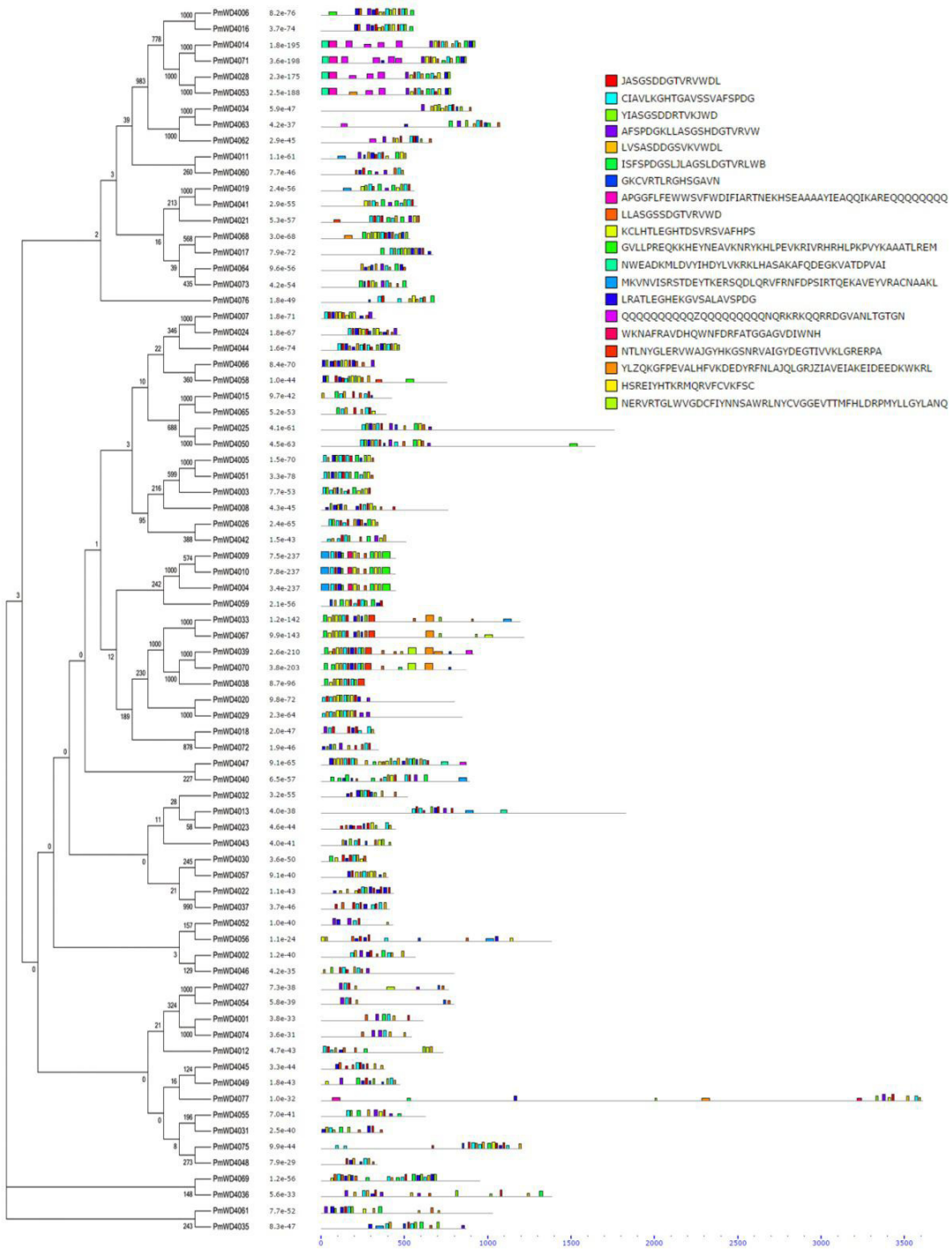


Figure 3 Classification of PmWD40 protein family and conversed motifs

Figure 4(on next page)

Multiple sequence alignment of PmWD40 protein family

To further understand the similarity between plum WD40 domains, we compared the sequences of all 77 WD40 domains (Figure 4), the result showed that 77 WD40 proteins were only 4.09% identity.

