Insect Science (2009) 16, 455–464, DOI 10.1111/j.1744-7917.2009.01275.x

Cloning, expression and localization of a trypsin-like serine protease in the spruce budworm, *Choristoneura fumiferana*

Wen-Ying He[1], Yi-Ping Zheng[2], Lin Tang[1], Si-Chun Zheng[1], Catherine Béliveau[3], Daniel Doucet[2], Michel Cusson[3] and Qi-Li Feng[1]

[1]*Guangdong Provincial Key Laboratory of Biotechnology for Plant Development, School of Life Sciences, South China Normal University, Guangzhou, China,* [2]*Great Lakes Forestry Center, Sault Ste Marie, ON and* [3]*Laurentian Forestry Centre, Canadian Forest Service, Québec, QC, Canada*

Abstract A trypsin-like molting-related serine protease cDNA (*Cf*MRSP) was cloned from the spruce budworm, *Choristoneura fumiferana*. The full-length *Cf*MRSP complementary DNA (cDNA) encoded a 43 kDa protein that contained a trypsin-like serine protease catalytic domain, but no clip domain. The C-terminal extension contained five cystein residues, which may allow the protein to form a homodimer through interchain disulfide bonds and regulate the activity of *Cf*MRSP. Phylogenetic tree analysis showed that *Cf*MRSP clusters with lepidopteran homologues such as serine protease 1 of *Lonomia obliqua*, hemolymph proteinase 20 (HP20), pattern recognition serine proteinase precursor (ProHP14) and a trypsin-like protein of *Manduca sexta*. Northern blot analysis of developmental expression of *Cf*MRSP indicated that its transcripts were found primarily in the epidermis and were produced during all of the tested stadia, from 4th instar larvae to pupae, but increased levels of *Cf*MRSP transcripts were always found after each molt. A high level of the protein was found in the epidermis by immunohistochemistry analysis. Altogether these data suggest that *Cf*MRSP plays a role in the epidermis during molting and metamorphosis.

Key words *Choristoneura fumiferana*, cuticle, metamorphosis, molt, serine protease

Introduction

The serine protease family is a conserved group of proteolytic enzymes containing a serine protease domain and these enzymes play critical roles in a variety of insect biological processes, such as immune responses, fertilization and embryonic development (Jiang & Kanost, 2000). A complex serine protease cascade in insect hemolymph is suggested to coordinate innate immune responses such as hemolymph coagulation and melanization. Prophenoloxidase (PPO) is proteolytically activated by a serine protease, to generate phenoloxidase (PO), which catalyzes the formation of quinones used for melanin synthesis (Kanost *et al*., 2004). Some members of this protein family consist of two functional domains, a regulatory amino-terminal clip domain and a catalytic serine protease domain at the carboxyl terminus (Jiang *et al*., 2003a, 2003b).

The growth of holometabolous insects is periodically interrupted by cyclic molting, which is initiated and regulated by a complex set of biochemical and cellular events. Chief among them is the ecdysone release followed by the expression of ecdysone receptors and ecdysis-related transcription factors (Riddiford *et al*., 2000). During molting and metamorphosis, insects shed the old cuticle to accommodate body size increase and morphological changes. The endocuticle layer located between the epidermic cells and the exocuticle layer must undergo degradation and separation from the epidermis (Zheng *et al*., 2003). New chitin and cuticular proteins are subsequently produced and deposited on the surface of the epidermis, forming

Correspondence: Q. L. Feng, School of Life Sciences, South China Normal University, Tianhe, Guangzhou, 510631, China. Tel/fax: +86 20 85215291; email: qlfeng@scnu.edu.cn

© 2009 The Authors
Journal compilation © Institute of Zoology, Chinese Academy of Sciences

a new cuticle layer (Hopkins *et al*., 2000). Insect epidermical cuticle is composed mainly of crystalline microfibres of chitin embedded into a protein matrix. Two main types of enzymes are suggested to be involved in the degradation of the endocuticle of the exoskeleton: proteases and chitinolytic enzymes, which degrade the cuticular proteins and chitin, respectively. Accumulating evidence indicates that serine proteases are involved in the degradation of old cuticle and/or melanization and sclerotization of the new cuticle. Enzymes capable of hydrolyzing cuticular constituents are found in the molting fluid (Samuels & Reynolds, 1993a, 1993b). Samuels *et al.* (1993a) purified and characterized several proteolytic enzymes from the molting fluid of *Manduca sexta*, including a cuticle-degrading protease: molting fluid protease 1 (MFP-I). MFP-1 was identified as a trypsin-like protein based on its substrate specificity and inhibition by soybean trypsin inhibitor. It degraded *M. sexta* cuticular proteins *in vitro*, producing peptide fragments ranging in size between 200 and 2 000 Da (Samuels *et al*., 1993b). It is not clear whether or not the proteases are secreted into molting fluid as inactive proenzymes at the onset of cuticle degradation during the early stage of ecdysis (Samuels & Paterson, 1995). In *Mamestra configurata*, hemolymph protease activity was comprised mostly of serine proteases the overall activity of which increased dramatically at the onset of the molt concomitant with a sharp decline in gene expression of serine protease inhibitor (serpin) (Chamankhah *et al*., 2003). Silencing of Lm-TSP (*Locusta migratoria manilensis* trypsin-like serine protease) led to dramatic reductions in cuticle-degrading activity of proteases in the molting fluid, resulting in molting defects (Wei *et al*., 2007). All of these reports indicate that proteases are involved in the cuticle degradation in the epidermis during the molt and metamorphosis in insects.

In this study, we initiated the characterization of a trypsin-like serine protease, the cDNA of which was cloned from the spruce budworm, *Choristoneura fumiferana*, one of the most widely distributed destructive forest insect pests in North America. Developmental expression of the gene was studied by Northern blot analysis. The gene was found to be expressed specifically and periodically in the epidermis during larval development.

Materials and methods

Insects

Third instar larvae of the spruce budworm (*Choristoneura fumiferana* Clem., Lepidoptera: Tortricidae) were reared on artificial diet at 22°C, under 70% relative humidity and a photoperiod of 12 h light and 12 h dark, until they reached the pupal or adult stage. Fifth instar larvae were selected after the insects underwent head capsule slippage in order to accurately determine their next developmental stages.

