

# Genome-wide identification and expression profiling of the C2H2-type zinc finger protein genes in the silkworm *Bombyx mori*

SongYuan Wu<sup>1,2</sup>, Xiaoling Tong<sup>1</sup>, ChunLin Li<sup>1</sup>, KunPeng Lu<sup>1</sup>, Duan Tan<sup>1</sup>, Hai Hu<sup>1</sup>, Huai Liu<sup>2</sup> and FangYin Dai<sup>1</sup>

<sup>1</sup> State Key Laboratory of Silkworm Genome Biology, Key Laboratory of Sericultural Biology and Genetic Breeding, Ministry of Agriculture, College of Biotechnology, Southwest University, Chong Qing, China

<sup>2</sup> College of Plant Protection, Southwest University, Chong Qing, China

## ABSTRACT

Cys2-His2 zinc finger (C2H2-ZF) proteins comprise the largest class of putative eukaryotic transcription factors. The zinc finger motif array is highly divergent, indicating that most proteins will have distinctive binding sites and perform different functions. However, the binding sites and functions of the majority of C2H2-ZF proteins remain unknown. In this study, we identified 327 C2H2-ZF protein genes in the silkworm, 290 in the monarch butterfly, 243 in the fruit fly, 107 in *elegans*, 673 in mouse, and 1,082 in human. The C2H2-ZF protein genes of the silkworm were classified into three main grouping clades according to a phylogenetic classification, and 312 of these genes could be mapped onto 27 chromosomes. Most silkworm C2H2-ZF protein genes exhibited specific expression in larval tissues. Furthermore, several C2H2-ZF protein genes had sex-specific expression during metamorphosis. In addition, we found that some C2H2-ZF protein genes are involved in metamorphosis and female reproduction by using expression clustering and gene annotation analysis. Among them, five genes were selected, *BGIBMGA002091* (*CTCF*), *BGIBMGA006492* (*fru*), *BGIBMGA006230* (*wor*), *BGIBMGA004640* (*lola*), and *BGIBMGA004569*, for quantitative real-time PCR analysis from larvae to adult ovaries. The results showed that the five genes had different expression patterns in ovaries, among which *BGIBMGA002091* (*CTCF*) gene expression level was the highest, and its expression level increased rapidly in late pupae and adult stages. These findings provide a basis for further investigation of the functions of C2H2-ZF protein genes in the silkworm, and the results offer clues for further research into the development of metamorphosis and female reproduction in the silkworm.

Submitted 10 March 2019

Accepted 30 May 2019

Published 5 July 2019

Corresponding authors

FangYin Dai, [fydai@swu.edu.cn](mailto:fydai@swu.edu.cn)

Huai Liu, [liuhuai@swu.edu.cn](mailto:liuhuai@swu.edu.cn)

Academic editor

Biswapriya Misra

Additional Information and  
Declarations can be found on  
page 15

DOI [10.7717/peerj.7222](https://doi.org/10.7717/peerj.7222)

© Copyright  
2019 Wu et al.

Distributed under  
Creative Commons CC-BY 4.0

**OPEN ACCESS**

**Subjects** Agricultural Science, Genetics, Genomics

**Keywords** C2H2-ZF, Silkworm, Characterization, Expression pattern, Functional association

## INTRODUCTION

Transcription factors (TFs) play an important role in regulating gene expression (*Latchman, 1990*). A TF is characterized by the DNA-binding domain; some TFs also contain other effector domains. TFs are involved in multiple biological processes,

including development (Sommer *et al.*, 1992), metabolism (Krebs *et al.*, 2014), apoptosis (Ma *et al.*, 2016), autophagy, (Chauhan *et al.*, 2013), and stemness maintenance (Lai *et al.*, 2014). The Cys2-His2 zinc finger proteins (C2H2-ZF) number more than 700 and are the largest family of TFs found in humans (Vaquerizas *et al.*, 2009; Weirauch & Hughes, 2011). In addition, zinc finger array and putative DNA-binding sequences are significantly differentiated, indicating that the functions of C2H2-ZF are diverse. The C2H2-ZF motif is composed of Cys-X<sub>2,4</sub>-Cys-X<sub>12</sub>-His-X<sub>3,4,5</sub>-His, and the motif interacts with zinc ions and folds into a finger-like structure. This finger-like structure is composed of  $\alpha$ -helix and antiparallel  $\beta$ -sheet (Klug & Schwabe, 1995; McCarty *et al.*, 2003).

In general, the binding sequence of the C2H2-ZF motif is according to the amino acid residue sequence of its  $\alpha$ -helical component. Regarding the first amino acid residue of the  $\alpha$ -helical region in the C2H2-ZF domain as position 1, the four positions -1, 2, 3, and 6 are the key amino acids for recognition and binding of DNA via interaction with the hydrogen donors and acceptors exposed in the major groove. The -1, 3, and 6 amino acid residues bind to base 3, base 2, and base 1, respectively, in the sense sequence of DNA, and the amino acid residue of position 2 binds to the fourth base of the antisense sequence (Klug, 2010; Wolfe, Nekludova & Pabo, 2000) (Fig. S1). Many TFs contain a group of tandem zinc fingers, and the prediction of their recognition sequence often depends on single zinc finger. However, the recognition sequences of individual zinc finger proteins may be modified due to adjacent zinc fingers. Therefore, it is difficult to analyze the recognition sequences of tandem zinc finger domains (Lam *et al.*, 2011; Ramirez *et al.*, 2008). In recent years, some bioinformatic analytical approaches have been made, along with chromatin immunoprecipitation followed by sequencing (ChIP-seq) data for tandem zinc finger proteins (Gupta *et al.*, 2014; Hamzeh-Mivehroud *et al.*, 2015; Jayakanthan *et al.*, 2009; Persikov & Singh, 2014; Sandelin *et al.*, 2004; Sander *et al.*, 2009; Wolfe *et al.*, 1999; Yang *et al.*, 2013). These methods provide important references for studying the molecular mechanism and function of zinc finger TFs. Nonetheless, the zinc finger family is numerous, and there is significant divergence among species. Zinc finger transcript factors participate in multiple biological processes by transcript regulation, and thus clarifying the function and molecular mechanism of zinc fingers is complicated. Therefore, further studies of the zinc finger protein family are important for understanding the function and evolution of TFs.

The silkworm *Bombyx mori* was domesticated from the Chinese wild silkworm *B. mandarina* over 5,000 years ago (Sun *et al.*, 2012). The silkworm has long been an economically important insect due to the production of silk. From the 19th century, the silkworm became a model for scientific discovery in microbiology, physiology, and genetics (Goldsmith, Shimada & Abe, 2005). Silkworms, which have undergone a long period of domestication and artificial selection, have a clearly understood genetic background. In addition, *B. mori* became a representative organism of Lepidoptera for basic research following the sequencing of the whole genome (Meng, Zhu & Chen, 2017). In this study, we identified C2H2-ZF proteins in the zinc finger gene family by genome-wide analysis, and we performed sequence and microarray analyses. Our findings should provide a basis for better understanding of the C2H2-ZF protein families, and the

results may provide important clues for further research concerning the functions of TFs in the silkworm and Lepidoptera.

## MATERIALS AND METHODS

### Genome-wide identification of the genes encoding C2H2-ZF proteins

The whole-genome protein sequences of *B. mori* were downloaded from the silkworm genome databases Kaikobase (<http://sgp.dna.affrc.go.jp/KAIKObase/>) (Shimomura *et al.*, 2009). The hidden Markov model (HMM) profile of the C2H2-ZF (PF00096) was downloaded from the Pfam database (<https://pfam.xfam.org>) (Sonnhammer, Eddy & Durbin, 1997). HMMER 3.0 software was downloaded from <http://www.hmmerr.org>, and the hmmsearch program was used to search for C2H2 zinc finger proteins in the silkworm genome, and the expected value was set to be less than 0.01 (Finn, Clements & Eddy, 2011). The candidate proteins containing C2H2 zinc finger domains were identified on the SMART website (<http://smart.embl-heidelberg.de/>) (Ponting *et al.*, 1999). The protein sequences of *Danaus plexippus* were downloaded from MonarchBase (<http://monarchbase.umassmed.edu>) (Zhan & Reppert, 2013), the proteins of *Drosophila melanogaster* were downloaded from flybase (<http://flybase.org>) (Gramates *et al.*, 2017), the protein sequences of *Caenorhabditis elegans* were downloaded from Wormbase (<http://www.wormbase.org>) and the protein sequences of *Homo sapiens* and *Mus musculus* were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov>). BLASTP searches against the non-redundant (nr) protein database were performed to annotate the C2H2-type zinc finger protein genes.

