Insect Molecular Biology

Insect Molecular Biology (2016) **25**(4), 413–421 doi: 10.1111/imb.12226

Roles of the *Laodelphax striatellus* Down syndrome cell adhesion molecule in Rice stripe virus infection of its insect vector

F. Zhang*†‡, Q. Li*†‡, X. Chen*†, Y. Huo*†, H. Guo*†, Z. Song*†, F. Cui§, L. Zhang*† and R. Fang*†

**State Key Laboratory of Plant Genomics, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China; †National Plant Gene Research Center, Beijing, China; ‡University of the Chinese Academy of Sciences, Beijing, China; and §State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing, China*

Abstract

The arthropod Down syndrome cell adhesion molecule (Dscam) mediates pathogen-specific recognition via an extensive protein isoform repertoire produced by alternative splicing. To date, most studies have focused on the subsequent pathogen-specific immune response, and few have investigated the entry into cells of viruses or endosymbionts. In the present study, we cloned and characterized the cDNA of *Laodelphax striatellus* Dscam (LsDscam) and investigated the function of *LsDscam* in rice stripe virus (RSV) infection and the influence on the endosymbiont *Wolbachia*. LsDscam displayed a typical Dscam domain architecture, including 10 immunoglobulin (Ig) domains, six fibronectin type III domains, one transmembrane domain and a cytoplasmic tail. Alternative splicing occurred at the N-termini of the Ig2 and Ig3 domains, the complete Ig7 domain, the transmembrane domain and the C-terminus, comprising 10, 51, 35, two and two variable exons, respectively. Potentially *LsDscam* could encode at least 71 400 unique isoforms and 17 850 types of extracellular regions. *LsDscam* was expressed in various *L. striatellus* tissues. Knockdown of *LsDscam* mRNA via RNA interference decreased the titres of both RSV and *Wolbachia*, but did not change the numbers of the extracellular symbiotic bacterium *Acinetobacter rhizosphaerae*. Specific Dscam isoforms may play roles in enhancing the infection of vector-borne viruses or endosymbionts.

Keywords: Down syndrome cell adhesion molecule, *Laodelphax striatellus*, rice stripe virus infection.

First published online 14 March 2016.

Correspondence: Dr Rongxiang Fang or Dr Lili Zhang, Institute of Microbiology, Chinese Academy of Sciences, No 3, Yard 1, West Beichen Road, Chaoyang, Beijing 100101, China. Tel.: + 86 10 64858245 (Rongxiang Fang); + 86 10 64861838 (Lili Zhang); fax: + 86 10 64858245; e-mails: fangrx@sun.im.ac.cn (Rongxiang Fang); zhangll@sun.im.ac.cn (Lili Zhang)

Introduction

The arthropod Down syndrome cell adhesion molecule (Dscam) has been the focus of much attention because of the extensive alternative splicing of this protein that generates tens of thousands of protein isoforms containing variable immunoglobulin (Ig) domain combinations (Schmucker et al., 2000; Watson et al., 2005; Brites et al., 2008; Jin et al., 2013; Ng et al., 2014; Li et al., 2015). Most arthropod Dscam homologues share similar domain architecture: nine Ig domains – four fibronectin type (FN) III domains – one Ig domain – two FNIII domains – transmembrane (TM) domain – cytoplasmic tail. Alternative splicing has been reported to occur at the N-termini of the Ig2 and Ig3 domains, the complete Ig7 domain, the TM domain and/or the cytoplasmic tail (Ng et al., 2014). For instance, Dscam protein in *Drosophila melanogaster* has 12, 48, 33 and two alternatives arising from the first four variable exon clusters, respectively, and produces 38 016 isoforms via alternative splicing, many more than the 15 016 genes encoded by the *D. melanogaster* genome (Schmucker et al., 2000). Commonly, the Ig-domain containing proteins are soluble or located on the cell surface, and are involved in ligand recognition, binding and adhesion processes (Garver et al., 2008; Dermody et al., 2009). Therefore, the arthropod Dscam was predicted to have the potential to recognize diverse ligands and pathogens. Over the past decade, accumulating evidence has been reported

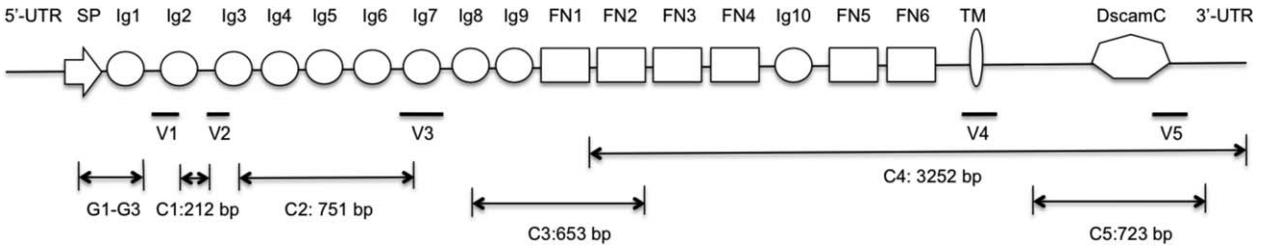

Figure 1. Cloning of the cDNA of the *Laodelphax striatellus Down syndrome cell adhesion molecule* (*LsDscam*) gene. C1–C5, five cDNA fragments identified from the annotated *L. striatellus* transcriptomes; G1–G3, three cDNA fragments identified from the *L. striatellus* genomic sequence; SP, cDNA fragment encoding the signal peptide; Ig1–10, 10 cDNA fragments encoding the corresponding immunoglobulin domains; FN1–6, six cDNA fragments encoding the corresponding fibronectin type III domains; TM, cDNA fragment encoding the transmembrane domain; DscamC, cDNA fragment encoding the DscamC domain; V1–V5, five variable regions.

to support this hypothesis (Syed Musthaq & Kwang, 2014; Armitage et al., 2015). For example, when challenged with different pathogens or elicitors, the *Anopheles gambiae* Dscam (AgDscam) could produce different splice form repertoires (Dong et al., 2006; Smith et al., 2011). The isoforms induced by a specific pathogen showed significantly stronger binding affinity to the pathogen itself (Watthanasurorot et al., 2011; Hung et al., 2013). Chou et al. (2009) revealed that acute white spot syndrome virus infection induced more isoforms, whereas long-term exposure to this virus resulted in some of these isoforms that might be able to specifically recognize the pathogen becoming predominant, indicating the stable inducible expression of pathogen-specific isoforms.

