Aquaculture 318 (2011) 54–60

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aqua-online

Identification of genes involved in immune response of Atlantic salmon (*Salmo salar*) to IPN virus infection, using expressed sequence tag (EST) analysis

Victoria Cepeda [a], Catalina Cofre [a], Ruth González [a], Simon MacKenzie [b], Rodrigo Vidal [a,*]

[a] Laboratory of Molecular Ecology, Genomics and Evolutionary Studies, Department of Biology, Faculty of Chemistry and Biology, Universidad de Santiago de Chile, Avenida Libertador Bernardo O'Higgins 3363 Estación Central, Santiago, Chile
[b] Unitat de Fisiologia Animal, Departament de Biologia Cellular, de Fisiologia i d'Immunologia, Universitat Autonoma de Barcelona (UAB), Barcelona, Spain

ARTICLE INFO

Article history:
Received 11 November 2010
Received in revised form 25 April 2011
Accepted 27 April 2011
Available online 4 May 2011

Keywords:
cDNA library
Expressed sequence tag (EST)
Infectious pancreatic necrosis virus (IPNV)
Immune-related genes
Atlantic salmon

ABSTRACT

Atlantic salmon is one of the most important aquaculture fish species in the world. A cDNA library was constructed from splenic leukocytes of Atlantic salmon challenged with infectious pancreatic necrosis virus, and sequenced to collect genomic information and identify genes involved in immune defense response. Sequencing of 1360 clones yielded 1043 high quality ESTs (expressed sequence tags) and these ESTs were assembled into 823 unigenes. BLAST analysis revealed that 203 unigenes were novel genes that had no significant matches to any protein sequences in the public databases. The remaining 586 unigenes were closely matched to the known genes or sequences deposited in public databases and 34 unigenes matched to unknown genes. Among them, GO and KEGG analyses identified several functional categories involved in the innate immune response, with immune response and cytokine activity being predominant. These results provided a useful resource for gene discovery for further research of this species.

1. Introduction

The Atlantic salmon (*Salmo salar*) is a very valuable commercial salmonid species with well-established markets in Europe, Japan and the United States. The major producers are Norway, Chile, the United Kingdom, and Canada, though Chile and Norway account for over 71% of aquaculture production (Asche and Bjorndal, 2011). As with other aquaculture species, intensive aquaculture of Atlantic salmon often faces disease problems especially in early life stages; this can limit stable production of the species (Hellberg et al., 2010). The continued success of this industry in the highly competitive world market requires additional research and development activities aimed at improving health management techniques.

To efficiently manage disease problems in Atlantic salmon aquaculture, information on how they respond to pathogen exposure is needed. At present our understanding of the immune response of salmonid in general is very limited.

Infectious pancreatic necrosis (IPN) is considered the most serious viral disease in terms of its impact on Atlantic salmon (*S. salar*) production in the European Union and Chile (Asche and Bjorndal, 2011. The aetiological agent, infectious pancreatic necrosis virus is a RNA virus member of the family Birnaviridae, which has a genome consisting of two linear double stranded RNA (dsRNA) segments (A and B) (Dobos, 1995; Roberts and Pearson, 2005). IPNV is globally widespread causing substantial mortality in salmonid fry. The disease also causes mortalities in Atlantic salmon post-smolts, with most incidences of disease occurring several weeks after transfer to sea water (Bowden et al., 2002).

Although vaccines against IPN based on inactivated virus or recombinant structural viral proteins are commercially available, protection is variable and not complete (Robertsen, 2011). Therefore, the application of genomic technologies to thoroughly characterize the immune response of Atlantic salmon to IPNV could improve our knowledge of this fish and provide for long-term enhancements in aquaculture production.

