Biochemical and Biophysical Research Communications **269,** 110–116 (2000)
doi:10.1006/bbrc.2000.2272, available online at http://www.idealibrary.com on IDEAL®

Comprehensive Gene Expression Profile of a Normal Human Liver

Taro Yamashita,*,† Shin-ichi Hashimoto,† Shuichi Kaneko,*,[1] Shigenori Nagai,† Nobuaki Toyoda,† Takuji Suzuki,† Kenichi Kobayashi,* and Kouji Matsushima†

**First Department of Internal Medicine & CREST, School of Medicine, University of Kanazawa, 13-1 Takara-Machi, Kanazawa, Ishikawa, 920-8641 Japan; and †Department of Molecular Preventive Medicine & CREST, School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan*

Received January 27, 2000

To investigate the gene expression profile of a normal human liver, we performed serial analysis of gene expression (SAGE), which allows the quantitative and simultaneous analysis of thousands of genes expressed in tissue. Polyadenylated RNA was obtained from a bulk normal human liver sample and SAGE was performed. Reverse transcriptase–polymerase chain reaction (RT-PCR) was also performed in each of 3 different normal liver samples to evaluate the validity of the profile in each individual. A total of 30,982 tags were sequenced, 8,596 of which were unique. The genes highly expressed in the normal liver were those encoding plasma proteins (>21.8% of total transcripts), cytoplasmic proteins (>8.6%), enzymes (>4.8%), protease inhibitors (>1.7%), complements (>1.1%), and coagulation factors (>0.75%). About 13.9% of all transcripts encoded genes not reported in GenBank thus far. This study identifies candidate genes to be examined in relation to various human liver diseases, including viral hepatitis, liver cirrhosis, and hepatocellular carcinoma.

The liver, the largest vital organ in the body, weighs about 1,500 grams and comprises one-fifteenth of the total adult body weight. There are 2×10^5 cells in each milligram of the liver and 1.7×10^5 of these are hepatic parenchymal cells. That is, there are about 2.6×10^{11} hepatic parenchymal cells in the normal adult liver. These large amounts of hepatocytes fulfill multiple functions, including plasma protein synthesis, carbohydrate metabolism, amino acid metabolism, lipid metabolism, drug metabolism, and bile acid secretion.

Abbreviations used: SAGE, serial analysis of gene expression; RT-PCR, reverse transcriptase–polymerase chain reaction; EST, expressed sequence tag.

[1] To whom correspondence and reprint requests should be addressed. Fax: 81-76-234-4250. E-mail: skaneko@medf.m.kanazawa-u.ac.jp.

Numerous studies have investigated the hepatic expression of various genes encoding plasma proteins, cytoplasmic proteins, and enzymes. Some genes are expressed throughout the liver lobule, whereas others exhibit heterogeneous expression patterns, and the involvement of transcriptional and posttranscriptional mechanisms in the regulation of specific genes has been speculated to reflect specialized function in specific hepatocyte subsets (1). Recent experiments have focused attention on the unique gene expression patterns in the liver lobule by means of Northern blot analysis and *in situ* hybridization. Several reports have suggested that, with respect to gene expression, the liver lobule can be divided into periportal domain and pericentral domain (2–6), and it is generally accepted that genes are expressed along a portcentral gradient in the liver (7, 8). These reports have demonstrated the importance of microenvironment in the expression of a limited number of genes in the liver. However, no comprehensive study of the expression of a large number of genes in the normal liver has been reported.

Serial analysis of gene expression (SAGE) is a recently developed technique for evaluating thousands of gene expressions quantitatively and simultaneously (9). SAGE can generate a transcript profile relying on 10 base cDNA sequences (SAGE tags) for gene identification and cloning. Thus far, several SAGE transcript profiles have been reported, including those for p53-induced apoptosis (10), and yeast transcriptome (11). In this study, we performed SAGE on a normal human liver to clarify the characteristics of its gene expression.

