

RESEARCH ARTICLE

# TGC repeat expansion in the *TCF4* gene increases the risk of Fuchs' endothelial corneal dystrophy in Australian cases

Abraham Kuot<sup>1</sup>, Alex W. Hewitt<sup>2,3</sup>, Grant R. Snibson<sup>2</sup>, Emmanuelle Souzeau<sup>1</sup>, Richard Mills<sup>1</sup>, Jamie E. Craig<sup>1</sup>, Kathryn P. Burdon<sup>1,3</sup>, Shiwani Sharma<sup>1</sup>\*

**1** Department of Ophthalmology, College of Medicine and Public Health, Flinders University, Adelaide, South Australia, Australia, **2** Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia, **3** Menzies Institute for Medical Research, University of Tasmania, Tasmania, Australia

\* These authors contributed equally to this work.

\* [shiwani.sharma@flinders.edu.au](mailto:shiwani.sharma@flinders.edu.au)



**OPEN ACCESS**

**Citation:** Kuot A, Hewitt AW, Snibson GR, Souzeau E, Mills R, Craig JE, et al. (2017) TGC repeat expansion in the *TCF4* gene increases the risk of Fuchs' endothelial corneal dystrophy in Australian cases. PLoS ONE 12(8): e0183719. <https://doi.org/10.1371/journal.pone.0183719>

**Editor:** Ralf Krahe, University of Texas MD Anderson Cancer Center, UNITED STATES

**Received:** December 16, 2016

**Accepted:** August 9, 2017

**Published:** August 23, 2017

**Copyright:** © 2017 Kuot et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** This work was funded by grants from the National Health and Medical Research Council (NHMRC), Australia, grant #1023911, Ophthalmic Research Institute of Australia, and the Flinders Medical Centre Foundation, South Australia, Australia. JEC and KPB are recipients of NHMRC practitioner and research fellowships, respectively. The funders had no role in study design, data

## Abstract

Fuchs' endothelial corneal dystrophy (FECD) is a progressive, vision impairing disease. Common single nucleotide polymorphisms (SNPs) and a trinucleotide repeat polymorphism, thymine-guanine-cytosine (TGC), in the *TCF4* gene have been associated with the risk of FECD in some populations. We previously reported association of SNPs in *TCF4* with FECD risk in the Australian population. The aim of this study was to determine whether TGC repeat polymorphism in *TCF4* is associated with FECD in the Australian population. In 189 unrelated Australian cases with advanced late-onset FECD and 183 matched controls, the TGC repeat polymorphism located in intron 3 of *TCF4* was genotyped using a short tandem repeat (STR) assay. The repeat length was verified by direct sequencing in selected homozygous carriers. We found significant association between the expanded TGC repeat ( $\geq 40$  repeats) in *TCF4* and advanced FECD ( $P = 2.58 \times 10^{-22}$ ; OR = 15.66 (95% CI: 7.79–31.49)). Genotypic analysis showed that 51% of cases (97) compared to 5% of controls (9) were heterozygous or homozygous for the expanded repeat allele. Furthermore, the repeat expansion showed stronger association than the most significantly associated SNP, rs613872, in *TCF4*, with the disease in the Australian cohort. This and haplotype analysis of both the polymorphisms suggest that considering both the polymorphisms together rather than either of the two alone would better predict susceptibility to FECD in the Australian population. This is the first study to report association of the TGC trinucleotide repeat expansion in *TCF4* with advanced FECD in the Australian population.

## Introduction

Fuchs' endothelial corneal dystrophy (FECD, MIM 136800) is a progressive, degenerative disease of the corneal endothelium [1]. The clinical hallmarks of the disease include the presence of microscopic outgrowths (guttatae), aberrant thickening of the Descemet's membrane [2], and corneal endothelial cell loss [3]; Descemet's membrane is the collagen rich basal lamina secreted by the corneal endothelium [4]. These pathological changes are accompanied by

collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

impaired ability of the corneal endothelium to pump excess fluid from the corneal stroma [1] which results in corneal oedema, pain, and ultimately vision loss, if left untreated [2]. Corneal transplantation, in the form of either penetrating or partial keratoplasty, is currently the only effective treatment for FECD [5].

The prevalence of FECD varies markedly across the world. In the USA, it affects ~4% of the population over the age of 40 [6] but is less frequent in Asian [7] and Middle-Eastern populations [8]. The prevalence has not been reported in Australia but corneal grafting for FECD accounted for ~26% (n = 395) of the corneal grafts (n = 1533) performed in 2014 [9], indicating that the disease is relatively common.

FECD is genetically a heterogeneous disease and manifests as two forms depending upon the age of onset. The rare, early-onset form [10] is typically inherited as an autosomal dominant disease with high penetrance and nearly uniform expressivity [2]. Mutations in the *COL8A2* (*collagen, type VIII, alpha 2*) gene account for some cases of early-onset FECD [11]; *COL8A2*, an extracellular matrix protein, is a major component of the Descemet's membrane [12]. The more common late-onset FECD typically occurs after the age of 40 [13] and can be a familial disease; the risk of the disease increases with age and female sex [6]. Familial late-onset FECD shows an autosomal dominant inheritance with high penetrance but variable expressivity [6]. Mutations in the *ZEB1* (*zinc-finger E-box binding homeobox 1*), *SLC4A11* (*solute carrier family 4, sodium borate transporter, member 11*), *LOXHD1* (*lipoxxygenase homology domains 1*) and *AGBL1* (*ATP/GTP binding protein-like 1*) genes cause the late-onset disease in a small number of familial, and/or unrelated cases [6, 8, 14–17]. Additionally, four chromosomal loci have been linked with familial late-onset disease but the causative genes are yet to be identified [17].

In 2010, Baratz and colleagues [18], through a genome-wide association study (GWAS) in white American cases, reported a highly significant association between single nucleotide polymorphisms (SNPs) in the *TCF4* (*Transcription factor 4*) gene and late-onset FECD; the intronic SNP rs613872 was the most significantly associated variant. Recently, our collaborative group through GWAS identified genome-wide significant association of SNPs in the *KANK4* (*KN motif- and ankyrin repeat domain-containing protein 4*) and *LAMC1* (*Laminin gamma-1*), and near *ATP1B1* (*Na<sup>+</sup>, K<sup>+</sup> transporting ATPase, beta-1 polypeptide*) genes with FECD [19] and revealed three additional loci involved in the disease. However, the *TCF4* locus remains the strongest associated locus with FECD [19].

