RESEARCH

Identification and Characterization of 30 K Protein Genes Found in *Bombyx mori* (Lepidoptera: Bombycidae) Transcriptome

Xiao-Feng Shi,[1,2] Yi-Nü Li,[1] Yong-Zhu Yi,[3] Xing-Guo Xiao,[2] and Zhi-Fang Zhang[1,4]

[1]The Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, 12 Zhongguancun South Street Beijing, 100081, China
[2]The College of Biological Sciences, China Agricultural University, Beijing, 100094, China
[3]The Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang, Jiangsu Province, 212018, China
[4]Corresponding author, e-mail: zhifangzhang@yahoo.com

Subject Editor: Marian Goldsmith

J. Insect Sci. *(2015)* 15(1): 71; DOI: 10.1093/jisesa/iev057

ABSTRACT. The 30 K proteins, the major group of hemolymph proteins in the silkworm, *Bombyx mori* (Lepidoptera: Bombycidae), are structurally related with molecular masses of ~30 kDa and are involved in various physiological processes, e.g., energy storage, embryonic development, and immune responses. For this report, known 30 K protein gene sequences were used as Blastn queries against sequences in the *B. mori* transcriptome (SilkTransDB). Twenty-nine cDNAs (*Bm30K-1–29)* were retrieved, including four being previously unidentified in the Lipoprotein_11 family. The genomic structures of the 29 genes were analyzed and they were mapped to their corresponding chromosomes. Furthermore, phylogenetic analysis revealed that the 29 genes encode three types of 30 K proteins. The members increased in each type is mainly a result of gene duplication with the appearance of each type preceding the differentiation of each species included in the tree. Real-Time Quantitative Polymerase Chain Reaction (Q-PCR) confirmed that the genes could be expressed, and that the three types have different temporal expression patterns. Proteins from the hemolymph was separated by SDS-PAGE, and those with molecular mass of ~30 kDa were isolated and identified by mass spectrometry sequencing in combination with searches of various databases containing *B. mori* 30K protein sequences. Of the 34 proteins identified, 13 are members of the 30 K protein family, with one that had not been found in the SilkTransDB, although it had been found in the *B. mori* genome. Taken together, our results indicate that the 30 K protein family contains many members with various functions. Other methods will be required to find more members of the family.

Key Words: gene family, phylogenetic analysis, temporal expression pattern

The most abundant proteins isolated from the silkworm, *Bombyx mori* (Lepidoptera: Bombycidae), hemolymph from the fifth instar to pupation stages have a molecular mass of ~30 kDa and very similar nucleotide and amino acid sequences. Therefore, this group of *B. mori* lipoproteins has been denoted the 30 K proteins (30KPs) and belong to the Lipoprotein_11 family (Gamo 1978, Tojo et al. 1980, Izumi et al. 1981, Zhu et al. 1986, Sakai et al. 1988, Mori et al. 1991, Kishimoto et al. 1999, Ogawa et al. 2005, Zhong et al. 2005, Hou et al. 2010). The proteins are mainly synthesized in the *B. mori* fat body and then secreted into the hemolymph during the last instar larval stage (Izumi et al. 1981). During pupation, 30KPs in the hemolymph accumulate to a great extent and then are gradually absorbed into oocytes (Chen and Yamashita 1990). Their synthesis is regulated by the juvenile hormone (Sakai et al. 1988, Mori et al. 1991, Ogawa et al. 2005), and, therefore, may be an ideal model for the study of the regulation of insect gene expression.

30KPs act as storage proteins during the growth and development of *B. mori*—representing ~35% (w/w) of the oocyte yolk protein—but are used only during embryonic development (Zhu et al. 1986). The 30KPs may act as storage units of amino acids for use in de novo synthesis of other proteins during embryonic development (Izumi et al. 1981, Mine et al. 1983).

30KPs may also inhibit programmed cell death (Rhee et al. 2002) and thereby prolong cell survival. Kim et al. found that the fraction isolated from hemolymph which contained 30KPs inhibits apoptosis in insect cells (Sf9) infected with baculovirus (AcNPV) most strongly (Kim et al. 2001). The 30KP LP1 has been expressed in *Escherichia coli*, and shown to inhibit apoptosis in insect and human cells when added into their culture medium and that recombinant LP1 also inhibits virally and chemically induced apoptosis, with a greater potency than that of *B. mori* larva hemolymph (Park et al. 2003). Kim et al. expressed the 30 K protein (30Kc6, GenBank accession number: X54735) in mammalian HEK293 cells and CHOK1 cells, and the expression of 30Kc6 inhibited apoptosis comparably to that of whole silkworm hemolymph (Kim et al. 2004). Two 30KPs have been expressed in baculovirus, and exhibit biological activities similar to the naturally occurring 30KPs isolated from the *B. mori*, including an inhibitory effect against H_2O_2-induced apoptosis (Yu et al. 2013).

As storage proteins, 30KPs are also involved in energy transport and metabolic processes, e.g., the release of diacylglycerol and the transport of steroids and hormones (Chapman 1980). In general, 30KPs can be conjugated to glucose, dextran, maltose, and glycoproteins and, as such, may be a defense mechanism (Ujita et al. 2005, Ueno et al. 2006). A *B. mori* innate immune system mechanism against fungi involves interaction of 30KPs binding to fungal β-glucans, which activates prophenoloxidase thereby interfering with the growth of hyphal so as to protect the *B. mori* from fungal infestation (Ujita et al. 2005).