Since the discovery of the Dscam-mediated pathogen-specific recognition system, most studies have focused on the effect of Dscam on insect immune responses. Watson et al. (2005) revealed that *Drosophila Dscam* was highly expressed in haemocytes and fat body cells, which constitute important cells of the insect immune system (Ramet et al., 2002; Tzou et al., 2002). Disturbing *Dscam* expression impaired the ability of haemocytes to phagocytose bacteria, indicating its function in activating and mediating phagocytosis (Watson et al., 2005; Dong et al., 2006; Watthanasurorot et al., 2011). Compared with the phagocytosis of pathogens (especially extracellular pathogens), receptor-ligand recognition could also mediate the entry into cells of some viruses, intracellular pathogens and endosymbionts (Miller et al., 2008; Cheng et al., 2010). At present, it is not known whether Dscam mediates the entry of intracellular pathogens/symbionts.

The small brown planthopper (SBPH), *Laodelphax striatellus*, is an economically important pest and vectors rice stripe virus (RSV), the causative agent of rice stripe diseases (Falk & Tsai, 1998). RSV can infect various *L. striatellus* tissues, be horizontally transmitted into the healthy plant or vertically transmitted to the offspring (Huo et al., 2014; Wu et al., 2014). Vertical transmission results in naturally existing RSV-infected *L. striatellus*. By analysing the microbiota within the whole body of SBPH, we found that *Wolbachia* endosymbionts constitute $>90\%$ of symbiotic bacteria. In this study, we cloned and characterized the *L. striatellus Dscam* (*LsDscam*), and investigated the role of the *Dscam* played in the infection level of persistent-infected virus and endosymbionts.

Results

Cloning and molecular characterization of LsDscam *cDNA*

The strategy employed to clone the cDNA of *LsDscam* is shown in Fig. 1. By searching the annotated *L. striatellus* transcriptomes [(Zhang et al., 2010) and two unpublished databases], five cDNA fragments that were predicted to encode Dscam homologues were identified (labelled as C1–C5). The longest fragment, C4, which was 3252 bp in length and located at the 3′ terminus, was used to search the databases using the BLASTX program (Altschul et al., 1997). The deduced amino acid sequence of this LsDscam fragment showed the highest level of identity to the *Megachile rotundata* Dscam (72.9%), which was subsequently used as a template to locate the other LsDscam fragments. The five located fragments covered a region of 5524 bp. The *Me. rotundata Dscam* cDNA was then used in a BLAST search against the *L. striatellus* genome (unpublished raw data) to identify the 5′-terminal fragments, of which three DNA fragments were identified (G1–G3). Primer pairs were designed according to the identified sequences and PCR was performed to fill the gaps. Rapid amplification of cDNA ends PCR (RACE-PCR) was performed to clone the 5′ and 3′ untranslated regions (UTRs), and overlapping PCR was performed to confirm the assembled full-length cDNA.

The *LsDscam* cDNA was 6.7 kb in length, containing an open reading frame of 5994 bp, a 5′ UTR of 361 bp and a 3′ UTR of 320 bp (Supporting Information Fig. S1). The deduced amino acid sequence consisted of

1997 residues with a calculated molecular mass of 220 kDa. By SignalP 4.1 analysis (Petersen et al., 2011), a signal peptide was found in the N-terminus with a cleavage site between amino acids 27 and 28 (Figs 1, S1). TransMembrane prediction using Hidden Markov Models (http://www.cbs.dtu.dk/services/TMHMM/) identified a TM helix from amino acids 1621 to 1643. These results suggested that LsDscam is a TM protein with an N-terminal extracellular region and a C-terminal cytoplasmic tail. Using SMART software, a total of 10 Ig domains and six FNIII domains were identified (Figs 1, S1). The domain architecture was: Ig 1–9, FNIII 1–4, Ig 10, FNIII 5–6, a TM domain and a cytoplasmic tail (Figs 1, S1).

BLAST analysis (Altschul et al., 1997) against sequence databases revealed that the full-length LsDscam amino acid sequence shared 75% identity to Dscam from *Me. rotundata*, *Apis mellifera*, *Monomorium pharaonis*, *Pogonomyrmex barbatus* and *Athalia rosae*, amongst others. Alignment of LsDscam with a set of *Me. rotundata* Dscam isoforms confirmed the four variable regions, located in the N-termini of Ig2 and Ig3, Ig7 and the TM region. Four sets of primer pairs V1-F/V1-R, V2-F/V2-R, V3-F/V3-R and V4-F/V4-R were designed for variable region sequencing (Table S1), and 100, 200, 200 and 10 randomly selected clones from each of the four types of cloned PCR products were sequenced and the deduced amino acid sequences were aligned using the CLUSTALW program. For the four variable regions listed above, 10, 51, 35 and two alternative sequences were identified, respectively (Fig. 2A–D). The 3′-UTR PCR revealed two alternative sequences located in the C-terminus (Fig. 2E). In theory, LsDscam may produce at least 71 400 isoforms ($10 \times 51 \times 35 \times 2 \times 2$), including 17 850 types of extracellular regions and two types of cytoplasmic regions. We then PCR amplified cDNAs covering the first three variable regions. Each of the 24 sequenced clones contained a set of Ig2, Ig3 and Ig7, indicating the production of LsDscam isoforms. The 6.0 kb LsDscam open reading frame was then PCR amplified to confirm expression of the complete gene. Two positive clones were obtained and each contained conserved regions and a set of domains: Ig2, Ig3, Ig7, a TM domain and a cytoplasmic tail.

