This is a repository copy of *Enrichment of pathogenic alleles in the brittle cornea gene, ZNF469, in keratoconus.*

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/87454/

Version: Accepted Version

**Article:**  
Lechner, J, Porter, LF, Rice, A et al. (10 more authors) (2014) Enrichment of pathogenic alleles in the brittle cornea gene, ZNF469, in keratoconus. Human Molecular Genetics, 23 (20). 5527 - 5535. ISSN 0964-6906

https://doi.org/10.1093/hmg/ddu253

---

**Reuse**  
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**  
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Enrichment of pathogenic alleles in the brittle cornea gene, *ZNF469*, in keratoconus

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Human Molecular Genetics</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>HMG-2014-D-00145.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>2 General Article - UK Office</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Complete List of Authors:**
- Willoughby, Colin; University of Liverpool, Faculty of Health & Life Sciences
- Lechner, Judith; Queen's University Belfast, Centre for Experimental Medicine
- Porter, Louise; Institute of Human Development, University of Manchester
- Rice, Aine; University of Leeds, Leeds Institute of Molecular Medicine
- Vitart, Veronique; University of Edinburgh, MRC Human Genetics Unit, IGMM
- Armstrong, David; Royal Victoria Hospital, Dept of Ophthalmology
- Schorderet, Daniel; Institute for Research in Ophthalmology, Department of Ophthalmology, University of Lausanne
- Munier, Francis; Jules-Gonin Eye Hospital, Ophthalmology
- Wright, Alan; University of Edinburgh, MRC Human Genetics Unit, IGMM
- Inglehearn, Chris; University of Leeds, Molecular Medicine
- Black, Graeme; University of Manchester, Institute of Human Development
- Simpson, David; Queen's University Belfast, Centre for Experimental Medicine
- Manson, Forbes; The University of Manchester, Institute of Human Development

**Key Words:**
- keratoconus, ZNF469, corneal thickness, Corneal Dystrophies, Hereditary, mutation
TITLE PAGE

Enrichment of pathogenic alleles in the brittle cornea gene, ZNF469, in keratoconus

Judith Lechner1,3, Louise F. Porter2,3, Aine Rice4, Veronique Vitart5, David J. Armstrong6, Daniel F. Schoderet7,8,9, Francis L. Munier7,10, Alan F. Wright5, Chris Inglehearn4, Graeme Black2, David A. Simpson1, Forbes Manson2,12, Colin E. Willoughby11,12,*.

*Corresponding Author:
Prof Colin E. Willoughby
Department of Eye and Vision Science,
Institute of Ageing and Chronic Disease,
Faculty of Health & Life Sciences,
University of Liverpool,
3rd floor, UCD Building,
Daulby Street,
Liverpool, L69 3GA,
United Kingdom
Tel: +44 (0)151 706 4070
Email: c.willoughby@liverpool.ac.uk
1Centre for Vision and Vascular Science, Queen’s University Belfast, Belfast, Northern Ireland, BT12 6BA, United Kingdom

2Institute of Human Development, The University of Manchester, Manchester Academic Health Science Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, M13 9WU, United Kingdom

3Joint First Authors

4Leeds Institute of Molecular Medicine, St. James's University Hospital, University of Leeds, Leeds LS9 7TF, United Kingdom

5MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, Scotland, United Kingdom

6Department of Ophthalmology, Royal Victoria Hospital, Belfast, Northern Ireland, BT12 6BA, United Kingdom

7IRO - Institute for Research in Ophthalmology, Sion, Switzerland

8Department of Ophthalmology, University of Lausanne, Lausanne, Switzerland

9Faculty of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

10Jules-Gonin Eye Hospital, Lausanne, Switzerland

11Department of Eye and Vision Science, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, L69 3GA, United Kingdom

12These authors contributed equally
ABSTRACT

Keratoconus, a common inherited ocular disorder resulting in progressive corneal thinning, is the leading indication for corneal transplantation in the developed world. Genome-wide association studies have identified common SNPs 100kb upstream of ZNF469 strongly associated with corneal thickness. Homozygous mutations in ZNF469 and PRDM5 genes result in brittle cornea syndrome type 1 and type 2 respectively. Brittle cornea syndrome is an autosomal recessive generalized connective tissue disorder associated with extreme corneal thinning and a high risk of corneal rupture. Some individuals with heterozygous PRDM5 mutations in these brittle corneal syndrome genes demonstrate a carrier ocular phenotype, which includes a mildly reduced corneal thickness, keratoconus and blue sclera. We hypothesized that heterozygous variants in PRDM5 and ZNF469 predispose to the development of isolated keratoconus. We found a significant enrichment of potentially pathologic heterozygous alleles in ZNF469 associated with the development of keratoconus (P=0.00102) resulting in a relative risk of 12.0. This enrichment of rare potentially pathogenic alleles in ZNF469 in 12.5% of keratoconus patients represents a significant mutational load, and highlights ZNF469 as the most significant genetic factor responsible for keratoconus identified to date.
INTRODUCTION