### Chromosomal distribution and domain analysis of the C2H2-ZF protein genes

The chromosomal sizes and gene locations were obtained from KAIKObase. The MapChart 2.32 software was used to draw the chromosomal distribution map (Voorrips, 2002). The analysis of conservative domains used the hmmscan program (Finn, Clements & Eddy, 2011). The well-characterized protein domain families with high-quality alignment data, Pfam-A (Sonnhammer, Eddy & Durbin, 1997), were downloaded from NCBI. The proteins of *B. mori*, *Danaus plexippus*, *Drosophila melanogaster*, *C. elegans*, *M. musculus*, and *H. sapiens* were searched against Pfam-A to annotate the domains, and the parameters are defaulted. The complete amino acid sequences of the identified BmZNF proteins were imported into MEGA7 (Kumar, Stecher & Tamura, 2016), and the MUSCLE program carried out multiple sequence alignments (Edgar, 2004). The alignment result was then used to build an unrooted phylogenetic tree by the neighbor-joining method with a bootstrap of 2,000 replicates (Saitou & Nei, 1987).

### Expression profiling analysis

The gene expression data were downloaded from a *B. mori* microarray data (Xia *et al.*, 2007). The multiple larval tissues examined from the third day of the fifth instar included testis, ovary, head, integument, fat body, midgut, hemocyte, Malpighian tubules, anterior/middle silk gland, and posterior silk gland. In addition, the 19 developmental time

points, from later larval to adult stages, included day 4 of the fifth instar (5th4d), 5th5d, 5th6d, 5th7d, the beginning of the wandering stage for spinning (W0), 12 h after wandering (W12h), W24h, W36h, W48h (spinning finished), beginning of pupation (P0), 1 day after pupation (P1), P2, P3, P4, P5, P6, P7, P8, and day 1 of the adult moth. In the microarray data analysis, if the signal intensity of the expression of a gene exceeded 400 units at a time point, this gene was considered to be expressed at that time point. In microarray analysis of metamorphosis, from 5th4d to adult, level of expression of genes in the silkworm larvae on day 3 of the fifth instar (5th3d) was used as a control. The ratios of the expression levels of genes were calculated by comparing gene expression with that in the 5th3d control. The expression array figures were drawn with the Pheatmap program of the R package (*R Development Core Team, 2015*).

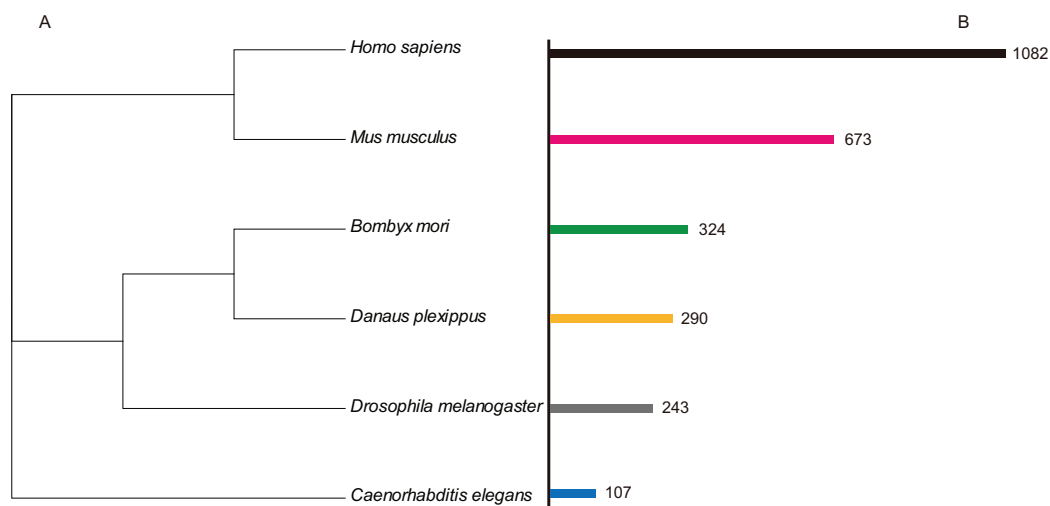
### Ovary dissect and qRT-PCR

The silkworm wild-type strain Dazao were obtained from the Silkworm Gene Bank of Southwest University, Chong Qing, China. The ovary of silkworm was removed using 0.7% normal saline solution, and then stored at  $-80^{\circ}\text{C}$  until RNA extraction was performed. Total RNA was isolated from the ovary using E.Z.N.A.<sup>®</sup> MicroElute Total RNA Kit (Omega Bio-tek, Norcross, GA, USA) according to the manufacturer's protocol. The primers that were used for qRT-PCR were list in [Table S1](#). qRT-PCR was performed using a CFX96<sup>™</sup> Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA) with SYBR Green qRT-PCR Mix (Bio-Rad, Hercules, CA, USA). The cycling parameters were as follows:  $95^{\circ}\text{C}$  for 4 min followed by 40 cycles of  $95^{\circ}\text{C}$  for 10 s and annealing for 30 s. The relative expression levels were analyzed using the classical  $R = 2^{-\Delta\Delta\text{Ct}}$  method. The gene for eukaryotic translation initiation factor 4A ([NP\\_001037376.1](#)) of silkworm was used as an internal control to normalize for equal sample loading.

## RESULTS

### Identification and chromosomal distribution of the BmZNF family

In order to identify C2H2-ZF protein genes in *B. mori*, the hmmsearch program in the HMMER3.0 package ([Finn, Clements & Eddy, 2011](#)) (<http://www.hmmmer.org>) was used to search the silkworm comprehensive genetic data sets ([Suetsugu et al., 2013](#)), and HMM profiles data were used for the C2H2-ZF motif (PF00096). The output data were confirmed by SMART ([Ponting et al., 1999](#)) (Simple Modular Architecture Research Tool, <http://smart.embl-heidelberg.de>) and manually modified. As a result, 327 unique silkworm C2H2 zinc finger protein genes were obtained, and these were named in an abbreviated format as BmZNF (*B. mori* zinc fingers C2H2-type) based on the principles for the nomenclature of C2H2 zinc finger proteins described by the HUGO Gene Nomenclature Committee (HGNC) ([Povey et al., 2001](#)). In addition, all of the BmZNF genes were analyzed using BLAST against the nr protein database (<http://www.ncbi.nlm.nih.gov/>). Using the same method, we identified 243 C2H2-ZF protein genes in *Drosophila melanogaster*, 290 genes in *Danaus plexippus*, 107 genes in *C. elegans*, 673 genes in



**Figure 1** The numbers of ZNF protein gene family member in the different species. (A) Indicated the species tree. (B) Indicated the C2H2-type zinc finger (ZNF) protein genes numbers in different species.

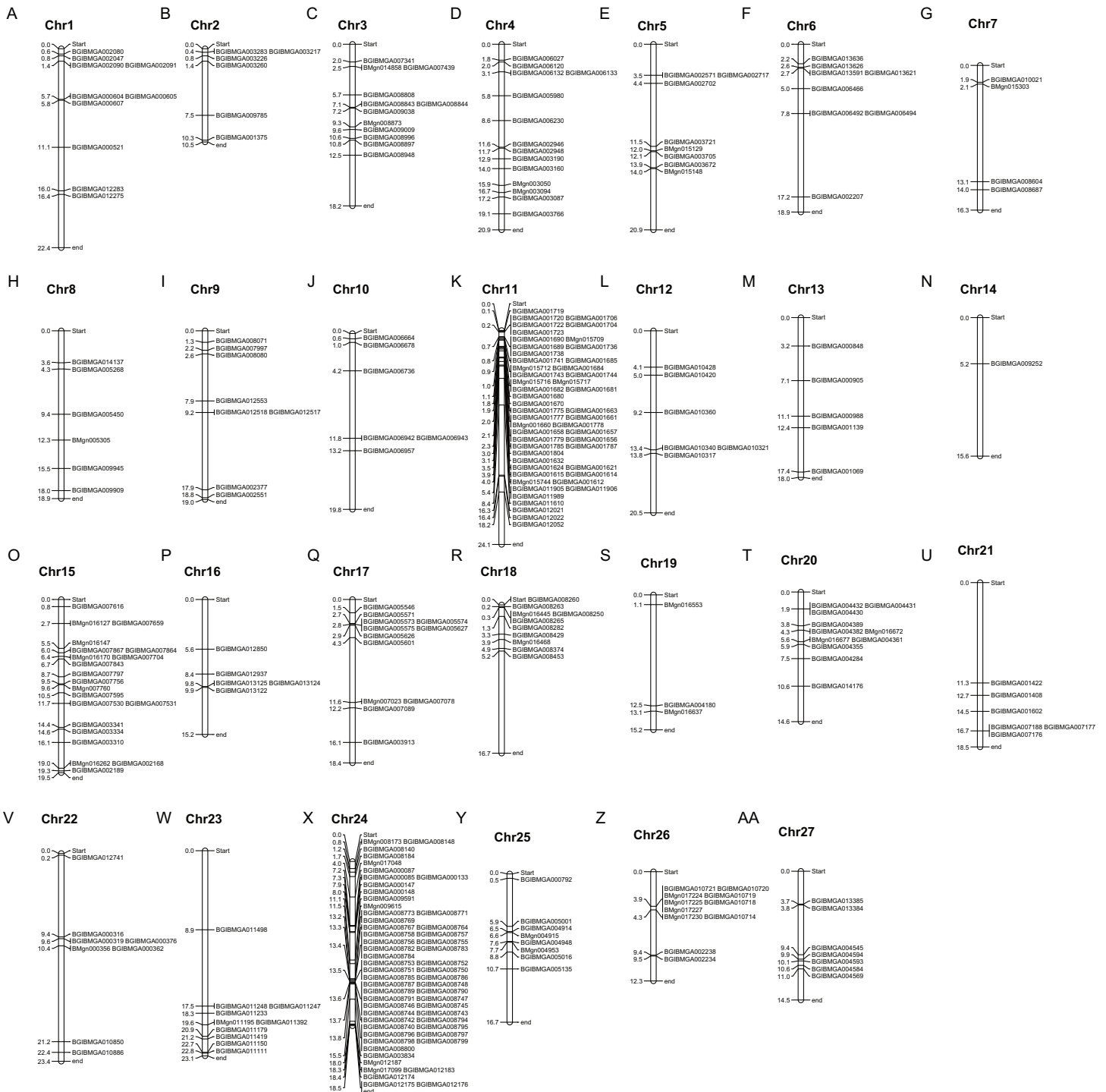
Full-size DOI: 10.7717/peerj.7222/fig-1

*M. musculus*, and 1,082 genes in *H. sapiens* (Fig. 1). The information of those genes was list in Table S2.

