Research Article

A syntenic coding region for vitelline membrane proteins in four lepidopteran insects

Yunmin Xu, Ziliang Zou, Xingfu Zha, Zhonghuai Xiang and Ningjia He

State Key Laboratory of Silkworm Genome Biology, Southwest University, Beibei, Chongqing 400715, China

Received on April 8, 2012; Accepted on July 22, 2012; Published on October 18, 2012

Correspondence should be addressed to Ningjia He; Phone: +86 23 6825 0797, Fax: +86 23 6825 1128, Email: hejia@swu.edu.cn

Abstract

The vitelline membrane is the inner layer of the eggshell, but the genomic information available for vitelline membrane proteins (VMPs) in Lepidoptera is limited. In the present study, we identified a syntenic coding region for VMPs in four lepidopteran genomes (*Bombyx mori*, *Manduca sexta*, *Danaus plexippus* and *Heliconius melpomene*) and four putative VMP coding genes located within it. RT-PCR results showed *Bombyx* VMP coding genes expressed prior to the early choriogenesis stage in follicles. Alignment analyses revealed that the vitelline membrane domain was shared between Lepidoptera and Diptera. However, the third cysteine residue conserved in dipteran VMPs was absent in those of Lepidoptera. In addition, another conserved region was identified in lepidopteran VMPs.

Introduction

The insect eggshell is composed of a specialized extracellular matrix which surrounds and shields the embryo from dehydration and infection. The eggshell has two major layers: the inner vitelline membrane layer (VM) and the external chorion layer (Margaritis *et al*. 1980, Mazur *et al*. 1980). Recently, several studies have demonstrated that the inner VM plays an important role in determining the embryonic pattern (Stevens *et al*. 2003, Zhang *et al*. 2009).

The vitelline membrane proteins (VMPs) are specifically synthesized and secreted prior to the choriogenesis stage in follicle cells (FCs). The VMPs and their coding genes were firstly identified in *Drosophila melanogaster* (Fargnoli & Waring 1982, Mindrinos *et al*. 1985). *D. melanogaster* VMPs are small proline-rich proteins with a highly conserved hydrophobic vitelline membrane domain (VM domain, 38 amino acids) in their C terminus (Scherer *et al*. 1988). Three VMP coding genes (*15a-1*, *15a-2* and *15a-3*) were subsequently identified in *Aedes aegypti*, based on the similarity to the VM domain in their proteins (Edwards *et al*. 1998, Lin *et al*. 1993). Among them was the VM26Ab (also called sV23) in *D. melanogaster* (Manogaran & Waring 2004, Savant & Waring 1989, Wu *et al*. 2010). Experimental evidence showed that sV23 is an abundant VMP, essential for female fertility. However, when comparing flies and mosquitoes, the genomic information of VMPs from non-dipteran insects is limited. In Lepidoptera, only two VMPs have been identified (Kendirgi *et al*. 2002, Xu *et al*. 2011) and it is still unclear whether these assemble into a functional network structure. In this study, we identified a syntenic coding region for VMPs in four lepidopteran genomes and four putative VMP coding genes located within it.

Materials and Methods

Insect preparation and tissue dissection

The silkworm strain *p50* (*DaZao*) used in this study was maintained at State Key Laboratory of Silkworm Genome Biology at Southwest University of China and the larvae were reared on fresh mulberry leaves at 25° C. Ovary tissues were dissected from the female pupae from Day 1 to Day 9 and the female moth on Day 1. For preparing follicles at different developmental stages, five proximal follicles, dissected from ovarioles at Day 6 pupa, were used as a sample for RT-PCR analysis according Kendirgi *et al* (2002) and Xu *et al* (2011).

RNA preparation and RT-PCR analysis

Total RNA was isolated from tissues using Trizol (Invitrogen, USA) and reverse transcribed to cDNA by

the AMV Reverse Transcriptase (Promega, USA). The primers used for RT-PCR detection are listed in Table 2S. The PCR parameters were set as following: one cycle of pre-denaturation at 94 °C for 3 min and 30 s, 25 cycles of denaturation at 94 °C for 45 s, annealing at 56 °C for 45 s, elongation at 72 °C for 45 s, and, finally, elongation at 72 °C for 10 min. The amplification of *BmActin A3* served as control.

