SECTION EDITOR: EDWIN M. STONE, MD, PhD

Gene Expression in Donor Corneal Endothelium

John D. Gottsch, MD; Gerami D. Seitzman, MD; Elliott H. Margulies, PhD; Amanda L. Bowers, BA; Allison J. Michels, BA; Saurabh Saha, MD, PhD; Albert S. Jun, MD, PhD; Walter J. Stark, MD; Sammy H. Liu, PhD

Objective: To report gene expression profiles of normal human corneal endothelium with microarray analysis and serial analysis of gene expression (SAGE).

Methods: Corneal endothelium was removed from normal human corneas obtained from eye banks. Total RNA was isolated and SAGE analysis was performed. The same RNA source was used to construct a complementary DNA library that was hybridized to microarrays containing 12 558 transcripts.

Results: A total of 9530 SAGE tags were sequenced, representing 4724 unique tags. Microarray analysis identified 542 distinct transcripts. A database of human corneal endothelial gene expression was compiled. Of the SAGE tags, 1720 matched known genes, 478 corresponded to expressed sequence tags, and 2526 had no known match to public databases. The 5 most abundantly expressed SAGE tags were cytochrome *c* oxidase subunit II, adenosine triphosphate synthase F_0 subunit 6, carbonic anhydrase XII, 12S ribosomal RNA, and ferritin, heavy polypeptide 1. Thirty-four percent of the transcripts (n=1616) were specific to the corneal endothelium, when compared with other publicly available SAGE libraries. The 5 most abundant unique tags were keratin 12, angiopoietinlike factor, annexin A8, and 2 tags with no match to the database. Many endothelial pump function enzymes were confirmed, including several plasma membrane Na^+/K^+ adenosine triphosphatases and a recently reported bicarbonate transporter.

Conclusions: Corneal endothelial gene expression profiles by the current analysis provide an understanding of endothelial metabolism, structure, and function; enable comparisons to diseased endothelium; and provide baseline data that may lead to the discovery of novel endothelial genes.

Arch Ophthalmol. 2003;121:252-258

THE ENDOTHELIUM is essential for corneal clarity. Fluid transport across this cell layer balances stromal hydration by an active electrolyte pump and the maintenance of a semipermeable membrane by cell-cell adhesion complexes.[1] Descemet membrane is derived from the endothelium. Assessing endothelial molecular activity has been technically limited to evaluating only a few proteins or cellular transcripts at one time. Recently, the capability to globally assess gene expression in a given tissue has become feasible. Microarray analysis is a relatively rapid technique that has been useful in providing gene expression profiles for a number of tissues.[2,3] A previous work described the use of complementary DNA (cDNA) microarray technology to provide a gene expression profile of the human cornea, characterizing some 1200 genes.[4] Microarrays, however, are limited by the finite number of oligonucleotide sequences localized on a chip and are not able to identify the expression of novel genes.

Serial analysis of gene expression (SAGE) is another method that provides quantitative and comprehensive gene expression profiles. SAGE depends on the generation of short sequence tags at a specific location within a transcript.[5,6] Through a series of standard enzymatic reactions, 10–base pair SAGE tags, which contain sufficient information to uniquely identify a gene, are generated, concatenated, and sequenced. By identifying genes corresponding to each tag and tabulating the frequency of each tag, the number of genes expressed and their expression level can be estimated. Novel genes can be suspected in a given tissue when tags cannot be matched to publicly available sequences. SAGE has the additional advantage of providing quantitative information about gene expression. In the present study, we sought to expand the gene expression profile of the corneal endothelium. In contrast to our previous report of

From the Center for Corneal Genetics, Cornea and External Disease Service, The Wilmer Eye Institute (Drs Gottsch, Seitzman, Jun, Stark, and Liu and Mss Bowers and Michels), and Johns Hopkins Oncology Center Molecular Genetics Laboratory, The Johns Hopkins School of Medicine (Dr Saha), Baltimore, Md; and Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Md (Dr Margulies).

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

a microarray analysis of the total cornea, herein we report the use of an expanded microarray chip and, for the first time, performance of SAGE analysis of normal donor corneal endothelium.

METHODS

MICROARRAY ANALYSIS

A cDNA library was constructed by standard methods from the endothelium of 15 pairs of intact donor corneas (mean ± SD age, 52 ± 12 years) from the Maryland Eye Bank, Baltimore. The endothelium was stripped from the stroma and stored in liquid nitrogen. Total RNA was isolated from corneal endothelium using monophasic phenol/guanidine isothiocyanate solution (TRIzol Reagent; Invitrogen, Carlsbad, Calif). Double-stranded cDNA was synthesized from 15 µg of messenger RNA according to manufacturer's protocol (Stratagene, Cedar Creek, Tex). *XhoI* and *EcoRI* linker-primer adapters were incorporated into the cDNA to create the restriction sites at the 5′ and 3′ ends of the cDNA. The cDNA was size selected (>1 kb) by gel filtration, ligated into the UniZAPXR vector (Stratagene), and packaged with the use of extract (Gigapack II; Stratagene). The packaged cDNA was titered, and the number of clones contained in the primary cDNA library was 1.0×10^6 plaque-forming units per milliliter.