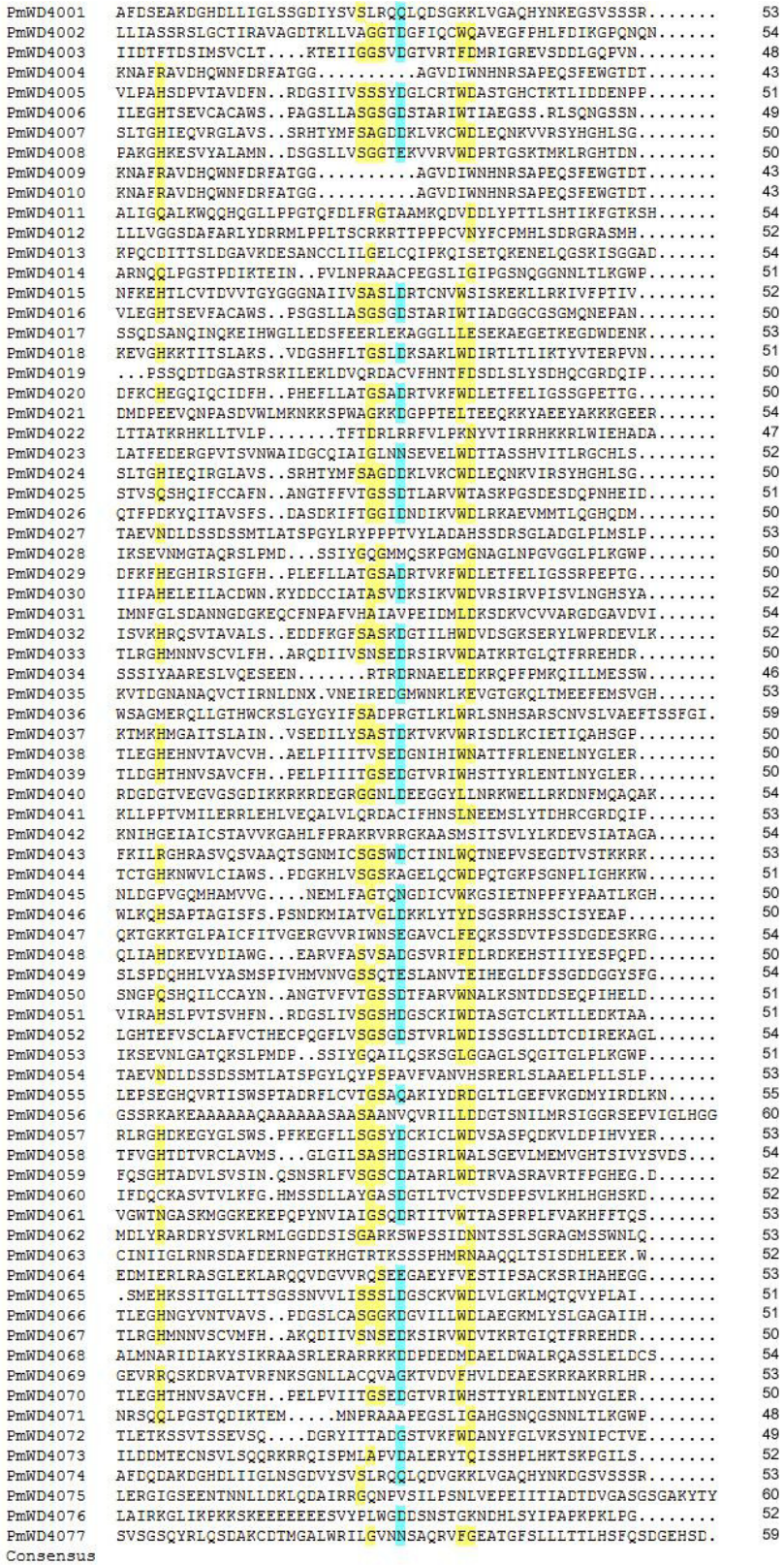


Figure 4 Multiple sequence alignment of PmWD40 protein family

Figure 5(on next page)

The intron-exon structures of *PmWD40* gene family

To further understand the structural diversity of *WD40* genes, we analyzed the exon/intron organization from the coding sequences of *PmWD40* genes (Figure 5).

