

The rat blood–brain barrier transcriptome

Bradley E Enerson and Lester R Drewes

Department of Biochemistry and Molecular Biology and The Center for Cell and Molecular Biology, Medical School Duluth, University of Minnesota, Duluth, Minnesota, USA

The blood–brain barrier (BBB) is the cellular interface between the circulating blood and neural environment, and is created by apposed endothelial cells and their intercellular tight junctions. Many aspects of how the BBB functions at the molecular level remain unresolved; therefore, we report for the first time a comprehensive gene expression profile of rat brain microvessels using serial analysis of gene expression (SAGE). We assembled a full and quantitative SAGE catalog containing 101,364 tags, of which 33% of the tags matched known genes, 51% matched expressed sequence tags (ESTs) in the Unigene database, and 16% of the tags were unassigned. The transcriptome catalog contains many new and novel transcripts among known BBB genes. A large complement of junctional proteins and an extensive assortment of facilitated carrier and ATP-dependent transporters are included. To identify microvessel-enriched transcripts, we compared the microvessel SAGE catalog to cortex and hippocampus SAGE catalogs. This resulted in identification of 864 genes, including several known for their abundant expression at the BBB, such as the transferrin receptor (TrnR). Sorting enriched genes based on function revealed groups that encode transporters (11%), receptors (5%), proteins involved in vesicle trafficking (4%), structural proteins (10%), and components of signal transduction pathways (17%). This genomic repertoire emphasizes the unique cellular phenotype existing within the brain and further implicates the BBB as a mediator between the brain and periphery. These results may provide a useful resource and reference point from which to determine the effects of different physiological, developmental, and disease processes on BBB gene expression.

Journal of Cerebral Blood Flow & Metabolism (2006) **26**, 959–973. doi:10.1038/sj.jcbfm.9600249; published online 23 November 2005

Keywords: cerebral vascular biology; endothelial cell; expressed sequence tags; gene expression profiling; microvessels; transporters

Introduction

The cells within the central nervous system are located in an environment protected from the influence of circulating blood constituents and peripheral tissues by the unique microvasculature of the brain. The impeditive nature of the brain's microvasculature shields the brain from variations in plasma levels of metabolites, neuroactive amino acids and peptides, xenobiotics, as well as blood-borne pathogens and toxins that might be present in the circulation. This physiological partition between the blood and brain is created by endothelial

cells and their intercellular tight junctions that form a cellular interface between the circulating blood and neural environment, and is referred to as the blood–brain barrier (BBB).

While the physiological, anatomical, and biochemical properties of the BBB have been studied for years and form the basis for our current understanding of BBB function, many aspects of how the BBB functions remain unresolved at the molecular level. Reports of global gene expression analysis in brain endothelial cells are few and not comprehensive. As gene expression is a function of development, aging, disease, and genetic makeup, a systematic determination of the brain microvessel transcriptome would greatly advance our understanding of how the BBB is affected by these events and processes.

The identification of genes expressed by the microvasculature is particularly valuable for a number of reasons. Physiologically, it is known that the brain endothelial cells mediate transport of nutrients, vitamins, and metabolic precursors to the brain (Drewes, 2001). Endothelial cells are also

Correspondence: Dr LR Drewes, Department of Biochemistry and Molecular Biology and The Center for Cell and Molecular Biology, Medical School Duluth, University of Minnesota Duluth, 1035 University Drive, Duluth, Minnesota 55812, USA.

E-mail: ldrewes@d.umn.edu

This research was supported by a grant from the NINDS (NS37764).

Received 30 August 2005; revised 11 October 2005; accepted 17 October 2005; published online 23 November 2005

endowed with carrier systems for inward movement of ions associated with osmoregulation and cerebral spinal fluid formation (Jones *et al*, 1992; Strange, 1992). The microvasculature possesses carrier systems (Golden and Pardridge, 2000) and catabolic or conjugating enzymes (Gherzi-Egea *et al*, 1995; Minn *et al*, 1991) that prevent drugs, toxins, and xenobiotics from entering the brain or facilitate efflux of metabolic byproducts. The brain endothelium also possesses specific receptors and signaling molecules that might be essential for communication between peripheral peptides or hormones and cells of the central nervous system (CNS) (Banks, 2001; Pardridge, 1986; Pardridge *et al*, 1981). These receptors might also be important for viral and pathogen entry to the brain. Identification of new proteins associated with the aforementioned physiological functions will greatly advance our understanding of cerebral vascular biology.

The endothelial cells of the BBB also play an important role during the initiation and progression of inflammatory responses, a prominent feature in CNS disease pathogenesis. It is well known that specific interaction between adhesion molecules on cerebral endothelial cells and their ligands on circulating immune cells plays an important role in regulating adhesion and transendothelial migration. However, at present, the mechanisms by which immune cells cross the BBB and gain access to the CNS have not been fully defined (Brown, 2001; Weller *et al*, 1996).

In diseased brain, the BBB represents a formidable obstacle for delivery of therapeutic modalities to the brain parenchyma; therefore, knowledge of new BBB-specific targets may aid in rational drug design. Additionally, BBB-specific ligands might be conjugated to pharmacologically active agents for brain-specific drug delivery (Li and Qian, 2002; Pardridge, 2001).

Recent advances in gene expression analysis make it possible to monitor the expression of thousands of genes simultaneously using high-throughput assays such as microarray hybridization and EST sequencing (Colantuoni *et al*, 2000). The complementary deoxyribonucleic acids (cDNAs) or oligonucleotides on most microarrays are derived from clones of known transcripts or cDNAs derived from ESTs. However, because endothelial cells represent one one-thousandth of the total brain volume, a unique endothelial cell-specific gene expressed at a moderate or low level might not be fully represented in gene or EST databases. Owing to this issue, the suppression subtractive hybridization method was recently used to generate a cDNA library containing clones for unique genes expressed in both human and rat brain microvasculatures (Li *et al*, 2001, 2002; Shusta *et al*, 2002). In this approach, messenger ribonucleic acid (mRNA) derived from liver and kidney was subtracted from brain microvessel mRNA and clones were sequenced and identified as being enriched in brain endothelium compared

with liver and kidney tissue. This work represents the first step in brain endothelial cell gene discovery.

Another approach to rapid gene discovery and high-throughput transcriptome analysis is the SAGE method (Velculescu *et al*, 1995). The SAGE procedure involves deriving short 14 base tags from the 3' end of cellular transcripts and randomly ligating these tags to form concatamers of 20–50 tags. Cloning and sequencing a few thousand of these concatamers yields a SAGE tag catalog that is comprehensive because it does not require *a priori* knowledge of any gene sequence and is quantitative because the tag frequency is proportional to transcript abundance. Identification of the tags is achieved by a comparison to computationally derived tags from known genes and ESTs in NCBI databases (SAGEmap) (Lash *et al*, 2000). Serial analysis of gene expression has been shown to be reliably quantitative, with relative tag counts correlating well to Northern blot analysis (Blackshaw *et al*, 2001). It is reproducible and relatively comprehensive (Trendelenburg *et al*, 2002). Serial analysis of gene expression libraries can be easily compared with other catalogs to define tissue-specific genes, or genes regulated under different physiological conditions, disease, or development. The usefulness of this method has been shown in studies of tumors (Argani *et al*, 2001; Porter *et al*, 2001; Untergasser *et al*, 2002; Zhang *et al*, 1997), hypoxia/ischemia (Trendelenburg *et al*, 2002), and seizures (Hendriksen *et al*, 2001). Serial analysis of gene expression has also been used to study gene expression changes during development (Gunnerson *et al*, 2002) and to identify tissue-specific genes or markers in specialized tissues such as endothelium (St Croix *et al*, 2000) and epithelium (Sharon *et al*, 2002).

In the work described here, we have used SAGE to identify a catalog of mRNAs expressed in rat brain microvessels. Furthermore, we determined genes enriched in microvessels by comparing our catalog to cortex and hippocampus SAGE catalogs. This work will serve as a useful reference to compare different physiological, developmental, and disease processes, and may aid target discovery for drug delivery to the brain.

Materials and methods

Isolation of Rat Brain Microvessels

All procedures using animals followed a protocol approved by the University of Minnesota Animal Care Committee. In all, 50 adult male Long-Evans rats (30 days old) were maintained on a defined, control diet consisting of 77% carbohydrates, 13% fat, and 10% protein (per calorie) (#TC00606, Harlan Teklad, Madison, WI, USA). Rat brain microvessels were isolated as described previously (Gerhart *et al*, 1988). Rat brain tissue was collected after rapid exsanguinations of rats under halothane anesthesia. The pia mater and associated vessels were

removed with forceps, and the cortical gray matter was excised and placed in cold M199 medium (Sigma, St Louis, MO, USA). The tissue was minced by stirring with a scalpel and homogenized in a Potter–Elvehjem tissue grinder. (Approximately 1 g of minced brain was removed at this point and frozen at -70°C for ribonucleic acid (RNA) extraction.) The homogenate was passed once through a $350\text{-}\mu\text{m}$ mesh polypropylene net, twice through $110\text{-}\mu\text{m}$ mesh nylon nets, and centrifuged (1000g, 10 mins). The pellets were resuspended in 20% dextran in M199 and centrifuged (2500g, 15 mins) in a swinging bucket rotor. The pellets were then combined and resuspended in cold medium, and the microvessels were collected on a $20\text{-}\mu\text{m}$ mesh nylon net. The microvessels were pooled in PBS and frozen at -70°C before RNA extraction.