Standard methods were used to recover phagemids by mass excision protocol (Stratagene). Approximately 1.6×10^6 plasmids were excised. The ratio of clones excised to the number of independent clones in the library was 1.6:1. Excised clones were used to transfect a large-volume bacterial SOLR cell culture (Stratagene), and plasmid preparations were performed by standard methods (QIAGEN, Valencia, Calif). Plasmids were digested by means of *Not*I restriction endonuclease (Gibco-BRL Life Technologies), extracted with phenol-chloroform, and precipitated with ethanol.

Biotin-labeled cRNAs were produced by in vitro transcription (Enzo Diagnostics, Farmingdale, NY), digested with DNase I (Gibco-BRL Life Technologies), and purified by means of RNeasy spin columns (QIAGEN). Analysis of biotin-labeled cRNAs by means of the HU95a microarray (Affymetrix, Santa Clara, Calif) was performed in duplicate (Research Genetics, Huntsville, Ala) by hybridizing the same labeled cRNA sample to 2 identical microarrays within a 6-week period.

GENERATION OF SAGE LIBRARIES

Three pairs of intact donor corneas (from a 79-year-old woman and 58- and 66-year-old men) were obtained from Maryland Eye Bank. The corneal endothelium of each donor eye was examined and photographed. All donor corneas were found to have normal endothelia and handled in an identical manner. The endothelium was stripped from the stroma and immediately stored in liquid nitrogen until use. Normal endothelial SAGE libraries were constructed according to the SAGE protocol[5,6] (http://www.sagenet.org/sage_protocol.htm). Total RNA from normal endothelium was isolated by direct lysis using monophasic phenol/guanidine isothiocyanate solution (Invitrogen). Messenger RNA was isolated from total RNA by standard methods (FastTrack 2.0; Invitrogen). Two micrograms of messenger RNA was reverse-transcribed to double-stranded cDNA (Superscript Choice Synthesis cDNA synthesis kit; Invitrogen) with a 5′-biotinylated oligo$(dT)_{18}$ primer (Integrated DNA Technologies, Coralville, Iowa). Double-stranded cDNA was digested with *Nla*III (New England BioLabs, Beverly, Mass); 3′ cDNAs were purified with magnetic beads and split into 2 equal pools, and biotinylated SAGE linkers 1 and 2 (Integrated DNA Technologies) were ligated to pools 1 and 2, respectively. The SAGE tags were released with the tagging enzyme *Bsm*FI (New England BioLabs) and blunt ends were synthesized using Klenow polymerase fragment. The tags from pools 1 and 2 were ligated to each other overnight at 16°C. A 1:200 dilution of the ligation product was amplified with 35 cycles of polymerase chain reaction (PCR). Precipitated PCR products were separated on 12% polyacrylamide gel, and only the 102-bp band containing ditags was isolated. The ditags were released from the linkers by digestion with *Nla*III and purified by means of streptavidin magnetic beads.[7] The products of the digestion were separated on a 12% polyacrylamide gel, and the 24- to 26-bp bands containing ditags were purified and used for self-ligation overnight at 16°C. Concatamers were run on an 8% polyacrylamide gel and a fraction from 500 bp to 1 kb was isolated and cloned into pZERO vector (Invitrogen) digested with *Sph*I. Ligation mixtures were electroporated into *Escherichia coli* strain TOP10 F′ (Invitrogen), and colonies were screened for inserts larger than 500 bp by means of colony PCR with M-13 forward and reverse primers. The PCR products from selected clones were sequenced by means of an automated sequencer (ABI 3700; Applied Biosystems, Foster City, Calif).

MICROARRAY DATA ANALYSIS

By means of the statistical methods used by Affymetrix, individual probe sets were scored as either absent, marginal, or present in the endothelial sample. Probe sets were considered to be represented in the sample if they were marginal or present in both replicate microarray experiments. Probe sets were mapped to UniGene clusters by parsing the description line with a Perl script. Updated probe set descriptions were obtained at http://www.NetAffx.com. Fifteen probe sets represented in the endothelial sample had no UniGene cluster match and could therefore not be included in the analysis with the SAGE data.

SAGE TAG MATCHING

The ehm tag mapping method[8] (http://genome.nhgri.nih.gov/ehmTagMapping) was used to match SAGE tags with specific UniGene clusters. This method is implemented through the use of several Perl scripts designed to extract tag-to-UniGene cluster information from the UniGene flatfiles available at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/UniGene). Briefly, the ehm tag mapping method extracts a SAGE tag from each sequence in a UniGene cluster, only if the orientation and 3′ end of the sequence can be confirmed by identifying poly(A) signals and/or tails. To minimize the extraction of SAGE tags from entries with potential sequencing errors, SAGE tags not representing at least 20% of all tags extracted from a given UniGene cluster are removed from the final ehm tag mapping flatfile. For the purpose of comparing SAGE data with Affymetrix data, low-complexity (NAAAAAAAAA) and unreliable tag-to-UniGene cluster matches (tag matches only observed once) were removed from the final ehm tag mapping. Occasionally, one tag will reliably match more than one UniGene cluster. Using the UniGene and Locus Link databases at National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov), functional annotations were assigned to SAGE tags with the use of the associated gene ontology category (http://www.geneontology.org/) as a guide.

