

Microarray Analysis of Gene Expression Patterns during Healing of Rat Corneas after Excimer Laser Photorefractive Keratectomy

Juan C. Varela,¹ Michael H. Goldstein,¹ Henry V. Baker,^{2,3} and Gregory S. Schultz¹

PURPOSE. To characterize changes over time in the genomic expression profile of rat corneas after excimer laser photorefractive keratectomy (PRK), in an effort to better understand the cellular response to injury and the dynamic changes that occur in gene expression patterns as a wound heals.

METHODS. The corneal gene expression profile of 1176 genes at 3 and 7 days after PRK was determined and compared with untreated corneal gene expression patterns by interrogating commercially available cDNA arrays with labeled target cDNA prepared from pooled total RNA harvested from the respective treatment group of adult male rats. The gene expression patterns were inferred based on the hybridization intensities of the probes on the cDNA arrays. The hybridization signals were globally normalized and filtered. The data were analyzed by using hierarchical and k-means clustering algorithms before and after normalization of variances.

RESULTS. Of the 1176 cDNA elements on the array, 588 consistently produced similar results in replicate experiments and comprised the data set analyzed in this work. In total, 73 genes were identified, with expression levels that differed by at least threefold at either 3 or 7 days after PRK. At 3 days after PRK, 70 genes were identified with expression levels that differed by more than threefold, compared with the expression level in untreated animals. The expression of 42 genes increased by threefold or more, whereas expression of 28 genes decreased by threefold or more. By day 7 after PRK, the number of genes displaying more than a threefold difference in expression pattern was reduced to 27 genes, 20 of which showed elevated levels, whereas 7 exhibited decreased levels. Hierarchical clustering of the 588 studied genes produced 10 clusters with correlation coefficients of 0.9 or greater. To determine whether any of the clusters were overrepresented by genes with related functions, the cumulative hypergeometric probability was calculated by obtaining the observed number of functionally related genes within each of the 10 clusters. Seven of the clusters were statistically overrepresented by one or more categories of functionally related genes, such as cell cycle regulators, transcription factors, and metabolic pathway genes. Clustering analysis of 56 genes generally considered to influ-

ence corneal wound healing produced 10 gene clusters with correlation coefficients of at least 0.9. Expression of 23 of these 56 genes increased at day 3, then decreased at day 7 to levels similar to those on day 0. These included several growth factors (*VEGF*, *FGF*, *IGF-I*), proteases (*PAI-1*, *PAI-2A*) and protease inhibitors (*TIMP-2* and *TIMP-3*). Expression of nine genes increased on both days 3 and 7 compared with expression on day 0 (e.g., *TGFB1*, *TGFBIR*, *M6P/IGFR-2*), and no genes decreased on both days 3 and 7, compared with day 0.

CONCLUSIONS. Microarray analysis of 1176 identified 588 genes with reproducible patterns of expression in rat corneas on days 3 and 7 after PRK and 73 genes with a threefold change in expression compared with untreated corneas. Hierarchical clustering of these 588 genes identified 10 clusters of genes with very similar patterns of expression. Clustering of genes with similar patterns of expression implies a common regulatory pathway for the genes within a cluster, and identifies potential new targets for regulating corneal wound healing. (*Invest Ophthalmol Vis Sci.* 2002;43:1772-1782)

Corneal wound healing is a complex biological process that involves the integrated actions of many genes that code for proteins with diverse functions, such as cytokines, growth factors, receptors, extracellular matrix proteins, integrins, proteases, inhibitors and cell cycle genes.^{1,2} When expression of these genes is regulated correctly, corneal wounds heal appropriately. However, when intrinsic or extrinsic factors disrupt the expression of these genes, corneal wounds may either fail to heal adequately or may heal with excessive scar formation. Thus, there is a need to understand, as broadly as possible, how the expression of key genes changes during corneal wound healing so that therapies can be logically designed to prevent complications from occurring.

Previous studies investigating gene expression during corneal wound healing have analyzed mRNA levels of a small number of selected growth factors, receptors, and extracellular matrix genes that are thought to regulate or participate in corneal wound healing. For example, quantitative reverse transcription-polymerase chain reaction (Q-RT-PCR) of rat corneas after photorefractive keratectomy (PRK) showed prolonged elevation of mRNAs for the three isoforms of transforming growth factor- β (*TGF β 1*, *TGF β 2*, *TGF β 3*), the *TGF β* type II receptor (*TGF β R-II*), and extracellular matrix proteins (types I, III, and VI collagen and fibronectin).³ Similar results were reported for mRNA levels for collagens I, III, IV, and V using semiquantitative RT-PCR in rat corneas after PRK.⁴ Analysis of mRNA levels for hepatocyte growth factor (HGF), keratinocyte growth factor (KGF), epidermal growth factor (EGF) and their receptors in mouse corneas after an epithelial scrape injury suggests that HGF and KGF also play important roles in regulating epithelial wound healing.⁵

Other groups have investigated changes in protein levels in corneas during wound healing. Immunohistochemical analysis of cat corneas after PRK indicated there were increased levels of *TGF β 1*, *TGF β 2*, and *TGF β 3* and the *TGF β R-I* and *TGF β R-II* receptor proteins in the subepithelial tissue of laser-ablated

From the ¹Institute for Wound Research and Department of Ophthalmology, University of Florida, Gainesville, Florida; ²Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, Florida; and ³University of Florida Genetics Institute, Gainesville, Florida.

Supported in part by National Institutes of Health Grant EY05587 and Research to Prevent Blindness.

Submitted for publication June 25, 2001; revised February 7, 2002; accepted February 15, 2002.

Commercial relationships policy: N.

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: Gregory S. Schultz, Institute for Wound Research and Department of ObGyn, 1600 SW Archer Road, University of Florida, Gainesville, FL 32610-0294; schultzg@obgyn.ufl.edu.

regions at 4 weeks after injury.⁶ Immunofluorescence staining of monkey corneas showed distinct changes in the amount and localization of fibronectin, type III collagen, type VII collagen, and keratin sulfate in ablated regions at different times during the 18 months after excimer laser ablation.⁷ Analyses of matrix metalloproteinases (MMPs) and the tissue inhibitors of metalloproteinases (TIMPs) in rat and rabbit corneas after PRK found increased levels of TIMP-2 and MMP-9, -7, and -13 proteins and mRNAs in migrating basal epithelial cells.⁸⁻¹² Immunofluorescence analysis of radial keratotomy (RK) wounds in human corneas detected several extracellular matrix proteins that were not present in the stroma of normal corneas, including collagen types III, VIII, and XIV; the $\alpha 1$ - $\alpha 2$ chains of type IV collagen; tenascin-C; and fibrillin-1.¹³ Immunohistochemical analysis of rat corneas after PRK ablation showed that macrophages moved into injured area within 36 hours after injury.¹⁴

These studies have provided important information about the expression of a few selected genes during healing of corneal wounds. However, the ability to simultaneously measure changes in expression of large numbers of genes during corneal wound healing is severely limited by conventional techniques of immunohistochemistry, enzyme analysis, Northern blot analysis, RNase protection assays, or Q-RT-PCR. In this study, we simultaneously analyzed expression of 588 genes in rat corneas on day 0 (nonablated), days 3 and 7 after excimer ablation, using membrane microarray technology. The expression patterns of these 588 genes were analyzed to identify genes with similar patterns of expression. In addition, the levels of expression of a selected subset of 56 genes thought to influence corneal wound healing were examined in greater detail.

