Thesis for the Degree of Master of Science

Molecular characterization of three genes encoding crustin-like peptides in the Morotoge shrimp, *Pandalopsis japonica*

by

Meesun Kim

Department of Marine Biology

The Graduate School

Pukyong National University

February 2012

Molecular characterization of three genes encoding crustin-like peptides in the Morotoge shrimp, *Pandalopsis japonica*

(물렁가시붉은새우로부터 세 개의 crustin 유사 유전자의 분자적 특성 연구)

Advisor: Prof. Hyun-Woo Kim

By

Meesun Kim

A thesis submitted in partial fulfillment of the requirements

for the degree of

Master of Science

In Department of Marine Biology, The graduate School,

Pukyong National University

February 2012

Molecular characterization of three genes encoding crustin-

like peptides

in the Morotoge shrimp, *Pandalopsis japonica*

A dissertation

by

Meesun Kim

Approved by :

_____

(Chairman) Dr. Chul-Woong Oh

_____ _____

(Member) Dr. Tae-Jin Choi **(Member) Dr. Hyun-Woo Kim**

February 24, 2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Molecular characterization of three genes encoding crustin-like peptides in the Morotoge shrimp, *Pandalopsis japonica*

Meesun Kim

Department of Mirine Biology, The graduate School,

Pukyong National University

Abstract

Antimicrobial peptides (AMPs) play an important role in the innate immune system, and members of the cationic AMP family are widely distributed in both vertebrates and invertebrates. Penaeidins and crustins are most widely studied in decapod crustaceans. Penaeidins are cationic AMPs of 5–7 kDa, characterized by a proline-rich amino-terminal domain and a cysteine-rich carboxyl terminus domain. In present study, three putative crustin-like full-length cDNAs (Pj-crus Ia, Ib and IIa) were isolated from the Pandalid shrimp, *Pandalopsis japonica* and their gene organizations were also determined. The full-length cDNAs of Pj-crus Ia, Ib and IIa consisted of 1135, 580 and 700 nucleotides encoding a protein with 109, 119 and 186 amino acids residues, respectively. All three crustins isolated in this study contained a WAP domain at the C-terminus formed by eight conserved cysteine residues for four disulfide bridges. Phylogenetic analysis by WAP domains revealed that Pj-crus Is belong to carcinin, which is mostly identified from Brachyura and Astacidea. By contrast, Pj-crus IIa clustered with typical crustin which are mainly found in Dendrobranchiata. Based on

the amino acid sequence similarities and domain organizations, the amino acid sequences of Pj-crus Ia, Ib and IIa showed that they belong to type I (Pj-crus Ia and Ib) and type II (Pj-crus IIa) crustin. Pj-crus Ia has four exons interrupted by three introns, whereas Pj-crus Ib and Pj-crus IIa have three exons. The expression profile data by end-point RT-PCR showed that three Pj-crus genes exhibit different tissue-distribution profiles. Pj-crus Ia was detected epithelia and brain and Pj-crus Ib was detected epithelia, brain and gill while Pj-crus IIa showed ubiquitous expression profile. This comparative study of putative crustacean AMPs will help us to extend our knowledge about the evolutional relationship and physiological roles of various arthropod AMPs.

Introduction

It is generally accepted that crustaceans lack the highly specific adaptive immune system shown in vertebrates. Their internal defenses rely mostly on innate immune responses (Rosa and Barracco 2010). Antimicrobial peptides (AMPs) are small protein molecules with less than 100 amino residues (~10 kDa), which are one of the major components of innate immune system of all kingdoms from bacteria to mammals (Boman, 1991). Currently, 15 different AMP family members have been identified in crustaceans (Rosa and Barracco, 2010). Three major families of AMPs, penaeidins, antilipopolysaccharide factors (ALFs) and crustins have been mostly studied (Destoumieux et al., 2000; Smith et al., 2008). Penaeidins are small cationic peptides (5.5-6.6 kDa) and have been identified only in Dendrobranchiata (Destoumieux et al., 1997). Their typical primary structure is composed of the N-terminal proline-rich domain (PRD) and the C-terminal cysteine-rich domain (CRD) forming three intradisulfide bonds (Destoumieux et al., 1997; Tassanakajon et al., 2010). Penaeidins can further classified into four distinct subgroups based on their amino acid similarity and their spectrums of activity vary, respectively (Chiou et al., 2007; Cuthbertson et al., 2002; Ho et al., 2004; Kang et al., 2007; O'Leary and Gross, 2006; Tassanakajon et al., 2010). ALFs (7-11 kDa)

are found in various decapods and characterized by highly hydrophobic N-terminal region and two characteristic cysteine residues (Beale et al., 2008; Liu et al., 2006; Rosa et al., 2008; Tassanakajon et al., 2010; Yue et al., 2010). ALFs exhibit multifunctional activities including inhibition of LPS-mediated hemolymph coagulation and antimicrobial activity against Gram-negative bacteria (Somboonwiwat et al., 2008). Major structural characteristics of crustin family is the whey-acidic protein (WAP) domain at the C-terminus, and these small peptides (7-14 kDa) show highest antimicrobial activity against Gram-positive bacteria (Amparyup et al., 2008b; Relf et al., 1999; Smith et al., 2008; Zhang et al., 2007).

First crustin family members were isolated from the shore crab *Carcinus maenas*. These peptides, named carcinins, exhibited antimicrobial activity against both Gram-positive and Gram-negative bacteria (Schnapp et al., 1996). Since the first discovery of carcinins, numerous crustin family members have been isolated from decapods crustaceans including penaeids (Antony et al., 2010; Jia et al., 2008; Rojtinnakorn et al., 2002; Rosa et al., 2007; Sun et al., 2010; Supungul et al., 2004; Vatanavicharn et al., 2009), caridae (Dai et al., 2009), anomura (Sperstad et al., 2009), achelata (Pisuttharachai et al., 2009), astacidae (Christie et al., 2007; Hauton et al., 2006; Jiravanichpaisal et al., 2007; Shi et al., 2009) and brachyura (Afsal et al., 2011; Imjongjirak et al., 2009; Mu et al., 2011; Sperstad et al., 2009; Yue et al., 2010). Until now, total XXX nucleotide sequences encoding crustin family

members have been reported in the GenBank database (http://www.ncbi.nlm.nih.gov).

General primary structure of crustin includes N-terminal signal sequence and WAP domain, which is composed of eight cysteine residues forming four disulfide bonds. Smith et al. (2008) divided crustin family members into three types based on the domain organizations; Type I, II and III (Fig. 1). Although all crustins contain N-terminal signal peptide and a C-terminal WAP domain, three types of crustins contain variable region between signal peptide and WAP domain. Type I crustins including the carcinins, have cysteine-containing region between signal sequence and WAP domain. In addition to all three region found in type I crustin, type II crustins posses additional glycine-rich region between signal sequence and cysteine-containing region, which is usually composed of 20 to 50 residues (Rosa et al., 2007). Type II crustins include penaedins, which are isolated mostly in Dendrobranchiata (Bartlett et al., 2002). Type III crustins exhibit unique domain structure in which glycine-rich and cysteine-containing regions are replaced by a short region containing proline and arginine residues between signal peptide and WAP domain (Fig. 1). Until now, Type III crustins have been isolated only in Dendrobranchiata including *Penaeus monodon* (Supungul et al., 2004), *Litopenaeus vanamei* (Jiménez-Vega et al., 2004), *Marsupenaeus japonicus* (Rojtinnakorn et al., 2002) and *Fenneropenaeus chinensis* (Jia et al., 2008).

Pandalopsis japonica is one of important decapod crustaceans in East Asian countries including Korea and Japan and its catch has been decreasing recently. In addition to its economic importance, this animal is easy to maintain and has been exhibited high survival rate after eyestalk ablation (ESA) as a good model system for crustacean physiology (Jeon et al., 2010; Lee et al., 2011). Since it inhabits in cold and deep water (4 – 6 °C), *P. japonica* is believed to exhibit different AMP profiles from other decapods living in warm and mid condition. The ultimate goal of this study is to identify and characterize different AMP genes and to determine relationship between structural and expressional characteristics and antimicrobial functions. In present study, we isolated three genes encoding crustin-like peptides named Pj-crus Ia, Pj-crus Ib and Pj-crus IIa from the morotoge shrimp, *P. japonica*. Their gene organization, primary structural characteristics, and tissue-distribution profiles were also determined. Finally, transcriptional changes of three Pj-crus genes were measured after immune challenge with both Gram-positive and Gram-negative bacterial strains. This study would expand our knowledge about defense mechanism of crustacean and give us basic information for healthy production of crustacean species.

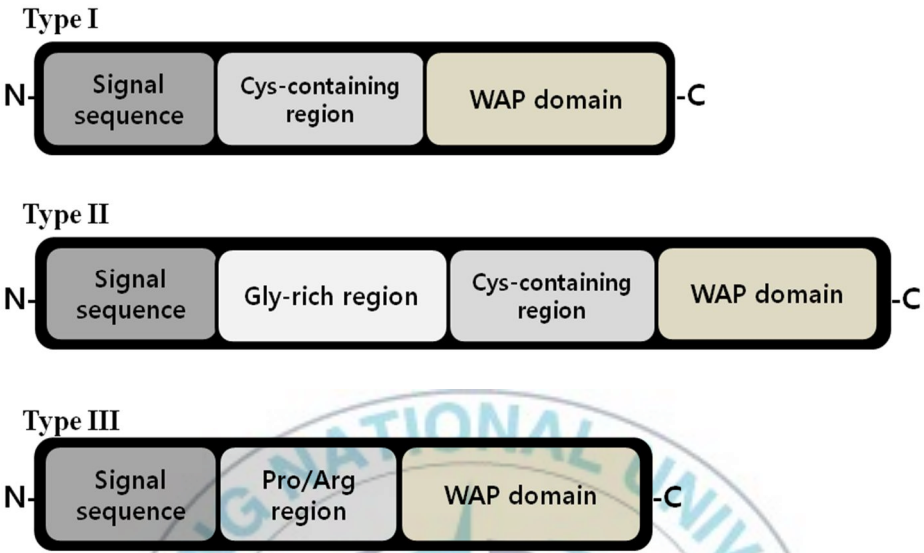

Fig. 1. Schematic representation of three crustin types. Both amino (N-) and carboxy (-C) termini are shown. (Smith et al., 2008).

Materials and Methods

Immune challenge and tissues collection

Live shrimps were purchased from a local seafood market, and held in a tank with 20 L of circulating aeration seawater for 3-5 days at 5 °C for acclimatization before experiments. The shrimps were randomly divided into three groups (PBS, *Vibrio parahaemolyticus* and *Staphylococus aureus*). Each group consisted of thirty individuals and bacterial challenge was carried out by injecting a 20 µl suspension of *V. parahaemolyticus* and *S. aureus* (4.8×10^7 and 4×10^8 cells per shrimp cells, respectively) in phosphate buffer saline (PBS: 140mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4) into the third abdominal segment of the cultured shrimps to introduce a bacterial challenge as described previously (Sun et al., 2010). Shrimp injected with 20 µl of PBS were used as controls. After being challenged for 24 and 48 hours, gill and epidermis of the three groups of shrimps, which contained among the bacterially challenged and unchallenged shrimps, were collected. Right after dissection, tissues were frozen in liquid nitrogen and stored at -80 °C until used for total RNA extraction .