Based on the genome assembly for the silkworm, we constructed a silkworm zinc finger C2H2-type protein gene distribution map. The members of the *BmZNF* gene family were widely and unevenly distributed on the chromosomes. As shown in Fig. 2, 312 of the 327 identified C2H2-type zinc finger protein genes were assigned to chromosomes 1–27 of the silkworm, and there was significant clustering on chromosomes 11 and 24. Chromosome 11 is about 24.1 mega bases (MB) in size, containing 50 ZNF genes, and chromosome 24 is about 18.5 MB, with 56 ZNF genes. The *BmZNF* genes on these two chromosomes account for about 1/3 of the total.

### PPIs domain analysis and phylogenetic tree of the BmZNF family

Cys2-His2 zinc finger proteins often contain some protein-protein interaction (PPI) domains; for example, BTB (Broad-Complex, Tramtrack, and Bric-a-brac) (Zollman *et al.*, 1994), the Krüppel-associated box (KRAB) (Mannini *et al.*, 2006), and SCAN (SRE-ZBP, CTfin51, AW-1, and Number 18 cDNA) domain (Schumacher *et al.*, 2000) are found in higher vertebrates. We used the hmmscan program in the HMMER package to perform the domain analysis. The C2H2-type zinc finger proteins of *B. mori*, *Danaus plexippus*, *Drosophila melanogaster*, *C. elegans*, *M. musculus*, and *H. sapiens* were compared to a well-characterized protein domain family in the HMMs database (Pfam-A) (Sonnhammer, Eddy & Durbin, 1997). There were 16, 13, 13, 3, 58, and 63 C2H2-ZF proteins containing the BTB domain (BTB-ZF-C2H2) in *B. mori*, *Danaus plexippus*, *Drosophila melanogaster*, *C. elegans*, *M. musculus*, and *H. sapiens*, respectively. In addition, the majority of C2H2-ZF family proteins are constructed of tandem zinc finger domains. Furthermore, KRAB- and SCAN-containing C2H2-ZF proteins were not found in any of the three insects and worm (Table S3). However, there were 497 and 303 ZNF proteins

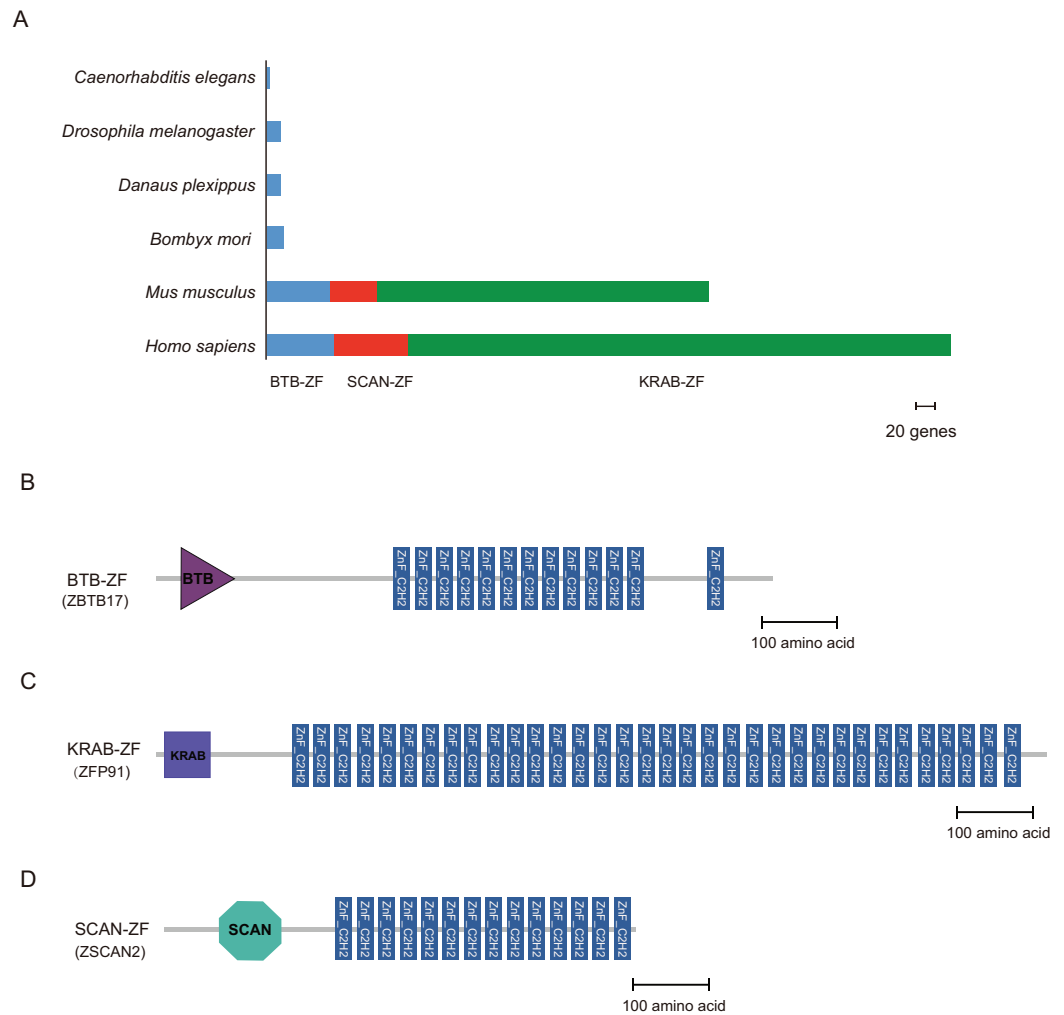


**Figure 2** Chromosomal distribution of *BmZNF* genes. (A–AA) The distribution map of *BmZNF* genes on silkworm chromosomes was drawn by MapChart software. The abbreviation chr represents a chromosome. The distance unit on chromosomes is MB (mega bases).

Full-size [DOI: 10.7717/peerj.7222/fig-2](https://doi.org/10.7717/peerj.7222/fig-2)

containing the KRAB domain in *H. sapiens* and *M. musculus*, respectively. And, the ZNF proteins containing the SCAN domain were found 67 in *H. sapiens* and 44 in *M. musculus*, respectively (Fig. 3).



