Journal of proteome research

Article

pubs.acs.org/jpr

First Proteome of the Egg Perivitelline Fluid of a Freshwater Gastropod with Aerial Oviposition

Jin Sun,[†,⊥] Huoming Zhang,[†,∥,⊥] Hao Wang,[‡] Horacio Heras,[§] Marcos S. Dreon,[§] Santiago Ituarte,[§] Timothy Ravasi,[∥] Pei-Yuan Qian,[‡] and Jian-Wen Qiu*[,†]

[†]Department of Biology, Hong Kong Baptist University, Hong Kong, China

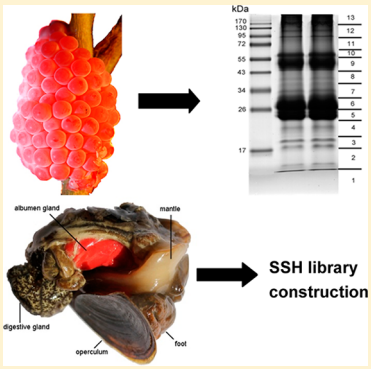
[‡]Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, China

[§]Instituto de Investigaciones Bioquímicas de La Plata (INIBIOLP), CONICET CCT La Plata — Universidad Nacional de La Plata (UNLP), 60 y 120, (1900) La Plata, Argentina

[∥]King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

Supporting Information

ABSTRACT: *Pomacea canaliculata* is a freshwater snail that deposits eggs on solid substrates above the water surface. Previous studies have emphasized the nutritional and protective functions of the three most abundant perivitelline fluid (PVF) protein complexes (ovorubin, PV2, and PV3) during its embryonic development, but little is known about the structure and function of other less abundant proteins. Using 2-DE, SDS-PAGE, MALDI TOF/TOF, and LC−MS/MS, we identified 59 proteins from the PVF of *P. canaliculata*, among which 19 are novel. KEGG analysis showed that the functions of the majority of these proteins are "unknown" (n = 34), "environmental information processing" (10), 9 of which are related to innate immunity, and "metabolism" (7). Suppressive subtractive hybridization revealed 21 PVF genes to be specific to the albumen gland, indicating this organ is the origin of many of the PVF proteins. Further, the 3 ovorubin subunits were identified with 30.2−35.0% identity among them, indicating their common origin but ancient duplications. Characterization of the PVF proteome has opened the gate for further studies aiming to understand the evolution of the novel proteins and their contribution to the switch to aerial oviposition.



KEYWORDS: *Pomacea canaliculata, apple snail, perivitelline fluid, proteomics, novel protein*

INTRODUCTION

Freshwater gastropods appear to have few egg deposition strategies. While some taxa such as viviparids and thiarids are ovoviviparous, the majority of freshwater gastropods lay fertilized eggs in a jelly mass on submerged objects.[1] Unlike most other freshwater gastropods, however, the family Ampullariidae (Caeogastropoda) has representatives that deposit eggs both below the water surface on submerged surfaces (genera *Asolene*, *Felipponea*, *Lanistes* and *Marisa*) and above the water surface on various solid objects such as rocks and vegetation stems (genera *Pila* and *Pomacea*) (Figure 1).[2,3] The change to aerial oviposition in ampullariids is a synapomorphy[3] and is associated with changes in a suite of morphological characters in the adult, such as longer respiratory siphon and increased lung size, which are believed to have contributed to their adaptation to adverse environmental conditions, leading to the success of the aerial depositing clade especially *Pomacea*, the most species abundant and widely distributed ampullariids.[3−6]

Embryos in aerial eggs are often exposed to stressful environmental conditions such as strong UV radiation, desiccation and thermal stress, as well as terrestrial predators due to their high nutritional value. Therefore the shift from underwater oviposition to aerial oviposition in ampullariids is expected to require drastic changes in the egg structure and composition. Indeed, ampullariids laying aerial eggs have a calcareous egg shell, whereas those laying eggs underwater have a gelatinous egg shell.[3] The perivitelline fluid (PVF) surrounding the developing embryo contains not only all the nutrients and energy donors for the embryonic development,[7] but also many macromolecules that protect the eggs from environmental stressors and predators.[8−10] Three major perivitellins forming glyco-lipoproteic complexes are present in the perivitelline fluid of *Pomacea canaliculata*. They were named ovorubin,[11] PV2 and PV3.[12] However, to avoid confusion with the major PVF proteins of other species of *Pomacea*, such as *Pomacea scalaris*, these proteins are hereafter renamed with a species specific prefix "Pc" as PcOvo, PcPV2 and PcPV3, respectively.

PcOvo, a glycoprotein with carotenoid astaxanthin as prosthetic group, is the major perivitellin in the PVF, accounting for 57.0% of the total protein content.[12] Perivitellin complex PcPV2 and PcPV3 accounts for 7.5 and 35.5% of the

Received: April 15, 2012
Published: June 28, 2012

   dx.doi.org/10.1021/pr3003613 | *J. Proteome Res.* 2012, 11, 4240−4248

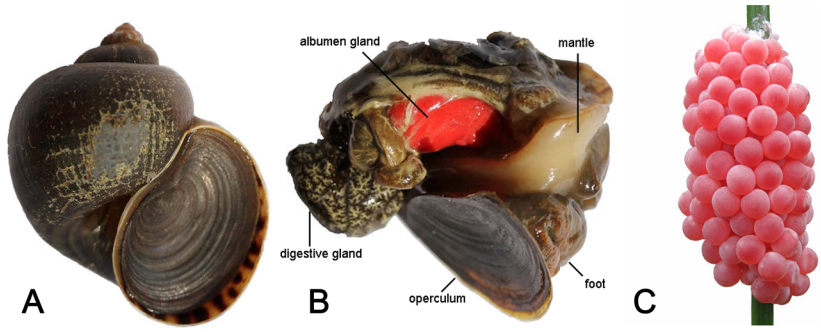


Figure 1. Pictures of a female *Pomacea canaliculata* with an intact shell (shell length 4.0 cm) (A); the same individual with the shell removed to show several body structures including albumen gland (B); an egg clutch (each approximately 2.5 mm in diameter) (C).

protein content, respectively. PcOvo is an oligomer of 300 kDa composed of at least three subunits of 35, 32, and 28 kDa. PcPV2 is an octamer of 400 kDa composed of four 98 kDa heterodimers consisting of 67 and 31 kDa subunits,[13] while PcPV3 is a heterogeneous fraction with subunits of 34, 29, and 26 kDa.[12] These perivitellin complexes not only are a source of energy and nutrients for the embryo, but also function in egg defense against predation. In this regard, PcOvo and PcPV3 provide for the conspicuous red coloration of the eggs and presumably advertise to visual-hunting predators the presence of egg defense (aposematic warning).[2] Besides, PcOvo belongs to the small Kunitz type serine protease inhibitors family. It displays high structural stability, which allows it to withstand the harsh/hostile conditions of the digestive tract of rats, playing an antinutritive and antidigestive role, decreasing the standard growth rate of this potential predator.[10] This defense is complemented by PcPV2, reported as a strong proteinaceous neurotoxin lethal to mice.[9]

