© Asian Fisheries Society
ISSN 0116-6514

Asian Fisheries Science 23 (2010):577-589

Oomics’ studies for genetic improvement of shrimp in China

Fuhua Li and Jianhai Xiang*

Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

Abstract

This paper mainly summarizes the progress of the research group in omics research on shrimp at the Institute of Oceanology, Chinese Academy of Sciences. With the increasing requirements for genetic improvement of shrimp in order to satisfy the sustainable development of the shrimp industry, we have undertaken research on the omics of shrimp, including transcriptomics, proteomics and genomics, which can provide important information or a platform for the further genetic improvement of shrimp.

Introduction

Production of shrimp aquaculture in China has increased very quickly since the end of the 20th century. However, sustainable development of the shrimp aquaculture industry has been facing difficult issues, including genetic improvement, disease control and remediation of polluted water environment. Therefore, development of knowledge-based technologies is urgently needed. Omics study can provide significant information about the genetic background, immune mechanisms and genetic markers of shrimp. Over the last decade, we have undertaken some studies on the omics of the Chinese shrimp, Fenneropenaeus chinensis and the Pacific white shrimp, Litopenaeus vannamei on DNA, RNA and protein levels. In this paper, we describe the major progress made in this area from our institute.

EST sequencing and data mining

Chinese shrimp, F. chinensis, is one of the most important mariculture species in China. However, its culture industry has been seriously affected by outbreaks of white spot syndrome virus (WSSV) since 1993. The mortality rate of WSSV-infected shrimp has been almost 100% in 3 to 10 days post-infection. Not much information about shrimp genetics and immunity was available at the time of its appearance.

*Corresponding author. E-mail address: jhxiang@qdio.ac.cn
Expressed sequence tags (EST) sequencing from cDNA libraries has been proven as an efficient approach to identify new genes with large scale sequencing as the most direct and effective approach to get detailed genetic information of the shrimp. EST data is very useful in genome annotation, comparative genomics, genetic mapping and marker-assisted selection for genetic breeding. Therefore in 2000, we constructed four cDNA libraries from different tissues, including cephalothorax, hemocytes, eyestalk and ovary. In total, 16000 ESTs were sequenced (Table 1).

Table 1. Sequenced EST numbers from different cDNA library

<table>
<thead>
<tr>
<th>Tissues for cDNA library</th>
<th>cephalothorax</th>
<th>hemocytes</th>
<th>ovary</th>
<th>eyestalk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequenced EST numbers</td>
<td>10446</td>
<td>3483</td>
<td>1517</td>
<td>1283</td>
</tr>
</tbody>
</table>

Through bioinformatic analysis, information on a number of functional genes was obtained. For example, the ESTs from the cephalothorax were clustered into 1399 contigs and 1720 singletons which represent about 3120 unique genes (Xiang et al. 2008). Among the discovered unique genes, 1373 EST clusters are known genes and 1747 EST clusters may represent novel genes that do not match any known genes from other species in the NCBI data bank. In order to find more information about the immune related genes, Shen et al. (2004) analyzed the ESTs from cephalothorax and found 47 contigs and 34 singletons responding to 428 ESTs which showed similarity to known immunity-related proteins. Considering the function of hemocytes in the immunity of shrimp, more immune-related genes were identified through bioinformatic analysis from the cDNA library of hemocytes (Dong and Xiang 2007). Among 2371 ESTs from the hemocytes cDNA library, 34 genes including 177 ESTs have been identified and these genes might be involved in defense or immune functions in shrimp, based on current knowledge. These genes are classified into five categories according to their putative functions in the shrimp immune system: 13 genes are different types of antimicrobial peptides (AMP, penaeidin, antilipopolysaccharide factor, etc.), 11 genes belong to prophenoloxidase system (prophenoloxidase, serine proteinase, serine proteinase inhibitor, etc.), five genes have high homology with clotting protein (lectin, transglutaminase, etc), three genes may be involved in inter-cell signal communication (peroxinectin, integrin etc.) and two genes have been identified to be chaperone proteins (Hsc70, thioredoxin peroxidase etc.). The aforementioned EST sequences information will enrich our understanding of the immune genes of *F. chinensis* and also help our research on the genetic improvement of shrimp.
In addition to ascertaining information on functional genes, EST sequence analysis can also help to obtain genetic markers such as microsatellite markers. Among the 10443 ESTs from the cephalothorax library, 229 SSRs (simple sequence repeats) were identified (Wang et al. 2005) and these could be developed into markers for pedigree analysis and genetic map construction to enhance the genetic breeding of shrimp.

