Identification and Characterization of C-type Lectins in *Ostrinia furnacalis* (Lepidoptera: Pyralidae)

Dongxu Shen, Lei Wang, Jiayue Ji, Qizhi Liu, and Chunju An

Department of Entomology, College of Plant Protection, China Agricultural University, Beijing 100193, P.R. China and 1Corresponding author; e-mail: anchunju@cau.edu.cn

Received 21 November 2017; Editorial decision 22 January 2018

Abstract

C-type lectins (CTLs) are a large family of calcium-dependent carbohydrate-binding proteins. They function primarily in cell adhesion and immunity by recognizing various glycoconjugates. We identified 14 transcripts encoding proteins with one or two CTL domains from the transcriptome from Asian corn borer, *Ostrinia furnacalis* (Gueneé; Lepidoptera: Pyralidae). Among them, five (OfCTL-S1 through S5) only contain one CTL domain, the remaining nine (OfIML-1 through 9) have two tandem CTL domains. Five CTL-Ss and six OfIMLs have a signal peptide and are likely extracellular while another two OfIMLs might be cytoplasmic. Phylogenetic analysis indicated that OfCTLs and OfIMLs had 1:1 orthologs in Lepidoptera, Diptera, Coleoptera and Hymenoptera species, but OfIMLs only clustered with immulectins (IMLs) from Lepidoptera. Structural modeling revealed that the 22 CTL domains adopt a similar double-loop fold consisting of β-sheets and α-helices. The key residues for calcium-dependent or independent binding of specific carbohydrates by CTL domains were predicted with homology modeling. Expression profiles assay showed distinct expression pattern of 14 CTLs: the expression and induction were related to the developmental stages and infected microorganisms. Overall, our work including the gene identification, sequence alignment, phylogenetic analysis, structural modeling, and expression profile assay would provide a valuable basis for the further functional studies of *O. furnacalis* CTLs.

Key words: C-type lectin, *Ostrinia furnacalis* (Gueneé), carbohydrate recognition domain, innate immunity

Unlike vertebrates, insects lack adaptive immune system and mainly depend on effective innate immune system to defend against the attack from bacteria, fungi and other pathogens or parasites (Kingsolver and Hardy 2012, Kanost and Jiang 2015). Insect innate immune response is induced by recognition of common components bearing on the microbial surfaces but absent from itself, known as pathogen-associated molecular patterns (PAMPs), by insect pattern recognition receptors (PRRs) (Steiner 2004, Takahashi et al. 2015). This type of recognition of non-self leads to the activation of intracellular signaling pathways to finally produce a battery of effector molecules for sequestering and eliminating the invading microorganisms (Basbous et al. 2011, Takahashi et al. 2015).

PRRs, soluble or membrane-bound, lack the binding specificity of antibodies and instead perform surveillance function by binding to polysaccharides, glycoproteins, and glycolipids which are common to the groups of microorganisms (Steiner 2004, Basbous et al. 2011). PRRs have been experimentally investigated in several insects. In *Manduca sexta*, hemolin, immulectin 1-4 (IML-1-4), peptidoglycan recognition protein 1 (PGRP1), β-1,3-glucanase-related protein 1-3 (βGRPs), microbe binding protein (MBP), and leureptin-1 were characterized as PRRs (Ladendorff and Kanost 1991, Ma and Kanost 2000, Yu et al. 2006, Zhu et al. 2010). In other arthropods, PGRPs, βGRPs, C-type lectins (CTLs), galectins, Leucine-rich repeat proteins (LRRPs), fibrinogen-related proteins (FREPs), hemocytins, Nimrods, scavenger receptors (SCRs), thioester proteins (TEPs), down syndrome cell adhesion molecule (DSCAM), Draper, and Eater were reported to function as PRRs (Christophides et al. 2004, Watson et al. 2005). Among these, CTLs comprise a large superfamily of proteins, and exist not only in insects, but in other invertebrates, vertebrates and plants (Zhang et al. 2013b, Dambuza and Gringhuis 2009).

CTLs, known as calcium-dependent lectins (or calcium-dependent carbohydrate binding proteins), are defined by the possession of at least one carbohydrate recognition domain (CRD) also called as CTL domain (Cambi et al. 2005, Geijtenbeek and Gringhuis 2009). Each CTL domain typically consists of 110–130 amino acid residues, and adopts a canonical fold composed of β-sheets, α-helices, and loops which is stabilized by two or three pairs of disulfide bonds (Weis et al. 1992, Zelensky and Gready 2005). CTL domain contains several conserved residues which are coordinated with a Ca²⁺ to form the basis of a primary sugar-binding site and, therefore, determine the sugar-binding specificity of CTL. CTLs with a
Glu-Pro-Asn (EPN) motif in the CTL domain are characteristic of mannose-binding and thus called mannose-type CTLs. Similarly, CTLs with a Gln-Pro-Asp (QPD) motif in the CTL domain are characteristic of galactose-binding and thus called galactose-type CTLs (Drickamer 1992, Zelensky and Greedy 2005). In human MinC (a macrophage inducible Ca2+-dependent CTL), Glu169, Asn171, Asn172, Asn193 and Asp194 are involved in calcium binding, Glu169 and Asn171 are also participated in the EPN motif (Furukawa et al. 2013). In rat mannose-binding lectin A, Glu185, Asn187, Glu193, Asn205, and Asp206 constitute the Ca2+-binding site 2, in which Glu185 and Asn187 belong to the EPN motif (Weis et al. 1992). Most CTLs only contain a single CTL domain. Some CTLs identified from lepidopteran insects, such as M. sexta IML-1, 4, Bombyx mori lipopolysaccharide-binding protein (LBP), MBP and LEL-1-3, and Helicoverpa armigera CTL-1-8, have two tandem CTL domains (Koizumi et al. 1999, Yu et al. 2006, Takase et al. 2009, Wang et al. 2012). CTLs with dual CTL domains are also reported in coleopteran insect Tribolium castaneum (Zou et al. 2007) and crustacean Fenneropenaeus chinensis (Xu et al. 2010). It is still unknown why dual CRDs arise in these CTLs.

