Comparative Transcriptome and Network Biology Analyses Demonstrate Antiproliferative and Hyperapoptotic Phenotypes in Human Keratoconus Corneas

Matthias Macé, Stéphane D. Gally, Angélique Erraud, José Enrique Mejía, Heather Etchevers, Michèle Allouche, Laurence Desjardins, Patrick Calvas, and François Malecaze

Purpose. To decipher the biological pathways involved in keratoconus pathophysiology by determining the patterns of differential gene expression between keratoconus and control corneas.

Methods. RNA was extracted from surgically removed corneas of 10 keratoconus patients and from normal corneas of 10 control patients who had undergone enucleation of an eye for ocular melanoma. Several hundred thousand RNA transcripts were assessed using exon microarrays. Statistical comparison and identification of differentially regulated and differentially spliced RNA transcripts was performed by comparing keratoconus cases and controls. In addition, relevant biological pathways were identified by information extraction using network biology.

Results. Eighty-seven genes showed significant differences in expression levels. Among these, 69 were downregulated in keratoconus patients, particularly partners of the transcription factor AP-1. The 18 overexpressed genes include mucins, keratins, and genes involved in fibroblast proliferation. In addition, 36 genes were shown to be differentially spliced, including 9 among those that were differentially expressed. Network biology and analysis using Gene Ontology descriptors suggest that many members of both groups belong to pathways of apoptosis and regulation of the balance between cellular differentiation and proliferation.

Conclusions. This work constitutes the first genome-wide transcriptome analysis of keratoconus patient corneas that include all currently known genes and exons. Differential expression suggests that mechanisms of cell loss resulting from antiproliferative and hyperapoptotic phenotypes may be responsible for the pathogenesis of keratoconus. Array information, experimental design, raw intensities, and processed log₂ ratios were deposited at the European Bioinformatic Institute’s ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/). The accession number is E-MEXP-2777. (Invest Ophthalmol Vis Sci. 2011;52:6181–6191) DOI:10.1167/iovs.10-70981

Keratoconus (KC) is a vision-threatening condition characterized by thinning and deformation of the cornea. It is one of the most common indications for corneal grafting in industrialized countries. The disease prevalence is approximately 1 in 2000, and familial aggregation, together with increased familial risk, suggests important genetic influences on its pathogenesis, but the etiology of KC is still poorly understood. Aside from genetic determinants, environmental stresses such as eye rubbing or atopy have been suggested as possible causes or aggravating factors in KC. To date, several loci for familial KC have been described. These have been mapped by genome-wide scans of varying resolutions to the chromosomal regions 2p24, 3p14-q13, 5q14.3-q21.1, and 16q22.2-q23.1, in familial studies, and to the chromosomal regions 4q31, 5q31, 9q34, 12p12, 14p11, 17q24, and 20q12, in affected-only linkage analyses. However, no genes have yet been identified as responsible for the development of most KC cases. Within these chromosomal regions, several candidate genes (COL6A1, SOD1, MMP9, MMP2, and COL8A1) have been excluded whereas mutations in VSX1 have been reported in a few KC patients. In addition, one differential expression study pointed to AQP5 encoding aquaporin 5. Taken together, these data suggest that KC is a complex disease involving multiple susceptibility loci.

Holistic approaches integrating various “omics” techniques are promising for discovering the molecular bases of genetic disorders with complex patterns of inheritance, such as KC. Variation in gene expression is an important mechanism underlying the susceptibility to complex diseases, and it has been established that the steady state abundance of mRNA transcripts for many human genes is a highly heritable, quantitative phenotypic trait. Recent technical advances for the profiling of virtually all human or mouse exons on genome-wide transcriptome microarrays have made the comprehension of molecular disease mechanisms possible. Thus far, most published analyses have assessed only the variations in levels of gene expression, without taking into account exon content. However, approximately 50% of all human genes are predicted to be alternatively spliced, and disruptions in the balance of multiple transcript isoforms have been shown to be at play in conditions ranging from Alzheimer’s disease to several types of...
cancer. In this study, we therefore sought to analyze the KC corneal transcriptome in a comparison with normal corneas, to identify differentially expressed or spliced genes that would highlight the pathophysiological pathways involved in KC. Network biology allowed us to assign some of these gene products to a small number of molecular cascades pertaining to proliferation, differentiation, and programmed cell death.

**METHODS**

**Ethics Statement**

This study was performed in accordance with French regulations and the Declaration of Helsinki. The samples were assigned a laboratory number and have remained anonymous throughout. The experimental protocol was approved by the relevant interregional ethics committee (CPP Sud-Ouest Outre-Mer No. 2). Written consent was obtained from all participants.

**Patients and Control Tissue Collection and Storage**

Ten corneas were collected from nonrelated patients during a penetrating keratoplasty procedure for advanced KC in the Centre National de Référence du Kératocône (Toulouse, France). All KC eyes included in the study were stage III (Amsler-Krumeich classification), with high refractive errors, severe loss of visual acuity, and absence of scarring. In addition all patients had contact lens intolerance, and none of them had worn contact lenses since at least 6 months. Patients were 34 ± 10.5 years old; six were men and four were women. Ten control corneas were obtained in the Institut Curie (Paris, France) from unrelated patients who underwent enucleation for choroidal melanomas strictly localized to the posterior pole of the eye. These patients did not receive antineoplastic treatment before surgery. None was wearing contact lenses. The patients were 59 ± 10.1 years old; five were men and five were women. In addition, for each control cornea, anatomic or topographic abnormality was ruled out by careful examination. Excised corneal buttons were 8 mm in diameter. They were processed according to identical standard procedures in the two centers and were immediately stored in 1.5-mL microtubes in liquid nitrogen until RNA extraction.