MATERIALS AND METHODS

Animal Models

Animal procedures were performed in accordance with the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research, and the animal protocol was approved by the University of Florida Animal Care and Use Committee. Rat corneas underwent excimer laser ablation as described previously.¹⁵ Briefly, a total of 30 adult Sprague-Dawley male rats (250 g) with normal eyes were anesthetized with an intraperitoneal injection of pentobarbital sodium (30 mg/kg). Eyelashes and whiskers surrounding the eye were removed from the visual field. A drop of proparacaine (0.05%) was applied to the eye, and the cornea was centered under the laser microscope. Bilateral ablation of the corneas was performed in a 4.4-mm treatment zone with an excimer laser (SVS Apex; Summit Technology, Waltham, MA). The corneal epithelium was ablated to a depth of 40 μm , followed by ablation of the stroma to a depth of 20 μm , for a total ablation depth of 60 μm . After excimer laser treatment, tobramycin (0.3%) drops were applied to the corneal surface to prevent infection. No postoperative topical steroid was administered. On days 3 and 7 after excimer laser ablation, 15 rats were randomly selected and killed by peritoneal injection of pentobarbital. Corneas were excised under an operating microscope and placed in solution (RNAlater; Ambion, Austin, TX), to stabilize the mRNA. Once harvested, the corneas were stored at 4°C not longer than 24 hours before RNA extraction. Thirty corneas were harvested at each time point. For each time point three pools of 10 corneas each were prepared, and total RNA was extracted, as described in the next section. An additional 30 untreated normal corneas were harvested, separated into three pools, and processed as described. Thus, nine RNA samples were generated comprising three independent preparations for each of the three experimental groups (nonablated corneas designated as day 0 and days 3 and 7 after ablation).

RNA Isolation from Rat Corneas

Total RNA was extracted from each group of 10 pooled corneas using guanidine isothiocyanate and phenol-chloroform according to standard methods. Briefly, pooled corneal tissue was homogenized in 0.75 mL RNA extraction solution (TRIzol; Gibco, Grand Island, NY) using a 5-mL frosted glass-on-glass tissue grinder (Duell; Konte Scientific Glassware, Vineland, NJ), RNA was extracted with chloroform, precipitated with isopropanol, washed with ethanol, and dissolved in RNase-free water treated with 0.1% diethylpyrocarbonate. The material was then treated with 1 U/ μL DNase I at 37°C for 30 minutes to remove any contaminating genomic DNA, and then extracted once with phenol-chloroform-isoamyl alcohol (25:24:1 pH 4.5), followed by extraction with chloroform, precipitated with 2 M sodium acetate and 95% ethanol, washed with 80% ethanol, and resuspended in RNase-free water. The concentration of RNA were measured spectrophotometrically at 260 nm (GeneQuant; Amersham Pharmacia Biotech, Uppsala, Sweden). Absorbance ratios of 280 to 260 nm were consistently more than 1.8 for all samples.

DNA Arrays and Interrogations

Rat cDNA expression arrays (Atlas 1.2; Clontech Laboratories, Palo Alto, CA) were used in this work to study gene expression patterns in corneas after PRK. Each DNA array was manufactured by spotting a nylon membrane with cDNA fragments representing 1176 known genes in addition to three plasmid and bacteriophage DNA sequences, which were included as a negative control, and 9 housekeeping gene sequences, which were included as a positive control. To minimize cross-hybridization and nonspecific binding, cDNA fragments bound on the array were selected, so that they ranged in size from 200 to 600 bp and they did not contain repetitive elements. All clones used to generate the immobilized gene probes were sequenced by the manufacturer to verify identity, and the size of PCR products generated from all clones by gene-specific primers was confirmed by gel electrophoresis.

For DNA array interrogation, the arrays were interrogated with $\alpha^{32}\text{P}$ -labeled cDNA prepared by reverse transcribing pools of total corneal RNA isolated. Using the manufacturer's protocol (Clontech Laboratories), 5 μg total RNA was reverse transcribed with gene-specific primers and reagents provided with the cDNA expression array. The radiolabeled target cDNAs were purified by column chromatography, denatured under basic conditions, neutralized, and added to 10-mL aliquots of hybridization solution (ExpressHyb; Clontech Laboratories) at 68°C, containing 100 $\mu\text{g}/\text{mL}$ heat-denatured and sheared salmon testes DNA (Sigma, St. Louis, MO), to obtain a final probe concentration of 5×10^6 cpm/mL. The radiolabeled cDNA hybridization solutions were applied to prehybridize the array membranes (30 minutes in ExpressHyb with salmon testes DNA at 68°C without labeled probe) and were hybridized overnight at 68°C in glass bottles with constant rotational mixing in a hybridization incubator. After hybridization, the membranes were washed four times with 140 mL $2\times$ SSC (300 mM sodium chloride, 330 mM sodium citrate, pH 7.0), and 1% SDS solution at 68°C for 30 minutes each, followed by one wash for 30 minutes in 140 mL 0.1% SSC and 0.5% SDS, at 68°C. Finally, the membranes were rinsed for 5 minutes in a solution of $2\times$ SSC at room temperature, and exposed to phosphor screens. Each sample was applied to a new membrane—that is, the membranes were not stripped and rehybridized.

Data Analysis

Data Acquisition. Hybridization signals on the cDNA expression arrays were scanned at a pixel resolution of 150 μm and digitized using a phosphorescent imager (Storm PhosphorImager; Molecular Dynamics Inc., Sunnyvale CA). The digitized data were analyzed using the image-analysis software associated with the cDNA arrays (Atlas Image, ver. 1.5; Clontech Laboratories). The signal intensity of each element on the array was determined and corrected by subtracting

background. Background intensities were determined by the image analysis software, based on the average background intensity from nine areas of each array where no DNA was present. The subtraction of the background generated the net signal intensity value for each of the elements on the array.

Global Normalization of Gene Spot Hybridization Intensities. To normalize for experimental variations between samples and membranes, such as differences in signal intensities due to differences in specific activities of labeled target cDNA, hybridization signals on each membrane were normalized using a global normalization method. Briefly, the net signal intensity of each element on the array was determined and multiplied by a scaling factor so that the average intensity of elements on the array was 12,000 phosphorescence image units after scaling. Once the hybridization signals of each element on each array were scaled to the same target intensity, the normalized intensities for each gene were compared from experiment to experiment.

Statistical Analysis

Variance Normalization and Hierarchical Clustering.

To aid in identifying patterns in changes of gene expression, we transformed the data set to normalize the variance observed for each set of gene expression observations. For each gene (element on the array), we calculated its average hybridization intensity and SD after global normalization, as described earlier. We then transformed the data set by subtracting the mean observation value from each observation and dividing by the SD: $(x - \text{mean of } x)/\text{SD of } x$. The mean transformed expression value of each gene in the transformed data set is 0 and the SD for each gene is 1.

