Comparative Biochemistry and Physiology, Part D 6 (2011) 150–157

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part D

journal homepage: www.elsevier.com/locate/cbpd

EST analysis on the gonad development related organs and microarray screen for differentially expressed genes in mature ovary and testis of *Scylla paramamosain*

Zhihua Zou, Ziping Zhang *[,1], Yilei Wang *, Kunhuang Han, Mingjun Fu, Peng Lin, Jia Xiwei

The Key Laboratory of Science and Technology for Aquaculture and Food Safety, Fisheries College, Jimei University, Xiamen 361021, China

ARTICLE INFO

Article history:
Received 26 September 2010
Received in revised form 18 December 2010
Accepted 27 December 2010
Available online 1 January 2011

Keywords:
Normalized cDNA library
Microarray
Testis
Ovary
Crab
Gene expression

ABSTRACT

A total of 5160 high quality ESTs (expressed sequence tags) averaging 357 bp were collected from normalized cDNA libraries created from testis, ovary and mixed organs of mud crab *Scylla paramamosain*. Clustering and assembly of these ESTs resulted in a total of 3837 unique sequences with 576 overlapping contigs and 3261 singletons. Comparisons with the GenBank non-redundant (Nr) protein database (BLASTx, *e*-values $<10^{-5}$) revealed putative functions or matched homologs from other organisms for 847 (22%) of the ESTs. Several gonad development related genes such as *cathepsin C, thioredoxin peroxidase, vitellogenin receptor precursor, 50S ribosomal protein L24* and *ubiquitin-conjugating enzyme E2 isoform 2* were identified from this EST project and demonstrated as gonad differential expression genes by rqRT-PCR. Sixty five different types of SSRs (simple sequence repeats) were identified from the total 411 EST-SSR motifs. A home-made cDNA microarray containing 5664 spots was developed and the hybridization results indicated that 39 unique transcripts were differentially expressed in testis and ovaries ($P<0.05$). The expression levels of eleven unique transcripts examined by rqRT-PCR were matched with microarray fairly. These results will provide a useful resource for functional genomic studies on the biology of reproduction of mud crab.

1. Introduction

Scylla paramamosain is a species of the family Portunidae that inhabits muddy bottoms in brackish water along the shoreline, mangrove areas, and river mouths (Macnae, 1969). Due to the high market demand of mud crab (Overton et al., 1997) and the advantage of large size, fast growth and high economic value, artificial culture of this species has developed rapidly over the last past 30 years (Wang et al., 2005). However, a shortage of juvenile crabs creates a bottleneck in overall production (Sugama and Hutapea, 1999). In order to improve the efficiency of the culture industry, the structure and function of many sex-related organs such as mandibular organ (Huang et al., 2003), X-organ, Y-organ (Ye et al., 2002) and androgenic gland (Ye et al., 2003), had been reported in detail and applied to the culture practice.

However, molecular mechanism of crustacean reproduction remains largely unknown. Although recent progress has brought new insights into this area and a few genes, such as Dmrt (Zhang and Qiu, 2010), Cdc2 kinase (Qiu and Liu, 2009), proliferating cell nuclear antigen (Zhang et al., 2010), 50S ribosomal protein L24 (Zhang et al., 2007), ubiquitin-conjugating enzyme E2r (Shen et al., 2009), leptin receptor (Jiang et al., 2010), and EJO1 (Ma et al., 2010), have now been cloned from gonad of crustacean. Genome database of crustacean is still limited comparing with model invertebrates such as sea urchin *Strongylocentrotus purpuratus* and nematode *Caenorhabditis elegans*. Thus, it is important to identify the genes that are differentially expressed in both testis and ovary at the transcriptome level to shed light on the mechanism of sex determination, sex differentiation, and gonad maturation, and to develop novel methods of reproduction control in the species.