Sequences downloaded and gene prediction

The genome sequences/predicted genes were downloaded from FlyBase (http://flybase.org/) and SilkDB (http://silkworm.swu.edu.cn/silkdb/). The putative VMPs in *B. mori* were employed as the query sequences to tBlastn against the other three lepidopteran Databases, Manduca Base (Agricultural Pest Genomics Resource Database: www.agripestbase.org), Monarch Base (http://monarchbase.umassmed.edu/) and Heliconius Database (http://etylus.bio.uci.edu/). The homologous coding regions for putative VMPs in *M. sexta*, *D. plexippus* and *H. melpomene* were identified and the putative VMP coding genes were predicted by GENSCAN (http://genes.mit.edu/GENSCAN.html). The putative lepidopteran VMPs were named according to their molecular weight.

Gene analysis and sequences alignment

The signal peptides of the VMPs were predicted by SMART (http://smart.embl-heidelberg.de/) and the PI/MW of the VMPs were predicted using ExPASy Compute pI/MW (http://web.expasy.org/compute_pi/). The hydrophilic/hydrophobic regions for VMPs were predicted by ExPASy ProtScale (http://web.expasy.org/protscale/). The alignments were performed using Clustalx, and the phylogenetic analysis was performed by MEGA 4.0.2 according to the VM domain.

Results and Discussion

The VMP coding region in *B. mori*

In a previous study, we identified a VMP (BmEP80, GenBank ID: HM012804) in *B. mori* (Xu *et al*. 2011). As noted earlier, three VMP coding genes (*VM26Ab*, *VM26Ac* and *VM26Aa*) clustered near to each other were identified in *D. melanogaster* (Alatortsev 2006) (Figure 1). It therefore seemed interesting to find out

