



Development of molecular resources for the Chinese horseshoe crab *Tachypleus tridentatus*

Hongjun Li¹, Ying Xia², Yiyong Rao⁴, Shuxi Liu¹, Jingjing Zhang^{1,5}, Lizhe Cai^{3,4,*}

¹Key Laboratory for Ecological Environment in Coastal Areas (SOA), National Marine Environmental Monitoring Center, Dalian 116023, PR China

²Dalian Environmental Monitoring Center, Dalian 116023, PR China

³Key Laboratory of the Ministry of Education for Coastal and Wetland Ecosystems, Xiamen University, Xiamen 361102, PR China

⁴College of the Environment and Ecology, Xiamen University, Xiamen 361002, PR China

⁵Dalian Ocean University, Dalian 116023, PR China

ABSTRACT: The Chinese horseshoe crab *Tachypleus tridentatus* is an endangered marine benthic species of great biomedical value and scientific significance. In this study, *T. tridentatus* hemocyte transcriptome was sequenced to rapidly develop molecular resources using the Illumina paired-end sequencing platform. Deep sequencing generated a total of 38 402 764 reads that consisted of 7.7 Gbp raw data, which were assembled into 38 759 unigenes with an average length of 668.6 bp. A total of 608 microsatellite loci and 71 959 high-confidence single nucleotide polymorphisms were discovered. All the assembled unigenes were annotated by running BLASTx similarity searches on Nr, Swiss-Prot, COG, and KEGG databases; of which 17 446 (45.0%) unigenes showed significant matches (E -value $< 1e^{-5}$) to known sequences in public databases. As determined through gene ontology annotation and KEGG pathway mapping, the functional annotation of the unigenes signified diverse biological functions and processes. Transcripts potentially involved in immune defense were identified among these genes. These transcripts included pattern recognition receptors, stress response genes, complement components, and immune effectors. This study provides important insights into the transcript sequences of *T. tridentatus*, which may facilitate studies on the conservation genetics and development of biomedical applications for *T. tridentatus*.

KEY WORDS: Chinese horseshoe crab · *Tachypleus tridentatus* · Endangered species · Transcriptome · RNA-seq · Microsatellite · SNP · Conservation genomics · Biomedicine

INTRODUCTION

Horseshoe crabs are perhaps the most striking 'living fossils' in the ocean, with the oldest fossils dating back to the Upper Ordovician 445 million yr ago (Rudkin et al. 2008). Four extant horseshoe crab species exist worldwide, including 3 Asian species: *Tachypleus tridentatus*, *T. gigas*, and *Carcinoscorpius rotundicauda*; and 1 American species: *Limulus polyphemus* (Botton 2009). Two horseshoe crab species (*T. tridentatus* and *C. rotundicauda*) are generally distributed in Chinese coastal waters (Liao &

Liu 2006). In recent years, the combined effects of overfishing, habitat destruction, and environmental deterioration have led to a significant loss of horseshoe crabs in China (Liao & Li 2001). Although horseshoe crabs are considered a 'Grade II Protected Animal' in the 'List of State Key Protected Wildlife' (Yang et al. 2009), they remain at increasing risk of extinction (Weng et al. 2012). Conservation and recovery of horseshoe crab populations are now prioritized throughout China, particularly in the Fujian and Guangxi provinces (Huang et al. 2003, Li & Hu 2011).

*Corresponding author: cailizhe@xmu.edu.cn

Similar to other invertebrate animals, horseshoe crabs rely completely on innate immunity, which is a sophisticated defense system that responds to common antigens on the surface of potential pathogens. The hemolymph of horseshoe crab contains soluble defense molecules, including clotting factors, proteinase inhibitors, lectins, antimicrobial peptides, and other humoral factors (Iwanaga 2002). Hence, horseshoe crabs have potential benefits for use in public health. Biomedical companies bleed horseshoe crabs to extract *Tachypleus* Amebocyte Lysate (TAL) for pathogenic endotoxin detection in injectable and implantable medical devices for humans. The TAL test is now a standard screening for endotoxin contamination in medical equipment. Overfishing for biomedical production of TAL is one of the main reasons behind the declining numbers of *T. tridentatus* in China.

In most cases, genetic events such as founder effects, inbreeding, and genetic drift threaten endangered species. Traditional conservation genetics address these issues with limited surveys on neutral marker variation. Through the numerous genome projects that have emerged, the genomes of hundreds of species have been sequenced. Clearly, conservation genetics can be transformed via genomic approaches (Allendorf et al. 2010). High-throughput genotyping technologies enable genome-wide surveys on endangered species (Stinchcombe & Hoekstra 2008). Simply increasing the number of neutral loci for screening increases the power and accuracy of estimating various important population parameters (e.g. kin relationship and inbreeding coefficient) (Allendorf et al. 2010). Additionally, an expanded genomic resource could provide a new list of target genes. For example, the expression level of deleterious genes that reduce biological fitness in a given population could be quantified. The study of evolutionary adaptation involves the detection and response of genetics to environmental changes (Luikart et al. 2003).

