2. Advances in Genomics and Genetics of Penaeid Shrimp

Javier Robalino¹, Robert W. Chapman², Enrique de la Vega³, Nuala A. O’Leary⁴, Danielle M. Gorbach⁵, Zhi-Qiang Du⁵, Max F. Rothschild⁵, Craig L. Browdy⁶, Gregory W. Warr⁴, Yannick Labreuche⁴

¹ Facultad de Ingeniería Marítima y Ciencias del Mar, Escuela Superior Politécnica del Litoral, Prosperina, Guayaquil, Ecuador
² Marine Resources Research Institute, South Carolina Department of Natural Resources, Charleston, SC, USA
³ CI OCEANOS, Cartagena, Colombia
⁴ Marine Biomedicine and Environmental Sciences Center, Medical University of South Carolina, Charleston, SC, USA
⁵ Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, Ames, IA, USA
⁶ Novus International Inc., Charleston, SC, USA

Abstract:

This chapter contains sections titled:

- Introduction
- EST Collection as an Approach to Gene Discovery in Shrimp
- Medium- to High-Throughput Studies of Differential Expression and Gene Discovery
- RNAi-Based Applications in Shrimp Aquaculture: From Reverse Genetics to Control of Diseases
- Markers, Genetic Maps, and Large Insert Genomic Libraries in Shrimp
- Analytical Challenges in Genomics and Genetics of Shrimp
- Concluding Remarks
- Acknowledgments
- References

Keywords: penaeid shrimp; EST collection, gene discovery; gene expression “snapshot”; gametogenesis; abiotic stress responses; differential expression cloning; RNAi-based applications, shrimp; targeted gene silencing; RNAi, and gene function; genetic markers in breeding
Introduction

Penaeid shrimp constitute one of the most important groups of species for aquaculture worldwide, ranking second overall in value in 2006 (FAO 2007). Sustainable development of this industry could greatly benefit from progress in our basic knowledge of genetics, genomics, and molecular immunology of shrimp. Recently, the application of high throughput molecular tools and approaches has led to significant developments in these fields. This chapter describes the current status of efforts to catalogue the transcriptome of shrimp through the collection, curation, data base development, and analyses of expressed sequence tags (ESTs) and through analysis of differential gene expression. Advances in the development of other genetic resources such as genetic maps and genomic libraries are also discussed. We emphasize how these tools can and are providing new molecular information about biological processes of relevance to shrimp aquaculture, such as immune responses and reproductive physiology. We then consider the contributions of reverse genetics through RNA-mediated interference (RNAi), which is now being used to test the involvement of specific genes in aquaculture-relevant traits. Proof-of-concept studies also demonstrate that RNAi is a promising approach to the development of antiviral therapies. Future developments in shrimp genetics and genomics that will further contribute to advancing biotechnological applications in aquaculture are also discussed throughout this chapter.

EST collection as an approach to gene discovery in shrimp

ESTs are short DNA sequences generated by large scale single-pass sequencing of randomly picked cDNA clones from libraries. They generally represent a single tissue or condition of interest at a given time thus providing a “snapshot” of the physiological status of a cell, tissue or an organism. EST sequencing projects can provide an efficient and rapid means for discovering new genes, alleles and polymorphisms, thereby providing data on gene expression and regulation, and for the development of genome maps. Genomic research by EST analysis has been conducted for several shrimp species including Penaeus monodon, Fenneropenaeus chinensis, Marsupenaeus japonicus, Litopenaeus setiferus, and Litopenaeus vannamei (Table 1). Most of these analyses have been small scale sequencing efforts conducted by individual laboratories studying mostly shrimp immunity and disease, and have contributed to the identification of a significant number of previously undescribed genes.

In contrast to small scale EST projects, when large sequencing efforts are conducted the rate of return of novel genes rapidly decreases as sequencing progresses. This is caused by a small number of genes producing large quantities of mRNA and appearing as highly redundant in the EST clones being sequenced. Therefore, for this kind of initiative it is recommended to use some form of normalization or redundancy subtraction to maximize the rate of gene discovery (Soares et al. 1994). We have used a direct redundancy subtraction method which involves subtracting highly redundant genes from arrayed cDNA libraries (E. de la Vega and N. O’Leary, unpublished). Briefly, the process involves producing a typical cDNA library, which is then plated onto large agar plates with antibiotic selection and single colonies are picked robotically into 384 well plates. The resulting
library is robotically spotted at high density onto nitrocellulose membranes and direct redundancy subtraction is achieved by sequencing 384 samples and identifying the most redundant clones. These clones are then used to produce probes which are hybridized to the membranes, highly redundant genes are thereby identified, and those clones which did not hybridize to the probes are then “cherry picked” and re-arrayed. This whole process yields a low redundancy library which increases the rate of gene discovery (Figure 1). This approach has been used to sequence more than 150,000 ESTs from hemocytes, gills, hepatopancreas, lymphoid organ, ventral nerve cord and eyestalk of *L. vannamei*, resulting in the identification of around 15,000 unigenes. All these ESTs are publicly available at the NCBI EST database (http://www.ncbi.nlm.nih.gov/dbEST) or at the Marine Genomics website (www.marinegenomics.org). This and other efforts have contributed, as of January 2009, a total of 179,032 ESTs for Penaeid shrimp (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html).

Although the comprehensive annotation of the now abundant EST data available for *L. vannamei* is at a very early stage, it is expected to yield a large number of novel genes involved in immunity, respiration, endocrinology, and digestion, among others, providing an opportunity to better understand the physiology of shrimp. In addition, large numbers of molecular markers such as single nucleotide polymorphisms (SNPs) or microsatellite markers are expected to be mined out of these sequences, making them available for shrimp breeding programs interested in using marker assisted selection. Some efforts in this respect have already been reported (see discussion later in this chapter) and show promise in the ability to use this EST information for marker development.