Although the above studies have greatly enhanced our understanding of the biochemical and biophysical properties of some of the PVF proteins in *P. canaliculata*, the amino acid sequences, except for the N-terminal sequences of PcPV2 and one subunit of PcOvo, have not been reported, which hinders further functional studies of these proteins. Besides, while these previous studies of *P. canaliculata* have focused on the three major perivitellin complexes, nothing is known about the presence of the less abundant proteins and their roles in the embryonic development. In fact, only few studies of the egg proteins have been conducted in Gastropoda. Among the few reports in this regard was a recent study of the bloodfluke planorb *Biomphalaria glabrata* using a combination of SDS-PAGE and MALDI-TOF/TOF techniques, which identified 20 proteins in the egg mass fluid (most of which was believed to be PVF), more than doubling the number of previously reported proteins in the egg mass fluid of this gastropod.[14] This study of *B. glabrata* revealed that the functions of maternally invested proteins in the eggs of a gastropod can be highly diverse. Using SDS-PAGE and LC−MS/MS, Aagaard et al.[15] identified several novel egg coating proteins from abalones *Haliotis* spp. and showed that positive selection of these reproductive proteins is involved in the evolution of abalone reproduction.

To provide a comprehensive profile of the proteome of *Pomacea canaliculata* PVF, we used a combination of 2-DE coupled with MALDI TOF/TOF and SDS-PAGE coupled with LC−MS/MS to determine protein components in PVF and suppressive subtractive hybridization to determine whether these PVF proteins are secreted directly by the albumen gland. In addition, we aimed to provide full length sequences for several abundant PVF proteins including PcOvo through cloning of their open reading frames (ORF).

MATERIALS AND METHODS

Collection and Maintenance of Apple Snails

Approximately 30 adults of *P. canaliculata* (25−35 mm shell length) were collected from a drainage channel in Yuen Long (22°15′ N, 114°10′ E) and maintained at 26 ± 1 °C in a 250-L laboratory aquarium at Hong Kong Baptist University. A canister filter was used to clean the water, and an electric air pump was used to provide aeration. The snails were fed with lettuce and carrot twice a week. Under these laboratory conditions, the snails mated and deposited eggs in clutches above water surface on the walls of the aquarium.

Protein Extraction

Three biological replicates of approximate 100 freshly deposited eggs (within 6 h of deposition) were collected. Egg shells were broken gently with a needle and PVF was extracted using a 10 μL pipet. The pooled PVF samples were centrifuged at 3000*g* for 10 min. The supernatant of each sample was then mixed with an equal volume of 8 M urea and further purified using 2D cleanup kit (Bio-Rad, Hercules, USA). The protein concentration was determined by the RC-DC kit (Bio-Rad), and each sample was divided for use in the following 2-DE and SDS-PAGE experiments.

2-DE and MALDI TOF/TOF

This experiment was conducted to separate and identify the proteins in the PVF samples. The procedures were modified from those previously described.[16] In brief, 0.3 mg protein samples were subject to isoelectric focusing using a 17 cm IPG strip (Bio-Rad) of pH 3−10. The focused strips were subject to the second dimension of electrophoresis using 12.5% SDS-PAGE and stained using colloidal coomassie blue (CCB).[17] Protein spots present on all triplicate samples were excised from the gels, destained twice with 50 mM NH_4HCO_3 in 50% methanol, washed twice with Milli-Q water, dried and rehydrated twice with 100% ACN and 100 mM NH_4HCO_3, respectively, and digested with sequencing grade trypsin (Promega, Madison, USA) in 50 mM NH_4HCO_3. The trypsin-digested peptides were analyzed using a MALDI TOF/TOF mass spectrometer (Autoflex III smartbeam, Bruker Daltonik, Bremen, Germany).[16]

SDS-PAGE and LC−MS/MS

Although 2-DE followed by MALDI TOF/TOF is a proven method for protein separation and identification, it suffers from

dx.doi.org/10.1021/pr3003613 | J. Proteome Res. 2012, 11, 4240−4248

low sensitivity and low reproducibility. Therefore, we also analyzed the protein samples by SDS-PAGE coupled with LC–MS/MS. Each biological replicate was divided into three technical replicates. The purified protein samples were each mixed with SDS-PAGE buffer (10% SDS, 10 mM dithiothreitol, 30% glycerol, 0.2 M Tris-HCl pH 6.8, and 0.05% bromophenol blue) at a 1:3 volume ratio. Proteins in each sample were separated by SDS-PAGE and stained by CCB. After destaining with Milli-Q water, each gel was cut into 13 slices with proteins of different mass (Figure 2). Slices with low abundance protein

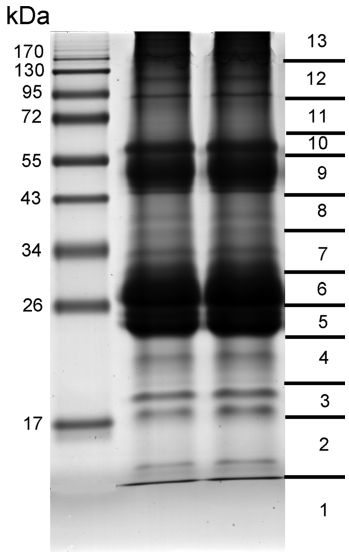


Figure 2. SDS-PAGE gel image of PVF showing two technical replicates and the horizontal slices for further LTQ-Orbitrap analysis.

bands were pooled from the technical replicates to increase the chance of identification for these proteins. Proteins underwent in-gel trypsin digestion as above.