Keratoconus (MIM 148300), a common bilateral, progressive corneal thinning disorder (1), is the leading indication for corneal transplantation in the developed world, accounting for 25% of the 2500 corneal transplants performed annually in the UK and a similar proportion of the 32000 grafts performed annually in the USA (2). Keratoconus usually arises in the teenage years and presents a significant health burden in work-age adults. The minimum incidence is 1 in 2000 but it is much more common in some ethnic groups (1, 3). There is strong evidence for a heritable component in the development of keratoconus (4, 5). Most studies describe autosomal dominant inheritance, with incomplete penetrance or variable expressivity (4). However, in a genetic modelling study in a multi-ethnicity population a major recessive genetic defect was the most parsimonious (6), although no recessive loci for keratoconus have been described to date.

The progressive corneal thinning associated with keratoconus (mean central corneal thickness 450-500 μm) (7) results in myopia and irregular corneal astigmatism. In healthy humans central corneal thickness (CCT) is a normally distributed quantitative trait with a mean of 536 μm ± 31 μm (8) which has an estimated heritability up to 95% (9). Genome-wide association studies (GWAS) in the healthy European and Asian populations have identified CCT-associated loci, with common SNPs upstream of zinc finger 469 (ZNF469 [MIM 612078]) the most strongly associated with CCT (10-14). Mutations in three genes (ZNF469, COL5A1 and COL8A2) close or within these identified loci are responsible for rare Mendelian conditions that affect the corneal structure: brittle corneal syndrome, Ehlers-Danlos syndrome and posterior polymorphous corneal dystrophy respectively (10-12).

Brittle cornea syndrome (BCS) is an autosomal recessive generalized connective tissue disorder associated with extreme corneal thinning (220-450 μm) and a high risk of corneal
Homozygous mutations in ZNF469 and PR domain-containing protein 5 (PRDM5 [MIM 614161]) genes result in brittle cornea syndrome type 1 (BCS1 [MIM 229200]) (17) and brittle cornea syndrome type 2 (BCS2 [MIM 614170]) (15) respectively. Some individuals with heterozygous PRDM5 mutations demonstrate a carrier ocular phenotype which includes a mildly reduced CCT (480-505μm), keratoconus and blue sclera (15). In one family with BCS2 there was a relationship between the severity and age of onset of keratoconus and PRDM5 mutational status. Family members with a homozygous PRDM5 mutation (deletion of exons 9–14) developed early and severe keratoconus whereas one heterozygous family member developed keratoconus which was clinically milder with a later onset (15). The relationship between the degree of CCT reduction and the presence of homozygous or heterozygous mutations in PRDM5 suggested a dosage effect (15). Although none of the CCT-associated loci have been mapped to PRDM5, a common SNP (rs10518367) which is 70 kb upstream of PRDM5 has been associated with CCT in the European population at the significance level of $P = 8.9 \times 10^{-5}$.

Given that rare ZNF469 and PRDM5 homozygous mutations result in the extreme corneal thinning disorder (BCS), and that there was evidence of a carrier ocular phenotype (keratoconus and corneal thinning) in some individuals with PRDM5 heterozygous mutations, and that common SNPs 100kb upstream of ZNF469 are strongly associated with CCT, we undertook Sanger sequencing of both genes in patients with isolated keratoconus; a common ocular disease characterised by progressive corneal thinning and ectasia.
RESULTS

Heterozygous PRDM5 mutations in carrier individuals from BCS2 families can result in keratoconus (15) and so the entire coding region and intron-exon junctions of PRDM5 (exons 1-16) were sequenced in an initial 96 unrelated European patients with keratoconus. This analysis failed to identify any pathogenic variants, although eight known SNPs were detected (rs146268537, rs74320998, rs343192, rs17051264, rs34666716, rs12499000, rs75893420 and rs55774575). We therefore proceeded to Sanger-sequence ZNF469 in the original cohort increased with additional keratoconus cases (total number of cases = 112) of unrelated European patients with keratoconus from three study centres (Belfast, Leeds and Lausanne).