### Medium to high throughput studies of differential expression and gene discovery

While ESTs can provide a “snapshot” of gene expression in a tissue of interest, it is not a cost-effective approach to identifying differential gene expression in multiple samples. Differential expression is often of particular interest as a first approach to understanding gene function. The presumption that expression of a gene is largely restricted to the cell or tissue where it is needed, at the time when it is needed, provides the rationale for the significant attention given to this issue. In shrimp of aquaculture significance, medium to high throughput methods for assessing changes in gene expression have been most intensively applied to studying the response to pathogens and to stimulators of the immune system.

For species where the expected gene content vastly exceeds the number of known genes (which is the case in shrimp), methods of assessing changes in gene expression that do not rely on previously known sequence information are particularly useful. Two widely used methods of this kind, Suppression Subtractive Hybridization (SSH, (Diatchenko et al. 1996)) and mRNA Differential Display (DD, (Liang and Pardee 1992)), are aimed at the isolation of cDNA fragments derived from differentially expressed genes. Comparative analyses of Expressed Sequence Tags (ESTs), isolated from tissues or cells subjected to conditions of interest, can also be used as a means of identifying genes potentially regulated by such conditions (see Table 1 for examples). All these methods are exploratory in nature, each subject to caveats and biases of their own, such that
independent experimental confirmation of differential expression by direct methods (such as quantitative reverse-transcription PCR) is essential. SSH and DD are especially limited in terms of direct discovery of new genes in understudied species, because they involve the isolation of short cDNA fragments, rather than full-length cDNAs. SSH and DD are also less amenable than traditional EST mining to high throughput sequencing, as the complexity of the cDNA pools obtained by these methods is generally low. Processes of interest to shrimp aquaculture that have been studied using differential expression cloning approaches include gametogenesis, abiotic stress responses, and immune responses (Table 2). Many of the sequences isolated from these studies do not allow identification of genes with predictable homology or function. Presumably, some of these sequences correspond to novel genes, while others represent poorly conserved and/or short regions from otherwise conserved genes.

A markedly different approach to study differential gene expression at high throughput is the use of microarray technology. Microarrays are arrangements of DNAs immobilized to solid supports in such a way that each element in the arrangement corresponds to a single known DNA sequence. Probing these arrangements with mRNA (or the corresponding cDNA) obtained from animals of interest allows a quantitative assessment of the expression of each gene whose sequence is represented on the microarray. Remarkably, thousands to millions of such DNA elements can be simultaneously evaluated due to the capacity of current technology to immobilize DNA on surfaces at extremely high densities. Microarrays with gene contents ranging from a few dozen to a few thousand genes have been generated and used to study expression in L. vannamei, P. monodon, L. stylirostris, M. japonicus, and F. chinensis. In Table 3 some of these studies are summarized, in terms of the nature of the microarrays developed, and in terms of their most significant findings potentially relevant to aquaculture.

New gene discovery and the identification of differentially expressed genes provide only an initial step towards better understanding shrimp biology, and towards using such information to improve shrimp aquaculture. For almost every shrimp gene considered in the studies described in tables 1, 2 and 3, the most critical experimental information necessary to understand its function is still lacking. However, for some of the genes found to be regulated during immune challenge, functional follow-up studies have started to define their roles in the defense from pathogens (see next section on shrimp RNAi studies). These types of investigations can begin to elucidate immune pathways, improving our fundamental understanding of shrimp immunity. Furthermore, such genes represent good candidates for markers of disease resistance in breeding efforts, and for the development of tools useful to monitor shrimp health in aquaculture settings. Thus, the discovery of genes that are differentially expressed in situations of interest, coupled with subsequent functional characterization of the gene products, provide important avenues to the development of biotechnological tools for aquaculture.

### RNAi-based applications in shrimp aquaculture: from reverse genetics to control of diseases

First described in nematodes (Fire et al. 1998) and subsequently in most eukaryotes studied so far, RNA interference (RNAi) is a highly conserved nucleic acid-based
mechanism, mediating sequence-specific targeted gene silencing. This machinery is initiated by double-stranded RNA (dsRNA), which is processed by a Dicer family member into small effector RNA duplexes (e.g. siRNAs, miRNAs, esiRNAs, etc). The small RNAs are incorporated onto a multimeric protein complex, the RNA-induced silencing complex (RISC) and related complexes, which mediate targeted degradation, translational repression, and other silencing phenomena by means of complementary base-pairing. Because dsRNA or short interfering RNAs (siRNAs) can be supplied exogenously to trigger specific gene silencing, RNAi has rapidly become the most widely used gene silencing tool in a broad variety of eukaryotic organisms [reviewed in (Campbell and Choy 2005)]. In this section, we review current RNAi-based research in shrimp and discuss future applications of this phenomenon in shrimp aquaculture.

**RNAi mechanisms in shrimp**

The existence of an intact RNAi machinery in shrimp was first proposed by Robalino and collaborators, who demonstrated in *L. vannamei* that *in vivo* administration of dsRNA induced a down-regulation of endogenous or viral gene expression in a sequence-specific manner (Robalino et al. 2005). Additional evidence came from the identification of possible RNAi pathway components in the black tiger shrimp *P. monodon*: Pem-AGO (Dechklar et al. 2008); Pm-Ago, another member of the Argonaute protein family (Unajak et al. 2006), and a Dicer nuclease, Pm-Dcr1 (Su et al. 2008). However, more work will be necessary to define the roles (if any) played by each of these components in shrimp RNAi. Nevertheless, these results strongly suggest the existence of functional RNAi in shrimp and have opened the possibility, for the first time, of using reverse genetic approaches to understand gene function in these organisms.

**Methods to trigger gene silencing by RNAi in shrimp**

In shrimp, targeted gene silencing has been experimentally induced by several methods: *in vivo* injection, *in vitro* delivery to primary cell cultures, and feeding bacteria carrying dsRNA *in vivo*.