Total RNA isolation and first-strand cDNA synthesis

Total RNA was extracted from various tissues using RNAiso Plus reagent according to the manufacturer's protocol (Takara Bio Inc., Japan). In order to eliminate possible contamination of genomic DNA, all isolated RNAs were treated with RNase free Dnase I (Takara Bio Inc., Japan). RNA quantity, purity and integrity were measured spectrophotometrically at 260 nm (Nanodrop Technologies, Inc., USA) and verified by electrophoresis on 1 % agarose gels. Isolated RNAs were aliquoted and stored at -70 °C until they were to be used. cDNA was synthesized in a reaction containing reverse transcriptase and oligo-dT primer. A 15 μl mixture containing 5 μg total RNA, 1 μl of 20 μM oligo-dT primer and 4 μl dNTPs was heated to 70 °C for 5 min and chilled on ice for 2 min. First-strand buffer (5 X, 5 μl), 2 μl 0.1 M DTT and RNaseOUT (Invitrogen Co., USA) were added to the mixture, which was incubated at 37 °C for 2 min. Finally, 1 μl M-MLV reverse transcriptase (Invitrogen Co., USA) was added and the mixture was incubated at 37 °C for 50 min. Synthesized cDNA was quantified, aliquoted and stored in -20 °C for later PCR reactions.

Cloning of the full length cDNAs of Pj-crus Ia, Ib and IIa

Bioinformatic analysis and the traditional PCR-based strategy were used for cloning of the full length sequence of Pj-crus Ia, Ib and IIa cDNAs.Total three cDNA contigs from EST database constructed from the *P. japonica* neuronal tissues exhibited high sequence similarity to crustins from decapods. Open reading frame (ORF) analysis identified start codon (ATG) and 5' UTRs but failed to identify stop condons and 3' UTRs. Rapid amplification of cDNA ends (RACE) technique was used to obtain full-length cDNA. All the primers used for each PCR were shown in table 1. All primers used this experiment were designed using IDTSciTools program (http://eu.idtdna.com/analyzer/Applications/ OligoAnalyzer/) and were synthesized by Bioneer Coperation, Korea (Table 1). 3' end of the full size of mRNA was identified by 3' RACE, which was described previously (Lee et al., 2011). Briefly, cDNA was synthesized using a reverse primer (Table 1). PCR was carried out with two sequence-specific primers and linker primers. The PCR reaction was performed in a total volume of 30 µl using S1000^{TM} Thermal cycler (Bio-Rad Laboratories, Inc., USA), PCR mixture containing 3 µl of 10 X PCR buffer, 2 µl of 2.5 mM dNTP mixture, 1 µl of 10 µM gene specific primer, 1 µl of 10 µM oligo-dT primer, 2 µl of template cDNA from gill, 0.5 µl of TakaRa Taq DNA polymerase (5 U/µl) (Takara Bio Inc., Japan). RT-PCR cycles were performed at 95 °C for 3 min followed by 30 cycles of 94 °C for 30 sec, 60 °C for 30 sec, 72 °C for 1 min, and a final cycle of 72 °C for 5 min. The amplified PCR products were analyzed by gel electrophoresis on 1.5 % agarose gel and stained

with ethidium bromide. The PCR products with expected size were purified using Gel Extraction Kit (Bioneer Co., Korea), ligated into TA plasmid vector with the pGEM-T Easy Cloning Kit (Promega, USA), Vectors containing cloned inserts were transformed into *E. coli* DH5α and incubated overnight at 37 ℃. Positive clones were then screened and sequenced in both directions using M13 forward and reverse primers (Table 1). RACE and original sequences were assembled by Computer-aided homology analysis and full length of each sequence was reconfirmed by RT-PCR using two sequence-specific primers, which aimed at each end of full sequence.

Determination of genomic organizations

Genomic DNA was extracted from the deep abdminal flexor muscle tissue using AccuPrep® Genomic DNA Extraction Kit according to the manufacturer's protocol (Bioneer Co., Korea). To obtain the genomic sequences of Pj-crus Ia, Ib and IIa, gene-specific primers were designed from the 5' UTR and 3' UTR of each gene and nested PCRs were carried out. The first PCR were performed in total volume of 30 μl containing 2 μl of 60ng genomic DNA as a template, 1 μl of 10 μM each forward and reverse primer (Table 1), 10 X Ex taq buffer, 2 μl of dNTP mixture (2.5mM each), 0.5 μl of Ex Taq DNA polymerase (5 U/μl) (Takara Bio Inc.,

Japan). PCR condition were performed at 94 ℃ for 5 min followed by 35 cycles of 94 °C for 30 sec, 60 °C for 30 sec, 72 °C for 2 min, and a final cycle of 72 °C for 5 min. For nested PCR, reaction contained 1 μl of first PCR product as a template, 1 μl of 10 μM each gene specific forward and reverse primer, 10 X Ex taq buffer, 2 μl of dNTP mixture (2.5mM each), 0.5 μl of Ex Taq DNA polymerase (5 U/μl) (Takara Bio Inc., Japan). Nested PCR conditions were same to the first reaction

The PCR products were analyzed on 1.5 %(w/v) TAE-agarose gel, and visualized by UV transillumination following ethidium bromide (2μg/ml) staining. The PCR products with expected size were purified using Gel Extraction Kit (Bioneer Co., Korea), and the DNA fragments were cloned into pGEM-T Easy vector (Promega, USA), Vectors containing cloned inserts were transformed into *E. coli* DH5α and incubated overnight at 37 °C. Positive clones were then screened and sequenced in both directions using M13 forward and reverse primers (Table 1), with an automated sequencer by a commercial service (Bioneer Co., Korea). Obtained genomic DNA sequences were assembled together with full-length cDNA sequence.

Computational analysis of three Pj-crus genes

The Pj-crus Ia, Pj-crus Ib and Pj-crus IIa sequences were analyzed using

the BLAST algorithm at the NCBI web site (http://blast.ncbi.nlm.nih.gov/Blast), and ORF was identified and deduced amino acid sequence was determined by ORF finder (http://www.ncbi.nlm.nih.gov/ projects/gorf/). Multiple amino acid sequence analysis was performed using ClustalW software (http://www.ebi.ac.uk/Tools/ clustalw2/index.html) and represented by GeneDoc program. The signalIP 3.0 program was utilized to predict the presence and location of signal peptide (http://www.cbs.dtu.dk/services/ SignalP/). Molecular mass and pI of Pj-crus genes were estimated by the compute pI/Mw tool (http://web.expasy.org/ compute_pi/). A phylogenetic diagram of crustins from crustacean was generated with 1,000 replications using Molecular Evolutionary Genetics Analysis (MEGA 4).

The mRNA expression level of significance of Pj-crus Ia, Ib and IIa in each tissue were evaluated statistically by comparing the means, using Sigma Plot. The results were subjected to analysis of *t*-test, and the level of statistically significant difference was set at $P<0.05$ (*).

Tissue distribution analysis of three Pj-crus genes

The mRNA expressions of Pj-crus Ia, Ib and IIa in various tissues, from heart, hemocytes, gonad, hepatopancreas, gill, extensor, flexor, epidermis, thoracic ganglia, abdominal ganglia and brain of unchallenged shrimps were

analyzed by PCR-based strategies. Tissue-specific expression profile was examined using end-point RT-PCR. Total RNAs from various tissues were isolated and quantified as described above. Genomic DNA was removed by treatment with RNase free DNase I (Takara Bio Inc., Japan). cDNA was synthesized as described above, except that random hexamers were used as primers for the reverse transcription.

Pairs of Pj-crus genes specific primers (Table 1) were used to amplify products of 262 bp, 290 bp and 286 bp, respectively. The 18S rRNA primers were used to serve as an internal control (Table 1). Primers were designed to land on each exon to eliminate possible gemonic contamination. Reaction (30 μl) contained 2 μl of cDNA (200ng), 1 μl of 10 μM gene specific forward primer and reverse primer, respectively (Table 1), 0.5 μl of TakaRa Taq DNA polymerase (5 U/μl) (Takara Bio Inc., Japan), 2 μl of dNTP mixture (2.5 mM each), and 3 μl of 10 X PCR buffer. PCR condition were performed at 94 °C for 3 min followed by 30 cycles of 94 °C for 30 sec, 60 °C for 30 sec, 72 °C for 30 sec, and a final cycle of 72 °C for 5 min. The sequence-specific primers for each target gene were designed using the IDTSciTools (http://eu. idtdna.com/analyzer/ Applications/OligoAnalyzer/). The PCR products were analyzed on 1.5 % (w/v) agarose-TAE gel, and visualized by UV transillumination following ethidium bromide (2μg/ml) staining.

Quantification of Pj-crus transcripts by immune challenge

The temporal expressions of Pj-crus Ia, Pj-crus Ib and Pj-crus IIa in gill and epidermis of shrimps challenged with *S. aureus*, *V. parahaemolyticus* and PBS as a positive control were determined by SYBR green quantitative real-time RT-PCR. The method for total RNA extrantion and the first-strand cDNA synthesis from differnt tissues was the same as the one used in tissue distribution analysis.

SYBR green quantitative real-time RT-PCR was performed on a DNA Engine Choromo4 Real-time Detector (Bio Bio-Rad Laboratories, Inc., USA) to investigate the expression of Pj-crus Ia, Ib and IIa. All the procedures for cDNA synthesis and PCR were same as used for end-point RT-PCR except that 40 cycles, rather than 30 cycles.Two 18S rRNA primers 18S rRNA F and 18S rRNA R (Table 1) were used to amplify a 18S rRNA gene fragment as the internal control for real-time RT-PCR. the standard curve testing was perfoermed using a series of 10-fold diluted samples, respectively for Pj-crus Ia, Ib, IIa and 18S rRNA. The slopes of standard cuves and PCR efficiency for the four genes were calculated and carried out to confirm that the fluorecent real-time PCR RT-PCR data are precise and trustworthy. Standard curves were constructed (Kim et al., 2008) to confirm the efficiency of primers and to quantify copy numbers were normalized to the 18S rRNA copy number

according to the equation: (actual copy numbers of Pj-crus Ia, Ib and IIa / actual copy number of 18S rRNA) X 10^{10}.

The real-time RT-PCR amplifications were carried out in total volume of 20 μl containing 10 μl 2 X SYBR Green premix Ex Taq™ (Takara Bio Inc., Japan), 3 μl of 300ng cDNA template, 1 μl of 10 μM each gene specific forward and reverse primer (Table 1). The PCR program was 94 °C for 1 min, followed by 40 cycles of 94 °C for 30 sec, 60 °C for 30 sec, 72 °C for 30 sec.