RESULTS

GENERATION OF SAGE DATA

A total of 9530 SAGE tags were sequenced from the normal corneal endothelium (**Table 1**), representing 4724

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

Table 1. Summary of the SAGE Libraries*

	Normal
Total SAGE tags excluding linkers, No.	9530
Unique tags, No.	4724
Linker contamination, No. (%)	7 (0.07)

Abbreviation: SAGE, serial analysis of gene expression.
*The SAGE tags were extracted from the *Phred*-analyzed sequence data (*.phd.1 files) using eSAGE version 1.2 such that sequences containing base calls with *Phred* quality values less than 20 (error rate of 1%) were not added to the database.

Table 2. Distribution of SAGE Tag Frequencies

% Abundance	Range of Tag Frequency	No. of Unique Tags
0.01	1	3843
>0.01-0.1	2-10	774
>0.1-0.5	11-46	88
>0.5-1	47-79	14
>1	>79	5

Abbreviation: SAGE, serial analysis of gene expression.

unique transcripts. Microarray analysis identified the expression of 542 distinct transcripts among 12558 total transcripts represented on the chip. A database of human corneal endothelial gene expression was compiled. Of the SAGE tags, 1720 (36%) matched known genes, 478 (10%) matched expressed sequence tags, and 2526 (54%) had no known match to publicly available databases. Of the 4724 unique tags, 3843 (81%) were represented by single copies, 774 (16%) by 2 to 10 copies, 88 (2%) by 11 to 46 copies, 14 (0.3%) by 47 to 79 copies, and 5 (0.1%) by greater than 79 copies (**Table 2**).

MOST ABUNDANTLY EXPRESSED SAGE TAGS AND FUNCTIONAL GROUPS

The 100 most abundant SAGE tags of human normal corneal endothelium are represented in **Table 3**. The percentage abundance for each of the genes ranged from 0.13% to 1.96%. The top 100 expressed transcripts were subdivided by cellular function. Ribosomal protein/RNA binding, transfer RNA, DNA binding, and gene regulation represented 28% of expressed tags; cellular metabolism and mitochondrial transcripts, about 18%; cell communication, cell growth/regulation, and cytoskeletal, 18%; expressed sequence tags, 8%; and genes with unknown functions, 7% (**Figure 1**). No-match tags composed 21% of the total expressed tags of the 100 most abundantly expressed genes.

To determine which genes were specific to the cornea, we performed a virtual subtraction between our SAGE and microarray data and 8 other publicly available SAGE libraries (see the "SAGE Tag Matching" subsection in the "Methods" section). Thirty-four percent of the transcripts (n = 1616) were identified as being specific to the corneal endothelium. The 5 most abundant cornea-specific tags were keratin 12, angiopoietinlike factor, annexin A8, and 2 tags with no match to the database.

Genes, identified by SAGE, representing previously recognized endothelial functions are listed in **Table 4**. Pump function, cytoskeletal, and cell adhesion proteins are listed under fluid transport. Collagens, glycosaminoglycans, and proteoglycans are listed under basement membrane proteins. The complete SAGE and microarray endothelial databases can be accessed at http://www.CorneaNet.net.

MICROARRAY ANALYSIS

Microarray detected 507 genes (**Figure 2**). Of these, 145 were also identified by SAGE. Three hundred sixty-two were detected by microarray only and 2164 transcripts were detected by SAGE only. Of the 100 most abundantly expressed transcripts detected by SAGE, 28 were also detected by microarray (Table 3).

COMMENT

In this study, SAGE and microarrays were used as complementary methods for assaying corneal endothelial gene expression. Since SAGE tags and Affymetrix probe sets could be mapped to corresponding UniGene clusters, we were able to compare the genes identified by each method. Because of the increased sensitivity of the SAGE method, greater than 4 times the number of expressed transcripts were detected when compared with microarray (2209 vs 507 UniGene clusters). There were 2164 UniGene clusters detected by SAGE that were not detected by microarray. One explanation for this bias is that the detection of a transcript by microarray relies on the statistical computation of data obtained from multiple, complex, and unpredictable in vitro hybridization assays, making it likely that numerous transcripts went undetected. Conversely, there were 362 UniGene clusters detected by microarray that were not detected by SAGE. Since the sensitivity of the SAGE method can be increased, simply by sequencing additional SAGE tags, it is likely that the number of UniGene clusters detected only by microarray will decrease as the number of total SAGE tags is increased.