The *TCF4* gene is located on chromosome 18 and codes for the helix-loop-helix transcription factor E2-2 [18]. Our group, through an independent replication study, reported the association of SNPs in *TCF4* with advanced FECD in white Australians, which demonstrated the contribution of this gene to the disease risk in the Australian population [20]. Similar independent studies by other groups have shown association of SNPs, mainly rs613872, in *TCF4* with advanced FECD including in the Indian and Chinese populations [21–24]. Subsequently, Wieben and colleagues [25] identified a significant association of thymine-guanine-cytosine (TGC) repeat expansion in intron 3 of *TCF4* with FECD in white Americans and reported the repeat expansion to be a stronger predictor of the disease than SNP rs613872; repeat lengths of >50 were found to be more frequent in cases compared to controls [25]. The *TCF4* TGC trinucleotide repeat was first identified by Breschel et al. in 1997 and named as the CTG18.1 locus [26]. Association of this repeat expansion in *TCF4* with the disease has been replicated since in independent white American, Indian, Chinese and Japanese case cohorts [24, 27–30]; TGC repeat length of  $\geq 40$  was associated in these populations. This cut-off is based on the initial discovery that indicated that expansion of the TGC repeat locus in 3% of subjects in white pedigrees was not associated with any known phenotypes [26] whereas the expanded alleles with repeat lengths >37 were reported to be unstable [26]. Additionally, Mootha et al. [31] found segregation of the expanded TGC repeats in *TCF4* with the disease in several affected families

with high or low penetrance indicating the importance of this repeat expansion in late-onset FECD. In this study, we aimed to determine whether TGC repeat polymorphism in *TCF4* is associated with late-onset FECD in the Australian population.

## Materials and methods

### Ethics statement, participant recruitment, and sample collection

The study was approved by the Southern Adelaide Clinical Human Research Ethics Committee, Southern Adelaide Local Health Network and Flinders University (South Australia, Australia), and the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital (Melbourne, Victoria, Australia). The research was conducted in accordance with the guidelines of the National Health and Medical Research Council, Australia, and adhering to the tenets of the revised Declaration of Helsinki. All participants underwent a complete ophthalmic examination including slit lamp examination, confocal specular microscopy, and funduscopy. Patients diagnosed with Grade 3–6 advanced late-onset FECD according to a modified Krachmer grading system [13] were recruited after obtaining written informed consent. Blood samples were collected from 189 participants through the Flinders Eye Clinic (Adelaide, South Australia, Australia) and the Royal Victorian Eye and Ear Hospital (Melbourne, Victoria, Australia). Control genomic DNA samples were from 183 unrelated, unaffected South Australian residents aged over 50 years recruited previously for use as controls in a variety of ocular genetic studies [32–34]. Genomic DNA from cases and controls was extracted using QIAamp DNA Blood Maxi kit (Qiagen Pty Ltd, Doncaster, Victoria, Australia) following the manufacturer's protocol.

### STR assay and DNA sequencing

Gene-specific PCR primers (Forward: 5' -CAGATGAGTTTGGTGTAAAGATG-3' , Reverse: 5' -ACAAGCAGAAAGGGGGCTGCAA-3' ) used for amplification of the TGC trinucleotide repeat region in the third intron in the *TCF4* gene were the same as reported by Wieben et al [25]. A 5'FAM label was added to the forward primer. The TGC repeat polymorphism in genomic DNA was genotyped using short tandem repeat (STR) assay as previously described [25, 31]. Forty nanograms of genomic DNA was used as template and PCR was performed in a 10µl volume using HotStar Plus Taq® DNA polymerase (Qiagen GmbH, Hilden, Germany). The enzyme was activated at 95°C for 6 minutes followed by 35 cycles of denaturation at 95°C for 1 minute, annealing at 64°C for 1 minute, and elongation at 68°C for 3 minutes. The final elongation was at 68°C for 7 minutes. Each PCR product was diluted 1:10 and 1µl of the diluted product mixed with 0.15µl of GeneScan™1200LIZ® Dye Internal Size Standard (Applied Biosystems, Foster City, CA) and 8.85µl of Hidi Formamide (Applied Biosystems), and electrophoresed on a 3130xL Genetic Analyser (Applied Biosystems) according to standard protocols. The TGC repeat alleles were manually called using Peak Scanner™ Software v1.0 (Applied Biosystems).

In selected cases and controls homozygous for the shortest or the longest repeat allele, Sanger sequencing was performed for validation of STR assay results. Amplification was performed as described above except unlabelled forward primer was used in a reaction volume of 20µl. The amplified products (5µl) were treated with 10U Exonuclease (New England Biolabs, Ipswich, MA, USA) and 2U Shrimp Alkaline Phosphatase (SAP; USB Corporation, Cleveland, Ohio, USA) at 37°C for 1 hour to digest and dephosphorylate remaining primer DNA and dNTPs; the reaction was terminated by incubation at 80°C for 20 minutes. Cleaned PCR products were sequenced using BigDye® Terminator v3.1 (Applied Biosystems) and forward primer, on a 3130xL Genetic Analyser (Applied Biosystems) according to standard protocols.

DNA Sequences were aligned to the *TCF4* reference sequence (NM\_001083962.1) retrieved from the Reference Human Genome sequence version 19 (hg19\_RefGen), using the Sequencher<sup>®</sup> software 5.0 (GeneCodes Corporation, Ann Arbor, MI). The TGC repeats were counted manually.

## Statistical analysis

Statistical analyses were performed using SPSS (Statistical Package for the Social Science; version 22) or in PLINK software [35]. Baseline characteristics of cases and controls—age and sex—were compared using Student's t-test and chi-square test, respectively. The trinucleotide repeat alleles in cases and controls were compared by chi-square test. Difference in the distribution of the TGC repeat lengths between cases and controls was assessed by Mann-Whitney U test. A trend test for association of genotypes with FECD was performed in PLINK [35]. Logistic regression analysis was also performed in PLINK with age and sex as covariates to explore their effects. Conditional analysis was performed on SNP rs613872 (previously genotyped and association reported [20]) and the expanded TGC allele in *TCF4* to determine if the SNP was independently associated with the disease. Haplotype analysis was also performed using PLINK to investigate the effect of haplotypes of the associated *TCF4* expanded TGC repeat locus and the SNP rs613872, on FECD.

