

Full Length Research Paper

Expression analysis and characteristics of Bm-LOC778477 gene, a hypothetical protein from silkworm pupae, *Bombyx mori*

Wei Yu, Aiwu Tan and Yaozhou Zhang*

Institute of Biochemistry, College of life science, Zhejiang Sci-Tec University, Hangzhou 310018, China.

Accepted 2 June, 2009

This is the first identification of a gene encoding Bm-LOC778477 protein (a “hypothetical protein”) from silkworm pupal cDNA library. The full-length Bm-LOC778477 cDNA contained a 546-bp ORF encoding 187 amino acids. Using bioinformatics, the Bm-LOC778477 gene was predicted on the W chromosome and 4 exons and 3 introns were determined when compared to the silkworm genome. The consensus splice sites of exon/intron junctions were consistent with Chambon’s rule. In addition, TATA box and initiator (Inr) were also predicated in the promoter region of the gene. To elucidate whether Bm-LOC778477 gene encoded a functional protein in silkworm, pGEX-4T-1-Bm-LOC778477 expression plasmid was constructed to express GST-LOC778477 fusion protein in *Escherichia coli* Rosetta. rBm-LOC778477 was used to generate anti- Bm-LOC778477 polyclonal antibody, which were used to determine the subcellular localization of Bm-LOC778477. Immunostaining indicated that Bm-LOC778477 protein could be found in both the cytoplasm and nucleus. Western blot analysis indicated that Bm-LOC778477 was specific expression during the pupa stage and there was seldom expression during the larvae stage. Therefore, we propose that Bm-LOC778477 may play an important role during pupa stage development. These results may lay an important foundation of further function studies of Bm-LOC778477 protein.

Key words: Bm-LOC778477, hypothetical protein, subcellular localization.

INTRODUCTION

Until June 2007, 1218 genomes including bacteria, archaea, viral and eukarya have been sequenced completely. Gene annotation indicates that ‘conserved hypothetical proteins’ often represent more than half of the potential protein-coding regions of a genome (Galperin and Koonin 2004; Roberts, 2004; Sivashankari, 2006). A 3× coverage draft sequence (Mita et al., 2004) and a draft sequence with 5.9× coverage of the silkworm genome (Xia et al., 2004) were published in 2004. The size of the genome is 428.7 Mb and the gene numbers are 21302. As a bioreactor for protein drugs, pupa is also a specific stage of holometabolic insects. In order to elucidate how many genes are expressed in this stage and what function of those genes are involved in silkworm metamorphosis,

1659 cDNAs were screened from the cDNA library of silkworm pupa constructed by our library (Zhang et al., 2007), of which 837 them were entered into Unigene database (NCBI). Due to the fact that many bioinformatics methods to predicate the function of genes in any given genome have difficulties, which shows more than 70% prediction accuracy (Bork, 2000), Roberts (2004) proposed to establish a database for those conserved hypothetical and hypothetical genes, and then to offer those proteins to the biochemical community as potential targets for their function research.

Bm-LOC778477 gene is a novel gene which firstly identified from silkworm pupal cDNA library and is referred to ‘hypothetical gene’ in the NCBI database. Based on the facts that no protein(s) encoded by this ‘gene’ has been reported yet and the deduced amino acids was referred to ‘hypothetical protein LOC778477’ in the NCBI RefSeq database, thus the gene was named Bm-LOC778477. However, three cDNA clones for this gene are obtained

*Corresponding author. E-mail: yaozhou@zstu.edu.cn. Tel: +86-571-86843199. Fax: +86-571-86843198.

from three different silkworm cDNA libraries, respectively, such as blood (GenBank Accession: CK527993), silk-gland (GenBank Accession: CN373876) and pupal (GenBank Accession: DN236950), suggesting that it is most likely a protein-encoding gene. It will be very helpful to determine whether the ORFs of the Bm-LOC778477 encode a functional protein and what the normal function of the encoded protein is in the silkworm. However, no proteins known by far have significant homology with it brings difficulty to find and determine the possible native Bm-LOC778477 protein as well as the normal function of this protein that may be identified in the nature.

In this study, we identified a silkworm LOC778477 complementary DNA (cDNA; Bm-LOC778477) from a pupal cDNA library (GenBank Accession: DQ813504). Using bioinformatics analysis, characterization of Bm-LOC778477 gene was predicted. Immunocytochemistry and Western blot analysis were also used to determine the subcellular localization, expression levels and tissue distribution of Bm-LOC778477 in various developmental stages of silkworm. Our results provide an important foundation for further studies on the functional roles played by Bm-LOC778477 in the silkworm.