Development and utilization of molecular tools have been popular in studies of phylogeny reconstruction (Xia 2000, Kamaruzzaman et al. 2011), population genetics (Xu et al. 2011, Rozihan et al. 2013, Weng et al. 2013), and gene expression patterns in horseshoe crabs (Ding et al. 2005). However, molecular resources for *T. tridentatus* are still limited, and only a few microsatellite loci, mitochondrial sequences, and expressed sequence tags are available in the database of the National Center for Biotechnology Information (NCBI; 21 January 2015). Studying the transcriptome of species by using next-generation sequencing technologies can efficiently generate

valuable genomic information at a reasonable cost (Metzker 2010). This method is particularly important for endangered species (Allendorf et al. 2010, Avise 2010) with scarce genomic resources. In the present study, RNA sequencing (RNA-seq, Illumina HiSeq™2000) technology is utilized to determine a significant portion of the *T. tridentatus* hemocyte transcriptome, from which thousands of potential molecular marker loci and immune-related genes can be found. Our results significantly deepen the pool of molecular resources available for this taxon and can be used as a guide for similar studies on endangered species.

MATERIALS AND METHODS

Crab collection and maintenance

Three Chinese horseshoe crabs (*Tachypleus tridentatus*) were collected from Oucuo (24°31'N, 118°13'E) near Xiamen, Fujian Province, in August 2013. The crabs were transferred to Xiamen University and were cultured in aerated natural seawater (30‰) at 28 to 30°C. Half of the tank water was renewed every day, and air stones were placed inside the tank to provide aeration. The crabs were fed with *Ruditapes philippinarum* clam meat chopped into small pieces. After 3 d of acclimation, the hemolymphs of *T. tridentatus* were collected using a syringe and then centrifuged at 2000 rpm for 10 min to harvest hemocytes for RNA extraction. The RNA samples were frozen immediately in liquid nitrogen and then stored at –80°C until utilization.

RNA extraction and sequencing

The total RNA was extracted from the *T. tridentatus* hemocytes by using an RNeasy pure tissue kit (Qiagen). The total RNA concentration and integrity were determined using an Agilent 2100 Bioanalyzer and a NanoDrop spectrophotometer (Thermo Scientific), respectively. RNA-seq library construction and sequencing (Illumina) were performed as described by Yang et al. (2014). After the mRNA was enriched by beads with Oligo(dT) primers, the mRNA was dissected into short fragments. Taking these short fragments as templates, a random hexamer primer was used to synthesize the first cDNA strand, and the second strand was then amplified. The double-stranded cDNA was purified using a polymerase chain reaction (PCR) extraction kit (Qiagen), and the short frag-

ments were connected using sequencing adaptors. Finally, a sequencing library was constructed via PCR amplification and sequenced by using the Illumina HiSeq2000 platform.

De novo transcriptome assembly

De novo assembly of the short reads from RNA sequencing was performed using Trinity (Grabherr et al. 2011). Trinity combined reads with a specific overlap length to form longer fragments called 'contigs.' The contigs were subjected to further sequence clustering to form longer sequences. Such sequences were termed 'unigenes.' Prior to assembly, raw reads were trimmed by stripping the adaptor sequences and ambiguous nucleotides by using SeqPrep (<https://github.com/jstjohn/SeqPrep>) and Sickle (<https://github.com/najoshi/sickle>), respectively. Those reads with quality scores of <20 and lengths <20 bp were removed. The resulting cleaned and filtered high-quality sequences were used in the subsequent assembly with the default settings.

Identification of simple sequence repeats and single nucleotide polymorphisms

Simple sequence repeat (SSR) mining was performed using the 'msatcommander' program (Faircloth 2008). The search criteria were set based on the number of repeat motifs: 7 for dinucleotides; 5 for trinucleotides, tetranucleotides, and pentanucleotides; and 4 for hexanucleotides. The single nucleotide polymorphism (SNP) discovery process was implemented using SAMtools (<http://samtools.sourceforge.net/>). Trinity-assembled transcripts were used as reference sequences. SNPs were determined as superimposed nucleotide peaks, where 2 or more reads contained polymorphisms at the variant allele with the default parameter. False positive SNPs were avoided because of sequencing errors (which may be monomorphic loci), and only those variants with a minimum variant count of 2 high quality (HQ) bases and a minimum site depth of 8 HQ bases were considered as putative SNPs.