Serial Analysis of Gene Expression

Total RNA was prepared from the pooled microvessels and from the minced brain sample using RNawiz (Ambion, Austin, TX, USA). The quantity and quality of the RNA was assessed by UV absorption spectrophotometry and by use of a RNA chip assay (RNA 6000 nanochip with Bioanalyzer, Agilent, Palo Alto, CA, USA). Serial analysis of gene expression libraries were generated from the RNA using the I-SAGE kit (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol; this is a slight modification of the protocol available from Velculescu and colleagues (www.sagenet.org). Restriction enzymes *Nla*III and *Bsm*FI were used for tag generation. Optimization of the polymerase chain reaction (PCR) amplification of the ditags resulted in a template dilution of 1/80, and PCR was performed for 27 cycles with 400 $50\text{-}\mu\text{L}$ reactions in parallel. Cloned concatamers were sequenced using the BigDye Terminator sequencing kit (PE/Applied Biosystems Foster City, CA, USA) in combination with an ABI377 DNA sequencer. Raw sequence trace files were analyzed using the Pfred algorithm (Ewing and Green, 1998; Ewing *et al*, 1998) and sequences with a low score were excluded. Tags were extracted from sequences using SAGE2000 software (www.sagenet.org). The catalogs generated by tag-to-gene mapping and contained in this report are publicly accessible in a searchable format at <http://fammed2/genomics/Transcriptome.htm>. A publicly available hippocampus SAGE catalog derived from young adult male Wistar rat brain (Datson *et al*, 2001) and containing approximately 30,000 different identified transcripts was used for comparison with our microvessel and cortex catalogs.

Real-Time Reverse Transcriptase (RT)-PCR

Primer pairs for real-time PCR were designed for selected cDNAs using Oligo 6.4 software (Molecular Biology Insights, Cascade, CO, USA) and are indicated in Supplementary Table 1. Complementary deoxyribonucleic acid was transcribed from total RNA using an Omniscript™ kit following the manufacturer's protocol (Qiagen, Valencia, CA, USA). Polymerase chain reaction reactions

were performed on a LightCycler™ instrument using the LightCycler™ DNA Master SYBR® Green I kit following the manufacturer's protocol (Roche, Indianapolis, IN, USA). Before quantitative measurements of cDNA were performed, PCR reaction conditions for each primer set were optimized and melting curve analysis was used to confirm that each primer pair produced a single product. Standard curves for candidate cDNAs were prepared from a series of five 10-fold serial dilutions of target cDNA. The quantity of DNA in each sample was normalized to 18S ribosomal RNA (rRNA). The normalized average ($n=3$) value from the microvessel samples was compared with the normalized average ($n=3$) cortex value.

Virtual Rot Analysis

The virtual Rot analysis was constructed as described previously (Blackshaw *et al*, 2001; Velculescu *et al*, 1997). EXCEL (Microsoft, Redmond, WA, USA) was used to plot different tag abundance levels against the cumulative fraction of total tags represented by that particular abundance level.

Results

Microvessel Isolation

Serial analysis of gene expression was performed on mRNA derived from microvessels to obtain a quantitative gene expression profile of the rat BBB. Serial analysis of gene expression was also performed on mRNA from rat cerebral cortex for use, in combination with other rat SAGE libraries, to identify genes enriched in the BBB.

The rat brain microvessel preparation is pure, as shown by lack of contaminating cells or tissues (Figure 1A). The microvessel preparation contains a mixture of cells, including endothelial cells and pericytes, that are together surrounded by a basement membrane. Astrocytic end feet attached to the basement membrane are also present in the preparation. Glucose transporter 1 (GLUT1), which is highly expressed in brain endothelium, was used as a marker to show the enrichment of the preparation (Pardridge *et al*, 1990). The low expression of S100beta, an astrocytic marker, in the microvessel fraction compared with the whole cortex fraction further establishes the overall purity of the microvessel preparation (Figure 1B) (Mrak *et al*, 1996). The levels of S100beta mRNA present in the sample indicate that transcripts derived from astrocytic end feet are present in our preparation, but represent a minor contribution to the SAGE library.

Serial Analysis of Gene Expression Library Statistics

Sequencing of SAGE tags from the microvessel and cortex SAGE libraries resulted in the assembly of catalogs containing 101,364 and 57,620 tags, respec-

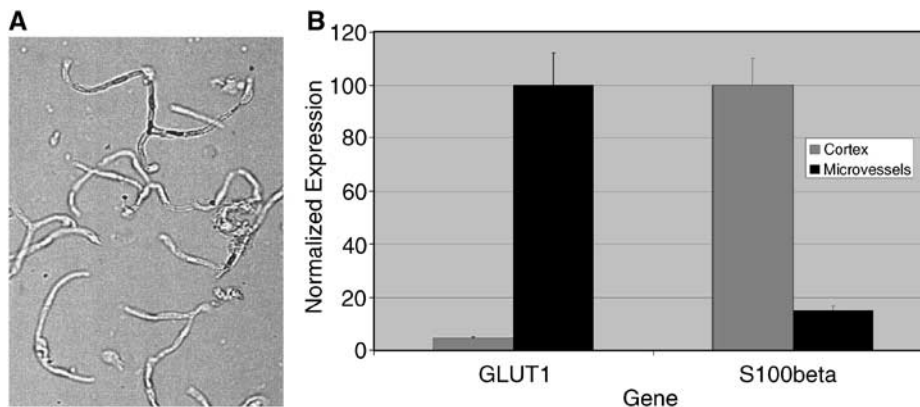


Figure 1 Isolated rat brain microvessels. (A) The microvessel preparation consists of segments of the brain microvasculature and contains endothelial cells and pericytes that together are surrounded by a basement membrane with remnants of astrocytic end feet attached. (B) Assessment of microvessel purity by real-time quantitative RT-PCR. Glucose transporter 1, which is highly expressed in brain endothelium, was used as a marker to show the enrichment of the preparation. S100beta, an astrocytic marker, is relatively abundant in the cortex sample. Glucose transporter 1, which is highly expressed in brain endothelium, was used as a marker to show the enrichment of the preparation (Pardridge *et al*, 1990). S100beta, an astrocytic marker, shows the overall purity of the preparation (B) (Mrak *et al*, 1996).

Table 1 Summary of SAGE library content

	Microvessel	Cortex
Total number of tags	101,364	57,620
Tags used in analysis ^a	77,648	45,294
Different transcripts	10,996	5407
Matched genes	33%	35%
ESTs	51%	51%
No match	16%	14%
<i>Tag frequency</i>		
> 100	51	41
51–100	81	76
21–50	423	216
11–20	863	405
6–10	1595	782
3–5	3605	1755
2	4379	2132
1	23,716	12,326

^aNumber of tags in catalog after removal of singletons and linker tags.

tively, per library (Table 1). Approximately 24% of the microvessel tags and 22% were observed only once. To minimize the possibility that these singleton tags were due to errors, the sequence data was scanned for low-quality sequence using the phred algorithm (Ewing and Green, 1998; Ewing *et al*, 1998). Regions of sequence with a low phred score (<20) were trimmed and removed before tag extraction, adding confidence that the singleton tags represent mRNAs that are expressed at very low levels. In this report, we have limited further analysis to tags observed at a frequency of 2 or greater.

After singleton tags and linker tags were removed, the individual tags for unique transcripts were combined. The total number of different transcripts was 10,996 for the microvessel library and 5407 for the cortex library. The unique tags were compared with the NCBI gene-to-tag mapping data set (Rat

Unigene Build 141, NCBI) to assign tags to specific genes (Lash *et al*, 2000). This analysis revealed that 33% of microvessel tags and 35% of cortex tags matched known genes, while 51% of both microvessel tags and cortex tags matched transcribed sequences (ESTs) in the Unigene database. The remaining 16% and 14% of tags in the microvessel and cortex libraries, respectively, could not be assigned to known genes or ESTs.

ROT Analysis

To determine whether the tags in the rat microvessel SAGE catalog were representative of an expected mRNA population, we performed a virtual Rot analysis (Supplementary Figure 1). A Rot analysis measures mRNA re-association kinetics to examine the relative fraction of mRNA composed of transcripts expressed at high, medium, and low levels. A virtual Rot performs the same type of analysis *in silico*. Our analysis of the relative fraction of total mRNA containing transcripts of different expression levels revealed several abundances classes that represented transcripts expressed at high, medium, or low levels spanning three orders of magnitude. When the virtual Rot of microvessel RNA was compared with an actual Rot analysis of mouse brain RNA, a strong similarity is revealed (Supplementary Figure 1).

