

Transcriptome of channel catfish (*Ictalurus punctatus*): initial analysis of genes and expression profiles of the head kidney

D. Cao, A. Kocabas, Z. Ju, A. Karsi, P. Li, A. Patterson and Z. Liu

The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Auburn University, Auburn, AL, USA

Summary

Analysis of expressed sequence tags (ESTs) is an efficient approach for gene discovery, expression profiling, and development of resources useful for functional genomics studies. As part of the transcriptome analysis in channel catfish (*Ictalurus punctatus*), we have conducted EST analysis using a cDNA library made from the head kidney. We analysed 2228 EST clones. Orthologues were established for 1495 (67.1%) clones representing 748 genes, of which 545 (36.5%) clones were singletons. The remaining 733 (32.9%) clones represent unknown gene clones, for which the number of genes has not yet been determined.

Keywords cDNA, expressed sequence tag (EST), functional genomics, gene expression, mapping, marker, microarray.

Introduction

Transcriptome refers to the entire RNA transcripts of an organism, in contrast to the genome and proteome of an organism, which refer to the entire genetic material and the entire protein composition, respectively. In contrast to the relatively stable genome, transcriptome is variable under factors that affect gene expression, such as development, physiology and environment. An efficient and informative approach to characterize a transcriptome is to analyse expressed sequence tags (ESTs, Adams *et al.* 1991). The EST analyses can provide information concerning gene expression profiles (Waterston *et al.* 1992; Hofte *et al.* 1993; Azam *et al.* 1996) and provide molecular reagents for functional genomics studies using cDNA microarray technology (Ermolaeva *et al.* 1998; Moch *et al.* 1999; van Hal *et al.* 2000). Additionally, ESTs are rich resources for type I polymorphic markers (Wang *et al.* 1998; Emahazion *et al.* 1999; Liu *et al.* 1999).

We have conducted EST analyses of pituitary, muscle and brain libraries in channel catfish (Karsi *et al.* 1998; Ju *et al.*

2000). Large-scale EST analysis is essential for adapting cDNA microarray technology to comparative functional genomics, particularly important for assessing the complex nature of gene expression involved in performance traits such as disease resistance. Gene cataloguing and profiling of the head kidney is an essential part of our effort in EST analysis of immune organs. The immune system of fish is similar to that of mammals (Ellis 1992) consisting of non-specific defence barriers and specific immune functions; the latter includes T- and B-cell mediated cellular and humoral immunity (Warr 1997). In addition to the thymus and spleen, which are also present in higher vertebrates, another lymphoid organ, the head kidney, exists in fish (Pichappan 1980; Chilmoczyk 1992) and is the functional counterpart of mammalian bone marrow. The objectives were to identify genes and their expression profiles in the head kidney, and to develop EST resources for functional genomic studies. As part of the transcriptome analysis of channel catfish, we report analysis of 2228 clones from the head kidney.

Methods

Tissue preparation and RNA isolation

All experimental fish were raised in troughs located in the hatchery of the Auburn University Fish Genetics Facility under the same conditions for 4 weeks prior to the initiation

Address for correspondence

Z. Liu, The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Auburn University, Auburn, AL 36849, USA. E-mail: zliu@acesag.auburn.edu

Accepted for publication 15 March 2001

of the experiments. To include all potential transcripts from immune organs, we collected tissues from non-challenged fish as well as from fish challenged with *Edwardsiella ictaluri*, the causative agent for the bacterial disease enteric septicaemia of catfish (ESC). At the start of the experiments, fish were divided into challenge and control groups. A total of 240 fish were challenged with *E. ictaluri* following the procedures of Dunham *et al.* (1993). Fingerlings were placed in a static 150-l tank containing *E. ictaluri* cells at a final concentration of 1.1×10^6 cells/ml for 1 h. The fish were then removed and stocked in a 1000 l tank. Mortality was monitored for 14 days. Dead fish were removed when observed, and clinical signs of ESC were recorded. Head kidney samples of 10 randomly selected fish were collected daily and frozen in liquid nitrogen until isolation of RNA. Pooled head kidney tissues were ground with a mortar/pestle and then homogenized with a hand-held tissue tearer (Model 985-370, Biospec Products Inc., Racine, WI, USA) in RNA extraction buffer following the guanidium thiocyanate method (Chomczynski & Sacchi 1987). Poly (A)⁺ RNA was purified from total cellular RNA using the Poly (A)⁺ Pure kit (Ambion, Austin, TX, Cat. #1915) according to the manufacturer's instructions.

cDNA library construction

A directional cDNA library of the head kidney was constructed using the pSPORT-1 SuperScript Plasmid Cloning System from Gibco/BRL. Construction of the cDNA library followed Gibco/BRL's instructions, except that the library was electroporated into ElectroMax DH12S cells. These cells are highly adapted to efficient electroporation and production of single-stranded phagemids (Gibco/BRL), features advantageous to development of normalized cDNA libraries. Over 10 million primary cDNA clones were obtained with an average insert size of 1.9 kb. The primary cDNA library was amplified once before colonies were picked for sequencing.

Plasmid preparation and sequencing analysis

The plasmid cDNA library was plated to a density appropriate for picking individual colonies. Random clones were grown in 1.5-ml LB medium overnight in 12 × 75 mm culture tubes. Plasmid DNA was prepared by alkaline lysis method (Sambrook *et al.* 1989) using the Qiagen Spin Column Mini-plasmid kits. Three microlitres of plasmid DNA (about 0.5–1.0 µg) were used in sequencing reactions. Chain termination sequencing (Sanger *et al.* 1977) was performed using cycleSeq-farOUT™ polymerase (Display Systems Biotech, Vista, CA, USA). The PCR profiles were: 95 °C for 30 s, 55 °C for 40 s, 72 °C for 45 s for 30 cycles. An initial 2 min denaturation at 96 °C and a 5-min extension at 72 °C were

always used. Sequences were analysed on an automatic LI-COR DNA Sequencer Long ReadIR 4200 or LI-COR DNA Analyzer Gene ReadIR 4200 (LI-COR, Lincoln, NB, USA).

Gene identity determination

BLAST searches (NCBI, Bethesda, MD, USA) were conducted to determine gene identity. Procedures for establishing orthologues are shown in Fig. 1. Matches were considered to be significant only when the probability (*P*) was less than 1×10^{-4} using BLASTN and BLASTX with all parameters at the defaults. The ESTs with significant similarities in BLASTN searches were considered an orthologue of known genes only when the similarities were not caused by simple sequences. Those ESTs failing to show significant similarities with BLASTN were further searched by BLASTX. Again, ESTs with significant similarities were evaluated for matches caused by simple amino acid matches or repeats. All ESTs that were not identified as orthologues of known genes were designated as unknown EST clones.

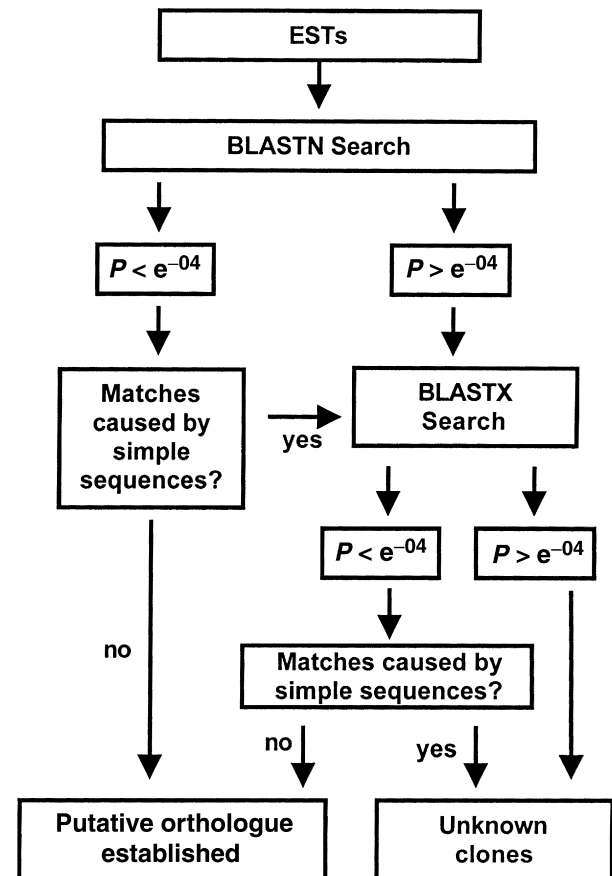


Figure 1 Schematic presentation of the procedures used for annotation of orthologues identified by BLASTN and BLASTX similarity comparisons.

Table 1 Expression profiles of orthologues established by BLASTN searches.

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk00121	BE468277	40 kDa ribosomal protein	9e-63	1
IpHdk01073	BE468278	Acidic ribosomal phosphoprotein P0	0.0	18
IpHdk00305	BE468279	Acidic ribosomal protein P1	6e-32	2
IpHdk01077	BE468367	Ribosomal protein large P2	2e-32	5
IpHdk01142	BE468368	Repeat family 3 gene	e-119	26
IpHdk01172	BE469789	Ribosomal protein L3	7e-89	11
IpHdk01166	BE468369	Ribosomal protein L4	1e-74	15
IpHdk00162	BE468370	Ribosomal protein L5	1e-46	5
IpHdk01777	BE468371	Ribosomal protein L5b	3e-04	5
IpHdk03103	BE468372	Ribosomal protein L6	6e-14	2
IpHdk00645	BE468373	Ribosomal protein L7	1e-18	2
IpHdk01616	BE468374	Ribosomal protein L7a	2e-85	14
IpHdk01008	BE468375	Ribosomal protein L8	e-126	3
IpHdk01435	BE469013	Ribosomal protein L9	1e-49	7
IpHdk01505	BE469244	Ribosomal protein L10A	e-104	3
IpHdk00224	BE469015	Ribosomal protein L11	1e-95	6
IpHdk01062	BE469016	Ribosomal protein L12	2e-57	6
IpHdk01050	BE469017	Ribosomal protein L14	3e-15	3
IpHdk01136	BE469018	Ribosomal protein L15	1e-58	10
IpHdk01038	BE469019	Ribosomal protein L17	e-135	8
IpHdk00287	BE469020	Ribosomal protein L18	e-104	5
IpHdk01185	BE469021	Ribosomal protein L18a	e-127	9
IpHdk01446	BE469022	Ribosomal protein L19	3e-72	5
IpHdk00105	BE469023	Ribosomal protein L21	4e-43	6
IpHdk01014	BE469024	Ribosomal protein L22	1e-82	3
IpHdk01602	BE469309	Ribosomal protein L24	4e-12	9
IpHdk01215	BE469025	Ribosomal protein L26	7e-97	1
IpHdk00299	BE469026	Ribosomal protein L27	1e-82	5
IpHdk01056	BE469027	Ribosomal protein L27a	3e-34	6
IpHdk01562	BE469278	Ribosomal protein L28	2e-07	5
IpHdk01290	BE469028	Ribosomal protein L29	5e-38	1
IpHdk02655	BE468313	Ribosomal protein L30	6e-54	3
IpHdk01122	BE469029	Ribosomal protein L31	8e-50	4
IpHdk01186	BE469030	Ribosomal protein L32	1e-45	7
IpHdk01952	BE469517	Ribosomal protein L34	4e-14	1
IpHdk01768	BE469421	Ribosomal protein L35	1e-20	5
IpHdk02538	BE468284	Ribosomal protein L35a	2e-10	7
IpHdk01934	BE469515	Ribosomal protein L36	9e-34	2
IpHdk00285	BE469031	Ribosomal protein L36a	1e-63	3
IpHdk01140	BE469032	Ribosomal protein L37	2e-18	2
IpHdk01625	BE469310	Ribosomal protein L37a	1e-17	4
IpHdk01011	BE469033	Ribosomal protein L38	3e-36	8
IpHdk01020	BE469034	Ribosomal protein L41	2e-25	12
IpHdk01527	BE469263	Ribosomal protein S2	e-100	16
IpHdk00297	BE469035	Ribosomal protein S3	e-125	9
IpHdk03030	BE469036	Ribosomal protein S4	e-148	10
IpHdk00302	BE468477	Ribosomal protein S5	e-108	4
IpHdk01047	BE469037	Ribosomal protein S6	6e-67	7
IpHdk01081	BE469038	Ribosomal protein S8	4e-77	5
IpHdk01143	BE469039	Ribosomal protein S9	2e-60	9
IpHdk01210	BE469040	Ribosomal protein S10	2e-42	9
IpHdk01138	BE469041	Ribosomal protein S11	e-147	4