Gene annotation, ontology, and pathway analysis

The resulting transcripts from Trinity assembly were used as queries for open reading frame (ORF) prediction. A set of utilities included in the Trinity

software was employed to extract the likely coding regions from the Trinity transcripts. Gene annotation was then performed on protein sequences using the predicted ORF. Sequence homology searches were performed using the BLASTP program against sequences in the NCBI non-redundant (Nr) protein, STRING (<http://string-db.org/>), and KEGG GENES databases (www.genome.jp/kegg/genes.html). The cutoff expected value (*E*-value) was set as $1e^{-5}$, and only the top hit result against known sequences was assigned as the annotation. Gene ontology (GO) analysis was processed using Blast2GO v.2.5.0 (Conesa et al. 2005), which is an automated tool to assign gene ontology terms. The final annotation file was generated after gene-identification mapping, GO-term assignment, annotation augmentation, and generic GO-Slim processes. All the annotated contigs were categorized with respect to a Level 2 biological process, cellular component, and molecular function. These contigs were used to determine the GO and COG (clusters of orthologous groups) terms and for further KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis.

RESULTS AND DISCUSSION

Sequencing and assembly

Illumina HiSeq™ 2000 platform sequencing was performed on the hemocyte cDNA library of multiple *Tachypleus tridentatus* individuals to obtain as many gene transcripts as possible and to deepen our understanding of *T. tridentatus* transcriptome biology. To the best of our knowledge, this study is the first report that uses a high-throughput sequencing approach to capture transcriptome sequences of *T. tridentatus*. High-throughput sequencing generated 38 402 764 transcriptomic reads that comprised 7 756 867 339 bp (7.7 Gbp) raw data, with 41.07 % GC percentage. The Q30 percentage was 83.61 %, which indicates the reliability of the sequencing data quality. The high-quality data were aligned and de novo assembled into 38 759 unigenes by using Trinity, with an average length of 668.6 bp and an N50 length of 1020 bp. Among these unigenes, 27 316 (70.5 %) were longer than 600 bp, and 6432 (16.6 %) were longer than 1000 bp (Fig. 1). The raw reads were submitted to the Sequence Read Archive at NCBI under the Accession Number SRP041203.

Traditional genetic analysis relies on limited genomic information. In recent years, mass-scale nucleic acid sequencing has become customary in next-

generation molecular technologies, facilitating studies of conservation genetics to enter the genomics era (Avisé 2010). We obtained 7.7 Gbp data in a single lane run by using the HiSeq™ 2000 platform. The average size of the unigenes in our present study was 668.6 bp, which is similar to those obtained by Sanger sequencing. With the emergence of next-generation sequencing technology, conservation genetics are transformed via genomic approaches (Kohn et al. 2006, Ekblom & Galindo 2011).

Annotation of unigenes

The 'getorf' function of the EMBOSS software is used to identify the ORFs of the assembled sequences. Among the 38 759 assembled unigene sequences, 29 262 (75.5%) had ORFs longer than 100 bp, with an average length of 366.5 bp (Fig. 2). To provide putative annotations for the assembled unigenes, BLAST similarity evaluation was employed to search against COG, GO, KEGG, Swiss-Prot, and Nr databases (Table 1). A total of 17 446 (45.0%) unigenes was matched with an E -value of $<1e^{-5}$ (Supplement 1 at www.int-res.com/articles/suppl/b024p017_supp.xls). The sequences without annotation information in this study represent a considerable percentage. Poor annotation efficiency is most likely caused by insufficient sequences being available in public databases from phylogenetically closely related species, especially species from the class Merostomata (subphylum Chelicerata). This assessment is consistent with the fact that most matches involve other invertebrate species, including tick (*Ixodes scapularis*), mite (*Metaseiulus occidentalis*), amphioxus

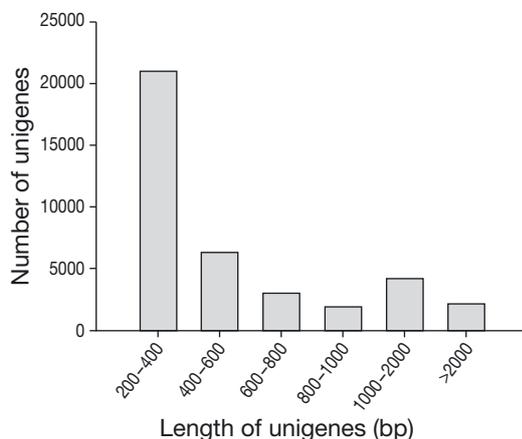


Fig. 1. Assembly length statistic of *Tachypleus tridentatus* transcriptome. Of the unigenes analyzed, 27316 (70.5%) were longer than 600 bp; 6432 (16.6%) of this subset were longer than 1000 bp