As a category of important PRRs, CTLs play various roles through recognizing and binding to various types of carbohydrates. Mammalian CTLs have been revealed to function in anti-fungal/bacterial immunity, homeostasis, antoinnimmunity, allergy, the recognition of dead cells and tumors, and complemet activation (Brown 2015, Drickamer and Taylor 2015, and references within). Insect CTLs perform relatively less functions and mainly participate in innate immune responses including phagocytosis (Jomori and Natori 1992), nodule formation (Koizumi et al. 1999), encapsulation and activation of prophenoloxidase (Yu and Kanost 2004). For example, M. sexta IML-1 and IML-4 can induce the agglutination of bacteria and yeast in a Ca2+-dependent manner (Yu and Kanost 2001, Yu et al. 2006). Recombinant CRD2 of IML-2 directly binds to Caenorhabditis elegans and enhances its encapsulation and melanization in vivo (Yu and Kanost 2004). B. mori LBP binds to a variety of Gram-negative bacteria and MBP principally binds to Gram-positive bacteria and yeast. They both trigger hemocyte aggregation and microbe clearance (Koizumi et al. 1999, Takase et al. 2009). Recombinant Drosophila melanogaster DL2 and DL3 (Drosophila lectin) can accelerate the agglutination of Escherichia coli in the presence of calcium and enhance cellular encapsulation and melanization by directly recruiting hemocytes (Ao et al. 2007). However, current knowledge about the function and binding mechanism of insect CTLs is still limited and incomplete.

Recent genome-wide analysis has helped to identify a number of genes encoding proteins with one or more CTL or CTL-like domains. Aedes aegypti, Anopheles gambiae, D. melanogaster, B. mori, M. sexta, and T. castaneum have 39, 25, 34, 23, 34, and 17 such genes, respectively (Christophides et al. 2002, Waterhouse et al. 2007, Zou et al. 2007, Rao et al. 2015a, Rao et al. 2015b). Development of high throughput sequencing and de novo assembly strategies resulted in the identification of more CTL or CTL-like transcripts from the transcriptomes of non-model insects. For example, 8, 9, and 24 transcripts for potential CTLs were found in H. armigera, Nilaparvata lugens, and Tenebrio molitor, respectively (Wang et al. 2012, Bao et al. 2013, Zhu et al. 2013). In the previous work, we have obtained a transcriptome dataset from an important insect pest, Asian corn borer, Ostrinia furnacalis (Gueneé; Lepidoptera: Pyralidae), and identified 14 possible CTL transcripts (Liu et al. 2014). After carefully manual validation, we clarified that 3 of 14 transcripts lacked one or more key amino acid residues critical for CTL structure or binding activity. Meanwhile, we identified another three potential CTL transcripts. Therefore, we finally obtained 14 O. furnacalis CTLs, designated as OFCTL-S1 through S5, and OFML1 through 9. In this study, we reported the identification, characterization, expression analysis of these 14 CTLs. Their sequence features and gene phylogeny relationships were investigated thoroughly, and their developmental and induced expression profiles were examined.

### Materials and Methods

#### Biological Materials

Asian corn borers (O. furnacalis) were reared on an artificial diet at 28°C under a relative humidity of 70–90% and a photoperiod of 16:8 (L:D) h (Zhang et al. 2013a). Beauveria bassiana (strain 252) was cultivated on potato dextrose agar (PDA) plates at 25°C and a relative humidity of 80%.

#### Identification and Feature Prediction of O. furnacalis CTLs

A transcriptome dataset containing 62,382 unigenes were obtained from O. furnacalis larvae with ‘next generation’ high-throughput sequencing (Liu et al. 2014). All unigene sequences were searched against CTLs from M. sexta and other insects, using BLASTX algorithm with an E-value cut-off of 10^-5. The BLASTX results were used to extract coding region sequences from corresponding unigene sequences. The predicted coding region sequences were further translated into peptide sequences using the Translate tool provided by the Swiss Institute Bioinformatics. Analysis of deduced amino acids sequences, including prediction of signal peptide, molecular weight, and isoelectric point, were performed in the Expert Protein Analysis System (EXPASY) (http://www.expasy.org). Conserved domains and transmembrane regions were predicted in SMART (http://smart.embl-heidelberg.de/smart/set_mode.cgi). The domain architectures were plotted with IBS 1.0 (http://ibs.biocuckoo.org/).

#### Sequence Alignment and Phylogenetic Analysis

Multiple sequences alignment of CTLs from O. furnacalis and other insects (https://www.ncbi.nlm.nih.gov/) were carried out using MUSCLE module of MEGA 6.0 (http://www.megasoftware.net/) at the following parameters: refine alignment, gap opening penalty = −2.9, gap extension penalty = 0, hydrophobicity multiplier = 1.2, Maximum iterations = 100, clustering method (for iterations 1, 2) = UPGMB, and Minimum diagonal length = 24. The resulted alignments were used to construct the phylogenetic tree using the neighbor-joining method. For neighbor-joining method, gaps were treated as characters, and statistics analysis was performed by bootstrap trails with 1,000 replications, Poisson model, uniform rates, and complete deletion of gaps or missing data (Tamura et al. 2013).

#### Structure Modeling of 22 O. furnacalis CTL Domains

The deduced amino acid sequences of the 22 O. furnacalis CTL domains were submitted to the Iterative Threading ASSEMBly Refinement (I-TASSER) server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) for the prediction of tertiary structure. The representative structural template was identified from PDB by a locally installed LOMETS meta-threading server. Models were constructed by iterative TASSER simulations. The generated PDB files were then visualized with Pymol Molecular Graphics System.

#### Expression Profile Analysis

To investigate the transcriptional changes of CTLs during the various stages of O. furnacalis, total RNA samples were individually prepared from three different stages including egg, larva, and pupa using TRNzol Reagent (TIANGEN, Beijing, China). One µg of RNA equally from five individual RNA samples in each stage was treated with DNase I (TIANGEN, Beijing,
China) and converted into first-strand cDNA from an oligo (dT) primer following the instructions for QuantScriptRT Kit (TIANGEN, Beijing, China). The cDNA products independently from three biological replicates were diluted 10-fold for use as template in RT-PCR experiments. Specific primers were designed and listed in Supplementary Table S1. *O. furnacalis* ribosomal protein L8 (rpL8) was used as an internal standard to adjust the template amounts in a preliminary PCR experiments. The thermal cycling conditions were 94°C for 3 min, then 28 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s followed by incubation at 72°C for 10 min. The PCR products were separated by electrophoresis on a 1.5% agarose gel.

To check the expression profiles of *O. furnacalis* CTLs under different inducement conditions, day 0 fifth instar larvae from the same batch were injected into the hemocoele with 2 μl of sterile phosphate-buffered saline (PBS) containing formaline-killed *Escherichia coli* DH5α (5 x 10⁵ cells/μl), dried *Micrococcus luteus* (0.5 μg/μl), *B. bassiana* suspension (3 x 10⁶ conidial/μl), *B. bassiana* conidia suspension was prepared as described previously (Zhang et al. 2013a), or sterile PBS as a control. After 24 h, each three larvae from challenged or control group were collected, and total RNA samples were individually prepared as described above. DNase I–treated RNA (1 μg) was converted into first-strand cDNA using FastQuant RT Kit (TIANGEN Biotechnology Co., Ltd, Beijing, China). The cDNA products were diluted 10-fold for use as template. The quantitative real-time PCR (qRT-PCR) was performed with SYBR Premix EX Taq (TaKaRa, Japan) on Applied Biosystems 7500 Real-Time PCR System (Life Technologies, Grand Island, NY), according to the manufacturer’s instructions. *O. furnacalis* ribosomal protein L8 (rpL8) was used as an internal standard to normalize the expression level. The thermal cycle conditions for qRT-PCR were 95°C for 2 min followed by 40 cycles of 95°C for 15 s, 55°C for 30 s and 72°C for 30 s followed by incubation at 72°C for 10 min. The PCR products were separated by electrophoresis on a 1.5% agarose gel.