To collect molting fluid, a small hole was made with a needle at the bottom of a 0.5-mL polymerase chain reaction (PCR) tube, which was then inserted into a 1.5-mL microfuge tube containing 1.5–2.0 μL 1× cold phosphate buffer saline. Old skin that had just separated from prepupae was collected and put into the 0.5 mL PCR tube, which was then centrifuged at 3 000 *g* for 5 s at 4°C. The buffered molting fluid was collected in the 1.5-mL tube and stored at −20°C until use for antibody production.

Cloning of CfMRSP cDNA

A complementary DNA (cDNA) clone encoding a trypsin-like serine protease was identified from a spruce budworm epidermis expression cDNA library by the antibodies generated against a mixture of the molting fluid proteins. Primers were designed based on the cloned cDNA sequence to amplify the 5′- and 3′-ends of the cDNA by rapid amplification of cDNA ends (RACE) PCR using the SMART^{TM} RACE cDNA Amplification Kit (Clontech, Mountain Virew, CA, USA) according to the manufacturer's instructions. For 3′-end RACE, primer A was used: 5′- ACGTCAAACTACAGTC AAAATAATGAAC-3′; for 5′-end RACE, primer B was used: 5′-CCTCCGCTACATCACCAGCGACAAGA-3′. The RACE PCR thermal cycle profile was as follows: denature for 2 min at 94°C, five cycles of 94°C for 30 s, 70°C for 30 s and 72°C for 3 min, followed by 25 cycles of 94°C for 30 s, 68°C for 30 s and 72°C for 3 min, and finally 72°C for 7 min. Amplified PCR products were agarose-gel purified, cloned into pMD-18T vector (Takara, Dalian, China) and sequenced.

Sequence comparison and phylogentic analysis

DNA sequence analysis and comparison were performed using the DNASTAR program and the BLAST program on the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/). Amino acid sequence alignment and phylogenetic analysis were carried out using the CLUSTAL W method (http://align.genome.jp/). Predictions of signal sequence and functional motifs were made using InterProscan (http://www.ebi.ac.uk/tools/InterProscan). The amino acid sequences for sequence alignment and phylogenetic tree included

© 2009 The Authors

Lonomia oblique serine protease 1 (*L. oblique* SP1, AAV91432, Veiga *et al*., 2005), *Manduca sexta* hemolymph proteinase 20 (*M. sexta* HP20, AAV91018, Jiang *et al*., 2005), *M. sexta* pattern recognition serine protease precursor (*M. sexta* PRSPP, AAR29602, Ji *et al*., 2004), *M. sexta* serine protease-like protein (*M. sexta* SPLP, AAO21508), *Tribolium castaneum* pattern recognition serine proteinase (*T. castaneum* PRSP, XP_967486), *Tenebrio molitor* modular serine protease zymogen (*T. molitor* MSPZ, BAG14264, Kim *et al*., 2008), *Culex quinquefasciatus* homologue (XP_001864236), *Aedes aegypti* homologue (XP_001655952), *Anopheles gambiae* str. PEST homologue (XP_321263) and *Nasonia vitripennis* homologue (XP_001607879).

Production of recombinant CfMRSP protein and antibody

A fragment of the cloned cDNA was expressed as a recombinant protein in a bacterial expression system. The 3′-end of the open reading frame (ORF) of the *Cf*MRSP cDNA was amplified by PCR and inserted into the pGEX-5x-3 expression vector (Promega, Madison, WI, USA) between the *Eco*RI and *Xho* I restriction sites, fusing with the glutathione S-transferase (GST) coding sequence at the 5′-end. The recombinant pGEX-*Cf*MRSP/GST plasmid DNA was used to transform DH5α *Escherichia coli* cells for protein expression. The recombinant *Cf*MRSP/GST was purified using a GST-glutathione affinity chromatography system from Amersham Bioscience (Piscataway, NJ, USA) following the manufacturer's protocol. The purified recombinant protein was injected into New Zealand rabbits. Antiserum was collected after three booster injections, each containing 100 μg of the recombinant protein.

SDS-PAGE and Western blotting

Protein samples were denatured by boiling for 5 min following addition of an equivalent volume of 2 × SDS (sodium dodecyl sulfate) sample buffer. SDS-PAGE (polyacrylamide gel electrophoresis) was performed in 10% polyacrylamide gels, which were then stained with Coomassie Brilliant Blue R-250. For Western blot analysis, proteins (20 μg per sample) were transferred from acrylamide gels to Hybond-P PVDF Transfer membranes (Amersham, Piscataway, NJ, USA). The membranes were blocked with 3% bovine serum albumen (BSA) in 1 × Tris-buffered saline Tween-20 (TBST) buffer for 2 h at room temperature, and then incubated with the anti- *Cf*MRSP/GST antibodies (1 : 1 000) at room temperature for 1 h. Goat anti-rabbit IgG conjugated with alkaline phosphatase was used as the secondary antibody at a dilution of 1 : 3 000. Nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate were used as substrates for color development.

Northern blotting

Total RNA was prepared from whole larvae or different tissues using the standard Trizol reagent (Ambion, Austin, TX, USA). An aliquot of 5 μg RNA per well was electrophoresed in formaldehyde denaturing gels and then blotted onto Nuclear N membrane (Amersham). A 930-bp DNA fragment of *Cf*MRSP was synthesized *in vitro* from linearized plasmid using α-^{32}P dCTP and the Random Primer DNA Labeling System's Kit (GE Healthcare, Vancouver, BC, Canada) and used as a probe. Pre-hybridization, hybridization, and post-hybridization washing were carried out as described by Béliveau *et al.* (2000).

Immunohistochemistry

Whole larvae at selected stages were fixed with 4% formaldehyde in 1 × phosphate buffered saline (PBS) for 24 h at 4°C and embedded in paraffin. Five-micrometer thick serial cross-sections were prepared for immunostaining. Deparaffined sections were blocked by incubation for 1 h in 1 × PBS containing 3% BSA and 1.5% sheep serum. The sections were then incubated overnight with anti-*Cf*MRSP/GST antibodies at 4°C. After washing three times with 1 × PBS buffer, the sections were incubated for 1 h with fluorescein-labeled sheep anti-rabbit IgG (Boshide, Wuhan, China) at a dilution of 1 : 500. After three washes in 1 × PBS buffer the sections were observed and photographed under a Leica fluorescence microscope (DMI 4000B).