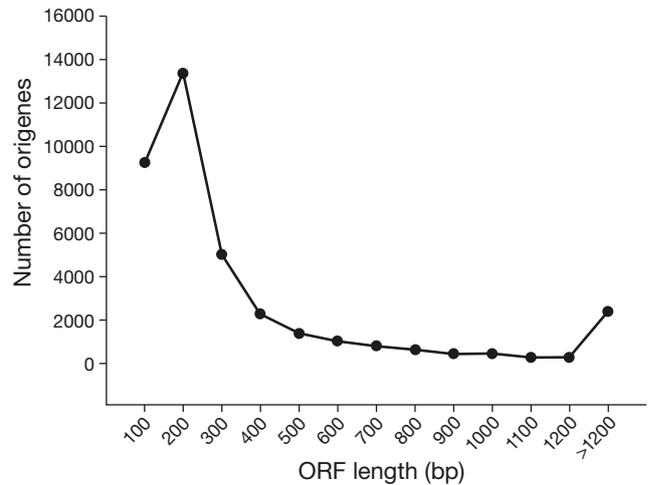


Fig. 2. Open reading frame (ORF) lengths of *Tachypleus tridentatus* transcriptome

(*Branchiostoma floridae*), oyster (*Crassostrea gigas*), crustacean (*Daphnia pulex*), hemichordate (*Saccoglossus kowalevskii*), polychaete (*Capitella teleta*), red flour beetle (*Tribolium castaneum*), sea hare (*Aplysia californica*), head lice (*Pediculus humanus*), and sea urchin (*Strongylocentrotus purpuratus*) (Fig. 3). Only 0.95% (166/17 446) of the BLAST hits matched species of the class Merostomata: *Tachypleus tridentatus* (86), *Carcinoscorpius rotundicauda* (40), and *Limulus polyphemus* (40). Moreover, given that the significance of the BLAST comparison depends partly on the query sequence length, short sequences would be hard to match with known genes (Hou et al. 2011). More than one-third of the assembled unigenes in the present study were not long (35.2% < 300 bp) and may have been too short to allow statistically significant matches. For sequences longer than 300 bp, the annotation rate was 50.3% (12 630/25 129), whereas the annotation rate for sequences longer than 1000 bp increased to 81.3% (5227/6432) (Table 2).

Table 1. Unigene annotation statistics of the *Tachypleus tridentatus* transcriptome. COG: clusters of orthologous groups; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; Swiss-Prot: Swiss protein; Nr: non-redundant

	Number	>300 bp	>1000 bp
Annotated databases			
COG	4979	2106	2166
GO	10067	4683	3023
KEGG	7426	3363	2498
Swiss-Prot	11192	4997	4302
Nr	17363	8215	5492
Sequences with annotation	17446	8258	5497
All sequences	38759	25129	6432

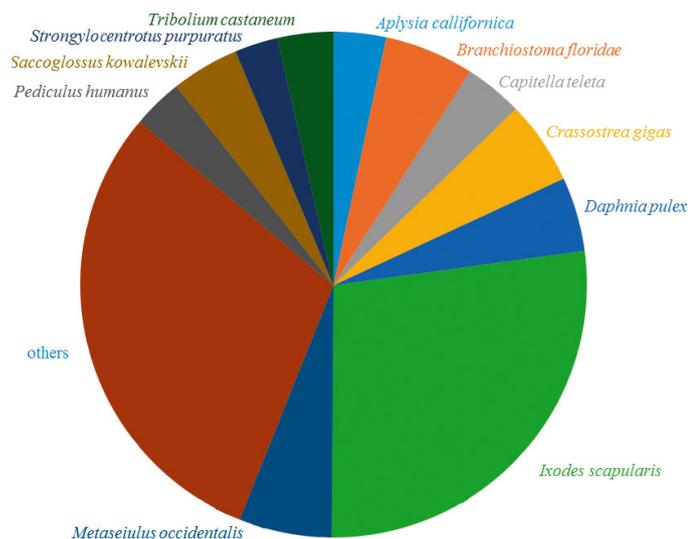


Fig. 3. Species matched to the annotated sequences of *Tachypleus tridentatus* by BLASTx. The top-match species included tick (*Ixodes scapularis*), mite (*Metaseiulus occidentalis*), amphioxus (*Branchiostoma floridae*), oyster (*Crassostrea gigas*), arthropod (*Daphnia pulex*), hemichordate (*Saccoglossus kowalevskii*), polychaete annelid (*Capitella teleta*), red flour beetle (*Tribolium castaneum*), sea hare (*Aplysia californica*), head lice (*Pediculus humanus*), and sea urchin (*Strongylocentrotus purpuratus*)

Molecular marker discovery

A total of 608 SSRs were identified from 38 759 unigene sequences examined. Dinucleotides were the most frequent type (506, 83.2%), followed by trinucleotides (87, 14.3%), and tetranucleotides (15, 2.5%). Based on the SSR motif distribution, the AT motif was the most frequent, which corresponds to 34.4% of the dinucleotide motifs. Potential SNPs were identified using SAMtool.