Hierarchical clustering of the variance in gene expression over the time course of the experiment was preformed using the Cluster algorithms developed by Eisen et al.¹⁶ Clusters calculates the distances in gene expression space of the expression values observed for each gene and constructs a hierarchical tree, placing the two genes with the most similar expression values next to one another until all branches of the tree are linked. In this study we used the Pearson (uncentered) correlation metric option and the average-linking method to generate the clusters. The output of Cluster was displayed for visual inspection using the TreeView program (available at <http://rana.stanford.edu/software>) also developed by Eisen et al.¹⁶

Cumulative Hypergeometric Probabilities. We calculated the probability of obtaining gene members of similar functional groups within the major clusters obtained by hierarchical clustering. We used the cumulative hypergeometric probability calculation to determine whether any of the 24 functional groups of genes as defined by the manufacturer (Clontech Laboratories) were significantly overrepresented in any of the 14 major clusters identified by hierarchical clustering.¹⁷ In general form, the cumulative hypergeometric probability calculation allows calculation of the probability of observing at least (k) genes from a functional group within a cluster of size (n) given (g) genes to choose from containing f genes of the functional group in question. The probability is calculated using the formula:

$$P = 1 - \sum_{i=0}^{k-1} \frac{\binom{f}{i} \binom{g-f}{n-i}}{\binom{g}{n}}$$

where f is the total number of genes within each functional group provided by the manufacturer (Clontech Laboratories), and g is the total number of genes present on the array (1176). To set the significance level, we applied a Bonferroni-type correction, because we tested for overrepresentation of each functional group. Therefore, the $P = 0.05$ significance level was divided by 24, which is the number of functional groups of genes present on the arrays. Accordingly probabilities greater than 0.00208 (0.05/24) ($-\log = 2.68$) were not considered statistically significant.

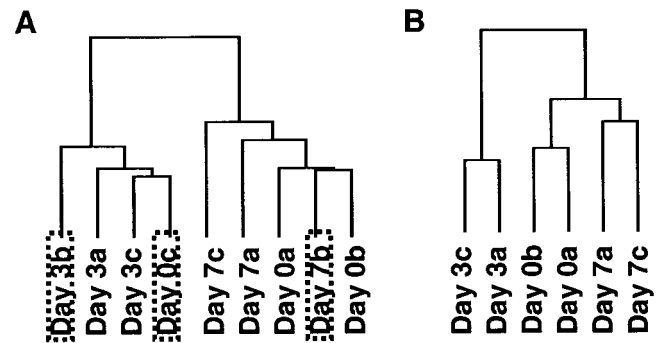


FIGURE 1. Dendrograms showing similarity of expression arrays. (A) The nine expression arrays (day 0 a-c, day 3 a-c, day 7 a-c) were hierarchically clustered with the algorithm Cluster and displayed with TreeView. The relatedness of the arrays is denoted by the distance to the node linking the arrays. *Boxed rectangles*: outliers. (B) Dendrogram of relatedness of arrays after removing the outliers boxed in (A).

K-Mean Clustering of Known Wound Healing Genes

Expression patterns of a set of 56 specific genes that are generally thought to play important roles in corneal wound healing based on previous publications were also studied.¹⁸⁻²² The expression patterns of these selected genes were analyzed using k-means clustering in the Cluster software. Briefly, k-means clustering divided the genes into a user-defined number of groups based on the similarities of their expression profiles. In this study, the data were clustered into five groups with 100 cycles. Subsequently, each of the gene clusters generated by k-means was hierarchically clustered as has been described, and the individual expression profiles of each gene within all clusters were plotted on line graphs.

RESULTS

Comparisons of Replicate Membrane cDNA Microarrays

DNA array experiments, as with any other type of experiment, are subject to noise and experimental artifact. We designed our experiments and analysis to filter out as much noise and variation due to uncontrolled experimental variables as possible. Three replicate experiments were conducted for each time point so that differences due to the controlled experimental variable, healing through a time course, would be reinforced, whereas differences due to uncontrolled experimental variables would be dampened. The first filter we applied to the data was the criterion that replicate measurements should yield similar results.

For each of the three time points, the three globally normalized arrays were compared to assess the uniformity and reproducibility of the data. For this purpose, we used the algorithm in Cluster to determine the similarity of the replicates to themselves at each time point. As shown by the dendrogram in Figure 1A, at each of the three time points, two of the arrays (day 0 a and b, day 3 a and c, day 7 a and c) yielded similar data, whereas one tended to be an outlier. In each case, the data from the outlying arrays (day 0 c, day 3 b, and day 7 b) were excluded from the data set. Therefore, the analyses presented are based on two replicate arrays at each time point (Fig. 1B).

In addition to assessing the overall similarities between replicate arrays, we assessed the gene-by-gene similarity of hybridization signal at each element on the arrays. Accordingly, the mean hybridization intensity from the replicate membranes were calculated. For the data of a gene to be included in the