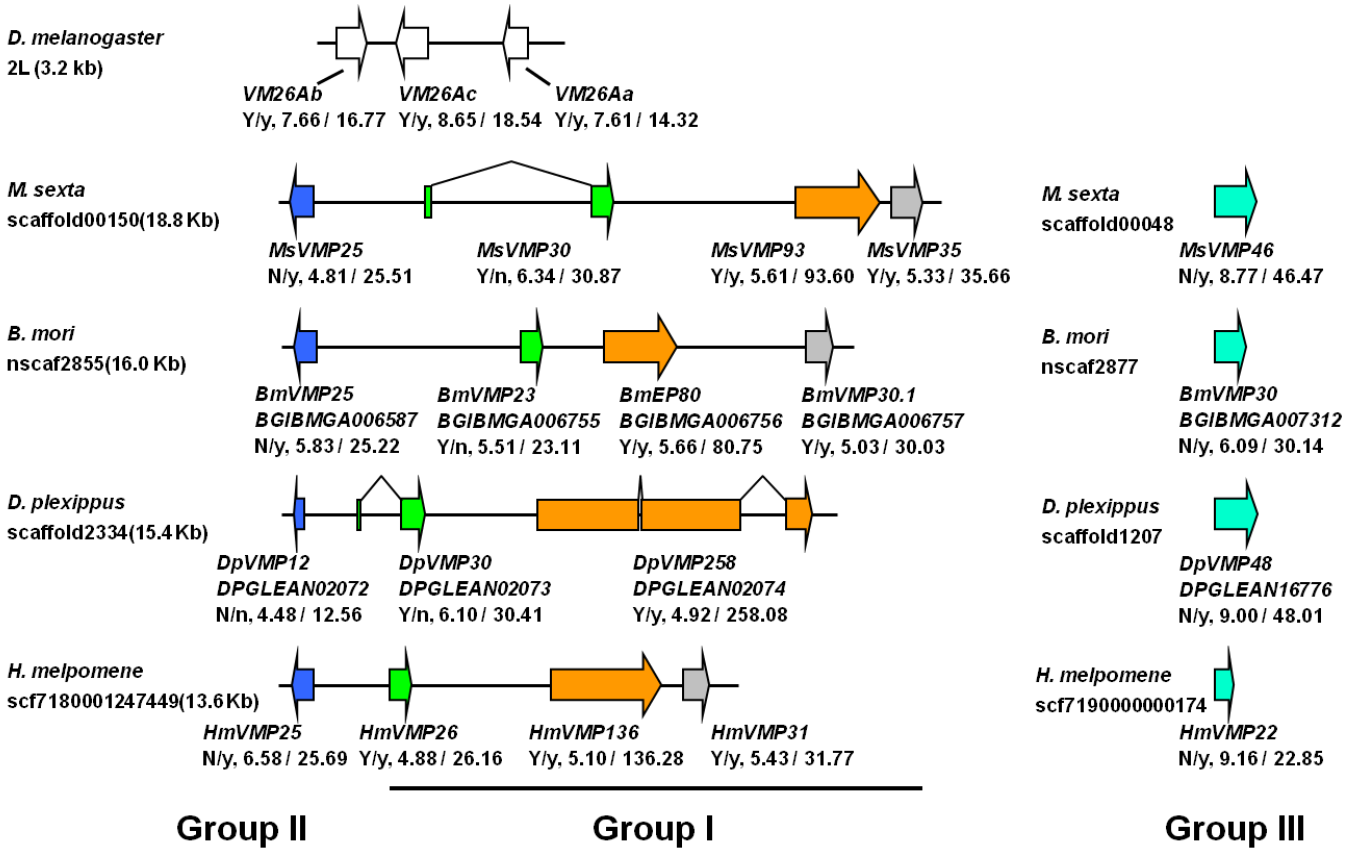

Figure 1. The VMPs coding regions in *D. melanogaster* and four Lepidopteran insects. The size and scaffold IDs for the VMPs coding regions are given. Each homologous gene among the four Lepidoptera is depicted in the same color and their NCBI GenBank accession numbers are listed in Table 1S. In "BmEP80, BGIBMGA006756, Y/y, 5.66/80.75'', BGIBMGA006756 is the ID number for BmEP80 in SilkDB, "Y/y" means that the VM domain/signal peptide is present in BmEP80, and "5.66/80.75" means that the PI/MW of BmEP80 is 5.66/80.75. The group numbers are indicated as I, II, III.

whether a similar VMP coding region is present in *B. mori*. The results showed that, like *D. melanogaster*, a VMP coding region located on the side of *BmEP80* was found in *B. mori*. In this region, three other putative VMP coding genes (*BmVMP25*, *BmVMP23* and *BmVMP30.1*) were identified (Figure 1). The putative silkworm VMP coding genes are intronless and encode small proteins no more than 30kDa in size.

The syntenic coding for VMPs in *M. sexta*, *D. plexippus* and *H. melpomene*

Recently, the genome sequencing of three lepidopteran insects (*M. sexta*, *D. plexippus* and *H. melpomene*) was completed (Dasmahapatra *et al*. 2012, Zhan *et al*. 2011) and the latter study reported that the *D. plexippus* genome shared prominent gene similarity and synteny with *B. mori*. In the present study, VMP coding regions homologous to *B. mori* were identified in *M. sexta*, *D. plexippus* and *H. melpomene*. In these regions (Figure 1), four, three and four putative VMP coding genes were identified in *M. sexta*, *D. plexippus* and *H. melpomene*, respectively. In addition, the order of the putative VMP coding genes was preserved among the four lepidopteran insects.