MATERIALS AND METHODS

Animals and materials

The *Bombyx mori* strain used in this study is the progeny of Jingsong×Haoyue. Silkworms were reared on mulberry leaves under standard conditions. Pupae, fifth instar larvae were dissected, frozen immediately in liquid nitrogen, and stored at -80°C . 4'-6-Diamidino-2-phenylindole (DAPI) was purchased from Sigma. Cy3-labeled goat anti-rabbit IgG was purchased from Proteintech Group Inc. Bm5 cells (a gift from Prof. ZF Zhang) were cultured in TC-100 medium (Sigma) supplemented with 10% fetal calf serum (FCS, Gibco BRL) at 27°C .

Bioinformatics analysis

The Bm-LOC778477 gene was identified in the process of analyzing the cDNA sequences from the *B. mori* pupa cDNA library constructed by our laboratory. The analysis of nucleotide and protein sequences was performed using the BLAST algorithm from NCBI (<http://www.ncbi.nlm.nih.gov/>). The characteristics of this gene were analyzed using DNASTar software.

Expression, purification of Bm-LOC778477 and antibody preparation

The Bm-LOC778477 cDNA (Accession No. DQ813504) was subcloned into prokaryotic expression vector pGEX-4T-1 (Amersham Biosciences, USA) which encodes a N-terminal GST-tag. The recombinant plasmid was named as pGEX-4T-1-Bm-LOC778477. *Escherichia coli* Rosetta (DE3) (Novagen, Germany) were transformed with the recombinant plasmid, the transformed *E. coli* were grown in LB medium, and recombinant protein expression was induced with IPTG. One liter LB media supplemented with ampicillin (100 $\mu\text{g}/\text{ml}$) and chloramphenicol (34 $\mu\text{g}/\text{ml}$) was used to inoculate 10 ml of fresh LB medium under the same condition until OD_{600} reached 0.6. Then IPTG was added to a final concentration of 0.5 mM

to induce the expression of GST-Bm-LOC778477. Then the culture was grown at 37°C for additional 4 h. The cell pellet was harvested by centrifuging at 6,000 g for 10 min at 4°C and resuspended in 50 ml lysis buffer (5 mM EDTA, 20 mM Tris, and 100 mM NaCl, pH 9.0). The bacteria were lysed by Scientz Ultrasonic homogenizer (Ningbo Scientz Biotechnology Co.Ltd., China) for 30 min on ice. The lysates were centrifuged at 10,000 g for 20 min at 4°C to remove supernatant. The precipitation was inclusion bodies and was frozen at -70°C .

In order to obtain the bioactive conformation of GST-Bm-LOC778477 fusion protein from the inclusion body, purification of the inclusion body was firstly solubilized with 8 M Urea and followed by refolding by dialysis. The inclusion bodies were washed by wash buffer I (5 mM EDTA, 20 mM Tris and 1% Trion X-100 (v/v), pH 9.0) and wash buffer II (5 mM EDTA, 20 mM Tris, and 4 mol/l Urea, pH 9.0), respectively, and then denatured in 30 ml of denaturation buffer (20 mM Tris, 5 mM EDTA, 8 mol/l Urea, 0.3 mM DTT, and 50 μM ZnCl_2 , pH 9.0) and allowed to stir for 5 h at 25°C . The suspension was centrifuged at 10,000 g for 20 min at 4°C . Then, protein concentration was determined by the Bradford method and BSA (sigma, USA) was used as the standard (Bradford, 1976). The solubilized inclusion bodies were diluted by denaturation buffer to a final concentration of 100 $\mu\text{g}/\text{ml}$. Refolding proteins were dialyzed extensively by refolding buffer (20 mM Tris, 5 mM EDTA, 0.1 mM GSH, 1 mM GSSG, 50 μM ZnCl_2 , the Urea concentration of refolding buffer I-II: 6, 4, 1 and 0 mol/L, respectively, pH 9.0) at 25°C (Ryu et al., 2006; Guise and West 1996; Tsumoto et al., 2003; Singh and Panda 2005). After that, the refolded proteins were centrifuged for 20 min at 10,000, and then the suspension was dialyzed against PBS (pH 7.4).