MATERIALS AND METHODS

Normal liver sample. A 31-year-old male normal human liver was resected surgically because of metastasis of colon cancer. The sample was obtained from the portion unaffected by the cancer and

0006-291X/00 $35.00

TABLE 1

Primers Used for RT-PCR

Albumin	Forward	5′-CGT CTG CCA AAC AGA GACT CA-3′
	Reverse	5′-AGA CAT CCT TTG CCT CAG CA-3′
α-1 antitrypsin	Forward	5′-TCA CCC ACG ATA TCA TCA CCA-3′
	Reverse	5′-TTG AGG AGC GAG AGG CAG TTA-3′
Aldolase B	Forward	5′-TAT GGC CAC CGT AAC AGC TCT-3′
	Reverse	5′-AAA AGT TGC TCC CTT TCA GCC-3′
Hemopexin	Forward	5′-TGG CCC ATT GCT CAT CAG T-3′
	Reverse	5′-AAT TGG GAC CAT GGA TGA GG-3′
Cytochrome P450 IIIA4	Forward	5′-TGA AAG AAA GTC GCC TCG AAG-3′
	Reverse	5′-CCA CCC TTT GGG AAT GAA CA-3′
1-8U gene	Forward	5′-ATC CCG ATT TGA CAA ATG CC-3′
	Reverse	5′-TCA CGT CGC CAA CCA TCT T-3′
Elongation factor 1 α	Forward	5′-TCA TTG ATG CCC CAG GAC A-3′
	Reverse	5′-ATT GCC ATC CTT ACG GGT GAC-3′
Translationally controlled tumor protein	Forward	5′-CCA CGA TGA GAT GTT CTC CGA-3′
	Reverse	5′-CCA TGC CAT CTG GAT TCA TGT-3′

frozen in liquid nitrogen. All procedures and risks were explained verbally and in a written consent form. The sample was sectioned and confirmed to be normal histologically. All laboratory data assessing hepatic function were within the normal range, including serum alanine aminotransferase, aspartate aminotransferase, γ-glutamyl transpeptidase, alkaline phosphatase, total bilirubin, albumin, prothrombin activity, glucose, cholesterol, and triglycerides (data are not shown). Serological tests of hepatitis B surface antigen, Hepatitis C virus antibodies, and Human immunodeficiency virus antibodies also showed negative. Neither heavy alcohol consumption nor the intake of chemical reagents was observed before surgical resection. Two other normal liver samples and cultured cells (HepG2) were used for RT-PCR. HepG2 cells were cultured and grown to 80–90% confluence in Dulbecco's modified minimum essential medium (Gibco BRL, Life Technologies, Rockville, MD) with 10% fetal bovine serum (Gibco BRL) under 5% CO_2/95% in a humidified incubator at 37°C.

TABLE 2

Highly Expressed Genes in the Liver

Tag sequence	Gene (GenBank Accession Number)	Count	Percent
AGAATAAGAG	Albumin (V00495)	1088	3.4774
TGGACGCGCT	Apolipoprotein A-I (M27875)	876	2.7998
TGGCCCCAGG	Apolipoprotein C-I (X00570)	772	2.4674
CTGGCCTCCC	Apolipoprotein C-III (X00567)	669	2.1382
CACCTAATTG	ATPase 6/8 (J01415)	468	1.4958
TGATTTCACT	Cytochrome oxidase 3 (J01415)	347	1.109
CCCATCGTCC	Cytochrome oxydase 2 (J01415)	346	1.1058
GGAAAAGTGG	Protease inhibitor 1, alpha-1 antitrypsin (K01396)	326	1.0419
ACTTTTTCAA	Cytochrome oxydase subunit 1 (J01415)	294	0.9396
CGACCCCACG	Apolipoprotein E (M12529)	294	0.9396
TTCATACACC	NADH dehydrogenase 4 (J01415)	292	0.9332
GCCGGGCCCT	S-protein, somatomedinB, vitronectin (X03168)	291	0.93
TGTACCTCAG	Orosomucoid-1 (M13692)	291	0.93
GTGGGCACCT	Plasma retinol binding protein (X00129)	274	0.8757
TGTGGAGAGC	Apolipoprotein A-II (M29882)	269	0.8597
GGCAACGGTA	α-1-microglobulin (X04494)	241	0.7702
CTAAGGTGGT	β-2-glycoprotein I, apolipoprotein H (X53595)	235	0.7511
TCTAAGTACC	Haptoglobin (M12387 X00442)	234	0.7479
CACTACTCAC	Cytochrome *b* (J01415)	227	0.7255
TAACCAAGAG	Transthyretin (K02091)	219	0.6999
CCCTGGGTTC	Ferritin light chain (M11147)	212	0.6775
CTAAGACTTC	16S rRNA, mitochondrial (J01415)	194	0.62
AGCCCTACAA	NADH dehydrogenase 3 (J01415)	189	0.604
ATTTGAGAAG	Cytochrome oxydase 1 (J01415)	163	0.5209
ACCCTTGGCC	NADH dehydrogenase 1 (J01415)	161	0.5145

Note. The top 25 transcripts expressed in the normal liver are listed. The tag sequence represents the 10-bp SAGE tag. Probable GenBank matches are listed. Additional and update of SAGE in normal human liver is available in our home page, http://www.prevent.m.u-tokyo.ac.jp.

