Genome-wide association study identifies three novel loci in Fuchs endothelial corneal dystrophy

Natalie A. Afshari1,*, Robert P. Igo Jr2,*, Nathan J. Morris2, Dwight Stambolian3, Shiwani Sharma4, V. Lakshmi Pulagam2, Steven Dunn5, John F. Stamler6, Barbara J. Truitt2, Jacqueline Rimmler7, Abraham Kuot4, Christopher R. Croasdale8, Xuejun Qin7, Kathryn P. Burdon4,9, S. Amer Riazuddin10, Richard Mills4, Sonja Klebe4,11, Mollie A. Minear7, Jiagang Zhao1, Elmer Balajonda12, George O. Rosenwasser13, Keith H. Baratz14, V. Vinod Mootha15, Sanjay V. Patil14, Simon G. Gregory7, Joan E. Bailey-Wilson16, Marianne O. Price17, Francis W. Price Jr17, Jamie E. Craig4, John H. Fingert6, John D. Gottsch10, Anthony J. Aldave18, Gordon K. Klintworth12,19, Jonathan H. Lass2,20,**, Yi-Ju Li7,21,** & Sudha K. Iyengar2,20,**

The structure of the cornea is vital to its transparency, and dystrophies that disrupt corneal organization are highly heritable. To understand the genetic aetiology of Fuchs endothelial corneal dystrophy (FECD), the most prevalent corneal disorder requiring transplantation, we conducted a genome-wide association study (GWAS) on 1,404 FECD cases and 2,564 controls of European ancestry, followed by replication and meta-analysis, for a total of 2,075 cases and 3,342 controls. We identify three novel loci meeting genome-wide significance (P<5×10^{-8}): KANK4 rs79742895, LAMC1 rs3768617 and LINC00970/ATP1B1 rs1200114. We also observe an overwhelming effect of the established TCF4 locus. Interestingly, we detect differential sex-specific association at LAMC1, with greater risk in women, and TCF4, with greater risk in men. Combining GWAS results with biological evidence we expand the knowledge of common FECD loci from one to four, and provide a deeper understanding of the underlying pathogenic basis of FECD.

1 Shiley Eye Institute, University of California, La Jolla, California 92039, USA. 2 Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio 44106, USA. 3 Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. 4 Department of Ophthalmology, Flinders Medical Centre, Adelaide, South Australia 5042, Australia. 5 Michigan Cornea Consultants, PC, Southfield, Michigan 48034, USA. 6 Department of Ophthalmology, University of Iowa, College of Medicine, Iowa City, Iowa 52242, USA. 7 Duke Molecular Physiology Institute, Duke University Medical Center, Durham, North Carolina 27701, USA. 8 Davis Duehr Dean Clinic, Madison, Wisconsin 53715, USA. 9 Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania 7000, Australia. 10 The Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287, USA. 11 Department of Pathology, Flinders Medical Centre, Flinders University, Adelaide, South Australia 5042, Australia. 12 Duke University Eye Center, Duke University Medical Center, Durham, North Carolina 27710, USA. 13 Central Pennsylvania Eye Institute, Hershey, Pennsylvania 17033, USA. 14 Department of Ophthalmology, Mayo Clinic, Rochester, Minnesota 55905, USA. 15 Department of Ophthalmology, University of Texas Southwesten Medical Center, Dallas, Texas 75235, USA. 16 Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health and Johns Hopkins University, Baltimore, Maryland 21224, USA. 17 Price Vision Group, Indianapolis, Indiana 46260, USA. 18 Stein Eye Institute, University of California, Los Angeles, California 90095, USA. 19 Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina 27710, USA. 20 Department of Ophthalmology and Visual Sciences, Case Western Reserve University and University Hospitals Hospitals Eye Institute, Cleveland, Ohio 44106, USA. 21 Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina 27710, USA. * These authors contributed equally to this work. ** These authors jointly supervised this work. Correspondence and requests for materials should be addressed to J.H.L. (email: Jonathan.Lass@UHhospitals.org) or to Y.J.L. (email: yiju.li@duke.edu) or to S.K.I. (email: ski@case.edu).
Corneal diseases are one of the most common causes of visual impairment and blindness worldwide, with Fuchs endothelial corneal dystrophy (FECD) being the most common indication for corneal transplantation in the United States. FECD is a bilateral, slowly progressive disorder of the corneal endothelium affecting 4% of population above the age of 40 years in the United States. This debilitating disorder becomes symptomatic in the fifth or sixth decade of life and affects two to four times more women than men. Clinical presentation of FECD varies, but often includes corneal oedema that results in a loss of corneal clarity, painful episodes of recurrent corneal erosions and severely impaired visual acuity. Medical management is often inadequate requiring corneal transplantation. Histopathological changes in FECD include loss of endothelial cells, thickening of Descemet membrane (DM) due to excessive accumulation of extracellular matrix, formation of central excrescences on DM (corneal guttata) and subepithelial bullae and scarring. FECD likely arises during neural crest differentiation, but the reason for loss of functioning endothelial cells is unknown.