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk00081	BE469042	Ribosomal protein S12	e-122	8
IpHdk00200	BE469043	Ribosomal protein S13	0.0	5
IpHdk01614	BE469290	Ribosomal protein S14	8e-66	3
IpHdk00250	BE469044	Ribosomal protein S15	e-106	3
IpHdk02607	BE468300	Ribosomal protein S16	2e-15	1
IpHdk01127	BE469045	Ribosomal protein S17	e-108	5
IpHdk01517	BE469256	Ribosomal protein S18 (Ke-3)	6e-85	9
IpHdk00331	BE468487	Ribosomal protein S19	2e-26	1
IpHdk00107	BE469046	Ribosomal protein S20	1e-18	3
IpHdk00276	BE469047	Ribosomal protein S21	1e-29	3
IpHdk02590	BE468297	Ribosomal protein S23	2e-44	2
IpHdk00440	BE468508	Ribosomal protein S24	2e-53	3
IpHdk01605	BE469286	Ribosomal protein S25	1e-14	1
IpHdk01631	BE469315	Ribosomal protein S27	3e-22	3
IpHdk01365	BE469049	Ribosomal protein S29	2e-31	4
IpHdk01205	BE469050	Ribosomal protein QM	3e-97	4
IpHdk02237	BE468550	37KD laminin receptor precursor/p40 ribosomal associated protein	2e-06	1
IpHdk03163	BE469051	Ubiquitin/ribosomal protein S27a	e-104	2
IpHdk00026	BE469052	Ubiquitin/ribosomal protein gene S27a-related	1e-10	1
IpHdk01415	BE469053	Hybrid protein (ubiquitin-like protein/ribosomal protein S30	4e-09	1
IpHdk00084	BE469054	Immunoglobulin heavy chain/ribosomal protein S8	5e-27	1
IpHdk01561	BE469277	Neoplasm-related C140 product/ribosomal protein L6	6e-14	1
IpHdk00358	BE468493	Translation initiation factor 2 eTIF2	5e-28	1
IpHdk00580	BE468539	Translation initiation factor 3, subunit 4	1e-33	1
IpHdk00111	BE469055	Translation initiation factor 3, subunit 5 (epsilon, 47 kDa)	3e-22	2
IpHdk01612	BE469288	Translation initiation factor eIF4A I	5e-31	1
IpHdk01315	BE469056	Translation initiation factor eIF4A II	2e-26	1
IpHdk02666	BE468316	Translation initiation factor eIF-4E	1e-39	2
IpHdk01281	BE469057	Translation initiation factor eIF5A	1e-25	1
IpHdk01178	BE469058	Translation initiation factor SUI1	2e-48	2
IpHdk02523	BE468281	Sui 1 iso1 protein	2e-36	2
IpHdk00050	BE469059	Translation elongation factor 1 alpha	0.0	9
IpHdk01188	BE469060	Translation elongation factor 2	e-141	5
IpHdk00587	BE468540	Nascent-polypeptide-associated complex alpha polypeptide	2e-51	2
IpHdk01687	BE469353	18S Ribosomal RNA	0	6
IpHdk01968	BE469521	28S Ribosomal RNA	e-114	4
IpHdk01514	BE469253	Beta-actin	0	44
IpHdk01229	BE469061	Coronin (actin binding protein 1 A)	8e-11	1
IpHdk02525	BE468282	Coronin (actin binding protein 1C)	3e-06	1
IpHdk02602	BE468299	Cofilin (small actin binding protein) 1, non-muscle	3e-32	3
IpHdk00103	BE469062	Alpha-tubulin	0.0	9
IpHdk02631	BE468307	Testis-specific alpha-tubulin	6e-59	1
IpHdk01925	BE469513	Beta tubulin	1e-45	3
IpHdk01697	BE469362	Keratin 8	3e-10	1
IpHdk01174	BE469063	Plasma gelsolin	3e-14	1
IpHdk00388	BE468500	Gelsolin-like capping protein (actin filament)	0.090	1
IpHdk01104	BE469064	Capping protein, beta3 isoform	1e-70	1
IpHdk02619	BE468303	Microglobulin precursor (b2m)	0.0	13
IpHdk00238	BE469065	Invariant chain-like protein	7e-08	1
IpHdk02247	BE468554	Cytokeratin type I	5e-09	1
IpHdk01508	BE469247	Plastin 1 (I isoform)	7e-08	3

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk02613	BE468301	Lymphocyte cytosolic protein 1 (L-plastin)	5e-18	2
IpHdk01109	BE469066	Procollagen, type V, alpha 2	9e-04	1
IpHdk02685	BE468319	Delta-crystallin	3e-15	1
IpHdkF00343	BE468489	Major vault protein MVP100	4e-09	1
IpHdkF00366	BE468496	Haemoglobin beta chain	e-129	10
IpHdk02585	BE468296	Alpha-globin IV (globin zeta)	1e-04	1
IpHdk00035	BE469067	Alpha-globin	1e-05	8
IpHdk01110	BE469068	Beta-globin	2e-07	7
IpHdk02110	BE469762	Danio rerio aA1 globin	5e-06	1
IpHdk01033	BE469069	Myosin II isoform	2e-60	1
IpHdk00504	BE468518	Myosin regulatory light chain 2	3e-90	1
IpHdk00532	BE468527	Myosin regulatory light chain	5e-71	1
IpHdk00232	BE469070	Smooth muscle myosin regulatory light chain	6e-80	2
IpHdk01206	BE469071	Phosphoethanolamine cytidylyltransferase 2 (CTP)	8e-18	1
IpHdk01646	BE469328	DNA methyltransferase	5e-86	1
IpHdk02550	BE468286	Glutathione S-transferase	5e-18	4
IpHdk02632	BE468308	Histone acetyltransferase	3e-08	1
IpHdk01155	BE469072	Methionine adenosyltransferase I, alpha	9e-23	1
IpHdk02614	BE468302	Methionine adenosyltransferase II, alpha	1e-07	1
IpHdk03140	BE469073	Beta-1,4-galactosyl-transferase VI	3e-16	1
IpHdk02186	BE488183	Dolichyl-diphospho-oligosaccharide- protein glycosyltransferase	3e-54	1
IpHdk02693	BE468324	Aldehyde dehydrogenase 2, mitochondrial	1e-15	1
IpHdk01504	BE469243	Aminolevulinate, delta-dehydratase	1e-12	1
IpHdk00255	BE469074	Carbonic anhydrase homologue	3e-07	1
IpHdk00073	BE469075	Isocitrate dehydrogenase 3	5e-36	1
IpHdk00301	BE468476	Lactate dehydrogenase	e-101	2
IpHdk03195	BE469076	Lactate dehydrogenase-A (Ldh-A)	e-133	1
IpHdk01560	BE469276	L-threonine 3 dehydrogenase	0.001	1
IpHdk00731	BE468846	Leucotriene A4 hydrolase	6e-8	2
IpHdk02568	BE468290	Malate dehydrogenase precursor	2e-10	1
IpHdkF00470	BE468510	Testicular 3-beta hydroxysteroid dehydrogenase	2e-23	1
IpHdk00322	BE468481	S-adenosylhomo-cysteine hydrolase	3e-33	1
IpHdk01555	BE469271	Ubiquitin C-terminal hydrolase UCH37	1e-31	1
IpHdk03136	BE469077	Vacuolar ATPase subunit A	3e-38	1
IpHdk01648	BE469330	ATPase, 9 kD	7e-08	1
IpHdk02170	BE469570	ATPase, 14 kDa	9e-14	1
IpHdk02244	BE468553	ATPase, 16 kDa	6e-36	1
IpHdk01248	BE469078	Integrin binding protein	8e-08	2
IpHdk01123	BE469079	ATPase, 21 kDa	2e-11	1
IpHdk02537	BE468283	ATPase, 31 kDa	2e-39	2
IpHdk01029	BE469080	ATPase subunit 8 and ATPase subunit 6	3e-99	3
IpHdk01385	BE469081	Calcium ATPase	5e-68	1
IpHdk00275	BE469082	ATP synthase beta-subunit	0.0	1
IpHdk01048	BE469083	ATP synthase, H + transporting mitochondrial FO complex, subunit c	3e-47	2
IpHdk00937	BE469702	Phosphatidylserine synthase 1	2e-06	1
IpHdk01515	BE469254	S-adenosylmethionine synthetase	1e-37	6
IpHdk01839	BE469459	Adenylosuccinate synthetase 1	9e-14	2
IpHdk02131	BE469571	Glutamine synthetase	8e-63	1
IpHdk02236	BE468549	Pseudouridine synthase 1	8e-08	1
IpHdk00518	BE468523	Methionine-tRNA synthetase	7e-15	2