SAGE. Total RNA was extracted from the homogenized sample using RNAzolB (TEL-Test, Inc., Friendswood, TX) according to the manufacturer's protocol. Polyadenylated RNA was extracted using a μMACS mRNA Isolation Kit (Miltenyi Biotec., GmbH Bergisch Gladbach, Germany). A total of 5.6 μg of polyadenylated RNA were extracted from 200 mg of the frozen sample, half of which was used for SAGE. The outline of the SAGE protocol has been described in a previous report (11). Briefly, polyadenylated RNA was converted to double-stranded complementary DNA (cDNA) by the SuperScript Choice System (Gibco BRL) with a 5′-biotinylated oligo dT_{18} primer. Double stranded cDNA was digested with *Nla*III and most 3′ primed fragments were bound to Dynabeads M-280 Streptavidin (Dynal A.S., Oslo, Norway). After ligation of the oligonucleotides containing recognition sites for *Bsm*FI, the linked cDNA was released from the beads by digestion with *Bsm*FI. Released tags were ligated to one another blunt-ended, concatemerized, and cloned into the *Sph*I site of the pZero-1 vector (Invitrogen Co., Carlsbad, CA). Samples were sequenced by an ABI PRISM 377 DNA Sequencer with a BigDye Terminator Cycle Sequencing Kit (PE Biosystems, Foster City, CA). Sequenced files were analyzed using SAGE 1.00 software. Tag sequences were examined by use of the advanced BLAST at NCBI 1. (www.ncbi.nlm.nih.gov/cgi-bin/BLAST).

RT-PCR. Total RNA was extracted using RNAzolB (TEL-Test, Inc.). Four hundred nanograms of total RNA was reverse transcribed in a total of 100 μL of reaction solution including 240 units of Moloney murine leukemia virus reverse transcriptase (Sawady Technology Co., Tokyo, Japan), 80 units of RNase inhibitor (Sawady Technology Co.), 50 mmol/L Tris-HCl, 125 mmol/L KCl, 8 mmol/L $MgCl_2$, 10 mmol/L dithiothreitol (Gibco BRL), 1 mmol/L deoxyribonucleotide triphosphates, and 2 mmol/L random hexamer (Promega Co., Madison, WI) for 1 hour at 42°C. PCR was performed in 20 μL of mixture containing 0.5 units of AmpliTaq DNA polymerase (PE Biosystems), 10 mmol/L Tris-HCl, 6.5 mmol/L $MgCl_2$, 1 mmol/L deoxyribonucleotide triphosphates, and 0.5 μmol/L sense and antisense primers by use of a Perkin–Elmer DNA thermal cycler. Primer sequences are listed in Table 1. The cycle profile of PCR was as follows. Denature at 94°C for 45 seconds, annealing at 57°C for 45 seconds, and extension at 72°C for 1 minute. Twenty-seven cycles of PCR amplification were performed to compare the gene expression levels of each sample. Thirty-six cycles of PCR amplification were performed to confirm the expression of 1-8U gene.

RESULTS

Unique tag species in SAGE tags of a normal liver. A total of 30,982 tags were cataloged from a normal liver using SAGE. These 30,982 tags corresponded to 8,596 unique different genes [Fig. 1A]. The tags expressed at levels of more than 10 counts accounted for 57.3% of the transcript mass, but for only 4.1% of the total number of the unique genes. Out of 8,596 unique transcripts, 2,986 (34.7%) encoded genes which were not reported thus far in GenBank. These unknown transcripts accounted for 13.9% of the transcript mass. The abundance of selected transcripts of SAGE tags in the normal liver was compared with those in HepG2 previously demonstrated by 3′ oriented ESTs (12), shown in Fig. 1B. In general, the abundance of SAGE tags was similar to that of ESTs in HepG2 in terms of the relative concentration of transcripts coding secreted proteins (shown in closed circles, $r^2 = 0.9563$; r, correlation coefficient). Transcripts coding the genes involved in cell growth, however, were more highly expressed in HepG2 as compared with those in the normal liver (shown in closed squares).