Table 1. Primers used for Pj-crus Ia, Pj-crus Ib and Pj-crus IIa

Primer	sequence (5'-3')	Description
linker oligo-dT	GCACGCTGTGAATGCTGCGACTACGATTTTTTT TTTTTTTTTT	cDNA synthesize
linker R1	ACGCTGTGAATGCTGCGACTAC	First reverse primer for 3'RACE
linker R2	GCTGTGAATGCTGCGACTACGA	Second reverse primer for 3'RACE
PJ-crus Ia F1	CTCAGAAGACTTGTGCTTCC	First Forward primer for 3'RACE Forward primer for first genomic DNA PCR
PJ-crus Ia F2	GACTTGTGCTTCCTAGTTCCTCAC	Second Forward primer for 3'RACE Forward primer for second genomic DNA PCR Forward primer for RT-PCR & Real-time PCR
Pj-crus Ib F1	CTCGTCTGAAGAAATATCTTTCCATCAC	First Forward primer for 3'RACE Forward primer for first genomic DNA PCR
Pj-crus Ib F2	CTTTCCATCACAAAATGGTCCGTCTGTTG	Second Forward primer for 3'RACE Forward primer for RT-PCR & Real-time PCR Forward primer for second genomic DNA PCR
Pj-crus IIa F1	GAGTACTTGGTATCTTGAAGTCACACG	First Forward primer for 3'RACE Forward primer for first genomic DNA PCR
Pj-crus IIa F2	CTTGGTATCTTGAAGTCACACGAAGACG	Second Forward primer for 3'RACE Forward primer for second genomic DNA PCR
Pj-crus Ia RT R1	CCAGATGCAGCACCTTGAGGAAAG	Reverse primer for RT-PCR & Real-time PCR
Pj-crus Ib RT R1	CCTGCAATAGCCATCGTGAGCACAGAG	Reverse primer for RT-PCR & Real-time PCR
Pj-crus IIa RT F1	ATGCAGCGTGCACGCACA	Forward primer for RT-PCR & Real-time PCR
Pj-crus IIa RT R1	CAGAGACAGAAGGCCTTCCTCCAAC	Reverse primer for RT-PCR & Real-time PCR
Pj-crus Ia gR1	ATACATATCAGACATTCATGGTTTAAGTGTTG	Reverse primer for first genomic DNA PCR
Pj-crus Ia gR2	CATATCAGACATTCATGGTTTAAGTGTTGAG	Reverse primer for second genomic DNA PCR
Pj-crus Ib gR1	CTATCTATCAAACCACACAATCAGCATAAC	Reverse primer for first genomic DNA PCR
Pj-crus Ib gR2	CAGCATAACAATTAATTTATATTATCCATTGA GCATC	Reverse primer for second genomic DNA PCR
Pj-crus IIa gR1	GACAGGAAATGGTTTAATACGAATGATG	Reverse primer for first genomic DNA PCR
Pj-crus IIa gR2	CGAATGATGTACATATTTGGTGACTCATTGAT G	Reverse primer for second genomic DNA PCR

18S rRNA F	ATGAGAGTGCTCAAAGCAGGCTACTC	Forward primer for RT-PCR & Real-time PCR
18S rRNA R	GGCGAATCGCTAGTCAGCATCGTT	Reverse primer for RT-PCR & Real-time PCR
M13F(-40)	CAGGAAACAGCTATGAC	Vector forward primer for PCR screening & sequencing
M13R(-20)	GTAAAACGACGGCCAG	Vector reverse primer for PCR screening & sequencing

Results and discussion

Isolation of cDNAs encoding Pj-crus Ia, Ib and IIa

As results of combinational sequence analysis of EST contigs and 3' RACE products, three different full length cDNAs encoding crustin-like prepropeptides were isolated. Single transcript of each sequence was reconfirmed by RT-PCR. Nomenclature of each cDNA was based on the similarity of deduced amino acid sequences and primary structure to the previously proposed classification (Smith et al., 2008). The first cDNA (1102 bp) encoded a protein of 109 amino acid residues and exhibited the highest sequence identity (31- 34 %) to crustins from to penaeid shrimp species including *Fenneropenaeus chinensis* (ACZ43783) and *Penaeus monodon* (ABW82154). It was named as Pj-crus I, since it contains cystein-containing region, and WAP domain, which is representative domain organization of type I crustins. Second cDNA (452 bp), named as Pj-crus Ib encodes a polypeptide with 119 amino acid residues and showed similar domain organization to Pj-crus Ia. It exhibits 45 % indentity to crustin from *P. monodon* (ACQ66005) and 42 % to the carcinin from *Portunus pelagicus* (ABM65762). The third cDNA (663 bp) encodes a longer protein with 186 amino acid residues and

addtional glycine-rich region was identified, which is characterized in type II crustins. It showed 47 % sequence identity to crustins from *P. japonicus* (ACU25384) and 46% identity with *Pacifastacus leniusculus* (ABP88044) and this cDNA sequence was named as Pj-crus IIa. The amino acid identity between Pj-crus Ia and Ib was 29% (44% similarity). Sequence similarity between Pj-crus Ia and IIa was partially as identified between Ib and IIa showing much lower similarity between two Pj-crus Is.

Typical polyadenylation signal sequence (AAUAAA) were detected in both Pj-crus Is whereas Pj-crus IIa lack the consensus sequence. Signal peptide prediction analysis showed that all three Pj-crus contained conserved motif and putative cleavage sites as signal peptide. Cleavage occurs between 20 and 21 (AFS-LP) for Pj-crus Ia (Fig.2), between 16 and 17 (ATA-SP) in Pj-crus Ib (Fig. 3) and 19 and 20 (AGA-EG) in Pj-crus IIa (Fig. 4), respectively. Those result support that isolated all three Pj-crus genes encode full length of secretory peptides. Total cysteine numbers were from 10 to 12. Eight cysteine residues were well conserved in all three Pj-crus forming WAP domain. Additional cysteines were also identified in cysteine-containing region. Four cysteine residues were identified in cysteine-containing region of Pj-crus Ia and Ib, whereas two cystein residues were identified in cysteine-containing region of Pj-crus IIa.These results suggest that Pj-crus may be folded into typical crustin peptides.

Gene organization of three Pj-crus genes

The genomic sequence of three Pj-crus genes were determined by PCR-based strategy. All three genomic DNA sequences at the exon/intron boundaries followed the canonical GT/AG splicing recognition rule (Zhang and Luo, 2003). In addtion, intron size were relatively small in all three Pj-crus genes from 173 bp to 412 bp. The 1789 bp of Pj-crus Ia gene (GenBank accession no. JQ004015) consisted of four exons interrupted by three introns (Fig. 2). The first splicing position was between first methionine and second methionine, second splicing occurs within cysteine-containing region, and third splicing occurs between second and third cysteine residues within WAP domain. Second type I crustin gene, Pj-crus Ib (888 bp; GenBank accession no. JQ004016) exibited the similar genomic organization to Pj-crus Ia (Fig. 3). Splicing positions were similar to Pj-crus Ia; within the cysteine-containing region and WAP domain. One major difference between Pj-crus Ia and Pj-crusIb is missed first exon containg single first methionine shown in Pj-crus Ia showing only three exons and two introns (Fig 2). Pj-crus IIa gene (1261 bp; GenBank accession no. JQ004017) showed different genomic organization from type I crustins (Fig. 4). Although it is consisted of same exon and intron numbers to Pj-crus Ib, splicing positions were considerably different from type I crustin genes. The first splicing position lied within signal sequence (between MQ and RA) and second splicing occurs within cysteine-

containing region. There is no scplicing within glycine-rich region and WAP domain (Fig. 4).

Since two types of Pj-crus genes exhibited different gene structure, we compared currently reported crustin gene organization if gene orgarnization may be one of criteria for crustin classification (Fig. 5). Similarily, the first splicing position was between first methionine and second lysine, second splicing occurs within cysteine-containing region, and third splicing occurs between second and third cysteine residues within WAP domain in *Scylla paramamosain* (Fig. 5D). Those results suggest that splicing position in type I crustins may be conserved even though first exon is missed. Until now only three genomic sequences of type I crustins have been known including our results and Dendrobranchiata may not have those genes.

To our knowledge, Pj-crus IIa is secondly isolated type II crustin cDNA from Pleocyemata and its gene organization was firstly reposted in our study. Another type II crustin-like cDNA was isolated from the red king crab, *Paralithodes camtschaticus* but it contained only five glycine residues have been found suggesting this may not belong to type II crustins (Sperstad et al., 2009). Two additional type II genes were reported from Dendrobranchiata, *P. monodon* (Fig. 5E and 5F). Crustin*Pm*5 has four exons interrupted by three introns, which has one more exon than Pj-crus IIa. However, as shown in Pj-crus IIa, no splicing occurs in WAP domain in Crustin*Pm*5. First splicing occurs right upstream of the

first methionine residue and third splicing site were wtihin cysteine-containing region, which are similar to Pj-crus IIa. By constrast, second splicing site lied within glycine-rich region, which is not found in Pj-crus IIa (Vatanavicharn et al., 2009). Crus-like*Pm*, second crustin-like gene from *P. monodon*, has two exons interrupted by one intron. Interestingly, there was no splicing position within ORF and splicing occurs only right 5' upstream of first methionine as shown in Crustin*Pm*5 (Amparyup et al., 2008b). Until now, one common character of type II crustin gene is no splicing site within WAP domain, whereas single splicing position is found in type I crustin. These result suggest that WAP domain without exon/intron boudary may be the critical characteristics of type II crustin but further study would be needed.