The fractions were dried using SpeedVac, and each fraction was reconstituted in 10 μL of 0.1% formic acid and analyzed three times using a LTQ-Orbitrap mass spectrometer[18] (Velos, Thermo Electron, San Jose, USA) coupled with a Proxeon EASY-nLC unit (Bruker Daltonik, Bremen, Germany). For each analysis, 3 μL sample was concentrated and desalted in a preconditioned column (0.3 × 50 mm) packed with C18 AQ (5 μm particles, 200 Å pore size, Michrom BioResources, Auburn, USA). The peptide separation was performed in a capillary column (0.1 × 150 mm, with C18 AQ of 3 μm particles and 200 Å pore size, Michrom BioResources). Mobile phase A (0.1% formic acid in H_2O) and mobile phase B (0.1% formic acid in ACN) were used to establish a 75-min gradient including 45 min from 100 to 65% A, 10 min from 65 to 20% A, and 20 min at 20% A. The LC was operated at a flow rate of 500 nL/min. Each sample was injected into the LTQ-Orbitrap through an ADVANCE CaptiveSpray Source (Michrom BioResources) with an electrospray potential of 1.5 kV. Temperature at the ion transfer tube was set to 160 °C. The LTQ-Orbitrap was set to perform data acquisition in the positive ion mode. A full MS scan (350–1600 m/z range) was acquired in the Orbitrap at a resolution of 60 000 (at 400 m/z) in a profile mode, a maximum ion accumulation time of 1 s and a target value of 1×10^6. Precursor ion charge state screening was activated. The linear ion trap was used to collect peptides, with the selection of the 10 most intense ions above a 500-count threshold for collision-induced dissociation (CID) in MS/MS, which was performed concurrently with a maximum ion accumulation time of 200 ms. Dynamic exclusion was activated for this process, with a repeat count of 2, exclusion duration of 45 s, and ±5 ppm mass tolerance. Precursors with a charge state of 1 and unassigned charge states were also excluded. The following parameters were set for CID: normalized collision energy = 35%, activation Q = 0.25, isolation width = 3.0, and activation time = 10 ms.

Database Search

The MS data were searched against a "target-decoy"[19] protein database based on our recent transcriptome analysis of *P. canaliculata*.[20] This database includes three components: 43 565 *P. canaliculata* protein sequences (target), their reversed sequences (decoy), and keratin sequences from *Homo sapiens* downloaded from the IPI database to allow for detection of experimental contamination. Data generated from MALDI TOF/TOF for each of the 2-DE protein spots were submitted to Mascot version 2.2 (Matrix Sciences Ltd., London, UK) with the following criteria: 75 ppm for the precursor and 0.3 Da for the fragment masses; fixed modification: carbamidomethyl (cysteine); variable modification: oxidation (methionine). Up to one missed trypsin cleavage was allowed. Protein identification was considered significant only when the ions score for peptide assignments was above the 95% confidence level.

For the mass spectrometry data generated from the LTQ Orbitrap, the MS raw files were first processed with Proteome Discoverer 1.0 (Thermo Fisher Scientific, San Jose, USA) to extract separate Mascot generic files. Database search was performed using Mascot version 2.2 (Matrix Sciences Ltd., London, UK) against the combined database as mentioned above. The search criteria were 5 ppm for precursor and 0.6 Da for fragments; fixed modification: carbamidomethyl (cysteine); variable modification: oxidation (methionine). Up to one missed trypsin cleavage was allowed. To achieve high confidence identification, peptide matches were filtered with an expectation value of less than 0.05 (expect ion score over 20) in the Mascot search. We then applied a none false discovery rate threshold in the final protein identification in each replicate; i.e., the threshold protein score in Mascot search for accepted identifications was dynamically assigned in order to eliminate all of the reversed hits in each replicate. Proteins containing only one matched spectrum were also removed. The abundance of each identified protein was measured by emPAI in Mascot, a label-free quantification method based on the number of sequenced peptides per protein.[21]

Identification of Genes Specifically Expressed in the Albumen Gland

A suppressive subtractive hybridization (SSH) cDNA library was constructed using a previously described method[22] to obtain the set of genes expressed only in the albumen gland (AG) of female *P. canaliculata*. The red colored AG tissue as well as other tissues including mantle, foot and digestive gland from three individuals were dissected and pooled. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, USA). Polysaccharides were removed from the RNA sample by adding a high salt solution (0.8 M sodium citrate and 1.2 M NaCl) in the RNA isopropanol precipitation step. Messenger RNA was enriched using a PolyATract mRNA Isolation System (Ambion, Austin, USA) according to the manufacturer's protocol. Synthesis of AG (tester) and other snail tissues

dx.doi.org/10.1021/pr3003613 | J. Proteome Res. 2012, 11, 4240–4248

(driver) cDNA and SSH were performed using a PCR-Select cDNA Subtraction Kit (Clontech, CLONTECH Laboratories, Palo Alto, USA) following the manufacture's protocol. The PCR products were ligated into a pMD18-simple T vector (TaKaRa, Dalian, China). All major colonies were picked up, and their inserts were sequenced. KEGG functional classification was performed using the Blastall software against the Kyoto Encyclopedia of Genes and Genomes database for both the PVF protein list and SSH gene list.

PcOvo Purification, Subunit Identification and Analysis

The PcOvo glyco-lipocarotenoprotein complex was purified according to Dreon et al.[23] The protein was first deglycosilated using trifluoromethansulfonic acid (TFMS, Sigma Chemical Co, St. Louis, USA) and then subjected to 2-DE analysis as described by Ituarte et al.[24] The protein spots were cut from the gel, in-gel digested with trypsin, and the fragments were analyzed by mass spectrometry in an ion trap (LCQ, ThermoQuest, San Jose, CA, USA) at the Proteomic Service, National Centre of Biotechnology, Madrid, Spain. The mgf files generated after MS/MS analysis were submitted to Mascot version 2.2 (Matrix Sciences Ltd., London, UK) to be searched against our *P. canaliculata* protein database as mentioned above with criteria of 75 ppm for the precursor and 0.3 Da for the fragment masses; fixed modification: carbamidomethyl (cysteine); variable modification: oxidation (methionine). Protein identification was considered significant only when the ions score for peptide assignments had to be above the 95% confidence level.

RACE cloning was performed as described in Wang and Qian[22] for PcOvo subunits. Sequence alignment was performed by using Clustal_X v 2.1 under the default settings and further manually corrected. SignalP 4.0 was used to check the translated amino acid sequences in order to determine whether they are secreted proteins.[25] Phylogenetic analysis was performed using MrBayes version 3.2.0[26] with the "mixed-model" approach and MEGA 5.0[27] to determine the sequence similarity of several protein homologues. Four chains were applied, with runs of 2.5 million generations, chains sampled every 100 generations, and a burn-in of 10 000 trees. The WAG model was applied as the best fitting substitution model in MrBayes. Amino acid substitution was calculated using the Jones–Taylor–Thornton model, and the maximum likelihood method was used to construct the tree. Phylogenetic test was calculated from 100 of bootstrap replicates in MEGA 5.0.

RESULTS

The 2-DE detected 26 ± 5 (mean ± standard deviation) protein spots in three gels (Figure 3). All of the 17 spots common to the three replicates were identified by MALDI TOF/TOF and database search (see Table 1 for a summary of the results, Supporting Information 1 for protein sequences, and Supporting Information 2 for details of protein identification). Most of them, except Pc123838, contained isoforms having identical peptide sequences but probably different post-translational modifications deduced from their different positions in the gel. These 17 spots belong to 8 proteins.