Microvessel Catalog

Analysis of the 50 most abundant tags in the microvessel catalog reveals that seven tags are derived from mitochondrial genes and several from transcripts of ribosomal proteins (Table 2). Other notable tags identified are SPARC (osteonectin), basigin (Bsg, CD147), and apolipoprotein E (ApoE),

Table 2 Catalog of the 50 most abundant tags in rat brain microvessels

N	SAGE tag	Count ^a	Gene symbol	Description	Genbank nucleotide
1	ATACTGACAC	1321	mt-Co3	Cytochrome <i>c</i> oxidase III, mitochondrial	NC_001665
2	GTGGCTCACA	751	Aco2	Aconitase 2	NM_024398
			Tmem19	Unknown (protein for MGC:72591)	NM_199098
3	ATAAGACATA	748	mt-Atp8	ATP synthase 8, mitochondrial	NC_001665
4	CACGCCTCTC	728	Hba-a1	Hemoglobin alpha, adult chain 1	NM_013096
			GloA	Globin, alpha	NM_001013853
5	TGGATCCTGA	589	Hbb	Hemoglobin β chain complex	BC058448
6	AATGGGAGGC	527	mt-Cytb	Cytochrome <i>b</i> , mitochondrial	NC_001665
7	AGCCGTCCTC	471	mt-Co2	Cytochrome <i>c</i> oxidase II, mitochondrial	NC_001665
8	TTGGTGAAGG	403	Tmsb4x	Thymosin, β 4	NM_031136
			Slit-1	Slit homolog 1 (<i>Drosophila</i>)	NM_022953
9	GGCTGCTGGG	357		Transcribed locus	BQ781374
10	CTAGTCTTTG	346	Rps29	Ribosomal protein S29 (Rps29)	NM_012876
11	GCTTCATCCA	345	Mbp	Myelin basic protein	NM_017026
12	CGAACTCTCA	295	Sparc	Secreted acidic cysteine-rich glycoprotein	NM_012656
13	CTAAGTAAAG	291		Transcribed locus	AW524827
14	ACCAGCCAGG	269	Apoe	Apolipoprotein E	NM_138828
15	ACGTCTCAAA	265	Tuba1	Tubulin, alpha 1	NM_022298
16	CCTGCTGTGT	245		Transcribed locus	AA925767
				Coat protein gamma-cop (predicted)	XM_216203
17	GTGACCACGG	242		Unknown	
18	GTGCTTTCAA	236		Transcribed locus	XM_576100
19	GGGATGGCCA	228		Nicotinamide nucleotide adenylyltransferase 2 (predicted)	XM_573475
20	TGGTTGCTGG	211	Hrasls3	HRAS-like suppressor	NM_017060
21	CTCTGACTTT	185	Bsg	Basigin	NM_012783
22	CAGAGCTTTA	184		Predicted: Insulin-like growth factor binding protein 7	BC086582
				Transcribed locus	AA819083
23	GATGCCCCCC	180	mt-Co1	Cytochrome <i>c</i> oxidase I, mitochondrial	NC_001665
24	CCAATGGTCC	177	Adcy6	Adenylate cyclase 6	NM_022217
			Cst3	Cystatin C	NM_012837
25	GGGTTGGGGA	171	Amph1	Amphiphysin 1	NM_022217
26	CTAAGGAAGT	161	Sparcl1	SPARC-like 1 (mast9, hevin)	NM_012946
27	GGAAGCCTCC	159		Similar to osteoclast inhibitory lectin	XM_342769
28	TGGGTTGTCT	158	Tpt1	Tumor protein, translationally controlled 1	NM_053867
				Transcribed locus	AI101070
29	AGGGGGAGGG	152	Verge	Vascular early response gene protein	NM_001003403
30	GAAGCAGGAC	147	Cfl1	Cofilin 1	NM_017147
			Bmp6	Bone morphogenetic protein 6	NM_013107
31	GGGAGGGGGG	147	Egfl7	EGF-like domain 7	NM_139104
32	AGCGATTCAA	145	mt-Nd3	NADH dehydrogenase 3, mitochondrial	NC_001665
33	GACTGACCCT	142		Transcribed locus	CA945877
34	TCAGGCTGCC	140	Fth1	Ferritin, heavy polypeptide 1	NM_012848
				Transcribed locus	AA800422
35	AATCAACCCG	137	Impa1	Inositol (myo)-1(or 4)-monophosphatase 1	NM_032057
36	GTGATGTGGC	133	Ptgds	Prostaglandin D2 synthase	NM_013015
37	AGTTTGCTGA	130		Transgelin 2 (predicted)	BC084703
38	CGGAAGGCGG	129	Rpl36	Ribosomal protein L36	NM_022504
39	GGATATGTGG	122	Egr1	Early growth response 1	NM_012551
40	TAGGTACAGG	122	mt-Atp6	ATP synthase 6, mitochondrial	NC_001665
41	GCCTTCCAAT	120	ddx5 gene	ddx5 gene	NM_001007613
42	TCCAATAAAG	116	Rplp1	Ribosomal protein, large, P1	NM_001007604
43	GTCGCTGAGA	112	CAR-XI	Carbonic anhydrase-related XI protein	NM_175708
44	TGCAGTAAA	111	Psap	Prosaposin	NM_013013
				Transcribed locus	XM_574375
45	CAAACATCCA	109	MT-RNR1	12 S ribosomal RNA, mitochondrial	NC_001665
46	GCTACTTCTG	108	B2m	β -2 microglobulin	NM_012512
47	ATGACTATTA	108	mt-Nd4	NADH dehydrogenase 4, mitochondrial	NC_001665
48	TGACAAACTG	107	Cox8a	Cytochrome <i>c</i> oxidase, subunit VIIIa	L48209
49	GGATTTGGCC	106	Rplp2	Ribosomal protein, large P2	XM_215116
				Similar to 60S acidic ribosomal protein P2	XM_573831
50	AGTTGGAAAC	105	Sv2b	Synaptic vesicle glycoprotein 2 b	NM_057207
				Transcribed locus	BC090353

^aTags sorted according to abundance and expressed in tags per 100,000.

which are known to be abundantly expressed at the BBB. Myelin basic protein (Mbp), although recognized as an oligodendroglial product, is also

relatively abundant in the brain microvessel catalog, a finding that corroborates a previous report that brain endothelial cells express this unique

transcript (Li *et al*, 2002). Transcripts for structural proteins, such as cofilin 1 and alpha tubulin, are also among the top 50 list.

The catalog was examined for the presence of genes associated with the BBB phenotype, including tight junction and adhesion proteins (Table 3), members of the Slc transporter families (Slc members; Table 4), and ATP-dependent transporters (ATP-binding cassette, ABC transporters; Table 5). Tags for transcripts encoding proteins of three major types of cell–cell junctional complexes are represented in the catalog, including tight junctions (claudins, occludin, junctional adhesion molecules, zona occludens proteins), adherins junctions (cadherins, protocadherins), and gap junctions (connexins). Several nutrient transporters expected in brain microvessels are abundantly expressed in the catalog and the variety of carrier proteins is remarkably extensive (Table 4). ATP-dependent transporters, including several that are implicated in drug transport into and from cells, are also expressed (Table 5).

Identification of Microvessel-Enriched Tags

To identify tags for genes that are enriched in microvessels, we compared the microvessel catalog to the cortex catalog and the rat hippocampus catalog generated by Datson *et al* (2001). Tags with an abundance ratio of at least 10 between the microvessel preparation and either the cortex catalog or the hippocampus catalog were considered enriched. Tags not found in a catalog were assigned an arbitrary value of 1 to avoid division by 0 in the fold difference calculation. In addition, for a gene to be considered enriched, a subjective threshold of expression was set after it was normalized to the total tag count in its respective catalog. A threshold of 50 tags per million was set as the cutoff. As the catalogs were of unequal sizes, to be considered for the enriched pool a microvessel tag had to be represented six times, a hippocampal tag five times, and a cortex tag four times in their respective catalogs. The results of this approach show that, of the 17,440 tags present in all catalogs, 864 are enriched in brain microvessels (Figure 2).

The 864 tags enriched in the microvessel SAGE catalog were annotated by comparisons to the UniGene reliable tag-to-gene-mapping data set. Only 31% of the microvessel-enriched tags correspond to known genes, while 61% matched transcribed sequences (ESTs) in the Unigene database (Figure 3A). Tags that did not match any sequence in the database represented 8% of the enriched pool. The assignment of a functional category for the proteins encoded by the mRNAs identified in our analysis was primarily based on the gene ontology database (<http://www.geneontology.org>).

Of the known genes in the enriched catalog, a large percentage of the tags identify mRNAs for

transporters, receptors, and proteins involved in signal transduction, vesicle trafficking, and the extracellular matrix (Figure 3B). Analysis of the 50 most abundant tags in the microvessel-enriched catalog (Table 6) reveals a number of expressed genes known from previous studies to be enriched in the brain endothelium. These include the transferrin receptor (TfR), GLUT1 (Slc2a1), claudin 5, cationic amino acid transporter-1 (CAT-1), organic anion transporter-1 (Oatp1c1), vonWillibrandt factor (vWF), and the VEGF receptor-1 (Flt-1).

Real-Time Polymerase Chain Reaction Analysis

Real-time PCR analysis was used to confirm microvessel enrichment of select genes (Figure 4). Analysis of six genes verified the enrichment and presence of β -hemoglobin, SPARC (osteonectin), GLUT1, monocarboxylate transporter 1 (MCT1), and CD36 in the microvessel preparation when compared with cortex. Cytochrome *c* oxidase II was determined to be more highly expressed in cortex than microvessels.