A total of 71 959 high-quality SNPs were identified, including 20 858 transitions and 46 446 transversions (Fig. 4). Among the 71 959 predicted SNPs, 43 671 (60.9%) were from contigs covered by 10 or more reads, thereby suggesting that a majority of the putative SNPs were covered at a sufficient sequencing depth and more likely represent actual SNPs. These microsatellite and SNP loci are valuable candidates for marker development in future conservation genetic studies on this species.

Applying molecular markers in a *T. tridentatus* conservation and recovery program is, as expected, a fruitful research area in relation to this endangered species. However, few genetic markers are currently available for this species (Li et al. 2009, Nishida & Koike 2010). High-throughput sequencing data

Table 2. KEGG biochemical mapping for *Tachypleus tridentatus*

KEGG category & subpathways	Unigene number	Enzyme number
Metabolism	2287	1143
Metabolism of cofactors and vitamins	143	89
Amino acid metabolism	409	227
Nucleotide metabolism	283	134
Metabolism of terpenoids and polyketides	24	14
Glycan biosynthesis and metabolism	197	94
Lipid metabolism	288	133
Xenobiotics biodegradation and metabolism	78	28
Energy metabolism	189	108
Carbohydrate metabolism	550	258
Metabolism of other amino acids	126	58
Genetic information processing	1757	934
Replication and repair	184	123
Translation	726	334
Transcription	262	161
Folding, sorting and degradation	585	316
Environmental information processing	910	299
Signal transduction	738	238
Signaling molecules and interaction	129	44
Membrane transport	43	17
Cellular processes	759	341
Cell growth and death	66	50
Cell motility	18	14
Transport and catabolism	602	232
Cell communication	73	45
Organismal systems	592	318
Nervous system	23	19
Excretory system	11	13
Sensory system	20	14
Digestive system	46	38
Circulatory system	48	19
Endocrine system	199	93
Immune system	146	91
Development	96	27
Environmental adaptation	3	4
Human diseases	269	207
Endocrine and metabolic diseases	7	9
Immune diseases	4	7
Infectious diseases: bacterial	8	8
Cancers: specific types	49	58
Infectious diseases: parasitic	37	29
Neurodegenerative diseases	119	51
Cardiovascular diseases	16	18
Cancers: overview	22	19
Infectious diseases: viral	7	8
Total	6574	3242

obtained through the Illumina platform provide a valuable resource to mine large numbers of gene-associated markers (Wei et al. 2011, Castoe et al. 2012).

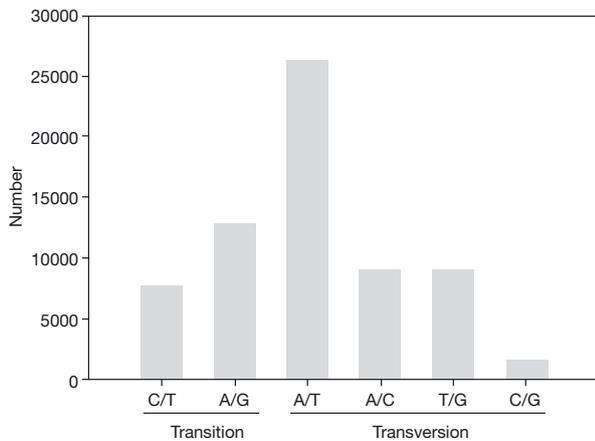


Fig. 4. Classification of single nucleotide polymorphisms (SNPs) identified from *Tachypleus tridentatus* transcriptome sequences. A total of 71959 high-quality SNPs were detected

GO, KEGG and COG classifications

After the initial BLAST searches, BLAST2GO analysis was conducted to categorize the known genes into Level 2 functional groups (Fig. 5). A total of 57 GO terms were assigned to 9967 unigenes, including 17 (29.8%) cellular component terms, 17 (29.8%) molecular function terms, and 23 (40.4%) biological process terms. For the cellular component, the genes

involved in the cell part and the cell were highly represented. For molecular function, binding was the most represented GO term, followed by catalytic activity. The biological process term mainly included cellular and metabolic processes.

