JVI Accepts, published online ahead of print on 22 August 2007
J. Virol. doi:10.1128/JVI.01006-07

White spot syndrome virus proteins and differentially expressed host proteins identified in shrimp epithelium by shotgun proteomics and cleavable isotope-coded affinity tag[§]

Jinlu Wu*, Qingsong Lin*, Teck Kwang Lim, Tiefei Liu, and Choy-Leong Hew[‡]

Address:
Department of Biological Sciences, National University of Singapore, Singapore 117543

Running Title:
Proteomic analysis of WSSV and cellular proteins

Corresponding author:
[‡]Professor Choy-Leong Hew
Department of Biological Sciences
National University of Singapore
Singapore 117543
Tel: 65-6516-2692
Fax: 65-6779-5671
E-mail: dbshead@nus.edu.sg

Word count:
Abstract: 232 words
Text: 4,945 words

* These authors contributed equally to this work.
§ Raw MS data is accessible *via* hyperlink: ftp://dbswjl:dbswjl@137.132.35.227/

ABSTRACT

Shrimp subcuticular epithelial cells are the initial and major targets of white spot syndrome virus (WSSV) infection. Proteomic studies of WSSV-infected subcuticular epithelium of *Peneaus monodon* were performed through two approaches: subcellular fractionation coupled with shotgun proteomics to identify viral and host proteins, and a time-course quantitative proteomic analysis using cleavable isotope-coded affinity tags (cICAT) to identify differentially expressed cellular proteins. Peptides were analyzed by off-line coupling of two-dimensional liquid chromatography with MALDI-TOF/TOF mass spectrometry. We identified 27, 20 and 4 WSSV proteins from cytosolic, nuclear and membrane fractions, respectively. 28 unique WSSV proteins with high confidence (total ion C.I. %>95) were observed, of which 11 were reported here for the first time, and 3 of these novel proteins were shown to be viral non-structural proteins by western blotting analysis. A first shrimp protein dataset containing 1999 peptides (ion score $\geqq$ 20) and 429 proteins (total ion score C.I. %>95) was constructed *via* shotgun proteomics. We also identified 10 down-regulated proteins and 2 up-regulated proteins from the shrimp epithelial lysate *via* cICAT analysis. This is the first comprehensive study of WSSV-infected epithelia by proteomics. These 11 novel viral proteins represent the latest addition to our knowledge of the WSSV proteome. Three proteomics datasets consisting of WSSV proteins, epithelial cellular

proteins and differentially expressed cellular proteins generated in the course of WSSV infection provide a new resource to further study WSSV-shrimp interactions.

INTRODUCTION

White spot syndrome virus (WSSV) is a catastrophic pathogen of cultured penaeid shrimps since its first appearance in the early 1990's (32). The initial and major target of this virus is shrimp epithelia, including subcuticular, stomach and gill. WSSV-infected epithelial cells show hypertrophied nuclei containing massive amounts of viruses (26). Genomic studies revealed that the virus consists of a dsDNA of about 300kbp with more than 180 predicted ORFs (9, 43, 54). So far, a majority of proteins encoded by predicted ORFs have not been detected and functions of many of these presumptive proteins remain elusive. Information on virus-host interactions is therefore very limited.

Proteomics has been demonstrated to be an important platform technology and has contributed to our understanding of virus-host interaction (4, 37). Shotgun 2D-LC-MS^2 (two dimensional-liquid chromatography-tandem mass spectrometry) is a promising approach for high throughput identification of proteins (21, 48). Cleavable isotope-coded affinity tags (cICAT) coupled with 2D-LC-MS^2 enables the pair-wise comparison of protein expression levels in uninfected and infected cells quantitatively (7, 15). Previous proteomic studies on WSSV had identified more than 40 viral structural proteins (19, 25, 41, 56), of which 33 were designated as

envelope proteins (25, 53). However, our knowledge of viral non-structural proteins and host cellular response during WSSV infection remains poor. To date, only a few non-structural proteins, which are encoded by highly conserved gene sequences, such as DNA polymerase (9), ribonucleotide reductase (27) and others (14, 16, 18, 27, 28, 47) have been confirmed by traditional gene cloning and immunoassays. Differential expression of host proteins was mainly investigated in the mRNA level using cDNA microarray and expressed sequence tag (EST) (11, 12, 17, 36, 39, 44). Only one investigation on the protein expression profile of stomach of WSSV-infected *Litopenaeus vannamei* using 2D gel electrophoresis and mass spectrometry has been reported (45). In the present study, we explored WSSV proteins and differentially expressed cellular proteins from WSSV-infected epithelium using shotgun and cICAT proteomics. We identified 28 viral proteins, containing 11 novel viral proteins, 3 of which were confirmed to be non-structural proteins. We also identified 10 down-regulated and 2 up-regulated cellular proteins. Their potential roles in virus-host interactions are discussed.

MATERIALS AND METHODS

Shrimp, virus and challenge. Virus inocula were prepared from stored hemolymph of WSSV-infected shrimp, and intramuscularly injected into black tiger shrimp (*Penaeus monodon,* body weight 10-15 g), while PBS vehicle was injected in parallel as controls (52). Challenge

dose was optimized to ensure 100% mortality within 5 days of infection. The virus infection was confirmed by PCR using a pair of specific primers for WSSV.

Abundance of viral proteins in different tissues. The stomach, gill and epithelium of cephalothorax subcuticule were sampled from moribund individuals 3 days post-infection (dpi) for the extraction of whole tissue lysates. Briefly, 0.5 g of each sample were homogenized with 1 ml pre-cooled lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris base, 2 mM tributylphosphine), and incubated on ice for 10 min. After centrifugation at 20,000 *g* for 20 min at 4°C, the supernatant was treated using a 2-D cleanup kit (GE Health-care), followed by dissolving in the denaturing buffer (0.1% SDS, 50 mM Tris pH 8.5) to keep the proteins in the same condition as in shotgun and cICAT analyses. The protein concentration was determined by RC DC protein assay (Bio-Rad) using bovine γ-globulin as a standard. Equal amounts of proteins (30 μg) were resolved by SDS-PAGE and the relative abundance of viral proteins was analyzed by western blotting.

Sample preparation for shotgun proteomic analysis. Cephalothorax subcuticular epithelium was sampled at 3 dpi. Cytosolic, membrane and nuclear fractions were sequentially isolated using a Qproteome Cell Compartment Kit (Qiagen) with the following procedures. 0.5 g of tissue were washed twice with ice-cold PBS and then homogenized with Buffer CE1. After incubation on ice for 10 min, the lysate was centrifuged at 1000 *g* for 20 min at 4°C. The supernatant

(cytosolic fraction) was transferred and stored on ice. Other two fractions were extracted using procedures described in the kit manual. Three fractions were then cleaned, re-dissolved in the denaturing buffer followed by the determination of protein concentrations as described previously. Fractionation efficiency was examined by western blotting before proceeding to trypsin digestion. 400 µg of proteins from each fraction were reduced with 2 mM triscarboxyethylphosphine (TCEP) and carbamidomethylated using 50 mM iodoacetamide. Porcine trypsin (Applied Biosystems) was added at an estimated enzyme-to-substrate ratio of 1:50 (wt/wt) and incubated overnight at 37°C. A strong-cation exchange column (Applied Biosystems) was used to remove SDS, trypsin and other reagents from the peptide digestion. Sep-Pak Cartridges (Waters) were used for peptide desalting.

Sample preparation for cICAT analysis. Cephalothorax subcuticular epithelia were collected at 6, 12, 24, and 72 hpi (hour post-infection) from both WSSV and PBS vehicle injected individuals for time-course cICAT analyses. To minimize deviations arising from individuals, the sample collected at each time point was a pool from 5 individuals for both WSSV- or PBS-injected group. The workflow for the sampling and the protein preparation is outlined in Fig. 1.

Whole tissue protein was extracted and quantitated using the same procedures as described in the previous part of abundance of viral proteins in different tissues. 100 µg of proteins from the

virus- and PBS-injected samples were reduced with 2 mM TCEP, labeled with heavy and light cICAT reagents (Applied Biosystems) respectively for 2 h at 37℃ in the dark. Labeled proteins were combined and digested with 25 µg trypsin at 37℃ for 16 h, followed by cation-exchange chromatography and avidin affinity purification.

