The major constituents of the venom gland of a braconid endoparasitoid, *Meteorus pulchricornis* (Hymenoptera: Braconidae).

Kakeru Yokoi*[1], Takeshi Sano*, Masahiro Suzuki, Toshiharu Tanaka[2], Chieka Minakuchi and Ken Miura**

Applied Entomology Laboratory, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8601, Japan

*, These two authors contributed equally to this work.

**, Corresponding author

Telephone: +81-52-789-4033

Fax: +81-52-789-4032

Email: k-miura@agr.nagoya-u.ac.jp

1, Present address

The National Agriculture and Food Research Organization (NARO) 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan

2, Present address

International Cooperation Center for Agricultural Education, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8601, Japan

Keywords

Endoparasitoids, virus like particle, *Meteorus pulchricornis, Mythimna separata*, RNA interference.

Abstract

The endoparasitoid *Meteorus pulchricornis* (Wesmael) introduces virus-like particles (MpVLP) and venom in its lepidopteran host larvae upon oviposition. These wasp-associated factors are considered to help impairing some host functions including immune defense. To obtain the major protein repertoire composing the MpVLP and venom that might support successful parasitism, we constructed a conventional cDNA library of the venom gland filament that produces both the MpVLP and venom, and sequenced cDNA clones arbitrarily in the 5' regions. The 5' ESTs obtained from 473 independent cDNA clones were grouped into 228 clusters. One hundred and five clusters were annotated with one or more GO terms by Blast2GO analyses. While the major repertoire of *M. pulchricornis* venom gland shared some constituents with those of the venom glands from other parasitoid wasp species, it was distinct from those encoded by PDVs. Twenty clusters selected for further analyses were fully sequenced and characterized again. Among them, 17 factors were subjected to quantitative RT-PCR analyses, and 12 factors were suggested to be highly adult-specific. Among the 17 factors, 10 were tested as RNA interference targets, and effective gene silencing was observed for five factors. The five factors included two components of MpVLP. The MpVLP deficient in the two factors prepared from the knockdown wasps was less effective than the wild-type MpVLP in terms of the inhibitory activity against host hemocyte spreading estimated in vitro.

Introduction

Parasitoids are known to occur in several insect taxonomical groups. They are especially common in Hymenoptera, and recent surveys suggest that the parasitoid wasps occupy 10 to 20% of all insect species (Pennacchio and Strand, 2006). The prosperity of this particular insect group may reflect their highly diversified strategies that are tailored to respective parasitoid species-host species configurations for successful parasitism. Parasitoids are generally divided into two distinct classes, namely idiobionts and koinobionts. The idiobionts halt host development upon their parasitism, and they are either ectoparasitoids or endoparasitoids. The koinobionts, whose host can develop even after the parasitism, include mainly endoparasitoids.

The parasitism of endoparasitoid wasps alters the physiological conditions of their host so that the host hemocoel provides a suitable environment for proper development of the wasp larvae (Vinson and Iwantsch, 1980). These alterations include the suppression of host immune defense, which is quite important for parasitoid eggs and larvae that would otherwise be excluded through encapsulation reactions exerted by host hemocytes (Fang et al., 2010; Lavine and Strand, 2002; Wang et al., 2013). Endoparasitoid wasps are considered to utilize several means either singly or in combination for this host regulation, such as venom, calyx fluid, polydnaviruses (PDVs), non-PDV-type viruses, as well as virus-like particles (VLPs) (Webb and Luckhart, 1994). Among these, PDVs associated with ichneumonid and braconid wasps are well-studied, partly because conventional methods of virology, such as sequencing of whole viral genomes, are applicable to these researches. The PDVs are injected into the host hemocoel upon oviposition and subsequently express encoded genes in several cell types, which eventually results in changes in host physiology (Kroemer and Webb,

2004). To date, genomic sequences of several PDV species have been reported (Desjardins et al., 2008).

Meanwhile, delineation of the repertoire of venom gland constituents has also been conducted recently, by employing distinct approaches from those for PDV studies. Crawford et al. (2008) have constructed the venom gland expressed sequence tag (EST) databases of parasitoid wasps *Microctonus hyperodae* (Loan) and *M. aethiopoides* (Loan), both of which lack PDVs or VLPs, and characterized respective ESTs in combination with proteomics approaches. More recently, Vincent et al. (2010) have reported the major venom components of another parasitoid wasp *Chelonus inanitus* (Linnaeus), which bears PDV, through the combination of random sequencing of cDNA clones and mass spectrometry analyses of venom proteins. Related reports have also been published thanks to the advances in omics analysis techniques (Burke and Strand, 2014; Wu et al., 2013). Details of venom proteins from endoparasitoid wasps are discussed in recent reviews by Asgari and Rivers (2011) and by Colinet et al. (2013).

The endoparasitoid, *Meteorus pulchricornis* (Wesmael) used in this study lacks PDVs and introduces VLP (MpVLP) and venom into its host while laying eggs (Suzuki and Tanaka, 2006). MpVLP literally exhibits virus-like morphology but lacks DNA or RNA. It is produced in a pair of the venom gland filament of early adults, and the venom reservoir becomes filled with a turbid fluid containing both MpVLP and venom by day 4 post adult emergence (Suzuki and Tanaka, 2006). MpVLP inhibits host hemocyte adhesion and spreading in vitro, and this is eventually followed by the induction of host hemocyte apoptosis while the venom seems not to be so active for these effects (Suzuki et al., 2008; Suzuki and Tanaka, 2006). The molecular mechanisms upholding these host hemocyte modulations are unknown at present. In addition, biochemical properties of MpVLP or venom components largely remain to be

determined.

In the present study, to obtain the major repertoire of the MpVLP and venom, we constructed a conventional cDNA library from *M. pulchricornis* venom gland filament that produces both the MpVLP and venom, and conducted random sequencing of cDNA clones. Individual sequences were clustered, and the sequence groups as well as the remaining singletons were subjected to in silico analyses. Some sequence groups were selected and examined further for expression as well as targets of gene knockdown by RNA interference (RNAi). Finally, functional analyses of two MpVLP constituents were performed by utilizing RNAi.

Materials and Methods

Insects

A solitary, larval endoparasitoid, *M. pulchricornis*, which is known to have a relatively wide host range, and its experimental host in our laboratory, *Mythimna separata* (Walker) were reared as described in our previous paper (Suzuki and Tanaka, 2007). The *M. pulchricornis* venom gland filament was dissected from mixed developmental stages of late pupae and early adults, rinsed in phosphate-buffered saline (PBS), quickly frozen in liquid N_2 and stored at -80 ^{o}C until use.

cDNA library construction

Total RNA was extracted from the pooled venom gland filaments using TRIZOL regent (life technology) according to the manufacturer's instruction. The RNA was dissolved

in nuclease-free water and determined spectrophotometrically. A full length clone-enriched cDNA library was constructed by utilizing a Creator SMART cDNA Library Construction Kit (Clontech). Total venom gland filament RNA (0.27 μg) was reverse-transcribed by PrimeScript reverse transcriptase (TAKARA) at 42 °C for 60 min. The reverse transcription reaction was primed with a modified oligo-(dT) primer (5'-ATTCTAGAGGCCGAGGCGGCCGACATG-$(dT)_{30}N_{-1}N$-3'), and the reaction mixture also contained a SMART IV oligo-ribonucleotide (5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3'), which provides an extended template for reverse transcriptase when the enzyme reaches the 5' end of each mRNA molecule. The addition of the modified oligo-(dT) primer and the SMART IV oligo-ribonucleotide in the reaction gave 1st strand cDNAs universal priming sites along with two distinct *Sfi*I restriction sites. An aliquot of reverse transcription reaction mixture was then subjected to Long-Distance PCR with Advantage2 DNA polymerase (Clontech) and a pair of universal primer: 5' PCR primer, 5'-AAGCAGTGGTATCAACGCAGAGT-3'; the modified oligo-(dT) primer used to prime the 1st strand cDNA synthesis. The thermal cycling conditions used were as follows. The sample was pre-denatured at 95 °C for 1 min, and this was followed by 23 cycles of 95 °C for 15 sec and 68 °C for 6 min. After confirming appreciable cDNA amplification by gel electrophoresis, the PCR products were subjected to proteinase K digestion. The reaction mixture was then extracted twice with phenol-chloroform-isoamyl alcohol, and the cDNA ethanol-precipitated and digested with *Sfi*I. Then the cDNA fragments possessing two distinct *Sfi*I-generated overhangs on both ends were size-fractionated on a CHROMA SPIN-400 column. Each fraction was checked by agarose gel electrophoresis, and the fractions containing cDNA with appropriate sizes combined and ligated directionally into pDNR-LIB vector that was

pre-digested with *Sfi*I. The ligation mixture was used to transform *Escherichia coli* DH5α, giving a titer of 6.0 x 10^4 cfu per microliter of the mixture.

Sequencing

To obtain 5' ESTs, pairs of glycerol stocks of randomly selected independent *E. coli* transformants were prepared in a 96-well plate format, and the copy plates were sent to TAKARA BIO Dragon Genomics Center for sequencing while the master plates were stored at -80 ºC in our laboratory for future use. At Dragon Genomics Center cDNA inserts were prepared from the glycerol stocks, and the 5' region of each cDNA clone was sequenced by the dideoxy chain termination method with a universal primer, pDNR-LIB_F_YAQ04. In the cases where full sequences were determined, plasmid DNA was prepared by a QIAquick Spin Miniprep Kit (QIAGEN), and the sequencing performed with an ABI Prism Dye Terminator Cycle Sequencing Kit (Applied Biosystems) and a Model 3130 DNA sequencer (Applied Biosystems). Sequencing primers for sense strands were synthesized based on respective ESTs and used for further sequencing by primer walking. The 3' regions of cDNA inserts were sequenced using DNR2010-R located on the vector arm. Primers used for sequencing are listed in Table 1.

Sequence analyses

The ESTs were grouped into clusters based on their 5' nucleotide sequences. A commercial software GENETYX ver. 9 (GENETYX corporation) was used for this grouping as well as the following sequence editing. Each nucleotide sequence was

subjected to one-by-one sequence comparison against the other sequences, and sequences showing high enough similarities were grouped into the same cluster. ESTs belonging to each cluster were aligned using Clustal W 1.83 algorithm (Thompson et al., 1994). One representative EST was selected for each cluster, translated into an amino acid sequence, and preliminarily loaded onto NCBI Blastx algorism. These nucleotide sequences were later loaded to Blast2GO (Conesa et al., 2005; Gotz et al., 2008). All analyses in the Blast2GO pipeline were performed under the default settings.