Microarray was able to detect 25% of the 100 most abundantly expressed transcripts identified by SAGE. Some of these frequently expressed SAGE tags had no matches and may represent novel genes that would not have been available for detection on a microarray chip. Many of the 100 most abundantly expressed genes (18%) in the endothelium are those concerned with energy production including mitochondrial gene expression and glycolysis.

Corneal clarity is dependent on endothelial regulation of stromal hydration. The endothelium is a semipermeable membrane that regulates hydration by providing a barrier to the leakage of fluid into the stroma and by active ion transport across endothelial membranes.[1,9-16] The fluid transport across the endothelium is believed to be mediated by multiple distinct pumps, one involving Na^+/K^+ adenosine triphosphatase and another, carbonic anhydrase. A sodium ion concentration gradient is established by basolaterally located Na^+/K^+ ATPase.[1,9-13] Carbon dioxide is thought to diffuse across the endothelial membrane and with water is converted

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

Table 3. Most Abundantly Expressed Genes in the Corneal Endothelium

SAGE Tag	Frequency*	Abundance, %†	ID‡	Description§	ID by Microarray‖
CCCATCGTCC	187	1.96	MITO	Cytochrome *c* oxidase subunit II	. . .
CACCTAATTG	119	1.25	MITO	ATP synthase F_0 subunit 6	. . .
CCACTGCACT	107	1.12	5338	Carbonic anhydrase XII	No
GTAAGTGTAC	102	1.07	MITO	12S ribosomal RNA	. . .
TTGGGGTTTC	98	1.03	62954	Ferritin, heavy polypeptide 1	Yes
GTGAAACCCC	79	0.83	262716	ESTs, weakly similar to *Homo sapiens* unnamed protein product	No
ACGGAACAAT	76	0.80	8272	Prostaglandin D_2 synthase (21 kd, brain)	Yes
CTAAGACTTC	74	0.78	MITO	16S ribosomal RNA	. . .
GGCCCAGGCC	71	0.75	575	Aldehyde dehydrogenase 3 family, member A1	Yes
GCATAATAGG	69	0.72	184108	Ribosomal protein L21	Yes
CCCTACCCTG	68	0.71	75736	Apolipoprotein D	Yes
GTGTGTTTGT	66	0.69	118787	Transforming growth factor β–induced, 68 kd	No
TACCATCAAT	59	0.62	169476	Glyceraldehyde-3-phosphate dehydrogenase	No
GCCCCTGCTG	54	0.57	195850	Keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Koebner/ Weber-Cockayne types)	Yes
GGGAAGCAGA	53	0.56	. . .	No match	. . .
CTCATAAGGA	52	0.55	294348	EST	. . .
CACAAACGGT	50	0.52	195453	Ribosomal protein S27 (metallopanstimulin 1)	Yes
GTGAAACCCT	49	0.51	51692	DKFZP434C091 protein	Yes
TTGGTCTTTG	48	0.50	66739	Keratin 12 (Meesmann corneal dystrophy)	Yes
TTCATACACC	46	0.48	MITO	NADH dehydrogenase subunit 4	. . .
TGTGTTGAGA	45	0.47	288036	tRNA isopentenylpyrophosphate transferase	Yes
CCTGTAATCC	44	0.46	274448	Hypothetical protein FLJ11029	. . .
CAACTAATTC	41	0.43	75106	Clusterin (complement lysis inhibitor, SP-40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)	Yes
GCTTTCTCAC	35	0.37	. . .	No match	. . .
GTGACCACGG	35	0.37	. . .	No match	. . .
TCGAAGCCCC	35	0.37	. . .	No match	. . .
CAAGCATCCC	34	0.36	. . .	No match	. . .
AGCCCTACAA	32	0.34	MITO	NADH dehydrogenase subunit 3	. . .
TAGGTTGTCT	31	0.33	326456	Hypothetical protein FLJ20030	. . .
TTGGCCAGGC	31	0.33	. . .	No match	. . .
TCAGATCTTT	29	0.30	108124	Ribosomal protein S4, X-linked	Yes
GAGGGAGTTT	28	0.29	76064	Ribosomal protein L27a	Yes
TGATTTCACT	27	0.28	MITO	Cytochrome *c* oxidase subunit III	. . .
GGATTTGGCC	27	0.28	119500	Ribosomal protein, large P2	Yes
TGGCGTACGG	27	0.28	. . .	No match	. . .
CACTACTCAC	26	0.27	MITO	Cytochrome *b*	. . .
TTCACTGTGA	25	0.26	621	Lectin, galactoside-binding, soluble, 3 (galectin 3)	Yes
TTGGTCCTCT	25	0.26	324406	Ribosomal protein L41	. . .
AAAACATTCT	25	0.26	. . .	No match	. . .
AACCCAGGAG	24	0.25	313159	ESTs, weakly similar to A46010 X-linked retinopathy protein [*Homo sapiens*]	. . .
AGGTCAGGAG	24	0.25	344177	ESTs	. . .
CCATTGCACT	24	0.25	344092	*Homo sapiens*, clone MGC:20188 IMAGE:4564707, mRNA, complete cds	. . .
CCAGAACAGA	24	0.25	334807	Ribosomal protein L30	. . .
TGAGCCTCGT	23	0.24	254105	Enolase 1(α)	. . .
ACTAACACCC	23	0.24	MITO	NADH dehydrogenase subunit 2	. . .
AGGTGGCAAG	23	0.24	. . .	No match	. . .
TGTGCTAAAT	22	0.23	250895	Ribosomal protein L34	. . .
ACACAGCAAG	22	0.23	. . .	No match	. . .
GGTCAGTCGG	22	0.23	. . .	No match	. . .
TTGGCCAGGA	22	0.23	. . .	No match	. . .
ACCCTTGGCC	21	0.22	MITO	NADH dehydrogenase subunit 1	. . .
AGAAAGATGT	21	0.22	78225	Annexin A1	No
GTAATCCTGC	21	0.22	. . .	No match	. . .
TAATAAAGGT	21	0.22	151604	Ribosomal protein S8	Yes
TGCATCTGGT	21	0.22	75410	Heat shock 70-kd protein 5 (glucose-regulated protein, 78 kd)	No
TGGTGTTGAG	21	0.22	275865	Ribosomal protein S18	. . .
AACCCGGGAG	20	0.21	. . .	No match	. . .
AGCACCTCCA	20	0.21	75309	Eukaryotic translation elongation factor 2	Yes
CTGGGTTAAT	20	0.21	298262	Ribosomal protein S19	. . .
GACCAGCTGG	20	0.21	74120	Adipose specific 2	Yes
GCTGACTCAG	20	0.21	105607	Bicarbonate transporter–related protein 1	. . .