2D LC separation of peptide mixtures. Peptide mixtures for both shotgun and cICAT proteomics were separated using an Ultimate™ dual-gradient LC system (Dionex-LC Packings) equipped with a Probot™ MALDI spotting device as follows. Peptide mixture was dissolved in 98% H_2O, 2% acetonitrile (ACN) with 0.05% trifluoroacetic acid (TFA) and injected into a 0.3 × 150-mm strong cation-exchange chromatography (SCX) column (FUS-15-CP, Poros 10S) (Dionex-LC Packings) for the first dimension separation. Mobile phases A and B were 5 mM KH_2PO_4 buffer, pH 3 + 5%ACN and 5 mM KH_2PO_4 buffer, pH3 + 5% ACN+ 500 mM KCl respectively. The flow-rate for the SCX column was 6 µl/min. Nine fractions were separated by step gradients of mobile phase B (unbound, 0-5, 5-10, 10-15, 15-20, 20-30, 30-40, 40-50, 50-100%). The eluted fractions were captured alternatively onto two 0.3 × 1-mm trap column (3-µm C18 PepMapTM, 100 Å) (Dionex-LC Packings) and washed with 0.05 % TFA followed by gradient elution to a 0.2 × 50-mm reverse-phase column (Monolithic PS-DVB) (Dionex-LC Packings). The mobile phase A and B used for the second-dimension separation were 98% H_2O, 2% ACN with 0.05% TFA and 80% H_2O, 20% ACN with 0.04% TFA respectively. The

gradient elution step was 0-60% mobile phase B in 15 min at a flow-rate of 2.7µl/min. The LC fractions were mixed with Matrix-assisted laser desorption/ionization (MALDI) matrix solution (7 mg/ml α-cyano-4-hydroxycinnamic acid and 130 µg/ml ammonium citrate in 75 % ACN) at a flow-rate of 5.4 µl/min through a 25-nl mixing tee (Upchurch Scientific) before spotting onto 192-well stainless steel MALDI target plates (Applied Biosystems, Foster City, CA) using a Probot Micro Fraction collector (Dionex-LC Packings), with a speed of 5 sec per well.

Mass spectrometry. The samples were analyzed by a MALDI TOF/TOF (time-of-flight) MS^2 (ABI 4700 Proteomics Analyzer, Applied Biosystems) as previously reported (5). For cICAT sample, ICAT pairs with normalized ratio changes (normalized against median ratio of all the ICAT pairs detected) ≥ 40% and the more intense peaks with S/N (signal/noise) ≥ 30 were selected as precursor ions. Singletons with S/N≥ 50 were also selected for MS^2 analysis.

Proteomic data analysis. The MS together with MS^2 spectra were searched using GPS Explorer ™ software version 3.0 and MASCOT 2.0 search engine (Matrix Science) allowing one missed cleavage (trypsin). Precursor error tolerance and MS^2 fragment error tolerance were set to 150 ppm and 0.4 Da respectively. Only fully tryptic peptides with seven amino acids or longer and a MS^2 score of above 20 were accepted for positive peptide identification.

For protein identification by shotgun proteomics, carbamidomethyl cysteine, N-terminal acetylation and pyroglutamation (E or Q), and methionine oxidation were selected as variable

modifications. A local database containing all predicted WSSV ORFs (2013 entries from three complete WSSV genome AF440570; AF332093 and AF369029) and the IPI human database v. 3.07 (62322 entries) (http://www.ebi.ac.uk/IPI/IPIhelp.html) was used for identification of WSSV proteins. Proteins with total ion score C.I % (percent confidence interval) above 95 were considered as positive identifications. Only a few proteins with a total ion score C.I.% less than 95 were checked manually and presented in the results for reference only. Identified novel viral proteins were further confirmed by western blotting and/or RT-PCR.

An NCBInr database (4111659 entries) was used for the host protein identifications, but only those peptides for which ion scores above 46 (corresponding to total ion score C.I.%>50) were accepted. All identified cellular proteins were further analyzed using an in-house software to group them into either "distinct" proteins (without sharing peptide(s) with other proteins) or "indistinct" proteins (sharing peptide(s) with multiple proteins) to clarify possibly ambiguous identifications. Unique and shared peptides are numbered and tabulated. If peptides matched to multiple members of a protein family, default setting criteria of MASCOT 2.0 software were used for selecting the one to be reported. In addition, we randomized the same NCBInr database using a pearl script downloaded from the Matrix Science website (http://www.matrixscience.com/help/decoy_help.html). The same MS and MS^2 data were searched against the randomized database (i.e. decoy database) using the same criteria and

software. The numbers of matches, including peptide and protein matches, were counted for the calculation of false positive rates.

For the analysis of differential expression of cellular proteins by cICAT, an NCBInr database (4111659 entries) was used for the search. Heavy and light cICAT-labeled cysteine, N-terminal acetylation, pyroglutamation (E or Q) as well as methionine oxidation were selected as variable modifications. Maximum peptide rank was set to one and minimum ion score (peptide) was set to 40. cICAT quantification was performed using GPS Explorer™ software v. 3.0 and normalized against median ratio obtained from all the ICAT peptide pairs detected in one sample. Cutoff values for protein up-regulation and down-regulation were set at H/L $\geq$1.4 and H/L$\leq$0.71, respectively.

Time course analysis of 11 novel WSSV protein genes by RT-PCR. Epithelium was collected in the same way as in the cICAT analysis, but with 3 additional time points: 2, 4 and 8 hpi. Total RNAs were isolated from specimens using an RNeasy mini kit (Qiagen). Isolated RNAs (5 μg) were reverse transcribed with SuperScript™ III Reverse Transcriptase and random hexamers (Invitrogen). The first-strand cDNA products were subjected to PCRs with primer sets shown in supplemental data I. The β-actin transcript was amplified and used as an internal control for RNA quality, efficiency of first strand cDNA synthesis, and loading amount for PCRs. One known early gene *wsv477* (14), two structural protein genes *wsv209* (encoding VP187) and *wsv465*

(encoding VP124) were included for comparison.

Antibody preparation and Western blotting. Full length *wsv051*, *-076*, *-294*, and *-477 genes* were cloned into pET-32a (+) vector (Novagen Inc.), expressed in *E. coli* BL21, and protein expressions were confirmed by both DNA sequencing and MALDI-TOF mass spectrometry. Purified fusion proteins were subjected to SDS-PAGE analysis prior to immunizing rabbits for polyclonal antibodies. Polyclonal antibodies against WSSV VP9, VP26 and VP28 were generated previously (28, 40). Anti-VP664 was kindly provided by Prof C.-F. Lo, National Taiwan University. The following antibodies were purchased: anti-actin (A2066, Sigma), anti-GAPDH (FL-335, Santa Cruz Biotechnology) and anti-Histone 2A (H-124, Santa Cruz Biotechnology). Western blotting was carried out as described elsewhere (28).

Computational annotation. Topology predictions were performed using the TMHMM predictor (www.cbs.dtu.dk/services/TMHMM/), and signal peptide predictions using SignalP3.0 (www.cbs.dtu.dk/services/SignalP/). Protein families, domains and functional sites were annotated by searching InterPro database (http://www.ebi.ac.uk/InterProScan/).

RESULTS

Identification of WSSV proteins by shotgun proteomics. Initial experiments were carried out to determine which of the three tissues contained higher concentrations of viral proteins. Western

blotting analysis showed that both the structural protein (VP28) and non structural protein (VP9) were more abundant in the subcuticular epithelium and the gill than in the stomach of WSSV-infected shrimp (Fig. 2a). The epithelium was hence chosen for subsequent proteomic analysis.

To improve the identification of viral proteins, cytosolic, nuclear and membrane fractions were isolated from the epithelia. The efficiency of isolation was examined by western blotting analysis. Results showed that VP28 and VP9 had the highest concentrations in the cytosolic fraction, and lowest in the membrane fraction (Fig. 2b). The cellular protein GAPDH, specific to the cytosolic fraction, was the most abundant in the cytosol, but could also be detected in the other two fractions. On the other hand, histone 2A, specific to the nuclear fraction, was the most abundant in the nuclear fraction. These results showed that subcellular fractionation could enrich cytosolic and nuclear proteins into their respective fractions.

A total of 27, 20 and 4 viral proteins were identified from the cytosolic, nuclear and membrane fractions, respectively. In total, 33 unique proteins were identified from the three fractions. 28 of these proteins showed high scores (total ion score C.I. % >95), while remaining 5 proteins with low scores (total ion score C.I. % ≤95%) were also listed (Table 1 and supplemental data II). Out of 28 unique proteins, 11 proteins were identified for the first time in this study (shown in bold letters), Of the remaining, there were 13 structural and 4 non-structural

proteins reported previously. Of 5 proteins with low scores, wsv226 is reported for the first time, while remaining four are previously reported structural proteins. All identified proteins, either with high or low ion scores, have distinct peaks of fragment ions in MS^2 spectra as found by manual examination. A total of 8 proteins were predicted to have transmembrane helix (TMH). Four proteins have signal peptides.

A dataset of host proteins identified by shotgun proteomics. To investigate the protein composition of shrimp epithelium, MS and MS^2 spectra from shotgun proteomics of epithelial cytosolic, nuclear and membrane fractions were combined and searched against an NCBInr database. 1999 peptides were identified with ion scores $\geqq$20, of which 1669 peptides were identified with ion score$\geqq$46 (supplemental data III). We identified 429 proteins with total ion score C.I. %>95, of which 144 were grouped into "distinct" proteins (supplemental data IVa), while remaining 285 proteins were grouped into "indistinct" proteins (supplemental data IVb). Using the randomized database, we identified 95 false positive matches at the peptide level (ion score$\geqq$46) and nine false positive matches at the protein level with total ion score C.I.%>95, which corresponds to a false positive rate of 5.7% (i.e. 95 in 1669) and 2.1% (i.e. 9 in 429) for peptide and protein identifications, respectively.

