

## **Novel member of Ras family proteins from Disk Abalone (*Haliotis discus discus*): Structural profiling and its transcriptional modulation under host pathologic conditions**

D. A. S. Elvitigala<sup>\*1,4</sup>, H. K. A Premachandra<sup>1</sup>,  
Q. Wan<sup>1</sup>, B. Lim<sup>2</sup>, H. Jung<sup>2</sup>,  
C. Y. Choi<sup>3</sup> and J. Lee<sup>1,2 \*\*</sup>

<sup>1</sup>Department of Marine Life Sciences, School of Marine Biomedical Sciences, Jeju National University, Jeju Self-Governing Province, 63243, Republic of Korea

<sup>2</sup>Fish Vaccine Development Center, Jeju National University, Jeju Special Self-Governing Province, 63243, Republic of Korea

<sup>3</sup>Division of Marine Environment and Bioscience, Korea Maritime University, Busan, 49112, Republic of Korea

<sup>4</sup>Center for Biotechnology, Department of Zoology, University of Sri Jayewardenepura, Gangodawila, Nugegoda, 10250, Sri Lanka.

<sup>\*</sup> Corresponding authors

<sup>\*</sup> E.mail - elvitigaladas@yahoo.com

<sup>\*\*</sup> Email – jehee@jejun.ac.ky

### **Abstract**

Among small GTPases, the Ras family proteins capture a remarkable place in dictating cellular proliferation, differentiation and survival in development of an organism. Major members of the Ras family include Ras (H-Ras, K-Ras, N-Ras), Rap1, and Rap2, all of which can act as oncogenes upon mutation. In the present study, a novel Ras family protein (AbRFP) was characterized from Disk Abalone (*Haliotis discus discus*), an economically important, edible marine gastropod; further analyzing its transcriptional profile in healthy and immune-challenged animals. The full-length cDNA of AbRFP is 2704 bp and it consists of an open reading frame of 552 bp, encoding a 184 amino acid peptide with a calculated molecular mass of ~21 kDa and isoelectric point of 8.63. The amino acid sequence resembles the characteristic features of typical Ras family proteins, including GTP/Mg<sup>2+</sup> binding sites and guanine nucleotide exchange factor (GEF) interaction sites, as predicted by the NCBI-conserved domain database server. Phylogenetic study of AbRFP showed the generally accepted relationships, with AbRFP exhibiting highest proximity to a Ras protein from Portuguese oyster. Quantitative real-time PCR detected ubiquitous AbRFP mRNA expression, with strongest levels in muscle along with mantle and the lowest level in hepatopancreas. The AbRFP transcriptional profile in gills of Abalone challenged with *Vibrio parahaemolyticus* or viral hemorrhagic septicemia virus (VHSV) demonstrated significant up-regulations ( $p < 0.05$ ) at 12 h and 24 h post-injection (p.i.), respectively. Moreover, significant elevation ( $p < 0.05$ ) of mRNA expression was detected in hemocytes at 72 h p.i. with *V. parahaemolyticus*. These findings suggest that AbRFP may play a role under pathological conditions in Disk Abalone.

**Key words:** Disk Abalone, Ras family protein, tissue distribution of mRNA, pathogen challenge, transcriptional modulation

## Introduction

Mariculture is a branch of aquaculture that is dedicated for cultivation of marine organisms, especially for human consumption as foods. This industry has increased in prominence due to marine organisms representing a sufficient substitution for the current terrestrial resources, regarding the fulfillment of daily nutritional requirements, which are becoming stressed by the ever-increasing global population. In addition, several marine invertebrates, such as shellfish are also frequently considered as a delicacy of the human diet, especially in East and Southeastern countries like China, Korea, and Japan. One such delicacy is the Abalone, a marine gastropod mollusk that is cultivated as an aquacrop and currently accounts for a considerable proportion of the yield in the commercial aquaculture industry worldwide. However, the marine snails, including Abalone, are sensitive to a wide range of environmental conditions, with negative effects on their survival and growth. In particular, pathogenic infections and toxicants, including carcinogens are known influential factors on marine organisms and have rendered aquacrops such as Abalones as bio-indicators with respect to their health and viability (van der Oost, *et al.*, 2003). Infections by numerous bacteria (Liu, *et al.*, 2000; Huang, *et al.*, 2001), viruses (Nakatsugawa, *et al.*, 1999) and parasites (Goggin and Perkinsus, 1995) have been recorded in Abalone crops. Under these pathogenic conditions, innate immune mechanisms allow Abalones to endure these pathogenic threats, at least up to a certain threshold level.

The Ras superfamily of immunomodulators consists of a large group of structurally- and functionally-related monomeric guanine nucleotide binding proteins, collectively known as the small GTP binding proteins. These membrane-associated proteins play important roles in conveying a wide range of extracellular signals to intracellular signaling cascades (receptor-mediated signaling), and mediate the traffic of small vesicles between different intracellular compartments (Ye and Carew, 2010) and cytoskeletal organization (Takai, *et al.*, 2001). Moreover, these small GTP-binding proteins are activated by GTP binding and inactivated by hydrolysis of the bound GTP to GDP (Takai, *et al.*, 2001). The kinetics of this process are modified by several classes of auxiliary proteins. Guanine nucleotide exchange factor (GEF) catalyzes the release of bound GDP, and dictates the capture of new GTP to re-activate the protein, whereas GTPase activating proteins (GAPs) stimulate the low intrinsic GTPase activity of the small GTP-binding protein. In addition, some of the effector proteins exert more regulatory control than others; for instance, the guanine

nucleotide dissociation inhibitor (GDI) strongly suppresses inactivation by physically blocking the nucleotide release (Takai, *et al.*, 2001). As such, a fine counterbalance of active and inactive small GTP binding proteins is essential for proper regulation of signaling cascades.

The Ras superfamily is divided into five main families according to the protein structure, namely Ras, Rho, Rab, Sar1/Arf and Ras-related nuclear protein (Ran) (Takai, *et al.*, 2001). Among these, the Ras family proteins (RFPs) include Ras (H-Ras, N-Ras, and K-Ras), Rap1 (Krev-1), and Rap2, which are the prominent members represented in the literature. The RFPs are known to play important roles during development. As critical molecular switches of cellular proliferation, differentiation, and survival (Barbacid, 1987; Konstantinopoulos, *et al.*, 2007), the RFPs are highly sensitive to mutation. Since these mutations can manifest as induction and promotion of tumorigenesis, the RFP genes are classified as oncogenes (Ye and Carew, 2010). RFPs are also known to facilitate exocytosis and endocytosis processes in cells by interacting with other families of proteins in the small G-protein superfamily, such as Ral and Rab (Une, *et al.*, 1991; Kikuchi and Williams, 1996; Urano, *et al.*, 1996).