The refolded proteins were filtered through a 0.45 μm filter (Millipore, USA) and then purified by GST-TrapFF column (Amersham, USA). The eluted GST-Bm-LOC778477 protein was dialyzed against thrombin cleavage buffer (PBS, pH 8.0). Then, the GST-Bm-LOC778477 protein solution was subsequently ultrafiltrated (Microcon YM-10, Millipore Corp) to concentrate and incubated with thrombin (Calbiochem[®], USA) for 16 h at 22°C . After digestion, the protein solution was loaded again on the GST-TrapFF column, pre-equilibrated with cleavage buffer, to purify Bm-LOC778477 by removing GST tag and undigested fusion protein. The molecular weight of Bm-LOC778477 protein was weighted by Q Trap LC/MS/MS system.

Polyclonal antibody preparation

Three milligram of rBm-LOC778477 protein was injected into a New Zealand rabbit. After 12 weeks, serum was collected at 4°C , 10,000 g for 20 min and the antiserum was passed through a Protein A HP chromatography column to obtain purified rabbit IgG. An ELISA was used to determine the polyclonal antibody titer, and the specificity of the polyclonal antibody was determined by Western blot analysis.

Expression analysis of Bm-LOC778477 by Western blot

For Western blot analysis, silkworm stages or tissues were frozen in liquid nitrogen, ground to a fine powder, and treated with protein extraction lysis buffer (100 mmol/l Tris-Cl, 100 mg/l PMSF, 5 $\mu\text{g}/\text{ml}$ Aprotin, 0.15 mol/l NaCl, 5 mol/l EDTA, 0.5 % NP-40(v/v), 1 mol/l DTT, 12 mmol/l Deoxycholate sodium, pH 7.4). After 30 min on ice, homogenates were centrifuged at 12,000 g for 15 min at 4°C . Protein concentrations were determined using the Bradford method. Equal amounts of proteins were separated by SDS-PAGE with 10% polyacrylamide gel and electrotransferred onto PVDF membranes (Millipore). Membranes were blocked overnight at 4°C in 3% BSA/PBS and simultaneously incubated with anti Bm-LOC778477 polyclonal antibody (1:750 dilutions) for 2 h at room temperature

```

1  ACATTTTTATTATAATCAGTGGTTAGCTATGTCAGATTCATCTGAAGATGAAGACCTAT
1  M S D S S E D E D L

61  CGCGTTTCAAAGACGTCGTCGACAACCTCTTTCGTTAAATCTCTATCCCAAACCTCGTGTTA
11  S R F K D V V D N S F V K S L S Q T R V

121 CAGGATCTAAAAACAATTCAAGGCATTGAAGAAAAGCCCGTCTCCCAACGTTACTTAGAAA
31  T G S K T I Q G I E E K P V S Q R Y L E

181 TATCAAGCCACTACAATGATGTGAAGGTTCCAGAAGAAATGCAAAAAAGAATTGGTGCTA
51  I S S H Y N D V K V P E E M Q K R I G A

241 AATTATCAGCTGTCATACAGAAAAATACAAGATTTGTAGATGTGGACCCTGTTCAACCAA
71  K L S A V I Q K N T R F V D V D P V Q P
    X

301 AAAAAACGCAAAAATAAAGGGAGGTGAAAAATTGTTCAAGAATTCAAATGGATTTTTATCTT