**Injection of dsRNA:** Several studies have shown in different shrimp species the ability of ectopic dsRNA to spread from the site of injection (pereiopod, tail muscle) to distant tissues to mediate a potent depletion of cognate mRNAs, thus facilitating reliable and reproducible gene silencing experiments (Table 4). However, several factors that are likely to affect silencing efficiency, such as target sequence selection and dsRNA dose, are still in need of detailed study in the different shrimp species. Altogether, injection of long dsRNA is to date the most widely used method for interrupting gene expression in shrimp *in vivo*.

Injections of siRNAs to induce knockdown of target gene expression have also been reported in shrimp (Table 4). Generated by cleavage of dsRNA by Dicer, siRNAs are 21–25-mer duplexes that confer sequence specificity to the silencing complexes. In mammals, siRNAs are extensively used because of their great specificity and efficacy, but also because these duplexes are small enough to bypass the induction of potent immune responses elicited by longer dsRNAs (Elbashir et al. 2001). In shrimp, important discrepancies in obtained results have been observed when using siRNAs *in vivo* (see...
Table 4). Results ranging from no silencing at all to highly effective silencing have been reported. The existence of different mechanisms for uptake of siRNA, selection and design methods of target sequences for RNAi or duration of siRNA activity in vivo have been hypothesized to explain these inconsistencies (Shekhar and Lu 2009). These issues warrant further investigation.

In vitro delivery to primary cell cultures: Few studies have demonstrated in shrimp the application of RNAi in transfection experiments (Tirasophon et al. 2005; Assavalapsakul et al. 2006; Dechklar et al. 2008). Because of the unavailability of clonal long-term cell lines in marine invertebrates (Rinkevich 2005), gene-specific dsRNA transfection experiments have been performed into primary cultures of lymphoid (Oka) cells, leading to down-regulation of endogenous or viral cognate mRNAs. Recently, Treerattrakool et al. described successful RNAi-mediated knockdown by dsRNA adjunction to P. monodon eyestalk ganglia and abdominal nerve cord explant cultures (Treerattrakool et al. 2008). These studies expand significantly the experimental capabilities for exploring gene function in shrimp, allowing the study of organ-specific phenomena outside of the context of the whole animal.

Feeding or bacteria-mediated RNAi: Establishing RNAi by bacterial feeding was first envisioned in Caenorhabditis elegans (Timmons and Fire 1998) and later in planarians (Newmark et al. 2003), plants (Tenllado et al. 2003), and insects (Walshe et al. 2008). In shrimp, progress has been made recently in the use of non-pathogenic bacteria to induce gene silencing. Sarathi and colleagues were the first to report the production of dsRNA using prokaryotic expression systems for RNAi purposes in P. monodon (Sarathi et al. 2008a). Later, these authors investigated in vivo whether bacterially expressed dsRNA derived from viral sequences could specifically interfere with viral infection following oral administration. Two different delivery approaches were tested: animals were fed with pellet feed coated either with inactivated bacteria containing overexpressed dsRNA, or with bacterially expressed dsRNA entrapped onto chitosan nanoparticles. Both feeding methodologies were conclusive, the best results being obtained when feed was coated with inactivated bacteria expressing dsRNA (Sarathi et al. 2008b).

RNAi as a tool to unravel gene function in vivo

As previously discussed, EST mining and differential expression cloning have been applied in shrimp to gain insight into gene function. Ultimately however, experimental manipulation of expression is necessary to directly address issues of function. In shrimp, RNAi is to date the only reliable tool for this type of experimentation. An increasing number of studies have resolved, in the recent past, gene functions involved in moulting, osmoregulation, reproduction, or glucose metabolism in shrimp by using gene-specific dsRNA/siRNA technology (Table 4). Although gene silencing by dsRNA can complicate immunology studies in shrimp due to non-specific stimulation of antiviral responses (Robalino et al. 2004), it is possible to perform carefully designed experiments to explore at least some aspects of their immune system (de la Vega et al. 2008; Shockey et al. 2008; Amparyup et al. 2009). These studies are of great relevance to biotechnology in aquaculture, as they pave the way to start defining the relationship between known genes and traits of commercial importance.
RNAi as a tool to promote antiviral protection in shrimp

The spread of infections due to geographically widespread viruses has led to huge economic losses in the shrimp farming industry, threatening its economic viability and long-term sustainability in some regions of the world (Lightner and Redman 1998; Lightner et al. 2006). To control the occurrence and spread of viral diseases, new strategies for disease prevention, such as exposure to inactivated virus or viral proteins, are currently being developed [reviewed in (Johnson et al. 2008)]. Because the RNAi machinery allows gene silencing in a highly sequence-specific manner with little or no risk of undesired off-target effects, injections of viral gene-specific dsRNA/siRNA into shrimp seem to be a more powerful and attractive tool to inhibit viral replication and/or protect shrimp from viral infections than other methods. Indeed, this strategy has proven to be effective against three unrelated viruses (Table 1) [see (Shekhar and Lu 2009) for a review]. Furthermore, as previously discussed, feeding of dsRNA has already been proven as a feasible approach to block viral pathogenesis in vivo (Sarathi et al. 2008b), opening the door to the development of dsRNA-based treatments applicable at a commercial scale.

Markers, genetic maps, and large insert genomic libraries in shrimp

The development of genetic markers for breeding purposes has been an area of significant interest among shrimp researchers for several years. Several types of markers, including microsatellites, Restriction Fragment Length Polymorphisms (RFLPs), Amplified Fragment Length Polymorphisms (AFLPs), and Randomly Amplified Polymorphic DNAs (RAPDs) have been developed, to different extents, in several Penaeid species [for examples see (Garcia et al. 1994; Xu et al. 1999; Wilson et al. 2002; Meehan et al. 2003)]. Here, we will focus mostly on how the new genomic resources previously described in this chapter have provided new opportunities for the identification of Single Nucleotide Polymorphisms (SNPs). Some significant efforts in the generation of genetic maps with large numbers of markers of diverse types will also be mentioned. Finally, we will briefly comment on the status of large DNA libraries for shrimp, a resource that will prove essential for the future of genomics in these species.