The RFPs show broad evolutionary conservation, and have been identified and characterized in various vertebrate and invertebrate species. Among invertebrates, the Ras protein has been characterized in shrimp (Une, *et al.*, 1991), blue mussel (Ciocan and Rotchell, 2005), fruit fly (Mozer, *et al.*, 1985), and silk worm (Ogura, *et al.*, 2009), while the Rap1 protein has been identified only in fruit fly (Asha, *et al.*, 1999) and Chinese white shrimp (Ren, *et al.*, 2012). Rap in shrimps are known to involve in the host immune system (Ren, *et al.*, 2012). Similarly, members of other families, such as Rab, Ran and Ras-like proteins, have been found to be involved in antibacterial and antiviral host defense processes (Han and Zhang, 2007; Rattanarojpong, *et al.*, 2007; Wu and Zhang, 2007; Ongvarrasopone, *et al.*, 2008; Wu, *et al.*, 2008; Zong, *et al.*, 2008).

In the present study, a Ras family protein was identified from a cDNA library of Disk Abalone (*Haliotis discus discus*) and designated as AbRFP. Herein, we describe the structural characterization of this first RFP recognized from a gastropode. The transcriptional profile of AbRFP was also determined for the various tissues of healthy animals and transcriptional modulation was detected in immune relevant organs (hemocytes and gills) of immune challenged animals with *Vibrio parahaemolyticus* and viral hemorrhagic septicemia virus (VHSV).

## **Materials and Methods**

### **Identification of the complete cDNA sequence of AbRFP**

The full-length cDNA sequence of AbRFP (contig no. 15699) was identified from the previously established Disk Abalone cDNA sequence database (De Zoysa, *et al.*, 2008) by using the Basic Local Alignment Search Tool (BLAST) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

### ***In-silico* characterization of AbRFP sequences**

The open reading frame (ORF) and amino acid sequence of AbRFP were derived by the DNAssist 2.2 program (version 3.0). Nucleotide sequences of Ras family proteins from other species were obtained by BLAST search and used for pairwise and multiple sequence alignments to AbRFP by the ClustalW2 program (Thompson, *et al.*, 1994). The phylogenetic relationship was determined by using the Molecular Evolutionary Genetics Analysis (MEGA) software, version 5, with the Neighbor-Joining method and bootstrapping values taken from 1000 replicates (Tamura, *et al.*, 2011). Characteristic protein signatures in the AbRFP sequence were predicted by the ExPASy-prosite server (<http://prosite.expasy.org>) and the NCBI conserved domain database (CDD) (Marchler-Bauer, *et al.*, 2011). Some physicochemical properties of AbRFP were determined by using the ExPASy protParam tool (<http://web.expasy.org/protparam>).

### **Computer simulation modeling of AbRFP**

The tertiary structure of AbRFP was modeled using the I-TASSER online structure assembly simulation server (Zhang, 2008; Roy, *et al.*, 2010), and the generated information of the model was constructed into a three-dimensional (3D) structure using the molecular visualizing software Swiss-Pdb viewer version 4.0.4 (Kaplan and Littlejohn, 2001).

### **Experimental animals and tissue collection**

Healthy Disk Abalones (*H. discus discus*), with an average size of 8 cm, were obtained from the “Youngsoo” Abalone farm (Jeju Island, Korea). Upon delivery to the laboratory, the Abalones were acclimated for one week by housing in a controlled environment that consisted of continuously filtered and aerated seawater with stable salinity ( $33\pm 1$  psu) and temperature ( $20\pm 1^\circ\text{C}$ ). During the acclimation period, the Abalones were fed daily with fresh marine seaweed (*Undaria pinnatifida*).

Hemolymph (1-2 mL/Abalone) was collected from the pericardial cavities of three healthy and unchallenged animals, using sterilized syringes; the respective samples were immediately centrifuged (3000 ×g, 4°C, 10 min) to harvest the hemocytes. In addition, tissues from adductor muscle, mantle, gill, hepatopancreas, digestive tract, and testis were collected from three animals. All samples were snap-frozen in liquid nitrogen and stored at -80°C until use.

### **Immune challenge experiment**

Abalones with an average weight of 50 g were selected for the immune challenge experiment. One group was challenged with *V. parahaemolyticus*, a Gram-negative bacterium, by injecting 100 µL of a  $1 \times 10^4$  cell/mL in saline solution intramuscularly. A second group was challenged with a Korean isolate of VHSV (FWando05) by intramuscular injection of 100 µL of a  $1 \times 10^8$  pfu/mL in saline solution. The VHSV was obtained from an infected olive flounder (*Paralichthys olivaceus*) and amplified by growth on the fathead minnow (FHW) cell line (in Dulbecco's minimum essential medium with antibiotics and 10% fetal bovine serum (FBS)) until extensive cytopathic effect was observed, at which point the supernatant was harvested by centrifugation (3000 × g, 4 °C, 10 min). A third group of un-injected Abalones served as the negative control. A fourth group was injected intramuscularly with saline alone and served as a control to normalize the potential effects caused by its use as the medium of the injected pathogens. At post-injection (p.i.) hours 3, 6, 12, 24, 48, 72, and 120, hemocytes and gills were collected from at least four animals of all groups.

### **Total RNA extraction and cDNA synthesis**

Total RNA was extracted from Abalone tissues pooled from 3 Abalones from each group by using the Tri Reagent<sup>TM</sup> (Sigma, St. Louis, MO, USA). After quantification by UV spectrophotometry (optical density at 260 nm), the total RNA samples were diluted up to 1 µg/µL and subjected to perform cDNA synthesis using the PrimeScript<sup>TM</sup> cDNA Synthesis Kit (TaKaRa Bio, Shiga, Japan). Finally newly synthesized cDNA was diluted 40-fold (total volume: 800 µL) and stored at -20°C until use.

### **AbRFP mRNA expression analysis by quantitative real-time PCR (qRT-PCR)**

qRT-PCR was performed by the Dice<sup>TM</sup> TP800 Real-Time Thermal Cycler System (TaKaRa) in 15 µL reaction volumes containing 4 µL of diluted cDNA, 7.5 µL of 2× TaKaRa ExTaq<sup>TM</sup> SYBR premix, 0.6 µL of AbRFP primers (Table 1), and 2.3 µL of ddH<sub>2</sub>O. The following thermal cycling conditions were used: 10 sec at 95°C, followed by 35 cycles of 5 sec at 95°C, 10 sec at 58°C and 20 sec at 72°C, and a final cycle of 15 sec at 95°C, 30 sec at 60°C and 15 sec at 95°C.