(continued)

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

Table 3. Most Abundantly Expressed Genes in the Corneal Endothelium (cont)

SAGE Tag	Frequency*	Abundance†	ID‡	Description§	ID by Microarray‖
CAGGTTTCAT	19	0.20	24395	Small inducible cytokine subfamily B (Cys-X-Cys), member 14 (BRAK)	
GCGAAACCCC	18	0.19	. . .	No match	. . .
GGATATGTGG	18	0.19	326035	Early growth response 1	No
TGGAAAGTGA	18	0.19	25647	v-*fos* FBJ murine osteosarcoma viral oncogene homologue	No
TGGCAGTCTG	18	0.19	288982	*Homo sapiens* cDNA: FLJ23475 fis, clone HIS13659	. . .
CCTCAGGATA	17	0.18	335919	EST, weakly similar to I55214 salivary proline-rich glycoprotein precursor, rat [*Rattus norvegicus*]	. . .
GGACCACTGA	17	0.18	119598	Ribosomal protein L3	Yes
AACCTGGGAG	16	0.17	. . .	No match	. . .
AGGTCCTAGC	16	0.17	226795	Glutathione-*S*-transferase π	No
ATCGCTTTCT	16	0.17	177486	Amyloid β (A4) precursor protein (protease nexin-II, Alzheimer disease)	Yes
CCTAGCTGGA	16	0.17	342389	Peptidylprolyl isomerase A (cyclophilin A)	. . .
GCAAGCCAAC	16	0.17	332048	ESTs, weakly similar to 810024J URF 4 [*Homo sapiens*]	. . .
GGGCTGGGGT	16	0.17	183698	Ribosomal protein L29	No
GTGAAGGCAG	15	0.16	77039	Ribosomal protein S34	No
TTGGTCAGGC	15	0.16	209159	ESTs	. . .
TTGGTGAAGG	15	0.16	75968	Thymosin, β 4, X chromosome	No
AAGGAGATGG	14	0.15	184014	Ribosomal protein L31	Yes
CCGACGGGCG	14	0.15	. . .	No match	. . .
GAAGCAGGAC	14	0.15	180370	Cofilin 1 (nonmuscle)	No
GAAGTCGGAA	14	0.15	. . .	No match	. . .
GCCTTCCAAT	14	0.15	76053	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 5 (RNA helicase, 68 kd)	No
GGTCCAGTGT	14	0.15	181013	Phosphoglycerate mutase 1 (brain)	No
ACAGATTTGA	13	0.14	41271	*Homo sapiens* mRNA full-length insert cDNA clone EUROIMAGE 1913076	. . .
AGTAGGTGGC	13	0.14	. . .	No match	. . .
ATAATTCTTT	13	0.14	539	Ribosomal protein S29	Yes
ATTTGAGAAG	13	0.14	MITO	Cytochrome *c* oxidase subunit 1	. . .
CCTGTGTTGG	13	0.14	. . .	No match	. . .
CGCCGCCGGC	13	0.14	182825	Ribosomal protein L35	Yes
GGCCCCTCAC	13	0.14	274313	Insulinlike growth factor binding protein 6	No
TCACCGGTCA	13	0.14	290070	Gelsolin (amyloidosis, Finnish type)	
TGCACGTTTT	13	0.14	169793	Ribosomal protein L32	Yes
GAAATACAGT	12	0.13	67201	5′(3′)-Deoxyribonucleotidase	
GACGACACGA	12	0.13	153177	Ribosomal protein S28	Yes
GAGACTCCTG	12	0.13	169902	Solute carrier family 2 (facilitated glucose transporter), member 1	No
GCCGAGGAAG	12	0.13	339696	Ribosomal protein S12	. . .
GCCTGTATGA	12	0.13	180450	Ribosomal protein S24	Yes
GGCAAGAAGA	12	0.13	111611	Ribosomal protein L27	Yes
GGTCACTGAG	12	0.13	347378	ESTs	. . .
TTTAACGGCC	12	0.13	. . .	No match	. . .