The ΔΔCT method (Livak and Schmittgen 2001) was used to estimate the relative expression levels of target genes. The 2−ΔΔCT was calculated with the 2−ΔΔCT (Livak and Schmittgen 2001).

### Results

#### Overview and General Properties of *O. furnacalis* CTLs

With next generation high-throughput sequencing, we obtained an *O. furnacalis* transcriptome dataset containing 62,382 unigenes. Among these, 14 ones were predicted to encode potential CTLs (Liu et al. 2014). After careful validation, we found three of them, unigene2023, unigene31526 and CL321.contig3, were possibly not real CTLs. Meanwhile, we identified another three transcripts, unigene4110, CL4301.contig1 and CL321.contig1, as potential CTLs. Thus, we still obtained totally 14 CTL transcripts, which encode proteins with single or dual CTL domains (Table 1). Five of the protein products (OfCTL-S1 ~ S5) only contain one CTL domain, and the remaining nine (OfIML-1 ~ 9) belong to the IML family with two tandem CTL domains (Fig. 1). None has other conserved structural units such as complement C1r/C1s (CUB) domain, complement control protein (CCP/Sushi) domain, epidermal growth factor (EGF)-like domain, and coagulation factor 5/8 C-terminal (FA58C) domain, which were reported in *B. mori* CTL-X1 ~ X5 and *M. sexta* CTL-X1 ~ X6 (Rao et al. 2015a, Rao et al. 2015b). *O. furnacalis* CTL-S1 ~ S5, IML-1 ~ 4, IML-7, and IML-8 have a secretion signal peptide at the amino-terminus and are likely secreted into plasma. *O. furnacalis* IML-5 and IML-6 lack the N-terminal secretion signal and may be located in the cytoplasm. It is unsure whether OfIML9 has a signal peptide because the translation start codon preceding the CTL domain is unavailable in its incomplete transcript (Table 1 and Fig. 1).

### Table 1. Structural features of 14 *O. furnacalis* CTLs

| Designated name | Unigene ID | Protein length (aa) | Signal peptide | Mr (kDa) | pI | Glycosylation sites | CRD numbers | Motif<sup>a</sup> |
|-----------------|------------|---------------------|----------------|---------|----|-------------------|-------------|----------------|
| OfCTL-S1        | CL4786.Contig1 | 220                | 1-19A<sup>b</sup>Q | 23.4 | 6.33 | 0 3 | 1 | QPD           |
| OfCTL-S2        | Unigene6545 | 223                | 1-19A<sup>b</sup>Q | 23.7 | 5.99 | 0 1 | 1 | QPD           |
| OfCTL-S3        | CL8286.Contig1 | 221               | 1-21A<sup>b</sup>Q | 23.1 | 8.47 | 0 6 | 1 | QPD           |
| OfCTL-S4        | Unigene9847 | 207                | 1-21A<sup>b</sup>Q | 21.3 | 6.20 | 1 5 | 1 | APQ           |
| OfCTL-S5        | CL321.Contig1 | 184                | 1-21T<sup>b</sup>V | 19.1 | 4.90 | 0 5 | 1 | EPN           |
| OfIML-1         | Unigene22572 | 301                | 1-19A<sup>b</sup>Q | 32.6 | 5.32 | 1 3 | 2 | APW / EPN    |
| OfIML-2         | Unigene14484 | 314                | 1-22G<sup>b</sup>R | 32.9 | 5.42 | 0 2 | 2 | EPD / EPN    |
| OfIML-3         | CL4301.Contig2 | 304               | 1-20S<sup>b</sup>N | 32.5 | 4.74 | 1 1 | 2 | VPL / EPN    |
| OfIML-4         | CL106.Contig3 | 328                | 1-21T<sup>b</sup>D | 33.9 | 5.13 | 1 2 | 2 | EDP / EDP    |
| OfIML-5         | Unigene7501 | 322                |                | 35.7 | 4.89 | N/A N/A | 2 | EDP / EDP    |
| OfIML-6         | Unigene2411 | 319                |                | 36.5 | 5.89 | N/A N/A | 2 | DIS / SPD    |
| OfIML-7         | CL1725.Contig1 | 307              | 1-20S<sup>b</sup>S | 32.3 | 6.20 | 3 3 | 2 | DVS / VPD    |
| OfIML-8         | Unigene5307 | 321                | 1-21S<sup>b</sup>Q | 34.1 | 4.98 | 2 0 | 2 | EPN / VPD    |
| OfIML-9         | Unigene4110 | >205               | N/A             | >23.2 | 4.44 | 2 2 | 1 | QPD           |

<sup>a</sup>The molecular weight (Mr) and isoelectric point (pI) are for mature proteins without signal peptides.

<sup>b</sup>OfIML-5 and OfIML-6 lack signal peptide, and, therefore, the predication for the numbers of glycosylation sites is not applicable (N/A).

<sup>c</sup>The canonical motifs (QPD and EPN) are in bold.
B. mori CTL-S4 ~ S6, -S8 ~ S10, and M. sexta CTL-S3 ~ S5 and -S8 (Rao et al. 2015b, Rao, 2015a), is observed in all 14 identified O. furnacalis CTLs. Numbers of putative N-linked and O-linked glycosylation sites are variant in each O. furnacalis CTL, from 1 in OfCTL-S2 to 6 in OfCTL-S3 and OfCTL-S4 (Table 1).

Conserved Residues and Structural Features of O. furnacalis CTLs

All CTLs contain several conserved amino acid residues which determine the sugar-binding specificity of CTL. Gln-Pro-Asp (QPD) and Glu-Pro-Asn (EPN) motif in the CTL domain suggest potential galactose and mannose binding specificities of CTLs, respectively (Drickamer 1992, Zelensky and Gready 2005). The QPD motif is found in the CTL domain in OfCTL-S1 through S3 and the second CTL domain in OfIML-4, OfIML-5, and OfIML-9 (Table 1, Fig. 2). Therefore, they may be galactose-type lectins. The EPN motif is found in the CTL domain in OfIML-1 through 3 (Table 1, Fig. 2). These O. furnacalis CTLs may be mannose-type lectins. OfIML-8 contains both motifs simultaneously, with the EPN motif in the first CTL domain and the QPD motif in the second CTL domain. However, other CTL domains, including OfCTL-S4, the first CTL domain in OfIML-1 through 7 and the second CTL domain in OfIML-6 and OfIML-7, contain the non-canonical motifs. Their binding specificities need to be further investigated.

