Article

Genome-Wide Identification and Characterization of WD40 Protein Genes in the Silkworm, *Bombyx mori*

Songzhen He, Xiaoling Tong, Minjin Han, Hai Hu and Fangyin Dai *

State Key Laboratory of Silkworm Genome Biology, Key Laboratory of Sericultural Biology and Genetic Breeding, Ministry of Agriculture, Southwest University, Chongqing 400715, China; szhe@swu.edu.cn (S.H.); xltong@swu.edu.cn (X.T.); minjinhan@126.com (M.H.); huhaiswu@163.com (H.H.)
* Correspondence: fydai@swu.edu.cn; Tel.: +86-23-6825-0793

Abstract: WD40 proteins are scaffolding molecules in protein-protein interactions and play crucial roles in fundamental biological processes. Genome-wide characterization of WD40 proteins in animals has been conducted solely in humans. We retrieved 172 WD40 protein genes in silkworm (*BmWD40s*) and identified these genes in 7 other insects, 9 vertebrates and 5 nematodes. Comparative analysis revealed that the WD40 protein gene family underwent lineage-specific expansions during animal evolution, but did not undergo significant expansion during insect evolution. The *BmWD40s* were categorized into five clusters and 12 classes according to the phylogenetic classification and their domain architectures, respectively. Sequence analyses indicated that tandem and segmental duplication played minor roles in producing the current number of *BmWD40s*, and domain recombination events of multi-domain *BmWD40s* might have occurred mainly after gene duplication events. Gene Ontology (GO) analysis revealed that a higher proportion of *BmWD40s* was involved in processes, such as binding, transcription-regulation and cellular component biogenesis, compared to all silkworm genes annotated in GO. Microarray-based analysis demonstrated that many *BmWD40s* had tissue-specific expression and exhibited high and/or sex-related expression during metamorphosis. These findings contribute to a better understanding of the evolution of the animal WD40 protein family and assist the study of the functions of *BmWD40s*.

Keywords: WD40 proteins; identification; characterization; multiple animals; silkworm

1. Introduction

WD40 proteins, also known as WD-repeat proteins or WD40-containing proteins, comprise one of the largest protein families in eukaryotes, but they are rare in prokaryotes [1,2]. WD40 protein was first identified and analyzed as a G protein β-subunit (bovine transducin β-subunit), containing repeated units of ~43 residues with conserved Glycine Histidine (GH) and tryptophan aspartate (WD) motifs [3]. Proteins with these repeated residue units (WD40 repeats or blades) were defined as WD40 proteins [1,4]. Generally, each WD40 repeat contains 40–60 residues with a WD motif and is able to fold into four anti-parallel β-strands [5]. A total of seven repeats constitute a canonical WD40 domain, which folds into a β-propeller structure [6–8]. Usually, a WD40 protein would contain 4–9 blades in one β-propeller [9].

WD40 proteins function as scaffolds for protein-protein interactions or provide platforms to recruit diverse molecules that form functional complexes [1,10,11]. This helps the proteins perform their bioactivities. WD40 proteins play vital roles in many basic biological processes, such as DNA replication, transcriptional regulation and damage responses [12–14], RNA processing and modification [15–18], protein degradation [19], histone recognition and modification [20,21], apoptosis [22–24] and signal transduction [25–27]. Many WD40 proteins are implicated in diseases. For instance, WDR36 was found to be mutated in primary open-angle glaucoma [28]; WDR45 is

implicated in neurodegeneration [29]; WDR62 is associated with human microcephaly [30,31]; and FBXW7 and TLE1 are involved in tumors [32–34]. BmPLA2 of silkworms influences fat-body metabolism. Knockdown of the *BmPLA2* gene resulted in reduced development and death [35].

Genome-wide systematic characterization of this protein family has been conducted in many species, including rice, foxtail millet, cucumber, *Arabidopsis* and humans [2,36–40]. In animals, this family has been thoroughly studied in humans, but little is known about WD40 genes in other species.

The silkworm (*Bombyx mori*) is an economically-important insect for silk production. It is also an ideal model species for microbiology, physiology, genetics and functional genomics analysis [41]. In the present study, we carried out a genome-wide identification of the WD40 protein gene family in multiple animals and performed gene expression pattern and sequence analysis of silkworm WD40 proteins. Our findings could facilitate the understanding of the WD40 protein family and provide important clues for subsequent research on its functions.

2. Results and Discussion

2.1. Identification of WD40 Proteins

To identify WD40 proteins in *Bombyx mori*, the hmmsearch program in the HMMER3.0 package (http://hmmer.org/) [42] was applied to search in the silkworm protein database using the Hidden Markov Model (HMM) profile for the WD repeat (PF00400). Redundant sequences were discarded manually. The presence of WD40 repeats was confirmed by CDD (Conserved Domain Database, http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [43], SMART (Simple Modular Architecture Research Tool, http://smart.embl-heidelberg.de/) [44] and WDSP (WD40-repeat Protein Structure Predictor, http://wu.scbb.pkusz.edu.cn/wdsp/predictor.jsp) [45] search-based domain analysis. As a result, 172 non-redundant silkworm WD40 proteins (*Bm*WD40) were retrieved, and each protein represented the typical product of a gene. The 172 *BmWD40* genes were numbered from *BmWD001–BmWD172* based on their order on chromosomes (Table S1).

In silico analysis showed that the *Bm*WD40 proteins varied greatly in length and physicochemical properties. The lengths of *Bm*WD40s ranged from 99 residues–3007 residues. The molecular weights varied from 11.061 kDa– 30.895 kDa. Isoelectric points ranged from 4.44–9.76. The number of WD40 repeats varied from 2–21 (Table S1), suggesting the existence of multiple, incomplete or atypical WD40 domains in some proteins. A total of 64 of the 172 *Bm*WD40s were DWD (Damaged DNA binding WD40) proteins (Table S1). These proteins are a subset of WD40 proteins containing a conserved DWD box in the WD40 repeats. They modulate numerous biological processes by ubiquitin-dependent proteolysis [46,47]. Basic information of these genes is summarized in Table S1.

WD40 protein genes in 7 other insect species, as well as 9 vertebrates and 5 nematodes were identified using the same approach, and the WD40 protein genes are listed in Table S2. The total numbers of WD40 protein genes present in the genomes of silkworm and seven other insect species were similar (Figure 1 and Table S2), indicating that the WD40 protein gene family did not undergo significant expansion during insect evolution. However, the number of WD40 protein genes in these surveyed insects (mean ~178) is larger than the number in nematodes (mean ~135) and smaller than that in vertebrates (mean ~256) (Figure 1 and Table S2). This suggests that the WD40 protein genes have undergone a lineage-specific expansion in the more highly-evolved groups of animals compared with the lower animals. Since WD40 proteins are pivotal regulatory proteins involved in a variety of physiological processes, we speculate that lineage-specific expansion of the animal WD40 gene family evolved to meet the more complex functional requirements of highly-evolved animals.
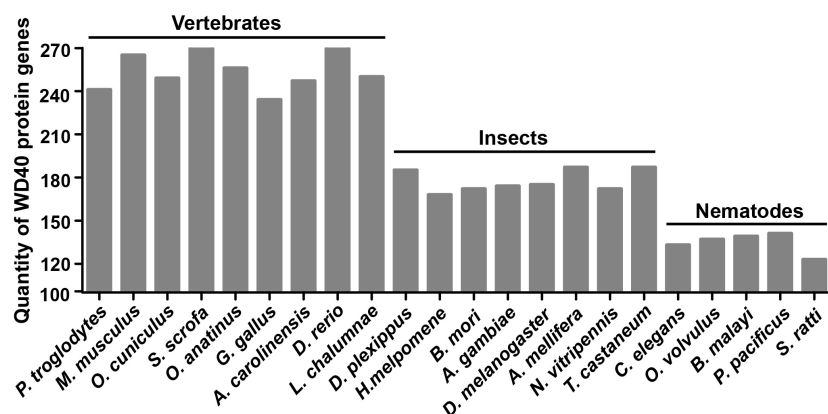
Figure 1. The number of WD40 protein genes in different species. WD40 protein genes in 8 insects, 9 vertebrates and 5 nematodes were identified. The total numbers of WD40 protein genes present in the genomes of the eight insects were similar. The number of WD40 protein genes in these insects (mean ~178) is larger than that in nematodes (mean ~135) and smaller than that in vertebrates (mean ~256).