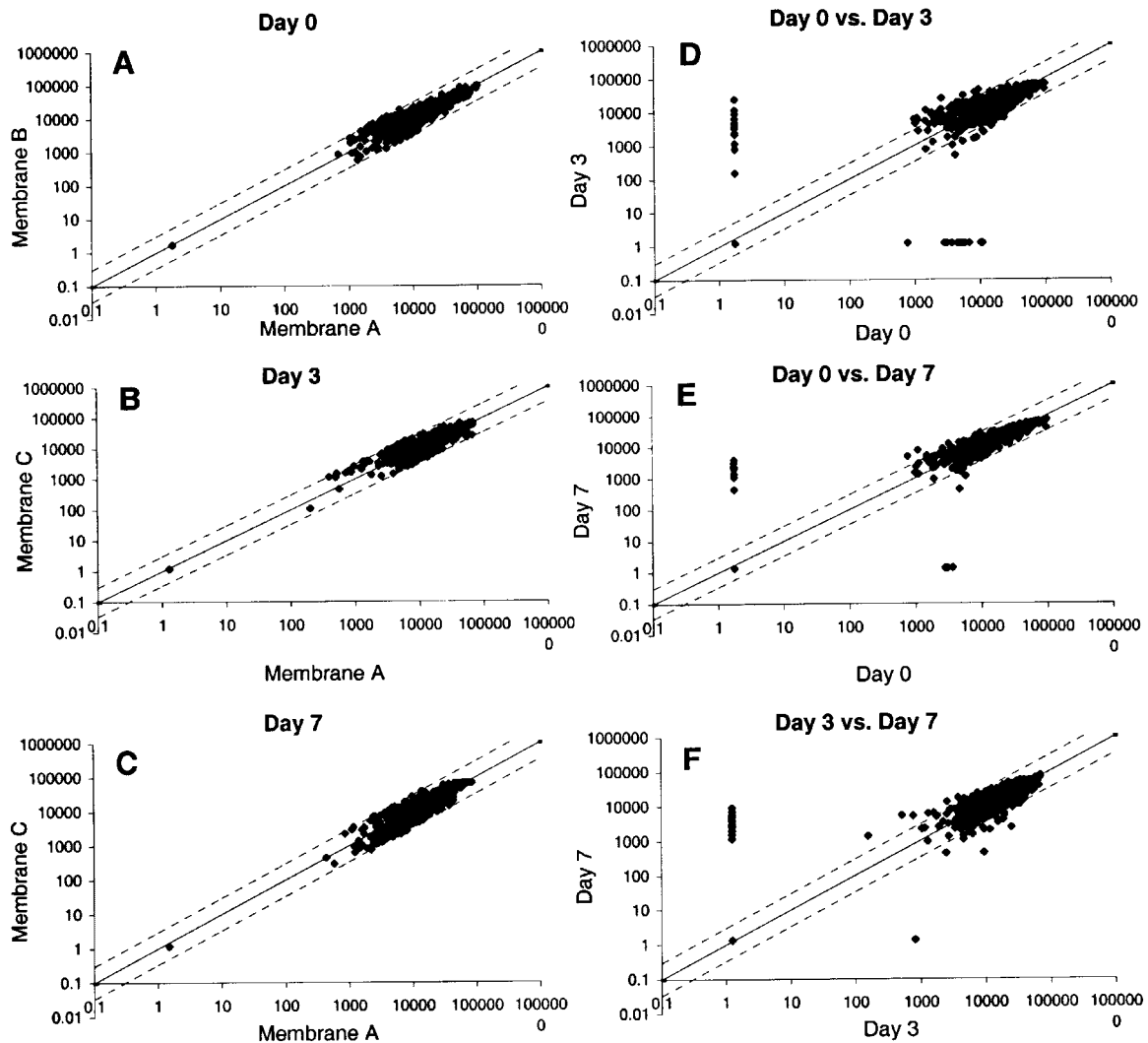


FIGURE 2. Scatterplot analysis of duplicate membranes and pair-wise comparisons between the three time points. Two-dimensional comparison of duplicate membranes for day 0 (A), day 3 (B) and day 7 (C). Scatterplot analyses of expression data for days 0, 3, and 7 are represented as the pair-wise comparisons of day 0 versus day 3 (D), day 0 versus day 7 (E), and day 3 versus day 7 (F) plotted on log-log axes. Each point represents the normalized expression levels of an individual gene on the two days. *Solid line*: predicted line of identity on which a gene would lie if no change occurred in the level of expression between the two days. *Dotted lines*: \pm threefold changes. The perpendicular distance of a point away from the diagonal line represents the degree to which a gene is differentially expressed between the two time points.

data set, both of its hybridization intensities at each of the three time points had to be within 50% of the observed mean hybridization intensity for that time point (e.g., consider two sets of observations for genes X and Y : $X_1 = 90$, $X_2 = 110$ and $Y_1 = 10$, $Y_2 = 190$. The mean observation value for both X and Y is 100. However, data for Y would be excluded, because the observations were outside the window defined by mean \pm 50% of mean [50–150]). After applying this second filter, we arrived at a data set comprising 588 genes. The entire unfiltered data set from all nine arrays can be viewed on our web site, <http://www.obgyn.ufl.edu/research/wound/default.htm>.

So-called housekeeping genes are generally believed to be expressed at similar levels under all conditions, although there are few or no data to support this view. Therefore, we were interested in assessing the expression of the nine housekeeping genes present on the arrays. Of these genes, phospholipase A_2 was severely depressed on day 3 (19-fold) recovering somewhat on day 7, but still depressed (2-fold) when compared with day 0. Expression of myosin heavy chain transcript was depressed twofold at day 3 but recovered by day 7. In contrast, hypoxanthine-guanine phosphoribosyltransferase displayed

higher levels of expression on day 3 (fivefold) recovering somewhat by day 7. Variations in expression of housekeeping genes have also been reported in postnatal rabbit sclera.²⁵ The other six housekeeping genes present on the arrays were expressed at similar levels at all three time points.

Pair-wise Comparisons

Figure 2A, 2B, and 2C show scatterplots of the hybridization intensities for the 588 selected genes at each of the three time points. Scatterplots comparing the levels of expression for each of the 588 genes on each day compared with the other two days are shown in Figures 2D, 2E, and 2F. The dotted lines are the boundaries for the \pm threefold change in levels of gene expression. As seen in Figure 2D, there was a change in the level of gene expression on day 3 compared with day 0. Specifically, expression of 42 genes increased at least threefold, and expression of 28 genes decreased at least threefold on day 3. On day 7 after PRK, the expression of only 20 genes increased at least threefold, and expression of seven genes decreased at least threefold compared with day 0 (Figure 2E).

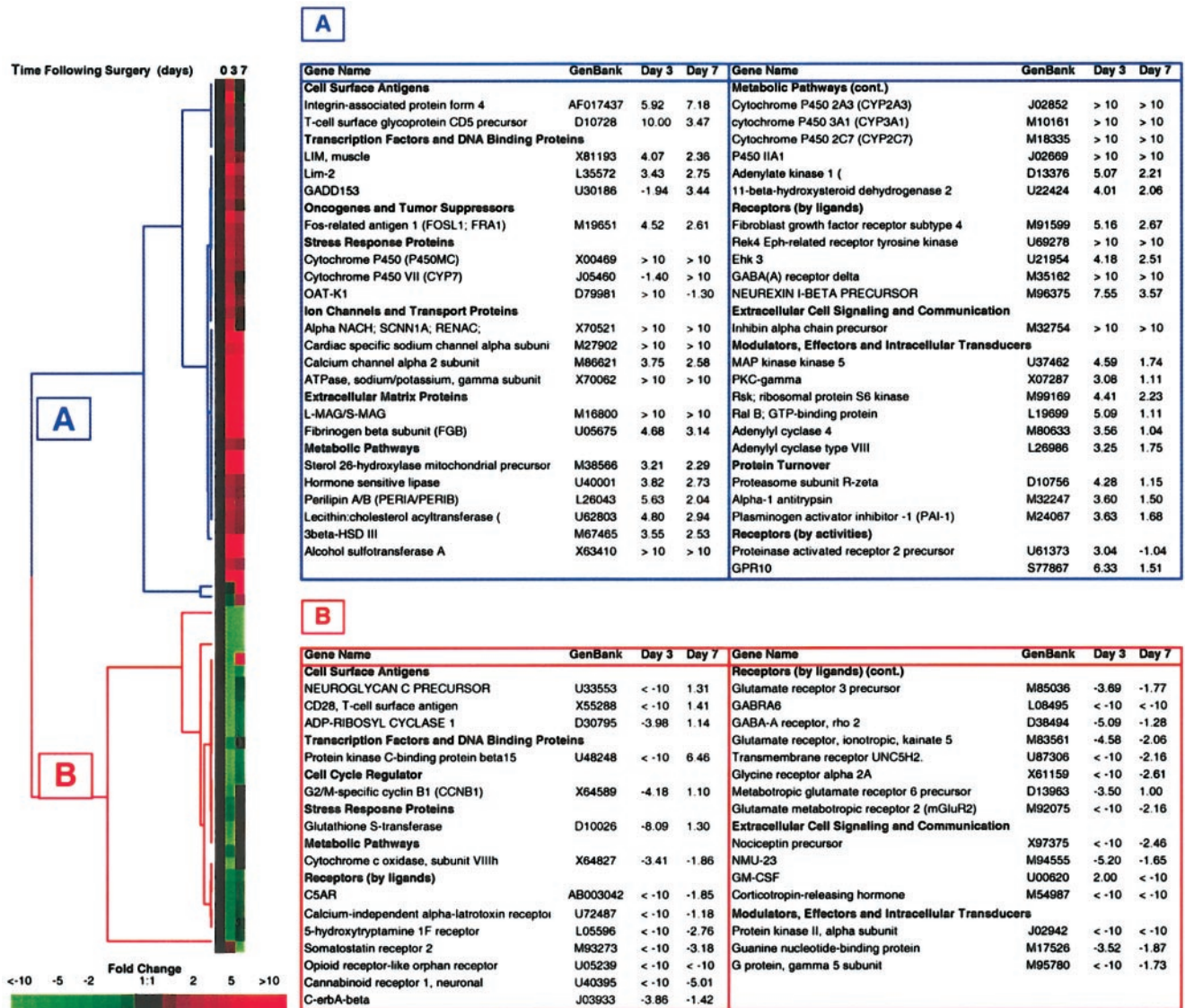


FIGURE 3. Differential expression analysis of detectable genes after PRK. *Left*: expression profiles of 73 differentially expressed genes (\pm threefold change) are graphically depicted as a continuum of color intensity from green (decrease) to black (no change) to red (increase), representing the average multiple of change for each gene across the three time points (days 0, 3, and 7). The level of expression for each gene on day 0 (normal corneas) was used as the reference value for comparison versus days 3 and 7. Two well-defined clusters (A, B) with distinct patterns of expression were identified. *Right*: all the genes belonging to each cluster. The genes are listed by functional categories with the corresponding multiples of changes and the accession numbers (GenBank is provided in the public domain by the National Center for Biotechnology Information, Bethesda, MD, and is available at <http://www.ncbi.nlm.nih.gov/Genbank>).