Sun et al. identified 10 *B. mori* 30KP genes (*Bmlp1–10*) (Sun et al. 2007), in which five have been expressed in *E. coli.* The crystal structure of Bmlp7 has been determined and shown to be a new member of the β-sheet superfamily (Yang et al. 2011). The *B. mori* 30KP sequences are similar to those of microvitellogenin in *Manduca sexta* (Wang et al. 1989). Zhang et al. found 73 genes for which their protein products may be members of the lepidopteran-specific Lipoprotein_11 family in 12 lepidopteran species, of which 46 are specific to the *B. mori.* Sequencing of these genes classified them into three groups: typical 30KPs, serine/threonine-rich (S/T-rich) 30KPs, and ENF peptide-binding proteins (ENF-BPs) (Zhang et al. 2012). The typical 30KP genes are mainly

expressed in the fat body, and the larval and pupal epidermis, whereas the ENF-BPs are mostly expressed in blood cells. In addition, the S/T-rich 30KPs are abundant in mature testis, indicating that they may be involved in the formation of *B. mori* sperm (Zhang et al. 2012).

In this study, the sequences encoding the C-terminal conserved domain or the entire coding sequence of known 30KP genes from GenBank were used to perform a local BLASTN against the SilkTransDB. Twenty-nine cDNA sequences (*Bm30K-1–29*) were retrieved, with four being previously unidentified. The expression patterns of these genes in the fat body of fifth-instar, spinning-stage larvae, and pupae were assessed by Q-PCR. Hemolymph collected from the fifth instar to pupa stages were separated by SDS-PAGE, and the identities of bands at 30 kDa were identified by mass spectrometry.

Materials and Methods

Materials. The *B. mori* strain JY-1 was housed at the Sericultural Research Institute, Chinese Academy of Agricultural Sciences (Jiangsu province, China) at 25°C, under 70–80% relative humidity (Lu 1991), and feed fresh mulberry leaves. Hemolymph was collected from larvae starting on the first day of the fifth instar, and collection continued until the second day of pupation at 1-d intervals. Each sample of hemolymph (from five larvae or pupae at least) was put into test tubes containing a small amount of phenylthiourea. After removing the hemocytes by centrifugation at 12,000 × *g* and 4°C for 10 min, the hemolymph was stored at –20°C. Each sample of fat bodies were scraped away from five dissected *B. mori* larvae, washed with physiological saline solution and DEPC (Diethypyrocarbonate) treated H_2O, and then stored at –80°C.

Identification and Phylogenetic Analysis of 30KP Genes. To retrieve *B. mori* cDNAs encoding the 30KPs in the SilkTransDB (http://124.17.27.136/gbrowse2/), a local BLASTN search was performed using, as the queries, the sequences of the C-terminal conserved domain or entire of known 30KP gene sequences. The SilkTransDB was constructed from high-throughput RNA-Seq data and as such serves as a transcriptome (Li et al. 2012).

The *B. mori* Genome Database (http://silkworm.genomics.org.cn/) was next searched to identify genomic DNA sequences corresponding to the cDNA sequence for each 30K protein. Amino acid sequence alignments and the phylogenetic tree were created using Clustal X (Thompson et al. 1997) and MEGA 4 (Tamura et al. 2007). The conserved domain of each sequence was retrieved using NCBI conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The SignalP 4.0 Server (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the signal peptide sequences (Petersen et al. 2011).

RNA Extraction and Q-PCR. Total RNA was extracted with TRIZOL reagent (Invitrogen, Beijing, China) according to the manufacturer's instructions. Total RNA (1μg) from each sample was used as the template for first-strand cDNA synthesis by M-MLV reverse transcriptase (Promega, Beijing, China) according to the manufacturer's instructions. PCR was individually performed using the reverse transcription products (equivalent to 10 ng of the starting RNA) as the templates and the primers listed in Supplementary Table S1. Q-PCR was performed using SYBR Green (Toyobo, Shanghai, China) according to the manufacturer's instructions. The actin expression level was used as the internal control. Each sample's quantitative assay was performed three times independently. To determine the relative expression of each gene, $2^{-\Delta ct}$ ($\Delta ct = ct_{target\ gene} - ct_{Aactin\ A3}$) was calculated (Livak and Schmittgen 2001).

SDS-PAGE of *B. mori* Hemolymph Proteins. *B. mori* hemolymph proteins (about 50μg) described earlier were boiled for 10 min according to Laemmli (1970), then separated through a polyacrylamide gel [12% (w/v) acrylamide], and stained with Coomassie Brilliant Blue R-250.

LC-ESI-MS/MS. Proteins of molecular mass ~30 kDa were extracted from the SDS-PAGE gel and subjected to LC-ESI-MS/MS using an Agilent 1100 HPLC system connected to an LTQ Orbitrap Linear Ion Trap Mass Spectrometer (Thermo Fisher Scientific Inc., Waltham, MA) at Jisizhuoyang Science and Technology Ltd., Beijing, China.

Protein Identification. All MS/MS spectra were searched with MASCOT. The search parameters were set as follows: taxonomy, as the sample; database, NCBI (http://www.ncbi.nlm.nih.gov/), *B. mori* Genome (http://silkworm.genomics.org.cn/), and SilkTransDB databases (http://124.17.27.136/gbrowse2/); enzyme, trypsin; fixed modifications, carbamidomethyl (C); variable modifications, oxidation (M); max missed cleavages, two mass tolerances for MS and MS/MS were 10 ppm and 0.5 Da.

Results

Identification of Genes Encoding 30Ks in the SilkTransDB. The known nucleotide sequences of five 30K proteins were used as BLASTN queries against nucleotide sequences in the SilkTransDB, and 29 sequences were retrieved (*Bm30K-1-Bm30K-29*) (Table 1); Of these sequences, 27 Open Reading Frames (ORFs) were complete. We could not complete the ORF sequences for *Bm30K-28* and *-29* by using RACE. Four genes *Bm30K-21, -27, -28*, and *-29* had not been previously predicted in the *B. mori* genome. The above 29 cDNA sequences have been submitted to GenBank, with accession numbers JN977519-JN977547.