The typical CTL domain is stabilized by three pairs of disulfide bonds between Cys-1 and Cys-2, Cys-3 and Cys-6, Cys-4 and Cys-5, although no all CTL domains contain all six cysteine residues (Weis et al. 1992, Furukawa et al. 2013). We aligned all 22 O. furnacalis CTL domain sequences including 5 ones from OfCTL-S1 through S5, 8 ones from the first CTL domain in OfCTL-1 through 8 (OfCTL-1A ~ -8A, the first CTL domain in OfCTL-9 is currently unavailable), and 9 ones from the second CTL domain in OfCTL-1 through 9 (OfCTL-1B ~ -9B) with well-studied human Mincle (PDB: 3WHD). These six Cysteine residues mentioned above are absolutely conserved in nine CTL domains (OfIML-1B through 9B). Cys-1 is present but Cys-2 is absent in the CTL domains in OfCTL-S1,
S2, -S5 and OfIML-2A. Both Cys-1 and Cys-2 are missing in the remaining CTL domains including OfCTL-S3, -S4, OfIML-1A, -3A through 8A (Fig. 2). Additionally, the cysteine residues in C(D/N)F(K/A)GC (shaded in grey in Fig. 2) of OfCTL-S1 through S3 may form a unique disulfide linkage.

CTLs usually recognize carbohydrates via a Ca$^2+$-mediated binding network (Zelensky and Gready 2005). In a canonical CTL, the EPN or QPD motifs on the loops of CTL cooperate with the Asn-Asp (ND) residues on the adjacent $\beta$-sheet to bind the calcium (Furukawa et al. 2013). For example, E169, N171, N172, N193, and D194 in human Mincle constitute the Ca$^2+$-bound sites (marked by red asterisks in Fig. 2) (Furukawa et al. 2013). The alignment of 22 O. furnacalis CTL domain with human Mincle indicated that CTL-Ss with the exception of CTL-S4 and most of the IML A domains lack these residues corresponding to those in Mincle whereas IML B domains other than OfIML-7B possess them (Fig. 2). We further performed structural modeling of 22 O. furnacalis CTL domains to explore the putative calcium and sugar binding sites. The overall structures of all 22 O. furnacalis CTL domains are predicted to be closely similar (Fig. 3). They all may bind to carbohydrates in a calcium-dependent manner since five of the models (-S5, -2B, -3B, -4A, and -8A) contain two Ca$^2+$ and the remaining 17 CTL domains contain one Ca$^2+$ (Table 2, Figs. 2 and 3). Further experiments are required to determine whether Ca$^2+$ indeed enhances the binding.

![Fig. 2. Sequence alignment of the 22 O. furnacalis CTL domains. Based on the domain predications with SMART, the amino acid sequences for 22 O. furnacalis CTL domains are aligned with that of human Mincle (PDB: 3WHD). The conserved Cys residues are numbered and shown in red. They are predicted to form three disulfide bonds (1–2, 3–6, 4–5) based on the determined structure in human Mincle. The regions C/D/NF/K/A/GC shaded in grey. The canonical QPD and EPN motifs are shaded in green and cyan, respectively. Residues involved in calcium binding and carbohydrate binding are in blue and purple, respectively. Residues involved in both Ca$^2+$ and sugar binding are in bold. The secondary structure elements ($\alpha$, $\alpha$-helix; $\beta$, $\beta$-strand; T, turn) of human Mincle are shown above or below the sequences. The asterisks above the sequences indicate the residues involved in calcium binding in human Mincle.](https://academic.oup.com/jinsectscience/article-abstract/18/2/24/4924850 by guest on 05 May 2019)
strength. On the other hand, three-dimensional structure modeling also allows us to predict the sugar binding specificities of 22 O. furnacalis CTL domains. The results from the above sequence alignment and molecular modeling are mostly consistent (Table 2, Figs. 2 and 3). The high C-scores of the models suggest that OfCTL-S1, -S3, IML-4B, and -5B may form stable complexes with galactose. They just contain the QPD motif. OfCTL-S5, IML-1B, -2B, -3B, and -8A has the EPN motif, and may bind to mannose tightly with C-scores ranging from 0.64 to 0.74 (Table 2). The exceptions are CTL-S2, IML-8B and -9B contain the QPD motif, but are predicted to bind galactose (Table 2 and Fig. 3). OfCTL-S4, IML-1A through -7A, -6B, and -7B have the non-canonical motifs, but also show the sugar binding potentials in structural modeling. More experiments will be required to validate the binding predictions. Beside, homology modeling also indicated that two amino acid residues (Glu116 in CTL-S5, Glu115 and Asp128 in IML-2B, Glu115 and Asp128 in IML-3B, Glu106