In addition to GO analysis for the annotated genes, KEGG pathway analysis was also performed, which is an alternative approach to categorize gene functions, placing emphasis on biochemical pathways. Enzyme commission numbers were assigned to 6574 unigenes involved in 215 different pathways. A summary of the sequences involved in these pathways is shown in Table 2. Among these 6574 sequences with KEGG annotation, 34.8% were classified into metabolism, where most were involved in carbohydrate, amino acid, and lipid metabolisms. The sequences classified as genetic information processing accounted for 26.7% of the KEGG annotated sequences. The well-represented metabolic pathways are translation, folding, sorting and degradation, and transcription. Environmental information processing was represented by 13.8% of the KEGG annotated sequences, including signal transduction, signaling molecules and interaction, and membrane transport. In addition, 24.6% of the sequences were classified into cellular processes, organismal systems, and human diseases.

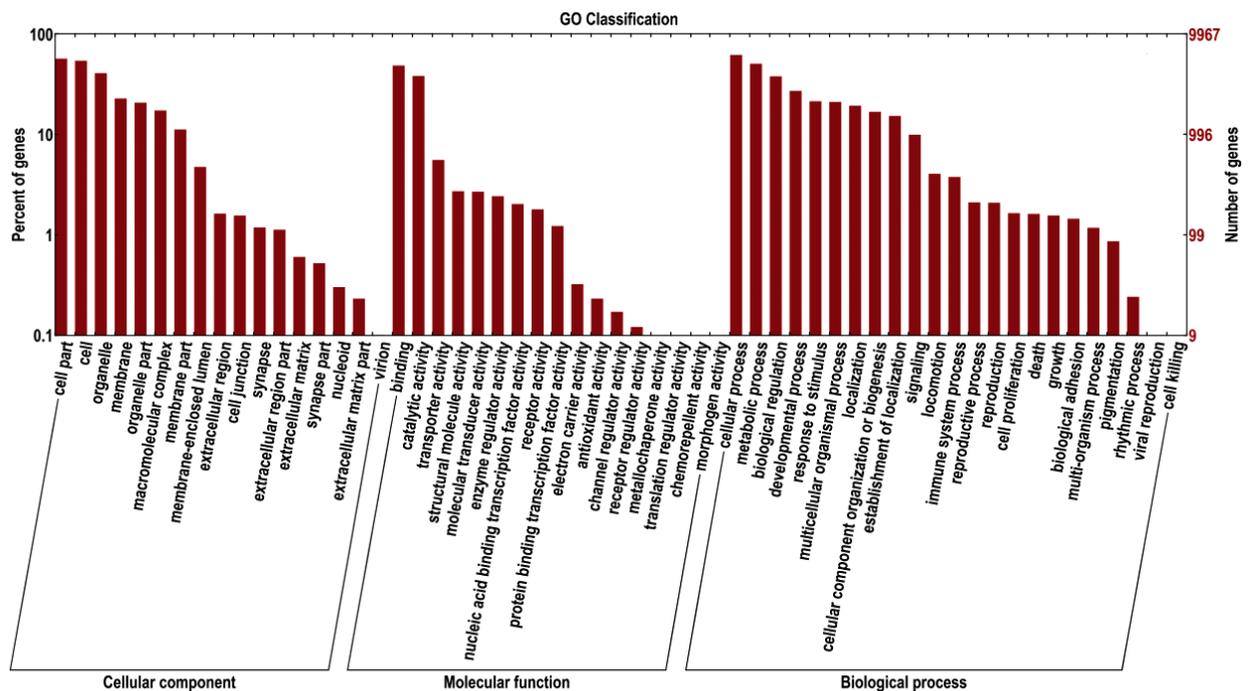


Fig. 5. Gene ontology (GO) classification of assembled unigenes of the *Tachypleus tridentatus* transcriptome on cellular component, molecular function, and biological process levels

