Science in China Series C: Life Sciences

www.scichina.com
life.scichina.com
www.springerlink.com

ESTs analyses of *Lampetra japonica* liver and comparation transcriptome with the jawed vertebrates

ZHU LiNa[1], DAI YaLi[2], MA Fei[1] & LI QingWei[1†]

[1] College of Life Sciences, Liaoning Normal University, Dalian 116029, China;
[2] College of Life Sciences, Anhui Agricultural University, Hefei 230036, China

A cDNA library was constructed from the liver of *Lampetra japonica*. 10077 ESTs were obtained by random selecting clones for sequencing. The results demonstrated that 8515 ESTs were assembled into 648 consensus sequences, represented 2210 unique transcripts, 47.06% of which were predicted as full length cDNAs. In addition, 1562 ESTs were singlets. Using the BLAST to align the assembled ESTs, we found that 93.9% (2053) transcripts shared similarity to sequences published in GenBank databases. The functional annotations to assembled ESTs showed that the genes, involved in immunology, blood coagulation and metabolism of jawed vertebrates, were highly expressed in the liver of *L. japonica*. Furthermore, 8 potential novel genes were identified. Further comparing liver transcriptome of *L. japonica* with *Fundulus heteroclitus*, *Mus musculus*, *Bos Taurus*, and *Homo sapiens* revealed that the genes of Chitinase and Polysaccharides metabolism were more highly expressed in *L. japonica* than the others, which implied that they may play an important role in immunity of *L. japonica*. In addition, using the TargetScan, we marked microRNA target within 3′ UTR of *L. japonica* liver transcriptome. The data indicated that some microRNA targets were homology with the targets embeded in human cancer genes. The result seems to provide a useful clue to the treatment of human cancer. Therefore, the present work will be an important resource for investigating the functional genomics and proteomics of *L. japonica* as well as evolution of vertebrates.

Lampetra japonica, liver, expressed sequence tag (EST), transcriptome comparison

Lamprey, a living jawless vertebrate, has been regarded as one of the most primitive groups of vertebrates, which filles the gap between invertebrates and vertebrates. Recent archaeological finds indicate that modern lampreyes, compared with the ancestors of 3.6 billion years ago, have little change in morphology structure. Indeed, the lamprey is living fossil[1]. Thus lamprey is the preferred material for revealing the origin and evolution of vertebrates. At present, most studies have been confined to the morphology structure and ecological distribution of the lamprey[2], whilst other aspects of genomics and proteomics are less known. The transcriptome analysis of sea lamprey intestine[3] and *Lampetra japonica* oral gland[4] have been published, but the biggest digestive gland-related genes in the liver have yet not been reported. Therefore, the analysis of *L. japonica* liver transcriptome will provide a theoretical foundation for origin and evolution of vertebrate liver.

Nevertheless, lamprey has not been intensively studied until the end of 19th century as a result of its unique lifestyle and morphology[5]. Many of such studies have been focused on the comparative morphology, especially emphasizing the early evolution of vertebrates. For example, some experiments demonstrated the homology of brain genes and pharyngeal structure related genes in the lamprey and jaw vertebrates[6−7], and discussed *Hox*

Received April 9, 2007; accepted November 7, 2007
doi: 10.1007/s11427-008-0002-3
†Corresponding author (email: liqw@263.net)
Supported by the National High-Tech R&D Program (863 Program) (Grant No. 2007AA09Z428), the National Natural Science Foundation of China (Grant Nos. 60305001 and 60575005), the Key Project of Science and Technology of the Educational Ministry of China (Grant No. 206032), and the Science and Technology Project of Dalian, China (Grant No. 2006E11SF068)

gene duplication mechanism[8]. As an important metabolism organ, liver has increasingly attracted much attention. Savina[9] compared the bioenergetic parameters of liver mitochondria in lamprey and frog during metabolic depression and activity. Rotchell[10] analyzed the catalytic properties of CYP1A enzyme in lamprey liver. Gamper[11] studied the reversible metabolic depression of lamprey liver during pre-spawning. In addition, liver also takes charge the front-line immune defense in the body. Recently, scientists identified a new type of variable lymphocyte receptors from sea lamprey lymphocyte-like cells, which use an unusual gene rearrangements process to produce receptor diversity. This research revealed lamprey has a unique adaptive immune responses[12]. To be brief, Lamprey imprint the evolutionary history from invertebrates to vertebrates, including the evolution information of the various functional genomes. Therefore, further studying on the lamprey liver genome and transcriptome has a very great theoretical and practical significance.

The two cDNA libraries of *L. japonica* oral gland[13] and liver have been constructed firstly in our lab. Now we have completed the bioinformatics analysis of the 1280 ESTs of oral gland[4]. Based on the cDNA library of the *L. japonica* liver, in the present work, we selected randomly cDNA clones for sequencing to construct EST library. After annotating the 10077 EST sequences by BLAST and Gene Ontology (GO), we also compared lamprey liver transcriptome with *Fundulus heteroclitus*, *Mus musculus*, *Bos Taurus* and *Homo sapiens*. Furthermore, we also predicted the microRNA target within 3′ untranslated region. The aim of this research is to obtain specific expression genes in lamprey liver and lay the foundation for future functional genomics research.