Discussion

The tight junctions between the endothelial cells that line the capillaries in the brain prevent paracellular movement of solutes between blood and brain. Therefore, it is not surprising that the endothelial cells are essential for the maintenance of metabolic homeostasis in the brain, and for serving as a conduit for metabolites and for communication between blood and brain. As endothelial cells account for only 1/1000 of total brain volume and are associated with the vasculature, the brain endothelium is often overlooked when global brain function is considered. In this report, we used SAGE analysis and conservatively identified a large number of genes specifically enriched in the cerebral capillaries. This catalog suggests a more prominent role for the brain microvasculature in neurological function than previously recognized. Knowledge of the unique genomic architecture of the BBB coupled with its distinctive physiologic properties further illustrates that the brain microvasculature is a unique and vital component of the CNS.

The brain microvessel SAGE catalog generated in this report is the first attempt at large-scale analysis of the BBB transcriptome. Serial analysis of gene expression was chosen over less-expensive and time-consuming hybridization-based gene expression technologies because it identifies unknown genes and produces information on absolute gene expression, corresponding well to Northern blot and real-time PCR expression data. In addition, SAGE catalogs from different tissues or different laboratories are easily comparable and shared. Here, we identify 34,712 unique tags in our catalog, of which

Table 3 Tight junction and adhesion tags present in microvessel SAGE catalog

N	SAGE tag(s) ^a	Count ^b	Gene symbol	Description, alias	Genbank ID
1	TGCTCGGGAG	60	Gja1	Gap junction membrane channel protein alpha 1, connexin 43	NM_012567
2	TGCACTGTTG	52	Cldn5	Claudin 5	NM_031701
3	GTGGAACCTCT TACACTCCCC CCCCCTCCCC	37	Pcdhgc3	Protocadherin gamma subfamily C, 3	NM_001012215
4	GCGCCTTCAA	19	Icam2	Intercellular adhesion molecule 2	NM_001007725
5	GGTGCCGTGC GACAAGTAAA	18	Esam	Endothelial cell adhesion molecule	NM_001004245
6	GGTGTAAACA	18	Cldn11	Claudin 11	NM_053457
7	TTAAAAAATA	15	Gja5	Gap junction membrane channel protein alpha 5, connexin 40	NM_019280
8	GAATAAACAG TGATGTTTGA CTGGTCTTTG	15	smagp	Small cell adhesion glycoprotein	NM_182817
9	TACAAGAAGT AGACCTTTT ACTGGCTCAG TGAATGATAC	13	Alcam	Activated leukocyte cell adhesion molecule	NM_031753
10	CACCAGGCTA	12	Cdh22	Cadherin 22	NM_019161
11	CATTCAAGAGT	11	Tjp2	Tight junction protein 2, zona occludens 2 (ZO-1)	U75916
12	GGTTTGGGGA	11	Jup	Junctional plakoglobin	NM_031047
13	AAAAAAGTTT ATCTCCAAGA	9	F11r	Junctional adhesion molecule 1	NM_053796
14	AGATGTAAGT	9	Cdh5_predicted	Cadherin 5 (predicted)	XM_226213
15	GAAGAAGAAT GAGAACGCCA	9	Pkp4_predicted	Plakophilin 4 (predicted)	XM_215733
16	GGCCTCCCTT	7	Gjb2	Gap junction membrane channel protein β 2, connexin 26	NM_001004099
17	GCTGTGGGTC	6	Icam5_predicted	Intercellular adhesion molecule 5, telencephalin (predicted)	XM_233737
18	TGTCAAGTTG CAGAATAATG	6	Adrm1	Adhesion regulating molecule 1	NM_031708
19	AAGAAGTCAC CTGCACAGAG	5	Nrcam	Neuron-glia-CAM-related cell adhesion molecule	NM_013150
20	TACGAAAATG CTACAAAAAG	5	Cdh11	Cadherin 11	XM_341639
21	TGGATTCATT	5	Pcdha13	Constant exon in protocadherin alpha gene cluster	NM_199503
22	ACTTCAGGGG	4	Tjp4_predicted	Tight junction protein 4 (peripheral) (predicted)	XM_236932
23	TGAGCTCTGG	4	Mcam	Melanoma cell adhesion molecule	NM_023983
24	TACATAATGC	3	Tjp1_predicted	Tight junction protein 1 (predicted), zona occludens 1 (ZO-1)	XM_218747
25	CTTCCCTGT	3	Catna1	Catenin (cadherin-associated protein), alpha 1, 102 kDa	NM_001007145
26	ACCATCTGTA	3	Ceacam9	CEA-related cell adhesion molecule 9	NM_053919
27	GTGCTGCAGA	3	Boc_predicted	Biregional cell adhesion molecule-related/downregulated by oncogenes (Cdon) binding protein (predicted)	XM_340986
28	TCACCTCCTT	3	Pkp2_predicted	Plakophilin 2 (predicted)	XM_213560
29	TTCAGCCTGG	2	Dscam	Down's syndrome cell adhesion molecule	NM_133587
30	GTGGGTCTG	2	NCAML1	Neural cell adhesion molecule L1	NM_017345

^aIn this catalog more than one tag was present for some high-abundance genes.

^bThe frequencies represent the total number of tags for each gene and are expressed in tags per 100,000.

10,996 are present at least twice. It should be noted that the depth of this catalog is such that low-abundance tags are not reliably quantitative, a common problem for all gene expression technologies. Among these, low-abundance transcripts are likely to be those that encode signal transduction molecules and transcription factors. It has been estimated that, to identify all transcripts present at a single copy per mammalian cell, a catalog of 650,000 tags would need to be sequenced

(Velculescu *et al*, 1999). In our study, the number of unique tags as a function of the total number of tags sequenced was plotted and subjected to non-linear regression analysis (Supplementary Figure 2). Extrapolation of the resulting curve reached a plateau of approximately 78,000 unique tags for expressed genes predicting the theoretical limit of unique transcripts.

While slight impurities exist in our microvessel preparation (Figure 1B), most tags derived from

Table 4 Tags identifying transporters in the microvessel SAGE catalog

<i>N</i>	<i>SAGE tag(s)^a</i>	<i>Count^b</i>	<i>Gene symbol</i>	<i>Description</i>	<i>Genbank ID</i>	<i>Function/substrate</i>
1	AGAAGGACCT TGTCTGGACA TGCTAAATCC	74	Slc2a1	Solute carrier family 2, member 1, GLUT1	NM_138827	Glucose
2	CCAACAAGAA	56	Slc16a6	Solute carrier family 16 (monocarboxylic acid transporters), member 6, MCT	XM_573214	
3	CTTCTGCAGA	49	Slc7a1	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 1, CAT-1	NM_013111	Lysine, arginine
4	AAGGTCATTT	48	Slco1c1	Solute carrier organic anion transporter family, member 1c1, Oatp1	NM_053441	Thyroid hormone, organic anions
5	AGGCTTTATG	39	Slc30a1	Solute carrier family 30 (zinc transporter), member 1, Znt1	NM_022853	Zinc efflux
6	GTGAGGTAGG CCAGGACAGT	38	Slc4a3	Solute carrier family 4, member 3, AE3	NM_017049	Anion (Cl ⁻ /HCO ₃ ⁻) exchanger
7	GGGGTTGGGG	29	Slc34a3	Solute carrier family 34 (sodium phosphate), member 3	NM_139338	Na ⁺ /Pi cotransporter
8	GATAAAACCA	28	Slc25a4	Solute carrier family 25 (mitochondrial adenine nucleotide translocator) member 4, ANT1	NM_053515	ADP/ATP exchange across mitochondrial inner membrane
9	GGGGGTGCTT CGTTAAAATA	25	Slc21a5	Solute carrier organic anion transporter family, member 1a4	NM_131906	
10	CTGAGCCTTG	24	Slc27a1	Solute carrier family 27 (fatty acid transporter), member 1, FATP1	NM_053580	Long chain fatty acids
11	ACTCTTAAGA	21	Slc9a3r2	Solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulator 2, NHERF-2	NM_053811	Sodium/hydrogen exchanger
12	TFTTCGCTTT TGGGAGGCAG AACCTTGTG ACAGTGAAGG	19	Slc3a2	Solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	NM_019283	Activators of dibasic and neutral amino acid transporters
13	GTCTCCTCTC TAGAAAAATG	17	Slc16a1	Solute carrier family 16 (monocarboxylic acid transporters), member 1, MCT1	NM_012716	H ⁺ coupled monocarboxylate, pyruvate/lactate
14	TGAATATGTC	15	Scl6a20	Solute carrier family 6, member 20, Xtrp3	NM_133296	Na ⁺ and Cl ⁻ coupled neurotransmitter
15	GGGAAAGCCT	14	Slc9a3r1	ERM-binding phosphoprotein, NHERF	NM_021594	Na ⁺ /H ⁺ exchange regulatory factor
16	CCGCCAGGT CGTTGCTAGG	14	Slc40a1	Solute carrier family 39 (iron-regulated transporter), member 1, IREG1	NM_133315	Proton-coupled divalent metal ion transporters
17	TGTGCTTATT TTATTGGTGT TGATTTCAAT	13	Slc38a2	Solute carrier family 38, member 2, SAT2 (system A amino acid transporter 2)	NM_181090	Na(+) dependent neutral amino acid
18	TTCATCTGTC	13	Slc20a1	Solute carrier family 20 (phosphate transporter), member 1, PiT1	NM_031148	Na(+) dependent phosphate symporter
19	AGAGCACTCA	13	Slc25a20	Solute carrier family 25 (carnitine/ acylcarnitine translocase), member 20	NM_053965	carnitine/ acylcarnitine translocase
20	ACCTCAGTGT ACTGTGATGG GAAACGCTGC	12	Slc20a2	Slc20a2 Solute carrier family 20, member 2, PiT2	NM_017223	Na(+) dependent phosphate transporter
21	GTGCTGACCC	11	Slc7a5	Slc7a5 Tumor-associated protein 1, LAT1	NM_017353	Large neutral amino acids
22	TTACTGTAGT	10	Slc2a3	Slc2a3 Solute carrier family 2 (facilitated glucose transporter), member 3, GLUT3	NM_017102	Glucose
23	TFTTGTAAAT CATTTTGTTC	9	Slc5a6	Slc5a6 Solute carrier family 5 (sodium-dependent vitamin transporter), member 6, SMVT Pentothenate, biotin, lipote	NM_130746	