**RNA Isolation from Whole Corneas**

Corneas were transferred from liquid nitrogen to 500 μL butter (RLT; RNAeasy Mini kit; Qiagen, Hilden, Germany)/β-mercaptoethanol (GE Healthcare, Pittsburgh, PA) solution, in a 2-mL microtube (Lysing Matrix D; MP Biomedicals, Irvine, CA) and kept on ice. This mix was then subjected to eight cycles of 20 seconds of shaking followed by 5 minutes of cooling at 4°C (FastPrep-24 System; MP Biomedicals). The supernatant was then retrieved, and total RNA was extracted and further purified (RNAeasy Mini kit and RNA-Free DNase Set; Qiagen), according to the manufacturer’s protocols. RNA quality was assessed on an RNA microarray (RNA 6000 Nano Chips; Agilent Technologies, Santa Clara, CA) on a bioanalyzer (model 2100; Agilent). RNA samples were immediately stored frozen at −80°C. RNAs with an RNA integrity number (RIN) ≥9 were selected for further analysis (i.e., above the 8.5 threshold recommended by Agilent). RNA concentration and purity were determined immediately before reverse transcription through measurement of A260/A280 ratios with a spectrophotometer (NanoDrop ND-1000; NanoDrop Technologies, Wilmington, DE).

**Microarray Transcriptome Hybridization**

We used an exon array platform (Human GeneChip Exon 1.0 ST Array; Affymetrix, Santa Clara, CA) designed to interrogate expression levels at both exon and gene levels, querying 1.4 million probesets. To minimize background and increase the sensitivity of the assay, ribosomal RNA was removed from 0.7 to 1 μg of total RNA from each preparation (RiboMinus Human/Mouse Transcriptome Isolation Kit; Invitrogen, Carlsbad, CA). The resulting RNA was then subjected to reverse transcription using random hexamers tagged with a T7 promoter sequence followed by second-strand cDNA synthesis using a DNA polymerase (GeneChip WT cDNA Synthesis Kit; Affymetrix). The resulting double-stranded cDNA was then used for amplification of antisense cRNA and cleaned (Gene Chip Sample Cleanup Module; Affymetrix). A second cycle of cDNA synthesis was then performed using random primers to reverse transcribe the cRNA into sense single-stranded DNA. This DNA was then fragmented, labeled, and hybridized to a human gene chip (Human Gene Chip Exon 1.0 ST Arrays; Affymetrix). Target labeling, array hybridization, washing, and staining were performed as described by the manufacturer (GeneChip Whole Transcript [WT] Sense Target Labeling; Affymetrix). Arrays were then hybridized, washed, and stained (GeneChip Hybridization, Wash and Stain Kit in a GeneChip Hybridization Oven 645 and a GeneChip Fluidics Station 450; Affymetrix). Arrays were then scanned (GeneChip Scanner 5000 7G; Affymetrix).

Array information, experimental design, raw intensities and processed log2 ratios are all available through ArrayExpress (http://www.ebi.ac.uk/arrayexpress/).

**In Silico Data Filtering for Probes, Signal Normalization, and Summarization**

Standard methods for outlier removal were used, including principal component analysis on the sample covariance matrix \((V = \sum \limits_{i=1}^{n} (y_i - \bar{y})(y_i - \bar{y})/n\) where \(\bar{y} = \sum \limits_{i=1}^{n} y_i/n\) and hierarchical clustering by defining a dissimilarity between the expression signatures between two samples as \(1 - |\rho|\), \(\rho\) being the sample correlation\(^22\) as implemented in the oneChannelGUI package\(^23\) of the R programming language. A hierarchical clustering tree is thus a dendrogram representing the pairwise similarity structure between arrays. A cutoff value is given as a proportion of the tree depth to remove too divergent arrays as potential artifacts.

After quantile normalization,\(^24\) data from the 20 probes were fitted to a global model of expression and probe affinities (model = chip effect + probe affinity + ϵ for the jth probe on the ith array). Expression levels were summarized using the robust multiplex average (RMA) model,\(^25\) which is based on the dependence of the measured intensity on the amount of material (chip effect), the probe affinity, and a reading error (measurement error). Model fitting was performed using the fast median polish algorithm.\(^26\) This expression summary provides an estimated value for the abundance of the transcript in the sample but does not provide a measurement of the reliability of this estimation. Exon arrays lack paired mismatch probes but instead comprise a separate pool of ~25,000 background probes that allows the computation of a detection above background (DABG) score by matching those probes to members of the background with the same GC content. It is used to discard poorly performing probesets. In addition to this probe-level summary, gene-level summaries (expression levels averaged across probes) are calculated as median expression of probes that (1) are not multiply targeted and (2) hit an exon with all probes in the probeset.

**Mapping to Annotation**

Probesets were mapped to the genome and to gene annotations using various X to Y functions included in the Exonmap R package\(^27\) and a local MySQL instance of the X:MAP database.\(^28\) This allowed discrimination between probesets hitting introns, transcripts, or genes as well as interrogation of known alternative splicings.

**Identification of Differentially Expressed Genes**

Linear modeling and tests for differential expression, adjusted for multiple testing, were performed with the Limma R package.\(^29\) The use of linear modeling allows the borrowing of information from all the transcripts to assist inference about each transcript individually. Briefly, Limma first adjusts a linear model over the systematic part of
the data, intending to estimate its variability (Imfit function). Then, a contrast step allows the fitted coefficients to be compared regardless of their number. For single-channel microarrays (e.g., Affymetrix exon), linear modeling is equivalent to ANOVA but with the fitting of a model for every gene. The fitted matrix and a contrast matrix are used to compute fold-changes (FCs) and t-statistics (makeContrasts function in Limma R). Change in expression (FC) was considered biologically relevant when the variation was twofold or greater (i.e., \( \log_2(FC) > 1 \)). Differential expression is then assessed by empirical Bayes statistics: the eBayes function is used to compute moderated t-statistics, moderated F-statistic, and log-odds of differential expression by empirical Bayes shrinkage of the standard errors toward a common value. Moderated means that the residual mean squares and degrees of freedom are moderated between probes. An adjustment for multiple testing was applied using the decideTests function. The Benjamini-Hochberg method for controlling the false discovery rate (FDR) was used.

The result was a list of significant probes within genes of interest. Last, we assessed whether for each of these genes all their probesets showed the same FC, thus reflecting gene-level differential expression. For this purpose, we computed for each gene the variance of the FC probesets are first converted to log2 space. For each probeset examination, we calculated the corrected log2 expression difference. This difference was considered biologically relevant when the variation was twofold or greater (i.e., \( \log_2(FC) > 1 \)). Differential expression is then assessed by empirical Bayes statistics: the eBayes function is used to compute moderated t-statistics, moderated F-statistic, and log-odds of differential expression by empirical Bayes shrinkage of the standard errors toward a common value. Moderated means that the residual mean squares and degrees of freedom are moderated between probes. An adjustment for multiple testing was applied using the decideTests function. The Benjamini-Hochberg method for controlling the false discovery rate (FDR) was used.

