Korean J Biol Sci 6: 323-326, 2002

Initial Gene Expression Profile of Rainbow Trout (*Oncorhynchus mykiss*) Intestine

Soonhag Kim*

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

Key Words:
Gene expression
Functional genomics
Intestine
Transcriptome
Rainbow trout
Oncorhynchus mykiss

One hundred and three random complementary DNA clones representing rainbow trout intestine were partially sequenced as an approach to analyze the transcribed sequences of its genome. Of the sequences generated, 66.0% of the ESTs were represented by 40 known genes. Thirty-five clones of unknown gene products potentially represented 34 novel genes. The most abundantly represented messages were the 28S ribosomal protein (16.5%) and beta actin (5.8%). The genes involved in ribosome formation (18%) accounted for the major gene expression. Development of EST panels representing the genes expressed in a particular tissue will be useful in determining the role of these genes in normal function and in response to developmental, hormonal, environmental and physiological changes.

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 those using Expressed Sequence Tags (EST), have not been extensive. Most 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 generated by single pass sequencing of the 5' and/or the 3' ends of cDNAs contained in selected cDNA libraries. Characterization of ESTs is a convenient and rapid way to discover new genes in various tissues and organisms. Additionally, it provides resources for a detailed profile of genes expressed in a tissue or cell type and for the development of gene chip methodologies. As of October 2002, database 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 *M. domesticus* (24.4%). The number of fish ESTs in dbEST, excluding the zebrafish *Danio rerio* (1.1%), was a total of 0.13% from channel catfish *Ictalurus punctatus*, Japanese medaka *Oryzias latipes*, winter flounder *Pleuronectes americanus*, common carp *Cyprinus carpio*, nile tilapia *Oreochromis niloticus*, and rainbow trout *O. mykiss*.

ESTs have become an important biological component of functional genomics. The most attractive aspect of functional genomics is the ability to determine gene expression changes on a global perspective. For instance, ESTs are used in hybridization studies where the intensity of hybridization signals vary 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 many rounds of differential display (Liang and Pardee, 1992). Obviously, a requirement step for functional genomics is the availability of large numbers of ESTs. In this study, we describe ESTs from rainbow trout intestine cDNA libraries.

Materials and Methods

Template preparation

The intestine cDNA library, prepared in λ ZapII vectors, was kindly provided by Dr. Joe Brunelli of Washington State University. The amplified libraries had a titer of approximately 1×10^7 plaque forming units (pfu) per microliter. About 1×10^7 pfu of the library was used for the conversion of the λ library into a plasmid library by using the mass *in vivo* excision procedure (Stratagene, La Jolla, CA). Briefly, the XL1-blue MRF' bacteria were infected with 10 million library phage (10 cells : 1 phage), and co-infected with the ExAssist helper phage at a 1:1 helper phage-to-cell ratio. Cells were incubated at 37℃ for phage absorption. The

* Tel: 82-2-380-1531, Fax: 82-2-354-1078
E-mail: kimsoonhag@hotmail.com

infected cells were then grown in 20 mL LB broth for 3 h in a shaker at 37℃. During this time, phagemids were excised and secreted into the medium. The cells were then heated to 70℃ for 20 min and removed by centrifugation. The excised phagemids in the supernatant were titered. SOLR cells (Stratagene, La Jolla, CA) 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.

DNA sequencing and data analysis

Double-stranded plasmid DNA templates were prepared using Qiagen Spin Column Mini-plasmid kits (Qiagen, Valencia, CA). Because 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 BigDyeTerminator chemistry (Perkin Elmer, Foster City, CA) and a T3 primer (5' AATTAACCCTCACTAAAGGG 3'). Sequencing products were analyzed on an ABI Prism™ 310 automated DNA sequencer (Perkin Elmer, Foster City, CA). Mini-preparation plasmid DNA (1 μL, about 200-500 ng) was used for all sequencing reactions. After an initial 2 min denaturation, the profiles for cycling were: 96℃ for 30 sec, 50 ℃ for 15 sec, 60℃ for 4 min for 25 cycles.

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 comparison utilizing 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).