We classified the lepidopteran VMPs into three groups. Group I contained the VMPs with conserved VM domains (described below) including BmEP80, MsVMP93, DpVMP258, HmVMP136, BmVMP23, MsVMP30, DpVMP30, HmVMP26, BmVMP30.1, MsVMP35 and HmVMP31. Four VMPs, BmVMP25, MsVMP25, DpVMP12 and HmVMP25 fell into group II. *BmVMP30* was the first identified lepidopteran VMP coding gene (Kendirgi *et al*. 2002). We found that *MsVMP46* (*M. sexta*), *DpVMP48* (*D. plexippus*) and *HmVMP22* (*H. melpomene*) genes were homologous to *BmVMP30*. In addition, transmembrane segments were predicted within the sequences of BmVMP30, MsVMP46,

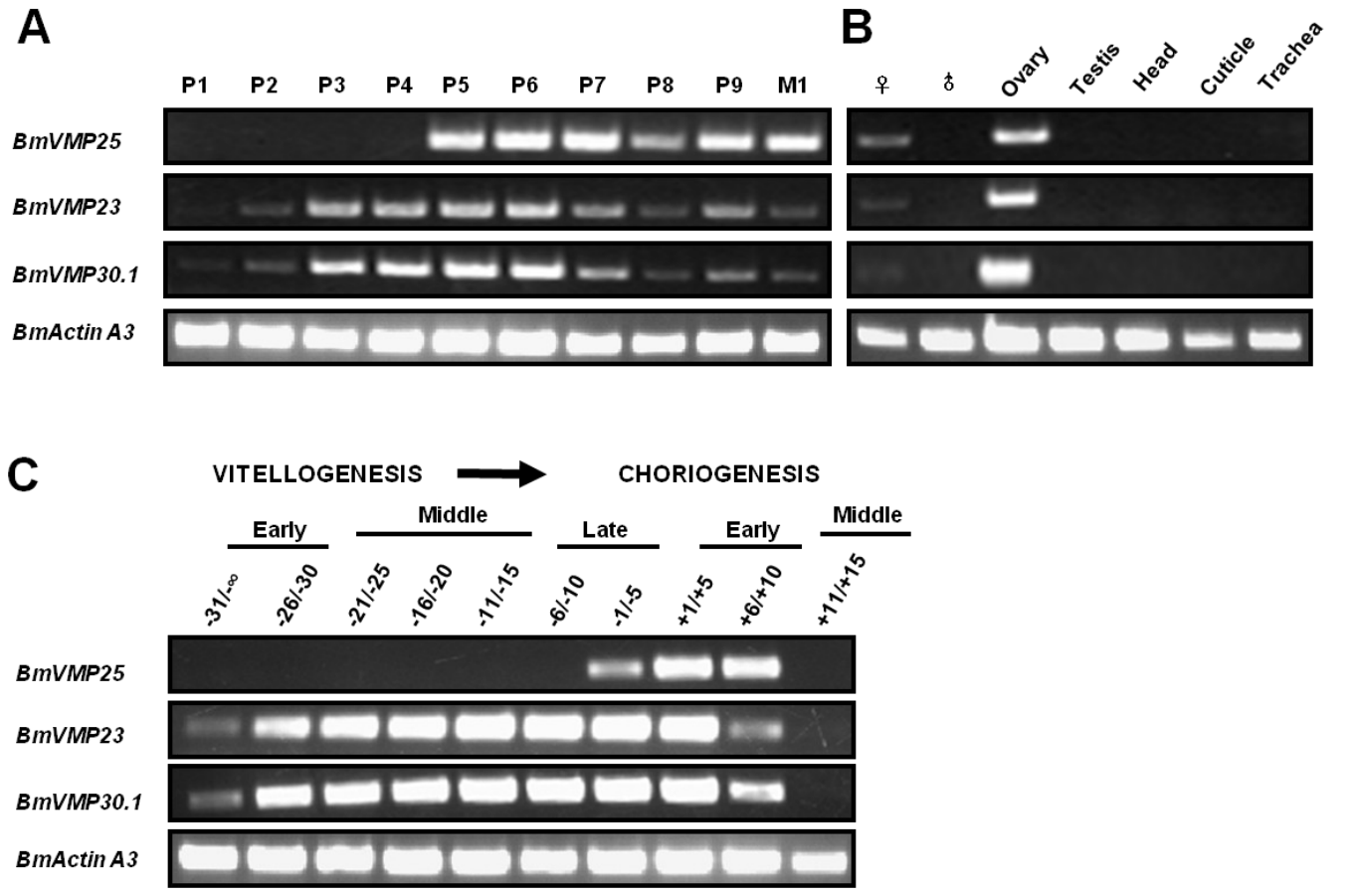

Figure 2. Expression patterns of *BmVMP25*, *BmVMP23* and *BmVMP30.1* in *B. mori*. The amplification of *BmActin A3* served as an internal control and the primers used for RT-PCR are listed in Table 2S. (A) Ovary tissues dissected from the pupal and moth stages were used for the RT-PCR analysis. P1-P9: pupal stage Day-1 to Day-9. M1: moth stage Day-1. (B) Tissue distribution of the *BmVMP25*, *BmVMP23* and *BmVMP30.1* transcripts at pupal stage Day-6. ♀: female, ♂: male. Ovary, head, epidermis, trachea and testis were dissected from pupa at Day-6. (C) The expression of *BmVMP25*, *BmVMP23* and *BmVMP30.1* in follicles at the different developmental stages. The follicles were dissected from ovarioles in Day-6 pupa. The earliest stage of choriogenesis for follicles was marked as +1 and the latest stage of vitellogenesis for follicles as -1.

DpVMP48 and HmVMP22 (predicted by SMART, data not shown). However, they have no sequence similarity to the VMPs in the group I and group II, so we classed the four sequences into group III.

Amino acid composition analysis showed that the dipteran VMPs and the lepidopteran VMPs in group II were rich in proline and alanine. Leucine and proline residues were abundant in the lepidopteran VMPs in group III. And the lepidopteran VMPs in group I bearing the VM domain were rich in hydrophilic serine residues.