1 Materials and methods

1.1 Materials

These *L. japonica* were collected from the Songhua River in Heilongjiang Province, China. PCR primer, enzymes, RNA extraction kits, RT-PCR kits, sequencing kit, pBluescript II SK(+) cloning vector, and *E. coli* JM109 Electro-cell were offered by TaKaRa Biotechnology (Dalian) Co., Ltd.

1.2 cDNA library construction and sequencing

Total RNA were extracted from the liver of *L. japonica* using Trizol Reagent. After being isolated and purified using Oligotex-$dT_{(30)}$, mRNA was used as template to synthesize first-strand of cDNA with Oligotex-dT Linker Primer, and further the double-strand of cDNA was synthesized. *Eco*R I adaptors were ligated to the cDNAs, which were digested with *Not* I and directionally cloned into the *Eco*R I and *Not* I sites of pBluescript II SK(+). Cloned cDNAs were electroporated into *E. coli* JM109 Electro-cell and amplified. The inserted clone fragments which are longer than 500 bp were picked randomly for 5′ sequencing. Moreover, sequences were sequenced using ABI 377 automated sequencer.

1.3 EST Assembly and identifying full-length cDNAs

ESTs were assembled using CAP3[14] to produce contigs and singlets. Sequences were grouped in the same cluster/contigs if there existed a minimum match of 25 bp and the identity greater than 75% in the overlap region. And then singlets were continually assembled until no new overlapping groups. In addition, with the default parameters values, the TargetIdentifier[15] (https://fungalgenome.concordia.ca/tools/TargetIdentifier.html) was used to identify full-length cDNAs.

1.4 Prediction of ORFs and UTR Analysis

OrfPredictor server (http://www.ncbi.nlm.nih.gov/projects/gorf/) was used to predict the ORFs of a batch of EST/ cDNA sequences. Based on the largest open reading frame, the 3′ UTR region sequences were obtained. Tar- getScan[16] software was applied to predict microRNAs conserved target sites among orthologous 3′ UTR.

1.5 Annotation and functional classification of the assembled ESTs

Homology searches were performed with BLASTN against the NCBI nonredundant database. Protein alignments were generated by BLASTX against the Uniprot database. To correlate the unique gene sequence of lamprey liver to a Gene Ontology, the software GoPipe[17] (http://www.fishgenome.org/bioinfo/), which is a tool for integrating BLAST results to streamlined GO annotation for batch sequences, was used to assign the unique gene to a biological process, molecular function, or cellular component. In addition, the sequences without BLASTX and BLASTN matches were further performed domain searches via InterProScan Server[18] (http://www.ebi.ac.uk/InterProScan/) against the Conserved Domain Database. If no homology domain could

be found in considered sequences, they may be considered as the firstly discovered ESTs/genes.

1.6 Comparative analysis of transcriptome

For comparing the trancriptomes, we downloaded liver ESTs of *F. heteroclitus* (33369), *M. musculus* (123756), *B. Taurus* (48787), *H. sapiens* (251224) from the NCBI database. Based on EST expression profile file and GO associational file, the tool GO-Diff[19] was used to identify functional differentiation of liver transcriptomes among *L. japonica* and the above four species, aiming to find these especially lower and higher expression genes in lamprey liver via homology search, molecular functions and biological processes. This goal of the research is to reveal the evolutionary adaptation to the environment and specific gene expression patterns of lamprey liver.

2 Results

2.1 cDNA Library Construction and EST Sequencing

The concentration of the obtained cDNA library of *L. japonica* was 4.6×10^6 pfu/mL and the average length of inserted fragments was 1.2 kb, which showed that the quality of cDNA library was excellent. The positive clones of the cDNA library were randomly picked for 5′ sequencing. After excluding the poor-quality sequences, vector sequences and pollution sequences, we obtained 10077 EST sequences.

2.2 Analysis of assembled ESTs and Identifying full-length cDNAs

Within the total 10077 ESTs, 8515 were assembled into 648 contigs, and 1562 remained as single sequences (singlets). The total number of contigs and singlets was 2210. The total length of contigs is 641486 bp with the average length of 990 bp. Most of the contigs consisted of two ESTs (Table 1). However, there were some large contigs comprised of thousands of ESTs corresponding to Blood plasma apolipoprotein LAL2, Fibrinogen, Vitellogenin, Ribosomal protein L4, Thymus thymic cells cDNA and Acidic mammalian chintinase (Table2). Based on the number of ESTs constituting homologous gene, the higher or lower gene expression can be roughly estimated. Assembly results seem to suggest that the above genes were much higher expressed in the liver of *L. japonica*.

To identify the full length cDNAs, TargetIdentifier webserver was used to predict local protein coding regions and potential start or stop codons. The results demonstrated that a total 1040 full-length cDNAs were predicted, including the 65.12% contigs (422) and 39.56% the single sequences (618), which indicated that nearly 50% of the transcripts were full-length cDNAs.

According to the distribution of ORF (Table 3), we found that the length of coding frame mostly ranged from 100 to 600 bp, and 627 transcripts had ORFs of 300 bp (100 aa) or greater. The results seemed to imply that most transcripts likely encoded proteins. In addition, our findings also displayed that 29 ORFs were larger than 1000 bp. Further homology searches demonstrated that the functions of these larger ORFs corresponded to ribosomal protein, fibrinogen, heat shock protein, acidic mammalian chitinase, bone marrow macrophage cDNA (Table 4). Moreover, the higher homology alignment score and the higher expectation (E) also confirmed the credibility of assembled ESTs.