FECD often clusters within families, with heritability estimated at ~39% (ref. 9). The understanding of genetics of FECD initially focussed on genetic linkage, as a consequence of the observation by Krachmer et al. that this disease showed a dominant mode of inheritance in families. Linkage and fine-mapping studies in families identified rare point mutations in COL8A2 (ref. 10), SLC4A11 (ref. 11) and TCF8 (ref. 12); these constituted the early genetic discoveries of Fuchs dystrophy. Most recently, a genome-wide association study (GWAS) including 280 cases and 410 controls in both discovery and replication samples found common genetic variation in TCF4 to be associated with Fuchs dystrophy, along with two other loci that did not reach genome-wide significance. The association with FECD, along with two other loci that did not reach genome-wide significance in 410 controls in both discovery and replication samples found common genetic variation in TCF4 to be associated with Fuchs dystrophy, along with two other loci that did not reach genome-wide significance. The association with TCF4 has been replicated by multiple investigators, including our group, and a subsequent publication suggested that a trinucleotide repeat expansion within intron 3 of the TCF4 gene is risk factor for this disease. Here, we report the discovery of three new loci associated with FECD, KANK4 (KN motif- and ankyrin repeat domain-containing protein 4), LAMC1 (laminin gamma-1) and LINC00970/ATPB1 (Na+, K+ transporting ATPase, beta-1 polypeptide), and confirm the strong effect of TCF4 strongly influential variant with good biologic plausibility, but does not fully explain disease occurrence in all patients because of the lack of penetrance in some families.

Properties of index variants. The most significant TCF4 SNP, rs784257, explained 21.9% of the variation in FECD in the discovery sample, whereas the top markers in the other three replicated genome-wide significant loci each explained between 0.9% and 1.5% of the variation in the discovery sample (Supplementary Table 8). Strong linkage disequilibrium (r² = 0.88) exists between rs784257 and the previously reported variant rs613872, in the third intron of TCF4 (ref. 13). Both SNPs are in moderate disequilibrium with the expanded form of the CTG18.1 trinucleotide repeat in intron 2 of TCF4 (r² = 0.47), consistent with the previously reported values of 0.65 and 0.31 between the triplet repeat and rs613872 in FECD cases and controls, respectively. The CTG18.1 repeat expansion is a strongly influential variant with good biologic plausibility, but does not fully explain disease occurrence in all patients because of the lack of penetrance in some families. Genetic risk scores were calculated for the discovery cohort based on the most significantly associated SNP from each of the four replicated regions. The area under the curve (AUC) for FECD cases versus controls was 0.782 (95% CI = 0.767, 0.797), showing strong predictive value compared to other complex traits analysed by GWAS (Fig. 2). Predictive value was primarily through genotypes at rs784257 in TCF4 (AUC = 0.750; 95% CI = 0.736–0.765); the AUC without rs784257 was substantially smaller but still significantly >0.5 (AUC = 0.606; 95% CI = 0.587–0.624). We also found that adding the three non-TCF4 SNPs to the model significantly increased the AUC (P = 7.7 × 10⁻¹⁸ by DeLong’s test).

Pairwise SNP × SNP interaction tests among the markers listed in Table 1 revealed no significant (P < 0.05 from logistic regression including an SNP × SNP interaction term) interactions between loci. The discovery data set was large enough to detect an interaction risk ratio of 1.4–1.8 with at least 80% power (Supplementary Table 9). Conditional analysis on the TCF4 marker rs784257 did not substantially alter the effect sizes of the other loci (Supplementary Table 10). Moreover, a sensitivity analysis using only FECD cases from the discovery cohort confirmed by histopathology (Supplementary Table 11) showed no overall reduction in effect sizes (Supplementary Table 12 and Supplementary Fig. 4), indicating that our classification of affected status is reliable.
variant rs3768617 confers a significantly greater elevated risk of FECD on women (OR = 1.52, 95% CI = 1.32, 1.72) than on men (OR = 1.16, 95% CI = 0.98, 1.34) (P value for heterogeneity, \( P_{\text{het}} = 0.013 \) by \( \chi^2 \) test; Supplementary Table 13 and Supplementary Figs 5 and 6), explaining 1.74% of the overall FECD risk for women in the sample, but only 0.23% of the risk for men. On the other hand, the TCF4 variant rs784257 shows higher risk of FECD on men (OR = 7.56, 95% CI = 5.96, 9.57) than women (OR = 5.06, 95% CI = 4.29, 5.96, \( P_{\text{het}} = 0.0063 \)). Sex-specific association analysis on the Flinders University replication cohort supported these findings, whereas results from the other two cohorts did not (Supplementary Fig. 7). Regardless, both rs3768617 and rs784257 showed significant sex heterogeneity of effect in meta-analyses combining discovery and replication samples (Supplementary Fig. 7). The discovery sample offered adequate power (~0.8) to detect the SNP × sex interaction effects observed for these two markers with significant P values (Supplementary Table 14). Herein, we present the first evidence of genetic contribution to the sex-specific risk for FECD.

**Functional significance of major association results.** To derive functional insights regarding the genes nearest the most significant SNPs from GWAS, we examined several biologic sources of information, including transcription factor, histone mark and expression quantitative trait loci (eQTLs) reference maps from public databases that did not include corneal tissue (HaploReg26; expression quantitative trait loci (eQTLs) reference maps from OrthoDB27). For each locus, we noted that TCF4, ATP1B1 and LAMC1 are highly expressed in corneal samples comprising only the corneal endothelium and DM (Supplementary Data 1). LINC00970, the lncRNA closest to ATP1B1, shows no expression, while KANK4 shows minimal expression. TCF4, a basic helix–loop–helix transcription factor, was detected in the nucleus of residual endothelial cells of Fuchs case samples by IHC (Supplementary Fig. 8). No genotype data were available from the sample donors, and, therefore, we could not explore the influence of associated variants on gene expression in ocular tissue.