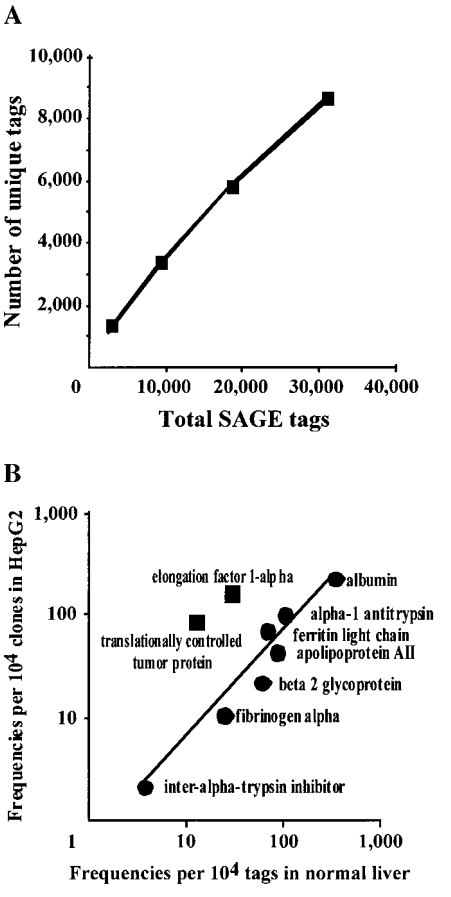

FIG. 1. (A) Number of unique SAGE tags. A total of 30,982 tags were obtained from a normal liver by SAGE. A proportional increase of unique tag numbers was observed relative to total tag numbers, revealing that the normal liver expressed more than 8,596 unique genes. (B) Relation between the abundance of selected transcripts in 3′ directed cDNA library from HepG2 and that of corresponding SAGE tags in a normal liver. The abundance of SAGE tags in a normal liver was compared with that of ESTs in HepG2 (12). Transcripts were chosen at random to represent a wide range of abundance. Frequencies in each transcript were calculated and plotted in logarithmic scale. As to the genes coding secreted proteins (shown in closed circle), a strong correlation was observed between the SAGE tag abundance in the normal liver and the EST abundance in HepG2 ($r^2 = 0.9563$, r, correlation coefficient). Genes coding elongation factor 1-α and translationally controlled tumor protein, which are involved in cell growth (shown in closed square), tended to be more highly expressed in HepG2.

Abundant transcripts in a normal liver. The expression of specific genes and production of their corresponding proteins have been extensively studied in the normal liver, but a systematic quantification of the expression levels of thousands of genes has not been published. The top 25 transcripts highly expressed in the normal liver are exhibited in Table 2. Almost all of these transcripts were genes encoding plasma proteins (>21.8% of total transcripts) and mitochondrial DNAs (>11.2%). The most abundant transcript encoding

TABLE 3

Cytoplasmic Protein

Tag sequence	Gene (GenBank Accession Number)	Count	Percent
CCCTGGGTTC	Ferritin light chain (M11147)	212	0.6775
GATCCCAACT	Metallothionein from cadmium-treated cells (V00594)	116	0.3707
TTGGTCCTCT	Ribosomal protein L41 (AB010874)	103	0.3292
TGTGTTGAGA	Elongation factor 1-alpha (X16869)	80	0.2556
TTGGGGTTTC	Ferritin heavy chain (L20941)	62	0.1981
ATAATTCTTT	Ribosomal protein S29 (U14973)	54	0.1725
TGCACGTTTT	Ribosomal protein L32 (X03342)	53	0.1693
GCCGAGGAAG	Ribosomal protein S12 (X53505)	51	0.163
CGCCGCCGGC	Ribosomal protein L35 (U12465)	50	0.1598
CACAAACGGT	Ribosomal protein S27 (U57847)	48	0.1534
GCCGTGTCCG	Ribosomal protein S6 (J03537)	43	0.1374
TAGGTTGTCT	Translationally controlled tumor protein (X16064)	43	0.1374
TCCTGCCCCA	Parathymosin (M24398)	38	0.1214
CTGGGTTAAT	Ribosomal protein S19 (M81757)	34	0.1086
CCGTCCAAGG	Ribosomal protein S16 (M60854)	31	0.099
CGCTGGTTCC	Ribosomal protein L11 (X79234)	29	0.0926
GGACCACTGA	Ribosomal protein L3 (X73460)	29	0.0926
CGCCGGAACA	Ribosomal protein L4 (D23660)	28	0.0894
GCAGCCATCC	Ribosomal protein L28 (U14969)	28	0.0894
TACAAACCTG	Metallothionein I-F (M13003)	28	0.0894
AGGCTACGGA	23 kDa highly basic protein (X56932)	26	0.0831
ACATCATCGA	Ribosomal protein L12 (L06505)	25	0.0799
AGCACCTCCA	Elongation factor 2 (Z11692)	25	0.0799
CCTAGCTGGA	Ribosomal protein L27a (U14968)	25	0.0799
GAGGGAGTTT	Ribosomal protein S26 (X69654)	24	0.0799