Single Nucleotide Polymorphisms (SNPs)

Panels of genetic markers have been developed for both L. vannamei and P. monodon, consisting primarily of microsatellites and AFLPs, although a limited number of SNPs have also been reported (Glenn et al. 2005). Little genomic data is currently available for shrimp species and most of the sequence data has originated from EST libraries, which commonly may contain many sequencing errors. Consequently, locating new polymorphisms is a slow and arduous task which is best completed through the use of computer predictions or large-scale sequencing techniques. Due to the limited alternatives, the primary method for in silico prediction of SNPs has been comparison of ESTs displaying sequence similarity, despite the challenges of distinguishing sequencing errors from true base differences. One pipeline used for such predictions is to cluster the available ESTs with CAP3 (Huang and Madan 1999) and predict SNPs using SNPidentifier (Gorbach et al. 2009), which has built-in quality control measures to compensate for the inherent errors in EST sequences. Thus far, this method has produced the most validated SNPs in shrimp (768 in L. vannamei) of any published process. When the same method
was tried using ESTs from nine other shrimp species and validated in vivo only in L. vannamei, the success rate of SNP prediction was decreased significantly (from 44% to 11%)(Gorbach et al. submitted)

A smaller-scale SNP identification method that has also been utilized is sequencing candidate genes. For a gene that is suspected to play a role in an important phenotype, primers are designed to sequence a region of the gene. When successful, one or more SNPs can be identified within the sequenced region. This process has identified at least 12 SNPs in L. vannamei and at least 3 SNPs in P. monodon (Glenn et al. 2005; Yu et al. 2006; Zeng et al. 2008). One of the SNPs discovered using this process has been placed on the P. monodon linkage map (Maneeruttanarungroj et al. 2006). Furthermore, sequences from various other resources can also be used to identify SNPs, e.g. short tandem repeats and AFLPs.

With the development of next generation sequencing platforms, such as the Roche Genome Sequencer FLX™ system, Illumina Genome Analyzer™ and Applied Biosystems SOLiD™ sequencing system, large-scale SNP discovery has become extremely fast and efficient, for a lower cost. Currently, none of these technologies have been applied to shrimp, but this is certainly the future for genomics in general and therefore where shrimp genomics should go in the ensuing years.

Genetic maps

Linkage maps based on microsatellite and AFLP markers have been published for both L. vannamei (Zhang et al. 2007) and P. monodon (Maneeruttanarungroj et al. 2006). A more complete linkage map for P. monodon has been constructed from AFLP markers (Staelens et al. 2008). This map identified some 44 male and 43 female linkage groups and, given the number of shrimp chromosomes (44, (Chow et al. 1990)), is likely to include the majority of chromosomal linkage groups in this species. A new linkage map for L. vannamei is currently under construction (Z.-Q. Du, D.M. Gorbach, and M.F. Rothschild, unpublished), which incorporates a large number of SNP markers, and may be further extended with the output from next generation sequencing technology.

Large insert size genomic libraries

Since the development of cloning methods for large DNA segments (O’Connor et al. 1989) Bacterial Artificial Chromosomes (BACs) have become the preferred means for construction of libraries containing most if not all of the genome. Usually BACs contain inserts of size 150-350 kbp, but can be as large as 750 kbp. There are several reasons for constructing such large insert libraries, which include reduced overlap of individual clones, reduction in the number of clones that must be maintained for complete genome coverage, efficiencies in DNA sequencing and genome mapping, and source material for the creation of transgenic organisms. Unfortunately, generating a stable large insert BAC library for shrimp has proven to be difficult. Several mid-sized (50 kbp) fosmid libraries have, however, been generated and at least one of these is publicly available (http://www.genome.clemson.edu/). While these tools are available, they have not been exploited and, thus, great opportunities are available to the community in this area.
Analytical challenges in genomics and genetics of shrimp

A well described characteristic of genomics studies is the generation of massively paralleled data sets, which will pose unprecedented analytical challenges. The well known problems are how to manage large quantities of data and mine these data for significant biologically meaningful signals (Warr et al. 2007). These are issues for all genomics studies regardless of species or the research area of interest, and the topic is frequently reviewed in the literature. Rather than rehash the issue, we chose to discuss a different topic that may be able to address some problems specific to shrimp aquaculture and potentially revise the way selective breeding in cultured species is prosecuted.

It is well known that phenotypic selection for quantitative traits in shrimp is hampered by the low observed heritabilities of certain desired traits (Perez-Rostro and Ibarra 2003) as opposed to typical values in agriculture species. This observation is consistent with what one might expect in a species, e.g. shrimp, which do not have a long history of domestication and inbreeding. For a trait like growth, which is likely influenced by a large number of genes and the environment, the absence of a long history of domestication reduces the probability that many loci will have gone to fixation by chance alone. This in turn makes it more difficult to detect the influence of individual genes that remain polymorphic in the population.

The high dimensionality of massive genomics data imposes a heavy burden on the analytical methods. How does one identify the important genetic markers when thousands to millions of genetic markers are available and only a small number of them are necessary for maximal information content? This is identical to the dimensionality reduction problem encountered in microarray analysis. Traditional linear algebra approaches (e.g. ANOVA) can only be useful for a limited number of markers, and cannot predict the phenotype when there are non-linear effects among alleles at a single locus or multigene interactions, which is almost certainly the case for most quantitative traits (cf. (Brockmann et al. 2000; Liu et al. 2007; Han et al. 2008; Ankra-Badu et al. 2009) for examples).