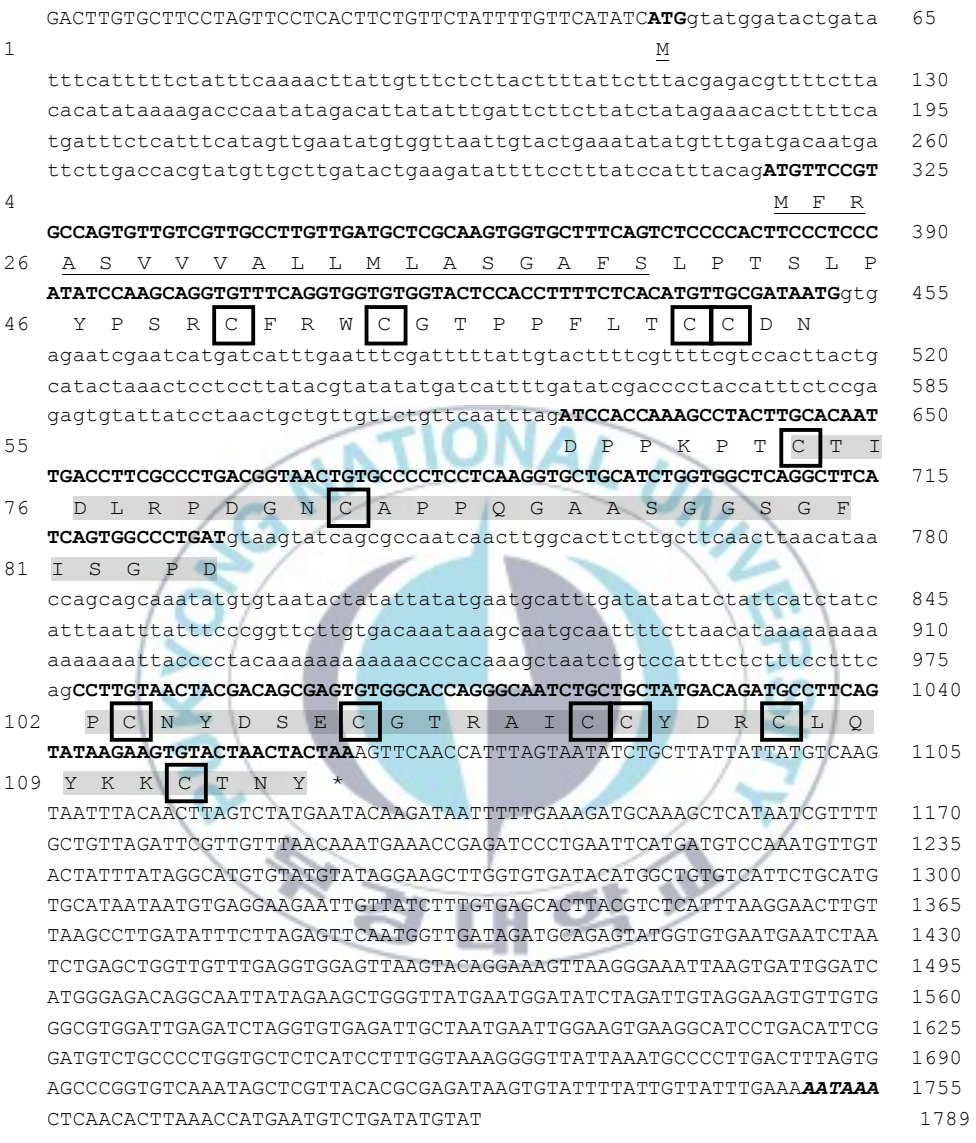
GACTTGTGCTTCCTAGTTCCTCACTTCTGTTCTATTTTGTTCATATCATGgtatggatactgata 65
M
tttcatttttctatttcaaaacttattgtttctcttacttttattctttacgagacgttttctta 130
cacatataaaagacccaatatagacattatatttgattcttcttatctatagaaacacttttca 195
tgatttctcatttcatagttgaatatgtggttaattgtactgaaatatatgtttgatgacaatga 260
ttcttgaccacgtatgttgcttgatactgaagatattttcctttatccatttacagATGTTCCGT 325
M F R
GCCAGTGTTGTCGTTGCCTTGTTGATGCTCGCAAGTGGTGCTTTCAGTCTCCCCACTTCCCTCCC 390
A S V V V A L L M L A S G A F S L P T S L P
ATATCCAAGCAGGTGTTTCAGGTGGTGTGGTACTCCACCTTTTCTCACATGTTGCGATAATGgtg 455
Y P S R C F R W C G T P P F L T C C D N
agaatcgaatcatgatcatttgaatttcgatttttattgtacttttcgttttcgtccacttactg 520
catactaaactcctccttatacgtatatatgatcattttgatatcgacccctaccatttctccga 585
gagtgtattatcctaactgctgttgttctgttcaatttagATCCACCAAAGCCTACTTGCACAAT 650
D P P K P T C T I
TGACCTTCGCCCTGACGGTAACTGTGCCCCTCCTCAAGGTGCTGCATCTGGTGGCTCAGGCTTCA 715
D L R P D G N C A P P Q G A A S G G S G F
TCAGTGGCCCTGATgtaagtatcagcgccaatcaacttggcacttcttgcttcaacttaacataa 780
I S G P D
ccagcagcaaatatgtgtaatactatattatatgaatgcatttgatatatatctattcatctatc 845
atttaatttatttcccggttcttgtgacaaataaagcaatgcaattttcttaacataaaaaaaaaa 910
aaaaaaattacccctacaaaaaaaaaaaacccacaaagctaatctgtccatttctctttcctttc 975
agCCTTGTAACTACGACAGCGAGTGTGGCACCAGGGCAATCTGCTGCTATGACAGATGCCTTCAG 1040
P C N Y D S E C G T R A I C C Y D R C L Q
TATAAGAAGTGTACTAACTACTAAAGTTCAACCATTTAGTAATATCTGCTTATTATTATGTCAAG 1105
Y K K C T N Y *
TAATTTACAACTTAGTCTATGAATACAAGATAATTTTTGAAAGATGCAAAGCTCATAATCGTTTT 1170
GCTGTTAGATTCGTTGTTTAACAAATGAAACCGAGATCCCTGAATTCATGATGTCCAAATGTTGT 1235
ACTATTTATAGGCATGTGTATGTATAGGAAGCTTGGTGTGATACATGGCTGTGTCATTCTGCATG 1300
TGCATAATAATGTGAGGAAGAATTGTTATCTTTGTGAGCACTTACGTCTCATTTAAGGAACTTGT 1365
TAAGCCTTGATATTTCTTAGAGTTCAATGGTTGATAGATGCAGAGTATGGTGTGAATGAATCTAA 1430
TCTGAGCTGGTTGTTTGAGGTGGAGTTAAGTACAGGAAAGTTAAGGGAAATTAAGTGATTGGATC 1495
ATGGGAGACAGGCAATTATAGAAGCTGGGTTATGAATGGATATCTAGATTGTAGGAAGTGTTGTG 1560
GGCGTGGATTGAGATCTAGGTGTGAGATTGCTAATGAATTGGAAGTGAAGGCATCCTGACATTCG 1625
GATGTCTGCCCCTGGTGCTCTCATCCTTTGGTAAAGGGGTTATTAAATGCCCCTTGACTTTAGTG 1690
AGCCCGGTGTCAAATAGCTCGTTACACGCGAGATAAGTGTATTTTATTGTTATTTGAAAAATAAA 1755
CTCAACACTTAAACCATGAATGTCTGATATGTAT 1789

Fig. 2. The nucleotides and deduced amino acid sequences of the Pj-crus Ia gene. Four exons (capitals) are interrupted by three introns (cases). The putative signal sequence is underlined, ORF is bolded and polyadenylation signal is written in italic and bolded. The stop codon (TAA) is indicated by an asterisk. 12 cysteines forming six putative disulfide bridges are boxed and WAP domain is shaded.

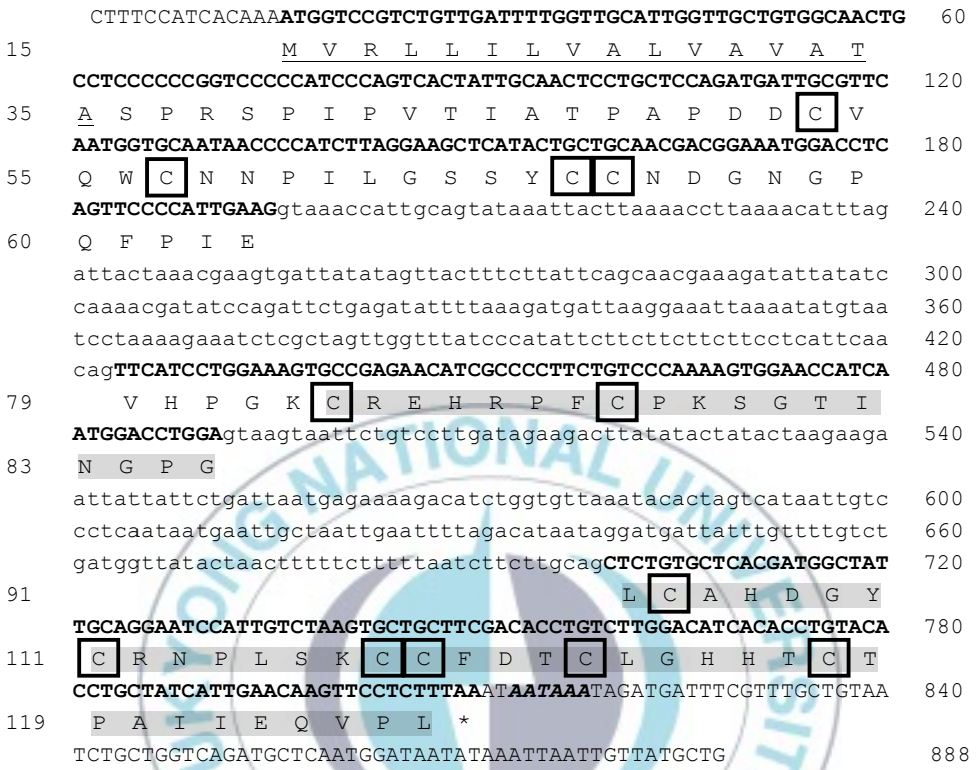

Fig. 3. The nucleotides and deduced amino acid sequences of the Pj-crus Ib gene. Three exons (capitals) are interrupted by two introns (cases). The putative signal sequence is underlined, ORF is bolded and polyadenylation signal is written in italic and bolded. The stop codon (TAA) is indicated by an asterisk. 12 cysteines forming six putative disulfide bridges are boxed and WAP domain is shaded.