A dataset of differentially expressed host proteins identified by cICAT. To understand the shrimp response to WSSV infection, subcuticular epithelia of WSSV- and PBS-injected

(uninfected control) shrimps were sampled in parallel at 6, 12, 24 and 72 hpi, and whole tissue lysates were then extracted and prepared for cICAT analysis. A total of 12 proteins showing differential expression were identified with 95% confidence level (Table 2, supplemental data V and VII). As previously noted, peptides from virus-infected group had a heavy tag (H), and peptides from uninfected group had a light tag (L). Out of the 12 proteins, 10 proteins (83.3% of differentially expressed proteins) were down-regulated (H/L≤0.71) and 2 proteins were up-regulated (H/L≥1.4). The mean and deviation of H/L ratio are shown in supplemental data VII. The expression of the clottable protein was identified to be unchanged at 6 and 12 hpi, up-regulated at 24 hpi but down-regulated at 72 hpi. Of the 10 down-regulated proteins, 1, 2, and 7 of them were identified at 12, 24 and 72 hpi respectively, showing an increasing number of down-regulated proteins with the time elapse of infection. For the 2 up-regulated proteins, up-regulation of hemocyanin was first detected at 12 hpi and lasted to 72 hpi, while up-regulation of arginine kinase was only detected at 72 hpi. Neither down-regulated nor up-regulated proteins were detected at 6 hpi.

Two peptides of tubulin group were detected at 24 hpi (Table 3, supplemental data VI and VII), one peptide was up-regulated (H/L=1.65), the other did not change significantly (H/L=0.97). All peptides of tubulin group identified at 72 hpi showed down-regulation (H/L<0.7). Same two peptides of ATPase group were detected at both 24 and 72 hpi (Table 3,

supplemental data V and VII), one was up-regulated at 24 hpi (H/L=2.28) and then down-regulated at 72 hpi (H/L=0.38), the other did not change significantly at these two time points.

Time course analysis of 11 novel WSSV protein genes by RT-PCR. To investigate temporal gene transcriptions of 11 novel viral proteins and to provide further evidence for their presence at the mRNA level, a time-course, semi-quantitative RT-PCR was performed with β-actin as an internal control and 3 other genes (*wsv209*, *-465* and *-477*) as positive controls. Out of 14 genes, 10 showed increasing mRNA levels with the time course of infection (group 1), one showed constant mRNA levels at all the time points (group2), while three (group 3) showed decreasing mRNA levels with the time course of infection (Fig. 3). On the other hand, six genes (*wsv076, -192, -143, -277, -343, -477*) commenced their expression at very early stage of the infection (2 hpi), while 2 genes (*wsv285, -294*) commenced at a later stage of the infection (24 hpi).

Western blotting. Western blotting analysis was performed to verify the presence of proteins in the virus-infected tissue and in the purified virus. Four proteins (wsv051, -076, -294, and a known non-structural viral protein wsv477 as positive control) were detected in the virus-infected whole tissue lysate and the cytosolic fraction, but were not detected in the purified virus and the uninfected cell lysate (Fig. 4). Positive controls using known viral structural proteins, including envelope protein VP28 and VP26 as well as nucleocapsid protein (VP664)

showed that these known structural proteins were present in the both tissue lysates and purified virus. The results indicated that the three novel proteins, *i.e.* wsv051, -076 and -294, might not be viral structural proteins.

DISCUSSION

The development of proteomic technologies has revolutionized our ability to analyze protein compositions of viruses and host protein changes on a global scale, but there are limitations to identify the low abundance proteins and dynamics of proteins in a single step procedure or approach (30). We used two complementary proteomic approaches, i.e. shotgun proteomics and cICAT to analyze the virus-infected cells. Coupling effective protein and peptide fractionation technologies with highly sensitive shotgun mass spectrometry enables identification of low abundance proteins, while a time-course, quantitative analysis using cICAT can reveal the dynamic alternations in the cellular protein profiles, deepening our understanding on how the cellular machinery is affected.

Comparing with the virion proteome, virus-infected cells have much more complex protein compositions. A major challenge for the identification of viral proteins from infected cells by shotgun proteomics is how to enrich the viral proteins while reducing the presence of high abundant cellular proteins (e.g. house keeping proteins) and to effectively separate the enormous

number of peptides. To enhance the viral protein identification, we undertook the following strategies: 1. to find a tissue type which contained relatively higher level of viral proteins; 2. to reduce the complexity of protein extracts by sequentially isolating three subcellular fractions; 3. to separate peptides by high performance 2D LC (48). Several commercial kits were tested to isolate cytosolic, nuclear and membrane fractions, none of which could completely eliminate the cross contamination among the three fractions (data not shown). Nonetheless, the subcellular fractionation did reduce the complexity of host cellular proteins (Fig. 2b). The proteomic analyses of these three fractions identified 28 unique viral proteins, 11 of these proteins were identified for the first time with high confidence. As these 11 proteins have never been previously identified from 1D and 2D protein gels of purified WSSV, they were tentatively assumed as candidates of non-structural proteins. Three of these 11 candidates have been confirmed to be the real viral non-structural proteins by western blotting analysis. We also tabulated the 5 proteins with low ion scores (Table 1) for reference only, as they all had the distinct MS^2 peptide fragment peaks (supplemental data II). Low ion scores might be related to the low abundance of these viral proteins in the sample, and possibly, also to other factors. Overall, this study represents an effective strategy to explore viral proteins from the infected host protein extract, showing its advantage over previous virion proteomics to explore the viral non-structural proteins.

In addition, we have carried out shotgun and iTRAQ proteomic studies of purified WSSV with the aim to assign the structural proteins into two categories, envelope and nucleocapsid proteins (25). A novel protein wsv143 reported in this study was also identified in the purified WSSV by shotgun proteomics. As shotgun proteomics is sensitive to detect proteins at low amount, and wsv143 was not identified in 1D and 2D gels of purified WSSV proteins, this protein may not be a viral structural protein but may be present in contaminated cell debris which was co-purified with WSSV particles.

A time-course RT-PCR analysis of the transcription of the 11 novel protein genes showed that they had different transcriptional profiles. Early transcribed genes might be involved in the viral replication and the modification of host cellular metabolism from the very beginning of infection (2 hpi). On the other hand, late transcribed genes might be involved in the virus assembly, maturation and release. Computational annotation of 11 proteins through InterPro database search revealed that three of these proteins have homologies with known functional domains, which are briefly stated below.

The wsv143 has homology to SOX (Sry box) proteins (protein family ID: PTHR10270), which contain a high-mobility group domain, allowing them to bind gene promoters or enhancers and function as transcription factors (49, 50). The wsv252 is the first member identified in the DUF1335 protein family at the protein level (43), and potentially may be involved in DNA

replication (23). The wsv343 matched to a P-loop containing nucleoside triphosphate hydrolases (NTPase) (SSF52540). Targeted reduction or point mutants of NTPase inhibit the *Toxoplasma gondii* proliferation (31) and viral DNA replication (20). Small molecules designed to inhibit the NTPase function may be promising anti-viral drugs (3, 8, 34, 38).

Human epithelial proteomics has provided powerful information on the relationships between biological molecules and disease mechanisms (58). Here we have generated a dataset of shrimp epithelial proteins through shotgun proteomics. This dataset is important not only for studying host-pathogen interactions but also for comparative biology. One problem encountered in generating such a dataset is that a same peptide sequence can be present in multiple different proteins or protein isoforms. Shared peptides therefore can lead to ambiguities in determining the identities of proteins (33). To avoid those ambiguities, we developed an in-house software, by which we classified identified proteins into two groups, *i.e.* "distinct" proteins containing peptides which are not shared with others and "indistinct" proteins containing peptides that are shared among multiple proteins. For those "indistinct" proteins, additional information such as molecular weight and gene sequence are needed for correct identification. This is a shortcoming of shotgun proteomics.

We further carried out a time-course cICAT analysis with the attempts to investigate the dynamic changes of host cellular proteins during WSSV infection. The combination of

chromatography and cICAT labeling can produce comprehensive information of the protein composition and provide accurate quantitation on a subset of proteins (15). The limited number of proteins identified as being differentially expressed in the present study was mainly due to the lack of sequence information of the shrimp genome. Once the genome of shrimp has been sequenced, we will perform database search again, expecting to identify many more changes in host proteins. Implications of these changes in protein abundance and presumed functional alterations in the context of virus pathogenesis will require further bioinformatics analysis and/or experimental investigations. However, we discuss here a partial selection of host proteins undergoing changes to highlight the importance for further validation and more in depth study of these proteins.