Different statistical and computational methods have been proposed to address the issue of dimension reduction, for instance, Bayesian Information Criteria and penalized likelihood (cf. (Manichaikul et al. 2008)), with the general conclusion that the penalized LOD score approach provides an accurate selection of QTLs and their interactions, at least in low dimensional space. Though promising, the approach taken by Manichaikul et al. (2008) is a rather brute force method which does not explore the global state space and could be more easily accomplished with Artificial Intelligence (AI) tools such as Genetic Algorithms and Genetic Programming optimization of Neural Networks (GPNN)(Ritchie et al. 2003; Motsinger et al. 2006). Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) have also been used for microarray and genotyping data (We note that SVMs using sigmoidal kernels are equivalent to a single layer ANN). The general conclusion from studies employing the AI tools is that ANNs can detect the influence of multiple genetic and environmental factors on phenotypic traits but can be computationally expensive as a dimensionality reduction tool. In our own work, SVMs using polynomial kernels have been more efficient in dimensionality reduction than ANNs, but ANNs generally provide superior classifications and models of system dynamics (R.W. Chapman, unpublished). Furthermore, advances in analytical methods in other species which are already using genomic selection are likely to be applicable for improving breeding systems in aquaculture species.
Concluding Remarks

The last decade of research and development in shrimp genomics and genetics has seen significant advancements. An abundance of sequence information from expressed genes is available in public databases, providing a first glimpse at the gene content of several Penaeid species. Researchers and aquaculture geneticists are only now beginning to exploit these resources, especially for the identification of genetic markers, candidate disease resistance genes, and genes linked to reproduction and other aquaculture-relevant processes. A growing number of initiatives are mining the available sequence data to implement markers and generate increasingly more extensive linkage maps. The number of shrimp genes for which at least some sequence information is now known is in the thousands, rather than the dozens, as was the case just a few years ago. The tools are also in place to refine the selection of candidate aquaculture-relevant genes through the characterization of two key aspects of the function of a gene: its expression and its loss-of-function phenotype. We can now measure the expression of thousands of these genes simultaneously, thanks to progress in transcriptomic methodologies, although much effort is still necessary within the community to meet the analytical challenges involved. Remarkably, it is now possible to block the expression of a gene in a targeted manner in vivo (using RNAi) a task that just a few years ago seemed only feasible in a few well-established animal models. A significant step forward in the control of viral diseases, one of the most important constraints to shrimp aquaculture, seems just around the corner. It is now possible to afford almost complete protection from highly pathogenic viruses by delivering dsRNA molecules that are produced by relatively simple methods. So, what is the future of shrimp genetics and genomics as it relates to aquaculture? We propose that future success in this area is closely linked to the ability of the community to work cooperatively, sharing information and resources in benefit of the overall advancement of the science. The relative lack of unified efforts and multilab resources such as EST databases, microarray platforms, and library repositories, are indicators of the need to strengthen the ties among shrimp researchers. The challenges imposed by the sheer magnitude of genomic projects will likely be met only by a unified community, with a demonstrated capacity to gain the most out of shared resources. Such an evolution of attitudes will likely make it feasible to support a full-genome sequencing project for a Penaeid shrimp in the near future.

References


Figure 1. Expected gene frequency distribution in an unmodified vs a depleted or modified cDNA library. Removing highly redundant gene sequences from a cDNA library results in a better rate of return of unique sequences with the same sequencing effort.
Table 1. Studies reporting characterization of ESTs in Penaeid shrimp

<table>
<thead>
<tr>
<th>Species</th>
<th>Tissues and conditions under study</th>
<th>#ESTs/uniqu e sequences</th>
<th>Principal findings/genes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pm</td>
<td>Testes</td>
<td>896/NA</td>
<td>Discovery of genes involved in reproductive maturation and sex determination. Description of testes specific genes.</td>
<td>(Leelatanawit et al. 2009)</td>
</tr>
<tr>
<td>Pm</td>
<td>lymphoid organ from Vibrio harveyi infected animals compared to non-infected animals</td>
<td>408 normal; 625 infected/NA</td>
<td>Discovery of genes differential expressed in response to bacterial challenge. Discussion of the cathepsin family (L and B). pmonodon.biotec.or.th/home.jsp</td>
<td>(Pongsomboon et al. 2008b)</td>
</tr>
<tr>
<td>Pm</td>
<td>vitellogenic ovaries</td>
<td>1051/559</td>
<td>Identification of sex-related genes. Further analysis of chromobox protein (CBX) which is preferentially expressed in ovaries.</td>
<td>(Freechaphol et al. 2007)</td>
</tr>
<tr>
<td>Pm</td>
<td>Postlarvae infected with WSSV</td>
<td>6,671 normal and 7,298 infected/9622</td>
<td>Gene discovery in postlarvae shrimp. xbio.lifescience.ntu.edu.tw/pm/</td>
<td>(Leu et al. 2007)</td>
</tr>
<tr>
<td>Pm</td>
<td>eyestalk, hepatopancrease, haematopoietic tissue, haemocyte, lymphoid organ, and ovary from normal, heat stressed, WSSV, YHV, and V. harveyi infected shrimp</td>
<td>10100/4845</td>
<td>Large scale gene discovery, GO analysis pmodon.biotec.or.th</td>
<td>(Tassanakajon et al. 2006)</td>
</tr>
<tr>
<td>Pm</td>
<td>Hemocytes from Vibrio harveyi infected animals compared to non-infected animals</td>
<td>1062/NA</td>
<td>Immune gene discovery with a focus on antimicrobial peptides.</td>
<td>(Supungul et al. 2004)</td>
</tr>
<tr>
<td>Fc</td>
<td>cephalothorax</td>
<td>10446/NA</td>
<td>Immune gene discovery. Immune genes including lectins, serine proteases, serpins, and lysozyme are discussed.</td>
<td>(Shen et al. 2004)</td>
</tr>
<tr>
<td>Pm</td>
<td>Hemocytes from non-infected animals</td>
<td>615/NA</td>
<td>Identification of immune genes expressed in normal hemocytes with focus on penaeidins, heat-shock proteins and anti-LPS protein.</td>
<td>(Supungul et al. 2002)</td>
</tr>
<tr>
<td>Mj</td>
<td>Hemocytes from WSSV-infected animals compared to non-infected animals</td>
<td>635 normal 370 WSSV/NA</td>
<td>Immune gene discovery</td>
<td>(Rojtinnakorn et al. 2002)</td>
</tr>
<tr>
<td>Lv, Lse</td>
<td>hemocytes and hepatopancreas</td>
<td>2045/268</td>
<td>Immune gene discovery</td>
<td>(Gross et al. 2001)</td>
</tr>
<tr>
<td>Pm</td>
<td>cephalothorax, eyestalk, and pleopod tissue</td>
<td>151/NA</td>
<td>Gene discovery with a discussion of tissue specific genes.</td>
<td>(Lehnert et al. 1999)</td>
</tr>
<tr>
<td>Species</td>
<td>Tissue Description</td>
<td>Gene Count</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------</td>
<td>------------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>Lv</td>
<td>Hemocytes, gills,</td>
<td>13656/7466</td>
<td>Large-scale gene discovery with focus on genes from immune related tissues.</td>
<td>(O’Leary et al. 2006)</td>
</tr>
<tr>
<td></td>
<td>hepatopancreas,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lymphoid organ,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eyestalk and ventral nerve cord</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lv</td>
<td>Abdominal muscle</td>
<td>311/NA</td>
<td>Profile of gene expression in shrimp muscle tissue.</td>
<td>(Cesar et al. 2008)</td>
</tr>
</tbody>
</table>