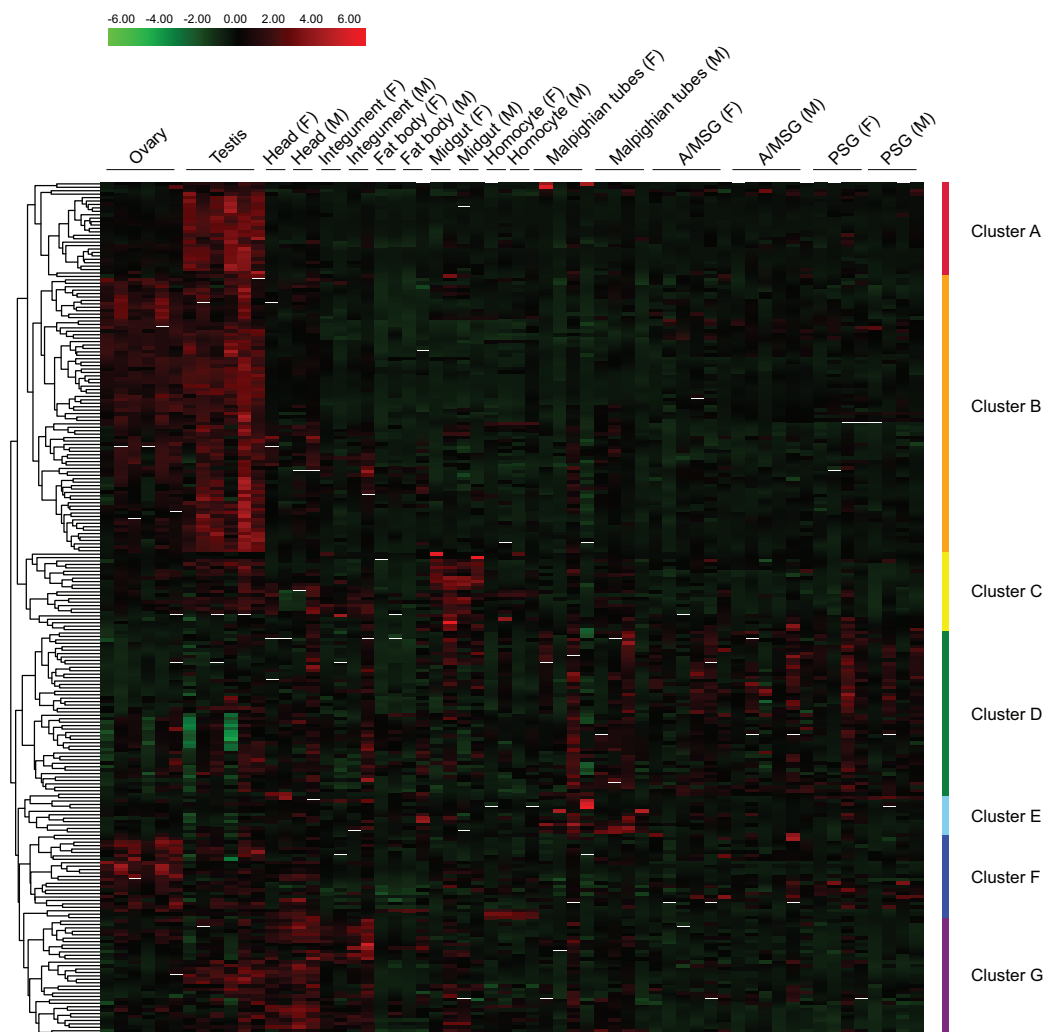


**Figure 3** Distribution of BTB-ZF, KRAB-ZF, and SCAN-ZF proteins in the different species. (A) Representation of BTB-ZF, KRAB-ZF, and SCAN-ZF proteins in selected sequenced genomes. Six species were represented, and each type of protein architecture was shown as bar segments. (B–D) We selected three representative proteins, (ZBTB17, ZFP91, and ZSCAN2), of the human, which are corresponding BTB-ZF, KRAB-ZF, and SCAN-ZF, respectively. The domains structure of those proteins is represented. [Full-size !\[\]\(fcc3264021d438d9732560e78099f674\_img.jpg\) DOI: 10.7717/peerj.7222/fig-3](https://doi.org/10.7717/peerj.7222/fig-3)

In order to analyze the derivation of C2H2-type zinc finger proteins in the silkworm, we used the complete amino acid sequences predicted by the identified silkworm C2H2-type zinc finger proteins to generate a phylogenetic tree of BmZNF protein genes via the neighbor-joining method (Saitou & Nei, 1987). The 327 genes were categorized into three main grouping clades; the clades II and III had two subclades. The five groups (I, II a, II b, III a, and III b) contained 17, 78, 61, 57, and 114 proteins, respectively (Fig. S2).

### Tissue expression profile of BmZNF protein genes

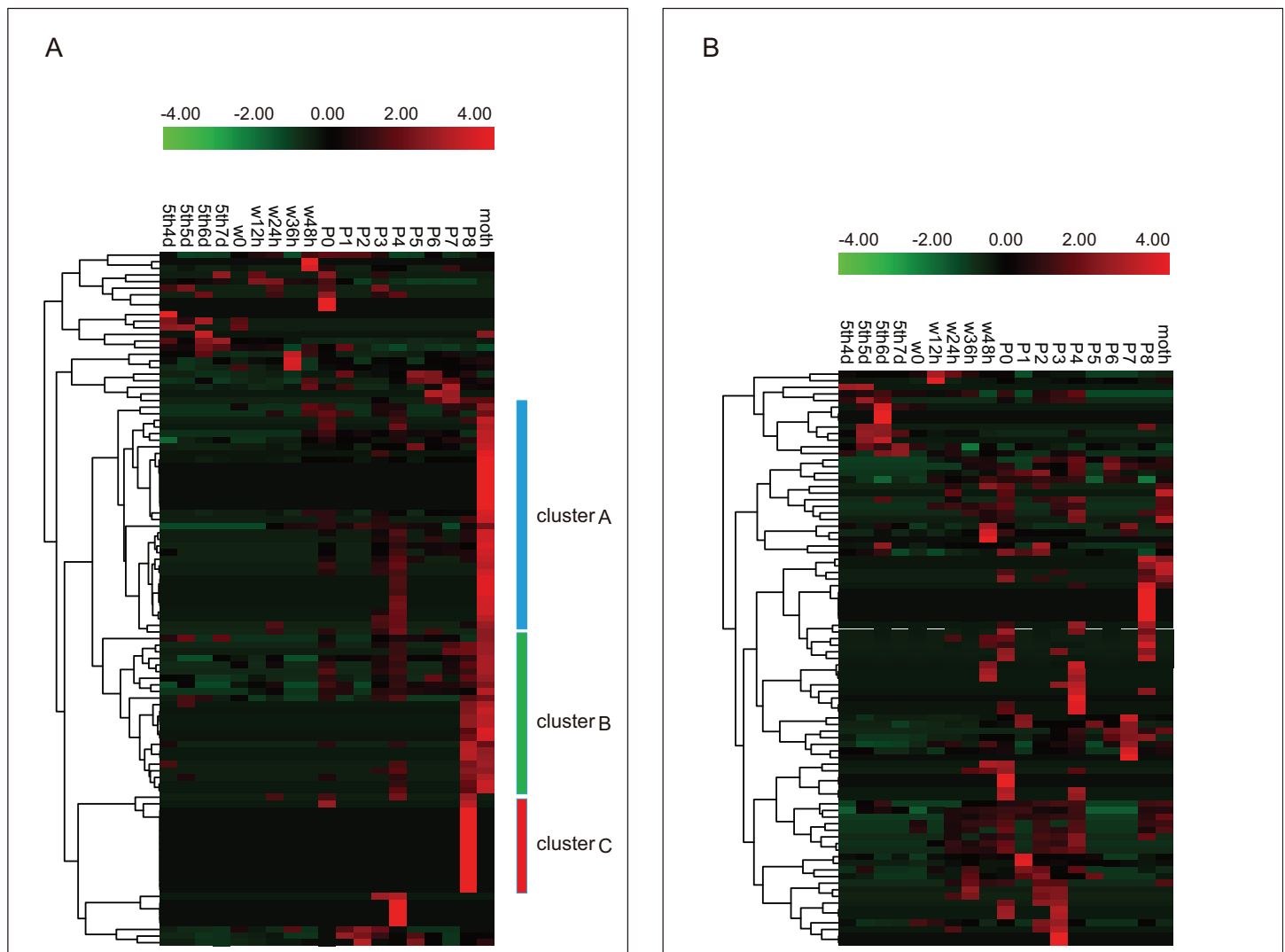
Based on reported microarray data for the silkworm (Xia et al., 2007), we investigated the expression of BmZNF protein genes in multiple tissues of silkworm larvae on day 3 of the fifth instar. We found that 249 of the 327 BmZNF genes were expressed in at least



**Figure 4** Expression profiling of *BmZNF* genes in multiple tissues of the silkworm. Each tissue sample was analyzed using at least two biological repeats. The heat map was constructed using the Pheatmap program of the R package. Clusters A–G are indicated by different colors. A/MSG, anterior/median silk gland; PSG, posterior silk gland; F, female; M, male. The expression profiling of *BmZNF* from the fourth day of the fifth instar to the adult stage. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02\_img.jpg\) DOI: 10.7717/peerj.7222/fig-4](https://doi.org/10.7717/peerj.7222/fig-4)

one larval tissue during this stage. Notably, the genes of cluster A were highly expressed in testis, and those of cluster F were highly expressed in ovary. In addition, the *BmZNF* genes of cluster B were specifically expressed in both ovary and testis. The genes of cluster C were highly expressed in the midgut of silkworm larvae. The genes in cluster D had relatively high expression levels in the silk gland. The genes of cluster E were highly expressed in the Malpighian tubules. Moreover, the genes of cluster G were highly expressed in the head, integument and testis of silkworms (Fig. 4). The results showed that C2H2-type zinc finger protein genes of the silkworm appeared to have significant tissue-specific expression. This suggests that *BmZNF* genes assume regulatory functions in the development of different tissues.





**Figure 5** Expression profiling of *BmZNF* genes during development stages of the silkworm. (A) A total of 105 *BmZNF* genes are expressed in females. The clusters A, B and C are genes that are highly expressed in the late pupal stage and the adult stage (B) 87 *BmZNF* genes are expressed in males. [Full-size !\[\]\(fcc3264021d438d9732560e78099f674\_img.jpg\) DOI: 10.7717/peerj.7222/fig-5](https://doi.org/10.7717/peerj.7222/fig-5)

### Developmental expression profile of *BmZNF* protein genes

Using the metamorphosis microarray data of the silkworm from the later larval to the adult stages, the developmental expression profile of the silkworm C2H2-type zinc finger protein genes was surveyed. There were 105 and 87 *BmZNF* protein genes that were expressed during silkworm metamorphosis in females and males, respectively (Table S4). Notably, the genes of cluster A were highly expressed on day 1 of the adult moth; cluster B genes were highly expressed at 8 days of the pupal phase and on day 1 of the adult moth, and the genes in cluster C were specifically expressed at 8 days of the pupal phase in female silkworms (Fig. 5A). Interestingly, the above expression pattern of *BmZNF* genes was significantly different in males (Fig. 5B).