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk01661	BE469343	Aminolevulinate synthase erythroid specific isoform	4e-59	4
IpHdk00223	BE469084	Spermine synthase	6e-06	2
IpHdk01467	BE469085	Cyclin-dependent kinase 2 (CDK2)	3e-48	1
IpHdk00505	BE468519	Nuclease diphosphate kinase B	1e-86	1
IpHdk02286	BE468560	Serine/threonine protein kinase EMK	7e-17	1
IpHdk01374	BE469086	Pyruvate kinase	9e-51	1
IpHdk00747	BE468861	Pyruvate dehydrogenase kinase 2	2e-11	1
IpHdk01494	BE469087	Casein kinase 2 alpha subunit	e-117	1
IpHdk00942	BE469572	Creatine kinase	e-139	1
IpHdk01221	BE469088	Glycogen synthase kinase 3 beta	3e-05	1
IpHdkF00303	BE468478	Mitogen-activated protein kinase p38 delta	1e-04	1
IpHdk01017	BE469089	Protein kinase (SEK1)	3e-13	1
IpHdk01967	BE469520	Protein kinase (Ste20-related kinase SMAK)	5e-21	1
IpHdk01655	BE469337	Tyrosine kinase (JAK1 kinase)	5e-18	1
IpHdk00502	BE468516	Phosphoglycerate kinase	3e-34	1
IpHdk01259	BE469090	Inorganic polyphosphatase	9e-20	1
IpHdk00282	BE469091	Brain glycogen phosphorylase B	1e-24	1
IpHdk01144	BE469092	Protein tyrosine phosphatase, receptor type, G	4e-59	1
IpHdkF00484	BE468514	Protein tyrosine phosphatase, non-receptor type 6	3e-4	1
IpHdk00526	BE468525	Protein phosphatase X	6e-58	1
IpHdk03049	BE469093	Protein phosphatase type 2C alpha 1	2e-45	1
IpHdk00141	BE469094	Dual specificity phosphatase	5e-13	3
IpHdk01266	BE469095	Collagenase 3	7e-06	1
IpHdk00577	BE468538	Collagenase-3 (gene A)	2e-05	1
IpHdk00640	BE468982	Procollagen (type III) N-endopeptidase	3e-56	1
IpHdk00552	BE468888	Neutrophil collagenase precursor (matrix metalloproteinase-8)	4e-16	1
IpHdk00474	BE468511	Disintegrin and metalloprotease domain 10 (ADAM10)	1e-06	1
IpHdk01086	BE469096	Matrix metalloproteinase 20 (enamelysin)	7e-06	2
IpHdk03004	BE469097	Microsomal signal peptidase (18 kd)	3e-35	1
IpHdk00908	BE469573	Signal peptidase complex 25 kDa subunit	5e-06	1
IpHdk01920	BE469512	26S protease subunit	2e-30	2
IpHdk01454	BE469098	Calcium-dependent protease, small subunit	4e-13	1
IpHdk00218	BE469099	Peptidase D	5e-15	1
IpHdk00326	BE468483	Serine protease-like protein	7e-06	1
IpHdk00387	BE468483	Proteasome (multiprotease complexes) subunit, alpha type 5	2e-42	1
IpHdk03021	BE470450	20S proteasome subunit alpha 1	5e-07	1
IpHdk01618	BE469292	Zea mays proteasome alpha	4e-07	1
IpHdk01262	BE469100	Proteasome (prosome, macropain) subunit, beta type 1	7e-98	1
IpHdk01153	BE469101	Nephrosin (secreted zinc endopeptidase)	3e-04	1
IpHdk01320	BE469102	Proteasome (prosome, macropain) subunit, alpha type 3	1e-37	1
IpHdk02213	BE468544	Cathepsin B	3e-04	1
IpHdk02184	BE469574	Cathepsin D	1e-44	1
IpHdk01973	BE469522	Cathepsin H	5e-06	1
IpHdk01164	BE469103	Cathepsin L	3e-51	1
IpHdk01559	BE469275	Cathepsin SS-cysteine proteinase)	2e-05	1
IpHdk00256	BE469104	Glutathione peroxidase 1	3e-07	1
IpHdk02230	BE468548	NADH ubiquinone reductase acyl carrier subunit	3e-13	2
IpHdk00713	BE468829	NADH-cytochrome b5 reductase	1e-09	1
IpHdk00528	BE468526	Thioredoxin reductase 1 (Txnrd1)	7e-08	1
IpHdk00429	BE468507	Thioredoxin (heat-stable oxidoreductase)	6e-04	2
IpHdk00624	BE468969	Ubiquinone oxidoreductase complex	2e-13	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk00120	BE469105	2-amino-3-ketobutyrate coenzyme A ligase	1e-07	1
IpHdk00936	BE469575	Aminoacylase	4e-04	1
IpHdk01455	BE469106	ADP/ATP translocase T1	1e-84	5
IpHdk01311	BE469107	ADP/ATP translocase T2	e-101	2
IpHdk01133	BE469108	Aldolase C	3e-90	4
IpHdk01135	BE469109	Alpha enolase	e-136	1
IpHdk03071	BE469110	cAMP specific phosphodiesterase 7B	3e-08	1
IpHdk01841	BE469461	Cu/Zn-superoxide dismutase	4e-28	1
IpHdk03138	BE469111	Cysteine desulphurase	8e-05	1
IpHdk01607	BE469287	Cytosolic IMP-GMP specific 5'-nucleotidase	1e-40	1
IpHdk03121	BE469112	Deacetylase	2e-52	1
IpHdk02147	BE469776	Ubiquitin-conjugating enzyme E2D 2	1e-12	1
IpHdk02291	BE468561	DEAD (aspartate-glutamate-alanine-aspartate) box polypeptide 3 (RNA helicase)	2e-08	1
IpHdk01181	BE469113	Nuclear RNA helicase	e-102	2
IpHdk01694	BE469359	Human protein (peptidyl-prolyl <i>cis/trans</i> isomerase) NIMA-interacting	3e-10	1
IpHdk02250	BE468555	Fucosidase, alpha-L-1	1e-06	1
IpHdk00201	BE469114	Gelatinase B	2e-27	5
IpHdk01932	BE469514	Human RuvB (<i>E. coli</i> homologue)-like 2 (DNA helicase)	1e-65	1
IpHdk01273	BE469115	Lysosomal acid lipase	4e-06	1
IpHdk01809	BE469453	Nuclear RNA helicase Bat1	2e-14	1
IpHdk00054	BE469116	Ornithine decarboxylase antizyme	4e-71	3
IpHdk02254	BE468556	Phosphoglucose isomerase	1e-06	1
IpHdk02659	BE468314	Phosphoglycerate mutase type B subunit	2e-91	1
IpHdkF00300	BE468475	Transaldolase	2e-21	4
IpHdk01132	BE469117	Transketolase	9e-32	1
IpHdk00296	BE469118	Triosephosphate isomerase	3e-14	1
IpHdk03160	BE469119	TXr.24 transposase	1e-10	1
IpHdk02662	BE468315	TXr.9 transposase	2e-20	1
IpHdk00515	BE468521	Ubiquitin conjugating enzyme	1e-83	1
IpHdkF00329	BE468486	Activin B	e-120	8
IpHdk00974	BE469576	Calpain 4	1e-15	1
IpHdk01235	BE469120	Follistatin	1e-15	1
IpHdk01718	BE469380	Growth hormone releasing hormone	1e-31	1
IpHdk01284	BE469121	Prepronerve growth factor homologue	2e-54	5
IpHdk01528	BE469264	Thioredoxin	1e-155	1
IpHdk00534	BE468528	Thymosin beta a	2e-05	2
IpHdk01261	BE469122	Thymosin beta-10	8e-05	1
IpHdkF00368	BE468722	<i>Oncorhynchus mykiss</i> thymosin beta	1e-10	1
IpHdk01453	BE469123	Alpha-prothymosin	7e-06	2
IpHdk00013	BE469124	Thymosin beta-4	7e-17	3
IpHdk01137	BE469125	C10 protein (chemokine)	7e-20	2
IpHdk00135	BE469126	Human antigen identified by antibodies 4F2, TRA10, TROP4 and T43	7e-06	1
IpHdk00557	BE468534	Human Sjogren's syndrome/scleroderma autoantigen 1 (SSSCA1)	1e-05	1
IpHdk01355	BE469127	<i>I. punctatus</i> clone Icpu-A-2 MHC class II antigen	0.0	5
IpHdk01497	BE469128	<i>I. punctatus</i> clone PG27 class G IgL chain, VJC region	0.0	3
IpHdk01558	BE469274	<i>I. punctatus</i> Ig rearranged H-chain, V-C-region	0.0	6
IpHdk00321	BE468480	<i>I. punctatus</i> Igh DNA for immunoglobulin heavy chain (μ)	e-176	3
IpHdk01689	BE469355	<i>I. punctatus</i> immunoglobulin light chain F class	7e-70	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk01083	BE470275	Immunoglobulin light chain constant region	1e-06	1
IpHdk00359	BE468494	<i>I. punctatus</i> MHC class I alpha chain Icpu-UAA	e-175	7
IpHdkF00374	BE468497	<i>I. punctatus</i> MHC class I alpha chain Icpu-UBA	e-152	1
IpHdkF00365	BE468495	<i>I. punctatus</i> MHC class I alpha chain Icpu-UCA	0.0	3
IpHdk01531	BE469266	Oncorhynchus kisutch non-classical MHC class I antigen (Onki-UAA)	2e-05	1
IpHdk01950	BE469516	<i>I. punctatus</i> MHC class II beta chain	e-105	1
IpHdk00691	BE468811	Immunoglobulin heavy chain	e-130	1
IpHdk01092	BE469129	Immunoglobulin light chain F class (IgL)	0.0	6
IpHdk03029	BE469130	Immunoglobulin light chain L2	1e-16	1
IpHdk00100	BE469131	Immunoglobulin light chain variable region	3e-33	2
IpHdk01905	BE469510	Immunoglobulin light chain variable region	4e-18	1
IpHdk01748	BE469404	LIM and senescent cell antigen-like domains 1	2e-10	1
IpHdk01350	BE469132	Monoclonal non-specific suppressor factor beta	3e-34	1
IpHdk01986	BE469524	Natural killer cell enhancement factor	4e-09	1
IpHdk00280	BE469133	PG4 class G IgL chain	0.0	4
IpHdk00565	BE468536	Proliferating cell nuclear antigen	3e-97	2
IpHdk01024	BE469134	Transplantation antigen	7e-82	9
IpHdk01346	BE469135	Type I rhodopsin gene, 5' flanking region/immuno-globulin heavy chain joining region fusion	2e-36	1
IpHdk01495	BE469136	High mobility group 1 protein (HMG-1)	8e-64	2
IpHdk02688	BE468321	High mobility group protein 2	1e-46	1
IpHdk01994	BE469525	High mobility group protein 17	5e-09	12
IpHdk02081	BE469577	High mobility group protein HMG-T	1e-46	2
IpHdk00191	BE469137	Small nuclear ribonucleoprotein B'	8e-20	1
IpHdk00642	BE468984	Small nuclear ribonucleoprotein D3 polypeptide	1e-43	1
IpHdk01647	BE469329	Karyopherin alpha 4 (importin alpha 3)	2e-10	1
IpHdk02011	BE469578	Importin beta	6e-86	2
IpHdk03084	BE469138	H1d-histone and H2B-histone	1e-10	1
IpHdk00620	BE468966	Variant histone H2A	3e-31	1
IpHdk01492	BE469139	Human non-histone chromosome protein	2e-18	1
IpHdk03102	BE469140	Histone H2A and H3	8e-57	1
IpHdk01716	BE469379	H1 histone family, member X	2e-08	1
IpHdk02107	BE469579	Heterogeneous nuclear ribonucleoprotein K	3e-23	1
IpHdk01401	BE469141	RAN binding protein 3	2e-05	1
IpHdk03176	BE470513	Single strand DNA-binding protein	2e-09	1
IpHdk03151	BE469142	Transmembrane protein E3-16	1e-07	1
IpHdk00220	BE469143	Annexin (Ca ²⁺ -regulated membrane binding proteins) max1	5e-14	2
IpHdk01101	BE469144	Annexin (Ca ²⁺ -regulated membrane binding proteins) max4	3e-19	1
IpHdk01329	BE469145	Annexin IV (p33/41)	3e-06	1
IpHdk02700	BE468325	Ceroid-lipofuscinosis, neuronal 5 (CLN5)	2e-04	1
IpHdk02563	BE468289	Clathrin heavy chain	3e-17	2
IpHdk00325	BE468482	Clathrin-associated AP-2 complex AP50 subunit	e-145	1
IpHdkF00344	BE468490	Epsilon-COP	2e-42	1
IpHdk02651	BE468312	Golgi stacking protein homologue GRASP55	1e-40	1
IpHdk02690	BE468322	Golgi-specific brefeldin A-resistance factor 1	6e-19	1
IpHdk02264	BE468558	26S proteasome-associated pad1	2e-45	1
IpHdk02267	BE468559	20S proteasome subunit C2	8e-36	1
IpHdk00426	BE468506	Muscle-specific beta 1 integrin binding protein	1e-05	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk01429	BE469146	Beta-1 integrin	2e-27	1
IpHdk02177	BE469580	Human adaptor-related protein complex 3, mu 2 subunit	2e-19	1
IpHdk02672	BE468317	Kinesin light chain	3e-49	3
IpHdk02507	BE468377	Dynein beta-heavy chain	3e-07	1
IpHdk03019	BE469147	Cytoplasmic dynein heavy chain	9e-82	2
IpHdk01264	BE469148	Dynein light chain	3e-11	2
IpHdk00503	BE468517	Dynein, axon, light chain 4	6e-31	1
IpHdk00226	BE469149	Uncoupling protein 2 (UCP2 gene)	3e-41	2
IpHdk00041	BE469150	Enterocyte differentiation associated factor	2e-13	1
IpHdk02220	BE468545	Neural precursor cell expressed, developmentally down-regulated 5	1e-22	1
IpHdk01247	BE469151	Human GW128 protein	9e-22	1
IpHdk02510	BE468280	Human uncharacterized haematopoietic stem/progenitor cells protein MDS027	7e-11	1
IpHdk00945	BE469581	Human uncharacterized haematopoietic stem/progenitor cells protein MDS029	1e-13	1
IpHdk01018	BE469152	Dendritic cell protein (GA17 protein)	3e-16	1
IpHdk01975	BE469523	TAN-1 (homologue of <i>Drosophila</i> Notch gene)	5e-09	1
IpHdk00541	BE468530	Angiopoietin-like 1 (angiogenesis protein)	0.001	1
IpHdk03053	BE469153	Ezrin	1e-52	1
IpHdk01535	BE469267	Human SH3 domain binding glutamic acid-rich protein homologue	3e-04	1
IpHdk00750	BE468864	Human SM22 alpha (involved in smooth muscle cell differentiation)	3e-16	2
IpHdk03036	BE469154	Intracisternal A-particle promoted placental gene	4e-04	1
IpHdk00543	BE468531	Human SMT3 (suppressor of mif two 3, yeast) homologue 1 (kinetochore protein)	2e-35	1
IpHdk00030	BE469155	Arsenite resistance protein	4e-11	1
IpHdk01209	BE469156	Arsenate resistance protein ARS2	3e-08	2