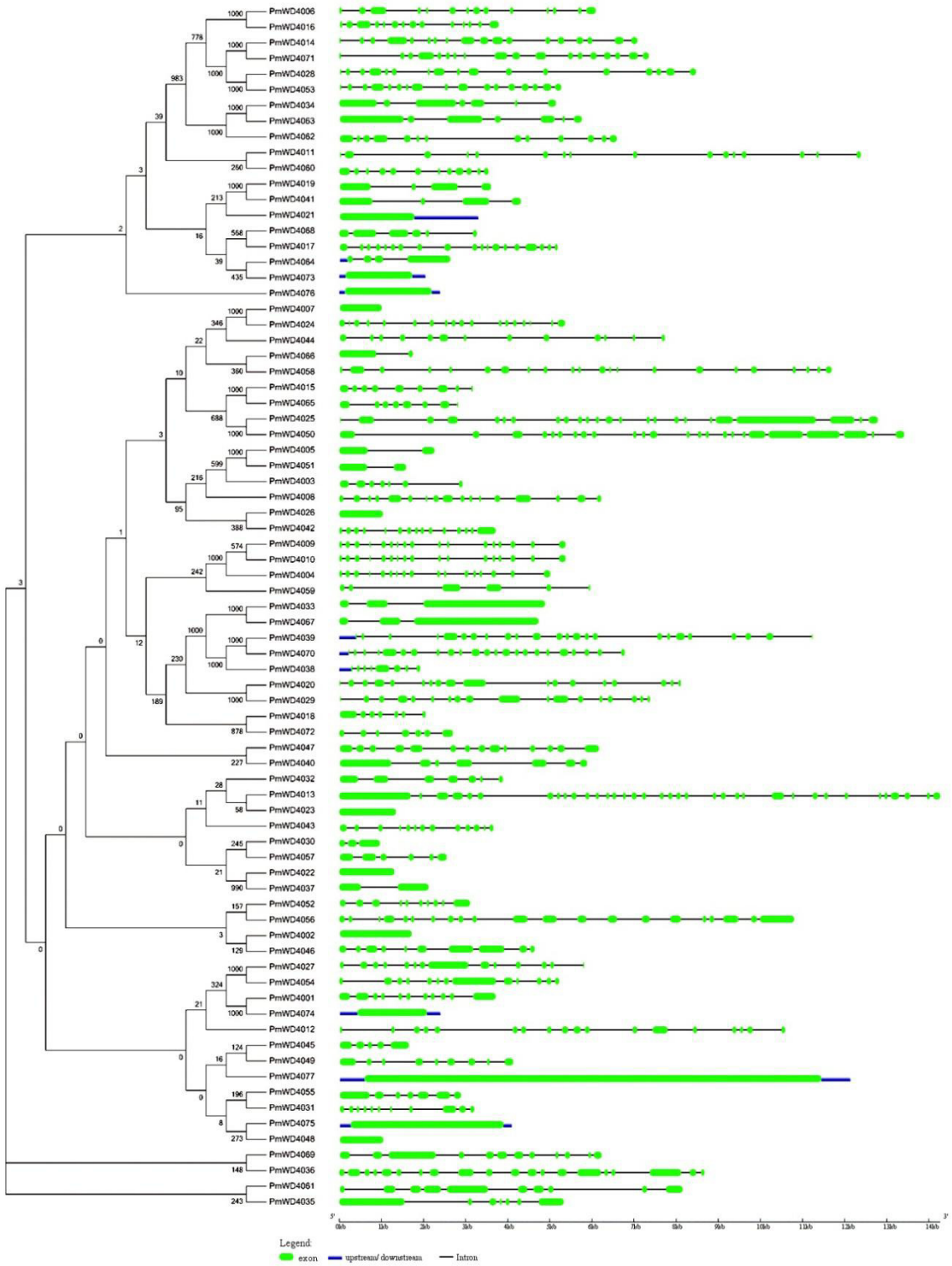


Figure 5 The intron-exon structures of *PmWD40* gene family

Figure 6(on next page)

Anthocyanin content in leaves of plum under different light illumination

The content of anthocyanin in plum leaves was analyzed and shown in Figure 6.