

## Sequences of 596 cDNA Clones (565,977 bp) of Japanese Flounder (*Paralichthys olivaceus*) Leukocytes Infected with Hirame Rhabdovirus

Takashi Aoki,<sup>1,\*</sup> Bo-Hye Nam,<sup>1</sup> Ikuo Hirono,<sup>1</sup> and Eiichi Yamamoto<sup>2</sup>

<sup>1</sup>Laboratory of Genetics and Biochemistry, Department of Aquatic Biosciences, Tokyo University of Fisheries, Konan 4-5-7, Minato, Tokyo 108-8477, Japan

<sup>2</sup>Tottori Prefectural Fisheries Experimental Station, Sakaiminato, Tottori 689-0602, Japan

**Abstract:** We have partially sequenced 785 sequences of 596 independent complementary DNA clones isolated from a cDNA library of Japanese flounder leukocytes infected with hirame rhabdovirus. These sequences consist of a total of 565,977 base pairs. The average size of the sequenced lengths was 721 bp. Of 596 clones, 386 (64.8%) were identified as previously reported genes by the BLASTN and BLASTX programs. About 30% of the identified clones could be recognized by only the BLASTX program. A total of 251 distinct genes were identified, and 181 of these genes are the first such genes reported from the teleostei. Approximately 27% of the identified Japanese flounder genes appear to be associated with cell division, cell structure or motility, and basic energy metabolism, 29% with gene or protein expression, and 17% with cell signaling, cell communication, and cell or organism defense. The most frequently identified expressed sequence tags of leukocytes of Japanese flounder were gelatinase b and ribosomal protein L23, which both had 1.34% prevalence.

**Key words:** Japanese flounder, *Paralichthys olivaceus*, cDNA, EST, leukocytes

### INTRODUCTION

Sequencing of complementary DNA libraries to generate large numbers of partial cDNA sequences or expressed sequence tags (ESTs) has been shown to be a powerful tool to identify homologues of reported genes of different animals and novel genes, to evaluate the gene expression profile in a given tissue, and to compare transcription levels of many genes between different tissues or cells (Adams et al., 1995; Banfi et al., 1996; Boguski, 1996; Okubo and Matsubara,

1997). The EST approach was described almost simultaneously in 1991 by Adams et al. (1991) and Okubo et al. (1991, 1992). Before the EST approach was described, Constanzo et al. (1983) and Wilcox et al. (1991) demonstrated the utility of randomly selected cDNA sequences for ascertaining the associations between genes and their functions. The database with the highest rate of growth has been a division of GenBank called dbEST, the EST database (Boguski et al., 1993). At present, more than 1.8 million ESTs derived from cDNA libraries of many organisms are known ([www.ncbi.nlm.nih.gov/dbEST](http://www.ncbi.nlm.nih.gov/dbEST)).

A cloned population of Japanese flounder, *Paralichthys olivaceus*, has been established (Yamamoto, 1995). Such a

Received February 12, 1999; accepted April 13, 1999

\*Corresponding author; telephone +81-3-5463-0556; fax +81-3-5463-0690; e-mail [aoki@tokyo-u-fish.ac.jp](mailto:aoki@tokyo-u-fish.ac.jp).

cloned population is very useful for molecular biological studies and genome analysis. Moreover, Japanese flounder is one of the most important food fish of both the fishing and aquaculture industries in Japan.

Leukocytes are important cells for the specific and non-specific immune systems. Several types of leukocyte cells express many kinds of biodefense-related peptides and proteins, including specific and nonspecific antimicrobial agents, and activators and regulators of the immune system. The cellular and protein levels of the leukocyte-related immune system of fish are of interest. However, there is not currently enough information on the genes and molecular mechanisms of fish biodefense-related molecules to understand the fish immune system, although molecular evolutionary studies of fish MHC, T cell receptor  $\alpha$  and  $\beta$ , and immunoglobulin genes have been performed.

Recently, we reported ESTs from the liver and spleen of apparently healthy Japanese flounder, from leukocytes infected with hirame rhabdovirus (Kimura et al., 1986) of Japanese flounder, and from the liver of medaka *Oryzias latipes* (Hirono and Aoki, 1997, Inoue et al., 1997, Nam et al., 1999). We have found many homologues of known genes from these ESTs. Most of these homologues were the first such genes cloned from fish. In this study, we sequenced and characterized further cDNA clones from a Japanese flounder cDNA library of leukocytes infected with hirame rhabdovirus because one of our interests is analysis of biodefense-related genes.

## MATERIALS AND METHODS

### Complementary DNA Library

We used the previously reported cDNA library of cloned Japanese flounder leukocytes infected with hirame rhabdovirus (Nam et al., 1999). The independent number of clones in this cDNA library was  $1.0 \times 10^7$ .

### Plasmid Preparation and DNA Sequencing

The template DNAs to be used for DNA sequencing were prepared by colony direct polymerase chain reaction (PCR) and a PCR product presequencing kit (Amersham, U.S.A.). Complementary DNA clones were sequenced using ThermoSequenase (Amersham, U.S.A.) with M13 forward or M13 reverse primers and an automated DNA sequencer LC4000 (Li-Cor, U.S.A.).

**Table 1.** Overview of This Study

Templates sequenced, <i>n</i>	596
ESTs, <i>n</i>	785
Putatively identified ESTs, <i>n</i>	439
Unidentified ESTs, <i>n</i>	346
Putatively identified clones, <i>n</i> (%)	386 (64.8)
Unidentified clones, <i>n</i> (%)	210 (35.2)
Putatively identified different genes, <i>n</i>	251
Total nucleotides in ESTs, bp	565,977
Average of sequenced length, bp	721

### Data Analysis

Each sequence was compared with all sequences available in GenBank using the BLAST 2.0 program (Altschul et al., 1990, 1997; [www.ncbi.nlm.nih.gov/dbEST](http://www.ncbi.nlm.nih.gov/dbEST)).

## RESULTS

We have partially sequenced 785 sequences of 596 independent cDNA clones isolated from a cDNA library of Japanese flounder leukocytes infected with hirame rhabdovirus (Table 1). These sequenced regions comprise 565,977 bp of DNA, with an average length of 721 bp. Of the 596 sequences, 386 (64.8%) demonstrated a significant similarity to previously reported genes according to the BLASTN and BLASTX programs. A total of 251 distinct genes were identified. Of this number, 181 genes (Table 2) had not previously been reported in the teleostei. Approximately 27% of the identified Japanese flounder genes appear to be associated with cell division, cell structure or motility, and basic energy metabolism, 29% with gene or protein expression, and 17% with cell signaling, cell communication, and cell or organism defense.