LsDscam *expression in different* L. striatellus *tissues*

Detailed expression profiling can aid investigations into the potential location of gene function. We thus tested *LsDscam* expression in different *L. striatellus* tissues, including the midgut and salivary glands, which are key tissues in RSV transmission, and the fat body, which plays roles in insect immunity. Quantitative Reverse-Transcription PCR (qRT-PCR) analysis indicated that *LsDscam* was expressed in all three tissues (Fig. 3). Healthy SBPH showed similar *LsDscam* expression levels in the midgut and fat body, whereas fivefold higher expression levels occurred in the salivary glands (Fig. 3). RSV infection significantly increased *LsDscam* expression in both the salivary glands and fat body, whereas RSV induction of *LsDscam* expression in the midgut was not significant (Fig. 3). The *LsDscam* expression profile was consistent with the RSV distribution in viruliferous *L. striatellus*, in which the virus titre was several-fold higher in the salivary glands and fat body than in the midgut [Fig. S2, also indicated in Fig. 4B, RSV burden in the double-stranded enhanced green fluorescent protein (*ds*GFP) group].

LsDscam-*deficient SBPHs show reduced RSV burden*

To determine the role of *LsDscam* in RSV infection, we generated *LsDscam*-deficient SBPH using RNA interference (RNAi). The double-stranded RNA (*ds*RNA) was designed to target the conserved domains (primers dscamT7-F/dscamT7-R, Table S1), which would result in deficiency of all *LsDscam* isoform mRNAs. Equal volumes of *LsDscam ds*RNA and e*gfp ds*RNA (abbreviated as *ds*Dscam and *ds*GFP, respectively) were delivered into the haemolymph of RSV-infected third-instar nymphal *L. striatellus* via microinjection, and the SBPHs were allowed to feed on healthy rice seedlings for 7 days. As *LsDscam* was expressed in various tissues that RSV colonized, we measured the virus burden in the whole body of the insect. qRT-PCR analysis indicated that compared with those injected with *ds*GFP, the *ds*Dscam-treated insects had significantly lower *LsDscam* mRNA abundance (Fig. 4A, LsDscam). *LsDscam* deficiency significantly decreased RSV burden, with a reduction rate of about 50% (Fig. 4A, RSV), suggesting that *LsDscam* might play a positive role in enhancing RSV accumulation. We further measured the effect of *LsDscam* on the RSV burden in individual tissues including the midgut, salivary glands and fat body. Consistent with the whole body results, RSV burden in both salivary glands and the fat body was significantly decreased by *LsDscam* deficiency (Fig. 4B). The virus burden in the midgut was also decreased, although the difference was not significant as determined by a Student's *t*-test (Fig. 4B).

LsDscam-*deficiency reduced the infection level of rice black streaked dwarf virus (RBSDV) and* Wolbachia *endosymbiont*

Our previous transcriptome analyses indicated the existence of *Wolbachia* endosymbionts within the *L. striatellus* body (Zhang et al., 2010). In this work, by deep sequencing of 16S ribosomal DNA, we found that the diversity of bacterial communities harboured by *L. striatellus* was low. *Wolbachia* was the only endosymbiont detected and accounted for more than 90% of the total