**Table 1.** Primers used in this study (5' → 3')

| Primer                 | Sequence                         | Gene target                  |
|------------------------|----------------------------------|------------------------------|
| <b>AbRFP-Forward</b>   | TGGATATTCTCGACACAGCAGGTC<br>AAGA | AbRFP                        |
| <b>AbRFP-Reverse</b>   | TGGACTGTGCGTGTGGGTAGAT           | AbRFP                        |
| <b>Ab-Ribo-Forward</b> | TCACCAACAAGGACATCATTTGTC         | Ribosomal protein<br>L5 gene |
| <b>Ab-Ribo-Reverse</b> | CAGGAGGAGTCCAGTGCAGTATG          | Ribosomal protein<br>L5 gene |

The accompanying Dice<sup>TM</sup> Real-Time System Software (version 2.00) automatically set the baseline. AbRFP expression was determined by the Livak ( $2^{-\Delta\Delta CT}$ ) method (Livak and Schmittgen, 2001). The same qRT-PCR profile was used for detection of the internal control gene, Abalone ribosomal protein L5 (GenBank ID: EF103443), with respective gene-specific primers (Table 1). Expression levels were analyzed in triplicate and the data are presented as mean  $\pm$  standard deviation (SD) of relative mRNA expression. The relative expression level of AbRFP at the 0h time point in negative controls was used as the baseline for comparative purposes. The significance ( $p < 0.05$ ) of differences between the experimental groups and negative control group were determined by two-tailed unpaired t-test.

## Results and Discussion

### Sequence profile and phylogenetic relationships of AbRFP

The complete AbRFP coding sequence was found to comprise 2703 nucleotides that include a 552 bp ORF encoding 184 amino acids. The 5'-untranslated region (UTR) was found to be composed of 164 nucleotides and the 3'-UTR was consisted of 1987 nucleotides. The molecular mass of AbRFP was predicted to be ~21.1 kDa, which is within the relative molecular mass range (20-25 kDa) of the small GTP binding proteins (Lundquist, 2006). The theoretical isoelectric point was determined to be 8.63. AbRFP was predicted to be an ortholog of Ras protein by the NCBI-CDD and ExPASy-prosite server; this prediction was based upon its content of several GTP/Mg<sup>2+</sup> binding sites (residues 12-18, 24-26, 56, 112, 113, 115, 116, 142 and 143), potential GEF interaction sites (residues 13, 14, 26-28, 30, 33, 36, 37, 50, 51, 53, 55-57, 59-61, 63, 65-67, 69, 98, 99 and 144), effector interaction sites (residues 21 and 33-37), putative GDI interaction sites (residues 7, 8, 55 and 56), and two switch regions (residues 29-36 (I) and 55-73 (II), respectively (Milburn, *et al.*, 1990) (Figure. 1).