Sequence variants of unknown significance (VUS) detected by Sanger sequencing were classified as potentially pathogenic alleles by filtering using (i) ethnically matched population specific control data from 784 individuals (outlined in the Methods); (ii) the data from dbSNP (Build 137), the May 2012 release of the 1000 Genomes (1KG) Project and the Exome Variant Server (EVS), NHLBI Exome Sequencing Project (ESP), with no allele having a minor allele frequency (MAF) of > 0.1%; and (iii) classified as damaging using the Sorting Intolerant from Tolerant (SIFT) program as outlined in the Methods and Figure 2.

From this stringently filtered sequencing data 12 potentially pathogenic non-synonymous heterozygous alleles were detected in the keratoconus cohort (Table 1) and 2 in-frame deletions: c.2904_2909delGTCGG; p.Ser969_Gly970del and c.9011_9025delTTCCCGGAACACCC; p.Leu3004_Thr3008del. In the keratoconus cohort following filtering there remained 15 non-synonymous classified as tolerated by SIFT which were absent from control data and had a MAF < 0.1% (Table 2). On the basis of poor conservation 6 of these variants were classified as polymorphisms leaving 9 variants of unknown significance. We detected 34 non-synonymous and 33 synonymous variants which
were observed in both cases and controls, had a MAF ≥ 0.1% or were common variants, and were deemed non-pathogenic (Supplemental Table 1, available online). Overall, this study identified 34 novel variants (29 non-synonymous and 5 synonymous) which have been submitted to the NCBI dbSNP (Supplemental Table 2, available online). The severity of keratoconus was graded using the Amsler-Krumeich classification (18, 19) and 10 individuals had grade stage III or above indicating severe disease (illustrated in Figure 1). Stages III and IV usually require surgical approaches for visual rehabilitation and 3 individuals required corneal transplantation.

As there was the possibility that the alleles identified in the keratoconus subjects represented chance events, we Sanger-sequenced the complete coding sequence of ZNF469 (13,203 bp) in 96 unaffected and unrelated European control samples (192 chromosomes) using the same experimental stringency as the case sequencing, and detected 1 non-synonymous heterozygous allele (c.1701G>T; p.Gln567His) deemed potentially pathogenic given our filtering criteria (Table 1). There was a statistically significant enrichment of potentially pathogenic ZNF469 alleles in the keratoconus subjects (14 variants) compared with the 96 European controls (1 variant); P=0.00102 (Odds Ratio 13.6, Relative Risk 12.0). The allele frequency differences make it impossible for the rare ZNF469 alleles to be in linkage disequilibrium with the common variant signal which is within a 53 kb linkage disequilibrium block 117kb away from the 5' end of ZNF46, and this has been replicated in diverse ancestries groups. The common variant, although strongly associated with corneal thickness, is not strongly associated with keratoconus (OR 1.25(95% CI 1.11-1.40)) (13.). Further functional studies and assays are required to confirm the pathogenicity of all alleles absent from ethnically matched controls and the population control data (dbSNP, EVS, 1KG).
DISCUSSION

Mutations in four genes, VSX1 (MIM 605020)(20, 21), SOD1 (MIM 147450)(21, 22), MIR184 (MIM 613146)(23) and ZEB1 (MIM 189909) (24), have been implicated in the pathogenesis of keratoconus in a minority of cases (< 4%) (21, 25). Two GWAS have been conducted in keratoconus cohorts which identified a SNP rs4954218, located near the RAB3GAP1 gene (MIM 602536)(26), and polymorphisms in HGF (MIM 142409)(27), associated with keratoconus susceptibility, but neither study reported genome-wide significant association. The identification of alleles that are predicted to be potentially pathogenic in 12.5% of keratoconus patients (11/112) makes ZNF469 the most significant genetic factor responsible for keratoconus identified to date. The small sizes of unrelated keratoconus cases cohorts favour a candidate genes approach so that while the p-value obtained for evidence of a burden of ZNF469 rare damaging variants in the keratoconus cases is statistically significant, it would not have been in a genome-wide context.