The study has so far revealed that the linear arrangements of VMP coding genes are conserved in Diptera and Lepidoptera. Further detailed sequence analysis uncovered distinctive properties between VMP coding genes in lepidopteran species and *D. melanogaster*. The VMP coding genes in *D. melanogaster* were intronless and encoding small proteins with the VM domain. However, *MsVMP30*, *DpVMP30* and *DpVMP258* genes had introns. *BmEP80*, *MsVMP93*, *DpVMP258* and *HmVMP136* genes encoded proteins with high molecular weight, and the VM domain was absent in eight putative lepidopteran VMPs (MsVMP25, BmVMP25, DpVMP12, HmVMP25, MsVMP46, BmVMP30, DpVMP48 and HmVMP22). Lepidopteran VMPs (group I and group III) and dipteran VMPs had different amino acid composition. Lepidoptera is a fast evolving insect order (Zhan *et al*. 2011). The complexity of lepidopteran VMP sequences might reflect the relatively fast evolution of the lepidopteran VMP coding genes.

The expression patterns of putative VMP coding genes in *B. mori*

As shown in Figure 2A and B, three putative VMP coding genes were expressed in an ovary-biased manner from the middle to the late pupal stages. According to the microarray data in larval stage (5th larvae day 3), the *BmVMP23* transcript was detectable in testis, midgut and malpighian tubule tissues during the larval stage (Xia *et al*. 2007) (Figure 1S). The expressions of three putative VMP coding genes were also detected in follicles during different developmental stages. As shown in Figure 2C, the *BmVMP23* and *BmVMP30.1* showed a nearly identical expression pattern, expressed in follicles from -30 to +10 (from early vitellogenesis to early choriogenesis). *BmVMP25* was expressed in follicles from -5 to +10 (from late vitellogenesis to early choriogenesis) and had a similar expression pattern to *BmVMP30* (GenBank ID: AF294885) and *BmEP80* (Kendirgi *et al*. 2002, Xu *et al*. 2011).

