Gene Expression Profile Studies of Human Keratoconus Cornea for NEIBank: A Novel Cornea-Expressed Gene and the Absence of Transcripts for Aquaporin 5

Yaron S. Rabinowitz,[1,2] *Lijin Dong,*[2,3] *and Graeme Wistow*[3]

Purpose. To increase the database of genes expressed in human cornea and to gain insights into the molecular basis of keratoconus (KC).

Methods. A cDNA library was constructed from KC corneas harvested at keratoplasty and used for expressed sequence tag (EST) analysis. Data were analyzed using grouping and identification of sequence tags (GRIST). Expression of selected clones was examined by RT-PCR.

Results. A total of 7680 clones was sequenced from the 5′ end. After bioinformatics analysis, 4090 clusters of clones, each potentially representing individual genes, were identified. Of these, 887 genes were represented by more than one clone. The five most abundant transcripts, represented by >60 clones each, were for keratin-12, TGFBI (BIGH3), decorin, ALDH3, and enolase 1, all known markers for cornea. Many other markers for epithelial, stromal, and endothelial genes were also present. One cluster of six clones came from an apparently novel gene (designated *KC6*) located on chromosome 18 at p12.3. RT-PCR of RNA from several human tissues detected KC6 transcripts only in cornea. In addition, no clones were observed for the usually prominent corneal epithelial cell marker aquaporin 5 (AQP5), a water channel protein. Semiquantitative RT-PCR confirmed that expression of AQP5 is much lower in KC cornea than in non-KC cornea.

Conclusions. This analysis increases the database of genes expressed in the human cornea and provides insights into KC. *KC6* is a novel gene of unknown function that shows cornea-preferred expression, whereas the suppression of transcripts for AQP5 provides the first clear evidence of a molecular defect identified in KC. (*Invest Ophthalmol Vis Sci.* 2005;46: 1239–1246) DOI:10.1167/iovs.04-1148

In the past decade, significant progress has been made in pinpointing genetic loci for corneal dystrophies.[1] This success has been accompanied by similar advances in developing techniques for gene therapy in the cornea.[2] Indeed, the exposed corneal epithelium is an unusually convenient potential site for such therapy. Despite these advances, few specific gene defects have been identified that clearly illustrate the pathogenetic pathways that cause any of the major types of corneal disorders.

Definition of the transcriptional repertoires of normal and diseased human cornea is an important step in understanding this tissue and in the development of therapies. The NEIBank (Bethesda, MD) project for ocular genomics has used expressed sequence tag (EST) analysis of cDNA libraries to examine gene expression profiles in several freshly dissected human eye tissues.[3,4] However, data for normal cornea are rather limited, at least partly because of the problems in obtaining fresh human corneal tissue, since donated corneas are typically held for as long as possible for use in transplantation. Some useful EST data for corneal gene expression have been obtained, most notably through the BodyMap project (http://bodymap.ims.u-tokyo.ac.jp/, a collaborative effort of the Institute of Molecular and Cell Biology, Osaka, Japan, and the University of Tokyo, Japan). This work identified 749 potentially independent gene clones from a corneal epithelium cDNA library[5] and 1460 from endothelium.[6] SAGE (serial analysis of gene expression) has also been performed on human endothelium[7] and on mouse whole cornea.[8]

Keratoconus (KC) is a noninflammatory corneal thinning disorder that leads to loss of visual acuity through ectasia, opacity, and irregular astigmatism and is a major reason for corneal transplantation. Although one form of the disease, associated with abnormal function of the retina, is attributable to mutation in the retinal transcription factor VSX1,[9] the genetic bases for most forms of KC remain poorly defined.[10] Identification of the major molecular defects in KC would open up possibilities for therapeutic intervention. Because early forms of the disease are easily detectable with videokeratography before the onset of clinical disease, it is potentially amenable to intervention by gene therapy at an early stage to retard its progression.[10] A microarray analysis comparing KC and myopic corneas has identified some genes that are apparently upregulated in KC,[11] but otherwise very little is known about the transcriptional repertoire of the KC cornea. In addition to *VSX1*, on chromosome 20 at p11.21, six other loci for KC have been mapped: 3p14-q13, 5q14.3-q21.1, 15q22.23-q24, 16q, 20q12, and 21q.[10,12–14] This genetic heterogeneity suggests that mutations in several different members of related pathways may converge on common targets responsible for the disease phenotype.

As an alternative approach to increase the coverage of cornea-expressed genes, we harvested fresh human corneas from patients with KC at the time of penetrating keratoplasty to make a high-quality cDNA library, which is used for EST analysis. The analysis provides a considerable increase in the database of known human cornea-expressed genes and also identifies some potentially interesting features of gene expression in KC. Because no comparable set of clones derived in the same way for normal cornea is yet available, it is not possible

From the [1]Cornea Genetic Eye Institute, Cedars-Sinai Medical Center, Los Angeles, California; and the [3]Section on Molecular Structure and Function, National Eye Institute, National Institutes of Health, Bethesda, Maryland.

[2]Contributed equally to the work and therefore should be considered equivalent authors.

Presented in part at the Association for Research and Vision in Ophthalmology meeting Fort Lauderdale, Florida, April 2004.

Supported in part by the Eye Birth Defects Research Foundation, Inc. and the Center for Keratoconus.

Submitted for publication September 27, 2004; revised November 24, 2004; accepted December 12, 2004.

Disclosure: **Y.S. Rabinowitz**, None; **L. Dong**, None; **G. Wistow**, None

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Graeme Wistow, Chief, Section on Molecular Structure and Function, National Eye Institute, Building 7, Room 201, National Institutes of Health, Bethesda, MD 20892-0703; graeme@helix.nih.gov.

to make detailed comparisons of levels of expression in KC. However, many genes of potential importance for normal corneal function were revealed, one usually abundant transcript was absent from KC, while a novel, cornea-preferred gene was identified in this collection of clones.