Abbreviations: ATP, adenosine triphosphate; cDNA, complementary DNA; EST, expressed sequence tag; mRNA, messenger RNA; NADH, nicotinamide adenine dinucleotide; SAGE, serial analysis of gene expression; tRNA, transfer RNA.

*Number of times tag was observed.

†Proportion of total tags in SAGE library.

‡UniGene cluster identifier as determined by the ehm tag mapping method. Tags matching mitochondrial transcripts (MITO) were identified from the complete human mitochondrial sequency (NC_001807). See "Methods" section for details.

§UniGene descriptor of gene corresponding with SAGE tag.

‖"Yes" denotes the detection of gene on Affymetrix chip. "No" denotes that the gene was represented on the Affymetrix chip but not detected. Ellipsis denotes it was not represented on the Affymetrix chip. See "Methods" section for details.

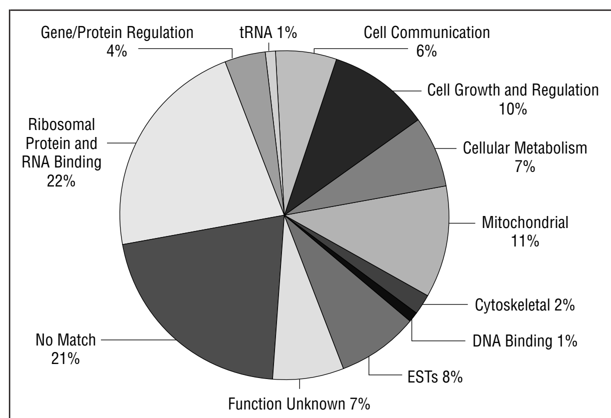

Figure 1. A functional group analysis of gene expression of human donor corneal endothelium showing the top 100 expressed transcripts. EST indicates expressed sequence tag; tRNA, transfer RNA.

by carbonic anhydrase to bicarbonate and H^+.[16,17] Protons may leave the cell via a Na^+/H^+ exchanger, and bicarbonate efflux has been proposed to be driven by a Cl^-/HCO_3 exchanger.[11,12] Several of these pump proteins were detected by SAGE (Table 4), including different Na^+/K^+ adenosine triphosphatases and carbonic anhydrase transcripts (Table 4). A recently discovered bicarbonate transporter that has been found in kidney, salivary glands, and others tissues[16] was identified by SAGE. Na^+/H^+ lysosomal proton transporter was identified in the corneal endothelium, but a plasma membrane-bound proton exchanger was not identified with this SAGE library of 10000 transcripts. A SAGE library sequenced to a greater depth may identify other pump function transcripts.

Another mechanism regulating fluid transport in the cornea is the endothelial barrier to fluid and electrolyte