|                                                                |                                             |                         |                  |                   |      |
|----------------------------------------------------------------|---------------------------------------------|-------------------------|------------------|-------------------|------|
|                                                                | GTGTGAGCTAAAGA                              | -164                    |                  |                   |      |
| GGTCAAAATCGTAGAC                                               | GATCGTGACGAAAGTG                            | TGCTTGTGTGCAAAGG        | CAACGAACATCGTCT  | GGCGTTTGTTTTATAT  | -150 |
| ATGTAAACAAGATAT                                                | ATCCTTGTTTCCATC                             | GGGGTGAGCCAGAAA         | GACGTCCAGCGGACG  | CACGTGAGGTCGAA    | -75  |
| <u>AT</u> GACCGAATACAAG                                        | CTTGTGGTGTTTGG                              | GCTGGAGGTGTAGGT         | AAGAGTGCCTTGACT  | ATCCAGCTCATACAA   | 75   |
| M T E Y K L V <sup>▲</sup> V G                                 | A G G <b>V</b> G K S A L T <u>I</u> Q L I Q |                         |                  |                   | 25   |
| AACCACCTTTGTAGAA                                               | GAGTATGATCCCACT                             | ATAGAGGATTCCTAC         | CGGAAACAGGTTGTG  | ATTGATGGGGAACG    | 150  |
| <b>N</b> H F <u>V</u> E E Y <u>D</u> P T <u>I</u> E            | D S Y R K Q V V I D G E T                   |                         |                  |                   | 50   |
| TGTCTACTGGATATT                                                | CTCGACACAGCAGGT                             | CAAGAGGAGTACAGC         | GCCATGAGGGACCAAG | TACATGCGGACGGGG   | 225  |
| C L L D <sup>▲</sup> <u>I</u> <sup>▲</sup> L D T A G Q E E Y S | A M R D Q Y M R T G                         |                         |                  |                   | 75   |
| GAGGGGTTTCTATGT                                                | GTGTTTGCGATTACC                             | AAGTCTTTGGAAGAT         | ATCAACCAAGTATAGA | GAACAGATTAAAGAGA  | 300  |
| E G F L C V F A V T K S F E                                    | D I N Q Y R E Q <b>I</b> K R                |                         |                  |                   | 100  |
| GTTAAGGATGCTGAC                                                | GAGGTGCCCATGGTT                             | TTGGTAGGAAATAAG         | GTAGATCTACCCACA  | CGCAGAGTCCTATTCA  | 375  |
| V K D A D E V P M V L <b>V</b> G N K                           | V D L P T R T V H S                         |                         |                  |                   | 125  |
| AAAGATGCGAAGCAG                                                | GTTGCCACAGCTAT                              | GGGATTCCTTACGTT         | GAAACATCAGCCAAAG | ACAAGGCAAGGTTGT   | 450  |
| K D A K Q V A D S Y G I P Y V                                  | E <b>T</b> S A K T R Q G V                  |                         |                  |                   | 150  |
| GATGACGCATTTTAC                                                | ACACTGTGTAGAGAA                             | ATCAGGAAATATAAA         | GAGAAGAGGGGAAAG  | GACAAGCGGAAACGA   | 525  |
| D D A F Y T L V R E I R K Y K                                  | E K K G K D K R K R                         |                         |                  |                   | 175  |
| AATAGAAGACGACTA                                                | TGCAAACTCTTCTGA                             | TCATGTGACATGTCA         | TGTGACCTGTGATGA  | CGTTATTTGTTTACC   | 600  |
| N R R R L C K L F                                              |                                             |                         |                  |                   |      |
| TGTTGTGGAAACACA                                                | GTGATCATTGATGTG                             | TATAAGGTGCCGAAT         | CGGATAGATAACATT  | CATTTTGTGTCTGTG   | 675  |
| TGTTAGTCCACTCGT                                                | CAGCGTCAATGTCAAC                            | TTGGCAGTGTTTTGT         | CATTGCTCATCTTTA  | CATAGCCAATAAGAT   | 750  |
| TCAAGTCTGTGTTTT                                                | TTTCTGTAAAAAATT                             | GTGAATTTTTAGATA         | TTTTACATTTTGTAT  | ATCTGTATACAGCTA   | 825  |
| CTGTTTTTCCCAATT                                                | CCTGTATGTGCAAGC                             | ATGATGTAGAGAAGC         | ATTGTGTTAAAAATG  | TTAAATAAACTGTGA   | 900  |
| TATTTCTTTTTGTGA                                                | CTTTACAGACAGCTA                             | <u>TTT</u> ATGATGTACAGT | AAACTAGTACATGAA  | GTGGATGTGCTGTAA   | 975  |
| TCTGTAGGAGAAACC                                                | ATGAATGAATTTAG                              | AAAGCAAAATGAACT         | TAGTCGGTATTATG   | AGACATTGCCATTGG   | 1050 |
| ATGATCATGATGATG                                                | TGGGGATATGAATGA                             | TATATATATCAGAAA         | TCAGTTGATGAAACA  | GAATGAATCCTGTGA   | 1125 |
| TTGGTGCCCGATGAA                                                | AAATTAATTTTGGAA                             | AGAAGGAAACGACAA         | TCTTGATGTAAAGGA  | GGTAATTCTGCCTCA   | 1200 |
| ATTGTTCAATTCATGA                                               | TATCATTAGTTGATT                             | TTGTAGGACTTGGAA         | TTAATTGATCTAAAT  | TCTTTTGTGATTGG    | 1275 |
| CATGAGTGTTCAT                                                  | GTACACCGCTGAATC                             | CTCCATACTTCGAT          | TACTTCTTACAGAA   | AGTTTTCATCTGAAA   | 1350 |
| TATTTCTGAGTTTGTG                                               | ATTGTAGAAGTACCT                             | CCTAATTGATTTCAG         | TTTTTCTCTTGAAT   | TGTAGGAGCTTGATAT  | 1425 |
| ATTGTACACAAAATA                                                | TACGCAGATGAGACT                             | TATAATGTATTCTAC         | AACACTTGAATGGTA  | AGAGTTCTCTCGGCT   | 1500 |
| GTAGCAGTCATGAAT                                                | TTTGTACAAAAAAGT                             | TTCTTAGTGTACAGT         | AAGGCATCATATGGC  | TAATGTACAGAGTGT   | 1575 |
| ATTATGATAGTTTTT                                                | TTAATTGTTTGTACC                             | TGTTTTCATTTTAT          | ATCAATATACTGTGA  | CAAAACCATCAATATG  | 1650 |
| TGCTAACCATCTTGG                                                | GTGCTTCAAGGATAA                             | TTCTTAACCTGTGACA        | GGCGGATCTGGT     | TTGTGCTTTAGAAAC   | 1725 |
| AGTTGACGGTGATAT                                                | GTAGCTGATTTCTAT                             | AACAACCGAAATGGT         | TTTGGAAATATGTCT  | CGAAATAAATTGCAAT  | 1800 |
| GAAATGAATGATTT                                                 | TTTTTCAGGTGTTTT                             | TTTCATGACATGTTG         | AATGACTTCTAGCTG  | TAGCCAACAAATTCG   | 1875 |
| TCAACTCAACATAAC                                                | ACTGATGAAGTAAAT                             | GTTTTGAAGAAGCTA         | GCAAAATGGATTGGG  | TTGTTTTCATATCTG   | 1950 |
| GTACCACAGATATAT                                                | ATTGCCTGTATCATG                             | CATCGATATCCTCAT         | TGTGTTTTTCATCAT  | TGAAAAGGAGGGGAAA  | 2025 |
| TTCTGCAGATCAGA                                                 | AACATTTTGAATAA                              | TATTTTTCATTACT          | AATCCGTCATTCTAA  | ACATGCATTTGTACCT  | 2100 |
| AAGTCGGTGTGAGT                                                 | ATCAATATGTGTGGC                             | AAGTCACTGACGTGT         | ACCAATTAAGATTAC  | TGATGTTGTGCTGTT   | 2175 |
| GAGATGTACAACCTG                                                | CAATTTGTTTGACGA                             | TGCTGGGAGAAATAGT        | AGCATCAGACAGAAAT | TTGTATCAGAGATCA   | 2250 |
| TTTGGATAAGGTGC                                                 | ATTGTCTAACGTTTG                             | ACATGCTGCTGTTTT         | GTTTTGGAACTATAC  | TATTTTCATAGATGTTG | 2325 |
| AAGAACACCCCAATT                                                | TTTTTCATGAGTCTA                             | GTAACTGTGTTAATC         | TTTTTCATAGAAAAA  | AATAGGTTGTTCTCG   | 2400 |
| ACTTCTTTTACATA                                                 | GTATAGTTCTATCTC                             | AATAAGGACATCTTT         | TCACATAGGTTTAT   | TCITTTGGATAAGGCAT | 2475 |
| TTATTTCTTTTGTAT                                                | AAGTTTAATTCACGT                             | CTTAGTAGGATAAAG         | GCAGGAGAAATATTC  | TTGCA             | 2539 |



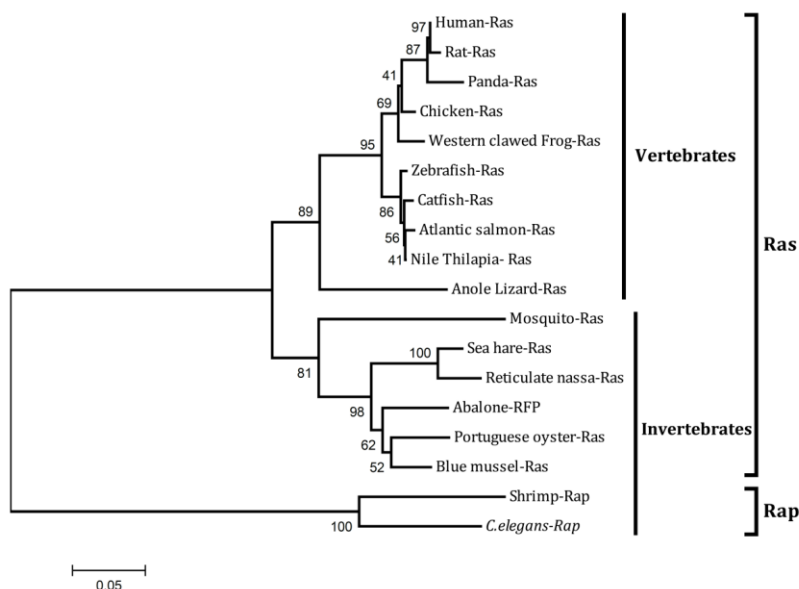
**Figure. 2.** Multiple sequence alignment of RFPs from different species. Conserved residues are indicated by asterisks. The catalytic residues, K<sub>16</sub> and Q<sub>61</sub> are denoted by red-patterned underlining. Amino acid residues for putative GEF interactions are denoted by gray shading. Well-conserved, putative GTP/Mg<sup>2+</sup> binding sites identified in AbRFP are boxed. Well-conserved residues for putative effector interactions are indicated by the ‘♦’ symbol. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The multiple sequence alignment also revealed that several of the GTP/Mg<sup>2+</sup> binding sites and GEF interaction sites in AbRFP had high evolutionary conservation, which further validated the *in silico* prediction of AbRFP identity as a Ras family GTP binding protein. Moreover, two catalytic residues (K<sub>16</sub> and Q<sub>61</sub>) that are characteristic of Ras proteins were found to be absolutely conserved in AbRFP, providing further evidence of its compatibility with Ras orthologs. In addition, AbRFP showed the highest levels of identity to the molluscan Ras orthologs, with 90.9% identity to Portuguese oyster and 89.8% identity to blue mussel (Table 2).