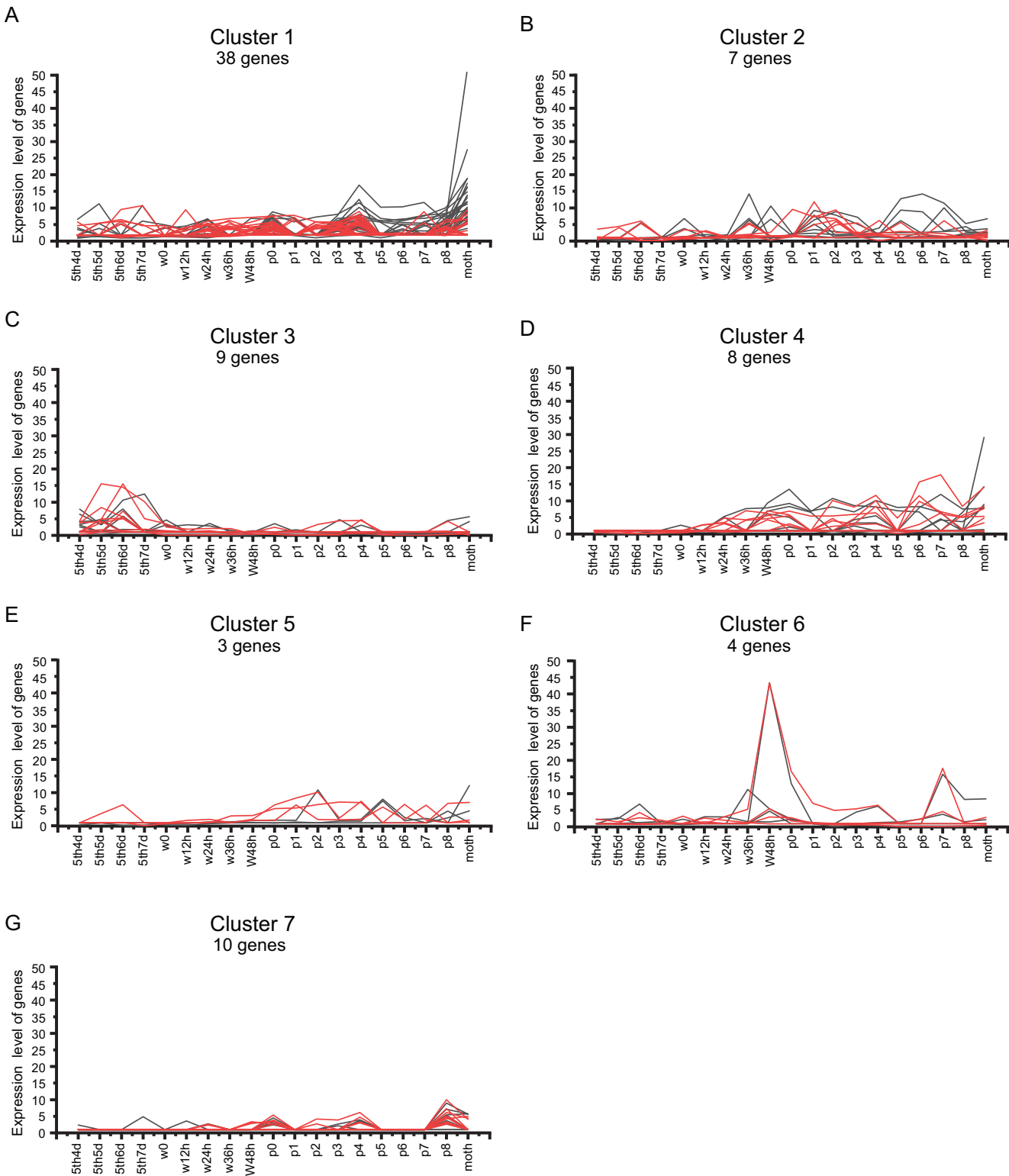
We analyzed the *BmZNF* genes that were expressed in both sexes. As shown in Fig. 6, 79 overlapping genes in females and males were grouped into seven clusters by *K*-means and *K*-medians clustering analysis using the MultiExperiment Viewer 4.90 software (Saeed *et al.*, 2003). The genes of cluster 1 had a wavelike expression pattern in the larval to adult stages, and these genes were up-regulated in the adult. The genes of cluster 2 had a similar fluctuating expression pattern as in cluster 1, but these genes were not remarkable up-regulated in the adult moth stage. Notably, the genes of cluster 3 were mainly up-regulation expressed during the later larval stages and had low expression levels from the wandering larval to adult stages. Interestingly, the genes of cluster 4 were up-regulation expressed from wandering larvae to adult, but there was a lower expression in the early larval stages. The genes of cluster 5 were up-regulation expressed from the beginning of pupation to the adult stage, with a low expression at the early larval and wandering larval stages, except for sw01846 (*BmZNF294*) which was up-regulation expressed at day 5 of the fifth instar to day 6 of fifth instar larvae in males. Furthermore, the genes of cluster 6 displayed peak expression at 48 h after the beginning of the wandering larval stage (W48h) and on day 7 of the pupal stage (P7), and the genes of cluster 7 displayed peak expression on P0 and P8. These times were during key periods of pupation and eclosion. The 79 overlapping genes exhibited seven different expression patterns. The expression patterns suggest that the gene groups may participate in different phases of metamorphosis. For example, genes of cluster 3 were expressed at the later larval stages. The genes of cluster 4 mainly contributed to pupation and eclosion. The genes of clusters 6 and 7 may function as activating factors in critical developmental stages of metamorphosis.

### Analysis of *BmZNF* gene expression in ovarian developmental period

From the above analysis, found that many *BmZNF* genes were specifically significantly up-regulated expression in females. We speculate that their up-regulated expression may be involved in the regulation of female reproduction. Five of these genes were selected, and their expression levels were investigated in the ovarian tissues of silkworm from day 3 of 5th instar to adult, by qRT-PCR experiment. The *BGIBMGA002091* (*CTCF*) and *BGIBMGA006492* (*fru*) gene shared a similar rising expression pattern, during the development of ovary. Moreover, the expression level of *BGIBMGA002091* (*CTCF*) is significantly higher than that of *BGIBMGA006492* (*fru*). The *BGIBMGA006230* (*wor*) and *BGIBMGA004640* (*lola*) gene have a saddle-shaped expression pattern in ovary, which were up-regulation expressed from the P0 to P3. The *BGIBMGA004569* gene have a wavelike expression pattern, and its expression level is the lowest among the five genes (Fig. 7).