The VM domain is shared between Lepidoptera and Diptera

Alignment of the VMPs in group I has enabled us to find a hydrophobic conserved region in their C terminus. This region was named "VM domain" according to the similarity to the VM domain in dipteran VMPs (Figure 3A). When comparing the VM domain between the VMPs in Lepidoptera and Diptera, we found that the second cysteine residue was conserved in all VMPs. All of the dipteran VMPs and nine of the lepi-

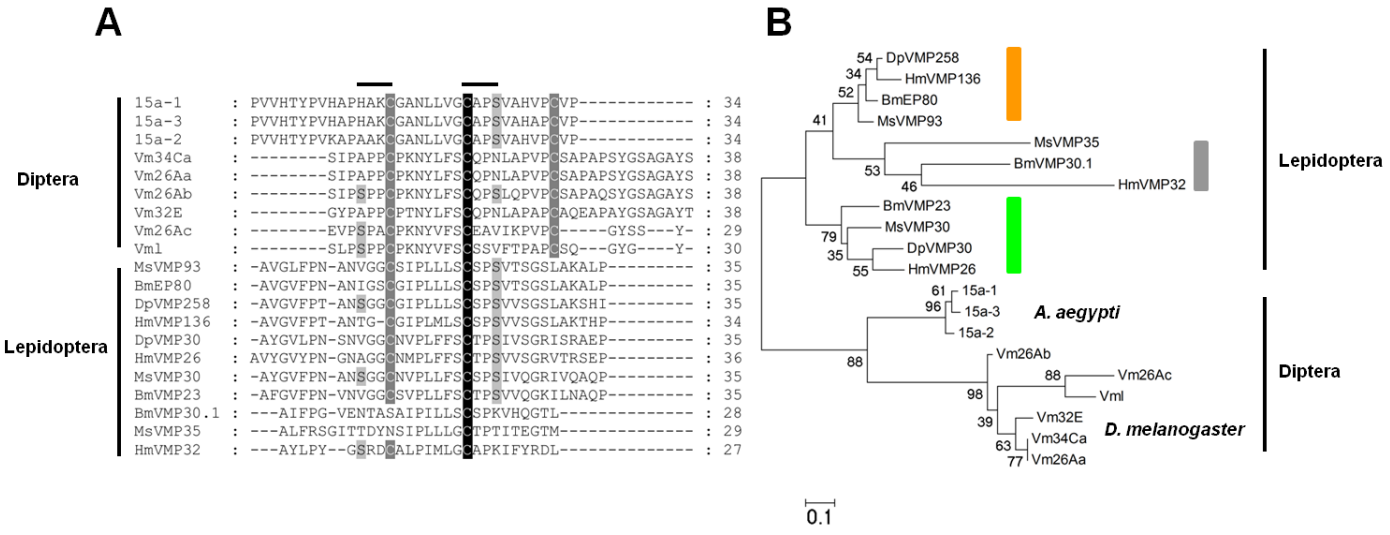

Figure 3. Alignment results of VM domain and phylogenetic analysis of VMPs. (A) The alignment of the VM domains for VMPs from *D. melanogaster* (Vm34Ca, Vm26Aa, Vm26Ab, Vm32E, Vm26Ac and Vml), *A. aegypti* (15a-1, 15a-2, 15a-3) and four Lepidoptera insects (*B. mori*, *M. sexta*, *D. plexippus* and *H. melpomene*). Three precisely spaced and conserved cysteines (CX_7CX_8C in Diptera) are highlighted. The bold line shows the non-classical redox active site CXXS and SXXC motif which remain in Diptera and Lepidoptera, respectively. (B) Phylogenetic analysis of dipteran and lepidoteran VMPs from group I. The color bars (same color as Figure 1) indicate the homologous genes among the four Lepidoptera.

dopteran VMPs had the first conserved cysteine residue. The VM domains for BmVMP30.1 and MsVMP35 were short and absent in the first cysteine residue. It is also interesting to note that the third cysteine residue which is conserved in dipteran VMPs was totally lost in their lepidopteran counterparts.