IpHdk02304	BE468562	CIRP (cold inducible RNA binding protein)	9e-04	2
IpHdk00197	BE469157	<i>Danio rerio</i> zfy1 (cold shock protein)	7e-73	1
IpHdk00545	BE468532	Cold-inducible gene product pAG4	4e-31	1
IpHdk00012	BE469158	Heat shock protein 70	0.0	2
IpHdk01726	BE469387	Heat shock protein hsp90 beta	e-103	2
IpHdk03059	BE469159	Heat shock factor binding protein 1	3e-26	1
IpHdk02594	BE468298	Cyclophilin (peptidyl-prolyl isomerase)	6e-35	1
IpHdk00089	BE469160	Cyclophilin	9e-06	1
IpHdk01225	BE469161	Cyclophilin D	2e-39	4
IpHdk02183	BE469582	Silica-induced gene 41 (Silg41)	9e-08	1
IpHdkF00337	BE468488	Ferritin heavy subunit	e-122	24
IpHdk02628	BE468305	Calmodulin	1e-79	1
IpHdk01796	BE469444	Calmodulin III	4e-96	3
IpHdk00566	BE468537	Calcium binding protein, intestinal	2e-14	1
IpHdk01213	BE469162	S100-like calcium binding protein	2e-07	2
IpHdk01779	BE469429	Ictacalcin (calcium binding protein)	8e-75	1
IpHdk00408	BE468505	Swelling dependent chloride channel	3e-08	1
IpHdk01080	BE469163	Voltage-dependent anion channel	1e-49	2
IpHdk02022	BE469583	Centrin, EF-hand calcium binding protein	1e-06	1
IpHdk01485	BE469164	Choline transporter-like protein 2 (CTL2)	1e-10	1
IpHdk00726	BE469583	sec61 gamma	5e-24	1
IpHdk03193	BE469165	Rab11B (vesicular transporter)	e-112	1
IpHdk02047	BE469584	ATP-binding cassette (ABC) protein	1e-12	1
IpHdk00512	BE468520	Human KDEL receptor	2e-92	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk01975	BE469523	<i>Ictalurus punctatus</i> NCC receptor protein 1	0.0	1
IpHdk00290	BE469166	Receptor for activated protein kinase C	0.0	6
IpHdk01353	BE469167	Leucocyte DNA binding receptor	4e-22	1
IpHdk02547	BE468285	<i>Danio rerio</i> odorant receptor gene cluster	6e-04	3
IpHdk02629	BE468306	Laminin receptor 1 (Lamr1)/ribosomal protein p40	9e-30	16
IpHdk02676	BE468318	Serine/threonine kinase receptor associated protein	6e-29	1
IpHdk01781	BE469430	Serine/threonine kinase receptor type1	1e-14	1
IpHdkF02084	BE469585	Growth factor receptor-bound protein 2	6e-52	1
IpHdk00476	BE468512	Pendulin	2e-20	1
IpHdk00478	BE468513	U6 snRNA-associated Sm-like protein LSM2	3e-47	1
IpHdk01270	BE469168	U5 snRNP-specific protein	7e-64	1
IpHdk00491	BE468515	Mouse clone SMT3B-g1/mSMT3 SMT3B (ubiquitin-like protein)	8e-45	1
IpHdk01873	BE469486	Apolipoprotein E (lipid metabolism)	4e-13	2
IpHdk02687	BE468320	H-Fatty acid binding protein	1e-15	1
IpHdk03187	BE469169	Arp2/3 protein complex (regulator of actin polymerization) subunit p20 (ARC20)	2e-80	1
IpHdk02642	BE468311	ARP2/3 complex subunit p21-Arc	2e-42	2
IpHdk01507	BE469246	ARP2/3 protein complex subunit 34	8e-11	1
IpHdk01098	BE469170	ARP2/3 protein complex subunit p41-Arc (ARC41)	1e-32	1
IpHdk03169	BE469171	Aryl hydrocarbon receptor-interacting protein	8e-23	1
IpHdk02575	BE468292	Cortical granule lectin	6e-04	1
IpHdk02046	BE469586	SKB1 [regulator of p21(Cdc42/Rac)-activated kinase]	9e-04	1
IpHdk02640	BE468310	Phosphatidylethanolamine binding protein	4e-06	1
IpHdk01046	BE469172	Proteasome activator subunit 1	7e-11	1
IpHdk01656	BE469338	Proteasome activator subunit 2	1e-12	1
IpHdk02584	BE468295	Protein kinase C inhibitor	4e-05	1
IpHdk01421	BE469173	RNAse L inhibitor	2e-24	1
IpHdk00751	BE468865	Small protein F (LSM6)	7e-32	1
IpHdk02554	BE468288	Syndecan (transmembrane heparan sulphate proteoglycan) binding protein	6e-05	1
IpHdk01003	BE469174	Ubiquitin	e-112	1
IpHdk03100	BE469175	Ubiquitin-like protein	2e-08	1
IpHdk00673	BE468796	Polyubiquitin	e-115	1
IpHdk00558	BE468535	Guanine nucleotide binding protein	6e-15	1
IpHdk00554	BE468533	Guanylate binding protein 2, interferon-inducible(GBP2)	0.019	1
IpHdk01051	BE469176	GTP-binding protein (mSara)	6e-33	1
IpHdk02012	BE469587	GTP-binding protein (rhoB)	6e-12	2
IpHdk00925	BE469588	GTP-binding protein (rhoC)	1e-99	2
IpHdk02229	BE468547	Cell division cycle 42	6e-79	1
IpHdk02325	BE468565	FK506 binding protein	1e-05	2
IpHdk01506	BE469245	14-3-3 protein beta	1e-28	3
IpHdk03077	BE469177	14-3-3 epsilon	e-129	1
IpHdk00317	BE468479	Apoptosis-related protein PNAS-2 (PNAS-2)	6e-25	1
IpHdk00635	BE468977	ARF-like protein 1	1e-40	1
IpHdk00767	BE468875	Rho GDP dissociation inhibitor alpha	3e-28	2
IpHdk02503	BE468376	GDP dissociation inhibitor 1	8e-88	1
IpHdk00193	BE469178	Transducin (G-protein) beta-2 subunit	4e-37	1
IpHdk00225	BE469179	Diazepam-binding inhibitor/acyl-CoA-binding protein	3e-04	1
IpHdk01406	BE469180	Diphtheria toxin resistance protein required for diphthamide biosynthesis	1e-10	1
IpHdk01267	BE469181	Fragile X-related protein	8e-61	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk03158	BE469182	Huntingtin interacting protein K	2e-10	1
IpHdk00391	BE468501	Prohibitin (membrane-bound chaperone)	3e-29	1
IpHdk01254	BE469183	<i>Danio rerio</i> t-complex (chaperonins) polypeptide 1	e-145	3
IpHdk00721	BE468837	Cctd mRNA for chaperonin containing TCP-1 delta	2e-69	3
IpHdk00717	BE468833	Up-regulated by 1,25-dihydroxyvitamin Cctd mRNAfor chaperonin containing D-3	4e-06	1
IpHdk00198	BE469184	Chaperonin Cct6	2e-23	3
IpHdk02321	BE468564	Chaperonin subunit 7 (eta) (Cct7)	3e-37	1
IpHdk01391	BE469185	Prefoldin 2 (chaperone)	9e-08	1
IpHdk01998	BE469526	Cellular nucleic acid binding protein	4e-06	1
IpHdk00768	BE468876	Cytotoxic granule-associated RNA-binding protein	1e-9	1
IpHdk01203	BE469186	DNA-binding protein, TAXREB107	1e-13	1
IpHdk01303	BE469187	TAR DNA binding protein	8e-08	1
IpHdk02623	BE468304	Oocyte-specific histone stem-loop binding protein	4e-14	1
IpHdk00264	BE469188	Poly(A)-binding protein	3e-35	1
IpHdk02553	BE468287	Poly(rC) binding protein 2	2e-45	1
IpHdk00383	BE468498	Purine-rich element binding protein B	1e-40	2
IpHdk01091	BE469189	RNA binding motif protein RBM1	5e-13	1
IpHdk01856	BE469474	Ro ribonucleoprotein-binding protein 1	3e-13	1
IpHdk02155	BE469589	DAZ-associated protein 1	4e-21	1
IpHdk00522	BE468524	Transcription factor BTF 3a	7e-76	2
IpHdk02583	BE468294	Transcription factor BTF 3b	3e-45	1
IpHdk00203	BE469190	Transcription factor TFIIB	6e-27	1
IpHdk02569	BE468291	RNA polymerase II elongation factor SIII, p15 subunit	6e-22	1
IpHdk01857	BE469475	Zinc finger transcription factor 207	5e-46	1
IpHdk01966	BE469519	Novel Sp1 like zinc finger transcription factor	7e-32	1
IpHdk00101	BE469191	Arginine/serine-rich 1 (alternate splicing factor)	4e-44	1
IpHdk00289	BE469192	Arginine/serine-rich 2 (splicing factor)	8e-15	1
IpHdk01954	BE469518	Splicing factor Prp8	5e-88	1
IpHdk02033	BE469590	Splicing factor 3a, subunit 266 kDa (Sf3a2)	2e-45	1
IpHdk02016	BE469591	Splicing factor 3a, subunit 360 kDa (SF3a3)	4e-31	2
IpHdk01638	BE469322	Ganglioside expression factor 2 (GEF-2)	6e-55	1
IpHdk01730	BE469391	Human splicing factor CC1.3	1e-05	1
IpHdk01043	BE469193	Myelin gene expression factor 2	2e-45	2
IpHdk00139	BE469194	Negative regulator of transcription 2	1e-37	1
IpHdk01009	BE469195	Nuclear transport factor 2	2e-18	1
IpHdk01280	BE469196	Transcriptional coactivator (CRSP34)	3e-51	1
IpHdk02188	BE469592	Transcriptional coactivator ALY	6e-34	1
IpHdk01349	BE469197	Single-strand D-D-box binding factor	4e-38	1
IpHdk01113	BE469198	Adenylate cyclase activating polypeptide	8e-10	1
IpHdk01513	BE469252	Adenylyl cyclase-associated protein (CAP)	8e-14	1
IpHdk00443	BE468509	Regulator of G-protein signalling 2	1e-05	1
IpHdk00106	BE469199	Y box protein 1	8e-92	10
IpHdk00114	BE469200	Suppressor of hairless protein 1	e-103	1
IpHdk01791	BE469439	Delta sleep inducing peptide (transcription regulator)	2e-08	1
IpHdk03050	BE469201	CALIFp (component of global regulator of RNA polymerase II transcription)	4e-35	1
IpHdk03190	BE469202	Sin3-associated polypeptide 18	8e-11	1
IpHdk00724	BE468840	c-myb	8e-69	1
IpHdk01220	BE469203	vav 1 oncogene	4e-04	1
IpHdk02635	BE468309	Block of proliferation 1 (Bop1)	7e-17	1
IpHdk00091	BE469204	Fau	3e-28	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
IpHdk00399	BE468503	flt4 protein (member of class III receptor tyrosine kinases)	0.16	1
IpHdk00742	BE468857	RAB5C, member of RAS oncogene family	4e-42	1
IpHdk00236	BE469205	Ras-related protein Rab7	1e-27	1
IpHdk01510	BE469249	RAD23b homologue (Rad23b)	1e-15	1
IpHdk03194	BE469206	Rap1b	0.0	2
IpHdk01330	BE469207	Rasa3 (RAS p21 protein activator 3)	4e-07	1
IpHdk01592	BE469299	Ras-related C3 botulinum toxin substrate 2	1e-65	3
IpHdk01244	BE469208	Hepatocellular carcinoma associated ring finger protein	3e-40	1
IpHdk03017	BE469209	Colin carcinoma laminin-binding protein	2e-79	1
IpHdk03124	BE469210	Ras-related GTP-binding protein involved in membrane traffic	5e-22	1
IpHdk03054	BE469211	Ras-related nuclear protein Ran	0.0	2
IpHdk02577	BE468293	Ras-related protein (Rac2)	6e-50	2
IpHdk02015	BE469593	v-Fos transformation effector	e-153	2
IpHdk00079	BE469212	LIM and SH3 domain-containing protein Lasp-1	2e-20	3
IpHdk03109	BE469213	von Hippel-Lindau binding protein 1	6e-06	1
IpHdk01373	BE469214	12S Ribosomal RNA	8e-71	1
IpHdk00327	BE468484	Cytochrome <i>b</i>	e-129	15
IpHdk01099	BE469215	Cytochrome <i>c</i>	2e-07	1
IpHdk02549	BE468363	NADH dehydrogenase subunit 1	5e-12	1
IpHdk01358	BE469216	NADH dehydrogenase (ubiquinone) 1 alpha	3e-10	1
IpHdk00357	BE468492	NADH dehydrogenase subunit 2	2e-43	4
IpHdk00540	BE468529	NADH dehydrogenase subunit 3	5e-29	2
IpHdk01337	BE469971	NADH dehydrogenase 5 and subunit 6 genes	2e-39	2
IpHdk01204	BE469217	NADH ubiquinone oxidoreductase subunit 4 L	4e-59	4
IpHdk03152	BE469218	NADH dehydrogenase (51 kd subunit)	5e-28	1
IpHdk01063	BE469219	Flavocytochrome <i>b</i> subunit	5e-15	1
IpHdk00328	BE468485	Cytochrome <i>c</i> oxidase subunit I	1e-74	10
IpHdk00356	BE468491	Cytochrome <i>c</i> oxidase II	2e-45	14
IpHdk00683	BE468803	Cytochrome <i>c</i> oxidase III	4e-46	7
IpHdk02238	BE468551	Cytochrome <i>c</i> oxidase subunit IV isoform 1	6e-21	2
IpHdk02206	BE468543	Cytochrome P-450	5e-06	1
IpHdk00755	BE468867	Antioxidant protein	1e-9	2
IpHdk00754	BE468866	t-RNA VAL gene	0.0	1
IpHdk02262	BE468557	<i>I. punctatus</i> tRNA-Thr, tRNA-Pro, D-loop; and tRNA-Phe genes	0.0	1
IpHdk01443	BE469220	Adrenal gland protein AD-005	3e-24	1
IpHdk01594	BE469301	<i>Danio rerio</i> LINE-like DNA	1e-15	2
IpHdk01746	BE469402	LINE-like DNA	8e-08	1
IpHdk01525	BE469261	<i>Heterodontus japonicus</i> HE1 SINE	1e-06	1
IpHdk00920	BE469594	<i>D. melanogaster</i> genomic scaffold	2e-05	1
IpHdk02036	BE469595	<i>D. melanogaster</i> genomic scaffold 142000013386035 section 15 of 105	1e-24	3
IpHdk03220	BE469221	<i>D. melanogaster</i> genomic scaffold 142000013386045 section 5 of 17	4e-13	1
IpHdk03032	BE469222	<i>D. melanogaster</i> genomic scaffold 142000013386046 section 2 of 16	9e-04	1
IpHdk00962	BE469596	<i>D. melanogaster</i> genomic scaffold 142000013386055	2e-23	1
IpHdk01390	BE469223	<i>D. melanogaster</i> genomic scaffold 142000013386055 section 20 of 63	1e-25	3
IpHdk02223	BE468546	Human 12p11-37.2-54.4 BAC RP11-1060J15	5e-09	1
IpHdk02243	BE468552	Human 12p13.3 BAC RPC11-543P15	3e-37	1