An increasing number of down-regulated proteins were detected with the time elapse of infection (0, 1, 2, 7 down-regulated proteins for 6, 12, 24, 72 hpi respectively). Farnesoic acid O-methyltransferase (FAMeT) was a down-regulated protein, which may be involved in gametogenesis, oocyte maturation, development and metamorphosis of the shrimp (13, 24). Down-regulation of FAMeT hints that WSSV infection might regulate, either directly or indirectly, host hormonal systems, growth and maturation.

Down-regulation of proapoptotic caspase adaptor protein (PACAP) was detected at 72 hpi. The PACAP has been demonstrated to exhibit specific binding to caspase-2 and -9, but not to

caspase-3, -4, -7, or -8 in 293T cells. Up-regulation of PACAP in human B cell lines triggered apoptosis (6). WSSV infection could cause a high rate of apoptosis in the shrimp lymphoid organ, but no apoptotic cells were observed in the epithelial cells in which WSSV massively replicated (51). It will be important to determine whether the down-regulation of PACAP in the epithelial cells plays a role in suppressing apoptosis and benefiting WSSV replication.

Ci-Tardbp (*Ciona intestinalis* trans-activation-responsive DNA binding protein) which may be involved in mRNA maturation and stabilization was also down-regulated. Sequence comparison with Ci-Tardbp and the longest isoforms of the human, *Drosophila* and *C. elegans* TDP43 (Tar DNA binding protein) shows that the highest degree of similarity among the four proteins resides in RRM1 and RRM2 (RNA recognition motifs, supplemental data VIII) (1, 46). TDP43 belongs to the family of heterogeneous nuclear ribonucleoproteins, characterized by binding to RNA, and in some cases DNA sequences through RRM, and participating in a variety of processes such as RNA transportation, stabilization and modification, premRNA splicing as well as transcriptional regulation (22, 46). It was demonstrated that the host cellular TDP43 could interact with the HIV Tat protein and modulate the HIV-1 gene expression (35, 50). How WSSV protein(s) interact(s) with the shrimp cellular TDP43 remains to be determined.

Down-regulation of elongation factor (EF), heat shock protein (Hsp70), tubulin and ATPase may represent defects in protein synthesis, folding and transportation. However, since peptides

identified are common in multiple isoforms of tubulin family, the cICAT study is unable to tell us as to which isoforms undergo changes in abundance. Therefore, we showed the H/L ratio of individual peptides identified, as an indication of the overall change of the tubulin family. The change of the ATPase family is shown in the same way.

WSSV infection causing up-regulation of hemocyanin has been investigated previously by transcriptional analysis (11, 12). Our cICAT results show for the first time that the protein level of hemocyanin was up-regulated from 12 hpi. Hemocyanin functions as an immunoglobulin superfamily molecule (57) and is one of the important host factors against pathogenic invasion (2). WSSV infection could induce the aggregation of hemocytes to sites of infection (42). Higher levels of hemocyanin in the epithelium may hint at an antivirus defense mechanism of the shrimp. It is also worth noting that the clottable protein was up-regulated at 24 hpi but dramatically down-regulated at 72 hpi. As the clottable protein is involved in type C coagulation and biodefense (10, 55), its up-regulation in the early stage of infection may indicate an innate immune response, but its down-regulation in the late stage of infection gives the first experimental data which may explain, at least partially, why hemolymph from shrimp at the late stage of WSSV infection does not coagulate.

In summary, we have developed a general strategy for comprehensive analysis of a large protein complex of virus-infected cells, and shown its successful application to study

WSSV-infected shrimp epithelia. We identified 11 novel viral proteins, 429 cellular proteins and 12 differentially expressed cellular proteins. Although the number of cellular proteins identified was limited due to the limitation of shrimp genome data, our research has opened a new venue to further study virus-host interactions, and provided useful data for comparison with different experimental approaches.

ACKNOWLEDGEMENTS

We thank Shashikant B. Joshi for critical reading of the manuscript and Zhuang Ying for her help on western blotting analysis. This work was supported by grants R-154-000-233-112 and R-154-000-255-112 from the Academic Research Council of the National University of Singapore to Choy Leong Hew.

REFERENCES

1. **Ayala, Y. M., S. Pantano, A. D'Ambrogio, E. Buratti, A. Brindisi, C. Marchetti, M. Romano, F. E. Baralle.** 2005. Human, Drosophila, and *C. elegans* TDP43: nucleic acid binding properties and splicing regulatory function. J. Mol. Biol. **348:**575-588.
2. **Bachère, E., E. Mialhe, D. Noël, V. Boulo, A. Morvan, and J. Rodriguez.** 1995. Knowledge and research prospects in the marine mollusk and crustacean immunology. Aquaculture **132:**17-32.
3. **Balzarini, J.** 2000. Effect of antimetabolite drugs of nucleotide metabolism on the anti-human immunodeficiency virus activity of nucleoside reverse transcriptase inhibitors. Pharmacol. Ther. **87:**175-187.
4. **Bernhard, O. K., R. J. Diefenbach, and A.L. Cunningham.** 2005. New insights into viral structure and virus-cell interactions through proteomics. Expert Rev. Proteomics. **2:**577-588.
5. **Bi, X., Q. Lin, T. W. Foo, S. Joshi, T. You, H. M. Shen, C. N. Ong, P. Y. Cheah, K. W. Eu, and C. L. Hew.** 2006. Proteomics analysis of colorectal cancer reveals alterations in metabolic pathways: mechanism of tumorigenesis. Mol. Cell Proteomics. **5:**1119-1130.
6. **Bonfoco, E., E. Li, F. Kolbinger, and N. R. Cooper**. 2001. Characterization of a novel proapoptotic caspase-2- and caspase-9-binding protein. J. Biol. Chem. **276:**29242-29250.
7. **Booy, A. T., J. D. Haddow, L. B. Ohlund, D. B. Hardie, and R. W. Olafson.** 2005.

Application of isotope coded affinity tag (ICAT) analysis for the identification of differentially expressed proteins following infection of Atlantic salmon (*Salmo salar*) with infectious hematopoietic necrosis virus (IHNV) or *Renibacterium salmoninarum* (BKD). J. Proteome Res. **4:**325- 334.

8. **Bretner, M., A. Baier, K. Kopanska, A. Najda, A. Schoof, M. Reinholz, A. Lipniacki, A. Piasek, T. Kulikowski, and P. Borowski.** 2005. Synthesis and biological activity of 1H-benzotriazole and 1H-benzimidazole analogues--inhibitors of the NTpase/helicase of HCV and of some related Flaviviridae. Antivir. Chem. Chemother. **16:**315-326.

9. **Chen, L. L., H. C. Wang, C. J. Huang, S. E. Peng, Y. G. Chen, S. J. Lin, W. Y. Chen, C. F. Dai, H. T. Yu, C. H. Wang, C. F. Lo, and G. H. Kou.** 2002. Transcriptional analysis of the DNA polymerase gene of shrimp white spot syndrome virus. Virology **301:**136-147.

10. **Chen, M. Y., K. Y. Hu, C. C. Huang, and Y. L. Song**. 2005. More than one type of transglutaminase in invertebrates? A second type of transglutaminase is involved in shrimp coagulation. Dev. Comp. Immunol. **29:**1003-1016.

11. **Dhar, A. K., A. Dettori, M. M. Roux, K. R. Klimpel, and B. Read.** 2003. Identification of differentially expressed genes in shrimp (*Penaeus stylirostris*) infected with white spot syndrome virus by cDNA microarrays. Arch Virol. **148:**2381-2396.

12. **Gross, P. S., T. C. Bartlett, C. L. Browdy, R. W. Chapman, G. W. Warr.** 2001. Immune

gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. Dev. Comp. Immunol. **25:**565–577.

13. **Gunawardene, Y. I. N. S., S. S. Tobe, W. G. Bendena, B. K. C. Chow, K. J. Yagi, and S. M. Chan.** 2002. Function and cellular localization of farnesoic acid O -methyltransferase (FAMeT) in the shrimp, *Metapenaeus ensis*. Eur. J. Biochem. **269:**3587–3595.

14. **Han, F., J. Xu, and X. Zhang.** 2007. Characterization of an early gene (wsv477) from shrimp white spot syndrome virus (WSSV). Virus Genes. **34:**193-198.

15. **Hansen, K. C., S. U. Gerold, R. J. Chalkley, J. Hirsch, M. A. Baldwin, and A. L. Burlingame.** 2003. Mass spectrometric analysis of protein mixtures at low levels using cleavable ^{13}C-Isotope-coded affinity tag and multidimensional chromatography. Mol. Cell. Proteomics. **2:**299-314.

16. **He, F., B. J. Fenner, A. K. Godwin, and J. Kwang.** 2006. White spot syndrome virus open reading frame 222 encodes a viral E3 ligase and mediates degradation of a host tumor suppressor via ubiquitination. J. Virol. **80:**3884-3892.