By mapping each of the cDNA sequences to the *B. mori* genome, one intron was found in each of the following genes: *Bm30K-1, -2, -5, -6, -7, -8, -10, -11, -12, -13, -14, -23*, and *-24*. The other genes contain no introns. Three genes (*Bm30K-12*, *-16*, and *-21*) are located on chromosome 22; three genes (*Bm30k-13*, *-18*, and *-28*) are located on chromosome 7; five genes (*Bm30K-14*, *-19*, *-22*, *-23*, and *-27*) are located on the chromosome 24: and the other 18 genes are located on chromosome 20.

B. mori 30KPs have been classified as the lepidopteran-specific Lipoprotein_11 family. According to their structural similarities, the 29 genes could be classified as ENF-BP, typical 30KPs, or S/T-rich 30KPs. Fifteen of the genes (*Bm30K-10, -7, -1, -6, -11, -2, -4, -3, -17, -15, -5, -25, -24, -8,* and *-9*) encode typical 30KPs and all located on the chromosome 20; three (*Bm30K-20, -26,* and *-29*) encode S/T-rich 30KPs; and the remaining 11 genes (*Bm30K-14, -19, -21, -22, -23, -13, -18, -12, -16, -27,* and *-28*) encoded ENF-BPs. All of the typical 30KP genes encode signal peptides, whereas those encoding the ENF-BPs do not. Moreover, the lengths of the encoded ENF-BPs are substantially longer than those of the typical 30KPs and the S/T-rich 30KPs.

Moreover, sequence analysis revealed that the retrotransposon Bm1 repetitive elements exist in 12 of the 30 K genes (*Bm30K-1, -3, -4, -5, -6, -7, -8, -10, -11, -20, -23, -24*) (Supplementary Data).

30KP Phylogenetic Tree. As shown by their amino-acid sequence alignment (Supplementary Fig. S1), the N-terminal region sequences of the proteins are quite different, resulting in their different molecular masses. The phylogenetic tree was constructed using the 30KP amino acid sequences of *B. mori*, *M. sexta, Spodoptera exigua*, and *Mythimna separate*, and the sequences clustered into three branches (Fig. 1). The tree architecture is similar to that built previously (Zhang et al. 2012). The typical 30KPs appear to be the oldest 30KPs and cluster as five groups. Furthermore, the 15 typical 30KPs in *B. mori* cluster as four groups, whereas the two typical 30KPs in *M. sexta* cluster as an single group. The ENF-BPs cluster as two groups, Cluster I contains Bm30K-14, -23, -13, -12, and Bm30K-21; Cluster II contains Bm30K-16, -18, -19, -22, -27, and -28. Four pairs of genes encoding ENF-BPs (*Bm30K-14*/*19*, *Bm30K-22*/*23*, *Bm30K-13*/*18*, *Bm30K-12*/*16*) are found on chromosomes 7, 22, and 24, and, for each pair, one gene belongs to Cluster I and the other to Cluster II (Fig. 2). Three of the genes encoding the newly identified 30 K proteins (*Bm30K-21, -27, and -28*) are located on chromosomes 22, 24, and Chr.7, respectively. These genes may represent duplications of *Bm30K-12, -19*, and *-18*, respectively. Finally, *Bm30K-20*, *Bm30K-26*, and the newly identified *Bm30K-29* are S/T-rich 30 K proteins, form an independent branch of the phylogenetic tree, and are all located on chromosome 20.

Table 1 The gene family encoding 30 K proteins in the silkworm, *Bombyx mori*

Gene name	Chr location	cDNA (ORF) length	Introns	Protein length	Signal peptide[a]	Conserved domain (aa position)	Subfamily	Corresponding gene name
Bm30K-1[d]	Chr 20	1,097 (256–1,026)	1	256	1–20 S*D	8–256	Typical 30KP	*Bmlp3*
Bm30K-2	Chr 20	2,234 (29–799)	1	256	1–17 A*A	8–256	Typical 30KP	*Bmlp7*
Bm30K-3[b,d]	Chr 20	1,425 (22–789)	0	255	1–17 A*N	8–255	Typical 30KP	*Bmlp9*
Bm30K-4[b,d]	Chr 20	908(20–787)	0	255	1–19 G*T	8–255	Typical 30KP	*Bmlp8*
Bm30K-5[d]	Chr 20	1,279 (35–790)	1	251	1–17 A*D	5–251	Typical 30KP	*Bmlp14*
Bm30K-6[d]	Chr 20	1,283 (53–844)	1	263	1–16 A*G	6–263	Typical 30KP	*Bmlp4*
Bm30K-7[b,d]	Chr 20	874(28–822)	1	264	1–16 A*G	6–264	Typical 30KP	*Bmlp2*
Bm30K-8[d]	Chr 20	918(30–818)	1	262	1–16 A*G	6–260	Typical 30KP	*Bmlp20*
Bm30K-9	Chr 20	871(4–804)	0	266	1–16 A*T	35–265	Typical 30KP	*Bmlp21*
Bm30K-10[d]	Chr 20	916(102–872)	1	256	1–19 A*T	17–256	Typical 30KP	*Bmlp1*
Bm30K-11[d]	Chr 20	822(6–806)	1	266	1–21 A*G	11–266	Typical 30KP	*Bmlp6*
Bm30K-12	Chr 22	1,426 (19–1,329)	1	436	NO	210–435	ENF-BP	*Bmlp44*
Bm30K-13	Chr 7	1,368 (15–1,325)	1	436	NO	204–435	ENF-BP	*Bmlp41*
Bm30K-14	Chr 24	1,509 (17–1,327)	1	436	NO	210–435	ENF-BP	*Bmlp37*
Bm30K-15[b]	Chr 20	1,156 (18–788)	0	256	1–23 A*S	16–255	Typical 30KP	*Bmlp13*
Bm30K-16[b]	Chr 22	967(65–925)	0	286	NO	54–285	ENF-BP	*Bmlp45*
Bm30K-17	Chr 20	801(13–777)	0	254	1–19 A*A	24–254	Typical 30KP	*Bmlp10*
Bm30K-18	Chr 7	1,330 (27–1,289)	0	420	NO	188–419	ENF-BP	*Bmlp42*
Bm30K-19	Chr 24	1,472 (58–1,323)	0	421	NO	189–420	ENF-BP	*Bmlp38*
Bm30K-20[b,d]	Chr 20	836 (2–817)	0	271	NO	41–268	S/T-rich 30KP	*Bmlp27*
Bm30K-21	Chr 22	807 (52–771)	0	239	NO	13–238	ENF-BP	
Bm30K-22[b]	Chr 24	1,049 (105–1,031)	0	308	NO	76–307	ENF-BP	*Bmlp39*
Bm30K-23[b,d]	Chr 24	1,362 (13–1,323)	1	436	NO	210–435	ENF-BP	*Bmlp40*
Bm30K-24[d]	Chr 20	1,303 (201–935)	1	244	1–18 A*E	23–244	Typical 30KP	*Bmlp17*
Bm30K-25	Chr 20	888(9–821)	0	270	1–23 A*S	13–270	Typical 30KP	*Bmlp15*
Bm30K-26	Chr 20	1,117 (6–926)	0	306	1–20 A*V	71–295	S/T-rich 30KP	*Bmlp32*
Bm30K-27	Chr 24	979(28–954)	0	308	NO	76–307	ENF-BP	
Bm30K-28	Chr 7	414	0	120	–	1–119 (partial)	ENF-BP	
Bm30K-29	Chr 20	555	0	176	–	7–175 (partial)	S/T-rich 30KP	
Bm30K-30[c]	Chr 20	771	0	256	1–15 A*D	4–253	Typical 30KP	*Bmlp5*