Construction of cDNA microarray and its application

The microarray technique has been proven to be a powerful tool to investigate the expression of thousands of genes in a single hybridization. In order to understand the interaction between shrimp and pathogens such as bacteria or virus, a cDNA microarray including 3136 spots was constructed in *F. chinensis* based on a large batch of the above ESTs. We used a microarray technology to study differentially expressed genes in WSSV-infected shrimp. Gene expression patterns in the cephalothorax of shrimp injected with WSSV 6 h after and moribund shrimp naturally infected by WSSV were analyzed. A total of 105 elements on the arrays showed a similar regulation pattern in artificially infected shrimp and naturally infected moribund shrimp; parts of the results were confirmed by semi-quantitative reverse transcriptase-polymerase chain reaction (RT-PCR). The up-regulated expression of immune-related genes including heat shock proteins (HSP70 and HSP90), trehalose-phosphate synthase (TPS) and ubiquitin C were observed when shrimp were challenged with WSSV. Genes including myosin LC2, ATP synthase A chain and arginine kinase were found to be down-regulated after WSSV infection (Wang et al. 2006).

To compare the gene expression pattern of shrimp to different pathogens such as bacteria and virus, we used microarray technique and analyzed the gene expression profiles of shrimp when they were challenged by WSSV and heat-inactivated *Vibrio anguillarum*, respectively. At 6 h post challenge (HPC), a total of 806 clones showed differential expression profile in WSSV-challenged samples but not in *Vibrio*-challenged samples. A total of 155 clones showed differential expression in the *Vibrio* challenged samples but not in WSSV-challenged samples. In total, 188 clones showed differential expression in both six and 12 HPC-WSSV and *Vibrio* samples. Most of the differentially expressed genes (185/188) were down-regulated in the samples of 12 h post WSSV challenge, whereas they were up-regulated in the samples at six and 12 h post *Vibrio* challenge and 6 h post WSSV challenge (Wang et al. 2008). The data will provide great insight for us in understanding the immune mechanism of shrimp responding to WSSV or *Vibrio*. 
Proteomic approach

Proteins are vital parts of living organisms as they are the main components of the physiological metabolic pathways of cells. Proteomics is the large-scale study of proteins, particularly their structures and functions. The expression profile at protein level can reflect the physiological state of animals. The outbreak of shrimp disease is usually accompanied by environmental stress such as a variation in temperature and/or dissolved oxygen. Hypoxia not only affects the growth of shrimp by reducing their molting frequency and retarding growth, but it also increases the susceptibility of shrimp to infectious disease. To better understand the mechanism of response to hypoxic stress in *F. chinensis*, a proteome research platform was developed in our laboratory and differentially expressed proteins of hepatopancreas in adult Chinese shrimp (between the control and hypoxia stressed groups) were screened. By a two-dimensional gel electrophoresis (2-DE) analysis, 67 spots showed obvious changes after hypoxia. Using LC-ESI-MS/MS, 51 spots representing 33 proteins were identified (Jiang et al. 2009). This study was the first analysis of differentially expressed proteins in the hepatopancreas of shrimp after hypoxia and it provides new insight for further study in hypoxic stress response of shrimp at the protein level.