The result was a list of significant probes within genes of interest. Last, we assessed whether for each of these genes all their probesets showed the same FC, thus reflecting gene-level differential expression. For this purpose, we computed for each gene the variance of the FC probesets are first converted to log2 space. For each probeset examined, the expression value is subtracted from the mean expression value of all constitutive aligning probesets to create a constitutive corrected log2 expression difference. This difference was calculated for each individual in the study, using microarray data from that subject only, and was then used to calculate the mean expression difference in each of the patient and control groups. The probeset SI value was then derived by subtracting the KC group mean from the control sample group mean. This value represents the change in exon inclusion (dI).

A t-test (two tailed, assuming unequal variance) of the means was performed for statistical significance. A df of \(-1\) indicated a twofold change in the expression of a probeset relative to the mean constitutive expression, with expression being higher in the KC patient group than in the control group.

The MidAS approach is a two-way ANOVA-based method measuring differences between the exon level and aggregating gene level signals, including an error term and possible interactions. First, probe logarithmic intensity error (PLIER) normalization is performed. This generates both exon-level signals and gene-level estimates which are robust against exon-level anomalous signals across samples. Under the null hypothesis of no alternative splicing at an exon level, the expectation is to observe a constant difference between the exon and the corresponding gene across all the samples.

**Identification of Differentially Spliced Variants**

Detection of alternative splicing events was performed using two parallel approaches, the Splicing Index and MidAS (microarray detection of alternative splicing). For both tests, given the much lower number of tests involved (equaling the number of probes per gene) compared with gene-level differential expression, \( P < 0.05 \) was applied.

In the Splicing Index (SI) method, the expression values of the probesets are first converted to log2 space. For each probeset examined, the expression value is subtracted from the mean expression value of all constitutive aligning probesets to create a constitutive corrected log2 expression difference. This difference was calculated for each individual in the study, using microarray data from that subject only, and was then used to calculate the mean expression difference in each of the patient and control groups. The probeset SI value was then derived by subtracting the KC group mean from the control sample group mean. This value represents the change in exon inclusion (dI). A t-test (two tailed, assuming unequal variance) of the means was performed for statistical significance. A df of \(-1\) indicated a twofold change in the expression of a probeset relative to the mean constitutive expression, with expression being higher in the KC patient group than in the control group.

The MidAS approach is a two-way ANOVA-based method measuring differences between the exon level and aggregating gene level signals, including an error term and possible interactions. First, probe logarithmic intensity error (PLIER) normalization is performed. This generates both exon-level signals and gene-level estimates which are robust against exon-level anomalous signals across samples. Under the null hypothesis of no alternative splicing at an exon level, the expectation is to observe a constant difference between the exon and the corresponding gene across all the samples.

**Functional Association and Network Biology Analyses**

The Cytoscape platform was used as a network visualization and analysis tool for differentially expressed and spliced transcripts. We specifically used several plugins for interaction retrieval, network statistics, and Gene Ontology (GO) enrichment (below).

**Annotation and Interactome Characterization**

To harvest protein–protein physical and functional interactions, we used the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database, which aggregates most of the available information on protein–protein interactions, organizing them by scoring and weighting. The database is queried for interaction matrices derived from data and text mining, including experimental data and predicted interactions. STRING querying is characterized by a unique scoring framework based on benchmarks of the different types of associations against a common reference set. This results into a single confidence score per prediction. For each type of interaction, a different algorithm is used to provide a score (listed in Ref. 34). To recover the proteins belonging to given metabolic maps and so presumed to interact in the same metabolic pathways, we also interrogated the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Networks were then represented by a cloud of nodes connected by edges. Node connectivity (number of edges of the node) was assessed using the Hubba plugin for Cytoscape.

**Identification of Molecular Pathways**

We used the BINGO plug-in from Cytoscape to identify enriched GO terms among the differentially expressed genes, as follows: When sampling \( K \) genes (set) out of \( R \) genes (an annotation reference set), we aimed to infer the probability that \( K \) or more of these genes belonged to a functional category shared by \( r \) of the \( R \) genes in the reference set. By performing a binomial test, sampling with replacement, we were able to provide an approximate \( P \) value. Only well-characterized genes (excluding hypothetical proteins) were included in the analysis. GO terms that were overrepresented in the analysis set were selected (five or more hits, binomial test \( P < 0.05 \) for the Benjamini-Hochberg [BH] correction). The enrichment score of each term cluster was computed as the geometric mean of each member’s \(-\log(P)\).

**Reverse Transcription and Real-Time Polymerase Chain Reaction Amplification**

In addition to the microarray experiment, quantitation of selected gene transcripts was performed by reverse transcription and real-time quantitative PCR amplification (RT-qPCR). Eight KC and 10 control corneas (of which five KC and five control corneas were shared with the microarray hybridizations) were used in this experiment. One microgram of each of the total RNA preparations was reverse transcribed into single-stranded cDNA using reverse transcriptase from a kit (SuperScript III; SuperScript VILO kit; Invitrogen). The following genes were selected to cover a range of FCs and gene locations within the interaction network studied: JUN, FOS, FOSB, BTF2, EGR1, MCL1, HSP90AA1, S100A6, MAT2A, ANO1, and KRT78. Three further genes, TBP, FBRS, and PHH1D1, were selected as standard baseline genes because they displayed some of the lowest differences between cases and controls in the microarray experiment. Relevant gene-specific PCR primers are listed in Supplementary Table S8 (http://www.iovs.orglookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental). qPCR was performed (DNA Master SYBR Green I reagents in a LightCycler 480; Hoffmann-La Roche, Basel, Switzerland). Thermal cycling conditions were as follows: 95°C for 5 minutes (denaturation); 40 cycles at 95°C for 15 seconds, 60°C for 10 seconds (amplification), and 72°C for 20 seconds; and the melting curve: 95°C for 10 seconds, 70°C for 20 seconds; and 97°C for 0.1 second qPCR data were analyzed using the comparative CT method using geometric averaging of the three internal control genes.

To assess the correlation between the FC levels from qPCR and from the microarrays (at both gene and probe levels), we computed Spearman’s rank correlation coefficient (Rho, \( \rho \)) statistics that estimate a rank-based measure of association.