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

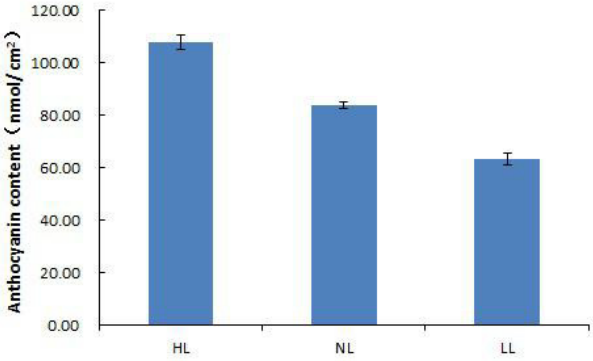


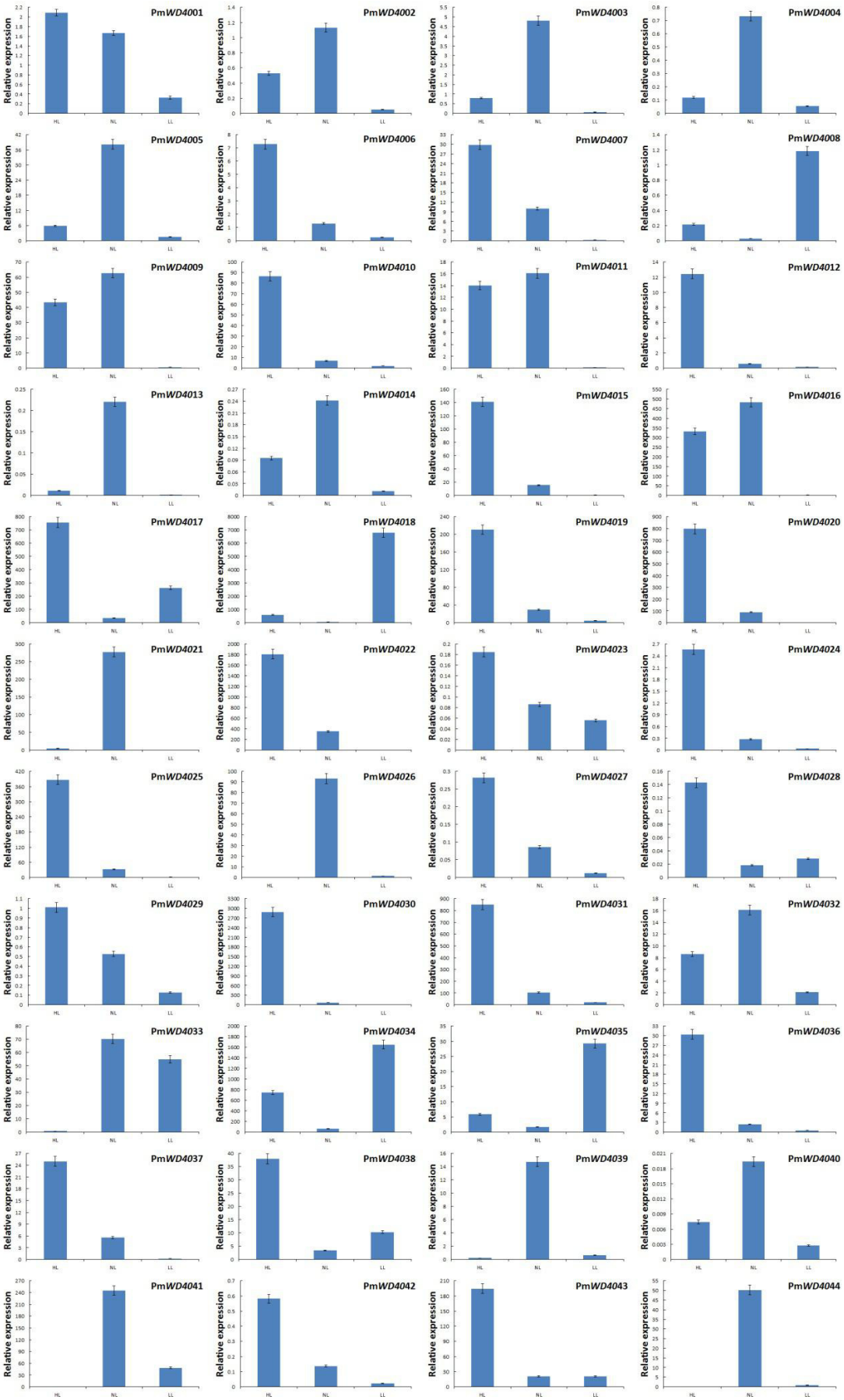
Figure 6 Anthocyanin content in leaves of plum under different light illumination