Note. Top 25 transcripts coding cytoplasmic proteins are listed. The tag sequence represents the 10-bp SAGE tag. Probable GenBank matches are listed.

plasma protein was the gene encoding albumin. Transcripts encoding apolipoproteins were also highly expressed, consistent with serum levels of the apolipoproteins. Other abundant transcripts encoded cytoplasmic proteins (>8.6%), enzymes (>4.8%), protease inhibitors (>1.7%), complements (>1.l%), and coagulation factors (>0.75%). Housekeeping genes such as those for cytoskeletal structure and protein synthesis constituted the majority of the genes encoding cytoplasmic proteins in the normal liver [Table 3]. Transcripts encoding ribosomal proteins and metal-binding proteins were abundant in cytoplasmic proteins. Among enzymes, it is noteworthy that not all of the most highly expressed genes were housekeeping genes involved in oxidative phosphorylation and energy synthesis [Table 4]. The most abundant transcript encoding enzyme was the gene encoding aldolase B, which is involved in hexose metabolism and expressed only in the liver. Also, transcripts encoding enzymes associated with amino acid metabolism, cholesterol synthesis, xenobiotics and drug metabolism were abundant in the normal liver. Other abundant transcripts encoding specific secreted proteins, including protease inhibitors, complements and coagulation factors, were listed [Table 5]. Transcripts encoding α-1 antitrypsin, antithrombin III, complement component 3, and fibrinogen, whose levels are known to be reduced in liver cirrhosis, were abundantly expressed in the normal liver. Together, these lists represent the characteristics of hepatic functions, including the synthesis of plasma proteins, the metabolism of various substrates, and the production of protease inhibitors, complements, and coagulation factors.

Validation of SAGE tag representation. We validated the fidelity of the SAGE data using RT-PCR in each individual [Fig. 2]. The 5 representative genes (albumin, alpha-1 antitrypsin, aldolase B, hemopexin, and cytochrome P450 IIIA4) identified by the expression of tags analyzed in this study were selected and confirmed to be expressed in each individual. The expression levels of each gene were correlated to the corresponding SAGE tag abundance. Genes encoding elongation factor 1-alpha and translationally controlled tumor protein were more highly expressed in HepG2 compared with those in the normal liver. In addition, genes encoding aldolase B, hemopexin, and cytochrome P450 IIIA4 were not detected in HepG2. These results were consistent with the estimation presented in Fig. 1B. The 1-8U gene, an interferon-inducible gene, whose expression has been newly identified in the normal liver by the expression of its SAGE tag, was confirmed by RT-PCR to be expressed in 3 normal liver samples and HepG2 as well.