For some clusters, the full-length cDNA sequences of representative clones were determined and loaded to the Blast2GO as described above. The sequences with no homologs after the Blast were further analyzed through the NCBI PSI-BLAST web server (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?CMD=Web&PAGE=Proteins&PROGRAM=blastp&RUN_PSIBLAST=on) (Altschul et al., 1997; Jones and Swindells, 2002). The analyses were carried out under the conditions with three times iteration and E-value < 0.005.

RNA extraction and quantitative RT-PCR (qRT-PCR)

Total RNA extraction, first strand cDNA synthesis, and the following qRT-PCR analyses were performed substantially as described in our previous paper (Koyama et al., 2015), except that *ribosomal protein L4* (*RPL4*) was adopted as a normalizer gene. Primers used for qRT-PCR are listed in Table 2. The EST cluster 01_E02 represents *RPL4* cDNA sequence.

RNAi

Double strand RNA (dsRNA) synthesis was conducted as described in our previous papers using a MEGAscript RNAi Kit (Ambion) (Yokoi et al., 2012a; Yokoi et al., 2012b). The dsRNA preparations were quantified spectrophotometrically, and the concentration was adjusted to 3 μg/μl in T buffer (10mM Tris-HCl, pH 8.0), and stored at -20^{o} C until use. Two hundred nanograms of dsRNAs were injected into day 2 pupae using a Nanoject II microinjector (Drummond Scientific Company). As negative controls, day 2 pupae were mock-treated by the injection of the same volumes of T buffer. These pupae were kept in a humid container at 25 ºC and allowed to develop until use. Later, total RNA was extracted from whole body of pooled animals, and qRT-PCR analyses were conducted. Sequences of T7 RNA polymerase promoter-tagged primer pairs used to prepare cDNA templates for dsRNA synthesis appear in Table 3.

Preparation of MpVLP, SDS-PAGE and N-terminal sequence determination

Day 4 wasps were dissected in ice-cold PBS, and the venom reservoirs were collected, homogenized and centrifuged at 2000 rpm for 10 min at 4 ºC in a microfuge. The supernatant was centrifuged at 15000 rpm for 30 min at 4 ºC to sediment MpVLP. The MpVLP pellet was washed once in PBS to remove venom fluid, and resuspended in 1 μl of PBS per each venom reservoir. The MpVLP preparation was separated by 12% SDS-PAGE, and the gel stained with SimplyBlue SafeStain (Invitrogen). Separated proteins on a gel were also electroblotted onto PDVF membrane, and the protein bands were stained, excised and subjected to N-terminal amino acid sequencing using a model Procise 494 HT (Applied Biosystems).

Observation of hemocytes

A hemocyte monolayer was prepared as follows. Fifth instar *M. separata* host larvae were anesthetized with CO_2, washed in 70% ethanol then in sterile water, and were placed on ice. Hemolymph from three or four larvae was collected on ice-cold Parafilm, and immediately diluted 22 times in SF900 medium (Gibco) containing 8% (v/v) of saturated phenylthiourea aqueous solution. Diluted hemolymph (30 μl) was put onto a slide-glass together with either MpVLP preparation or PBS, incubated for 30 min at room temperature and fixed with 4% paraformaldehyde in PBS for 10 min. Where MpVLP was included, one female-equivalent amount of MpVLP in 1μl of PBS was added to 30 μl hemolymph while as a negative control 1 μl of PBS was used. MpVLP was prepared from either knockdown or wild-type wasps as described in the above section. The fixed hemocyte monolayer was washed three times in PBS and subsequently permeabilized with PBS containing 0.2% (w/v) Triton X-100 for 10min. Three times washing in PBS was followed by a blocking procedure with PBS containing 2% (w/v) BSA (PBSB) for 30min at room temperature. The hemocyte monolayer was then overlaid with PBSB containing fluorescein-labeled phalloidin (0.2mg/ml, Biotium Inc.) for 20 min. After three times washing in PBS, the monolayer was observed under a microscope equipped with fluorescent optics (Olympus model BX-41).

Knockdown and in vivo parasitization experiments

Day 2 pupae were injected with dsRNA of clusters 01_D09 and 01_E04 either singly or

in combination. Control, mock-treated pupae underwent T buffer-only injection. The pupae were thereafter kept in humid containers at 25 ºC, and the emerged adult wasps were given 4th instar larvae of host *M. separata* and allowed to lay eggs at day 4 post emergence. Parasitized larvae were reared and observed afterwards, and mature wasp larvae emergence and the following pupation were recorded.

Results

cDNA library construction and quality evaluation

A cDNA library of *M. pulchricornis* venom gland filament was constructed as described in the Materials and Methods. To determine the proportion of clones bearing cDNA inserts, which represents an index of cDNA library qualities, inserts were amplified from randomly selected 14 *E. coli* transformants by PCR with a primer pair encompassing the cloning site (DNRLIB-F and DNRLIB-R in Table 1) and analyzed by agarose gel electrophoresis. All clones had cDNA inserts, the sizes of which ranged from 0.6 to 1.7 kbp (data not shown). Then, to examine whether or not the cDNAs were inserted in an expected direction into the vector plasmid, four clones were selected, and the plasmid DNA was prepared and sequenced on both strands using the primer pair described above. Sequencing results indicated that all four cDNA inserts were inserted in the expected direction (data not shown). Thus, we considered that most clones of the library had directionally inserted cDNAs. Then, we proceeded to sequencing in a 96-well plate format.

Sequencing, clustering and annotation/curation of *M. pulchricornis* venom gland

constituents

First, 5' regions of 480 independent cDNA clones were sequenced. All 480 clones had inserts with reasonable lengths (an average read was 552 nts) and proper quality, except seven clones that showed poor sequence qualities. Then, the remaining 473 ESTs were grouped through one-by-one comparison into 228 clusters, 167 of which were singletons. The largest cluster was composed of 25 ESTs. Six clusters comprised at least 10 ESTs while 14 clusters had five to nine ESTs. These clustering results are summarized in Fig. 1.

The representative ESTs from 228 clusters were preliminarily translated into amino acid sequences and manually loaded to NCBI Blastx. The results were manually checked and annotated/curated even when E-values were high. This procedure is termed 'preliminary Blast analyses (PBA)' in this study. Among the 473 ESTs (228 clusters), 38 ESTs of either rRNA or of relatively low sequence qualities were excluded, and the remaining 435 ESTs comprising 219 clusters were deposited in GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases and appear with the consecutive accession numbers of FY736475-FY736909. Then, the representative sequences of the 219 clusters were subjected to Blast2GO analyses. Among the 219 sequences, while 94 sequences had no Blast hits, 20 sequences had single or multiple Blast hits without gene ontology (GO) terms, and the remaining 105 sequences were annotated with one or more GO terms. Detailed results of Blast2GO analyses appear in supplemental table 1, and the distributions of annotated GO terms in three distinct classification criteria, namely, biological process, molecular function and cellular component, are illustrated in Fig. 2. In the biological process terms, the sum of cellular and metabolic process accounted for over a half of all terms, and this was

followed by localization and single-organism processes (Fig. 2A). The other terms were allocated to smaller numbers of, up to seven, sequences. As to the molecular function terms, binding-, catalytic activity- and structural molecule activity-related sequences occupied a great portion (Fig. 2B). Subsequently, transporter activity was assigned to seven sequences, and enzyme regulator activity and nucleic acid binding transcription factor activity only to a single sequence, respectively. Finally, the cellular component part consists largely of three terms, cell, organelle and macromolecular complex, while membrane, extracellular matrix and extracellular region were allocated only to one or two sequences (Fig. 2C).

Properties of selected sequence clusters for further analyses

Based on both PBA and Blast2GO results, we chose 20 clusters for further analyses, and the representative clones of these clusters were fully sequenced. The selected clusters were assigned to six functional groups, which we defined and might be involved in host regulation by wasp-associated factors (Table 4). The six groups of putative functions are: cell adhesion/fusion (AD); apoptosis induction/cell disruption (AP); cell motility-related (CM), immune suppression (IS); functional venom protein (VE); not defined (ND). Two clusters were assigned as AD group members. 01_F07 cluster encoded CD63 antigen homolog in this wasp species. CD63 is a member of tetraspanin superfamily (Yáñez-Mó et al., 2009). Cluster 01_D07-representative sequence was revealed to have no appreciable homologs upon Blast2GO analyses, but it showed a partial similarity to laminin gamma1 from *Anopheles gambiae* with a low hit score in PBA. Two were classified into AP group. Cluster 05_C10 encoded proteasome assembly chaperone 2, and cluster 05_D08 showed a partial homology to

caspase activation and recruitment domain (CARD) containing proteins in PBA, but with no significant hit in Blast2GO. Two clusters encoding small GTPase or related proteins were categorized into CM group, namely cytoplasmic fmr1-interacting protein (03_B02 cluster) and ras-like GTP-binding protein rho1 isoform x2 (03_D03 cluster). In addition, 01_C03 cluster, which showed a moderate similarity to rabaptin in PBA but with no hit in Blast2GO, was added to this CM group. Three clusters belonged to IS category. 01_A02 cluster showed a partial similarity to macroglobulin/complement-like protein in PBA. This cluster was the largest of all, being composed of 25 ESTs. PBA showed that 01_B04 cluster and 01_B05 cluster were partially related to serine protease inhibitors (serpins) with low hit scores. Four clusters represented VE category. 02_E05 cluster was hyaluronidase. 02_D09 cluster encoded a protein showing a partial similarity to chitinase, while 04_A10 cluster was annotated with a GO term, bvpp41b protein, which possesses metalloendopeptidase activity. 01_D09 cluster, which was composed of the second largest numbers (16) of ESTs, showed a partial homology to hemolysin in PBA, and revealed to be a component of MpVLP by N-terminal amino acid sequencing of the mature protein (Fig. 3A). We analyzed six clusters from ND category because ESTs of these clusters were frequently found in the venom gland cDNA library. Some of these clusters encoded proteins showing a little similarity to known proteins in our PBA as indicated in Table 4. Of note is that 01_E04 cluster, which was composed of 5 ESTs, was also found to be an MpVLP component (Fig. 3B).