The LC–MS/MS detected 355 unique peptides from a total of 32 428 peptides. These unique peptides were assigned to 59 proteins (see Table 1 for a summary of the results and Supporting Information 1 for protein sequences). These included all of the 8 proteins detected by 2-DE. Analysis of

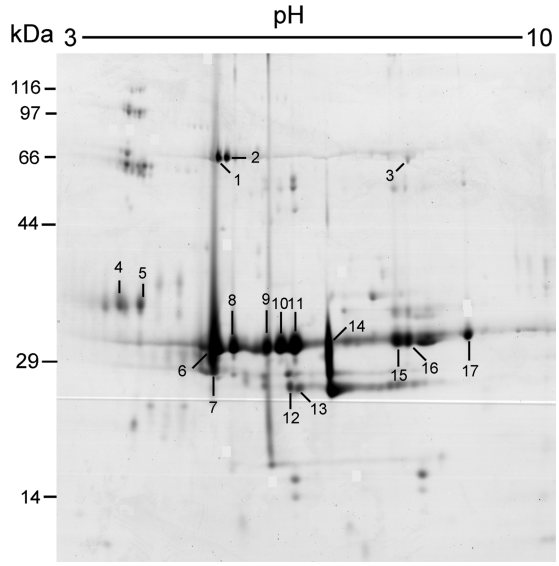


Figure 3. Representative 2-DE gel image showing the PVF protein spots selected for mass spectrometry analysis. The identified proteins are listed in Table 1.

these 59 proteins using emPAI, a label-free protein quantification method, revealed a high dynamic range of abundance (0.02–48.61%), indicating our analytical approach is sensitive enough to detect very low abundance proteins. Among these 59 proteins, 19 (including 3 PcOvo subunits, see next paragraph) are novel with no homologue in NCBI nonredundant protein database. KEGG functional analysis classified the 59 proteins into "unknown" (n = 34), "environmental information processing" (10), "metabolism" (7), "organismal systems" (2), cellular processes (3), and others (3) (Figure 4A). Among the "environmental information processing" cluster, 9 proteins (e.g., thioester-containing protein, transferrin, melanotransferrin) are related to the innate immunity.

AG specific genes were revealed by SSH methodology. A total of 209 AG specific clones were sequenced, which contained 44 nonredundant sequences. The unique SSH sequences were matched to our transcriptome database in order to obtain longer sequences. The matched unigene names in the transcriptome (see Table 1) were annotated by adding "SSH" as a prefix. The purified PcOvo was analyzed by 2-DE (Supporting Information 3), and its major spots were analyzed by MS/MS (Supporting Information 4) and searched against our *P. canaliculata* protein database, which revealed that this protein complex comprised at least three protein subunits: SSH9 (annotated as PcOvo1), SSH4 (annotated as PcOvo2) and SSH8 (annotated as PcOvo3). A multiple alignment of the full-length ORF of these three subunits obtained by RACE showed that they shared two conserved motifs, GXSWPR and one N-link glycosylation site NXTL (Figure 5). KEGG analysis of the 44 unique SSH sequences showed that most (41) of the genes are of "unknown" function, two are "organism systems" and one is "metabolism" (Figure 4B). Besides the three PcOvo full-length proteins, 7 full-length sequences were obtained for other PVF proteins, among which 9 had predicted signal peptides. These 10 full-length sequences were deposited in NCBI GenBank (accession numbers: JQ818214–JQ818223). A self-BLAST under a threshold of 1.0×10^{-5} revealed 31 pairs of homologues among the PVF proteins (Table 2).

dx.doi.org/10.1021/pr3003613 | J. Proteome Res. 2012, 11, 4240–4248

Table 1. Proteins Identified from the PVF of *Pomacea canaliculata*

unigenes[a]	protein name[b]	no. of unique peptides	no. of accepted peptides	coverage %[c]	avg %[d]	2-DE spot no.
SSH9	PcOvo1 [SF]	28	7259	87	48.61 ± 4.10	6,7
SSH2	novel protein [SF]	20	4231	66	9.50 ± 2.07	15,16,17
SSH20	novel protein	12	2821	45	7.15 ± 1.30	9,10,11
SSH4	PcOvo2 [SF]	19	3497	71	4.96 ± 0.46	6,8
SSH8	PcOvo3 [SF]	28	6803	56	4.72 ± 0.96	12,13,14
SSH208	Perforin	47	2303	74	4.31 ± 1.42	1,2
SSH95	novel protein [SF]	19	1222	63	3.37 ± 1.92	
SSH140	ovomucoid/protease inhibitor	6	169	73	2.39 ± 0.44	
Pc109422	cell adhesion protein	11	170	31	1.17 ± 0.35	
Pc66440	thioester-containing protein	4	10	57	1.09	
SSH3	novel protein	2	988	66	1.05 ± 0.58	
Pc123838	apoptosis-inducing factor [F]	35	354	60	0.84 ± 0.61	3
Pc108510	thioester-containing protein	5	49	44	0.81 ± 0.15	
SSH115	tachylectin-like protein	18	688	76	0.78 ± 0.35	
SSH36	C1q domain containing protein [SF]	7	477	33	0.77 ± 0.08	
Pc379	alpha-2-macroglobulin-like protein	3	7	21	0.64	
Pc54251	novel protein	2	28	66	0.43 ± 0.05	
Pc99555	Transferrin	5	23	54	0.42 ± 0.30	
Pc119627	Melanotransferrin	6	24	37	0.36 ± 0.06	
Pc75576	novel protein	2	5	21	0.36	
Pc111579	thioester-containing protein	5	36	38	0.35 ± 0.29	
Pc46253	novel protein	1	4	40	0.35	
Pc52282	scavenger receptor cysteine-rich protein	2	9	25	0.33 ± 0.16	
SSH25	neurotrypsin-like	3	393	52	0.32 ± 0.06	
Pc94087	FG-GAP repeat protein	4	46	32	0.31 ± 0.14	
Pc49292	melanotransferrin-like	2	13	45	0.29 ± 0.14	
Pc47034	novel protein	2	8	42	0.28 ± 0.13	
Pc52841	tachylectin-like protein	2	21	79	0.25 ± 0.05	
Pc59410	novel protein	2	5	42	0.25 ± 0.12	
Pc65629	novel protein	1	19	19	0.24 ± 0.18	
SSH122	novel protein	1	5	20	0.23 ± 0.02	
SSH111	novel protein	2	3	23	0.22	
Pc29496	similar to Niemann-Pick disease type C2	2	95	26	0.19 ± 0.03	
Pc112854	thioester-containing protein	4	10	31	0.17 ± 0.15	
Pc53884	cysteine rich transmembrane BMP regulator	1	5	35	0.16 ± 0.03	
Pc86328	Ubiquitin	2	3	33	0.13 ± 0.05	
Pc117840	15-hydroxyprostaglandin dehydrogenase	2	10	18	0.13 ± 0.04	
Pc119142	transmembrane protease serine	3	5	24	0.13	
Pc71049	alpha-2-macroglobulin	1	4	18	0.12 ± 0.02	
SSH218	tachylectin-related protein	3	106	27	0.12 ± 0.02	
Pc101904	Transferrin	1	2	18	0.12	
Pc124918	Melanotransferrin	6	23	27	0.12 ± 0.03	
Pc87491	telomeric repeat-binding factor 2-interacting protein	1	3	14	0.12 ± 0.02	
Pc17674	beta actin	2	5	32	0.12 ± 0.02	
SSH6	novel protein	1	196	24	0.12 ± 0.02	
SSH42	scavenger receptor cysteine-rich protein	1	36	14	0.11 ± 0.02	
SSH14	peptidoglycan recognition protein S1L [SF]	3	12	18	0.10 ± 0.09	
SSH94	novel protein	1	7	10	0.09 ± 0.02	
Pc101185	Copper/zinc superoxide dismutase	1	3	17	0.09 ± 0.02	
SSH51	kunitz-like protease inhibitor [SF]	3	51	33	0.09 ± 0.08	4,5
Pc100845	Gap-Pol polyprotein-like	1	6	5	0.08 ± 0.01	
Pc120886	Chitotriosidase	3	19	22	0.08 ± 0.01	
Pc114721	selenium-dependent glutathione peroxidase	1	5	10	0.08 ± 0.01	
Pc94566	novel protein	1	4	10	0.07	
Pc106229	aldehyde dehydrogenase	1	2	8	0.06	
SSH86	Calcium-binding protein [SF]	1	115	7	0.05 ± 0.01	
Pc120959	aldehyde oxidase	1	4	9	0.04 ± 0.01	
Pc122290	lysosome-associated membrane glycoprotein 1	1	1	5	0.03	
Pc34343	novel protein	1	6	1	0.02	