Brittle cornea syndrome is a rare recessive connective tissue disorder associated with consanguinity, with most BCS patients originating from countries in the Middle East and North Africa(16). ZNF469, the gene for BCS1, was originally mapped to chromosome 16q24 in a single large Palestinian family and a homozygous frameshift mutation, (c.9527delG) predicted to result in a premature termination codon (p.Gln3178ArgfsX23), was subsequently reported (17). Five further homozygous ZNF469 mutations have been reported in the literature: a founder mutation in five Tunisian patients (c.5934delA) predicted to result in a premature termination codon (p.Gly1983AlafsX16), p.Gln1392X (Syrian origin) (28), p.Phe717SerfsX14(15), p.Gln1757X (15) and one homozygous missense mutation (p.Cys3339Tyr) in a consanguineous Norwegian family (29). Homozygous mutations in
PRDM5 result in BCS2 and in some families PRDM5 heterozygous gene carriers display an ocular carrier state which includes a mildly reduced central corneal thickness (CCT; 480-505μm), keratoconus and blue sclera (15). Families harboring homozygous and heterozygous PRDM5 mutations show a gene dosage relationship in terms of the degree of CCT reduction and the severity and age of onset of keratoconus(15). There is no data available of CCT measurement or corneal topography for heterozygous carriers of ZNF469 mutations in BCS1 families. Our data for ZNF469 in the keratoconus population mirrors that seen in the ocular phenotype of PRDM5 heterozygous carriers. Heterozygous ZNF469 pathological alleles result in progressive corneal thinning and ectasia producing the keratoconus phenotype. The majority of potentially pathogenic alleles in the keratoconus cohort were missense variants likely to have a less deleterious effect on protein function than the ZNF469 truncating mutations commonly associated with BCS1. This further supports a gene dosage phenomenon wherein homozygous, severely deleterious ZNF469 mutations result in an early onset severe and visually devastating ocular phenotype (extreme corneal thinning, ectasia and spontaneous rupture) (15), whereas heterozygous, missense ZNF469 mutations result in corneal thinning, ectasia and keratoconus. Further functional studies and cell based assays are required to interrogate the molecular pathology and mutational mechanisms associated with these potentially pathogenic ZNF469 alleles.

ZNF469 is a 3925 amino acid evolutionarily poorly conserved C2H2 zinc finger (C2H2-ZNF) protein of unknown function(30). C2H2-ZNF genes constitute the largest class of transcription factors in humans making up ~2% of all the human genes and represent the second largest gene family in humans (30). The first identified members of the C2H2-ZNF family were xenopus TFIIIA and drosophila Kruppel and thus genes of this family are often called zinc finger genes of the TFIIIA or Kruppel type(30-32). Most C2H2-ZNF genes code for transcription factors which can bind DNA, RNA, DNA-RNA hybrids and proteins (32).
The physiological role of ZNF469 is not well established but there is evidence that ZNF469 regulates extracellular matrix development and maintenance (15). ZNF469 shows 30% sequence similarity to the helical parts of COL1A2 (MIM 120160), COL1A1 (MIM 120150) and COL4A1 (MIM 120130), all of which are highly expressed in the cornea (17). The cornea is composed of 70% collagen, mostly collagen type I, and there is evidence of a dysregulation of collagen homeostasis in the keratoconic cornea (33, 34). Corneal thinning has been reported in osteogenesis imperfecta (35) which results from mutations in COL1A1 or COL1A2.

We have identified an enrichment of potentially pathogenic alleles in ZNF469 in patients with keratoconus. Further work is required to determine the functional impact of these variants and the pathways regulated by ZNF469 which are involved in the development of keratoconus. Identifying genes responsible for keratoconus may also provide insights into the genetic basis for the normal variation in CCT. Decreased CCT has been proposed as a risk factor for primary open angle glaucoma (POAG [MIM 137760]), the leading cause of irreversible blindness worldwide affecting more than 60 million people (36). Individual patients with a thin cornea have a substantially increased risk for developing POAG (37, 38), and glaucoma patients with a thin CCT have an increased severity and more rapid progression of visual field loss (39). The genetic basis of CCT may provide insights into the development of glaucoma. Common SNPs near ZNF469 are the strongest CCT-associated loci although the functional role of these SNPs is not known (10-13). The role ZNF469 plays in the development of POAG and maintenance of CCT in normal subjects has not been determined. Combining resequencing with GWAS has yielded success in identifying rare disease associated variants (40, 41). Our study establishes the significant role ZNF469 plays in the development of keratoconus.
MATERIALS AND METHODS

All studies adhered to the tenets of the Declaration of Helsinki, and were approved by the relevant institutions with all participants giving written informed consent.