METHODS

Tissue Procurement

For library construction, seven keratoconus (KC) corneal host buttons, approximately 7.5 mm in size, were removed with a Barron Hessburg trephine at the time of penetrating keratoplasty. A piece of tissue was sent for histopathology, and the rest was immediately immersed in preservative (RNAlater; Ambion, Austin, TX). All patients had moderate to advanced keratoconus, with some having moderate corneal scarring. None of the patients had worn contact lenses for ≥3 months before undergoing cornea transplantation. The procedure for obtaining the tissues was within the tenets of the Declaration of Helsinki. Before corneal transplants the patients not only consented to having their corneas removed for therapeutic reasons but also signed an additional consent form approved by the Western Institutional Review Board (http://www.irb.com), allowing their corneas to be used for genetic research. Two additional corneas, one with keratoconus and one essentially normal cornea with mild endothelial polymorphous dystrophy were collected in the same way and used in the RT-PCR experiments. In addition, 9-mm diameter samples of normal (myopic) and KC corneal epithelium were obtained at the time of photorefractive (PRK) or phototherapeutic keratectomy (PTK) and were also used for RT-PCR. The myopic patients ranged from −6 D to −9 D and were soft contact lens wearers but had not worn their contact lenses for at least 2 weeks before the procedure.

cDNA Library Construction

Seven KC corneal buttons were pooled, and total RNA was extracted (RNAzol; Tel-Test Inc., Friendswood, TX). A portion (40 μg) of total RNA was used for cDNA synthesis. Poly(A)+ RNA was prepared with an oligo-dT cellulose affinity column.

Oligo-dT primed cDNA was synthesized at Bioserve Biotechnology (Laurel, MD), using a DNA synthesis program (Superscript II; Invitrogen, Carlsbad, CA), as described previously.[15] The cDNA was run over a resin column (Sephacryl S-500 HR; Invitrogen) to fractionate cDNA larger than 500 bp. The first two 35-μL fractions, containing cDNA, were directionally cloned in *Not*I/*Sal*I sites in the pCMVSPORT6 vector (Invitrogen) to make cDNA sublibraries with the code designations *od* and *oe*.

cDNA Sequencing and Bioinformatics

Methods for sequencing and bioinformatics analysis are described in detail elsewhere.[15,16] Briefly, randomly picked clones were sequenced from the 5′ end at the NIH Intramural Sequencing Center (NISC). GRIST (Grouping and Identification of Sequence Tags) was used to analyze and assemble the data and to display the results in Web page format.[16] Clusters of sequences were also examined (SeqMan II; DNAstar, Madison, WI) to check the assembly of clusters and to examine alternative transcripts. Sequences were also searched through genome resources at the National Institutes of Health (http://www.ncbi.nlm.nih.gov/) and the University of California at Santa Cruz (http://genome.ucsc.edu/). Keywords for protein function were extracted from GO (Gene Ontology) annotations.[17]

RT-PCR Detection of *AQP5* and *KC6* Transcripts

Corneal samples were stabilized (RNAlater; Ambion) immediately after harvesting and individually processed for total RNA isolation. Total RNA was isolated (RNeasy Mini Kit; Qiagen GmbH, Hilden, Germany), according to manufacturer-recommended procedures. Briefly, 30 mg or less of each corneal tissue was first homogenized in a guanidinium-isothiocyanate–containing solution. Subsequently, homogenized tissue suspension was added to a mini silicone column, and bound RNA was washed and eluted in 30 μL of deionized water. The concentration of total RNA was then determined by its optical density. Reverse transcriptase (RT) reactions were performed using 50 ng total RNA and random hexamers (Sensiscript RT Kit; Qiagen) by manufacturer-recommended procedures. Individual cDNA pools obtained from RT reactions were then used as templates for PCR detection of AQP5 and KC6 expression using the following primers: (1) *AQP5*: TCCACTGACTCCCGCCGCACCAGC; AQP5 reverse: *KC6*TCAGCGGGTGGTCAGCTCCATGGT; (2) 5′ forward: AGAATCCAGGGGAAGATGAAGCAGC; KC6 5′ reverse: *KC6*GAGGAAGCATAGGTTGAATGATCTG; and (3) 3′ forward: ATTGTGTATACATGGAGGTGGGATG; KC6 3′ reverse; AGGTCAAGCAATCTAAGCTGCATAG.

RESULTS AND DISCUSSION

Initial quality control sequencing of the *od* and *oe* sublibraries showed that they were very similar in composition. Subsequently, an approximately equal number of clones were sequenced from both sublibraries, and the data were pooled. The combined *od*/*oe* libraries yielded 7680 sequence reads of average length over 600 bp. The data contained approximately 2% contamination with mitochondrial genome sequences and <1% rRNA clones. After analysis by GRIST, 4090 clusters of clones, each potentially representing an individual transcribed gene, were identified. This is the largest set of cornea-derived clones currently available and can be viewed in its entirety on the cornea page of the NEIBank Web site (http://www.neibank.nei.nih.gov), which also includes a GRIST-analyzed representation of the BodyMap data.

Abundantly Expressed Genes

Table 1 shows the most abundant transcripts, genes represented by eight or more cDNA clones. The five most abundant clones are all for known corneal markers, keratin-12, TGFBI (BIGH3), decorin, ALDH3, and enolase 1.[1,5,18–23] Genes expressed in all three regions of the cornea are represented, including keratin-12 and -3 from the epithelium, decorin and keratocan from the stroma,[24] and ovary-specific protein from the endothelium.[6] Several other keratins and proteoglycans known to function in the cornea were also abundant, including keratin-3 and -5, keratocan, and lumican. Previously, in a specific effort to examine changes in gene expression in KC, microarray analysis was used to compare expression of 5600 genes using cDNA from KC and myopic corneas.[11] In this study, increased relative expression of keratins 6 and 13 was noted in the KC samples, but these are both absent from the cDNA library collection.