dx.doi.org/10.1021/pr3003613 | *J. Proteome Res.* 2012, 11, 4240−4248

Table 1. continued

[a]Genes obtained from SSH were renamed using SSH as prefix. [b]Proteins were annotated by BLAST against the NCBI Nonredundant Protein Database. A protein with no homologue in this database is annotated as "novel protein". "S" at the end of the annotation indicates a predicted signal peptide, and "F" a protein with full amino acid sequences. [c]For mature proteins only. [d]This column shows the mean and standard deviation of emPAI value based on three replicates.

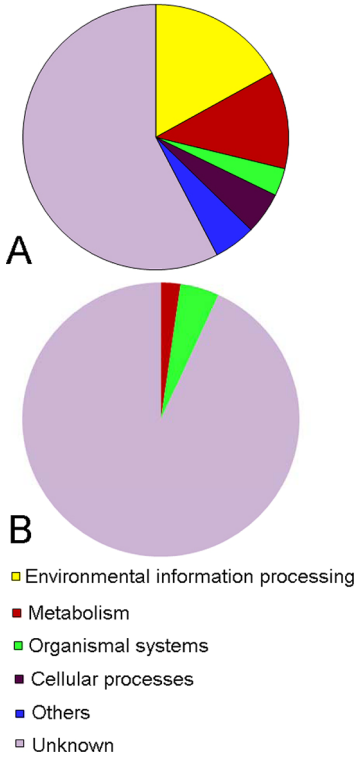


Figure 4. KEGG functional classification of PVF proteins (A) and albumen gland specific genes (B).

Phylogenetic analysis of six paralogous proteins with 23.3–38.7% identity including the three PcOvo subunits revealed the presence of two clades (Figure 6), with the three PcOvo subunits clustered in the same clade.

DISCUSSION

Identification of PVF Proteins

Through this study, we have identified 59 proteins from the PVF of *P. canaliculata*, greatly expanding the list of known PVF proteins in gastropods, demonstrating the powerfulness of our combined approach using conventional (2-DE, SDS-PAGE) and state-of-the-art (MALDI TOF/TOF, LC–MS/MS) proteomics techniques. This catalogue of PVF proteins is by far the most comprehensive among those of gastropods. The results of our study support the conclusions from a recent study in *Biomphalaria glabrata* suggesting that gastropod eggs may be heavily invested with proteins of various functions, and that proteomic techniques are of great potential to unveil the

Table 2. Thirty One Pairs of PVF Protein Homologues Based on Self-BLAST with a Threshold *e*-Value of 1.0×10^{-5}

PVF proteins	PVF proteins	% homology	*e*-value
SSH4	SSH8	35.00	3.0×10^{-31}
SSH9	SSH8	32.66	1.0×10^{-29}
SSH95	SSH2	36.72	2.0×10^{-29}
Pc119627	Pc124918	33.33	1.0×10^{-25}
SSH9	SSH4	30.18	2.0×10^{-23}
SSH20	SSH9	38.71	4.0×10^{-22}
SSH20	SSH8	38.32	5.0×10^{-18}
SSH20	SSH4	33.06	3.0×10^{-17}
SSH95	SSH8	23.30	5.0×10^{-17}
SSH2	SSH8	27.27	3.0×10^{-15}
SSH2	SSH4	30.34	4.0×10^{-13}
SSH122	SSH95	86.67	7.0×10^{-13}
SSH2	Pc75576	44.44	6.0×10^{-12}
SSH42	SSH25	57.45	7.0×10^{-12}
SSH111	SSH6	53.42	1.0×10^{-11}
Pc66440	Pc111579	52.17	9.0×10^{-11}
SSH2	SSH9	33.65	9.0×10^{-11}
Pc101904	Pc124918	45.76	3.0×10^{-10}
SSH95	Pc75576	34.48	2.0×10^{-9}
SSH9	SSH95	24.37	2.0×10^{-9}
SSH115	SSH218	32.61	1.0×10^{-8}
Pc49292	Pc99555	42.31	3.0×10^{-8}
SSH42	Pc52282	43.18	6.0×10^{-8}
SSH218	Pc52841	60	8.0×10^{-8}
Pc124918	Pc99555	38.46	2.0×10^{-7}
SSH140	SSH3	100	2.0×10^{-7}
SSH9	Pc119513	25	1.0×10^{-6}
SSH95	SSH4	24.79	2.0×10^{-6}
SSH20	SSH2	25	2.0×10^{-6}
Pc52841	SSH115	66.67	2.0×10^{-6}
Pc75576	SSH8	27.59	3.0×10^{-6}