Results

Cloning and sequence analysis of CfMRSP

A partial sequence encoding a trypsin-like serine protease was identified from a *C. fumiferana* epidermis expression cDNA library screened by the antibodies generated against molting fluid proteins. A full-length cDNA sequence was assembled after 5′- and 3′-terminal fragments were amplified by RACE PCR using RNA isolated from the 6th larval epidermis. The full-length cDNA encoded a trypsin-like serine protease. This cDNA (*C. fumiferana* molting-related serine protease, *Cf*MRSP,

© 2009 The Authors

```
   1  CGGCACGAGGCTGTCCTCAAGCGGGTCGCGGTTTATGGTCATAACTCGTACAGAATAAGAGGATTACACA
  71  ATCCAAATTATGTATTTATACGCCGTGATATTAATAGTTTCAGGTTTATTTTTATCTCTAGGATCCGTTG
                     M  Y  L  Y  A  V  I  L  I  V  S  G  L  F  L  S  L  G  S  V  G        21
 141  GAGCCGGTCAAACTGACGATAATGAGGAAATTAAGCAATTTTACTCTAATATAACGTCAAACTACAGTCA
        A  G  Q  T  D  D  N  E  E  I  K  Q  F  Y  S  N  I  T  S  N  Y  S  Q             44
 211  AAATAATGAACAACCAGAATACGACGACGAAGATGATTTCGCCAGAGTAGCAATCGACGGGGAACCTGTG
       N  N  E  Q  P  E  Y  D  D  E  D  D  F  A  R  V  A  I  D  G  E  P  V              67
 281  CCACCATTCCATATATTATATCCGTCACCTGGATCCCAGAGCATACCGGCATCTATCGGCGCGCAGCCAA
      P  P  F  H  I  L  Y  P  S  P  G  S  Q  S  I  P  A  S  I  G  A  Q  P  T            91
 351  CAACAAACGGCCACTCCATGTTGTGGCACGTCGGTGTATACAGCAAGCGCTTCACTCCATACATGCAGAT
        T  N  G  H  S  M  L  W  H  V  G  V  Y  S  K  R  F  T  P  Y  M  Q  I            114
 421  TTGTGGTGGTACCCTGGTTACCATGAAAGCTGTCATTTCAGCGGCGCATTGCTTCTGGAGTGACGTTGAA
       C  G  G  T  L  V  T  M  K  A  V  I  S  A  A [H] C  F  W  S  D  V  E             137
 491  GGCGCAATGCCAGCAGAGGATTTTGCGGTGGGCACCGGCAAGCGGTTCCGACCTTGGGATTTTCTCCAAG
      G  A  M  P  A  E  D  F  A  V  G  T  G  K  R  F  R  P  W  D  F  L  Q  D           161
 561  ACAACTTTAGCCAGATGTCACAAGTGTCAGCGATCCACATCCCAGTCCGATACCATGACGCAGCCACCAG
        N  F  S  Q  M  S  Q  V  S  A  I  H  I  P  V  R  Y  H  D  A  A  T  S            184
 631  CTACCAGGAGGATATTGCCGTCTTACTTCTGGCTCAGGAGTTCACGGTCAGTCAGGCTGTCCGTCCGGCC
       Y  Q  E [D] I  A  V  L  L  L  A  Q  E  F  T  V  S  Q  A  V  R  P  A             207
 701  TGTGTCAACTTCGAAGAAGATTTTGACAACGAGCAATTGAAGGAAGGAAACTTAGGAACTATCCTGGGTT
      C  V  N  F  E  E  D  F  D  N  E  Q  L  K  E  G  N  L  G  T  I  L  G  W           231
 771  GGGGGCTGACAGAGGAAGATGGAGATCCCTCACAAATACTCAACTTCCTGGAGCTGCCCACCGTCAGTAT
        G  L  T  E  E  D  G  D  P  S  Q  I  L  N  F  L  E  L  P  T  V  S  I            254
 841  TGAGGAGTGTCTGAAAAGCGCTGACGCGAGCTTCCTCCGCTACATCACCAGCGACAAGATCTGCGCAGGC
       E  E  C  L  K  S  A  D  A  S  F  L  R  Y  I  T  S  D  K  I  C  A  G             277
 911  GTCACTACGGGTAAAGCCCTCTGTCGTGGCGACAGCGGCGGCGGACTGGTGTTTACGGAGCAGAGAGGAT
      V  T  T  G  K  A  L  C  R  G  D [S] G  G  G  L  V  F  T  E  Q  R  G  S           301
 981  CCGTGCCCGTCCCGTACCTGCGCGGGGTGGCGTCCACCGCGCCGCGCAACGAGCACCACTGCAACGCGTA
        V  P  V  P  Y  L  R  G  V  A  S  T  A  P  R  N  E  H  H  C  N  A  Y            324
1051  CGCCCTCGCTGCACTTACGCATGTTCACGCGCACCGAGCGTTCCTGCGACGATACATCCCAAACTTGGAA
       A  L  A  A  L  T  H  V  H  A  H  R  A  F  L  R  R  Y  I  P  N  L  E             347
1121  GACGAGTGCGAGAGGCATTACCCTTCAGCAGATGCCCTGAAAGTTAGGTACAGCAACAATGAACCTGCCG
      D  E [C] E  R  H  Y  P  S  A  D  A  L  K  V  R  Y  S  N  N  E  P  A  V           371
1191  TCTGCCAGCCAGTTCAAGTGACGTGCAATTGTTTCTGCCAAAATGCTACGACTTCTTAAAAAGAGACTGT
       [C] Q  P  V  Q  V  T [C] N [C] F [C] Q  N  A  T  T  S  ***                      389
1261  TTTATATGACCCTCGTCTGCGGCTTAGTTAACACACTGGTTAGCGCAGTAGCGCTGCAAATTTTGCATGA
1331  GTTTATATGTTTTAAAATAACGCCGATAAAGTAATAAAATCTAGAAAAATCGAAAAAAAAAAAAAAAAAA
1401  AAAAAAAAAAA
```

Fig. 1 Nucleotide and deduced amino acid sequences of *Choristoneura fumiferana* trypsin-like molting-related serine protease complementary DNA (cDNA) (*Cf*MRSP). The predicted signal peptide is underlined. The putative *N*-glycosylation sites at positions 37, 41 and 162 are shown in bold. The amino acids of the putative catalytic triad are indicated by white boxes. Five cystein residues in the C-terminal extension are indicated by grey boxes. The stop codon is indicated by stars. The numbers on the left refer to the nucleotide positions, while the numbers on the right refer to the amino acid positions. The GenBank accession number for this cDNA sequence is FJ628776.