Comparison of the multiple increase in expression between days 3 and 7 after PRK showed an intermediate result, with 24 genes having increased expression on day 7 compared with their levels on day 3, whereas expression of 16 genes decreased at least threefold (Figure 2F). Overall, the scatterplots indicate that the level of expression of 70 genes increased or decreased at least threefold on day 3 after PRK, compared with noninjured normal corneas. By day 7, expression of many of these genes had returned to levels similar to those of day 0. Twenty-seven genes were differentially expressed on day 7 at more than three times the level on day 0.

Hierarchical Cluster Analysis

Analysis of the 588 selected genes identified 73 genes with expression that varied more than threefold at at least one time point after PRK when compared with the level of expression

observed before injury. Figure 3 shows the result of hierarchical clustering of these 73 genes at each of the three time points. Because the level of expression of each gene on day 0 was set as the reference value for that gene, ratios were 1 in all cases. If a gene decreased in expression compared with the reference condition, its expression was plotted in green, whereas, if its expression increased, its value was plotted in red. The expression profile of each gene was graphically depicted as a continuum of color luminescence from saturated green (<10-fold decrease) to black (no change) to saturated red (>10-fold increase), representing the average multiple of change for each gene (depicted as a horizontal color bar) across the three time points (days 0, 3, and 7). Visual inspection of the TreeView image in Figure 3 indicates that there were two well-defined clusters of genes: A and B. Of the 44 genes in cluster A, 42 are upregulated and only 2 showed a

decrease in level of expression at day 3. At day 7, 38 remain upregulated and only 6 showed a level of expression similar to that at day 0. In cluster B, 28 genes showed a decrease in level of expression, whereas only one showed a very slight increase at day 3. By day 7, 22 remained downregulated, 6 returned to levels of expression similar to day 0, and 1 showed an increase in expression. In summary, most of the genes in cluster A showed an increase in level of expression at day 3 and remained upregulated at day 7, whereas almost all the genes in cluster B showed a decrease in expression at day 3 and remain downregulated at day 7.

A more in-depth look into clusters A and B reveals the prevalence of certain functional groups in those clusters. Very interesting patterns emerged from this observation. For example, cluster A contained 11 genes from the metabolic pathways functional group, whereas cluster B contains only 1. This may reflect a general increase in metabolism during healing of the corneal injuries. Included in the group in cluster A from the metabolic pathways group were four cytochrome P450 genes (*2A3*, *3A1*, *2C7*, and *2A1*). Also of importance is the presence in cluster A of three genes from the protein turnover group including the general inhibitor of serine proteases, α -1 antitrypsin, and the more specific serine protease inhibitor, plasminogen activator inhibitor-1 (*PAI-1*) along with the genes for fibrinogen β subunit and myelin-associated glycoprotein precursor belonging to the extracellular matrix proteins functional group. Similarly, cluster B contains 15 genes from the receptor (by ligands) functional group, whereas cluster A contains only 5. Genes with decreased expression from the receptor (by ligands) group in cluster B included two subunits of the γ -aminobutyric acid (GABA) receptor, glutamate metabotropic receptors 2 and 6, as well as glutamate receptor 3 and somatostatin receptor 2. Figure 2 (right) shows a complete list of all the genes from clusters A and B with their respective multiples of increases and functional group designation.

It is important to point out that the justification for identifying genes that show large increases or decreases in expression (i.e., threefold induction or reduction) after PRK ablation is based on the assumption that genes that show large changes in expression play important roles in corneal wound healing. However, it is also important to examine expression of genes that show relatively smaller changes in expression, because it is likely that many genes that play key roles in regulating corneal wound healing do so by regulating expression of other genes or by altering the activities of cascading pathways. Therefore, small changes in expression of some key regulator genes can generate effects in cells that are amplified many times more than the actual change in their level of transcription. To better evaluate genes with smaller changes in expression, we analyzed the changes in expression patterns of all 588 genes by variance normalization and hierarchical clustering analysis.

Variance Normalization and Hierarchical Clustering of the Selected 588 Genes

Variance normalization was used to detect similarities in the expression patterns of genes that show the same pattern of variation in their gene expression profiles. After variance normalization, the transformed data set for each gene has a mean of 0 and a SD of 1. After hierarchical cluster analysis of the variance-normalized data set, genes that are expressed at different levels but that vary in similar ways in their expression pattern cluster together. Hierarchical clustering of the variance-normalized expression of all 588 genes identified 10 clusters of genes (Figs. 4) with similar expression patterns. The cluster sizes ranged from 18 genes for cluster F to 109 genes for cluster A. The correlation coefficients of the expression pat-

terns of the genes in each of the 10 clusters were greater than 0.9. Genes that show similar patterns of expression may be regulated by common regulatory networks.

The expression of 213 genes increased by day 3 and then decreased by day 7 (clusters E-H). Conversely, 255 genes decreased expression by day 3 (clusters A-D), and of those 223 genes increased expression by day 7 (clusters A, C, D). The genes that were upregulated included those involved in translation, metabolic pathways, protein turnover, trafficking and targeting, and immune system proteins among others. Most of the genes that were downregulated belonged to the extracellular cell signaling and communication, receptors (by ligands), modulators, effectors and intracellular transducers, and cell surface antigens groups. The complete database of the expressed genes and their levels and patterns of expression is available in our Web site at the address provided earlier.