Patients

Clinically affected keratoconus patients of European ethnicity were recruited as part of ongoing studies from Belfast (Belfast Health and Social Care Trust, UK), Leeds (St. James’s University Hospital, Leeds, UK) and Lausanne (Jules-Gonin Eye Hospital, Lausanne and Institute for Research in Ophthalmology, Sion, Switzerland); and genomic DNA was extracted from peripheral blood leukocytes using commercial kits. The diagnosis of keratoconus was performed by an experienced ophthalmologist based on well-established clinical signs on slit-lamp biomicroscopy and cycloplegic retinoscopy; and a confirmatory videokeratographic map obtained using the Topographic Modelling System-1 (Computed Anatomy Inc, NY, USA), Orbscan II (Bausch & Lomb, Salt Lake City, UT, USA) or the Pentacam (Oculus, Wetzlar, Germany) (20, 27). Slit-lamp biomicroscopy was used to identify the key features of keratoconus including stromal corneal thinning, Vogt’s striae, and Fleischer rings in affected individuals. The oil droplet sign and scissoring of the red reflex were assessed by retinoscopy performed with a fully dilated pupil. Patients were considered as having keratoconus if they had at least one clinical sign of the disease in conjunction with a confirmatory videokeratography map (20, 27). The severity of keratoconus was graded using the Amsler-Krumeich classification (18, 19):
<table>
<thead>
<tr>
<th>Stage</th>
<th>Amsler-Krumbein classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Eccentric corneal steppening</td>
</tr>
<tr>
<td></td>
<td>Myopia and/or astigmatism &lt;5.00D</td>
</tr>
<tr>
<td></td>
<td>Mean central K readings &lt;48.00D</td>
</tr>
<tr>
<td>II</td>
<td>Myopia and/or astigmatism 5.00 to 8.00D</td>
</tr>
<tr>
<td></td>
<td>Mean central K readings &lt;53.00D</td>
</tr>
<tr>
<td></td>
<td>Absence of scarring</td>
</tr>
<tr>
<td></td>
<td>Minimal corneal thickness &gt;400 μm</td>
</tr>
<tr>
<td>III</td>
<td>Myopia and/or astigmatism 8.00 to 10.00D</td>
</tr>
<tr>
<td></td>
<td>Mean central K readings &gt;53.00D</td>
</tr>
<tr>
<td></td>
<td>Absence of scarring</td>
</tr>
<tr>
<td></td>
<td>Minimal corneal thickness 300 to 400 μm</td>
</tr>
<tr>
<td>IV</td>
<td>Refraction not measurable</td>
</tr>
<tr>
<td></td>
<td>Mean central K readings &gt;55.00D</td>
</tr>
<tr>
<td></td>
<td>Central corneal scarring</td>
</tr>
<tr>
<td></td>
<td>Minimal corneal thickness 200 μm</td>
</tr>
</tbody>
</table>

**Ethnically matched population specific control data.**

All affected and control individuals were of European ethnicity and population specific control data was obtained from three sources: (1) a total of 96 unrelated individuals (192 chromosomes) without ocular disease (aged 60 and over) from the Northern Irish population (U.K.) underwent Sanger sequencing; (2) exome data from 275 non-glaucomatous individuals from the Manchester population (U.K.), which are effectively ethnically identical to the Leeds population; and (3) normative control data for 413 normal individuals from Lausanne (Swiss population; European) was obtained from The CoLaus Study (http://www.colaus.ch/) (42).
DNA Sequencing and Statistical Analysis

PCR primers for amplification of the 16 exons and flanking intron sequences of PRDM5 were designed using Primer3 (v. 0.4.0) software ([http://frodo.wi.mit.edu/primer3/](http://frodo.wi.mit.edu/primer3/)) and are listed in Supplementary Table S3. PCR and Sanger sequencing of ZNF469 was undertaken with primers identical to those previously used by Christensen et al. (29) (personal communication) with adapted conditions (Supplementary Table S4). Sequencing results were analyzed manually using the sequence analysis software SeqScape 2.1.1 (Applied Biosystems, USA). Identified sequence variants were described according to the guidelines published by the Human Genome Variation Society. Variants were annotated in accordance with Ensembl transcript ENST00000437464 or NCBI NM_001127464.1 (Build GRCh37/hg19). The sequence variants were initially classified as variants of unknown significance (VUS) and then passed through a series of filtering steps shown in Figure 2.