HL：High-light；NL：Normal-light；LL：Low-light

Figure 7(on next page)

Expression analysis of *PmWD40* gene family in different illumination

T he level of gene expression in the high-light treatment was h igher than that in normal-light treatment, the lowest level of gene expression was detected in the low-light treatment. Data of gene expression (Figure 7) showed that 27 genes that conform to the trend (*PmWD4001*、*-06*、*-07*、*-10*、*-12*、*-15*、*-19*、*-20*、*-22*、*-23*、*-24*、*-25*、*-27*、*-29*、*-31*、*-36*、*-37*、*-42*、*-48*、*-53*、*-56*、*-57*、*-62*、*-64*、*-65*、*-70*and *-72*).



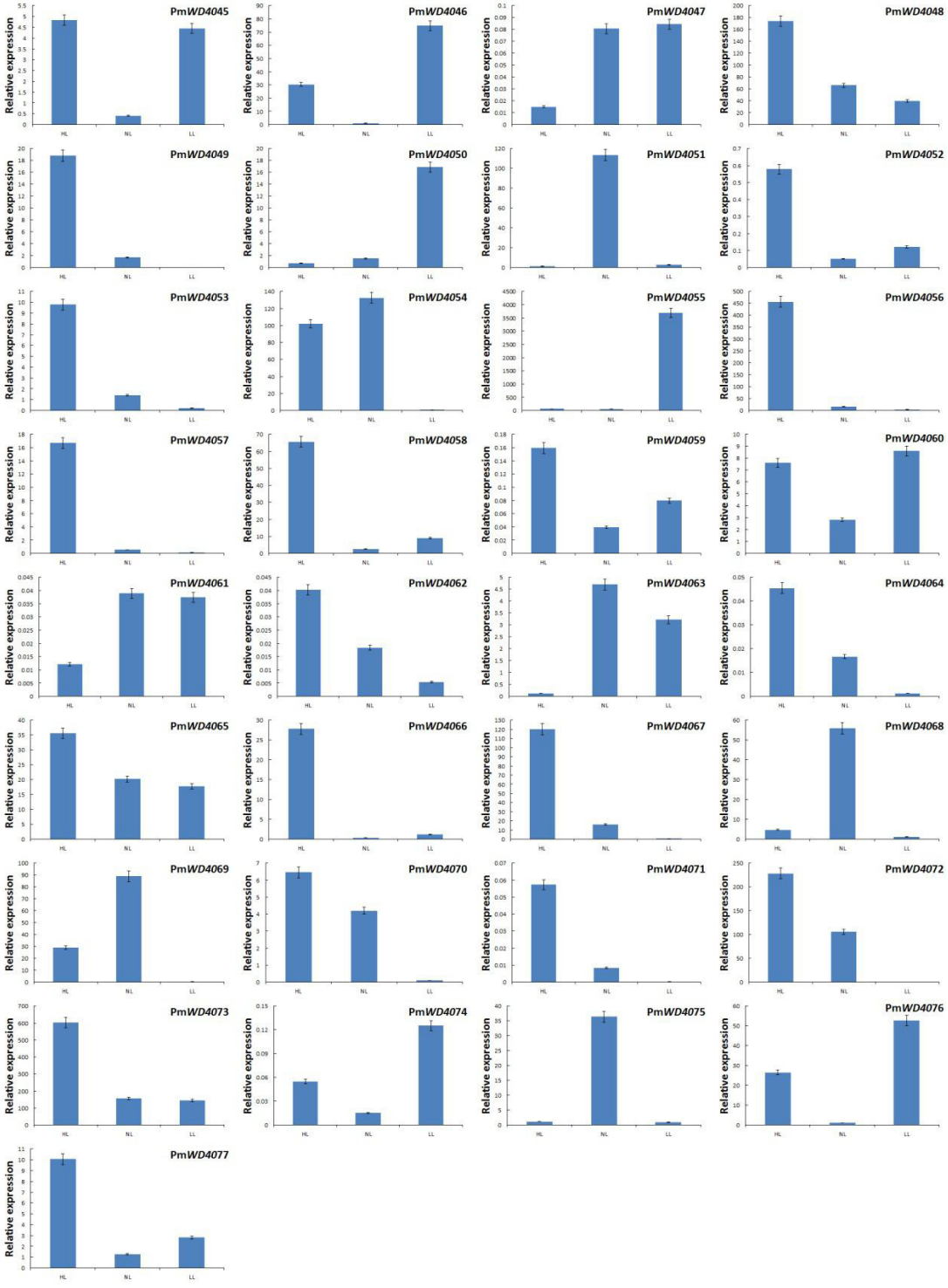


Figure 7 Expression analysis of *PmWD40* gene family in different illumination

HL：High-light；NL：Normal-light；LL：Low-light

Table 1(on next page)

The *WD40* gene family in *Prunus mume*

The details of the *WD40* gene family, such as locusnames, are provided inTable1.