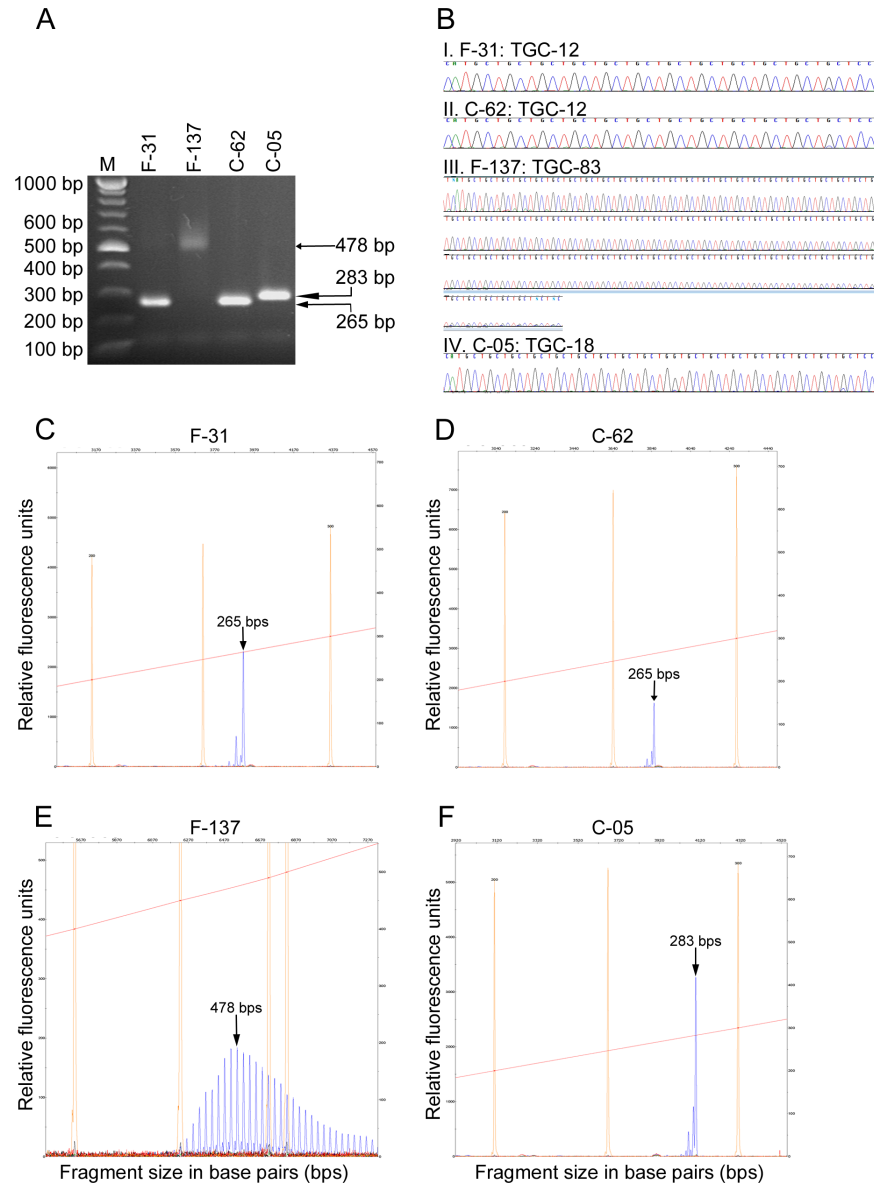
## Results

In this study, we determined an association between the trinucleotide repeat polymorphism in *TCF4* gene and FECD in white Australian cases by screening a total of 189 unrelated cases with advanced late-onset disease and 183 controls for the TGC repeat polymorphism in the gene. The TGC repeat was individually genotyped in each case and control by the previously reported STR assay [25]. The TGC repeats in selected homozygous cases and controls carrying the shortest repeat length (4 cases and 4 controls with 12 repeats each), and the longest repeat lengths (3 cases with 76, 83 and 84 repeats, respectively; 3 controls with 18 repeats each) were directly sequenced to confirm the repeat lengths. Size fractionation of the amplified repeat region showed products of the expected sizes in both cases and controls (Fig 1A) and confirmed the repeat lengths detected by the STR assay (Fig 1B). Sequence chromatograms and STR electropherogram traces of representative cases and controls are shown in Fig 1B–1F.

The characteristics of the case and control cohorts are given in Table 1. Females comprised 69% of cases and 72% of controls and the difference was not statistically significant ( $p = 0.555$ ). A relatively higher percentage of females than males in the case cohort is consistent with the reported higher prevalence of FECD in females [36]. The controls were significantly older than cases ( $p = 0.023$ ) by design to reduce the likelihood of yet to manifest disease.

In accordance with denotation of the *TCF4* TGC repeat in previous studies [27, 31, 37], we dichotomised the TGC repeat alleles such that repeat length of  $\geq 40$  was considered as an expanded allele and  $< 40$  as a non-expanded allele. As shown in Table 1, the expanded allele is relatively rare in controls, but occurs on 28% of chromosomes in cases ( $p = 2.58 \times 10^{-22}$ ; OR = 15.66 (95% CI: 7.79–31.49)). We found that cases carrying alleles with  $\geq 40$  TGC repeat lengths in *TCF4* are at 15 times greater risk of developing the disease than non-carriers. This is consistent with the findings of previous studies in white cases with FECD [25, 31], which also showed that those carrying expanded alleles of TGC repeats in this gene are at a greater risk of the disease than those carrying non-expanded alleles.

Next, we assessed the distribution of the TGC repeat lengths in the study cohort. In cases, median repeat length was 53 and ranged from 11 to 115 whereas in controls it was 18 and ranged from 11 to 83 (Fig 2). The difference in the distribution of the repeat length in cases compared to controls was statistically significant ( $p = 0.0005$ ).



**Fig 1. Analysis of *TCF4* TGC repeat polymorphism in homozygous FECD cases and controls.** PCR was performed on genomic DNA from cases and controls carrying the shortest or the longest *TCF4* TGC repeat length. Data from two representative cases (F-31 and F-137 with 12 and 83 repeats, respectively) and controls (C-62 and C-05, with 12 and 18 repeats, respectively) are shown. **A.** Agarose gel electrophoresis of the PCR amplified repeat region from homozygous FECD cases and controls. Sizes of the products are shown on the right and correspond with the expected sizes (F-31 and C-62, 265 bp each; F-137, 478 bp; C-05, 283 bp). Sizes of DNA markers are indicated on the left. **B.** Sequencing chromatograms of FECD-affected (F-31 and F-137) and control (C-62 and C-05) individuals homozygous for the shortest (**B.I and B.II**) and the longest (**B.III and B.IV**) repeat alleles are shown. TGC repeat length in each individual was calculated by subtracting 230 bps, corresponding to the DNA region flanking the repeat region amplified during PCR, from the detected PCR product size and dividing the difference by 3 (number of nucleotides of the repeat). **C-F.** Electropherograms showing the sizes of TGC repeat alleles in FECD cases F-31 (**C**) and F-137 (**E**) and in two control individuals C-62 (**D**) and C-05 (**F**) detected by STR assay. The peaks representing the *TCF4* TGC repeat fragments are indicated by arrows. Multiple peaks seen in panel E are due to variation in product size when large repeats are amplified. X-axis, fragment sizes in base pairs; Y-axis, relative fluorescence units; orange peaks, sizes of internal standards; red line across the electropherograms, slope threshold for peak start/end.