[a]The amino acid position of signal peptide and its cleavage site predicted by SignalP 4.0.
[b]The genes whose starting or ending positions of open-reading frames different than those predicted.
[c]The protein identified by mass spectrometry in our study.
[d]The gene containing Bm1 elements.

Two *M. sexta* protein sequences, which appear in the tree earlier than the *B. mori* ENF-BPs and S/T-rich 30KPs, are closely related to the *B. mori* typical 30KPs. However, they form a cluster independent of the four *B. mori* typical 30KP clusters. The proteins from *S. exigua* and My. separate appeared at the same time in evolutionary history as the *B. mori* ENF-BPs did, and they have the same junction, indicating that they may be most closely related to the ENF-BPs.

Verifying the Expression Level of Genes Encoding the *B. mori* 30 K Proteins by Q-PCR. Q-PCR was used to verify the expression level of *B. mori* 30KPs at different development stages. In general, mRNA expression of the 29 genes in the fat body gradually increased from the fifth-instar to the late-spinning stages, and then decreased during pupation (Fig. 3). On the seventh day of the fifth-instar stage (Fig. 3A), the levels of the 29 mRNAs differed significantly. For the typical 30KPs, the transcriptional levels of *Bm30K-1*, *-2*, *-6*, *-7*, *-10*, and *-11* were the highest, whereas those of *Bm30K-4*, *-5*, *-15*, *-17*, *-25* much lower, and those of *Bm30K-3*, *-8*, *-9*, and *-24* even lower. All of the mRNA levels of the ENF-BPs and S/T-rich 30KPs were lower than those of the typical 30KPs at all assayed stages. During the late spinning stage (Fig. 3B), the expression patterns of the 30KP genes in the fat body are basically the same as those of the seventh day of the fifth-instar. But by the second day of pupation (Fig. 3C), the transcriptional levels of all Bm30K genes had decreased except that for *Bm30K-1*.

Mass Spectrometry of the 30KPs. *B. mori* hemolymph from the first day of the fifth instar to the second day of pupation was collected and subjected to SDS-PAGE. Abundant amounts of protein at ~30 kDa were found (not shown). Additionally, the amounts of 30 kDa proteins gradual accumulated from the fifth instar until the earliest day of pupation, after which the amount of 30 kDa protein decreased, a finding consistent with a previous study (Zhang et al. 2012).

The portion of the gels containing the 30 kDa proteins were excised and subjected to LC-ESI-MS/MS. Thirty-four proteins were identified when MS-determined sequences were used as queries for searches in the NCBI (http://www.ncbi.nlm.nih.gov/), *B. mori* Genome (http://silkworm.genomics.org.cn/), and SilkTransDB databases (http://124.17.27.136/gbrowse2/) (Supplementary Table S2). Of the identified proteins, 13, Bm30K-1, -2, -10, -11, -6, -3, -7, -25, -24, -8, -5, -9, and Bmlp5 (PROTEIN:BGIBMGA004456-PA) are 30KPs, and in which 12 are encoded by the 29 genes found in our study. The 13th protein Bmlp5 is a previously identified *B. mori* member of the Lipoprotein_11 family and is a typical 30KP. Six other proteins (the 27-kDa glycoprotein, antichymotrypsin-2, sex-specific storage-protein 2, trypsin, cytospin-A, and synaptonemal complex protein 1) with molecular mass of ~30 kDa and high scores were also identified. An additional 15 proteins were tentatively identified, but their scores were such that the reliability of their identification was in doubt.

