



OPEN

RNA sequencing of corneas from two keratoconus patient groups identifies potential biomarkers and decreased NRF2-antioxidant responses

Vishal Shinde^{1,6}, Nan Hu^{1,6}, Alka Mahale², George Maiti¹, Yassine Daoud³, Charles G. Eberhart⁴, Azza Maktabi², Albert S. Jun³, Samar A. Al-Swailem² & Shukti Chakravarti^{1,5}✉

Keratoconus is a highly prevalent (1 in 2000), genetically complex and multifactorial, degenerative disease of the cornea whose pathogenesis and underlying transcriptomic changes are poorly understood. To identify disease-specific changes and gene expression networks, we performed next generation RNA sequencing from individual corneas of two distinct patient populations - one from the Middle East, as keratoconus is particularly severe in this group, and the second from an African American population in the United States. We conducted a case: control RNA sequencing study of 7 African American, 12 Middle Eastern subjects, and 7 controls. A Principal Component Analysis of all expressed genes was used to ascertain differences between samples. Differentially expressed genes were identified using Cuffdiff and DESeq2 analyses, and identification of over-represented signaling pathways by Ingenuity Pathway Analysis. Although separated by geography and ancestry, key commonalities in the two patient transcriptomes speak of disease - intrinsic gene expression networks. We identified an overwhelming decrease in the expression of anti-oxidant genes regulated by NRF2 and those of the acute phase and tissue injury response pathways, in both patient groups. Concordantly, NRF2 immunofluorescence staining was decreased in patient corneas, while KEAP1, which helps to degrade NRF2, was increased. Diminished NRF2 signaling raises the possibility of NRF2 activators as future treatment strategies in keratoconus. The African American patient group showed increases in extracellular matrix transcripts that may be due to underlying profibrogenic changes in this group. Transcripts increased across all patient samples include Thrombospondin 2 (*THBS2*), encoding a matricellular protein, and cellular proteins, *GAS1*, *CASR* and *OTOP2*, and are promising biomarker candidates. Our approach of analyzing transcriptomic data from different populations and patient groups will help to develop signatures and biomarkers for keratoconus subtypes. Further, RNA sequence data on individual patients obtained from multiple studies may lead to a core keratoconus signature of deregulated genes and a better understanding of its pathogenesis.

Keratoconus (KCN) is a condition where the cornea develops bilateral ectasia, becomes progressively thin and protrudes conically. The patient develops astigmatism, myopia, corneal scarring, with eventual loss of vision¹⁻⁵. The major form of keratoconus is asyndromic, where the cornea alone is affected. However, syndromic types of KCN also exist, and are associated with Down, Leber congenital amaurosis, Turner, Marfan and Ehlers-Danlos syndromes⁶. Isolated KCN affects individuals in adolescence with an incidence and prevalence of 13.3/100,000

¹Department of Ophthalmology, NYU Langone Medical Center, New York, NY, USA. ²King Khaled Eye Specialist Hospital, Riyadh, Saudi Arabia. ³Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁴Ophthalmology and Oncology Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁵Department of Pathology, NYU Langone Medical Center, New York, NY, USA. ⁶These authors contributed equally: Vishal Shinde and Nan Hu. ✉e-mail: Shukti.Chakravarti@nyulangone.org

and 265/100,000, respectively, in subjects of European ancestry^{7,8}, with a higher reported prevalence in populations from India, Pakistan and Saudi Arabia^{9,10}. The prognosis of KCN depends on the severity and management of the disease¹¹. There are no curative treatments but UV crosslinking of collagen fibrils is used to slow the progressive loss of corneal shape^{12,13}. Other disease management options include contact lens use, corneal inserts or “intacs” and, in severe cases, cornea transplantation.

Both the corneal epithelium and stroma harbor pathogenic changes in KCN, and these have been considered as the initial causal steps^{14–16}. Microscopy and biochemical studies reported various changes as important disease-specific pathologies. The epithelium appears thickened in some studies, with breaks in Bowman’s membrane. The thinner stroma shows reduced cell density^{17–19} as well as altered ultrastructure of collagen fibrils and proteoglycans^{20,21}. Studies also cite increases in acid esterases, acid phosphatases, acid lipases, cathepsin B and G, matrix metalloproteinases and decreases in tissue inhibitors of metallo-proteases (TIMP)^{22–27}. Recent proteomic studies, including from this laboratory, show changes consistent with loss of epithelial integrity, altered stromal ECM proteins, reduced hydroxylated collagens, and changes in integrated stress response and TGF β signaling^{28–32}. These diverse changes are consistent with KCN being inherited as a complex multifactorial disease with genetic and environmental influences that ultimately cause loss of corneal integrity^{6,33,34}. Genetic etiologies have been pursued through genome-wide association studies of central corneal thickness and keratoconus, implicating *VSX1*, *LOX*, *ZNF469*, *SOD1*, *TGFBI*, *FOXO1*, *FNDC3B*, *ZFN469*, *COL5A1* and *AKAP13*^{6,33,35}. A number of environmental stressors, eye rubbing, contact lens wear, dry climate and UV exposure, are suspected to induce or exacerbate the disease as well^{36–38}. In addition, biomarkers for keratoconus can help early diagnosis or predict disease progression. Thus, a recent study has identified prolactin-inducible protein as a promising predictive biomarker for keratoconus³⁹. RNA sequence data can serve a valuable starting point for development of additional biomarkers for keratoconus.

We opted to use next generation sequencing of total RNA from patient corneas to provide an unbiased and comprehensive view of the transcriptome and disease pathogenesis. To date, two studies have reported transcriptomic analyses of corneas obtained during transplantation^{40,41}. The first, compared KCN with non-KCN corneas in patients undergoing grafting for other conditions, and the second compared a smaller set of KCN corneas with control donor corneas from deceased individuals. Although these studies cited changes in members of the WNT, TGF β and ECM pathways, deeper regulatory networks were not identified.

Here we performed transcriptomic analyses of all protein-coding genes on patients from two distinct populations – African American and Middle Eastern subjects from United States and Saudi Arabia respectively. We compared these with non-diseased donor corneas, to identify common, generalizable pathogenic features. Keratoconus is particularly severe in the Middle East⁹, and therefore, we included patients from this region to increase the likelihood of finding major pathogenic changes. We identified genes related to acute phase tissue injury and NRF2-regulated oxidative stress response as core regulatory networks disrupted in keratoconus. We observed upregulations in ECM genes primarily in African American patients, suggesting underlying profibrogenic changes in this group. We detected a few genes, upregulated across all samples, which include matricellular *THBS2*, and cellular protein-encoding genes *GAS1*, *CASR* and *OTOP2*.

Results

Donor and patient samples. We recruited keratoconus patients of African American ancestry (KC) from Baltimore (17–69 years old) and of Middle Eastern ancestry (KJ) from Saudi Arabia (18–36 years old). (Supplemental Table S1). As controls for the KC samples, we used corneas from deceased African American donors (25–75 years old) for the KC samples. For the KJ corneas, we used European American donor corneas for two reasons. First, corneal tissue donation in Saudi Arabia is extremely rare making them unavailable for research⁴². Second, Middle Eastern individuals are genetically much closer to individuals of European than African ancestry⁴³.

A majority of the samples retained the epithelial layer but were missing all or part of the single-cell layered endothelium. Paraffin sections stained with H&E showed the corneas to have an attached epithelium and a mostly intact stroma. The KJ sample showed an irregular and slightly thickened epithelium, but otherwise intact epithelia and stroma (Fig. 1). This indicated that the quality and epithelial retention were similar for KCN and donor corneas, although the former were placed immediately at -80°C or processed for RNA extraction, while donor corneas were extracted one week after their receipt in Optisol at 4°C . Further analyses of RNA quality, read depth, yield of transcripts and genes showed no difference between donor and KC samples indicating that Optisol-storage did not have a detectable negative impact (discussed below).

RNA quality, sequence quality and patient-donor differences. The average RNA yield and RIN value among all 36 samples was 1.88 μg and 8.02, respectively (Supplemental Table S2). Across all samples, we detected 15,197–17,950 transcripts encoded by 11,286–13,025 unique protein-coding genes per sample with FPKM ≥ 1 , and 6,900–9,155 transcripts encoded by 6,507–8,559 unique protein coding genes per sample at FPKM ≥ 5 (Table 1). For this study, we focused only on coding transcripts to identify gene-level changes between the patient groups. The data have been deposited in Gene Expression Omnibus at NCBI and are accessible through GEO Series accession number GSE151631 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE151631>). To assess sequencing and analysis accuracy, the same RNA preparation for each of the eight samples (7 KC, 1 donor) was sequenced twice and the results show very high correlation between these technical replicates (Supplemental Fig. S1).