**Table 2. Structural features of the 22 O. furnacalis CTL domain models**

| Domain | Ca2+ | Putative Ca2+ coordinators | Motif | Sugar* | Putative sugar binding residues | Template | C-scoreb |
|--------|------|-----------------------------|-------|--------|---------------------------------|---------|---------|
| OfCTL-S1 | 1 | [48,50,54,168] | QPD  | GQ2 | 69,125,127,128,142,155,156,157,163 | 4C9F | 0.52 |
| -S2 | 1 | [47,49,53,172] | QPD  | MMA | 129,131,132,142,159,160,161 | 1KWU | 0.53 |
| -S3 | 1 | [40,42,46,159] | QPD  | GQ2 | 61,117,119,120,133,146,147,148,154 | 4CF9 | 0.58 |
| -S4 | 1 | [43,45,49,167] | APQ  | PI  | 100,101,111,111,112,122,123,124,125,126 | 2ORK | 0.25 |
| -S5 | 2 | [49,51,55,145] | EPN  | TRE | 72,104,106,116,122,132,133,138 | 4ZRV | 0.65 |
| OfIML-1A | 1 | [40,42,61,121] | APW  | GAL | 69,125,127,128,142,155,156,157,163 | 4C9F | 0.59 |
| -1B | 1 | [43,45,49,142] | EPN  | TRE | 71,108,110,112,116,122,129,130,135 | 4KZV | 0.72 |
| OfIML-2A | 1 | [38,40,44,126] | EPN  | TRE | 71,108,110,112,116,122,129,130,135 | 4KZV | 0.72 |
| -2B | 2 | [43,45,49,140] | EPN  | TRE | 73,107,109,111,115,121,127,128,133 | 4KZV | 0.64 |
| OfIML-3A | 1 | [40,42,61,122] | VPL  | BM3 | 89,91,93,97,99,109,111 | 2OJR | 0.51 |
| -3B | 2 | [43,45,49,141] | EPN  | TRE | 72,107,109,111,126,129,134 | 4ZRV | 0.69 |
| OfIML-4A | 1 | [40,42,61,121] | EPN  | TRE | 72,107,109,111,126,129,134 | 4ZRV | 0.66 |
| -4B | 1 | [43,45,49,143] | EPN  | TRE | 72,106,108,110,121,129,130,131,134 | 4KZV | 0.64 |
| OfIML-5A | 1 | [45,47,51,131] | EPN  | TRE | 72,106,108,110,121,129,130,131,134 | 4ZRV | 0.66 |
| -5B | 1 | [43,45,49,145] | EPN  | TRE | 72,106,108,110,121,129,130,131,134 | 4ZRV | 0.69 |
| OfIML-6A | 1 | [40,42,61,122] | DIS  | MAN | 64,98,100,106,118,119,121 | 2OJR | 0.44 |
| -6B | 1 | [42,44,48,137] | SPN  | TRE | 71,105,107,111,117,123,124,129 | 4ZRV | 0.59 |
| OfIML-7A | 1 | [45,47,51,126] | DVS  | NGA | 96,98,101,105,113,114,115 | 1BCH | 0.34 |
| -7B | 1 | [42,44,48,139] | VPD  | MBG | 26,64,65,96,98,101,105,115,119,121 | 4CF9 | 0.25 |
| OfIML-8A | 1 | [40,42,61,128] | VPD  | MBG | 105,107,109,113,125,126 | 1AFA | 0.34 |
| -8B | 1 | [43,45,49,137] | VPD  | MBG | 105,107,109,113,125,126 | 2OJR | 0.74 |
| OfIML-9B | 1 | [43,45,49,136] | VPD  | MBG | 71,105,107,112,118,124,125,130 | 4ZRV | 0.66 |

EPN and QPD motifs are in bold. The residues involved in both calcium and sugar binding are in bold.

*GQ2, a viral ligand to human herpesvirus 6 (HHV-6); MMA, methyl α-D-mannopyranoside; PI, phosphatidylinositol; TRE, trehalose dimycolate; GAL, β-D-galactose; MAN, α-D-mannose; BM3, N-acetyl-α-D-mannosamine; NGA, N-acetyl-D-galactosamine; MBG, methyl-β-galactose.

bC-score is the confidence score representing the quality of the generated models. It is calculated based on the Z-score of LOMETS threading alignments and the convergence of I-TASSER simulations. C-score ranges 0–1, where a higher score indicates a more reliable prediction. Ca2+ sites with C-score > 0 and sugar sites with C-score > 0.2 are shown in the table. C-scores shown here are for putative sugar binding sites.

Fig. 3. Structural models of O. furnacalis CTL-S1 (A), IML-8A (B), and IML-8B (C). The tertiary structures predicted from I-TASSER server are exhibited as cartoons. The calcium ions are shown as red spheres and indicated by red arrows. The mannose-type EPN motif and galactose-type QPD motif are shown as blue sticks and indicated by blue arrows. The Cys residues for the formation of disulfide bonds are represented as orange sticks.
in IML-4A, Glu$^{106}$ in IML-5A, and Glu$^{103}$ in IML-8A) are involved in both calcium and carbohydrate binding (Table 2 and Fig. 3).

**Phylogenetic Analysis of *O. furnacalis* CTLs**

To investigate the evolutionary relationships between CTLs of *O. furnacalis* and other insects, we retrieved 77 CTL protein sequences from 19 insect species by BLASTP searches in GenBank, VectorBase, and Insect Inmate Immunity Database. Since no CTLs with other complicated structural unites than CTL domains are currently identified in *O. furnacalis*, we did not include *M. sexta* CTL-Xs and their orthologs, which contain CTL domain(s) together with other conserved domains in the alignment. The phylogenetic tree (Fig. 4) shows that all CTLs are generally divided into two categories based on the numbers of CTL domains. OfCTL-S1 through S5 formed tight, monophyletic groups with their respective orthologs from the other species in Lepidoptera and Diptera. Moreover, the orthologs of OfCTL-S1 and OfCTL-S4 were also found in the red flour beetle *T. castaneum* (Coleoptera) (Fig. 4).

Nine *O. furnacalis* IMLs are grouped with the CTLs with dual CTL domains, with a bootstrap value of high to 99 (Fig. 4). All the orthologs in this group are from lepidopteran insects including *B. mori*, *M. sexta*, *H. armigera*, *Hyphantria cunea*, *Danaus plexippus*, *Danaus plexippus*, and *Pieris rapae*. Unlike MsIML-1, 18, 19 and BmIML-1, 4–6 which had 1:1 orthologs, no *O. furnacalis* IML formed 1:1 orthologous group. OfIML-4 and OfIML-5, as well as OfIML-6 and OfIML-7, showed lineage-specific expansions (Fig. 4). They may evolve by multiple gene duplication. Additionally, we constructed the phylogenetic tree based on the alignment of *O. furnacalis* CTL domains alone, instead of the alignment of entire protein sequences from 19 insect species by BLASTP searches in GenBank, VectorBase, and Insect Inmate Immunity Database. Since no CTLs exist in *O. furnacalis* currently identified in other insects, we did not included the orthologs in different insects, some immunection genes show extensive lineage-specific expansions. For instance, MsIML-3, 4, 6–8, IML-9–12, and IML-13–17 and *H. armigera* IML-1, 2, and 6 are four groups of related IMLs (Fig. 4) (Wang et al. 2012, Rao et al. 2015a, Rao et al. 2015b). Unlike CTL-Ss which normally have clear orthologs in different insects, some immunection genes show extensive lineage-specific expansions. For instance, *M. sexta* IML-3, 4, 6–8, IML-9–12, and IML-13–17 and *H. armigera* IML-1, 2, and 6 are four groups of related IMLs (Fig. 4) (Wang et al. 2012, Rao et al. 2015a, Rao et al. 2015b). As the formation of IMLs with dual CTL domains, it may be a result of the duplication of CTL-S genes because the first CTL domains in *B. mori* and *M. sexta* IMLs seem to evolve from common ancestors of BmCTL-S1 and MsCTL-S7, respectively (Rao et al. 2015a, Rao et al. 2015b). We speculated that the event of entire gene duplications and sequence divergence resulting in the appearance of tandem CTL domains in moths and butterflies took place before the radiation of Lepidoptera. In addition, there exists another type of IMLs containing one or two CTL domains along with other structural units, such as *D. melanogaster* furrowed, uninflatable, and contactin (Stork et al. 2008, Xie et al. 2012, Chin and Mlodzik 2013), *B. mori* CTL-X1–5 (Rao et al. 2015b), *M. sexta* CTL-X1–6 (Rao et al. 2015a), and *T. castaneum* CTL-X1–5 (Zou et al. 2007). However, we did not identify any similar IMLs in *O. furnacalis* transcriptome. The possible reason is that the assembly quality of current transcriptome is not high enough to find the long extensions for encoding the other structural modules. We believed that there also existed similar complex CTL-Xs in *O. furnacalis*, but the identification requires more sequence resources.