2.2. Chromosomal Distribution and Gene Structure of the BmWD40 Family

To determine the distribution and contexts of the *BmWD40* genes family on silkworm chromosomes, a "*BmWD40s* distribution map" was constructed based on the chromosomal coordinates extracted from the silkworm genome database (http://sgp.dna.affrc.go.jp/KAIKObase/). We found that the distribution of *BmWD40* genes is widespread and uneven on the chromosomes. As shown in the "*BmWD40s* distribution map", 160 *Bm*WD40s were assigned to 27 silkworm chromosomes (excluding chromosome 16) (Figure 2 and Table S1), while the other 12 *BmWD40s* did not have location information on chromosomes. Chromosome 17 has the largest number of *BmWD40* genes (14 genes; the ratio of gene quantity to chromosome size is 0.76), followed by chromosomes 4 (12 genes, ratio = 0.57), while lesser numbers of genes were found on chromosome 7 (two genes, ratio = 0.06) (Figure 2). The distribution of *BmWD40* genes exhibits a positional accumulation pattern on certain chromosomes. Relatively high densities of *BmWD40* genes were observed at the top of chromosomes 3, 13 and 17, as well as at the bottom of chromosomes 4, 19 and 24. Relatively low densities of *BmWD40* genes were also found in some other chromosomal regions, including the top of chromosomes 1, 2, 7, 8, 20, 24 and 27 and at the bottom of chromosomes 8, 11 and 26 (Figure 2). Previous studies have shown that the distributions of the WD40 proteins genes on chromosomes in plants and human were also extensive and uneven [37–40]. These findings suggest that the pervasive-uneven distribution is a universal characteristic of the WD40 gene family in plants and animals.

Genome duplication events, including tandem duplication and segmental duplication, are thought to contribute in the evolution and expansion of gene families [48,49]. Previous studies indicated that the expansion and distribution pattern of the WD40 gene family in *Arabidopsis* [38], rice [37] and foxtail millet [39] were due to both tandem duplication and segmental genome duplication events. Zou et al. speculated that this distribution pattern of WD40 proteins genes in humans was related to different degrees of segmental duplication on different chromosomes [40]. However, cucumber studies revealed that tandem duplication and segmental duplication did not play critical roles in the expansion of cucumber WD40 genes [38], due to the absence of the recent whole-genome duplication events and tandem duplications [50]. There is no evidence of whole genome duplication events in the silkworm. In this study, 12 (~7%) *BmWD40* genes were identified to be tandem repeats (tandem duplication of genes) based on sequence alignment and chromosomal distribution (Figure 2), and only two *BmWD40* genes (*BmWD005* and *BmWD101*) were related to potential segmental duplication events. These data suggest that tandem and segmental duplication played a minor role in the number of *BmWD40* genes. Additional study is needed to determine the mechanism behind the expansion and distribution pattern of the *BmWD40* gene family.
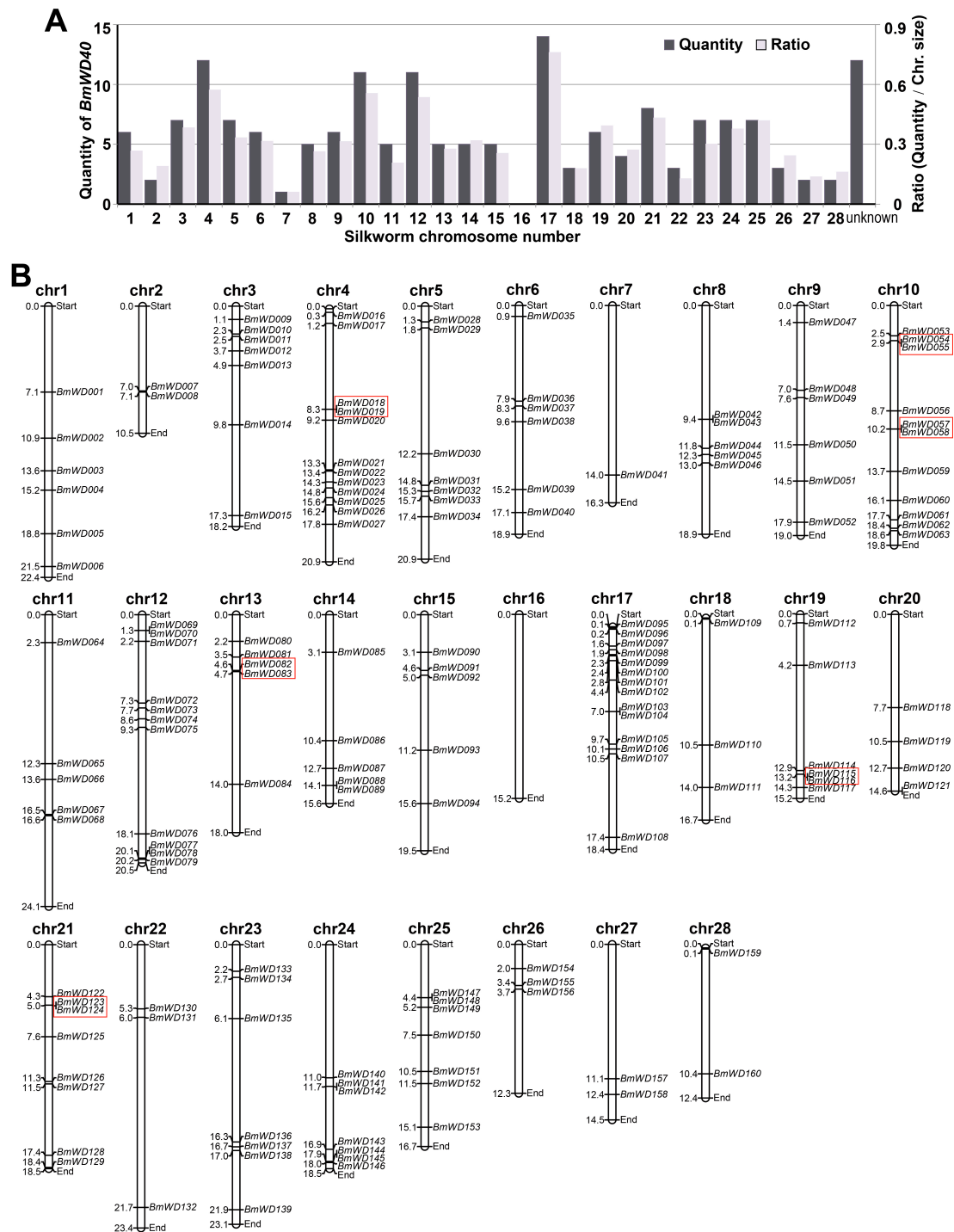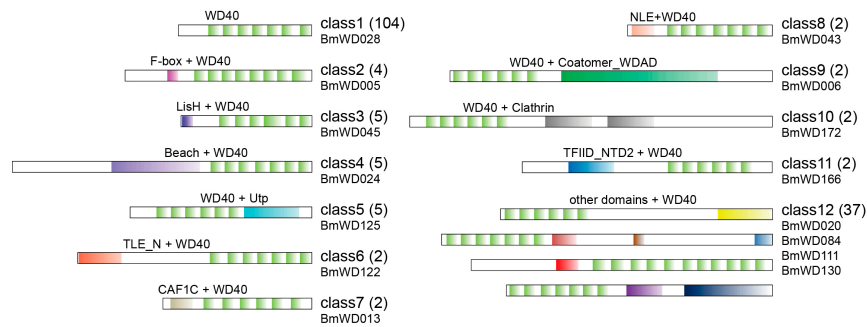
Figure 2. Distribution of the *BmWD40* genes on chromosomes. (**A**) Number of *BmWD40s* on each silkworm chromosome and the ratio of gene number to chromosome size to show the distribution abundance; (**B**) "*BmWD40s* distribution map" on silkworm chromosomes. Tandem duplicated genes are marked by red boxes. The chr represents chromosome. Chromosomal distances are given in Mb (Mega base).

We also investigated gene structures and found that the intron sequences within the Open Reading Frame (ORF) of the *BmWD40* genes varied greatly and the number of introns ranged from 0–56. The longest *BmWD40* gene was *BmWD11* with about a 103.6-kb genomic sequence, while the shortest one was only 417 bp (*BmWD064*) (Table S1).

2.3. Phylogenetic Classification and Domain Architectures of BmWD40s

To study the evolutionary relationships of the WD40 family in silkworm, 172 identified *Bm*WD40 protein sequences were used to construct a phylogenetic tree according to the Neighbor-Joining (NJ) method [51]. As shown in Figure 3, the *Bm*WD40s were categorized into five main distinct clusters (Clusters I–V) containing 27, 18, 30, 25 and 72 proteins, respectively. There were three sub-clusters (Clusters Va, Vb and Vc) in Cluster V, suggesting the complex and differentiated structure of *Bm*WD40s.