**Table 2.** Percentage similarities and identities of AbRFP with Ras proteins of other species

| Species (common name)                                      | Accession number   | Identity, % | Similarity, % | Amino acids |
|------------------------------------------------------------|--------------------|-------------|---------------|-------------|
| <i>Crassostrea angulata</i> (Portuguese oyster)            | ACU33971           | 90.9        | 94.1          | 184         |
| <i>Mytilus edulis</i> (blue mussel)                        | AY679522           | 89.8        | 91.9          | 184         |
| <i>Aplysia californica</i> (sea hare)                      | NP001191472        | 86.9        | 89.5          | 191         |
| <i>Nassarius reticulatus</i> (reticulate nassa)            | DQ198150           | 86.1        | 90.9          | 187         |
| <i>Rattus norvegicus</i> (Norway rat)                      | NM_031515          | 81.9        | 91.0          | 188         |
| <i>Mus musculus</i> (house mouse)                          | NM_021284          | 81.9        | 91.0          | 188         |
| Kirsten murine sarcoma virus                               | CAA80675           | 81.4        | 90.4          | 188         |
| <i>Rana catesbeiana</i> (bullfrog)                         | BT081877           | 80.1        | 90.9          | 186         |
| <i>Xenopus (Silurana) tropicalis</i> (Western clawed frog) | NM001017003        | 79.4        | 87.8          | 189         |
| <i>Ictalurus furcatus</i> (blue catfish)                   | GU587856           | 79.0        | 85.1          | 189         |
| <i>Taeniopygia guttata</i> (zebra finch)                   | NM001245553        | 78.8        | 87.3          | 189         |
| <i>Gallus gallus</i> (chicken)                             | NM205292           | 78.8        | 87.3          | 189         |
| <i>Homo sapiens</i> (human)                                | ENST00000311189    | 78.8        | 86.8          | 189         |
| <i>Salmo salar</i> (Atlantic salmon)                       | NP001135104        | 76.9        | 86.2          | 189         |
| <i>Danio rerio</i> (zebra fish)                            | NP001018465        | 76.9        | 85.6          | 189         |
| <i>Oreochromis niloticus</i> (Nile tilapia)                | XM003440405        | 76.9        | 85.6          | 189         |
| <i>Ailuropoda melanoleuca</i> (panda)                      | XP002929527        | 76.8        | 86.8          | 189         |
| <i>Anolis carolinensis</i> (anole lizard)                  | ENSACAP00000003000 | 76.8        | 85.6          | 194         |
| <i>Caenorhabditis elegans</i> (roundworm)                  | Ensemble-F17C8.4.1 | 44.2        | 58.1          | 211         |
| <i>Drosophila melanogaster</i> (fruit fly),                | Ensemble-          | 22.9        | 39.0          | 201         |
| kappa-B Ras                                                | FBpp0088074        |             |               |             |

The phylogenetic analysis identified two main clusters among the Ras family proteins, namely Ras and Rap (Figure.3). Under the Ras clade, AbRFP was located among a sub-cluster comprised of Portuguese oyster and blue mussel proteins, which had a substantial bootstrap value (62). Moreover, this sub-cluster occurred within the clade formed by the molluscan species, which further indicated the higher proximity of AbRFP to the invertebrate orthologs. The Rap clade contained a separate cluster of invertebrate proteins, forming an out-group. These overall findings indicate a common invertebrate ancestral origin of AbRFP and further reveal its evolutionary relationship with Ras counterparts, rather than Rap.



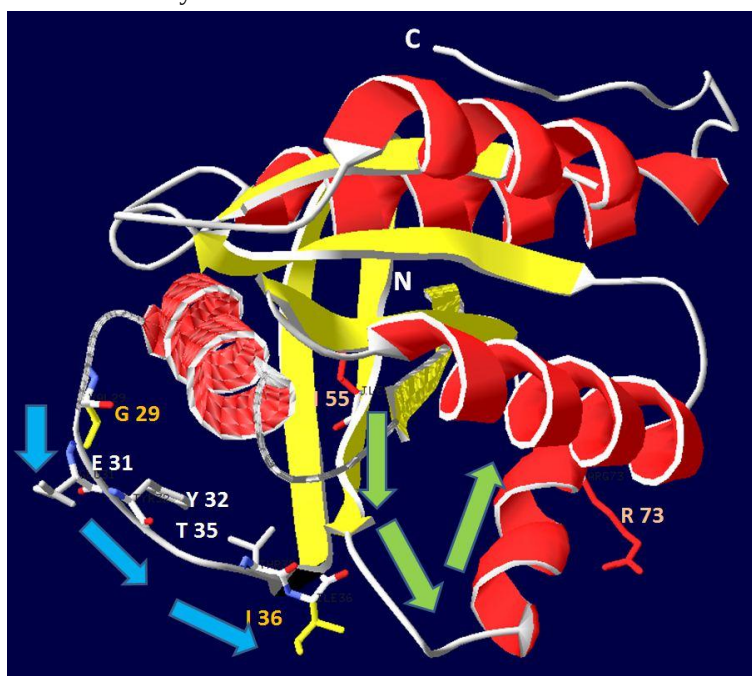
**Figure. 3.** Phylogenetic tree of Ras proteins of different species. Bootstrap values are shown on the left next to the corresponding lineages, and the identities of the major clusters are indicated on the right. The respective GenBank accession numbers of the various Ras orthologs can be found in Table 1, except for mosquito-Ras protein (No. XP001845948) and *C. elegans* Rap (NP501549).