Analysis of the abundance of particular ESTs of Japanese flounder leukocytes infected with hirame rhabdovirus was determined using the current results combined with the data of Nam et al. (1999), who reported ESTs of 300 clones. The most abundant identified ESTs of Japanese flounder leukocytes infected with hirame rhabdovirus were gelatinase b and ribosomal protein L23 (Table 3). The genes expressed and the abundance of identified ESTs from the leukocytes and liver of Japanese flounder were different (Table 3). The most frequently identified ESTs of leukocytes of Japanese flounder were gelatinase b and ribosomal protein L23,

**Table 2.** List of Identified ESTs

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
<i>Cell division (7)*</i>							
WG2-11R	AU050193	Apoptosis-related protein TFAR15	<i>Homo sapiens</i>	AF022385	640	148	76.4
WB6-12F	AU050219	<i>c-fos</i>	<i>Tetraodon fluviatilis</i>	U53520	780	205	88.8
WA7-23F	AU050761	Cycline-like protein	<i>Arabidopsis thaliana</i>	AL021768	487	78aa	30.8
WF6-16R	AU050412	Helicase p54	<i>Xenopus laevis</i>	X92421	692	395	76.2
WB7-18F	AU050498	High mobility group protein HMG-1	<i>Oncorhynchus mykiss</i>	X02666	632	504	82.1
WG7-11R	AU050204	Ras-related protein rap 1a	<i>Homo sapiens</i>	A08691	598	534	87.3
WE3-8R	AU050051	14-3-3 Protein $\epsilon$	<i>Homo sapiens</i>	U28936	665	625	76.5
<i>Cell signaling/cell communication (35)†</i>							
WA8-20F	AU050568	Adenylyl cyclase type 9	<i>Mus musculus</i>	U30602	694	181	66.3
WD2-8R	AU050044	ADP-ribosylation factor	<i>Bos taurus</i>	J03275	611	383	82.5
WA12-23R	AU050765	ASM-like phosphodiesterase 3a	<i>Mus musculus</i>	Y08135	661	300	66.3
WF9-14F	AU050304	Calmodulin	<i>Oryzias latipes</i>	D10364	725	399	92.2
WB1-17F	AU050441	Carrier protein c2	<i>Caenorhabditis elegans</i>	X76116	926	230aa	53.0
WH10-9F	AU050122	CD18 (integrin $\beta$ 2)	<i>Gallus gallus</i>	X71786	713	112aa	39.3
WA4-22R	AU050689	CD20	<i>Homo sapiens</i>	X07203	621	111aa	27.9
WA8-9R	AU050077	CD42B	<i>Homo sapiens</i>	M22403	728	194aa	31.4
WH9-10F	AU050173	CD45	<i>Homo sapiens</i>	Y00062	877	250aa	61.6
WF11-9F	AU050108	CD49e (integrin $\alpha$ V)	<i>Xenopus laevis</i>	L10191	572	270	61.9
WE12-15R	AU050349	CD63	<i>Mus musculus</i>	S43511	1008	184aa	23.9
WG6-23F	AU050791	Channel-associated protein of synapse-110	<i>Homo sapiens</i>	U32376	721	378	76.5
WB7-12F	AU050220	Dopamine receptor D1B	<i>Gallus gallus</i>	B55886	860	56aa	41.1
WE6-8F	AU050054	Fc- $\gamma$ -1/ $\gamma$ -2 receptor	<i>Cavis porcellus</i>	D13691	801	99aa	28.3
WA11-20F	AU050570	G-CSF-receptor	<i>Mus musculus</i>	A34898	361	27aa	48.1
WG7-22F	AU050739	G-protein binding protein Rab27	<i>Homo sapiens</i>	U38654	922	589	77.1
WG1-23F	AU050788	G-protein coupled thrombin re- ceptor	<i>Rattus norvegicus</i>	M81642	470	284	60.2
WG5-22F	AU050738	Intercellular adhesion molecule-3 (ICAM-3, CD50)	<i>Homo sapiens</i>	S28904	989	79aa	31.6
WE9-23R	AU050783	Kupffer cell receptor	<i>Mus musculus</i>	D88577	726	135aa	36.3
WC3-9R	AU050091	Laminin receptor	<i>Mus musculus</i>	J02870	574	464	82.5
WE3-8F	AU050050	Leptin receptor	<i>Rattus novergicus</i>	D84126	801	61aa	36.1
WC6-17F	AU050458	Muscarinic acetylcholine receptor	<i>Gallus gallus</i>	A55019	939	50aa	36.0
WC6-9R	AU050092	Phospholipid hydroperoxide glu- tathione peroxidase	<i>Sus scrofa</i>	S80257	801	280	68.6
WC7-9F	AU050093	Platelet proteoglycan core pro- tein	<i>Homo sapiens</i>	X17042	801	64aa	35.9
WB7-22F	AU050700	Polymeric immunoglobulin re- ceptor	<i>Mus musculus</i>	U06431	750	128aa	32.8
WF6-21R	AU050668	Prostaglandin F2a receptor	<i>Mus musculus</i>	U26663	664	95aa	33.7
WH9-20F	AU050616	Protein arginine N-methyltrans- ferase 3	<i>Homo sapiens</i>	AF059531	356	56aa	41.1
WC7-21F	AU050641	Retinoid X receptor epsilon	<i>Danio rerio</i>	U29942	950	774	78.6
WD6-13F	AU050267	Semaphorin C	<i>Mus musculus</i>	X85992	889	242	55.6

Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WH9-20R	AU050617	Thrombin receptor	<i>Papio hamadryas</i>	AF028727	788	72aa	37.5
WA12-11R	AU050181	Thrombospondin	<i>Bos taurus</i>	AB005287	751	567	79.4
WB2-17R	AU050442	Thromboxane A2 receptor	<i>Bos taurus</i>	Q95125	806	101aa	52.5
WB11-23F	AU050770	Transforming growth factor- $\beta$	<i>Oncorhynchus mykiss</i>	X99303	878	280	80.7
WF1-9F	AU050106	Tumor necrosis factor receptor 1	<i>Mus musculus</i>	X59238	747	102aa	44.1
WB1-20R	AU050573	Tyrosine kinase	<i>Gallus gallus</i>	X59669	799	164aa	28.7
<i>Cell structure/motility (23)‡</i>							
WF11-19F	AU050559	Actin-related protein AcTr66B	<i>Drosophila melanogaster</i>	X71789	552	306	72.5
WB12-21R	AU050634	Actin-binding protein	<i>Homo sapiens</i>	X53416	908	700	77.6
WB8-21F	AU050631	Actin-related protein	<i>Homo sapiens</i>	AF006083	563	464	87.4
WC6-10F	AU050143	$\alpha$ 5-Tubulin	<i>Gallus gallus</i>	X08061	858	582	86.8
WF6-18F	AU050515	$\alpha$ -Actinin	<i>Caenorhabditis elegans</i>	AF016687	977	296aa	48.3
WF1-14F	AU050296	$\alpha$ -Tubulin	<i>Oncorhynchus keta</i>	X66973	725	602	89.4
WB2-11F	AU050182	$\beta$ -8-Tubulin	<i>Zea mays</i>	L10636	751	242	78.5
WD3-20R	AU050585	$\beta$ -Actin	<i>Sparus aurata</i>	X89920	594	516	94.2
WB11-10F	AU050140	$\beta$ -Tubulin	<i>Pleuronectes americanus</i>	X74492	993	465	73.8
WB1-8F	AU050031	Brain myosin II isoform	<i>Rattus norvegicus</i>	Z31721	798	542	84.7
WD2-12R	AU050234	Calponin	<i>Homo sapiens</i>	D83735	962	797	75.3
WC10-22F	AU050710	Capping protein $\alpha$ -2 isoform	<i>Gallus gallus</i>	M80589	817	816	75.6
WE9-12R	AU050254	Cytochrome <i>c</i> oxidase III	<i>Cyprinus carpio</i>	X17006	1101	550	74.9
WF8-8F	AU050064	Cytochrome oxidase 1 & 2, tRNA-Ser, tRNA-Asp	<i>Paralichthys olivaceus</i>	AB000675	699	523	98.1
WD5-15R	AU050335	Fibrinogen-like protein	<i>Mus musculus</i>	M16238	969	598	66.6
WC10-23F	AU050774	Microtubule-associated protein 1A	<i>Rattus norvegicus</i>	U05784	669	362	80.7
WE5-12R	AU050249	Myosin IB heavy chain	<i>Gallus gallus</i>	S31926	845	33aa	69.7
WH4-22F	AU050748	Myosin light chain	<i>Rattus norvegicus</i>	X05566	581	541	83.9
WE5-8R	AU050053	Myosin-IA	<i>Drosophila melanogaster</i>	U07595	564	406	59.9
WH11-23R	AU050798	Nuclear pore complex	<i>Rattus norvegicus</i>	L31840	1039	327	67.3
WA11-21R	AU050627	Synaptobrevin-like protein	<i>Mus musculus</i>	X96737	395	278	77.0
WH7-17R	AU050484	Tubulin-folding cofactor D	<i>Bos taurus</i>	U61233	841	196	73.5
WB8-20F	AU050575	Vinculin	<i>Homo sapiens</i>	M33308	679	434	77.2
<i>Cell/organism defense (19)§</i>							
WD8-16F	AU050396	$\beta_2$ -Microglobulin	<i>Oncorhynchus mykiss</i>	L63533	1046	360	63.3
WF7-19R	AU050557	Carbonic anhydrase V	<i>Mus musculus</i>	I59261	830	56aa	60.7
WF1-22F	AU050727	DNA repair protein	<i>Saccharomyces cerevisiae</i>	S51470	749	191aa	27.7
WH9-18F	AU050524	Heat shock protein 10	<i>Gallus gallus</i>	AF031309	581	304	71.4
WE8-16R	AU050407	HLA class II hitocompatibility antigen	<i>Homo sapiens</i>	P04233	621	81aa	35.8
WC7-15R	AU050326	Immunoglobulin- $\delta$	<i>Ictalurus punctatus</i>	U67437	858	50aa	40.0
WC1-22F	AU050704	Immunoglobulin- $\kappa$	<i>Homo sapiens</i>	X13078	627	105aa	37.1
WE8-14F	AU050293	Immunoglobulin light chain	<i>Oncorhynchus mykiss</i>	X68519	783	100aa	63.0
WB7-16R	AU050378	Interferon-induced 54 kDa protein	<i>Homo sapiens</i>	I59087	840	202aa	22.3
WC7-22F	AU050709	MHC class I antigen	<i>Triakis scyllium</i>	AF034342	760	198aa	40.4
WG2-16F	AU054150	MHC class IA	<i>Poecilia reticulata</i>	Z54086	1038	641	75.5
WC3-10F	AU051420	MHC class II B	<i>Morone saxatilis</i>	L33965	932	414	85.5

Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WH6-23R	AU057960	T cell receptor- $\delta$	<i>Raja eglanteria</i>	U75770	738	229aa	28.4
WD10-22R	AU050717	Thioredoxin	<i>Mus musculus</i>	X77585	501	248	69.8
WA2-8F	AU050023	Transferrin	<i>Paralichthys olivaceus</i>	D88801	582	470	98.1
WD1-19F	AU050545	Transferrin receptor	<i>Gallus gallus</i>	X55348	843	714	63.9
WA2-22F	AU050687	Transplantation antigen	<i>Salmo trutta</i>	U18488	1100	581	86.1
WB7-17R	AU050445	Tumor suppressor	<i>Homo sapiens</i>	AF061836	946	82aa	72.0
WD2-16R	AU050391	Xeroderma pigmentosum group C protein HHR23A	<i>Homo sapiens</i>	D21235	917	199	70.9
<i>Gene/protein expression (65)<sup>ll</sup></i>							
WC6-22F	AU050708	Acidic ribosomal protein PO	<i>Rattus norvegicus</i>	Z29530	798	727	81.6
WD8-14F	AU050283	ATP-dependent protease	<i>Homo sapiens</i>	U02389	771	311	74.6
WB9-16F	AU050379	bZIP transcription factor MafA	<i>Coturnix coturnix japonica</i>	AF034693	878	68aa	38.2
WB9-16R	AU050380	Carboxypeptidase N small sub-unit	<i>Homo sapiens</i>	X14329	834	434	75.6
WG10-20R	AU050609	CCAAT/Enhancer binding protein beta	<i>Homo sapiens</i>	S35336	857	48aa	85.4
WF1-22R	AU050728	Collagenase 4	<i>Xenopus laevis</i>	L76275	926	147aa	36.1
WC10-24R	AU050807	Cyclophilin-like protein CyP-60	<i>Homo sapiens</i>	U37219	666	602	72.3
WE2-15F	AU050340	Cys2/His2 zinc finger protein	<i>Rattus norvegicus</i>	U27186	973	38aa	55.3
WE4-16R	AU050404	Cystein proteinase	<i>Nicotiana tabacum</i>	Z13964	869	141aa	47.5
WG6-20F	AU050606	Dolichol monophosphate man-nose synthase	<i>Homo sapiens</i>	AF007875	942	705	73.3
WC9-18F	AU050500	Elongation factor 1 $\alpha$	<i>Danio rerio</i>	X77689	563	162	76.5
WG9-15R	AU050358	Gelatinase-b	<i>Cynops phyllorhogaster</i>	D82052	676	447	68.5
WA6-18R	AU050488	Glia maturation factor	<i>Homo sapiens</i>	AB001106	806	427	76.7
WE7-18F	AU050511	Glycyl tRNA synthetase	<i>Homo sapiens</i>	D30658	691	688	71.5
WD4-9F	AU050097	Lon protease-like protein	<i>Homo sapiens</i>	X76040	932	178aa	79.8
WB8-12F	AU050222	MC3 mRNA for protease	<i>Mus musculus</i>	X70303	831	671	78.1
WB5-12F	AU050218	Procollagen C-proteinase enhancer	<i>Homo sapiens</i>	L33799	1032	71aa	36.6
WE1-10F	AU050153	Proteasome $\alpha_2$ subunit	<i>Carassius auratus</i>	AB013342	400	115aa	55.7
WB4-18R	AU050496	Proteasome subunit LMP7	<i>Danio rerio</i>	AF032390	476	128	83.6
WE11-12R	AU050256	Proteasome subunit p31	<i>Homo sapiens</i>	D38047	916	247	73.7
WA6-17R	AU050432	Proteasome Z subunit	<i>Mus musculus</i>	D83585	630	169aa	56.2
WE1-20F	AU050589	Regulatory factor X-associated protein	<i>Homo sapiens</i>	Y12812	778	35aa	80.0
WG4-22F	AU050737	Ribosomal protein	<i>Mus musculus</i>	M85235	711	576	76.7
WD7-21F	AU050650	Ribosomal protein L10	<i>Rattus norvegicus</i>	X93352	620	580	81.7
WE4-11R	AU050190	Ribosomal protein L13	<i>Rattus norvegicus</i>	P41123	847	210aa	84.8
WG2-14R	AU050306	Ribosomal protein L17	<i>Homo sapiens</i>	X52839	471	359	83.8
WH10-18F	AU050525	Ribosomal protein L19	<i>Mus musculus</i>	M62952	792	583	83.2
WA8-10R	AU050130	Ribosomal protein L23	<i>Rattus norvegicus</i>	X58200	687	479	79.7
WF3-11R	AU050194	Ribosomal protein L28	<i>Homo sapiens</i>	S55915	800	359	72.1
WC10-12R	AU050229	Ribosomal protein L37a	<i>Homo sapiens</i>	L22154	777	266	85.0
WH7-20F	AU050614	Ribosomal protein L41	<i>Homo sapiens</i>	AF026844	361	91	92.3

Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WC11-17R	AU050463	Ribosomal protein L5	<i>Rattus norvegicus</i>	X06148	1101	775	78.7
WA5-23R	AU050760	Ribosomal protein L6	<i>Mus musculus</i>	X82987	432	422	70.6
WG9-10F	AU050164	Ribosomal protein L7	<i>Rattus norvegicus</i>	M17422	606	492	78.9
WB4-20R	AU050574	Ribosomal protein L7a	<i>Gallus gallus</i>	X62640	679	645	82.0
WC8-10F	AU050144	Ribosomal protein L8	<i>Xenopus laevis</i>	U00920	867	757	83.9
WA5-21R	AU050626	Ribosomal protein P2	<i>Bos taurus</i>	U17836	430	301	68.1
WG6-13R	AU050271	Ribosomal protein S11	<i>Xenopus laevis</i>	X78805	673	497	80.3
WH10-10F	AU050174	Ribosomal protein S12	<i>Tilapia nilotica</i>	U94500	822	460	88.7
WB9-17R	AU050448	Ribosomal protein S16	<i>Homo sapiens</i>	M60854	726	446	80.7
WF4-14R	AU050299	Ribosomal protein S18	<i>Homo sapiens</i>	X69150	522	459	83.4
WF6-17R	AU050476	Ribosomal protein S22	<i>Xenopus laevis</i>	M34706	477	408	78.9
WE9-20R	AU050593	Ribosomal protein S23	<i>Rattus norvegicus</i>	X77398	505	436	84.2
WE12-15F	AU050348	Ribosomal protein S26	<i>Rattus norvegicus</i>	X02414	676	356	80.9
WG4-15R	AU050355	Ribosomal protein S5	<i>Homo sapiens</i>	P97461	426	267	85.8
WH4-10F	AU050168	Ribosomal protein S8	<i>Mus musculus</i>	X73829	852	628	80.6
WA8-20R	AU050569	Ribosomal protein S9	<i>Rattus norvegicus</i>	X66370	678	592	81.4
WB3-9R	AU050083	RNA cyclase	<i>Drosophila melanogaster</i>	P56175	724	223aa	53.8
WA4-8F	AU050025	RNA polymerase II initiation factor $\alpha$	<i>Rattus norvegicus</i>	X65948	649	587	77.3
WB4-16F	AU050375	Scaffold attachment factor B	<i>Homo sapiens</i>	L43631	876	94aa	34.0
WA3-10F	AU050125	Seryl-tRNA synthetase	<i>Mus musculus</i>	M74012	570	705	71.1
WE10-14R	AU050295	Signal peptidase	<i>Homo sapiens</i>	2208369A	535	94aa	61.7
WD10-23F	AU050778	SnRNP associated protein B	<i>Mus musculus</i>	X17567	611	504	79.2
WE2-10F	AU050154	snRNP core protein sm D2	<i>Homo sapiens</i>	U15008	471	356	79.5
WC5-8F	AU050038	Suil translation initiation factor	<i>Homo sapiens</i>	L26247	801	361	78.9
WF12-21F	AU050671	TATA binding protein associated factor	<i>Homo sapiens</i>	AF040701	932	30aa	70.0
WG4-11R	AU050201	TFIIF- $\alpha$	<i>Drosophila melanogaster</i>	Q05193	742	42aa	52.2
WE1-21F	AU050654	Transcription enhance factor	<i>Mus musculus</i>	L13853	458	39aa	97.4
WF8-20R	AU050601	Transcriptional regulator Id2	<i>Gallus gallus</i>	AF049135	882	380	80.3
WA11-12F	AU050213	Translational controlled 40-kDa polypeptide	<i>Mus musculus</i>	X06406	1101	596	82.0
WD1-17F	AU050464	Translational regulatory factor $\alpha$ NAC	<i>Mus musculus</i>	U22151	462	434	79.7
WC10-16R	AU050388	Ubiquitin conjugating enzyme UBCH5	<i>Homo sapiens</i>	X78140	632	446	78.0
WD12-10F	AU050151	Ubiquitin I	<i>Gallus gallus</i>	M11100	767	273	90.8
WB4-17R	AU050444	Ubiquitin/ribosomal protein L40	<i>Rattus norvegicus</i>	X82636	483	398	86.7
WD11-16F	AU050398	Zinc finger protein ZNF183	<i>Homo sapiens</i>	X98253	919	180aa	83.9
<i>Metabolism (29)¶</i>							
WA12-17F	AU050439	17 $\beta$ -Hydroxysteroid dehydrogenase type 3	<i>Homo sapiens</i>	U05659	821	523	58.1
WF4-22F	AU050729	Adenylosuccinate synthetase	<i>Mus musculus</i>	L24554	927	646	70.4
WH12-22F	AU050756	Aryl hydrocarbon receptor nuclear translocator isoform b	<i>Oncorhynchus mykiss</i>	U73841	826	227	77.1

Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WC2-17F	AU050452	ATPase	<i>Schistosoma mansoni</i>	Z29947	965	673	61.4
WB12-21F	AU050633	ATPase 6	<i>Carassius auratus</i>	AB006953	673	23aa	95.6
WA4-10F	AU050126	C-SRC-kinase	<i>Homo sapiens</i>	X59932	839	782	76.6
WD7-12F	AU050238	Calcium-calmodulin-dependent protein kinase	<i>Rattus norvegicus</i>	L42810	680	42aa	97.6
WC5-22R	AU050707	Dihydroxyacetone acyltransferase	<i>Homo sapiens</i>	AF043937	949	269aa	35.3
WG5-23F	AU050790	Geranylgeranyltransferase type I	<i>Rattus norvegicus</i>	L24116	716	681	68.9
WF3-12F	AU050259	Glutamate dehydrogenase GLUD	<i>Mus musculus</i>	X57024	761	338	78.1
WH4-16F	AU050424	Glycerol 3-phosphate dehydrogenase	<i>Rattus norvegicus</i>	AB002558	832	503	72.0
WE1-22F	AU050718	H <sup>+</sup> -ATPase	<i>Bos taurus</i>	J03835	743	451	81.1
WD8-15R	AU050339	Leukotriene A-4	<i>Homo sapiens</i>	J02959	928	249aa	45.4
WB2-15R	AU050313	Lysozyme II	<i>Oncorhynchus mykiss</i>	X59491	831	258	72.5
WA4-17F	AU050428	Malate dehydrogenase	<i>Mus musculus</i>	M29462	640	282	73.0
WD7-19F	AU050548	Mitochondrial ATPase	<i>Salmo salar</i>	D84148	675	116	86.2
WH5-17F	AU050482	NADH-ubiquinone oxidoreductase subunit 1	<i>Salmo salar</i>	D84148	975	192aa	84.4
WC1-21F	AU050635	Neutrophil NADPH oxidase factor 1	<i>Mus musculus</i>	Q09014	623	112aa	46.4
WA1-21F	AU050621	Ornithine decarboxylase anti-enzyme	<i>Xenopus laevis</i>	D32141	506	295	66.8
WC2-8F	AU050036	Phospholipase C	<i>Homo sapiens</i>	X14034	801	623	61.8
WD8-20R	AU050587	Phospholipase D	<i>Rattus norvegicus</i>	AB000779	297	44aa	81.8
WD4-16F	AU050393	Prosaposin	<i>Gallus gallus</i>	AB003471	885	536	61.6
WD7-8R	AU050048	Protein kinase inhibitor P58	<i>Mus musculus</i>	U28423	671	476	67.6
WF2-21F	AU050664	Saposin protein A-D	<i>Homo sapiens</i>	M32221	962	527	60.9
WB6-9F	AU050085	Serine/threonine phosphatase 6	<i>Homo sapiens</i>	X92972	735	29aa	55.2
WE10-19F	AU050552	Steroid dehydrogenase	<i>Anas platyrhynchos</i>	AB009304	838	696	64.2
WC10-8F	AU050041	Transaldolase	<i>Homo sapiens</i>	L19437	801	392	65.6
WC4-16F	AU050383	Tyrosine and tryptophan hydroxylase	<i>Drosophila melanogaster</i>	M77518	818	404	73.3
WC3-8F	AU050037	Vitamin D3 25-hydroxylase	<i>Homo sapiens</i>	X59812	802	125aa	36.8
<i>Unclassified (73)</i>							
WE6-16F	AU050406	146-kDa Nuclear protein	<i>Xenopus laevis</i>	Y08997	986	178aa	89.3
WH4-21F	AU050679	Ubiquitin-activating enzyme	<i>Homo sapiens</i>	X56976	990	726	76.7
WC10-12F	AU050228	Ras-related GTPase ARP 1	<i>Rattus norvegicus</i>	X78603	782	202aa	74.3
WG6-19R	AU050562	bup = 5' of bmi-1 proviral insertion locus	<i>Mus musculus</i>	S54914	903	476	59.2
WB9-11F	AU050183	C32A3.3	<i>Caenorhabditis elegans</i>	Z48241	796	39aa	51.3
WC3-23R	AU050772	Calcium-binding protein SDF4	<i>Mus musculus</i>	D50461	705	378	78.8
WD11-14R	AU050286	Carcinoma-associated protein EMS1	<i>Mus musculus</i>	A48063	896	67aa	64.2
WH4-11F	AU050208	Chromosome 7 STS SWSS3576	<i>Homo sapiens</i>	G30917	798	192	69.3
WF2-15R	AU050350	Clone 079B16aE1	<i>Fugu rubripes</i>	Z89320	838	127	81.9

Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WA3-21R	AU050623	Clone RES4-22C	<i>Homo sapiens</i>	AB000461	907	200	61.5
WC9-11R	AU050186	Clotting factor XIIIa	<i>Homo sapiens</i>	M22001	840	85aa	34.1
WC5-14R	AU050279	Cysteine-rich intestinal protein	<i>Homo sapiens</i>	M13018	582	60aa	81.7
WD6-12R	AU050237	D1007.8	<i>Caenorhabditis elegans</i>	AF003151	772	40aa	57.4
WF12-8F	AU050066	D1043.a	<i>Caenorhabditis elegans</i>	Z99771	671	193aa	31.1
WF1-12R	AU050258	1,25-Dihydroxyvitamin D-3 up-regulated protein	<i>Homo sapiens</i>	Q03422	827	251aa	27.9
WA4-19R	AU050534	Emerin	<i>Homo sapiens</i>	X82434	824	179aa	32.4
WE3-22F	AU050719	ETS-related protein TEL	<i>Mus musculus</i>	Y07915	886	155aa	42.6
WA12-8F	AU050029	F13D12.1	<i>Caenorhabditis elegans</i>	Z49217	767	80aa	38.6
WB3-23R	AU050767	F38H4.7	<i>Caenorhabditis elegans</i>	Z77660	358	101aa	39.6
WF4-20F	AU050597	Fractionated X-irradiation-induced 29 thymoma	<i>Mus musculus</i>	U38252	937	148aa	39.9
WD11-9F	AU050100	Fragile-chorion membrane protein	<i>Drosophila melanogaster</i>	A43742	720	73aa	32.9
WD9-20R	AU050588	gp25L2 protein	<i>Homo sapiens</i>	X90872	683	352	76.4
WC6-8F	AU050039	GS3686	<i>Homo sapiens</i>	AB000115	501	75aa	40.0
WD1-12R	AU050232	GTP binding protein	<i>Homo sapiens</i>	E02609	870	547	81.5
WC2-23R	AU050771	GTPase activating protein	<i>Homo sapiens</i>	M64788	747	382	64.1
WE9-13F	AU050269	Heat shock factor binding protein 1	<i>Homo sapiens</i>	AF068754	885	61aa	91.8
WB6-24R	AU050805	HLA-B-associated transcript 3	<i>Homo sapiens</i>	A35098	719	67aa	47.8
WH4-10R	AU050169	HuR RNA binding protein	<i>Homo sapiens</i>	U38175	343	275	79.6
WE11-10F	AU050157	Hypothetical 12.5 kDa protein	<i>Caenorhabditis elegans</i>	U21320	195	27aa	59.0
WD10-12F	AU050240	Hypothetical 16.2 kDa protein	<i>Saccharomyces cerevisiae</i>	S37791	679	63aa	52.4
WA4-23R	AU050759	Hypothetical 76.2 kDa protein ZK632.7	<i>Caenorhabditis elegans</i>	S40939	643	197	66.0
WF6-19F	AU050554	Hypothetical protein	<i>Homo sapiens</i>	X76487	353	26aa	88.5
WE9-10F	AU050155	Hypothetical protein MJ1065	<i>Methanococcus jannaschii</i>	H64432	886	241aa	40.7
WE8-14R	AU050294	Hypothetical protein x	<i>Arabidopsis thaliana</i>	S46330	959	43aa	46.5
WH6-21R	AU050681	Inducible protein	<i>Homo sapiens</i>	L47738	670	665	76.4
WG3-16R	AU050418	KIAA0058	<i>Homo sapiens</i>	D31767	481	20aa	75.0
WC4-17R	AU050455	KIAA0356	<i>Homo sapiens</i>	AB002354	1018	29aa	55.2
WH12-18F	AU050527	KIAA0427	<i>Homo sapiens</i>	AB007887	805	188aa	31.4
WD12-14F	AU050287	KIAA0652	<i>Homo sapiens</i>	AB014552	834	154aa	77.9
WC5-21R	AU050638	Leucin-rich neuronal protein	<i>Homo sapiens</i>	AF053356	869	208aa	54.3
WH10-20R	AU050618	LGN protein	<i>Homo sapiens</i>	U54999	782	185aa	33.0
WA6-16F	AU050373	Ly-9	<i>Homo sapiens</i>	L41621	873	176aa	27.8
WG7-23F	AU050792	Manic fringe	<i>Mus musculus</i>	AF015769	858	475	68.4
WA10-18R	AU050491	Matrix metalloproteinase inhibitor	<i>Rattus norvegicus</i>	L31884	694	297	66.0
WD6-17R	AU050466	Metalloproteinase inhibitor	<i>Rattus norvegicus</i>	U14526	589	79aa	68.4
WF7-18F	AU050516	Myotubularin	<i>Homo sapiens</i>	U46024	1019	162aa	39.5
WD5-17F	AU050465	Nascent polypeptide-associated complex $\alpha$ subunit	<i>Homo sapiens</i>	X80909	460	144aa	63.9
WD8-15F	AU050338	Nbr1	<i>Mus musculus</i>	U73039	830	19aa	84.2