Another marker for epithelial cells that appeared in the top 50 cDNAs is desmoglein 1 (DSG1), which is surprisingly abundant. The current version of Unigene contains only 35 ESTs for this gene (Hs.2633) from all sources, mostly from cell lines or tumors (http://www.ncbi.nlm.nih.gov/UniGene; provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD). In the KC cornea collection alone there are nine clones for DSG1 plus another group of 4 ESTs located just 3′ to the gene in the human genome that are most likely derived from longer 3′ UTRs for the same gene. The BodyMap collection also contains two 3′ clones for DSG1. DSG1 is a calcium-binding transmembrane glycoprotein of the cadherin superfamily that is a component of desmosomes and is associated with the most differentiated cells within the cutaneous epithelium.[25] A single clone for the related DSG2 is also present. The previous microarray analysis of KC cornea found expression of a third member of the family, DSG3, to be

TABLE 1. The Most Abundantly Represented Genes in the KC cDNA Library

Gene Name	GenBank ID	Clones (*n*)
K12 keratin	D78367	283
Transforming growth factor-beta–induced gene product (*BIGH3*)	M77349	101
Decorin variant A	AF138300	101
Aldehyde dehydrogenase type III (*ALDHIII*)	M74542	95
Enolase 1, (alpha) (*ENO1*)	NM_001428	61
TRPM-2/clusterin	M64722	46
Keratin 3 (*KRT3*),	NM_057088	35
Eukaryotic translation elongation factor 1 alpha 1 (*EEF1A1*),	NM_001402	32
Keratin type II (58 kDa)	M21389	31
Ferritin heavy-chain subunit	AF088851	29
Keratan sulfate proteoglycan	AF063301	25
NAD(P)H:menadione oxidoreductase	J03934	23
Transmembrane protein (*THW* gene)	AJ251830	23
Angiopoietin-like factor (*CDT6*),	NM_021146	22
Aldehyde dehydrogenase 1 family, member A1 (*ALDH1A1*)	NM_000689	21
Lumican (*LUM*),	NM_002345	19
Prosaposin	D00422	18
Lipocortin II	D00017	17
Prostaglandin D2 synthase	NM_000954	13
Polyubiquitin UbC	AB009010	13
Tripartite motif-containing 29 (*TRIM29*)	NM_012101	13
ASPIC (acidic secreted protein in cartilage)/*CRTAC1*	AJ276171	12
Collagen, type XVII, alpha 1 (*COL17A1*)	NM_000494	12
Apolipoprotein D	J02611	12
Transketolase (tk)	L12711	12
Paired box gene 6 (*PAX6*)	NM_001604	12
Ribosomal protein L4	D23660	12
Pyruvate kinase, muscle (*PKM2*)	NM_002654	11
Desmoplakin	J05211	11
HSP27	AB020027	11
Gelsolin	NM_000177	11
GAPDH	AY007133	11
Connexin 43 (*GJA1,* Cx43)	M65188	11
Glutamine synthase	X59834	10
Actin, gamma 1 (*ACTG1*),	NM_001614	10
Transmembrane protein BRI (*BRI*)	AF152462	10
Lactate dehydrogenase A (*LDHA*)	NM_005566	10
LOC283120 (*LOC283120*)	XM_208516	9
Hsp89-alpha-delta-N	AF028832	9
SERPINB5	NM_002639	9
Aspartate beta-hydroxylase (*ASPH*)	NM_004318	9
Calcium-activated chloride channel-2 (*bCLCA2*)	AF043977	9
Desmoglein 1 (*DSG1*)	AF097935	9
Ewings sarcoma *EWS*-Fli1 oncogene	AF327066	9
Non-integrin laminin-binding protein	M36682	9
Connexin 26 (*GJB2*) gene	AF479776	8
DEAD (Asp-Glu-Ala-Asp) box polypeptide 5 (*DDX5*)	NM_004396	8
MEN1 region clone epsilon/beta	AF001893	8
Syndecan-1 (*SDC-1*)	AJ551176	8
Ribosomal protein L3 (*RPL3*)	NM_000967	8
Ovary-specific acidic protein	AF329088	8

Genes represented by eight or more cDNAs are shown, along with a representative GenBank entry and the number of clones seen in the EST analysis.

elevated in KC compared with myopic cornea,[11] although again no clones for this gene were found in the cDNA collection. Such differences may be explained by cross-hybridization of related genes in the array experiment or insufficiently deep sequencing of the library to detect rare transcripts.

Among the most abundant cDNAs from KC cornea are those for cartilage acidic protein-1 (CRTAC1) also known as acidic secreted protein in cartilage-1 (ASPIC1) or CEP-68. This is a protein associated with the differentiated state of chondrocytes.[26] Six of the 12 clones are apparently full length and, of these, two make use of an alternative promoter, giving a variant 5′ end, and also include two novel alternatively spliced exons (Fig. 1). As a result, the predicted protein sequence resulting from these transcripts has a different, longer N-terminal domain containing several cysteine residues. Part of this alternative domain actually makes use of an alternative reading frame in a common part of what is the second exon of canonical CRTAC1. Because the function of the protein in chondrocytes is unclear and, because neither the known nor novel N-terminal domains have obvious similarity to other functional domains, it is not possible to predict what effects this might have. The *CRTAC1* gene is located on chromosome 10 at q24.2, which is not close to any mapped loci for KC, and CRTAC1 is also fairly abundant in the BodyMap collection (three clones) suggesting that it is common to both normal and KC cornea.