Cumulative Hypergeometric Probability

Genes that are functionally related tend to be governed by the same or overlapping sets of genetic regulatory networks. The utility of cluster analysis with respect to gene expression data is based on the tacit assumption that genes subject to the same genetic regulatory networks cluster in gene expression space. We sought to determine whether any of the clusters we identified in this study were overrepresented by genes of a particular functional group. The arrays used in this study (Clontech Laboratories) contained 1176 genes in total that were classified into 24 functional groups. We calculated the cumulative hypergeometric probability of obtaining at least the number of genes of each functional group in each of the 10 clusters we obtained. Because we tested each of the 24 functional categories for overrepresentation, we applied a Bonferroni-type correction ($0.05/24$) to the standard 0.05 significance level. Thus, probabilities lower than 0.00208 ($-\log = 2.68$) were considered statistically significant. As shown in Table 1, five functional groups were significantly overrepresented in 7 of the 10 clusters identified in this study (transcription factors and DNA-binding proteins; modulators effectors, and intracellular transducers; cell cycle regulators; oncogenes and tumor suppressors; and metabolic pathways). For example, 24 genes belonging to the transcription factors and DNA-binding proteins functional group are present on the 1176 gene array. Nine of the 71 genes in cluster J of Figure 3 are members of the transcription factors and DNA-binding proteins functional group. The probability of a cluster of 71 genes' (the size of cluster J) containing at least 9 genes from the transcription factors and DNA-binding proteins functional group, which totals 24 genes on the array of 1176 genes, is 4.0×10^{-6} ($-\log = 5.39$). Thus, transcription factors and DNA binding proteins are overrepresented in cluster J at levels higher than would be expected by mere chance. Similarly, 6 of the 83 genes in cluster D are members of the cell cycle regulators functional group. Cell cycle regulators functional group consists of a total of 19 genes present on the array. The probability of finding 6 of the 19 cell cycle regulator genes among the 83 genes in the cluster is $P = 1.3 \times 10^{-3}$ ($-\log = 2.88$). The chance probability of finding 16 of the 125 metabolic pathways genes among the 68 genes in cluster G is even lower, at $P = 1.2 \times 10^{-3}$ ($-\log = 2.91$). Therefore, the general pattern of gene expression shown in cluster G of a peak in expression on day 3, followed by a decrease to levels similar to day 0 corneas, is indicative of a coordinated physiological change in general metabolic pathway activity during the early phase of corneal healing.

TABLE 1. Cumulative Hypergeometric Distribution Probability

Cluster	Genes in Cluster (<i>n</i>)	Functional Group (Total Genes on Array)	Genes from Functional Group In Cluster (<i>n</i>)	<i>P</i> (−log 10)
A	109	Modulators, effectors, and intracellular transducers (196)	36	5.11
B	32	Modulators, effectors, and intracellular transducers (196)	15	4.29
D	83	Transcription factors and DNA-binding Proteins (24) Cell cycle regulators (19) Oncogens and tumor suppressors (49)	8 6 10	3.89 2.88 2.82
F	18	Modulators, Effectors, and Intracellular transducers (196)	9	2.99
G	68	Metabolic pathways (125)	16	2.91
I	49	Metabolic pathways (125)	14	3.52
J	71	Transcription factors and DNA-Binding proteins (24)	9	5.39

growth factor (*VEGF*), fibrinogen- β , and TIMP-2 and -3, and the proteases α_1 -antitrypsin, plasminogen activator inhibitor I and 2A (*PAI-1* and *PAI-2A*). These data suggest that in the early stages of wound healing, there was elevated expression of several genes, such as cytokines and proteases, which probably participate in the inflammatory response and enhance migration of corneal cells.²⁴

Growth factors and their receptors play key roles in corneal wound healing.^{24,25} In this study, levels of TGF β (*TGFB*) mRNA increased on days 3 and 7, whereas levels of *EGF-R* mRNA increased on day 3 and decreased on day 7 which is similar to the upregulation of *TGFA* mRNA reported in rat corneas after epithelial debridement.²⁶ The TGF β system has been implicated in promoting corneal scar formation. Levels of *TGFB1* (cluster K5A), and *TGFBIR* (cluster K5B) mRNAs increased at days 3 and 7, which generally agrees with studies that have reported increased levels of TGF β s and receptor mRNAs and protein in rat corneas after PRK.^{6,15,27} Expression of the platelet-derived growth factor (PDGF) B chain was in-

creased at day 7, whereas expression of the PDGF A chain decreased, which is consistent with previous studies that found that the PDGF BB dimer was more effective than the PDGF AA dimer in accelerating epithelial regeneration in rabbit corneas after anterior keratectomy.²⁸ Levels of PDGF- α receptor mRNA increased on days 3 and 7 compared with day 0, whereas the levels of PDGF- β receptor mRNA decreased on day 3 then increased on day 7 to a level similar to that on day 0. It is likely that prolonged upregulation of profibrotic growth factors and their receptors is a factor in generation of the subepithelial haze and regression after PRK, and suppression of these growth factors through pharmacologic means may be of therapeutic benefit after PRK.

DISCUSSION

Healing of corneal wounds is a complex process that requires the integrated actions of numerous genes. Corneal scarring can