91  K K R K I K G G V K L F K N S N G F L S

361 GTGATGACATTAAGGACATATCCACTGAGAAACATAATTGTGAAGCGAAACTGTTAAAGT
111 C D D I K D I S T E K H N C E A K L L K

421 ATAGTAAAGTTGGTCCTGACTCAGATGACATAGAGAAACTACATGCAGTTACAGTCAGTG
131 Y S K V G P D S D D I E K L H A V T V S
    X

481 GAGAATATGTAATCAAAAAGAAGAAATCAAATGTTGAAATCTAGAAGAAAAAGAGAAAT
151 G E Y V L S K E E I K C W K S R R K E K

541 TATATAAATACAAGGTGTCACAAAATAGTGATGTTTAGTAGCTTTAGAATAAACTTGT
171 L Y K Y K V S Q N S D V L V A L D *

601 TTCACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```

Figure 1. Complete nucleotide and deduced amino acid sequence of Bm-LOC778477 gene. The nucleotides and amino acids are numbered along the left margin and the exon/intron consensus splice sites are consistent with Chambon rules of GU-AG; The asterisk (*) indicates the stop codon TAA; the classical polyadenylation signal is marked by box; the underline stands for the initiator (Inr); (X) stands for splice sites of 4 exons of Bm-LOC778477 gene.

with gentle shaking. Then horseradish peroxidase-conjugated anti-rabbit total IgG from goat (Sigma-Aldrich, USA) secondary antibodies were added and allowed to incubate for 2 h at room temperature. Immunoblotting bands were detected with DAB.

Immunocytochemistry of Bm-LOC778477 protein

Bm5 cells were cultured overnight on glass cover slips, washed for 10 min three times in PBS, and fixed (PBS, 4% poly-formaldehyde, 0.1% Triton X-100, pH 7.4) at room temperature for 10 min. The fixed cells were blocked in 3% BSA/PBST at room temperature for 2 h followed by three 10-min washes in PBST (0.05% Tween-20 in PBS). Cells were then incubated with antiserum containing anti-Bm-LOC778477 polyclonal antibody (diluted 1:100 in blocking buffer) at 4°C overnight. The negative serum was obtained from the rabbit immunizing with the antigen. After three times 10-min washes in PBST, cells were incubated with Cy3-labeled goat anti-rabbit IgG (Proteintech Group Inc., USA) (diluted 1:1,000) at 37°C for 2 h and were then washed twice for 10 min in PBST. Cells were then incubated with 4'-6-diamidino-2-phenylindole (1 g/ml in PBS) at room temperature for 10 min. After washing once with PBST, cells were examined under a Nikon Confocal laser scanning microscopy (Eclipse TE2000-E, Japan); images were analyzed using EZ-C1 software.

RESULTS

Characteristics of Bm-LOC778477