2.3 MicroRNA target analysis within 3′ untranslated region (3′ UTR)

The TargetScan software was used to predict microRNA target sites and their corresponding microRNAs.

Table 1 Distribution of contigs containing EST

No. of contigs containing EST	No. of contigs	Contig (%)	No. of ESTs	EST (%)
2	263	40.59	526	5.22
3	103	15.90	309	3.07
4	59	9.10	236	2.34
5	27	4.17	135	1.34
6	26	4.01	156	1.55
7	14	2.16	98	0.97
8	9	1.39	72	0.71
9	12	1.85	108	1.07
10	7	1.08	70	0.69
>10	128	19.75	8367	83.03
Total	648	100	10077	100

Table 2 EST numbers of high overlapping groups

Assembly EST	No. of EST	Length (bp)	Similar protein
Contig401	3255	1077	Blood plasma apolipoprotein LAL2
Contig35	777	5385	Fibrinogen
Contig547	669	6050	Vitellogenin
Contig277	265	3425	Ribosomal protein L4
Contig168	202	2958	Thymus thymic cells cDNA
Contig582	191	3454	Acidic mammalian chitinase

Table 3 Distribution of ORF

Length of ORF (bp)	No. of transcripts	%
>1000	29	2.79
1000－600	87	8.37
600－300	511	49.13
300－100	405	38.94
<100	8	0.77
Total	1040	100

Table 4 Function prediction of ORF over 1000 bp

No.	Length (bp)	Similar protein	E	Score
Contig502	4955	Ubiquitin/ribosomal protein S27a fusion protein	3.00E-62	239
Contig65	2452	Fibrinogen alpha-1 chain precursor	0	817
Contig187	1791	40S ribosomal protein S4	1.00E-138	493
Contig119	1731	Heat shock protein HSP 90-alpha	0	827
Contig582	1575	Acidic mammalian chitinase precursor	1.00E-109	400
Contig504	1508	Betaine-homocysteine methyltransferase	0	638
Contig515	1488	Fibrinogen beta chain	0	914
Contig252	1392	Ribosomal protein S6	1.00E-112	411
Contig34	1371	Serpin precursor	0	686
Contig208	1347	Bone marrow macrophage cdna	0	800
Contig447	1323	Serpin precursor	4.00E-87	325
Contig410	1318	Alpha-2 tubulin	0	810
Contig439	1302	*S*-adenosylhomocysteine hydrolase	0	765
Contig35	1299	Fibrinogen gamma chain precursor	0	796
Contig325	1241	Ribosomal protein L5a	1.00E-129	463
Contig308	1224	Serpind1-prov protein	6.00E-37	157
Contig250	1218	Eukaryotic translation initiation factor 3	1.00E-84	317
Contig470	1215	Chromosome 3 SCAF14679	0	635
Contig296	1185	Translationally-controlled tumor protein	1.00E-64	250
Contig401	1164	Blood plasma apolipoprotein LAL2 precursor	5.00E-78	295
Contig514	1155	Opticin	1.00E-58	138
Contig218	1125	Ribosomal protein L17	6.00E-79	298
Contig360	1125	Beta actin	0.00E+00	553
Contig518	1095	Fibrinogen gamma polypeptide	9.00E-58	144
Contig321	1059	Blood plasma apolipoprotein LAL2 precursor	5.00E-06	55
Contig595	1050	ADP/ATP translocase 1	1.00E-147	523
Contig280	1020	MGC81823 protein	1.00E-145	518
Contig22	1014	Cytochrome P450, family 4	1.00E-102	373
Contig21	1005	Serum lectin precursor	0	661

The results showed that 25500 target sites were complementary to 2097 microRNAs of the other species within the 611 of 791 3′ UTRs. The further analyses indicated that 23107, 1895, 498 of 25500 microRNA target sites were respectively matched to the animal, *C. elegans* and virus. In addition, our findings also demon-

strated that microRNA target sites of different species were highly conserved. For example, the miR-7, miR-92, miR-18 family of 17 species, miR-19a family of 16 species, miR-16, miR-17 family of 15 species, and miR-15a, miR-20, miR-29, miR-100, miR-103, miR-135, miR-181, miR-106 family of 14 species were respectively matched to the same target site in 3′ UTR of lamprey.

2.4 Annotation of ESTs and discovering of novel genes

The results of BLASTX search showed that 2053 (93.9%) of all transcripts were similar to protein sequences published in the nonredundant protein database. Of the 2053 transcripts, 1686 tanscripts shared significant similarity with known sequences, and 367 transcripts matched to the speculatedor unknown functional protein sequences. Furthermore, the results also indicated that many transcripts of lamprey liver corresponded to the same protein product. The representative instance was that 123 transcripts corresponded to various ribosomal protein subtypes, namely Macrophages cDNA (51), NADH oxidase dehydrogenation (38), Thymus cells cDNA (32), Blood plasma apolipoprotein (28), the enzyme associated with chitin metabolism (22). In addition, it was also interesting to find several immunity associated factors namely T-cell receptor (TCR), B cell receptor (BCR), Major histocompatibility complex (MHC class I, MHC class II), Complement factor (C3), Complement factor B (CFB), Transferrin and Heat shock protein in the lamprey liver ESTs.