GenBank accession no.: FJ628776) contained 1 411 base pairs (bp), including an open reading frame of 1 179 bp, a 5′-end untranslated region (UTR) of 79 bp and a 3′-end UTR of 153 bp (Fig. 1). The deduced peptide sequence contained 389 amino acids, with a predicted molecular mass of 43 kDa and a isoelectric point (pI) of 4.73. A putative signal peptide consisting of 22 amino acid residues was predicted at the N-terminus, the cleavage of which would generate a protein of 40.5 kDa. Three putative N-glycosylation sites at positions 37, 41 and 162 were identified. The peptide contained a trypsin-like serine protease (or proteinase-catalytic) domain and a catalytic triad found in many serine proteases (Rypniewski *et al*., 1994) (Figs. 1 and 2A). *Cf*MRSP

© 2009 The Authors

does not possess an N-terminal clip-domain, a low-density lipoprotein receptor repeat (LDLr), an RXR motif, 7C region or a Sushi domain. The C-terminal extension contained five Cys residues, which may form a novel domain structure and allow the protein molecules to form homodimers through interchain disulfide bonds. Alignment of the trypsin-like serine protease-reactive domain and the C-terminal extension of the *Cf*MRSP protein with those regions of the top 10 serine protease homologues in other insect species revealed that the region of the trypsin-like serine protease domain is conserved, while the C-terminal extension greatly varies in length (Fig. 2A). Phylogenetic analysis by comparing the serine protease reactive domain and the C-terminal extension showed that *Cf*MRSP was clustered together closely with *Lonomia obliqua* serine protease 1 (AAV91432), which was identified from the cuticular bristles and tegument of the caterpillar. This cluster was grouped with three serine proteases from *M. sexta*, including, hemolymph proteinase 20 (MsHP20) (AAV91018, Jiang *et al*., 2005), pattern recognition serine

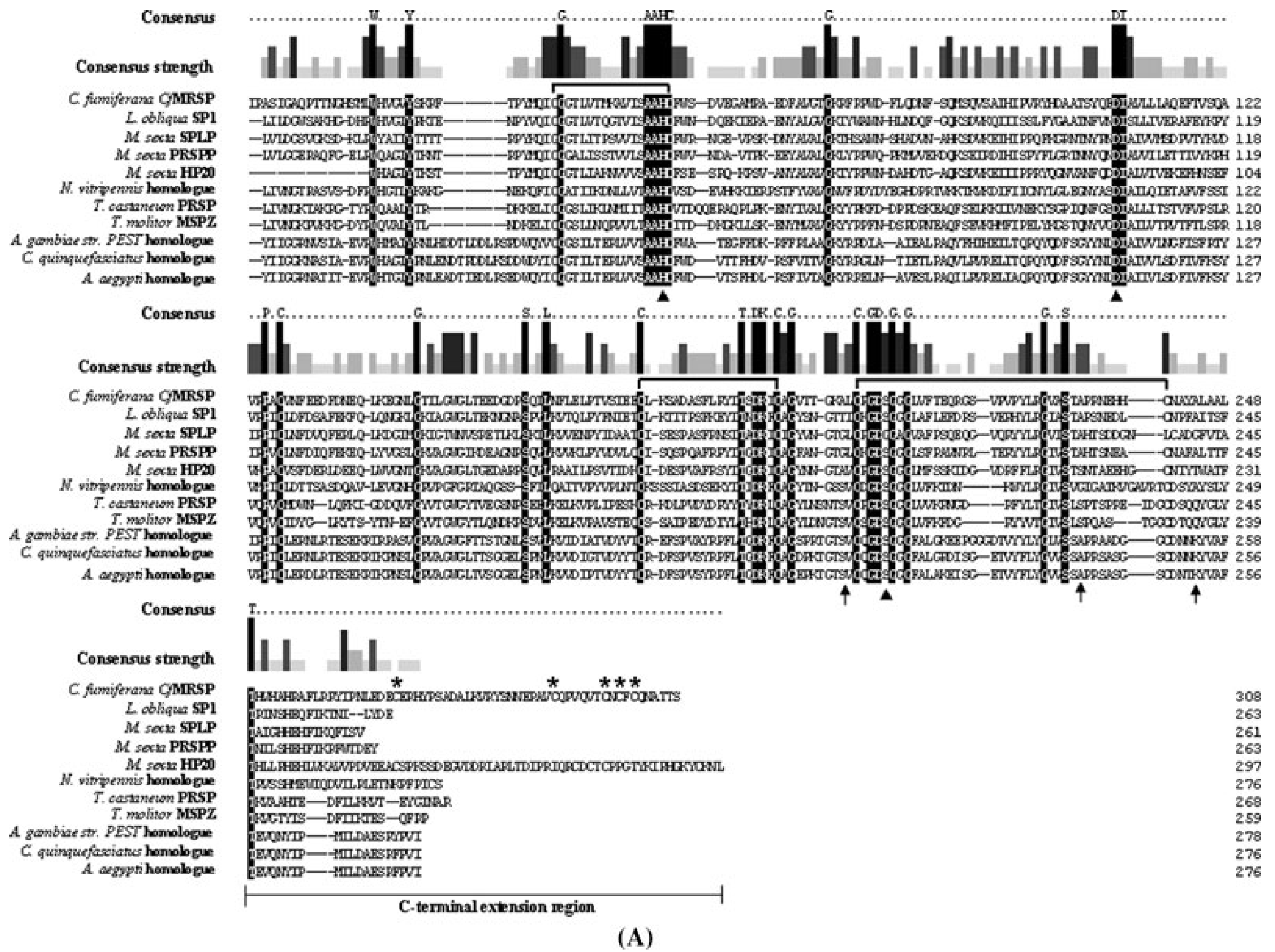

Fig. 2 (A) Alignment of the trypsin-like serine protease reactive (proteinase-catalytic) domain and the C-terminal extension of the *Cf*MRSP (FJ628776) protein with the similar regions of the top 10 serine protease homologues in other insect species (see Materials and methods). Numbers on the right indicate the positions of amino acid residues. Consensus residues that are identical among all sequences are indicated in white letters in black background and are shown at the top of the panels. Consensus strength is shown by the scale columns above the sequence panels. The Asp, His and Ser residues of the catalytic triad are identified by arrowheads. The positions of three amino acid residues responsible for specificity in the pattern recognition serine protease in *M. sexta* (Ji *et al*., 2004) are identified by arrows. Three pairs of Cys residues (Cys115 and Cys131, Cys257 and Cys275, and Cys285 and Cys321) identified as forming disulfide bonds are connected by a line. Five cystein residues in the C-terminal extension are indicated by stars. (B) Phylogenetic tree analysis of *Cf*MRSP with its homologues from other insect species, in which only the trypsin-like serine protease reactive domain and the C-terminal extension are compared and used for analysis.