Pm, Penaeus monodon; Fc, Fenneropenaeus chinensis; Mj, Marsupenaeus japonicus; Lv, Litopenaeus vannamei; Lse, Litopenaeus setiferus; NA, not reported
<table>
<thead>
<tr>
<th>Species</th>
<th>Tissues and conditions under study</th>
<th>Method</th>
<th># unique sequences identified</th>
<th>Principal findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ls</td>
<td>Hepatopancreas from shrimp experimentally infected with WSSV vs. uninfected shrimp</td>
<td>DD</td>
<td>32</td>
<td>One of the earliest indications of the challenges of assigning function to novel sequences in shrimp</td>
<td>(Astrofsky et al. 2002)</td>
</tr>
<tr>
<td>Pm</td>
<td>Hepatopancreas from shrimp surviving a WSSV outbreak vs. uninfected shrimp</td>
<td>DD</td>
<td>NA</td>
<td>PrmA, a C-type lectin with apparent antiviral activity in a heterologous (non shrimp) virus-host system</td>
<td>(Luo et al. 2003)</td>
</tr>
<tr>
<td>Mj</td>
<td>Hemocytes from shrimp stimulated by heat-killed microbes vs. shrimp not stimulated</td>
<td>SSH</td>
<td>77 with homology</td>
<td>Diverse functional groups of genes appear regulated by heat-killed microbes, including protease inhibitors</td>
<td>(He et al. 2004)</td>
</tr>
<tr>
<td>Mj</td>
<td>Hemocytes from shrimp surviving a WSSV outbreak vs. uninfected shrimp</td>
<td>SSH</td>
<td>30</td>
<td>Diverse functional categories including chaperones, lectins, and protease inhibitors enriched in animals surviving an outbreak</td>
<td>(He et al. 2005)</td>
</tr>
<tr>
<td>Mj</td>
<td>Hepatopancreas from shrimp surviving a WSSV outbreak vs. uninfected shrimp</td>
<td>SSH</td>
<td>31</td>
<td>Diverse functional categories including lectins, glucan binding proteins, and small GTPases enriched in animals surviving an outbreak</td>
<td>(Pan et al. 2005)</td>
</tr>
<tr>
<td>Ls</td>
<td>Hemocytes from animals surviving a bacterial challenge compared to non-surviving animals</td>
<td>SSH</td>
<td>184</td>
<td>Increased expression of several antimicrobials (e.g. penaeidin, lysozyme, cryptdin-like) co-relating with survival to bacterial infection</td>
<td>(de Lorgeril et al. 2005)</td>
</tr>
<tr>
<td>Pm</td>
<td>Hemocytes from Vibrio harveyi infected animals compared to non-infected animals</td>
<td>DD</td>
<td>24 with homology</td>
<td>Induction of expected immune effectors (e.g. lysozyme, ALF, transglutaminase), and of Argonaute, a gene of the RNAI pathway</td>
<td>(Somboonwiwat et al. 2006)</td>
</tr>
<tr>
<td>Lv</td>
<td>Hemocytes, gills, or hepatopancreas from shrimp induced with heat-inactivated microbes, dsRNA, or WSSV compared to mock treated animals. Same tissues from shrimp infected with WSSV at permissive and non-permissive temperatures</td>
<td>SSH</td>
<td>3,231 from both SSH and standard cDNA libraries</td>
<td>A wide range of genes with potential roles in immunity, including antimicrobials, signaling factors, transcription factors, regulators of apoptosis, were suggested to be regulated by immune stimuli</td>
<td>(Robalino et al. 2007)</td>
</tr>
<tr>
<td>Lv</td>
<td>Hepatopancreas from uninfected members of a shrimp family selected based on its reduced susceptibility to WSSV, compared to a family with high susceptibility</td>
<td>SSH</td>
<td>193, with 40 matches to known genes</td>
<td>Diverse functional categories of genes suggested to be enriched in WSSV-resistant family. These included lysozymes, cathepsins, lectins, and other potential antimicrobials</td>
<td>(Zhao et al. 2007)</td>
</tr>
<tr>
<td>Organ</td>
<td>Description</td>
<td>Method</td>
<td>Result</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>--------</td>
<td>--------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Pm</td>
<td>Hemocytes from shrimp subjected to osmotic, hypoxic, or thermal stress, compared to non-stressed animals</td>
<td>SSH</td>
<td>176, with 58 matches to known genes</td>
<td>Some known immune factors and, strikingly, many retrotransposon-related sequences are regulated during abiotic stress (de la Vega et al. 2007a)</td>
<td></td>
</tr>
<tr>
<td>Me</td>
<td>Hepatopancreas from shrimp at different stages of ovarian development</td>
<td>DD</td>
<td>15 clones with homology</td>
<td>Vitellogenin and some functionally diverse enzymes regulated during ovarian maturation (Wong et al. 2008)</td>
<td></td>
</tr>
<tr>
<td>Pm</td>
<td>Testes from broodstock vs juvenile shrimp</td>
<td>SSH</td>
<td>80 with homology</td>
<td>Discovery of a progestin receptor membrane component 1 gene (Leelatanawit et al. 2008)</td>
<td></td>
</tr>
</tbody>
</table>