An open reading frame (ORF) of 546 bp encodes a polypeptide of 187 amino acids with a predicted molecular weight of 21.26 kDa and a theoretical isoelectric point of 8.94 (Figure 1). The complete cDNA of Bm-LOC778477 has a 27bp 5' untranslated region (UTR), a canonical polyadenylation signal sequence (AATAAA) which overlaps the terminal codon TAA and a 21poly (A) tail downstream of the TAA (Zarudnaya et al., 2003). The ACATTT sequence in the 5'-UTR of Bm-LOC778477 gene was considered as initiator (Inr), which was identical to the Inr consensus sequences (T-C-A_n-1-G/T-T-C/T) of *Drosophila melanogaster* (Kadonaga, 2002).

Bm-LOC778477 gene was in the Scaffold001541 of silk worm genome, including 4 exons and 3 introns, where the exon/intron splice sites consensus was consistent with Chambon rules (Mount, 1982). Besides, the Bmb007964 gene encoding reverse transcriptase located in the W

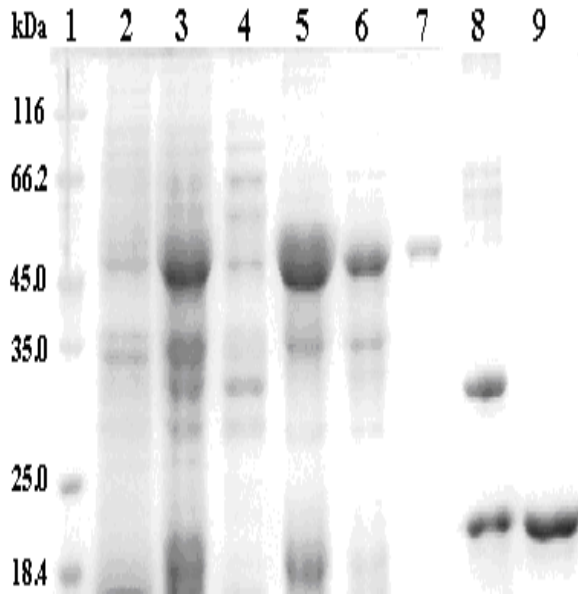


Figure 2. Expression and purification of fusion protein GST-Bm-LOC778477. Lane 1, low molecular weight protein marker; lane 2, protein from uninduced *E. coli* Rosetta (DE3)-pGEX-4T-1-Bm-LOC778477; lane 3, protein from Rosetta (DE3)-pGEX-4T-1-Bm-LOC778477 after inducing with IPTG; lane 4, soluble protein fraction from the sonic extract of induced *E. coli* Rosetta (DE3)-pGEX-4T-1-Bm-LOC778477; lane 5, insoluble protein fraction from the sonic extract of induced *E. coli* Rosetta (DE3)-pGEX-4T-1-Bm-LOC778477; lane 6, protein of inclusion body solubilized in 8 M Urea; Lane 7, purified fusion protein GST-Bm-LOC778477; lane 8, cleavage of GST fusion tag with thrombin; lane 9, purified fusion protein Bm-LOC778477.

chromosome was also on the Scaffold001541 (Abe et al., 2005). Based on these facts Bm-LOC778477 gene was predicted on the W chromosome.

Prokaryotic expression of Bm-LOC778477 and preparation of polyclonal antibody

The Bm-LOC778477 was subcloned into pGEX-4T-1 expression vector and then induced by IPTG in *E. coli* Rosetta (DE3). The molecular weight of the expressed fusion protein GST-Bm-LOC778477, as determined by electrophoresis, was 48 kDa, which was consistent with the calculated value (26 kDa of GST plus 21.26 kDa of Bm-LOC778477). The fusion protein existed mainly in the form of inclusion body (Figure 2). The Bm-LOC778477 protein after separation from the cleaved GST peptide is shown in Figure 2. The molecular weight of Bm-LOC778477 protein was approximately 21.363 kDa, approaching to the theoretical MW of 21.407 kDa, due to two additional N-terminal amino acids (G and S) (Figure 3).

Using digested and purified fusion protein rBm-LOC778477, anti-Bm-LOC778477 polyclonal antibody was gene-