The CTL domain adopts a characteristic double-loop fold consisting of five $\beta$-sheets and two $\alpha$-helices: the closely located N- and C-termini ($\beta1$, $\beta5$) make the overall domain a loop, and

**Expression Profiles of *O. furnacalis* CTL Genes**

We analyzed the mRNA levels of *O. furnacalis* CTLs in various development stages using semi-quantitative RT-PCR methods. As shown in Fig. 5A, the 14 *O. furnacalis* CTLs exhibited distinct expression patterns. The transcripts of OfCTL-S5, IML-4, IML-5, and IML-8 were detected in all examined developmental stages, including egg, larval and pupal stages. OfCTL-S1 was also detected in all stages, but with obviously higher level in the stage of egg and young larva and low level in the stage of the fifth instar larva and pupa. OfCTL-S2 was clearly expressed in all tested developmental stages other than egg. IML-6, -7 and -9 were expressed in the second to fifth instar larval stage and pupal stage, but not in eggs and the first instar larva. CTL-S3 mRNA was detectable in the old larva and pupae, and the mRNA level was lower in the pupal stage. The remaining three IMLs exhibited irregular expression patterns: OfCTL-S4 was detected in egg and pupa; IML-2 in egg and the fifth instar larva; and IML-3 in egg and the fourth instar larva (Fig. 5A).

To check the *O. furnacalis* CTL expression profiles after exposure to microbial elicitors, we randomly selected eight CTLs and analyzed their transcript level after *O. furnacalis* larvae were injected with *E. coli*, *M. luteus*, *B. bassiana*, or PBS as a control. The result from quantitative RT (qRT)-PCR assay indicated that the expression patterns of these eight CTLs could be divided into three groups: the transcript levels of IML-8, CTL-S2, and CTL-S5 remained unchanged; CTL-S4 decreased significantly upon challenge; and CTL-S1, IML-2, IML-3, and IML-9 increased after infection (Fig. 5B). There still existed some differences although CTL-S1, IML-2, IML-3, and IML-9 mRNA levels all increased upon challenge: *E. coli*-injection only caused the significant increase for IML-9; *M. luteus*-injection led to the significant up-regulation for both CTL-S1 and IML-2; *B. bassiana*-injection resulted in the significant increase for all four transcripts other than CTL-S1 (Fig. 5B).

**Discussion**

CTLs act as a category of important PRRs, and play various physiological roles in animals including human and insects (Brown 2015, Dambuzo and Brown 2015). In this study, we have identified 14 CTLs (OfCTL-S1 – S5, and OfIML-1 – 9) from *O. furnacalis* transcriptome. We performed comprehensive sequence analysis for these 14 CTLs and investigated their expression profiles in different developmental stages and upon different infections. CTLs are characterized by processing one or more carbohydrate-recognition domains (also known as CTL domains) (Cambi et al. 2005, Geijtenbeek and Gringhuis 2009). Among 14 identified *O. furnacalis* CTLs, 5 ones only has a single such domain, and 9 contain two tandem CTL domains (Table 1 and Fig.1). So far, CTLs with one CTL domain have been identified from almost all animals including human (Furukawa et al. 2013), rats (Weiss et al. 1992), crustaceans (Zhang et al. 2013b) and insects (Rao et al. 2015a, Rao et al. 2015b). For example, *D. melanogaster*, *A. gambiae*, *B. mori*, *M. sexta*, and *T. castaneum* have 34, 25, 12, 9, and 10 genes encoding proteins with one CTL domain, respectively (Christophides et al. 2002, Waterhouse et al. 2007, Zou et al. 2007, Rao et al. 2015a, Rao et al. 2015b). We only identified five transcripts for encoding proteins with one CTL domain. Therefore, it is highly possible there are other unidentified transcripts to encode *O. furnacalis* CTL-Ss. At least, there should be one CTL-S for ortholog in CTL-S5 group since other insects than *O. furnacalis* have clearly 1:1 orthologous relationship in this group (Fig. 4). Compared to the wide distribution of short CTL-Ss, IML family with two tandem CTL domains was found mainly in lepidopteran insects (Yu and Kanost 2001). Six, 19, and eight such CTLs have been reported in *B. mori*, *M. sexta*, and *H. armigera*, respectively (Wang et al. 2012, Rao et al. 2015a, Rao et al. 2015b). Unlike CTL-Ss which normally have clear orthologs in different insects, some immunection genes show extensive lineage-specific expansions. For instance, IML-3, 4, 6–8, IML-9–12, and IML-13–17 and *H. armigera* IML-1, 2, and 6 are four groups of related IMLs (Fig. 4) (Wang et al. 2012, Rao et al. 2015a, Rao et al. 2015b). As the formation of IMLs with dual CTL domains, it may be a result of the duplication of CTL-S genes because the first CTL domains in *B. mori* and *M. sexta* IMLs seem to evolve from common ancestors of BmCTL-S1 and MsCTL-S7, respectively (Rao et al. 2015a, Rao et al. 2015b). It is speculated that the event of entire gene duplications and sequence divergence resulting in the appearance of tandem CTL domains in moths and butterflies took place before the radiation of Lepidoptera. In addition, there exists another type of IMLs containing one or two CTL domains along with other structural units, such as *D. melanogaster* furrowed, uninflatable, and contactin (Stork et al. 2008, Xie et al. 2012, Chin and Mlodzik 2013), *B. mori* CTL-X1–5 (Rao et al. 2015b), *M. sexta* CTL-X1–6 (Rao et al. 2015a), and *T. castaneum* CTL-X1–5 (Zou et al. 2007). However, we did not identify any similar IMLs in *O. furnacalis* transcriptome. The possible reason is that the assembly quality of current transcriptome is not high enough to find the long extensions for encoding the other structural modules. We believed that there also existed similar complex CTL-Xs in *O. furnacalis*, but the identification requires more sequence resources.
the region between β2 and β3 strands forms the second loop also called the long loop region (Zelensky and Gready 2005, Furukawa et al. 2013, Feinberg et al. 2016). Canonical CTL domain has the long loop region and the compact one lacks it. All identified 22 O. furnacalis CTL domains possess the long loop region. It suggests that they all belong to canonical CTLs. On the other side, the long loop region is reported to be involved in Ca2+-dependent carbohydrate binding and in domain-swapping dimerization of some CTLs (Liu and Eisenberg 2002, Zelensky and Gready 2005). Structural modeling also indicated that 22 CTL domains with long loop region had potential to bind one or two calcium ions, suggesting they might function in a calcium-mediated manner. Ca2+-binding site 1 (Ca-1) close to the N- and C-terminus is present in all 22 CTL domains (Fig. 3). It is consistent with the previous report that Ca-1 is a highly conserved structural feature in virtually all CTL domains (Zelensky and Gready 2003, 2005). However, only 6 out of 22 CTL domains (CTL-S5, IML-2B, -3B, -4A, -5A and -8A) were predicted to have Ca2+-binding site 2 (Ca-2), and no O. furnacalis CTL domain has other potential Ca2+-binding sites (Ca-3 or Ca-4) observed in human macrophage C-type lectin (MCL) and Mincle (Furukawa et al. 2013). The region surrounding Ca-2 in O. furnacalis CTL domains is more similar to that in human DC_SIGNR than to that in MCL and Mincle. The roles of calcium-binding sites are various in individual CTL, either involved in carbohydrate binding or just playing structural roles. For example, the calcium-binding site in human tetranectin is known not to bind carbohydrates, but interact with kringle domain-containing proteins leading to changes in the long loop region conformation (Nielbo et al. 2004). The functional impact of calcium-binding sites in O. furnacalis CTL domains is still unclear.