© 2009 The Authors

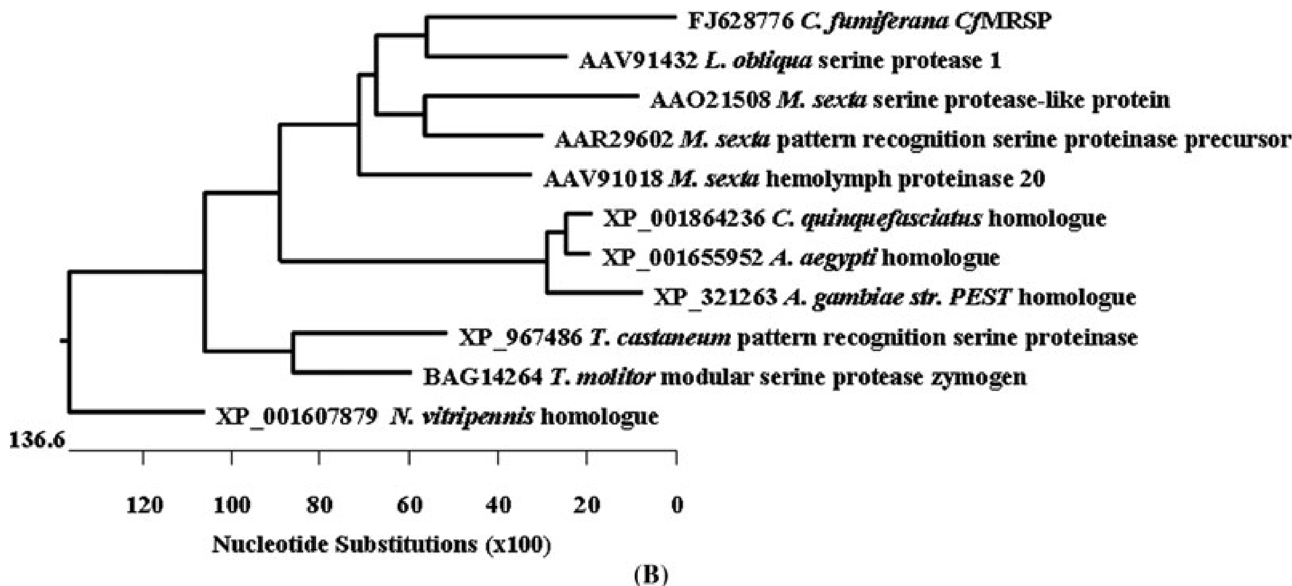

Fig. 2 Continued.

proteinase precursor (MsProHP14) (AAR29602, Ji *et al*., 2004), and a serine protease-like protein (AAO21508) (Fig. 2B). Pattern recognition serine proteinases from different species, including *M. sexta*, *Tribolium castaneum* and *Tenebrio molitor*, were not closely clustered together when the proteinase-catalytic domain and the C-terminal extension were compared.

In vitro *expression of CfMRSP*

A fragment encoding the first 308 amino acids of the *Cf*MRSP cDNA was expressed in *E. coli*, generating a recombinant protein fused with GST (Fig. 3). The *Cf*MRSP/GST fusion protein had an apparent molecular mass of 60 kDa, as determined by SDS-PAGE. Given that the predicted molecular mass of GST is 26 kDa, the observed size of the recombinant *Cf*MRSP protein was 34 kDa. If we add to this value the predicted size of the 81 amino acid protein of *Cf*MRSP not included in the construct, we obtain 42.9 kDa, which is in agreement with the mass predicted based on the cDNA. Antibodies were generated against the recombinant *Cf*MRSP/GST protein. The antibodies could immunologically recognize the recombinant protein in Western blots (Fig. 3B), but did

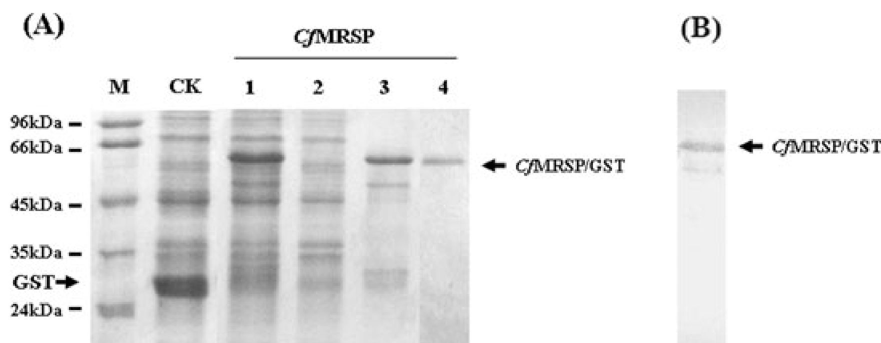

Fig. 3 (A) Sodium dodecyl sulfate – polyacrylamide gel electrophoresis analysis of expression and purification of the recombinant *Cf*MRSP protein in a bacterial expression system. The recombinant *Cf*MRSP protein fused with glutathione S-transferase (GST) tag was 60 kDa. M: molecular mass of standard protein. CK: control proteins expressed in pGEX-5x-3 vector alone. Lane 1: total protein; Lane 2: soluble fraction; Lane 3: insoluble fraction. Lane 4: purified protein. (B) Western blot showing the immunoreaction of the antibody with the recombinant *Cf*MRSP-GST protein.