The full-length nucleotide sequences of the 20 factors were determined and deposited in GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases with the consecutive accession numbers of AB701637-AB701656. The full nucleotide sequences were loaded to Blast2GO analysis. The results are shown in supplemental

table 2. Six sequences had multiple Blast hits with GO terms, and the results did not differ much from those of corresponding EST clusters. The other 14 sequences had no hits and were subsequently loaded to PSI-BLAST analyses. PSI-BLAST analysis showed that 01_D07 was similar to excinuclease ABC subunit C, and 01_B04 was similar to antitrypsin-like serine protease inhibitor (supplemental table 2) while the other sequences had no hits.

Expression analyses

Seventeen factors out of 20, the checkboxes of which in Table 4 are marked, were subjected to qRT-PCR analyses. The analyses were performed to estimate the relevance to adult-specific functions of factors. RNA was prepared from a pool of whole body of day 2 pupae whose venom glands were quite small in volume and indistinguishable from other fluffy tissues under a dissecting microscope. RNA was also extracted from a pool of whole body of day 0 adults (6 days post pupation) when the venom glands develop well but the reservoirs still appear to be vacant. The mRNA levels of respective factors were determined at the two distinct developmental stages, and the adult versus pupal ratios of mRNA levels were calculated (Table 5). While the mRNA levels of all factors tested were found to increase during the pupal-adult development, the degree of increment differed greatly among factors. Five factors showed only a shallow increase: 01_F07 (CD63 antigen); 05_D08 (CARD containing protein); 03_B02 (cytoplasmic fmr1-interacting protein); 03_D03 (Ras-like GTP-binding protein rho1 isoform x2); 04_A10 (bvpp41b protein). These factors exhibited relatively high mRNA levels even in day 2 pupae and showed only a marginal increase during this developmental period, suggesting that these factors were not venom

gland-specific. The other 12 factors (marked with ** or * in Table 5) showed a massive increase, ranging from 250- to 25000-fold, during the pupa-adult transition, suggesting that they were adult-specific and might be associated with the venom gland functions. For these factors, the mRNA levels at day 2 pupal stage were quite low, thus resulting in higher adult/pupal ratios.

Gene knockdown by RNAi

Among 17 factors described above, ten factors were selected for gene knockdown experiments (Table 4), based on the following reasons. We did not exclude factors showing only modest increases during the pupal-adult transition for this test: 01_F07 (CD63 antigen); 05_D08 (CARD containing protein); 04_A10 (bvpp41b protein). CD63 antigen was selected because its frequency in the EST database was relatively high (3 ESTs) and because we supposed its role in MpVLP-host cell fusion although the localization of the protein was uncertain. The factor related to CARD was chosen for its suggestive apoptosis-inducing potential that might match the reported nature of MpVLP (Suzuki and Tanaka, 2006) whereas localization of this protein in MpVLP has not been established. Bvpp41b protein, which is related to ADAM metallopeptidase (Primakoff and Myles, 2000) was selected since we considered it might shred some receptor proteins located on the surfaces of immune-related host cells. As for factors with adult-specific natures, proteasome assembly chaperone 2 (05_C10) was selected by the same reason for the CARD containing protein. Rabaptin-related protein (01_C03) was chosen among CM category members because of its adult-specific expression profiles. Three clusters, macroglobulin/complement-like protein (01_A02), the factor partially related to matrilin (01_E06) and one factor that did not shown any similarities

to known proteins (01_B10), were selected since these factors constituted the top three largest clusters. The remaining two were clusters revealed to encode the constituents of MpVLP through N-terminal amino acid sequencing, namely hemolysin-like (01_D09) and another factor with no homology to known proteins (01_E04).

Day 2 pupae were injected with dsRNA of the ten factors described above, and the mRNA levels at 96 h post injection (day 0 adult) were compared to those of mock-treated controls. The results are shown in Table 6. dsRNA-mediated gene silencing was effective for a half of factors tested, namely rabaptin-related (01_C03), hemolysin-like MpVLP component (01_D09), matrilin-related (01_E06) and two factors without any known homologs (01_B10 and 01_E04), the latter of which is an MpVLP component. The RNAi did not work effectively for the remaining five factors. The results suggest that this wasp species is somewhat amenable to gene silencing by dsRNA injection. We did not pursue further the reasons for unsuccessful RNAi for some genes.

Knockdown of MpVLP components and phenotype assays

Our previous studies revealed that the MpVLP inhibits the attachment and spreading of host adhesive hemocytes during an early phase of incubation in vitro, and this ultimately leads to hemocyte apoptosis (Suzuki et al., 2008; Suzuki and Tanaka, 2006). Here, we generated the MpVLP deficient in the two components (01_D09 and 01_E04) by either single or combined gene knockdown, and performed phenotype assays focusing on the inhibitory effects of MpVLP on host hemocyte spreading.

The full nucleotide and deduced amino acid sequences of 01_D09 and 01_E04 are shown in Fig. 3A and 3B. The ORF of 01_D09 encoded a polypeptide of 198

amino acid residues with a calculated molecular mass of 21, 641 daltons. The ORF of 01_E04 encoded a polypeptide of 181 amino acid residues with a molecular mass of 20, 006 daltons. These two polypeptides shared some physicochemical properties: both were rich in charged residues such as K, D and R; both were rich in S and G residues near the C-termini; both had alkaline isoelectric points; both had hydrophobic stretches near the N-termini. N-terminal amino acid sequencing of the mature polypeptides in MpVLP revealed that the two polypeptides were processed to smaller forms, with molecular masses of 11, 139 (for 01_D09) and 10, 951 (for 02_E04) daltons. Interestingly, the both mature polypeptides are likely to be generated by the proteolytic cleavage after QKR sequence, suggesting an occurrence of common processing enzymes. dsRNA treatment of either 01_D09 or 01_E04 effectively reduced the targeted mRNA levels (Fig. 3C). We also confirmed the effects of single as well as double knockdown of these genes at the protein level (Fig. 3D). The electrophoretic mobility of the two polypeptides appeared to be smaller than expected from the molecular masses of mature polypeptide moieties. This may suggest the occurrence of O-linked glycosylation since the both are rich in S residues.

Thus, after confirming that effective knockdown occurred, we proceeded to the characterization of MpVLP deficient in 01_D09 and 01_E04. MpVLP was prepared from day 4 knockdown adults along with control, wild-type ones, and the effects on the spreading of host larvae hemocytes, namely plasmatocytes and granular cells, were examined during an early phase of incubation (Fig. 4). PBS-treated, control plasmatocytes and granular cells spread well, and the former exhibited a conspicuous protrusion of both lamellipodia and filopodia. Upon exposure to wild-type MpVLP, the spreading of lamellipodia of both hemocyte species was clearly impaired, resulting in smaller shapes in both cases. The projection of plasmatocyte filopodia appeared to

be still maintained after the exposure to wild-type MpVLP, resulting in spiniferous. When the MpVLP deficient in both 01_D09 and 01_E04 were examined, the repressing effects on the lamellipodia of plasmatocytes were attenuated, but those found for granular cells remained unchanged. Given these results, we considered that these two MpVLP factors were involved in the inhibition of plasmatocyte spreading but not in that of granular cells. Single gene knockdown experiments were also conducted, and the results were that 01_D09 single knockdown seemed to result in only marginal recovering effects on plasmatocyte spreading (data not shown). We did not examine long-term effects of the two factors on host hemocytes, such as apoptosis induction, in this study.

The effects of 01_D09 and 01_E04 knockdown were also examined by in vivo parasitization experiments. Knockdown wasps of either 01_D09, 01_E04 or both were allowed to lay eggs into the host larvae, and the status of parasitism was followed until the emergence of mature wasp larvae from the hosts. The results are summarized in Fig. 5, showing that the knockdown, even the double knockdown, did not affect the rates of successful parasitism significantly.

Discussion

In the present study, we constructed a small-scale EST database of the *M. pulchricornis* venom gland filament, and characterized the respective EST cluster through annotation/curation by PBA, Blast2GO and PSI-BLAST. The results of GO term distribution in the EST database showed that the major transcripts in *M. pulchricornis* venom gland had cellular- or metabolic-process related functions in ‘biological process’, binding, catalytic activity and structural molecule activity in ‘molecular function’, and

cell, organelle macromolecular complex in ‘cellular component’. The endoparasitoid wasps grow in host larvae by utilizing host nutritious resources (Asgari and Rivers, 2011). The components of *M. pulchricornis* venom gland that were predicted to work in metabolic process might contribute to the alternation of host metabolism and the conversion of host nutritious resource into those parasitoid larvae can utilize, or just a reflection of the process functioning in this particular wasp tissue. On the other hand, the endoparasitoid wasps in host larvae should circumvent the attack of the host immune reactions. Encapsulation is one of the representative cellular immune reaction of insects against foreign bodied that are unable to be engulfed by a hemocyte, such as parasitoid wasp eggs or parasitic nematodes (Lavine and Strand, 2002). Multiple hemocyte types function coordinately in this immune reaction, and thus some cellular processes should be needed for the proper capsule formation. The mRNAs in *M. pulchricornis* venom gland annotated as function related to cellular processes might disturb the host cellular signals, leading to the inhibition of encapsulation reaction.

Clusters 01_B04 (supplemental table 2) and 04_E12 (supplemental table 1) were shown to have some similarities to serpin and serine protease stubble, respectively, by PSI-BLAST and Blast2GO. One of the major humoral immune responses in arthropods including insects is melanization by phenoloxidases (Cerenius et al., 2008; Cerenius and Soderhall, 2004; Yokoi et al., 2015; Yu and Kanost, 2004; Zou et al., 2010). The invasion of foreign bodies activates the extracellular serine protease cascade composed of several serine proteases and serpins, and eventually leads to the melanin deposition on the surfaces of invaders. These serine protease/serpin-related proteins may function in repressing the host melanization responses. A serine protease like venom protein that inhibits melanization of host hemolymph has been reported in *Cotesia rubecula* (Marshall) (Asgari et al., 2003). Moreover, proteins annotated as

serine proteases or serpins were found in venoms of other parasitoid species (Colinet et al., 2013). The results by others and us suggest that the modulation of host melanization responses by serine proteases and serpins are conserved between many parasitoid wasp species.