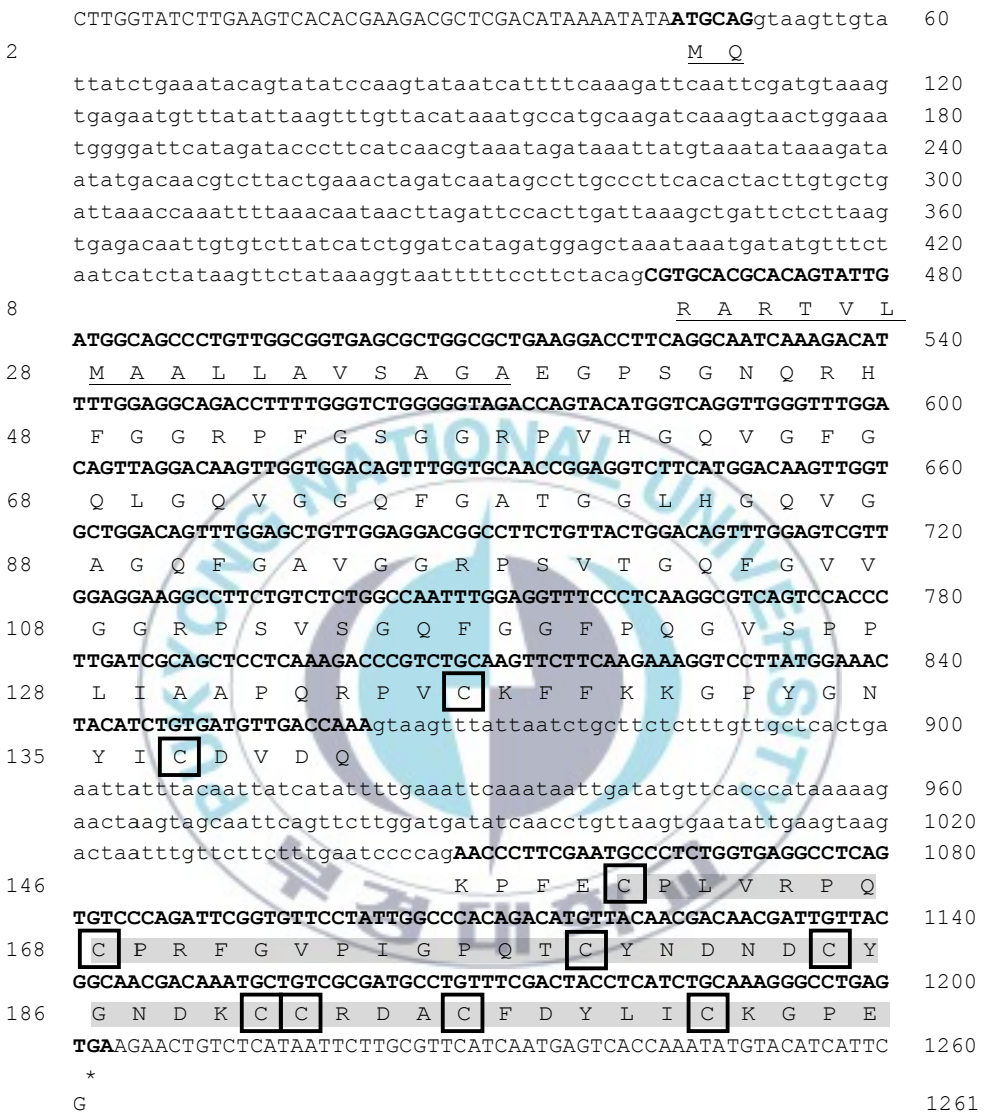
CTTGGTATCTTGAAGTCACACGAAGACGCTCGACATAAAATATAATGCAGgtaagttgta 60
2 M Q
ttatctgaaatacagtatatccaagtataatcattttcaaagattcaattcgatgtaaag 120
tgagaatgtttatattaagtttgttacataaatgccatgcaagatcaaagtaactggaaa 180
tggggattcatagataccttcatcaacgtaaatagataaattatgtaaatataaagata 240
atatgacaacgtcttactgaaactagatcaatagccttgcccttcacactacttgtgctg 300
attaaaccaaattttaaacaataacttagattccacttgattaaagctgattctcttaag 360
tgagacaattgtgtcttatcatctggatcatagatggagctaaataaatgatatgtttct 420
aatcatctataagttctataaaggtaatttttccttctacagCGTGCACGCACAGTATTG 480
8 R A R T V L
ATGGCAGCCCTGTTGGCGGTGAGCGCTGGCGCTGAAGGACCTTCAGGCAATCAAAGACAT 540
28 M A A L L A V S A G A E G P S G N Q R H
TTTGGAGGCAGACCTTTTGGGTCTGGGGGTAGACCAGTACATGGTCAGGTTGGGTTTGGA 600
48 F G G R P F G S G G R P V H G Q V G F G
CAGTTAGGACAAGTTGGTGGACAGTTTGGTGCAACCGGAGGTCTTCATGGACAAGTTGGT 660
68 Q L G Q V G G Q F G A T G G L H G Q V G
GCTGGACAGTTTGGAGCTGTTGGAGGACGGCCTTCTGTTACTGGACAGTTTGGAGTCGTT 720
88 A G Q F G A V G G R P S V T G Q F G V V
GGAGGAAGGCCTTCTGTCTCTGGCCAATTTGGAGGTTTCCCTCAAGGCGTCAGTCCACCC 780
108 G G R P S V S G Q F G G F P Q G V S P P
TTGATCGCAGCTCCTCAAAGACCCGTCTGCAAGTTCTTCAAGAAAGGTCCTTATGGAAAC 840
128 L I A A P Q R P V C K F F K K G P Y G N
TACATCTGTGATGTTGACCAAAgtaagtttattaatctgcttctctttgttgctcactga 900
135 Y I C D V D Q
aattatttacaattatcatattttgaaattcaaataattgatatgttcacccataaaaag 960
aactaagtagcaattcagttcttggatgatatcaacctgttaagtgaatattgaagtaag 1020
actaatttgttcttctttgaatccccagAACCCTTCGAATGCCCTCTGGTGAGGCCTCAG 1080
146 K P F E C P L V R P Q
TGTCCCAGATTCGGTGTTCCTATTGGCCCACAGACATGTTACAACGACAACGATTGTTAC 1140
168 C P R F G V P I G P Q T C Y N D N D C Y
GGCAACGACAAATGCTGTCGCGATGCCTGTTTCGACTACCTCATCTGCAAAGGGCCTGAG 1200
186 G N D K C C R D A C F D Y L I C K G P E
TGAAGAACTGTCTCATAATTCTTGCGTTCATCAATGAGTCACCAAATATGTACATCATTC 1260
*
G 1261

Fig. 4. The nucleotides and deduced amino acid sequences of the Pj-crus IIa. Three exons (capitals) interrupted by two introns (cases). The putative signal sequence is underlined and ORF is bolded. The stop codon (TGA) is indicated by an asterisk. 10 cysteines forming five disulfide bridges are boxed and WAP domain is shaded.

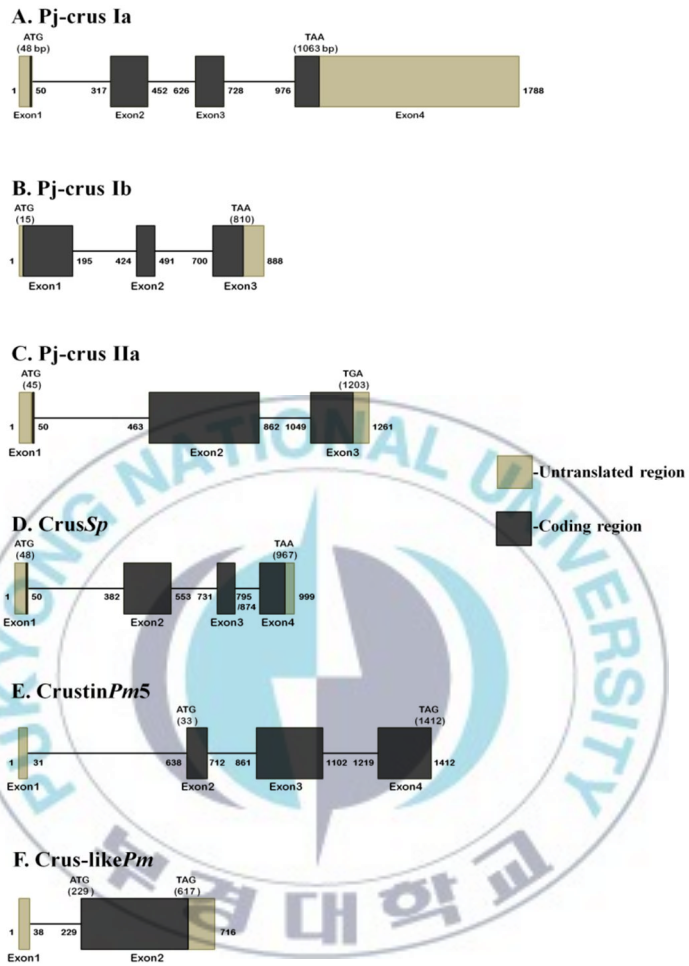

Fig. 5. Schematic representation of genomic organization among Pj-crus Ia, Pj-crus Ib and Pj-crus IIa. The boxes represent the exons, and the dark grey boxes indicate ORF. UTRs are shown as light grey boxes and the exons are connected by introns (lines). Crustin*Pm*5 (FJ380049), Crus-like*Pm* (EF654659) and Crus*Sp* (EU161288) are crustin genes from *P. monodon (Amparyup et al., 2008b; Imjongjirak et al., 2009; Vatanavicharn et al., 2009)*.

Analysis of primary structure of three crustins from P. japonica

In order to estimate structural and functional characteristics, multiple alignment was carried out with three Pj-crus sequences and representative crustin family members from various decapods (Fig. 6). Multiple alignment results showed that all crustins can be classified by four distinct domains; signal sequence region, glycine-rich region, cysteine-containing region, and WAP domain. All type I crustins including Pj-crus Ia and Pj-crus Ib were consisted of signal sequence, cysteine-containing region, and WAP domain whereas type II crustins including Pj-crus IIa, were composed of signal sequence, glycine-rich region, cysteine-containing region, and WAP domain (Fig. 6).

All compared crustins exhibit conserved signal sequence and cleavage sites suggesting all crustins are secretory proteins (Fig. 6). As shown in other type I curstins, Pj-crus Ia and Ib lack glycine rich domain. Although all type II crustins contain Glycine-rich region, the size and number of glycine residues vary individually. Pj-crus IIa exhibited a longer glycine rich region (98 residues) but numbers of glycine residues was 30, which is similar numbers of other crustins (20 ~ 50 residues). Repetitive VGGGLG motif found in type II crustins from Dendrobranchiata was not identified in Pj-crus IIa (Bartlett et al., 2002; Rosa et al., 2007; Supungul et al., 2004; Supungul et al., 2008; Vargas-Albores et al., 2004; Zhang et al., 2007). This finding may not be limited to Pj-crus IIa since one type II

crustin from *P. monodon* (EF523614) contained 22 glycine residues at the glycine-rich region but did not exhibit VGGGLG motifs (Bartlett et al., 2002). Although the functions of VGGGLG motifs is not clear, high number of glycine residues, as the smallest amino acid, might render the glycine-rich region flexible (Smith et al., 2008) but its biological implications are not clear and further study is needed.

All crustins contain cysteine-containing region but numbers of cysteine residue varies (Fig. 6). Four cysteine residues have been found in the cysteine-containing region of Pj-crus Ia and Ib, whereas two cysteines were identified in Pj-crus IIa (Fig. 6). Four cysteine residues appear to be common character of most crustins and Supungul et. al suggested that 12 cysteines (4 from cysteine-containing region and 8 from WAP domain) are predicted to form six disulfide bonds (Bartlett et al., 2002; Supungul et al., 2008). Tassannakajon et. al reviewd that conserved spacing of four cystein-containg region is C-$(X)_{2-3}$- C-$(X)_{7-8}$-CC (Tassanakajon et al., 2010). We searched all crustins available in GenBank database and found that residues between second and third cysteine at the cysteine-containing region varies from 7 to 12 suggesting C-$(X)_{2-3}$- C-$(X)_{7-12}$-CC. For example, twelve cysteine residues were identified in crustins from *Homarus americanus*, *Homarus gammarus*, *Pacifastacus leniusculus*, *Procambarus clarkii* (GenBank accession no. ABM92333, CAH10349, ABP88043, AEB54630, respectively). Nine resiudes found in the corresponding region of Pj-crus Ib suggesting spacing between second and third cysteine residues vary individually.

It it interesting to know only two cysteine residues exist in Pj-crus IIa, however, two cysteine residues at the cysteine-containing region was not limited to Pj-crus IIa. There are three cystein residues at the cysteine-containing region in the crustin from *Portunus trituberculatus* (GenBank accession no. ACO07303). Until now, the crystal structure of crustins are not determined and Effects of various cysteine residues and spacing within cysteine-containing region on overall structure of crustin and its biological implications are still not known.

All crustins including three crustins from *P. japonica* contained WAP domain backboned by a four disulfide bonds (Grütter et al., 1988). WAP domain was originally identified in the whey acidic protein (WAP), omwaprin, antileukoproteinase and elastase-inhibitor proteins, which are different in biological functions (Nair et al., 2007; Seemüller et al., 1986; Wiedow et al., 1990). By searching against the PROSITE database (http://prosite.expasy.org/), the position of eight conserved cysteine residues which build up four disulfide bonds is C_1-X_n-C_2-X_n-C_3-X5-C_4-X5-C_5-C_6-X3, X5-C_7-X3, X4-C_8. Disulfide bonds occur in C1-C6, C2-C7, C3-C5, and C4-C8. WAP domain from Pj-crus Ia exhibited conserved cysteine numbers and spacing except for residues between C7 and C8. Conserved cysteine numbers and spacing of WAP domain was described by to Smith et al. (2008) and they proposed that three or four residues exist between C7 and C8 of most WAP domains. However, crustins from most decapods including Pj-crus Is and IIa showed five residues in the corresponding region (Zhang et al., 2007).

Exceptionally, a type II crustin from *M. rosenbergii* contains six residues between C7 and C8 and seven residues exist at the corresponding position of crustin from *P. camtschaticus*. Besides five residues between C7 and C8, Pj-crus Ib exhibited unusual six residues between C4 and C5, whereas other five residues exist in other crustins (Fig. 6). Unlike type I and II crustins, cysteine spacing and numbers followed the rule of WAP domain of most type III crustin. Although it is still not clear what is the biological implications of those variations in residues between cysteine residues, different residue numbers would affects on regional structure backboned by four disulfide bonds and further study is needed.