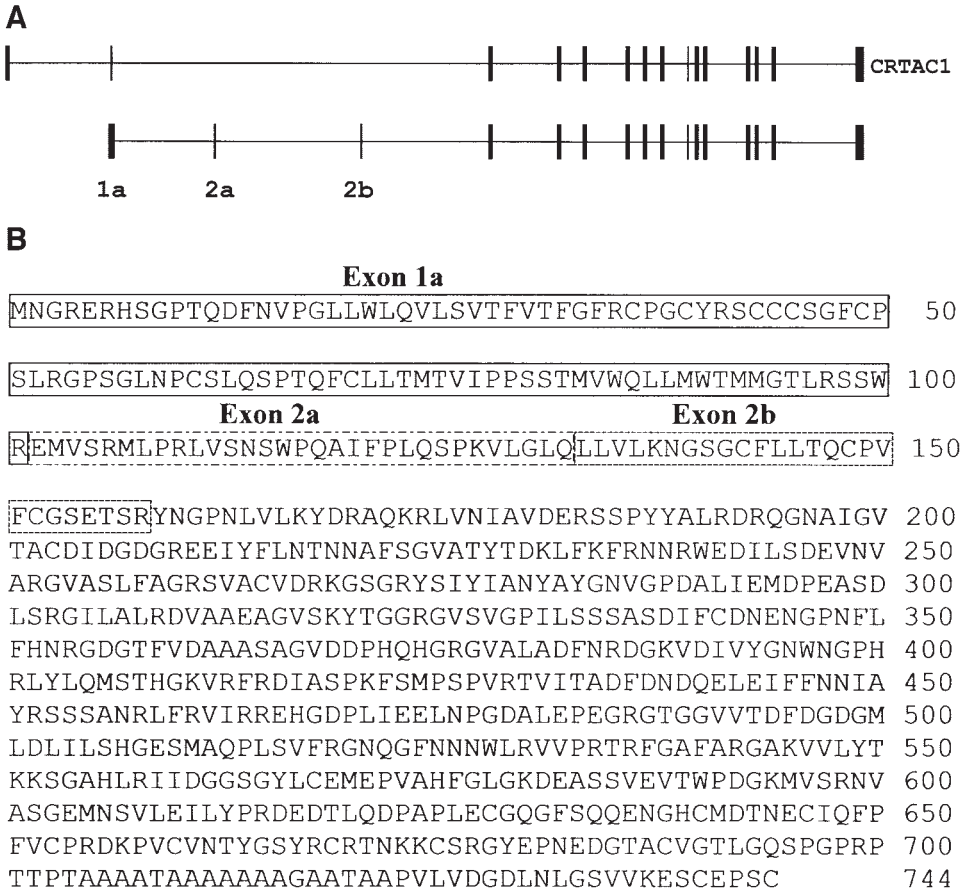

FIGURE 1. A novel splice variant of the *CRTAC1* gene from KC cornea. (**A**) Diagram of *CRTAC1* gene structure showing the novel alternative splice pattern revealed by clones from KC cornea. *Top line*: canonical *CRTAC1* gene with vertical blocks indicating exons. *Bottom line*: the variant expressed in KC cornea with the three variants of alternative exons numbered 1a, 2a, and 2b. (**B**) Predicted amino acid sequence of the variant CRTAC1 protein arising from the variant transcript. The variant N-terminal domain is *boxed* to show the regions that arise from the alternative exons 1a, 2a, and 2b. The remainder of the protein is the same as canonical CRTAC1.

Collagens

Collagen is a major structural component of the cornea, and several clinical reports suggest an association between keratoconus and connective tissue disorders that may involve defects in collagen.[10] Table 2 lists the 10 collagen genes represented in the present KC cornea analysis. Two that are abundant in this collection, collagen-17A1 (12 clones) and -12A1 (4 clones), and two represented by single clones—collagen-18A1 and -21A1—have not been described in the cornea. The BodyMap data for normal corneal epithelium contains an additional collagen, 5A3, that is absent from the KC data set. The KC data set also lacks clones for collagens-2, -7, and -8, which have been identified in normal cornea by immunochemical and peptide analysis[27] or for collagen-4A1, -11A1, or -16A1 that were identified by hybridization of cDNA library-derived probes to a commercial microarray.[28] Some of these differences could be due to insufficiently deep sequencing of the KC cornea cDNA library that would have detected some rare transcripts, to errors in identification in earlier analyses, or to actual differences in the collagen expression profile of KC. All these differences necessitate further study to determine their significance.

TABLE 2. Collagen Genes Identified in the KC Cornea cDNA Library

Gene Description	GenBank No.	Clones (*n*)
Collagen, type XVII, alpha 1	NM_000494	12
Collagen, type I, alpha 2	NM_000089	6
Collagen, type V, alpha 2	J04478	4
Collagen, type XII, alpha 1	NM_004370	4
Collagen, type VI, alpha 2	AY029208	3
Collagen, type IV, alpha 3	NM_000091	1
Collagen, type IV, alpha 6	NM_033641	1
Collagen, type VI, alpha 3	NM_004369	1
Collagen, type XXI, alpha 1	AF330693	1

Transcription-Related Genes

Supplementary Table S1, at www.iovs.org/cgi/content/full/46/4/1239/DC1, lists cDNAs from the KC collection that according to GO annotation[17] are connected with transcription. Among the most abundant cDNAs for proteins with a well-established role in gene expression is PAX6, one of the key determinants of eye development.[29] PAX6 is able to exert tissue-specific effects through alternative splicing of the gene transcript that gives rise to different versions of the DNA-binding domain in the protein. Indeed, the alternatively spliced exon 5a is used in half of the human KC cornea clones, which is consistent with previous observations which showed approximately equal levels of *PAX6* mRNA, with and without the alternative exon in bovine cornea.[30] Four of six apparently full-length PAX6 clones (of a total of 12 clones) from KC cornea come from the far upstream "A" promoter, which has previously been shown to be used primarily in cerebellum.[31] This may indicate some tissue preference in the expression mechanism for this important gene in cornea.