Our EST database included some putative toxic factors. 01_F04 cluster gene was annotated as venom acid phosphatase acph-1-like (supplemental table 1). In *Apis cerana* (Fabricius), the orthologous gene was expressed in a venom grand specific-manner, and the proteins showed venom acid phosphatase activities (Kim and Jin, 2014). Venom phosphatase proteins were also found in some other parasitoid wasps (Colinet et al., 2013). This implies that parasitoid wasps including *M. pulchricornis* use the phosphatase proteins as a toxic factor against host larvae.

02_E05 clusters encoded hyaluronidase. Hyaluronidase is found as a component of *Apis mellifera* (Linnaeus) venom (King and Wittkowski, 2011). Hyaluronidase degrades the major glycosaminoglycan hyaluronan, and is involved in several physiological/pathological processes. To date, hyaluronidase gene has not been found in the other parasitoid wasp transcriptomes, implying its putative functions may be *M. pulchricornis*-specific. The hyaluronidase of this species might have a role as an erosive molecule because it showed a high-level expression in adults.

The cluster 01_F07 encoded CD63 antigen, which is a member of tetraspanin family adhesion molecules (Berditchevski and Odintsova, 2007). Tetraspanin proteins have been reported to regulate cell morphology, motility, invasion and fusion. For example, murine tetraspanin variant CD9 is shown to be responsible for sperm-oocyte fusion (Kaji et al., 2000). We had speculated that the wasp tetraspanin was located in envelope-like structure of MpVLP and had a role in MpVLP-target cell fusion. However, subsequent analyses demonstrated that *M. pulchricornis* tetraspanin was not

highly adult-specific, and RNAi was not successful for this gene. Tetraspanins are also found in the venom gland major constituents of parasitoid wasps *M. hyperodae* and *M. aethiopoides* (Crawford et al., 2008).

The singleton cluster 05_C10 encodes proteasome assembly chaperone 2 protein, which is a TNF-related protein. TNF is a cytokine, and by binding to its cognate membrane receptor, it induces apoptosis of target cells (Nagata, 1997). This *M. pulchricornis* factor also has a similarity to clast3 protein. Clast3 was first cloned from activated murine B cells, and has been shown to induce growth retardation, polyploidy and multinucleation when overexpressed in transfected culture cells (Bahar et al., 2002). The wasp counterpart may function as a death inducer or disruptor of proper cell cycle progression in host tissues. The expression of this factor was highly adult-specific, while the RNAi did not work well. In addition, our EST library contained two more factors possessing putative apoptosis-related functions. 05_A07 cluster (supplemental table 1) encoded death associated protein and 05_D08 cluster encoded a protein partially homologous to CARD containing protein. These factors may also interact with host hemocytes to induce apoptosis.

The present study revealed that the venom grand repertoire contained appreciable numbers of small GTPases or related proteins, namely Ras-related Rab-40c (01_A09 cluster, in supplemental table 1), cytoplasmic fmr1-interacting protein (03_B02 cluster), Ras-related Rab-11a (05_F12 cluster, in supplemental table 1) and Ras-like GTP-binding protein rho1 isoform x2 (03_D03 cluster). These factors are also found in venoms of other parasitoid wasps (Colinet et al., 2013), and may function in dysregulating the cytoskeleton of host cells. In fact, Rac-1-deficient *Drosophila* fails to encapsulate foreign materials (Williams et al., 2006). Among them, the mRNA amount of cytoplasmic fmr1-interacting protein (03_B02) and Ras-like GTP-binding

protein rho1 isoform x2 (03_D03) was measured. However, these genes were not expressed in an adult-specific fashion.

We note here that the major constituents of *M. pulchricornis* venom gland are apparently different from major factors encoded by PDVs, such as protein tyrosine phosphatases, IκB-like inhibitors of NFκB transcription factors as well as C-type lectins (reviewed in Desjardins et al., 2008). Meanwhile, as mentioned above, it is also notable that *M. pulchricornis* venom gland repertoire and those reported for non-VLP-bearing parasitoid wasps (Crawford et al., 2008; Colinet et al., 2013; Vincent et al., 2010) share a fraction of putative erosive enzymes or protease inhibitors, which are presumably venom constituents. The venom of *M. pulchricornis* does not give a significant impact on host hemocyte spreading and apoptosis induction (Suzuki et al., 2008; Suzuki and Tanaka, 2006). Taken together, it might be reasonable to consider that the venom constituents shared by *M. pulchricornis* and other non-VLP bearing wasps may play a role in host regulation other than the modulation of host hemocytes.

We found in the present study that the two MpVLP components were involved in regulating host hemocyte behaviors in vitro. These two proteins share some physicochemical characteristics as discussed above, and we propose that these two factors are somewhat functionally redundant and act in an additive manner, since only the double knockdown resulted in a well recognizable phenotype. On the other hand, contrary to the in vitro results, the knockdown of the two proteins did not affect the rates of successful parasitism by the wasps in vivo. This may suggest that these two proteins are not indispensable for the establishment of parasitism between *M. pulchricornis* an *M. separata*, while we do not exclude the possibility that the proteins could help parasitism under less favorable conditions, for example, when parasitizing other host species.

In closing, the functions of most factors found in this study have yet to be determined. More comprehensive studies that utilize both genetic- and proteomics-based approaches are to be needed for better understanding of host regulation measures employed by this parasitoid wasp species.

Acknowledgements

We thank Drs. H. Mori, Y. Nakamatsu and O. Namba for their helpful technical advice. This work was supported in part by JSPS KAKENHI Grant Numbers 23658047 and 25450486 to KM.

References

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402

Asgari S, River DB (2011) Venom proteins from endoparasitoid wasps and their role in host-parasite interactions. Annu Rev Entomol 56: 313-335

Asgari S, Zhang G, Zareie R, Schmidt O. (2003) A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph. Insect Biochem Mol Biol 33: 1017-1024

Berditchevski F, Odintsova E (2007) Tetraspanins as regulators of protein trafficking. Traffic 8: 89-96

Burke GR, Strand MR (2014). Systematic analysis of a wasp parasitism arsenal. Mol Ecol 23: 890-+

Cerenius L, Lee BL, Soderhall K (2008) The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol 29: 263-271.

Cerenius L, Soderhall K (2004) The prophenoloxidase-activating system in invertebrates. Immunol Rev 198: 116-126

Colinet D, Mathe-Hubert H, Allemand R, Gatti J-L, Poirie M (2013) Variability of venom components in immune suppressive parasitoid wasps: From a phylogenetic to a population approach. J Insect Physiol 59: 205-212

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676

Crawford AM, Brauning R, Smolenski G, Ferguson C, Barton D, Wheeler TT, McCulloch A (2008) The constituents of *Microctonus* sp. parasitoid venoms. Insect Mol Biol 17: 313-324

Desjardins CA, Gundersen-Rindal DE, Hostetler JB, Tallon LJ, Fadrosh DW, Fuester RW, Pedroni MJ, Haas BJ, Schatz MC, Jones KM, *et al.* (2008) Comparative genomics of mutualistic viruses of Glyptapanteles parasitic wasps. Genome Biol 9: R183

Fang Q, Wang L, Zhu J, Li Y, Song Q, Stanley DW, Akhtar Z-r. Ye G (2010) Expression of immune-response genes in lepidopteran host is suppressed by venom from an endoparasitoid, *Pteromalus puparum*. Bmc Genomics 11:

Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36: 3420-3435

Jones DT, Swindells MB (2002) Getting the most from PSI-BLAST. Trends Biochem Sci 27: 161-164

Kaji K, Oda S, Shikano T, Ohnuki T, Uematsu Y, Sakagami J, Tada N, Miyazaki S, Kudo A (2000) The gamete fusion process is defective in eggs of Cd9-deficient mice. Nat Genet 24: 279-282

Kim BY, Jin BR (2014) Molecular characterization of a venom acid phosphatase Acph-1-like protein from the Asiatic honeybee *Apis cerana*. J Asia-Pacific Entomol 17: 695-700

King TP, Wittkowski KM (2011) Hyaluronidase and hyaluronan in insect venom allergy. Int Archs of Allergy Immunol 156: 205-211

Koyama H, Kato D, Minakuchi C, Tanaka T, Yokoi K, Miura K (2015) Peptidoglycan recognition protein genes and their roles in the innate immune pathways of the red flour beetle, *Tribolium castaneum*. J Invertebr Pathol 132: 86-100

Kroemer JA, Webb BA (2004) Polydnavirus genes and genomes: emerging gene families and new insights into polydnavirus replication. Annu Rev Entomol 49: 431-456

Lavine MD, Strand MR (2002) Insect hemocytes and their role in immunity. Insect Biochem Mol Biol 32: 1295-1309

Nagata S. (1997) Apoptosis by death factor. Cell 88: 355-365.

Pennacchio F, Strand MR (2006) Evolution of developmental strategies in parasitic hymenoptera. Annu Rev Entomol 51: 233-258

Primakoff P, Myles DG (2000) The ADAM gene family: surface proteins with adhesion and protease activity. Trends Genet 16: 83-87

Suzuki M, Miura K, Tanaka T (2008) The virus-like particles of a braconid endoparasitoid wasp, *Meteorus pulchricornis*, inhibit hemocyte spreading in its noctuid host, *Pseudaletia separata*. J Insect Physiol 54: 1015-1022

Suzuki M, Tanaka T (2006) Virus-like particles in venom of *Meteorus pulchricornis* induce host hemocyte apoptosis. J Insect Physiol 52: 602-613

Suzuki, M, Tanaka T (2007) Development of *Meteorus pulchricorni*s and regulation of its noctuid host, *Pseudaletia separata*. J Insect Physiol 53: 1072-1078

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680

Vincent B, Kaeslin M, Roth T, Heller M, Poulain J, Cousserans F, Schaller J. Poirié M, Lanzrein B, Drezen JM, *et al.* (2010). The venom composition of the parasitic wasp

Chelonus inanitus resolved by combined expressed sequence tags analysis and proteomic approach. BMC Genomics 11: 693

Vinson S, Iwantsch G (1980) Host Regulation by insect parasitoids. Quart Rev Biol 55: 143-165

Wang L, Fang Q, Qian C, Wang F, Yu X-Q, Ye G. (2013) Inhibition of host cell encapsulation through inhibiting immune gene expression by the parasitic wasp venom calreticulin. Insect Biochem Mol Biol 43, 936-946

Webb BA, Luckhart S (1994) Evidence for an early immunosuppressive role for related *Campoletis sonorensis* venom and ovarian proteins in *Heliothis virescen*s. Arch Insect Biochem Physiol 26: 147-163

