J. Fish. Sci. Tech. 6(1), 41−44, 2003

JOURNAL OF
FISHERIES SCIENCE
AND TECHNOLOGY

〈Notes〉

Initial Transcriptome Profile of Rainbow Trout (*Oncorhynchus mykiss*) Liver

Soonhag Kim*

National Genome Research Institute, Korea National Institute of Health, Seoul 122-701, Korea

(Received December 2002, Accepted March 2003)

Ninety nine random complementary DNA clones from rainbow trout (*Oncorhynchus mykiss*) liver cDNA library were partially sequenced as one approach to analyze the transcribed sequences of its genome. Of the sequence generated, 64.0% of the ESTs were represented by 29 known genes. Thirty six clones of the unknown gene products potentially represent 31 unique genes. Serum albumin (16.1%) was the most abundant in the liver. The structural genes in the liver (19%) were the highly expressed functional category. This research is helpful to understand tissue specific gene expression profile and basic relationship between tissue and functional categories of the genes.

Key words: EST, Expression, Gene, Liver, Transcriptome, Trout, *Oncorhynchus mykiss*

Rainbow trout (*Oncorhynchus mykiss*) is one of the most important seafood protein sources in the world. A great deal of research has been focused on the ecology, physiology, and biochemistry of rainbow trout. Nevertheless, molecular genetic studies of rainbow trout, such as the Expressed Sequence Tags (EST), have not been extensively studied. The characteristics of an organism are determined by the genes expressed within it (Velculescu et al., 1995). Several important advances such as serial analysis of gene expression (SAGE) (Velculescu et al., 1995) and EST analysis (Adams et al., 1991) have provided technological possibilities for examination of gene expression at the whole genome level.

ESTs are partial cDNA sequences corresponding to mRNAs and generated from randomly selected library clones (Adams et al., 1991). ESTs are also generated by single pass sequencing of both/either the 5′ and/or the 3′ ends of cDNAs contained in selec-ted cDNA libraries. Characterization of ESTs is a convenient and rapid mechanism for discovery of new genes in various organisms. Additionally, it provides resources for a detailed profile of genes expressed in a tissue or cell type and development of gene chip technology.

In October 2002, data-base EST (dbEST) in NCBI contained 13,209,117 entries. Among the 330 species present, the most extensively studied species were the human, *Homo sapiens* (43.9%) and mouse, *Mus musculus* and *Mus domesticus* (24.4%). The number of fish ESTs in dbEST, excluding the zebrafish, *Danio rerio* (1.1%), was a total of 0.13% from various fish species: Channel catfish, *Ictalurus punctatus,* Japanese medaka, *Oryzias latipes;* Winter flounder, *Pleuronectes americanus;* Common carp, *Cyprinus carpio;* Nile tilapia, *Oreochromis niloticus;* Rainbow trout, *O. mykiss.*

The most attractive aspect of functional genomics is its ability to determine gene expression changes in a global view. ESTs have been an important biological component of functional genomics. For instance, ESTs are used in hybridization studies where the intensity of hybridization signals varies with the quantity of messenger RNAs harvested under different physiological conditions. The simultaneous hybridization of a large number of clones permits information to be obtained on the expression of numerous genes. This is equivalent to running tens of thousands of Northern blots or equivalent to many rounds of differential display

*Corresponding author: kimsoonhag@hotmail.com

(Liang and Pardee, 1992). Obviously, a necessary step for functional genomics is the availability of large numbers of ESTs. In this study, we describe intial ESTs information from rainbow trout liver cDNA library.

The liver cDNA library, prepared in λ ZapII vectors, was kindly provided by Dr. Joe Brunelli of Washington State University. The amplified library had a titer of approximately ten million plaque forming units (pfu) per microliter. About 10 million pfu of this library were used for the conversion of the λ libraries into plasmid libraries by using the mass *in vivo* excision procedure (Stratagene La Jolla, CA, USA). The XL1-blue MRF′ bacteria were infected with 10 million of phage (10 cells : 1 phage), and also infected with the ExAssist helper phage at a 1:1 helper phage-to-cell ratio. Cells were incubated at 37℃ for phage absorption. The infected cells were then grown in 20 mℓ LB broth for 3 hours in a shaker at 37℃. During this time, phagemids were excised and secreted into medium. The cells were then heated to 70℃ for 20 minutes and removed by centrifugation. The excised phagemids in the supernatant were titered. SOLR cells (Stratagene, La Jolla, CA, USA) were infected with 10^7 pfu of phagemids and incubated at 37℃ for 15 min. The cells were then plated onto LB plates containing ampicillin at 100 μg/mℓ. Colonies were picked and cultured in LB liquid culture for plasmid preparation.