**Results**

**RNA Extraction and Data Quality Control**

On average, 2.3 \( \mu \)g of total RNA was recovered from each of 20 individual whole corneas: 10 KC corneas and 10 control
One hundred sixteen transcripts, represented by 794 probes, displayed significant differential expression levels at one or more probesets, with absolute $\log_2(\text{FC}) > 1$ and FDR-adjusted $P < 0.05$. Among these transcripts, 87 displayed a probeset FC variance low enough (<1) to be considered differentially expressed at the gene level. The most significant differential annotations are listed in Table 1, together with their FCs and levels of significance. Of note, most of the differently expressed transcripts meeting these statistical criteria were downregulated in KC (69 downregulated versus 18 upregulated genes). Interestingly, although considered to be a housekeeping gene, ACTB belongs to this list. A subset of downregulated genes—FOS, JUN, FOSB, MYC, and CDKN1A—are involved in cell cycle regulation and varied most significantly in expression between KC and control corneas. Figure 1 presents the relative distribution of transcript expression levels in KC and control corneas for overexpressed (Fig. 1A) and underexpressed (Fig. 1B) genes.

Only 18 genes displayed upregulation by this analysis. Among these are several genes involved in the extracellular matrix and the epithelial cell cytoskeleton (PTCH2, KRT5, KRT78, and LYPD3), in the stress response (HSP90AA1 and ALDH1A3), or encoding mucins (MUC4 and MUC16).

Markers considered to be melanoma-specific (42–44) were absent from the differentially expressed transcripts, consistent with our screening of the control corneas.

In parallel, a weighted correlation network analysis using the WGCNA R package (1) identified a module of 134 transcripts (Supplementary Table S4, http://www iovs org/lookup/suppl DOI 10 1167/ iovs 10 70981/DCSupplemental) that are co-expressed in correlation with the disease status across the microarray samples ($P = 9 \times 10^{-4}$) and globally decreased in KC (average $\log_2(\text{FC}) = -0.68$). The correlation network and individual gene analyses are convergent, in that 33 transcripts in the set of 69 individually underexpressed transcripts belong also to the WGCNA module.