TABLE 4

Enzyme

Tag sequence	Gene (GenBank Accession Number)	Count	Percent
CTCCAGAATA	Aldolase B (K01177)	145	0.4634
TAAGCCCCGC	4-Hydroxyphenylpyruvate dioxygenase (X72389)	74	0.2365
GAGGCCAAGA	Glutathione *S*-transferase subunit 1 (M25627)	70	0.2237
AGCCTCCCGG	L-Alanine:glyoxylate aminotransferase (X53414)	69	0.2205
AGTGTGTGGA	Cytochrome P450 IIE1 (J02843)	62	0.1981
GTGGGTTGGC	Aldehyde dehydrogenase 2 (K03001)	50	0.1598
TTTTCTCTGG	Cytochrome P450 IIIA4 (M13785)	50	0.1598
TACCCTGGAA	Alcohol dehydrogenase 4 class II (X56419)	48	0.1534
AACGTGCAGG	Argininosuccinate synthetase (X01630)	45	0.1438
CGCCACAAAG	UDP glycosyltransferase 2 family (J05428)	43	0.1374
CACCGCTGCA	11-β-hydroxysteroid dehydrogenase (M76665)	42	0.1342
ATCTGGAGCA	Alcohol dehydrogenase 1 class I (M12963)	39	0.1246
CAGCTTTCTA	Betaine:homocysteine methyltransferase (U50929)	34	0.1086
TGTTAGAACT	3-Hydroxy-3-methylglutaryl coenzyme A synthase (X83618)	31	0.099
CCCTTTCGAC	Cytochrome P450 IVA (S67580)	30	0.0958
TTGGAGATCT	NADH:ubiquinone oxidoreductase MLRQ subunit (U94586)	29	0.0926
GCTTAACCTG	Glutamate dehydrogenase (M20867)	28	0.0894
GCTTTGATGA	Epoxide hydrolase (L25880)	24	0.0767
ACCTTGTGCC	Sorbitol dehydrogenase (U67243)	22	0.0703
CAGGATTGTG	UDP-glucuronosyltransferase 2B4 (AF064200)	22	0.0703
CAAAGATTAT	Alcohol dehydrogenase 2 class I (M21692)	21	0.0671
GCCTGCTGGG	Glutathione peroxidase 4 (X71973)	21	0.0671
GCTTAATGTT	Catalase (X04096)	21	0.0671
GGGCGCTGTG	Ubiquinol-cytochrome *c* reductase (D55636)	21	0.0671
ACTATTTCCA	Fructose 1,6-bisphosphatase (M19922)	20	0.0639

Note. The top 25 transcripts coding enzymes are listed. The tag sequence represents the 10-bp SAGE tag. Probable GenBank matches are listed. Mitochondrial transcripts were excluded.

DISCUSSION

To our knowledge, this is the first study that describes the gene expression profile of a normal human liver. It was clarified that more than 21.8% of the transcripts encoded various plasma proteins in the normal liver. The number of transcripts encoding albumin was almost equal to that of all transcripts encoding ribosomal proteins. Also, the number of transcripts encoding all apolipoproteins was almost twice that of the transcripts encoding enzymes. These data reveal the massive production of plasma proteins in the normal liver. Compared with the expression profiles of human skeletal muscle (13), peripheral blood monocytes, macrophages, and dendritic cells (14, 15), the normal liver shows extensively elevated gene expression of plasma proteins. The enzymes highly expressed in the normal liver could be categorized into the following three groups based on their characteristics: (1) enzymes associated with hexose and lipid metabolism [aldolase B, 3-hydroxy-3-methylglutaryl coenzyme A synthase]; (2) enzymes catalyzing drugs and xenobiotics [cytochrome P450, glutathione *S*-transferase]; and (3) enzymes involved in the biosynthesis of urea [argininosuccinate synthetase, glutamate dehydrogenase]. Reduction of these enzyme activities is known to cause hyperammonemia, instability of blood glucose level, and elevation of serum bilirubin. A recent report revealed that enzyme activities catalyzing drugs were decreased according to the severity of liver disease (16). Another report has suggested the possibility that allelic variants of genes encoding enzymes for the catalyzation of xenobiotics might be responsible for the efficiency of carcinogenic reagent metabolism (17, 18). Although transcription levels are not necessarily in parallel with the protein amounts, it would seem beneficial to assess the genetic polymorphism of these genes by utilizing this gene expression profile in the diagnosis of various liver diseases.

To date, only one study has described the gene expression profile of hepatocytes [the large-scale sequencing of a 3′ directed cDNA library from HepG2] (12). In investigating the gene expression profile of hepatocytes, the advantages of using HepG2 as a resource lie in the homogeneity of cell population and the good condition of extracted RNA. Yet different gene expression profiles exist between HepG2 and the normal liver, as shown in our SAGE and RT-PCR results. Previous reports have revealed the importance of heterotypic cell to cell interactions in stabilizing liver-specific functions in isolated hepatocytes (19–23). With this in mind, we could conclude that our results faithfully describe the physiological gene expression profile of the normal human liver.