Table 1 (Contd.)

Clone no.	Accession no.	Identity	P-value	No. of clones
lpHdk01190	BE469225	Human cDNA DKFZp434K0	5e-58	1
lpHdk03080	BE469226	Human cDNA DKFZp434K0521	5e-06	1
lpHdk00685	BE468805	Human cDNA FLJ10309 fis, clone NT2RM2000287	4e-15	1
lpHdk01317	BE469227	Human cDNA FLJ20541 fis, clone KAT11364	2e-07	1
lpHdk02692	BE468323	Human CGI-110 protein	8e-50	1
lpHdk01799	BE469447	Human CGI-136 protein	2e-08	1
lpHdk00257	BE469228	Human CGI-141 protein	1e-17	1
lpHdk00401	BE468504	Human CGI-27 protein	6e-26	1
lpHdk01844	BE469464	Human CGI-35 protein	5e-33	1
lpHdk00398	BE468502	Human chromosome 5 clone CITB-H1-2018B2	7e-04	1
lpHdk01553	BE469269	Human chromosome 7 cloneRP11-351B12	5e-14	1
lpHdk01697	BE469362	Human chromosome 7 clone RP11-264E12	2e-11	1
lpHdk03155	BE470502	Human cDNA: FLJ21016 fis, clone CAE05735	1e-19	1
lpHdk00594	BE468542	Human chromosome 8 map 8p23-p22 clones CTB-164D9, etc.	2e-11	2
lpHdk02307	BE468563	Human clone RP11-340F1 from 7p1415	2e-11	1
lpHdk00268	BE469229	Human clone RP4-561L24	3e-04	1
lpHdk00182	BE469230	Human DNA sequence from clone 1052 M	3e-14	1
lpHdk03223	BE469231	Human DNA sequence from clone 316G316G12	2e-15	1
lpHdk02091	BE469597	Human HSPC307	2e-42	1
lpHdk01727	BE469388	Human hypothetical protein (LOC51318)	2e-08	1
lpHdk00591	BE468541	Human hypothetical protein FLJ10099	2e-11	1
lpHdk03055	BE470464	Human hypothetical protein FLJ20059	2e-09	1
lpHdk03165	BE469232	Human PRO2446	2e-45	1
lpHdk01301	BE470358	Homologue of human and rodent ESTs AA454028, AA9255224 and AA023712	1e-07	1
lpHdk03069	BE469233	Human protein × 0004	5e-09	1
lpHdk01516	BE469255	Human putative 28 kd protein	9e-04	1
lpHdk02105	BE469759	sec13-like protein	3e-36	1
lpHdk00516	BE468522	<i>Ictalurus punctatus</i> NCCR-1 gene	4e-04	1
lpHdk01313	BE469234	KIAA0045 gene	2e-14	1
lpHdk00034	BE469235	KIAA0103 gene	3e-09	1
lpHdk00902	BE469598	KIAA0174 gene	7e-27	1
lpHdk01382	BE469236	KIAA1249	3e-04	1
lpHdk01234	BE469237	KIAA1388 protein	6e-24	1
lpHdk01893	BE469503	Mouse 10–11 day cDNA AK013196	5e-10	1
lpHdk00229	BE469238	Mouse brain cDNA, clone MNCb-2831	1e-12	1
lpHdk02043	BE469599	<i>X. laevis</i> cDNA clone 19E1-2	9e-61	1
lpHdk00043	BE469239	<i>X. laevis</i> cDNA clone 5O2	4e-40	2