Relative to calcium-binding sites, the characteristic motif, such as mannosetype EPN motif or galactose-type QPD motif, is possibly more important for carbohydrate binding of CTLs. For example, the EPN motif in rat MBP-A (E185, P186, N187) provides carbonyl

Fig. 4. Phylogenetic analysis of insect CTLs. The amino acid sequences of 91 CTLs from 20 insect species were examined. Lepidopteran-specific CTLs were indicated with bracket. The branches specific for CTL-S1 through -S6 are shaded. Numbers at the nodes were bootstrap values as percentage. Only bootstrap values greater than 70 were shown. Of, O. furnacalis; Aa, A. aegypti; Ag, A. gambiae; Ap, Antheraea pernyi; Bm, B. mori; Bt, Bombus terrestris; Cq, Culex quinquefasciatus; Dm, D. melanogaster; Dp, D. plexippus; Ha, H. armigera; Hc, H. cunea; Hv, Heliothis virescens; Mr, Megachile rotundata; Ms, M. sexta; Nv, Nasonia vitripennis; Pr, P. rapae; Ppy, Papilio polytes; Pxy, Papilio xuthus; Pwy, Plutella xylostella; and Tc, T. castaneum.
Expression profile analysis of *O. furnacalis* CTLs. (A) Expression profiles of *O. furnacalis* CTLs at different stages of development. RNA was extracted from the whole bodies collected from eggs, first-instar (L1), second-instar (L2), third-instar (L3), fourth-instar (L4), fifth-instar (L5) larvae, and pupae. The *rpL8* was used as an internal control. (B) Expression profiles of *O. furnacalis* CTLs upon microbial challenge. Day 0, fifth instar larvae were infected with PBS, *E. coli*, *M. luteus*, or *B. bassiana*. RNA was prepared from the whole bodies 24 h after injection. qRT-PCR was used to assess the transcript change of *OfCTLs* with *rpL8* as an internal standard to indicate a consistent total mRNA amount. The bars represent mean ± SD (n = 3). Bars labeled with different letters are significantly different (one-way ANOVA, followed by the Newman-Keus test, P < 0.05).

Replacing the EPN sequence in MBP-A with a QPD sequence was enough to switch the specificity to galactose (Drickamer 1992). Twelve out of 22 *O. furnacalis* CTL domains contain canonical EPN or QPD motif (Tables 1 and 2), suggesting these *O. furnacalis* CTLs have the potential to binding carbohydrates. Additionally, it is worthy of note that the remaining 10 CTL domains contain unusual motifs in the long loop region such as EPD, APQ motif (Tables 1 and 2). Human MCL also has an EPD motif, and it still shows the binding affinity to trehalose dimycolate (Furukawa et al. 2013). Sarcophaga lectin from the flesh fly *Sarcophaga peregrina* contains an NPD motif, and exhibits the galactose-binding property (Kawaguchi et al. 1991). It is more surprising that blood dendritic cell antigen 2 (BDCA-2) contains an EPN motif, but bind galactose-terminated glycans instead of predicted mannose (Jegouzo et al. 2015). Langerin contains an EPN motif, but bind both mannose and galactose (Feinberg et al. 2011). Therefore, we infer that there exist non-canonical sugar-binding sites in CTLs. Another possibility is that contacts outside of the primary sugar-binding site in extended or secondary binding sites also contribute to the sugar binding of CTLs. Further experiments are undergoing to determine the binding specificities of *O. furnacalis* CTLs.

The developmental and induction profiles of the CTL genes reflect their functional importance to some extent (Fig. 5). For instance, *CTL-S4* transcripts were detected in eggs and pupae, but not in the larvae, suggesting that it may be involved in the development. Its mRNA levels surprisingly decreased after immune challenge (Fig. 5). The orthologous *B. mori* *CTL-S5* also showed down-regulation in the *E. coli* or *S. aureus* injected fat bodies (Rao et al. 2015b). However, the function of *B. mori* *CTL-S5* is currently unknown. On the other side, immulections have been reported to participate in the immune responses including agglutination and/or encapsulation of microorganisms (Zelensky and Gready 2005). The expression of most IMLs is stimulated after microbial challenge (Yu and Kanost 2004, Yu et al. 2006). *O. furnacalis* IML-2, IML-3, and IML-9 were consistently up-regulated after infection from microorganism, but IML-8 remained unchanged upon infection (Fig. 5). So far, we only investigated the preliminary function of *O. furnacalis* IML-2. It exhibited the promising potential to bind and agglutinate *E. coli* (data not shown). The functional studies for other *O. furnacalis* CTLs are undergoing.
Acknowledgments
This work was supported by National Natural Science Foundation of China (31672361) and National Pearl-Industry Technology Research System (CARS-29). We thank Dr. Kanglai He, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, for kindly providing O. furnacalis eggs, and Dr. Weiguo Fang from Zhejiang University for generously providing B. bassiana strain.

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

References Cited
Ao, J., E. Ling, and X. Q. Yu. 2007. *Drosophila* C-type lectins enhance cellular encapsulation. Mol. Immunol. 44: 2541–2548.