Discussion

Four proteins of 30KPs were purified by chromatography techniques from *B. mori* larval hemolymph by Izumi et al. (1981). Earlier, five genes had been cloned from the mid-fifth instar fat body and expressed at high levels. These genes have similar nucleotide sequences (Sakai et al. 1988). Sun et al. (2007) found 10 genes encoding 30KPs in the *B. mori* genome and named them *Bmlp1–10* according to their similarity with the previously reported 30KP proteins. The sequences of the proteins encoded by the 10 genes contain 246 to 271 residues, with their molecular masses between 28 and 31 kDa, and their pI values between 6.1 and 8.4. The sequence similarity of the 10 proteins is > 60%. After searching the whole *B. mori* genomic sequence using the 30KP sequences as queries, Zhang et al. found 46 genes that might encode *B. mori* Lipoprotein_11 family proteins (*Bmlp1–46*) for which 22 had no

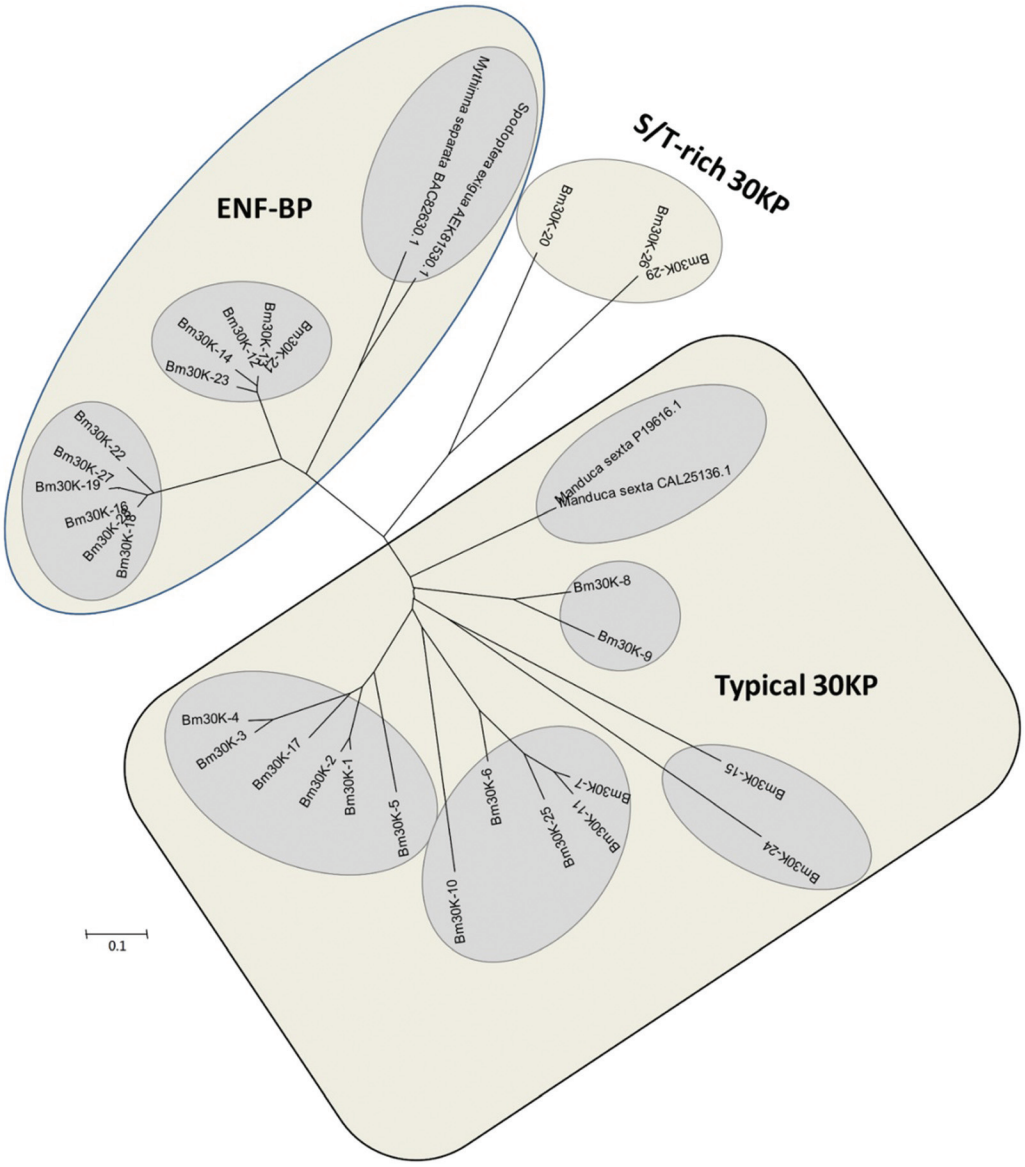

Fig. 1. Phylogenetic tree of sequences of Bm30K proteins and the homologous proteins in other lepidopteron insects. The protein sequences were aligned using Clustal X 1.83. The phylogenetic tree was constructed using neighbor-joining method and with 1,000 bootstrap replicates and displayed with MEGA 4.0 program. Bm: *Bombyx mori*; *Manduca sexta, Spodoptera exigua*, and *Mythimna separate* were three species which have homologous proteins of Bm30K. The typical 30KP, S/T-rich 30KP and ENF-BP represent the three types of 30 K protein gene family, respectively.

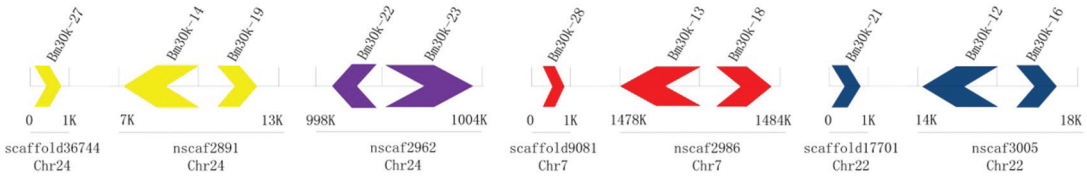

Fig. 2. Genome locations of ENF-BPs of silkworm 30 K protein genes on Chr. 24, 7, and 22. The Fig. 2 was drawn according to the Fig. 3C from Zhang (2012). The different colors referred to different gene pairs (*Bm30K-14/19, 22/23, 13/18, 12/16*). *Bm30K-21, 27, 28* were novel ENF-BPs members. The direction of gene transcription was indicated by the arrow.