Several other transcription factors represented have specific roles in epithelial tissues. One of these, ELF3 (also known as ESX, ERT, EPR-1, or ESE-1), is represented by 6 cDNA clones. ELF3 is an *ets* domain transcription factor specific to epithelial cells and is associated with terminal differentiation in the epidermis.[32] A second member of the same superfamily, EHF (also known as ESE-3 or ESEJ), is represented by two clones, whereas the collection contains a third clone that probably

comes from a longer 3′ untranslated region (UTR) of the same gene. EHF is also associated with epithelial cells but has different target gene specificity from ELF3.[33] Three other *ets* domain factors—ETV4, ETV5, and ERF—are represented by single clones. ZFP67/hcKrox, a zinc finger protein that plays a role in expression of type 1 collagen genes and has an important role in epidermal differentiation,[34] is represented by three clones. The forkhead domain protein FOXC1, which is the locus for various glaucoma phenotypes including primary congenital glaucoma, autosomal dominant iridogoniodysgenesis anomaly, and Axenfeld-Rieger anomaly,[35] is represented by one clone.

VSX1, a transcription factor of the retina and the locus of a form of KC associated with retina anomalies,[9] is not represented in the KC collection, but this could certainly be due to insufficiently deep sequencing to detect rare transcripts. Among the sequence collections from other NEIBank human eye cDNA libraries, only one VSX1 clone is found and that is in retina.

Apoptosis

Apoptosis, or programmed cell death, appears to be elevated in KC compared with normal donor corneas or those with stromal dystrophies.[36] Supplementary Table S2, at www.iovs.org/cgi/content/full/46/4/1239/DC1, lists the 67 apoptosis-related gene transcripts identified in this study, some of which are present at high levels. Clusterin (46 clones) is highly expressed in cornea and other eye tissues and, among many other possible roles, has been implicated in protection from apoptosis, although its role in cornea and elsewhere is not clear.[23] The second most abundant protein represented by cDNAs in this list is PERP (also known under several aliases including TP53 apoptosis effector, THW and PIGPC1). PERP (23 clones) is a tetraspanin protein of the plasma membrane that is a target gene for p53 and is highly upregulated during p53-mediated apoptosis.[37] Two other related tetraspanins that are also associated with apoptosis, EMP-1 and EMP-2,[38] are also present in the KC data set.

Caspases are cysteine proteases that are key initiators or effectors of the apoptotic program.[39] In the KC cornea data set, there are three clones for the effector or "executioner" caspase, caspase-6, that exerts its effects by cleavage of lamin A and other substrates.[40,41] Of the three clones, two are essentially full-length and clearly come from the longer, α form of caspase-6. There are also single clones for another effector, caspase-3, and for caspase-4.

A Novel Gene: *KC6*

Among the abundantly expressed genes identified in the KC data set is a novel gene of unknown function. For convenience, this has been given the temporary designation *KC6*, reflecting its discovery as a group of six ESTs from the KC library. A more formal name awaits further characterization. The gene for *KC6* is on human chromosome 18 at q12.3, close to the gene for phosphoinositide-3-kinase, class 3 (*PIK3C3*). The only other ESTs in the current version of dbEST that appear to come from the same gene are from embryonic stem cells (accession numbers CD654078 and CN413612). The ESTs from the KC library define a gene with at least six exons, whereas the partial sequence of EST CN413612 suggests the possibility of an additional upstream exon. This gene exhibits alternative splicing and alternative 3′ ends (with alternative polyadenylation signals; Fig. 2) and has all the hallmarks of a polymerase II–dependent gene that encodes an mRNA, with the notable exception that there is no evidence of any significant open reading frame in the transcribed sequences. EST CN413612 from human ES cells, which may represent a partial transcript of the same gene, contains a 5′ *Alu*-like sequence that contains

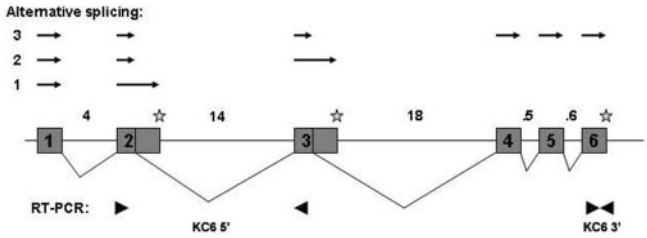

FIGURE 2. *KC6*, a novel gene from KC cornea. The six exons defined by cDNAs from KC cornea are shown approximately to scale, together with alternative splicing patterns that have been observed so far. The *arrows* above the gene figure illustrate the three alternative splice forms identified. *Arrowheads* below the gene diagram show the relative positions of primers used for RT-PCR. The sequence of alternative splice form 3, including sequence from all six exons has GenBank accession number AY762618.

a short ORF similar to some *Alu*-derived sequences in GenBank (not shown; http://www.ncbi.nlm.nih.gov/Genbank; provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD), but overall there is no clear evidence that this novel gene encodes a protein and it may instead belong to the largely mysterious class of noncoding RNAs.[42] The program Repeatmasker (used to delineate repetitive sequences in the genome; http://ftp.genome.washington.edu/cgi-bin/RepeatMasker/ provided in the public domain by the University of Washington Genome Center, Seattle, WA) detects short stretches of LINE-like sequence[43] in some parts of the *KC6* gene, suggesting a possible relationship with retroviral-like sequences. No clear mouse orthologue is apparent, but examination of the mouse genome reveals the presence of four ESTs, also from ES cells, that map to an equivalent region of the mouse genome (on mouse chromosome 18), close to the orthologous *PIK3C3* gene, and this region of the mouse genome also contains LINE-related sequences. This is under further investigation. The sequence of the splice variant encompassing all six exons observed in the KC data set has been submitted to GenBank (accession number: AY762618).