Results and discussion

Gene and orthologue identification

To identify genes expressed in the channel catfish head kidney and establish their expression profiles, we sequenced 2228 ESTs from a non-normalized cDNA library of the channel catfish head kidney. Of the 2228 clones, 1495 (67.1%) were identified as orthologues of known genes from other organisms. The remaining 733 (32.9%) clones could not be identified by similarity comparisons.

Gene annotations were conducted using BLASTN for nucleotide similarity comparisons followed by BLASTX for amino acid similarity comparisons (Fig. 1). BLASTN comparisons established 1248 clones as orthologues of known genes, representing 529 unique transcripts (Table 1). In some cases, significant similarities existed between the EST clones and sequences in GenBank, but examination of the homologous region indicated that the significant similarities were the result of simple sequences. In those cases, the EST clones were regarded as unknown clones.

Table 2 Orthologues of known genes identified by BLASTX searches.

Clone no.	Accession no.	Putative identity	P-value	F
IpHdk03217	BE470528	Human hypothetical gene/a <i>C. elegans</i> protein encoded in cosmid K12D12(Z49069)	2e-96	1
IpHdk00608	BE468960	Ribosomal protein L5	7e-91	2
IpHdk00104	BE470171	Leucine aminopeptidase	2e-86	1
IpHdk01438	BE470408	ARL-6 interacting protein	2e-84	2
IpHdk01068	BE470268	Human CGI-10 protein	2e-83	1
IpHdk01405	BE470393	Neutrophil cytosolic factor 4	1e-79	1
IpHdk00770	BE468878	Isoleucyl-tRNA synthetase	4e-78	1
IpHdk00385	BE468732	Cromodomains helicase DNA binding protein 2	4e-77	1
IpHdk01193	BE470313	Uridine monophosphate synthase	1e-76	2
IpHdk00227	BE470217	DNA Polymerase iota	1e-74	1
IpHdk02306	BE468944	ERO1-like protein	5e-73	1
IpHdk03148	BE470498	Vitamin k-dependent plasma protein S (alpha)	3e-68	1
IpHdk02522	BE468343	NADH dehydrogenase 1 beta subcomplex, 5 (16kd, SGDH)	3e-65	1
IpHdk01946	BE469534	p47-phox/leukaemia-related protein	8e-64	1
IpHdk00915	BE469696	Lysozyme G (1,4-beta-n-acetylmuramidase)	9e-64	3
IpHdk02146	BE469775	Reverse transcriptase-like protein	2e-63	1
IpHdk01368	BE470380	Phospholipid scramblase 1	4e-62	1
IpHdk01223	BE470325	Translation initiation factor EIF-2B gamma	5e-62	1
IpHdk01516	BE469255	Putative 28 kDa protein	6e-59	1
IpHdk00261	BE470231	KIAA1475 protein/deep orange protein of <i>Drosophila</i>	4e-55	1
IpHdk01904	BE469509	Human hypothetical protein HSPC014	7e-55	3
IpHdk02034	BE469730	Heterogeneous nuclear ribonucleoprotein A0	3e-54	1
IpHdk00693	BE468812	Human KIAA1286 protein	1e-53	1
IpHdk01250	BE470338	Disulphide isomerase-related protein	9e-53	1
IpHdk02248	BE468920	tRNA selenocysteine associated protein	1e-52	1
IpHdk00524	BE468882	92 Kd type IV collagenase precursor (matrix metalloproteinase 9)	2e-50	1
IpHdk01308	BE470361	Mitochondrial processing peptidase-beta	3e-50	1
IpHdk01363	BE470379	Annexin III	8e-50	1
IpHdk01554	BE469270	Sirtuin type 5	3e-49	1
IpHdk00595	BE468903	RNA binding motif protein 4	5e-47	2
IpHdk00452	BE468762	Haeme-oxygenase (HO)	6e-47	1
IpHdk01942	BE469531	<i>C. elegans</i> hypothetical protein C13C4.4	1e-46	1
IpHdk00953	BE469710	Human KIAA0804 protein	2e-46	1
IpHdk00679	BE468799	Lysozyme precursor	6e-45	2
IpHdk00914	BE469695	<i>Drosophila</i> CG6567 gene	1e-44	1
IpHdk02169	BE469783	Papillomavirus regulatory factor PRF-1	4e-42	1
IpHdk03201	BE470523	Human hypothetical protein AL133097 similar to interleucin 17 receptor	4e-42	1
IpHdk01028	BE470252	Human HDCMC29P	5e-42	1
IpHdk02181	BE469787	Dihydroxyvitamin D3-induced protein	6e-41	1
IpHdk03057	BE470466	WW domain binding protein 2	6e-41	2
IpHdk00186	BE470466	Corneal wound healing related protein	3e-40	1
IpHdk00538	BE468886	Human WW domain binding protein-2	4e-39	1
IpHdk02331	BE468954	Similar to <i>S. cerevisiae</i> kti12 protein	7e-39	1
IpHdk00772	BE468880	Human KIAA0849 protein	8e-39	1
IpHdk02110	BE469762	Haemoglobin alpha chain	1e-38	1
IpHdk00514	BE468788	Lamina associated polypeptide 1C short splice form	2e-38	1
IpHdk00384	BE468731	54 Kd nuclear RNA-binding protein	2e-38	1
IpHdk01389	BE470388	Gelatinase B	3e-38	1
IpHdk01074	BE470271	Cytochrome c oxidase polypeptide VB (VI)	6e-38	1
IpHdk02296	BE468939	Xenopus RPA interacting protein alpha	3e-37	1

Table 2 (Contd.)

Clone no.	Accession no.	Putative identity	P-value	F
IpHdk01283	BE470349	MEGF6 protein (protein with multiple EGF-like motifs)	4e-37	1
IpHdk01058	BE470263	Zinc finger protein 147 (estrogen-responsive protein)	4e-37	1
IpHdk01282	BE470348	CD63 antigen (AD 1 antigen)	6e-37	2
IpHdk02308	BE468945	<i>Drosophila</i> CG10050 gene	7e-37	1
IpHdk03221	BE470530	<i>C. elegans</i> hypothetical protein ZC308.1	2e-36	1
IpHdk02674	BE468450	Lanthionine synthetase C-like protein 1	6e-36	1
IpHdk02219	BE468912	Putative chitinase precursor	1e-35	1
IpHdk02580	BE468394	Sex-regulated protein Janus-a	4e-34	1
IpHdk01409	BE470396	Human hypothetical protein DKFZP564M2423	5e-34	2
IpHdk00382	BE468730	<i>Drosophila</i> CG12792 gene	2e-33	1
IpHdk03013	BE470447	Ring finger protein (C3HC4 type)	2e-33	1
IpHdk00172	BE470196	Arginase	4e-33	2
IpHdk00732	BE468847	Ribosomal protein L35a	9e-33	2
IpHdk02187	BE469788	Estrogen-responsive finger protein	2e-32	1
IpHdk02086	BE469754	Human KIAA1495 protein (leucine-rich neuronal protein)	4e-32	1
IpHdk02048	BE469737	Fetuin-like protein IRL685	2e-31	1
IpHdk02009	BE469725	Zinc-binding protein A33	1e-30	1
IpHdk00093	BE470165	Adrenodoxin/proferredoxin	2e-30	1
IpHdk02604	BE468405	Retinol dehydrogenase type III	2e-30	1
IpHdk02038	BE469731	Centrin 3	3e-30	1
IpHdk01240	BE470333	Tyrosylprotein sulphotransferase 2	3e-30	1
IpHdk00420	BE468746	Human KIAA1194 protein	3e-30	1
IpHdk00739	BE468854	Human hypothetical protein FLJ20446	6e-30	1
IpHdk01295	BE470355	Collagenase 3 precursor (matrix metalloproteinase-13)	9e-30	5
IpHdk00286	BE470240	Cold inducible RNA-binding protein	1e-29	1
IpHdk00095	BE470166	Eosinophil peroxidase	2e-29	2
IpHdk00497	BE468779	Ankyrin repeat-containing protein	3e-29	1
IpHdk01936	BE469527	Hepatocyte growth factor activator	3e-29	1
IpHdk00510	BE468785	Human HSPC280	4e-29	1
IpHdk02153	BE469778	SRP20 (arginine/serine-rich 3 splicing factor)	5e-29	1
IpHdk02143	BE469773	IgD B-cell receptor-associated protein (BAP)	5e-29	1
IpHdk00123	BE470177	Human KIAA0969 protein	5e-29	1
IpHdk03042	BE470461	Human KIAA1438 protein	1e-28	1
IpHdk00980	BE469720	Dopamine receptor/systatin B	2e-28	1
IpHdk00521	BE468881	Cytochrome <i>c</i> oxidase polypeptide VIIa-heart precursor	3e-28	1
IpHdk02279	BE468928	Human adrenal gland protein AD-002	1e-27	1
IpHdk01222	BE470324	Estrogen-responsive B box protein	2e-27	1
IpHdk02565	BE468386	Human gene HSPC337	3e-27	1
IpHdk02539	BE468356	Glycosyl-phosphatidyl-inositol-anchored protein	2e-26	1
IpHdk02557	BE468379	TBX1 protein	3e-26	2
IpHdk01456	BE470415	S100-like calcium binding protein	1e-25	1
IpHdk01381	BE470386	<i>Drosophila</i> CG3224 gene	2e-25	1
IpHdk00453	BE468763	Human KIAA1324 protein	2e-25	1
IpHdk02686	BE468460	SH3 domain-containing adapter protein	4e-25	1
IpHdk03166	BE470506	Transcription factor	5e-25	1
IpHdk00439	BE468755	Synaptogyrin 2 (cellugyrin)	5e-25	1
IpHdk00320	BE468701	<i>N</i> -methyl-D-aspartate receptor glutamate-binding subunit	5e-25	1
IpHdk01801	BE469448	Invariant chain-like protein 2	9e-25	3
IpHdk03007	BE470443	Human KIAA1460 protein	6e-24	1
IpHdk00507	BE468783	Inhibitor of apoptosis protein	7e-24	1
IpHdk00406	BE468743	Bromodomain containing protein	2e-23	1
IpHdk03173	BE470511	Molybdopterin-synthase large subunit	3e-23	1