Expressed Sequence Tags (EST) expression (Zhang et al. 2012). For this study, we used the sequences encoding the C-terminal conserved domain or the entire coding sequences of known 30KP genes as Blastn queries to search the SilkTransDB, and retrieved 29 30KP genes, of which four genes (*Bm30K-21, -27, -28,* and *-29)* had not been identified previously. The other 25 sequences are consistent with those reported previously (Zhang et al. 2012), with only the starting or ending positions of eight open-reading frames different than those predicted (the detailed information is listed in Table 1).

Our sequence and phylogenetic characterizations support the observation that *B. mori* Lipoprotein_11 family members can be divided into three subfamilies: the typical 30KPs (*Bmlp-1–24*), the ENF-BPs

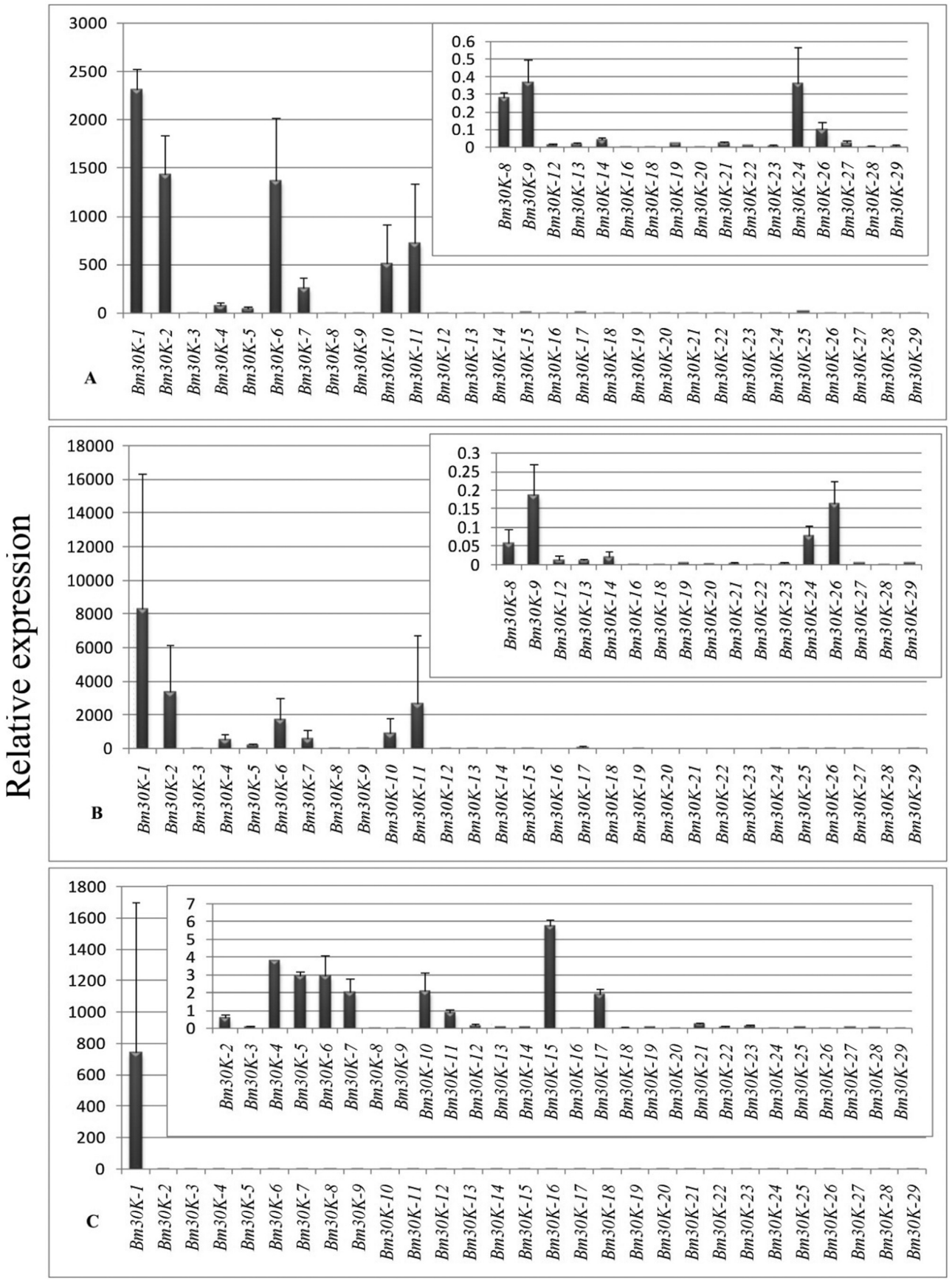

Fig. 3. The expression pattern of Bm30K genes in fat body at the fifth-instar, spinning, and pupation. (**A**) Fat body at the seventh day of fifth-instar; (**B**) Fat body at the late-spinning stage; (**C**) Fat body at the second day of pupation.

(*Bmlp-37–46*), and the S/T-rich 30KPs (*Bmlp-25–36*). Zhang et al. reported that 24 typical *B. mori* 30KP genes can be divided into four gene clusters, with all found in the nscaf2795 region of chromosome 20 (Zhang et al. 2012). The 15 typical 30KP genes that we identified are also all located in this region. In addition, 8 of the 10 S/T-rich 30KP genes are found within 120 kbps of nscaf2795 (2,230–2,350 k) (Zhang et al. 2012). The three S/T-rich 30KP genes found in our study are all confined to this region.