17. **He, N., Q. Qin, and X. Xu.** 2005. Differential profile of genes expressed in hemocytes of white spot syndrome virus-resistant shrimp (*Penaeus japonicus*) by combining suppression subtractive hybridization and differential hybridization. Antiviral Res. **66:**39-45.

18. **Hossain, M. S., S. Khadijah, and J. Kwang.** 2004. Characterization of ORF89-a latency-related gene of white spot syndrome virus. Virology **325:**106-115.

19. **Huang, C., X. Zhang, Q. Lin, X. Xu, Z. Hu, and C. L. Hew**. 2002. Proteomic analysis of shrimp white spot syndrome viral proteins and characterization of a novel envelope protein VP466. Mol. Cell. Proteomics. **1:**223-231.

20. **Kahn, J. S., and M. Esteban.** 1990. Identification of the point mutations in two vaccinia virus nucleoside triphosphate phosphohydrolase I temperature-sensitive mutants and role of this DNA-dependent ATPase enzyme in virus gene expression. Virology. **174**:459-471.

21. **Kislinger, T., B. Cox, A. Kannan, C. Chung, P. Hu, A. Ignatchenko, M. S. Scott, A. O. Gramolini, Q. Morris, M. T. Hallett, J. Rossant, T. R. Hughes, B. Frey, and A. Emili.** 2006. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell **125:**173-186.

22. **Krecic, A. M., and M. S. Swanson.** 1999. hnRNP complexes: composition, structure, and function. Curr Opin Cell Biol. **11:**363-371.

23. **Krell, P. J.** 1996. Passage effect of virus infection in insect cells. Cytotechnology. **20:**125–137.

24. **Laufer, H., D. W. Borst, F. C. Baker, M. Carrasco, M. Sinkus, C. C. Reuter, L. W. Tsai, and D. A. Schooley.** 1987. Identification of a juvenile hormone-like compound in crustacean. Science. **235:**202–205.

25. **Li, Z., Q. Lin, J. Chen, J. L. Wu, T. K. Lim, S. S. Loh, X. Tang, and C. L. Hew.** 2007. Shotgun identification of structural proteome of shrimp white spot syndrome virus and iTRAQ differentiation of envelope and nucleocapsid subproteomes. Mol. Cell Proteomics. [Epub ahead of print]

26. **Lightner, D.V.** 1996. A handbook of pathology and diagnostic procedures for diseases of penaeid shrimp. World Aquaculture Society, Baton Rouge.

27. **Lin, S. T., Y. S. Chang, H. C. Wang, H. F. Tzeng, Z. F. Chang, J. Y. Lin, C. H. Wang, C. F. Lo, and G. H. Kou.** 2002. Ribonucleotide reductase of shrimp white spot syndrome virus (WSSV): expression and enzymatic activity in a baculovirus/insect cell system and WSSV-infected shrimp. Virology **304:**282-290.

28. **Liu, Y., J. Wu, J. Song, J. Sivaraman, and C. L. Hew.** 2006. Identification of a novel nonstructural protein, VP9, from white spot syndrome virus: its structure reveals a ferredoxin fold with specific metal binding sites. J. Virol. **80:**10419-10427.

29. **Lo, W. Y., K. F. Liu, I. C. Liao, and Y. L. Song.** 2004. Cloning and molecular characterization of heat shock cognate 70 from tiger shrimp (*Penaeus monodon*). Cell Stress

Chaperones **9:**332–343.

30. **Maxwell, K. L., and L. Frappier.** 2007. Viral proteomics. Microbiol. Mol. Biol. Rev. **71:**398-411.

31. **Nakaar, V., B. U. Samuel, E. O. Ngo, and K. A. Joiner.** 1999. Targeted reduction of nucleoside triphosphate hydrolase by antisense RNA inhibits. J. Biol. Chem. **274:**5083-5087.

32. **Naylor, R. L., R. J. Goldburg, H. Mooney, M. Beveridge, J. Clay, C. Folke, N. Kautsky, J. Lubchenco, J. Primavera, and M. Williams.** 1998. ECOLOGY: Nature's 10 Subsidies to Shrimp and Salmon Farming. Science **282:**883-884.

33. **Nesvizhskii, I. A., and R. Aebersold.** 2005. Interpretation of shotgun proteomic data-the protein inference problem. Mol Cell Proteomics. **4:**1419-1440.

34. **Neyts, J., A. Meerbach, P. McKenna, and E. De Clerq.** 1996. Use of the yellow fever virus vaccine strain 17D for the study of strategies for the treatment of yellow fever virus infections. Antiviral Res. **30:**125-132.

35. **Ou, S. H., F. Wu, D. Harrich, L. F. Garcia-Martinez, and R. B. Gaynor.** 1995. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. J. Virol. **69:**3584-3596.

36. **Pan, D., N. He, Z. Yang, H. Liu, and X. Xu.** 2005. Differential gene expression profile in hepatopancreas of WSSV-resistant shrimp (*Penaeus japonicus*) by suppression subtractive

hybridization. Dev. Comp. Immunol. **29:**103-112.

37. **Ravichandran, V., and E. O. Major.** 2006. Viral proteomics: a promising approach for understanding JC virus tropism. Proteomics **6:**5628-5636.

38. **Ray, A. S.** 2005. Intracellular interactions between nucleos(t)ide inhibitors of HIV reverse transcriptase. AIDS Rev. **7:**113-125.

39. **Reyes, A., M. Salazar, and C. Granja.** 2007. Temperature modifies gene expression in subcuticular epithelial cells of white spot syndrome virus-infected *Litopenaeus vannamei*. Dev Comp Immunol. **31:**23-29.

40. **Tang, X., J. Wu, J. Sivaraman, and C. L. Hew.** 2007. Crystal structure of two major envelope proteins VP26 and VP28 from white spot syndrome virus (WSSV) shed light on their evolutionary relationship. J. Virol. in press.

41. **Tsai, J. M., H. C. Wang, J. H. Leu, H. H. Hsiao, A. H. J. Wang, G. H. Kou, and C. F. Lo.** 2004. Genomic and proteomic analysis of thirty-nine structural proteins of shrimp white spot syndrome virus. J. Virol. **78:**11360-11370.

42. **van de Braak, C. B., M. H. Botterblom, E. A. Huisman, J. H. Rombout, and W. P. van der Knaap.** 2002. Preliminary study on haemocyte response to white spot syndrome virus infection in black tiger shrimp *Penaeus monodon*. Dis. Aquat. Org. **51:**149-155.

43. **Van Hulten, M. C. W., J. Witteveldt, S. Peters, N. Kloosterboer, R. Tarchini, M. Fiers, H. Sandbrink, R. K. Lankhorst, and J. M. Vlak.** 2001. The white spot syndrome virus DNA genome sequence. Virology **286:**7-22.

44. **Wang, B., F. Li, B. Dong, X. Zhang, C. Zhang, and J. Xiang.** 2006. Discovery of the genes in response to white spot syndrome virus (WSSV) infection in *Fenneropenaeus chinensis* through cDNA microarray. Mar Biotechnol. **8:**491-500.

45. **Wang, Hao-Ching, Han-Ching Wang, J. H. Leu, G. H. Kou, A. H. Wang, and C. F. Lo.** 2007. Protein expression profiling of the shrimp cellular response to white spot syndrome virus infection. Dev. Comp. Immunol. in press.

46. **Wang, H.-Y., I-F. Wang, J. Bose, and C.-K. J. Shena.** 2004. Structural diversity and functional implications of the eukaryotic TDP gene family. Genomics **83:**130-139.

47. **Wang, Z., H. K. Chua, A. A. Gusti, F. He, B. Fenner, I. Manopo, H. Wang, and J. Kwang.** 2005. RING-H2 protein WSSV249 from white spot syndrome virus sequesters a shrimp ubiquitin-conjugating enzyme, PvUbc, for viral pathogenesis. J Virol. **79:**8764-8772.

48. **Wienkoop, S., M. Glinski, N. Tanaka, V. Tolstikov, O. Fiehn, and W. Weckwerth.** 2004. Linking protein fractionation with multidimensional monolithic reversed-phase peptide chromatography/mass spectrometry enhances protein identification from complex mixtures even in the presence of abundant proteins. Rapid Commun. Mass Spectrom. **18:**643–650.

49. **Wilson, M., and P. Koopman.** 2002. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr. Opin. Genet. Dev. **12:**441–446.

50. **Wissmuller, S., T. Kosian, M. Wolf, M. Finzsch, and M. Wegner.** 2006. The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors. Nucleic Acids Res. **34:**1735–1744.

51. **Wu, J. L., and K. Muroga.** 2004. Apoptosis does not play an important role in the resistance of 'immune' *Penaeus japonicus* against white spot syndrome virus. J. Fish Dis. **27:**15-21.