Presently, only 411 nucleotide sequences of *S. paramamosain* are available in the NCBI Genbank (http://www.ncbi.nlm.nih.gov, as of April 2010). A previous study in our laboratory using an amplified fragment length polymorphism (AFLP) technique with 52 primer combinations, has successfully identified molecular markers related to the sex locus of *S. paramamosain* (Wang et al., 2004). In addition, an important EST project of the porcelain crab, *Petrolisthes cinctipes*, obtained 4024 unique transcripts from 7 libraries derived from different organs has been reported (Stillman et al., 2006). Similarly, in Decapoda, ESTs from several other species such as *Eriocheir sinensis* (Jiang et al., 2009), *Callinectes sapidus* (Coblentz et al., 2006), *Uca pugilator, Carcinus maenas, Homarus americanus* (http://www.ncbi.nlm.nih.gov/sites/entrez), *Penaeus monodon* (Tassanakajon et al., 2006), *Marsupenaeus japonicus* (Yamano and Unuma, 2006), *Litopenaeus vannamei* (O'Leary et al., 2006), *L. setiferus* (Gross et al., 2001), and *L. stylirostris* (de Lorgeril et al., 2005) have already been published. However, few of these studies focused on sex or gonad

* Corresponding authors. Wang is to be contacted at Tel.: +86 592 6182723; fax: +86 592 6181420. Zhang, Tel.: +1 512 245 0358; fax: +1 512 245 1922.
E-mail addresses: zhangziping@hotmail.com (Z. Zhang), ylwang@jmu.edu.cn (Y. Wang).
[1] Current address: Department of Chemistry and Biochemistry, Texas State University, San Marcos, TX, 78666, USA.

1744-117X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbd.2010.12.003

development related organs. In this study, 3 normalized cDNA libraries were constructed from testis, ovary and pooled-tissue. A large number of ESTs were sequenced from cDNA clones of these cDNA libraries.

cDNA microarray provides a platform to investigate the expression levels of thousands of genes under various experimental conditions. Recently, by using zebrafish Affymetrix microarrays, Small et al. (2009) identified a large number of genes (5899) demonstrating differences in transcript abundance between male and female somatic and gonadal tissues of *Danio rerio*. Karoonuthaisiri et al. (2009) also used microarray to simultaneously examine expression patterns of 4992 transcripts in ovaries and testes of *P. monodon*, and several transcripts were found to be differentially expressed during *P. monodon* ovarian development. These studies have demonstrated the great potential of microarray analysis in identifying differentially expressed genes.

As a first step in understanding the molecular mechanism of crustacean reproduction at the transcriptome level, this study documents our development of a large-scale EST database and first cDNA microarray focused on sex or gonad development related organs of mud crab. This project shall provide a useful molecular resource and genomic tool for further investigation of the mechanism of sex determination, sex differentiation, and gonad development for crustacean.

2. Materials and methods

2.1. Animal and organ preparation

Fifteen male and fifteen female crabs (*S. paramamosain*, male 68–82 mm, and female 69–88 mm in body length) were purchased from the local market in Xiamen, Fujian, China (24°.4′ N, 118°.1′ E) for constructing all of the three cDNA libraries. According to the gonad-somatic index [GSI = (gonad weight/body weight) × 100], samples were collected from 6 different organs (testis, ovary, eyestalk, ejaculatory duct, vas deferens, and androgenic gland) at different gonad developing stages (testis: GSI ≤ 0.1, 0.1 < GSI < 0.3, and GSI ≥ 0.3; ovary: GSI ≤ 1, 3 < GSI < 8, and GSI ≥ 10). These organs were dissected, snap frozen in liquid nitrogen, and store at −80 °C until use. For ovary or testis cDNA libraries, only ovary or testes RNAs from all of the crabs were mixed equally. For pooled organ cDNA library, 6 different organs' (testis, ovary, eyestalk, ejaculatory duct, vas deferens, and androgenic gland) RNA from all of the crabs were mixed equally.