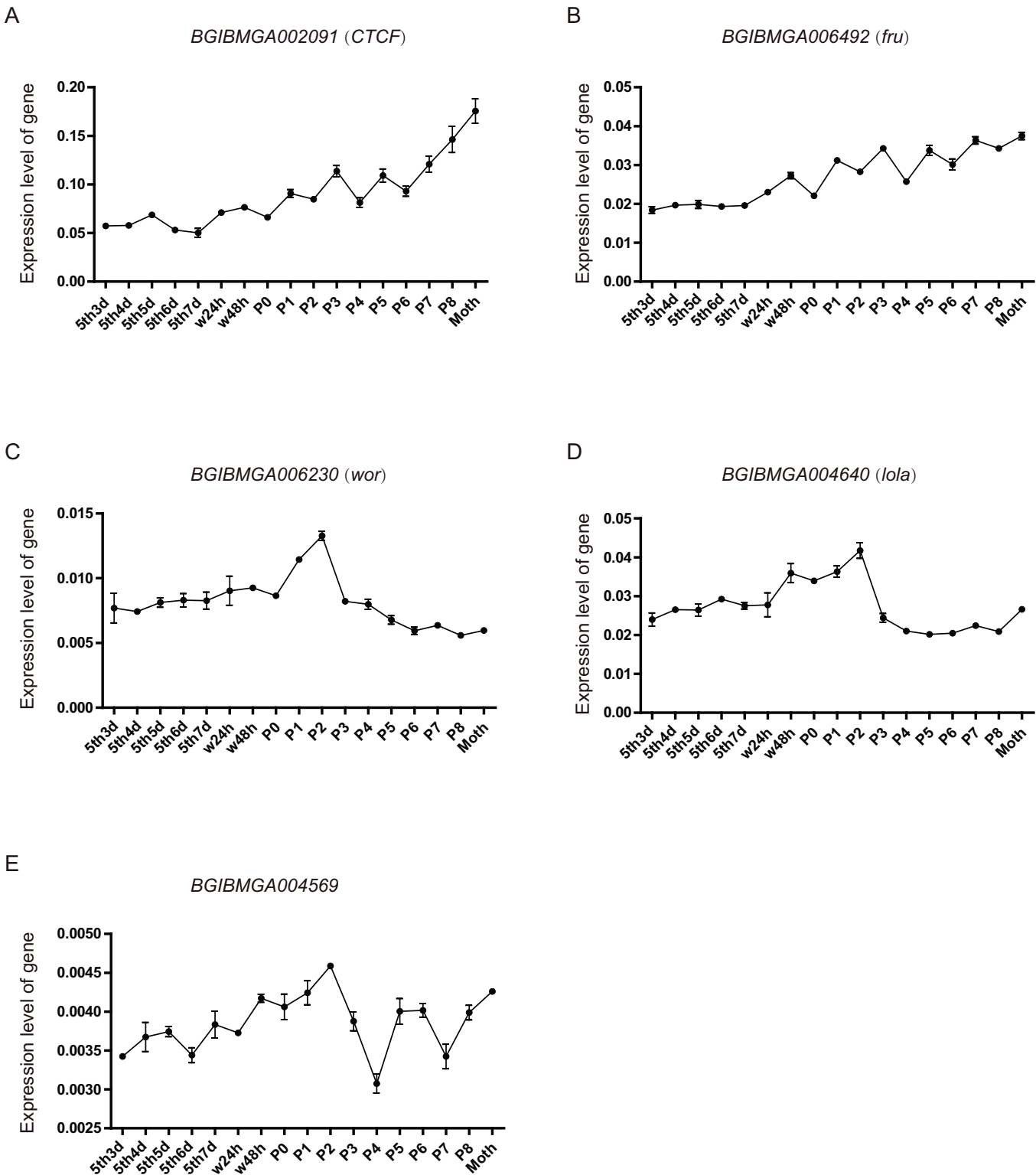
## DISCUSSION

The C2H2-ZF proteins are the most common and diverse type of TFs. In this study, we identified 327, 290, 243, 107, 673, and 1,082 C2H2-type zinc finger protein genes, accounting for approximately 1.9%, 1.9%, 1.6%, 0.4%, 2.6%, and 3.2% of the total gene number in *B. mori*, *Danaus plexippus*, *Drosophila melanogaster*, *C. elegans*, *M. musculus*,



**Figure 6** Cluster analysis of *BmZNF* gene expression in both sexes. (A–G) The *K*-mean/*K*-means cluster method was used to analyze the shared genes that are expressed in males and females. Clusters 1–7 represent different expression patterns of *BmZNF* genes. Red indicates male, black indicates female.

Full-size DOI: [10.7717/peerj.7222/fig-6](https://doi.org/10.7717/peerj.7222/fig-6)



**Figure 7** The expression pattern of ZNF gene in ovary of silkworm. (A–E) The five genes, *BGIBMGA002091 (CTCF)*, *BGIBMGA006492 (fru)*, *BGIBMGA006230 (wor)*, *BGIBMGA004640 (lola)*, and *BGIBMGA004569*, were detected the expression level in silkworm ovary, during day 3 of the fifth instar to adult, by qRT-PCR.

Full-size DOI: [10.7717/peerj.7222/fig-7](https://doi.org/10.7717/peerj.7222/fig-7)

and *H. sapiens*, respectively. This indicates that the ZNF protein genes occupy a large proportion of higher taxa' genome, and with the increase of species complexity, the number of genes in the ZNF family is also increasing. Meanwhile, these genes show obvious differentiation. Even among insect species, more than 100 ZNF genes in *Drosophila* have no orthologous genes in *B. mori* or monarch butterflies (Table S5). Moreover, the zinc finger proteins often contain PPI domains besides the tandem zinc finger domain. We identified 16, 13, 13, 3, 58, and 63 zinc finger proteins containing the BTB domain (BTB-ZF) in *B. mori*, *Danaus plexippus*, *Drosophila melanogaster*, *C. elegans*, *M. musculus*, and *H. sapiens*, respectively. The C2H2-ZF proteins containing KRBA and SCAN are common in mammals but were not found in the three insects and worms. This implies that the BTB domain may be more primitive in evolution compared to the other two PPI domains. Furthermore, these C2H2 zinc finger proteins also harbor other DNA binding domains such as Homeobox (PF00046.28), HTH\_psq (PF05225.15), GAGA (PF09237.10), and BolA (PF01722.17). This implies that the mechanism by which these proteins bind DNA sequences may be more complicated. In addition, the 243 C2H2-type zinc finger protein genes of *Drosophila melanogaster* produced 556 C2H2-ZF proteins due to the fact that most zinc finger protein genes have alternative splicing patterns. Nevertheless, due to incomplete protein data, the total number of proteins translated by C2H2-type zinc finger protein genes is unclear in *B. mori* and *Danaus plexippus*. However, those species have dozens more C2H2-ZF protein genes than *Drosophila melanogaster*; we may surmise that the silkworm and monarch produce more C2H2-ZF proteins than the fruit fly. Previous research has shown that alternative splicing patterns of genes function in different tissues or during different periods (Venables, Tazi & Juge, 2012). This makes the study of zinc finger proteins more complicated. In the future, studies concentrated on a particular period and tissue of the focal organism will be necessary to clarify the functions and evolution of C2H2-ZF protein genes.

The analysis of the tissue expression data shows that C2H2-ZF protein genes exhibit tissue-specific expression patterns in silkworm larvae. Interestingly, over 40% of the *BmZNF* genes were expressed in gonad tissue, in which 28 *BmZNF* genes were specific to the testis, 24 *BmZNF* genes were significantly up-regulated in ovary, and 80 *BmZNF* genes were highly expressed in both male and female gonad tissues. The *BGIBMGA007530* gene is orthologous to *tramtrack*, which plays an important role in regulation of the development of the compound eye (Xiong & Montell, 1993), nerves (Badenhorst, 2001), and trachea (Araujo, Cela & Llimargas, 2007) in fruit flies; a ubiquitous TF *CTCF* (*BGIBMGA002091*) that binds to insulators and domain boundaries (Filippova et al., 1996) was highly expressed in the gonads of both sexes. The *BGIBMGA008740* gene is orthologous to the *grauzone* (*grau*) gene that is expressed in the ovary, and the loss of its function can lead to infertility in *Drosophila*. The expression of this gene in the silkworm was specific to the testis. The *ovo* (*BGIBMGA000988*) gene is an important factor in female germline differentiation (Andrews et al., 2000; Xue et al., 2014). The *BGIBMGA008808* is orthologous to *bowli*, which contributes to regulation of limb development (Ibeas & Bray, 2003). The *BGIBMGA002702* is orthologous to *ken*, which is involved in genital formation in flies (Lukacsovich et al., 2003). The latter three genes were highly expressed in the ovary of the silkworm. These zinc finger proteins, which as mentioned above have been

**Table 1** C2H2 type zinc finger encoding genes and the phenotypes associated with their mutation.

Silkworm symbol	Gene ID	Fly symbol	Phenotypes in fly
<i>BmZNF158</i>	BGIBMGA007530	<i>tramtrack</i>	Compound eye, nerves and trachea
<i>BmZNF4</i>	BGIBMGA002091	<i>CTCF</i>	Gonads
<i>BmZNF273</i>	BGIBMGA008740	<i>grau</i>	Female sterility
<i>BmZNF279</i>	BGIBMGA008800	<i>grau</i>	Female sterility
<i>BmZNF141</i>	BGIBMGA000988	<i>ovo</i>	Female germline differentiation
<i>BmZNF20</i>	BGIBMGA008808	<i>bowl</i>	Limb development
<i>BmZNF45</i>	BGIBMGA002702	<i>ken</i>	Genital
<i>BmZNF306</i>	BGIBMGA013385	<i>dati</i>	Courtship
<i>BmZNF56</i>	BGIBMGA006492	<i>fru</i>	Courtship
<i>BmZNF2</i>	BGIBMGA002047	<i>luna</i>	Larval death
<i>BmZNF71</i>	BGIBMGA008080	<i>noc</i>	Wing and leg
<i>BmZNF34</i>	BGIBMGA006230	<i>wor</i>	Neurogenesis
<i>BmZNF325</i>	BGIBMGA004640	<i>lola</i>	Lethal

well researched in *Drosophila melanogaster*, have an important relationship with tissue development. Therefore, the other *BmZNF* protein genes that are highly expressed in the tissues may also play important roles in the regulation of development. This provides a reference for subsequent functional studies.

Period expression analysis of *BmZNF* genes from fifth-instar 4-day larva to the adult stage shows that some genes have differential expression patterns in the two sexes. These differentially expressed genes may be involved in the development of sexual dimorphism. In females, we found that a group of genes was highly expressed near the adult stage. The genes of cluster III were highly expressed at day 8 of the pupa, the last stage of pupation. For example, the *BGIBMGA013385* is orthologous to the *dati* gene of this cluster; loss function of this gene can cause rejection of male courtship, resulting in female functional sterility in *Drosophila* (Schinaman et al., 2014). The genes of cluster I including *BGIBMGA006492*, an ortholog of *fru* that contributes to courtship behavior (Manoli et al., 2005), were highly expressed at day 1 of the adult moth. Furthermore, the genes of cluster II were highly expressed at day 8 of the pupal stage and day 1 of the adult moth. The gene *BGIBMGA008800* (a paralogue of *BGIBMGA008740*) in this cluster is orthologous to the *grau* gene, of which dysfunction results in female sterility, meiosis II arrest, and egg activation defects (Chen et al., 2000). Based on these results, we can infer that these gene clusters may be related to mating, ovary development and egg laying in female moths.