Our lead SNP in KANK4, rs79742895, was associated with promoter and enhancer mark enrichment in multiple tissues (Supplementary Data 2), conceivably altering motifs for transcription factors, hypermethylated in cancer1 and Y in Yang1. The cellular function of KANK4 is poorly understood, but mutations in KANK proteins lead to steroid-resistant nephrotic sydrome27. Modelling of mutations via knockdowns illustrate that ~50% of kank2 morphant zebrafish develop periortibial oedema and kidney disease, which the authors ascribe to loss of kidney function. Knockdown analysis of the KANK3 orthologue in Caenorhabditis elegans suggests a role in cell-to-cell contact and tissue integrity28. Distinct from TCF4, KANK4 immunostaining revealed that localization of KANK4 is mainly within the endothelial cytoplasm in both control and FECD samples (Supplementary Fig. 8). KANK family proteins may have a role in the regulation of actin stress fibres, which are hypothesized to hold the endothelial cell nuclei in place through cellular adhesion of the endothelial cytoplasmic layer29.

rs79742895, in KANK4, was also associated with an eQTL for the LINE-1 containing transposase domain containing 1 (LITD1) gene 23.84 kb distal to KANK4 (ref. 30). Similar to KANK4, LITD1 is only weakly expressed in the corneal endothelium (Supplementary Data 1). LITD1 is a RNA-binding protein involved in preserving pluripotency of stem cells via post-transcriptional regulation, showing high expression in pluripotent cells31. Functions of moieties that interact with LITD1 can be divided into three categories: RNA post-transcriptional modification, protein synthesis and gene expression. Emami et al.31 reported that LAMC1 is involved in protein synthesis under control of LITD1. No other variants with \( r^2 \geq 0.8 \) were located in linkage disequilibrium with rs79742895 and identified by HaploReg.

The strongest functional evidence for rs3768617 in LAMCI is with eQTLs. Eight entries with significant eQTLs were found to map within LAMCI or LAMC2, the vast majority in brain or nerve tissue (Supplementary Data 2). Thus, this variant, or others in linkage disequilibrium, appear to self-regulate the expression levels of LAMCI and LAMC2, both of which have a key role as cell adhesion proteins in basement membranes, such as DM32.
conduction. Analogous defects likely occur in the corneal matrix deposition between cells that upset electrical signal membrane formation, leading to inappropriate extracellular knockout model in cardiomyocytes derived from human Na\textsubscript{FGF2} stained cells (Supplementary Fig. 8), which retained cytoplasmic and endothelium and DM in FECD. Strengthening this hypothesis, in concomitant decrease of KANK4 cases, a decline in the number of endothelial cells resulted laminins are asymmetrically expressed on the endothelial side of DM\textsuperscript{33}. Knockouts of this gene cause failure of ureteric bud development, or reduction in kidney size, with an insufficiency of water transport postnatally\textsuperscript{34}. An important feature of this model is the absence of basement membranes in the kidney early in development, and reduction in levels of integrin-\textalpha{6} and FGF2, key elements of cellular growth. A similar LAMC1-knockout model in cardiomyocytes derived from human embryonic stem cells also exhibited lack of proper basement membrane formation, leading to inappropriate extracellular matrix deposition between cells that upset electrical signal conduction\textsuperscript{35}. Analogous defects likely occur in the corneal endothelium and DM in FECD. Strengthening this hypothesis, IHC of control corneal endothelial samples pinpointed KANK4 and LAMC1 protein expression within the cytoplasm. In FECD cases, a decline in the number of endothelial cells resulted in concomitant decrease of KANK4 and LAMC1 positively stained cells (Supplementary Fig. 8), which retained cytoplasmic expression of the proteins.

rs1200114 maps to an intergenic region between ATP1B1 and LINCO00970, which does not show expression in the corneal endothelium. Conversely, ATP1B1 shows high expression in the endothelium (Supplementary Data 1). This SNP shows no significant eQTLs in GTEx, but does show strong enrichment for enhancer and promoter marks, particularly in numerous brain tissues (Supplementary Data 2). The \beta-subunit of the multimeric Na\textsuperscript{+},K\textsuperscript{−}-ATPase has a key function in structural integrity of the protein, and its functional maturity. It is a regulator of fluid balance, ion transport and maintains the cellular homeostatic equilibrium\textsuperscript{36}. Lack of ATP1B1 may lead to hypertonicity within the cornea and cause the corneal oedema seen in patients suffering from FECD.

Recognizing that ATP1B1 represented our most significant GWAS signal after TCF4, we sought additional supporting evidence at the gene expression level. Unbiased mining of proteomic data by Poulsen et al.\textsuperscript{37} corroborated the expression of LAMC1 and ATP1B1 in disease-relevant tissues. By IHC using LAMC1 and ATP1B1 antibodies in three pairs of full-thickness corneas from FECD cases and controls, we confirmed expression of these proteins in the corneal endothelium (Supplementary Fig. 9).

### Discussion

We identified three novel loci, KANK4, LAMC1 and ATP1B1, for FECD, and demonstrated cellular expression of TCF4. Our results support a multifactorial model for disease aetiology with strong predictive power (AUC = 0.78 from an oligogenic risk-score model), although variants in TCF4 alone do a commendable job of predicting risk in Caucasians. We confirm the strong association of variants such as rs613872 in TCF4, initially discovered by Baratz et al.\textsuperscript{13}, and extend these results by demonstrating sex-specific effects. In contrast, we observed only modest association at loci previously identified through family-based and candidate-gene studies\textsuperscript{10–12,28–40}.