TABLE 5

Protease Inhibitor, Complement and Coagulation Factor

Tag sequence	Gene (GenBank Accession Number)	Count	Percent
GGAAAAGTGG	Protease inhibitor 1, α-1 antitrypsin (K01396)	326	1.0419
GGCAGAGTAG	Antithrombin III (X68793)	94	0.3004
GACTCTTCAG	α-1-antichymotrypsin (K01500)	32	0.1022
CCTTCCATTA	Inter-α-globulin-inhibitor, H1 polypeptide (X63652)	30	0.0958
AGCTCCTTAA	α 2 plasmin inhibitor (D00174)	24	0.0767
GAAGTGAAGG	Inter-α-trypsin inhibitor (X07173)	12	0.0383
TCTCCCATAC	Heparin cofactor II (M12849)	12	0.0383
CTCTCCAAAC	Plasma protease-C1-inhibitor (M13656)	9	0.0287
GACACCGGAT	Protein phosphatase-1 inhibitor (U48707)	8	0.0255
GTTGTCTTTG	Complement C3 (K02765)	125	0.3995
TTGGGATGGG	Complement factor H (M65292)	63	0.2013
TTCTGTGCTG	Complement C1r (M14058)	61	0.1949
AACACAGCCT	Complement C4A (K02403)	42	0.1342
ACTGAAAGAA	Complement C1s (X06596)	20	0.0639
CAAGCTGTTT	Complement C8 beta subunit (M16973)	13	0.0415
GCCTTTGCTT	Complement C5 (M57729)	12	0.0383
ACCAGGGAGA	Complement C8 α subunit (M16974)	7	0.0223
GAATTTCCCA	Complement C2 (X04481)	7	0.0223
TTAAATGGAA	Fibrinogen, α (T53910)	76	0.2429
ACGGAAAGGA	Fibrinogen, β (M64983)	63	0.2013
TGTTCCGCCT	Prothrombin, coagulation factor II (J00307)	35	0.1118
TTCCAAGGAA	Kininogen (K02566)	28	0.0894
ATAACTGTTG	Fibrinogen-like 1, FGL1 (D14446)	16	0.0511
CCCCGGAGGT	Coagulation factor X (K01886)	11	0.0351
GAACATTTTG	Coagulation factor IX (M11309)	4	0.0129

Note. The transcripts coding protease inhibitors, complements and coagulation factors are listed. The tag sequence represents the 10-bp SAGE tag. Probable GenBank matches are listed.

We found a tag encoding the 1-8U gene, which belongs to the interferon inducible gene family, in the normal human liver. 1-8U gene expression has not previously been reported in the liver, and its expression was also confirmed by RT-PCR in two other normal livers and HepG2. The 1-8U gene has been de-

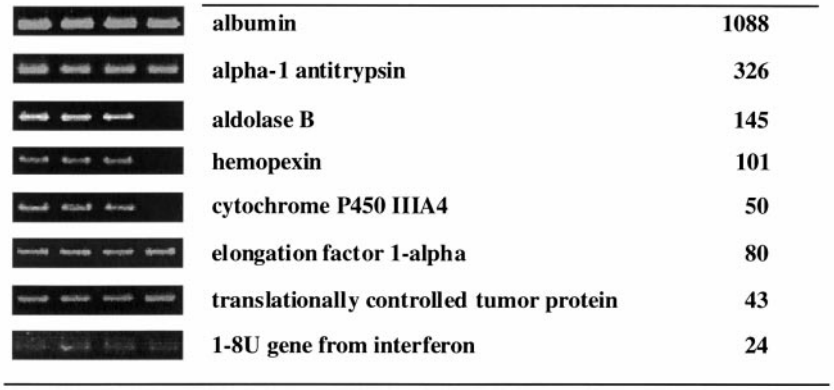

FIG. 2. RT-PCR analysis. According to the abundance of SAGE tag numbers, the 5 representative genes (albumin, α-1 antitrypsin, aldolase B, hemopexin, cytochrome P450 IIIA4) were chosen to determine the universality of gene expression in each individual by RT-PCR. The expression levels of each gene were apparently relative to the abundance of SAGE tags. Aldolase B, hemopexin, and cytochrome P450 IIIA4 gene expressions were not detected in HepG2. The gene expression levels of elongation factor 1-α and translationally controlled tumor protein were slightly increased in HepG2 compared with normal liver samples. The expression of 1-8U gene, an interferon-inducible gene, was newly identified in both normal liver samples and HepG2.

tected and cloned in T98G neuroblastoma cell line after interferon treatment (24), but its function remains to be identified (25). Out of 8,596 unique tags, we found 2,986 whose sequences did not correspond to any genes reported in GenBank thus far. These transcripts accounted for 13.9% of the total tag numbers, revealing that at least 13.9% of the transcripts expressed in the normal liver were unknown genes. We expect that the analysis of these novel genes will provide further understanding of hepatic function.