In addition, we analyzed the genes expressed in both sexes. Cluster analysis of period expression was performed for 79 genes expressed in both sexes. We found that 45 *BmZNF* genes were expressed from the terminal larval stage to the adult stage, and 34 *BmZNF* genes were expressed at a specific stage. The different patterns of expression of these genes imply that they may perform different tasks during metamorphosis. For example, the genes of cluster 3 were notable for high expression at the later stages of larval development followed by low levels of expression. The *BGIBMGA002047* gene is orthologous to the



*luna* gene of cluster 3, a gene that belongs to the Krüppel-like TF family. The *luna* gene plays an important role in *Drosophila* development; mutations of the gene can cause larval death (De Graeve et al., 2003). These genes from cluster 3 may be necessary for larval stage development. Combined with the observation that the genes of cluster 4 were expressed from the wandering larva to the adult stage and the lack of expression in the early larval stages, suggests that these genes may be crucial for pupation and eclosion. The gene *BGIBMGA008080* is orthologous to the *noc* gene in cluster 4, which plays an important role in wing and leg development in the fruit fly (Weihe et al., 2004). The information of genes mentioned above was list in Table 1.

The analysis of period expression identifies the time when these *BmZNF* genes are active and provides a reference for future research on the functions of these genes. Five interesting genes were selected for investigate their expression in the ovaries of 5th instar larvae and adults. The qRT-PCR experiment indicated that the five genes, which up-regulated in females according to microarray data, were expressed in ovaries. This further suggests that these genes may regulate female ovarian development, in silkworm.

## CONCLUSIONS

This study is a comprehensive identification and expression analysis of C2H2 type zinc finger proteins in the silkworm *B. mori*. The results serve as a new resource for studying zinc finger protein TFs that are involved in the development of the silkworm and other species of Lepidoptera. The gonad and metamorphosis associated genes analyzed by our screening can provide references for further functional research of these genes.

## ACKNOWLEDGEMENTS

We would like to thank our reviewers and the editor of PeerJ for their comments, which have improved the manuscript tremendously.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

The research was supported by the National Natural Science Foundation of China (No. 31830094, No. 31472153), the Hi-Tech Research and Development 863 Program of China Grant (No. 2013AA102507), Fundamental Research Funds for the Central Universities (No. XDJK2019C046), and fund of the China Agriculture Research System (No. CARS-18-ZJ0102). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 31830094, 31472153.

Hi-Tech Research and Development 863 Program of China: 2013AA102507.

Fundamental Research Funds for the Central Universities: XDJK2019C046.

China Agriculture Research System: CARS-18-ZJ0102.

## Competing Interests

The authors declare that they have no competing interests.

## Author Contributions

- SongYuan Wu conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Xiaoling Tong authored or reviewed drafts of the paper, approved the final draft.
- ChunLin Li analyzed the data, approved the final draft.
- KunPeng Lu performed the experiments, approved the final draft.
- Duan Tan analyzed the data, approved the final draft.
- Hai Hu approved the final draft.
- Huai Liu authored or reviewed drafts of the paper, approved the final draft.
- FangYin Dai conceived and designed the experiments, authored or reviewed drafts of the paper, approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

Raw data is available in the [Supplemental Files](#).

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.7222#supplemental-information>.

## REFERENCES

- Andrews J, Garcia-Estefania D, Delon I, Lu JN, Mevel-Ninio M, Spierer A, Payre F, Pauli D, Oliver B. 2000. OVO transcription factors function antagonistically in the *Drosophila* female germline. *Development* 127:881–892.
- Araujo SJ, Cela C, Llimargas M. 2007. Tramtrack regulates different morphogenetic events during *Drosophila* tracheal development. *Development* 134(20):3665–3676 DOI 10.1242/dev.007328.
- Badenhorst P. 2001. Tramtrack controls glial number and identity in the *Drosophila* embryonic CNS. *Development* 128(20):4093–4101.
- Chauhan S, Goodwin JG, Chauhan S, Manyam G, Wang J, Kamat AM, Boyd DD. 2013. ZKSCAN3 is a master transcriptional repressor of autophagy. *Molecular Cell* 50(1):16–28 DOI 10.1016/j.molcel.2013.01.024.
- Chen B, Harms E, Chu TY, Henrion G, Strickland S. 2000. Completion of meiosis in *Drosophila* oocytes requires transcriptional control by Grauzone, a new zinc finger protein. *Development* 127:1243–1251.
- De Graeve F, Smaldone S, Laub F, Mlodzik M, Bhat M, Ramirez F. 2003. Identification of the *Drosophila* progenitor of mammalian Kruppel-like factors 6 and 7 and a determinant of fly development. *Gene* 314:55–62 DOI 10.1016/s0378-1119(03)00720-0.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5):1792–1797 DOI 10.1093/nar/gkh340.

- Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, Neiman PE, Collins SJ, Lobanenkov VV. 1996. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian *c-myc* oncogenes. *Molecular and Cellular Biology* **16**(6):2802–2813 DOI [10.1128/mcb.16.6.2802](https://doi.org/10.1128/mcb.16.6.2802).
- Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* **39**(suppl):W29–W37 DOI [10.1093/nar/gkr367](https://doi.org/10.1093/nar/gkr367).
- Goldsmith MR, Shimada T, Abe H. 2005. The genetics and genomics of the silkworm, *Bombyx mori*. *Annual Review of Entomology* **50**:71–100.
- Gramates LS, Marygold SJ, Dos Santos G, Urbano JM, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, Falls K, Goodman JL, Hu Y, Ponting L, Schroeder AJ, Strelets VB, Thurmond J, Zhou P, The FlyBase Consortium. 2017. FlyBase at 25: looking to the future. *Nucleic Acids Research* **45**(D1):D663–D671 DOI [10.1093/nar/gkw1016](https://doi.org/10.1093/nar/gkw1016).
- Gupta A, Christensen RG, Bell HA, Goodwin M, Patel RY, Pandey M, Enuameh MS, Rayla AL, Zhu C, Thibodeau-Beganny S, Brodsky MH, Joung JK, Wolfe SA, Stormo GD. 2014. An improved predictive recognition model for Cys2-His2 zinc finger proteins. *Nucleic Acids Research* **42**(8):4800–4812 DOI [10.1093/nar/gku132](https://doi.org/10.1093/nar/gku132).
- Hamzeh-Mivehroud M, Moghaddas-Sani H, Rahbar-Shahrouziasl M, Dastmalchi S. 2015. Identifying key interactions stabilizing DOF zinc finger–DNA complexes using in silico approaches. *Journal of Theoretical Biology* **382**:150–159 DOI [10.1016/j.jtbi.2015.06.013](https://doi.org/10.1016/j.jtbi.2015.06.013).
- Ibeas JMD, Bray SJ. 2003. Bowl is required downstream of Notch for elaboration of distal limb patterning. *Development* **130**(24):5943–5952 DOI [10.1242/dev.00833](https://doi.org/10.1242/dev.00833).
- Jayakanthan M, Muthukumar J, Chandrasekar S, Chawla K, Punetha A, Sundar D. 2009. ZifBASE: a database of zinc finger proteins and associated resources. *BMC Genomics* **10**(1):421 DOI [10.1186/1471-2164-10-421](https://doi.org/10.1186/1471-2164-10-421).
- Klug A. 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry* **79**(1):213–231 DOI [10.1146/annurev-biochem-010909-095056](https://doi.org/10.1146/annurev-biochem-010909-095056).
- Klug A, Schwabe JW. 1995. Protein motifs 5. Zinc fingers. *FASEB Journal* **9**(8):597–604 DOI [10.1096/fasebj.9.8.7768350](https://doi.org/10.1096/fasebj.9.8.7768350).
- Krebs CJ, Zhang D, Yin L, Robins DM. 2014. The KRAB zinc finger protein RSL1 modulates sex-biased gene expression in liver and adipose tissue to maintain metabolic homeostasis. *Molecular and Cellular Biology* **34**(2):221–232 DOI [10.1128/mcb.00875-13](https://doi.org/10.1128/mcb.00875-13).
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**(7):1870–1874 DOI [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Lai KP, Chen JW, He M, Ching AKK, Lau C, Lai PBS, To KF, Wong N. 2014. Overexpression of ZFX confers self-renewal and chemoresistance properties in hepatocellular carcinoma. *International Journal of Cancer* **135**(8):1790–1799 DOI [10.1002/ijc.28819](https://doi.org/10.1002/ijc.28819).
- Lam KN, Van Bakel H, Cote AG, Van Der Ven A, Hughes TR. 2011. Sequence specificity is obtained from the majority of modular C2H2 zinc-finger arrays. *Nucleic Acids Research* **39**(11):4680–4690 DOI [10.1093/nar/gkq1303](https://doi.org/10.1093/nar/gkq1303).
- Latchman DS. 1990. Eukaryotic transcription factors. *Biochemical Journal* **270**:281–289.
- Lukacsovich T, Yuge K, Awano W, Asztalos Z, Kondo S, Juni N, Yamamoto D. 2003. The ken and barbie gene encoding a putative transcription factor with a BTB domain and three zinc finger motifs functions in terminalia development of *Drosophila*. *Archives of Insect Biochemistry and Physiology* **54**(2):77–94 DOI [10.1002/arch.10105](https://doi.org/10.1002/arch.10105).