Our data favour LAMC1 as increasing risk in women, whereas TCF4 increases risk in men. Alleles at both loci are quite common, and their precise role in mediating sex-specific risk remains to be elucidated. Taken together, these results suggest that these proteins have essential roles in the corneal endothelium; for instance, water transport, appropriate basement membrane maintenance, as well as tissue integrity and cell-to-cell contact. Disruption of these activities may lead to pathogenic processes characteristic of FECD. The primary function of the corneal endothelium is to maintain a state of relative dehydration, with hindrance of routine fluid transport leading to corneal oedema. We propose that all three genes regulate fluid transport, with ATP1B1 having the most obvious role as it encodes a subunit of the sodium-potassium plasma membrane pump. LAMC1 also has a role in normal basement membrane deposition, as shown in knockout models of the kidney and cardiomyocytes, possibly leading to thickening of the DM. Although functional data supporting an effect of rs3768617 on LAMC1 expression exists in GTEx, this database does not include the cornea. Limited data on the function of KANK4 means that its function can currently only be inferred by analogy to its paralogues, KANK2 and KANK3, and is suggestive of maintenance of cell-to-cell contact and tissue integrity. In FECD, endothelial cells reduce in number and change shape to maintain regular contact with each other and to preserve the endothelial barrier to fluid entry into the corneal stroma\textsuperscript{41,42}. Identification of additional genetic risk factors adds to our understanding of FECD, thus enabling development of novel therapies for this disabling disease.

### Table 1 | Major results from the FECD genome-wide association study.

| Gene    | SNP     | Chr | Pos | All. | Discovery cohort (n = 3,968) | Replication Cohorts (n = 1,449) | Combined (n = 5,417) |
|---------|---------|-----|-----|-----|-----------------------------|--------------------------------|------------------------|
|         |         |     |     |     | RAF case                     | OR (95% CI)                    | OR (95% CI)             |
| KANK4   | rs79742895 | 1   | 62782860 | C/T | 0.0889 | 0.0347 | 2.07 (1.69, 2.54) | 2.18 (1.36, 3.50) | 2.09 (1.73, 2.52) | 0.19 |
| ATP1B1  | rs1200014 | 1   | 169060489 | G/A | 0.4277 | 0.3411 | 1.43 (1.29, 1.57) | 1.53 (0.50, 0.72) | 1.46 (1.34, 1.58) | 0.99 |
| LAMC1   | rs3768617 | 1   | 183092500 | A/G | 0.3553 | 0.4214 | 0.75 (0.68, 0.82) | 1.08 (0.91, 1.27) | 1.25 (0.16, 1.36) | 0.02 |
| CFH     | rs2374700 | 1   | 196662947 | T/C | 0.4287 | 0.5309 | 1.31 (1.19, 1.44) | 2.10 (0.79, 1.07) | 2.08 (0.75, 0.88) | 0.24 |
| SLC2A2  | rs12223242 | 11  | 7727101 | A/G | 0.4826 | 0.5445 | 0.77 (0.70, 0.85) | 2.10 (0.91, 1.27) | 1.25 (0.16, 1.36) | 0.02 |
| TCF4    | rs784257  | 18  | 53397199 | A/G | 0.4779 | 0.5779 | 5.77 (5.05, 6.60) | 3.89 (3.30, 4.59) | 4.94 (4.45, 5.48) | 0.001 |

All, alleles with reference allele first; CI, confidence interval; Chr, chromosome; OR, odds ratio; Qp, P value from Cochran’s Q test for heterogeneity; Pos, hg37 physical map position; RAF case and RAF ctrl, reference allele frequency in cases and controls, respectively.

### Figure 2 | ROC curves for FECD status.

Curves are shown from genetic risk scores based on all replicated loci (black), rs784257 in TCF4 only (blue) and all loci except rs784257 (magenta).
Methods

Samples and phenotypes. Informed consent was obtained from all participants, in accordance with the Declaration of Helsinki, and all procedures were conducted after approval from the Institutional Review Boards at all participating centres.

FECD study subjects in the discovery sample were recruited through the FECD Genetics Multi-center Study, based at Case Western Reserve University, and at Duke University Eye Center (DUUCE)\(^5\).\(^6\).\(^7\).\(^8\).\(^9\). The phenotype criteria are listed in Supplementary Table 1. The sample originally comprised 2,710 individuals, of whom 2,089 (1,412 FECD cases and 677 controls) were included in association analysis (Supplementary Table 2 and see below). In addition, we included 1,879 FECD control participants from the AREDS Refractive Error Substudy (dBsGp accession phs000001.v3.p1), originally collected for an age-related macular degeneration clinical trial, but subsequently included in a GWAS study on refractive error; these subjects had no age-related macular degeneration, previous cataract surgery or evidence of corneal dystrophy\(^1\).\(^2\). Twenty-nine AREDS samples were submitted to both studies by both groups to ensure that genotype clustering was compatible between samples originally submitted by the AREDS group and those submitted to both studies by both groups to ensure that genotype clustering was compatible between samples originally submitted by the AREDS group and those submitted to both studies by both groups.

Supplementary Table 3 summarizes the characteristics of the discovery samples.

Severity of FECD was assessed in both discovery samples using a modified Krachmer grading scale\(^2\) (Supplementary Table 2). Discovery samples had undergone penetrating keratoplasty or Descemet stripping endothelial keratoplasty in at least one eye, or had at least grade 3 FECD by the modified Krachmer scale in at least one eye (although most individuals were symmetric). Because keratoplasty is an unreliable indicator of FECD severity (228 of 1,254 cases (18.2%) with no corneal oedema and a grade 4 or below\(^2\)), we conducted a sensitivity analysis restricting the case definition to include only FECD cases with histopathological verification of severe FECD (see below). Based on discussions between all three centres, the following inclusion criteria for controls have been established to determine subject eligibility from each ascertainment site: (1) 60 years of age or older, (2) normal cornea by slit-lamp biomicroscopy with no epithelial, stromal or endothelial abnormalities with the exceptions of (a) arcus senilis and (b) lamellar holes, (b) cornea scar from infection or penetrating trauma; (c) pterygium; inactive superficial vascularization of the epithelium or subepithelial zone; (3) previous cataract, glaucoma or retina–vitreous surgery was allowed, as long as the preoperative record indicated no evidence by slit-lamp examination of guttata or FECD and postoperative exam also showed no evidence; and (4) previous laser surgery (sclerostomy, trabeculostomy or pan retinal photocoagulation) was allowed with same caveats as intraocular surgery.