Table 4 Continued

<i>N</i>	<i>SAGE tag(s)^a</i>	<i>Count^b</i>	<i>Gene symbol</i>	<i>Description</i>	<i>Genbank ID</i>	<i>Function/substrate</i>
24	GAGAGGAGAC	9	Slc9a1	Slc9a1 Solute carrier family 9, member 1, NHE1	NM_012652	Na ⁺ /H ⁺ exchanger
25	TACAAAGCCA GAGGAAACCA	8	Slc38a3	Slc38a3 Solute carrier family 38, member 3, amino acid transport N system, SN1 Glutamine, asparagine histidine, Na ⁺ /H ⁺ dependent	NM_145776	
26	CCCAGCTTTT	8	Slc5a5	Slc5a5 Solute carrier family 5 (sodium iodide symporter), member 5, NIS	NM_052983	Na ⁺ iodide
27	GAGCCCTGCT	8	Asna1	ArsA arsenite transporter, ATP-binding, homolog 1 (bacterial) (predicted)	XM_213848	arsenite efflux
28	TTCTGTGTGG	8	Slc1a6	Slc1a6 Solute carrier family 1 (high affinity aspartate/glutamate transporter), member 6, EAAT4	NM_032065	Aspartate/glutamate
29	AGTCGTTTGA AAGATGTGTT	8	Slc6a1	Slc6a1 GABA transporter protein, GAT1	NM_024371	Na ⁺ (and Cl ⁻) dependent transporter of gamma-aminobutyric acid
30	CTGCTGATGC CAGTGGGTGG	7	Slc35a1	LOC287642 Galactose transporter, (putative Slc35a1)	NM_199081	UDP-galactose
31	CTGGAGCTGG	7	Slc25a10	Solute carrier family 25 (mitochondrial carrier; dicarboxylate transporter), member 10	NM_133418	Mitochondrial dicarboxylic acids
32	CCCCTTATTT CGCCCTCCCT	7	Slc15a3	Peptide/histidine transporter PHT2	NM_139341	Peptide/histidine
33	GGCTGAATGC	6	Slc12a2	Solute carrier family 12, member 2 (hypothetical)	NM_031798	Na ⁺ /K ⁺ chloride
34	ATGCTTTTGA GACATAGCCC	6	Slc6a8	Choline transporter	NM_017348	Na ⁺ (and Cl ⁻) creatine/choline
35	TGACAGGCAC TCCATCCAGG	5	Slc12a5	Solute carrier family 12, (potassium-chloride transporter) member 5	NM_134363	K ⁺ -Cl ⁻ cotransporter
36	TTTCACCCCA	5	Slc5a2	Solute carrier family 5 (sodium/glucose cotransporter), member 2, SGLT2	NM_022590	Na ⁺ /glucose
37	AACTTTGAAA AGTGGGGAGA	5	Slc13a5	Nact Sodium-coupled citrate transporter	NM_170668	Na ⁺ /citrate
38	GCGCCTCTGT	5	Slc38a1	Solute carrier family 38, member 1, SAT1	NM_138832	Na ⁺ dependent glutamine
39	GGAGAGGAGG CCCAAATCAA	4	Slc6a13	Solute carrier family 6 (neurotransmitter transporter, GABA), member 13, GAT3	NM_133623	Na ⁺ dependent GABA
40	GTCGTCCTCT	4	Slc17a7	Solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 7	NM_053859	Na ⁺ dependent phosphate cotransporter

^aIn this catalog more than one tag was present for some high abundance genes.

^bThe frequencies represent the total number of tags for each gene and are expressed in tags per 100,000.

parenchymal brain tissue (nonvascular cells) are also present in the cortex and hippocampus catalogs. Therefore, when the microvessel catalog is compared with the brain tissue catalogs, the derived list represents genes that are unique or highly expressed in only microvessel-associated cells, including primarily endothelial cells (Table 6). Although this *in silico* method towards defining the microvessel-enriched set of tags was accomplished using a heuristic approach, the validity of this approach is shown by the presence of a number of genes previously identified as enriched in brain endothelium. Most well-defined BBB

markers are present in our set of enriched tags. These genes include: glucose transporter (GLUT1), P-glycoprotein (MDR1a, data not shown), CAT-1, TfR, MCT1, vWF, and Oatp1c1 (Slco1c1). A number of these known markers were confirmed by quantitative RT-PCR. One gene identified as enriched, cytochrome *c* oxidase II, mitochondrial, was expressed at higher levels in cortex than microvessels. This most likely results from one or more genes sharing a common SAGE tag and shows the need to verify specific SAGE data by other methods. A β -hemoglobin tag in the catalog is most likely derived from transcripts extracted from reticulocytes that

Table 5 ATP-binding cassette SAGE tags in the microvessel catalog

<i>N</i>	<i>SAGE tag(s)^a</i>	<i>Count^b</i>	<i>Gene symbol</i>	<i>Description, alias</i>	<i>Genbank ID</i>	<i>Function/Substrate</i>
1	TGGGAGTAGG	49	Abca2	ATP-binding cassette, subfamily A (ABC1), member 2	NM_024396	Lipid transport/metabolism
2	ATGGCACCAG	45	Abcd3	ATP-binding cassette, subfamily D (ALD), member 3, PMP70	NM_012804	Peroxisomal import of fatty acid/fatty acyl coA
3	TCTCCTTAGC GGGCACAAAA AGAGCTTATT	38	Abcc9	ATP-binding cassette, subfamily C (CFTR/MRP), member 9, SUR2	NM_013040	Receptor, K transport channel
4	GCTCCAAATA GGGACTTGAG	24	Abcf1	ATP-binding cassette, subfamily F (GCN20), member 1	AF293383	mRNA translation
5	TGAACTGTGT GTGGCATCCC	19	Abcg2	ATP-binding cassette, subfamily G (WHITE), member 2, BCRP1	NM_181381	Drugs, xenobiotics efflux
6	ATCATTATCT TATGAAAATG	16	Abcb1	ATP-binding cassette, subfamily B (MDR/TAP), member 4, P-gp	NM_133401	Drugs, xenobiotics efflux
7	CGAGATGTTA CTGGGATTCA	16	Abcb6	ATP-binding cassette, subfamily B (MDR/TAP), member 6	NM_080582	Transporter of iron/sulfur from mitochondria to cytosol
8	ACTTTGGGGG TTTGCTAAAC	14	Abcc5	ATP-binding cassette, subfamily C (CFTR/MRP), member 5, MRP5	NM_053924	Drug, putative glutathione (GSH) conjugate (GS-x) pump
9	CAGGTGCAAT	10	Abcb8	ATP-binding cassette, subfamily B (MDR/TAP), member 8 (predicted)	NM_001007796	Transport of phospholipids into mitochondrial membranes
10	AACGAGAAAA	9	Abca7	ATP-binding cassette, subfamily A, member 7	NM_207598	Transmembrane lipid transport
11	CCACATTCGG	6	Abcg1	ATP-binding cassette, subfamily G (WHITE), member 1	NM_053502	Transfer of cellular cholesterol to HDL
12	GCCTTGCCTT GAAACCAAAA	2	Abcc1	ATP-binding cassette, subfamily C (CFTR/MRP), member 1, MRP1	NM_022281	Multispecific anion transporter
13	AGCAGCAGCA	2	Abcc3	ATP-binding cassette, subfamily C (CFTR/MRP), member 3, MRP3	NM_080581	Organic anions, drugs bile salt efflux
14	GCCTGTGTGT	2	Abcc4	ATP-binding cassette, subfamily C (CFTR/MRP), member 4, MRP4	NM_133411	Organic anion transporter, putative efflux pump for cAMP/cGMP
15	ACTCCCTGAG	2	Abcc6	ATP-binding cassette, subfamily C (CFTR/MRP), member 6, MRP6	NM_031013	Organic anions amphipatics efflux
16	GCTACGGAAA	2	Abcc8	ATP-binding cassette, subfamily C (CFTR/MRP), member 8, MRP8	NM_013039	Regulator of K ⁺ channels and insulin release
17	TAGGTTTATG	2	Abce1	ATP-binding cassette, subfamily E (OABP), member 1 (predicted)	XM_341669	RNase inhibitor, RNA turnover

^aIn this catalog more than one tag was present for some high abundance genes.