Table 2 (Contd.)

Clone no.	Accession no.	Putative identity	P-value	F
IpHdk03198	BE470522	Guanine nucleotide binding protein (gamma-5 subunit)	5e-23	1
IpHdk00228	BE470218	Mitoxantrone-resistance associated protein	6e-23	2
IpHdk01425	BE470402	Signal peptidase (12 kDa)	7e-23	1
IpHdk03210	BE470526	FYN/EVH1 domain binding protein	1e-22	1
IpHdk00722	BE468838	Acidic ribosomal protein P2	1e-22	1
IpHdk02677	BE468452	Cytochrome c oxidase polypeptide VIIA-liver form	5e-22	1
IpHdk00736	BE468851	DEAD/Helicase box polypeptide 21	1e-21	1
IpHdk01151	BE470297	ATPase inhibitor precursor	2e-21	1
IpHdk02591	BE468398	Human KIAA0826 protein	2e-21	1
IpHdk02302	BE468943	Transthyretin precursor	3e-21	1
IpHdk03056	BE470465	Serum inducible protein kinase	6e-20	1
IpHdk00422	BE468747	Human HSPC036 protein	6e-20	1
IpHdk00202	BE470211	Adenylyl cyclase-associated protein (CAP 1)	4e-20	1
IpHdk02149	BE469777	Protoporphyrinogen oxidase	4e-20	1
IpHdk02167	BE469781	Homeodomain interacting protein kinase 1	4e-20	2
IpHdk03008	BE470444	Human HSPC095	1e-19	1
IpHdk03107	BE470484	Ras homologue gene family, member G (rho G)	4e-19	1
IpHdk02699	BE468468	Chain F, cytochrome Bc1 Complex	7e-19	1
IpHdk02287	BE468932	Homologue of mammalian thyroid peroxidase	7e-19	1
IpHdk00951	BE469709	Human hypothetical protein FLJ10856	1e-18	1
IpHdk01845	BE469465	<i>Drosophila</i> CG10343 gene	2e-18	1
IpHdk00432	BE468753	Annexin A3	3e-18	1
IpHdk01902	BE469507	Acyl-coenzyme A binding protein	7e-18	1
IpHdk01888	BE469499	Cathepsin G/natural killer cell protease 1	2e-17	1
IpHdk00719	BE468835	Plasmodium falciparum hypothetical protein CO39OW	3e-17	1
IpHdk00775	BE468997	Caspase-10 (interleukin-1b-converting enzyme 2)	6e-17	1
IpHdk00947	BE469707	DNA-binding protein BMI-1 (zinc finger protein)	6e-17	1
IpHdk02698	BE468467	Human KIAA1376 protein	9e-17	1
IpHdk00154	BE470187	Collagenase 4 precursor (matrix metalloproteinase-3)	1e-16	2
IpHdk00588	BE468901	Profilin	4e-16	1
IpHdk00791	BE469010	Human unnamed protein AK023972	4e-16	1
IpHdk01094	BE470280	Human KIAA0582 protein	9e-16	1
IpHdk00940	BE469704	High affinity immunoglobulin gamma Fc receptor I	1e-15	1
IpHdk00347	BE468714	Splicing factor, arginine/serine-rich 11	1e-15	1
IpHdk03067	BE470472	Human CGI-145 protein	1e-15	1
IpHdk01597	BE469304	Putative G protein-coupled receptor	5e-15	1
IpHdk00180	BE470202	Methionyl tRNA synthetase	5e-15	2
IpHdk00216	BE470216	<i>Drosophila melanogaster</i> CG7671	1e-14	1
IpHdkF00468	BE468769	Centaurin-alpha2 protein	1e-14	2
IpHdk01149	BE470296	GAP-associated tyrosine phosphoprotein p62	1e-14	1
IpHdk00293	BE470242	Interleucin-2 receptor gamma	3e-14	1
IpHdk00161	BE470190	Arginyl-tRNA synthetase	5e-14	1
IpHdk01084	BE470276	Human hypothetical protein AK001663	5e-14	1
IpHdk01860	BE469478	Growth modulatory factor granulin-3	6e-14	1
IpHdk03184	BE470518	Human CGI-55 protein	9e-14	1
IpHdk02163	BE469780	Human hypothetical protein BM-004 (AF208846)	1e-13	1
IpHdk03003	BE470441	Human hypothetical protein AF208846/onzin	1e-13	2
IpHdk02179	BE469786	Serum amyloid P component	2e-13	1
IpHdk02615	BE468413	Coronin homologue	3e-13	1
IpHdk01241	BE470334	L-threonine 3-dehydrogenase	3e-13	1
IpHdk02542	BE468358	Interferon inducible Mx protein	3e-13	1

Table 2 (Contd.)

Clone no.	Accession no.	Putative identity	P-value	F
IpHdk03095	BE470478	<i>Xenopus laevis</i> NO27	4e-13	1
IpHdk01395	BE470389	Ribosomal protein L39	6e-13	1
IpHdk00350	BE468715	MHC class II transactivator CIITA	1e-12	1
IpHdk00776	BE468998	Enhancer-of-split and hairy-related protein 2	3e-12	1
IpHdk00653	BE468993	Formin-like protein	1e-11	1
IpHdk02678	BE468453	Human KIAA0101 gene	2e-11	1
IpHdk01145	BE470293	Human hypothetical protein FLJ11296	2e-11	1
IpHdk00409	BE468744	NADH dehydrogenase 1 beta subcomplex, 3 (12 kDa, B12)	3e-11	1
IpHdk01785	BE469433	Adenylyl cyclase-associated protein (CAP 2)	3e-11	1
IpHdkF00312	BE468698	Fizzy1	3e-11	1
IpHdk00060	BE470148	Spermidine/spermine N1-acetyl transferase	4e-11	1
IpHdk03180	BE470516	Lysosomal membrane glycoprotein 1	4e-11	1
IpHdk00076	BE470157	Interferon-induced GTP-binding protein Mx3	6e-11	1
IpHdk01238	BE470331	Lectin, chain A, mannose-specific agglutinin (Lectin)	6e-11	1
IpHdk01249	BE470337	TGF- β inducible early growth response 2	9e-11	1
IpHdk01861	BE469479	CD94-B protein/killer cell lectin-like receptor	1e-10	1
IpHdk01158	BE470298	CXC chemokine K60 protein	1e-10	1
IpHdk02598	BE468402	Unnamed human protein BAB15362	2e-10	1
IpHdk01042	BE470256	B-cell receptor-associated protein 37	2e-10	1
IpHdk02144	BE469774	Human hypothetical protein DKFZp434K114.1	2e-10	1
IpHdk02078	BE469749	LDL-phospholipase A2	2e-10	1
IpHdk00083	BE470159	Nuclear RNA helicase	3e-10	1
IpHdkF00477	BE468773	Peripheral benzodiazapine receptor	3e-10	1
IpHdk00390	BE468734	Human HSPC071 protein	3e-10	1
IpHdk00392	BE468735	Beta-spectrin III	5e-10	1
IpHdk03114	BE470488	Tom (target of myb 1) protein	8e-10	1
IpHdk01251	BE470339	Ataxia-telangiectasia group D-associated protein	1e-09	1
IpHdk01901	BE469506	Lysosome-associated membrane glycoprotein 1 precursor	1e-09	1
IpHdk03206	BE470525	<i>Drosophila</i> CG8447	1e-09	1
IpHdk02139	BE469771	Human hypothetical protein FLJ10460	2e-09	1
IpHdkF00332	BE468704	Phospholipase D3	3e-09	1
IpHdk02282	BE468930	Sema domain, and GPI membrane anchor (semaphorin) 7 A	3e-09	1
IpHdk02245	BE468918	Phospholipase A2 inhibitor gamma subunit B	5e-09	1
IpHdk03023	BE470452	<i>Drosophila</i> Pebble gene	5e-09	1
IpHdk02669	BE468446	Neurochondrin-2	6e-09	1
IpHdk01619	BE469293	Immunoglobulin heavy chain	1e-08	1
IpHdk00603	BE468956	Nuclear factor 7	3e-8	1
IpHdk03058	BE470467	CC chemokine-1	3e-08	1
IpHdk01265	BE470344	Poly(A)-binding protein	5e-08	1
IpHdk01321	BE470363	Human KIAA0685 gene	2e-07	1
IpHdk00749	BE468863	Leucocyte surface antigen CD53	2e-7	1
IpHdk01889	BE469500	Thyroid receptor interacting protein 7	2e-07	1
IpHdk00637	BE468979	Human hypothetical protein 384D8-7	3e-7	1
IpHdk00680	BE468800	Small inducible cytokine subfamily A, member 21	3e-7	1
IpHdk00919	BE469698	Human unnamed protein AK022732	4e-07	1
IpHdk02285	BE468931	Mouse unnamed protein	5e-07	1
IpHdk01325	BE470367	Chloride intracellular channel protein 1	6e-07	1
IpHdk01134	BE470292	<i>Drosophila</i> CG14977 gene	7e-07	1
IpHdk01408	BE470395	H-REV107 protein-related protein	7e-07	1
IpHdk02633	BE468421	Mouse FISH protein	1e-06	1
IpHdk01093	BE470279	Wiskott-aldrich syndrome-associate protein	1e-06	1
IpHdk01022	BE470250	Human CGI-83 protein <i>Drosophila</i> CG12375 protein	1e-06	1

Table 2 (Contd.)