rated in an immunized New Zealand rabbit and was purified by Protein A HP column chromatography (Amersham). The anti-Bm-LOC778477 polyclonal antibody titer, as determined by ELISA, was 1:32,000, and western blot analysis indicated that the antibody reached specifically with Bm-LOC778477 (Figure 2).

Expression analysis of Bm-LOC778477

In order to understand whether Bm-LOC778477 gene encode a hypothetical protein, ELISA and Western blot analyses were used to determine Bm-LOC778477 protein levels in total protein extracts from hemocyte, silk gland and pupae. The result of indirect ELISA indicated that Bm-LOC778477 protein maybe express in the hemocyte, silk gland and pupae. The Western blot results showed that Bm-LOC778477 expression was difficult to detect in the hemocyte and silk gland but increased significantly in pupa (Figure 4).

Subcellular localization of Bm-LOC778477

Subcellular localization of Bm-LOC778477 will offer a clue for biological function analysis of Bm-LOC778477 protein. Using software analysis tools (PSORT[□], <http://psort.nibb.ac.jp/>), Bm-LOC778477 protein was located in the nucleus (P: 89%). The treated cells were examined under a Nikon ECLIPSE TE2000-E Confocal Microscope, and images were analyzed using EZ-C1 software. Cy3-labeled goat anti-rabbit IgG emit red fluorescence when stimulated with light having a wavelength of 550 nm, and DAPI-stained nuclei emit red fluorescence when stimulated with light having a wavelength of 353 nm. Our results indicate that Bm-LOC778477 is located in both the cytoplasm and nucleus of Bm5 cells (Figure 5).

DISCUSSION

Bm-LOC778477 gene was a novel gene screened out from the silkworm pupal cDNA library. There was only experimental evidence for mRNA level but no protein(s) level up to now. Therefore, Bm-LOC778477 protein remained to refer as a "hypothetical protein". Protein similarity alignment with the protein from others species published in databases including NCBI, EBI, DDBJ, PDB etc. indicated there were only 5 proteins from other species including *Aedes aegypti* (Accession No.EAT36792 and EAT36794), *Tribolium castaneum* (Accession No.XP_967083) and *D. melanogaster* (Accession No.AAN09597) which have similarity with Bm-LOC77847777 (Accession No.ABG67690). So we speculated that Bm-LOC77847777 gene was insect-specific gene.

Silkworm was holometabolism insect, and the transformation of pupae into adult need 8-13 days. Two other cDNA clones of Bm-LOC778477 gene (Accession No. CK527993 and CN373876, respectively) were also ob-

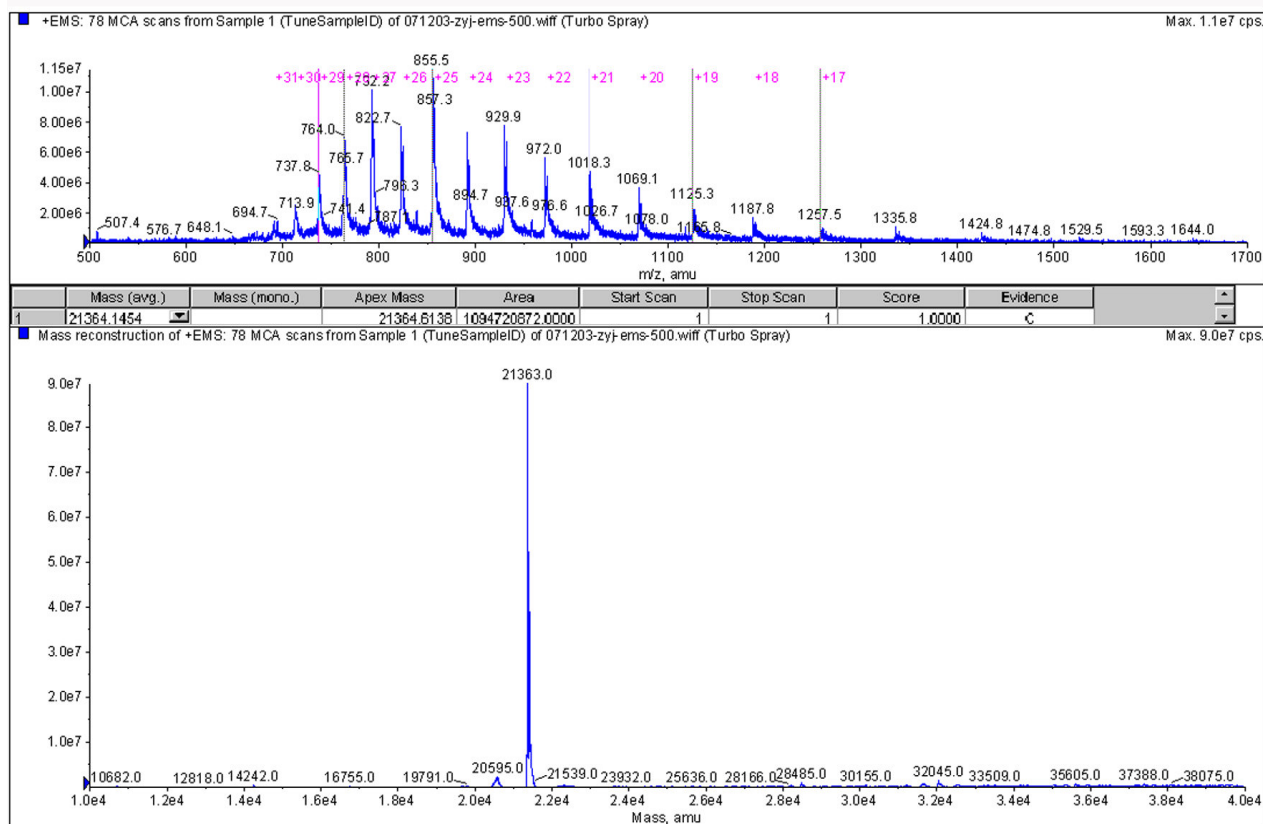


Figure 3. Mass spectrometry analysis of the Bm-LOC778477 protein. The MW of Bm-LOC778477 protein was 21.363 kDa.

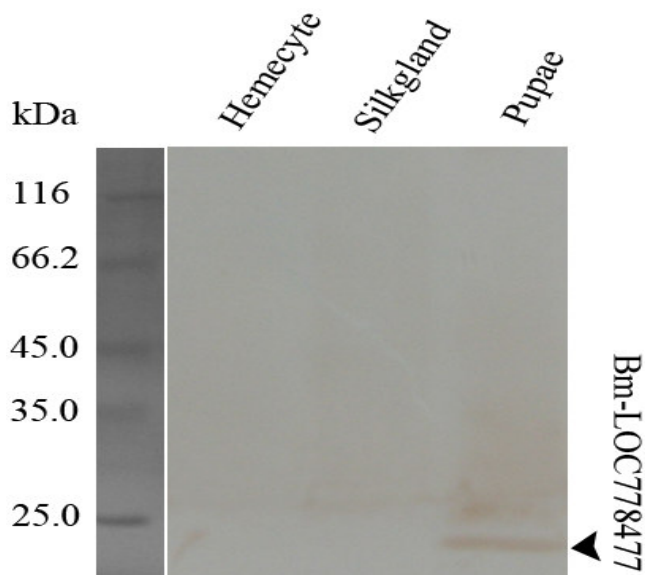