RT-PCR was used to verify the expression of *KC6* in KC cornea and to examine its tissue preference (Fig. 3). Two different primer sets, one for the sixth exon defined by the identified *KC6* clones and one that crosses the intron between the first and second exons defined by the *KC6* ESTs, were used. These were tested on freshly obtained samples of epithelium from KC and non-KC myopic human cornea. Primers for the epithelial transcription factor ELF3/ESX were used as positive controls. While ELF3/ESX transcripts were detected at similar levels in both cornea samples and in retina, brain, kidney, and heart, *KC6* transcripts were detected in both KC and "normal" cornea but not in any other tissue tested, suggesting that *KC6* is cornea preferred or even cornea specific.

Aquaporin 5

In addition to the unexpected presence of *KC6*, the KC EST analysis revealed an unexpected absence. Aquaporin 5 (AQP5) is a member of a large family of integral membrane proteins whose principal function is to serve as water channels.[44] AQP5 is expressed in salivary and lacrimal glands and in corneal epithelium, while the related AQP1 is expressed in corneal endothelium.[45] Water is a major component of the corneal stroma and cell layers and indeed, deletion of AQP5 in genetically modified mice causes corneal swelling.[46] AQP5 cDNA is present (two clones) in the BodyMap corneal epithelium collection[5] and is also prominent in SAGE of mouse cornea.[8] The human KC cornea cDNA collection contains seven clones for

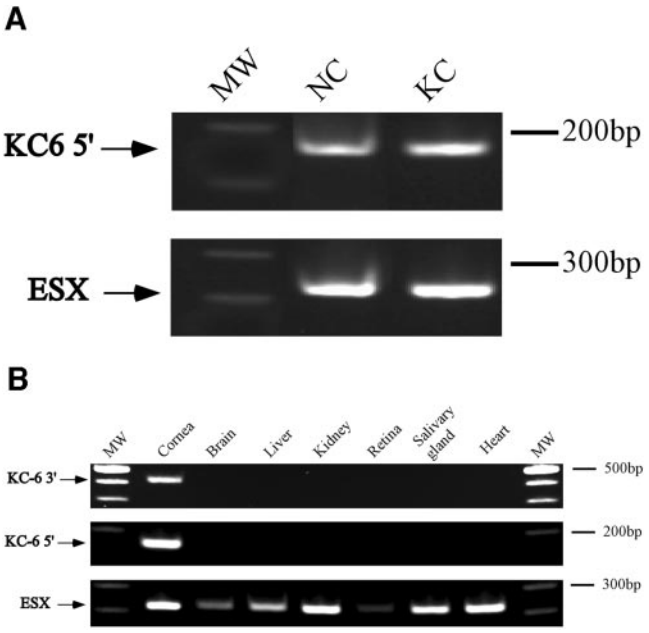

FIGURE 3. KC6 expression is cornea preferred. (**A**) RT-PCR detection of KC6 expression in epithelial samples of non-KC cornea (NC) and KC cornea. *Top*: results for intron-spanning *KC6* specific primers from the 5′ region of the gene, as shown in Fig. 2. *Bottom*: results for ELF3/ESX (ESX). (**B**) Expression of KC6 and ELF3/ESX (ESX) in human cornea (non-KC), retina, brain, kidney, and heart. *Top panels*: results for two different primer sets located near the 3′ and 5′ ends of the *KC6* gene as shown in Figure 2. *Bottom panel*: results for ESX as in part (**A**).

the endothelial marker AQP1 and one clone for AQP3, which is known to be expressed in conjunctiva.[45] However, it contains no clones for AQP5, even though many other epithelial markers are present. This striking absence was confirmed by semiquantitative RT-PCT analysis of samples from KC and non-KC (posterior polymorphous dystrophy with normal epithelium) cornea, again using the transcription factor ELF3/ESX as a moderately abundant marker for epithelium (Fig. 4). ESX levels were indistinguishable in both KC and non-KC cornea, but although AQP5 was strongly amplified from non-KC cornea, it was barely detectable in KC. The PCR bands obtained were cloned and sequenced for conformation of their identity. Indeed, sequence of the open reading frame of AQP5 amplified from KC cornea showed no mutations and normal splicing. It seems likely that the gene for AQP5 is normal in KC but that expression of the gene is suppressed in KC, at least at the stage of disease progression at which transplantation is indicated. At this point, it is unknown whether suppression of AQP5 expression is a major contributor to the phenotype of KC or whether it is one of a number of late-stage downstream effects. It is possible that the genetic heterogeneity of KC reflects multiple genes in a common pathway, one of whose targets is AQP5. More detailed analysis of KC by microarray is underway to search for other affected genes.

To exert a phenotype, reduced gene expression should be associated with reduced levels of protein. Attempts were therefore made to examine protein for AQP5 in normal and KC cornea. However, from the samples at our disposal, no reactive bands for AQP5 were observed in Western blot of either normal or KC, although as a control, another epithelial membrane protein, connexin 50, could be detected (not shown). Several different commercial antisera were tested with the same result. All these antisera are targeted to the same C-terminal peptide of AQP5. It is known that AQP0 (the lens protein MIP) loses its C-terminal peptide with age,[47] and it is possible the same thing has happened to AQP5 in the adult human corneal samples we have tested. New antibodies are currently being developed to target other regions of the AQP5 protein to investigate this further.

CONCLUSIONS

The KC cornea cDNA library is an excellent source of clones for genes expressed in human cornea and greatly expands the representation of such genes in the databases. Further sequencing of the same library would certainly reveal even more of the transcriptional repertoire of this tissue. However, the analysis to date has already identified approximately 4000 cornea-expressed genes and provides new candidates for genes whose expression may be affected in KC.