### Predicted tertiary structure of AbRFP

In order to predict the tertiary structure of AbRFP in 3D space, a model was generated using threading alignment and an assembly simulation approach. Reliability of the structure prediction was affirmed by several parameters calculated by the I-TASSER server. Substantial percentage sequence identities found among the whole template chains Substantial percentage sequence identity of the templates in the threading aligned region with the query sequence (ranging 78-42% and 82-39%, respectively) along with

the normalized Z-score values of the threading alignments exceeding 1 were found to be significant evidences among them. In addition, the algorithmic calculation of TM score (measurement of the similarity of topologies of two protein structures) was  $0.90 \pm 0.06$ , indicating a high degree of confidence for the model.

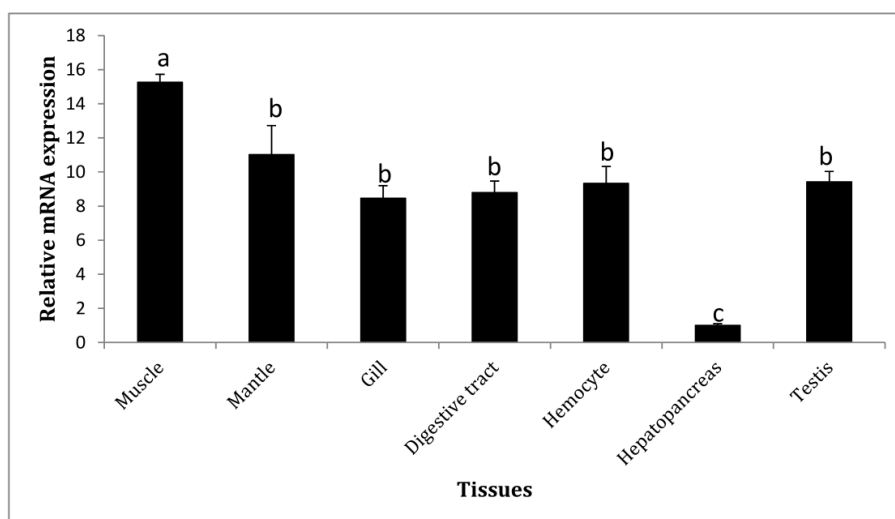
The generated AbRFP model resembled the typical 3D crystal structure of human Ras protein, consisting of six  $\beta$  strands forming a central sheet with five  $\alpha$  helices distributed on both sides of the  $\beta$  sheet (Figure. 4) (Brunger, *et al.*, 1990; Wittinghofer and Pai, 1991). Interestingly, the switch I and switch II regions formed an overlapping portion of AbRFP, which agreed with their prominent presence in the human Ras protein. Collectively, the tertiary structure model of AbRFP validated the derived primary structure, further affirming its compatibility with Ras proteins among the different members of the Ras family.



**Figure. 4.** Predicted 3D structure of AbRFP. The  $\alpha$  helices shown in red and the  $\beta$  strands are shown in yellow. The putative switch I and switch II regions are indicated by blue and green arrows, respectively. The starting and ending positions are indicated by the respective amino acid pairs: in switch I, Val 29 to Ile 36; in switch II, Ile 55 to Arg 73. N and C represent the amino and carboxy terminals of the protein, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### Tissue-specific mRNA expression profile of AbRFP

In order to determine the transcriptional profile of AbRFP under normal physiological conditions, qRT-PCR was performed using gene-specific primers and cDNA synthesized from different tissues of healthy Abalones. The relative mRNA expression fold was calculated using the Abalone ribosomal gene L5 as a reference gene, and the expression in each tissue was further normalized to that in hepatopancreas. AbRFP was detected as ubiquitously expressed in all tissues examined (Figure. 5), which suggests that this protein functions in key signaling pathways that are generally active in Abalones.



**Figure. 5.** Tissue-specific profile of AbRFP mRNA, as determined by qRT-PCR. Error bars represent the SD (n=3). Data with different letters are significantly different ( $p < 0.05$ ) among the different tissues.

However, the mRNA expression levels differed among the tissues. Compared to the tissue with the lowest level (hepatopancreas), muscle and mantle showed the highest levels, and gill, digestive tract, hemocytes and testis showed moderate levels. This observation does not agree with the profiles reported for other RFPs, but it may reflect the differential involvement of RFPs in various signaling mechanisms. Muscle tissues are important in movement and locomotion of animals, even in the relatively sedentary Abalone. Moreover, the cytoskeleton is an important component of muscle cells and plays a significant role in regulating the development of mechanical tension and properties related to its movements (Gunst and Zhang, 2008). Accordingly, the abundant

transcript level of AbRFP may suggest a potential role in signaling pathways related to cytoskeletal organization, which has already been demonstrated as a functional property of other small GTP binding proteins (Takai, *et al.*, 2001). Meanwhile, the strong mRNA expression of AbRFP in mantle tissue may suggest its role in signaling pathways related to secretory functions. In particular, AbRFP may mediate cell signaling cascades related to trafficking of small vesicles between different intracellular compartments, which has already been shown for other molluscan GTP binding proteins (Ye and Carew, 2010), thereby facilitating secretion of shell components.

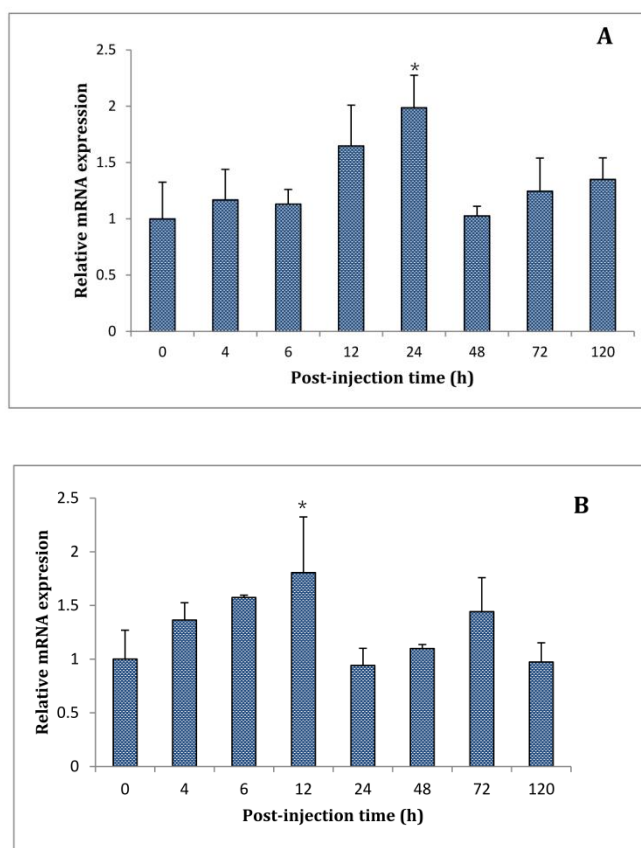
The Rap gene in Chinese white shrimp has been previously shown to be constitutively transcribed in all tissues analyzed, with the highest levels occurring in hemocytes, gill, stomach and hepatopancreas and lowest levels in heart and intestine (Ren, *et al.*, 2012). Similar to the tissue-specific expression profile detected for AbRFP, Ran, another RFP, was ubiquitously expressed in all shrimp tissues examined, including gill, intestine, hepatopancreas, muscle, and hemolymph (Han and Zhang, 2007). Furthermore, a similar transcriptional profile was also detected with respect to the ‘Ran’ homologue in *Haliotis diversicolor* Abalone, with appreciable expression in hemocytes, gills, digestive glands, foot, muscle, and mantle margin (Wu, *et al.*, 2011). In striking contrast to the AbRFP expression profile detected in healthy tissues, another Ran isoform from shrimp was previously detected as intensively expressed in hepatopancreas, as compared to all of the other shrimp tissues analyzed (Han, *et al.*, 2012). Thus, it is possible that the AbRFP may play a unique role in hepatopancreas.