A total of 1,449 samples (671 FECD cases and 778 controls) from three study centres were included in the replication phase (Supplementary Table 5). Three centres participated: the University of Iowa (Iowa City, IA, USA), Flinders University (Adelaide, SA, Australia) and Johns Hopkins (JHU) (Baltimore, MD, USA). All three centres required controls to be 50 years or older, and cases to have Krachmer grade 1 (ref. 5) (modified Krachmer grade 2) FECD or higher (Supplementary Table 2). We used the grading scheme provided by the replication cohort to determine case-control status for cases used for the replication cohort. Detailed below are the grading criteria used by each cohort. Because of the scarcity of available replication samples, the case definition is in general less strict than for the discovery sample. The University of Iowa cohort comprised 113 patients with FECD and 113 control subjects, all Caucasian ancestry. All participants underwent detailed ophthalmic evaluation that included slit-lamp biomicroscopy. Patients were considered to have FECD if they had more than 12 central guttata or if they had histopathologic confirmation following corneal transplantation (keratoplasty) for corneal oedema\(^2\). Control subjects were examined by a board-certified ophthalmologist and did not have signs of FECD. A sample of ~10 ml blood was collected from each study participant.

The Flinders University cohort comprised 190 patients with FECD. Additionally, genomic DNA from 282 unrelated, unaffected South Australian residents aged over 50 years, recruited previously for use as controls in ocular genetics studies in our laboratory, was available for this study\(^4\).\(^6\).\(^7\).\(^8\).\(^9\).

The cases were Caucasian Australian patients diagnosed with modified Krachmer scale grade 2 or above FECD who were recruited by ophthalmologists at the Flinders Eye Clinic (Adelaide, SA, Australia) and the Royal Victorian Eye and Ear Hospital (Melbourne, Victoria, Australia). Venous blood samples were collected from 190 patients diagnosed using QIAGEN DNA Blood Maxi Kit (Qiagen, Doncaster, Victoria, Australia) according to the manufacturer’s protocol.

This study was approved by the Human Research Ethics Committees of the Flinders Medical Centre/Flinders University of South Australia (Adelaide, SA, Australia) and the Royal Victorian Eye and Ear Hospital. The research was conducted in accordance with the National Health and Medical Research Council, Australia guidelines.

The JHU cohort consisted of 368 Caucasian FECD cases and 380 ethnically matched control subjects. All participants underwent a detailed ophthalmic evaluation and were genotyped using the well-characterized TCF4 SNP rs613872, which was not specifically selected for follow-up, but genotype data were available for the SNP in the Flinders University sample. rs2296292 and rs20560 in the Flinders University sample. rs2274700 failed to genotype.

For the Flinders University sample, DNA was extracted using the Gentra Puregene Blood Kit (Qiagen, Santa Clara, CA, USA).

Genotyping. For the discovery samples was performed on the Illumina HumanOmni2.5-4v1_H array at the Center for Inherited Disease Research (CIDR, Baltimore, MD, USA). DNA samples on 96-well plates were balanced by FECD status, collection site (Case Western Reserve University multi-center versus DUUCE) and sex.

A total of 2,841 FECD case–control samples, including duplicates, were submitted for genotyping, of which 2,818 passed initial genotyping QC by CIDR. The 1,922 AREDS control DNA samples, including 29 cross-study duplicate pairs and 10 internal duplicate pairs, were processed separately at CIDR, but genotype calling was performed with both datasets together.

Eighteen SNPs were chosen for follow-up based on the genetic structure and annotation information for the most significant results from the Omni12.5 M and imputed GWAS within each region yielding P = 5 × 10\(^{-8}\) for association for at least one marker. The LD structure was characterized, using genotype data from our FECD controls only, among the five SNPs with the most significant associations from the Omni12.5 M, and at least five SNPs from the imputed data. From each region we chose two or more SNPs for replication: the most significant SNP plus one or more additional SNPs and, where available, highly significant eonic variants. Follow-up SNPs were chosen to minimize LD within association regions; SNPs within often run on a chromosome peak were excluded (see below). Positive disease status was determined with the Krachmer scale grade \(\geq 2\) FECD in 18 SNPs in or near KANK4 (rs12082238, rs12058446, rs79742895), APTIP1B (rs12010114, rs1200115), LAMCI (rs2296922, rs7368617, rs1413386, rs20560, rs20561), ATP2B1 (rs2274700), SLC24A2 (rs12223324, in PDDC1; rs4963153; rs138714, in PNPLA2) and TCF4 (rs1639715, rs1639713). The well-characterized SNP rs613872 was not specifically selected for follow-up, but genotype data were available for the SNP in the Flinders University sample. rs2296292 and rs20560 in LAMCI were chosen specifically because they were coding variants (A992A and R1376R, respectively).

For the University of Iowa replication sample, DNA was extracted using the Gentra Puregene Blood Kit (Qiagen, USA). Samples were submitted to LGC Genomics (Petaluma, CA, USA) for genotyping using the KASP assay, an end-point polymerase chain reaction (PCR) assay using two competitive, allele-specific forward primers and a common reverse primer. The average call rate over 17 assays was 99.5%; rs2274700 failed to genotype.

For the Flinders University sample, replication data for the SNP rs613872 were obtained for 190 FECD cases and 282 unaffected controls. All the SNPs were genotyped in cases and controls using IPLEX GOLD chemistry (Sequenom, Herston, QLD, Australia) on a MassARRAY Compact Spectrometer (Sequenom).