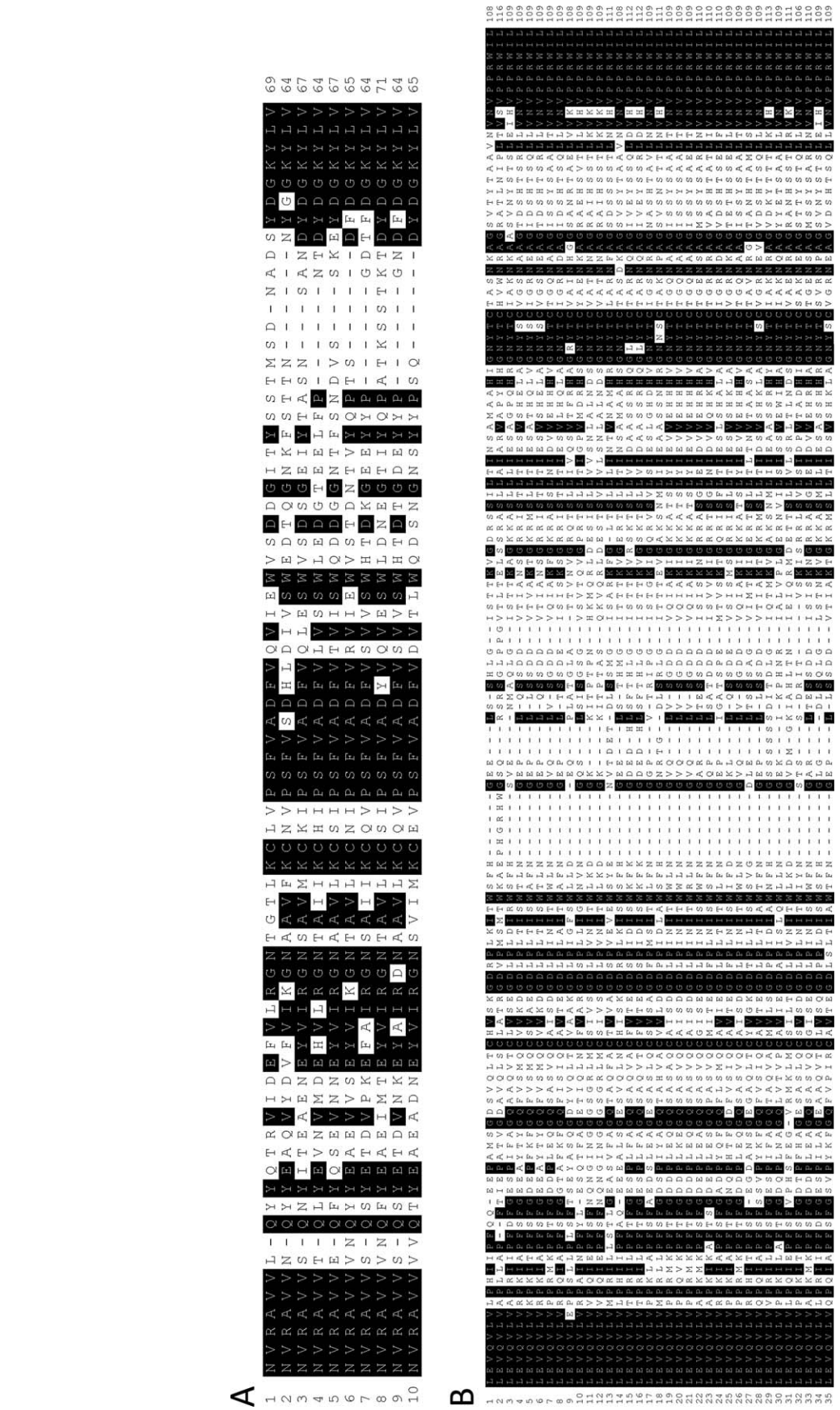

Figure 2. Sequence alignment of each of the *Laodelphax striatellus* Down syndrome cell adhesion molecule (LsDscam) variable regions. (A) Variable region 1, the N-terminus of the immunoglobulin 2 (Ig2) domain; (B) variable region 2, the N-terminus of the Ig3 domain; (C) variable region 3, the entire Ig7 domain; (D) variable region 4, the transmembrane domain; (E) variable region 5, the LsDscam C-terminus. The variable isoforms are numbered along the left margin. The total number of amino acids for each isoform is indicated on the right. Conserved amino acids are shaded in black.

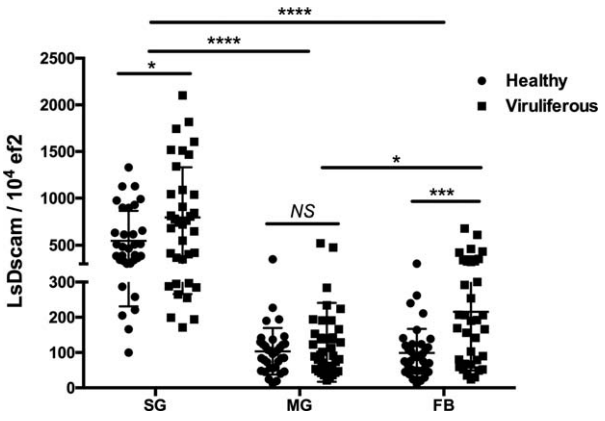

Figure 3. *Laodelphax striatellus Down syndrome cell adhesion molecule* (*LsDscam*) mRNA distribution in different *L. striatellus* tissues. In healthy *L. striatellus* (Healthy), similar *LsDscam* expression levels were found in the midgut (MG) and fat body (FB), whereas higher expression was evident in the salivary glands (SG). Rice stripe virus infection (Viruliferous) significantly increased *LsDscam* expression in the SG and FB, whereas its influence on *LsDscam* expression in MG was not significant (*NS*). Each dot represents the corresponding tissue obtained from one third-instar nymphal *L. striatellus*. Both mean and SD were calculated from three independent experiments, with 10–12 mRNA samples per experiment. *, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$.

bacteria (Fig. 4C: a–e). *Acinetobacter rhizosphaerae*, an extracellular symbiotic bacteria, was present in all of the tested *L. striatellus* samples (Fig. 4C: f).

Similar to RSV, *Wolbachia* also infects various *L. striatellus* tissues. To investigate the influence of *LsDscam* on the *Wolbachia* infection level, the SBPHs were treated with either *ds*Dscam or *ds*GFP according to the methods described above. Seven days after feeding on healthy rice seedlings, SBPHs were collected and qRT-PCR was performed to determine the *Wolbachia* burden in the whole body of *L. striatellus*. Compared with the *ds*GFP-treated SBPHs, the *ds*Dscam-treated SBPHs showed a significantly lower burden of *Wolbachia* (Fig. 4A: *Wolbachia*). By contrast, *LsDscam*-deficiency did not change the number of the extracellular *Ac rhizosphaerae* (Fig. 4A: *Ac. rhizosphaerae*). These results suggested that *LsDscam* might play a role in enhancing the infection of intracellular pathogens/symbionts.

We further tested the effect of *LsDscam* on the infection level of RBSDV, a fijivirus that is also transmitted by *L. striatellus* in a persistent propagative manner (Hajano et al., 2015). As RBSDV is not vertically transmitted, fully infected SBPH were first generated according to the methods described previously (Hajano et al., 2015), and then treated with *ds*Dscam or *ds*GFP. Virus burden was analysed 7 days later. Similar to the results obtained from RSV-infected *L. striatellus*, the *ds*Dscam-treated SBPHs showed a significantly lower RBSDV burden when compared with the *ds*GFP-treated insects (Fig. 4D). These results further confirmed the function of *LsDscam* in enhancing infection of intracellular pathogens.

Discussion

This study sequenced and characterized *LsDscam* and investigated its effect on a vertically acquired virus and intracellular symbiont. Our findings indicate that *LsDscam* plays a positive role in enhancing the infection of both RSV and *Wolbachia* endosymbionts.

The *L. striatellus* genome has not been completely sequenced. As restricted alternative splicing has been detected in some tissues (Watson et al., 2005; Dong et al., 2006), we extracted RNA from the SBPH whole body to maximize the capture of alternative exons. Using Reverse Transcription PCR (RT-PCR) and RACE-PCR, we found that LsDscam has at least 10 Ig2 variants, 51 Ig3 variants, 35 Ig7 variants, two TM variants and two cytoplasmic tails (Fig. 2), suggesting that there are at least 71 400 possible combinations. LsDscam has identical domain components to the classic arthropod Dscam orthologues: one signal peptide, 10 Ig domains, six FNIII domains, one TM domain and one cytoplasmic tail. A phylogenetic tree based on the protein sequences of the conserved region from domain Ig8 to FNIII6 was constructed to evaluate the molecular evolutionary relationships of LsDscam with other Dscams. In the constructed phylogenetic tree (Fig. S3), LsDscam firstly clustered with homologues from arthropods, then clustered with homologues from vertebrates. The relationships displayed in the phylogenetic tree are largely consistent with traditional taxonomy.