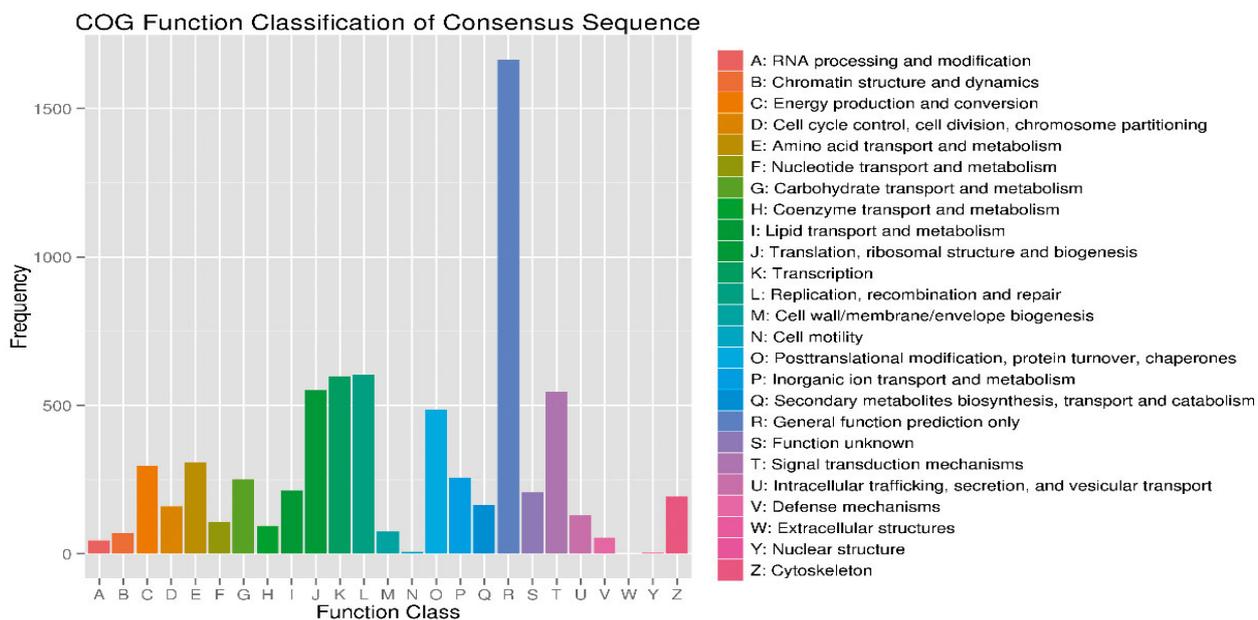


Fig. 6. Classification according to clusters of orthologous groups (COG)

Similarly, COG-ore-liner putative proteins were functionally classified into at least 25 molecular families (Fig. 6), such as cellular structure, biochemistry metabolism, molecular processing, signal transduction, gene expression, and immune defense. All these families corresponded to the categories observed in the GO analysis.

Identification of functional genes involved in immunity

Hemocytes circulating in horseshoe crab hemolymph are highly sensitive to Gram-negative bacterial endotoxins (Nakamura et al. 1988). Exposure of hemocytes to lipopolysaccharide activates the intracellular clotting system, which participates in both hemostasis and defense against invading microorganisms. The circulating hemolymph in horseshoe crab contains various biologically active substances, including lectins, clotting factors, proteinase inhibitors, and microbial peptides, which all contribute to a self-defense system against microorganisms (Shigenaga et al. 1990).

Horseshoe crabs are living fossils, spanning over 450 million years of evolution. The identification of a number of stress and immune-related transcripts (Supplement 2 at www.int-res.com/articles/suppl/b024p017_supp.xls) is of great interest to researchers, since horseshoe crabs have been used as a model system for studies in behavioral ecology, physiology,

development, and defense mechanisms. Both GO and KEGG analyses identified transcripts potentially involved in responses to environmental pressure and stimuli. The GO annotation obtained 37 transcripts related to stimulus responses (GO: 0050896), and KEGG analysis resulted in 106 transcripts classified as immune system responses. These transcripts included pattern recognition receptors, stress response genes, complement components, and immune effectors. Among these genes, stress response genes such as heat shock proteins (HSPs) were highly expressed. HSPs are a family of proteins that are produced by cells in response to stressful conditions (Kregel 2002). Several types of Toll receptors were identified, which are an ancient family of pattern recognition receptors playing a key role in innate immunity (Kopp & Medzhitov 1999).

Acknowledgements. This research was supported by grants from public science and technology research funds for marine projects (201305030, 201505012) and Young Marine Science Foundation of the State Oceanic Administration (2013102).

LITERATURE CITED

- Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nat Rev Genet* 11:697–709
- Awise JC (2010) Perspective: conservation genetics enters the genomics era. *Conserv Genet* 11:665–669
- Botton M (2009) The ecological importance of horseshoe