Figure 5. Sequence alignment for the three ovorubin subunits using Clustal_X. Several symbols above the amino acids are used to mark the conserved positions: "*" indicates positions with a single, fully conserved residue; ":" indicates that one of the following "strong" groups is fully conserved; and "." indicates that one of the following "weaker" groups is fully conserved. There are two conserved motifs: NXTL (an N-link glycosylation site) and GXSWPR.

dx.doi.org/10.1021/pr3003613 | *J. Proteome Res.* 2012, 11, 4240–4248

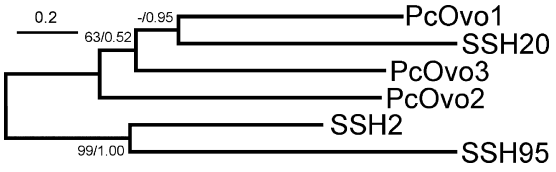


Figure 6. Phylogeny of six paralogous PVF proteins, including three subunits of PcOvo, i.e., PcOvo1, PcOvo2, and PcOvo3. Percentages (to the left of the line near the nodes) of 100 bootstrap resamplings in the maximum likelihood analysis in MEGA 5.0 and probability values (to the right of the line near the nodes) in MrBayes 3.2 are shown to indicate the stability of the tree.

composition of acellular components that are refractory to transcriptomic analysis.[14]

Of the two protein identification approaches used in this study, the LC−MS/MS approach was clearly superior in sensitivity as demonstrated by the high dynamic range of detected proteins and the detection of significantly more proteins. The 2-DE method, however, was useful in separating protein isoforms by different isoelectric points for further analysis of PcOvo subunits. The presence of protein isoforms is probably due to post-translational modification such as phosphorylation and glycosylation as previously demonstrated for PcOvo.[24] Protein phosphorylation has been documented in PcOvo, which was suggested to provide a phosphorus reserve during *P. canaliculata* embryonic development.[24] In addition, PcOvo is heavily glycosylated.[28]

Functions of PVF Proteins

Of the *P. canaliculata* PVF proteins of known functions, the class "environmental information processing" is the most abundant (10 of 25), and most of them have been reported to be involved in innate immunity responses in other mollusks. For example, expression of a thioester-containing gene in the scallop *Chlamys farreri* was triggered by microbial challenge.[29] A C1q-domain-containing gene in *C. farreri* was upregulated by lipopolysaccharides (LPS) treatment, and the expressed protein was able to bind to the LPS of *E. coli*.[30] Peptidoglycan recognition protein, which can bind to peptidoglycan (the major component of bacterial cell wall), is often upregulated by bacterial challenge, and therefore it can mediate host responses to bacterial infections.[31] Transferrin and melanotransferrin can bind iron to cause a low free iron concentration in the environment, reducing the survival of bacteria.[32]

In addition to the above defense proteins against bacteria, *P. canaliculata* PVF contains chitotriosidase, a chitinase that can cleave chitin, the polymer present in coatings of many pathogens such as fungi and exoskeleton of arthropods.[33] The presence of chitinases and *N*-acetyl glylcosaminidases in gastropods has only been reported in the digestive tract of *Helix*.[34] Interestingly, chitinases are also present in plant seeds like barley seed where they are involved in embryo resistance to fungal and insect attack.[35] It is possible that it plays a similar role in *P. canaliculata* eggs. Overall, the presence of these diverse innate immunity related proteins in the PVF of *P. canaliculata* is consistent with a recent discovery of 16 defense or immunity related proteins in egg mass fluid of *Biomphalaria glabrata*.[14] They may protect the developing embryos from bacteria and fungal infection. However, this class of proteins, except peptidoglycan recognition protein, was not detected among the SSH sequences, suggesting innate immunity proteins production in the PVF may not be restricted to genes expressed in the AG of *P. canaliculata*. Nevertheless, most of the novel proteins (11 of 19), including PcOvo, were expressed specifically in the AG, further supporting the importance of AG in terms of maternal investment in egg defense, as previously observed.[9,10,23,36]

Unexpectedly, an apoptosis-inducing factor (AIF) was found in the PVF, detected by both 2-DE and LC−MS/MS approaches. AIF is a mitochondrial protein that triggers chromatin condensation and DNA degradation in a cell in order to induce programmed cell death; it is also involved in the control of early morphogenesis in humans.[37] Its presence in the apple snail PVF, a cell-free fluid, shows for the first time that this protein can be extracellular. However, it is not known whether this protein participates in the apple snail embryonic development or is a defense factor in concert with PcOvo and PcPV2 proteins[9,10] and the other potential defensive proteins identified in the present study. Another PVF protein, annotated as Niemann−Pick disease type C2 (NPC2) protein, was also found in the chicken egg white.[38] NPC2 is believed to be a lysosomal cholesterol-binding glycoprotein.[39] Cholesterol, an essential structural component of cell membrane, is hydrophobic and therefore must be bound to proteins before being transported into a cell. NPC2 may thus be involved in cholesterol (or other sterols) transport during apple snail embryonic development.

Copper−zinc superoxide dismutase is an antioxidant defense enzyme found in almost all eukaryotic cells. The presence of this protein in the PVF may indicate its involvement in embryonic antioxidant defense system. However, its quantity is too low (0.09%) to exert a significant protective effect. Previous studies have shown that PcOvo, which contains the potent hydrophobic antioxidant astaxanthin, is thermal stable and present throughout the embryonic development. This perivitellin is the main egg protein component, and it is thought to maintain the optimal condition for embryo development, including provision of lipid-phase antioxidants.[7,8,40]

Although several studies have examined the biochemical and biophysical properties of PcOvo, PcPV2 and PcPV3, their full protein sequences, except the N-terminal sequences of PcPV2, were unknown. The present study confirmed PcOvo to comprise at least three polypeptides (PcOvo1, PcOvo2 and PcOvo3) and provided their full length sequences. In addition, the full length sequences of 10 other PVF proteins were also obtained. As a whole, these results have given basic information for further functional analysis of these PVF proteins, such as expression assay in *E. coli*.

