Comparative Biochemistry and Physiology, Part D 5 (2010) 116–123

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology, Part D

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

Identification of immune-relevant genes by expressed sequence tag analysis of head kidney from grass carp (*Ctenopharyngodon idella*)

Feng Liu [a,1], Dapeng Wang [b,c,1], Jianjun Fu [a], Gaoyuan Sun [b], Yubang Shen [a], Lingli Dong [b], Bing Zhang [b], Songnian Hu [b,*], Jiale Li [a,d,*]

[a] Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources Certificated by Ministry of Education, Shanghai Ocean University, Shanghai 201306, PR China
[b] Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, PR China
[c] Graduate University of Chinese Academy of Sciences, Beijing 100049, PR China
[d] Aquaculture Division, E-Institute of Shanghai Universities, Shanghai Ocean University, Shanghai 201306, PR China

ARTICLE INFO

Article history:
Received 20 November 2009
Received in revised form 11 February 2010
Accepted 15 February 2010
Available online 20 February 2010

Keywords:
Grass carp
cDNA library
Head kidney
Immune-related gene
FcRγ
Immunoglobulin light chains

ABSTRACT

Grass carp is the third largest aquaculture species in global production. However, genomic research of this species has been limited. To identify immune-related genes in grass carp, a normalized full-length cDNA library was constructed from head kidney tissues, and 6432 randomly selected clones were sequenced. 5289 high quality expressed sequence tags (EST) were generated and assembled into 2687 unigenes. Among them, 1585 unigenes showed significant similarity with known sequences in public databases, whereas the remaining 1102 unigenes appeared to be novel sequences with unknown functions. In particular, 136 immune-related genes were identified to encode immunoglobulins, FcRγ, IFN-related proteins, various CD markers, MHCs, complements and other important immune-related factors; a majority of these genes are reported in grass carp for the first time. Sequence analysis indicated that grass carp has at least three subtypes of immunoglobulin light chains, namely L1, L2 and L3. Furthermore, FCRγ was found to broadly express in different tissues. Our study constitutes the first EST analysis of lymphatic tissue in grass carp, and could pave the way for further research of immune-related genes and functional genomics in grass carp.

1. Introduction

Grass carp (*Ctenopharyngodon idella*) is one of the most important farmed fish species in China, with a culture history dating back to the 7Pth[P] century (Tang Dynasty) (Cai 1991). Because of its herbivorous, low protein diet and excellent growth traits, grass carp is highly suitable for aquaculture in developing countries, and has been introduced into 115 countries for aquaculture and biological control against aquatic plants (Cudmore and Mandrak, 2004). According to FAO (Food and Agriculture Organization of the United Nations), the yield of grass carp reached more than 4 million metric tons in 2006, accounting for the third biggest contributor to the global aquaculture production (FAO 2008). Despite its favourable growth traits, grass carp is susceptible to several pathogens, including *Aeromonas hydrophila* that causes the grass carp bacterial hemorrhage disease and heavy losses in aquaculture every year (Huang et al., 1983). In recent years, several functional genomics studies in grass carp have been reported. For example, 1411 and 567 ESTs have been found in the intestine and hepatopancreas, respectively (Chen et al., 2006; Zhang et al., 2007). In addition, various immune factors such as IgM (Wang et al., 2008), MHC (Hao et al., 2007), microglobulin (Hao et al., 2006) tumor necrosis factors and related proteins (Xu et al., 2008b), complements (Li et al., 2007), chemotatic factor receptors (Chang et al., 2007), toll-like receptors (Su et al., 2009), source of immunodominant MHC-associated peptides (Chang et al., 2005; Xu et al., 2008a), and LEAP-2 (Liu et al., 2010) have been cloned and analyzed. However, fewer than 700 grass carp nucleotide sequences can be obtained in GenBank to date. The vast majority of functional genes remain uncharacterized in grass carp.