The VM domain in lepidopterans possessed the first/second conserved cysteine. Wu *et al*. (2010) reported that the substitution mutations at the first or third cysteines were tolerated in sV23 in *D. melanogaster*. Conversely, the females with a substitution at the second position of sV23 were sterile and laid collapsed eggs. In the present study, we found that the second cysteine was conserved and the third cysteine was absent in the lepidopteran VM domain, indicating that the second cysteine residue might be critical for the function of lepidopteran VMPs.

As shown in Figure 3A, within the VM domain, non-classical redox active site motifs (CXXS) for thioredoxin (Fomenko & Gladyshev 2002) in the second cysteine and SXXC for protein disulfide isomerase (PDI) (Lith et al. 2005) in the first cysteine were identified. Whether these motifs function in isomerizing disulfide bonds to form a stable VM network is needed to be confirmed by future studies.

A phylogenetic analysis of VMPs revealed that the lepidopteran VMPs were isolated from the dipteran VMPs (Figure 3B). Within the lepidopteran VMPs, the homologous VMPs (depicted with the same color in Figure 1) were clustered together.

Other conserved regions in lepidopteran VMPs

In contrast to the sequence diversity of the central region, BmEP80, MsVMP93, DpVMP258 and HmVMP136 have another conserved hydrophilic region in their N terminus (region 1, Figure 2S). Manogaran & Waring (2004) reported that the hydrophobic N terminal region of sV23 was essential for the assembly of a functional VM network in *D. melanogaster*. Selective removal of the N terminus of sV23 produced flaccid and infertile eggs with a soluble VM network. A conserved hydrophilic region in the N terminus of BmEP80, MsVMP93, DpVMP258 and HmVMP136 indicated that this region might be essential for the function of VMPs in Lepidoptera. But, the region 1 was hydrophilic, so it might have a different function with the hydrophobic N terminal region in sV23.

Conclusion

Seventeen putative VMP coding genes were identified in four lepidopteran insects. The present study revealed that the linear arrangements of VMP coding genes were conserved in Diptera and Lepidoptera. *Bombyx* putative VMP coding genes were expressed prior to the early choriogenesis stage in follicles. The presence of the vitelline membrane domain in non-dipteran insects was identified for the first time. The first/second cysteine residues and the non-classical redox active site motifs in VM domain were evolutionarily conserved between Diptera and Lepidoptera. The demonstration of synteny among lepidopteran insects will facilitate the functional investigations of VMPs.

Conflicts of interest

The authors have no conflicts of interest.

Acknowledgements

This work was supported by research grants from the National Basic Research Program of China (Grant No.2012CB114600), Science Fund for Distinguished Young Scholars of Chongqing (Grant No.cstc2011jjjq0010), and the Fund for Doctor in Southwest University (Grant No. kb2010015).

References

Alatortsev VE 2006 New genes for vitelline membrane proteins in Drosophila. *Mol Biol* **40** 330-332.

Dasmahapatra KK, James R. Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Hughes DST, Ferguson LC, Martin SH, Salazar C, Lewis JJ, Adler S, Ahn SJ, Baker DA, Baxter SW, Chamberlain NL, Chauhan R, Counterman BA, Dalmay T, Gilbert LE, Gordon K, Heckel DG, Hines HM, Hoff KJ, Holland PWH, Jacquin-Joly E, Jiggins FM, Jones RT, Kapan DD, Kersey P, Lamas G, Lawson D, Mapleson D, Maroja LS, Martin A, Moxon S, Palmer WJ, Papa R, Papanicolaou A, Pauchet Y, Ray DA, Rosser N, Salzberg SL, Supple MA, Surridge A, Tenger-Trolander A, Vogel H, Wilkinson PA, Wilson D, Yorke JA, Yuan F, Alexi L. Balmuth AL, Eland C, Gharbi K, Thomson M, Gibbs RA, Han Y, Jayaseelan JC, Kovar C, Mathew T, Muzny DM, Ongeri F, Pu L, Qu J, Thornton RL, Worley KC, Wu Y, Linares M, Blaxter ML, Ffrench-Constant RH, Joron M, Kronforst MR, Mullen SP, Reed RD, Scherer SE, Richards S, Mallet J, Owen McMillan W & Jiggins CD 2012 Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* doi:10.1038/nature11041.