Evolution of PVF Proteins

Members of the family Ampullariidae exhibit a wide range of egg color, size, and oviposition location,[2,3] which implies the presence of novel proteins and evolution of protein orthologues in this clade of gastropods in association with their different reproductive habits. Therefore, our results showing that many of the identified proteins are novel or of "unknown" function may indicate that at least some of them are unique to *P. canaliculata,* or at least the Ampullariidae. This hypothesis is further supported by a tBLASTn analysis against the ESTs (source: NCBI) of *Aplysia californica* and *B. glabrata*, two model species of gastropods with the most abundant genomic resources but which are quite distantly related (*Pomacea* belongs to superorder Caenogastropoda, whereas *Biomphalaria* and *Aplysia* both belong to superorder Heterobranchia), which revealed 19 of the 59 proteins to have no positive hits (*e*-value

dx.doi.org/10.1021/pr3003613 | *J. Proteome Res.* 2012, 11, 4240−4248

$>1.0 \times 10^{-5}$). Of the 20 proteins detected in the *B. glabrata* egg mass fluid,[14] only two similar proteins (protease inhibitor and Cu/Zn superoxide dismutase) were also found in the PVF of *P. canaliculata*. Comparison between the protein orthologues in ampullariid genera/species with different egg laying habits may thus recover important information about the evolution of this clade of gastropods. Furthermore, identification of these PVF proteins may also contribute to an understanding of the speciation process in Gastropoda.[15]

A self BLAST of our PVF protein list reveals as many as 31 pairs of homologues (threshold *e*-value: 1.0×10^{-5}), which indicates the occurrence of several gene duplication events[41] during the evolution of *Pomacea*. To further examine this possibility, we conducted a phylogenetic analysis of six paralogues including three PcOvo subunits, which shows a low similarity (30.2−35.0%) among each other, indicating that the gene duplications for these PcOvo sequences might have occurred early in the evolution of *Pomacea*.

In summary, our study has provided a comprehensive list of PVF proteins from the eggs of *P. canaliculata*, to our knowledge the most complete one in gastropods. The analytical approach employed in our study can be applied to determine the PVF protein composition of other genera/species of ampullariids. Comparison across several species of ampullariids representing different egg laying habits may allow for testing hypotheses about the molecular basis of reproductive behavior in this lineage of Gastropoda. Further studies should also be conducted to examine the arrangement of PVF genes in the genome of ampullariids to better understand their function and evolution.

■ ASSOCIATED CONTENT

Supporting Information

SI 1: Sequences and NCBI accession numbers of proteins identified from the perivitelline fluid of *Pomacea canaliculata*. **SI 2**: A list of proteins identified from the PVF of *Pomacea canaliculata* by using 2-DE. **SI 3**: 2-DE gel image showing the subunits of the purified PcOvo of *Pomacea canaliculata*. **SI 4**: A list of proteins identified from the subunits of the purified PcOvo of *Pomacea canaliculata*. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 852-34117055. Fax: 852-34115995. E-mail: qiujw@hkbu.edu.hk.

Author Contributions

⊥These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

H.H. and S.I. are members of Carrera del Investigador CONICET, Argentina. M.S.D. is a member of Carrera del Investigador CICBA, Argentina. This study was supported by a HKBU postgraduate scholarship to J.S. and a grant (HKBU FRG2/10-11/009) to J.W.Q.

■ ABBREVIATIONS

2-DE, two-dimensional electrophoresis; ACN, acetonitrile; AG, albumen gland; AIF, apoptosis-inducing factor; CCB, colloidal coommassie blue; CID, collision-induced dissociation; IPG, immobilized pH gradient; KEGG, Kyoto encyclopedia of genes and genomes; LC−MS/MS, liquid chromatography coupled with tandem mass spectrometry; LPS, lipopolysaccharides; LTQ-Orbitrap, linear trap quadropole-orbitrap; MALDI TOF/TOF, matrix-assisted laser desorption/ionization-tandem mass spectrometry; NPC2, Niemann-Pick disease type C2; ORF, open reading frame; PVF, perivitelline fluid; RACE, rapid amplification of cDNA ends; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SSH, suppressive subtractive hybridization

■ REFERENCES

(1) Dillon, R. T. *The Ecology of Freshwater Mollusks*; Cambridge University Press, Cambridge, U.K., 2000.

(2) Heras, H.; Dreon, M. S.; Ituarte, S.; Pollero, R. J. Egg carotenoproteins in neotropical Ampullariidae (Gastropoda: Arquitaenioglossa). *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2007**, *146*, 158−167.

(3) Hayes, K. A.; Cowie, R. H.; Jørgensen, A.; Schultheiß, R.; Albrecht, C.; Thiengo, S. C. Molluscan models in evolutionary biology: apple snails (Gastropoda: Ampullariidae) as a system for addressing fundamental questions. *Am. Malacol. Bull.* **2009**, *27*, 47−58.

(4) Cowie, R. H. Apple snails as agricultural pests: their biology, impacts, and management. In *Molluscs As Crop Pests*; Barker, G. M., Ed.; CABI Publishing:Wallingford, New Zealand, 2002; pp 145−192.

(5) Jørgensen, A.; Kristensen, T. K.; Madsen, H. A molecular phylogeny of apple snails (Gastropoda, Caenogastropoda, Ampullariidae) with an emphasis on African species. *Zool. Scr.* **2008**, *37*, 245−252.

(6) Hayes, K. A.; Cowie, R. H.; Thiengo, S. C. A global phylogeny of apple snails: Gondwanan origin, generic relationships, and the influence of outgroup choice (Caenogastropoda: Ampullariidae). *Biol. J. Linn. Soc.* **2009**, *98*, 61−76.

(7) Heras, H.; Garin, C. F.; Pollero, R. J. Biochemical composition and energy sources during embryo development and in early juveniles of the snail *Pomacea canaliculata* (Mollusca: Gastropoda). *J. Exp. Zool.* **1998**, *280*, 375−383.

(8) Dreon, M. S.; Schinella, G.; Heras, H.; Pollero, R. J. Antioxidant defense system in the apple snail eggs, the role of ovorubin. *Arch. Biochem. Biophys.* **2004**, *422*, 1−8.

(9) Heras, H.; Frassa, M. V.; Fernández, P. E.; Galosi, C. M.; Gimeno, E. J.; Dreon, M. S. First egg protein with a neurotoxic effect on mice. *Toxicon* **2008**, *52*, 481−488.

(10) Dreon, M. S.; Ituarte, S.; Heras, H. The role of the proteinase inhibitor ovorubin in apple snail eggs resembles plant embryo defense against predation. *PLoS One* **2010**, *5*, e15059.

(11) Cheesman, D. F. Ovorubin, a chromoprotein from the eggs of the gastropod mollusc *Pomacea canaliculata*. *Proc. R. Soc. London, Ser. B* **1958**, *149*, 571−587.

(12) Garín, C. F.; Heras, H.; Pollero, R. J. Lipoproteins of the egg perivitelline fluid of *Pomacea canaliculata* snails (Mollusca: Gastropoda). *J. Exp. Zool.* **1996**, *276*, 307−314.

(13) Frassa, M. V.; Ceolín, M.; Dreon, M. S.; Heras, H. Structure and stability of the neurotoxin PV2 from the eggs of the apple snail *Pomacea canaliculata*. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1492−1499.