Figure 3. Phylogenetic classification of *Bm*WD40 proteins. MEGA6 (Molecular Evolutionary Genetics Analysis version 6.0, http://www.megasoftware.net/) was used to construct the phylogenetic tree using the neighbor-joining method with 1000 bootstrap replicates. The tree was divided into five main distinct clusters (Cluster I–V), and Cluster V was divided into three sub-clusters (Clusters Va, Vb and Vc).

To get more details of the complex protein domain structures, the domain architectures of the *Bm*WD40s were annotated. The *Bm*WD40s were categorized into 12 classes based on their domain compositions (Figure 4 and Table S1). A total of 104 *Bm*WD40s with only the WD40 domain were grouped into Class 1. The other *Bm*WD40s with additional functional domains were categorized into Classes 2–12 as follows: 4 *Bm*WD40s containing the F-BOX and U-BOX domain were grouped into Class 2; 5 *Bm*WD40s comprising the LisH domain were identified as Class 3; 5 *Bm*WD40s with the Beige/BEACH domain were categorized into Class 4; 5 *Bm*WD40s with the Utp domain were identified as Class 5; *Bm*WD40s with the TLE_N domain, subunit C of the CAF1 complex domain, the NLE (NUC) domain, the coatomer WD-associated region, the clathrin domain and the TFIID_NTD2 domain were categorized into Classes 6–11, respectively. For simplicity, the remaining 37 *Bm*WD40s with only one member for each domain-architecture were put into Class 12 (Figure 4 and Table S1). Apart from the domain architectures conserved between animal and plants, we also found potential animal-specific architectures according to previous reports [40], including Class 6 (TLE_N + WD40) and

four architectures in Class 12 (Dynein_IC2 + WD40, striatin + WD40, HELP + WD40, NACHT + WD40) (Figure 4 and Table S1).


Figure 4. Domain architectures of *Bm*WD40 proteins. The protein structure is identified by WDSP, SMART and CDD. This organizes the *Bm*WD40s into 12 classes. The WD40 repeats are colored in light green, and other domains are filled in with other colors separately. The number of members in each class is shown in parentheses. The name of the displayed protein is given below.

Domain recombination events and gene duplication events might have happened in the evolution of multi-domain protein gene families. Determining the order of these two events will promote our understanding of the evolutionary history of the gene families. Accordingly, we used the amino acid sequences of WD40 domains from the identified *Bm*WD40 proteins to construct a phylogenetic tree. Figure 5 shows that many members of Class 1 (with only the WD40 domain) are separately grouped with the members of other classes (with multi-domains) in the phylogenetic tree, and only a few multi-domain *Bm*WD40 proteins from the same classes (Class 2, 5–7 and 11) group together into one sub-clade. These results suggest that the multi-domain *Bm*WD40 protein genes may have evolved from the genes with only the WD40 domain within the same sub-clade. The domain recombination events might have happened mainly after the gene duplication events in the evolution of the multi-domain *Bm*WD40 protein genes.

2.4. Gene Ontology and KEGG Analysis of the BmWD40s

To survey the functions of the *BmWD40* genes, Gene Ontology (GO) annotation was performed using Blast2GO software and also obtained from two silkworm genome databases (http://www.silkdb.org and http://sgp.dna.affrc.go.jp/KAIKObase/). Compared to a wide range of silkworm genes annotated in GO, a higher proportion of *BmWD40s* played roles in binding, transcription regulator activity, anatomical structure formation, cellular component organization and cellular component biogenesis (Figure 6). Consistent with the essential protein roles, the analysis revealed that most *BmWD40s* (165, ~95.9%) participated in binding (Figure 6, Table S3). The results also showed that the *BmWD40s* were involved in diverse biological processes and predominantly participated in cellular metabolic processes (42, ~24.4%), followed by primary metabolic processes (39, ~22.7%) and macromolecule metabolic processes (35, ~20.3%) (Table S3).

To ascertain whether the *BmWD40s* categorized into the five distinct clusters (Cluster I–V) of the phylogenetic tree had functional preference, respectively, GO analysis was also performed with the genes of the five clusters. The results revealed that the *BmWD40s* of the five distinct clusters had a variety of biological functions and were involved in various physiological processes with a degree of functional specificity. For example, the gene (*BmWD015*) participating in multicellular organismal processes and developmental processes belonged to Cluster I, and the *BmWD40s* involved in structural molecule activity (*BmWD006* and *BmWD155*) and molecular transducer activity (*BmWD014*) are only found in Cluster III. The genes implicated in transporter activity (*BmWD057* and *BmWD162*), antioxidant activity (*BmWD148*) and transcription regulator activity (*BmWD053, BmWD093* and *BmWD166*) were all grouped into Cluster V (Figure S1).
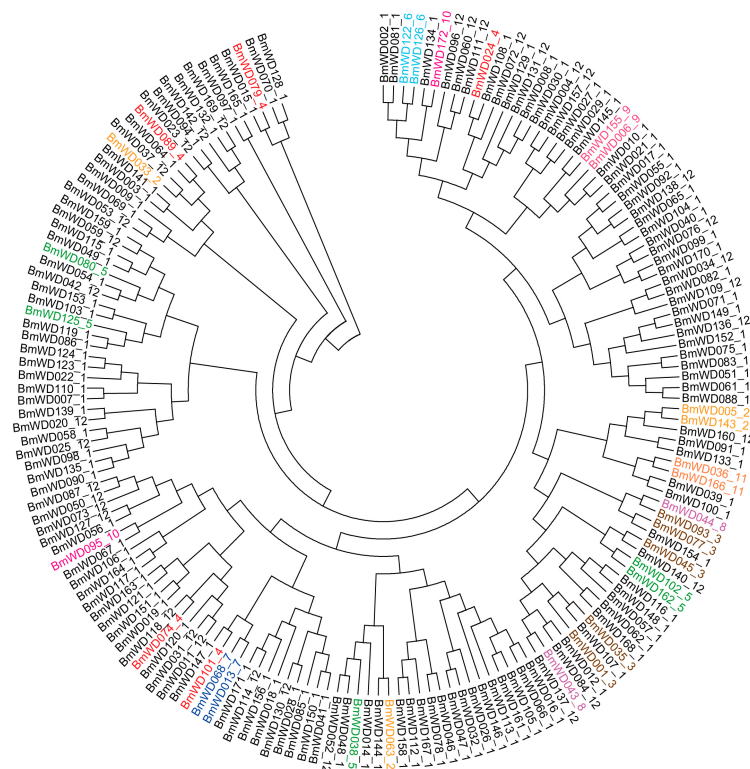
Figure 5. WD40 domain sequence-based phylogenetic tree of *Bm*WD40 proteins. MEGA6 was used to construct the phylogenetic tree using the neighbor-joining method with 1000 bootstrap replicates. Multi-domain *Bm*WD40 proteins from each class (i.e., from Class 2–Class 11) with multiple members for each domain-architecture are indicated by the same colors, respectively.
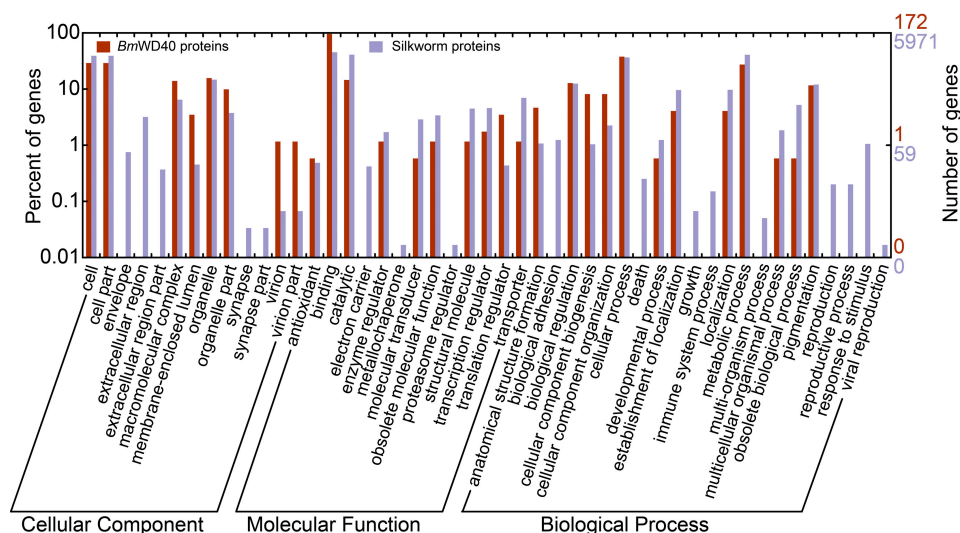
Figure 6. Gene Ontology (GO) categories of *BmWD40* genes and genes with GO annotations obtained from SilkDB. This analysis was visualized with WEGO (Web Gene Ontology Annotation Plotting, http://wego.genomics.org.cn/). Details are provided in Supplementary File Table S3.