<https://doi.org/10.1371/journal.pone.0183719.g001>



**Table 1. Characteristics of the FECD case and control cohorts, and dichotomised distribution of the *TCF4* TGC repeat alleles in cases and controls.** The age and sex between cases and controls were compared using the Student's t-test and chi-square test, respectively. Expanded allele counts between cases and controls were compared using chi-square test.

Description	Case	Control	p-value
N Participants	189	183	-
N Females (%)	130 (69%)	131 (72%)	0.555
Mean age in years ± SD	69.9 ± 11.2	76.6 ± 8.7	0.023
(Age range in years)	(32–93)	(42–96)	
N TGC repeat alleles with repeat length of <40/≥40	271/107	357/9	2.58 × 10 <sup>-22</sup>

N = number; SD = standard deviation.

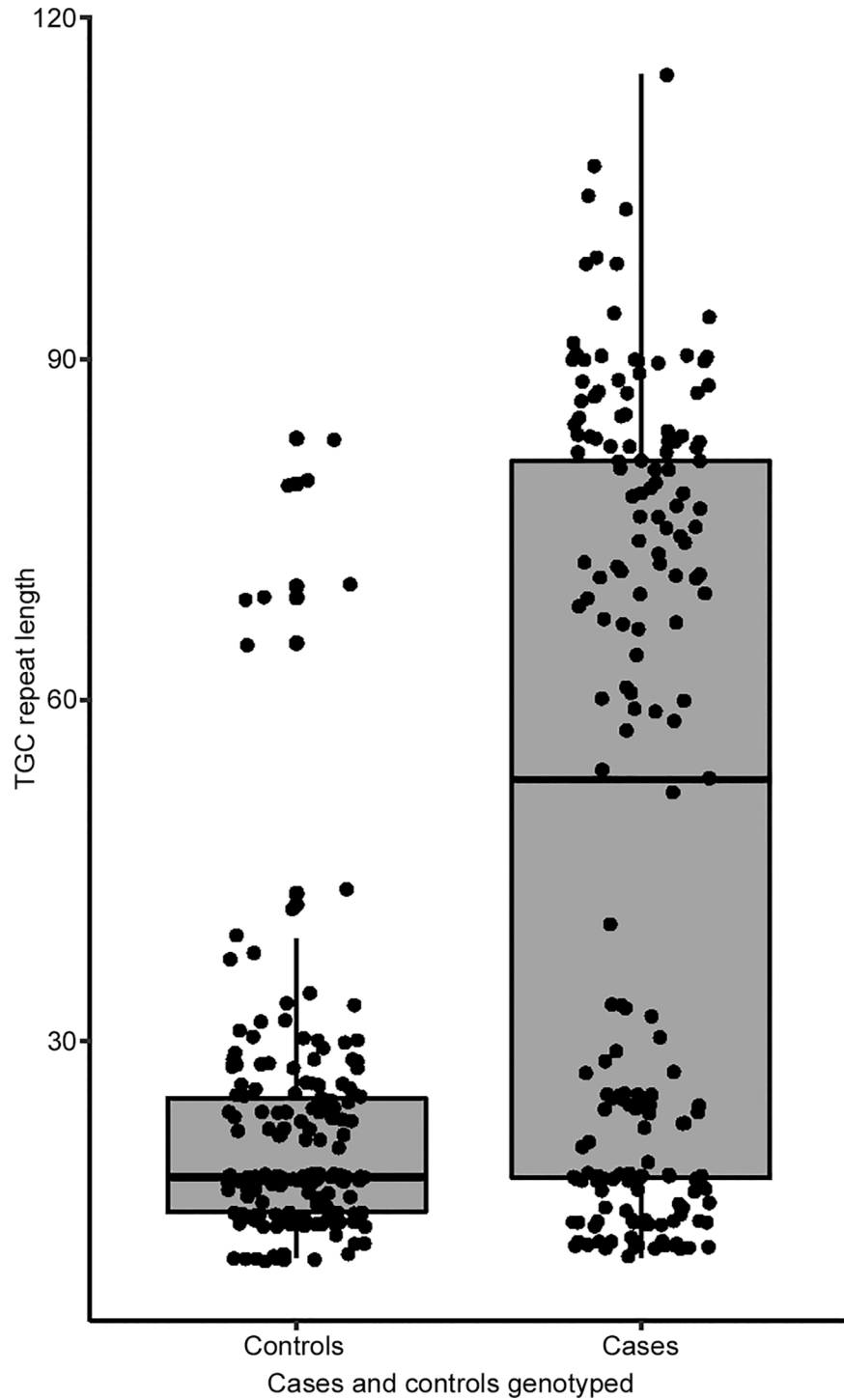
<https://doi.org/10.1371/journal.pone.0183719.t001>

Considering the genotype of the TGC repeat at the individual level, 87 cases (46%) were heterozygous and 10 cases (5%) homozygous for the expanded repeat allele; the remaining cases (n = 92; 49%) were homozygous for the non-expanded repeat allele (Fig 3). Of the cases homozygous for the expanded repeat, 7 cases carried both alleles with the same repeat lengths. Of those homozygous for the non-expanded repeat, 34 cases (18%) carried two alleles with different repeat lengths, and the remaining 58 cases (30.7%) carried both alleles with the same repeat lengths. In contrast, the majority of the controls were homozygous for the non-expanded repeat allele (n = 174; 95%) and a small proportion were heterozygous for the expanded repeat allele (n = 9; 5%). None of the controls carried the expanded repeat allele in a homozygous state (Fig 3). Of the homozygous non-expanded repeat carrying controls, the majority, 154 (84%), carried different repeat lengths on the two alleles; only 20 controls (11%) carried both alleles with the same repeat lengths. Overall, we observed a significant genotypic association between FECD and the expanded repeat allele ( $p = 3.92 \times 10^{-22}$ ) in the Australian population.

Next, we evaluated the effect of age and sex on the observed association through logistic regression analysis which revealed that the association of the expanded TGC repeat with the disease is independent of age and sex ( $p = 2.09 \times 10^{-14}$ ; OR = 18.26 (95% CI: 8.67–38.46)). To determine any inter-dependence of association of the repeat expansion and the most associated SNP rs613872 in *TCF4* with FECD, we performed conditional analysis on the SNP. As evident from the reduced p-value and odds ratio, this analysis showed that association of the TGC repeat with the disease is partially independent of the SNP rs613872 ( $p = 9.74 \times 10^{-10}$ ; OR = 10.76 (95% CI: 5.02–23.05)), and that using both polymorphisms together can better predict FECD susceptibility in Australian cases than either of the two alone. This was confirmed by the haplotype analysis of the associated *TCF4* TGC repeat locus and SNP rs613872, which revealed significant association with FECD with overall p-value of  $7.74 \times 10^{-32}$  (Table 2). A similar observation has been reported in non-Australian populations in an independent study [31].