Figure 4. Western blot analysis of the expression of Bm-LOC778477 in different silkworm tissues. The protein samples used to detect in this assay were hemocyte, silk gland and pupae of 144 h after pupation. Equal amounts of protein were separated by 10% SDS-PAGE. There was a specific band on the PVDF membrane in the sample of pupae. The dilution of primary antibody and second antibody is 1:750 and 1:1,000, respectively.

tained from blood and silk gland cDNA library, however, Bm-LOC778477 protein was only detected in the pupae by immunoblot. Thus, we inferred that Bm-LOC778477 may play an important role in metaphase of metamorphosis. Using SignalP 3.0 Server, there was no signal peptide sequence in the Bm-LOC778477 protein; however, Bm-LOC778477 protein distributed in the cytoplasm and nucleus of Bm5 cells by immunocytochemistry, thus, the mechanism of how the Bm-LOC778477 protein entered into nucleus was unclear. Therefore, the task to understand the mechanism of the Bm-LOC778477 protein of how to enter into nucleus will be helpful to further study or character the function of Bm-LOC778477 protein.

ACKNOWLEDGMENTS

This work was supported by National Basic Research Program of China National Basic Research Program of China (863 Program) (No. 2007AA100504, 2007AA021-703), (973 Program) (No. 2005CB121006), National Key Technologies R&D Program (2006BAI01B04), National Natural Science Foundation of China (No.30740015, No.30600323, No.30670447) and Qianjiang Talent Plan of Zhejiang Province No.QJD0702014. The authors thank ZF Zhang researcher for his kindly gifting of Bm5 cell strains.

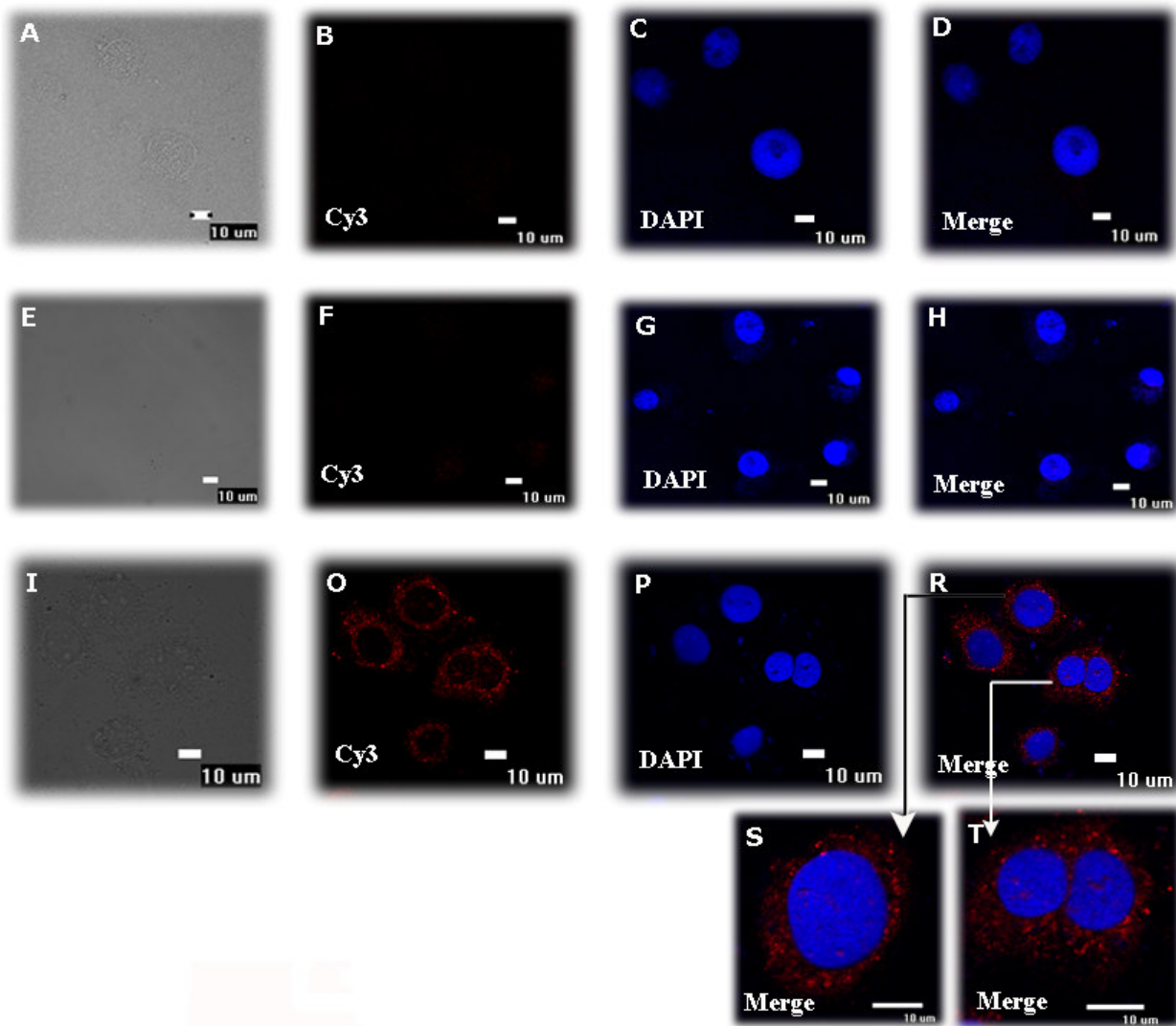


Figure 5. Subcellular localization of Bm-LOC778477 with Cy3-labeled goat anti-rabbit IgG and DAPI. A-D and E-H: negative control group, using negative rabbit serum; I-R: Experimental group, using anti-Bm-LOC778477 polyclonal antibody; A, E and I: Bm5 cells at the transmission light; B, F and O: Bm-LOC778477 subcellular localization as indicated by Cy3-labeled second antibody; C, G and P: DAPI staining; D, H, R, S and T: merged image; D: merged image of B and C; H: merged image of F and G; R, S and T: merged image of O and P; the magnification of A to R are 400 \times , S and T are 1,600 \times ; The scale bar for both panels is 10 μ m.

REFERENCES

- Abe H, Mita K, Yasukochi Y, Oshiki T, Shimada T (2005). Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. *Cytogenet Genome Res.* 110: 144-151.
- Bork P (2000). Powers and pitfalls in sequence analysis: the 70% hurdle. *Genome Res.* 10: 398-400.
- Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Galperin MY, Koonin EV (2004). Conserved hypothetical proteins: prioritization of targets for experimental study. *Nucleic acids res.* 32: 5452-5463.
- Guisse AD, West SM (1996). Chaudhuri JB. Protein folding *in vivo* and renaturation of recombinant proteins from inclusion bodies. *Mol. Biotechnol.* 6: 53-64.
- Kadonaga JT (2002). The DPE, a core promoter element for transcription by RNA polymerase II. *Exp. Mol. Med.* 34: 259-264.
- Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikummar G, Shimomura M, Nagamura Y, Shin-I T, Abe H, Shimada T, Morishita S, Sasaki T (2004). The genome sequence of silkworm, *Bombyx mori*. *DNA Res.* 11: 27-35.
- Mount SM (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* 10: 459-472.
- Roberts RJ (2004). Identifying protein function—a call for community action. *PLoS Biol.* 2: E42.
- Ryu EK, Cho KJ, Kim JK, Harmer NJ, Blundell TL, Kim KH (2006). Expression and purification of recombinant human fibroblast growth factor receptor in *Escherichia coli*. *Protein Exp. Purif.* 49: 15-22.
- Singh SM, Panda AK (2005). Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 99: 303-310.
- Sivashankari S (2006). Shanmughavel P. Functional annotation of hypothetical proteins—A review. *Bioinformatics.* 1: 335-338.

- Tsumoto K, Ejima D, Kumagai I, Arakawa T (2003). Practical considerations in refolding proteins from inclusion bodies. *Protein Exp. Purif.* 28: 1-8.
- Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, Zhao P, Zha X, Cheng T, Chai C, Pan G, Xu J, Liu C, Lin Y, Qian J, Hou Y, Wu Z, Li G, Pan M, Li C, Shen Y, Lan X, Yuan L, Li T, Xu H, Yang G, Wan Y, Zhu Y, Yu M, Shen W, Wu, D, Xiang Z, Yu J, Wang J, Li R, Shi J, Li H, Li G, Su J, Wang X, Li G, Zhang Z, Wu Q, Li J, Zhang Q, Wei N, Xu J, Sun H, Dong L, Liu D, Zhao S, Zhao X, Meng Q, Lan F, Huang X, Li Y, Fang L, Li C, Li D, Sun Y, Zhang Z, Yang Z, Huang Y, Xi Y, Qi Q, He D, Huang H, Zhang X, Wang Z, Li W, Cao Y, Yu Y, Yu H, Li J, Ye J, Chen H, Zhou Y, Liu B, Wang J, Ye J, Ji H, Li S, Ni P, Zhang J, Zhang Y, Zheng H, Mao B, Wang W, Ye C, Li S, Wang J, Wong GK, Yang H (2004). Biology Analysis Group. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science*, 306: 1937-1940.
- Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM (2003). Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic Acids Res.* 31: 1375-1386.
- Zhang YZ, Chen J, Nie ZM, Lü ZB, Wang D, Jiang CY, He PA, Liu LL, Lou YL, Song L, Wu XF (2007). Expression of open reading frames in silkworm pupal cDNA library. *Appl. Biochem. Biotechnol.* 136: 327-343.