The *BmWD40* genes were also classified into reference pathways in the Kyoto Encyclopedia of Genes and Genomes database (KEGG, http://www.kegg.jp/) [52]. The results showed that the *BmWD40* genes were also involved in numerous biological pathways (Figure 7, Table S4). A total of seven KEGG pathways, including RNA transport (bmor03013), ribosome

biogenesis in eukaryotes (bmor03008), ubiquitin-mediated proteolysis (bmor04120), spliceosome (bmor03040), mRNA surveillance pathway (bmor03015), regulation of autophagy (bmor04140) and phototransduction (bmor04745), were significantly enriched (corrected *p*-value less than 0.05) for the *BmWD40* genes.
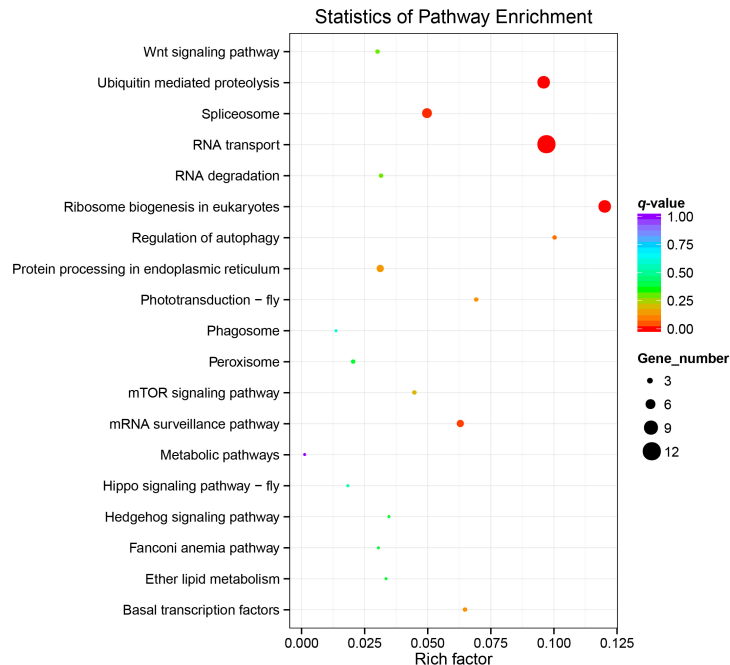
Figure 7. Scatterplot of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. The "Rich factor" is the ratio of the *BmWD40* gene number to the total annotated gene number in a certain pathway. The size and color of the dots represent the range of the gene number and the *q*-value, respectively. Details are provided in Supplementary File Table S4.

2.5. Spatial and Temporal Expression Profile of the BmWD40s

The gene expression profile can also provide a strong indication about the function and biological activity of a gene. Highly- and ubiquitously-expressed WD40 protein genes might originate early in evolution and regulate the transcription of a broad set of genes. In contrast, other genes of this family could be involved in diverse physiological processes in tissue- or stage-specific manners [39,40].

To view the spatial expression profile of *BmWD40* genes, we used the microarray gene expression dataset from SilkDB, which includes normalized gene expression levels across 10 silkworm larval tissues. Then, 130 genes were detected as expressed in at least one tissue. A total of 22 genes was expressed in all investigated tissues, and these were involved in basic molecular functions including transcription regulator, translation regulator, catalytic activity, structural molecule activity and binding (Figure S2, Table S5).

A large number of *BmWD40s* showed tissue-specific gene expression features. Most *BmWD40s* showed high expression in the gonads, and a considerable proportion of these genes exhibited testis-specific high expression (red box in Figure 8); some of them were also highly expressed in silk glands (blue box in Figure 8). *BmWD012*, *BmWD065* and *BmWD086* were specifically highly expressed in the Posterior Silk Gland (PSG); *BmWD093* and *BmWD136* were highly expressed in the Anterior/Median Silk Gland (A/MSG). Furthermore, *BmWD041*, *BmWD042* and *BmWD060* exhibited Malpighian-tubule-specific high expression and showed higher expression in males. *BmWD083* and *BmWD098* were mainly expressed in the midgut. *BmWD087* and *BmWD151* were highly expressed in the integument (Figure 8). These diverse expression features indicate that certain *Bm*WD40 proteins participate in special physiological functions in specific tissues.
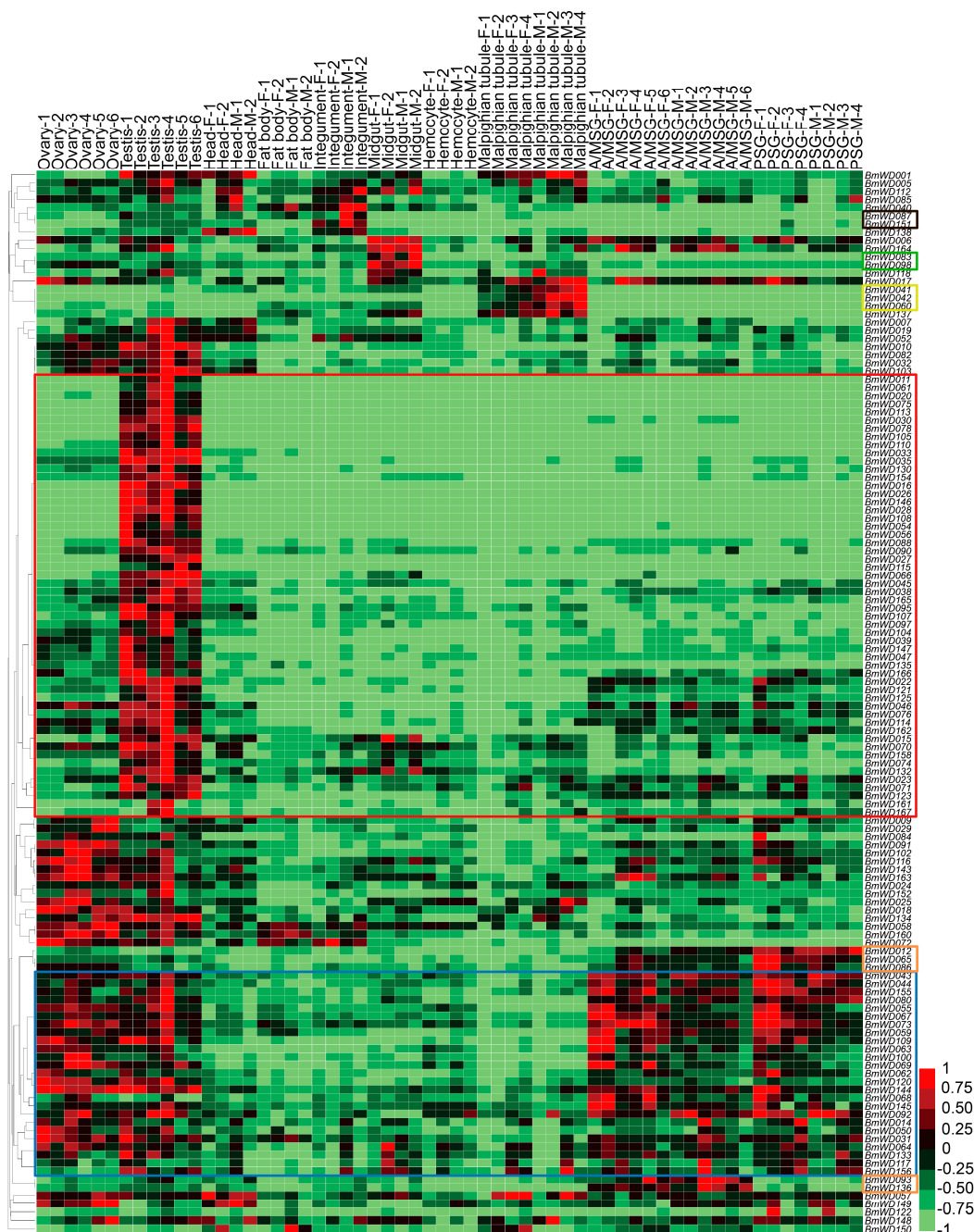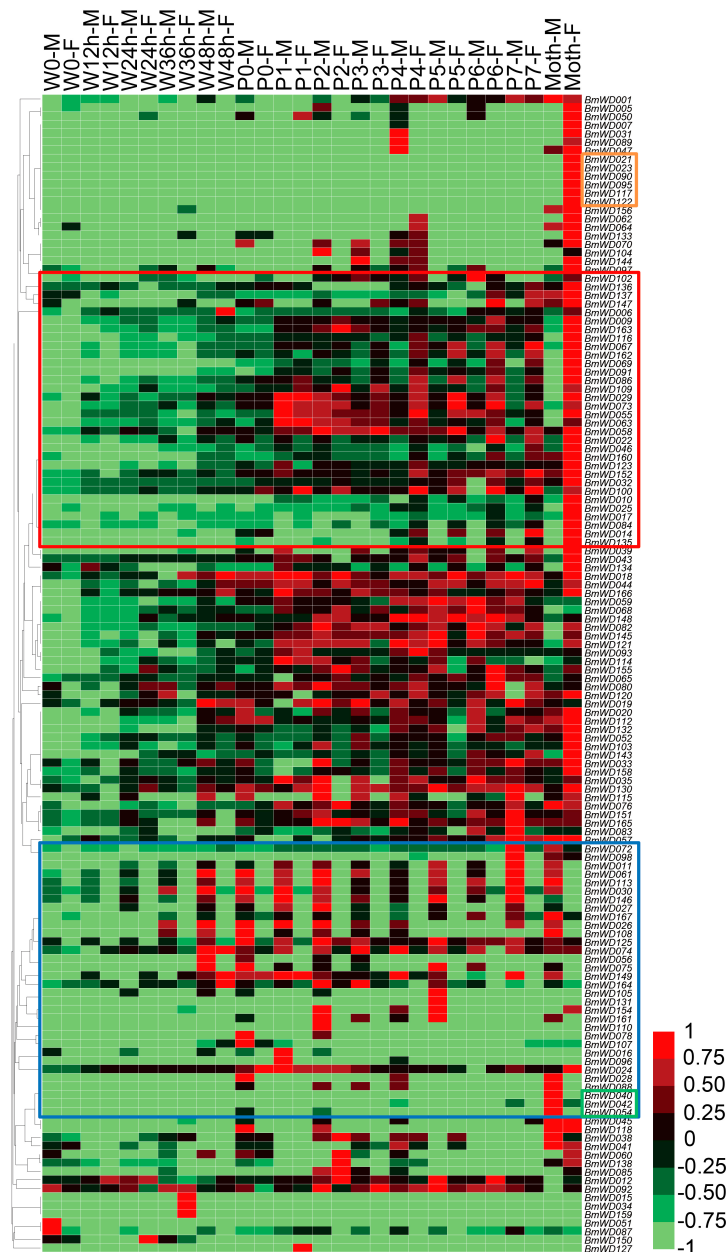
Figure 8. Microarray-based gene expression profiling of *BmWD40* genes in multiple tissues of silkworm larvae. The microarray gene expression dataset in multiple larval tissues of the silkworm on Day 3 of the fifth instar was downloaded from the SilkDB database. Repeated samples of each tissue are indicated with different numerals. A/MSG: Anterior/Median Silk Gland, PSG: Posterior Silk Gland, F: Female, M: Male. Red box indicates highly-expressed testis-specific genes. Blue box indicates the genes with high expression in gonad and silk glands. Orange boxes indicate the genes with silk-gland-specific high expression. Yellow box indicates genes that showed Malpighian-tubule-specific high expression. Green box indicates genes that were mainly expressed in the midgut. Black box indicates genes that were highly expressed in the integument.