If the sequence variants VUS were present in the ethnically matched population specific control data they were excluded. The remaining sequence variants VUS were required to have a minor allele frequency (MAF) of < 0.1% in the data from dbSNP (Build 137), the May 2012 release of the 1000 Genomes (1KG) Project and the Exome Variant Server (EVS), NHLBI Exome Sequencing Project (ESP). Following this the remaining non-synonymous alleles were filtered using SIFT which can identify if an amino acid substitution influences protein function resulting in a phenotypic change; classified as damaging or tolerated (44). SIFT can distinguish between functionally neutral and deleterious amino acid changes in mutagenesis studies and on human polymorphisms (45, 46). There are alternative prediction tools that use a combination of methods based on sequence homology, protein structure information and physicochemical properties of amino acids for prediction (44). Given that ZNF469 is a poorly characterised protein with no verified structural homologs, the SIFT
algorithm was applied as SIFT computes a combined score derived from the distribution of
amino acid residues observed at a given position in the sequence alignment and the estimated
unobserved frequencies of amino acid distribution calculated from a Dirichlet mixture and
does not rely on structural or physiochemical information (44). The conservation of the
affected amino acid across species was analysed using Homologene
(http://www.ncbi.nlm.nih.gov/homologene/) and multiple sequence alignment with ClustalX
(http://www.clustal.org/) visualised with GeneDoc software
(http://www.nrbsc.org/gfx/genedoc/).

The collapsing method (47) was used to compare the frequency of remaining potentially
pathogenic alleles between case and control subjects with the level of significance set to $p < 0.05$.
“This method involves collapsing genotypes across variants and applying a univariate
test which is powerful for analysing rare variants (47). Specifically, each individual was
assigned an indicator variable that takes the value 1 if the subject carries at least one
potentially pathogenic variant and zero otherwise. Whether the proportions of individuals
with index variable 1 differ significantly in cases and controls was tested using a Fisher exact
test on the corresponding contingency table of indicator variable counts.” The estimated odds
ratio (OR), relative risk (RR), 95% confidence interval (CI) and Fisher exact p-value were
calculated using JavaStat (http://statpages.org/ctab2x2.html).
Web Resources

1000 Genomes Project: http://browser.1000genomes.org/index.html

Exome Variant Server, NHLBI Exome Sequencing Project, Seattle, WA:
http://evs.gs.washington.edu/EVS/


Fischer exact test calculations: http://statpages.org/ctab2x2.html

Sorting Int intolerant from Tolerant (SIFT) http://sift.jcvi.org/

Online Mendelian Inheritance in Man (OMIM): http://www.omim.org/

Homologene (http://www.ncbi.nlm.nih.gov/homologene/)

ClustalX (http://www.clustal.org/)

GeneDoc software (http://www.nrbsc.org/gfx/genedoc/).
ACKNOWLEDGMENTS

We are grateful to the patients and families who participated in this study. We would like to thank Eyvind Rødahl from the University of Bergen, Norway for providing the sequence of the ZNF469 primers. Special thanks to Drs S. George S, D.G. Frazer, U. Donnelly and J.E. Moore (Belfast Health and Social Care Trust, UK) for assistance with DNA collection and recruitment. This work was supported by Northern Ireland Research and Development Office RRG [grant number 4.46 (C.E.W.)]; Fight for Sight [grant number 1787 (C.E.W., J.L.)]; and the NIHR Manchester Biomedical Research Centre (L.F.P., F.D.M, G.C.B). JL is a Fight for Sight (UK) PhD student. We are grateful for access to exome sequence data from the CoLaus cohort, which was sequenced as part of a partnership between the Wellcome Trust Sanger Institute, the CoLaus principal investigators and the Quantitative Sciences Department of GlaxoSmithKline.

CONFLICT OF INTEREST STATEMENT

None declared.
REFERENCES


LEGENDS TO FIGURES

Figure 1: Corneal topography of the 28 year old European patient from the UK with a c.3119A>C (p.Lys1040Thr) ZNF469 pathogenic allele (Table1) using Pentacam corneal topography; OD indicates right eye and OS left eye. The topography shows the anterior corneal steepening (upper images) associated with keratoconus with a large cone centrally in the right eye and paracentrally in the left eye with mean central K readings of 72.4 D (OD) and 57.3 D (OS) respectively. The lower images demonstrate the associated corneal thinning underlying the cones with a minimum corneal thickness of 318 μm (OD) and 438 μm (OS). The keratoconus is Stage III in both eyes (Amsler-Krumeich classification) with a best corrected Snellen acuity of 6/36 right and 6/24 left. The patient subsequently underwent a deep anterior lamellar keratoplasty (corneal transplant) in the right eye.