With the development of EST sequencing and the analysis methods in bioinformatics, new genes can be extensively identified. In present study, we found 18 putative transcripts without BLASTX and BLASTN matches. Further InterProScan searches revealed that 4 of 18 putative transcripts contained at least one known domain. Moreover, 8 of the remained 14 transcripts had the ORFs longer than 300 bp. These above results seemed to imply that the 8 transcripts, i.e., contig68, contig600, LLLNSEM008FB09, LLLNSEM011FG11, LLLNSEM055FD08, LLLNSEM134FD08, LLLNSEM-120FF10, LLLNSEM064FD09, may be the firstly discovered novel genes. The present findings provide the useful information for the novel gene cloning and further functional analysis.

2.5 Assignment of GO Terms to the EST Assembly

To address the issue of function annotation, based on the BLASTX results, these transcript sequences were annotated using Gopipe software. The results indicated that 664 sequences were assigned to the biological process, 764 sequences were involved in molecular function, and 565 sequences were categorized to the cellular components. We further found, in the biological process, the order of the abundantly expressive categories were “physiological processes” (664), “metabolism” (526), “cell growth” (220), and “transport” (158) (Figure 1), whilst in the molecular function category “binding” was the most (356) (Figure 2), and the “intracellular” was the most expressed in cellular components (427) (Figure 3).

2.6 Comparative analysis of transcriptome

Comparing the liver transcriptome with the other species

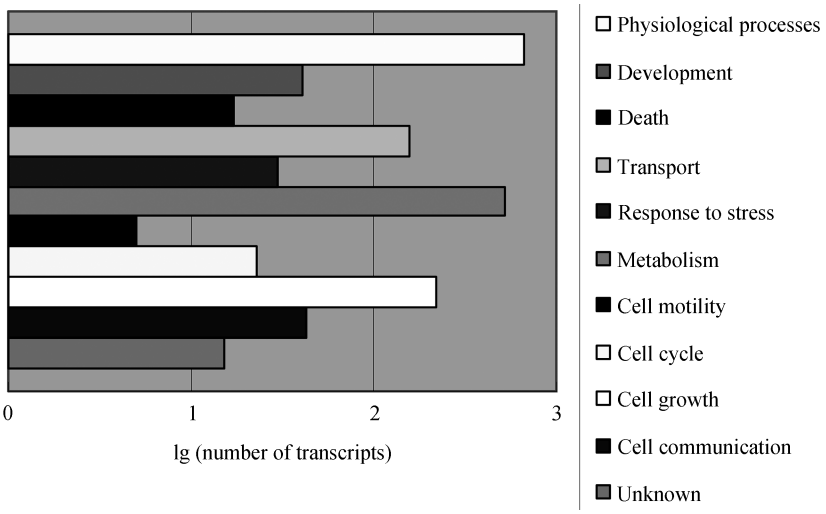

Figure 1 Classification by biological process.

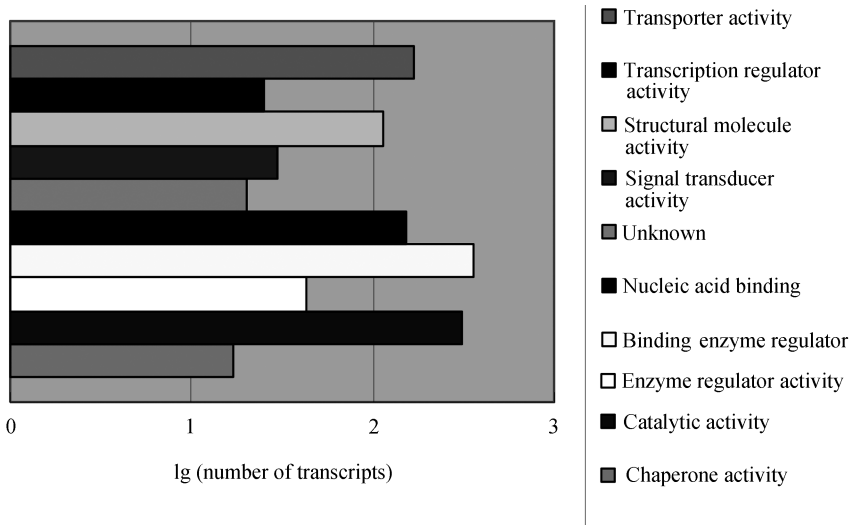

Figure 2 Classification by molecular function.

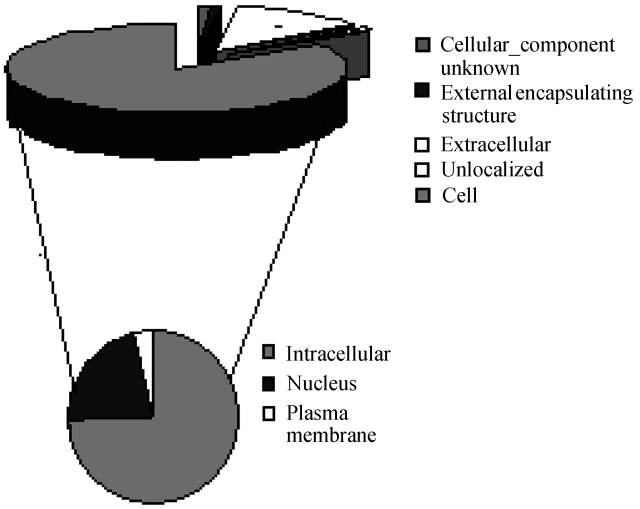

Figure 3 Classification by cell component.