Edwards MJ, Severson DW & Hagedorn HH 1998 Vitelline envelope genes of the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **28** 915-925.

Fargnoli J & Waring GL 1982 Identification of vitel-

line membrane proteins in *Drosophila melanogaster*. *Dev Biol* **92** 306-314.
Fomenko DE & Gladyshev VN 2002 CxxS: fold-independent redox motif revealed by genome-wide searches for thiol/disulfide oxidoreductase function. *Protein Sci* **11** 2285-2296.
Kendirgi F, Swevers L & Iatrou K 2002 An ovarian follicular epithelium protein of the silkworm (*Bombyx mori*) that associates with the vitelline membrane and contributes to the structural integrity of the follicle. *FEBS Lett* **524** 59-68.
Lin Y, Hamblin MT, Edwards MJ, Barillas-Mury C, Kanost MR, Knipple DC, Wolfner MF & Hagedorn HH 1993 Structure, expression, and hormonal control of genes from the mosquito, *Aedes aegypti*, which encode proteins similar to the vitelline membrane proteins of *Drosophila melanogaster*. *Dev Biol* **155** 558-568.
Lith MV, Hartigan N, Hatch J & Benham AM 2005 PDILT, a divergent testis-specific protein disulfide isomerase with a non-classical SXXC motif that engages in disulfide-dependent interactions in the endoplasmic reticulum. *J Biol Chem* **280** 1376-1383.
Manogaran A & Waring GL 2004 The N-terminal prodomain of sV23 is essential for assembly of a functional vitelline membrane network in Drosophila. *Dev Biol* **270** 261-271.
Margaritis LH, Kafatos FC & Petri WH 1980 The eggshell of *Drosophila melanogaster*. Fine structure of the layers and regions of the wild-type eggshell. *J Cell Sci* **43** 1-35.
Mazur GD, Regier JC & Kafatos FC 1980 The silkmoth chorion: morphogenesis of surface structures and its relation to synthesis of specific proteins. *Dev Biol* **76** 305-321.
Mindrinos MN, Scherer LJ, Garcini FJ, Kwan H, Jacobs KA & Petri WH 1985 Isolation and chromosomal location of putative vitelline membrane genes in *Drosophila melanogaster*. *The EMBO J* **4** 147-153.
Savant SS & Waring GL 1989 Molecular analysis and rescue of a vitelline membrane mutant in Drosophila. *Dev Biol* **135** 43-52.
Scherer LJ, Harris DH & Petri WH 1988 Drosophila vitelline membrane genes contain a 114 base pair region of highly conserved coding sequence. *Dev Biol* **130** 786-788.
Stevens LM, Beuchle D, Jurcsak J, Tong X & Stein D 2003 The Drosophila embryonic patterning determinant torsolike is a component of the eggshell. *Curr Biol* **13** 1058-1063.
Wu T, Manogaran AL, Beauchamp JM & Waring GL 2010 Drosophila vitelline membrane assembly: a critical role for an evolutionarily conserved cysteine in the "VM domain" of sV23. *Dev Biol* **34** 360-368.
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 Biol* **8** 1-13.
Xu Y, Fu Q, Li S & He N 2011 Silkworm egg proteins at the germ-band formation stage and a functional analysis of BmEP80 protein. *Insect Biochem Mol Biol* **41** 572-581.
Zhan S, Merlin C, Boore JL & Reppert SM 2011 The monarch butterfly genome yields insights into long-distance migration. *Cell* **147** 1171-1185.
Zhang Z, Stevens LM & Stein D 2009 Sulfation of eggshell components by pipe defines dorsal-ventral polarity in the Drosophila embryo. *Curr Biol* **19** 1200-1205.