We used the microarray gene expression data from late larval to adult stages of silkworm to study the temporal expression profiles of the *BmWD40* genes during metamorphosis. A total of 136 *BmWD40s* was found to be expressed during silkworm metamorphosis (Table S6). Most *BmWD40s* exhibited

high expression in the pupal and adult stages (Figure 9), indicating that these genes were involved in silkworm metamorphosis. The hierarchical clustering graph showed that the expression pattern of many genes displayed sexual dimorphism (Figure 9). Most of the genes in the box outlined in red were specifically highly expressed in female, whereas the expression of most genes in the box outlined in blue were male-specific. Moreover, *BmWD040*, *BmWD042* and *BmWD054* exhibited male-moth-specific high expression and *BmWD021*, *BmWD023*, *BmWD090*, *BmWD095*, *BmWD117* and *BmWD122* exhibited female-moth-specific expression.

Figure 9. Microarray-based gene expression profiling of *BmWD40* genes during silkworm metamorphosis. One hundred thirty six *BmWD40* genes were detected to be expressed during silkworm metamorphosis. M: Male, F: Female, W: Wandering (W0: beginning of Wandering for spinning, W12 h: 12 h after Wandering), P: Pupation (P0: beginning of Pupation, P1: one day after Pupation), h: hour. Most genes in the red box were specifically highly expressed in female, whereas most genes in the blue box were male-specific. Genes in the green box exhibited male-moth-specific high expression. Genes in the orange box showed female-moth-specific expression.

We found that many genes, such as *BmWD011*, *BmWD061*, *BmWD075* and *BmWD113*, were expressed specifically in the testis of larvae and in male individuals during metamorphosis (Table S6, Figures 8 and 9), indicating that these genes may be required for testicular development or spermatogenesis. Among these, four genes, *BmWD016*, *BmWD105*, *BmWD113* and *BmWD161*, which belonged to Class 1 of the *BmWD40* family, encoded dynein intermediate chain proteins. The *dynein Intermediate Chain* (*IC*) genes are associated with sperm ultrastructure, and gene defects may bring about anomalies in the sperm flagella [53,54]. Accordingly, we propose that the spatiotemporally-specific expression of the four *IC* genes may be necessary for silkworm spermatogenesis or the formation and maintenance of spermatozoa flagella.

Our findings are also supported by functional studies of orthologs in other species. *WDR63*, *WDR78* [55] and *IFT140* [56], which are orthologs of the three highly-expressed testis-specific genes *BmWD027*, *BmWD054* and *BmWD104*, respectively, are also particularly highly expressed in the testis of mice, and they are related to sperm formation, sperm function and male fertility in mice. Knockout of *Pex7* [57] (ortholog of *BmWD132*, which is particularly highly expressed in the testis of silkworm) in mice leads to defects in plasmalogen biosynthesis and results in very-long-chain fatty acids-induced (VLCFA-induced) degeneration and apoptosis of spermatocytes. Furthermore, *WDR45* [29], *WDR47* [58], *WDR48* [59] and *GNB5* [60,61] (orthologs of *BmWD103*, *BmWD001*, *BmWD160* and *BmWD112*, which have relatively high expression in the heads of silkworm) are related to neurogenesis and brain development in humans, mice and nematodes. *FBXW7* [33] (ortholog of *BmWD005*, which is highly expressed in the midgut of silkworm) is involved in colorectal cancer. *WDR37* [62] (ortholog of *BmWD150*, which is highly expressed in the Malpighian tubules, an insect organ functionally homologous to mammalian kidneys) is associated with kidney function and chronic kidney disease.

2.6. Hotspot Residues of Protein-Protein Interactions

WD40 proteins can participate in protein-protein interactions with the help of certain major residues (hotspot residues) on their surfaces [9,11,63]. Therefore, we predicted potential hotspot residues on the *Bm*WD40 proteins using WDSP (http://wu.scbb.pkusz.edu.cn/wdsp/predictor.jsp). The results are shown in Table S1. To investigate the potential protein that interacts with WD40, we also analyzed protein-protein interaction networks based on the STRING database (Search Tool for Recurring Instances of Neighbouring Genes, http://string-db.org/), and proteins with potential interaction with the *Bm*WD40 proteins are listed in Table S7.

3. Materials and Methods

3.1. Genome-Wide Identification of WD40 Proteins

The whole-genome protein sequences of silkworm were downloaded from two silkworm genome databases, SilkDB (http://www.silkdb.org) and KAIKObase (http://sgp.dna.affrc.go.jp/KAIKObase/). The HMM profiles for the WD repeat (PF00400) were downloaded from the Pfam database (http://pfam.sanger.ac.uk/) [64]. The hmmsearch program (E-value $\leq 1 \times 10^{-1}$, score ≥ 0) of the HMMER3.0 software (http://hmmer.org/) was used for the identification of WD40 proteins, and redundant sequences were discarded manually. The protein sequences of other species were downloaded from Ensembl (http://www.ensembl.org), and the same approach was used to identify WD40 proteins in 7 other insects: *Danaus plexippus* (Lepidoptera), *Heliconius melpomene* (Lepidoptera), *Anopheles gambiae* (Diptera), *Drosophila melanogaster* (Diptera), *Apis mellifera* (Hymenoptera), *Nasonia vitripennis* (Hymenoptera) and *Tribolium castaneum* (Coleoptera); as well as 9 vertebrates, including *Pan troglodytes* (Primates), *Mus musculus* (Rodentia), *Oryctolagus cuniculus* (Lagomorpha), *Sus scrofa* (Artiodactyla), *Ornithorhynchus anatinus* (Monotremata), *Gallus gallus* (Aves), *Anolis carolinensis* (Reptilia), *Danio rerio* (Cypriniformes) and *Latimeria chalumnae* (Coelacanthiformes);

and 5 nematodes, including *Caenorhabditis elegans*, *Onchocerca volvulus*, *Brugia malayi*, *Pristionchus pacificus* and *Strongyloides ratti*.