The distribution of Dscam mRNA in different tissues may aid elucidation of its potential function. Although in most studied arthropod systems, *Dscam* has been demonstrated to be highly expressed in both neurones and in immune-related tissues (Armitage et al., 2015), some studies have also indicated *Dscam* expression in tissues or cell types other than the immune or nervous systems (Celotto & Graveley, 2001; Dong et al., 2006; Chou et al., 2009). For example, *Litopenaeus vannamei Dscam* was expressed in 10 shrimp tissues, and the mRNA levels in the heart and midgut were comparable to those in the lymphoid organ and haemocytes (Chou et al., 2009). Upon *Plasmodium* invasion of the *An. gambiae* midgut epithelium, the *AgDscam* gene in the corresponding tissue responded by producing parasite-specific isoform repertoires (Dong et al., 2006). At present, the function that *LsDscam* expressed in tissues not belong to the immune or nervous systems remain unclear. Compared with the function of Dscam in immunity, its role in non-immune responses is less well understood (Syed Musthaq & Kwang, 2014; Armitage et al., 2015). In this study, the transcripts of *LsDscam* were found to be significantly higher in the salivary glands than in the fat body (the immune-related tissue in insects). Within persistent-infected SBPH, the salivary

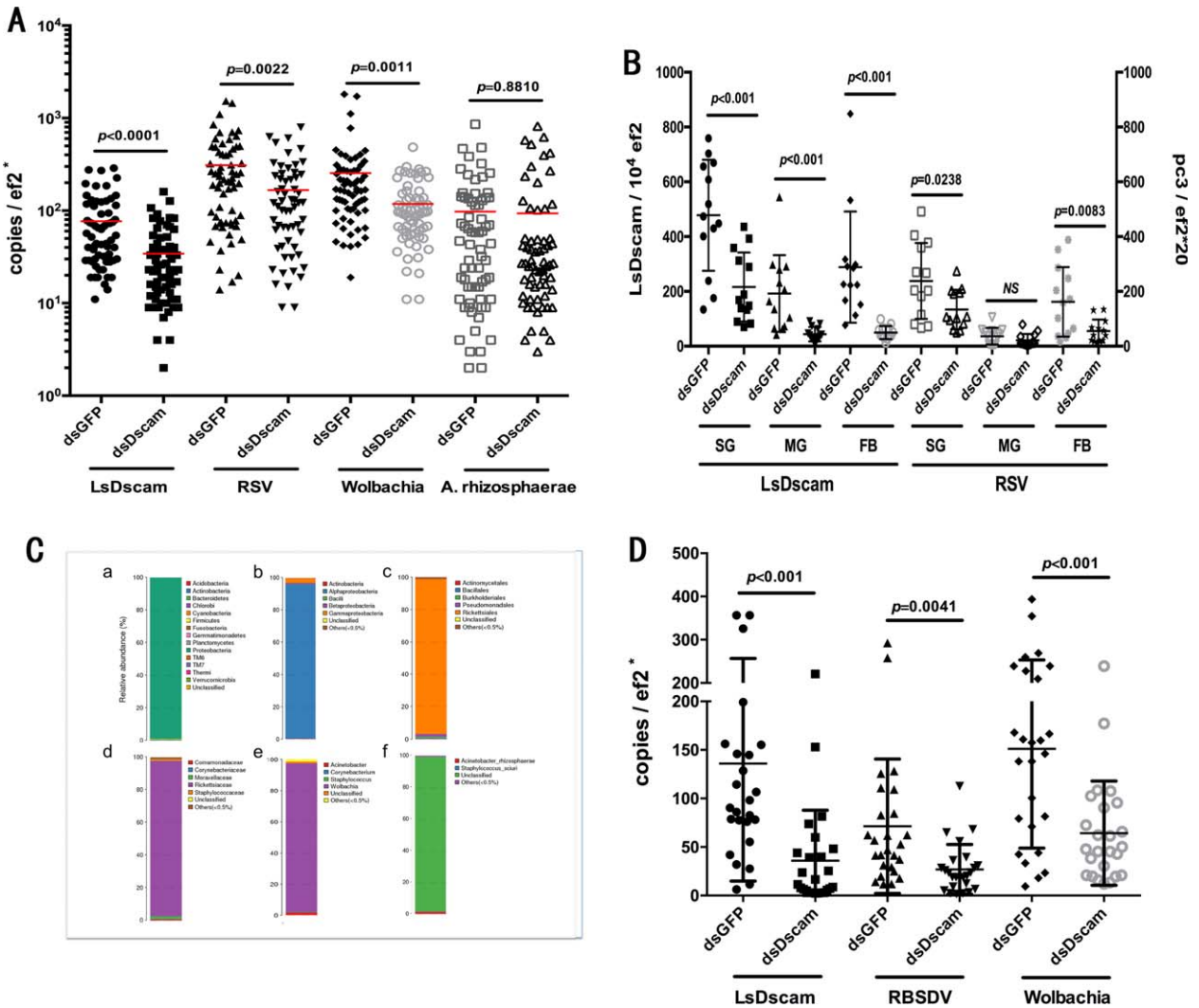

Figure 4. Effect of *Laodelphax striatellus Down syndrome cell adhesion molecule* (*LsDscam*)-deficiency (double-stranded LsDscam, *ds*Dscam) on rice stripe virus (RSV), *Wolbachia*, rice black streaked dwarf virus (RBSDV) and *Acinetobacter rhizosphaerae* burden in the small brown planthopper (SBPH). (A) *LsDscam*-deficiency (*ds*Dscam) decreased both RSV and *Wolbachia* burden in the SBPH whole body, but did not affect the *Ac. rhizosphaerae* burden. Each dot represents one third-instar nymph. Data were collected from four independent experiments. Double-stranded enhanced green fluorescent protein (dsGFP) was used as a RNA interference (RNAi) control. (B) *LsDscam*-deficiency (*ds*Dscam) significantly decreased RSV burden in the salivary glands (SG) and fat body (FB); it also decreased RSV burden in the midgut (MG) although not significantly (*NS*). Each dot represents the corresponding tissue from one third-instar nymphal *L. striatellus*. The mean and SD were calculated from two independent experiments. (C) The composition of symbiotic bacteria within the *L. striatellus* whole body. From (a) to (f): phylum (a), class (b), order (c), family (d), genus (e) and species (f) level compositions of each sample. *Wolbachia* belongs to Proteobacteria (shown in green in a), Alphaproteobacteri (shown in blue in b), Rickettsiales (shown in orange in c), Rickettsiaceae (shown in purple in d) and Wolbachia (shown in purple in e). *Ac. rhizosphaerae* (shown in red in f) belongs to Proteobacteria (shown in green in a), Gammaproteobacteri (shown in orange in b), Pseudomonadales (shown in purple in c), Moraxellaceae (shown in green in d) and *Acinetobacter* (shown in red in e). (D) *LsDscam*-deficiency (*ds*Dscam) decreased RBSDV burden in the SBPH whole body. Each dot represents one RBSDV-infected SBPH. Data were collected from three independent experiments. dsGFP was used as an RNAi control. **LsDscam* mRNA level shown as *LsDscam* copies/10^4 ef2; RSV burden shown as *pc3* copies/100 ef2; *Wolbachia* burden shown as 16S copies/ef2; *Ac. rhizosphaerae* burden shown as 16S copies/10^3 ef2; RBSDV burden shown as *S4* copies/10 ef2. ef2, elongation factor; pc3, nucleoprotein (Protein pc3); 16S, 16S rRNA; S2, genome segment 2.