© 2009 The Authors

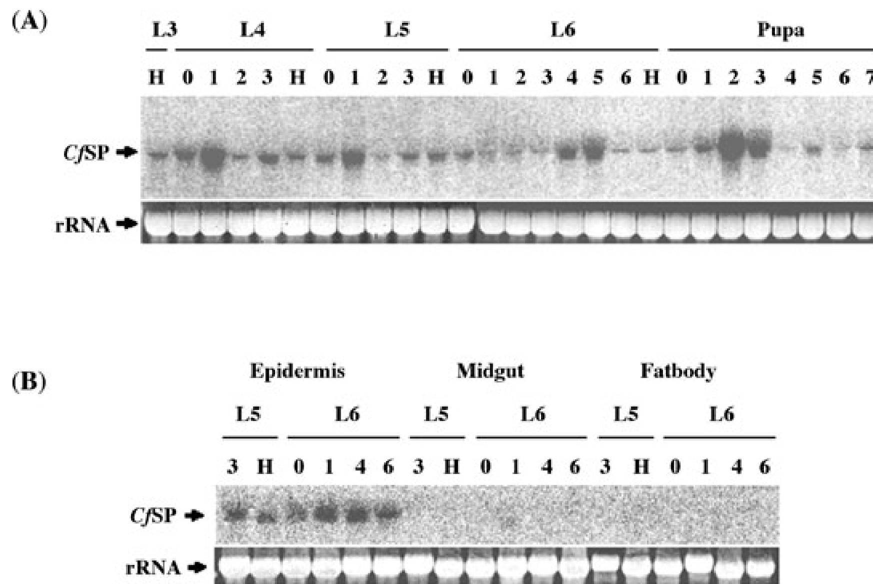

Fig. 4 Northern blot analysis of developmental (A) and tissue-specific (B) expression of *Cf* MRSP. (A) RNA was collected from whole larvae from the late 3rd instar to day 7 of the pupal stage. (B) RNA was isolated from the midgut, epidermis and fat body at indicated days of the 5th and 6th instars. Recombinant RNA (rRNA) was stained by ethidium bromide to verify equal loading of total RNA (20 μg). H: head capsule slippage stage.

not cross-react with five *C. fumiferana* GSTs identified previously (Feng *et al.*, 1999; Zheng *et al.*, 2007; Huang *et al.*, 2009) (data not shown).

Developmental and tissue-specific expression of CfMRSP

Developmental expression of the *Cf* MRSP gene in 4th to 6th instar larvae and pupae was examined using Northern blot analysis (Fig. 4A). *Cf* MRSP transcripts were detected in all of the tested stadia from 4th instar larvae to pupae. Northern blots revealed that *Cf* MRSP expression fluctuated depending on developmental stage; for example, relatively high levels of *Cf* MRSP transcripts were detected on day 1 (D1) of the 4th and 5th instar larvae, on days 4 and 5 of 6th instar or on days 2 and 3 of pupae, although the expression levels at these time points were to some extent different (Fig. 4A). Among the three tissues examined in 4th and 5th instars, *Cf* MRSP transcripts were abundant in the epidermis, while they were not detected in the fat body and midgut. The pattern of developmental and tissue-specific expression of *Cf* MRSP indicated that its expression was induced in the epidermis after each molt, 1–4 days after ecdysis. In the case of 6th instars, the rise in transcript abundance also coincided with the onset of metamorphosis.

Immunohistochemical localization of CfMRSP

Localization of *Cf* MRSP in the 6th instar larva was examined using immunohistochemical analysis (Fig. 5). At the 6th instar feeding stage, *Cf* MRSP protein was detected predominantly in the epidermis, although the protein was also detected in the fat body, muscles and midgut, but at relatively lower levels.

Discussion

The functions of serine proteases in the melanization cascade during the innate immune response of insects have been well studied (Jiang & Kanost, 2000; Kanost *et al.*, 2004; Jiang *et al.*, 2005; Tong *et al.*, 2005). Insects respond to injury or invasion of micro-organisms by activating prophenoloxidase (PPO) through the action of serine proteinases. Activated phenoloxidase hydroxylates monophenols, converting them to *O*-quinones via *O*-diphenols. Quinones polymerize to form melanin at the injury site or around the invading organisms to kill the invaders. In addition, quinones have been shown to be involved in sclerotization of the cuticle (Sugumaran *et al.*, 1992).

However, the roles of serine proteases in insects are not limited to immune responses. Accumulating

© 2009 The Authors

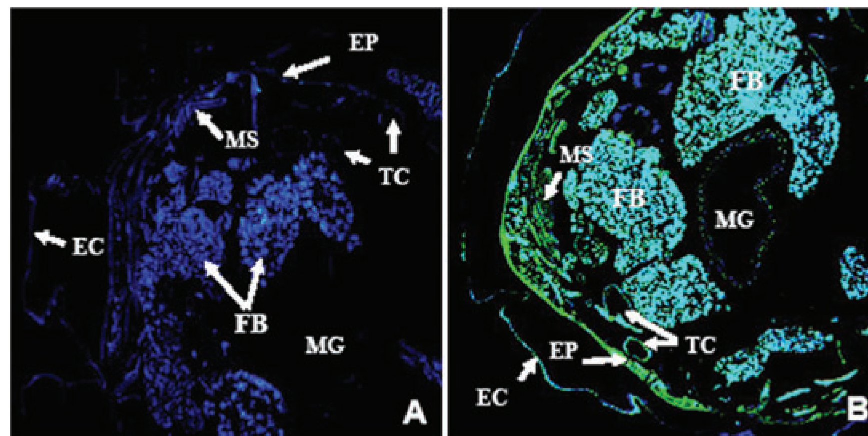

Fig. 5 Immunohistochemical staining of cross-section of *C. fumiferana* larvae on day 6 (prepupal) of the 6th instar stage. Five-micron cross-sections were immunostained with anti-*Cf*MRSP/GST antiserum, followed by goat anti-rabbit IgG conjugated with streptavidin biotin complex – fluorescein isothiocyanate (SABC-FITC) and counter-stained with 4′,6-diamidino-2-phenylindole (DAPI). The sections were observed by fluorescence microscopy and the photographs were taken as double exposure with DAPI filters (A) and fluorescein (B). The bright green fluorescence indicates the presence of *Cf*MRSP protein. EC: exocuticle; EP: epidermis; FB: fat body; MS: muscle; MG: midgut; TC: trachea. The bars represent 50 μm.

experimental evidence suggests that some serine proteases play important roles in the processes of molting and metamorphosis. Insects periodically shed their old cuticle and tracheae so as to allow increases in body size as well as morphological and structural changes. In the context of these processes, cuticular chitin-binding proteins in the integument and tracheae must be degraded prior to ecdysis. After each molt, the newly produced cuticle undergoes melanization and sclerotization. Chamankhah *et al*. (2003) found that, in *M. configurata* hemolymph, serine protease activity increased significantly at the onset of the molt, while expression of four serpin isoforms decreased sharply, suggesting that these serine proteases and serpins may be involved in the regulation of molting. Wei *et al*. (2007) isolated a trypsin-like serine protease, *Lm*tsp, from *L. manilensis manilensis*, and found that silencing of *Lm*tsp by RNA interference led to dramatic reductions in protease and cuticle-degrading activity in the molting fluid, resulting in molting defects from the 4th to 5th instars, as well as between the 5th instar and the adult stage. These reports suggest that trypsin-like serine proteases may play important roles in degrading cuticular proteins and/or in melanization and sclerotization during molting and metamorphosis.