Ls, *Litopenaeus stylirostris*; Pm, *Penaeus monodon*; Mj, *Marsupenaeus japonicus*; Lv, *Litopenaeus vannamei*; Me, *Metapenaeus ensis*; DD, differential display; SSH, suppression subtractive hybridization; NA, not reported
<table>
<thead>
<tr>
<th>Species</th>
<th>Estimated gene content (unique sequences)</th>
<th>Tissue and conditions studied</th>
<th>Principal findings</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ls</td>
<td>84</td>
<td>Hepatopancreas from shrimp experimentally infected with WSSV vs uninfected shrimp</td>
<td>Several potential pattern recognition proteins, such as lectins and LPS/glucan binding protein, induced by WSSV</td>
<td>(Dhar et al. 2003)</td>
</tr>
<tr>
<td>Fc</td>
<td>1,578 unique cDNAs plus 1,536 SSH clones</td>
<td>Whole cephalothorax from shrimp experimentally infected with WSSV vs uninfected shrimp, and from naturally infected shrimp vs uninfected wild animals</td>
<td>Diverse groups of genes regulated during experimental and natural WSSV infection, including chaperones and genes involved in metabolism and cell structure</td>
<td>(Wang et al. 2006)</td>
</tr>
<tr>
<td>Lv</td>
<td>2,469</td>
<td>Hepatopancreas from shrimp experimentally infected with WSSV vs uninfected shrimp</td>
<td>Induction of some known antimicrobials and repression of oxidative stress genes and of the immune transcription factor STAT</td>
<td>(Robalino et al. 2007)</td>
</tr>
<tr>
<td>Pm</td>
<td>NA</td>
<td>Hemocytes from shrimp exposed to either osmotic, hypoxic, or thermal stress</td>
<td>A complex response to abiotic stress, which included regulation of known immune factors (e.g. crustin, lysozyme, transglutaminase), as well as changes in mRNAs corresponding to retrotransposons</td>
<td>(de la Vega et al. 2007b)</td>
</tr>
<tr>
<td>Pm</td>
<td>2,028</td>
<td>Hemocytes from YHV infected shrimp vs mock infected animals</td>
<td>Known immune genes with complex patterns of temporal regulation. Cathepsin L highly induced in YHV infected hemocytes</td>
<td>(Pongsomboon et al. 2008a)</td>
</tr>
<tr>
<td>Mj</td>
<td>2,036</td>
<td>Hemocytes from peptidoglycan stimulated shrimp vs animals not stimulated</td>
<td>Known immune factors such as antimicrobial proteins respond to peptidoglycan stimulation</td>
<td>(Fagutao et al. 2008)</td>
</tr>
<tr>
<td>Fc</td>
<td>3,114 elements</td>
<td>Hepatopancreas, hemocytes, gills, and lymphoid organ from shrimp infected with WSSV vs mock infected, and from animals stimulated with heat-killed Vibrio anguillarum vs mock stimulated</td>
<td>Very functionally diverse groups of genes responsive to either immune stimulus, but also an overlapping response was observed</td>
<td>(Wang et al. 2008a)</td>
</tr>
</tbody>
</table>