**Table 1. Transcripts with Differentially Levels of Expression between KC and Control Corneas**

<table>
<thead>
<tr>
<th>HUGO Gene Symbol</th>
<th>$\log_2(\text{FC})$</th>
<th>$P$</th>
<th>Gene Name</th>
<th>Change Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRT78</td>
<td>1.21</td>
<td>1.62 $\times 10^{-5}$</td>
<td>Keratin 78-keratin 5B</td>
<td>2.31</td>
</tr>
<tr>
<td>MUC4</td>
<td>0.49</td>
<td>7.46 $\times 10^{-6}$</td>
<td>Mucin 4</td>
<td>1.40</td>
</tr>
<tr>
<td>S100A6</td>
<td>0.27</td>
<td>7.92 $\times 10^{-6}$</td>
<td>S100 calcium binding protein A6 (calcyclin)</td>
<td>1.21</td>
</tr>
<tr>
<td>ROR2</td>
<td>0.12</td>
<td>9.57 $\times 10^{-5}$</td>
<td>Receptor tyrosine Kinase 10-like orphan receptor 2</td>
<td>1.09</td>
</tr>
<tr>
<td>SQSTM1</td>
<td>-0.1</td>
<td>4.90 $\times 10^{-5}$</td>
<td>Sequestosome 1</td>
<td>0.93</td>
</tr>
<tr>
<td>INSIG1</td>
<td>-0.15</td>
<td>4.59 $\times 10^{-5}$</td>
<td>Insulin-induced gene 1</td>
<td>0.90</td>
</tr>
<tr>
<td>HSPB1</td>
<td>-0.16</td>
<td>4.87 $\times 10^{-5}$</td>
<td>Heat shock 27kDa protein 1</td>
<td>0.90</td>
</tr>
<tr>
<td>NQO1</td>
<td>-0.2</td>
<td>1.51 $\times 10^{-5}$</td>
<td>NAD(P)H dehydrogenase, quinone 1</td>
<td>0.87</td>
</tr>
<tr>
<td>DDX3X</td>
<td>-0.2</td>
<td>9.25 $\times 10^{-6}$</td>
<td>DEAD (ASPGL-ALAASP) box polypeptide 3, X-linked</td>
<td>0.87</td>
</tr>
<tr>
<td>KRT5</td>
<td>-0.24</td>
<td>6.86 $\times 10^{-7}$</td>
<td>Keratin 5A</td>
<td>0.85</td>
</tr>
<tr>
<td>HSP90AA1</td>
<td>-0.27</td>
<td>4.64 $\times 10^{-5}$</td>
<td>Heat shock protein 90kDa alpha (cytosolic), class A member 1</td>
<td>0.83</td>
</tr>
<tr>
<td>FTH1</td>
<td>-0.32</td>
<td>3.45 $\times 10^{-7}$</td>
<td>Ferritin, heavy polypeptide 1</td>
<td>0.80</td>
</tr>
<tr>
<td>RNF39</td>
<td>0.32</td>
<td>1.10 $\times 10^{-6}$</td>
<td>Ring finger protein 39</td>
<td>0.78</td>
</tr>
<tr>
<td>DDX5</td>
<td>0.39</td>
<td>1.33 $\times 10^{-6}$</td>
<td>Dead (Asp-Glu-Ala)-Box polypeptide 5</td>
<td>0.76</td>
</tr>
<tr>
<td>ACTB</td>
<td>-0.45</td>
<td>5.12 $\times 10^{-6}$</td>
<td>Actin, Beta</td>
<td>0.73</td>
</tr>
<tr>
<td>SAT1</td>
<td>-0.59</td>
<td>1.56 $\times 10^{-6}$</td>
<td>Spermidine/Spermine N1-Acetyltransferase</td>
<td>0.66</td>
</tr>
<tr>
<td>DNAJB1</td>
<td>-0.59</td>
<td>1.65 $\times 10^{-6}$</td>
<td>Dnaj (Hsp40) Homolog, Subfamily B, Member 1</td>
<td>0.66</td>
</tr>
<tr>
<td>ZNF750</td>
<td>-0.61</td>
<td>1.54 $\times 10^{-6}$</td>
<td>Hypothetical Protein LOC79755</td>
<td>0.66</td>
</tr>
<tr>
<td>THBS1</td>
<td>-0.65</td>
<td>6.00 $\times 10^{-6}$</td>
<td>Thrombospondin 1</td>
<td>0.65</td>
</tr>
<tr>
<td>MYC</td>
<td>-0.64</td>
<td>9.27 $\times 10^{-6}$</td>
<td>V-Myc myelocytomatosis viral oncogene homolog (avian)</td>
<td>0.64</td>
</tr>
<tr>
<td>HES1</td>
<td>-0.71</td>
<td>1.13 $\times 10^{-5}$</td>
<td>Hairy and enhancer of split 1, (Drosophila)</td>
<td>0.64</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>-0.76</td>
<td>2.79 $\times 10^{-6}$</td>
<td>Cyclin-dependent kinase inhibitor 1a (P21, Cip1)</td>
<td>0.59</td>
</tr>
<tr>
<td>MAT2A</td>
<td>-0.83</td>
<td>3.16 $\times 10^{-6}$</td>
<td>Methionine adenosyltransferase II, Alpha</td>
<td>0.56</td>
</tr>
<tr>
<td>ZFP3G1L</td>
<td>-0.87</td>
<td>6.85 $\times 10^{-6}$</td>
<td>Zinc finger protein 36, C3h Typ.10-Like 1</td>
<td>0.59</td>
</tr>
<tr>
<td>MCL1</td>
<td>-1.01</td>
<td>9.20 $\times 10^{-6}$</td>
<td>Myeloid cell leukemia sequence 1 (Bcl2-related)</td>
<td>0.50</td>
</tr>
<tr>
<td>ID1</td>
<td>-1.08</td>
<td>3.59 $\times 10^{-6}$</td>
<td>Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein</td>
<td>0.47</td>
</tr>
<tr>
<td>ODCl</td>
<td>-1.08</td>
<td>3.57 $\times 10^{-6}$</td>
<td>Ornithine decarboxylase 1</td>
<td>0.47</td>
</tr>
<tr>
<td>SLC20A1</td>
<td>-1.08</td>
<td>5.56 $\times 10^{-6}$</td>
<td>Solute carrier family 20 (phosphate transporter), Member 1</td>
<td>0.47</td>
</tr>
<tr>
<td>NUAK2</td>
<td>-1.16</td>
<td>9.61 $\times 10^{-5}$</td>
<td>Nuak family, Snf1-like kinase, 2</td>
<td>0.45</td>
</tr>
<tr>
<td>BTG2</td>
<td>-1.52</td>
<td>1.24 $\times 10^{-5}$</td>
<td>Btg family, member 2</td>
<td>0.35</td>
</tr>
<tr>
<td>UBC (RP527A)</td>
<td>-1.84</td>
<td>1.85 $\times 10^{-5}$</td>
<td>Ubiquitin C</td>
<td>0.28</td>
</tr>
<tr>
<td>ZFP36</td>
<td>-1.92</td>
<td>1.02 $\times 10^{-5}$</td>
<td>Zinc finger protein 36, C3h type, Homolog (Mouse)</td>
<td>0.26</td>
</tr>
<tr>
<td>DUSP1</td>
<td>-2.19</td>
<td>2.02 $\times 10^{-5}$</td>
<td>Dual specificity phosphatase 1</td>
<td>0.22</td>
</tr>
<tr>
<td>JUN</td>
<td>-3.13</td>
<td>9.88 $\times 10^{-5}$</td>
<td>V-Jun sarcoma virus 17 oncogene homolog (Avian)</td>
<td>0.11</td>
</tr>
<tr>
<td>FOXB</td>
<td>-3.69</td>
<td>1.87 $\times 10^{-5}$</td>
<td>Fbj Murine osteoscarcoma viral oncogene homolog B</td>
<td>0.08</td>
</tr>
<tr>
<td>FOXS</td>
<td>-4.2</td>
<td>2.95 $\times 10^{-5}$</td>
<td>V-Fos Fbj murine osteoscarcoma viral oncogene homolog</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Gene names (HGNC [HUGO Gene Nomenclature Committee] at the European Bioinformatics Institute) are provided alongside their respective FC, and the P value from Student’s $t$-test. A more complete table (including annotation) is provided as Supplementary Table S1 (http://www iovs org/lookup/suppl DOI 10 1167/ iovs 10 70981/DCSupplemental).
**FIGURE 1.** Levels of expression of the differentially expressed genes. The boxplots represent dispersion among samples. KC, keratoconjunctivitis samples; ctrl, control samples. (A) Transcripts underexpressed in KC, ordered by decreasing absolute FC. (B) Transcripts overexpressed in KC, ordered by increasing absolute FC.
KC Patients Express Different Splice Isoforms of Certain Transcripts

We next set out to identify genes that were differentially spliced in the comparison between the KC and control groups. Thirty-six candidates were identified with a threshold of \( P = 0.05 \) with either the Splicing Index or MiDAS methods (Table 2, and see Methods for details on significance levels). MUC4, encoding a membrane-tethered sialomucin of the ocular surface epithelium\(^6\), is remarkable among these, because it displays profound changes in transcript regulation at both the gene (overexpression in KC) and splice regulation levels.

Network Biology and GO Annotations

An interaction network was inferred to test whether underexpressed genes were involved in one or more biological signaling pathways. This network was obtained after a protein–protein interaction matrix was retrieved from a high-confidence STRING database query (score \( \geq 0.7 \)). The representation of weighted protein interactions from STRING provides a high-level view of functional linkage, enhancing the analysis of modularity in biological processes. Without the addition of further molecular partners, we were able to cluster most of the both differentially expressed and differentially spliced proteins into a single, highly connected network (Fig. 2). A core of those genes seems central to this network, as they are the most connected nodes. In decreasing order of connectivity, they are JUN, MYC, FOS, PTGS2, CDKN1A, ODC1, HSP90AA1, ALDH3A1, ANXA1, and NQO1.