Most notable in this category is the virtual absence of AQP5 transcripts in the KC corneas. The library was made from seven pooled KC corneas and RT-PCR of an additional KC samples confirms that mRNA for this important water channel protein is barely detectable in the diseased cornea. A water channel defect would be a plausible contributor to the stromal thinning characteristic of KC and although *AQP5* is located on human chromosome 12 at q13, which has not been identified as a locus for KC, it is possible that the identified disease loci correspond to genes in pathways that converge on AQP5. However, deletion of *AQP5* in mice produces corneal swelling,[46] rather than stromal thinning as in human KC patients. Although this could reflect a species difference in expression and behavior of aquaporins between human and rodents eye,[48] it is also likely that other genes are involved in the KC phenotype, which is more complex that a simple *AQP5*-null.

The abundance in the KC library of cDNAs for genes involved in apoptosis is striking and is at least consistent with a high level of programmed cell death associated with the disease. However several of the abundant apoptosis transcripts, such as clusterin and PERP are also abundant in the smaller BodyMap collection of cDNAs (72 and 5 clones respectively) and it remains to be seen whether there is any disease association in any of these expressed genes. Patterns of collagen and keratin expression in KC show some similarities and some differences with those described in the literature for normal cornea while the roles of several proteins such as the desmogleins and CRTAC1 in the cornea are also of potential interest. Clearly, a much larger survey of expressed genes in normal cornea would be valuable for further comparisons, however the EST analysis described herein provides ample candidates for future study.

Last, the expression of *KC6* reveals an unexpected new marker for cornea. So far, this mysterious gene seems to have a preference for expression in cornea but is also expressed in

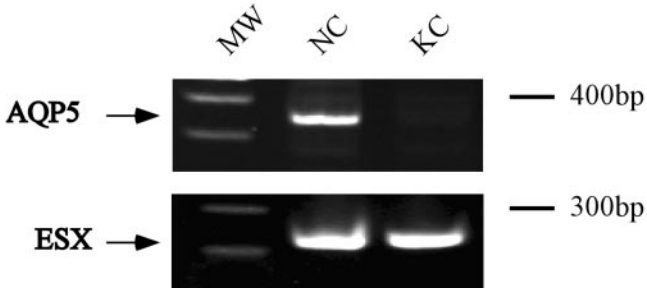

FIGURE 4. AQP5 expression is suppressed in KC cornea. RT-PCR of AQP5 and ELF3/ESX (ESX) for individual epithelial samples of KC and non-KC (NC) human cornea.

embryonic stem cells. The corneal epithelium is known to have populations of stem cells that respond to corneal wounding and to the normal loss of epithelial cells by differentiation and replacement of the lost cells. No molecular markers for these stem cells have yet been identified.[49] They are thought be principally localized to the corneal limbus, which is missing from the KC button samples. If *KC6* does have an association with stem cell populations in the cornea, it suggests that some of these cells are more widely distributed. Efforts to fully define the expression and role of *KC6* are in progress.

Acknowledgments

The authors thank James Gao and Patee Buchoff for their work on bioinformatics for NEIBank.