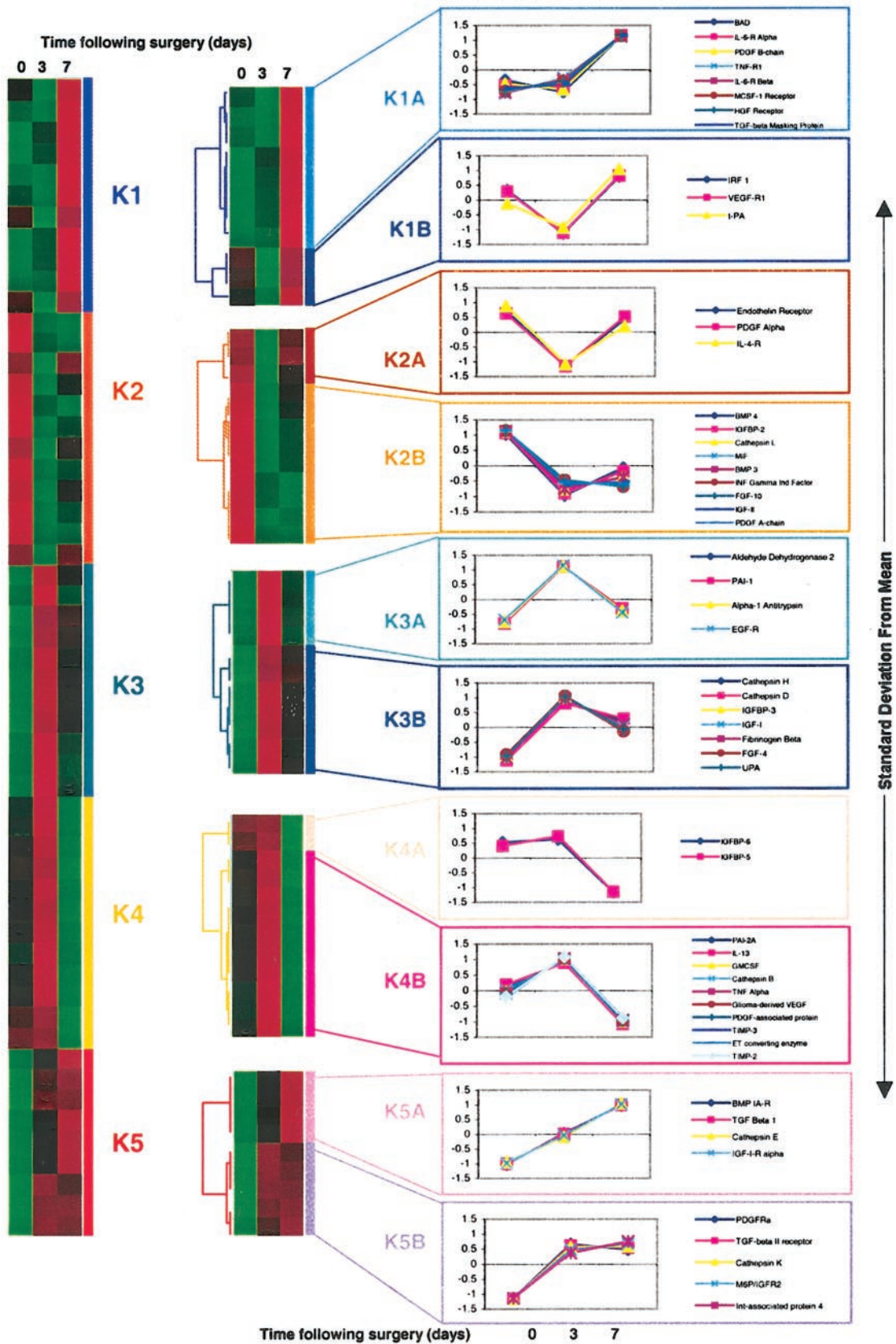


FIGURE 5. Expression profiles of genes previously identified to play roles in corneal wound healing. The expression profiles of 56 genes presumed to play important roles in wound healing were analyzed by k-means and hierarchical clustering. K-means clustering divided the genes into five separate groups based on the similarities of their expression profiles. Each group was then hierarchically clustered into 10 clusters. All 56 genes changed expression levels at days 3 or 7 after PRK, with the expression of 32 genes increasing at day 3. Included in this group are *TNFA*, glioma-derived VEGF, *TGFB1*, *TGFB1R*, *EGFR*, *TIMP2*, *TIMP3*, and α_1 -antitrypsin among others. This indicates that the process of wound healing is most active at day 3 after PRK.

severely limit visual function, yet there are no agents currently available that safely limit the development of corneal haze. With the increase in PRK and LASIK procedures, the need to limit corneal scarring is not restricted to trauma or infection. Thus, there is a need to better understand which factors induce corneal haze and to design agents that selectively reduce their action. Because the phenotype of a tissue is the manifestation of the expressed genotype of the cells, analysis of the expressed genotype of cells within the cornea during wound healing can provide important information about how cells interact at a molecular level. Microarray technology enables a large number of genes to be simultaneously analyzed, which enables broad insights to be made into the general trends and patterns of gene expression.¹⁶ Previous reports of cDNA microarray gene analyses have been performed on *Drosophila*,²⁹ yeast,^{30,31} the nematode *Caenorhabditis elegans*,³² tumor samples,^{31,33} embryonic stem cells,³⁴ or cells in culture.^{35,36} This study is the first report of a cDNA microarray analysis of the cornea after PRK.

In this study, the whole cornea was analyzed during the early phase of wound healing, rather than just the response of a single cell type (e.g., fibroblasts or epithelial cells). One limitation of this experimental design, however, is that the measured values represent the average expression of a gene for all the cells in the cornea and does not permit assessing the contribution of different cell types in the cornea. Another limitation of the use of cDNA microarrays is that the values represent the steady state levels of mRNAs, and there are other well-known cellular mechanisms that regulate the activity of gene products and occur after transcription, including translational regulation, proteolytic activation of latent proteins, or covalent modifications, such as phosphorylation reactions. However, the most common mechanism for regulating the action of a gene product is to regulate the transcription of the gene. In spite of these limitations, cDNA microarray analysis is the most powerful experimental technique to study changes in expression of large numbers of genes. When combined with hierarchical clustering, patterns of gene expression can identify genes that may be regulated by common transcriptional regulatory pathways. It is also important to recognize that corneal cells *in vivo* are exposed, not only to autocrine and paracrine corneal factors such as TGF α ³⁷ and IL-1 α ³⁸ that can regulate corneal gene expression after injury, but also are exposed to exocrine proteins that are present in tears that can influence corneal gene expression. For example, EGF,³⁹ TGF- β 1,⁴⁰ TGF- β 2,⁴¹ VEGF,⁴⁰ and plasmin⁴² are all present in physiologically relevant levels in normal tears, and their levels typically change after a corneal injury. Thus, the results of analysis of corneal gene expression during healing after PRK injury probably reflect regulation by autocrine, paracrine, and exocrine factors.

Analyzing the cornea at days 0, 3, and 7 after PRK provided insight into the magnitude and time course of changes in gene expression at different stages of corneal wound healing. Microarray analysis of the expression of the selected 588 genes showed that at day 3, when the epithelial injury was not fully closed, 70 genes were differentially expressed (a \pm threefold change) compared with day 0, which shows that a relatively large number of genes were induced or repressed at this early stage of corneal healing. At day 7 when the epithelial wound typically had closed but the stromal injury was still actively healing, 27 genes were differentially expressed (\pm threefold change) compared with day 0. Thus, there was a greater change in overall gene expression on day 3 compared with day 7 when the number of differentially expressed (\pm threefold change) genes decreased by 62% (70 vs. 27 differentially expressed genes). This suggests that corneal cells were undergo-

ing more phenotypic changes during the early phase of healing (day 3) than during the later phase (day 7).

A more extensive analysis of the 73 genes that had a threefold change or more in expression revealed important observations about the patterns of gene expression during corneal wound healing. A substantial number of metabolic pathway genes increased in expression threefold or more at day 3, including genes for metabolic enzymes, such as lipases (e.g., hormone sensitive lipase), dehydrogenases (e.g., 11- β -hydroxysteroid dehydrogenase 2), and transferases (e.g., lecithin-cholesterol acyltransferase, alcohol sulfotransferase A). The up-regulation of these genes is probably a result of an increased need for cellular energy during wound healing, to allow cells to proliferate, migrate, and synthesize extracellular matrix. A total of 11 metabolic pathway genes showed at least a threefold increase at day 3. Also, three genes belonging to the protein turnover functional group had induction of threefold or more at day 3 (α ₁-antitrypsin, *PAI-1*, and proteasome subunit R- ζ). Turnover of proteins by specific proteases is essential during wound healing, with proteases playing essential roles in migration, angiogenesis, apoptosis, and extracellular matrix formation.

Among the most striking observations was the extreme similarity in the pattern of expression of the genes within the same hierarchical cluster in Figure 4. Presented on the right side the figure are all the line plots showing the expression profiles for all the genes within each cluster. The correlation coefficients from these clusters ranged from 0.930 to 0.997 suggesting that the genes within each cluster are "moving together" through gene expression space and therefore could be controlled by common regulatory elements. This is an important finding, because it could help in the understanding of the molecular environment during corneal wound healing.

In summary, this is the first report of cDNA microarray analysis performed on healing rat corneas after PRK. The selected 588 genes self-assembled into 10 clusters, based on the similarities of their expression profiles. The grouping of these genes into the same clusters suggests that these genes interact with each other and/or may have common elements of regulation during the process of corneal wound healing. In addition, the expression patterns of 56 genes that are believed to play a part in corneal wound healing provided insight that showed increases on days 3 and/or 7. To further expand our understanding of corneal wound healing, experiments are currently being conducted to examine the expression of 12,626 genes and expressed sequence tags (ESTs) over a broader time course after PRK.

Acknowledgements

The authors thank Clear Vision (Lakewood, CO) for donation of the excimer laser.

References

- Schultz GS. Modulation of corneal wound healing. In: Krachmer JH, Mannis MJ, Holland EJ, eds. *Cornea: Fundamentals of Cornea and External Disease*. Vol. 1. St. Louis: Mosby; 1997:183-198.
- Cameron JD. Corneal response to injury. In: Krachmer JH, et al., eds. *Cornea: Fundamentals of Cornea and External Disease*. Vol. 1. St. Louis: Mosby; 1997:163-182.
- O'Kane S, Ferguson MW. Transforming growth factor betas and wound healing. *Int J Biochem Cell Biol*. 1997;29:63-78.
- Power WJ, Kaufman AH, Merayo-Llodes J, Arranategui-Correa V, Foster CS. Expression of collagens I, III, IV, and V mRNA in excimer wounded rat cornea: analysis by semi-quantitative PCR. *Curr Eye Res*. 1995;14:879-886.
- Wilson SE, Chen L, Mohan RR, Liang Q, Liu J. Expression of HGF, KGF, EGF and receptor messenger RNAs following corneal epithelial wounding. *Exp Eye Res*. 1999;68:377-397.