In addition to cysteine numbers and spacing (Fig. 6), lysine residue neighboring fifth cysteine residue is well conserved in WAP domain of all the crustins (-C-XX-D-XX-C-XXXX-K-CC-X-D-) (Jiravanichpaisal et al., 2007; Ranganathan S, 1999). Aspartic acid is also well conserved in type II crustins forming the conserved 'DKCC' sequences (-C-XX-D-XX-C-XXXD-K-CC-X-D-) (Tassanakajon et al., 2010). Exceptionally, both aspartic acid and lysine resiudes were identified in all three type I crustins from *E. sinensis* (GenBank Aceesion number: ACF25907, ACF25908, ACR77767). Those exceptional example is also shown in Pj-crus Ia, in which both residues lack in the corresponding region (Fig. 6). Although biological roles of Lysine and Aspartic acid residues are not clear, WAP domain is key factor for antimicrobial activity (Zhang et al., 2007) and their conserved cysteine spacing and residue neighboring conserved disulfide bonds may

play important role in exerting antimicrobial activity. At least, those two residues may be key residues for discriminating type II crustins from other crustins.

Isoelectric point (pI) of Pj-crus Ia and Pj-crus Ib were neutral (6.2 and 6.46) while Pj-crus IIa's pI value appears to be true cationic (8.73). Although crustins are known as a cationic peptides, pI values of known crustacean crustins varies from 5.0 to 8.7 (Smith et al., 2008). In order to know which domain contributes most in determining their pI values and what is implicated in biological functions, domain pI values of known crustins were also analyzed (Table 2). Although their overall pI values were from neutral to cationic, no significant contribution of each domain in determining pI value was detected and domain pI values varied in each crustin (Table 2). These results suggest that pI value may not be prerequisitive for antimicrobial functions of crustacean crustins.

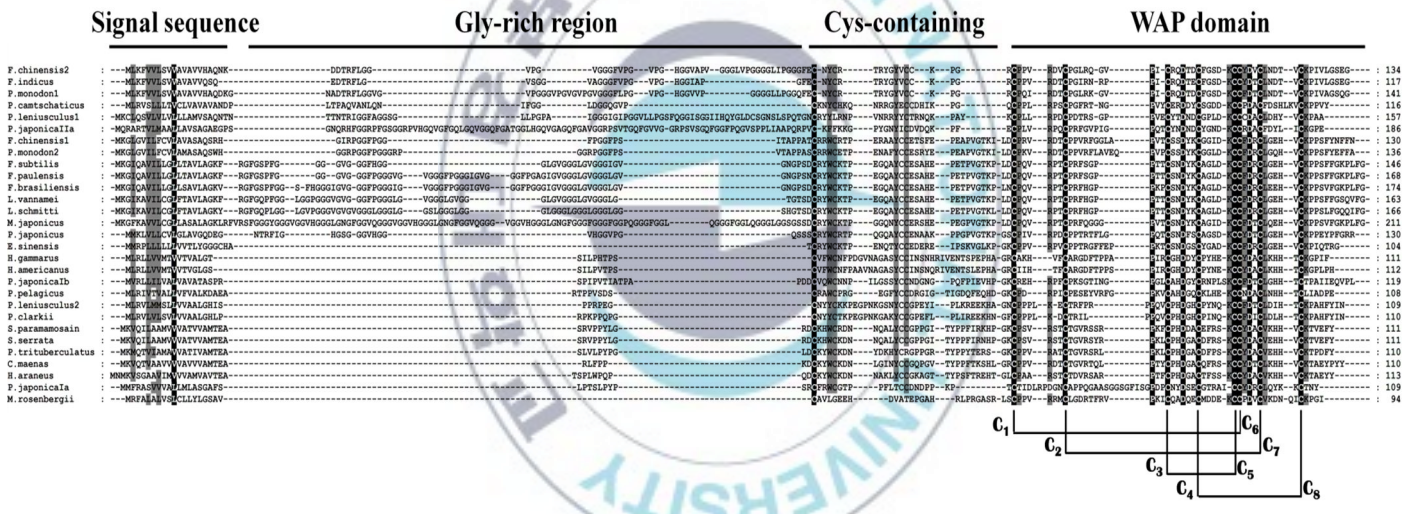
Signal sequence
Gly-rich region
Cys-containing
WAP domain
C1
C2
C3
C4
C5
C6
C7
C8

Fig. 6. Amino acid sequence multiple alignment of *P. japonica* Pj-crus Ia, Pj-crus Ib and Pj-crus IIa with crustin family in crustaceans. Black boxes are indicated similarity of amino acid sequences. The GenBank accession number : *Carcinus maenas*, CAD20734; *Eriocheir sinensis*, ACR77767; *Fenneropenaeus chinensis*, AAX63903, AAZ76017; *Farfantepenaeus paulensis*, ABM63361; *Farfantepenaeus brasiliensis*, ABQ96197; *Farfantepenaeus subtilis*, ABO93323; *Fenneropenaeus indicus*, ACV84092; *Homarus americanus*, ABM92333; *Homarus gammarus*, CAH10349; *Hyas araneus*, ACJ06763; *Litopenaeus schmitti*, ABM63362; *Litopenaeus vannamei*, AAS59736; *Macrobrachium rosenbergii*, ABQ41252; *Marsupenaeus japonicus*, BAD15064; *Pacifastacus leniusculus*, ABP88044, ABP88043; *Pandalopsis japonica*, JQ004015, JQ004016, JQ004017; *Paralithodes camtschaticus*, ACJ06765; *Panulirus japonicus*, ACU25382; *Penaeus monodon*, ABV25094, ABW82154; *Portunus pelagicus*, ABM65762; *Portunus trituberculatus*, ACO07303; *Procambarus clarkii*, AEB54630; *Scylla serrata*, ADW11096; *Scylla paramamosain*, ABY20728

Table 2. Predicted domain pI values of crustins

Peptide	Species	Accession number	Mature peptide	Proline -rich region	Glycine -rich region	Cysteine-containing region	WAP domain	Crustin type	Reference
Pj-crus Ia	*Pandalopsis japonica*	JQ004015	6.2			7.88	4.86	I	Present study
Pj-crus Ib	*Pandalopsis japonica*	JQ004016	6.46			4.66	7.74	I	Present study
Pj-crus IIa	*Pandalopsis japonica*	JQ004017	8.73		12.3	8	4.86	II	Present study
crustin antimicrobial peptide	*Scylla paramamosain*	ABY20728	8.99			9.38	8.5	I	(Imjongjirak et al., 2009)
crustin-like protein precursor	*Homarus americanus*	ABM92333	6.94			6.46	7.08	I	(Christie et al., 2007)
Crustin	*Farfantepenaeus paulensis*	ABM63361	8.29		11	4.96	8.5	II	(Rosa et al., 2007)
crustin-like peptide type 3	*Marsupenaeus japonicus*	BAD15064	8.48		9.99	6.73	8.20	II	(Rattanachai et al., 2004)
crustin-like antimicrobial peptide	*Penaeus monodon*	ABV25095	8.30		5.38	9.20	6.13	II	(Amparyup et al., 2008b)
single WAP domain-containing protein isoform 2	*Penaeus monodon*	ACF28467	8.74	12.3			6.23	III	(Amparyup et al., 2008a)
single WAP domain protein	*Litopenaeus vannamei*	AAS17722	8.97	12.6			6.19	III	(Jiménez-Vega et al., 2004)

Phylogenetic analysis of Pj-crus Ia,Pj-crus Ib and Pj-crus IIa

Phlylogenetic analysis was carried out using various crustins from decapods, including crabs, lobsters, shrimp, and crayfish. Since N-terminal regions of each crustin vary in length and amino acid composition, only WAP domain was used for the analysis. Phylogenetic analysis showed that type II crustins exhibited significant difference from type I and III crustins. We failed to identify distinct two subgroups; type I and III, by the analysis of WAP domain. Although there was no type III genomic sequence, those results corresponds to our genomic organization analysis of WAP domain. One splicing site was identified within WAP domain of type I crustins, whereas no splicing occurs within type II WAP domains suggesting this can reconfirm that the splicing site within WAP domain may be the key for classifying crustins. All type I crustins were originated from Pleocyemata, whereas most type II crustins were from Dendrobranchiata. Two type I crustins from *E. sinensis* were clustered together with type II crustins. Type III crustins were from both Pleocyemata and Dendrobranchiata. These results suggest that various crustins may have generated from gene duplications but their variation may occur within species according to various pathogens.

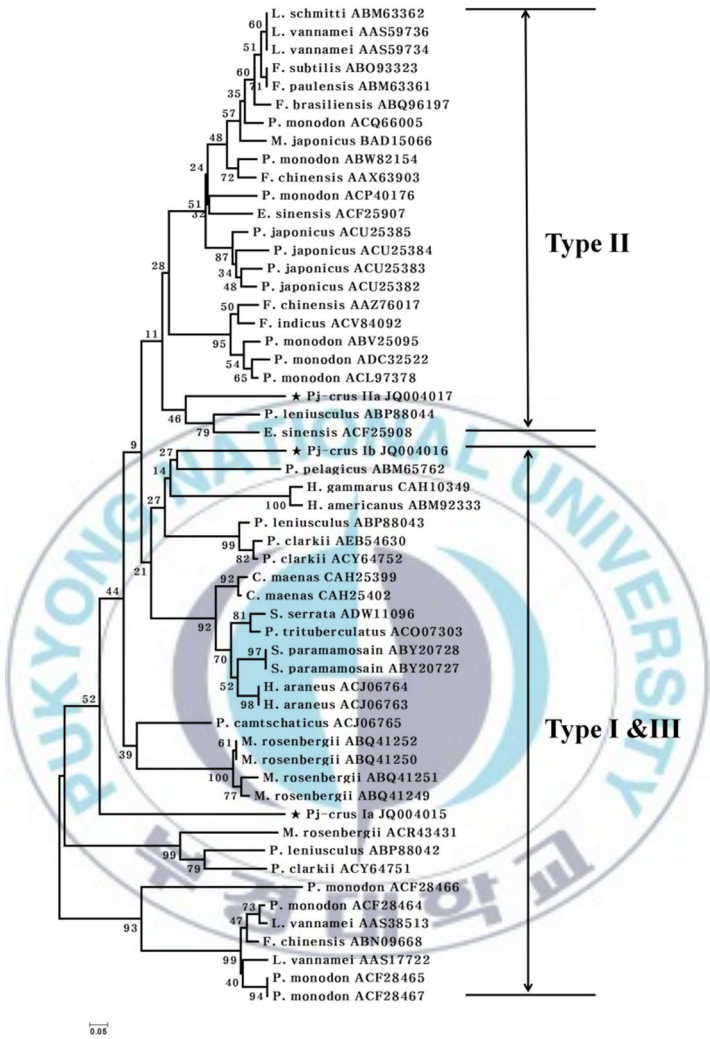

Fig. 7. Phylogenetic analysis of Pj-crus Ia, Pj-crus Ib and Pj-crus IIa with other decapod crustins. Diagram was generated and presented by minimum-evolution method with MEGA 4.0 program. Bootstrap replications were 1000. Crustins from *P. japonica* are marked by asterisks. Abbreviated species names and GenBank Accession Numbers were shown.