^bThe frequencies represent the total number of tags for each gene and are expressed in tags per 100,000.

might have been entrapped in the microvessel preparation.

This SAGE catalog intersects well with previous BBB gene discovery efforts. For example, in a recent report, Li *et al* (2002) identified 69 clones for genes selectively enriched in the rat brain microvasculature by the suppressive subtractive hybridization method, using driver cDNA from liver and kidney. Of the 36 known genes identified, 75% match genes were identified by tags in the total microvessel catalog (Table 2) and 42% are found in the microvessel-enriched catalog (Table 6). The lack of full concordance between the data sets is to be expected due to inherent differences in the methodologies. Our microvessel-enriched catalog also contains many of the pan endothelial markers defined by SAGE analysis of human colorectal

endothelium (St Croix *et al*, 2000). Recently, a SAGE analysis was conducted on endothelial cells isolated from human brain microvessels (Madden *et al*, 2004). Several cell-type-specific transcripts were identified in these other endothelial cell catalogs and their abundance correlated closely with the rat microvessel tag abundance reported here (Table 7). One evident difference is that the glucose transporter, GLUT1, was much more abundant in brain microvessels and endothelial cells than in colon endothelial cells. This might be related to different functions of endothelial cells in the respective tissues.

We examined the full microvessel SAGE tag catalog for genes associated with the BBB phenotype, including transporters, adhesion molecules, and ABC cassette containing transporters. Blood-brain barrier transport processes have been well documented at the biochemical and molecular levels (Drewes, 2001). The microvessel SAGE tag catalog generated in this report includes known BBB transporters (GLUT1, MCT1, CAT-1, Oatp1c1, and others) and cloned transporters that have not previously been identified at the BBB (Gerhart *et al*, 1997; Pardridge *et al*, 1990). For example, our analysis shows for the first time that the sodium-dependent vitamin transporter (SMVT, NM_130746), a specific transporter for pantothenate, biotin, and lipoate, is present in the microvessel catalog as well as the microvessel-enriched catalog (Prasad *et al*, 1998). In addition, our data suggest that the sodium/iodide symporter, Slcs5a5 (NIS), is present and enriched at the BBB. Interestingly, although iodide is actively transported from blood to brain, there are no reports of Slcs5a5 expression at BBB (Davson and Hollingsworth, 1973). Finally, while functional evidence exists for the abluminal expression of a sodium/glucose transporter at the BBB, the responsible transporter has not been identified (Lee *et al*, 1997). A SAGE tag for the mRNA encoding the low-affinity Na/glucose transporter, SGLT2, is also present and enriched in the microvessel SAGE catalog; its existence is supported by previous reports indicating low levels of transcript in rat brain (You *et al*, 1995).

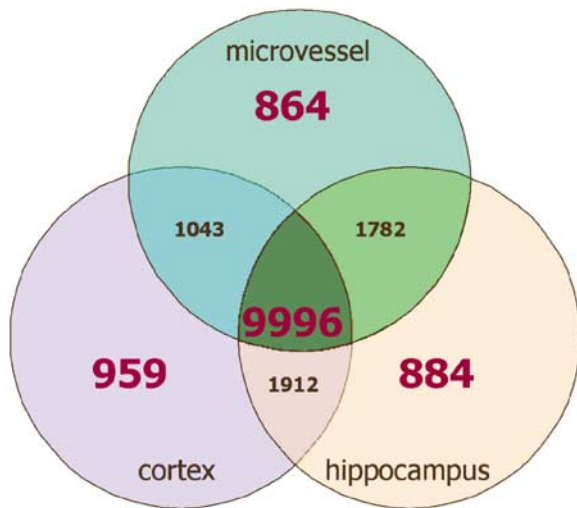


Figure 2 Venn diagram comparing the rat microvessel, cortex, and hippocampus SAGE libraries. Tags with an abundance ratio of at least 10 between the microvessel preparation and either the cortex library or the hippocampus library were considered enriched. Tags not found in a library were assigned an arbitrary value of 1 to avoid division by 0 in the fold difference calculation. In addition, for a gene to be considered enriched, a threshold of 50 tags per million was set as the cutoff.

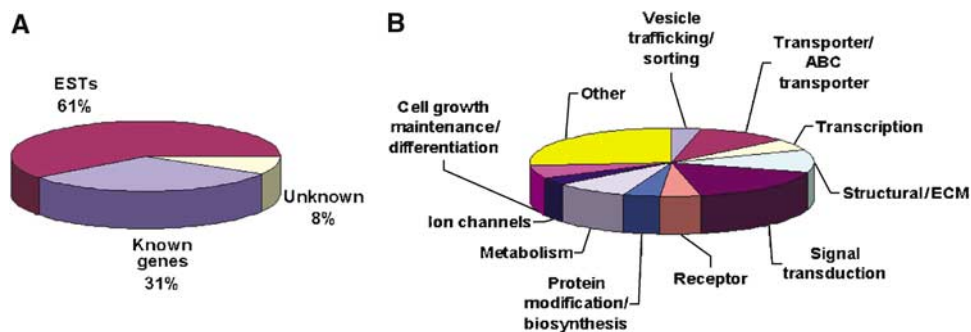


Figure 3 Distribution of microvessel-enriched genes based on functional classification. (A) Distribution of the 864 microvessel-enriched genes based on matches to known genes or ESTs in the Unigene database. (B) The assignment of a functional category to the proteins encoded by the mRNAs of known genes identified in (A).

Table 6 Catalog of the 50 most abundant tags enriched in rat brain microvessels^a

<i>N</i>	<i>MV SAGE Tag</i>	<i>MV TPM</i>	<i>CX TPM</i>	<i>HP TPM</i>	<i>Symbol</i>	<i>Description</i>	<i>Genbank ID</i>
1	TGGATCCTGA	5811	347	52	Hbb	Hemoglobin β chain complex	BC058448
2	AGCCGTCCCT	4647	0	0	mt-Co2	Cytochrome <i>c</i> oxidase II, mitochondrial	NC_001665
3	CTAAGTAAAG	2871	35	65		Transcribed sequence	AW524827
4	GTGCTTTCAA	2328	223	39		Transcribed sequence	CA505636
5	GGAAGCCTCC	1569	0	0		Similar to osteoclast inhibitory lectin	XM_342769
6	AGGGGGAGGG	1500	87	0	Verge	Vascular early response gene protein	NM_001003403
7	AGTTTGCTGA	1283	0	65		Transgelin 2 (predicted)	BC084703
8	CCCTTATAGA	848	35	26	Tfrc	Transferrin receptor	XM_340999
9	CAGGCTTCGT	661	0	0		Phospholipid transfer protein (predicted)	XM_215939
10	GCGGGGGGAG	651	0	0	Tsc2	Tuberous sclerosis 2	NM_012680
			0	0	Pkd1	Polycystic kidney disease 1 homolog	AF277452
11	ATTATTGTAT	631	0	52		Transcribed sequence	CB705117
12	GAGGGAGAGG	612	0	0		Tensin like C1 domain containing phosphatase (predicted)	XM_235710
13	ATGGGTGAGA	582	52	26		Transcribed sequence	AI111918
14	GCCTCCAAGA	562	52	0	Pkd1	Plasminogen activator, tissue	NM_013151
15	AGAAGGACCT	552	0	0	Slc2a1	Solute carrier family 2, member 1	NM_138827
16	GTGGAACCT	513	0	0	Cldn5	Claudin 5	NM_031701
17	CTTCTGCAGA	483	0	0	Slc7a1	Solute carrier family 7, member 1	NM_013111
18	AAGGTCATTT	474	0	39	Slco1c1	Solute carrier organic anion transporter family, member 1c1	NM_053441
19	ACTCAGTGGG	454	0	0		palmdelphin (predicted)	XM_227623
20	GCAAAGAACC	424	35	39	Csrp2	Cysteine- and glycine-rich protein 2	NM_177425
21	ACTGAAGCAA	404	0	26	Scarb1	Scavenger receptor class B, member 1	NM_031541
22	TAGCTGAAAA	395	0	26	Nr4a1	Immediate early gene transcription factor NGFI-B	NM_024388
23	CGTTTGTCCG	395	0	0	Lamb2	Laminin, β 2	NM_012974
24	AGCGACTGTG	375	0	0		Transcribed sequence	AI030602
						Transcribed sequence	AI716117
25	CTGTGGCCTG	365	0	0	Vwf	Von Willebrand factor	XM_342759
26	CCAGAGTCAG	365	0	0		Transcribed sequence	NM_001009620
27	AGGCTTTATG	355	0	26	Slc30a1	Slc30a1 solute carrier family 30, member 1	NM_022853
28	AGAGCTTATT	355	0	0	Abcc9	ATP-binding cassette, subfamily C (CFTR/MRP), member 9	NM_013040
29	GGGAAGGGGG	345	0	0		COP9 (constitutive photomorphogenic) homolog, subunit 7a (predicted)	XM_232351
						Transcribed sequence	AI136016
30	TTTGACCCCC	345	0	0	Acvrl1	Activin A receptor type II-like 1	NM_022441
31	ATGGGGAAAGT	345	0	0		Transcribed sequence	AI170144
32	GCTTCTGGGG	345	0	0		Transcribed sequence	AI230388
33	ATGAATGTAT	335	0	0	Flt1	FMS-like tyrosine kinase 1	NM_019306
34	CTGGAGCATC	335	0	0		Transcribed sequence	CO402891
35	GACAGGAACG	335	0	0		Transcribed sequence	AI100789
36	GGTCAGTCGG	335	0	0	Ap1gbp1	AP1 gamma subunit binding protein 1	XM_573165
37	GCGCCTGTGA	326	0	0		Usher syndrome 1C binding protein 1 (predicted)	XM_214324
38	ATCATAGCCT	316	17	26	Ihpk1	Inositol hexaphosphate kinase 1	XM_576469
39	CTTGGTGCCG	316	0	0		Transcribed sequence	AA799725
40	CCAACACTTT	306	0	0	Fn1	Fibronectin 1	NM_019143
41	GCAAAATGGA	306	0	0	Fgf3	Fibroblast growth factor 3	NM_130817
42	TGCCACCCG	306	0	0	Baiap2	Brain-specific angiogenesis inhibitor 1-associated protein 2	NM_057196
43	TGCTACCCTT	306	0	0	Acta2	Smooth muscle α -actin	M22757
44	CTGAGGATGT	296	0	0	Etl	ETL protein	NM_022294
45	TACAGAAGGA	296	0	0		Transcribed sequence	CR477624
46	TGTGGTGGAG	296	0	0		Ubiquitin-specific protease 24 (predicted)	XM_233260
47	TGGTGAATAC	276	0	0	Cxcl12	Chemokine (C-X-C motif) ligand 12	NM_022177
48	TGAACACCGA	266	0	0	Tgm2	Tissue-type transglutaminase	NM_019386
						Transcribed sequence	AI710446
49	GACAGATGGA	266	0	0	Pthr1	Parathyroid hormone receptor 1	NM_020073
50	GTCCAAGGA	266	35	0	Dci	Dodecenoyl-coenzyme A delta isomerase	NM_017306