Double-stranded plasmid DNA templates were prepared by using the Qiagen Spin Column Miniplasmid kits (Qiagen, Valencia, CA, USA). Since the cDNAs were directionally cloned into the λ ZapII vectors, single run sequencing on the 5′ ends of each cDNA clone was performed by the Sanger dideoxy termination method (Sanger et al., 1977) utilizing the BigDyeTerminator chemistry (Perkin Elmer, Foster City, CA, USA) and a T3 primer (5′ AATTAACCCTCACTAAAGGG 3′). Sequencing products were analyzed on an ABI Prism™ 310 automated DNA sequencer (Perkin Elmer, Foster City, CA, USA). Mini-preparation plasmid DNA (1 μℓ, about 200−500 ng) was used for all sequencing reactions. The profiles for cycling were: 96℃ for 30 sec., 50℃ for 15 sec., 60℃ for 4 min. for 25 cycles. An initial two minutes of extra denaturation at 96℃ were always used.

Classification of the clones was based on their similarity to NCBI dbEST entries. In general, 400 to 450 nucleotides following the EcoRI adapter sequence were taken for data base comparisions with the BLASTN and BLASTX servers (NCBI, Bethesda, MD). Matches were considered to be significant only when the probability (P) was less than 0.001. Comparisons and alignments of the EST sequences were performed with the Gene Tool 1.0 program package (Edmonton, AB, Canada).

Single-pass sequencing from the 5′ end of the cDNA was performed for 99 clones from rainbow trout non-normalized liver library. The vector-derived sequence and ambiguous sequences were removed from the collected EST sequences prior to computer analyses. The average length of edited ESTs obtained was thus 420 bp. Among sequence generated from rainbow trout liver cDNA library, 64 % of ESTs had homologies with known genes through NCBI Blast search (Table 1). Multiple clones were sequenced for 8 genes ranging from 2−17 clones per gene. They represent products of 29 genes. Thirty-six percent of ESTs were from the unknown genes representing the products of 34 genes. Comparison of the sequences of the unknown clones with each other revealed that one gene was represented by two independent cDNA clones and two genes by three cDNA clones, while the others are probably distinct genes.

Nine of 29 known genes from rainbow trout liver cDNA library have previously been reported in rainbow trout, encoding ferritin middle subunit, transferrin, complement factor Bf-1, -2, and C3, apolipoprotein A-I-1, G-protein, mitochondrion, and beta globin Currently, approximately 300 genes have been released in NCBI dbEST. Most of our ESTs are new discovery in rainbow trout. However, the high abundance of the major transcripts in the liver tissues necessitates construction of the subtracted cDNA library for discovery of the rainbow trout genes and for cataloguing the ESTs for functional genomics study in the rainbow trout.

The total number of genes in the human is believed to be approximately 26,000 to 38,000 and that of a fruit fly is about 13,600 (Paabo, 2001). In any given cell type, approximately 10,000 genes are expressed at levels of from 200,000 copies to 1 copy or less per cell (Patanjali et al., 1991). Here, several

Table 1. Rainbow trout *(Oncorhynchus mykiss)* liver ESTs

Clone #	Accession number	Most homologous to	Putative Identification	Overlap (bp)	Number of genes
RTL 1	BE669038	X52397.1	Serum albumin	418	16
RTL 27		AF042218.1	Apolipoprotein A-I-1	433	7
RTL 10		L24433.1	Complement component C3	355	4
RTL 3		D89083.1	Transferrin	413	3
RTL 2		S77386.1	Ferritin middle subunit	386	2
RTL 51	BE669039	S77332.1	Ferritin heavy subunit	202	2
RTL 13	BE669041	AB029337.1	glyceraldehyde-3-phosphate dehydrogenase	315	2
RTL 40	BE669045	X99385.1	Pentraxin	408	1
RTL 6	BE669040	AJ007505.1	Catalase	329	1
RTL 8		NC001717.1	Mitochondrion	334	1
RTL 34		AF089860.1	Complement factor Bf-2	344	1
RTL 9		AF089861.1	Complement factor Bf-1	253	1
RTL 22	BE669042	NM009147.1	SEC23A	191	1
RTL 25	BE669043	NM000098.1	Carnitine palmitoyltransferace II	142	1
RTL 3	BE669044	U55177.1	Carbonic anhydrase homologoue	183	1
RTL 42	BE669046	AB025356.1	Activin	379	1
RTL 48		U45967.1	G-protein(ras)proto-oncogene	78	1
RTL 49	BE677951	AF153429.1	CDC42 PROTEIN	140	1
RTL 52	BE669047	AJ236884.1	Retinol binding protein	180	1
RTL 53	BE669048	U48394.1	Ribosomal protein S5	153	1
RTL 72	BE669049	NM001011.1	Ribosomal protein S7	403	1
RTL 55	BE669050	AB023582.1	ATP synthase beta subunit	319	1
RTL 68	BE669051	MN000690.1	Aldehyde dehydrogenase 2	308	1
RTL 81	BE669052	NM016100.1	Acetyltransferase	147	1
RTL 85	BE669053	AF087432.1	Fibrinogen gamma chain precusor	41	1
RTL 87	BE669054	AF216873.1	Acetyl-CoA synthetase	147	1
RTL 90		D82926.1	Beta globin	409	1
RTL 98	BE669055	L47669.1	/translation elongation factor 1	101	1
RTL 101	BE669056	AB026119.1	Glutathione S-transferase	349	1
RTL 4	BE669058		Unknown		3
RTL 18	BE669062		Unknown		2
RTL 28	BE669065		Unknown		3
			Unknown		28
Total					99