Table 5 Gene function of *L. japonica* liver transcription expression more than other species

GO No.	GO function	*Fundulus* (times)	*Mus* (times)	*Bos* (times)	*Homo* (times)
272	Polysaccharide catabolism	62.74	72.45	44.91	79.49
44247	Cellular polysaccharide catabolism	124.53	71.90	65.14	123.96
4568	Chitinase activity	123.57	160.53	70.03	246.01
5529	Sugar binding	81.77	21.34	22.41	23.94
6030	Chitin metabolism	123.57	142.70	70.03	166.65
6032	Chitin catabolism	123.57	160.53	70.03	246.01
6040	Amino sugar metabolism	41.19	36.69	28.01	35.14
6041	Glucosamine metabolism	41.19	45.87	28.01	41.00
6043	Glucosamine catabolism	123.57	160.53	70.03	172.20
6044	*N*-acetylglucosamine metabolism	41.19	45.87	28.01	44.16
6046	*N*-acetylglucosamine catabolism	123.57	160.53	70.03	246.01
46348	Amino sugar catabolism	123.57	160.53	70.03	172.20
8061	Chitin binding	123.57	1284.28	120.05	469.65

showed that, in *L. japonica* liver, the lower expressed sequences were the genes associated with RNA splicing, DNA repair and cell apoptosis negative regulation, whereas polysaccharides hydrolytic activity proteins were the higher expressed, including chitin metabolism enzymes and *N*-acetylglucosamine metabolism enzymes. As shown in Table 5, the expression level of these genes was significantly higher than other species. The number of above-mentioned functional sequences were 22, and 12 of which were homologous to Acidic mammalian chitinase, their expression levels were respectively 123, 142, 70 and 166 times higher than *F. heteroclitus*, *M. musculus*, *B. Taurus* and *H. sapiens*. In addition, conserved domain search demonstated that the chitinase of *L. japonica* liver has the similar domain to human, which belonged to polysaccharide hydrolysis 18 family.

3 Discussion

3.1 cDNA Library construction and EST sequencing

L. japonica has been regarded as one of the most primitive groups in vertebrate, belonging to *Vertebrata Cyclostomata Petromyzoniformes*, which is also the important material for the study of animal evolution. Currently *L. japonica* is very scarce species and in severely danger situation. However, there are still lacking effective protection measures. Herein establishing cDNA library from different tissues will not only provide useful information to the evolutionary genetics, but also make efforts in the preservation of genetic resources.

3.2 Annotation of *L. japonica* liver transcriptome

According to the number of ESTs in contigs and homologous proteins, we can roughly estimated the protein expression level. Among the genes with assigned functions, the most abundant respresentation was Ribosomal protein, Blood plasma apolipoprotein LAL2, Fibrinogen, Vitellogenin, NADH oxidase dehydrogenation, Macrophage cDNA, Thymus cells cDNA and Acidic mammalian chitinase. Furthermore, we found some genes related to the jaw vertebrate immune system, such as above-mentioned Macrophages cDNA, Thymus cells cDNA, Chitinase, TCR, BCR, MHC I, MHC II, Complement factors C3, C2, CFB, Interferon, Transferrin, Heat shock protein (Table 6). Liver is rich in macrophages, which can be used to remove and swallow harmful substances such as microorganisms from blood and absorption with the intestinal by the means of phagocytosis, isolation, and elimination of the invasion and endogenous antigen. Immature T cells grows in the thymus and thymosin can enhance the ability of inducing B-cell mature[20]. Chitin is a kind of animal fiber with a variety of biological activity, including strengthening immune function, and improving disease defense mechanism. Therefore, chitin metabolism may be associated with immunity[21]. The high expression of Macrophages cDNA, Thymus cells cDNA, and Chitinase suggested that they appear to play an important role in *L. japonica* defense.

Complement system is the central support of innate immune system, and its evolutionary history can be traced back to echinoderms (sea urchin), invertebrate (echinoderms, ascidians) and cyclostomata. They have the most primitive complement system (lectin pathway). As the very conservative widespread innate immune system, interferon is well known to be the cell functional regulation systems emerged firstly in host defense response. Besides, transferrin, an important factor on inhibiting bacterial growth, has antibacterial and germicidal self-protection function[22]. *L. japonica,* mainly depending on other fish flesh for food, is semi-parasitic animals, and its potential nonspecific immunity mechanism plays an important role in preventing infection which is the more effective way to remove and degrade pathogenic microorganisms and other harmful substances. Therefore, it is likely that the nonspecific immunity of *L. japonica* plays a greater role in the immune system.