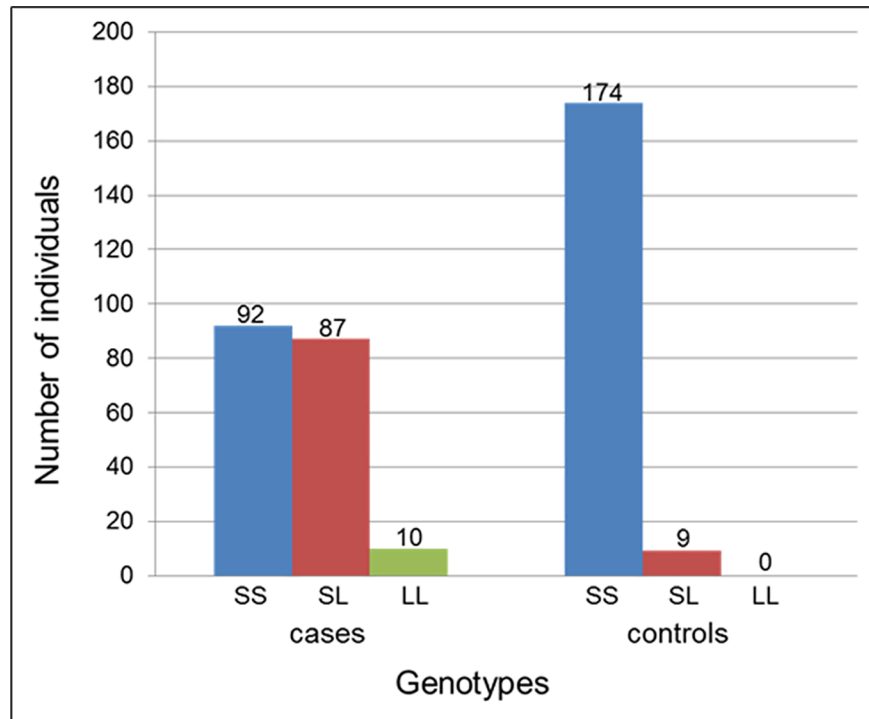
## Discussion

In this study, we determined the association between FECD and the expanded TGC repeat polymorphism in *TCF4* in Australian cases. The study reveals that the expanded TGC repeat is significantly associated with the disease in the white Australian population and is more significantly associated ( $P = 2.58 \times 10^{-22}$ ; OR = 15.66 (95% CI: 7.79–31.49)) than the rs613872 SNP ( $P = 5.25 \times 10^{-15}$ ; OR = 4.05 (95% CI: 2.82–5.83)) observed in our previous study [20]. Conditional analysis of the rs613872 SNP with the expanded TGC repeat suggests that these polymorphisms in *TCF4* are partially independently associated with FECD, implying that each can independently contribute to the pathogenesis of the disease. This study replicates the findings



**Fig 2. Distribution of TGC repeat lengths in the *TCF4* gene in FECD cases and controls.** Median repeat length in cases = 53; range: 11–115, and median repeat length in controls = 18; range: 11–83. The box represents the second and third quartiles and the line in the middle indicates median. The lower and upper whiskers represent the limits of the first and fourth quartiles, respectively. The dots represent individual data points for controls (n = 183) and cases (n = 189).

<https://doi.org/10.1371/journal.pone.0183719.g002>



**Fig 3. Distribution of genotypes of the TCF4 TGC repeat alleles in FECD cases and controls.** The numbers of individuals with each of the three possible genotypes of the dichotomised repeat alleles are shown. S represents short (<40 repeats; non-expanded) and L long (≥40 repeats; expanded) allele. SS represents homozygous non-expanded, LL homozygous expanded, and SL heterozygous with one non-expanded and one expanded allele.

<https://doi.org/10.1371/journal.pone.0183719.g003>

by other groups in American and Chinese cases with FECD [25, 27, 31]. It also suggests that considering both, the rs613872 SNP and the expanded TGC repeat polymorphism together can better predict susceptibility to FECD than either of the two alone, as shown by the haplotype analysis (*p*-value of  $7.74 \times 10^{-32}$ , Table 2).

FECD is the first aging-related ocular disease to be associated with trinucleotide repeat expansion. Pathogenic expansion of trinucleotide repeat sequences has been reported in several neurodegenerative and neuromuscular diseases such as Friedreich’s ataxia and myotonic dystrophy type 1 [38–40]. Pathogenic expansion of both alleles of the GAA repeat present in the first intron of the *Frataxin* gene has been reported in the majority of cases with Friedreich’s ataxia; it contributes to the disease by prolonging transcription, resulting in a significant decrease in Frataxin protein levels [38, 41]. The disease can also occur in a minority of patients as a result of expansion of one GAA repeat allele and the presence of a point mutation in the

**Table 2. Association between common haplotypes of genotyped polymorphisms (TGC repeat locus, L/S, and rs613872, G/T) in the TCF4 gene and FECD.** f = frequency; OR = Odds ratio; CI = Confidence Interval; L = long repeat allele; S = short repeat allele.

Haplotype	Over-all p-value = $7.74 \times 10^{-32}$			
	f cases	f controls	P-value	OR (95% CI)
LG	0.26	0.02	$9.59 \times 10^{-24}$	19.98 (9.08–44)
SG	0.21	0.14	0.009	1.68 (1.13–2.5)
LT	0.03	0.003	0.0002	113.3 (0.32–3.99× 10 <sup>4</sup> )
ST	0.50	0.84	$3.34 \times 10^{-26}$	0.12 (0.08–0.19)

<https://doi.org/10.1371/journal.pone.0183719.t002>



second allele [41]. Regardless of the genetic abnormality, transcription inhibition causes loss-of-function of *Frataxin*, and consequently pathogenesis of Friedreich ataxia [42]. The FECD-associated expanded TGC repeat in *TCF4* is also an intronic repeat. However, our differential gene expression analysis data (unpublished) and reported studies [28, 43] have shown that expression levels of *TCF4* mRNA in the corneal endothelium between FECD cases and controls and between carriers of risk and non-risk alleles of SNP rs613872 are similar indicating that transcription inhibition likely is not involved in the pathogenesis of FECD.