genes account for the major transcriptional activities of the liver Eight genes were represented with more than one cDNA clone in the liver cDNA library as indicated in Table 1. Serum albumin was the most highly expressed gene (16.1%) in the liver cDNA library, followed by apolipoprotein A-I-1 (7.0%). Serum albumin is one of the most abundant proteins in salmon (*Salmo salar*) liver, representing 1.6 % of all clones in a liver cDNA library (Byrnes and Gannon, 1990).

Forty four percent of the ESTs identity and functions could not be determined. Rainbow trout liver known as ESTs are functionally classified into 6 groups based on predicted or known functions (Fig. 1). The structural genes (19%), such as apolipoprotein A1, complement component C3, ferritin middle subunit, ferritin heavy subunit, and beta globin, were the most highly expressed in the liver. The genes related to regulatory account for 17%, enzymes for 11%, transport/binding proteins for 9%, and ribosomal RNAs for 2%. Proteins categorized as micellaneous (not included in the categories of sturcture, enzymes, regulatory, ribosomal RNA, and transport/binding protein) account for 5% of the major transcriptional activity.

One of the advantages of EST analysis using non-normalized libraries is the ability to produce expression profiles. The gene expression profiles in

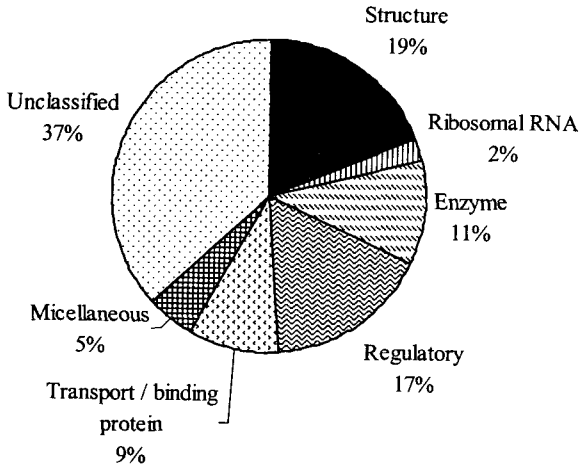

Fig. 1. Gene expression profiles in the rainbow trout liver as revealed by the expressed sequence tag (EST) analysis.

the liver provide valuable information for strength of promoters involved in transcription of these genes although gene copy numbers are not known. Assuming that a gene is present in a genome as a single copy, identification of strong and tissue-promoter may be useful in transgenic studies and genetic engineering. If the EST analysis is for the purpose of EST cataloguing for the development of bio-component, repeated sequencing of highly expressed genes is not desirable. Normalized cDNA libraries (Patanjali et al., 1991; Sasaki et al., 1994) or the subtracted cDNA libraries are therefore needed for characterization of a large number of unique ESTs. The subtracted cDNA libraries of rainbow trout liver were developed in our laboratory using the highly expressed genes found more than one cDNA clone as subtractants. Approximately 39% of 3,000 colonies randomly picked from the non-normalized liver cDNA library hybridized to the subtractants (data not shown). This approach has allowed as reducing redundancy in the identification of ESTs.

References

Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde, R.F. Moreno, A.R. Kerlavage, W.R. McCombie and J.C. Venter. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. Science, 252, 1651−1656.

Byrnes, L. and F. Gannon. 1990. Atlantic salmon (*Salmo salar*) serum albumin: cDNA sequence, evolution, and tissue expression. DNA Cell Biol., 9, 647−655.

Liang, P. and A.B. Pardee. 1992. Differential display of eukaryotic mRNA by means of the polymerase chain reaction. Science, 257, 967−971.

Paabo, S. 2001. The human genome and our view of ourselves. Science, 291, 1219−1220.

Patanjali, S.R., S. Parimoo and S.M. Weissman. 1991. Construction of a uniform-abundance (normalized) cDNA library. Proc. Nat'l. Acad. Sci. USA, 88, 1943−1947.

Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat'l. Acad. Sci. USA, 74, 5463−5467.

Sasaki, Y.F., D. Ayusawa and M. Oishi. 1994. Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. Nucl. Acids Res., 22, 987−992.

Velculescu, V.E., L. Zhang, B. Vogelstein and K.W. Kinzler. 1995. Serial analysis of gene expression. Science, 270, 484−487.