For genotyping of the Johns Hopkins University sample, the PCR was performed in 5-µl volumes containing 10 ng genomic DNA, 2.5 µl SNP genotyping master mix (TaqMan; Applied Biosystems, Foster City, CA, USA) and 0.125 µl genotyping assay mix (TaqMan; Applied Biosystems). Reactions for all nine SNPs were amplifed independently in a thermocycler (9700; Applied Biosystems). The cycling conditions consisted of 2-min incubation at 95 °C and denaturation at 95 °C for 10 min, followed by 40 cycles of 10 s at 95 °C and 1-min elongation at 72 °C with a final 10-min extension at 72 °C. Amplified products were analysed for the enrichment of specific alleles (ABI 7900HT Sequence Detection System; Applied Biosystems).

For the Flinders University sample, we genotyped for the replication SNPs were obtained for 190 FECD cases and 282 unaffected controls. All the SNPs were genotyped in cases and controls using IPLEX GOLD chemistry (Sequenom, Herston, QLD, Australia) on a MassARRAY Compact Spectrometer (Sequenom).

GWAS quality control. Genotype data were subject to intensive quality control by the Quality Assurance/Quality Control analysis team at the University of Washington Genetics Coordinating Center. Genotype concordance between duplicate samples was high; the median discordance rate over 102 duplicate pairs was 1.5 × 10\(^{-3}\). All samples passing quality control had call rates above 98%, and the median call rate was 99.7%.

Relatedness between pairs of samples was estimated by means of the method-of-moments approach\(^20\) implemented in the R package SNPRelate, on a subset of 124,583 autosomal SNPs in low linkdase disequilibrium. A total of 15 samples were omitted on account of unresolved identity issues, including disagreement of annotated Mendelian sex and of unexpected and unresolved relationships. Samples were classified as ‘related’ if their estimated kinship coefficient was >1/12 (equivalent to half the expected value for first cousins). Association analysis
were run by the same logistic regression model as described above, except that sex (Supplementary Fig. 2).

The area under the curve (AUC) only at the most significantly associated SNP at each of the six candidate gene (Supplementary Table 6) was performed by the inverse-variance weighting method, compared using Fisher’s exact test. Meta-analysis over three replication samples was carried out at the University of Washington Genetics Coordinating Center. Individual significance of association was measured using a Wald test. In testing X-linked SNPs, the approach of Patterson et al. was omitted from association analyses. Anomalies of the X chromosome were included a total of 2,089 unique, unrelated samples of European descent from the discovery sample. Genetic sex was determined by means of X-chromosome heterozygosity and the operating characteristic (ROC) curves were computed, using the R package pROC. Sex-specific genetic risk scores were calculated by the same procedure, using effect estimates derived from Supplementary Table 13.

Pairwise SNP-SNP interactions were tested among pairs of SNPs from Fig. 1 by including two genotypes and an interaction term in the association model for the discovery sample.

**Gene expression and IHC.** Surgical specimens of corneal endothelium and Descemet’s membrane (CE + DM) complex were obtained from patients undergoing keratoplasty because of advanced FECD and placed directly into RNAlater Solution (Life Technologies, Mulgrave, VIC, Australia). The RNA from these tissues was used for sequencing. Equivalent control specimens were obtained through the Eye Bank of Sydney, Australia (Eye Bank of Sydney, Centrelink, Redfern, Sydney, NSW, Australia). The specimens were placed in RNAlater Solution immediately after surgery or dissection, and stored at 4 °C for at least two days. Then RNAlater solution was removed and the specimen stored at −80 °C for later gene expression analysis.

Normal, eye-bank-derived, post-mortem corneas in Optisol GS from three unrelated subjects were obtained for gene expression analysis from Eversholt Ohio (Cleveland, OH, USA). The storage solution was discarded and tissues were placed into RNAlater Solution and frozen on receipt. For each of the six loci from the initial GWAS, the set of GENCODE genes (accessed through Ensembl Biomart, http://www.ensembl.org) within and overlapping an interval of 1 Mb centred on the most significant marker was included in the analysis. RNA was quality checked with the Bio-Rad Experian/Agilent Bio-analysers. Samples passing initial quality control were subject to rRNA depletion using RiboZero Gold (Epicentre, San Diego, CA, USA) according to the manufacturer’s instructions. rRNA-depleted fractions were used as input for library preparation using ScriptSeq Gold Low Input Library Preparation Kit (Epicentre). All libraries were quantified using the Illumina PCR-freeT assay (Life Technologies, Carlsbad, CA, USA) and RNA integrity was checked with the Bio-Rad Experion/Agilent Bio-analysers. Samples passing initial quality control were subject to Illumina cDNA synthesis using ScriptSeq Gold Low Input Library Preparation Kit (Epicentre).

The primary immunohistochemical analysis, formalin-fixed, paraffin-embedded sections, 4–5 μm thick, of five FECD-affected and five normal corneas were used. Immunohistochemical labelling on sections of affected and normal corneas was performed as follows. Primary antibodies were applied to each section and incubated for 1.5 h in a chamber at 37 °C. The dilutions were as follows: ATP1B1 (LifeSpan Biosciences; LS-B235), 1:1,000; LAMC1 (LifeSpan Biosciences; LS-B7354), 1:50; LAMC2 (LifeSpan Biosciences; LS-0027), 1:50; KANK4 (Abcam; ab121410), 1:100. We used goat anti-rabbit horseradish peroxidase antibody (Abcam; ab6721) diluted 1:400 and goat anti-mouse horseradish peroxidase antibody (Fisher Scientific; PI14306) diluted 1:200 as the secondary antibodies. Sections were mounted with DAB Quanto (Thermo Fisher Scientific; TA-060-QHDX), were counterstained with haematoxylin and were mounted in Cytoseal 60 (Thermo Fisher Scientific; 1001000).