RNA-mediated gain-of-function mechanism has been shown to underlie myotonic dystrophy type 1 [38] that is caused by expansion of the CTG repeat in the 3'UTR of the *DMPK* (*Dys-trophia myotonica protein kinase*) gene [38]. The expanded CTG repeat has been reported to affect alternative splicing of the *DMPK* gene [38, 44]. The expanded *DMPK* (CUG)<sub>n</sub> mRNA is transcribed normally but fails to undergo translation and is thus retained in the nucleus as hairpin structures [45–47]. The hairpin structures exhibit a toxic dominant gain-of-function abnormality by sequestration and accumulation of RNA-binding regulatory proteins, such as the alternative splicing regulator muscle blind-like 1 (MBNL1) and CUG triplet repeat RNA-binding protein 1 (CUG-BP1) [48], and formation of nuclear foci [49–53]. The aberrant nuclear foci cause cellular toxicity that contributes to the disease pathogenesis [44]. Interestingly, Mootha et al. [28] and Du et al. [54] reported the presence of RNA nuclear foci in a proportion of corneal endothelial cells in some patients with FECD carrying the *TCF4* expanded TGC repeat. The expanded mRNA co-localised with sequestered MBNL1 protein in nuclear foci leading to mis-splicing of MBNL1-regulated transcripts. These reports implicated RNA toxicity and mis-splicing in the pathogenesis of FECD, and possibly a shared disease mechanism between FECD and myotonic dystrophy type 1 disease [28, 54]. The recent study by Wieben et al. [55] found mis-splicing of several genes in the corneal endothelium of FECD cases carrying expanded TGC repeat in *TCF4*, and validates these findings [28, 54]. Thus, RNA mis-splicing induced by TGC trinucleotide repeat expansion in the *TCF4* gene is the likely mechanism of pathogenesis in patients with FECD carrying the expanded repeats [55]. However other yet to be identified mechanisms may underlie the disease in patients carrying non-expanded repeats in this gene.

The STR analysis employed for detection of TGC repeat lengths in this study, is unable to detect very large repeat expansions of several hundred repeats, which is a limitation of the study. Previous studies have reported the prevalence of very large TGC repeat lengths in *TCF4* in ~5% of FECD cases [37]. Assuming similar prevalence of very large repeat expansions in the Australian cohort, nine cases, and one of the 20 controls found to carry the same non-expanded repeat length on both the alleles may carry a very large repeat expansion that was not detected in this study. This missing information would further strengthen the overall findings of the study.

In conclusion, this study provides independent evidence for the association of the expanded TGC repeat polymorphism in the *TCF4* gene with advanced FECD. This is the first study to report an association of the repeat polymorphism in *TCF4* with FECD in the Australian population and further confirms the contribution of this gene to FECD susceptibility in Australia. The findings suggest a partially independent association of the expanded TGC repeat and SNP rs613872 in *TCF4* with FECD in Australian cases, and indicate that considering both the polymorphisms together would better predict susceptibility to FECD in the Australian population.

## Acknowledgments

We sincerely thank all the participants whose DNA samples were used in this study.

## Author Contributions

**Conceptualization:** Shiwani Sharma.

**Data curation:** Abraham Kuot.

**Formal analysis:** Abraham Kuot, Kathryn P. Burdon.

**Funding acquisition:** Grant R. Snibson, Jamie E. Craig, Kathryn P. Burdon, Shiwani Sharma.

**Investigation:** Abraham Kuot.

**Project administration:** Emmanuelle Souzeau, Shiwani Sharma.

**Resources:** Alex W. Hewitt, Grant R. Snibson, Richard Mills, Jamie E. Craig.

**Supervision:** Jamie E. Craig, Kathryn P. Burdon, Shiwani Sharma.

**Validation:** Abraham Kuot, Kathryn P. Burdon, Shiwani Sharma.

**Writing – original draft:** Abraham Kuot.

**Writing – review & editing:** Alex W. Hewitt, Grant R. Snibson, Emmanuelle Souzeau, Richard Mills, Kathryn P. Burdon, Shiwani Sharma.

## References

1. Kannabiran C. Genetics of corneal endothelial dystrophies. *Journal of Genetics*. 2009; 88(4):487–94. PMID: [20090209](#)
2. Klintworth G. Corneal dystrophies. *Orphanet J Rare Dis*. 2009; 4(1):7.
3. Kenney MC, Labermeier U, Hinds D, Waring GO 3rd. Characterization of the Descemet's membrane/posterior collagenous layer isolated from Fuchs' endothelial dystrophy corneas. *Exp Eye Res*. 1984; 39(3):267–77. PMID: [6389163](#)
4. Adamis AP, Filatov V, Tripathi BJ, Tripathi RAmC. Fuchs' endothelial dystrophy of the cornea. *Surv Ophthalmol*. 1993; 38(2):149–68. PMID: [8235998](#)
5. Price MO, Gorovoy M, Benetz BA, Price FW Jr, Menegay HJ, Debanne SM, et al. Descemet's Stripping Automated Endothelial Keratoplasty Outcomes Compared with Penetrating Keratoplasty from the Cornea Donor Study. *Ophthalmology*. 2010; 117(3):438–44. <https://doi.org/10.1016/j.ophtha.2009.07.036> PMID: [20031230](#)
6. Riazuddin SA, Zaghoul NA, Al-Saif A, Davey L, Diplas BH, Meadows DN, et al. Missense Mutations in TCF8 Cause Late-Onset Fuchs Corneal Dystrophy and Interact with FCD4 on Chromosome 9p. *Am J Human Genet*. 2010; 86(1):45–53.
7. Santo RM, Yamaguchi T, Kanai A, Okisaka S, Nakajima A. Clinical and histopathologic features of corneal dystrophies in Japan. *Ophthalmology*. 1995; 102(4):557–67. PMID: [7724173](#)
8. Vithana EN, Morgan PE, Ramprasad V, Tan DTH, Yong VHK, Venkataraman D, et al. SLC4A11 mutations in Fuchs endothelial corneal dystrophy. *Hum Mol Genet*. 2008; 17(5):656–66. <https://doi.org/10.1093/hmg/ddm337> PMID: [18024964](#)
9. Williams KA, Keane MC, Galettis RA, Jones VJ, Mills RAD, Coster DJ. The Australian Corneal Graft Registry 2015 Report. Adelaide: Flinders University, Ophthalmology; 2015 2015. Report No.
10. Biswas S, Munier FL, Yardley J, Hart-Holden N, Perveen R, Cousin P, et al. Missense mutations in COL8A2, the gene encoding the  $\alpha 2$  chain of type VIII collagen, cause two forms of corneal endothelial dystrophy. *Hum Mol Genet*. 2001; 10(21):2415–23. PMID: [11689488](#)
11. Mok JW, Kim HS, Joo CK. Q455V mutation in COL8A2 is associated with Fuchs' corneal dystrophy in Korean patients. *Eye*. 2008; 23(4):895–903. <https://doi.org/10.1038/eye.2008.116> PMID: [18464802](#)
12. Kabosova A, Azar DT, Bannikov GA, Campbell KP, Durbeej M, Ghohestani RF, et al. Compositional Differences between Infant and Adult Human Corneal Basement Membranes. *Investig Ophthalmol Vis Sci*. 2007; 48(11):4989–99.
13. Krachmer JH, Purcell JJ Jr., Young CW, Bucher KD. Corneal endothelial dystrophy. A study of 64 families. *Arch Ophthalmol*. 1978; 96(11):2036–9. PMID: [309758](#)
14. Mehta JS, Vithana EN, Tan DTH, Yong VHK, Yam GHF, Law RWK, et al. Analysis of the Posterior Polymorphous Corneal Dystrophy 3 Gene, TCF8, in Late-Onset Fuchs Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci*. 2008; 49(1):184–8. <https://doi.org/10.1167/iovs.07-0847> PMID: [18172091](#)