Ls, Litopenaeus stylirostris; Fc, Fenneropenaeus chinensis; Lv, Litopenaeus vannamei; Pm, Penaeus monodon; Mj, Marsupenaeus japonicus; NA, not reported
<table>
<thead>
<tr>
<th>Species</th>
<th>Target gene</th>
<th>RNAi-based application</th>
<th>Phenotype</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shrimp gene physiological function studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pm</td>
<td>Argonaute (Pm-AGO)</td>
<td>dsRNA transfection into Oka cells</td>
<td>impaired RNAi ability</td>
<td>(Dechklar et al. 2008)</td>
</tr>
<tr>
<td>Lv</td>
<td>putative farnesol acid O-methyltransferase (LvFAOMeT)</td>
<td>dsRNA injection into the 5th pereiopod</td>
<td>role in molting lethal phenotype induced</td>
<td>(Hui et al. 2008)</td>
</tr>
<tr>
<td>Lac</td>
<td>crustacean hyperglycemic hormone (CHR)</td>
<td>dsRNA injection into abdominal body cavity</td>
<td>decrease in hemolymph glucose levels</td>
<td>(Lugo et al. 2006)</td>
</tr>
<tr>
<td>Lv</td>
<td>putative ion transport peptide (LvITP)</td>
<td>dsRNA injection into the 5th pereiopod</td>
<td>role in osmo-regulatory function lethal phenotype induced</td>
<td>(Tiu et al. 2007)</td>
</tr>
<tr>
<td>Lv</td>
<td>Hemocyanin</td>
<td>dsRNA/siRNA intramuscular injection</td>
<td>role in immune function against bacterial and fungal infections</td>
<td>(Robalino et al. 2005)</td>
</tr>
<tr>
<td>Me</td>
<td>Molt-inhibiting hormone (MeMIH-B)</td>
<td>dsRNA injection into the pereiopod</td>
<td>gonad-stimulatory function</td>
<td>(Tiu and Chan 2007)</td>
</tr>
<tr>
<td>Pm</td>
<td>Gonad-inhibiting hormone (Pm-GIH)</td>
<td>dsRNA injection into the pereiopod - incubation in explant culture</td>
<td>gonad-inhibitory function</td>
<td>(Treeratt rakool et al. 2008)</td>
</tr>
<tr>
<td>Pm</td>
<td>Vitellogenin receptor (VgR)</td>
<td>dsRNA intramuscular injection</td>
<td>role in the processing of vitellogenin</td>
<td>(Tiu et al. 2008)</td>
</tr>
<tr>
<td><strong>Host-pathogen interaction studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lv</td>
<td>Anti-lipopolysaccharide factor (LvALF)</td>
<td>dsRNA intramuscular injection</td>
<td>role in immune function against bacterial and fungal infections</td>
<td>(de la Vega et al. 2008)</td>
</tr>
<tr>
<td>Lv</td>
<td>Crustin (LvABP1)</td>
<td>dsRNA intramuscular injection</td>
<td>role in antibacterial response</td>
<td>(Shockey et al. 2008)</td>
</tr>
<tr>
<td>Pm</td>
<td>Prophenoloxidases (PmproPO1,2)</td>
<td>dsRNA intramuscular injection</td>
<td>role in antibacterial response</td>
<td>(Amparyup et al. 2009)</td>
</tr>
<tr>
<td>Mj</td>
<td>Rab-GTPase (PjRab)</td>
<td>siRNA intramuscular injection</td>
<td>increased viral replication</td>
<td>(Wu et al. 2008)</td>
</tr>
<tr>
<td>Pm</td>
<td>Small GTPase protein (PmRab7)</td>
<td>dsRNA intramuscular injection</td>
<td>role in the endosomal trafficking pathway</td>
<td>(Gongyarra sopone et al. 2008)</td>
</tr>
<tr>
<td>Mj</td>
<td>- Transglutaminase (TGase) - Clotting protein (CP)</td>
<td>dsRNA intramuscular injection</td>
<td>role in immune function against bacterial and fungal infections</td>
<td>(Maningas et al. 2007)</td>
</tr>
<tr>
<td>Mj</td>
<td>Caspase (PjCaspase)</td>
<td>siRNA intramuscular injection</td>
<td>role in virus-induced apoptosis</td>
<td>(Wang et al. 2008b)</td>
</tr>
<tr>
<td>Mj</td>
<td>β-integrin</td>
<td>siRNA intramuscular injection</td>
<td>cellular receptor for WSSV infection.</td>
<td>(Li et al. 2007)</td>
</tr>
<tr>
<td>Lv</td>
<td>Caspase-3 homologue (Cap-3)</td>
<td>dsRNA intramuscular injection</td>
<td>role in virus-induced apoptosis</td>
<td>(Rijiravich et al. 2008)</td>
</tr>
<tr>
<td>Pm</td>
<td>Dicer 1 (Pm Dcr1)</td>
<td>dsRNA</td>
<td>increased</td>
<td>(Su et al. 2008)</td>
</tr>
<tr>
<td>Source</td>
<td>dsRNA/SiRNA</td>
<td>Method</td>
<td>Result</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>YHV binding protein (pmYRP65)</td>
<td>dsRNA transfection into Oka cells</td>
<td>Inhibition of YHV infection (Assavala et al. 2006)</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>YHV-protease</td>
<td>dsRNA intramuscular injection</td>
<td>Inhibition of YHV replication (Tirasophon et al. 2007)</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>coding region of a protease gene (YHV)</td>
<td>dsRNA intramuscular injection</td>
<td>Inhibition of YHV multiplication in infected shrimp (Yodmuang et al. 2006)</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>Vp28 (WSSV)</td>
<td>siRNA transfection in insect cells - intramuscular injection</td>
<td>Silencing of homologous genes in a heterologous expression system - siRNAs failed to induce sequence-specific antiviral immunity (Westenberg et al. 2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Mj</strong></td>
<td>Vp28 (WSSV)</td>
<td>siRNA intramuscular injection</td>
<td>Reduction in viral DNA production of infected animals (Xu et al. 2007)</td>
<td></td>
</tr>
<tr>
<td><strong>Lv</strong></td>
<td>Vp19 (WSSV)</td>
<td>dsRNA/siRNAs intramuscular injection</td>
<td>Higher survival rates after dsRNA injection - siRNAs failed to induce antiviral immunity (Robalino et al. 2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Lv</strong></td>
<td>predicted protease gene (TSV)</td>
<td>dsRNA intramuscular injection</td>
<td>Protection from TSV infection (Robalino et al. 2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>Vp28 (WSSV)</td>
<td>bacterially expressed dsRNA oral administration</td>
<td>Protection from WSSV infection (Sarathi et al. 2008b)</td>
<td></td>
</tr>
<tr>
<td><strong>Pm</strong></td>
<td>Vp28 (WSSV)</td>
<td>bacterially expressed dsRNA intramuscular injection</td>
<td>Protection from WSSV infection (Sarathi et al. 2008a)</td>
<td></td>
</tr>
</tbody>
</table>
Pm, Penaeus monodon; Lv, Litopenaeus vannamei; Lsc, Litopenaeus schmitti; Me, Metapenaeus ensis; Mj, Marsupenaeus japonicus; Fc, Fenneropenaeus chinensis
References