All libraries were quantified using the Illumina HiSeq 2500. The reads were mapped to the reference human genome (hg38) using tophat (https://ccb.jhu.edu/software/tophat/index.shtml). Read counts were performed for features in the GenCode version 21 database using htseq-count (https://www-huber.embl.de/users/anders/HTSeq/doc/count.html). The expression level was normalized for each gene by calculating reads per kilobase per million reads. We defined ‘high expression’ as >10 reads per kilobase per million reads averaged over the three replicates.

For the primary immunohistochemical analysis, formalin-fixed, paraffin-embedded sections, 4–5 μm thick, of five FECD-affected and five normal corneas were used. Immunohistochemical labelling on sections of affected and normal corneas was performed as follows. Primary antibodies were applied to each section and incubated for 1.5 h in a chamber at 37 °C. The dilutions were as follows: ATP1B1 (LifeSpan Biosciences; LS-B235), 1:1,000; LAMC1 (LifeSpan Biosciences; LS-B7354), 1:50; LAMC2 (LifeSpan Biosciences; LS-0027), 1:50; KANK4 (Abcam; ab121410), 1:100. We used goat anti-rabbit horseradish peroxidase antibody (Abcam; ab6721) diluted 1:400 and goat anti-mouse horseradish peroxidase antibody (Fisher Scientific; PI14306) diluted 1:200 as the secondary antibodies. Sections were mounted with DAB Quanto (Thermo Fisher Scientific; TA-060-QHDX), were counterstained with haematoxylin and were mounted in Cytoseal 60 (Thermo Scientific, 3812-4). Images were captured on an Olympus BX-2 microscope (Olympus, Tokyo, Japan) fitted with a Canon EOSD 20.2 Megapixel Digital Camera (Canon, Tokyo, Japan) using DSR Remote Pro Imaging Software (Breeze Systems, Surrey, UK).

Labelling on sections of affected and normal corneas for confirmatory immunohistochemical analysis was performed as above, but without any antigen retrieval. Sections were hybridized with the mouse anti-human ATP1B1 (clone M17–P5–F11, catalogue no. PIEMA3–930, Thermo Scientific Pierce, supplied by SABiosciences; 1:1,000 dilution) or rat anti-human LAMC1 (catalogue no. MAB2692, Chemicon International, CA, USA) overnight at room temperature. Monoclonal antibody (1:50) at 4 °C overnight. This was followed by hybridization with the NovoLink Polymer complex reagent (Leica Microsystems, Bannockburn, IL, USA). Primary antibody binding was detected with Liquid DAB + substrate Chromogen System (K3648; Dako Australia). Sections were counterstained with haematoxylin and mounted in DePeX (Merck KGaA, Darmstadt, Germany). Labelling was imaged on an Olympus BX50 microscope fitted with QImaging Micropublisher RTV 5 Megapixel Digital Camera using QCapture Imaging Software (Olympus).
Data availability. We have deposited all genotype data supporting our findings from the discovery cohort in the Database of Genotypes and Phenotypes (dbGaP), with accession code phs000421.v1.p1. Other data that support our findings are available from the authors by request; see author contributions and their published references for specific data sets.