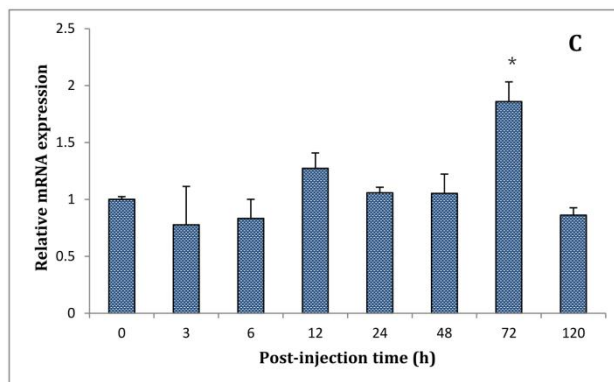
### **Immune response of AbRFP transcription upon pathogen exposure**

In invertebrates, the Ras protein-related signaling pathways are known to be involved in immune modulatory activities (Ragab, *et al.*, 2011). Therefore, in order to evaluate the potential transcriptional responsiveness of AbRFP to immune stimulation by pathogen exposure, gills and hemocytes were collected from immune-challenged animals. These organs are considered as potent immune regulators of marine invertebrate species, as they exert host defense mechanisms against invading pathogens (Philipp, *et al.*, 2012).

As shown in graph A of Figure 6, after healthy Abalones were challenged with VHSV, the transcript level of AbRFP at 24 h time point was significantly up-regulated in gill ( $p < 0.05$ ), after which the expression gradually returned to basal level. However, at none of the time points after VHSV injection, AbRFP transcription was significantly up-regulated in hemocytes (data not shown). Viruses can orchestrate evasion mechanisms against host innate immune responses, which may compromise the host defense system (Iannello, *et*

*al.*, 2006; John, 2009). Thus, the lack of a significant transcriptional response of AbRFP in hemocytes upon VHSV induction may reflect an evasion mechanism of the virus that obstructs an immune signaling mechanism otherwise mediated by RFPs. As previously shown in Chinese white shrimp, infection with the white spot syndrome virus (WSSV) leads to significant up-regulations of Rap transcription in hemocytes at 2 h, 6 h, and 12 h p.i. ( $p < 0.05$ ) and in gill at 6 h p.i. following an initial down-regulation at 2 h p.i. (Ren *et al.*, 2012). Moreover, Ran gene transcription in shrimp hepatopancreas was significantly enhanced upon WSSV injection, and the transcriptional response was maintained in virus-resistant animals over three series of injections given over a period of four weeks (Han and Zhang, 2007). In addition, another study of shrimp Ran transcriptional response to WSSV infection showed significantly enhanced levels, but only at 6 h p.i. ( $P < 0.05$ ) (Han, *et al.*, 2012), and this result complies with the AbRFP mRNA expression pattern detected in gill tissue (Figure. 6A).





**Figure 6.** Relative transcriptional profile of AbRFP in gill upon stimulation with (A) VHSV, (B) *V. parahaemolyticus*, and in hemocytes (C) upon stimulation with *V. parahaemolyticus*, as determined by qRT-PCR. The relative expression was calculated by the  $2^{-\Delta\Delta CT}$  method, using Abalone ribosomal protein gene L5 as the reference gene, and normalized to the corresponding saline-injected controls for each time point. The relative expression fold at 0 h post-injection of un-injected control was set as the baseline for comparisons. Error bars represents the SD (n=3); \*  $p < 0.05$ .

Upon challenge with *V. parahaemolyticus*, AbRFP transcription was significantly ( $p < 0.05$ ) up-regulated in gill and hemocytes at 12 h and 72 h p.i., respectively (Figure. 6B and 6C). Similarly, a previous study of Chinese white shrimp showed that *V. anguillarum* infection significantly elevates the transcription of Rap in hemocytes at 2 h p.i. ( $p < 0.01$ ) but has no effect on the transcription in gill (Ren, *et al.*, 2012).

Collectively, the above observations of AbRFP transcriptional response to experimental immune challenges suggest that AbRFP may play an important role in Abalone host defense, in addition to its general roles in mediating signaling mechanisms contributing to the normal physiologic condition.

### Acknowledgement

This research was supported by Golden Seed Project, Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS).

## References

- Asha, H., de Ruiter, N.D., Wang, M.G. and Hariharan, I.K. (1999). The Rap1 GTPase functions as a regulator of morphogenesis in vivo. *The EMBO Journal* 18, 605-615.
- Barbacid, M. (1987). Ras genes. *Annu Reviews of Biochemistry* 56, 779-827.
- Brunger, A.T., Milburn, M.V., Tong, L., deVos, A.M., Jancarik, J., Yamaizumi, Z., Nishimura, S., Ohtsuka, E. and Kim, S.H. (1990). Crystal structure of an active form of RAS protein, a complex of a GTP analog and the HRAS p21 catalytic domain. *Proceedings National Academy of Science U S A*, 87, 4849-4853.
- Ciocan, C.M. and Rotchell, J.M. (2005). Conservation of cancer genes in the marine invertebrate *Mytilus edulis*. *Environmental Science and Technology*, 39, 3029-3033.
- De Zoysa, M., Pushpamali, W.A., Whang, I., Kim, S.J. and Lee, J. (2008). Mitochondrial thioredoxin-2 from Disk Abalone (*Haliotis discus discus*): Molecular characterization, tissue expression and DNA protection activity of its recombinant protein. *Comparative Biochemistry Physiology B Biochemistry Molecular Biology* 149, 630-639.
- Eisenmann, D. M., Wnt signaling. (2005). *WormBook*, ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.7.1, <http://www.wormbook.org>.
- Goggin, R.J.G. and Perkins, L. (1995). A Protistan parasite of Abalone in Australia. *Marine Freshwater Research* 46, 639-646.
- Gunst, S.J. and Zhang, W. (2008). Actin cytoskeletal dynamics in smooth muscle: a new paradigm for the regulation of smooth muscle contraction. *American Journal of Physiology. Cell Physiology* 295, 576-587.
- Han, F., Wang, X. and Wang, Z. (2012). Molecular characterization of a Ran isoform gene up-regulated in shrimp immunity. *Gene* 495, 65-71.
- Han, F. and Zhang, X. (2007). Characterization of a ras-related nuclear protein (Ran protein) up-regulated in shrimp antiviral immunity. *Fish & Shellfish Immunology* 23, 937-944.
- Huang, C.Y., Liu, P.C. and Lee, K.K.. (2001). Withering syndrome of the small Abalone, *Haliotis diversicolor supertexta*, is caused by *Vibrio parahaemolyticus* and associated with thermal induction. *Z Naturforsch C*. 56, 898-901.