In order to gain an understanding of the function of the shrimp lymphoid organ at the protein level, we analyzed the proteome of the lymphoid organ in healthy *F. chinensis* through a 2-DE analysis based proteomic approach. A total of 95 spots representing 75 protein entries were identified by liquid chromatography tandem mass spectrometry (LCeMS/MS) with both an online and in-house database. Among them, approximately 36% of proteins related to cytoskeleton are noticeable. Subsequently, a comparative proteomic approach was employed to investigate the differentially expressed proteins in lymphoid organ of *V. anguillarum*-challenged *F. chinensis*. At 24 h post-injection (hpi), 17 differentially expressed protein spots were successfully identified, including four up-regulated protein spots (representing four proteins: cathepsin L, protein similar to squid CG16901-PC, protein kinase C and protein similar to T-complex Chaperonin 5 CG8439-PA), and 13 down-regulated protein spots (representing nine proteins: actin, beta-actin, cytoplasmic actin CyII, alpha tubulin, beta tubulin, protein similar to proteasome delta, vacuolar ATP synthase subunit B, elongation factor 2, carboxypeptidase B) (Zhang et al. 2010). This data may help us to understand the function of the lymphoid organ and the molecular immune mechanism of shrimp in response to pathogen infection.
Immune related genes

Based on the information obtained through EST analysis, cDNA microarray hybridization and proteomic analysis, we have focused on several categories of genes involved in the immune functions of shrimp.

Identification of important genes involved in humoral immune pathways

(1) Pattern recognition receptors (PRRs),

PRRs are proteins expressed by cells of the innate immune system to identify pathogen-associated molecular patterns (PAMPs), which are associated with microbial pathogens or cellular stress. Therefore, PRRs play very important roles in innate immunity. We have identified a series of PRRs in *F. chinensis*, including lectin, toll, LGBP and tetraspanin which are key factors to recognizing different pathogens and triggering a series of immune reactions.

Lectin is regarded as a molecule potentially involved in immune recognition and phagocytosis through opsonization in crustaceans. Knowledge on lectin at a molecular level would help us to understand its regulation mechanism in crustacean immune systems. A novel C-type lectin gene (*Fclectin*) was cloned from hemocytes of *F. chinensis* by 3' and 5' rapid amplification of cDNA ends (RACE) PCR. The deduced amino acid sequence contains a putative signal peptide of 19 amino acids and two carbohydrate recognition domains/C-type lectin-like domains (CRD1 and CRD2). Expression profiles of *Fclectin* gene were greatly modified after bacteria, LPS or WSSV challenge (Liu et al. 2007). This data indicated that C-type lectin might play an important role in the recognition of different pathogens in innate immunity of shrimp and would therefore be helpful in controlling shrimp disease.

Toll-like receptors (TLRs) are a class of proteins that play a key role in the innate immune system. They are single, membrane-spanning, non-catalytic receptors that recognize structurally conserved molecules derived from microbes. Once these microbes have breached physical barriers such as the skin or intestinal tract mucosa, they are recognized by TLRs and immune cell responses are activated. In *F. chinensis*, we cloned the full-length cDNA of Toll (FcToll), which encodes a putative protein of 931 amino acids. The predicted protein consists of an extracellular domain with a potential signal peptide, 16 leucine-rich repeats (LRR), two LRR-C-terminal (LRR-CT) motifs and two LRR-N-terminal (LRR-NT) motifs, followed by a transmembrane segment of 23 amino acids and a cytoplasmic Toll/Interleukin-1R (TIR) domain of 139...
residues. Genomic structure of FcToll gene contains five exons and four introns. Phylogenetic analysis revealed that it belongs to the insect-type invertebrate Toll family. Transcripts of FcToll gene were constitutively expressed in various tissues, with predominant levels in the lymphoid organ. Real-time PCR assays demonstrated that expression patterns of FcToll were distinctly modulated after bacterial or viral stimulation with significant enhancement after 5 h post-\textit{V. anguillarum} challenge but markedly reduced levels immediately after WSSV exposure (Yang et al. 2008). These results suggest that FcToll might be involved in innate host defense especially against the pathogen \textit{V. anguillarum}.