The identification of the genes expressed within the cells of a given tissue is a basic step in the determination of the gene function and the analysis of tissue physiology. One efficient approach to the characterization of gene transcripts is the expressed sequence tag (EST). Random sequencing through EST projects has played a key role in identifying genes involved in the immune system of commercially important fish species (Feng et al., 2009; He et al., 2004; Rhodes et al., 2009; Tsoi et al., 2004). Moreover, the availability of many Atlantic salmon EST libraries (Adzhubei et al., 2007; Koop et al., 2008; Rise et al., 2004) provides valuable resources for comparative analysis. Currently, the information of immune response of Atlantic salmon to IPNV is limited to studies utilizing a limited number of candidate genes analyzed by real time PCR technique (Haugland et al., 2005; Larsen et al., 2004; McBeath et al., 2007).

In this work we generated ESTs from one spleen normalized cDNA library to improve our understanding of the immune response of Atlantic salmon to IPNV infection. This information can be used to identify the molecular mechanism for disease resistance of cultured Atlantic salmon.

* Corresponding author. Tel.: +56 2 7181120; fax: +56 2 6812108.
E-mail address: ruben.vidal@usach.cl (R. Vidal).

0044-8486/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.aquaculture.2011.04.045

2. Materials and methods

2.1. Viruses

The virus was obtained inoculating CHSE-214 cell monolayers (cat # CRL-1681, ATCC, Manassas, VA, USA) with SP and HE strains (originated from clinical cases of IPN) (Espinoza et al., 1985; Isshiki, et al., 2001) at a multiplicity of infection of 0.1 pfu/mL, in minimum essential medium (MEM, Gibco Invitrogen Co., Grand Island, NY, USA) supplemented with 2% fetal bovine serum (FBS, Gibco Invitrogen Co.) and antibiotics (100 IU/L penicillin and 100 mg/mL streptomycin, Gibco Invitrogen Co.) (MEM 2%). After 1 h of adsorption, the medium was removed containing excess virus and MEM 2% was added and allowed to passed the infection at 15 °C until the monolayer was observed with evident cytopathic effect (CPE) (48 to 72 h post infection). The viral inoculated was titrated by plaque assay on CHSE-214 cells to determine the number of plaque-forming units per mL of inoculum (pfu/mL) (Jashes et al., 1996). The $TCID_{50}$ was calculated by the Kärber method (Kärber, 1931).

2.2. IPNV challenge and tissue preparation

Thirty-two *S. salar* of approximately 40 g were acclimated for 7 days and fed daily with an amount of commercial feed pellet corresponding to 1% of their body weight for 1 h, after which the water was changed by using the technique of trap, to avoid toxic levels of ammonium that could affect gene expression. Eight groups of healthy fish (with normal feeding and behavior) were placed in eight individual tanks. To rule out the viral presence, blood was taken from the caudal vein of each fish to detect the presence of IPNV by RT-PCR (Rodriguez Saint-Jean et al., 2001).

The salmonids were subjected to two challenge tests, in the first instance being infected by immersion with 1×10^5 pfu/mL of virus, considered as day zero when the challenge test started. The water temperature was 11 ± 1 °C.

Thirteen days post first challenge test, the fish was reinfected with a dose of 2×10^4 pfu/mL virus. Seven days post-second infection liver, head kidney, gills and muscle were collected and stored at −80 °C immediately. In addition, spleen was collected to perfuse in order to obtain splenic leukocytes, resulting in 70×10^6 cells, which were stored at −20 °C in RNA Protect (cat # 76526, Qiagen GmbH, Hilden, Germany) for the construction of the cDNA library. In parallel, viral presence was assessed by analysis of IPNV detection by RT-PCR (Rodriguez Saint-Jean et al., 2001). Moreover, IPN diagnosis was established by means of gross pathology, histopathology and verified with immunohistochemistry (Evensen and Rimstad, 1990).