52. **Wu, J. L., K. Suzuki, M. Arimoto, T. Nishizawa, and K. Muroga**. 2002. Preparation of an inoculum of white spot syndrome virus for challenge tests in *Penaeus japonicus*. Fish Pathol. **37:**65–69.

53. **Xie, X., L. Xu, and F. Yang.** 2006. Proteomic analysis of the major envelope and nucleocapsid proteins of white spot syndrome virus. J. Virol. **80:**10615-10623.

54. **Yang, F., J. He, X. Lin, Q. Li, D. Pan, X. Zhang, and X. Xu.** 2001. Complete genome sequence of the shrimp white spot bacilliform virus. J. Virol. **75:**11811-11820.

55. **Yeh, M. S., C. J. Huang, J. H. Leu, Y. C. Lee, and I. H. Tsai.** 1999. Molecular cloning and characterization of a hemolymph clottable protein from tiger shrimp (*Penaeus monodon*). Eur. J. Biochem. **266:**624-633.

56. **Zhang, X., C. Huang, X. Tang, Y. Zhuang, and C. L. Hew.** 2004. Identification of structural proteins from shrimp white spot syndrome (WSSV) by 2 DE-MS. Proteins **55:**229-235.

57. **Zhang, Y., S. Wang, A. Xu, J. Chen, B. Lin, and X. Peng.** 2006. Affinity proteomic approach for identification of an IgA-like protein in *Litopenaeus vannamei* and study on its agglutination characterization. J. Proteome Res. **5:**815-821.

58. **Zhao, H., B. K. Adler, C. Bai, F. Tang, and X. Wang.** 2006. Epithelial proteomics in multiple organs and tissues: similarities and variations between cells, organs, and diseases. J. Proteome Res. **5:**743-755.

Figure legends

FIG. 1. Workflow for the sample collection and sample preparation.

FIG. 2. Relative abundance of viral proteins in different tissues of infected shrimp. (a) Equal amounts (30 µg) of whole lysate proteins extracted from stomach, gill and subcuticular epithelia of WSSV-infected shrimp were resolved by SDS-PAGE, then probed with antibodies against β-actin, VP28 (WSSV structural protein) and VP9 (WSSV nonstructural protein) respectively, showing that the concentrations of both the structural and the non-structural proteins were relatively higher in the epithelial lysate. (b) Equal amounts (10 µg) of cytosolic, membrane and nuclear proteins from the epithelium were resolved by SDS-PAGE, probed with antibodies against Histone 2A, GAPDH, VP28 and VP9 respectively, showing that the subcellular fractionation can enrich viral proteins while reducing the nuclear specific protein Histone 2A in the cytosolic fraction.

FIG. 3. Temporal transcription of the 11 novel viral protein genes. mRNAs were extracted from subcuticular epithelia sampled at 0, 2, 4, 6, 8, 12, 24 and 72 hpi. cDNAs were synthesized using random hexamers and normalized to an equal concentration by *β-actin.* As the time-elapse of infection, 10 of 14 genes show increasing mRNA levels (group 1), one maintains a constant mRNA level (group 2), while 3 show decreasing mRNA levels (group 3).

FIG. 4. Western blotting analysis of novel viral proteins. Three novel proteins were present in the total cell lysate and the cytosolic fraction of WSSV-infected epithelia, but were not detected in the purified virus and the uninfected cell lysate by western blotting using antibodies against wsv051, -076, -294 and -477 (a control of viral non-structural protein). Viral structural proteins including the capsid protein VP664 and the envelope proteins (VP28, VP26) were present in the infected cell lysate and the purified virus.

Supplemental data I: Primer sets used for RT-PCR

Supplemental data II: Details of WSSV protein identification by shotgun proteomics. The information of each identification including protein name (i.e. predicted ORF number in China isolate/Taiwan isolate), mass, score, sequence (matched peptides highlighted in bold red), peptide features and the MS^2 spectrum of the peptide with highest peptide score are shown. For single peptide-based identifications, the data of fragment assignments are provided as well.

Supplemental data III: List of peptides identified by shotgun proteomics. MS and MS^2 spectra from shotgun analyses of cytosolic, nuclear and membrane fractions were combined and searched against the NCBInr database. Peptide number, peptide sequence, peptide presence in identified protein(s), protein accession number, observed peptide mass, delta mass (=observed peptide mass-calculated peptide mass), peptide ion score and peptide ion score C. I. % are listed.

Supplemental data IV: List of shrimp epithelial cellular proteins identified by shotgun proteomics. A protein identified without sharing peptide with other proteins was grouped into distinct proteins (supplemental data IVa consisting of 144 proteins), while a protein sharing peptide(s) with other protein (s) was grouped into indistinct proteins (supplemental data IVb consisting of 285 proteins). Protein name, accession number, mass, pI, matched peptide number, total ion score and total ion C.I. % are presented.

Supplemental data V: The details of identification of differentially expressed cellular proteins by cICAT. The information of each identification including protein name, accession number, mass, score, sequence (matched peptides shown in bold red), peptide features and one MS^2 spectrum of the peptide with the highest peptide score are shown. For single peptide-based identifications, the data of fragment assignments are provided as well.

Supplemental data VI: The details of peptides identified in tubulin and ATPase groups by cICAT proteomics.

Supplemental data VII: The details of all proteins and peptides identified by cICAT proteomics, including the means and deviations of H/L ratios.

Supplemental data VIII: Sequence alignment of ci-Tardbp with homologs. Human (hTDP43), *Drosophila* (dmTDP43), *C. elegans* (ceTDP43) and the identified ci-Tardbp protein sequences were retrieved from the NCBI database (Genbank accession no. NP031401,

BAA34421, NP495921 and BAE06721, respectively) for multiple alignment using the program ClustalW (www.ebi.ac.uk/clustalw) with default parameters. RNA recognition motifs, RRM1 and RRM2, were underlined with thin dark lines, while the RNP (ribonucleoprotein) consensus sequences 2 and 1 present in each RRM were underlined with thick grey lines. Amino acids identical or similar among the four proteins are depicted in dark or grey, respectively.