gland harbours the highest RSV burden (Fig. 4B) and is the most active tissue in transmitting RSV. One possibility is that high *LsDscam* abundance in the salivary glands would facilitate a rapid accumulation of the virus.

Whereas most of the recognition by haemocytes mediated phagocytosis, recognition by receptors in both haemocytes and non-immune cells could also mediate the entry of intracellular pathogens/symbionts. When *L. striatellus*, the major vector for RSV (Falk & Tsai, 1998), becomes infected with this virus, the virus can infect the midgut epithelium and disseminate into the haemolymph. From the haemolymph, RSV can further invade various tissues, including the fat body, salivary glands, testis and ovaries (Wu et al., 2014). In general,

arthropod haemolymph is hostile to pathogens because it contains both circulating haemocytes that mediate phagocytosis, and antibacterial proteins/peptides produced by both the haemocytes and the fat body cells (Johns et al., 1998; Fogaca et al., 1999; Ceraul et al., 2002). However, many vector-borne pathogens can survive in the haemolymph for a period longer than the time required for them to pass through the haemolymph and reach the next tissues. The *L. striatellus* haemolymph was also found to be a relatively benign environment for RSV. The virus was able to survive in the haemolymph for more than 24 h (Fig. S4), which was much longer than the time taken for a microbe to pass through the haemolymph. In this case, and in non-immune-related tissues, recognition by Dscam on the cell surface might mediate virus entry into cells rather than mediating virus clearance via phagocytosis.

Future studies will aim to identify the LsDscam isoforms that are specific to RSV and *Wolbachia*. The present study has provided a new insight that the arthropod Dscam, following the recognition of some vector-borne viruses and intracellular symbionts, is able to facilitate the entry of microbes into cells.

Experimental procedures

Virus, small brown planthopper and host plant

Both the RSV-free (healthy) and RSV-infected (viruliferous) *L. striatellus* used in this study were originally captured in Jiangsu Province, China, and were maintained in our laboratory. All plants used for *L. striatellus* rearing were grown inside a growth incubator at 25 °C with a photoperiod of 16 h light and 8 h dark. To ensure a high offspring infection rate, viruliferous female imagoes were cultured separately and 15% of the corresponding offspring were tested for RSV infection through dot-enzyme-linked immunosorbent assay using RSV-specific monoclonal antibodies (provided by Dr Xueping Zhou, Institute of Biotechnology, Zhejiang University; Zhou et al., 2004). An insect population with a predicted infection rate of 100% was used in the experiments.

Full-length LsDscam *cDNA cloning*

Total RNA was extracted from the *L. striatellus* whole body. For each RNA sample, a pool of five healthy SBPHs and five viruliferous SBPHs were grouped. cDNAs were synthesized according to the protocols provided by the manufacturer [SuperScript® VILO™ Master Mix, Invitrogen (cat no. 11755050), Carlsbad, CA, USA].

The partial *LsDscam* cDNA sequences that were obtained from the annotated *L. striatellus* transcriptomes [(Zhang et al., 2010) and two unpublished databases] and the *L. striatellus* genomic sequence (unpublished raw data) were used as templates to design primers for gap filling. The primer pairs were: GAP-1F/GAP-1R, GAP-2F/GAP-2R, GAP-3F/GAP-3R and GAP-4F/GAP-4R. Three rounds of 5′-RACE were performed to amplify overlapping fragments according to the protocols provided by the manufacturer [FirstChoice® RLM-RACE Kit, Ambion (cat. no. 1700), Waltham, MA, USA]. Two rounds of 3′-RACE were performed to clone the 3′ end. RACE-PCR primers specific for the 5′ end were 5R1, 5R2 and 5R3, and primers specific for the 3′ end were 3R1 and 3R2 (Table S1). All PCR products were cloned into the pGEM-T vector (Promega, Madison, WI, USA) for sequencing.

After obtaining the full-length *LsDscam* cDNA, the variable regions were predicted from alignments with a set of *Me. rotundata* Dscam isoforms. Four primer pairs, V1-F/V1-R, V2-F/V2-R, V3-F/V3-R and V4-F/V4-R (Table S1), were designed to amplify the four variable regions. Primer pair V1-F/V3-R (Table S1) was used to amplify the *LsDscam* fragment containing the first three variable regions, and primer pair Dscamtotal-F/Dscamtotal-R (Table S1) was used to amplify the full-length *LsDscam*. All PCR products were cloned into the pGEM-T vector for sequencing.

Tissue collection

SBPHs were chilled on ice before being dissected in phosphate-buffered saline (PBS). Three tissue types were isolated: midgut, salivary glands and fat body. Both the midgut and salivary glands were washed twice in PBS to remove any contaminating virus from the haemolymph.

qRT-PCR to measure LsDscam *expression in different* L. striatellus *tissues*

For quantitative analysis of *LsDscam* expression in different tissues, healthy or viruliferous *L. striatellus* were dissected and the tissues were collected according to the protocol described above. RNAs were extracted from the tissues of single insects. RT-PCR and SYBR-Green based quantitative PCR (qPCR) were performed according to the protocols provided by the manufacturer (Hercules, CA, USA). *L. striatellus* elongation factor *ef2* (primers listed in Table S1) was amplified as an internal control for the loading of cDNA isolated from different samples. H_2O was used as a negative control.