To determine the relationships between all 36 samples, we conducted principal component analysis (PCA) using 10,652 expressed genes with FPKM ≥ 5 in at least one sample (Fig. 2). The first principal component (PC1) accounted for 31.2% of the expression variance and separated all control healthy tissues from diseased corneal tissues. The second principal component (PC2) accounted for 14.5% of the expression variance and was largely

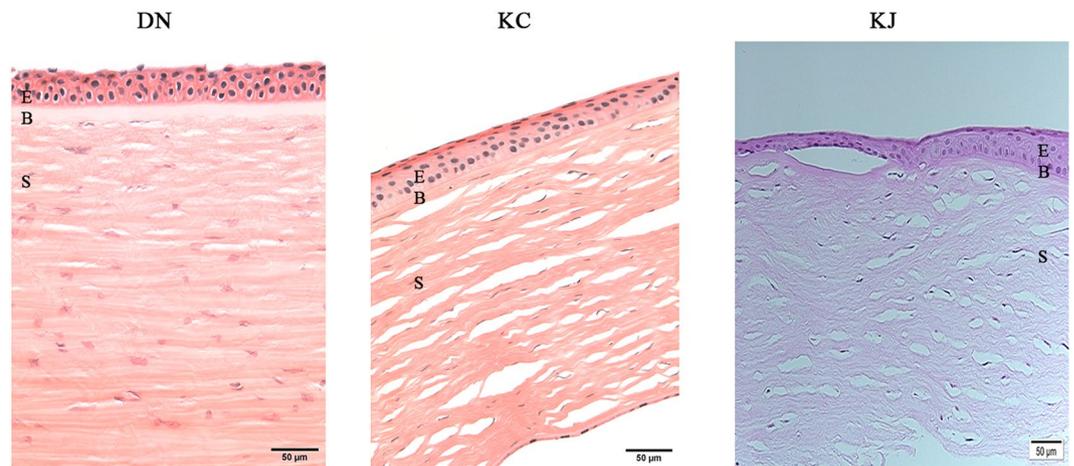


Figure 1. Histology of patient and donor cornea sections. The figure shows similar retention of epithelial layers in DN and KC corneas, only KJ sample showed irregular and thickened epithelial cell layer in H&E staining of paraffin embedded sections. KC: keratoconus cornea from an African American patient; KJ: keratoconus cornea from a Middle Eastern patient; E: Stratified squamous epithelium; B: Bowman's Layer; S: stroma; K: keratocytes.

explained by one outlier (KC406) which was nevertheless consistent with its replicate. A scree plot shows the variance as represented by each principal component (Supplemental Fig. S1B). Pairwise comparison of all samples is shown in Supplemental Table S3. The mean distance between technical duplicates is small (2.57 ± 0.97), while comparisons of cases to controls show ~30-fold larger mean distances: 74.99 ± 27.96 and 73.15 ± 18.36 between the 4 DN (control) versus the 7 KC, and the 3 LE (control) versus the 12 KJ samples, respectively. Additionally, mean distances between cases and controls are 2 and 4.8 standard deviations greater than distances within cases or controls. Repeat PCA using a smaller set of 4,787 differentially expressed genes (DEG) with FPKM ≥ 5 in all samples showed similar relationships albeit with a significantly smaller separation along the minor axis (Supplemental Fig. S2). Keratoconus associated with atopy, hay fever or vernal keratoconjunctivitis, may represent a subtype within keratoconus; in our patient groups there were 4 such cases in the KJ and one in the KC group. However, these five showed no clustering or clear separation from other cases: mean distances between the allergy and non-allergy patients was 11.55 ± 7.53 , and, within the non-allergy group was 14.55 ± 8.26 . In addition, there were no DEGs that were significantly different between the allergy and non-allergy KCN, and larger sample size in the future may reveal gene signatures for these subtypes of KCN (Supplemental Table S4).

Differentially expressed genes (DEGs) in patient and donor groups. We tested expression differences between donor and patient corneas of 13,397 genes^{44,45}, retaining those with $q \leq 0.01$, ≥ 2 or ≤ -2 -fold change and with genes expressed at FPKM ≥ 5 in at least one sample. This filtered set contained 819 (232 upregulated, 587 downregulated) and 993 (213 upregulated, 780 downregulated) unique protein-coding DEGs in the Baltimore KC and Saudi Arabian KJ groups, respectively (Supplemental Table S4). Analysis using a second independent method showed high concordance, 80–81% of the DEGs in each canonical pathway (discussed later) were identical (Supplemental Table S5)⁴⁶.

Upon comparing the DEGs in the two patient groups, we detected 53 commonly up-regulated and 380 commonly down-regulated genes, and an additional 179 and 160 uniquely up-regulated genes in KC and KJ, respectively (Table 2). Only a few genes showed opposite expression trends in the two groups: *LGALS7* was elevated in KJ but decreased in KC, while 12 genes (*CD34*, *CHST6*, *COLGALT2*, *DKK2*, *KDR*, *KERA*, *MAMDC2*, *MME*, *NANOS*, *PDGFD*, *RECK*, *STEAP4*) were decreased in KJ and increased in KC. Thus, patient corneas from two distinct populations demonstrate highly concordant differences. This is further supported by similar distributions of DEGs when categorized by cellular localization and annotated functions of the encoded proteins (Supplemental Fig. S3A,B). All DEGs with FDR adjusted p -value for the two KCN groups are shown in two volcano plots in Fig. 3; we detected greater similarities in down regulated than upregulated genes in the two patient groups (detailed list in Supplemental Table S6). For example, both volcano plots show significant decreases in *ATF3*, *CA3*, *FOSL1*, *PLAUR* and *UBAP1* gene expression. Many of these genes are connected to stress response as discussed later.

Fewer similarities in most increased transcripts in the two patient groups. There are few overlaps in the ten most elevated transcripts between the two groups of samples. Most increased transcripts unique to KC corneas include 7 that encode ECM and matricellular proteins (Fig. 3A, Supplemental Table S6A). By contrast, the ten most increased DEG in KJ samples (Supplemental Table S6B) include genes that regulate inflammation (*S100A9*, *S100A8*, *LGALS9C*, *LCN2*), oxidative deamination of histamine and allergic response (*AOC1*), eye development and chloride channel functions (*CLIC*), and *ADAMTS14*, encoding a collagenase required for the production of assembled mature collagen fibrils⁴⁷.

Sample ID	# Genes with FPKM ≥ 1	# Genes with FPKM ≥ 5	# Transcripts with FPKM ≥ 1	# Transcripts with FPKM ≥ 5
LE2C	12,111	7,648	16,734	8,152
LE4C	11,286	6,613	15,197	6,986
LE6C	11,876	7,615	16,371	8,075
DN5487	11,411	6,507	15,440	6,900
DN5511	11,601	7,000	15,972	7,565
DN76411	11,864	7,182	16,279	7,695
DN76507	11,799	7,321	16,228	7,923
KC276_2	12,265	7,843	16,971	8,457
KC366_2	11,591	7,409	15,860	7,918
KC369_2	11,761	7,768	16,240	8,258
KC388_2	11,476	7,457	15,839	8,089
KC395_2	12,442	8,240	17,242	8,738
KC400_2	11,994	7,822	16,509	8,291
KC406_2	12,948	8,462	17,887	9,057
KJ04	12,093	7,751	16,665	8,373
KJ05	11,740	7,868	16,405	8,540
KJ06	11,860	7,867	16,295	8,393
KJ09	11,784	7,563	16,437	8,119
KJ10	12,132	7,773	16,970	8,391
KJ11	11,753	7,659	16,417	8,339
KJ12	12,137	8,056	16,946	8,648
KJ13	11,859	7,972	16,519	8,566
KJ14	11,946	7,971	16,684	8,534
KJ17	11,862	7,925	16,600	8,529
KJ22	11,873	7,767	16,775	8,499
KJ25	12,094	7,925	16,758	8,547
Samples not used for differential expression analysis				
DN373	12,179	7,698	16,976	8,109
DN401	11,749	7,548	16,315	8,074
DN401_2	11,743	7,587	16,337	8,132
KC276	12,342	8,010	17,102	8,583
KC366	11,727	7,608	16,063	8,075
KC369	11,861	7,932	16,394	8,428
KC388	11,722	7,974	16,173	8,637
KC395	12,491	8,310	17,306	8,842
KC400	12,076	8,012	16,711	8,507
KC406	13,025	8,559	17,950	9,155

Table 1. Transcript and gene counts for each patient and control sample.

The few highly upregulated genes shared by both patient groups require further studies as these may harbor promising biomarkers for keratoconus. Thrombospondin 2 (*THBS2/TSP2*), a matricellular protein-encoding gene was significantly increased in both patient groups (Supplemental Table S6A). *THBS2* is involved in cell-ECM interactions and collagen fibrillogenesis as evidenced by thin corneas in *Thbs2* null-mice^{48–50}. Within the eye, it is present in the lens epithelium, injured corneal epithelium and the stroma⁵¹. Other significant DEGs increased in both sample sets, with FPKM ≥ 5 , include genes encoding cellular proteins that regulate growth (*GAS1*), Calcium sensing (*CASR*), exocytosis (*RAB3D*), cytoskeleton (*KRT78*) and proton-channel (*OTOP2*) (Fig. 3A,B, Supplemental Table S4).

In contrast, the ten most decreased DEGs (Supplemental Table S6C) show considerable overlaps between the two patient groups and point to shared pathogenesis. *SLC2A3* encodes a glucose transporter; its decrease may be upstream of nutrient deprivation-related stress. Decreased transcripts for *PLAUR*, involved in plasmin formation, *CHI3L1*, encoding a chitinase-like protein, and possibly a promoter of tissue remodeling, may both be part of a tissue injury response program. The gene encoding cyclic AMP dependent transcription factor ATF3, decreased in both groups (Fig. 3A,B, Supplemental Table S6C), regulates cell proliferation and differentiation, and is a stress response gene that interacts with JUN and FOS; additionally, related genes, *JUND* and *FOLSL1* are both downregulated in the keratoconus samples. Decreased expression of *NR4A1* encoding a steroid-retinoid hormone receptor, known to promote apoptosis may also be part of a dysregulated injury response. Other notable decreases with FPKM ≥ 5 in the KC cases include *SERPIN2*, *PPP1R15A* (*GADD34*), *KRT17*, *APOBEC3A*, *PTGS2*, *GADD45A*, *GADD45B*, *TM4SF1*, *NEDD9* and *PIM1*. Among the KJ samples similar decreases were seen for *PPP1R15A*, *PIM1*, *NEDD9* and *GADD45B* (Supplemental Table S4).