FIG. 1, Page 7

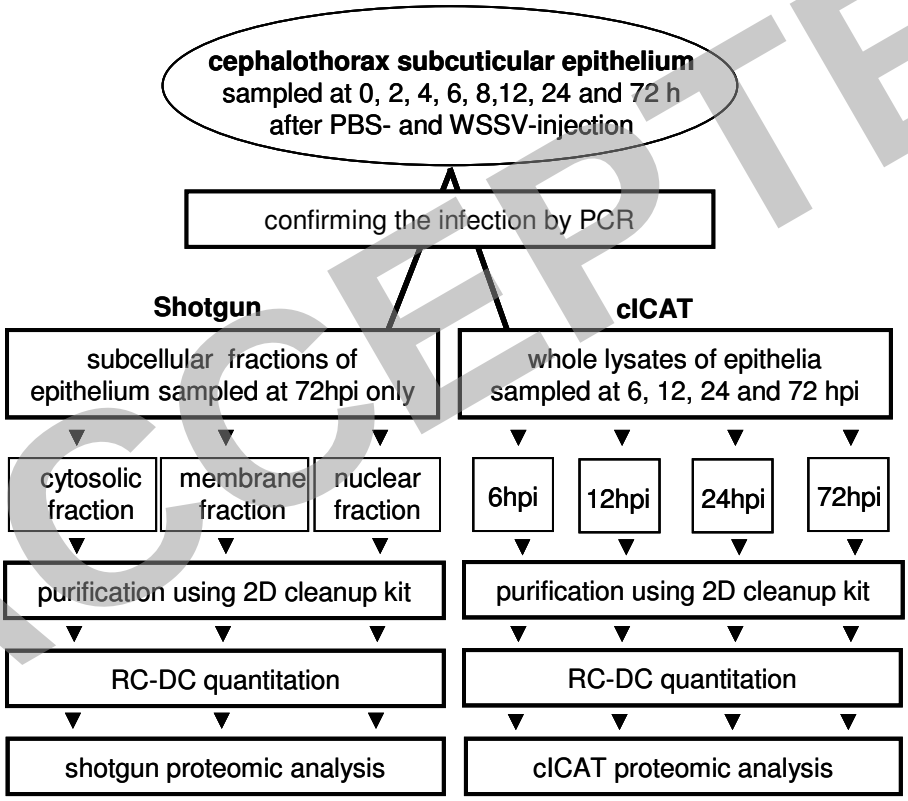

FIG. 2, Page 12

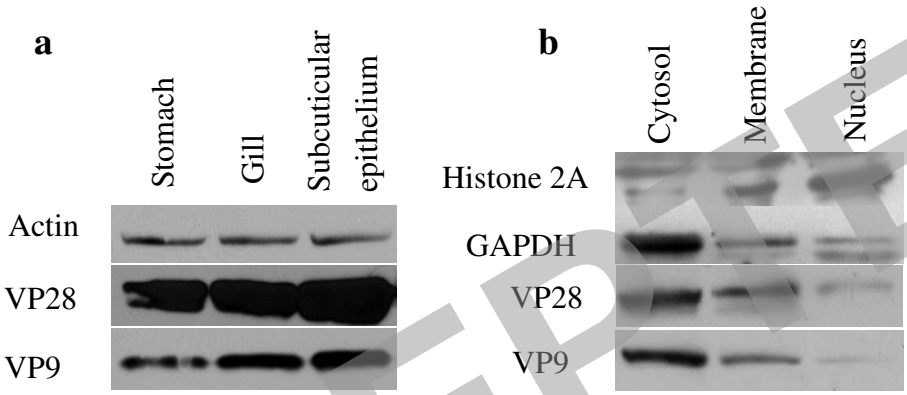

FIG. 3, Page 15

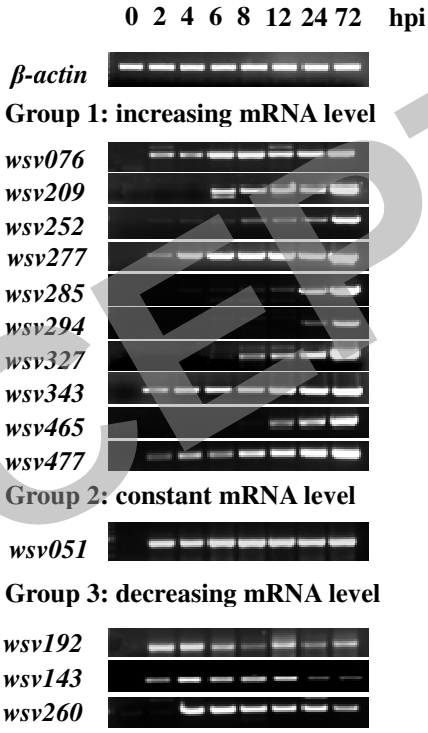

FIG. 4, Page 16

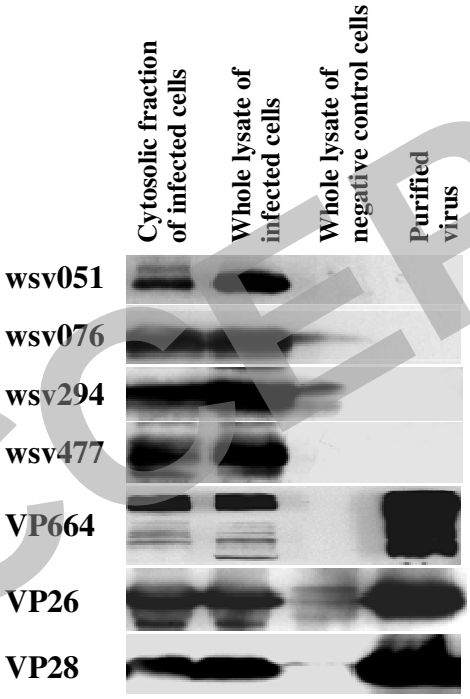

Table 1. Page 12

List of WSSV proteins identified by shotgun proteomics. Cytosolic, nuclear and membrane fractions were isolated from WSSV-infected subcuticular epithelium for shotgun 2D-LC-MS^2 analysis. The number of peptides, best ion score, total ion score and the total ion score C.I. % for the identification of WSSV proteins are indicated. The closer the C.I. % value is to 100%, the more likely the protein is correctly identified. Proteins listed below the dashed line (total ion score C.I.%<95) are not considered as confident identifications. Proteins shown in bold letters were identified for the first time. Details of MS data were shown in supplemental data II.

Predicted ORF	Accession GI No.	Fraction[a]	No. of Peptides matched	Best Ion Score[b]	Total Ion Score	Total Ion C.I. %	Predicted structure and function	Protein name & Reference
wsv277	**17158380**	**C N**	**14**	C 150	**1131**	**100**	**No hits**	**this study**
wsv360	17158462	C M N	11	N 105	582	100	No hits	VP664 (14)
wsv421	17158523	C M N	4	C 112	301	100	SP[c]	VP28 (14)
wsv230	17158334	C	3	C 151	287	100	No hits	VP9 (33)
wsv051	**17158155**	**C**	**2**	**C 94**	**152**	**100**	**No hits**	**this study**
wsv209	17158313	C	3	C 87	174	100	No hits	VP187 (32)
wsv294	**17158396**	**C N**	**2**	**C 90**	**144**	**100**	**No hits**	**this study**
wsv327	**17158429**	**C**	**1**	**C 103**	**103**	**100**	**No hits**	**this study**
wsv332	17158434	N	2	N 57	111	100	No hits	VP75(14)
wsv252	**17158355**	**C N**	**1**	**N 99**	**99**	**100**	**TMH[d]**	**this study**
wsv188	17158292	C	1	C 85	85	100	1TMH	RR2 (17)
wsv465	17158566	C M N	2	C 48	85	100	No hit	VP136B (14)
wsv343	**17158445**	**C N**	**2**	**N 50**	**83**	**100**	**NTPase**	**this study**
wsv076	**17158180**	**C**	**2**	**C 52**	**81**	**100**	**No hits**	**this study**
wsv289	17158391	C	1	C 44	44	100	No hits	VP190 (16)
wsv285	**17158388**	**N**	**1**	**N 61**	**61**	**100**	**No hits**	**this study**
wsv477	17158578	C	2	C 38	58	100	No hits	WSV477 (21)
wsv216	17158320	C M N	1	N 57	57	100	1TMH, SP	VP124 (35)
wsv002	17158106	N	1	N 53	53	100	SP	VP24(14)
wsv308	17158410	N	1	N 50	50	100	No hits	VP51C(14,15)
wsv011	17158115	C N	2	C 25	45	100	3TMH, SP	VP53A(14)
wsv143	**17158247**	**C N**	**1**	**C 34**	**42**	**99**	**HMG domain**	**this study**
wsv172	17158276	C N	2	C 22	43	99	1TMH	RR1 (17)
wsv254	17158357	N	1	N 43	43	98	No hits	VP36b(14,15)

wsv415	17158517	C N	1	C 42	42	98	No hits	VP60B(14,15)
wsv192	**17158296**	**C**	**1**	**C 33**	**41**	**98**	**RNP1, 1TMH,**	**this study**
wsv260	**17158363**	**C N**	**1**	**C 33**	**37**	**96**	**D/E/S rich[e]**	**this study**
wsv390	17158492	C	1	C36	36	95	No hits	VP38B(14)
wsv226	**17158330**	**C**	**2**	**C 31**	**37**	**94**	**No hits**	**this study**
wsv284	17158387	C	1	C 29	29	72	1TMH, SP	VP13(14)
wsv311	17158413	C	1	C 27	27	59	1TMH, SP	VP26(14)
wsv001	17158566	C N	1	C 27	25	36	No hits	VP180(14,15)
wsv242	17158346	N	1	N 23	23	0	No hits	VP41B(14,15)

a. Capital letters C, M and N in this column indicate that the protein was identified in cytosolic, membrane and nuclear fractions, respectively.

b. Capital letters in this column shows that the proteins were best identified in the cytosol (C), membrane (M) or nucleus (N) with the highest ion score, total ion score and total ion score C.I. %.

c. SP represents signal peptide, which was predicted using SignalP3.0 software.

d. TMH represents transmembrane helix, which was predicted using TMHMM-2.0 software.

e. D/E/S are single letter representations of amino acids, i.e. Asp/Glu/Ser.

Table 2, Page 14

Summary of differentially expressed cellular proteins identified by a time-course cICAT analysis. Pairs of epithelial lysates from uninfected and WSSV-infected *Peneaus monodon* at 6, 12, 24 and 72 hour post-infection were prepared for cICAT analyses. A protein identified with a value of H/L ratio less than 0.71 is regarded as down-regulated, while a protein with a value of H/L ratio greater than 1.4 is regarded as up-regulated. Details are shown in the supplemental data V and VII.