RNAi

A 500-bp *LsDscam* fragment was PCR amplified using the primers dscamT7-F and dscamT7-R (Table S1). Then, dsRNA was synthesized using a commercial kit (MEGAscript T7 kit, Ambion cat. no. AM1334) and purified by phenol : chloroform extraction and isopropanol precipitation. For *LsDscam* silencing in the viruliferous *L. striatellus* whole body, the *LsDscam ds*RNA was microinjected into the insect haemolymph as described previously (Huo et al., 2014). The third-instar nymphal *L. striatellus* was injected with 23.6 nl dsRNA at 1 ng/nl. Then, *gfp* dsRNA, which was used as negative control, was microinjected following the same protocol. The insects were cultured in new chambers with healthy rice seedlings for 7 days. To assess the RSV burden in different tissues, SBPHs were dissected and tissues including the gut, salivary glands and fat body were collected separately. RNA was extracted from a single SBPH or the tissue from a single SBPH. qPCR primers for assessing *LsDscam* silencing efficiency were dscam-F/dscam-R (Table S1). The RSV nucleoprotein (Protein *pc3*), the Wolbachia *16S rRNA* and the *Ac. rhizosphaerae 16S rRNA* were PCR amplified

to assess the number of corresponding microbes. The 16S primer sequences were designed to locate at variable regions to ensure species specific PCR amplification. The sequences of the qPCR primers, pc3-F/pc3-R, wol16S-F/wol16S-R and arh16S-F/arh16S-R, are listed in Table S1.

16S rDNA sequence analysis

Third-instar *L. striatellus* were first surface-sterilized with 70% ethanol for 3–5 s and then rinsed with sterilized water for three times. DNA was extracted from pools of 20 viruliferous third-instar *L. striatellus* (in total, 80 viruliferous *L. striatellus*). The genomic DNA was extracted according to the manufacturer's instructions [Blood & Cell Culture DNA Mini Kit, Qiagen (cat. no. 13323), Valencia, CA, USA]. The V3-V4 variable region of the bacterial 16S rRNA was amplified as described by Fadrosh et al. (2014) and sequenced on the Illumina MiSeq platform (San Diego, California, USA). Sequences with a quality score lower than 20 were discarded (Fadrosh et al., 2014). Short reads were merged into tags using FLASH (Fast Length Adjustment of Short Reads, v. 1.2.11; Magoc & Salzberg, 2011), and clean tags were grouped into operational taxonomic units (OTUs) at the species level by clustering them based on a sequence similarity score of 97% using the USEARCH program v. 7.0.1090 (Edgar, 2013). OTUs were used in BLAST searches against the databases, and the phylum, class, order, family, genus and species level compositions of each sample was annotated.

Sequence analysis

Homology searches of the cDNA and protein sequences of LsDscam were conducted using the BLAST algorithm at the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignments of LsDscam and other Dscams were performed using the CLUSTALW program (http://www.ch.embnet.org/software/ClustalW.html). The phylogenetic tree was constructed based on the deduced amino acid sequence of LsDscam and other Dscams using the neighbour-joining algorithm and MEGA 6.06 software (www.megasoftware.net).

Microinjection of RSV into the L. striatellus *haemolymph*

Purified RSV particles in PBS buffer were adjusted to 0.1 ng/nl. Then, 23.6 nl of the virus solution was delivered into the third-instar *L. striatellus* haemolymph via microinjection. The insects were cultured in new chambers with healthy rice seedlings, and at different time points postinjection, SBPHs were collected and RNA was extracted from the insect whole body to measure RSV titres.

Statistical analysis

Graphing and statistical analysis were performed using PRISM 6.0 software (GraphPad Software Inc., La Jolla, CA, USA). Results are expressed as the mean ± SD. The significance of the differences between the mean values of the groups was evaluated using Student's *t*-tests.

Acknowledgements

This work was supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (no. XDB11040100) and the Major State Basic Research Development Program of China (973 Program) (no. 2014CB138401).

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Armitage, S.A., Peuss, R. and Kurtz, J. (2015) Dscam and pancrustacean immune memory - a review of the evidence. *Dev Comp Immunol* **48**: 315–323.

Brites, D., McTaggart, S., Morris, K., Anderson, J., Thomas, K., Colson, I. *et al.* (2008) The Dscam homologue of the crustacean *Daphnia* is diversified by alternative splicing like in insects. *Mol Biol Evol* **25**: 1429–1439.

Celotto, A.M. and Graveley, B.R. (2001) Alternative splicing of the *Drosophila* Dscam pre-mRNA is both temporally and spatially regulated. *Genetics* **159**: 599–608.

Ceraul, S.M., Sonenshine, D.E. and Hynes, W.L. (2002) Resistance of the tick *Dermacentor variabilis* (Acari: Ixodidae) following challenge with the bacterium *Escherichia coli* (Enterobacteriales: Enterobacteriaceae). *J Med Entomol* **39**: 376–383.

Cheng, G., Cox, J., Wang, P., Krishnan, M.N., Dai, J., Qian, F. *et al.* (2010) A C-type lectin collaborates with a CD45 phosphatase homolog to facilitate West Nile virus infection of mosquitoes. *Cell* **142**: 714–725.

Chou, P.H., Chang, H.S., Chen, I.T., Lin, H.Y., Chen, Y.M., Yang, H.L. *et al.* (2009) The putative invertebrate adaptive immune protein *Litopenaeus vannamei* Dscam (LvDscam) is the first reported Dscam to lack a transmembrane domain and cytoplasmic tail. *Dev Comp Immunol* **33**: 1258–1267.

Dermody, T.S., Kirchner, E., Guglielmi, K.M. and Stehle, T. (2009) Immunoglobulin superfamily virus receptors and the evolution of adaptive immunity. *PLoS Pathog* **5**: e1000481.

Dong, Y., Taylor, H.E. and Dimopoulos, G. (2006) AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. *PLoS Biol* **4**: e229.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998.

Fadrosh, D.W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R.M. *et al.* (2014) An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**: 6.

Falk, B.W. and Tsai, J.H. (1998) Biology and molecular biology of viruses in the genus Tenuivirus. *Annu Rev Phytopathol* **36**: 139–163.