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

Table 4. Genes Identified by SAGE and Roles in Corneal Endothelial Function

Description	Tag	Tag Frequency
Pump Function		
Bicarbonate transporter-related protein 1	GCTGACTCAG	20
ATPase, Na^+/K^+ transporting, α1 polypeptide	TAGCTCTATG	4
ATPase, Na^+/K^+ transporting, β3 polypeptide	TAGGATGGGG	1
ATPase, Na^+/K^+ transporting, β1 polypeptide	TTCTAACATA	3
Carbonic anhydrase XII	CCACTGCACT	107
ATPase, H^+ transporting, lysosomal (vacuolar proton pump), subunit 1	GGAAAGTGAC	2
ATPase, H^+ transporting, lysosomal (vacuolar proton pump), 16 kd	CGCAGTGTCC	1
ATPase, H^+ transporting, lysosomal (vacuolar proton pump), 21 kd	TTTGGGGCTG	1
Potassium channel, subfamily K, member 6 (TWIK-2)	CCTGTGATCC	1
Cytoskeletal Function		
Tight junction protein 2 (zona occludens 2)	TATTGACAAC	1
Actinin, α4	TAATATTTTT	1
SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily d, member 2	TGGAGGCCAG	2
Actin cross-linking factor	TTGCTGTGTG	1
Actin filament-associated protein	GTGATACGTT	1
Actin, γ1	CTAGCCTCAC	3
ARP1 (actin-related protein 1, yeast) homologue A (centractin α)	GACCTGCGGC	1
Microsomal glutathione-*S*-transferase 2	AAGGTAATAT	1
Gap junction protein, α1, 43 kd (connexin 43)	TGTTCTGGAG	10
Gap junction protein, β2, 26 kd (connexin 26)	GTTTCCAAAA	1
Gap junction protein, β6 (connexin 30)	ATACGCTTAA	6
Gap junction protein, β2, 26 kd (connexin 26)	ACTGTGGTAG	1
Intermediate filament protein syncoilin	ACTAACTGTG	2
Plectin 1, intermediate filament binding protein, 500 kd	TTCCACTAAC	1
Keratin 12 (Meesmann corneal dystrophy)	TTGGTCTTTG	48
Keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Koebner/Weber-Cockayne types)	GCCCCTGCTG	54
Keratin 14 (epidermolysis bullosa simplex, Dowling-Meara, Koebner)	GATGTGCACG	3
Keratin 15	TAATAAAGAA	1
Keratin 17	CTTCCTTGCC	1
Keratin 18	CAAACCATCC	1
Keratin 4	GGCAGAGAAG	1
Keratin 6A	AAAGCACAAG	1
Cadherin 13, H-cadherin (heart)	CACACACACA	1
Cadherin 1, type 1, E-cadherin (epithelial)	TGTGGGTGCT	1
Catenin (cadherin-associated protein), α-like 1	AAACTGATTG	1
Protocadherin γ subfamily B, 6	GGCTTTGGAG	1
Catenin (cadherin-associated protein), α-like 1	AAACTGATTG	1
Basement Membrane Proteins		
Fibrous proteins		
Collagen, type I, α2	TTTGGTTTTC	1
Collagen, type VIII, α2	TTAACTGTAT	9
Collagen, type VI, α1	TTGCTGACTT	5
Collagen, type V, α3	CCAGGTGGTT	1
Collagen, type XIII, α1	GCAAAAAAAA	2
Procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase, Ehlers-Danlos syndrome type VI)	AGAGCAAACC	1
Collagen, type XVII, α1	GATATGTTAT	1
Collagen, type IX, α3	AAGGAGCGGG	1
Procollagen (type III) N-endopeptidase	GGCCAGGTGG	2
Collagen, type IV, α6	GTGTCAGTTT	1
Collagen, type I, α1	TGGAAATGAC	1
Collagen, type XII, α1	AGGAGCCTCA	1
Collagen, type VII, α1 (epidermolysis bullosa, dystrophic, dominant and recessive)	GTGCTGATTC	1
Adhesive proteins		
Laminin receptor 1 (67 kd, ribosomal protein SA)	GAAAAATGGT	7
Laminin, γ2 (nicein [100 kd], kalinin [105 kd], BM600 [100 kd], Herlitz junctional epidermolysis bullosa)	GCAACTTAGA	1
Glycosaminoglycans		
Syndecan 2 (heparan sulfate proteoglycan 1, cell surface-associated, fibroglycan)	TGGCCTAATA	4
Heparan sulfate (glucosamine) 3-*O*-sulfotransferase–1	AGCCAAAAAA	1
Transforming growth factor β receptor III (betaglycan, 300 kd)	AAACTGACAG	1
Proteoglycans		
Lumican	TTATGTTTAA	2
	TTAACAATTC	1
Decorin	ACTTATTATG	10

Abbreviations: ATPase, adenosine triphosphatase; SAGE, serial analysis of gene expression.

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.

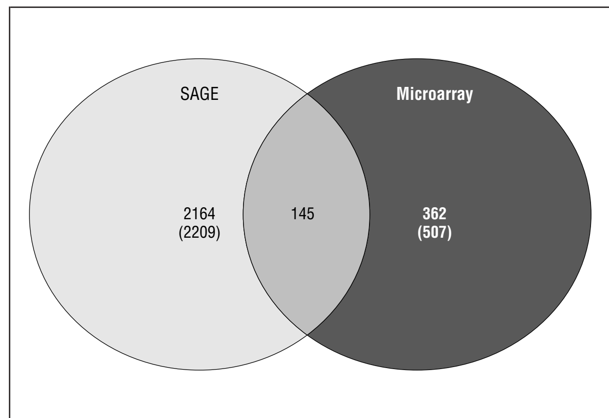

Figure 2. Venn diagram representation of UniGene clusters identified by serial analysis of gene expression (SAGE) and microarray. Numbers in parentheses are the total number identified by each method.

transport into the stroma.[17] This endothelial barrier is a function of cytoskeletal proteins and intercellular adhesion structures. Protein components of this barrier were identified by SAGE and include tight and gap junctions (actins and connexin), cytoplasmic intermediate filaments (keratins and vimentin), and Ca^{2+}-dependent cell-cell adhesion proteins (cadherin and catenin) (Table 4).