L. japonica has shown some characteristics of the adaptive immune system[23−25]. They have an adaptive immune response, in which the receptor is divergent to the jaw vertebrate recombinant antigen receptor. Interestingly, they use an unusual gene rearrangement process to produce receptor diversity[12]. Lymphocyte-like cells in the intestine of the sea lamprey were isolated, but no sequences with significant similarity to MHC and T cell receptor were found. The lamprey was considered to reach a stage for the emergence of adaptive immunity[3]. Although the above various factors were found in our EST sets, the similarity was still very low. Significantly, the heat shock protein was relatively abundant. More and more evidence showed that heat shock proteins participated in specific and non-specific immune response. Besides, the heat shock protein, as a molecular chaperone with antigen information, can promote spe-

Table 6 *L. japonica* liver immune-related genes

No. of transcripts	Similar protein	No. of transcripts	Similar protein
51	Bone marrow macrophage cDNA	1	Killer immunoglobulin
32	Thymus thymic cells cDNA	1	Interferon gamma inducible protein
18	NOD-derived CD11c +ve	1	HLA-B associated transcript 1
12	Acidic mammalian chitinase	1	MHC class I antigen
7	Heat shock protein	1	MHC class II antigen
6	Proteasome 26S ATPase	1	Lymphotoxin
5	ABC transporter	1	Killer cell lectin-like receptor
4	Complement C3 precursor	1	B-cell translocation gene 1 protein
2	Macrophage stimulating protein	1	B-cell receptor CD22 precursor
2	LMPX(7) of lamprey	1	CD3 epsilon-associated signal transducer
2	Complement factor B	1	Dendritic cell protein (GA17 protein)
1	CD9-like protein	1	Complement regulatory GPI anchor protein
1	Cd63-prov protein	1	Complement receptor 2
1	Thymosin	1	Complement receptor 1
1	T-cell receptor beta chain	1	Complement factor H-related 5
1	Proteasome subunit beta type 4	1	Complement C2 precursor
1	Proteasome subunit beta type 7	1	CD81 antigen
1	Proteasome subunit alpha type 2	1	CD68 antigen
1	Proteasome subunit alpha type 3	1	Transferrin

cific immune response and stimulate the body's immune response, but sometimes heat shock proteins can also activate immune reaction[26−27]. Due to the lack of B cell, T cell receptor and MHC antigen in *L. japonica*, it was reasonable to estimate that *L. japonica* tried to make use of heat shock protein to carry out antigen targeting or activating immune reaction in the early adaptive immunity stages.

The other representatives which expressed abundantly were involved in protein synthesis and energy metabolism, such as Blood plasma apolipoprotein LAL2, Ribosomal protein, NADH dehydrogenation oxidase, Cytochrome oxidase and ATP synthase. Previous studies had identified that Blood plasma apolipoprotein LAL1 and LAL2 were abundant in lamprey blood[28]. Here, our results demonstated that they were highly expressed in the liver. Apolipoprotein, playing the major role in transportting, dissolving cholesterol and phospholipids, was controlled by the apolipoprotein E gene which was secreted in the liver. Moreover, it also was responsible for inducing macrophage and cholesterolste accumulation, and significantly promoting macrophage proliferation. Mouse liver SAGE data analysis showed that apolipoprotein mRNA expression was higher than 3.8%. Furthermore, the genes which encoded plasma proteins expressed highly in *L. japonica* liver. In addition, our results also indicated that 123 transcripts corresponded to the Ribosomal protein.

The coagulation factors including fibrinogen α, β, γ chain coagulation factor VIIa and IX, as well as prothrombin, expressed large. Because the liver was the tissue where coagulation factors were synthesized, it was reasonable that the fibrinogen was highly expressed[29]. Therefore, our findings also implied that the *L. japonica* might have a perfect coagulation mechanism.

3.3 MicroRNA target analysis

MicroRNA, one of the small molecular RNA families, is a very small section of non-coding RNA sequences, which can mediate negative post-transcriptionregulation. Nevertheless, the majority of animal microRNA do not lead to the degradation of target mRNA[30]. With the corresponding microRNA targets, we tried to investigate the potential regulation.

Volinia et al. from the Ohio State University analysed 540 cancer samples of lung, chest, stomach, prostate, colon and pancreas, and discovered that most microRNA overexpressed in cancer, including miR-17-5p, miR-20a, miR-21, miR-92, miR-106a and miR-155 family[31]. Our data also identified the targets of miR-92, miR-17, miR-20 and miR-106 microRNA family. Adult B-cell chronic lymphatic leukemia (CLL) patients often present the absence or downregulation of miR-15a and miR-16-1 cluster[30]. Additionally, we identified the targets of miR-15a and miR-16 family. The microRNA families putatively involved in the regulation of cancer genes are similar in the *L. japonica* and human. Therefore, *L. japonica* can become a new kind of animal

model for studying human cancer and diseases.

3.4 Comparison of tanscriptome

Comparisons between different genomes can obtain a lot of evolutionary and genetic information, providing important insight into the origin of vertebrates[32]. *L. japonica* represents an important transitional phase from invertebrate to vertebrate. Comparing the *L. japonica* liver transcriptome with the integrated liver transcriptome of *F. heteroclitus*, *M. musculus*, *B. taurus* and *H. sapiens* can relatively trace certain genes change in the evolutionary process and provide clues on the origin and evolution of vertebrates. By comparison, we found that the genes involved in RNA splicing, DNA repair, negative regulation and cell apoptosis expressed less in *L. japonica* liver than those in the others. This result showed that some mentioned aspects of evolution mechanism in liver were perfect gradually. On the other hand, the higher expressed genes in *L. japonica* liver were involved in polysaccharide hydrolysis activity, including chitin metabolism and glucosamine metabolism and so on. These genes expressed about 100 times in considered tissue than the other species. Moreover, chitinase expressed higher in BLASTX annotation, which suggested it must play an important role.