Figure 2: Hierarchical flow diagram of filtering process performed on sequence variants identified in ZNF469 in keratoconus cohort by Sanger sequencing.
TABLES

Table 1: Potentially Pathologic \textit{ZNF 469} Alleles Identified in Keratoconus and Control Subjects.
<table>
<thead>
<tr>
<th>Nucleotide Change</th>
<th>Amino Acid Change</th>
<th>Present in 1KG data (MAF (%))</th>
<th>Present in EVS data (MAF (%))</th>
<th>rs number</th>
<th>SIFT prediction (SIFT Score)</th>
<th>Amsler-Krumeich Classification</th>
<th>Corneal Transplantation</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.290C&gt;T</td>
<td>p.Pro97Leu</td>
<td>No</td>
<td>No</td>
<td>rs273585617</td>
<td>Damaging (0)</td>
<td>Stage III</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.337G&gt;A</td>
<td>p.Glu113Lys</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>Damaging (0)</td>
<td>Stage II</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.2063C&gt;A</td>
<td>p.Thr688Asn</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>Damaging (0)</td>
<td>Stage IV</td>
<td>Stage III</td>
</tr>
<tr>
<td>c.2699C&gt;G</td>
<td>p.Pro900Arg</td>
<td>No</td>
<td>No</td>
<td>rs273585618</td>
<td>Damaging (0.02)</td>
<td>Stage III</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.2699C&gt;T</td>
<td>p.Pro900Arg</td>
<td>No</td>
<td>No</td>
<td>rs273585618</td>
<td>Damaging (0)</td>
<td>Stage II</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.2904_2909delGTCGGG</td>
<td>p.Ser969_Gly970</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>Stage IV</td>
<td>Stage III</td>
</tr>
<tr>
<td>c.3119A&gt;C</td>
<td>p.Lys1040Thr</td>
<td>No</td>
<td>No</td>
<td>rs273585619</td>
<td>Damaging (0)</td>
<td>Stage III</td>
<td>Stage III</td>
</tr>
<tr>
<td>c.4363G&gt;T</td>
<td>p.Ala1455Ser</td>
<td>No</td>
<td>No</td>
<td>rs116532825</td>
<td>Damaging (0.02)</td>
<td>Stage I</td>
<td>Stage III</td>
</tr>
<tr>
<td>c.5464C&gt;A</td>
<td>p.Pro1822Thr</td>
<td>Yes (NA)</td>
<td>No</td>
<td>rs74032866</td>
<td>Damaging (0.04)</td>
<td>Stage I</td>
<td>Stage I</td>
</tr>
<tr>
<td>c.6095C&gt;A</td>
<td>p.Ser2032Tyr</td>
<td>No</td>
<td>No</td>
<td>rs273585623</td>
<td>Damaging (0.05)</td>
<td>Stage II</td>
<td>Stage III</td>
</tr>
<tr>
<td>c.8912G&gt;T</td>
<td>p.Gly2971Val</td>
<td>No</td>
<td>No</td>
<td>rs273585625</td>
<td>Damaging (0.04)</td>
<td>Stage III</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.9011_9025delTTCCGGGAACACCC</td>
<td>p.Leu3004_Thr3008</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>Stage III</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.9047C&gt;T</td>
<td>p.Thr3016Met</td>
<td>No</td>
<td>No</td>
<td>rs273585626</td>
<td>Damaging (0.02)</td>
<td>Stage II</td>
<td>Stage I</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>---</td>
<td>---</td>
<td>-------------</td>
<td>---------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>c.11615C&gt;T</td>
<td>p.Pro3872Leu</td>
<td>No</td>
<td>No</td>
<td>rs273585630</td>
<td>Damaging</td>
<td>(0.03)</td>
<td>Stage II</td>
</tr>
<tr>
<td>c.1701G&gt;T</td>
<td>p.Gln567His</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>Damaging</td>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>

**NORMAL CONTROLS**

---

a ZNF469 Ensembl transcript ENST00000437464 or NCBI NM_001127464.1 (Build GRCh37/hg19)

b Positions with normalized probabilities less than 0.05 are predicted to be damaging, those greater than or equal to 0.05 are predicted to be tolerated

c Stages described in Methods(18, 19)

NA: not available or applicable
Table 2: ZNF 469 Sequence Variants of Unknown Significance (VUS) Identified in Keratoconus and Control Subjects.