The chromosomal locations, gene lengths, probes and predicted protein lengths were obtained from SilkDB and KAIKObase. Theoretical molecular weights and isoelectric point values were calculated online at http://web.expasy.org/protparam/. Three online programs, CDD (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [43], WDSP (http://wu.scbb.pkusz.edu.cn/wdsp/predictor.jsp) [45] and SMART (http://smart.embl-heidelberg.de/) [44], were used to confirm the WD40 repeats in the predicted proteins. The predicted WD40 protein genes were manually annotated by online BLASTP searches against the non-redundant protein database in NCBI (https://www.ncbi.nlm.nih.gov/).

3.2. Chromosomal Distribution and Gene Structure

Based on the chromosomal coordinates extracted from the KAIKObase, the WD40 genes were plotted onto chromosomes according to their order of physical position, and the "*BmWD40* distribution map" was displayed by MapChart [65]. The adjacent genes, which belonged to the same subfamily and were located within the same or neighboring genomic region, were regarded as tandem repeats [66]. To identify the *BmWD40s* related to potential segmental duplication events, we checked all the genes in the silkworm segmental duplication dataset of Zhao's previous report [67], which contains the composition of genes in the silkworm segmental duplication regions. The exon-intron structures of the genes were determined by comparing protein coding sequences of *BmWD40s* with their corresponding genomic sequence.

3.3. Phylogenetic Analysis and Domain Architecture Annotation

The complete amino acid sequences or WD40 domain sequences of the identified *Bm*WD40 proteins were imported into MEGA6 [68], and multiple sequence alignments were carried out by the MUSCLE program [69]. The alignment result was then subjected to build an unrooted phylogenetic tree by the neighbor-joining method [51] with a bootstrap of 1000 replicates. Three online programs, WDSP [45], CDD [43] and SMART [44], were used to determine the domain architectures of the silkworm WD40 proteins.

3.4. Gene Ontology Annotation and Pathway Enrichment

Gene Ontology (GO) analysis was performed using Blast2GO software (Version 4.0.7, BioBam Bioinformatics S.L., Valencia, Spain) with default settings [70]. GO annotation of genes was also obtained from the two silkworm genome databases (SilkDB and KAIKObase). KOBAS v2.0 (http://kobas.cbi.pku.edu.cn) [71] was used to test the statistical enrichment of *BmWD40* genes in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. KEGG pathways with a corrected *p*-value less than 0.05 were considered as enriched.

3.5. Expression Profiling Using Microarray Data

To elucidate the tissue-specific expression profile of *BmWD40* genes, the microarray gene expression data of 10 larval tissues on Day 3 of the fifth instar, namely head, ovary, testis, fat body, integument, midgut, Malpighian tubule, hemocyte, Anterior/Middle Silk Gland (A/MSG) and Posterior Silk Gland (PSG), were retrieved from SilkDB (http://www.silkdb.org). Each tissue sample was analyzed using at least two biological repeats. The microarray data were analyzed using a previously-described protocol [72]. Raw microarray data were normalized using a linear normalization method, in which four confirmed house-keeping genes (encoding proteasome β subunit, eIF 4A, eIF 3A subunit 5 and eIF-3 subunit 4) were used. The normalized values were analyzed using one-way Analysis Of Variance (ANOVA) across all investigated tissues ($p < 0.001$). The filtered data were used to examine differential gene expression with log2 (Fold-change) of ± 1 and $p < 0.01$ (unpaired *t*-tests). Genes with normalized signal intensity values less than 400 were defined as "unexpressed".

To elucidate the expression pattern of *BmWD40* genes during silkworm metamorphosis, we analyzed the microarray gene expression data at 14 developmental time points from late larval to adult stages, including the spinning stage (beginning of Wandering for spinning (W0), 12 h after Wandering (W12 h), W24 h, W36 h and W48 h), pupal stage (beginning of Pupation (P0), 1 day after Pupation (P1), P2, P3, P4, P5, P6 and P7) and adult. Each sample was analyzed using two biological repeats. Microarray hybridization was performed as previously reported [72], and the microarray data were processed as described in a previous report [73]. Briefly, gene expression on Day 3 of the fifth Larval instar (L5D3) was used as a control. If the signal intensity of a gene exceeded 400 units at a time point, the gene was considered to be expressed. Then, the ratio of the signal intensities of each *BmWD40* gene at each developmental time point to that in the L5D3 control was used to detect dynamic changes in the expression levels.

The hierarchical clustering graph showing spatiotemporal expression profiles was generated based on average linkage clustering and a Pearson correlation distance metric using the Heat map Illustrator (HemI) software package (version 1.0.3.3, Huazhong University of Science and Technology, Wuhan, Hubei, China) [74].

4. Conclusions

WD40 proteins play key roles in dissimilar biochemical mechanisms and various biological processes. We conducted a genome-wide analysis and identified the numbers of WD40 protein genes in 9 vertebrates (mean ~256), 8 insects (mean ~178) and 5 nematodes (mean ~135). The WD40 protein family has undergone lineage-specific expansions in animal evolution. Sequence analyses of the silkworm suggested that tandem and segmental duplication did not play a major role in producing the current number of *BmWD40s* and that the gene duplication events occurred predominantly prior to the domain recombination events in the evolution of the multi-domain *BmWD40s*. Microarray-based expression profiling revealed the tissue-specific, stage-specific, or sex-dimorphic expressions of the *BmWD40* genes. These findings increase our understanding of the evolution of the WD40 protein family in animals and provide detailed knowledge of the silkworm WD40 protein family. The exact functions of these differentially-expressed *BmWD40* genes during silkworm development require future characterization.

Author Contributions: Songzhen He and Fangyin Dai conceived of the study. Songzhen He performed the study. Songzhen He and Minjin Han analyzed the data. Hai Hu contributed materials. Songzhen He wrote the paper. Xiaoling Tong, Minjin Han and Fangyin Dai edited and revised the manuscript.

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

Abbreviations

WD	Tryptophan-Aspartic Acid (W-D)
LisH	Lis Homology
BEACH	Beige and Chediak–Higashi
CAF-1	Chromatin Assembly Factor-1
VLCFA	Very-Long-Chain Fatty Acids