Chitinase, encoded by single gene, is a digestive enzyme that breaks down glycosidic bonds in chitin. Abundantly expressed chitinase in *L. japonica* liver was homologous to the acidic mammalian chitinase, namely AMCase. The chitinase, firstly found the gastrointestinal tract and lung of human, may play a key role in digestion and defense[33]. In the same year, Suzuki reported that AMCase was expressed in bovine liver which may defense crustacean pathogenic microorganisms[34]. The AMCase expressed in peokaryotic system can hydrolyse chitin in the cell walls of *T. rubrum* and *C. albicans*, which implies its potential antifungal function[35]. Ov-CHI-1 was a chitinase as a type of chitinase, Ov-CHI-1 was specifically expressed in the larvae stage of *Onchocerca volvulus*. Various evidence showed that this gene expressed only in 22% of normal individuals, but expressed in 42%—62% of infected individuals. Meanwhile IgG3 and peripheral mononuclear cells (PBMC) expressed increasedly[36]. A favorable evidence is that the macrophages are expressed abundantly in *L. japonica* liver. In addition, the immune function of decomposed chitin had been recognized internationally. Nishimura[37] injected multi-porous deacetylated chitin to mice intraperitoneally, and the results showed that the substance could stimulate macrophages to produce interleukin 1 and enhance the activity of macrophages. Glucosamine which was the decomposed products of chitin could also activate natural killer cells (NK) and lymphokine-activated killer cells (LAK). Chitosan could induce partial macrophage proliferation, and increase the activity of immune function[38]. In a word, the chitin was immune-related.

In summary, we speculated the chitinase immunity metabolic pathway in *L. japonica* liver as followed. Firstly, the chitinase perfects the "disrupting task" to the cell wall of pathogenic microorganisms. Next, chitosan stimulates the macrophage proliferation in livers, which can be used to remove or swallow harmful substances via nonspecific immunity. Abundant expression of chitinase in *L. japonica* liver may be correlated with its specific lifestyle. Chitin exists widespreadly in cell wall of low plant and the shell of fungus, shrimp, crab, insects and crustaceans. As known, lamprey larvae mainly feed on planktonic, lives in freshwater streams, while the majority of adult lamprey are the semi-parasitic animals. They attach to other fish and feed on their blood. In dependent life, they eat plankton. Thus the pathogens from blood and plankton may be the sources of chitin.

Considered the MHC-I, MHC-II, C3, CFB, Interferon, Transferrin, Heat shock protein immune-related factor and chitinase, Macrophage cDNA, we primely estimated that nonspecific immunity may occupy an important position in the *L. japonica* immune mechanisms. Immunity should be related with chitin metabolism. In addition, with certain factors of the adaptive immune being found in this article, *L. japonica* could have the ability of adaptive immune. Therefore, we speculated that adaptive immune of *L. japonica* might be close but different from the jaw vertebrate.

4 Conclusion

For the special status of *L. japonica* in phylogenetic evolution, it is very significance to construct the cDNA library of *L. japonica* for further elucidating the functional genome of *L. japonica* and the evolution of vertebrates. In the present work, the cDNA library of *L. japonica* liver was constructed. 10077 ESTs were obtained by random selecting clones for sequencing. Bioinformatics analyses of the EST sequences revealed the pro-

tein expression characteristics and the potential mechanisms of higher expressed protein in *L. japonica* liver. Further analysises of microRNA target within 3′ UTR of *L. japonica* liver transcriptome demonstrated that some microRNA targets were homologous with the counterparts in human cancer gene regulation, which may provide some useful clews for the treatment of human cancer. Moreover, by comparing liver transcriptome of *L. japonica* with the other four species, it could be found that the genes of chitinase and polysaccharides metabolism were more highly expressed in *L. japonica* than in the others, which implied that they may play an important role in immunity of *L. japonica*. According to the present results, we further speculated the unique immune response model and evolution stage of *L. japonica*. We have systematically studied the two transcriptomes of oral glands and liver in *L. japonica*. In conclusion, our work will have very important academic significance and application foreground, as well as provide the foundation for further functional gene cloning and proteintomics study.