Bao, Y. Y., L. Y. Qu, D. Zhao, L. B. Chen, H. Y. Jin, L. M. Xu, J. A. Cheng, and C. X. Zhang. 2013. The genome- and transcriptome-wide analysis of innate immunity in the brown planthopper, *Nilaparvata lugens*. BMC Genomics. 14: 160.

Babous, N., F. Coste, P. Leone, R. Vincentelli, J. Royet, C. Kellenberger, and A. Roussell. 2011. *Drosophila* peptidoglycan-recognition protein LF interacts with peptidoglycan-recognition protein LC to downregulate the IMD pathway. EMBO Rep. 12: 327–333.

Brown, G. D. 2015. C-type lectins in innate antifungal immunity: a key to the therapeautic future? MYCOSES. 58:3: 1.

Cambi, A., M. Koopman, and C. G. Figdor. 2005. How C-type lectins detect pathogens. Cell. Microbiol. 7: 481–488.

Chin, M. L., and M. Mlodzik. 2013. The *Drosophila* selectin furrowed mediates intercellular planar cell polarity interactions via frizzled stabilization. Dev. Cell. 26: 455–468.

Christophides, G. K., E. Zdobnov, C. Barillas-Mury, E. Birney, S. Blandin, G. Dimopoulos, et al. 2002. Immunity-related genes and gene families in *Anaoplophora glycinae*. Science. 298: 159–165.

Christophides, G. K., D. Vlachou, and F. C. Kafatos. 2004. Comparative and functional genomics of the immune system in the malaria vector *Anaoplophora glycinae*. Immunol. Rev. 198: 127–148.

Dambuza, I. M., and G. D. Brown. 2015. C-type lectins in immunity: recent developments. Curr. Opin. Immunol. 32: 21–27.

Drickamer, K. 1992. Engineering galactose-binding activity into a C-type mannos-binding protein. Nature. 360: 183–186.

Drickamer, K., and M. E. Taylor. 2015. Recent insights into structures and functions of C-type lectins in the immune system. Curr. Opin. Struct. Biol. 34: 26–34.

Feinberg, H. M., E. Taylor, N. Razi, R. McBrde, Y. A. Knirel, S. A. Graham, K. Drickamer, and W. I. Weis. 2011. Structural basis for length receptor recognition of diverse pathogen and mammalian glycans through a single binding site. J. Mol. Biol. 405: 1027–1039.

Feinberg, H. N. D, Ramburuth, S. A., Jegouzo, K. M., Jacobson, R. Djurhuus, T. B. Poulsen, W. I. Weis, M. E. Taylor, and K. Drickamer. 2016. Binding sites for acylated trehalose analogs of glycolipid ligands on an extended carbohydrate recognition domain of the macrophage receptor mincle. J. Biol. Chem. 291: 21222–21233.

Furukawa, A., J. Kamishikiryo, D. Mori, K. Toyonaga, Y. Okabe, A. Toji, R. Kanda, Y. Miyake, T. Ooe, S. Yamasaki, et al. 2013. Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL. Proc. Natl. Acad. Sci. U. S. A. 110: 17438–17443.

Geijtenbeek, T. B. H., and S. I. Gringhuis. 2009. Signalling through C-type lectin receptors: shaping immune responses. Nat. Rev. Immunol. 9: 463–479.

Jegouzo, S. A., H. Feinberg, T. Dungarwalla, K. Drickamer, W. I. Weis, and M. E. Taylor. 2015. A novel mechanism for binding of galactose-terminated glycans by the C-type carbohydrate recognition domain in blood dendritic cell antigen 2. J. Biol. Chem. 290: 16759–16771.

Jomori, T., and S. Natori. 1992. Function of the lipopolysaccharide-binding protein of *Periplaneta americana* as an opsonin. FERS Lett. 296: 283–286.
Weis, W. I., K. Drickamer, and W. A. Hendrickson. 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. Nature. 360: 127–134.

Xie, G., H. Zhang, G. Du, Q. Huang, X. Liang, J. Ma, and R. Jiao. 2012. Uif, a large transmembrane protein with EGF-like repeats, can antagonize Notch signaling in Drosophila. Plos One. 7: e36362.

Xu, W. T., X. W. Wang, X. W. Zhang, X. F. Zhao, X. Q. Yu, and J. X. Wang. 2010. A new C-type lectin (FcLec3) from the Chinese white shrimp Fenneropenaeus chinensis. Amino Acids. 39: 1227–1239.

Yu, X. Q., and M. R. Kanost. 2001. A family of C-type lectins in Manduca sexta. Adv. Exp. Med. Biol. 484: 191–194.

Yu, X. Q., and M. R. Kanost. 2004. Immulecin-2, a pattern recognition receptor that stimulates hemocyte encapsulation and melanization in the tobacco hornworm, Manduca sexta. Dev. Comp. Immunol. 28: 891–900.

Yu, X. Q., E. Ling, M. E. Tracy, and Y. Zhu. 2006. Immulecin-4 from the tobacco hornworm Manduca sexta binds to lipopolysaccharide and lipoteichoic acid. Insect Mol. Biol. 15: 119–128.

Zelensky, A. N., and J. E. Gready. 2003. Comparative analysis of structural properties of the C-type lectin-like domain (CTLD). Proteins. 52: 466–477.

Zelensky, A. N., and J. E. Gready. 2005. The C-type lectin-like domain superfamily. Febs J. 272: 6179–6217.

Zhang, X. W., Q. Ren, H. W. Zhang, K. K. Wang, and J. X. Wang. 2013a. A C-type lectin could selectively facilitate bacteria clearance in red swamp crayfish, Procambarus clarkii. Fish Shellfish Immunol. 35: 1387–1394.

Zhang, M., F. Zhou, Y. Chu, Z. Zhao, and C. An. 2013b. Identification and expression profile analysis of antimicrobial peptide/protein in Asian corn borer, Ostrinia furnacalis (Gueneé). Int. J. Biol. Sci. 9: 1094–1102.

Zhu, Y., E. J. Ragan, and M. R. Kanost. 2010. Leureptin: a soluble, extracellular leucine-rich repeat protein from Manduca sexta that binds lipopolysaccharide. Insect Biochem. Mol. Biol. 40: 713–722.

Zhu, J. Y., P. Yang, Z. Zhang, G. X. Wu, and B. Yang. 2013. Transcriptomic immune response of Tenebrio molitor pupae to parasitization by Scleroderma guani. Plos One. 8: e54411.

Zou, Z., J. D. Evans, Z. Lu, P. Zhao, M. Williams, N. Sumathipala, C. Hetru, D. Hultmark, and H. Jiang. 2007. Comparative genomic analysis of the Tribolium immune system. Genome Biol. 8: R177.