Table 1 The *WD40* gene family in *Prunus mume*

Number	Name	Accession	Protein				Chr.	Location
			Length (aa)	MW (Da)	pI	Instability index (II)		
1	*PmWD4001*	XP_008225471	614	66062.17	8.56	47.46	1	6321988-6325700
2	*PmWD4002*	XP_008230996	567	63275.97	7.56	32.54	1	11904441-11906144
3	*PmWD4003*	XP_008232154	299	32994.07	6.59	33.36	1	13443750-13446661
4	*PmWD4004*	XP_008234484	451	51404.30	9.45	45.17	1	15948020-15953033
5	*PmWD4005*	XP_008236866	321	35060.25	6.74	30.00	1	18528859-18531118
6	*PmWD4006*	XP_008238593	562	63272.69	5.44	29.13	1	19563514-19569604
7	*PmWD4007*	XP_008246350	330	37023.73	9.53	36.84	1	23177695-23178687
8	*PmWD4008*	XP_016652838	763	83295.09	6.41	37.54	1	25148660-25154879
9	*PmWD4009*	XP_008218563	450	51264.12	9.45	44.70	2	73976-79352
10	*PmWD4010*	XP_008218559	451	51363.25	9.45	44.62	2	73976-79352
11	*PmWD4011*	XP_008219078	513	58131.38	6.97	38.80	2	3432932-3445299
12	*PmWD4012*	XP_016647902	734	80859.72	6.12	53.20	2	5590839-5601411
13	*PmWD4013*	XP_008219585	1831	204174.56	5.42	37.42	2	6808229-6822498
14	*PmWD4014*	XP_008219867	924	101890.34	6.37	54.75	2	8422560-8429637
15	*PmWD4015*	XP_008220347	425	46906.19	6.89	43.05	2	10248086-10251239
16	*PmWD4016*	XP_008220285	558	62394.27	4.89	36.25	2	10555467-10559238
17	*PmWD4017*	XP_008220685	673	74506.94	6.41	48.31	2	12444740-12449903
18	*PmWD4018*	XP_008220972	326	35873.52	6.84	28.81	2	13845587-13847622
19	*PmWD4019*	XP_008221109	562	62913.43	5.78	50.23	2	14624323-14627912
20	*PmWD4020*	XP_008221114	803	87248.38	7.86	30.41	2	14652774-14660890
21	*PmWD4021*	XP_008221806	590	66945.99	6.02	42.59	2	18330280-18333564
22	*PmWD4022*	XP_008221944	436	47399.85	8.99	48.11	2	19245606-19246916
23	*PmWD4023*	XP_008222478	448	50453.61	7.21	48.94	2	23746059-23747405
24	*PmWD4024*	XP_008223206	480	52958.35	9.18	48.16	2	31657152-31662512
25	*PmWD4025*	XP_016647785	1761	198372.16	6.42	49.69	2	32308287-32321062
26	*PmWD4026*	XP_008223841	344	37875.20	6.89	34.72	2	36681141-36682175
27	*PmWD4027*	XP_008224052	768	83766.49	6.33	52.47	2	38353576-38359380
28	*PmWD4028*	XP_008224605	776	84044.69	6.09	44.02	2	42054622-42063077
29	*PmWD4029*	XP_008224985	850	93233.31	7.59	50.51	3	1752654-1760017
30	*PmWD4030*	XP_008225331	279	30818.08	5.97	55.33	3	3344967-3345930
31	*PmWD4031*	XP_008225323	372	40922.78	7.43	36.31	3	3592102-3595287
32	*PmWD4032*	XP_008225587	523	57942.41	6.07	46.33	3	4882114-4885983
33	*PmWD4033*	XP_008225617	1195	134347.47	7.60	31.72	3	5020656-5025523
34	*PmWD4034*	XP_008226626	905	101201.84	5.59	55.22	3	11861155-11866300
35	*PmWD4035*	XP_008226666	872	95750.40	6.64	48.84	3	12126631-12131958
36	*PmWD4036*	XP_016649058	1387	153146.20	5.81	44.17	4	306730-315398