Results

Single-pass sequencing from the 5' end of the cDNA was performed for 103 clones from rainbow trout non-normalized intestine library. The vector-derived sequence

Table 1. Rainbow trout intestine ESTs

Clone #	Accession number	Most homologous to	Putative identification	Overlap (bp)	Number of gene
RTI 4		U34341.1	28S rRNA	430	17
RTI 16		AF012125.1	Beta actin	357	6
RTI 13		NC001717.1	Mitochodrion	419	5
RTI 9		AF125208.1	Cytochrome b	370	2
RTI 11	BE669055	AF184170.1	Elongation factor 1	343	2
RTI 96		U20946.1	MHC class II beta	401	2
RTI 3	BG360520	NM014761.1	KIAA0174	303	1
RTI 5	BG360521	AJ005016.1	ATP binding cassette (ABC) trasporter	258	1
RTI 71	BG360522	NM001916.1	Cytochrome c1	214	1
RTI 17	BG360523	U60205.1	Methyl sterol oxidase	107	1
RTI 18		U90321.1	Cathepsin D	359	1
RTI 37	BG360524	X99730.1	Cathepsin L-like cysteine protease	75	1
RTI 45	BG360525	NM001814.1	Cathepsin C	44	1
RTI 89	BG360526	NM001908.1	Cathepsin B	62	1
RTI 27	BG360527	AB027708.1	chaperonin containing T complex (CCT)	275	1
RTI 34	BG360528	U25632.1	CCT-2	48	1
RTI 35	BG360529	AL031320.6	Human clone RP1 20N2	84	1
RTI 42	BG360530	AF231347.1	Adenine nucleotide translocase 2	101	1
RTI 46	BG360531	B10364.1	Calmodulin	314	1
RTI 48		D86625.1	Ferritin H1	406	1
RTI 50	BG360532	AF074094.1	5' external transcribed spacer	283	1
RTI 51	BG360533	M85235.1	L7 ribosomal protein	157	1
RTI 52	BG360534	AF157110.1	L plasitin	54	1
RTI 74	BG360535	AB014875.1	T plastin	66	1
RTI 54	BG360536	NM006854.2	KDEL receptor 2	106	1
RTI 56	BG360537	U29159.1	Polyubiquitin	107	1
RTI 61	BG360538	NM000520.1	Hexosaminidase A	57	1
RTI 63	BG360539	Z29555.2	L3 ribosomal protein	391	1
RTI 70	BE669047	AJ236884.1	Retinol binding protein	347	1
RTI 72	BG360540	U25721.1	Secreted acidic cysteine rich glycoprotein	402	1
RTI 80	BG360541	X92522.1	Type II cytokeratin	346	1
RTI 81	BG360542	D86957.1	Septin 5	41	1
RTI 82	BG360543	AF025803.1	Cyclophilin 1	185	1
RTI 84	BG360544	NM014860.1	KIAA0764	51	1
RTI 85	BG360545	AF015811.1	Lysophosphatidic acid acyltransferase	173	1
RTI 91	BG360546	NM007990.1	Monoclonal nonspecific suppresor factor	61	1
RTI 94	BG360547	X73278.1	Dipeptidylpeptidase IV	136	1
RTI 95	BG360548	S66606.1	prolactin II	40	1
RTI 106	BG360549	NM004155.1	Protease inhibitor 9	26	1
RTI 19	BG360550	NM017052.1	Sorbitol dehydrogenase	44	1
RTI 32	BG360561		Unknown		2
			Unknown		**34**
Total					103

The blank in the accession number indicates that genes had been reported previously in rainbow trout in NCBI dbEST. The bold letter **Unknown** represents singleton unknown ESTs (the e-value of sequence similarity > 0.001) and were omitted from the list (Accession number from BG360551 to BG360584).

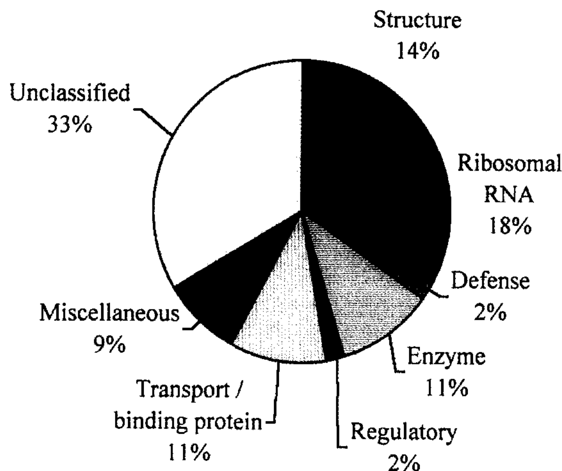

Fig. 1. Gene expression profiles in the rainbow trout intestine as revealed by EST analysis.