MV = microvessel, CX = cortex, HP = hippocampus, TPM = tags per million.

^aValues for tag counts are expressed in tags per million. Tags sorted according to abundance; when possible BLAST analysis was used to verify identity.

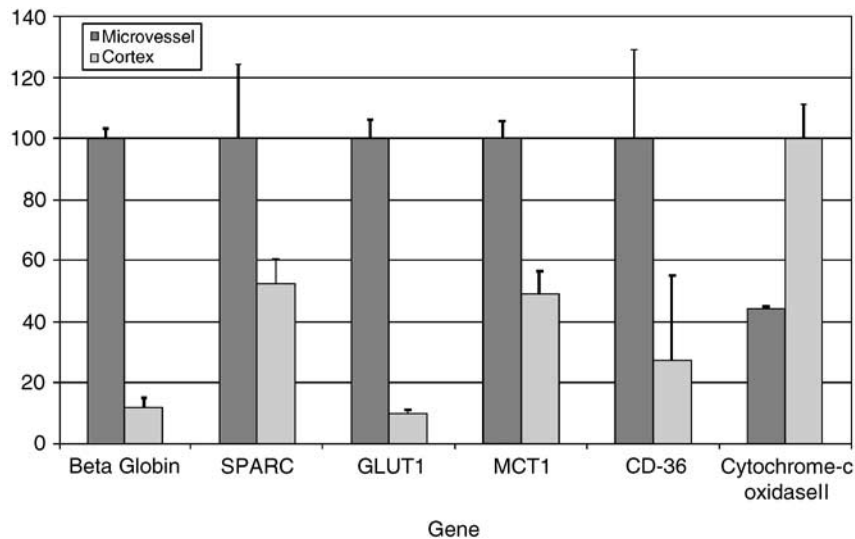


Figure 4 Confirmation of selected genes identified by SAGE to be enriched in rat brain microvessels. Real-time quantitative PCR was performed on a LightCycler™ instrument to quantify selected genes in microvessel mRNA and cortex mRNA preparations. Standard curves for candidate cDNAs were prepared from a series of four to six 10-fold serial dilutions of target cDNA. The quantity of DNA in each sample was normalized to 18S ribosomal RNA. The average normalized value from the samples showing larger amounts of target was set to 100. Error bars indicate standard deviation, $n = 3$ per group.

Table 7 Comparison of cell-type-specific gene expression levels between rat brain microvessels and endothelial cells from human brain and colon

Gene	Specificity	Rat brain MV	Human brain EC ^a	Human colon EC ^a
Hevin	EC	81	51, 99	161
vWF	EC	18	12, 53	35
Tie2/TEK	EC	12	2, 4	4
CD34	EC	29	3, 10	5
CD14	Hematopoietic	0	1, 2	1
CK8	Epithelial	3	0, 0	1
GLUT1	Brain EC	28	8, 37	0
GFAP	Glial	4	0, 0	0

MV = microvessels, EC = endothelial cells, tag frequencies per 50,000 total tags.

^aData from two SAGE catalogs for human brain EC (Madden *et al*, 2004) and one for human colon EC (St Croix *et al*, 2000).

Junctional complexes present between endothelial cells are the main structures responsible for the barrier function of brain capillaries. The barrier consists of intrinsic, that is, integral membrane proteins, and frequently is associated with peripheral proteins that make up the junctional complex. The constituents of tight junctions are mostly present in the full microvessel catalog (Table 2) and include occludin, claudins, and junctional adhesion molecules. Also present are various cadherin proteins that exist in adherens-type junctions along with peripheral molecules such as ZO-1 and ZO-2 that might be found in both types of junctions. As occludin, ZO-1, and ZO-2, junctional proteins that are associated with the characteristic tight junctions of the BBB, are reportedly present in astrocytes (Bauer *et al*, 1999; Duffy *et al*, 2000; Howarth and Stevenson, 1995), they are represented

in the full microvessel catalog, but not in the enriched catalog (Table 6).

Tags encoding transcripts for ATP cassette containing proteins are present in the microvessel SAGE catalog (Table 5). Among these ABC transporters are members such as *p*-glycoprotein that are characterized to deter the entry of drugs and xenobiotics into the brain. Others such as ABCG2 (BCRP1) and ABCc5 (MRP5) are much less well characterized.

The physiological functions of the BBB are also revealed by examining the functional categorization of BBB enriched genes (Figure 3B). In fact, nearly 40% of the 864 BBB enriched genes of known function encode transporters, receptors, and signal transduction proteins. The large group of mRNAs for proteins involved in signaling pathways show how the BBB may function in mediating signals from the

periphery to the CNS, which is possibly a significant and common occurrence (Ek *et al.*, 2001). For example, the enrichment of type I iodothyronine deiodinase in the microvessel catalog suggests a role for the brain endothelium in potentiating the availability of thyroid hormones to the brain. In addition, components of pathways that regulate endothelial cell growth and differentiation are also present in the pool of microvessel-enriched tags such as activin receptor-like kinase 1 (Table 6). Activin receptor-like kinase 1 is an endothelial-specific receptor of the TGF- β receptor family that has been implicated in angiogenesis. Mutations in this gene cause hereditary hemorrhagic telangiectasia (Lamouille *et al.*, 2002; Lux *et al.*, 1999). Also, present is flt-1, another angiogenesis-related signaling system.

It is not surprising that structural proteins and proteins that are involved in cell adhesion are strongly represented in the catalog of microvessel-enriched genes. Some of the proteins encoded by these genes, such as laminin and fibronectin, form the basement membrane that surrounds the endothelial cells and pericytes. Cytoskeletal genes are also represented in the data set, including vimentin, tubulin, and smooth muscle α -actin.

The molecular portrait of the BBB revealed by this initial quantitative and comprehensive analysis of its transcriptome uncovers numerous features of this cellular interface between blood and brain. First, the genomic repertoire emphasizes the unique cellular phenotype existing within the brain and sheds light on potential novel functions linking the brain and periphery. The BBB transcriptome catalog also identifies a number of tags for transcripts encoding known proteins that might be targets for affecting brain drug delivery. These transporters are known for their efflux properties and may contribute significantly to the regulation of the brain extracellular fluid composition and the effectiveness by which drugs penetrate the brain. Finally, these findings provide a useful resource and reference point for future studies in which the effects of different physiological, developmental, and disease processes on BBB gene expression are examined.