6. Kaji Y, Mita T, Obata H, et al. Expression of transforming growth factor beta superfamily and their receptors in the corneal stromal wound healing process after excimer laser keratectomy [letter]. *Br J Ophthalmol*. 1998;82:462-463.
7. Sundar RN, Geiss MJ, Fantes M, et al. Healing of excimer laser ablated monkey corneas: an immunohistochemical evaluation. *Arch Ophthalmol*. 1990;108:1604-1610.
8. Azar DT, Hahn TW, Jain S, Yeh YC, Steller-Stevensen WG. Matrix metalloproteinases are expressed during wound healing after excimer laser keratectomy. *Cornea*. 1996;15:18-24.
9. Ye HQ, Azar DT. Expression of gelatinases A and B, and TIMPs 1 and 2 during corneal wound healing. *Invest Ophthalmol Vis Sci*. 1998;39:913-921.
10. Lu PC, Ye H, Maeda M, Azar DT. Immunolocalization and gene expression of matrilysin during corneal wound healing. *Invest Ophthalmol Vis Sci*. 1999;40:20-27.
11. Ye HQ, Maeda M, Yu FS, Azar DT. Differential expression of MT1-MMP (MMP-14) and collagenase III (MMP-13) genes in normal and wounded rat corneas. *Invest Ophthalmol Vis Sci*. 2000;41:2894-2899.
12. Azar DT, Pluznik D, Jain S, Khoury JM. Gelatinase B and A expression after laser in situ keratomileusis and photorefractive keratectomy. *Arch Ophthalmol*. 1998;116:1206-1208.
13. Ljubimov AV, Alba SA, Burgeson RE, et al. Extracellular matrix changes in human corneas after radial keratotomies. *Exp Eye Res*. 1998;67:265-272.
14. O'Brien TP, Li Q, Ashraf MF, Matteson DM, Stark WJ, Chan CC. Inflammatory response in the early stages of wound healing after excimer laser keratectomy. *Arch Ophthalmol*. 1998;116:1470-1474.
15. Chen C, Michelini-Norris B, Stevens S, et al. Measurement of mRNAs for TGFβs and extracellular matrix proteins in corneas of rats after PRK. *Invest Ophthalmol Vis Sci*. 2000;41:4108-4116.
16. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA*. 1998;95:14863-14868.
17. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. *Nat Genet*. 1999;22:281-285.
18. Gillitzer R, Goebeler M. Chemokines in cutaneous wound healing. *J Leukoc Biol*. 2001;69:513-521.
19. Rumalla VK, Borah GL. Cytokines, growth factors, and plastic surgery. *Plast Reconstr Surg*. 2001;108:719-733.
20. Lawrence WT. Physiology of the acute wound. *Clin Plast Surg*. 1998;25:321-340.
21. Bennett NT, Schultz GS. Growth factors and wound healing: biochemical properties of growth factors and their receptors. *Am J Surg*. 1993;165:728-737.
22. Bennett NT, Schultz GS. Growth factors and wound healing. II: role in normal and chronic wound healing. *Am J Surg*. 1993;166:74-81.
23. Moe TK, Ziliang J, Barathi A, Beuerman RW. Differential expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β-actin and hypoxanthine phosphoribosyltransferase (HPRT) in postnatal rabbit sclera. *Curr Eye Res*. 2001;23:44-50.
24. Imanishi J, Kamiyama K, Iguchi I, Kita M, Sotozono C, Kinoshita S. Growth factors: importance in wound healing and maintenance of transparency of the cornea. *Prog Retinal Eye Res*. 2000;19:113-129.
25. Schultz G, Chegini N, Grant M, Khaw P, MacKay S. Effects of growth factors on corneal wound healing. *Acta Ophthalmol*. 1992;202(suppl):60-66.
26. Zieske JD, Takahashi H, Hutcheon AE, Dalbone AC. Activation of epidermal growth factor receptor during corneal epithelial migration. *Invest Ophthalmol Vis Sci*. 2000;41:1346-1355.
27. Mita T, Yamashita H, Kaji Y, et al. Functional difference of TGF-beta isoforms regulating corneal wound healing after excimer laser keratectomy (Letter). *Exp Eye Res*. 1999;68:513-519.
28. Stern ME, Waltz KM, Buererman RW, et al. Effect of platelet-derived growth factor on rabbit corneal wound healing. *Wound Repair Regeneration*. 1995;3:59-65.
29. Bryant Z, Subrahmanyam L, Tworoger M, et al. Characterization of differentially expressed genes in purified Drosophila follicle cells: toward a general strategy for cell type-specific developmental analysis. *Proc Natl Acad Sci USA*. 1999;96:5559-5564.
30. Lopez MC, Baker HV. Understanding the growth phenotype of the yeast *gcr1* mutant in terms of global genomic expression patterns. *J Bacteriol*. 2000;182:4970-4978.
31. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406:747-752.
32. Hill AA, Hunter CP, Tsung BT, Tucker-Kellogg G, Brown EL. Genomic analysis of gene expression in *C. elegans*. *Science*. 2000;290:809-812.
33. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403:503-511.
34. Kelly DL, Rizzino A. DNA microarray analyses of genes regulated during the differentiation of embryonic stem cells. *Mol Reprod Dev*. 2000;56:113-123.
35. Iyer VR, Eisen MB, Ross DT, et al. The transcriptional program in the response of human fibroblasts to serum [see comments]. *Science*. 1999;283:83-87.
36. Collier HA, Grandori C, Tamayo P, et al. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc Natl Acad Sci USA*. 2000;97:3260-3265.
37. Khaw PT, Schultz GS, MacKay SL, et al. Detection of transforming growth factor-alpha messenger RNA and protein in human corneal epithelial cells. *Invest Ophthalmol Vis Sci*. 1992;33:3302-3306.
38. Strissel KJ, Rinehart WB, Fini ME. Regulation of paracrine cytokine balance controlling collagenase synthesis by corneal cells. *Invest Ophthalmol Vis Sci*. 1997;38:546-552.
39. van Setten G-B, Tervo T, Viinikka L, Perheentupa J, Tarkkanen A. Epidermal growth factor in human tear fluid: a minireview. *Int Ophthalmol*. 1991;15:359-362.
40. Vesaluoma M, Teppo AM, Gronhagen-Riska C, Tervo T. Release of TGF-beta 1 and VEGF in tears following photorefractive keratectomy. *Curr Eye Res*. 1997;16:19-25.
41. Kokawa N, Sotozono C, Nishida K, Kinoshita S. High total TGF-beta 2 levels in normal human tears. *Curr Eye Res*. 1996;15:341-343.
42. van Setten G-B, Salonen E-M, Vaheri A, et al. Plasmin and plasminogen activator activates in tear fluid during corneal wound healing after anterior keratectomy. *Curr Eye Res*. 1989;8:1293-1298.