1 Janvierg P. Palaeontology: Modern look for ancient lamprey. Nature, 2006, 443: 921－924
2 Zhang K K, Guo W Y, Li X D, et al. Analysis of fatty acids composition in *Lampetra japonica* muscle. J Fish Sci China (in Chinese), 2000, 7(1): 116－117
3 Mayer W E, Uinukool T, Tichy H, et al. Isolation and characterization of lymphocyte-like cells from a lamprey. Proc Natl Acad Sci, 2002, 99(22): 14350－14355
4 Gao Q, Feng Y, Wu Y, et al. Expressed sequence tags (ESTs) analysis of the oral gland of *Lampetra japonica*. Acta Genet Sin (in Chinese), 2005, 32(10): 1045－1052
5 Hardisty M W. The skeleton. In: Hardisty M W, Potter I C, eds. The Biology of the Lampreys. London: Academic Press, 1981. 333
6 Murakami Y, Ogasawara M, Satoh N, et al. Compartments in the lamprey embryonic brain as revealed by regulatory gene expression and the distribution of reticulospinal neurons. Brain Res Bull, 2002, 57: 271－275
7 McCauley D W, Bronner-Fraser M. Importance of *SoxE* in neural crest development and the evolution of the pharynx. Nature, 2006, 441: 750－752
8 Takio Y, Pasqualetti M, Kuraku S, et al. Evolutionary biology: Lamprey *Hox* genes and the evolution of jaws. Nature, 2004, 416: 386－387
9 Savina M V, Emelyanova L V, Belyaeva E A. Bioenergetic parameters of lamprey and frog liver mitochondria during metabolic depression and activity. Comp Biochem Physiol, 2006, 145: 296－305
10 Rotchell J M, Steventon G B, Bird D J. Catalytic properties of CYP1A isoforms in the liver of an agnathan (*Lampetra flu*6*iatilis*) and two species of teleost. Comp Biochem Physiol, 2000, 125: 203－214
11 Gamper N L, Savina M V. Reversible metabolic depression in hepatocytes of lamprey (*Lampetra flu*6*iatilis*) during pre-spawning: Regulation by substrate availability. Comp Biochem Physiol, 2000, 127: 147－154
12 Pancer Z, Amemiya C T, Ehrhardt G A, et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature, 2004, 430: 174－180
13 Li X X, Feng Y, Xiao R, et al. The construction of *Lampetra japonica* salivary gland. J Liaoning Normal Univ (in Chinese), 2004, 27(1): 73－75
14 Huang X, Madan A. CAP3: A DNA sequence assembly program. Genome Res, 1999, 9: 868－877
15 Min X J, Butler G, Storms R, et al. TargetIdentifier: A webserver for identifying full-length cDNAs from EST sequences. Nucleic Acids Res, 2005, 33: 669－672
16 Lewis B P, Burge C B, Bartel D P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell, 2005, 120(1): 15－20
17 Chen Z Z, Xue C H, Zhu S, et al. GoPipe: Streamlined gene ontology annotation for batch anonymous sequences with statistics. Biochem Biophys, 2005, 32(2): 187－191
18 Zdobnov E M, Apweiler R. InterProScan—An integration platform for the signature-recognition methods in InterPro. Bioinformatics, 2001, 17(9): 847－848
19 Chen Z Z, Wang W, Liu J J, et al. GODiff: Mining functional differentiation between EST-based transcriptomes. BioMed Central Bioinformatics, 2006, 7(1): 72
20 Wang F, Zuo M X. Human and Animal Physiology (in Chinese). Beijing: Higher Education Press, 2001
21 Chen P, Gao W P. Chitosan health function. Chin J Food Hygiene (in Chinese), 2001, 13(3): 38－39
22 Long H. Research advances on 10 nonspecific immune molecules of aquatic animals. J Yangtze Univ (in Chinese), 2005, 2(11): 67－72
23 Good R A, Finstad J, Litman G W. The biology of lampreys. II. Immunology, 1972, 1: 405－432
24 Hagen M, Filosa M F, Youson J H. The immune response in adult sea lamprey (*Petromyzon marinus* L.): The effect of temperature. Comp Biochem Physiol, 1985, 1: 207－210
25 Perey D Y, Finstad J, Pollara B, et al. Evolution of the immune response. VI. First and second set skin homograft rejections in primitive fishes. Lab Investig, 1968, 1: 591－597
26 Srivastava P K, Menoret A, Basu S. Heat shock proteins come of age: Primitive functions acquire new roles in an adaptive world. Immunity, 1998, 8: 657－665
27 Srivastava P. Interaction of heat shock proteins with peptidesand antigen presenting cells: Chaperoning of the innate andadaptive immune responses. Ann Rev Immunol, 2002, (20): 395－425

28 Pontes M, Xu X, Graham D, et al. cDNA sequences of apolipoproteins from lamprey. Biochemistry, 1987, 26(6): 1611－1617
29 Chen T G. Bioinformatics study on transcriptome of human fetal liver aged 22 weeks of gestation and genome of SARS-CoV(BJ-01). Dissertation for Doctoral Degree. Beijing: Academy of Military Medical Sciences (in Chinese), 2005. 91－100
30 Urora E K. Oncomirs－MicroRNAs with a role in cancer. Nat Rev Cancer, 2006, 6(4): 259－269
31 Volinia S, Calin G A, Liu C G, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci, 2006, 103(7): 2257－2261
32 Wu S F, Zhu Y P, He F C. Progress in the comparison of transcriptome and proteome. Prog Biochem Biophys (in Chinese), 2005, 32(2): 99－105
33 Boot R G, Blommaart E F C, Swart E, et al. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. J Biol Chem, 2001, 276(9): 6770－6778
34 Suzuki M, Morimatsu M, Yamashita T, et al. A novel serum chitinase that is expressed in bovine liver. FEBS Lett, 2001, 506(2): 127－130
35 Chen L, Shen Z, Liu Y F, et al. Cloning of acidic mammalian chitinase gene mutant and its sequence analysis. Med J Chinese PLA (in Chinese), 2006, 31(3): 231－233
36 Yang W, Gillian E, James M, et al. Human immune responses to infective stage larval-specific chitinase of filarial parasite, *Onchocerca volvulus*, Ov-CHI-1. Filaria J, 2003, 2(1): 6－20
37 Nishimura K, Nishimura S, Seo H, et al. Macrophage activation with multi-porous beads prepared from partially deacetylated chitin. J Biomed Mater Res, 1986**,** 20(9): 1359－1372
38 Xiao H. Immunology and antitumor activities of chitin/chitosan and its derivatives. Marine Sci (in Chinese), 1999, 3(3): 30－32