2.3. RNA extraction and cDNA library construction

Total RNA from leukocytes was isolated from each sample using RNeasy Lipid tissue mini kit (Qiagen, Germany) according to the manufacturer's instructions. A cDNA library was constructed using the library construction kit (Creator SMART, Clontech, USA) and normalized with the Trimmer-Direct cDNA normalization kit according to the manufacturer's protocol (Evrogen, Russia). We determined the number of primary recombinants and the average size of inserts from 100 clones by PCR using T7 vector primers to test the quality of the constructed cDNA library.

2.4. Sequences, annotation and data analysis

A total of 1360 individual colonies were sequenced using the ET dye terminator method (Amersham, USA) and the primer T7 (5′-TAA TAC GAC TCA CTA TAG GG-3′), providing sequences from the upstream end of the cDNA clones. The sequencing products were analyzed on ABI 3730 XL automated sequencers (Applied Biosystems, Framingham, MA, USA). The raw sequences were analyzed with the Phred software for base calling and assignment of quality scores (Ewing et al., 1998; Ewing and Green, 1998). Vector, contaminant and adaptor sequences were trimmed using the SeqClean (Chen et al., 2007) and Lucy (Chou and Holmes, 2001) software, while repetitive elements were masked with RepeatMasker (Tarailo-Graovac and Chen, 2009). Short sequences (<100 bp in length) were removed using our customized Perl program. High quality ESTs were submitted to the clustering and assembly using CAP3 (Huang and Madan, 1999). The resulting contigs and singletons (unigenes) were compared with the nonredundant (nr) protein database using BLASTX (Altschul et al., 1997), at the National Center for Biotechnology Information (NCBI; Bethesda, MD, USA). *E* values less than 10^{-5} were considered significant. The unmatched unigenes were searched in the NCBI nucleotide (nt) database using BLASTN (Altschul et al., 1990) with a cutoff of $1e^{-10}$. Unigenes that do not matched to the NCBI nr and nt databases were submitted to dbEST and analyzed with the InterProScan program (Zdobnov and Apweiler, 2001) to assign putative functions.

The totality of unigenes was submitted for Gene Ontology (GO) annotation to the online version of the BLAST2GO v1 program (Gotz et al., 2008) (April, 2009). BLAST2GO allows the selection of a significance level for the False Discovery Rate (FDR) which was used as a cut-off at 0.05% probability level. The data presented herein represent the level 2 analysis, illustrating general functional categories. Metabolic pathway was performed using the tools supplied by the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) (April, 2009). The data were processed using the bi-directional best hit method (forward and reverse reads) to assign orthologs.

3. Results

A total of 1.360 randomly selected clones were sequenced from the 5′ end, resulting in the characterization of 1.043 (76.69%) ESTs that were longer than 100 bp after the elimination of vector sequences (Table 1). A larger fraction of the sequences ranged from 400 to 700 bp and the average readable EST length on which the following analysis was based was about 539.6 bp. These ESTs were assembled into 192 contigs and 631 singletons. The average length of a readable contig was 565.8 bp and the average GC content from this library was 46.03%. The high quality ESTs were deposited in the GenBank (accession codes GW914325–GW915367). The distribution of contig sizes varied from two-to three (99.6%) to four-to-five (0.3%) ESTs sequences by cluster.

All the unigenes were used to search the public databases using BLASTX and BLASTN. The multiple annotations provided greater assurance about gene description and frequency of annotation than in a single database. We found that there were 586 (56.18%) unigenes that showed homology with genes of known function (mainly Atlantic salmon), 34 (3.25%) matched to unknown genes (hypothetical proteins) and the remaining 203 (19.46%) had no significant matches to known sequences in public database. It is not clear at present why these genes were transcribed and what their possible roles were.

Table 1
Summary statistics of ESTs generated from *Salmo salar*.

Description	Value
Total number of clones sequenced	1,360
High-quality ESTs	1,043
Average length of high-quality ESTs (bp)	539.6
Number of contigs	192
Number of singletons	631
Total number of unigenes	823
Average length of unigenes (bp)	565.8