Table 2. Continued

EST	Accession no.	Putative identification	Species	Data base accession	Sequenced length (bp)	ID length	% ID
WC12-8F	AU050042	Neuroendocrine-specific protein C	<i>Homo sapiens</i>	L10335	800	558	64.5
WC5-21F	AU050637	Neuronal protein	<i>Felis domestica</i>	X79682	677	266	61.3
WH8-19F	AU050565	Osteoclast stimulating factor	<i>Homo sapiens</i>	U63717	839	645	74.6
WD2-11R	AU050187	Pleckstrin	<i>Homo sapiens</i>	X07743	872	747	67.1
WF3-9R	AU050107	Regulator of G-protein signaling A28-RGS 14P	<i>Mus musculus</i>	U70427	693	287	68.6
WE6-12F	AU050250	RW1	<i>Mus musculus</i>	AF060565	670	69aa	31.9
WD9-12R	AU050239	S-rex/NSP in neurons	<i>Rattus norvegicus</i>	U17604	782	209aa	44.5
WA3-23F	AU050758	SH3-domain-containing protein p52	<i>Gallus gallus</i>	Z50798	694	188aa	44.7
WE4-19F	AU050551	Shwannoma-associated protein	<i>Mus musculus</i>	AF026124	864	324	57.7
WG8-13R	AU050273	Similar to <i>S. cerevisiae</i> Lpg15p	<i>Caenorhabditis elegans</i>	AF036706	792	25aa	64.0
WH5-11R	AU050210	Similarity to a C2H2-type zinc finger protein	<i>Caenorhabditis elegans</i>	AF016437	713	53aa	66.0
WE7-14R	AU050292	Src kinase-associated phosphoprotein 55-related protein	<i>Homo sapiens</i>	AF072166	936	217aa	59.4
WG8-16F	AU050420	T27E9.4	<i>Caenorhabditis elegans</i>	Z81486	721	47aa	44.7
WH6-9F	AU050120	Tom-1B	<i>Gallus gallus</i>	Y08741	779	203aa	26.6
WB11-12R	AU050224	Translationally controlled tumor protein	<i>Gallus gallus</i>	D26312	537	138	72.5
WE6-22R	AU050721	Transmembrane protein E3-16	<i>Gallus gallus</i>	AF017985	594	40aa	55.0
WB10-10F	AU050139	Transmembrane protein rnp21.4	<i>Rattus norvegicus</i>	X92097	930	610	82.5
WB7-9R	AU050087	Tspan-3	<i>Homo sapiens</i>	AF054840	413	77aa	46.8
WD11-14F	AU050286	Unknown	<i>Homo sapiens</i>	U84971	862	85aa	36.5
WB2-15F	AU050312	Unknown	<i>Fugu rubripes</i>	AF056116	668	79aa	63.3
WB10-17R	AU050449	Unknown gene product	<i>Homo sapiens</i>	AC004382	782	201aa	43.8
WE3-12R	AU050247	Unnamed protein	<i>Homo sapiens</i>	AB000459	1099	46aa	71.7
WA9-22R	AU050694	Vesicle-associated membrane protein/synaptobrevin binding protein	<i>Aplysia californica</i>	U36779	710	320	62.5
WF11-22R	AU050736	Wiskott-Aldrich Syndrome protein interacting protein	<i>Homo sapiens</i>	AF031588	724	58aa	60.3
WB4-15F	AU050314	ZK1251.9	<i>Caenorhabditis elegans</i>	Z68218	765	178aa	42.1

\*Cell division: DNA synthesis/replication, apoptosis, cell cycle, chromosome structure.

†Cell signaling/cell communication: cell adhesion, channel/transport proteins, effector/modulators, hormone/growth factors, intracellular transducers, metabolism, protein modification, receptors.

‡Cell structure/motility: cytoskeletal, extracellular matrix, microtubule-associated proteins/motors.

§Cell/organism defense: homeostasis, immunology.

¶Gene/protein expression: RNA synthesis, embryonic development, protein synthesis.

¶Metabolism: amino acid, cofactor, energy/TCA cycle, lipid, nucleotide, protein modification, sugar/glycolysis, transport.

**Table 3.** Redundancy of Sequencing of ESTs of Japanese Flounder

Leukocytes*			Liver†		
number	%	Genes	number	%	Genes
13	1.34	Gelatinase-b	15	4.29	Apolipoprotein A-I
		Ribosomal protein L23	11	3.14	Complement C3
10	1.03	Collagenase	7	2.00	Cystatin
8	0.83	$\beta$ -Actin	6	1.71	Fibrinogen B- $\beta$ -subunit
		Brain myosin II isoform	5	1.43	Fibrinogen $\alpha$
		HMG-1	4	1.14	Complement regulatory plasma protein
		Osteoclast stimulating factor			elongation factor 1 $\alpha$
7	0.72	H <sup>+</sup> -ATPase	3	0.86	fibrinogen $\gamma$ -subunit
		Ribosomal protein L28			L-SF
		WASP interacting protein			Transferrin
6	0.62	$\beta_2$ -Microglobulin	2	0.57	6-Phosphofructo-2-kinase
		Calmodulin			Acidic ribosomal phosphoprotein
		Pleckstrin			Ceruloplasmin
		Ribosomal protein L41			Complement C7
5	0.52	Fibrinogen-like protein			Complement C9
		HuR RNA binding protein			Cytochrome <i>c</i> oxidase I
		Malate dehydrogenase			Fibrinogen $\alpha$ -subunit
4	0.41	Thrombospondin			Intestinal and liver tetraspan membrane protein
		Actin-related protein			PHKA1
		Adenylosuccinate synthetase			Plasminogen
		<i>c-fos</i>			Elastase
		Retinoid X receptor- $\epsilon$			Ribosomal protein L15
		Ribosomal protein S16			Ribosomal protein S3a
		Ribosomal protein S8			Ribosomal protein S6
		Src kinase-associated phosphoprotein 55			Thrombin
		Tumor necrosis factor receptor I			Transglutaminase
		Transferrin receptor			Tripsinogen
		Transmembrane protein rnp21.4			

\*Leukocytes (983 clones); this study and Nam et al., 1999.

†Liver (350 clones); Inoue et al., 1997.

which were both prevalent at 1.34%. Interestingly, the ribosomal proteins in leukocytes and liver of Japanese flounder were also different in prevalence.

The relations between the homology to known genes and the distances of the ESTs from the ends of the cDNAs are shown in Table 4.

## DISCUSSION

In this study, we analyzed Japanese flounder leukocyte ESTs using the BLASTN and BLASTX programs in BLAST 2.0.

Some of the EST DNA sequences had no significant homology to DNA sequences in GenBank, although they may have significant homology to the deduced amino acid sequence of known genes. Approximately 30% of the identified clones were detected by only the BLASTX program, which compared sequences at the amino acid level (Table 2). In addition, the predicted amino acid sequences of some of ESTs were conserved for only certain specific regions that might be important for determining the characteristics or functions of their respective proteins. For example, the T cell receptors have some highly conserved regions in their amino acid sequences. The sequence of the clone WB9-24R

**Table 4.** Regions of Known Genes Homologous to the Identified ESTs

Homologous region*	%†
1 ~ 100	56.7
101 ~ 200	17.5
201 ~ 300	9.5
301 ~ 400	8.2
401 ~ 500	3.6
501 ~ 600	3.1
601 ~ 700	0.9
>650	0.5

\*Distance of a start nucleotide position of homologous region from sequenced cDNA end.