References

1. Neer, E.J.; Schmidt, C.J.; Nambudripad, R.; Smith, T.F. The ancient regulatory-protein family of WD-repeat proteins. *Nature* **1994**, *371*, 297–300. [CrossRef] [PubMed]
2. Hu, X.J.; Li, T.; Wang, Y.; Xiong, Y.; Wu, X.H.; Zhang, D.L.; Ye, Z.Q.; Wu, Y.D. Prokaryotic and Highly-Repetitive WD40 Proteins: A Systematic Study. *Sci. Rep.* **2017**, *7*, 10585. [CrossRef] [PubMed]
3. Fong, H.K.; Hurley, J.B.; Hopkins, R.S.; Miake-Lye, R.; Johnson, M.S.; Doolittle, R.F.; Simon, M.I. Repetitive segmental structure of the transducin beta subunit: Homology with the *CDC4* gene and identification of related mRNAs. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 2162–2166. [CrossRef] [PubMed]
4. Van der Voorn, L.; Ploegh, H.L. The WD-40 repeat. *FEBS Lett.* **1992**, *307*, 131–134. [CrossRef]
5. Andrade, M.A.; Perez-Iratxeta, C.; Ponting, C.P. Protein repeats: Structures, functions, and evolution. *J. Struct. Biol.* **2001**, *134*, 117–131. [CrossRef] [PubMed]
6. Wall, M.A.; Coleman, D.E.; Lee, E.; Iniguez-Lluhi, J.A.; Posner, B.A.; Gilman, A.G.; Sprang, S.R. The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell* **1995**, *83*, 1047–1058. [CrossRef]
7. Lambright, D.G.; Sondek, J.; Bohm, A.; Skiba, N.P.; Hamm, H.E.; Sigler, P.B. The 2.0 A crystal structure of a heterotrimeric G protein. *Nature* **1996**, *379*, 311–319. [CrossRef] [PubMed]
8. Sondek, J.; Bohm, A.; Lambright, D.G.; Hamm, H.E.; Sigler, P.B. Crystal structure of a G-protein beta gamma dimer at 2.1A resolution. *Nature* **1996**, *379*, 369–374. [CrossRef] [PubMed]
9. Smith, T.F. Diversity of WD-repeat proteins. *Subcell. Biochem.* **2008**, *48*, 20–30. [PubMed]
10. Li, D.; Roberts, R. WD-repeat proteins: Structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci.* **2001**, *58*, 2085–2097. [CrossRef] [PubMed]
11. Xu, C.; Min, J. Structure and function of WD40 domain proteins. *Protein Cell* **2011**, *2*, 202–214. [CrossRef] [PubMed]
12. Wakasugi, M.; Kawashima, A.; Morioka, H.; Linn, S.; Sancar, A.; Mori, T.; Nikaido, O.; Matsunaga, T. DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair. *J. Biol. Chem.* **2002**, *277*, 1637–1640. [CrossRef] [PubMed]
13. Znaidi, S.; Pelletier, B.; Mukai, Y.; Labbe, S. The Schizosaccharomyces pombe corepressor Tup11 interacts with the iron-responsive transcription factor Fep1. *J. Biol. Chem.* **2004**, *279*, 9462–9474. [CrossRef] [PubMed]
14. Jennings, B.H.; Pickles, L.M.; Wainwright, S.M.; Roe, S.M.; Pearl, L.H.; Ish-Horowicz, D. Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol. Cell* **2006**, *22*, 645–655. [CrossRef] [PubMed]
15. Ajuh, P.; Sleeman, J.; Chusainow, J.; Lamond, A.I. A direct interaction between the carboxyl-terminal region of CDC5L and the WD40 domain of PLRG1 is essential for pre-mRNA splicing. *J. Biol. Chem.* **2001**, *276*, 42370–42381. [CrossRef] [PubMed]
16. Vander Kooi, C.W.; Ren, L.; Xu, P.; Ohi, M.D.; Gould, K.L.; Chazin, W.J. The Prp19 WD40 domain contains a conserved protein interaction region essential for its function. *Structure* **2010**, *18*, 584–593. [CrossRef] [PubMed]
17. Ren, L.; McLean, J.R.; Hazbun, T.R.; Fields, S.; Vander Kooi, C.; Ohi, M.D.; Gould, K.L. Systematic two-hybrid and comparative proteomic analyses reveal novel yeast pre-mRNA splicing factors connected to Prp19. *PLoS ONE* **2011**, *6*, e16719. [CrossRef]
18. Yan, C.; Hang, J.; Wan, R.; Huang, M.; Wong, C.C.; Shi, Y. Structure of a yeast spliceosome at 3.6—Angstrom resolution. *Science* **2015**, *349*, 1182–1191. [CrossRef] [PubMed]
19. Higa, L.A.; Wu, M.; Ye, T.; Kobayashi, R.; Sun, H.; Zhang, H. CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol.* **2006**, *8*, 1277–1283. [CrossRef] [PubMed]
20. Song, J.J.; Garlick, J.D.; Kingston, R.E. Structural basis of histone H4 recognition by p55. *Genes Dev.* **2008**, *22*, 1313–1318. [CrossRef] [PubMed]
21. Suganuma, T.; Pattenden, S.G.; Workman, J.L. Diverse functions of WD40 repeat proteins in histone recognition. *Genes Dev.* **2008**, *22*, 1265–1268. [CrossRef] [PubMed]
22. Adams, J.M.; Cory, S. Apoptosomes: Engines for caspase activation. *Curr. Opin. Cell Biol.* **2002**, *14*, 715–720. [CrossRef]
23. Yuan, S.; Yu, X.; Topf, M.; Ludtke, S.J.; Wang, X.; Akey, C.W. Structure of an apoptosome-procaspase-9 CARD complex. *Structure* **2010**, *18*, 571–583. [CrossRef] [PubMed]

24. Reubold, T.F.; Wohlgemuth, S.; Eschenburg, S. Crystal structure of full-length Apaf-1: How the death signal is relayed in the mitochondrial pathway of apoptosis. *Structure* **2011**, *19*, 1074–1083. [CrossRef] [PubMed]

25. Chen, R.H.; Miettinen, P.J.; Maruoka, E.M.; Choy, L.; Derynck, R. A WD-domain protein that is associated with and phosphorylated by the type II TGF-beta receptor. *Nature* **1995**, *377*, 548–552. [CrossRef] [PubMed]

26. Gaudet, R.; Bohm, A.; Sigler, P.B. Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin. *Cell* **1996**, *87*, 577–588. [CrossRef]

27. Datta, P.K.; Moses, H.L. STRAP and Smad7 synergize in the inhibition of transforming growth factor beta signaling. *Mol. Cell. Biol.* **2000**, *20*, 3157–3167. [CrossRef] [PubMed]

28. Gallenberger, M.; Meinel, D.M.; Kroeber, M.; Wegner, M.; Milkereit, P.; Bosl, M.R.; Tamm, E.R. Lack of WDR36 leads to preimplantation embryonic lethality in mice and delays the formation of small subunit ribosomal RNA in human cells in vitro. *Hum. Mol. Genet.* **2011**, *20*, 422–435. [CrossRef] [PubMed]

29. Ozawa, T.; Koide, R.; Nakata, Y.; Saitsu, H.; Matsumoto, N.; Takahashi, K.; Nakano, I.; Orimo, S. A novel WDR45 mutation in a patient with static encephalopathy of childhood with neurodegeneration in adulthood (SENDA). *Am. J. Med. Genet. A* **2014**, *164A*, 2388–2390. [CrossRef] [PubMed]

30. Bilguvar, K.; Ozturk, A.K.; Louvi, A.; Kwan, K.Y.; Choi, M.; Tatli, B.; Yalnizoglu, D.; Tuysuz, B.; Caglayan, A.O.; Gokben, S.; et al. Whole-exome sequencing identifies recessive *WDR62* mutations in severe brain malformations. *Nature* **2010**, *467*, 207–210. [CrossRef] [PubMed]

31. Nicholas, A.K.; Khurshid, M.; Desir, J.; Carvalho, O.P.; Cox, J.J.; Thornton, G.; Kausar, R.; Ansar, M.; Ahmad, W.; Verloes, A.; et al. WDR62 is associated with the spindle pole and is mutated in human microcephaly. *Nat. Genet.* **2010**, *42*, 1010–1014. [CrossRef] [PubMed]

32. Wang, X.; Zhang, J.; Zhou, L.; Sun, W.; Zheng, Z.G.; Lu, P.; Gao, Y.; Yang, X.S.; Zhang, Z.C.; Tao, K.S.; et al. Fbxw7 regulates hepatocellular carcinoma migration and invasion via Notch1 signaling pathway. *Int. J. Oncol.* **2015**, *47*, 231–243. [CrossRef] [PubMed]

33. Zhan, P.; Wang, Y.; Zhao, S.; Liu, C.; Wang, Y.; Wen, M.; Mao, J.H.; Wei, G.; Zhang, P. FBXW7 negatively regulates ENO1 expression and function in colorectal cancer. *Lab. Investig.* **2015**, *95*, 995–1004. [CrossRef] [PubMed]

34. Ramasamy, S.; Saez, B.; Mukhopadhyay, S.; Ding, D.; Ahmed, A.M.; Chen, X.; Pucci, F.; Yamin, R.; Wang, J.; Pittet, M.J.; et al. Tle1 tumor suppressor negatively regulates inflammation in vivo and modulates NF-kappaB inflammatory pathway. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1871–1876. [CrossRef] [PubMed]

35. Orville Singh, C.; Xin, H.H.; Chen, R.T.; Wang, M.X.; Liang, S.; Lu, Y.; Cai, Z.Z.; Miao, Y.G. BmPLA2 containing conserved domain WD40 affects the metabolic functions of fat body tissue in silkworm, *Bombyx mori*. *Insect Sci.* **2016**, *23*, 28–36. [CrossRef] [PubMed]

36. Van Nocker, S.; Ludwig, P. The WD-repeat protein superfamily in *Arabidopsis*: Conservation and divergence in structure and function. *BMC Genom.* **2003**, *4*, 50. [CrossRef] [PubMed]

37. Ouyang, Y.; Huang, X.; Lu, Z.; Yao, J. Genomic survey, expression profile and co-expression network analysis of OsWD40 family in rice. *BMC Genom.* **2012**, *13*, 100. [CrossRef] [PubMed]

38. Li, Q.; Zhao, P.; Li, J.; Zhang, C.; Wang, L.; Ren, Z. Genome-wide analysis of the WD-repeat protein family in cucumber and *Arabidopsis*. *Mol. Genet. Genom.* **2014**, *289*, 103–124. [CrossRef] [PubMed]