Full-length cDNA library is an important resource for gene identifying and functional genome sequencing. Several methods have been developed to construct cDNA libraries enriched for full-length cDNAs, including RNA oligo ligation to the 5′ end of mRNA (Suzuki et al., 1997), 5′cap affinity selection via eukaryotic initiation factor (Edery et al., 1995), and 5′ cap biotinylation followed by biotin affinity selection (Carninci et al., 1996). However, these methods are all laborious and prone to mRNA degradation. The SMART (switching mechanism at the 5′ end of RNA transcript) method was developed in recent years, and has been used for rapid construction of full-length

* Corresponding authors. S. Hu is to be contacted at Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, No. 7 Beitucheng West Road, Chaoyang, Beijing 100029, PR China. Tel./fax: +86 10 82995362. J. Li, Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources Certificated by Ministry of Education, Shanghai Ocean University, 999 Huchenghuan Road, Shanghai 201306, PR China. Tel./fax: +86 21 61900401.
E-mail addresses: husn@big.ac.cn (S. Hu), jlli2009@126.com (J. Li).
[1] These authors contributed equally to this work.

1744-117X/$ – see front matter.
doi:10.1016/j.cbd.2010.02.003

cDNA library (Zhu et al., 2001; Wellenreuther et al., 2004; Ling et al., 2007). Furthermore, in conjunction with the duplex-specific nuclease (DSN) normalization method, SMART can be used to generate the normalized full-length library where highly abundant cDNAs are partially degraded to facilitate identification of rare transcripts (Zhulidov et al., 2004; Xie et al., 2007).

Immunoglobulins, composed of light chains and heavy chains, play critical roles in adaptive immune response (Kasahara et al., 2004). Distinct light chains have been identified in mammals, birds and fishes. Birds appear to have only one type of light chain, λ chain, whereas mammals have both λ and κ chains; in comparison, *Heterodontus francisci* (elasmobranch) has I, II, and III subtypes of light chains (Rast et al., 1994). In teleosts, several subtypes of light chains such as L1, L2, L3, G and F have been reported (Ghaffari and Lobb, 1997; Haire et al., 2000; Timmusk et al., 2000; Tomana et al., 2002; Ishikawa et al., 2004). As a subunit of the high affinity immunoglobulin E receptor, FcRγ is also present in the low affinity Fc receptor for IgG in mammals, and may also be shared by FcμR and FcδR (Fujiki et al., 2000). FcRγ activates immune response mediated by lymphocytes, and is involved in sensitization response and resistance to parasites (Kinet, 1999). In mammal NK cells, FcRγ appears to exclusively function in low affinity Fc receptor for IgG. However, research in channel catfish has showed that FcRγ may function in the immune response mediated by NK cells via combination between IgM and FcμR (Shen et al., 2003), indicating functional divergence between fish and mammalian FcRγs. Factors of the clusters of differentiation (CDs) localize on the surface of leukocytes, and play diverse roles in signaling cascades and immune responses. Among them, CD40 has been reported in several fish species (Park et al., 2005), whereas no study has been published on fish CD11/CD18 dimer.

In the current study, a normalized full-length cDNA library was constructed from head kidney tissues of grass carp treated with *A. hydrophila* by using SMART coupled with DSN. Subsequently, sequencing, EST assembly and annotation led to identification of 136 immune-related genes. Furthermore, sequence analyses of IgL and FcRγ were performed, and the expression pattern of FcRγ in different tissues was investigated using real-time PCR.

2. Materials and methods

2.1. Materials

Grass carp aged 6 and 18 months and weighing about 20 g and 150 g, respectively, were raised in circulating water systems at 23 °C for 4 weeks prior to experiment. The dissolved oxygen concentration was maintained above 5 mg/mL by an oxygen supply system. 30 fish for each age group were i.m. injected with formalin-killed *A. hydrophila* S2 (kindly provided by the Aquatic Pathogen Collection Centre of Ministry of Agriculture, China) at a dose of 2.0×10^6 cells (suspended in 200 μL PBS) per 100 g body mass. Head kidney samples from 10 fish for each age group were collected on days 1, 3 and 7 post injection respectively. Samples were quickly frozen in liquid nitrogen and stored in a −80 °C ultra-low freezer until RNA isolation. All these pooled procedures were done in order to achieve expressed genes from fishes at different ages or at different status after injection.