Williams MJ, Wiklund ML, Wikman S, Hultmark D (2006) Rac1 signalling in the *Drosophila* larval cellular immune response. J Cell Sci 119: 2015-2024

Wu SF, Sun FD, Qi YX, Yao Y, Fang Q, Huang J, Stanley D, Ye GY (2013) Parasitization by *Cotesia chilonis* influences gene expression in fatbody and hemocytes of *Chilo suppressalis*. Plos One 8:

Yokoi K, Hayakawa Y, Kato D, Minakuchi C, Tanaka T, Ochiai M, Kamiya K, Miura K (2015). Prophenoloxidase genes and antimicrobial host defense of the model beetle, *Tribolium castaneum*. J Invertebr Pathol 132: 190-200

Yáñez-Mó M, Barreiro O, Gordon-Alonso M, Sala-Valdés M, Sánchez-Madrid F (2009) Tetraspanin-enriched microdomains: a functional unit in cell plasma membranes. Trends Cell Biol: 19 434-446

Yokoi K, Koyama H, Ito W, Minakuchi C, Tanaka T, Miura K (2012a) Involvement of NF-κB transcription factors in antimicrobial peptide gene induction in the red flour beetle, *Tribolium castaneum*. Dev Comp Immunol 38: 342-351

Yokoi K, Koyama H, Minakuchi C, Tanaka T, Miura K (2012b) Antimicrobial peptide gene induction, involvement of Toll and IMD pathways and defense against bacteria in the red flour beetle, *Tribolium castaneum*. Results Immunol 2: 72-82

Yu XQ, Kanost MR (2004) Immulectin-2, a pattern recognition receptor that stimulates hemocyte encapsulation and melanization in the tobacco hornworm, *Manduca sexta*. Dev Comp Immunol 28: 891-900

Zou Z, Shin SW, Alvarez KS, Kokoza V, Raikhell AS (2010). Distinct Melanization Pathways in the Mosquito *Aedes aegypti*. Immunity 32: 41-53

Figure legends

Fig. 1. Summary of clustering *M. pulchricornis* venom gland ESTs based on nucleotide sequences. 5' regions of randomly selected 480 independent cDNA clones were sequenced, and clustered except seven ESTs qualified as poor sequence quality. The clusters are categorized based on composing EST numbers, from 1 to 25. Numerals associated with portions of the pie chart represent cluster numbers composing respective categories while numerals in parentheses are total numbers of ESTs belonging to those categories. Areas of pieces in the pie chart are proportional to cluster numbers.

Fig. 2. Distributions of GO terms of representative venom gland EST sequences. The distribution in terms of biological process (A), molecular function (B) and cellular component level terms (C) are shown in each pie chart. Numbers in parentheses are those of sequences assigned to respective terms. Note that the sum of numbers appearing in parentheses in (A) and (C) are larger than that of sequences given GO terms (105) since some sequences are assigned to multiple terms. Detailed results of Blast2GO are shown in supplemental table 1.

Fig. 3. Sequences of 01_D09 and 01_E04 clones and effects of dsRNA injection at the mRNA and protein levels. Full nucleotide and predicted amino acid sequences are shown for 01_D09 (A) and 01_E04 (B). Thin underlines indicate the positions of primers used for qRT-PCR, while thicker underlines show those used for the preparation of T7 RNA polymerase promoter sequence-tagged cDNA templates. Broken lines represent the N-terminal amino acid sequences determined using the corresponding mature proteins. The dsRNAs were injected into day 2 pupae, and gene knockdown

confirmed at both the mRNA and protein levels. Mock-treated pupae (T buffer-only injected) served as controls. Determination of mRNA levels was done for 01_D09, 01_E04 dsRNA-, or mock-treated animals at 96 h post injection (day 0 adult) (C). The determination was performed with four to six animals for each knockdown category and independently repeated three times. Bars represent means ± S.D. For the test at the protein level, either the single or combined knockdown of the two genes was conducted. MpVLP was prepared from pools (around 12 animals) of dsRNA-injected day 4 wasps along with mock-treated ones, and the protein patterns compared by 12% SDS-PAGE (D). Arrowheads indicate the positions of repressed protein bands. The positions of molecular weight makers are shown on the right.

Fig. 4. Impact of MpVLP on host adhesive hemocytes in vitro and effects of knockdown of its components encoded by 01_D09 and 01_E04. The MpVLPs prepared from either knockdown or wild-type wasps were examined for the ability to inhibit host hemocyte spreading. Hemocytes were allowed to spread in the presence of wild-type (middle panels) or double knockdown MpVLP (lower panels). As a negative control, PBS was included instead of the MpVLP suspension (upper panels). Hemocytes were stained for F-actin with fluorescein-labeled phalloidin, observed under a fluorescent microscope and photographed. Three pictures are shown for each category. Plasmatocytes are marked with asterisks while granular cells with dots. Hemocytes with ambiguous morphology are left unmarked. Horizontal bars represent 50 μm.

Fig. 5. Effects of MpVLP component knockdown on successful parasitism of progeny. Host larvae parasitized either by knockdown or mock-treated wasps were reared, and

the success of parasitism was estimated by the emergence of mature wasp larvae and the following pupation. Experiments were done using around 17 host larvae for each knockdown category or controls, and repeated three times independently. Each bar represents a mean ± S.D. of successful parasitism (%).