- crabs in estuarine and coastal communities: a review and speculative summary. In: *Biology and conservation of horseshoe crabs*. Springer, New York, NY, p 45–63
- Castoe TA, Poole AW, de Koning AJ, Jones KL and others (2012) Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* 7:e30953
 - Conesa A, Götz S, García-Gómez JM, Terol J and others (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676
 - Ding JL, Tan KC, Thangamani S, Kusuma N and others (2005) Spatial and temporal coordination of expression of immune response genes during *Pseudomonas* infection of horseshoe crab, *Carcinoscorpius rotundicauda*. *Genes Immun* 6:557–574
 - Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1–15
 - Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol Ecol Resour* 8:92–94
 - Grabherr MG, Haas BJ, Yassour M, Levin JZ and others (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29: 644–652
 - Hou R, Bao ZM, Wang S, Su HL and others (2011) Transcriptome sequencing and de novo analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. *PLoS ONE* 6:e21560
 - Huang Q, Lin NF, Chen YL, You H, Lai XX (2003) Suggestions on the planning for a horseshoe crab nature reserve in Pingtan. *Fujian Environ* 20:35–38 (in Chinese with English Abstract)
 - Iwanaga S (2002) The molecular basis of innate immunity in the horseshoe crab. *Curr Opin Immunol* 14:87–95
 - Kamaruzzaman BY, John BA, Zaleha K, Jalal KCA (2011) Molecular phylogeny of horseshoe crab. *Asian J Biotechnol* 3:302–309
 - Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *Trends Ecol Evol* 21:629–637
 - Kopp EB, Medzhitov R (1999) The Toll-receptor family and control of innate immunity. *Curr Opin Immunol* 11:13–18
 - Kregel KC (2002) Invited review: heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. *J Appl Physiol* 92:2177–2186
 - Li QZ, Hu MH (2011) Status of Chinese horseshoe crab population and conservation advance in Beihai, Guangxi. *Mar Environ Sci* 30:131–134 (in Chinese with English Abstract)
 - Li QZ, Li Q, Liu J, Ni LH, Kong LF (2009) Isolation and characterization of microsatellite loci in the horseshoe crab (*Tachypleus tridentatus*). *Conserv Genet* 10:1879–1881
 - Liao YY, Li XM (2001) Present situation of horseshoe crab resources in the sea area of China and tactics of preservation. *Resour Sci* 23:53–57 (in Chinese with English Abstract)
 - Liao YY, Liu JX (2006) Species and distribution of horseshoe crab in Asia sea area. *J Trop Oceanogr* 26:85–90 (in Chinese with English Abstract)
 - Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat Rev Genet* 4: 981–994
 - Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46
 - Nakamura T, Furunaka HTTM, Miyata T, Tokunaga F and others (1988) Tachyplesin, a class of antimicrobial peptide from the hemocytes of the horseshoe crab (*Tachypleus tridentatus*). Isolation and chemical structure. *J Biol Chem* 263:16709–16713
 - Nishida S, Koike H (2010) Isolation and characterization of compound microsatellite loci in the tri-spine horseshoe crab (*Tachypleus tridentatus*). *Conserv Genet Resour* 2: 31–33
 - Rozihan M, Ismail E, John BA, Jalal KCA, Mohd-Adnan A (2013) Genetic diversity of horseshoe crab (*Tachypleus gigas*) in Malaysia revealed using microsatellite markers. *Asian J Anim Vet Adv* 8:63–72
 - Rudkin DM, Young GA, Nowlan GS (2008) The oldest horseshoe crab: a new xiphosurid from Late Ordovician Konservat-Lagerstätten deposits, Manitoba, Canada. *Palaeontology* 51:1–9
 - Shigenaga T, Muta T, Toh Y, Tokunaga F, Iwanaga S (1990) Antimicrobial tachyplesin peptide precursor. cDNA cloning and cellular localization in the horseshoe crab (*Tachypleus tridentatus*). *J Biol Chem* 265:21350–21354
 - Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100: 158–170
 - Wei W, Qi XQ, Wang LH, Zhang YX and others (2011) Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics* 12: 451
 - Weng CH, Xie YJ, Xiao ZQ, Huang LM and others (2012) Distribution and resource of Chinese horseshoe crab (*Tachypleus tridentatus*) in Fujian and other coast water of China. *Chinese J Zool* 47:40–48 (in Chinese with English Abstract)
 - Weng ZH, Xie YJ, Xiao ZQ, Wang ZY, Gui JF (2013) Microsatellite and mitochondrial DNA analysis of the genetic structure of Chinese horseshoe crab (*Tachypleus tridentatus*) in southeast China coast. *Afr J Biotechnol* 12: 2088–2099
 - Xia X (2000) Phylogenetic relationship among horseshoe crab species: effect of substitution models on phylogenetic analyses. *Syst Biol* 49:87–100
 - Xu Q, Chen F, Shin PK, Cheung SG, Chen Y, Ke CH (2011) AFLP analysis of genetic variation among three natural populations of horseshoe crab *Tachypleus tridentatus* along Chinese coast. *Chin J Oceanology Limnol* 29: 284–289
 - Yang MC, Hsieh HL, Huang H, Chen CA (2009) Phylogeography, demographic history, and reserves network of horseshoe crab, *Tachypleus tridentatus*, on the South and East China seaboards. In: *Biology and conservation of horseshoe crabs*. Springer, New York, NY, p 163–181
 - Yang Q, Sun FY, Yang Z, Li HJ (2014) Comprehensive transcriptome study to develop molecular resources of the copepod *Calanus sinicus* for their potential ecological applications. *BioMed Res Int* 2014:493825, doi:10.1155/2014/493825