2.2. Construction of cDNA library and DNA sequencing

Total RNA was isolated from pooled samples using Trizol reagent (Invitrogen, USA). The cDNAs were synthesized with SMARTP^{TMP} PCR cDNA synthesis Kit (Clontech, USA), and normalized by using TRIMMER cDNA normalization kit (Evrogen, Switzerland). The products were used as templates for successive long distance (LD) PCR amplifications, which were carried out using Advantage 2 PCR kit (Clontech, USA) with the following program: 95 °C for 1 min, followed by 25 cycles of 95 °C for 15 s, 64 °C for 20 s and 72 °C for 4 min, 1 cycle of 72 °C for 5 min. The PCR products were subjected to electrophoresis, and purified using a gel extraction kit (Qiagen, Germany). Different sizes of cDNAs were ligated into pGEM-T vector (Promega, USA) respectively before electroporation into ElectroMAXTM DH10BTM cells (Invitrogen, USA). Individual clones were randomly picked and cultured in 96-well plates containing 2×YT medium for 16 h at 37 °C. Plasmids were isolated using the alkaline lysis method and stored at −20 °C. Sequencing was conducted on 3730XL DNA automated sequencer (ABI, USA) using primer T7 (5′- TAATACGACTCACTATAGGG-3′). All operations followed the manufacturers' protocols.

2.3. EST assembly, annotation, and bioinformatic analysis

The raw sequences were analyzed for base calling and quality assessment using the Phred software. Low quality sequences and the sequences of vector and adaptor were trimmed off by using the Q13 and cross-match programs, respectively (Ewing and Green, 1998; Ewing et al., 1998). High quality EST sequences were assembled using the Phrap software (http://www.phrap.org/) after short sequences (length <100 bp) had been filtered. The contigs were examined using the Consed program, and the misassembled sequences were corrected manually (Gordon et al., 1998). Unigenes including contigs and singletons were searched against the NCBI non-redundant (nr) protein database using the BLASTX program (Altschul et al., 1997). Unmatched sequences were searched against the NCBI nucleotide (nt) databases using BLASTN (Altschul et al., 1990). The cutoff E-values were 1.0×10^{-10} for BLASTX and BLASTN. The best hits were regarded as the annotations of unigenes, and sequences with the same annotation were considered as the same unigene.

For function assignment, unigenes were searched against the Universal Protein Resource (Uniprot) database (http://www.uniprot.org/) by BLASTX with a threshold of 1.0×10^{-5}. GO terms were assigned to unigenes using the Uniprot2GO program (Ashburner et al., 2000). Unigenes with GO terms were further classified into specific function categories and submitted to the Kyoto Encyclopedia of Genes and Genomes (KEGG) online service (http://www.genome.jp/kegg/) for Pathway analysis (Kanehisa et al., 2008). To identify the microsatellite markers, we used the program SciRoKo version 3.3 (http://www.kofler.or.at/bioinformatics/SciRoKo/index.html) that can easily do the whole genome microsatellite search and investigation. Both perfect and mismatched motifs were identified with a minimum number of repeat units of five and a minimum size of repeats of two.

To compare the functional distributions of unigenes between grass carp and other species, 4885 EST sequences from head kidney of channel catfish were downloaded from NCBI dbEST database (http://www.ncbi.nlm.nih.gov/dbEST/). These EST were clustered into 3027 unigenes. As a control, all unigenes of zebrafish, including 51,481 sequences, were collected from NCBI unigene database (http://www.ncbi.nlm.nih.gov/unigene). GO annotation analyses were performed for unigenes from channel catfish and zebrafish respectively, following the same procedures performed in grass carp.

FcRγ and light chains of immunoglobulin were selected for amino acid sequence alignment and phylogenetic analysis by using the ClustalW 2.0 software (Larkin et al., 2007) and the neighbour-joining (NJ) method with the MEGA 4.0 software (Kumar et al., 2008), respectively.

2.4. Expression analysis of the FcRγ gene in different tissues

Three grass carp, aged one year and weighing about 100 g, were obtained from Nanhui fish farm in Shanghai, China. All fish were healthy and acclimatized for 14 days in a circulating water system at 23 °C. Liver, trunk kidney, head kidney, spleen, intestine, gill, skin and blood samples were collected, and total RNA was isolated from these tissues using TRIZOL (invitrogen). All RNA samples were treated with RNase–free DNase (Takara, Japan).