A previous study showed that none of the encoded ENF-BP nucleotide sequences have signal peptides, suggesting that the ENF-BPs are intracellular proteins and performed a conserved function (Matsumoto et al. 2003). In the *B. mori* genome, the four ENF BP gene pairs (*Bm30K-14/Bm30K-19, Bm30K-22/Bm30K-23, Bm30K-13/Bm30K-18, Bm30K-12/Bm30K-16*) are each found together on a single chromosome but in different direction, the three new identified members (*Bm30K-21, Bm30K-27 and Bm30K-28*) are located on chromosomes 22, 24, and 7, respectively. Given their evolutionary distances, *Bm30K-21, -27, and -28* appear to more closely related to one of a gene pair members and may arise later than did the four gene pairs, which suggests *Bm30K-21, -27, and -28* may have arisen by gene duplication of one of a gene pair members. These observations also indicate that the Lipoprotein_11 family has been enlarging with time.

Our phylogenetic study indicates that the typical 30KPs may be the earliest members of the 30KP family and can be clustered into five groups. The ENF-BPs and S/T-rich 30KPs may have arisen later. The S/T-rich 30KPs are an independent branch of the phylogenetic tree, with no homologous genes found in the other lepidoptera, suggesting that the S/T-rich 30KPs may be unique to the *B. mori.* Two homologous *M. sexta* proteins appear to have arisen earlier than the *B. mori* ENF-BPs and S/T-rich 30KPs and are closely related to the typical 30KPs. The proteins from *S. exigua* and My. separate seem to have emerged during the same period as the *B. mori* ENF-BPs and are most closely related to the *B. mori* ENF-BPs. The *B. mori* 30KPs and their homologs in other species may have originated from a common ancestor, and then increased in number through gene duplication. So, they may have different functions, and we suggested the differentiation of the gene families may have occurred earlier than the differentiation of species.

We verified the expression level of the 29 genes that we had found by the Blastn search by Q-PCR of RNA from the *B. mori* fat body during the critical development stages of *B. mori.* Expression levels of the genes differed significantly: all typical 30KP genes, except for *Bm30K-3, -8, -9,* and *-24*, were highly expressed at assayed stages. Conversely, transcription of the ENF-BP and S/T-rich 30KP genes was minimal at all three stages. ENF-BP genes are well expressed in the hemocyte, whereas S/T-rich 30KP genes are expressed in larva at the fourth and fifth instar stages (Zhang et al. 2012). Obviously, the structures and expression patterns of the ENF-BP and S/T-rich 30KP genes are different from those of typical 30KP genes, suggesting that the proteins may have different roles.

Typical 30KP are synthesized in the fat body and then secreted into the hemolymph. They have an N-terminal signal peptide of ~20 residues, which guide them through the membrane. Additionally, their synthesis in the fat body is regulated by the juvenile hormone (Sakai et al. 1988, Mori et al. 1991). ENF-BPs are highly expressed in the hemocyte (Zhang et al. 2012), which explains why we did not observe the existence of the identified ENF-BPs by MS, because the hemocyte was omitted in the samples. S/T-rich 30KPs are expressed in the larva at the fourth and fifth instar stages (Zhang et al. 2012). Again, we did not find the existence of S/T-rich 30KPs by MS, possibly because the times at which we sampled the time point were inappropriate, or because they are not secreted storage proteins. Because 30 kDa proteins of the hemolymph are mainly typical 30KPs, their amounts were greatest at the early pupation stage. However, their mRNA levels were highest at the end of the fifth instar, whereas their mRNA levels had decreased by pupation, which supports the idea that the typical 30KPs act mainly as storage proteins. However, because some of the 30KPs lack signal peptides, they may be not secreted into the hemolymph, and because only the proteins present in the hemolymph were subjected to MS/MS, those lacking a signal peptide would not have been identified in the study.

The *B. mori* 30KP family is very large, with gene sequences and expression patterns that differ substantially, suggesting that the proteins have various functions. By continuing to search the *B. mori* genome for hypothetical 30KPs, searching the transcriptome database, and characterization of 30 kDa proteins by MS, the 30KP family will be fully described. Such comprehensive approaches are needed to fully characterize 30KP gene family.

Supplementary Data

Supplementary data are available at *Journal of Insect Science* online.

Acknowledgments

This work was supported by the National Basic Research Program of China (2012CB114600); and the National Natural Sciences Foundation of China (No. 31070139) and the "863" Project (No.2011AA100603).

References Cited

Chapman, M. 1980. Animal lipoproteins: chemistry, structure, and comparative aspects. J. Lipid Res. 21: 789–853.

Chen, Y., and O. Yamashita. 1990. Nonselective uptake of different 30 kDa plasma proteins by developing ovaries of the silkworm, Bombyx mori. J. Seric. Sci. Jpn. 59: 202–209.

Gamo, T. 1978. Low molecular weight lipoproteins in the haemolymph of the silkworm, *Bombyx mori*: Inheritance, isolation and some properties. Insect Biochem. 8: 457–470.

Hou, Y., Y. Zou, F. Wang, J. Gong, X. Zhong, Q. Xia, and P. Zhao. 2010. Comparative analysis of proteome maps of silkworm hemolymph during different developmental stages. Proteome Sci. 8: 45.

Izumi, S., J. Fujie, S. Yamada, and S. Tomino. 1981. Molecular properties and biosynthesis of major plasma proteins in *Bombyx mori*. Biochim. Biophys. Acta (BBA) Protein Struct. 670: 222–229.