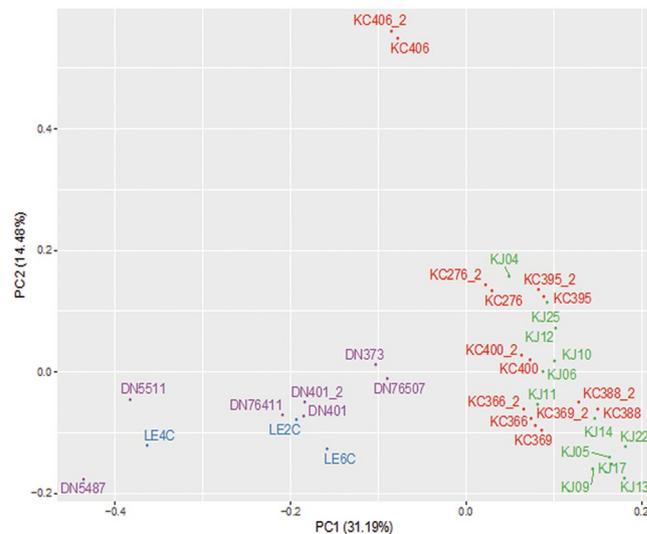


Figure 2. Principal component analysis (PCA) of patient and donor samples. PCA was performed on 19 Keratoconus (7 KC and 12 KJ) and 10 control (7 DN and 3 LE) corneas using 10,652 genes with FPKM value ≥ 5 in at least one sample. Technical duplicates of 7 KC and 1 DN samples (KC*_2 and DN401_2) appear close together. The major separation along PC1 was between KCN and controls. KC and KJ samples cluster together along PC1. African American donor (DN) and KC patient samples are labeled purple and red, respectively; Caucasian donors (LE) are in blue and Middle Eastern patients (KJ) in green. Euclidean distance between samples in a pairwise comparison of samples in PC1 and PC2 is shown in Supplemental Table S3.

	Up-regulated in KJ	No change in KJ	Down-regulated in KJ
Up-regulated in KC	53	167	12
No change in KC	159	12,031	388
Down-regulated in KC	1	206	380

Table 2. Comparison of differentially expressed genes in African American (KC) and Middle Eastern (KJ) patients.

Decreased acute phase and NRF2-mediated oxidative stress in keratoconus. To identify the biological processes potentially important to corneal functions and perturbed in KCN, we used DAVID Bioinformatic Resources and Ingenuity Pathway Analyses, yielding five significant canonical pathways in each patient group (Fig. 4)^{52,53}. First, acute phase (tissue injury) responses and IL-6 signaling related DEGs were significant in both patient groups. Second, NRF2-regulated oxidative stress response DEGs were significantly decreased in the KJ samples, with a similar but non-significant trend in KC samples. Third, two networks, often active in arthritis, that essentially regulate immune signals, inflammation and ECM, were decreased in KJ. Finally, PPAR signaling, RAR and ECM-related DEGs were significant in KC. In the following we examine these pathways and their implications for KCN.

Multiple acute phase and tissue injury-response genes were decreased in both patient groups (Fig. 4). These included *CEBPB/TCF5*, a transcription factor that regulates immune and inflammatory responses, and *IL1A*, a pleiotropic cytokine that regulates hematopoiesis, inflammatory and injury responses. Other decreases from this network included *IL6R*, (interleukin 6 receptor subunit alpha), *IL6ST* (a signal transducer for IL6), *SOCS3*, *SOCS8* (suppressors of cytokine signaling), and *SAA1*, encoding an apolipoprotein.

NRF2 (nuclear factor erythroid2-related factor 2) is a transcription factor that promotes the induction of antioxidants and other components that control reactive oxygen species (ROS) to counteract their harmful effects⁵⁴. NRF2-mediated antioxidant protection is important to the eye and NRF2-boosting treatments are being considered for ocular wound healing and macular degeneration⁵⁵. Multiple antioxidant genes, positively regulated by NRF2, were decreased in KJ, and to a lesser extent in KC (Fig. 5A). These NRF2 target genes include glutathione S-transferase (*GSTM3*), thioredoxin reductase (*TXNRD1*), producers of reducing factors (*GCLC*, *GCLM*), a transcription regulator (*FOSL1*), chaperone-encoding genes (*DNAJA1*, *DNAJB4*, *DNAJB9*, *UBAP1*) involved in protein-folding, degradation and trafficking, and, *HMOX1*, a heme oxygenase that regulates iron metabolism. NRF2 is itself regulated post-transcriptionally, which may explain why we did not detect it by differential transcriptome analyses. Under homeostatic conditions, NRF2 protein is retained in the cytoplasm and maintained at low levels through its rapid turnover by KEAP1, which mediates interactions with CUL3 leading to ubiquitin proteasomal degradation of NRF2⁵⁴. Elevated ROS leads to disassociation of the KEAP1-NRF2 complex, escape of NRF2 from ubiquitination and degradation (Fig. 5B). A few studies have reported expression of NRF2 target genes, or NRF2 itself in cultured corneal cells, but these did not examine NRF2 in tissue layers of the cornea.

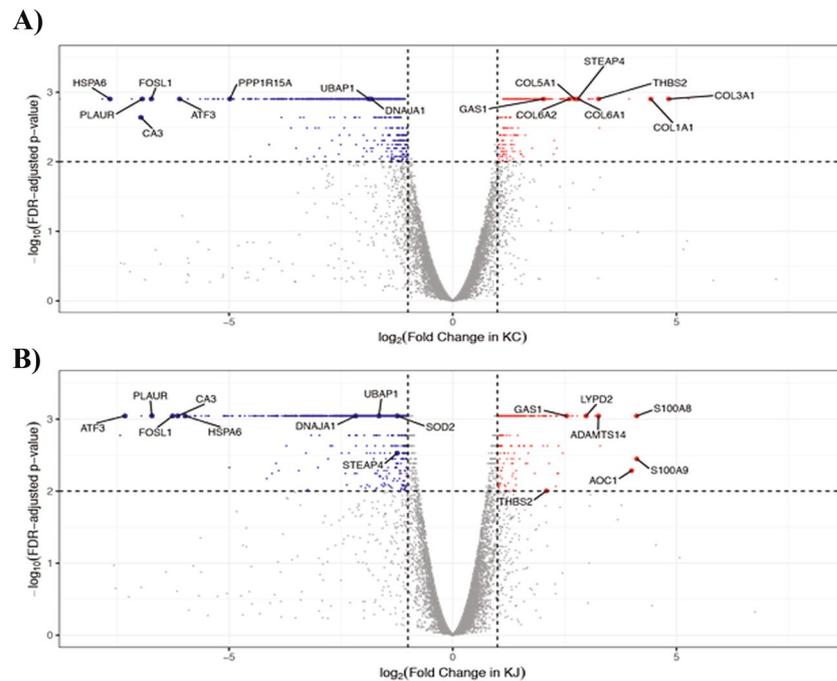


Figure 3. Differentially expressed genes (DEGs) in patient groups. **(A)** DEGs in KC. **(B)** DEGs in KJ. Among successfully tested genes, 8,762 and 8,990 genes have \log_2 (fold change) between -8 and 8 in KC and KJ, respectively. Down-regulated genes (q value ≤ 0.01 , and fold change ≤ -2 , and FPKM ≥ 5 in control or patient groups) are labeled in blue; up-regulated genes (q value ≤ 0.01 , and fold change ≥ 2 , and FPKM ≥ 5 in control or patient groups) are labeled in red; genes not significantly changed are in grey.

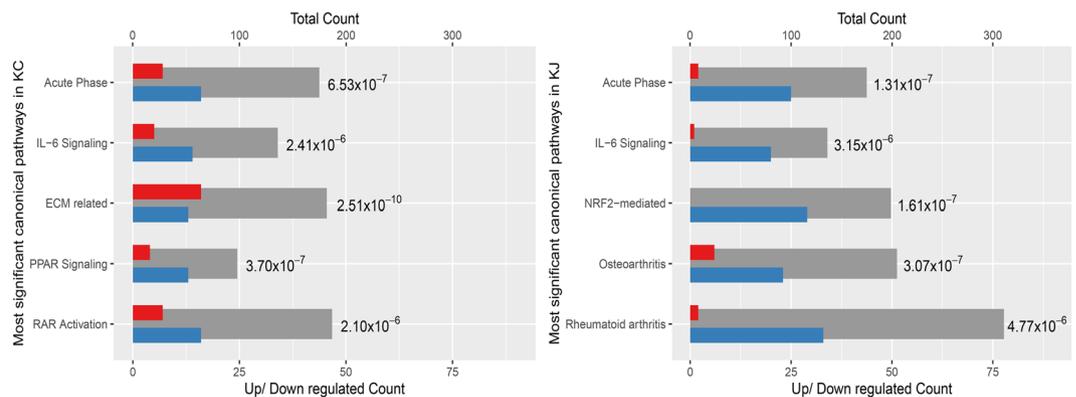


Figure 4. Canonical pathways significantly altered in KC and KJ. Grey bars show total number of genes in the pathway (scale on top X axis), with blue and red bars showing the numbers of down-regulated and up-regulated genes (scale on bottom X axis). The p value next to each bar was calculated for that pathway by the IPA Core Analysis. “Acute Phase” is the Acute Phase Response Signaling pathway; “NRF2-mediated” is NRF2-mediated Oxidative Stress Response pathway; “ECM related” is the Hepatic Fibrosis / Hepatic Stellate Cell Activation pathway; “Rheumatoid Arthritis” is the Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis pathway; and, “Osteoarthritis” includes genes associated with osteoarthritis.