The full length cDNA of lipopolysaccharide and β-1,3-glucan binding protein (LGBP), a pattern recognition protein (PRP) was cloned from \textit{F. chinensis} using homology cloning and RACE techniques. Two putative integrin binding motifs, RGD (Arg-Gly-Asp) and a potential recognition motif for β-1,3-linkage of polysaccharides were observed in the deduced amino acid sequence. The cloned LGBP gene was mainly expressed in hemocytes and hepatopancreas. A significant enhancement of LGBP transcription appeared at 6 h post-injection in response to bacterial infection. When shrimp were challenged by bacteria (\textit{V. anguillarum} and \textit{Staphylococcus aureus}), the transcript of LGBP increased significantly at 6 h post challenge which indicated that LGBP plays an important role in the recognition to bacteria pathogen (Liu et al. 2009).

Besides the aforementioned three types of PRRs, a potential new type of PRRs, tetraspanins/TM4SF was focused. The tetraspanins/TM4SF is a widely expressed superfamily in eukaryotic organisms. Three members of the tetraspanins/TM4SF superfamily were cloned from \textit{F. chinensis}. The deduced amino acid sequences of the three proteins have typical motifs of the tetraspanins/TM4SF superfamily. Phylogenetic analysis of the proteins together with the known tetraspanins of invertebrates and vertebrates revealed that they belong to different tetraspanin subfamilies: CD9, CD63 and tetraspanin-3. CD9, CD63 and tetraspanin-3 showed apparently different tissue distributions in shrimp. The CD9 gene (\textit{FcCD9}) was specifically expressed in hepatopancreas. The highest expression of the CD63 gene (\textit{FcCD63}) was detected in nerves, epidermis and hearts; low expression was detected in hemocytes, ovaries, gills, hepatopancreas and stomachs, and no expression was detected in intestines, muscles and lymphoid organs. Compared with \textit{FcCD9} and \textit{FcCD63}, the tetraspanin-3 gene (\textit{FcTetraspanin-3}) was more broadly expressed and its highest expression was detected in the intestine. All three tetraspanins were markedly up-regulated in the live WSSV-challenged shrimp tissues (Wang et al. 2010). Additionally, dsRNA interference was utilized to examine the functional role of \textit{FcTetraspanin-3} in response to WSSV infection. Results found a decrease of the viral copy number in the tetraspanin knockdown shrimp. These results suggested that \textit{FcTetraspanin-3} might play an
important role in response to WSSV infection and the large extracellular loop (LEL) of tetraspanin might mediate the entry of WSSV (unpublished data). These data suggested that the three cloned members of TM4SF superfamily in Chinese shrimp may play a key role in the route of WSSV infection.

(2) Antimicrobial peptides (AMP)

Antimicrobial peptides (also called host defense peptides) are an evolutionarily conserved component of the innate immune response and are found among all classes of life. Antimicrobial peptides are a unique and diverse group of molecules that are divided into subgroups based on their amino acid composition and structure. Since the first antimicrobial peptide was identified in *Litopenaeus vannamei* in 1987, different groups of antimicrobial peptides including penaeidin, anti-lipopolysaccharide factor (ALF), crustin and lysozyme were found in shrimp. In *F. chinensis*, we have identified the genes encoding different antimicrobial peptides, penaeidin (Kang et al. 2004), ALF (Liu et al. 2005) and crustin (Zhang et al. 2007a; Liu et al. 2008). The recombinant protein of penaeidin (Li et al. 2005) and crustin (Zhang et al. 2007a,b) in *Escherichia coli* or yeast showed apparent antimicrobial activity. Lysozyme activity changed after a bacteria challenge showed that lysozme is also an important antimicrobial effector in shrimp (Yao et al. 2008). In summary, antimicrobial peptides should play crucial roles in a shrimp’s immunity to bacterial infection.