Several findings in the present study suggest that *Cf*MRSP may be involved in molting and metamorphosis. First, *Cf*MRSP was first identified by using the antiserum generated against a mixture of the molting fluid proteins to screen an expression cDNA library made with RNA from the epidermis of *C. fumiferana* 6th instar larvae, indicating that this protein was present in the molting fluid. *Cf*MRSP has a predicted signal peptide and it is possible that the mature protein is secreted into the molting fluid after synthesis. Second, the gene was expressed specifically and periodically in the epidermis and a high level of the protein was found by immunohistochemistry in the epidermal cells. Third, the *Cf*MRSP transcripts were detected at all of the tested stages from 4th instar larvae to pupae, but higher levels of *Cf*MRSP transcripts were always found at variable time points after each of the larval molts, it is possible that the *Cf*MRSP protein is produced and stored as an inactive zymogen during the intermolt phases (Jiang & Kanost, 2000; Kanost *et al*., 2004; Kaji *et al*., 2009) and become active during next molting. Another possibility is that *Cf*MRSP may not be directly involved in cuticular protein degradation in the epidermis during molting, because cuticular protein degradation should take place before the shedding of the old cuticle. As the pattern of changes in *Cf*MRSP transcript abundance was consistent with the events of melanization and sclerotization of the newly formed cuticle after each molt, it is possible that this gene is involved in the melanization and/or sclerotization of the new cuticle.

The *Cf*MRSP protein is most homologous to *L. oblique* serine protease 1 (AAV91432, Veiga *et al*., 2005), which was identified from the cuticular bristles and tegument of the caterpillar. This homology between *Cf*MRSP and *L. obliqua* SP1 implies that these proteins may play their functions in degradation of cuticular proteins or melanization and/or sclerotization of new cuticle of the

© 2009 The Authors
Journal compilation © Institute of Zoology, Chinese Academy of Sciences,

epidermis. This subgroup separated from the hemolymph proteases, such as the hemolymph protease (HP) 20 (Jiang *et al*., 2005), pattern recognition serine proteinase precursor (ProHP14) (Ji *et al*., 2004) and serine protease-like protein of *M. sexta*. *Cf*MRSP, is expressed predominantly in the epidermis, whereas MsHP20 and MsHP14 are expressed predominantly in hemocytes and in the fat body, respectively. Both MsHP20 and MsHP14 are involved in the immune response in the hemolymph, but the *Cf*MRSP protein may play a role in the epidermis. Thus, *Cf*MRSP appears to be a serine protease different from the hemolymph proteases MsHP20 and MsHP14. One of the remarkable features of *Cf*MRSP is that its C-terminal extension contained five Cys residues. Because there is no clip domain in the *N*-terminal region of *Cf*MRSP, this C-terminal novel domain may provide a regulatory mechanism for the activity of *Cf*MRSP, for example by forming homodimer through interchain disulfide bonds.

At the RNA level, *Cf*MRSP was observed to be present as a single transcript specifically expressed in the epidermis. However, positive signals were also detected in the fat bodies and the muscles at low levels. Although the possibility that the antibodies can cross-react with *C. fumiferana* GSTs cannot be completely ruled out, at least five previously isolated *C. fumiferana* GSTs were immunologically reactive (data not shown). An alternative and more likely possibility is that the antibody could recognize other serine proteases, given that the large number of prophenoloxidase-activating proteases and hemolymph proteases are present in insects (Jiang & Kanost, 2000; Gorman & Pakewitz, 2001; Christophides *et al*., 2002; Jiang *et al*., 2005). Whether the protein was transported to other tissues after the gene is transcribed in the epidermis or the antibody-recognized proteins are in fact other serine proteases needs to be further investigated and clarified.

Acknowledgments

This research was supported by the National Natural Science Foundation (Grant No.: 30571258) and the Guangdong Natural Science Foundation (Grant No.: 5005941) of China and Genome Canada Grant through the Ontario Genomics Institute of Canada.

References

Béliveau, C., Laforge, M., Cusson, M. and Bellemare, G. (2000) Expression of a *Tranosema rostrale* polydnavirus gene in the spruce budworm, *Choristoneura fumiferana*. *Journal of General Virology*, 81, 1871–1880.

Chamankhah, M., Braun, L., Visal-Shah, S., O'Grady, M., Baldwin, D., Shi, X., Hemmingsen, S.M., Alting-Mees, M. and Hegedus, D.D. (2003) *Mamestra configurata* serpin-1 homologues: cloning, localization and developmental regulation. *Insect Biochemistry and Molecular Biology*, 33, 355–369.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Müller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P. and Kafatos, F.C. (2002) Immunity-related genes and gene families in *Anopheles gambiae*. *Science*, 298, 159–165.

Feng, Q.L., Davey, K.G., Pang, A.S.D., Primavera, M., Ladd T.R., Zheng S.-C., Sohi, S.S., Retnakaran, A. and Palli, S.R. (1999) Glutathione S-transferase from the spruce budworm, *Choristoneura fumiferana*: identification, characterization, localization, cDNA cloning, and expression. *Insect Biochemistry and Molecular Biology*, 29, 779–793.

Gorman, M.J. and Pakewitz, S.M. (2001) Serine proteases as mediators of mosquito immune responses. *Insect Biochemistry and Molecular Biology*, 31, 257–262.