Kim, E. J., W. J. Rhee, and T. H. Park. 2001. Isolation and characterization of an apoptosis-inhibiting component from the hemolymph of *Bombyx mori*. Biochem. Biophys. Res. Commun. 285: 224–228.

Kim, E. J., W. J. Rhee, and T. H. Park. 2004. Inhibition of apoptosis by a *Bombyx mori* gene. Biotechnol. Progr. 20: 324–329.

Kishimoto, A., H. Nakato, S. Izumi, and S. Tomino. 1999. Biosynthesis of major plasma proteins in the primary culture of fat body cells from the silkworm, *Bombyx mori*. Cell Tissue Res. 297: 329–335.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.

Li, Y. N., G. Z. Wang, J. Tian, H. F. Liu, H. P. Yang, Y. Z. Yi, J. H. Wang, X. F. Shi, F. Jiang, B. Yao, et al. 2012. Transcriptome analysis of the silkworm (*Bombyx mori*) by high-throughput RNA sequencing. PLoS One 7: e43713.

Livak, K. J., and T. D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ method. Methods 25: 402–408.

Lu, H. 1991. The sericultural science in China. Shanghai Scientific and Technical Publishers, Shanghai, China.

Matsumoto, Y., Y. Oda, M. Uryu, and Y. Hayakawa. 2003. Insect cytokine growth-blocking peptide triggers a termination system of cellular immunity by inducing its binding protein. J. Biol. Chem. 278: 38579–38585.

Mine, E., S. Izumi, M. Katsuki, and S. Tomino. 1983. Developmental and sex-dependent regulation of storage protein synthesis in the silkworm, *Bombyx mori*. Dev. Biol. 97: 329–337.

Mori, S., S. Izumi, and S. Tomino. 1991. Structures and organization of major plasma protein genes of the silkworm *Bombyx mori*. J. Mol. Biol. 218: 7–12.

Ogawa, N., A. Kishimoto, T. Asano, and S. Izumi. 2005. The homeodomain protein PBX participates in JH-related suppressive regulation on the expression of major plasma protein genes in the silkworm, *Bombyx mori*. Insect Biochem. Mol. Biol. 35: 217–229.

Park, H. J., E. J. Kim, T. Y. Koo, and T. H. Park. 2003. Purification of recombinant 30 K protein produced in *Escherichia coli* and its anti-apoptotic effect in mammalian and insect cell systems. Enzyme Microb. Tech. 33: 466–471.

Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8: 785–786.

Rhee, W. J., E. J. Kim, and T. H. Park. 2002. Silkworm hemolymph as a potent inhibitor of apoptosis in Sf9 cells. Biochem. Biophys. Res. Commun. 295: 779–783.

Sakai, N., S. Mori, S. Izumi, K. Haino-Fukushima, T. Ogura, H. Maekawa, and S. Tomino. 1988. Structures and expression of mRNAs coding for major plasma proteins of *Bombyx mori*. Biochim. Biophys. Acta 949: 224–232.

Sun, Q., P. Zhao, Y. Lin, Y. Hou, Q. Y. Xia, and Z. H. Xiang. 2007. Analysis of the structure and expression of the 30 K protein genes in silkworm, *Bombyx mori*. Insect Sci. 14: 5–14.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596–1599.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.

Tojo, S., M. Nagata, and M. Kobayashi. 1980. Storage proteins in the silkworm, *Bombyx mori*. Insect Biochem. 10: 289–303.

Ueno, Y., N. He, M. Ujita, K. Yamamoto, Y. Banno, H. Fujii, and Y. Aso. 2006. Silkworm midgut proteins interacting with a hemolymph protease inhibitor, CI-8. Biosci. Biotech. Biochem. 70: 1557–1563.

Ujita, M., Y. Katsuno, I. Kawachi, Y. Ueno, Y. Banno, H. Fujii, and A. Hara. 2005. Glucan-binding activity of silkworm 30-kDa apolipoprotein and its involvement in defense against fungal infection. Biosci. Biotech. Biochem. 69: 1178–1185.

Wang, X. Y., K. D. Cole, and J. H. Law. 1989. The nucleotide sequence of a microvitellogenin encoding gene from the tobacco hornworm, *Manduca sexta*. Gene 80: 259–268.

Yang, J. P., X. X. Ma, Y. X. He, W. F. Li, Y. Kang, R. Bao, Y. Chen, and C. Z. Zhou. 2011. Crystal structure of the 30 K protein from the silkworm *Bombyx mori* reveals a new member of the beta-trefoil superfamily. J. Struct. Biol. 175: 97–103.

Yu, W., Q. Li, Y. Yao, Y. Quan, and Y. Zhang. 2013. Two Novel 30 K proteins overexpressed in baculovirus system and their antiapoptotic effect in insect and mammalian cells. Int. J. Genomics 2013: 323592.

Zhang, Y., Z. Dong, S. Liu, Q. Yang, P. Zhao, and Q. Xia. 2012. Identification of novel members reveals the structural and functional divergence of lepidopteran-specific Lipoprotein_11 family. Funct. Integr. Genomics 12: 705–715.

Zhong, B. X., J. K. Li, J. R. Lin, J. S. Liang, S. K. Su, H. S. Xu, H. Y. Yan, P. B. Zhang, and H. Fujii. 2005. Possible effect of 30 K proteins in embryonic development of silkworm *Bombyx mori*. Acta Biochim. Biophys. Sin. 37: 355–361.

Zhu, J., L. S. Indrasith, and O. Yamashita 1986. Characterization of vitellin, egg-specific protein and 30 kDa protein from *Bombyx* eggs, and their fates during oogenesis and embryogenesis. Biochim. Biophys. Acta (BBA) Gen. Subjects 882: 427–436.

Received 17 February 2014; accepted 23 April 2015.