39. Mishra, A.K.; Muthamilarasan, M.; Khan, Y.; Parida, S.K.; Prasad, M. Genome-wide investigation and expression analyses of WD40 protein family in the model plant foxtail millet (*Setaria italica* L.). *PLoS ONE* **2014**, *9*, e86852. [CrossRef] [PubMed]

40. Zou, X.D.; Hu, X.J.; Ma, J.; Li, T.; Ye, Z.Q.; Wu, Y.D. Genome-wide Analysis of WD40 Protein Family in Human. *Sci. Rep.* **2016**, *6*, 39262. [CrossRef] [PubMed]

41. Goldsmith, M.R.; Shimada, T.; Abe, H. The genetics and genomics of the silkworm, *Bombyx mori*. *Annu. Rev. Entomol.* **2005**, *50*, 71–100. [CrossRef] [PubMed]

42. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* **2011**, *39*, W29–W37. [CrossRef] [PubMed]

43. Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; et al. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, D225–D229. [CrossRef] [PubMed]

44. Ponting, C.P.; Schultz, J.; Milpetz, F.; Bork, P. SMART: Identification and annotation of domains from signalling and extracellular protein sequences. *Nucleic Acids Res.* **1999**, *27*, 229–232. [CrossRef] [PubMed]

45. Wang, Y.; Hu, X.J.; Zou, X.D.; Wu, X.H.; Ye, Z.Q.; Wu, Y.D. WDSPdb: A database for WD40-repeat proteins. *Nucleic Acids Res.* **2015**, *43*, D339–D344. [CrossRef] [PubMed]

46. Angers, S.; Li, T.; Yi, X.; MacCoss, M.J.; Moon, R.T.; Zheng, N. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **2006**, *443*, 590–593. [CrossRef] [PubMed]

47. Hua, Z.; Vierstra, R.D. The cullin-RING ubiquitin-protein ligases. *Annu. Rev. Plant. Biol.* **2011**, *62*, 299–334. [CrossRef] [PubMed]

48. Zhang, L.; Lu, H.H.; Chung, W.Y.; Yang, J.; Li, W.H. Patterns of segmental duplication in the human genome. *Mol. Biol. Evol.* **2005**, *22*, 135–141. [CrossRef] [PubMed]

49. Shoja, V.; Zhang, L. A roadmap of tandemly arrayed genes in the genomes of human, mouse, and rat. *Mol. Biol. Evol.* **2006**, *23*, 2134–2141. [CrossRef] [PubMed]

50. Huang, S.; Li, R.; Zhang, Z.; Li, L.; Gu, X.; Fan, W.; Lucas, W.J.; Wang, X.; Xie, B.; Ni, P.; et al. The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* **2009**, *41*, 1275–1281. [CrossRef] [PubMed]

51. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [PubMed]

52. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef] [PubMed]

53. Ogawa, K.; Takai, H.; Ogiwara, A.; Yokota, E.; Shimizu, T.; Inaba, K.; Mohri, H. Is outer arm dynein intermediate chain 1 multifunctional? *Mol. Biol. Cell.* **1996**, *7*, 1895–1907. [CrossRef] [PubMed]

54. Guichard, C.; Harricane, M.C.; Lafitte, J.J.; Godard, P.; Zaegel, M.; Tack, V.; Lalau, G.; Bouvagnet, P. Axonemal dynein intermediate-chain gene (DNAI1) mutations result in situs inversus and primary ciliary dyskinesia (Kartagener syndrome). *Am. J. Hum. Genet.* **2001**, *68*, 1030–1035. [CrossRef] [PubMed]

55. Young, S.A.; Miyata, H.; Satouh, Y.; Kato, H.; Nozawa, K.; Isotani, A.; Aitken, R.J.; Baker, M.A.; Ikawa, M. CRISPR/Cas9-Mediated Rapid Generation of Multiple Mouse Lines Identified Ccdc63 as Essential for Spermiogenesis. *Int. J. Mol. Sci.* **2015**, *16*, 24732–24750. [CrossRef] [PubMed]

56. Zhang, Y.; Liu, H.; Li, W.; Zhang, Z.; Shang, X.; Zhang, D.; Li, Y.; Zhang, S.; Liu, J.; Hess, R.A.; et al. Intraflagellar transporter protein (IFT27), an IFT25 binding partner, is essential for male fertility and spermiogenesis in mice. *Dev. Biol.* **2017**, *432*, 125–139. [CrossRef] [PubMed]

57. Brites, P.; Mooyer, P.A.; El Mrabet, L.; Waterham, H.R.; Wanders, R.J. Plasmalogens participate in very-long-chain fatty acid-induced pathology. *Brain* **2009**, *132*, 482–492. [CrossRef] [PubMed]

58. Kannan, M.; Bayam, E.; Wagner, C.; Rinaldi, B.; Kretz, P.F.; Tilly, P.; Roos, M.; McGillewie, L.; Bar, S.; Minocha, S.; et al. WD40-repeat 47, a microtubule-associated protein, is essential for brain development and autophagy. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E9308–E9317. [CrossRef] [PubMed]

59. Dahlberg, C.; Kowalski, J.; Juo, P. WDR-20 and WDR-48 regulate the deubiquitinating enzyme USP-46 to promote the abundance of glutamate receptors in the ventral nerve cord of *C. elegans*. *Mol. Biol. Cell.* **2013**, *299*, 3444–3456.

60. Lodder, E.M.; De Nittis, P.; Koopman, C.D.; Wiszniewski, W.; Moura de Souza, C.F.; Lahrouchi, N.; Guex, N.; Napolioni, V.; Tessadori, F.; Beekman, L.; et al. GNB5 Mutations Cause an Autosomal-Recessive Multisystem Syndrome with Sinus Bradycardia and Cognitive Disability. *Am. J. Hum. Genet.* **2016**, *99*, 704–710. [CrossRef] [PubMed]

61. Jones, J.M.; Huang, J.D.; Mermall, V.; Hamilton, B.A.; Mooseker, M.S.; Escayg, A.; Copeland, N.G.; Jenkins, N.A.; Meisler, M.H. The mouse neurological mutant flailer expresses a novel hybrid gene derived by exon shuffling between Gnb5 and Myo5a. *Hum. Mol. Genet.* **2000**, *9*, 821–828. [CrossRef] [PubMed]

62. Kottgen, A.; Pattaro, C.; Boger, C.A.; Fuchsberger, C.; Olden, M.; Glazer, N.L.; Parsa, A.; Gao, X.; Yang, Q.; Smith, A.V.; et al. New loci associated with kidney function and chronic kidney disease. *Nat. Genet.* **2010**, *42*, 376–384. [CrossRef] [PubMed]

63. Smith, T.F.; Gaitatzes, C.; Saxena, K.; Neer, E.J. The WD repeat: A common architecture for diverse functions. *Trends Biochem. Sci.* **1999**, *24*, 181–185. [CrossRef]

64. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [CrossRef] [PubMed]

65. Voorrips, R.E. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [CrossRef] [PubMed]

66. Shiu, S.H.; Bleecker, A.B. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in *Arabidopsis*. *Plant Physiol.* **2003**, *132*, 530–543. [PubMed]

67. Zhao, Q.; Zhu, Z.; Kasahara, M.; Morishita, S.; Zhang, Z. Segmental duplications in the silkworm genome. *BMC Genom.* **2013**, *14*, 521. [CrossRef] [PubMed]

68. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]

69. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef] [PubMed]

70. Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Williams, T.D.; Nagaraj, S.H.; Nueda, M.J.; Robles, M.; Talon, M.; Dopazo, J.; Conesa, A. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **2008**, *36*, 3420–3435. [CrossRef] [PubMed]

71. Xie, C.; Mao, X.; Huang, J.; Ding, Y.; Wu, J.; Dong, S.; Kong, L.; Gao, G.; Li, C.-Y.; Wei, L. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **2011**, *39*, W316–W322. [CrossRef] [PubMed]

72. Xia, Q.; Cheng, D.; Duan, J.; Wang, G.; Cheng, T.; Zha, X.; Liu, C.; Zhao, P.; Dai, F.; Zhang, Z.; et al. Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori*. *Genome Biol.* **2007**, *8*, R162. [CrossRef] [PubMed]

73. Cheng, D.; Qian, W.; Meng, M.; Wang, Y.; Peng, J.; Xia, Q. Identification and Expression Profiling of the BTB Domain-Containing Protein Gene Family in the Silkworm, *Bombyx mori*. *Int. J. Genom.* **2014**, *2014*, 865065. [CrossRef] [PubMed]

74. Deng, W.; Wang, Y.; Liu, Z.; Cheng, H.; Xue, Y. HemI: A toolkit for illustrating heatmaps. *PLoS ONE* **2014**, *9*, e111988. [CrossRef] [PubMed]