After loading this network into Cytoscape, the BINGO plugin was used to further investigate whether specific GOs or pathways were overrepresented in the whole network. Within the first cluster of GO terms relative to biological function (enrichment score \( = 3.080 \)), we filtered out several redundant terms (Supplementary Table S5, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental), mainly within categories related to apoptosis regulation. On the other hand, within clusters 3, 4, and 5 (scores 2.43, 2.22, and 2.09, respectively), there were many cell-cycle/proliferation-oriented categories. More than half of the genes had GO terms within these categories and corresponding GO term \( P \)

<table>
<thead>
<tr>
<th>HUGO Gene Symbol</th>
<th>SI</th>
<th>SI ( P )</th>
<th>MiDAS ( P )</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTGS2</td>
<td>1.58</td>
<td>( 2.50 \times 10^{-05} )</td>
<td>( 3.93 \times 10^{-02} )</td>
<td>Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)</td>
</tr>
<tr>
<td>FOSL2</td>
<td>1.24</td>
<td>( 8.35 \times 10^{-05} )</td>
<td>( 3.71 \times 10^{-02} )</td>
<td>FOS-like antigen 2</td>
</tr>
<tr>
<td>MYL6</td>
<td>1.35</td>
<td>( 1.86 \times 10^{-04} )</td>
<td>( 3.80 \times 10^{-02} )</td>
<td>Myosin, light chain 6, alkali, smooth muscle and non-muscle</td>
</tr>
<tr>
<td>HSPB1</td>
<td>2.72</td>
<td>( 2.32 \times 10^{-04} )</td>
<td>( 5.14 \times 10^{-02} )</td>
<td>Heat shock 27kDa protein 1</td>
</tr>
<tr>
<td>BAG1</td>
<td>1.18</td>
<td>( 5.86 \times 10^{-04} )</td>
<td>( 4.82 \times 10^{-02} )</td>
<td>BCL2-associated athanogene</td>
</tr>
<tr>
<td>ALDH3A1</td>
<td>1.16</td>
<td>( 8.26 \times 10^{-05} )</td>
<td>( 3.46 \times 10^{-02} )</td>
<td>Aldehyde dehydrogenase 3 family, member A1</td>
</tr>
<tr>
<td>DDX5</td>
<td>1.40</td>
<td>( 6.13 \times 10^{-05} )</td>
<td>( 3.30 \times 10^{-02} )</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 5</td>
</tr>
<tr>
<td>ODC1</td>
<td>3.00</td>
<td>( 8.16 \times 10^{-05} )</td>
<td>( 3.26 \times 10^{-02} )</td>
<td>Ornithine decarboxylase 1</td>
</tr>
<tr>
<td>SDC1</td>
<td>1.74</td>
<td>( 2.65 \times 10^{-05} )</td>
<td>( 3.60 \times 10^{-02} )</td>
<td>Syndecan 1</td>
</tr>
<tr>
<td>SLC2A1</td>
<td>1.19</td>
<td>( 2.25 \times 10^{-04} )</td>
<td>( 4.06 \times 10^{-02} )</td>
<td>Solute carrier family 2 (facilitated glucose transporter), member 1</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>1.04</td>
<td>( 2.99 \times 10^{-05} )</td>
<td>( 2.77 \times 10^{-02} )</td>
<td>Cyclin-dependent kinase inhibitor 1A (p21, Cip1)</td>
</tr>
<tr>
<td>ID1</td>
<td>1.31</td>
<td>( 2.26 \times 10^{-04} )</td>
<td>( 2.34 \times 10^{-02} )</td>
<td>Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein</td>
</tr>
<tr>
<td>SAT1</td>
<td>1.30</td>
<td>( 3.23 \times 10^{-04} )</td>
<td>( 2.21 \times 10^{-02} )</td>
<td>Spermide/spermine N1-acetyltransferase 1</td>
</tr>
<tr>
<td>LDLR</td>
<td>1.72</td>
<td>( 1.14 \times 10^{-06} )</td>
<td>( 1.86 \times 10^{-02} )</td>
<td>Low density lipoprotein receptor (familial hypercholesterolemia)</td>
</tr>
<tr>
<td>H3F3B</td>
<td>1.23</td>
<td>( 8.80 \times 10^{-05} )</td>
<td>( 3.64 \times 10^{-02} )</td>
<td>H3 histone, family 3B (H3.3B)</td>
</tr>
<tr>
<td>TPT1</td>
<td>1.82</td>
<td>( 1.88 \times 10^{-04} )</td>
<td>( 3.57 \times 10^{-02} )</td>
<td>Tumor protein, translationally-controlled 1</td>
</tr>
<tr>
<td>RRAS2</td>
<td>1.77</td>
<td>( 2.82 \times 10^{-04} )</td>
<td>( 4.21 \times 10^{-02} )</td>
<td>Related RAS viral (r-ras) oncogene homolog 2</td>
</tr>
<tr>
<td>BHLHJ2/BHLHE40</td>
<td>1.54</td>
<td>( 6.83 \times 10^{-04} )</td>
<td>( 4.04 \times 10^{-02} )</td>
<td>Basic helix-loop-helix domain containing, class B, 2/class E basic helix-loop-helix protein 40</td>
</tr>
</tbody>
</table>

Table 2. Genes Differentially Spliced between Keratoconus and Control Corneas

Gene names (HGNC/HUGO Gene Nomenclature Committee at the European Bioinformatics Institute) are provided alongside their respective Splicing Index (SI) or MiDAS (Microarray Detection of Alternative Splicing) readouts. A more complete table (including annotation) is provided as Supplementary Table S2 (http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental).
waves, after correction, were below $10^{-6}$, meaning these categories were significantly overrepresented relative to their frequency in a randomized sample of expressed transcripts. Further examination showed that KC patients had significantly lower expression of genes that fell into a limited number of KEGG signaling pathways: ErbB (CDKN1A, JUN, and MYC), MAP kinases (FOS, DUSP1, JUN, HSPB1, and MYC), and focal adhesion (ACTG1, ACTB, CDKN1A, JUN, and THBS1) molecular cascades.

In addition, we mapped the FCs in expression between cases and controls (Supplementary Table S6, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental) over the loci identified in previously published linkage analyses.\textsuperscript{2–8} In agreement with the genetic heterogeneity of these studies and their difficulty in reproducing each other’s results, no obvious candidate genes for KC were distinguishable, as their FC ranged from −0.45 to +0.58. We also scrutinized the previously hypothesized candidate genes for KC (COL6A1, SOD1, MMP9, MMP2, and AQP5). None of them showed significance at the Bonferroni-corrected threshold of $10^{-6}$.