References

1. Klintworth GK. The molecular genetics of the corneal dystrophies: current status. *Front Biosci.* 2003;8:687–713.
2. Jun AS, Larkin DF. Prospects for gene therapy in corneal disease. *Eye.* 2003;17:906–911.
3. Wistow G. A project for ocular bioinformatics: NEIBank. *Mol Vis.* 2002;8:161–163.
4. Tomarev SI, Wistow G, Raymond V, Dubois S, Malyukova I. Gene expression profile of the human trabecular meshwork: NEIBank sequence tag analysis. *Invest Ophthalmol Vis Sci.* 2003;44:2588–2596.
5. Nishida K, Adachi W, Shimizu-Matsumoto A, et al. A gene expression profile of human corneal epithelium and the isolation of human keratin-12 cDNA. *Invest Ophthalmol Vis Sci.* 1996;37:1800–1809.
6. Sakai R, Kinouchi T, Kawamoto S, et al. Construction of human corneal endothelial cDNA library and identification of novel active genes. *Invest Ophthalmol Vis Sci.* 2002;43:1749–1756.
7. Gottsch JD, Seitzman GD, Margulies EH, et al. Gene expression in donor corneal endothelium. *Arch Ophthalmol.* 2003;121:252–258.
8. Norman B, Davis J, Piatigorsky J. Postnatal gene expression in the normal mouse cornea by SAGE. *Invest Ophthalmol Vis Sci.* 2004;45:429–440.
9. Heon E, Greenberg A, Kopp KK, et al. VSX1: a gene for posterior polymorphous dystrophy and keratoconus. *Hum Mol Genet.* 2002;11:1029–1036.
10. Rabinowitz YS. The genetics of keratoconus. *Ophthalmol Clin North Am.* 2003;16:607–620.
11. Nielsen K, Birkenkamp-Demtroder K, Ehlers N, Orntoft TF. Identification of differentially expressed genes in keratoconus epithelium analyzed on microarrays. *Invest Ophthalmol Vis Sci.* 2003;44:2466–2476.
12. Brancati F, Valente EM, Sarkozy A, et al. A locus for autosomal dominant keratoconus maps to human chromosome 3p14–q13. *J Med Genet.* 2004;41:188–192.
13. Tang YM, Taylor K, Li X, Picornell Y, Rabinowitz YS, Yang H. A two-stage genome-wide scan identifies a new susceptibility locus for autosomal dominant keratoconus in a large Caucasian pedigree. *Am J Hum Genet*. 2003;73(suppl):1822.
14. Hughes AE, Dash DP, Jackson AJ, Frazer DG, Silvestri G. Familial keratoconus with cataract: linkage to the long arm of chromosome 15 and exclusion of candidate genes. *Invest Ophthalmol Vis Sci.* 2003;44:5063–5066.
15. Wistow G, Bernstein SL, Wyatt MK, et al. Expressed sequence tag analysis of adult human lens for the NEIBank Project: over 2000 non-redundant transcripts, novel genes and splice variants. *Mol Vis.* 2002;8:171–184.
16. Wistow G, Bernstein SL, Touchman JW, et al. Grouping and identification of sequence tags (GRIST): bioinformatics tools for the NEIBank database. *Mol Vis.* 2002;8:164–170.
17. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25:25–29.
18. Kim HS, Yoon SK, Cho BJ, Kim EK, Joo CK. *BIGH3* gene mutations and rapid detection in Korean patients with corneal dystrophy. *Cornea.* 2001;20:844–849.
19. Li W, Vergnes JP, Cornuet PK, Hassell JR. cDNA clone to chick corneal chondroitin/dermatan sulfate proteoglycan reveals identity to decorin. *Arch Biochem Biophys.* 1992;296:190–197.
20. Abedinia M, Pain T, Algar EM, Holmes RS. Bovine corneal aldehyde dehydrogenase: the major soluble corneal protein with a possible dual protective role for the eye. *Exp Eye Res.* 1990;51:419–426.
21. Cooper DL, Baptist EW, Enghild JJ, Isola NR, Klintworth GK. Bovine corneal protein 54K (BCP54) is a homologue of the tumor-associated (class 3) rat aldehydrogenase (RATALD) gene. *Gene.* 1991;98:201–207.
22. Zieske JD, Bukusoglu G, Yankauckas MA, Wasson ME, Kettmann HT. α-Enolase is restricted to basal cells of stratified squamous epithelium. *Dev Biol.* 1992;151:18–26.
23. Kinoshita S, Adachi W, Sotozono C, et al. Characteristics of the human ocular surface epithelium. *Prog Retin Eye Res.* 2001;20:639–673.
24. Liu CY, Shiraishi A, Kao CW, et al. The cloning of mouse keratocan cDNA and genomic DNA and the characterization of its expression during eye development. *J Biol Chem.* 1998;273:22584–22588.
25. Wu H, Stanley JR, Cotsarelis G. Desmoglein isotype expression in the hair follicle and its cysts correlates with type of keratinization and degree of differentiation. *J Invest Dermatol.* 2003;120:1052–1057.
26. Benz K, Breit S, Lukoschek M, Mau H, Richter W. Molecular analysis of expansion, differentiation, and growth factor treatment of human chondrocytes identifies differentiation markers and growth-related genes. *Biochem Biophys Res Commun.* 2002;293:284–292.
27. Zimmermann DR, Fischer RW, Winterhalter KH, Witmer R, Vaughan I. Comparative studies of collagens in normal and keratoconus corneas. *Exp Eye Res.* 1988;46:431–442.
28. 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.
29. Wawersik S, Purcell P, Maas RL. Pax6 and the genetic control of early eye development. *Results Probl Cell Differ.* 2000;31:15–36.
30. Jaworski C, Sperbeck S, Graham C, Wistow G. Alternative splicing of Pax6 in bovine eye and evolutionary conservation of intron sequences. *Biochem Biophys Res Commun.* 1997;240:196–202.
31. Okladnova O, Syagailo YV, Mossner R, Riederer P, Lesch KP. Regulation of PAX-6 gene transcription: alternate promoter usage in human brain. *Brain Res Mol Brain Res.* 1998;60:177–192.
32. Oettgen P, Alani RM, Barcinski MA, et al. Isolation and characterization of a novel epithelium-specific transcription factor, ESE-1, a member of the ets family. *Mol Cell Biol.* 1997;17:4419–4433.
33. Kas K, Finger E, Grall F, et al. ESE-3, a novel member of an epithelium-specific ets transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J Biol Chem.* 2000;275:2986–2998.
34. Widom RL, Culic I, Lee JY, Korn JH. Cloning and characterization of hcKrox, a transcriptional regulator of extracellular matrix gene expression. *Gene.* 1997;198:407–420.
35. Nishimura DY, Swiderski RE, Alward WL, et al. The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25. *Nat Genet.* 1998;19:140–147.
36. Kim WJ, Rabinowitz YS, Meisler DM, Wilson SE. Keratocyte apoptosis associated with keratoconus. *Exp Eye Res.* 1999. 69:475–481.
37. Ihrie RA, Attardi LD. Perp-etrating p53-dependent apoptosis. *Cell Cycle.* 2004;3:267–269.
38. Wilson HL, Wilson SA, Surprenant A, North RA. Epithelial membrane proteins induce membrane blebbing and interact with the P2X7 receptor C terminus. *J Biol Chem.* 2002;277:34017–34023.
39. Thornberry NA, Lazebnik Y. Caspases: enemies within. *Science.* 1998;281:1312–1316.

40. Alnemri ES. Mammalian cell death proteases: a family of highly conserved aspartate specific cysteine proteases. *J Cell Biochem*. 1997;64:33–42.
41. Slee EA, Adrain C, Martin SJ. Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis. *J Biol Chem.* 2001;276:7320–326.
42. Morey C, Avner P. Employment opportunities for non-coding RNAs. *FEBS Lett.* 2004;567:27–34.
43. Ostertag EM, Kazazian HH Jr. Biology of mammalian L1 retrotransposons. *Annu Rev Genet.* 2001;35:501–538.
44. Agre P, Kozono D. Aquaporin water channels: molecular mechanisms for human diseases. *FEBS Lett.* 2003;555:72–78.
45. Verkman AS. Role of aquaporin water channels in eye function. *Exp Eye Res.* 2003;76:137–143.
46. Thiagarajah JR, Verkman AS. Aquaporin deletion in mice reduces corneal water permeability and delays restoration of transparency after swelling. *J Biol Chem.* 2002;277:19139–19144.
47. Alcala J, Maisel H. Biochemistry of lens plasma membranes and cytoskeleton. In: Maisel H, ed. *The Ocular Lens*. New York: Marcel Dekker, Inc.; 1985:169–222.
48. King LS, Agre P. Man is not a rodent: aquaporins in the airways. *Am J Respir Cell Mol Biol.* 2001;24:221–223.
49. Boulton M, Albon J. Stem cells in the eye. *Int J Biochem Cell Biol.* 2004;36:643–657.