37	*PmWD4037*	XP_008228475	414	44902.13	8.14	38.13	4	4396763-4398882
38	*PmWD4038*	XP_016649234	269	30746.79	5.57	27.28	4	4947383-4949304
39	*PmWD4039*	XP_008228579	920	104100.63	4.95	33.56	4	4951898-4963139
40	*PmWD4040*	XP_008228644	893	99515.16	6.20	33.20	4	5604633-5610504
41	*PmWD4041*	XP_008229590	578	64475.45	5.98	45.55	4	15376411-15380722
42	*PmWD4042*	XP_008230548	511	56514.05	8.18	45.92	4	20725224-20728936
43	*PmWD4043*	XP_008230874	426	46723.72	6.03	35.64	4	22724411-22728053
44	*PmWD4044*	XP_008231471	474	52818.19	8.74	32.43	5	2013064-2020784
45	*PmWD4045*	XP_008232410	383	41457.88	6.16	28.24	5	12896479-12898113
46	*PmWD4046*	XP_008233498	799	86711.95	6.40	55.61	5	19935281-19939920
47	*PmWD4047*	XP_008234564	875	96195.99	6.37	40.86	5	24779991-24786140
48	*PmWD4048*	XP_008234637	342	38541.30	5.05	52.56	5	25093939-25094967
49	*PmWD4049*	XP_008234949	475	52968.44	8.07	41.79	6	255840-259954
50	*PmWD4050*	XP_008235203	1647	184806.32	6.34	47.10	6	1250721-1264120
51	*PmWD4051*	XP_008235296	317	34790.30	8.62	28.58	6	1655174-1656762
52	*PmWD4052*	XP_008235396	432	47016.42	5.57	34.19	6	2210656-2213743
53	*PmWD4053*	XP_008235489	779	85211.51	6.78	53.23	6	2585993-2591248
54	*PmWD4054*	XP_016650868	799	86920.71	6.00	48.67	6	3117178-3122398
55	*PmWD4055*	XP_008235576	629	69389.96	5.98	40.92	6	3195747-3198621
56	*PmWD4056*	XP_008235875	1384	148757.52	6.87	52.13	6	4755411-4766210
57	*PmWD4057*	XP_008236022	405	45696.73	4.58	36.80	6	5406535-5409070
58	*PmWD4058*	XP_008236777	758	83903.68	5.52	37.07	6	9984575-9996249
59	*PmWD4059*	XP_008236879	377	40907.99	7.14	29.62	6	10974125-10980067
60	*PmWD4060*	XP_008236956	501	55805.09	9.13	52.32	6	11626071-11629615