- Ma X, Huang M, Wang Z, Liu B, Zhu Z, Li C. 2016. ZHX1 inhibits gastric cancer cell growth through inducing cell-cycle arrest and apoptosis. *Journal of Cancer* 7(1):60–68 DOI 10.7150/jca.12973.
- Mannini R, Riviaccio V, D'Auria S, Tanfani F, Ausili A, Facchiano A, Pedone C, Grimaldi G. 2006. Structure/function of KRAB repression domains: structural properties of KRAB modules inferred from hydrodynamic, circular dichroism, and FTIR spectroscopic analyses. *Proteins: Structure, Function, and Bioinformatics* 62(3):604–616 DOI 10.1002/prot.20792.
- Manoli DS, Foss M, Villella A, Taylor BJ, Hall JC, Baker BS. 2005. Male-specific fruitless specifies the neural substrates of *Drosophila* courtship behaviour. *Nature* 436(7049):395–400 DOI 10.1038/nature03859.
- McCarty AS, Kleiger G, Eisenberg D, Smale ST. 2003. Selective dimerization of a C2H2 zinc finger subfamily. *Molecular Cell* 11(2):459–470 DOI 10.1016/s1097-2765(03)00043-1.
- Meng X, Zhu F, Chen K. 2017. Silkworm: a promising model organism in life science. *Journal of Insect Science* 17(5):97 DOI 10.1093/jisesa/iex064.
- Persikov AV, Singh M. 2014. De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins. *Nucleic Acids Research* 42(1):97–108 DOI 10.1093/nar/gkt890.
- Ponting CP, Schultz J, Milpetz F, Bork P. 1999. SMART: identification and annotation of domains from signalling and extracellular protein sequences. *Nucleic Acids Research* 27(1):229–232 DOI 10.1093/nar/27.1.229.
- Povey S, Lovering R, Bruford E, Wright M, Lush M, Wain H. 2001. The HUGO gene nomenclature committee (HGNC). *Human Genetics* 109(6):678–680 DOI 10.1007/s00439-001-0615-0.
- R Development Core Team. 2015. *R: a language and environment for statistical computing*. Vol. 1. Vienna: The R Foundation for Statistical Computing, 12–21. Available at <http://www.R-project.org/>.
- Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF, Joung JK. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nature Methods* 5(5):374–375 DOI 10.1038/nmeth0508-374.
- Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003. TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* 34(2):374–378 DOI 10.2144/03342mt01.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4):406–425.
- Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research* 32(90001):D91–D94 DOI 10.1093/nar/gkh012.
- Sander JD, Zaback P, Joung JK, Voytas DF, Dobbs D. 2009. An affinity-based scoring scheme for predicting DNA-binding activities of modularly assembled zinc-finger proteins. *Nucleic Acids Research* 37(2):506–515 DOI 10.1093/nar/gkn962.
- Schinaman JM, Giesey RL, Mizutani CM, Lukacsovich T, Sousa-Neves R. 2014. The KRUPPEL-like transcription factor DATILOGRAFO is required in specific cholinergic neurons for sexual receptivity in *Drosophila* females. *PLOS Biology* 12(10):e1001964 DOI 10.1371/journal.pbio.1001964.
- Schumacher C, Wang H, Honer C, Ding W, Koehn J, Lawrence Q, Coulis CM, Wang LL, Ballinger D, Bowen BR, Wagner S. 2000. The SCAN domain mediates selective

oligomerization. *Journal of Biological Chemistry* 275(22):17173–17179

DOI 10.1074/jbc.m000119200.

- Shimomura M, Minami H, Suetsugu Y, Ohyanagi H, Satoh C, Antonio B, Nagamura Y, Kadono-Okuda K, Kajiwara H, Sezutsu H, Nagaraju J, Goldsmith MR, Xia Q, Yamamoto K, Mita K. 2009.** KAIKObase: an integrated silkworm genome database and data mining tool. *BMC Genomics* 10(1):486 DOI 10.1186/1471-2164-10-486.
- Sommer RJ, Retzlaff M, Goerlich K, Sander K, Tautz D. 1992.** Evolutionary conservation pattern of zinc-finger domains of *Drosophila* segmentation genes. *Proceedings of the National Academy of Sciences of the United States of America* 89(22):10782–10786 DOI 10.1073/pnas.89.22.10782.
- Sonnhammer EL, Eddy SR, Durbin R. 1997.** Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins: Structure, Function, and Genetics* 28(3):405–420 DOI 10.1002/(sici)1097-0134(199707)28:3<405::aid-prot10>3.0.co;2-l.
- Suetsugu Y, Futahashi R, Kanamori H, Kadono-Okuda K, Sasanuma S, Narukawa J, Ajimura M, Jouraku A, Namiki N, Shimomura M, Sezutsu H, Osanai-Futahashi M, Suzuki MG, Daimon T, Shinoda T, Taniai K, Asaoka K, Niwa R, Kawaoka S, Katsuma S, Tamura T, Noda H, Kasahara M, Sugano S, Suzuki Y, Fujiwara H, Kataoka H, Arunkumar KP, Tomar A, Nagaraju J, Goldsmith MR, Feng Q, Xia Q, Yamamoto K, Shimada T, Mita K. 2013.** Large scale full-length cDNA sequencing reveals a unique genomic landscape in a lepidopteran model insect, *Bombyx mori*. *G3: Genes, Genomes, Genetics* 3(9):1481–1492 DOI 10.1534/g3.113.006239.
- Sun W, Yu H, Shen Y, Banno Y, Xiang Z, Zhang Z. 2012.** Phylogeny and evolutionary history of the silkworm. *Science China Life Sciences* 55(6):483–496 DOI 10.1007/s11427-012-4334-7.
- Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. 2009.** A census of human transcription factors: function, expression and evolution. *Nature Reviews Genetics* 10(4):252–263 DOI 10.1038/nrg2538.
- Venables JP, Tazi J, Juge F. 2012.** Regulated functional alternative splicing in *Drosophila*. *Nucleic Acids Research* 40(1):1–10 DOI 10.1093/nar/gkr648.
- Voorrips RE. 2002.** MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* 93(1):77–78 DOI 10.1093/jhered/93.1.77.
- Weihe U, Dorfman R, Wernet MF, Cohen SM, Milan M. 2004.** Proximodistal subdivision of *Drosophila* legs and wings: the elbow-no ocelli gene complex. *Development* 131(4):767–774 DOI 10.1242/dev.00979.
- Weirauch MT, Hughes TR. 2011.** A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. *Subcellular Biochemistry* 52:25–73 DOI 10.1007/978-90-481-9069-0\_3.
- Wolfe SA, Greisman HA, Ramm EI, Pabo CO. 1999.** Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *Journal of Molecular Biology* 285(5):1917–1934 DOI 10.1006/jmbi.1998.2421.
- Wolfe SA, Nekludova L, Pabo CO. 2000.** DNA recognition by Cys2His2 zinc finger proteins. *Annual Review of Biophysics and Biomolecular Structure* 29:183–212.
- Xia Q, Cheng D, Duan J, Wang G, Cheng T, Zha X, Liu C, Zhao P, Dai F, Zhang Z, He N, Zhang L, Xiang Z. 2007.** Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori*. *Genome Biology* 8(8):R162 DOI 10.1186/gb-2007-8-8-r162.
- Xiong WC, Montell C. 1993.** Tramtrack is a transcriptional repressor required for cell fate determination in the *Drosophila* eye. *Genes & Development* 7(6):1085–1096 DOI 10.1101/gad.7.6.1085.

- Xue R, Hu X, Cao G, Huang M, Xue G, Qian Y, Song Z, Gong C. 2014.** Bmovo-1 regulates ovary size in the silkworm, *Bombyx mori*. *PLOS ONE* **9(8)**:e104928 DOI [10.1371/journal.pone.0104928](https://doi.org/10.1371/journal.pone.0104928).
- Yang JH, Li JH, Jiang S, Zhou H, Qu LH. 2013.** ChIPBase: a database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data. *Nucleic Acids Research* **41(D1)**:D177–D187 DOI [10.1093/nar/gks1060](https://doi.org/10.1093/nar/gks1060).
- Zhan S, Reppert SM. 2013.** MonarchBase: the monarch butterfly genome database. *Nucleic Acids Research* **41(D1)**:D758–D763 DOI [10.1093/nar/gks1057](https://doi.org/10.1093/nar/gks1057).
- Zollman S, Godt D, Prive GG, Couderc JL, Laski FA. 1994.** The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **91(22)**:10717–10721 DOI [10.1073/pnas.91.22.10717](https://doi.org/10.1073/pnas.91.22.10717).