Tissue distribution analysis

The expression of three Pj-crus mRNAs in various tissues were determined by end-point RT-PCR (Fig. 8). Strong 18S rRNA PCR products in every tissue confirmed cDNA were properly synthesized. End-point RT-PCR results showed that Pj-crus Is are expressed in tissue-specific manners, whereas expression pattern of Pj-crus IIa is rather ubiquitous. Although both Pj-crus Ia and Ib were expressed in epidermis and brain, only Pj-crus Ib expression was detected in gill. Pj-crus I transcripts were not detected in the other tissues. By constrast, transcription of Pj-crus IIa was ubiquitous (Fig. 8). Interestingly, no Pj-crus expression was detected in hemocytes, which is different from previous studies (Brockton et al., 2007; Hauton et al., 2006; Relf et al., 1999; Supungul et al., 2008).

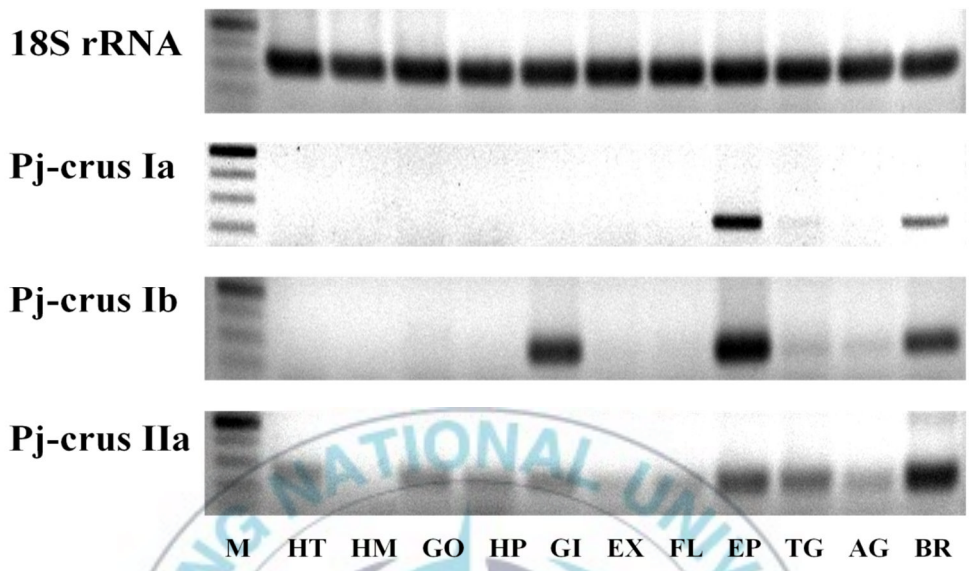

Fig. 8. Expression of Pj-crus Ia, Pj-crus Ib and Pj-crus IIa in various tissues. End-point RT-PCR was carried out for 30 cycles and products were separated on 1.5% agarose gel. the 18S rRNA was used as a control. M, size marker; HT, heart; HM, hemocyte; GO, Gonad; HP, hepatopancrease; GI, gill; EX, extensor muscle; FL, flexor muscle; EP, epidermis; TG, thoracic ganglia; AG, abdominal ganglia; BR, brain

Quantitative analysis of Pj-crus Ia, Ib and IIa mRNA transcription after immune challenge

In order to know effects of bacterial infection on each Pj-crus gene transcription, quantitative RT-PCR analysis was carried out after bacterial challenge (Fig. 9). One Gram-positive (*S. aureus*) and one Gram-negative (*V. parahaemolyticus*) strains were challenged and transcriptional levels were measured at first, second, and third days post injection. PBS buffer injection was considered as control group. Although relatively high transcription of Pj-crus Ia was detected in epidermis, no significant transcriptional change was detected (Fig. 9). Instead, Pj-crus Ia transcription was up-regulated in gill by 3.5-folds and 4.57-folds at 24 and 48 hours after *S. aureus* challenge, respectively (Fig. 9). No transcriptional change was detected against Gram-negative bacteria, *V. parahaemolyticus*. The Pj-crus Ib mRNA expression in epidermis was not induced but in the gill, which injected *S. aureus* after 24 hours and *V. parahaemolyticus* after 72 hours, were induced and increased 10.37-folds and 3.92-folds compared with the control group, respectively (Fig. 9). These results showed that Pj-crus Ib can respond not only to Gram positive but also Gram negative bacteria. The Pj-crus IIa mRNA expression in epidermis was not induced but in the gill, which injected *S. aureus* after 24 hours and *V. parahaemolyticus* after 72 hours, were induced and

increased 26.52-folds and 2.2-folds compared with the control group, respectively (Fig. 9). It is still not clear how Pj-crus genes were induced, whereas no specific response has been detected in epidermis. One possible explanation is that injected pathogens may interact directly with epithelial cell of gill since hemolymph circulates and pass through gill to exchange gas. Since we failed to detect induction of Pj-crus gene after bacterial injection (data not shown), we exclude possibility that those induced Pj-crus transcripts were not originated from hemocytes. Previous studies showed that hemocytes are major production sites for crustins but our data showed little expression in hemolymph suggesting different crustins exhibit different expression profile. In contrast, in the gills, all three crustin mRNA expression were significantly induced and increased by injecting *S. aureus* after 24 hours, whereas mRNA expression which is injected *V. parahaemolyticus* showing a small increase after 72 hours compared with post-injection of *S. aureus*.

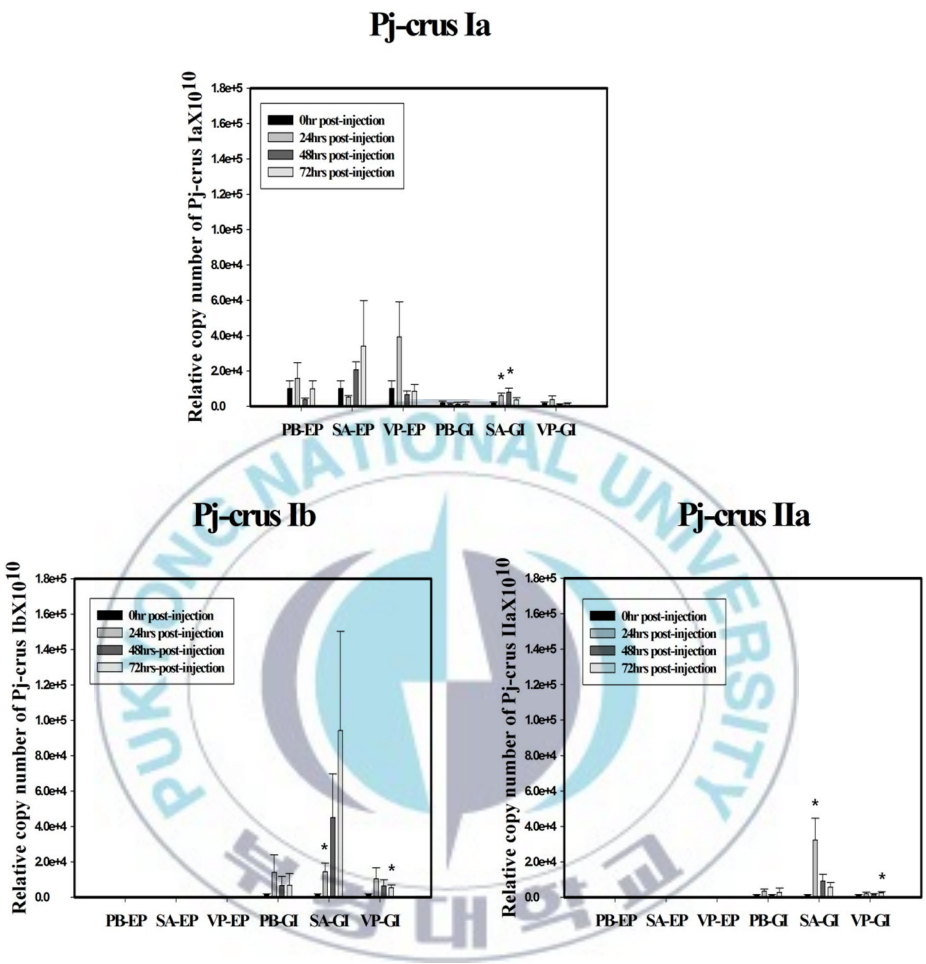

Fig. 9. Relative copy number of Pj-crus Ia, Ib and IIa between epidermis and gill from *P. japonica*. PB; PBS, SA; Staphylococcus aureus, VP; Vibrio parahaemolyticus, EP; epidermis, GI; gill. Copy numbers were normalized by the number of 18S rRNA. Statistical significance was accepted only when P <0.05 (*).

References

Afsal, V.V., Antony, S.P., Sathyan, N., Philip, R., 2011. Molecular characterization and phylogenetic analysis of two antimicrobial peptides: Anti-lipopolysaccharide factor and crustin from the brown mud crab, *Scylla serrata*. Results in Immunology.

Amparyup, P., Donpudsa, S., Tassanakajon, A., 2008a. Shrimp single WAP domain (SWD)-containing protein exhibits proteinase inhibitory and antimicrobial activities. Developmental & Comparative Immunology 32, 1497-1509.

Amparyup, P., Kondo, H., Hirono, I., Aoki, T., Tassanakajon, A., 2008b. Molecular cloning, genomic organization and recombinant expression of a crustin-like antimicrobial peptide from black tiger shrimp *Penaeus monodon*. Molecular Immunology 45, 1085-1093.

Antony, S.P., Bright, S.I.S., Philip, R., 2010. Molecular characterization of a crustin-like, putative antimicrobial peptide, Fi-crustin, from the Indian white shrimp, *Fenneropenaeus indicus*. Fish & Shellfish Immunology 28, 216.

Bartlett, T.C., Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S., Warr, G.W., 2002. Crustins, homologues of an 11.5-kDa antibacterial peptide, from two species of penaeid shrimp, *Litopenaeus vannamei* and *Litopenaeus setiferus*. Marine Biotechnology 4, 278-293.

Beale, K.M., Towle, D.W., Jayasundara, N., Smith, C.M., Shields, J.D., Small, H.J., Greenwood, S.J., 2008. Anti-lipopolysaccharide factors in the American lobster

Homarus americanus: Molecular characterization and transcriptional response to Vibrio fluvialis challenge. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 3, 263-269.

Boman, H.G., 1991. Antibacterial peptides: key components needed in immunity. Cell 65, 205-207.

Brockton, V., Hammond, J.A., Smith, V.J., 2007. Gene characterisation, isoforms and recombinant expression of carcinin, an antibacterial protein from the shore crab, *Carcinus maenas*. Molecular Immunology 44, 943-949.

Chiou, T.T., Lu, J.K., Wu, J.L., Chen, T.T., Ko, C.F., Chen, J.C., 2007. Expression and characterisation of tiger shrimp *Penaeus monodon* penaeidin (mo-penaeidin) in various tissues, during early embryonic development and moulting stages. Developmental & Comparative Immunology 31, 132-142.

Christie, A.E., Rus, S., Goiney, C.C., Smith, C.M., Towle, D.W., Dickinson, P.S., 2007. Identification and characterization of a cDNA encoding a crustin-like, putative antibacterial protein from the American lobster *Homarus americanus*. Molecular Immunology 44, 3333-3337.

Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S., 2002. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics 54, 442-445.

Dai, Z.M., Zhu, X.J., Yang, W.J., 2009. Full-length normalization subtractive hybridization: a novel method for generating differentially expressed cDNAs.

Molecular biotechnology 43, 257-263.

Destoumieux, D., Bulet, P., Loew, D., Van Dorsselaer, A., Rodriguez, J., Bachère, E., 1997. Penaeidins, a new family of antimicrobial peptides isolated from the shrimp *Penaeus vannamei* (Decapoda). Journal of Biological Chemistry 272, 28398.

Destoumieux, D., Munoz, M., Bulet, P., Bachère*, E., 2000. Penaeidins, a family of antimicrobial peptides from penaeid shrimp (Crustacea, Decapoda). Cellular and Molecular Life Sciences 57, 1260-1271.

Grütter, M.G., Fendrich, G., Huber, R., Bode, W., 1988. The 2.5 A X-ray crystal structure of the acid-stable proteinase inhibitor from human mucous secretions analysed in its complex with bovine alpha-chymotrypsin. The EMBO journal 7, 345.

Hauton, C., Brockton, V., Smith, V.J., 2006. Cloning of a crustin-like, single whey-acidic-domain, antibacterial peptide from the haemocytes of the European lobster, *Homarus gammarus*, and its response to infection with bacteria. Molecular Immunology 43, 1490-1496.

Ho, S.H., Chao, Y.C., Tsao, H.W., Sakai, M., Chou, H.N., Song, Y.L., 2004. Molecular cloning and recombinant expression of tiger shrimp *Penaeus monodon* penaeidin. Fish Pathology 39, 15-24.

Imjongjirak, C., Amparyup, P., Tassanakajon, A., Sittipraneed, S., 2009. Molecular cloning and characterization of crustin from mud crab *Scylla paramamosain*.

Molecular biology reports 36, 841-850.

Jeon, J.M., Lee, S.O., Kim, K.S., Baek, H.J., Kim, S., Kim, I.K., Mykles, D.L., Kim, H.W., 2010. Characterization of two vitellogenin cDNAs from a Pandalus shrimp (*Pandalopsis japonica*): Expression in hepatopancreas is down-regulated by endosulfan exposure. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 157, 102-112.

Jia, Y.P., Sun, Y.D., Wang, Z.H., Wang, Q., Wang, X.W., Zhao, X.F., Wang, J.X., 2008. A single whey acidic protein domain (SWD)-containing peptide from fleshy prawn with antimicrobial and proteinase inhibitory activities. Aquaculture 284, 246-259.

Jiménez-Vega, F., Yepiz-Plascencia, G., Söderhäll, K., Vargas-Albores, F., 2004. A single WAP domain-containing protein from *Litopenaeus vannamei* hemocytes. Biochemical and biophysical research communications 314, 681-687.

Jiravanichpaisal, P., Lee, S.Y., Kim, Y.A., Andren, T., Soderhall, I., 2007. Antibacterial peptides in hemocytes and hematopoietic tissue from freshwater crayfish *Pacifastacus leniusculus*: characterization and expression pattern. Developmental & Comparative Immunology 31, 441-455.

Kang, C.J., Xue, J.F., Liu, N., Zhao, X.F., Wang, J.X., 2007. Characterization and expression of a new subfamily member of penaeidin antimicrobial peptides (penaeidin 5) from *Fenneropenaeus chinensis*. Molecular Immunology 44, 1535-1543.

Kim, K.S., Kim, B.K., Kim, H.J., Yoo, M.S., Mykles, D.L., Kim, H.W., 2008. Pancreatic lipase-related protein (PY-PLRP) highly expressed in the vitellogenic ovary of the scallop, *Patinopecten yessoensis*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 151, 52-58.

Lee, S.O., Jeon, J.M., Oh, C.W., Kim, Y.M., Kang, C.K., Lee, D.S., Mykles, D.L., Kim, H.W., 2011. Two juvenile hormone esterase-like carboxylesterase cDNAs from a Pandalus shrimp (*Pandalopsis japonica*): Cloning, tissue expression, and effects of eyestalk ablation. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology.

Liu, H., Jiravanichpaisal, P., Soderhall, I., Cerenius, L., Soderhall, K., 2006. Antilipopolysaccharide factor interferes with white spot syndrome virus replication in vitro and in vivo in the crayfish *Pacifastacus leniusculus*. Journal of virology 80, 10365.

Mu, C., Zheng, P., Zhao, J., Wang, L., Qiu, L., Zhang, H., Gai, Y., Song, L., 2011. A novel type III crustin (CrusEs2) identified from Chinese mitten crab *Eriocheir sinensis*. Fish & Shellfish Immunology.

Nair, D.G., Fry, B.G., Alewood, P., Kumar, P.P., Kini, R.M., 2007. Antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins. Biochemical Journal 402, 93.

O'Leary, N.A., Gross, P.S., 2006. Genomic structure and transcriptional regulation of the penaeidin gene family from *Litopenaeus vannamei*. Gene 371, 75-83.

Pisuttharachai, D., Fagutao, F.F., Yasuike, M., Aono, H., Yano, Y., Murakami, K., Kondo, H., Aoki, T., Hirono, I., 2009. Characterization of crustin antimicrobial proteins from Japanese spiny lobster *Panulirus japonicus*. Developmental & Comparative Immunology 33, 1049-1054.

Ranganathan S, S.K., Shaw DC, Nicholas K., 1999. <The whey acidic protein family-A new signature motif and three-dimensional structure by comparative modeling.pdf>. J Mol Graph Model 17, 13.

Rattanachai, A., Hirono, I., Ohira, T., Takahashi, Y., Aoki, T., 2004. Cloning of kuruma prawn *Marsupenaeus japonicus* crustin-like peptide cDNA and analysis of its expression. Fisheries Science 70, 765-771.

Relf, J.M., Chisholm, J.R.S., Kemp, G.D., Smith, V.J., 1999. Purification and characterization of a cysteine-rich 11.5-kDa antibacterial protein from the granular haemocytes of the shore crab, *Carcinus maenas*. European Journal of Biochemistry 264, 350-357.

Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y., Aoki, T., 2002. Gene expression in haemocytes of kuruma prawn, *Penaeus japonicus*, in response to infection with WSSV by EST approach. Fish & Shellfish Immunology 13, 69-83.

Rosa, R.D., Bandeira, P.T., Barracco, M.A., 2007. Molecular cloning of crustins from the hemocytes of Brazilian penaeid shrimps. FEMS microbiology letters 274, 287-290.

Rosa, R.D., Barracco, M.A., 2010. Antimicrobial peptides in crustaceans. ISJ 7,

262-284.

Rosa, R.D., Stoco, P.H., Barracco, M.A., 2008. Cloning and characterisation of cDNA sequences encoding for anti-lipopolysaccharide factors (ALFs) in Brazilian palaemonid and penaeid shrimps. Fish & Shellfish Immunology 25, 693-696.

Schnapp, D., Kemp, G.D., Smith, V.J., 1996. Purification and Characterization of a Proline-rich Antibacterial Peptide, with Sequence Similarity to Bactenecin-7, from the Haemocytes of the Shore Crab, *Carcinus Maenas*. European Journal of Biochemistry 240, 532-539.

Seemüller, U., Arnhold, M., Fritz, H., Wiedenmann, K., Machleidt, W., Heinzel, R., Appelhans, H., Gassen, H.G., Lottspeich, F., 1986. The acid-stable proteinase inhibitor of human mucous secretions (HUSI-I, antileukoprotease):: Complete amino acid sequence as revealed by protein and cDNA sequencing and structural homology to whey proteins and Red Sea turtle proteinase inhibitor. FEBS letters 199, 43-48.

Shi, X.Z., Zhang, R.R., Jia, Y.P., Zhao, X.F., Yu, X.Q., Wang, J.X., 2009. Identification and molecular characterization of a Spätzle-like protein from Chinese shrimp (*Fenneropenaeus chinensis*). Fish & Shellfish Immunology 27, 610-617.

Smith, V.J., Fernandes, J.M.O., Kemp, G.D., Hauton, C., 2008. Crustins: enigmatic WAP domain-containing antibacterial proteins from crustaceans. Developmental & Comparative Immunology 32, 758-772.

Somboonwiwat, K., Bachère, E., Rimphanitchayakit, V., Tassanakajon, A., 2008. Localization of anti-lipopolysaccharide factor (ALFPm3) in tissues of the black tiger shrimp, *Penaeus monodon*, and characterization of its binding properties. Developmental & Comparative Immunology 32, 1170-1176.

Sperstad, S.V., Haug, T., Paulsen, V., Rode, T.M., Strandskog, G., Solem, S.T., Styrvold, O.B., Stensvåg, K., 2009. Characterization of crustins from the hemocytes of the spider crab, *Hyas araneus*, and the red king crab, *Paralithodes camtschaticus*. Developmental & Comparative Immunology 33, 583-591.

Sun, C., Du, X.J., Xu, W.T., Zhang, H.W., Zhao, X.F., Wang, J.X., 2010. Molecular cloning and characterization of three crustins from the Chinese white shrimp, *Fenneropenaeus chinensis*. Fish & Shellfish Immunology 28, 517-524.

Supungul, P., Klinbunga, S., Pichyangkura, R., Hirono, I., Aoki, T., Tassanakajon, A., 2004. Antimicrobial peptides discovered in the black tiger shrimp *Penaeus monodon* using the EST approach. Diseases of aquatic organisms 61, 123-135.

Supungul, P., Tang, S., Maneeruttanarungroj, C., Rimphanitchayakit, V., Hirono, I., Aoki, T., Tassanakajon, A., 2008. Cloning, expression and antimicrobial activity of crustinPm1, a major isoform of crustin, from the black tiger shrimp *Penaeus monodon*. Developmental & Comparative Immunology 32, 61-70.

Tassanakajon, A., Amparyup, P., Somboonwiwat, K., Supungul, P., 2010. Cationic antimicrobial peptides in penaeid shrimp. Marine Biotechnology 12, 487-505.

Vargas-Albores, F., Yepiz-Plascencia, G., Jiménez-Vega, F., Ávila-Villa, A., 2004.

Structural and functional differences of *Litopenaeus vannamei* crustins. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 138, 415-422.

Vatanavicharn, T., Supungul, P., Puanglarp, N., Yingvilasprasert, W., Tassanakajon, A., 2009. Genomic structure, expression pattern and functional characterization of crustinPm5, a unique isoform of crustin from *Penaeus monodon*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 153, 244-252.

Wiedow, O., Schröder, J., Gregory, H., Young, J., Christophers, E., 1990. Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence. Journal of Biological Chemistry 265, 14791.

Yue, F., Pan, L., Miao, J., Zhang, L., Li, J., 2010. Molecular cloning, characterization and mRNA expression of two antibacterial peptides: Crustin and anti-lipopolysaccharide factor in swimming crab *Portunus trituberculatus*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 156, 77-85.

Zhang, J., Li, F., Wang, Z., Xiang, J., 2007. Cloning and recombinant expression of a crustin-like gene from Chinese shrimp, *Fenneropenaeus chinensis*. Journal of biotechnology 127, 605-614.

Zhang, L., Luo, L., 2003. Splice site prediction with quadratic discriminant analysis using diversity measure. Nucleic acids research 31, 6214.


PUKYONG NATIONAL UNIVERSITY
부경대학교