and ambiguous sequences were removed from the collected EST sequences prior to computer analyses. The average length of edited ESTs obtained was approximately 420 bp. Sixty-eight clones from the rainbow trout intestine cDNA library were from known genes and 35 clones were from unknown genes. Redundant clones were identified for 7 genes ranging from 2-17 clones per gene (Table 1). Thus, the 68 known ESTs represented gene products from 40 genes and the 35 clones of unknown genes represented products of 34 genes. Only one sequence from the unknown clones was represented by two independent cDNA clones. Ten of 40 known genes from rainbow trout intestine cDNA library have been reported previously in rainbow trout, encoding 28S rRNA, cytochrome b, mitochondrion, beta actin, cathepsin D, ferritin H1, secreted acidic cystein rich glycoprotein, type II cytokeratin, prolactin II, and MHC class II beta.

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 genes account for the major transcriptional activities of the intestine. Seven genes were represented with more than one cDNA clone in the intestine cDNA library as indicated in Table 1. The most abundant were 28S ribosomal protein (16.5%) and beta actin (5.8%). We classified the 68 clones from rainbow trout intestine cDNA library into 7 functional categories based on predicted or known functions through sequence similarity of annotated sequences (Fig. 1). The genes involved in ribosomal structure (18%) accounted for the major gene expression in the intestine. These genes included 28S and L7 and L3 ribosomal proteins. The genes involved in structural functions accounted for 14%, enzymes for transport/binding proteins for 11%, regulatory proteins for 2%, and defense for 2%. Proteins categorized as "miscellaneous" (not included in the categories of structure, enzymes, regulatory, defense, and binding protein) accounted for 9% of the major transcriptional activity. Forty four percent of the ESTs identity and functions could not be determined. The gene diversity was the highest in genes involved in enzymes in intestine representing 11 different genes.

Discussion

In an attempt to discover novel genes and analyze gene expression profiles from rainbow trout intestine, we randomly cloned and sequenced about 103 ESTs from non-normalized intestine cDNA library. Seventy-four unique gene products were found in intestine. Among the genes found 10 had already been reported from rainbow trout. Currently, approximately 300 rainbow trout genes have been released in NCBI dbEST. Most of our ESTs are new discovery in 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). However, high abundance of the major transcripts in the intestine tissue necessitates the construction of a subtracted cDNA library for discovery of rainbow trout genes and for cataloguing the ESTs for functional genomic studies in rainbow trout.

One of the advantages of EST analysis using non-normalized libraries is the ability to produce gene expression profiles. Several genes highly expressed in rainbow trout intestine, such as ribosomal RNAs, beta actin, elongation factor, cytochrome b, and RTI 32, provide valuable information on strength of promoters involved in transcription of these genes even if gene copy numbers are not known. Assuming that a gene is present in the genome as a single copy, identification of strong and tissue specific promoters may be useful in transgenic studies and genetic engineering. If the EST analysis is for the purpose of EST cataloguing for 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 subtracted cDNA libraries are therefore needed for characterization of large numbers of unique ESTs. Subtracted cDNA libraries of rainbow trout intestine were developed in our laboratory using highly expressed genes whose cDNA clones were found more than once as subtractants. Approximately 23% of 3000 colonies randomly picked from the non-normalized intestine cDNA library hybridized to the subtractants (data not shown). This approach has allowed to reduce redundancy in the identification of ESTs.

References

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, and Venter JC (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252: 1651-1656.

Bishop JO, Morton JG, Rosebach M, and Richardson M (1974)

Three abundance classes in Hela cell messenger RNA. *Nature* 250: 199-204.
Liang P and Pardee AB (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 SR, Parimoo S, and Weissman SM (1991) Construction of a uniform-abundance (normalized) cDNA library. *Proc Natl Acad Sci USA* 88: 1943-1947.
Sanger F, Nicklen S, and Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467.
Sasaki YF, Ayusawa D, and Oishi M (1994) Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. *Nucleic Acids Res* 22: 987-992.
Velculescu VE, Zhang L, Vogelstein B, and Kinzler KW (1995) Serial analysis of gene expression. *Science* 270: 484-487.

[Received October 11, 2002; accepted November 12, 2002]