(14) Hathaway, J. J. M.; Adema, C. M.; Stout, B. A.; Mobarak, C. D.; Loker, E. S. Identification of protein components of egg masses indicates parental investment in immunoprotection of offspring by *Biomphalaria glabrata* (Gastropoda, Mollusca). *Dev. Comp. Immunol.* **2010**, *34*, 425−435.

(15) Aagaard, J. E.; Yi, X.; MacCoss, M. J.; Swanson, W. J. Rapidly evolving zona pellucida domain proteins are a major component of the vitelline envelope of abalone eggs. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 17302−17307.

(16) Sun, J.; Zhang, Y.; Thiyagarajan, V.; Qian, P.-Y.; Qiu, J.-W. Protein expression during the embryonic development of a gastropod. *Proteomics* **2010**, *10*, 2701−2711.
(17) Neuhoff, V.; Arold, N.; Taube, D.; Ehrhardt, W. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **1988**, *9*, 255−262.
(18) Gan, C. S.; Guo, T.; Zhang, H.; Lim, S. K.; Sze, S. K. A comparative study of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus SCX-IMAC-based methods for phosphopeptide isolation/enrichment. *J. Proteome Res.* **2008**, *7*, 4869−4877.
(19) Käll, L.; Storey, J. D.; MacCoss, M. J.; Noble, W. S. Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. *J. Proteome Res.* **2007**, *7*, 29−34.
(20) Sun, J.; Wang, M.; Wang, H.; Zhang, H.; Thiyagarajan, V.; Qian, P.-Y.; Qiu, J.-W. *De novo* assembly of the transcriptome of an invasive snail and its multiple ecological applications. *Mol. Ecol. Resour.* (submitted).
(21) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4*, 1265−1272.
(22) Wang, H.; Qian, P.-Y. Involvement of a novel p38 mitogen activated protein kinase in larval metamorphosis of the polychaete *Hydroides elegans* (Haswell). *J. Exp. Zool., Part B* **2010**, *314*, 390−402.
(23) Dreon, M. S.; Heras, H.; Pollero, R. J. Metabolism of ovorubin, the major egg lipoprotein from the apple snail. *Mol. Cell. Biochem.* **2003**, *243*, 9−14.
(24) Ituarte, S.; Dreon, M. S.; Pasquevich, M. Y.; Fernández, P. E.; Heras, H. Carbohydrates and glycoforms of the major egg perivitellins from *Pomacea* apple snails (Architaenioglossa: Ampullariidae). *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **2010**, *157*, 66−72.
(25) Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785−786.
(26) Ronquist, F.; Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572−1574.
(27) Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731−2739.
(28) Dreon, M. S.; Heras, H.; Pollero, R. J. Characterization of the major egg glycolipoproteins from the perivitelline fluid of the apple snail *Pomacea canaliculata*. *Mol. Reprod. Dev.* **2004**, *68*, 359−364.
(29) Zhang, H.; Song, L.; Li, C.; Zhao, J.; Wang, H.; Gao, Q.; Xu, W. Molecular cloning and characterization of a thioester-containing protein from Zhikong scallop *Chlamys farreri*. *Mol. Immunol.* **2007**, *44*, 3492−3500.
(30) Zhang, H.; Song, L.; Li, C.; Zhao, J.; Wang, H.; Qiu, L.; Ni, D.; Zhang, Y. A novel C1q-domain-containing protein from Zhikong scallop *Chlamys farreri* with lipopolysaccharide binding activity. *Fish Shellfish Immunol.* **2008**, *25*, 281−289.
(31) Dziarski, R. Peptidoglycan recognition proteins (PGRPs). *Mol. Immunol.* **2004**, *40*, 877−886.
(32) Ritchie, R. F.; Palomaki, G. E.; Neveux, L. M.; Navolotskaia, O.; Ledue, T. B.; Craig, W. Y. Reference distributions for the negative acute-phase serum proteins, albumin, transferrin and transthyretin: a practical, simple and clinically relevant approach in a large cohort. *J. Clin. Lab. Anal.* **1999**, *13*, 273−279.
(33) Van Eijk, M.; van Roomen, C. P.; Renkema, G. H.; Bussink, A. P.; Andrews, L.; Blommaart, E. F.; Sugar, A.; Verhoeven, A. J.; Boot, R. G.; Aerts, J. M. Characterization of human phagocyte-derived chitotriosidase, a component of innate immunity. *Int. Immunol.* **2005**, *17*, 1505−1512.
(34) Kimura, A. Purification and physicochemical properties of chitinase extracted from digestive tract of *Helix peliomphala*. *Tokushima J. Exp. Med.* **1977**, *24*, 1−8.
(35) Swegle, M.; Kramer, K. J.; Muthukrishnan, S. Properties of barley seed chitinases and release of embryo-associated isoforms during early stages of imbibition. *Plant Physiol.* **1992**, *99*, 1009−1014.
(36) Dreon, M. S.; Lavarias, S.; Garin, C. F.; Heras, H.; Pollero, R. J. Synthesis, distribution, and levels of an egg lipoprotein from the apple snail *Pomacea canaliculata* (Mollusca: Gastropoda). *J. Exp. Zool.* **2002**, *292*, 323−330.
(37) Joza, N.; Susin, S. A.; Daugas, E.; Stanford, W. L.; Cho, S. K.; Li, C. Y. J.; Sasaki, T.; Elia, A. J.; Cheng, H. Y. M.; Ravagnan, L.; Ferri, K. F.; Zamzami, N.; Wakeham, A.; Hakem, R.; Yoshida, H.; Kong, Y.-Y.; Mak, T. W.; Zuniga-Pflucker, J. C.; Kroemer, G.; Penninger, J. M. Essential role of the mitochondrial apoptosis-inducing factor in programmed cell death. *Nature* **2001**, *410*, 549−554.
(38) Mann, K. The chicken egg white proteome. *Proteomics* **2007**, *7*, 3558−3568.
(39) Liou, H. L.; Dixit, S. S.; Xu, S.; Tint, G. S.; Stock, A. M.; Lobel, P. NPC2, the protein deficient in Niemann-Pick C2 disease, consists of multiple glycoforms that bind a variety of sterols. *J. Biol. Chem.* **2006**, *281*, 36710−36723.
(40) Dreon, M. S.; Ceolín, M.; Heras, H. Astaxanthin binding and structural stability of the apple snail carotenoprotein ovorubin. *Arch. Biochem. Biophys.* **2007**, *460*, 107−112.
(41) Clark, M. S.; Thorne, M. A. S.; Toullec, J.-Y.; Meng, Y.; Guan, L. L.; Peck, L. S.; Moore, S. Antarctic krill 454 pyrosequencing reveals chaperone and stress transcriptome. *PLoS One* **2011**, *6*, e15919.