References

- Argani P, Iacobuzio-Donahue C, Ryu B, Rosty C, Goggins M, Wilentz RE, Murugesan SR, Leach SD, Jaffee E, Yeo CJ, Cameron JL, Kern SE, Hruban RH (2001) Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: identification of a new pancreatic cancer marker by serial analysis of gene expression (SAGE). *Clin Cancer Res* 7:3862–8
- Banks WA (2001) Anorectic effects of circulating cytokines: role of the vascular blood–brain barrier. *Nutrition* 17:434–7
- Bauer H, Stelzhammer W, Fuchs R, Weiger TM, Danninger C, Probst G, Krizbai IA (1999) Astrocytes and neurons express the tight junction-specific protein occludin *in vitro*. *Exp Cell Res* 250:434–8
- Blackshaw S, Fraioli RE, Furukawa T, Cepko CL (2001) Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell* 107:579–89
- Brown KA (2001) Factors modifying the migration of lymphocytes across the blood–brain barrier. *Int Immunopharmacol* 1:2043–62
- Colantuoni C, Purcell AE, Bouton CM, Pevsner J (2000) High throughput analysis of gene expression in the human brain. *J Neurosci Res* 59:1–10
- Datson NA, van der Perk J, de Kloet ER, Vreugdenhil E (2001) Expression profile of 30,000 genes in rat hippocampus using SAGE. *Hippocampus* 11:430–44
- Davson H, Hollingsworth JR (1973) Active transport of 131-I across the blood–brain barrier. *J Physiol* 233:327–347
- Drewes LR (2001) Molecular architecture of the brain microvasculature: perspective on blood–brain transport. *J Mol Neurosci* 16:93–8; discussion 151–7
- Duffy HS, John GR, Lee SC, Brosnan CF, Spray DC (2000) Reciprocal regulation of the junctional proteins claudin-1 and connexin43 by interleukin-1beta in primary human fetal astrocytes. *J Neurosci* 20:RC114
- Ek M, Engblom D, Saha S, Blomqvist A, Jakobsson PJ, Ericsson-Dahlstrand A (2001) Inflammatory response: pathway across the blood–brain barrier. *Nature* 410:430–1
- Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8:186–94
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175–85
- Gerhart DZ, Broderius MA, Drewes LR (1988) Cultured human and canine endothelial cells from brain microvessels. *Brain Res Bull* 21:785–93
- Gerhart DZ, Enerson BE, Zhdankina OY, Leino RL, Drewes LR (1997) Expression of monocarboxylate transporter MCT1 by brain endothelium and glia in adult and suckling rats. *Am J Physiol* 273:E207–13
- Gherzi-Egea JF, Leininger-Muller B, Cecchelli R, Fenstermacher JD (1995) Blood–brain interfaces: relevance to cerebral drug metabolism. *Toxicol Lett* 82-83:645–53
- Golden PL, Pardridge WM (2000) Brain microvascular P-glycoprotein and a revised model of multidrug resistance in brain. *Cell Mol Neurobiol* 20:165–81
- Gunnarsen JM, Augustine C, Spirkoska V, Kim M, Brown M, Tan SS (2002) Global analysis of gene expression patterns in developing mouse neocortex using serial analysis of gene expression. *Mol Cell Neurosci* 19:560–573
- Hendriksen H, Datson NA, Ghijsen WE, van Vliet EA, da Silva FH, Gorter JA, Vreugdenhil E (2001) Altered hippocampal gene expression prior to the onset of spontaneous seizures in the rat post-status epilepticus model. *Eur J Neurosci* 14:1475–84
- Howarth AG, Stevenson BR (1995) Molecular environment of ZO-1 in epithelial and non-epithelial cells. *Cell Motil Cytoskelet* 31:323–32
- Jones HC, Keep RF, Butt AM (1992) The development of ion regulation at the blood–brain barrier. *Prog Brain Res* 91:123–31
- Lamouille S, Mallet C, Feige JJ, Bailly S (2002) Activin receptor-like kinase 1 is implicated in the maturation phase of angiogenesis. *Blood* 100:4495–501

- Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg RL, Riggins GJ, Altschul SF (2000) SAGEmap: a public gene expression resource. *Genome Res* 10: 1051–1060
- Lee WJ, Peterson DR, Sukowski EJ, Hawkins RA (1997) Glucose transport by isolated plasma membranes of the bovine blood–brain barrier. *Am J Physiol* 272:C1552–7
- Li H, Qian ZM (2002) Transferrin/transferrin receptor-mediated drug delivery. *Med Res Rev* 22:225–50
- Li JY, Boado RJ, Pardridge WM (2001) Blood–brain barrier genomics. *J Cereb Blood Flow Metab* 21:61–8
- Li JY, Boado RJ, Pardridge WM (2002) Rat blood–brain barrier genomics. II. *J Cereb Blood Flow Metab* 22: 1319–1326
- Lux A, Attisano L, Marchuk DA (1999) Assignment of transforming growth factor beta1 and beta3 and a third new ligand to the type I receptor ALK-1. *J Biol Chem* 274:9984–92
- Madden SL, Cook BP, Nacht M, Weber WD, Callahan MR, Jiang Y, Dufault MR, Zhang X, Zhang W, Walter-Yohrling J, Rouleau C, Akmaev VR, Wang CJ, Cao X, St Martin TB, Roberts BL, Teicher BA, Klinger KW, Stan RV, Lucey B, Carson-Walter EB, Laterra JJ, Walter KA (2004) Vascular gene expression in nonneoplastic and malignant brain. *Am J Pathol* 165:601–8
- Minn A, Ghersi-Egea JF, Perrin R, Leininger B, Siest G (1991) Drug metabolizing enzymes in the brain and cerebral microvessels. *Brain Res Brain Res Rev* 16: 65–82
- Mrak RE, Sheng JG, Griffin WS (1996) Correlation of astrocytic S100 beta expression with dystrophic neurites in amyloid plaques of Alzheimer's disease. *J Neuropathol Exp Neurol* 55:273–9
- Pardridge WM (1986) Receptor-mediated peptide transport through the blood–brain barrier. *Endocr Rev* 7:314–30
- Pardridge WM (2001) Brain drug targeting and gene technologies. *Jpn J Pharmacol* 87:97–103
- Pardridge WM, Boado RJ, Farrell CR (1990) Brain-type glucose transporter (GLUT-1) is selectively localized to the blood–brain barrier. Studies with quantitative Western blotting and *in situ* hybridization. *J Biol Chem* 265:18035–40
- Pardridge WM, Frank HJ, Cornford EM, Braun LD, Crane PD, Oldendorf WH (1981) Neuropeptides and the blood–brain barrier. *Adv Biochem Psychopharmacol* 28:321–8
- Porter DA, Krop IE, Nasser S, Sgroi D, Kaelin CM, Marks JR, Riggins G, Polyak K (2001) A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res* 61:5697–702
- Prasad PD, Wang H, Kekuda R, Fujita T, Fei YJ, Devoe LD, Leibach FH, Ganapathy V (1998) Cloning and functional expression of a cDNA encoding a mammalian sodium-dependent vitamin transporter mediating the uptake of pantothenate, biotin, and lipoate. *J Biol Chem* 273:7501–6
- Sharon D, Blackshaw S, Cepko CL, Dryja TP (2002) Profile of the genes expressed in the human peripheral retina, macula, and retinal pigment epithelium determined through serial analysis of gene expression (SAGE). *Proc Natl Acad Sci USA* 99:315–20
- Shusta EV, Boado RJ, Mathern GW, Pardridge WM (2002) Vascular genomics of the human brain. *J Cereb Blood Flow Metab* 22:245–52
- St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, Lal A, Riggins GJ, Lengauer C, Vogelstein B, Kinzler KW (2000) Genes expressed in human tumor endothelium. *Science* 289:1197–202
- Strange K (1992) Regulation of solute and water balance and cell volume in the central nervous system. *J Am Soc Nephrol* 3:12–27
- Trendelenburg G, Prass K, Priller J, Kapinya K, Polley A, Muselmann C, Ruscher K, Kannbley U, Schmitt AO, Castell S, Wiegand F, Meisel A, Rosenthal A, Dirnagl U (2002) Serial analysis of gene expression identifies metallothionein-II as a major neuroprotective gene in mouse focal cerebral ischemia. *J Neurosci* 22: 5879–5888
- Untergasser G, Koch HB, Menssen A, Hermeking H (2002) Characterization of epithelial senescence by serial analysis of gene expression: identification of genes potentially involved in prostate cancer. *Cancer Res* 62:6255–62
- Velculescu VE, Madden SL, Zhang L, Lash AE, Yu J, Rago C, Lal A, Wang CJ, Beaudry GA, Ciriello KM, Cook BP, Dufault MR, Ferguson AT, Gao Y, He TC, Hermeking H, Hiraldo SK, Hwang PM, Lopez MA, Luderer HF, Mathews B, Petroziello JM, Polyak K, Zawel L, Kinzler KW *et al* (1999) Analysis of human transcriptomes. *Nat Genet* 23:387–8
- Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. *Science* 270:484–7
- Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Jr, Hieter P, Vogelstein B, Kinzler KW (1997) Characterization of the yeast transcriptome. *Cell* 88:243–51
- Weller RO, Engelhardt B, Phillips MJ (1996) Lymphocyte targeting of the central nervous system: a review of afferent and efferent CNS-immune pathways. *Brain Pathol* 6:275–88
- You G, Lee WS, Barros EJ, Kanai Y, Huo TL, Khawaja S, Wells RG, Nigam SK, Hediger MA (1995) Molecular characteristics of Na(+)-coupled glucose transporters in adult and embryonic rat kidney. *J Biol Chem* 270:29365–71
- Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, Vogelstein B, Kinzler KW (1997) Gene expression profiles in normal and cancer cells. *Science* 276:1268–72

Supplementary information is available on the Journal of Cerebral Blood Flow & Metabolism website (<http://www.nature.com/jcbfm>).