- Iannello, A., Debbeche, O., Martin, E., Attalah, L.H., Samarani, S. and Ahmad, A. (2006). Viral strategies for evading antiviral cellular immune responses of the host. *Journal of Leukocyte Biology*. 79, 16-35.
- John, A.L. (2009). Viral evasion of interferon stimulated genes-review. *Biosciencehorizons*. 2, 212-224.
- Kaplan, W. and Littlejohn, T.G. (2001). Swiss-PDB Viewer (Deep View). *Briefings in Bioinformatics*. 2, 195-197.
- Kikuchi, A. and Williams, L.T. (1996). Regulation of interaction of ras p21 with RalGDS and Raf-1 by cyclic AMP-dependent protein kinase. *Journal of Biological Chemistry* 271, 588-594.
- Konstantinopoulos, P.A., Karamouzis, M.V. and Papavassiliou, A.G. (2007). Post-translational modifications and regulation of the RAS superfamily of GTPases as anticancer targets. *Nature Reviews in Drug Discovery*. 6, 541-55.
- Liu, P.C., Chen, Y.C., Huang, C.Y. and Lee, K.K. (2000). Virulence of *Vibrio parahaemolyticus* isolated from cultured small Abalone, *Haliotis diversicolor supertexta*, with withering syndrome. *Letters in Applied Microbiology*. 31, 433-437.
- Livak, K.J. and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods*. 25, 402-8.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C. Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C. and Bryant, S.H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*. 39, D225-9.
- Milburn, M.V., Tong, L., deVos, A.M., Brunger, A., Yamaizumi, Z., Nishimura, S. and Kim, S.H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* 247, 939-945.
- Mozer, B., Marlor, R., Parkhurst, S. and Corces, V. (1985). Characterization and developmental expression of a *Drosophila* ras oncogene. *Molecular Cell Biology* 5, 885-889.

Nakatsugawa, T., Nagai, K., Hiya, T., Nishizawa, K. and Muroga, A. (1999). A virus isolated from juvenile Japanese black Abalone *Nordotis discus discus* affected with amyotrophia. *Diseases in Aquatic Organisms*. 36, 159-161.

Ogura, T., Tan, A., Tsubota, T., Nakakura, T. and Shiotsuki, T. (2009). Identification and expression analysis of ras gene in silkworm, *Bombyx mori*. *PLoS One* 4, e8030.

Ongvarrasopone, C., Chanasakulniyom, M., Sritunyalucksana, K. and Panyim, S. (2008). Suppression of PmRab7 by dsRNA inhibits WSSV or YHV infection in shrimp. *Marine Biotechnology* (NY) 10, 374-381.

Philipp, E.E., Kraemer, L., Melzner, F., Poustka, A.J., Thieme, S., Findeisen, U., Schreiber, S. and Rosenstiel, P. (2012). Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. *PLoS One* 7, e33091.

Ragab, A., Buechling, T., Gesellchen, V., Spirohn, K., Boettcher, A.L. and Boutros, M. (2011). *Drosophila* Ras/MAPK signalling regulates innate immune responses in immune and intestinal stem cells. *The EMBO Journal*: 30, 1123-1136.

Rattanarojpong, T., Wang, H.C., Lo, C.F. and Flegel, T.W. (2007). Analysis of differently expressed proteins and transcripts in gills of *Penaeus vannamei* after yellow head virus infection. *Proteomics*: 7, 3809-3814.

Ren, Q., Zhou, J., Jia, Y.P., Wang, X.W., Zhao, X.F. and Wang, J.X. (2012). Cloning and characterization of Rap GTPase from the Chinese white shrimp *Fenneropenaeus chinensis*. *Developmental and Comparative Immunology* 36, 247-252.

Roy, A., Kucukural, A. and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5, 725-738.

Takai, Y., Sasaki, T. and Matozaki, T. (2001). Small GTP-binding proteins. *Physiology Reviews* 81, 153-208.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S., (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*. 28, 2731-2739.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,

position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-4680.

Une, C., Gronberg, A., Axberg, I., Jondal, M. and Orn, A., (1991). Phospholipase C treatment of certain human target cells reduces their susceptibility to NK lysis without affecting binding or sensitivity to lytic granules. *Cellular Immunology* 133, 127-137.

Urano, T., Emkey, R. and Feig, L.A. (1996). Ral-GTPases mediate a distinct downstream signaling pathway from Ras that facilitates cellular transformation. *EMBO Journal* 15, 810-816.

van der Oost, R., Beyer, J. and Vermeulen, N.P. (2003). Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environmental Toxicology and Pharmacology* 13, 57-149.

Wittinghofer, A. and Pai, E.F. (1991). The structure of Ras protein: a model for a universal molecular switch. *Trends in Biochemical Sciences* 16, 382-387.

Wu, L., Wu, X. and Du, M. (2011). Identification and expression localization of a Ran homologue in mollusc Abalone, *Haliotis diversicolor supertexta*. *Fish & Shellfish Immunology* 30, 986-91.

Wu, W. and Zhang, X., (2007). Characterization of a Rab GTPase up-regulated in the shrimp *Penaeus japonicus* by virus infection. *Fish & Shellfish Immunology* 23, 438-445.

Wu, W., Zong, R., Xu, J. and Zhang, X. (2008). Antiviral phagocytosis is regulated by a novel Rab-dependent complex in shrimp *Penaeus japonicus*. *Journal of Proteomics Research* 7, 424-431.

Ye, X. and Carew, T.J. (2010). Small G protein signaling in neuronal plasticity and memory formation: the specific role of ras family proteins. *Neuron* 68, 340-361.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40.

Zong, R., Wu, W., Xu, J. and Zhang, X. (2008). Regulation of phagocytosis against bacterium by Rab GTPase in shrimp *Marsupenaeus japonicus*. *Fish & Shellfish Immunology* 25, 258-263.