A notable function of the endothelium is its contribution to a basal lamina and to Descemet membrane. Transcripts that may code for proteins contributing to these structures are fibrous proteins including collagens and elastin transcripts (Table 4). Adhesive protein transcripts detected were those for laminin, fibronectin, and fibrillin. Glycosaminoglycans noted were hyaluronan, chondroitin sulfate, heparan, and keratan. Proteoglycans identified were lumican, decorin, and betaglycan (transforming growth factor β receptor III). Integrins as linker proteins whose external domains bind to the extracellular matrix were also detected by SAGE.

This study is the first, to our knowledge, to apply SAGE and microarray techniques to comprehensively assess gene expression in the human corneal endothelium. The results expand our understanding of endothelial metabolism, structure, and function. The full SAGE dataset can be accessed at the National Center for Biotechnology Information gene expression Omnibus repository with accession numbers GSM1652 for normal corneal endothelium. We have also added these gene expression profiles to an online database, http://www.CorneaNet.net, so that theoretic comparisons with other SAGE analyses in corneal diseases such as corneal dystrophies and graft rejection can be performed, possibly suggesting further areas of laboratory study.

Submitted for publication April 4, 2002; final revision received August 30, 2002; accepted October 23, 2002.

This study was supported in part by a grant from Research to Prevent Blindness Inc, New York, NY (Dr Gottsch) and by the Helen and Raymond Kwok Research Fund (Dr Stark).

Corresponding author: John D. Gottsch, MD, The Wilmer Eye Institute, The Johns Hopkins Hospital, 600 N Wolfe St, Baltimore, MD 21287 (e-mail: jgottsch@jhmi.edu).

REFERENCES

1. Geroski DH, Mamoru M, Yee RW, Edelhauser HF. Pump function of the human corneal endothelium: effects of age and cornea guttata. *Ophthalmology.* 1985; 92:759-763.
2. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet.* 1999;21(suppl):20-24.
3. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet.* 1999;21(suppl):10-14.
4. Jun AS, Liu SH, Koo EH, Do DV, Stark WJ, Gottsch JD. Microarray analysis of gene expression in human donor corneas. *Arch Ophthalmol.* 2001;119:1629-1634.
5. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science.* 1995;270:484-487.
6. St Croix B, Rago C, Velculescu V, et al. Genes expressed in human tumor endothelium. *Science.* 2000;289:1192-1202.
7. Powell J. Enhanced concatemer cloning: a modification to the SAGE (Serial Analysis of Gene Expression) technique. *Nucleic Acids Res.* 1998;26:3445-3446.
8. Margulies EH, Kardia SLR, Innis JW. A comparative molecular analysis of developing mouse forelimbs and hindlimbs using serial analysis of gene expression (SAGE). *Genome Res.* 2001;11:1686-1698.
9. Geroski DH, Kies JC, Edelhauser HF. The effects of ouabain on endothelial function in human and rabbit corneas. *Curr Eye Res.* 1984;3:331-338.
10. Geroski DH, Edelhauser HF. Quantitation of Na/K ATPase pumpsites in the rabbit corneal endothelium. *Invest Ophthalmol Vis Sci.* 1984;25:1056-1060.
11. Bonanno JA, Giasson C. Intracellular pH regulation in fresh and cultured bovine corneal endothelium, I: Na^+/H^+ exchange in the absence and presence of HCO_3^-. *Invest Ophthalmol Vis Sci.* 1992;33:3058-3067.
12. Bonanno JA, Giasson C. Intracellular pH regulation in fresh and cultured bovine corneal endothelium, II: Na^+:HCO_3^- cotransport and Cl^-/HCO_3^- exchange. *Invest Ophthalmol Vis Sci.* 1992;33:3068-3079.
13. Stiemke MM, Roman RJ, Palmer ML, Edelhauser HF. Sodium activity in the aqueous humor and corneal stroma of the rabbit. *Exp Eye Res.* 1992;55:425-433.
14. Silverman DN, Gerster R. The detection and localization of carbonic anhydrase in the rabbit cornea. *Exp Eye Res.* 1973;17:129-136.
15. Kuang K, Xu M, Koniarek JP, Fischbarg J. Effects of ambient bicarbonate, phosphate, and carbonic anhydrase inhibitors on fluid transport across rabbit corneal endothelium. *Exp Eye Res.* 1990;50:487-493.
16. Parker MD, Ourmozdi EP, Tanner MJA. Human BTR1, a new bicarbonate transporter superfamily member and human AE4 from kidney. *Biochem Biophys Res Commun.* 2001;282:1103-1109.
17. Kim EK, Geroski DH, Holley GP, Urken SI, Edelhauser HF. Corneal endothelial cytoskeleton changes in F-actin with aging, diabetes, and after cytochalasin exposure. *Am J Ophthalmol.* 1992;114:329-335.

Downloaded from www.archophthalmol.com at Johns Hopkins University, on October 29, 2007
©2003 American Medical Association. All rights reserved.