Validation of Gene Microarray Data

Relative quantitative real-time polymerase chain reaction (qPCR) was conducted as an independent technique to validate the expression level changes in the microarray experiment. Results in Figure 3 show a good level of correlation between the three empiric assessment of expression as shown by rho values $\sim 0.9$. Even considering as a replicate the samples used only for the qPCR experiment, the same levels of correlation are observed (data not shown).

DISCUSSION

Using full human transcriptomic microarray, our study is the first comparison of expression of all known genes between control and KC whole corneas. The microarray chosen is a multitarget array including several probesets per exon and four neighbor replicates per probeset, allowing a better confidence in gene-level expression estimates and as such also provides the possibility of exploring alternative RNA splicing. As a matter of fact, quantitative RT-PCR validated the microarray data, even though this last method reflects the expression of only
immunity53), we have transposed this approach to KC. The microarray chips targeting 5600 genes with a limited number of the differentially spliced ones (Supplementary Table S7, http://the annotation database we used (STRING), is supposed to replicate the results of the others, perhaps due to a limited segment of the eye was free of neoplastic extensions; and (3) corneas excluding previously treated subjects; (2) the anterior (1) collection was achieved in conditions similar to those of KC melanoma. They were considered to be a proper control since whole corneas. As controls, we used corneas from patients in vivo48 and in vitro49 keratocytes. Two studies used restricted microarray chips targeting 5600 genes with a limited number of probe sets,50,51 and one study directly targeted the expression of 164 apoptotic genes.49 However, none of these studies replicate the results of the others, perhaps due to a limited number of corneal samples, as well as the limited probe set distribution and redundancy on microarray chips used in those studies compared with currently used microarrays (e.g., the Affymetrix Exon array).

Our own study on human KC tissue was conducted on whole corneas. As controls, we used corneas from patients who underwent enucleation for posterior pole (i.e., chorioidal) melanoma. They were considered to be a proper control since (1) collection was achieved in conditions similar to those of KC corneas excluding previously treated subjects; (2) the anterior segment of the eye was free of neoplastic extensions; and (3) melanoma extension is blood borne,50 cornea is avascular, and melanoma transmission was not observed after grafting cornea from a melanoma donor.51 Consistently, we did not observe overexpression of any melanoma-specific transcript.

In this context, our aim was to reassess the genetic component of the KC, taking into account the following hypotheses: a condition characterized by genetic heterogeneity, with various loci linked in different populations, and a complex phenotypic trait whose transmission appears to deviate from a Mendelian model. Because transcriptomics have helped identify molecular pathways and alternatively spliced variants involved in other complex diseases (e.g., in oncology15,20,52 and autoimmune53), we have transposed this approach to KC. The interplay of both experimental and predicted interactions of the annotation database we used (STRING), is supposed to offer a more accurate approximation of the interactome than GO alone. Nevertheless, this analysis infers statistical links between KC phenotype and gene expression and requires experimental validation, perhaps in animal models. We identified mice mutants for 29 of the differentially expressed genes and 27 of the differentially spliced ones (Supplementary Table S7, http://www iovs.org/lookup/suppl/doi:10.1167/iovs.10-70981/-/DC Supplemental) in the Mouse Genome Informatics database (http://www.informatics.jax.org/ Jackson Laboratory, Bar Harbor, ME). Some of them display corneal clinical phenotypes: The Jun mutant shows eye opacity and increased incidence of corneal inflammation, significant in the light of a recent study also implicating a downstream JUN effector, JNK2, in abnormal corneal barrier response to dry eyes.54 The THBS1 (thrombospondin 1) mutant displays abnormal corneal epithelial morphology, and the Ldlr (low density lipoprotein receptor) mutant also has dry eyes. However, their relevance remains to be tested, especially in human disease: JUN is a human proto-oncogene implicated in carcinogenesis, whereas LDLR mutations lead to autosomal dominant hypercholesterolemia.

The majority (69/87) of the genes highlighted were downregulated in KC corneas. They include several genes globally involved in cellular proliferation and the prevention of differentiation (e.g., the AP-1 transcription factor partners FOS, JUN, and FOSB)55 or MYC, required for proliferation or stem cell mobilization.56 Confirming previous studies, we did not find any significant change regarding putative candidate genes previously proposed (MMP2, MMP9, COL6A1, SOD1, and AQ5). Another original finding is that the β-actin ACTB is downregulated, thus confirming that, at least in pathologic conditions, it cannot be considered a good reference gene for expression in cornea. GO terms enrichment analysis demonstrated, independent of the algorithms used, that these downregulated genes are more widely involved in so-called (cellular) developmental processes or cell differentiation. First, we observed a redundancy of GO terms in the first three annotation clusters (GO:0032502, developmental process; GO:0048869, cellular development; and GO:0030154, cell differentiation). Their congruence strengthens the conclusion that some KC corneal cells are maintained in a less differentiated, proliferative state.57,58 Simultaneously, decreased protection from apoptosis may be implied by the terms GO:0043066 (negative regulation of apoptosis) or GO:0048468 (cell development). The “gene population background”-based enrichment analysis from DAVID’s web-based annotation tool,59 which uses GO-Elite60 gave similar results. The weighted correlation network analysis showed significant co-expression of modules containing intermediaries consistent with the highlighted pathways. Finally, GO enrichment analysis gave roughly the same
results when applied to the $P < 0.01$ and $P < 0.05$ differentially expressed gene lists. Moreover, the qPCR results show that for all genes selected for validation, the direction and magnitude of changes were consistent with the results obtained from the microarray analysis.

Finally, we failed to replicate previous differential expression experiments on KC versus normal corneas. None of the differentially expressed genes found in these studies overlapped each other and with our results. Nevertheless, we performed a network biology analysis combining the highlighted genes from these and our studies, as well as three genes involved in communication between the epithelium and the stroma: HGF, KGF, and EGF. (Supplementary Fig. 59, http://www.iovs.orglookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental). We found that although we did not highlight the same genes, the same biological pathways seem to be involved. The dataset from Stachs et al. comes from freshly isolated keratocytes, while the one of Nielsen et al. comes from fresh epithelial cells. Our own dataset composed of both epithelial and stromal components is linked with both datasets, showing that these different studies are congruent. In addition, the links observed between the majority of our highlighted genes and Nielsen et al.’s ones by means of EGF, KGF, and EGF, considered to be involved in epithelium/stroma signaling, are further arguments that there is probably an interplay between both cellular layers in KC.