Figure

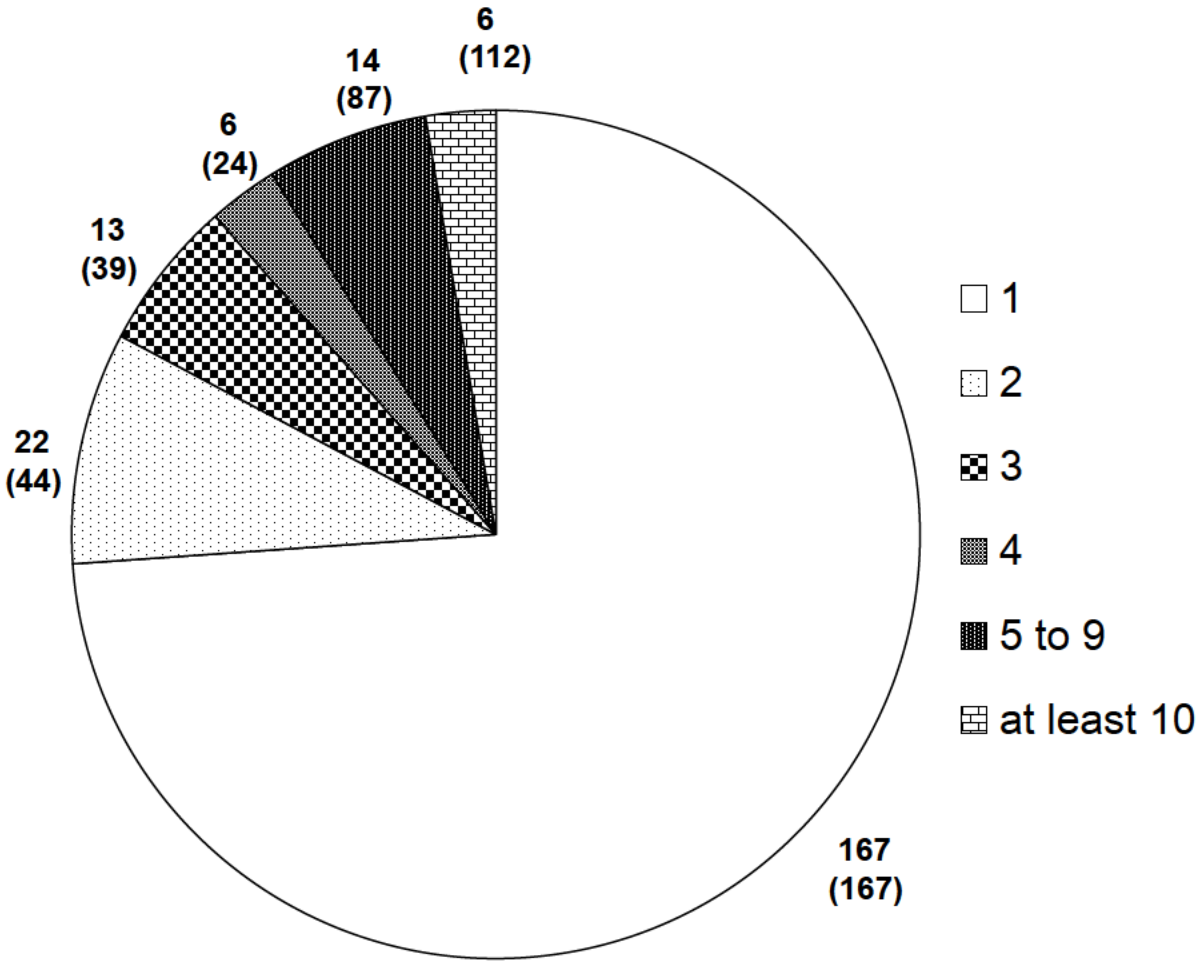

Fig.1 Yokoi et al.,

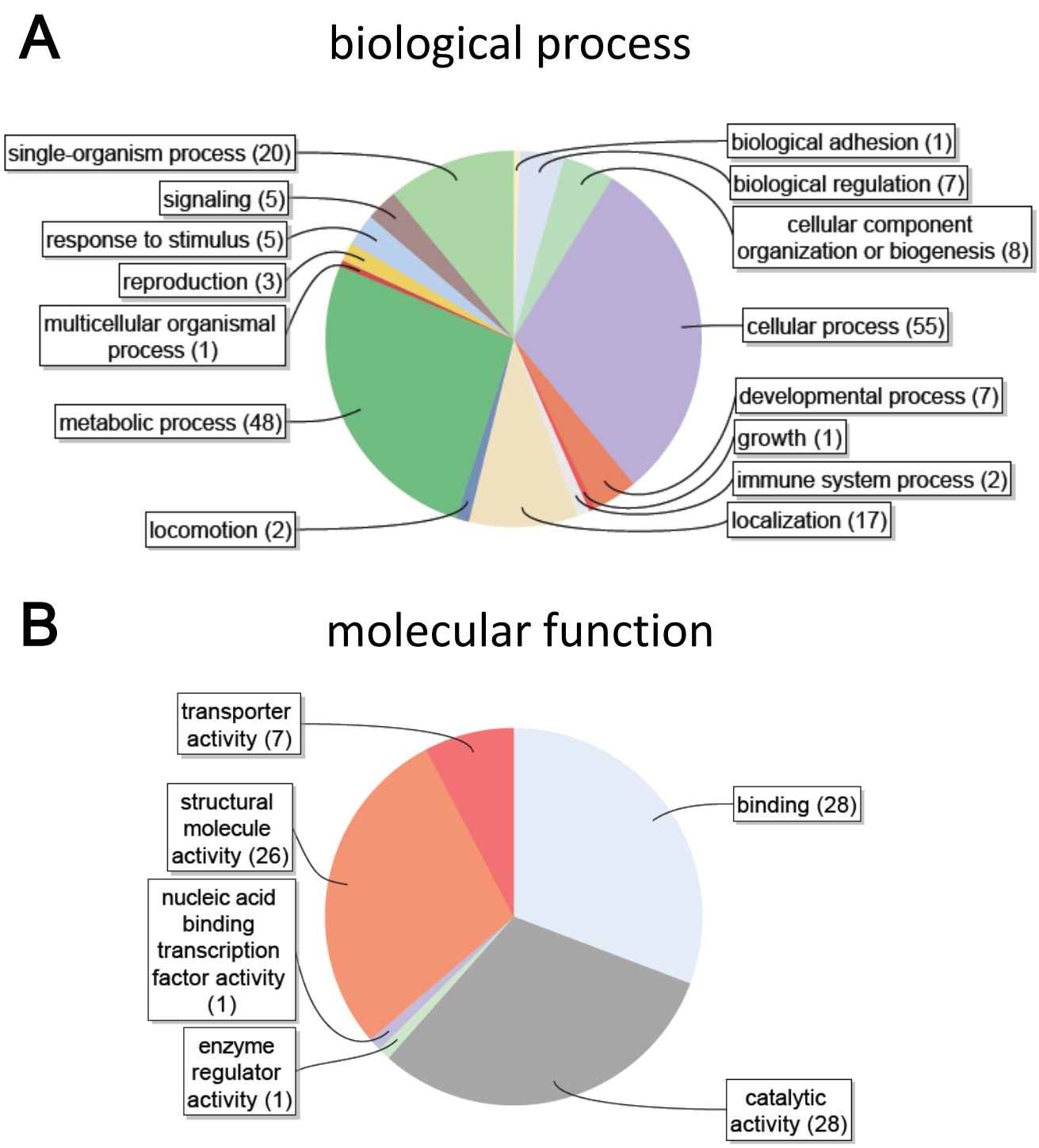

Fig. 2 Yokoi et al.,

C

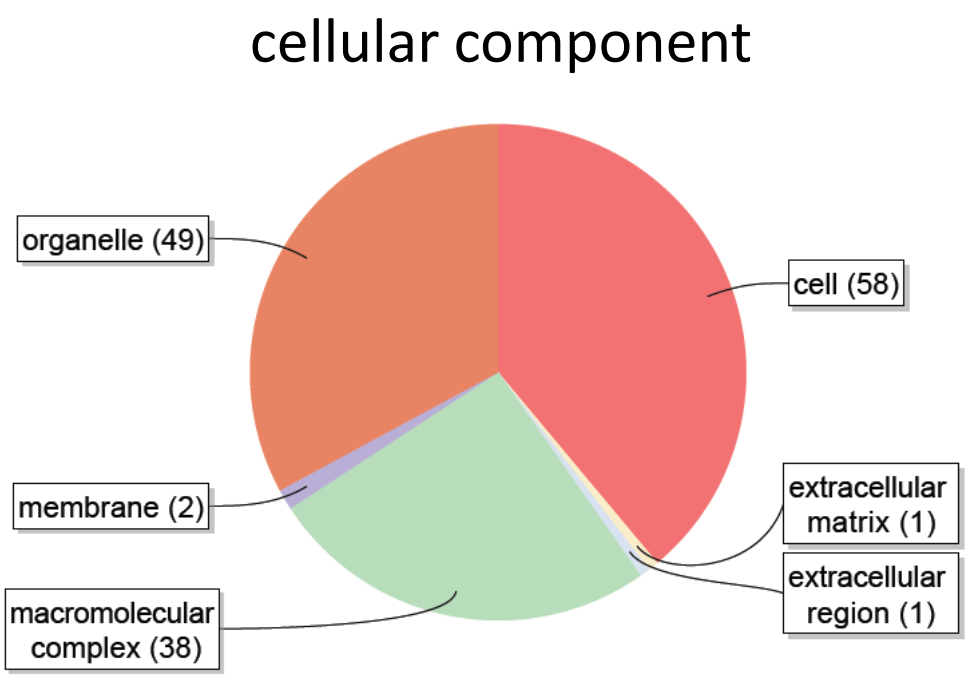

Fig.2 Yokoi et al., (continued)

A

```
  1  GACTATTCCTCAGAAATTTGTGAGTCCAATTTGAAAATATTCTCAGACCTGAAAAATGGC     60
                                                          M  A

 61  GATCAACATCAAAACGTCATTAGTAATTCTACTGGGGTTGACATGGATCTCATCCGTAAT    120
      I  N  I  K  T  S  L  V  I  L  L  G  L  T  W  I  S  S  V  I

121  AAACGCGAGTCCAGTCGAAGAGTTGAAGACTGAAGTAAATCCCAACGCTGAAGCTGCACT    180
      N  A  S  P  V  E  E  L  K  T  E  V  N  P  N  A  E  A  A  L

181  GCCAGTGAAGACCAAGAGAGACGCATCACCGTTAGATGCAGATTTCCCTCTTCAACTGGT    240
      P  V  K  T  K  R  D  A  S  P  L  D  A  D  F  P  L  Q  L  V

241  AAACAGCAAACTCATGAAATCAATCAGAGATCTCTCAGCATTAGGACTTTTTGATGGCGC    300
      N  S  K  L  M  K  S  I  R  D  L  S  A  L  G  L  F  D  G  A

301  ATTGGAAAATATCATTGGTCAAGTGACACATAATATACAAAAACGTGGTACACGATCACA    360
      L  E  N  I  I  G  Q  V  T  H  N  I  Q  K  R  G  T  R  S  H

361  CGCATCTCGAAAGTTTCAAAATTTCCAACGCAAATGGGGTGGAAGAGCGAAGAAGCTGTC    420
      A  S  R  K  F  Q  N  F  Q  R  K  W  G  G  R  A  K  K  L  S

421  CACACCCTGGAGACGTCAAAAACCGAAAACTCAATACAATCATGATGGTAGCGTAGATGC    480
      T  P  W  R  R  Q  K  P  K  T  Q  Y  N  H  D  G  S  V  D  A

481  AGATGATGATATTTCATTGAGGTATCATCCATCAAGAGCTGGCTCGCTCGAATCGCTCGG    540
      D  D  D  I  S  L  R  Y  H  P  S  R  A  G  S  L  E  S  L  G

541  AACTGATTCAACGGGTGGGTACAATTCGGGTGGCTATGATTCCAACTCTGATGCGGATTC    600
      T  D  S  T  G  G  Y  N  S  G  G  Y  D  S  N  S  D  A  D  S

601  TATCGCTGGTAATAGCCTCGATTCTAGTGACTATTCATCTAGTTTTTGGTGAATAGCTTG    660
      I  A  G  N  S  L  D  S  S  D  Y  S  S  S  F  W  *

661  GAAAGAGATTTATAACTTTCAATCCCTGATGATTGGAATGAATATCTCAACTTGCATCAA    720
721  ATCAGAATCATGAAATTGCACATAAAATATTCATCACTGATAAGTCTCATGAATATGAAA    780
781  TAAAAGAATACAAAATATCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA    830
```

Fig.3 Yokoi et al.,

B

```
  1  GATATTTTTCAAGAAACTTCGATCTGAACTCGCAGATACAAAAGTTGTTATAAAAATGGA     60
                                                        M  D

 61  TTTCAAGACAGTTGTTATACTGTCGGCATTTGTTATTGTGTCCATCGGGGCATCTGCAAA    120
      F  K  T  V  V  I  L  S  A  F  V  I  V  S  I  G  A  S  A  N

121  TCCTCCGCAAGAAAAAGCTCCTGAAAGCAATTTGAATTCAGTAGACACCGCAGCAGTAAA    180
      P  P  Q  E  K  A  P  E  S  N  L  N  S  V  D  T  A  A  V  K

181  AACTAAGAGAGAGTCTGCACTTCCAGAGAAAGATCAGGAGATCCAACAAGAAAATAAAGA    240
      T  K  R  E  S  A  L  P  E  K  D  Q  E  I  Q  Q  E  N  K  E

241  ATTAGAACATGACGATGGTATTCCACTGCATCTTCAAGCTGTACATGAAATACAAAAAAG    300
      L  E  H  D  D  G  I  P  L  H  L  Q  A  V  H  E  I  Q  K  R

301  AGCCTTCAAAAATTTCAAATTGAATCGGCTGAGTGGAAGACTAAGGAAGGCGCAAAATAA    360
      A  F  K  N  F  K  L  N  R  L  S  G  R  L  R  K  A  Q  N  K

361  GCTGCGAAGTATAGTTAAAAAATTGCGCAAAGCACAGGATGACGTTGACAGATATACACT    420
      L  R  S  I  V  K  K  L  R  K  A  Q  D  D  V  D  R  Y  T  L

421  CTTACAGACAAAAGCTCAAGCAACAGTGGATAAATACAAACGGAAAATAAATCAGTCTGG    480
      L  Q  T  K  A  Q  A  T  V  D  K  Y  K  R  K  I  N  Q  S  G

481  ATCGAAAAGTCGATCTCGGTCTTCTGGATCGAATCGTGGCGATGACGGATATGGATCGGA    540
      S  K  S  R  S  R  S  S  G  S  N  R  G  D  D  G  Y  G  S  D

541  TTCAGGCAGTGGAGGATATGGAGGCGGTGAACCTTCGGCACCATATAATCAAGATTATTA    600
      S  G  S  G  G  Y  G  G  G  E  P  S  A  P  Y  N  Q  D  Y  *

601  GTTGAATTAGTGTGAAAGGAACGATGATTATTCTGCAGAGGGAAAGAGGACCCTCTTCTG    660
661  TGAATAGACAATACCACTGATGACCGAATGTGTCCACGAATACGAATAGATTTCTCAGGT    720
721  CAGCACTACTAAATAAAACTTTTCCAAATATTCCATCTTCAATTATTGTGTCACTTTTAT    780
781  TCTTATCATTGATAACATGCCTAAAAAATTGAAGAAATATCTGGATATGTATAAATAAAA    840
841  ATTATTTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA    879
```