In conclusion, we performed SAGE on the normal human liver and identified its characteristic gene expression. This normal gene expression profile provides the candidate genes to be used in the investigation of various human liver diseases such as viral hepatitis, liver cirrhosis, and hepatocellular carcinoma.

ACKNOWLEDGMENTS

The authors thank Drs. Victor E. Velculescu, Lin Zhang, Bert Vogelstein, and Kenneth W. Kinzler for their help in SAGE analysis. They also thank Dr. Masashi Unoura for helpful discussions and Drs. Kouji Konishi and Masahiko Tsuji for their valuable help in providing human liver samples.

REFERENCES

1. Yamauchi, M., Potter, J. J., and Mezey, E. (1988) *Hepatology* **8,** 243–247.
2. Kuo, F. C., and Darnell, J. E., Jr. (1991) *Mol. Cell. Biol.* **11,** 6050–6058.
3. Jungermann, K. (1995) *Histochemistry* **103,** 81–91.
4. Jungermann, K., and Kietzmann, T. (1996) *Annu. Rev. Nutr.* **16,** 179–203.
5. Gupta, S., Rajvanshi, P., Sokhi, R. P., Vaidya, S., Irani, A. N., and Gorla, G. R. (1999) *J. Biol. Chem.* **274,** 2157–2165.
6. Massimi, M., Lear, S. R., Williams, D. L., Jones, A. L., and Erickson, S. K. (1999) *Hepatology* **29,** 1549–1555.
7. Lamers, W. H., Geerts, W. J., Jonker, A., Verbeek, F. J., Wagenaar, G. T., and Moorman, A. F. (1997) *Hepatology* **26,** 398–406.
8. Christoffels, V. M., Sassi, H., Ruijter, J. M., Moorman, A. F., Grange, T., and Lamers, W. H. (1999) *Hepatology* **29,** 1180–1192.
9. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) *Science* **270,** 484–487.
10. Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W., and Vogelstein, B. (1997) *Nature* **389,** 300–305.
11. Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Jr., Hieter, P., Vogelstein, B., and Kinzler, K. W. (1997) *Cell* **88,** 243–251.
12. Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara K. (1992) *Nat. Genet.* **2,** 173–179.
13. Welle, S., Bhatt, K., and Thornton, C. A. (1999) *Genome Res.* **9,** 506–513.
14. Hashimoto, S., Suzuki, T., Dong, H., Yamazaki, N., and Matsushima, K. (1999) *Blood* **94,** 837–844.
15. Hashimoto, S., Suzuki, T., Dong, H., Nagai, S., Yamazaki, N., and Matsushima K. (1999) *Blood* **94,** 845–852.
16. Furlan, V., Demirdjian, S., Bourdon, O., Magdalou, J., and Taburet, A. M. (1999) *J. Pharmacol. Exp. Ther.* **289,** 1169–1175.
17. Hu, X., Xia, H., Srivastava, S. K., Pal, A., Awasthi, Y. C., Zimniak, P., and Singh, S. V. (1998) *Cancer Res.* **58,** 5340–5343.
18. Hu, X., Herzog, C., Zimnik, P., and Singh, S. V. (1999) *Cancer Res.* **59,** 2358–2362.
19. Fraslin, J. M., Kneip, B., Vaulont, S., Glaise, D., Munnich, A., and Guguen-Guillouzo, C. (1985) *EMBO J.* **4,** 2487–2491.
20. Mesnil, M., Fraslin, J., Piccoli, C., Yamasaki, H., and Guguen-Guillouzo, C. (1987) *Exp. Cell Res.* **173,** 524–533.
21. Sibille, J. C., Kondo, H., and Aisen, P. (1988) *Hepatology* **8,** 296–301.
22. Goulet, F., Normand, C., and Morin, O. (1988) *Hepatology* **8,** 1010–1018.
23. Bhatia, S. N., Balis, U. J., Yarmush, M. L., and Toner, M. (1999) *FASEB J.* **13** 1883–1900.
24. Friedman, R. L., Manly, S. P., McMahon, M., Kerr, I. M., and Stark, G. R. (1984) *Cell* **38,** 745–755.
25. Lewin, A. R., Reid, L. E., McMahon, M., Stark, G. R., and Kerr, I. M. (1991) *Eur. J. Biochem.* **199,** 417–423.