(3) Transcription Factors

Rel/NF-kB transcription factors play key roles in regulating the AMP gene expression in Toll and Imd pathways. In shrimp, different types of AMP were reported. In order to know how the transcription of AMP was regulated in shrimp, we identified two different Rel/NF-kB transcription factors including relish and dorsal. Relish homolog (FcRelish) was cloned from *F. chinensis*. The full length cDNA of FcRelish consists of 2,157 bp, including 1,512 bp open reading frame, encoding 504 amino acids. FcRelish had the highest expression levels in the hemocytes and lymphoid organ. Both *V. anguillarum* and *Micrococcus lysodeikticus* stimulation to shrimp can affect the transcription profile of FcRelish. Silencing of FcRelish through dsRNA interference can greatly change the transcription profile of AMP (Li et al. 2009b). Furthermore, another transcription factor, Dorsal homolog (FcDorsal), was cloned from *F. chinensis*. The full length cDNA of FcDorsal consists of 1627 bp, revealing a 1071 bp open reading frame encoding 357 amino acids. Spatial expression profiles showed that FcDorsal had the highest expression level in the hemocytes and lymphoid organ (Oka). The expression profiles in the hemocytes and lymphoid organ were apparently modulated when shrimp were stimulated by bacteria or WSSV. Both Gram-positive (G+) bacteria (*M. lysodeikticus*) and Gram-negative (G−) bacteria (*V. anguillarum*) injection to shrimp
caused the up-regulation of FcDorsal at the transcription level. DsRNA approach was used to study the function of FcDorsal and the data showed that FcDorsal was related to the transcription of penaeidin 5 in shrimp (Li et al. 2010). Through comparison of the expression profiles between FcDorsal and FcRelish in shrimp responsive to WSSV challenge, we speculate that FcDorsal and FcRelish might play different roles in shrimp immunity.

**Antioxidant enzymes**

The main function of antioxidant enzymes is neutralizing free radicals. When the organism encounters pathogen infection or environmental stress, its NADPH-oxidase will activate, in turn enhancing the glycolytic reactions that will increase the consumption of oxygen and induce the production of a mass of reactive oxygen species (ROS) such as superoxide anion (O$_2^-$), hydrogen peroxide (H$_2$O$_2$) and hydroxyl radical (OH$^-$). Though ROS can kill foreign invaders, the mass accumulation of these reactive molecules in animals will cause serious host cell damage. So the rapid elimination of these excessive ROS is essential for the proper functioning of cells and the survival of organisms. To protect against toxicity and eliminate ROS, organisms have evolved protective enzymatic systems including superoxide dismutase, catalase and many kinds of peroxidases. In *F. chinensis*, we identify several oxidant enzymes including mMnSOD (Zhang et al. 2007c), peroxiredoxin (Zhang et al. 2007d) and catalase (Zhang et al. 2008) that are closely related to the immune response of shrimp to pathogens.

**Chaperone proteins**

When studying the differential expressions of genes or proteins of shrimp after pathogen infection or stress through cDNA microarray or proteomic approach, we found that chaperone genes showed apparent variation at transcription or translation level. This indicates that chaperone proteins play a very important role in shrimp. We have cloned different chaperon protein genes including HSC70 (Jiao et al. 2004), HSP70 (Luan et al. 2010), HSP90 (Li et al. 2009a), glucose regulated protein 78 (GRP78) (Luan et al. 2009) and calreticulin (Luan et al. 2007). Their expression of different genes was analyzed. The data further indicated that chaperone proteins were essential for the immunity homeostasis in shrimp.