References
1. Eye Bank Association of America. 2015 Eye Banking Statistical Report (Eye Bank Association of America, Washington, DC, 2016).
2. Fuchs, E. Dystrophia epithelialis corneae. Arch. Clín. Exp. Ophthalmol., 76, 478–508 (1910).
3. Musch, D. C., Nizioł, L. M., Stein, J. D., Kamyar, R. M. & Sugar, A. Prevalence of corneal dystrophies in the United States: estimates from claims data. Invest. Ophthalmol. Vis. Sci. 52, 6995–6993 (2011).
4. Gotschj, J. D. et al. Inheritance of a novel COL2A2 mutation defines a distinct early-onset subtype of Fuchs corneal dystrophy. Invest. Ophthalmol. Vis. Sci. 36, 1934–1939 (2000).
5. Krämer, J. H., Purell, Jr J. J., Young, C. W. & Bucher, D. K. Corneal endothelial dystrophy. A study of 64 families. Arch. Ophthalmol. 96, 2036–2039 (1978).
6. Afsahi, N. A., Pittard, A. B., Siddiqui, A. & Klintworth, G. K. Clinical study of Fuchs corneal endothelial dystrophy leading to penetrating keratoplasty: a 30-year experience. Arch. Ophthalmol. 124, 777–780 (2006).
7. Klintworth, G. K. Corneal dystrophies. Orphanet J. Rare Dis. 4, 7 (2009).
8. Bahn, C. F. et al. Classification of corneal endothelial disorders based on neural crest origin. Ophthalmology 91, 558–563 (1984).
9. Louttit, M. D. et al. A multi-center study to map genes for Fuchs’ endothelial corneal dystrophy: baseline characteristics and heritability. Cornea 31, 26–35 (2012).
10. Biwas, S. et al. Missense mutations in COL2A2, the gene encoding the α2 chain of type VIII collagen, cause two forms of corneal endothelial dystrophy. Hum. Mol. Genet. 10, 2415–2423 (2001).
11. Vithana, E. N. et al. Mutations in sodium-borate cotransporter SLCA11 cause recessive congenital hereditary endothelial dystrophy (CHED2). Nat. Genet. 38, 755–757 (2006).
12. Riazuddin, S. et al. Missense mutations in TCFC8 cause late-onset Fuchs corneal dystrophy and interact with FCD4 on chromosome 9p. Am. J. Hum. Genet. 86, 45–53 (2010).
13. Baratz, K. H. et al. E2-2 protein and Fuchs’ corneal dystrophy. N. Engl. J. Med. 369, 1035–1044 (2013).
14. Igo, Jr R. P. et al. Differing roles for TCFC8 and COL2A2 in central corneal thickness and Fuchs endothelial corneal dystrophy. PLoS ONE 7, e46742 (2012).
15. Li, Y. et al. Replication of TCFC8 through association and linkage studies in late-onset Fuchs endothelial corneal dystrophy. PLoS ONE 6, e18044 (2011).
16. Wieren, E. D. et al. A common trinucleotide repeat expansion within the transcription factor 4 (TCFC) E2-2 gene predicts Fuchs corneal dystrophy. PLoS ONE 7, e90083 (2012).
17. Stambolian, D. et al. Meta-analysis of genome-wide association studies in five cohorts reveals common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive error. Hum. Mol. Genet. 22, 2754–2764 (2013).
18. Riazuddin, S. et al. Replication of the TCFC4 intrinsic variant in late-onset Fuchs corneal dystrophy and evidence of independence from the FCD2 locus. Invest. Ophthalmol. Vis. Sci. 52, 2825–2829 (2011).
19. Thalamountu, A. et al. Association of TCFC4 gene polymorphisms with Fuchs corneal dystrophy in the Chinese. Invest. Ophthalmol. Vis. Sci. 52, 5573–5578 (2011).
20. Kuot, A. et al. Association of TCFC4 and CLU polymorphisms with Fuchs’ endothelial dystrophy and implication of CLU and TGFβ1 proteins in the disease process. Eur. J. Hum. Genet. 20, 632–638 (2012).
21. Mootha, V. V., Gong, X., Ku, H.-C. & Xing, C. Association and familial segregation of TCFC8.1 trinucleotide repeat expansion of TCFC8 gene in Fuchs’ endothelial corneal dystrophy. Invest. Ophthalmol. Vis. Sci. 55, 33–42 (2014).
22. Du, J. et al. RNA toxicity and missplicing in the common eye disease Fuchs endothelial corneal dystrophy. J. Biol. Chem. 290, 5979–5990 (2015).
23. Mootha, V. V. et al. TCFC4 triplet repeat expansion and nuclear RNA foci in Fuchs’ endothelial corneal dystrophy. Invest. Ophthalmol. Vis. Sci. 56, 2003–2011 (2015).
24. Joshi, M. A. & Barnett, J. C. Genetic risk prediction in complex disease. Hum. Mol. Genet. 20, R182–R188 (2011).
25. Delong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845 (1988).
26. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 40, D930–D934 (2012).
27. Gee, H. Y. et al. KANK deficiency leads to podocyte dysfunction and nephrotic syndrome. J. Clin. Invest. 125, 2375–2384 (2015).
Acknowledgements
We would like to acknowledge the strong scientific know-how and mentorship provided by Dr Gordon Klintworth, who passed away during the writing of this paper. We are incredibly grateful for the wisdom he imparted. We are also grateful to the patients who provided us with information and samples, and made this study possible. Finally, we would like to show appreciation to the many funding agencies who supported this research, including National Eye Institute Grants R01 EY023196, R01 EY016482, R01 EY016514, R01 EY016835, R21 EY015145 and P30 EY11373; Research to Prevent Blindness; the Beulah and Florence Usher Endowment; NHMRC (Australia) Grant No 1031362; Ophthalmic Research Institute of Australia; Flinders Medical Centre Foundation; the Ohio Lions Eye Research Foundation; and the intramural research program of the National Human Genome Research Institute, National Institutes of Health (NIH). Genotyping was performed by the Centre for Inherited Disease Research under Grant 1X01HG006619. This publication was made possible by the Clinical and Translational Science Collaborative of Cleveland, 4ULTR000439, from the National Centre for Advancing Translational Sciences (NCATS) component of the NIH and NIH Roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. This article was prepared while M.A.M. was employed at Duke University; she is currently serving as an American Association for the Advancement of Science (AAAS) Science & Technology Policy Fellow at the NIH. She worked on this article in her personal capacity, and the opinions expressed in this article are her own and do not necessarily reflect the views of the NIH, the Department of Health and Human Services or the United States government.

Author contributions
N.A.A., R.P.I., G.K.K., J.H.L., Y.J.L. and S.K.I. conceived the project. N.A.A., D.S., C.R.C., E.B., G.O.R., K.H.B., V.V.M., S.V.P., J.E.C., A.J.A., G.K.K., J.H.L. and S.K.I. performed clinical phenotyping. R.P.I., N.J.M., S.S., B.J.T., J.R., A.K., X.Q., K.P.B., S.A.R. and Y.J.L. performed statistical analyses. D.S., S.S., V.L.P., S.D., J.F.S., B.J.T., A.K., C.R.C., K.P.B., S.A.R., S.K., M.A.M., J.Z., E.B., S.G.G., J.E.B.-W., M.O.P., F.W.P., J.E.C. and G.K.K. did the genotyping and molecular laboratory work. N.A.A., R.P.I., Y.J.L. and S.K.I. drafted the paper. D.S., S.S., C.R.C., K.P.B., M.A.M., K.H.B., J.E.B.-W., M.O.P., A.J.A. and J.H.L. critically reviewed the manuscript.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Afshari, N. A. et al. Genome-wide association study identifies three novel loci in Fuchs endothelial corneal dystrophy. Nat. Commun. 8, 14898 doi: 10.1038/ncomms14898 (2017).

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