Fogaca, A.C., da Silva, P.I., Miranda, M.T., Bianchi, A.G., Miranda, A., Ribolla, P.E. *et al.* (1999) Antimicrobial activity of a bovine hemoglobin fragment in the tick *Boophilus microplus*. *J Biol Chem* **274**: 25330–25334.

Garver, L.S., Xi, Z. and Dimopoulos, G. (2008) Immunoglobulin superfamily members play an important role in the mosquito immune system. *Dev Comp Immunol* **32**: 519–531.

Hajano, J.U., Wang, B., Ren, Y., Lu, C. and Wang, X. (2015) Quantification of southern rice black streaked dwarf virus and rice black streaked dwarf virus in the organs of their vector and nonvector insect over time. *Virus Res* **208**: 146–155.

Hung, H.Y., Ng, T.H., Lin, J.H., Chiang, Y.A., Chuang, Y.C. and Wang, H.C. (2013) Properties of *Litopenaeus vannamei* Dscam (LvDscam) isoforms related to specific pathogen recognition. *Fish Shellfish Immunol* **35**: 1272–1281.

Huo, Y., Liu, W., Zhang, F., Chen, X., Li, L., Liu, Q. *et al*. (2014) Transovarial transmission of a plant virus is mediated by vitellogenin of its insect vector. *PLoS Pathog* **10**: e1003949.

Jin, X.K., Li, W.W., Wu, M.H., Guo, X.N., Li, S., Yu, A.Q. *et al*. (2013) Immunoglobulin superfamily protein Dscam exhibited molecular diversity by alternative splicing in hemocytes of crustacean, *Eriocheir sinensis*. *Fish Shellfish Immunol* **35**: 900–909.

Johns, R., Sonenshine, D.E. and Hynes, W.L. (1998) Control of bacterial infections in the hard tick *Dermacentor variabilis* (Acari: Ixodidae): evidence for the existence of antimicrobial proteins in tick hemolymph. *J Med Entomol* **35**: 458–464.

Li, D., Yu, A.Q., Li, X.J., Zhu, Y.T., Jin, X.K., Li, W.W. *et al*. (2015) Antimicrobial activity of a novel hypervariable immunoglobulin domain-containing receptor Dscam in *Cherax quadricarinatus*. *Fish Shellfish Immunol* **47**: 766–776.

Magoc, T. and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

Miller, J.L., de Wet, B.J., Martinez-Pomares, L., Radcliffe, C.M., Dwek, R.A., Rudd, P.M. *et al*. (2008) The mannose receptor mediates dengue virus infection of macrophages. *PLoS Pathog* **4**: e17.

Ng, T.H., Chiang, Y.A., Yeh, Y.C. and Wang, H.C. (2014) Review of Dscam-mediated immunity in shrimp and other arthropods. *Dev Comp Immunol* **46**: 129–138.

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

Ramet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B. and Ezekowitz, R.A. (2002) Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*. *Nature* **416**: 644–648.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M. *et al*. (2000) *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**: 671–684.

Smith, P.H., Mwangi, J.M., Afrane, Y.A., Yan, G., Obbard, D.J., Ranford-Cartwright, L.C. *et al*. (2011) Alternative splicing of the *Anopheles gambiae* Dscam gene in diverse *Plasmodium falciparum* infections. *Malar J* **10**: 156.

Syed Musthaq, S.K. and Kwang, J. (2014) Evolution of specific immunity in shrimp - a vaccination perspective against white spot syndrome virus. *Dev Comp Immunol* **46**: 279–290.

Tzou, P., De Gregorio, E. and Lemaitre, B. (2002) How Drosophila combats microbial infection: a model to study innate immunity and host-pathogen interactions. *Curr Opin Microbiol* **5**: 102–110.

Watson, F.L., Puttmann-Holgado, R., Thomas, F., Lamar, D.L., Hughes, M., Kondo, M. *et al*. (2005) Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309**: 1874–1878.

Watthanasurorot, A., Jiravanichpaisal, P., Liu, H., Soderhall, I. and Soderhall, K. (2011) Bacteria-induced Dscam isoforms of the crustacean, *Pacifastacus leniusculus*. *PLoS Pathog* **7**: e1002062.

Wu, W., Zheng, L., Chen, H., Jia, D., Li, F. and Wei, T. (2014) Nonstructural protein NS4 of Rice stripe virus plays a critical role in viral spread in the body of vector insects. *PLoS ONE* **9**: e88636.

Zhang, F., Guo, H., Zheng, H., Zhou, T., Zhou, Y., Wang, S. *et al*. (2010) Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). *BMC Genomics* **11**: 303.

Zhou, Y.J., Liu, H.J., Wang, G.Z., Huang, X., Cheng, Z.B., Chen, Z.X. *et al*. (2004) Immuno-detection of Rice stripe virus carried by brown planthopper. *Jiangsu Agric Sci* **1**: 50–51.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Primers used in this study.

Figure S1. Nucleotide and predicted amino acid sequences of *Laodelphax striatellus* Down syndrome cell adhesion molecule (LsDscam). Signal peptide (bold), domains (underlined) and variable regions (V1–V5, circled) are indicated.

Figure S2. Rice stripe virus (RSV) distribution in different *Laodelphax striatellus* tissues. Each dot represents the corresponding tissue obtained from one third-instar nymphal *L. striatellus*. Both mean and SD were calculated from two independent experiments. MG, midgut; SG, salivary glands; FB, fat body.

Figure S3. Phylogenetic tree of Down syndrome cell adhesion molecule (Dscam; $n = 17$). The tree was constructed based on the deduced amino acid sequence of *Laodelphax striatellus* Dscam (LsDscam) and other Dscams by the neighbour-joining algorithm using MEGA 6.06 software. The scale bar corresponds to 0.1 estimated amino-acid substitutions per site.

Figure S4. Rice stripe virus (RSV) survival in the *Laodelphax striatellus* haemolymph. An equal volume of RSV solution was delivered into each *L. striatellus* haemolymph via microinjection. After 0, 1, 24, 48 or 96 h, RSV numbers were calculated. Each dot represents one third-instar nymphal *L. striatellus*. The mean and SD were calculated from two independent experiments.