Fig.3 Yokoi et al., (continued)

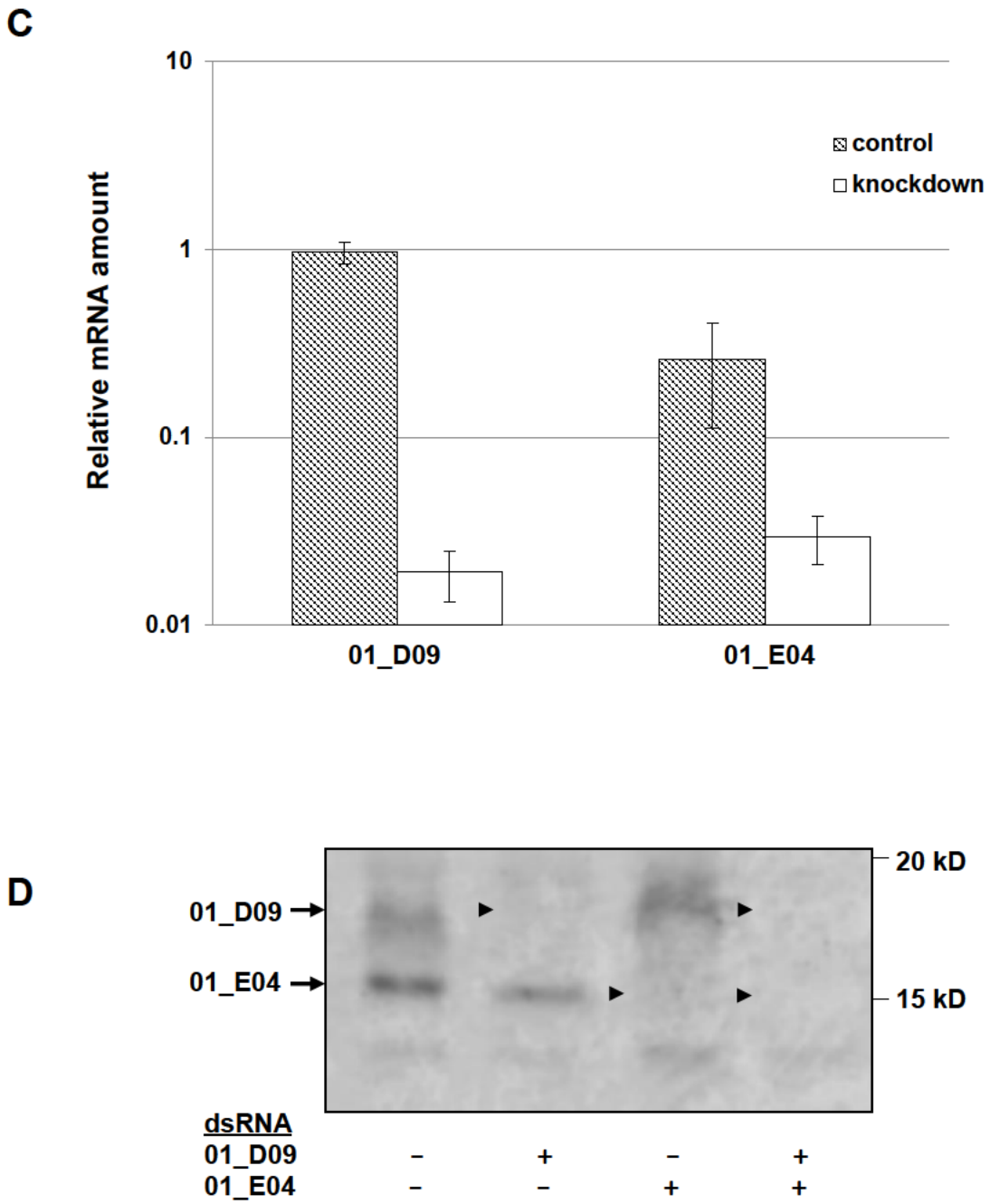

Fig.3 Yokoi et al., (continued)

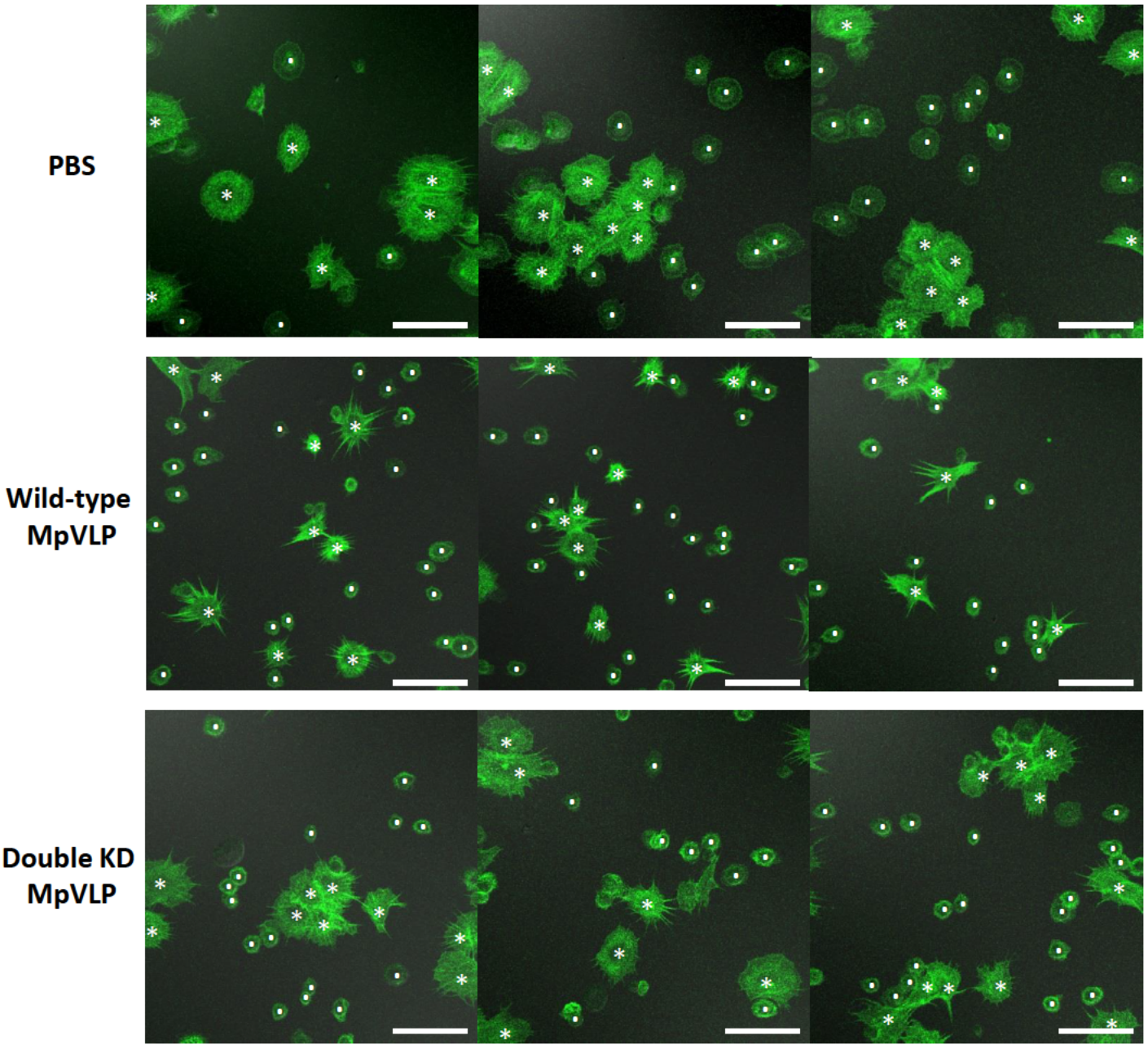

Fig.4 Yokoi et al.,

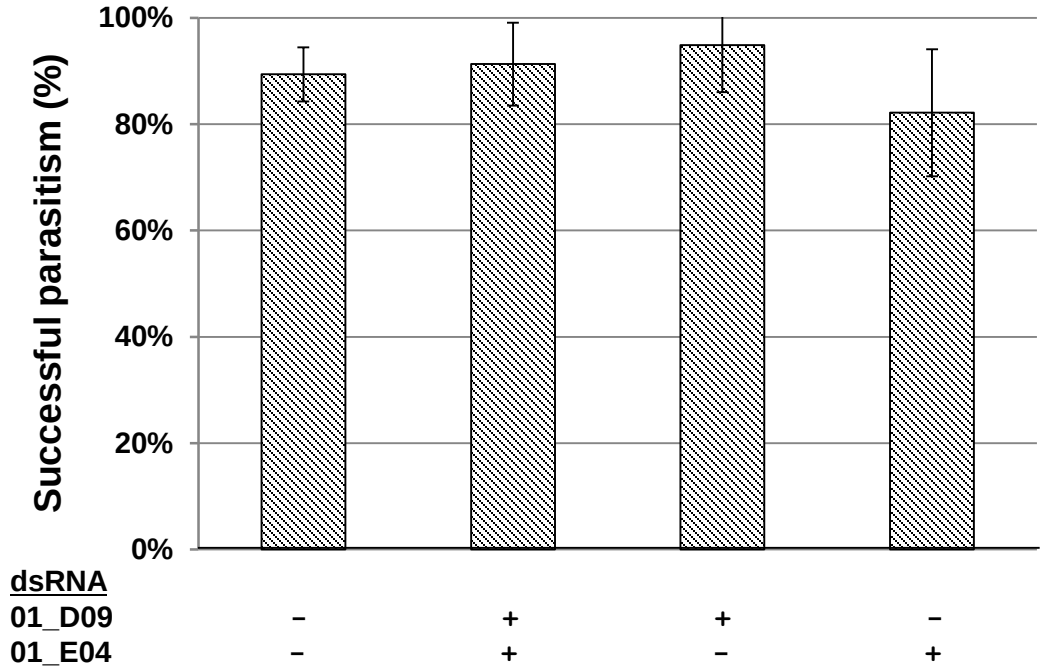

Fig.5 Yokoi et al., (continued)