**Construction of genetic map**

A primary genetic linkage map is a prerequisite to detailed genetic studies in any organism. High-density linkage maps are necessary for efficient mapping of quantitative trait loci (QTLs) to complement marker-assisted selection (MAS) and for comparative genome mapping. Due to the worldwide importance of *L. vannamei* in
shrimp aquaculture, a genetic linkage map was constructed using amplified fragment length polymorphism (AFLP) and microsatellite markers. One hundred and eight select AFLP primer combinations and 30 polymorphic microsatellite markers produced 2071 markers that were polymorphic in either of the parents and segregated in the progeny. Of these segregating markers, 319 were mapped to 45 linkage groups of the female framework map covering a total of 4134.4 cM and 267 markers were assigned to 45 linkage groups of the male map covering a total of 3220.9 cM. A sex-linked microsatellite marker was mapped on the female map with 6.6 cM to sex and a LOD of 17.8. Two other microsatellite markers were also linked with both 8.6 cM to sex and LOD score of 14.3 and 16.4 (Zhang et al. 2007e). The genetic maps presented here will serve as a basis for the construction of a high-resolution genetic map, quantitative trait loci (QTLs) detection, marker-assisted selection (MAS) and comparative genome mapping.

Construction of BAC library and genome sequencing

To facilitate gene cloning and characterization, genome analysis, physical mapping and molecular selective breeding of marine shrimp, we have developed the techniques to isolate high-quality megabase-sized DNA from hemocyte nuclear DNA of female shrimp and constructed a bacterial artificial chromosome (BAC) genomic library for the species. The library was constructed in the Hind III site of the vector pECCBAC1 and it consists of 101760 clones arrayed in 265, 384-well microtiter plates; the average insert size is 101 kb and the genome is covered approximately fivefold. To characterize the library, 92160 clones were spotted onto high-density nylon filters for hybridization screening. A set of 18 pairs of overgo probes designed from eight cDNA sequences of *L. vannamei* genes were used in hybridization screening and 35 positive clones were identified (Zhang et al. 2010). These results suggest that the shrimp BAC libraries will provide a useful resource for screening of genomic regions of interest candidate genes, gene families or large-sized synthetic DNA regions. In addition, the libraries will promote future work on comparative genomics, physical mapping and large-scale genome sequencing in the species. Genetic information especially whole genome sequence is necessary to better understand and improve shrimp. However, the research on shrimp genome is still very limited; until Sept. 2010, there were only about 170000 EST sequences and a small quantity of genes of *L. vannamei* published in NCBI. Next-generation sequencing (NGS) technologies have been demonstrated as a feasible way to sequence large eukaryotic genomes with high-throughput and rapid, cost-effective de novo assembly and re-sequencing. Using NGS technologies, different insert-size pair-end and mate-pair libraries were constructed in *L. vannamei* and sequenced by Solexa. A total of 125-gigabase data were generated successfully with an average reads length 80 bp; another 390-megabase data were obtained by GS FLX
(454). With the Kmer frequency of Solexa data and C-value, we predicted that the genome size of *L. vannamei* was about 2.4-gigabase, suggesting that more than 51× coverage data of the genome size of *L. vannamei* has been obtained (Xiang, reported at PAG XIX, San Diego, USA 2011).

**Selective breeding**

Through seven generations of selective breeding in *L. vannamei*, a selected line called “KeHai No. 1” was obtained which showed an increase of 12-42% in growth rate when cultured at different densities. The selected line can achieve superior growth at higher culture density than unselected shrimp. During the selective breeding, microsatellite markers were used for pedigree identification which is a proven selective breeding strategy.

In summary, the discovery and research on a batch of immune related genes of shrimp provides a possible opportunity to develop disease control strategies and also provides a platform for marker selection for disease resistant shrimp. The construction of a genetic map, genome sequencing and the establishment of a selective line provide the basement for QTL mapping, marker assistant selective breeding and whole genome selective breeding. The improvement of shrimp is necessary to satisfy the requirements of the shrimp aquaculture industry.

**Acknowledgements**

The research has been supported by the Natural Science Foundation of China, the Ministry of Science and Technology, and the Chinese Academy of Sciences. We extend our appreciation to Dr. Cheng-Sheng Lee for his generous invitation to participate in the 2009 AIP Shrimps Conference and to Meredith Brooks for editing this paper for publication.

**References**