Thus, we localized KEAP1 and NRF2 proteins in control DN and keratoconus corneas by immunofluorescence (IF) staining (Fig. 6A,B) in 3 different DN and KCN corneas (Supplemental Fig. S4A,B). We find that both KEAP1 and NRF2 are primarily present in the corneal epithelial layers. In KCN corneas, KEAP1 staining is generally decreased with areas of focal increase in the epithelium (Fig. 6A). IF staining of NRF2 itself is decreased in KCN; in particular DN corneas show more nuclear NRF2 staining (Fig. 6B). Earlier we reported increased KEAP1 and CUL3 proteins in a comparative proteomics of KCN and donor corneas⁵⁶.

Validations of RNA seq findings. We validated select findings from our transcriptome analyses in three ways: testing RNA expression by qRT-PCR, protein localization and presence by IF staining of donor and keratoconus corneal sections, and against the literature. We selected *ICAM1* (cell adhesion), *TUBB2A*, *TUBB3*

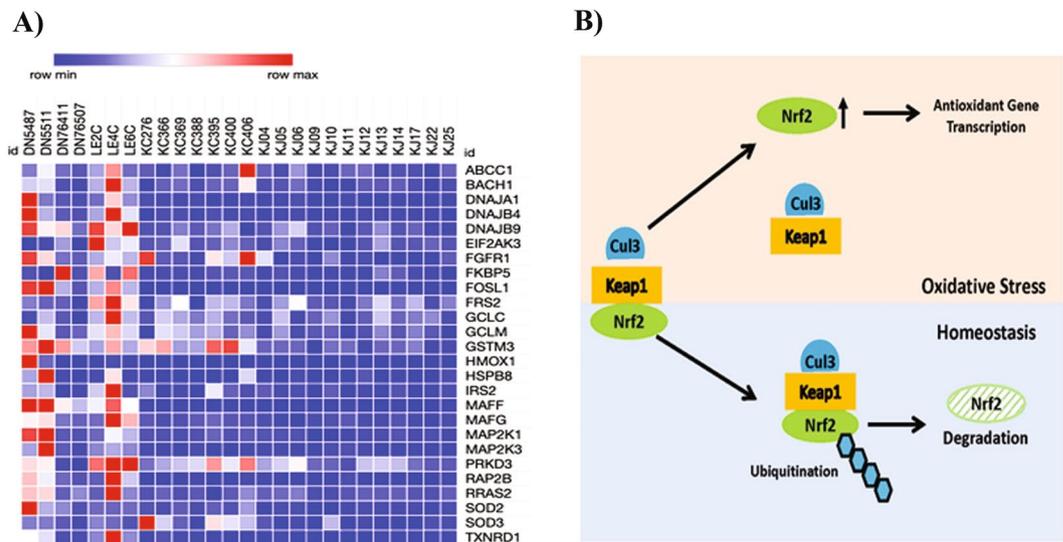


Figure 5. NRF2 target genes significantly altered in KCN corneas. **(A)** Decreased NRF2 target gene expression in KC and KJ RNA seq. **(B)** Mechanism of NRF2 regulation by KEAP1; the latter binds to NRF2 and CUL3 for NRF2-ubiquitination and degradation to maintain low levels of NRF2 under homeostatic conditions. Under oxidative stress, KEAP1 dissociates from NRF2 allowing its increase and upregulation of target genes.

(cytoskeletal), *SOD2* (belonging to the NRF2-mediated antioxidant pathway), *RXRA* (nuclear retinoid receptor) and *MFAP3L* (microfibril associated protein3 like) for testing by qRT-PCR because RNA sequencing showed robust expression at $\text{FPKM} \geq 5$, which in our experience, allows reliable detection by qRT-PCR (Supplemental Fig. S5 and Table S7). In agreement with the RNA Seq data, qRT-PCR detected elevated *MFAP3L* and decreased *TUBB3* expression in KCN. We detected robust expression of *RXRA* by qPCR, but it was not significantly different between DN and KCN groups. However, *RXRA* by RNA seq was elevated in both the KC (mean 35 FPKM) and the KJ (mean 43 FPKM) patient groups compared to donor samples (17 FPKM). Expression *GAS1* by qRT-PCR was slightly increased in KCN, without reaching significance, although by RNA seq, it was markedly increased in both patient groups (Supplemental Table S3). *HSP40/DNAJA1*, detected as markedly decreased by RNA seq in patients, was decreased by qPCR as well without reaching significance. Taken together, the qPCR data is in general agreement with the RNA seq results.

We selected *DNAJA1* and *GAS1* for localization of their respective proteins in the cornea by IF staining. *DNAJA1* (Hsp40 member a1), unknown to the eye, belongs to the heat shock protein family; other members of this family are known to serve molecular chaperone functions in the lens and the cornea⁵⁷. In agreement with the RNA seq and qRT-PCR results, *DNAJA1/HSP40* IF staining was decreased in KCN (Fig. 7A). *GAS1* is known to be expressed in the retinal pigment epithelium (RPE) and the cornea, regulates cell growth and eye development, and *GAS1*-null mice which are viable at birth develop microphthalmia⁵⁸. *GAS1* shows very little IF staining of KCN corneas (Fig. 7B) compared to uniform strong staining of all donor corneal epithelial layers. While IF staining is not a reliable indication of protein levels, qualitatively decreased staining of KCN corneal sections contrasts the *GAS1* transcript-level increases in both patient groups (Fig. 3). It remains to be seen whether decreased *GAS1* staining is due to 1) an actual decrease protein synthesis, or 2) shedding or loss of cell surface or vesicular *GAS1* from the corneal layers. IF-staining of *RXRA* showed a dramatic increase in nuclear staining in epithelial layers of KCN corneas (Fig. 7C). This is consistent with increases in its transcript.

Overall our transcriptome results also agree with findings from earlier studies. For example, *CTGF*, *SMAD7*, *TGFB1* gene expressions were decreased in both KC and KJ samples and were also reported to be decreased as a part of the Hippo, Wnt and *TGF β* signaling pathways⁴⁰. Consistent with ER stress identified as important by our earlier proteomic studies^{30,31,56}, here we detected changes in translation-related factors (*EIF1*, *EIF1B*, *EIF4E*) and proteasomal degradation components (*UBAP1*).

Discussion

Using deep transcriptome analysis of keratoconus and control corneas from African American and Saudi Arabian patients, we respectively identified 819 and 993 differentially expressed genes. As tissue donation is an extremely rare practice in Saudi Arabia, we selected donor controls of European Ancestry as previously it has been shown that these individuals are genetically closer to Middle Eastern individuals than Africans⁴³. In addition, on average donor groups were older than the patient groups, however, there were no age-dependent distribution of samples in the principal component analysis. Remarkably, despite differing ancestry and geography (ecology), and the use of generally older control corneas from an eye bank, patient transcriptomes were clustered and clearly separated from control transcriptomes by principal component analysis. Previous transcriptomic analyses have also examined corneas from keratoconus patients^{40,41}; in one study, keratoconus corneas were compared to non-keratoconus corneas from patients undergoing cornea grafting for other conditions. These patients had overt inflammation, infection and other eye diseases, and were likely reasons why case versus control corneas