<table>
<thead>
<tr>
<th>Nucleotide Change</th>
<th>Amino Acid Change</th>
<th>rs number</th>
<th>Present in 1KG data (MAF (%))</th>
<th>Present in EVS data (MAF (%))</th>
<th>SIFT prediction</th>
<th>Classification(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KERATOCONUS COHORT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Non-Synonymous Variants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.77G&gt;C</td>
<td>p.Ser26Thr</td>
<td>rs273585616</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.1627G&gt;A</td>
<td>p.Gly543Ser</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism (Ser common)</td>
</tr>
<tr>
<td>c.2297G&gt;A</td>
<td>p.Arg766Gln</td>
<td>rs144492145</td>
<td>Yes (0.05)</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>c.3236G&gt;A</td>
<td>p.Arg1079Gln</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism (Gln in marmoset)</td>
</tr>
<tr>
<td>c.4394C&gt;T</td>
<td>p.Pro1465Leu</td>
<td>rs369382753</td>
<td>No</td>
<td>Yes (0.02)</td>
<td>Tolerated</td>
<td>Polymorphism (Leu in rat and mouse)</td>
</tr>
<tr>
<td>c.4826G&gt;C</td>
<td>p.Arg1609Pro</td>
<td>rs273585621</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism (Pro in dog)</td>
</tr>
<tr>
<td>c.5060G&gt;A</td>
<td>p.Arg1687Lys</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>Variant</td>
<td>Residue Before/After</td>
<td>Ensembl ID</td>
<td>TDB 1</td>
<td>TDB 2</td>
<td>TDB 3</td>
<td>TDB 4</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>c.5597A&gt;T</td>
<td>p.Gln1866Leu</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.6007G&gt;A</td>
<td>p.Glu2003Lys</td>
<td>rs273585622</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism (Lys in gorilla)</td>
</tr>
<tr>
<td>c.6725C&gt;A</td>
<td>p.Ser2242Tyr</td>
<td>rs273585624</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.7527G&gt;C</td>
<td>p.Glu2509Asp</td>
<td>rs199519673</td>
<td>No</td>
<td>Yes (0.090)</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.7747G&gt;A</td>
<td>p.Glu2583Lys</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.7847G&gt;A</td>
<td>p.Arg2616Gln</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.9835A&gt;G</td>
<td>p.Thr3279Ala</td>
<td>rs273585627</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
<tr>
<td>c.11101G&gt;A</td>
<td>p.Gly3701Ser</td>
<td>rs273585629</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
</tr>
</tbody>
</table>

**Synonymous Variants**

<table>
<thead>
<tr>
<th>Variant</th>
<th>Residue Before/After</th>
<th>Ensembl ID</th>
<th>TDB 1</th>
<th>TDB 2</th>
<th>TDB 3</th>
<th>TDB 4</th>
<th>TDB 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.99G&gt;A</td>
<td>p.Pro33Pro</td>
<td>rs273585631</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.720G&gt;A</td>
<td>p.Glu240Glu</td>
<td>rs273585632</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.2478G&gt;T</td>
<td>p.Pro826Pro</td>
<td>rs273585634</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.6453T&gt;C</td>
<td>p.Asp2151Asp</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.10843C&gt;T</td>
<td>p.Leu3615Leu</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
</tbody>
</table>

**NORMAL CONTROL COHORT**

**Non-Synonymous Variants**

<table>
<thead>
<tr>
<th>Variant</th>
<th>Residue Before/After</th>
<th>Ensembl ID</th>
<th>TDB 1</th>
<th>TDB 2</th>
<th>TDB 3</th>
<th>TDB 4</th>
<th>TDB 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.10115C&gt;T</td>
<td>p.Pro3372Leu</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.11252G&gt;A</td>
<td>p.Arg3751Lys</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Tolerated</td>
<td>Polymorphism (Lys in dog)</td>
<td></td>
</tr>
<tr>
<td>Synonymous Variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.30G&gt;A</td>
<td>p.Pro10Pro</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
<tr>
<td>c.4281C&gt;A</td>
<td>p.Leu1427Leu</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>VUS</td>
<td></td>
</tr>
</tbody>
</table>

*If predicted to be tolerated using SIFT the conservation of the residue was assessed and if poorly conserved the variant was classified as a polymorphism*
ABBREVIATIONS

IKG  1000 Genomes

BCS  Brittle cornea syndrome

CCT  Central corneal thickness

CI   Confidence interval

ESP  Exome sequencing project

EVS  Exome variant server

GWAS Genome-wide association study

MAF  Minor allele frequency

OR   Odds ratio

POAG Primary open angle glaucoma

RR   Relative risk

VUS  Variant of unknown significance
Figure 1: Corneal topography of the 28 year old European patient from the UK with a c.3119A>C (p.Lys1040Thr) ZNF469 pathogenic allele using Pentacam Corneal Topography. 1200x829mm (65 x 65 DPI)
Figure 2: Hierarchical flow diagram of filtering process performed on sequence variants identified in *ZNF469* in keratoconus cohort by Sanger sequencing.

355x266mm (96 x 96 DPI)