Table.1 Primers used for sequence determination

Primer name	Sequence (5' to 3')
01_B04-WF1	CAATTGCAGCGATGGCTTT
02_A04-WF1	GAAGTGGTTCACCTCAAG
02_G09-WF1	GCATCAGATGCCCTCAT
02_A06-WF1	TGCAAAGGCAGCAGAAGA
03_A06-WF1	AGTCTAACAGGCGCAACT
03_A09-WF1	GATGACGTTGACAGATATACAC
03_B02-WF1	CGCCAATCCATCCATCAT
03_D03-WF1	TGAAACAGGAGCCAGTTAA
03_E02-WF1	CCTCACCGATCCATCATT
04_B10-WF1	CTTCATATGCCCTCGTTAA
05_F05-WF1	TTAAGGGTCCAAAGGTGAA
pDNR-LIB_F_YAQ04	GCATACATTATACGAAGTTATCAGTCG
DNRLIB-F	AGCGAGTCAGTGAGCGAGGAA
DNRLIB-R	TAGGGAAACAGCTATGACCATGT
DNR2010-R	TGGTGAGAATCCAAGCACTAGT

Primers used for further sequencing of respective 5’ ESTs by primer walking were synthesized based on determined sequences. The first six digits of the primer names indicate target ESTs. pDNR-LIB_F_YAQ04 was initially used to obtain 480 independent 5’ ESTs. DNRLIB-F and DNRLIB-R were used to amplify cDNA inserts. DNR2010-R was used to sequence 3’ regions of inserts.

Table 2. Gene-specific primers used for qRT-PCR

EST	Forward (5' to 3')	Reverse (5' to 3')
01_C03	AAATCCATATGCTGCGCCATCA	TATCAGCCGCATCGTGGAACA
01_E02*	GGTCACCAAACATCGGCTGAA	CCACCGCGACACATATTACCA
02_A06	GGAAGTAATGCAATGGCTAGTG	ACACCAGAGCTGCGCTCATAA
02_D09	TGACTCAAGGGATAGCAATGAG	ACAGACGCGTATCTTCTTCAAG
02_F07	CAGGAAACGACTGTAGAACCTT	GCTCCAACATCTAGGTTCATGT
02_G09	TCGATGCGGATTTCGCAGGAA	CATCTGATGCGTCTCGATGATA
03_A09	TCAGTAGACACCGCAGCAGTA	TGCAGTGGAATACCATCGTCAT
03_B02	GTTGGACCACCACCAGCTAAT	CACCAAACAATTCCTCGACAGT
03_D01	CATCGACAATCTGCCAAATCCT	CCGATCCTTGGTTAATGTCTTC
03_D03	TGAACTCGCGCTTTGGGATAC	TCTAGTGAATCGGGACTGTCAA
03_G06	TGGCATTCACCTTCATGTCACT	CCGGATAGCAGAATCCGTCAA
03_H10	GTGTCACTGGAGACCAATGCA	ATGGGTTGTTTCCAGCAGATTC
04_A07	AAGCTGCACTGCCAGTGAAGA	CCTAATGCTGAGAGATCTCTGA
04_A10	GCTCAGAGAGTTGCTTGAAATG	GTTTAGGGACATCCAGTGATCT
05_C10	GCAGTAGGAAATGTTGGACAAC	AGGATCTGCACCAAGAATAGGA
05_D08	GCCAAGCAGGTGCTACAGAAA	GATTCGACCCACTGACATAGAA
05_D10	GATTCGTGCCGCATATGATGAT	CAACATCCAAAGAACGCGATGA
05_G01	CGATGTAGTAAGTGACGTGCAT	GTGCTTTAACAGCCTCGTCTGAT

*, 01_E02 encodes ribosomal protein L4 and was used as a normalizer of mRNA quantitation.

Table 3. Sequence of primers for synthesizing T7 promoter-tagged cDNA

Target EST	Forward (5' to 3')	Reverse (5' to 3')
04_A07	TTGACATGGATCTCATCCGTAAT	CACTAGAATCGAGGCTATTACC
03_A09	TCGATCTGAACTCGCAGATACA	ACACATTCGGTCATCAGTGGTA
04_A10	AAGACAGGCATGGTCACTTGTT	CATGGTTGGGCTGAAGCACAA
05_D10	GCAGTTCGCGAGAATTATTGCA	CGATGAACGTAGTTGAGACTCT
05_C10	ATTGTGCACTGCCAGCGATGT	CAATGGCGCTGGATTACCAAAT
05_D08	GGAGACACTGAAGGAATTCCTA	CGGTTCAGGCAAACAATCAAGA
02_F07	CAACAGGACTTGCGGCAAGTT	TACGAGGAAGATGTGGATGCAT
03_G06	TTGAGCAAGTTGGACAGACAAG	CTCGCCGATGAAGCAGAACTT
03_H10	GATACACAGAAGCGCAAGTTCA	GGTGGTCCCATCGGTTAAAGT
01_C03	TTCGCTCCACTCCGACTGAAT	CTCAGGTTCTGGATCGCTGTT

Only gene-specific portions of sequences are shown. T7 promoter sequence (5'-TAATACGACTCACTATAGGG-3') is indeed on the 5' halves.

Table.4. Clusters selected for further analysis.

PF	Cluster	Blast2GO top hit (PBA results)	EST nos. in cluster	Representative EST accession number	qRT-PCR	RNAi
AD	01_D07	no homology (laminin)	10	FY736890	✔	
	01_F07	CD63 antigen	3	FY736866	✔	✔
AP	05_C10	proteasome assembly chaperone 2	1	FY736854	✔	✔
	05_D08	no homology (CARD containing protein)	1	FY736864	✔	✔
CM	01_C03	no homology (rabaptin)	3	FY736498	✔	✔
	03_B02	cytoplasmic fmr1-interacting protein	1	FY736661	✔	
	03_D03	Ras-like GTP-binding protein rho1 isoform x2	1	FY736684	✔	
IS	01_A02	no homology (macroglobulin/complement-like protein)	25	FY736619	✔	✔
	01_B04	no homology (serine protease inhibitor)	1	FY736488		
	01_B05	no homology (serine protease inhibitor)	1	FY736489		
VE	01_D09[§]	no homology (hemolysin-like)	16	FY736740	✔	✔
	02_D09	no homology (chitinase)	5	FY736597	✔	
	02_E05	hyaluronidase	4	FY736632	✔	
	04_A10	bvpp41b protein	1	FY736743	✔	✔
ND	01_B03	no homology (methyl-accepting chemotaxis protein)	9	FY736682	✔	
	01_B10	no homology	21	FY736731	✔	✔
	01_D06	no homology (iron transporter)	8	FY736563	✔	
	01_E04[§]	no homology	5	FY736656	✔	✔
	01_E06	no homology (matrilin)	22	FY736716	✔	✔
	02_A04	no homology (ferredoxin)	7	FY736561		

Sequence clusters were chosen from respective (cell adhesion/fusion (AD), apoptosis induction/cell disruption (AP), cell motility-related (CM), immune suppression (IS), functional venom protein (VE) and not defined (ND)) group of putative function (PF) regarding host regulation. Six clusters from the category ND are selected accordingly to their cluster size. Annotations are based on Blast2GO results, and if any, direct homologs are shown. In the cases that Blast2Go did not give any hits, annotations/curations of our PBA (preliminary Blast analyses) are shown in

parentheses, if any. §, ESTs that encode MpVLP components. Clusters were examined by qRT-PCR or RNAi where check-marked.

Table.5. Changes of mRNA levels during pupal-adult development.

Cluster	PF	Blast2GO result (PBA result)	Day 2 pupa	Day 0 adult	A/P
01_D07	AD	no homology (laminin)	0.00021	0.33	1524.83**
01_F07	AD	CD63 antigen	0.069	0.17	2.42
05_C10	AP	proteasome assembly chaperone 2	3.11E-06	0.078	24901.5**
05_D08	AP	no homology (CARD containing protein)	0.017	0.064	3.74
01_C03	CM	no homology (rabaptin)	7.40E-05	0.58	7783.81**
03_B02	CM	cytoplasmic fmr1-interacting protein	0.014	0.026	1.85
03_D03	CM	ras-like gtp-binding protein rho1 isoform x2	0.45	1.22	2.71
01_A02	IS	No homology (macroglobulin/complement-like protein)	0.0021	0.53	249.65*
01_D09§	VE	no homology (hemolysin-like)	0.0034	0.97	283.45*
02_D09	VE	no homology (chitinase)	0.00034	0.21	637.7*
02_E05	VE	hyaluronidase	4.70E-05	0.34	7267.76**
04_A10	VE	bvpp41b protein	0.01	0.053	5.15
01_B03	ND	no homology	0.00034	0.41	1234.93**
01_B10	ND	no homology	0.00077	3.54	4613.28**
01_D06	ND	no homology (iron transporter)	0.0001	0.033	319.55*
01_E04§	ND	no homology	0.00052	0.26	500.65*
01_E06	ND	no homology (matrilin)	8.50E-05	1.4	16477.67**

mRNA levels were determined relative to that of *RPL4* in day 2 pupae as well as in day 0 adults (corresponds to day 6 post pupation) by qRT-PCR, and the ratios of adult/pupa (A/P) calculated. §, ESTs encoding MpVLP components. *, 100- to 1000-fold increment. **, over 1000-fold increment.

Table.6. Degree of target gene knockdown by RNAi.

Cluster	PF	Blast2GO result (PBA result)	Relative target mRNA amount after RNAi
01_F07	AD	CD63 antigen	1.599
05_C10	AP	proteasome assembly chaperone 2	1.145
05_D08	AP	no homology (CARD containing protein)	1.283
01_C03	CM	no homology (rabaptin)	0.145*
01_A02	IS	no homology (macroglobulin/complement-like protein)	1.044
01_D09§	VE	no homology (hemolysin-like)	0.019*
04_A10	VE	bvpp41b protein	0.831
01_B10	ND	no homology	0.026*
01_E04§	ND	no homology	0.112*
01_E06	ND	no homology (matrilin)	0.025*

The degree of gene knockdown was estimated by calculating the dsRNA-treated/mock-treated ratios of mRNA levels, and the values shown in the rightmost column. §, ESTs encoding MpVLP components. Potent levels of knockdown are marked with asterisks.