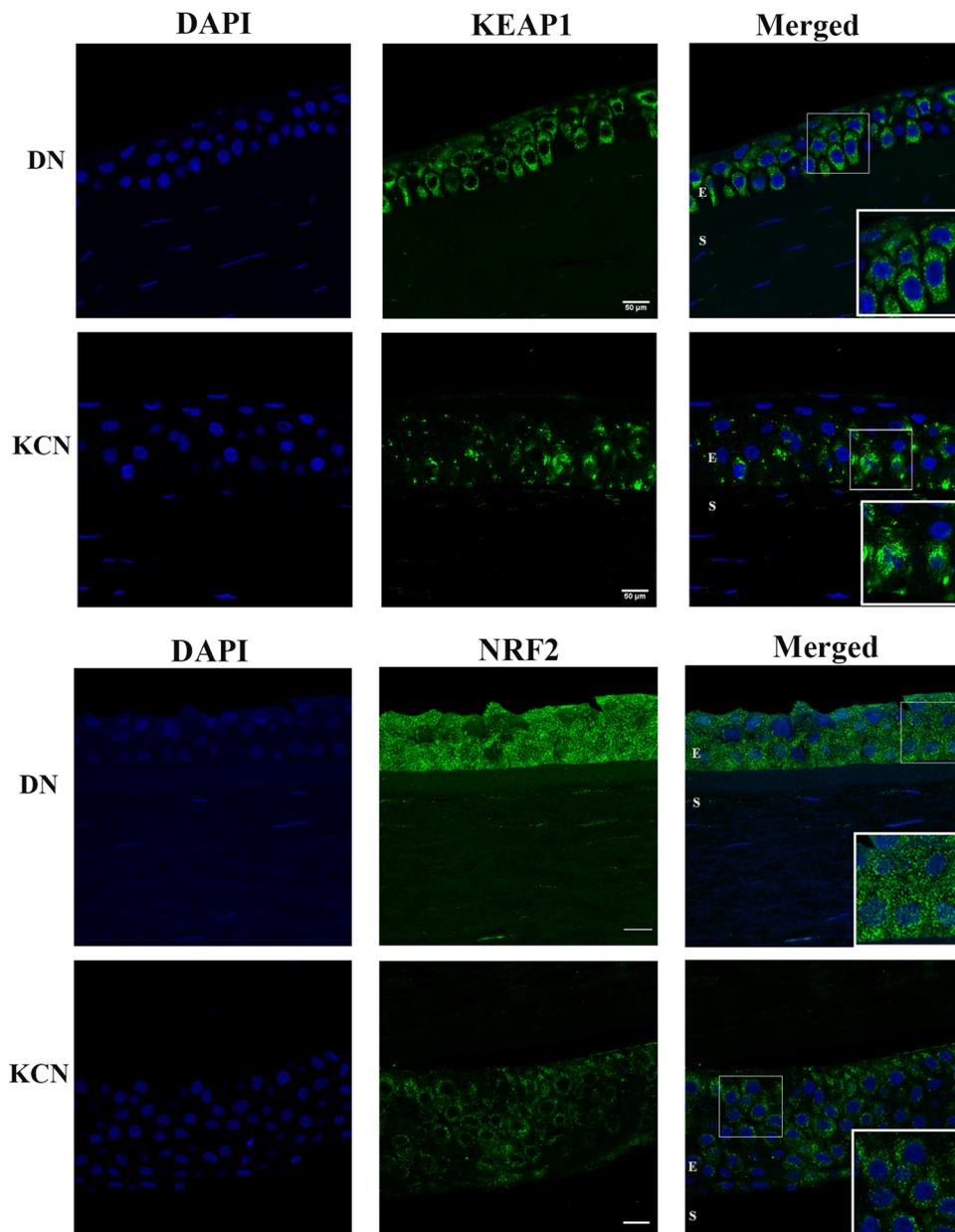


Figure 6. KEAP1 and NRF2 immunostaining in DN and KCN corneas. **(A)** KEAP1 staining is decreased in KCN corneas, with focally increased staining in some basal epithelial cells (inset), whereas in DN corneas KEAP1 shows staining of all epithelial layers (inset). **(B)** NRF2 shows very little to no staining of KCN corneas and these were all cytoplasmic (inset), while DN sections show stronger staining of epithelial cells and some nuclear staining (inset) DAPI nuclear staining shown in blue. IF staining of additional KCN and DN cornea sections are shown in Supplemental Fig. S4. Scale bar: 50 μm .

were poorly distinguished. The internal consistency of our data, therefore, allowed us to demonstrate decreases in NRF2-mediated oxidative stress and acute phase tissue injury responses as core pathogenic changes in keratoconus in both patient groups. *HMOX1* is an important NRF2-downstream antioxidant target gene that catabolizes heme and resolves iron imbalance and toxicity, and decrease in its transcript in both keratoconus patient groups affirms NRF2 dysfunctions. Reduced *HMOX1* may also contribute to the iron lines or Fleischer's Ring observed in keratoconus corneas, and the long-held belief of iron imbalance in this disease^{3,59}. Additional support for NRF2 impairments in KCN comes from our earlier proteomic study of pooled KCN and donor corneas where NRF2-mediated oxidative stress response proteins were decreased in the epithelium³⁰. Given that samples were pooled in that earlier study, we were not fully appreciative of the significance of NRF2 in keratoconus at that time. In a recent proteomic study of individual corneas from keratoconus patients we detected decreases in proteins that regulate complement pathways and tissue injury responses, and other proteins encoded by NRF2-target genes, emphasizing a failure in protective mechanisms in the cornea⁵⁶.

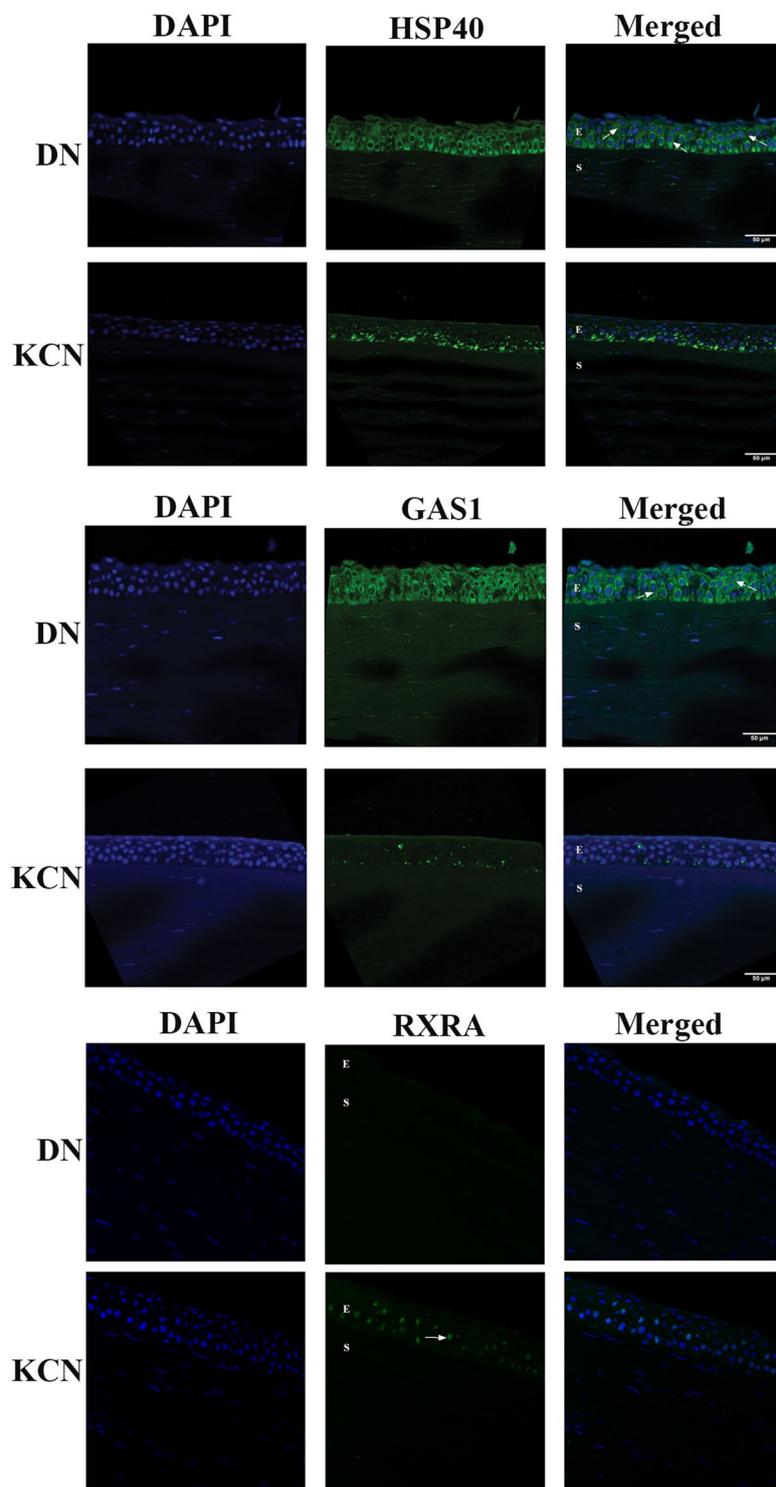


Figure 7. Immunofluorescence staining of selected markers in corneal sections. (A) DNAJA1/Hsp40 is weaker in the KCN, with some staining of basal epithelial cells and more uniform staining of all epithelial layers (E) in DN (arrows). (B) GAS1 staining was almost non-existent in KCN but show uniform staining of all epithelial layers (arrows) and weak stromal (S) staining in DN. These represent one of two KCN and DN cornea samples. (C) RXRA showed a dramatic increase in nuclear staining in epithelial layers of KCN corneas Scale bar: 50 μm.

The healthy cornea routinely encounters and resolves reactive oxygen species (ROS) produced in response to restricted nutrient supply, UV exposure, mechanical injury and infections⁶⁰. This ability to dissipate ROS was suspected to be malfunctioning in keratoconus, causing oxidative damage and consequent degenerative changes in corneal cells and the ECM. A “cascade hypothesis” was proposed to tie this oxidative damage to abnormal stromal ECM and keratocyte cell death, as observed in keratoconus²³. Therefore, it is pathogenetically significant that

we have now identified NRF-2 signal disruptions as an important early regulatory step in the poor antioxidant functions in keratoconus. Concordantly, sulforaphane, a NRF-2 activator reduces oxidative stress in an *ex vivo* cell culture model of Fuchs dystrophy⁶¹ and corneal damage in a rabbit model of corneal thinning⁶².