†(Number of clones/416 identified clones) × 100.

had significant homology to the T cell receptor at the amino acid sequence level in certain regions, although the homology of the whole amino acid sequence between *Raja eglanteria* TCR- $\delta$  and the WB9-24R clone was only 28.4% (Figure 1, Table 2). However, the DNA sequence of WB9-24R had no significant homology to any of the DNA sequences in GenBank. These results indicate that comparisons of amino acid sequences of ESTs and examination of certain specific conserved regions are important for the identification of homologues among known genes.

The average size of the sequenced lengths of human ESTs is approximately 300 bp (Adams et al., 1995). Such a sequence length is typically enough to identify the ESTs of mammals because many genes, cDNAs, and DNA segments have been reported and registered in the DNA databank. Consequently, the homology between these ESTs and known DNA sequences is usually high. However, a sequence length of 300 bp may not be informative enough for identification of fish ESTs. Previously, Gong et al. (1994) reported ESTs from five species of fish. The sequence lengths of their ESTs were 200 to 300 bp, and approximately 31% of the ESTs were identified. In this study, we were able to identify more than 60% of the ESTs obtained because the average sequence length was 721 bp. Our results of the relation between the identification ratio and EST sequence length suggest that at least 650 bp should be sequenced for the identification of fish ESTs (Table 4). In addition, sequencing from both ends of the clone is also important for identification of fish ESTs because some conserved regions of certain proteins occurred at the amino-terminal half, whereas others were present at the carboxy-terminal half.

```

WB9-24:47  MGRSVLRCREYKTSVYVYLXWYRQDSDSAPQVLLRRCARGWSDDEDIPQKEY--ESRTNQTWYETITINKLTLADSAIYCA 300a
EGEB+ C V+ Y YL WKR DS P+++LRK + + D R+ + +T+ T++ L L'DSA+YCA
Raja eglanteria 6  MGRSWTFSCYTDYATRYTYLWYRHHQDSRPFILRKYSSGEEKADPAQARSTQLQTRKRLVSLTSGLELS DSAIYCA 94a

```

**Figure 1.** Comparison of the deduced amino acid sequence of WB9-24R and the amino acid sequence of *Raja eglanteria* T cell receptor- $\delta$ . T cell receptor and immunoglobulin-like motifs are shown in boldface. The numbering of the WB9-24R sequence (a) is based on the nucleotide sequence position with respect to the end of the cDNA. The numbering of the *Raja eglanteria* sequence (b) is based on the amino acid sequence positions reported in GenBank.

We identified 251 different genes, including those associated with cell division, cell structure or motility, basic energy metabolism, gene or protein expression, cell signaling, cell communication, and cell or organism defense. More than 70% of these genes are reported here for the first time from teleostei. We also found several immune-related genes, such as immunoglobulin- $\delta$ , immunoglobulin  $\kappa$ , interferon-induced 54-kDa protein, lysozyme, T cell receptor  $\delta$ , TGF- $\beta$ , and so on. The information reported in this article should be useful for researchers in the field of fisheries sciences and aid them in analyzing genes and proteins of their particular interest.

Together, these results suggest that EST analysis is a powerful and useful technique for the identification and characterization of fish genes.

## ACKNOWLEDGMENTS

This study was supported in part by a grant from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF97L00902).

## REFERENCES

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., and Venter, J.C. (1991). Complementary DNA sequencing: Expressed sequence tags and the human genome project. *Science* 252:1651-1656.
- Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., et al. (1995). Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 377(suppl):3-173.

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, E.W. (1990). Basic local alignment search tool. *J Mol Biol* 215:403–410.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- Banfi, S., Borsani, G., Rossi, E., Bernard, L., Guffanti, A., Rubboli, F., Marchitello, A., Giglio, S., Coluccia, E., Zollo, M., Zuffardi, O., and Ballabio, A. (1996). Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching. *Nature Genet* 13:167–174.
- Boguski, M.S. (1991). The turning point in genome research. *Trends Biochem Sci* 20:295–296.
- Boguski, M.S., Lowe, T.M.J., and Tolstoshev, C.M. (1993). dbEST-database for expressed sequence tags. *Nature Genet* 4:332–333.
- Constanzo, F., Castagnali, L., Dente, L., Arcari, P., Smith, M., Costanzo, P., Raugeri, G., Izzo, P., and Pietropaolo, T.C. (1983). Cloning of several complementary DNA segments coding for human liver proteins. *EMBO J* 2:57–61.
- Hirono, I., and Aoki, T. (1997). Expressed sequence tags of medaka (*Oryzias latipes*) liver mRNA. *Mol Mar Biol Biotechnol* 6:348–353.
- Inoue, S., Nam, B.-H., Hirono, I., and Aoki, T. (1997). A survey of expressed genes in Japanese flounder (*Paralichthys olivaceus*) liver and spleen. *Mol Mar Biol Biotechnol* 6:378–382.
- Kimura, T., Yoshimizu, M., and Gorie, S. (1986). A new rhabdovirus isolated in Japan from cultured himare (Japanese flounder) *Paralichthys olivaceus* and ayu *Plecoglossus altivelis*. *Dis Aquat Org* 1:209–217.
- Nam, B.-H., Yamamoto, E., Hirono, I., and Aoki, T. (1999). A survey of expressed genes in the leukocytes of Japanese flounder, *Paralichthys olivaceus*, infected with hirame rhabdovirus. *Dev Comp Immunol* (in press).
- Okubo, K., and Matsubara, K. (1997). Complementary DNA sequence (EST) collections and the expression information of the human genome. *FEBS Lett* 403:225–229.
- Okubo, K., Hori, N., Matoba, R., Niiyama, T., and Matsubara, K. (1991). A novel system for large-scale sequencing of cDNA by PCR amplification. *DNA Seq* 2:137–144.
- Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet* 2:173–179.
- Wilcox, A.S., Khan, A.S., Hopkins, J.A., and Sikela, J.M. (1991). Use of 3' untranslated sequences of human cDNA for rapid chromosome assignment and conversion to STSs: implications for an expression map of the genome. *Nucleic Acids Res* 19:1837–1843.
- Yamamoto, E. (1995). Studies on sex-manipulation and production of cloned populations in hirame flounder, *Paralichthys olivaceus* (Temminck et Schlegel) (in Japanese). *Bull Tottori Pref Fish Exp Stn* 34:1–145.