2.2. RNA extraction and normalized cDNA library construction

Total RNAs were extracted as described in our previous studies (Zhang et al., 2007) and kept in absolute ethanol at −80 °C. mRNA was further isolated using the Oligotex Direct mRNA Mini Kit (Qiagen Inc., CA, USA) according to the manufacturer's handbook. The concentration of RNAs was determined by measuring the absorbance at 260 nm in a spectrophotometer. cDNA synthesis was performed using the Creator™ SMART™ cDNA Library Construction Kit (Clontech Laboratories, Inc., CA, USA), which generates first-strand cDNA flanked with asymmetric adapter sequences comprising Sfi IA and Sfi IB restriction sites. Then the SMART-cDNA group was normalized using the TRIMMER cDNA Normalization Kit (Innovative Biotechnology Company, Moscow, Russia) which employs a combination of specific suppression PCR-based approach and Duplex-Specific Nuclease (DSN) enzyme to remove double strand cDNA from complex mixtures of nucleic acids. Afterward, the normalized cDNAs were digested with sfi I enzyme and size fractionated using a gel filtration column, and cDNA fractions longer than 500 bp were ligated into the pDNR-LIB vector and subsequently transformed into DH5α electro-competent *E. coli* cells. Three cDNA libraries, one from the pooled organ and the other two from the testis and ovary respectively, were constructed.

2.3. Inserted cDNA sequencing

PCR assay was employed using M13 vector specific primers to determine the single positive clones and to assess the average insert size. Target clones suspended in water were inoculated with 2 mL LB/Chloromycetin broth at 37 °C for 12–16 h.

Plasmids were isolated in 96-well format on a plate centrifuge using the MultiScreen PCR 96 Filter Plate (Millipore Corporate, MA, USA). As well known that a long 3′ UTR would hamper gene identification whereas the 5′ ESTs were more useful for most of them contain longer coding sequences. Inserts were sequenced directly from the plasmids in 96-well format from the 5′ end. Sequencing PCR reactions contained 4 μL of the purified plasmid, 4 μL DYEnamic ET reagent premix (Amersham Biosciences Corp, NJ, USA), and 2 μL of M13 forward vector primer (1 μM; 5′-TAC GAT TTA GGT GAC ACT ATA G-3′). Reaction conditions were performed according to the user manual. Sequence reaction products were purified by ethanol precipitation and analyzed on a MegaBACE500 automated DNA sequencer (GE Healthcare, CA, USA).

2.4. EST assembly, clustering and bioinformatics analysis

Phred (Ewing and Green, 1998; Ewing et al., 1998) and cross_match (University of Washington, Seattle, WA, USA) were used to perform base calling, trim contaminant sequences at the 5′, remove chimeric ESTs, mask repeat sequences and low complexity regions, and discard EST sequences shorter than 100 bp. Chimeric EST sequences were removed by the CAP3 program (Huang and Madan, 1999). The cleaned ESTs were assembled and clustered into unique transcripts including contigs and singletons with the CAP3 program. Closely related sequences were first grouped into clusters. The ESTs within each cluster were then compared, and where overlapping regions of identical or very nearly identical sequence were found, the sequences were arranged as contigs. A contig represents a putative transcript randomly chosen for sequencing more than once. Singlets, representing transcripts sequenced only once, are ESTs that do not fit into clusters, or if in clusters, do not form part of contigs. Non-redundant nucleotide sequences of ESTs were compared with non-redundant (nr) protein and nucleotide databases deposited in NCBI network database (as of April 2010), using both BLASTX and BLASTN for annotation, by setting the significant cut-off value of the expectation value (E-value) below 1.0×10^{-5}. We also submitted the unique transcripts to the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) and Clusters of Orthologous Groups of proteins (COGs) online server to acquire pathway and protein classification information (Tatusov et al., 2003; Tatusov et al., 1997).

2.5. Identification of EST-SSRs

All unique sequences were sent to the following web site (http://www.gramene.org/db/searches/ssrtool) for microsatellite searching. The minimum number of nucleotide repeats was set at eight for dinucleotides, five for trinucleotides, four for tetranucleotides, and three for both penta- and hexanucleotides.

2.6. Microarray printing, hybridization and obtaining data sets

Insert DNA of 5664 cDNA clones which represent 3837 unique genes of *S. paramamosain* derived from normalized cDNA libraries constructed in this project, as well as 2 positive control sequences (fragments of 18S and β-actin of *S. paramamosain*) and 1 negative control sequence of intergenic rigion, were PCR amplified and purified