Our results implicate both known and novel pathways as perhaps playing a key role in the pathophysiology of KC. It is noteworthy to mention here that many of the selected genes have elsewhere been associated with cancer. Nevertheless, they are also primarily involved in developmental processes and cell proliferation. Thus, it is not surprising that a change in mRNA expression could be associated with a disease suspected to derive from cell loss. In fact, apoptosis has been observed in KC corneas by direct confocal microscopy and histology showing decreased cell densities in the three corneal layers, as well as biochemical or expression studies. This is in full agreement with the overrepresentation of GO term annotation in our dataset. The primary cause of apoptosis in KC remains unknown, but it may be a secondary consequence of another molecular defect. Several reports have suggested a decreased resistance to environmental aggressions by hypersensitivity to oxidative stress in cultivated keratocytes from KC corneas, as well as in vivo. Cultured KC fibroblasts also may have an inherent, hypersensitive response to oxidative stressors that involves mitochondrial damage and the expression of intermediate filaments such as vimentin or tenascin or the proinflammatory TGF-β and IL-1. Apoptosis in KC may also result from mechanical trauma or be secondary to the production of cytokines after atopy.

At this point, the cellular population affected by these network alterations remains uncertain. This lack of hypothesis led us to perform, in a first approach, a full-thickness corneal study. The recovered amounts of mRNAs were largely in favor of epithelial cells, and this could have biased the results and consequently their interpretation. Nevertheless, one can imagine that both epithelial cells and keratocytes are key players. Keratocytes are specialized fibroblasts, derived from neural crest mesenchyme in common with the cartilaginous stromal layer behind the eyeball, that produce collagen fibers and proteoglycans constituting the corneal stroma. Throughout life, most keratocytes are in a quiescent state. By the end of eye development, a keratocyte network, interconnected through dendritic processes, is established. Keratocyte apoptosis, either of quiescent or actively dividing cells, is a process of great interest for corneal growth and remodeling. Previous reports have speculated that keratocytes may be involved in KC development through increased catabolism, either misregulation of metalloprotease activity or a modification of collagen subtype composition. We suggest here the novel hypothesis that keratocyte implication would be mediated by a tissue-specific misregulation of apoptosis that may be due to less redundant antiapoptotic pathways in the cornea. In a healthy cornea, programmed cell death is a rare occurrence, but immediately after an injury, keratocytes directly below the injury site undergo apoptosis when the basal membrane is broken. In the following steps of the healing and scarring process, this cell loss is counteracted by mitoses among the remaining adjacent keratocytes. In KC corneas, impairment of apoptotic signaling pathways on minor corneal injuries sustained through life may affect appropriate keratocyte proliferation, itself necessary for the recovery of the normal cell density. Keratocyte-restricted deficiencies in the regulation of apoptosis may lead to the gene expression differences we have observed in this study. This could provide an explanation for the loss of cell density and corneal thinning observed in KC. Interestingly, a recent study underlined the role of TWIST2, a bHLH transcription factor, in keratocyte proliferation in mouse leading to a cellontional thinning. The epithelium, with cells that are continuously renewed, may also be affected by the antiproliferative and hyperapoptotic phenotypes. The 18 overexpressed genes we observed to be significantly differentially expressed included KRT78, ROR2, S100A6, MUC4, and KRT78 is a structural protein of epithelial cells whose cross-linking helps withstand mechanical and chemical stresses. ROR2 encodes a nuclear orphan receptor in the noncanonical Wnt pathway that appears to trigger, among other effects, the maintenance and proliferation of stem cells. S100A6 (calycin) is involved in the fibroblast cell cycle and overexpressed during wound healing after corneal injury. In addition, sialomucin (MUC4) is expressed by the corneal epithelium, which raises the question of how epithelial cells may also be involved in the pathogenesis of KC. We found increased expression of the whole transcript population of MUC4 in KC corneas. The apical location of mucins could suggest their general involvement in mechanisms of response to epithelial damage. In addition, among mucins, MUC4 displays tissue-specific expression patterns—notably, a conjunctival-type expression pattern is observed on the corneal surface in limbal stem cell deficiency. Both mucin-4 (the product of the MUC4 gene) and calycin are implicated in corneal protection and wound healing by the constitution of the lacrimal film. Mutations in other mucin genes are already implicated in human pathology, notably the dry eye syndromes (Supplementary Table S6, http://www.iovs.orglookup/suppl/doi:10.1167/iovs.10-70981/-/DCSupplemental). Alternatively sialomucin is also an activator partner of the epidermal growth factor ErbB2 receptor, potentially implicating its overexpression in growth factor signaling pathways leading to either differentiation or increased cell proliferation. Each of these overexpressed transcripts therefore could be either pathogenic or simply deregulated as a consequence of other causal molecular mechanisms. Overexpression of MUC4 may be a protective response or could be causative in KC, and a corneal-specific mouse overexpression model would be welcome to resolve this question.

Another differentially spliced molecule is annexin A1 (ANXA1). This calcium-dependent phospholipid-binding protein has no allelic variants described, regulates phospholipase A2 activity, and promotes membrane fusion. It has been shown to be upregulated in KC corneal epithelial cells relative to normal epithelial cells. ANXA1 is located within the chromosomal interval 9q21, which has been suggestively linked to KC using nonparametric linkage. However, it is the only differentially expressed gene that showed consensus with the
chromosomal regions linked to KC so far. This observation confirms that expression analysis is an interesting complementary approach to linkage analyses, and that our having included the keratocyte population from whole KC versus normal corneas not only found this difference but has implicated broad functional pathways in addition to individual molecules.

Altogether these results cannot exclude that, whatever the pathways involved, KC could result from a distortion of the cross-talk between epithelial and stromal cells of the cornea. In the future, more precise, layer specific expression studies should be conducted. Such expression results may be combined with the latest possibilities for genome-wide association studies to identify genomic variants that correlate with both expression and KC phenotype.

**Acknowledgments**

The authors thank Nathalie Marsaud (Genotoul Biopuces, Toulouse Biochip Facility) for help and advice about the experimental protocol and Didier Laborie and the Bioinfo-Genotoul team at the Toulouse Bioinformatics Facility.

**References**

40. Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe.