15. Riazuddin SA, Vithana EN, Seet L-F, Liu Y, Al-Saif A, Koh LW, et al. Missense mutations in the sodium borate cotransporter SLC4A11 cause late-onset Fuchs corneal dystrophy. *Hum Mutat.* 2010; 31(11):1261–8. <https://doi.org/10.1002/humu.21356> PMID: 20848555
16. Riazuddin SA, Parker DS, McGlumphy EJ, Oh EC, Iliff BW, Schmedt T, et al. Mutations in LOXHD1, a recessive-deafness locus, cause dominant late-onset Fuchs corneal dystrophy. *Am J Hum Genet.* 2012; 90(3):533–9. <https://doi.org/10.1016/j.ajhg.2012.01.013> PMID: 22341973
17. Riazuddin SA, Vasanth S, Katsanis N, Gottsch JD. Mutations in AGBL1 cause dominant late-onset Fuchs corneal dystrophy and alter protein-protein interaction with TCF4. *Am J Hum Genet.* 2013; 93(4):758–64. <https://doi.org/10.1016/j.ajhg.2013.08.010> PMID: 24094747
18. Baratz KH, Tosakulwong N, Ryu E, Brown WL, Branham K, Chen W, et al. E2-2 protein and Fuchs's corneal dystrophy. *N Engl J Med.* 2010; 363(11):1016–24. <https://doi.org/10.1056/NEJMoa1007064> PMID: 20825314
19. Afshari NA, Igo RP, Morris NJ, Stambolian D, Sharma S, Pulagam VL, et al. Genome-wide association study identifies three novel loci in Fuchs endothelial corneal dystrophy. *Nature Communications.* 2017; 8:14898. <https://doi.org/10.1038/ncomms14898> PMID: 28358029
20. Kuot A, Hewitt AW, Griggs K, Klebe S, Mills R, Jhanji V, et al. Association of TCF4 and CLU polymorphisms with Fuchs' endothelial dystrophy and implication of CLU and TGFBI proteins in the disease process. *Eur J Hum Genet.* 2012; 20(6):632–8. <https://doi.org/10.1038/ejhg.2011.248> PMID: 22234156
21. Riazuddin SA, McGlumphy EJ, Yeo WS, Wang J, Katsanis N, Gottsch JD. Replication of the TCF4 intronic variant in late-onset Fuchs corneal dystrophy and evidence of independence from the FCD2 locus. *Invest Ophthalmol Vis Sci.* 2011; 52(5):2825–9. <https://doi.org/10.1167/iovs.10-6497> PMID: 21245398
22. Li Y-J, Minear MA, Rimmler J, Zhao B, Balajonda E, Hauser MA, et al. Replication of TCF4 through Association and Linkage Studies in Late-Onset Fuchs Endothelial Corneal Dystrophy. *PLoS One.* 2011; 6(4):e18044. <https://doi.org/10.1371/journal.pone.0018044> PMID: 21533127
23. Thalamuthu A, Khor CC, Venkataraman D, Koh LW, Tan DTH, Aung T, et al. Association of TCF4 Gene Polymorphisms with Fuchs Corneal Dystrophy in the Chinese. *Invest Ophthalmol Vis Sci.* 2011; 52(8):5573–8. <https://doi.org/10.1167/iovs.11-7568> PMID: 21659310
24. Nanda GG, Padhy B, Samal S, Das S, Alone DP. Genetic Association of TCF4 Intronic Polymorphisms, CTG18.1 and rs17089887, With Fuchs' Endothelial Corneal Dystrophy in an Indian Population. *Invest Ophthalmol Vis Sci.* 2014; 55(11):7674–80. <https://doi.org/10.1167/iovs.14-15297> PMID: 25342617
25. Wieben ED, Aleff RA, Tosakulwong N, Butz ML, Highsmith WE, Edwards AO, et al. A Common Trinucleotide Repeat Expansion within the Transcription Factor 4 (TCF4, E2-2) Gene Predicts Fuchs Corneal Dystrophy. *PLoS One.* 2012; 7(11):e49083. <https://doi.org/10.1371/journal.pone.0049083> PMID: 23185296
26. Breschel TS, McInnis MG, Margolis RL, Sirugo G, Corneliussen B, Simpson SG, et al. A novel, heritable, expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1. *Hum Mol Genet.* 1997; 6(11):1855–63. PMID: 9302263
27. Xing C, Gong X, Hussain I, Khor CC, Tan DT, Aung T, et al. Trans-ethnic Replication of Association of CTG18.1 Repeat Expansion of TCF4 Gene with Fuchs Corneal Dystrophy in Chinese Implies Common Causal Variant. *Invest Ophthalmol Vis Sci.* 2014; 55(11):7073–8. <https://doi.org/10.1167/iovs.14-15390> PMID: 25298419
28. Mootha VV, Hussain I, Cunnusamy K, Graham E, Gong X, Neelam S, et al. TCF4 Triplet Repeat Expansion and Nuclear RNA Foci in Fuchs' Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci.* 2015; 56(3):2003–11. <https://doi.org/10.1167/iovs.14-16222> PMID: 25722209
29. Nakano M, Okumura N, Nakagawa H, Koizumi N, Ikeda Y, Ueno M, et al. Trinucleotide Repeat Expansion in the TCF4 Gene in Fuchs' Endothelial Corneal Dystrophy in Japanese. *Invest Ophthalmol Vis Sci.* 2015; 56(8):4865–9. <https://doi.org/10.1167/iovs.15-17082> PMID: 26218914
30. Thalamuthu A, Khor CC, Venkataraman D, Koh LW, Tan DT, Aung T, et al. Association of TCF4 gene polymorphisms with Fuchs' corneal dystrophy in the Chinese. *Invest Ophthalmol Vis Sci.* 2011; 52(8):5573–8. <https://doi.org/10.1167/iovs.11-7568> PMID: 21659310
31. Mootha VV, Gong X, Ku HC, Xing C. Association and familial segregation of CTG18.1 trinucleotide repeat expansion of TCF4 gene in Fuchs' endothelial corneal dystrophy. *Invest Ophthalmol Vis Sci.* 2014; 55(1):33–42. <https://doi.org/10.1167/iovs.13-12611> PMID: 24255041
32. Dimasi DP, Burdon KP, Hewitt AW, Fitzgerald J, Wang JJ, Healey PR, et al. Genetic investigation into the endophenotypic status of central corneal thickness and optic disc parameters in relation to open-angle glaucoma. *Am J Ophthalmol.* 2012; 154(5):833–42.e2. <https://doi.org/10.1016/j.ajo.2012.04.023> PMID: 22840486