Hopkins, T.L., Krchma, L.J., Ahmad, S.A. and Kramer, K.J. (2000) Pupal cuticle proteins of *Manduca sexta*: characterization and profiles during sclerotization. *Insect Biochemistry and Molecular Biology*, 30, 19–27.

Huang, Y.-F., Krell, P.J., Ladd, T.R., Feng, Q. and Zheng, S.-C. (2009) Cloning, characterization and expression of two glutathione *S*-transferase cDNAs in the spruce budworm, *Choristoneura fumiferana*. *Archive of Insect Biochemistry and Physiology*, 70, 44–56.

Ji, C., Wang, Y., Guo, X., Hartson, S. and Jiang, H.A. (2004) Pattern recognition serine proteinase triggers the prophenoloxidase activation cascade in the tobacco hornworm, *Manduca sexta*. *Journal of Biological Chemistry*, 279, 34101–34106.

Jiang, H. and Kanost, M.R. (2000) The clip-domain family of serine proteinases in arthropods. *Insect Biochemistry and Molecular Biology*, 30, 95–105.

Jiang, H., Wang, Y., Yu, X.-Q., Zhu, Y. and Kanost, M.R. (2003a) Prophenoloxidase-activating proteinase-2 (PAP-2) from hemolymph of *Manduca sexta*: a bacteria-inducible serine proteinase containing two clip domains. *Journal of Biological Chemistry*, 278, 3552–3561.

Jiang, H., Wang, Y., Yu, X.-Q., Zhu, Y. and Kanost, M.R. (2003b) Prophenoloxidase-activating proteinase-3 (PAP-3) from *Manduca sexta* hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. *Insect Biochemistry and Molecular Biology*, 33, 1049–1060.

© 2009 The Authors

Jiang, H., Wang, Y., Gu, Y.L., Guo, X.P., Zou, Z.H., Scholz, F., Trenczek, T. and Kanost, M. (2005) Molecular identification of a bevy of serine protease in *Manduca sexta* hemolymph. *Insect Biochemistry and Molecular Biology*, 35, 931–943.

Kaji, K., Tomino, S. and Asano, T. (2009) A serine protease in the midgut of the silkworm, *Bombyx mori*: protein sequencing, identification of cDNA, demonstration of its synthesis as zymogen form and activation during midgut remodeling. *Insect Biochemistry and Molecular Biology*, 39, 207–217.

Kanost, M.R., Jiang, H. and Yu, X. (2004) Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunology Review*, 198, 97–105.

Kim, C.H., Kim, S.J., Kan, H., Kwon, H.M., Roh, K.B., Jiang, R., Yang, Y., Park, J.W., Lee, H.H., Ha, N.C., Kang, H.J., Nonaka, M., Soderhall, K. and Lee, B.L. (2008) Three-step proteolytic cascade mediates the activation of the peptidoglycan-induced Toll pathway in an insect. *Journal of Biological Chemistry*, 283, 7599–7607.

Riddiford, L.M., Cherbas, P., Truman, J.W. (2000) Ecdysone receptors and their biological actions. *Vitamin and Hormones*, 60, 1–73.

Rypniewski, W.R., Perrakis, A., Vorgias, C.E. and Wilson, K.S. (1994) Evolutionary divergence and conservation of trypsin. *Protein Engineering*, 7, 57–64.

Samuels, R.I., Charnley, A.K. and Reynolds, S.E. (1993a) A cuticle degrading proteinase from the moulting fluid of the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry and Molecular Biology*, 23, 607–614.

Samuels, R.I., Charnley, A.K. and Reynolds, S.E. (1993b) An aminopeptidase from the moulting fluid of the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry and Molecular Biology*, 23, 615–620.

Samuels, R.I. and Paterson, I.C. (1995) Cuticle degrading proteinase from insect moulting fluid and culture filtrates of entomopathogenic fungi. *Comparative Biochemistry and Physiology*, 110B, 661–669.

Samuels, R.I. and Reynolds, S.E. (1993a) Molting fluid enzymes of the tobacco hornworm, *Manduca sexta*: Timing of proteolytic and chitinolytic activity in relation to pre-ecdysial development. *Archive of Insect Biochemistry and Physiology*, 24, 33–44.

Samuels, R.I. and Reynolds, S.E. (1993b) Moulting fluid enzymes of the tobacco hornworm, *Manduca sexta*: Inhibitory of 20-hydroxyecdysone on the activity of the cuticle degrading enzyme MFP-1. *Insect Physiology*, 39, 633–637.

Sugumaran, M., Semensi, V., Kalyanaraman, B., Bruce, J.M. and Land, E.J. (1992) Evidence for the formation of a quinone methide during the oxidation of the insect cuticular sclerotizing precursor 1, 2-dehydro-N-acetyldopamine. *Journal of Biological Chemistry*, 267, 10355–10361.

Tong, Y., Jiang, H. and Kanost, M.R. (2005) Identification of plasma proteases inhibited by *Manduca sexta* serpin-4 and serpin-5 and their association with components of the prophenoloxidase activation pathway. *Journal of Biological Chemistry*, 280, 14932–14942.

Veiga, A.B., Ribeiro, J.M., Guimarães, J.A. and Francischetti, I.M. (2005) A catalog for the transcripts from the venomous structures of the caterpillar *Lonomia obliqua*: identification of the proteins potentially involved in the coagulation disorder and hemorrhagic syndrome. *Gene*, 355, 11–27.

Wei, Z., Yin, Y.P., Zhang, B.S.H., Wang, Z.H.K., Peng, G.X., Cao, Y.Q. and Xia, Y.X. (2007) Cloning of a novel protease required for the molting of *Locusta migratoria manilensis*. *Development, Growth and Differentiation*, 49, 611–621.

Zheng, S.-C., Deng, H.-M., Ladd, T.R., Tomkins, B., Krell, P.J. and Feng, Q. (2007) Clonging and characterization of two glutathione *S*-transferase cDNAs in the spruce budworm, *Choristoneura fumiferana*. *Archive of Insect Biochemistry and Physiology*, 66, 146–157.

Zheng, Y.P., Retnakaran, A., Krell, P.J., Arif, B.M., Primavera, M. and Feng, Q.L. (2003) Temporal, spatial and induced expression of chitinase in the spruce budworm, *Choristoneura fumiferana*. *Journal of Insect Physiology*, 49, 241–247.

Accepted April 13, 2009

© 2009 The Authors
Journal compilation © Institute of Zoology, Chinese Academy of Sciences,