61	*PmWD4061*	XP_008237351	1031	112846.51	7.56	40.35	6	15916139-15924274
62	*PmWD4062*	XP_008238207	666	75377.61	6.76	48.50	7	2889792-2896366
63	*PmWD4063*	XP_008239070	1076	120179.30	5.72	45.75	7	10325897-10331641
64	*PmWD4064*	XP_008239374	508	55828.98	5.90	40.62	7	12238630-12241249
65	*PmWD4065*	XP_016651149	394	42370.62	7.22	38.11	7	12669460-12672266
66	*PmWD4066*	XP_008240184	327	36250.28	6.75	31.68	7	16057328-16059075
67	*PmWD4067*	XP_008240283	1218	136907.52	6.69	33.48	7	16523322-16528054
68	*PmWD4068*	XP_008240787	535	59182.73	6.28	40.38	8	4919046-4922313
69	*PmWD4069*	XP_008241369	956	106126.55	6.81	39.37	8	9169264-9175477
70	*PmWD4070*	XP_008242602	877	100133.24	6.60	39.62	8	15026969-15033749
71	*PmWD4071*	XP_008242692	874	96114.21	6.80	52.33	8	15356288-15363638
72	*PmWD4072*	XP_008242709	347	37838.35	5.55	44.10	8	15431608-15434309
73	*PmWD4073*	XP_008245266	523	56688.18	6.28	29.09	Un	28274-35262
74	*PmWD4074*	XP_008243924	546	58364.14	7.28	41.74	Un	81696-86665
75	*PmWD4075*	XP_008244008	1207	133326.65	5.88	43.57	Un	510628-521653
76	*PmWD4076*	XP_008244270	683	77844.33	5.69	43.19	Un	588268-592847
77	*PmWD4077*	XP_008235353	3612	402874.69	5.77	42.77	Un	1986938-2004042

Table 2(on next page)

The details of the putative WD40 domain sequences t;mso-f.

The amino acid sequences shown in Table 2 are the core sequence of PmWD40 proteins.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27106v1 | CC BY 4.0 Open Access | rec: 10 Aug 2018, publ: 10 Aug 2018

Table 2 The details of the putative WD40 domain sequences

Motif number	Length (aa)	Sequence
1	15	JASGSDDGTVRVWDL
2	21	CIAVLKGHTGAVSSVAFSPDG
3	15	YIASGSDDRTVKJWD
5	15	LVSASDDGSVKVWDL
9	15	LLASGSSDGTVRVWD