33. Awadalla MS, Burdon KP, Souzeau E, Landers J, Hewitt AW, Sharma S, et al. Mutation in TMEM98 in a large white kindred with autosomal dominant nanophthalmos linked to 17p12-q12. *JAMA Ophthalmol*. 2014; 132(8):970–7. <https://doi.org/10.1001/jamaophthalmol.2014.946> PMID: 24852644
34. Burdon KP, Macgregor S, Bykhovskaya Y, Javadiyan S, Li X, Laurie KJ, et al. Association of polymorphisms in the hepatocyte growth factor gene promoter with keratoconus. *Invest Ophthalmol Vis Sci*. 2011; 52(11):8514–9. <https://doi.org/10.1167/iovs.11-8261> PMID: 22003120
35. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901
36. Hamill CE, Schmedt T, Jurkunas U. Fuchs endothelial cornea dystrophy: a review of the genetics behind disease development. *Semin Ophthalmol*. 2013; 28(5–6):281–6. <https://doi.org/10.3109/08820538.2013.825283> PMID: 24138036
37. Vasanth S, Eghrari AO, Gapsis BC, Wang J, Haller NF, Stark WJ, et al. Expansion of CTG18.1 Trinucleotide Repeat in TCF4 Is a Potent Driver of Fuchs' Corneal Dystrophy. *Invest Ophthalmol Vis Sci*. 2015; 56(8):4531–6. <https://doi.org/10.1167/iovs.14-16122> PMID: 26200491
38. Orr HT, Zoghbi HY. Trinucleotide repeat disorders. *Annu Rev Neurosci*. 2007; 30:575–621. <https://doi.org/10.1146/annurev.neuro.29.051605.113042> PMID: 17417937
39. Brusco A, Gellera C, Cagnoli C, Saluto A, Castucci A, Michielotto C, et al. Molecular genetics of hereditary spinocerebellar ataxia: mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families. *Arch Neurol*. 2004; 61(5):727–33. <https://doi.org/10.1001/archneur.61.5.727> PMID: 15148151
40. Nelson DL, Orr HT, Warren ST. The Unstable Repeats—Three Evolving Faces of Neurological Disease. *Neuron*. 2013; 77(5):825–43. <https://doi.org/10.1016/j.neuron.2013.02.022> PMID: 23473314
41. Kozlowski P, de Mezer M, Krzyzosiak WJ. Trinucleotide repeats in human genome and exome. *Nucleic Acids Res*. 2010; 38(12):4027–39. <https://doi.org/10.1093/nar/gkq127> PMID: 20215431
42. Santos R, Lefevre S, Sliwa D, Seguin A, Camadro JM, Lesuisse E. Friedreich ataxia: molecular mechanisms, redox considerations, and therapeutic opportunities. *Antioxidants and redox signaling* 2010; 13(5):651–90. <https://doi.org/10.1089/ars.2009.3015> PMID: 20156111
43. Oldak M, Ruskowska E, Udziela M, Oziębło D, Bińczyk E, Ścieżyńska A, et al. Fuchs Endothelial Corneal Dystrophy: Strong Association with rs613872 Not Paralleled by Changes in Corneal Endothelial TCF4 mRNA Level. *BioMed research international*. 2015; 2015:640234. <https://doi.org/10.1155/2015/640234> PMID: 26451375
44. Echeverria GV, Cooper TA. RNA-binding proteins in microsatellite expansion disorders: mediators of RNA toxicity. *Brain Res*. 2012; 1462:100–11. <https://doi.org/10.1016/j.brainres.2012.02.030> PMID: 22405728
45. Miller JW, Urbinati CR, Teng-Umnuy P, Stenberg MG, Byrne BJ, Thornton CA, et al. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J*. 2000; 19(17):4439–48. <https://doi.org/10.1093/emboj/19.17.4439> PMID: 10970838
46. Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, Harper PS, et al. Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet*. 2002; 11(7):805–14. PMID: 11929853
47. Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, et al. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet*. 2001; 10(19):2165–70. PMID: 11590133
48. Konieczny P, Stepniak-Konieczna E, Sobczak K. MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res*. 2014; 42(17):10873–87. <https://doi.org/10.1093/nar/gku767> PMID: 25183524
49. Zoghbi HY, Orr HT. Pathogenic mechanisms of a polyglutamine-mediated neurodegenerative disease, spinocerebellar ataxia type 1. *J Biol Chem*. 2009; 284(12):7425–9. <https://doi.org/10.1074/jbc.R800041200> PMID: 18957430
50. Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell*. 2014; 56(2):311–22. <https://doi.org/10.1016/j.molcel.2014.08.027> PMID: 25263597
51. Goodwin M, Mohan A, Batra R, Lee KY, Charizanis K, Fernandez Gomez FJ, et al. MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. *Cell Rep*. 2015; 12(7):1159–68. <https://doi.org/10.1016/j.celrep.2015.07.029> PMID: 26257173
52. Li Y, Polak U, Bhalla AD, Rozwadowska N, Butler JS, Lynch DR, et al. Excision of Expanded GAA Repeats Alleviates the Molecular Phenotype of Friedreich's Ataxia. *Mol Ther*. 2015; 23(6):1055–65. <https://doi.org/10.1038/mt.2015.41> PMID: 25758173

53. O'Rourke JR, Swanson MS. Mechanisms of RNA-mediated disease. *J Biol Chem*. 2009; 284(12):7419–23. <https://doi.org/10.1074/jbc.R800025200> PMID: 18957432
54. Du J, Aleff RA, Soragni E, Kalari K, Nie J, Tang X, et al. RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. *J Biol Chem*. 2015; 290(10):5979–90. <https://doi.org/10.1074/jbc.M114.621607> PMID: 25593321
55. Wieben ED, Aleff RA, Tang X, Butz ML, Kalari KR, Highsmith EW, et al. Trinucleotide Repeat Expansion in the Transcription Factor 4 (TCF4) Gene Leads to Widespread mRNA Splicing Changes in Fuchs' Endothelial Corneal Dystrophy. *Invest Ophthalmol Vis Sci*. 2017; 58(1):343–52. <https://doi.org/10.1167/iovs.16-20900> PMID: 28118661