Clone no.	Accession no.	Putative identity	P-value	F
IpHdk01586	BE469280	Arginine/serine-rich splicing factor 2	3e-06	1
IpHdk02215	BE468910	Heat shock protein 108	7e-06	1
IpHdk01795	BE469443	Protein-tyrosine kinase	7e-06	1
IpHdk01432	BE470405	Transposase	9e-06	1
IpHdk02513	BE468334	Parathyroid hormone	1e-05	1
IpHdk00075	BE470156	Transmembrane protein TMP	1e-05	1
IpHdk02130	BE469765	Ribonucleoprotein (RNA-binding protein)	1e-05	1
IpHdk00682	BE468802	P19 protein, homologous to p16ink4	1e-5	1
IpHdk00434	BE468754	Spi-1/PU.1 transcription factor	1e-05	1
IpHdk01596	BE469303	<i>Drosophila</i> CG12746 gene	1e-05	1
IpHdk01291	BE470353	MMTV receptor variant 2	5e-05	1
IpHdk00267	BE470233	CAGH44	7e-05	1
IpHdk01288	BE470352	Human KIAA0127 gene	7e-05	1
IpHdk02546	BE468362	Small inducible cytokine b10 precursor	8e-05	1

All unknown EST clones after BLASTN searches were further analysed by BLASTX searches. The ESTs with significant similarities ($P < 1 \times 10^{-4}$) to known proteins were evaluated to determine if the significant similarities were caused by simple amino acid matches. The BLASTX searches identified orthologues of an additional 247 clones representing 219 unique genes (Table 2). Thus, a total of 748 orthologues of known genes were identified in this EST analysis. The remaining 733 EST clones represent unknown genes. Because most of our ESTs were long with very few unambiguous readings, a large proportion of these unknown EST clones may represent novel genes, but the actual number of genes is unknown at this time.

The proportion of known genes obtained in head kidney was comparable with that from channel catfish brain (Ju *et al.* 2000). The fact that the majority of EST clones (over two-thirds) could be identified by similarity comparisons suggests that high-quality EST analysis is an efficient way for gene annotation in less-well studied species. Additionally, some of the unknown clones may be resolved with an analysis using variable stringency parameters. On the other hand, caution should be exercised in functional annotations as a homologous gene may not be the counterpart gene in different organisms. Towards this end, the 219 genes identified by BLASTX searches were sorted according to their significance of similarities (Table 2).

A large number of clones showed significant similarities to known sequences of unknown function from model systems such as human, mouse, zebrafish, *Drosophila*, *C. elegans*, and *Xenopus laevis*. Although functions are not yet known, their conservation in fish demonstrated the existence of many gene families through evolution. Once a

gene is characterized in any one of these species, comparative functional genomics will allow annotation to these orthologous genes.

Expression profile and the most abundantly expressed genes

Of the 1495 EST clones identified by BLASTN and BLASTX, 545 (36.5%) were singletons. Although redundancy will increase as the number of sequenced clones increases, the high percentage of singletons indicated that the complexity and coverage of this channel catfish head kidney cDNA library was good. As shown in Fig. 2, multiple clones were sequenced for 205 genes ranging from 2 to 44 clones per gene. The most abundant cDNA was β -actin, which was sequenced 44 times, accounting for almost 2% of the transcripts in the head kidney. The top 10 most abundant

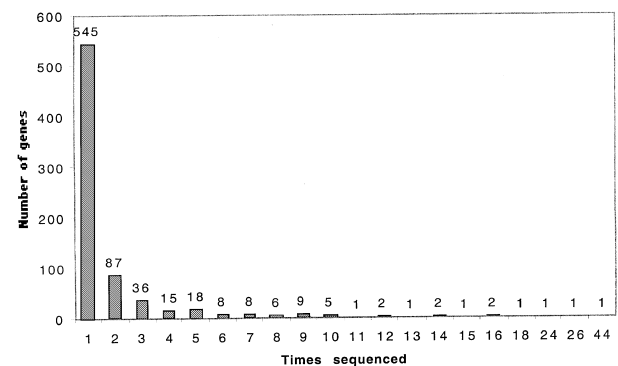


Figure 2 Sequencing profiles of 1495 known gene clones of the head kidney of channel catfish (*Ictalurus punctatus*). While 545 genes (36.5%) were singletons, the rest of the known genes were sequenced 2–44 times.

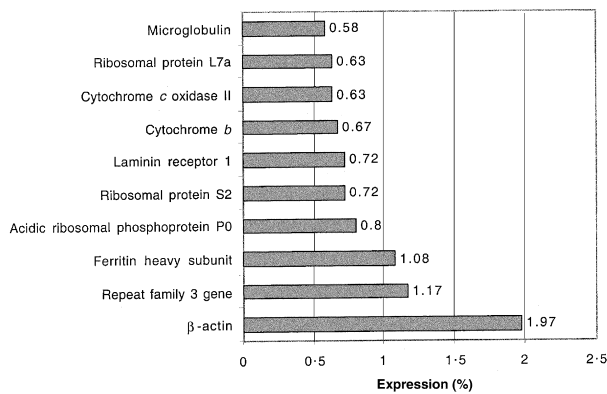


Figure 3 The 10 most abundantly expressed genes in the head kidney tissue of channel catfish, given as percentage expression.

ESTs were: β -actin (1.97%), repeat family 3 gene (1.17%), ferritin heavy subunit (1.08%), acidic ribosomal phosphoprotein P0 (0.8%), ribosomal protein S2 (0.72%), laminin receptor 1 (0.72%), cytochrome *b* (0.67%), cytochrome *c* oxidase II (0.63%), ribosomal protein L7a (0.63%), and microglobulin (0.58%). These top 10 genes accounted for almost 9% of the 2228 sequenced clones (Fig. 3). Among the top 10 most abundantly expressed genes, five were related to ribosomal proteins (P0, L7a, S2, repeat family 3 gene that showed high homology to S2, and laminin receptor 1 that showed homology to p40 ribosomal associated protein).

In conclusion, this work identified 748 known genes and 733 novel clones from the channel catfish head kidney. Expression profiles of these genes were revealed by their frequency in a cDNA library. The library was made using tissues collected before and after challenge with *E. ictaluri*, the causing agent of enteric septicaemia of catfish. This experiment design was adopted in order to include all potential transcripts in the cDNA library, particularly those induced after infection. Further studies using cDNA microarrays are needed to identify the differentially expressed transcripts after disease infection.

Acknowledgements

This project was supported by a USDA NRICGP (#98-35205-6738) and by the Auburn University Competitive BioGrant (Biogrant J. Liu 99). We appreciate the support of Auburn University Department of Fisheries and Allied Aquacultures, College of Agriculture, and the Vice President for Research for their matching funds to USDA NRI Equipment Grants (98-35208-6540, 99-35208-8512).

References

- Adams M.D., Kelley J.M., Gocayne J.D. *et al.* (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651–6.
- Azam A., Paul J., Sehgal D., Prasad J., Bhattacharya S. & Bhattacharya A. (1996) Identification of novel genes from *Entamoeba histolytica* by expressed sequence tag analysis. *Gene* **181**, 113–6.
- Chilmoczyk S. (1992) The thymus in fish: development and possible function in the immune response. *Annual Review of Fish Disease* **2**, 181–200.
- Chomczynski P. & Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry* **162**, 156–9.
- Dunham R.A., Brady Y. & Vinitnantharat S. (1993) Response to challenge with *Edwardsiella ictaluri* by channel catfish, *Ictalurus punctatus*, selected for resistance to *E. ictaluri*. *Journal of Applied Aquaculture* **3**, 211–22.
- Ellis A.E. (1992) Fish immunology. In: *Encyclopedia of Immunology* (Eds by L.M. Roitt & P.J. Delves), Vol. III, pp. 536–9. San Diego, CA, USA.
- Emahazion T., Jobs M., Howell W.M., Siegfried M., Wyoni P., Prince J.A. & Brooks A.J. (1999) Identification of 167 polymorphisms in 88 genes from candidate neurodegeneration pathways. *Gene* **238**, 315–24.
- Ermolaeva O., Rastogi M., Pruitt K.D., Schuler G.D., Bittner M.L., Chen Y., Simon R., Meltzer P., Trent J.M. & Boguski M.S. (1998) Data management and analysis for gene expression arrays. *Nature Genetics* **20**, 19–23.
- van Hal N.L., Vorst O., van Houwelingen A.M., Kok E.J., Peijnenburg A., Aharoni A., van Tunen A.J. & Keijer J. (2000) The application of DNA microarrays in gene expression analysis. *Journal of Biotechnology* **78**, 271–80.
- Hofte H., Desprez T., Amselem J., Chiappello H., Caboche M., Moisan A., Jourjon M.F., Charpentreau J.L., Berthomieu P. & Guerrier D. (1993) An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. *Plant Journal* **4**, 1051–61.
- Ju Z., Karsi A., Kocabas A., Patterson A., Li P., Cao D., Dunham R. & Liu Z.J. (2000) Transcriptome analysis of channel catfish (*Ictalurus punctatus*): genes and expression profile from the brain. *Gene* **261**, 373–82.
- Karsi A., Li P., Dunham R. & Liu Z.J. (1998) Transcriptional activities in the pituitaries of channel catfish before and after induced ovulation by injection of carp pituitary extract as revealed by expressed sequence tag analysis. *Journal of Molecular Endocrinology* **21**, 121–9.
- Liu Z.J., Karsi A. & Dunham R. (1999) Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. *Marine Biotechnology* **1**, 437–47.
- Moch H., Schraml P., Bubendorf L., Mirlacher M., Kononen J., Gasser T., Mihatsch M.J., Kallioniemi O.P. & Sauter G. (1999) High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. *American Journal of Pathology* **154**, 981–6.

- Pichappan R.M. (1980) Review on the phylogeny of splenic structure and function. *Development and Comparative Immunology* **4**, 395–416.
- Sambrook J., Frisch E.F. & Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger F., Nicklen S. & Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74**, 5463–7.
- Wang D.G., Fan J.B., Siao C.J. *et al.* (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**, 1077–82.
- Warr G.W. (1997) The adaptive immune system of fish. In: *Fish Vaccinology* (Eds by R. Gudding, A. Lillehaug, P.J. Midtlyng & F. Brown), pp. 15–21. Dev. Biol. Stand. Basel, Karger.
- Waterston R., Martin C., Craxton M. *et al.* (1992) A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genetics* **1**, 114–23.