The two patient groups do harbor some differences as discussed below. First, certain growth factor and inflammation gene-networks are decreased in the Saudi Arabian group, although there are more conjunctivitis cases in this group. The connection between inflammation and keratoconus is not clear however. Historically, KCN is described as a non-inflammatory corneal ectasia, although tear fluid IL-6 increases, TNF- α decreases⁶³ and increases⁶⁴ suggest otherwise. The African American cases show significant changes in RAR/Vitamin A signaling, PPAR and ECM-related networks. Interestingly, increased nuclear staining of RXRA in KCN corneas and increases in its transcript in our RNA seq data may tie in with its potential candidate gene stature. Increased transcripts for ECM-related genes, encoding COL1A1, COL3A1 and FN, potentially implicate fibrosis-related changes in the African American group, that we will investigate in a larger set of patients. In a proteomic study of pooled corneas, although on a different set of patients, we noted decreases in some of these same ECM proteins³⁰. However, the relationship between transcript level changes in ECM genes and ECM protein accumulation in a tissue is often not direct, as it is influenced by feedback regulation, secretion and correct assembly of ECM proteins, and their degradation and remodeling. A vast body of work has investigated the TGF β -ECM remodeling axis in the cornea in the context of wound healing and fibrosis^{65–69}. Along these lines, we have previously probed TGF β signals in cultured stromal cells from keratoconus corneas and detected increases in non-canonical TGF β -SMAD 1/5/8 signals, and poor activation of AKT, that controls cell survival and their metabolic and biosynthetic capabilities³¹. Yet others have reported decreased regulatory SMAD7 in cultured keratoconus stromal cells⁷⁰, which is consistent with our observations that *SMAD7* transcript is decreased in African American cases. Other transcriptomic studies have also reported gene expression decreases in members of the TGF β -SMAD pathway⁴⁰. Finally, this study shows changes in multiple epithelial genes, providing credence to an existing hypothesis of epithelial damage as a potential driver of stromal degeneration in keratoconus.

Importantly, a small set of genes (*THBS2*, *GAS1*, *OTOP2*, and others) were upregulated across patients in both groups compared to controls. These may yield early biomarkers for keratoconus. In addition, a combination of genes significantly upregulated in the KJ samples only, *S100A8*, *S100A9*, *ADAMTS14*, *LYPD2* and *AOC1* may yield distinguishing biomarkers for subtypes of keratoconus. This will require additional careful phenotyping of disease and its progression, and studies of larger groups for transcript and protein-level changes.

Our study provides three major concepts in keratoconus. First, the degenerative ECM thinning phenotype of the cornea is an outcome of loss of cellular functions that is geared to counteract oxidative stress and tissue injury. Second, the NRF2-regulated gene network has a significant role in this cellular response to oxidative stress in the cornea. Finally, using our approach, analysis of transcriptomic data from different populations and patient groups will help to develop signatures and biomarkers for different subtypes of keratoconus.

Materials and Methods

Patient and donor samples. This study was conducted following the Principles of the Declaration of Helsinki on Biomedical Research involving human subjects. We obtained written informed consents from all the participants following a protocol approved by the Institutional Review Board of the Johns Hopkins Medical Institutions, Baltimore, MD and the King Khaled Eye Specialist Hospital (KKESH), Riyadh, Saudi Arabia. Further analyses were performed by study team members and additional guidance of an approved protocol from the Institutional Review Board at NYU Langone Health. A total of 7 and 12 corneas were obtained from patients undergoing cornea transplantation at the Wilmer Eye Institute (WEI) at Johns Hopkins and KKESH, respectively (Supplemental Table S1). The patients from WEI and KKESH were of African American and Saudi Arabian ancestry, respectively, and referred to as KC and KJ throughout the study. Keratoconus diagnosis was performed by trained physicians at both sites and included slit-lamp biomicroscopy, fundus evaluations and corneal topography measurements by Pentacam. Patients at both sites were graded as severe, with central keratometry (K) > 52 diopters (D) or not measurable (WEI) or K > 55D (KKESH). The Lions Eye Institute for Transplant and Research, Florida, provided control corneas, deemed unsuitable for transplants, and without keratoconus or other inflammatory diseases, obtained from deceased individuals of African American (DN) and European American (LE) ancestry. An additional 2 control corneas of unknown ancestry were obtained from WEI (DN373 and DN401). Donor corneas with an intact limbal ring packed in Optisol solution at 4°C were shipped to Baltimore for these studies, with death to Optisol storage time being approximately 9.5 hours. The donor corneas were received within a week after their procurement and processed immediately for RNA extraction or paraffin-embedding.

Tissue preparation. Samples received from patients had the central cornea only without peripheral limbal tissue. Central cornea halves received from KKESH patients were snap frozen in liquid nitrogen and stored at –80°C until shipment on dry ice to WEI for RNA extraction and sequencing. At WEI, patient cornea halves obtained immediately after surgery were placed directly in TRIzol and brought to the laboratory for RNA extraction. The limbal ring was removed from donor corneas, halved, and one half used for RNA extraction.

RNA extraction, quantification, sequencing and analysis. Total RNA was isolated from individual corneas by homogenization in 1 ml TRIzol Reagent (Life Technologies -15596–026) at room temperature and chloroform-extracted. The RNA in the aqueous phase was precipitated using 0.5 ml of 100% isopropanol and collected by centrifugation at 12,000 g, washed with 75% ethanol and resuspended in RNase-free water.

A fluorescence-based (Ribogreen, Life technologies) method was used for RNA quantification. cDNA synthesis, fragmentation and library preparation and sequencing were performed by MacroGen Inc. using the TruSeq stranded RNA library kit. A template size check was performed by running the samples on an Agilent

Technologies 2100 Bioanalyzer with a DNA chip. The samples were sequenced at 101 bp using the Illumina HiSeq 2500 platform. Raw images were generated using the HiSeq Control software v2.2.38, base calling using an integrated primary analysis software, Real time Analysis v1.18.61.0, and converted into FASTQ files using the Illumina package bcl2fastq v1.8.4. The sequencing library size, total reads, GC content and Q20/Q30 values were comparable for all samples and demonstrated high quality (Supplemental Table S2). After cufflink quantification (file “isoforms.fpkm_tracking” and file “genes.fpkm_tracking”), we obtained transcript-level and gene-level expression values (Table 1); expressed transcripts were used to obtain summed FPKM (fragments per kilobase of transcripts per million mapped reads) counts per gene.

Alignment of reads against the human reference genome hg19 was performed using the STAR 2.6.0a software. During alignment, we removed non-canonical junctions (option: `-outFilterIntronMotifs RemoveNoncanonical`) and generated XS strand attributes for splice alignments (option: `-outSAMstrandField intronMotif`). Gene expression quantification was completed using Cufflinks 2.2.1⁴⁵ with the following settings: 1) all annotated rRNA and mitochondrial transcripts were ignored (option: `-mask-file`); 2) bias detection and correction algorithm was enabled (option: `-frag-bias-correct`); 3) reads mapping to multiple locations in the genome were weighted during an initial estimation procedure (option: `-multi-read-correct`); 4) UCSC hg19 annotation file was supplied (option: `-GTF-guide`). We first checked for contamination, adaptor sequences, or other overrepresented sequences in the raw reads (FASTQ files) of all 36 sequenced samples by FastQC 0.11.7 and found no evidence of these. Next, differential expression analysis was conducted using Cuffdiff with bias correction and multi-reads corrections⁴⁴. Genes identified as significantly different were confirmed by reanalysis using DESeq2 (version: 1.24.0) with default parameters⁴⁶. In DESeq2, for final statistical significance of a gene, we used Benjamini-Hochberg adjusted $p \leq 0.01$, absolute fold change ≥ 2 , and FPKM ≥ 5 in at least one group.

FPKM value for summed transcripts was used as each gene's FPKM value. We defined significant differential expression by q value ≥ 0.01 (FDR adjusted p), $|\text{fold change}| \geq 2$, and FPKM ≥ 5 in control or patient groups. Next, we used Ingenuity Pathway Analysis (IPA) software (Qiagen) on these significant differentially expressed genes (DEG) to identify enriched biological pathways and networks among these genes.

For Principal Component Analysis (PCA) of sequenced samples, we used two definitions: a) genes expressed at FPKM ≥ 5 in *any one* sample, and b) genes expressed at FPKM ≥ 5 in *all* samples. PCA was performed using the “prcomp()” function in R “stats” Package version 3.5.0 after expression values were $\log_2(\text{FPKM} + 1)$ transformed. The coordinates for each sample in the principal components (PC1 and PC2) were used to calculate the Euclidian distance between samples by the function “pointDistance()” in R “raster” Package version 3.0–2.

Gene expression assays by qRT-PCR. cDNA was prepared from each RNA sample using a cDNA Reverse transcription kit (Biorad). Each cDNA (20 ng) was subjected to qRT-PCR using Applied Biosystems TaqMan assays for selected genes on a One Step Plus instrument (Applied Biosystems). The number of cycles (Ct) needed to reach the midpoint of the linear phase was noted and all observations were normalized against the housekeeping gene *GAPDH*.

Immunofluorescence (IF) staining. Paraffin-embedded cornea sections were stained with H&E (NYU School of Medicine Center for Biospecimen Research & Development). For immunofluorescence (IF), slides were blocked with 10% animal serum or BSA in PBS, followed by overnight incubation at 4 °C in primary antibody, washed three times with Tris-buffered saline, incubated with a secondary antibody for 2 hours and the nuclei counterstained with DAPI. The following primary antibodies were used from Biorbyt Research Products: NRF2 (orb128433, 1:50), KEAP1 (orb48426, 1:100), Hsp40 (orb520080, 1:100), and GAS1 (orb414757, 1:100) and RXRA antibody were used from Novus Biologicals (NBP2–75653, 1:50). A fluorescently conjugated anti-rabbit secondary antibody (Invitrogen, Cat# A21206) was used at a concentration of 5 $\mu\text{g}/\text{ml}$. Images were acquired with a Zeiss LSM 700 microscope.