Protein Name	Gene bank Accession	Peptide count 6/12/24/72[a]	Total Ion Score 6/12/24/72	Total Ion Score C.I. % 6/12/24/72	Avg ICAT Ratio (H/L): 6/12/24/72
Down-regulated proteins					
elongation factor-2	37704007	-/-/-/1[b]	-/-/-/54[c]	-/-/-/98.5[d]	-/-/-/0 [e]
aldehyde dehydrogenase	66514094	-/-/-/1	-/-/-/62	-/-/-/99.8	-/-/-/0.3
Serum albumin	62113341	-/-/1/-	-/-/90/-	-/-/100/-	-/-/0.4/-
heat shock protein 70	123585	-/-/-/1	-/-/-/50	-/-/-/95.3	-/-/-/0.4
farnesoic acid O-methyltransferase	85677401	-/1/-/-	-/55/-/-	-/98.5/-/-	-/0.5/-/-
Ci-Tardbp	70571316	-/-/1/1	-/-/48/64	-/-/96.9/99.8	-/-/1.5/0.4
beta-hemoglobin	29446	-/-/2/-	-/-/105/-	-/-/100/-	-/-/0.5/-
PACAP protein	18204192	-/-/-/1	-/-/-/49	-/-/-/94.7	-/-/-/0.6
elongation factor 1	22128323	-/-/-/1	-/-/-/84	-/-/-/100	-/-/-/0.7
clottable protein	6601498	2/3/1/3	109/254/69/195	100/100/100/100	0.8/1/1.8/0.6
Up-regulated proteins					
arginine kinase	46401522	-/-/1/1	-/-/75/100	-/-/100/100	-/-/1.4/1.7
hemocyanin	7414468	-/2/2/2	-/126/146/226	-/100/100/100	-/1.7/1.6/2.1

a: sampling time points, i.e. 6, 12, 24 and 72 hours post-infection

b: peptide numbers detected at 6, 12, 24 and 72 hours post-infection. "-" represents no peptide of the protein was detected at the time point.

c: total ion scores of proteins detected at 6, 12, 24 and 72 hours post-infection.

d: total ion score C.I. % of the proteins detected at 6, 12, 24 and 72 hours post-infection.

e: values of H/L ratios detected at 6, 12, 24 and 74 hours post-infection. Zero means that only the peptide labeled with light tag was detected.

Table 3, Page 15

List of peptides of tubulin and ATPase groups identified by cICAT. Details are shown in supplemental data VI.

Peptide sequences	Mr(expt)	Mr(calc)	Score	Expect	H/L
Tubulin group identified at 24 hours post-infection (supplemental data VI-1)					
SIQFVDWCPTGFK + ICAT_heavy	1762.88	1762.88	56	0.0055	0.97
AYHEQLSVAEITNACFEPANQMVK + ICAT_heavy	2928.43	2928.42	83	1.3e-005	1.65
Tubulin group identified at 72 h post-infection (supplemental data VI-2)					
EIVHLQTGQCGNQIGTK + ICAT_light	2052.04	2052.04	132	2.8e-010	0.60
SIQFVDWCPTGFK + ICAT_light	1753.83	1753.85	73	0.00022	0.56
RSIQFVDWCPTGFK + ICAT_light	1909.95	1909.95	52	0.029	0.39
AVCMLSNTTAIAEAWAR + ICAT_light	2034.00	2034.00	66	0.0013	0.50
TIQFVDWCPTGFK + ICAT_light	1767.86	1767.87	68	0.00069	0.53
ATPase group identified at 24 hours post-infection (supplemental data VI-3)					
IPIFSAAGLPHNEIAAQICR + ICAT_heavy	2356.27	2356.28	124	9.8e-010	1.41
KDHSDVSNQLYACYAIGK + ICAT_heavy	2247.11	2247.10	45	0.072	2.28
ATPase group identified at 72 hours post-infection (supplemental data VI-4)					
IPIFSAAGLPHNEIAAQICR + ICAT_light	2347.24	2347.25	125	1.5e-009	0.64
KDHSDVSNQLYACYAIGK + ICAT_light	2238.08	2238.07	70	0.00045	0.38

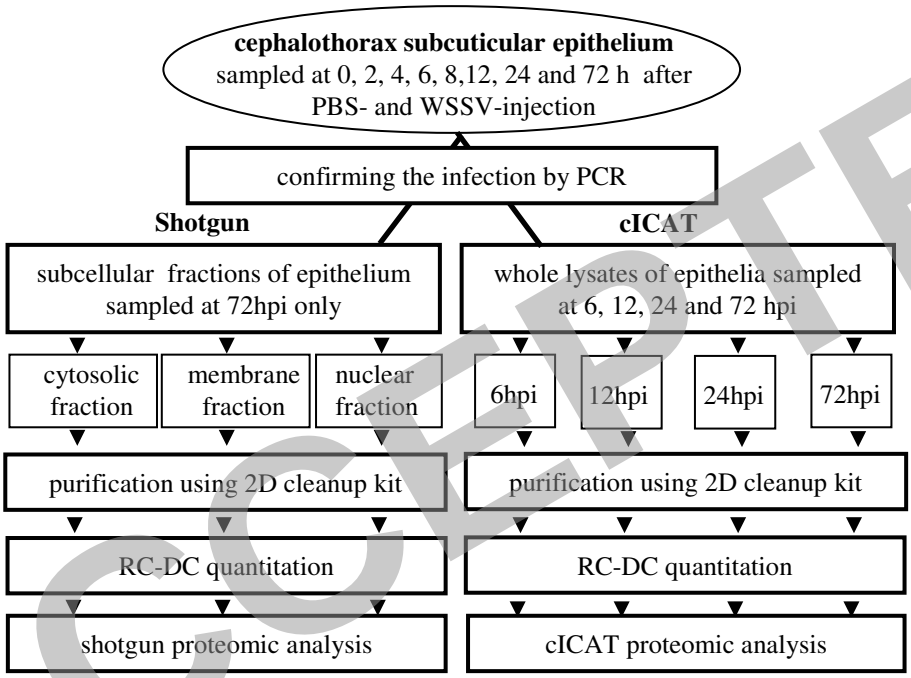

Fig. 1

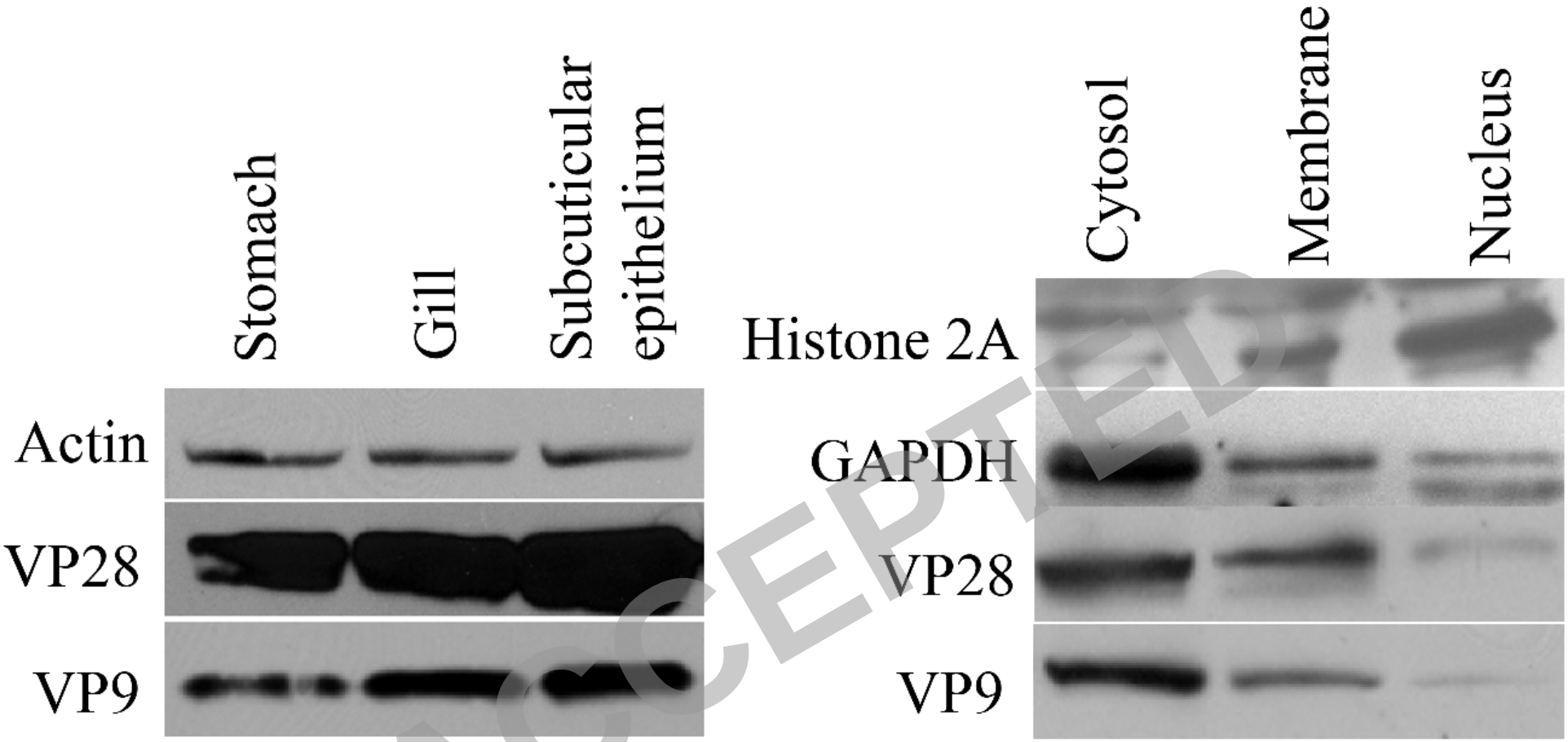

Fig. 2