Received: 19 December 2019; Accepted: 17 April 2020;

Published online: 18 June 2020

References

- Klintworth, G. K. & Damms, T. Corneal dystrophies and keratoconus. *Curr. Opin. Ophthalmol.* **6**, 44–56 (1995).
- Krachmer, J. H., Feder, R. S. & Belin, M. W. Keratoconus and related noninflammatory corneal thinning disorders. *Surv. Ophthalmol.* **28**, 293–322, 0039-6257(84)90094-8 [pii] (1984).
- Bron, A. J. Keratoconus. *Cornea* **7**, 163–169 (1988).
- Rabinowitz, Y. S. Keratoconus. *Surv. Ophthalmol.* **42**, 297–319 (1998).
- Soiberman, U., Foster, J. W., Jun, A. S. & Chakravarti, S. Pathophysiology of Keratoconus: What Do We Know Today. *open. Ophthalmol. J.* **11**, 252–261, <https://doi.org/10.2174/1874364101711010252> (2017).
- Chang, H. Y. & Chodosh, J. The genetics of keratoconus. *Semin. Ophthalmol.* **28**, 275–280, <https://doi.org/10.3109/08820538.2013.825295> (2013).
- Kennedy, R. H., Bourne, W. M. & Dyer, J. A. A 48-year clinical and epidemiologic study of keratoconus. *Am. J. Ophthalmol.* **101**, 267–273, [https://doi.org/10.1016/0002-9394\(86\)90817-2](https://doi.org/10.1016/0002-9394(86)90817-2) (1986).
- Godefrooij, D. A., de Wit, G. A., Uiterwaal, C. S., Imhof, S. M. & Wisse, R. P. Age-specific Incidence and Prevalence of Keratoconus: A Nationwide Registration Study. *Am. J. Ophthalmol.* **175**, 169–172, <https://doi.org/10.1016/j.ajo.2016.12.015> (2017).
- Assiri, A. A., Yousuf, B. I., Quantock, A. J. & Murphy, P. J. Incidence and severity of keratoconus in Asir province, Saudi Arabia. *Br J Ophthalmol* **89**, 1403–1406, 89/11/1403 [pii] <https://doi.org/10.1136/bjo.2005.074955> (2005).
- Jonas, J. B., Nangia, V., Matin, A., Kulkarni, M. & Bhojwani, K. Prevalence and associations of keratoconus in rural maharashtra in central India: the central India eye and medical study. *Am. J. Ophthalmol.* **148**, 760–765, <https://doi.org/10.1016/j.ajo.2009.06.024> (2009).
- Wagner, H., Barr, J. T. & Zadnik, K. Collaborative Longitudinal Evaluation of Keratoconus (CLEK) Study: methods and findings to date. *Cont. Lens Anterior Eye* **30**, 223–232, <https://doi.org/10.1016/j.clae.2007.03.001> (2007).

12. Jhanji, V., Sharma, N. & Vajpayee, R. B. Management of keratoconus: current scenario. *Br J Ophthalmol* **95**, 1044–1050, bjo.2010.185868 [pii] <https://doi.org/10.1136/bjo.2010.185868> (2011).
13. Wollensak, G., Spoerl, E. & Seiler, T. Riboflavin/ultraviolet-a-induced collagen crosslinking for the treatment of keratoconus. *Am J Ophthalmol* **135**, 620–627, S0002939402022201 [pii] (2003).
14. Poulliquen, Y. *et al.* Acute corneal edema in pellucid marginal degeneration or acute marginal keratoconus. *Cornea* **6**, 169–174 (1987).
15. Pahuja, N. *et al.* Differential Molecular Expression of Extracellular Matrix and Inflammatory Genes at the Corneal Cone Apex Drives Focal Weakening in Keratoconus. *Invest. Ophthalmol. Vis. Sci.* **57**, 5372–5382, <https://doi.org/10.1167/iovs.16-19677> (2016).
16. Srivastava, O. P., Chandrasekaran, D. & Pfister, R. R. Molecular changes in selected epithelial proteins in human keratoconus corneas compared to normal corneas. *Mol Vis* **12**, 1615–1625, v12/a185 [pii] (2006).
17. Efron, N. & Hollingsworth, J. G. New perspectives on keratoconus as revealed by corneal confocal microscopy. *Clin Exp Optom* **91**, 34–55, CXO195 [pii] <https://doi.org/10.1111/j.1444-0938.2007.00195.x> (2008).
18. Ku, J. Y., Niederer, R. L., Patel, D. V., Sherwin, T. & McGhee, C. N. Laser scanning *in vivo* confocal analysis of keratocyte density in keratoconus. *Ophthalmology* **115**, 845–850, S0161-6420(07)00623-9 [pii] <https://doi.org/10.1016/j.ophtha.2007.04.067> (2008).
19. Niederer, R. L., Perumal, D., Sherwin, T. & McGhee, C. N. Laser scanning *in vivo* confocal microscopy reveals reduced innervation and reduction in cell density in all layers of the keratoconic cornea. *Invest. Ophthalmol. Vis. Sci.* **49**, 2964–2970, <https://doi.org/10.1167/iovs.07-0968> (2008).
20. Meek, K. M. *et al.* Changes in collagen orientation and distribution in keratoconus corneas. *Invest Ophthalmol Vis Sci* **46**, 1948–1956, 46/6/1948 [pii] <https://doi.org/10.1167/iovs.04-1253> (2005).
21. Mikula, E. *et al.* Axial mechanical and structural characterization of keratoconus corneas. *Exp. Eye Res.* **175**, 14–19, <https://doi.org/10.1016/j.exer.2018.05.019> (2018).
22. Fukuchi, T., Yue, B. Y., Sugar, J. & Lam, S. Lysosomal enzyme activities in conjunctival tissues of patients with keratoconus. *Arch. Ophthalmol.* **112**, 1368–1374 (1994).
23. Kenney, M. C. & Brown, D. J. The cascade hypothesis of keratoconus. *Cont Lens Anterior Eye* **26**, 139–146, S1367-0484(03)00022-5 [pii] [https://doi.org/10.1016/S1367-0484\(03\)00022-5](https://doi.org/10.1016/S1367-0484(03)00022-5) (2003).
24. Lema, I. & Duran, J. A. Inflammatory molecules in the tears of patients with keratoconus. *Ophthalmology* **112**, 654–659, S01616420(04)019189 [pii] <https://doi.org/10.1016/j.ophtha.2004.11.050> (2005).
25. Lema, I. *et al.* Inflammatory response to contact lenses in patients with keratoconus compared with myopic subjects. *Cornea* **27**, 758–763, <https://doi.org/10.1097/ICO.0b013e31816a3591.00003226-200808000-00003> [pii] (2008).
26. Sawaguchi, S. *et al.* Alpha-1 proteinase inhibitor levels in keratoconus. *Exp. Eye Res.* **50**, 549–554 (1990).
27. Sawaguchi, S., Yue, B. Y., Chang, I., Sugar, J. & Robin, J. Proteoglycan molecules in keratoconus corneas. *Invest. Ophthalmol. Vis. Sci.* **32**, 1846–1853 (1991).
28. Nielsen, K. *et al.* Proteome profiling of corneal epithelium and identification of marker proteins for keratoconus, a pilot study. *Exp Eye Res* **82**, 201–209, S0014-4835(05)00184-3 [pii] <https://doi.org/10.1016/j.exer.2005.06.009> (2006).
29. Balasubramanian, S. A., Wasinger, V. C., Pye, D. C. & Willcox, M. D. Preliminary identification of differentially expressed tear proteins in keratoconus. *Mol. Vis.* **19**, 2124–2134 (2013).
30. Chaerkady, R. *et al.* The keratoconus corneal proteome: Loss of epithelial integrity and stromal degeneration. *J. Proteom.* **87**, 122–131, S1874-3919(13)00264-9 [pii] <https://doi.org/10.1016/j.jpro.2013.05.023> (2013).
31. Foster, J. *et al.* Transforming growth factor beta and insulin signal changes in stromal fibroblasts of individual keratoconus patients. *PLoS One* **9**, e106556, <https://doi.org/10.1371/journal.pone.0106556> (2014).
32. Foster, J. W. *et al.* Integrated Stress Response and Decreased ECM in Cultured Stromal Cells From Keratoconus Corneas. *Invest. Ophthalmol. Vis. Sci.* **59**, 2977–2986, <https://doi.org/10.1167/iovs.18-24367> (2018).
33. Nielsen, K., Hjortdal, J., Pihlmann, M. & Corydon, T. J. Update on the keratoconus genetics. *Acta Ophthalmol.* <https://doi.org/10.1111/j.1755-3768.2012.02400.x> (2012).
34. Szczotka-Flynn, L. *et al.* Disease severity and family history in keratoconus. *Br J Ophthalmol* **92**, 1108–1111, 92/8/1108 [pii] <https://doi.org/10.1136/bjo.2007.130294> (2008).
35. Bykhovskaya, Y., Margines, B. & Rabinowitz, Y. S. Genetics in Keratoconus: where are we? *Eye Vis.* **3**, 16, <https://doi.org/10.1186/s40662-016-0047-5> (2016).
36. Wachtmeister, L., Ingemansson, S. O. & Moller, E. Atopy and HLA antigens in patients with keratoconus. *Acta Ophthalmol.* **60**, 113–122 (1982).
37. Chwa, M. *et al.* Increased stress-induced generation of reactive oxygen species and apoptosis in human keratoconus fibroblasts. *Invest. Ophthalmol. Vis. Sci.* **47**, 1902–1910, 47/5/1902 [pii] <https://doi.org/10.1167/iovs.05-0828> (2006).
38. Yeniad, B., Alparslan, N. & Akarçay, K. Eye rubbing as an apparent cause of recurrent keratoconus. *Cornea* **28**, 477–479, <https://doi.org/10.1097/ICO.0b013e31818d37fa00003226-200905000-00023> [pii] (2009).
39. Sharif, R. *et al.* Prolactin-Induced Protein is a novel biomarker for Keratoconus. *Exp. Eye Res.* **179**, 55–63, <https://doi.org/10.1016/j.exer.2018.10.015> (2019).
40. Kabza, M. *et al.* Collagen synthesis disruption and downregulation of core elements of TGF-beta, Hippo, and Wnt pathways in keratoconus corneas. *Eur. J. Hum. Genet.* **25**, 582–590, <https://doi.org/10.1038/ejhg.2017.4> (2017).
41. Khaled, M. L. *et al.* Differential Expression of Coding and Long Noncoding RNAs in Keratoconus-Affected Corneas. *Invest. Ophthalmol. Vis. Sci.* **59**, 2717–2728, <https://doi.org/10.1167/iovs.18-24267> (2018).
42. Alanazi, L. F. *et al.* Attitude, beliefs and awareness towards corneal donation in Saudi Arabia. *Saudi J. Ophthalmol.* **33**, 121–129, <https://doi.org/10.1016/j.sjopt.2019.02.003> (2019).
43. Belkadi, A. *et al.* Whole-exome sequencing to analyze population structure, parental inbreeding, and familial linkage. *Proc. Natl Acad. Sci. USA* **113**, 6713–6718, <https://doi.org/10.1073/pnas.1606460113> (2016).
44. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* **31**, 46–53, nbt.2450 [pii], <https://doi.org/10.1038/nbt.2450> (2013).
45. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511–515, nbt.1621 [pii] <https://doi.org/10.1038/nbt.1621> (2010).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. *Genome Biol.* **15**, 550, <https://doi.org/10.1186/s13059-014-0550-8> (2014).
47. Bekhouche, M. & Colige, A. The procollagen N-proteinases ADAMTS2, 3 and 14 in pathophysiology. *Matrix Biol.* **44–46**, 46–53, <https://doi.org/10.1016/j.matbio.2015.04.001> (2015).
48. Chatterjee, A., Oh, D. J., Kang, M. H. & Rhee, D. J. Central corneal thickness does not correlate with TonoLab-measured IOP in several mouse strains with single transgenic mutations of matricellular proteins. *Exp. Eye Res.* **115**, 106–112, <https://doi.org/10.1016/j.exer.2013.06.017> (2013).
49. Bornstein, P., Kyriakides, T. R., Yang, Z., Armstrong, L. C. & Birk, D. E. Thrombospondin 2 modulates collagen fibrillogenesis and angiogenesis. *J. Invest. Dermatol. Symp. Proc.* **5**, 61–66, <https://doi.org/10.1046/j.1087-0024.2000.00005.x> (2000).
50. Armstrong, D. J., Hiscott, P., Batterbury, M. & Kaye, S. Corneal stromal cells (keratocytes) express thrombospondins 2 and 3 in wound repair phenotype. *Int. J. Biochem. Cell Biol.* **34**, 588–593 (2002).
51. Hiscott, P. *et al.* Thrombospondin 1, thrombospondin 2 and the eye. *Prog. Retin. Eye Res.* **25**, 1–18, <https://doi.org/10.1016/j.preteyres.2005.05.001> (2006).

52. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57, nprot.2008.211 [pii] <https://doi.org/10.1038/nprot.2008.211> (2009).
53. Jimenez-Marin, A., Collado-Romero, M., Ramirez-Boo, M., Arce, C. & Garrido, J. J. Biological pathway analysis by ArrayUnlock and Ingenuity Pathway Analysis. *BMC Proc.* **3**(Suppl 4), S6, <https://doi.org/10.1186/1753-6561-3-S4-S6> (2009).
54. Ma, Q. Role of nrf2 in oxidative stress and toxicity. *Annu. Rev. Pharmacol. Toxicol.* **53**, 401–426, <https://doi.org/10.1146/annurev-pharmtox-011112-140320> (2013).
55. Batliwala, S., Xavier, C., Liu, Y., Wu, H. & Pang, I. H. Involvement of Nrf2 in Ocular. *Diseases. Oxid. Med. Cell. Longev.* **2017**, 1703810, <https://doi.org/10.1155/2017/1703810> (2017).
56. Shinde, V. *et al.* Mapping Keratoconus Molecular Substrates by Multiplexed High-Resolution Proteomics of Unpooled Corneas. *OmicS: a journal of integrative biology*, <https://doi.org/10.1089/omi.2019.0143> (2019).
57. Piatigorsky, J. Enigma of the abundant water-soluble cytoplasmic proteins of the cornea: the “refracton” hypothesis. *Cornea* **20**, 853–858 (2001).
58. Lee, C. S., May, N. R. & Fan, C. M. Transdifferentiation of the ventral retinal pigmented epithelium to neural retina in the growth arrest specific gene 1 mutant. *Dev. Biol.* **236**, 17–29, <https://doi.org/10.1006/dbio.2001.0280> (2001).
59. Loh, A., Hadziahmetovic, M. & Dunaief, J. L. Iron homeostasis and eye disease. *Biochim Biophys Acta* **1790**, 637–649, S0304-4165(08)00245-6 [pii] <https://doi.org/10.1016/j.bbagen.2008.11.001> (2009).
60. Marchitti, S. A., Chen, Y., Thompson, D. C. & Vasilou, V. Ultraviolet radiation: cellular antioxidant response and the role of ocular aldehyde dehydrogenase enzymes. *Eye Contact Lens* **37**, 206–213, <https://doi.org/10.1097/ICL.0b013e3182212642> (2011).
61. Ziaei, A., Schmedt, T., Chen, Y. & Jurkunas, U. V. Sulforaphane decreases endothelial cell apoptosis in fuchs endothelial corneal dystrophy: a novel treatment. *Invest. Ophthalmol. Vis. Sci.* **54**, 6724–6734, <https://doi.org/10.1167/iovs.13-12699> (2013).
62. Liu, R. & Yan, X. Sulforaphane protects rabbit corneas against oxidative stress injury in keratoconus through activation of the Nrf-2/HO-1 antioxidant pathway. *Int. J. Mol. Med.* **42**, 2315–2328, <https://doi.org/10.3892/ijmm.2018.3820> (2018).
63. Jun, A. S. *et al.* Subnormal cytokine profile in the tear fluid of keratoconus patients. *Plos One* **6**, e16437, <https://doi.org/10.1371/journal.pone.0016437> (2011).
64. Lema, I., Brea, D., Rodriguez-Gonzalez, R., Diez-Feijoo, E. & Sobrino, T. Proteomic analysis of the tear film in patients with keratoconus. *Mol. Vis.* **16**, 2055–2061 (2010).
65. Mohan, R. R. & Wilson, S. E. *Ex vivo* human corneal epithelial cells express membrane-bound precursor and mature soluble epidermal growth factor (EGF) and transforming growth factor (TGF) alpha proteins. *Exp Eye Res* **68**, 129–131, S0014-4835(98)90568-1 [pii] <https://doi.org/10.1006/exer.1998.0568> (1999).
66. Saika, S. *et al.* TGFbeta2 in corneal morphogenesis during mouse embryonic development. *Dev Biol* **240**, 419–432 S0012-1606(01)90480-4 [pii], <https://doi.org/10.1006/dbio.2001.0480> (2001).
67. Zieske, J. D., Hutcheon, A. E., Guo, X., Chung, E. H. & Joyce, N. C. TGF-beta receptor types I and II are differentially expressed during corneal epithelial wound repair. *Invest. Ophthalmol. Vis. Sci.* **42**, 1465–1471 (2001).
68. Jester, J. V., Huang, J., Petroll, W. M. & Cavanagh, H. D. TGFbeta induced myofibroblast differentiation of rabbit keratocytes requires synergistic TGFbeta, PDGF and integrin signaling. *Exp Eye Res* **75**, 645–657, S001448350292066X [pii] (2002).
69. He, J. & Bazan, H. E. Epidermal growth factor synergism with TGF-beta1 via PI-3 kinase activity in corneal keratocyte differentiation. *Invest Ophthalmol Vis Sci* **49**, 2936–2945, doi:49/7/2936 [pii] <https://doi.org/10.1167/iovs.07-0900> (2008).
70. Priyadarsini, S., McKay, T. B., Sarker-Nag, A. & Karamichos, D. Keratoconus *in vitro* and the key players of the TGF-beta pathway. *Mol. Vis.* **21**, 577–588 (2015).

Acknowledgements

This study was supported by funds from King Khaled Eye Specialist Hospital (SC and SAS) and NIH/NEI EY 026104 (SC). The funding organizations had no role in the design or conduct of this research.

Author contributions

V.S. Conducted the research and prepared figures. N.H. Analysis of RNA seq data and prepared figures. A.M. Phenotyped and provided samples, critiqued manuscript. G.M. Provided technical advice. Y.D. Clinical evaluation and patient cornea samples. C.E. Provided donor cornea sections for immunostaining. A.M. Clinical evaluation and provided patient cornea samples. A.J. Clinical evaluation, provided patient corneas and critiqued manuscript. S.S. Clinical evaluation and provided patient corneas. S.C. Designed the research and wrote the main manuscript. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-66735-x>.

Correspondence and requests for materials should be addressed to S.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020