Title: Role of the \textit{TCF4} gene intronic variant in normal variation of corneal endothelium.

Authors: David A. Mackey$^1$, Nicole M. Warrington$^2$, Alex W. Hewitt$^3$, Sandra K Oates$^1$, Seyhan Yazar$^1$, Alla Soloshenko$^1$ Geoffrey J Crawford$^1$ Jenny A Mountain$^4$, Craig E Pennell$^2$.

1. Lions Eye Institute and the Centre of Ophthalmology and Visual Science, The University of Western Australia, Perth, Western Australia
2. School of Women's and Infants' Health, The University of Western Australia, Perth, Western Australia
3. Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, University of Melbourne, Victoria
4. Telethon Institute for Child Health Research, The University of Western Australia, Perth, Western Australia

Financial Disclosure: Startup funds were provided by the University of Western Australia and the Australian Foundation for the Prevention of Blindness.

Conflict of Interest: The authors have no proprietary or commercial interest in any materials discussed in this article.

Running head: \textit{TCF4} gene in normal variation of corneal endothelium.

Address for correspondence:
Prof. David Mackey
Lions Eye Institute,
2 Verdun Street,
Nedlands 609 Western Australia
Abstract

**Objective:** To identify early features of Fuchs’s endothelial dystrophy (FED) in carriers of the rs613872(G) transcription factor 4 gene (*TCF4*) aged 20-21 years.

**Design:** Prospective cohort study

**Participants:** People aged 20-21 years previously enrolled in the Western Australia Pregnancy (Raine) Cohort.

**Methods:** Specular microscopy was performed using a non-contact specular microscopy (EM-3000, Tomey, Nagoya, Japan). Individual genotype data were extracted from the genome-wide Illumina 660 Quad Array. Analysis of the association between the rs613872 risk allele in *TCF4* and specular microscopy measurements was conducted.

**Main Outcome Measures:** Association between the rs613872 risk allele and corneal endothelial cell density as well as the coefficient of variation in cell shape.

**Results:** Genotype and specular microscopy data were available for a total of 445 participants (46% female). The median cell density was 2851 and 2850 cells/mm² in the right and left eyes, respectively. No significant differences between inter-eye variability in endothelial cell density were seen (right eye:left eye correlation = 0.64); however, a significant difference in variability of endothelial cell density between males and females was observed (male OD:2839±124 cells/mm² OS:2845±124 cells/mm² vs female OD:2838±134 cells/mm² OS:2842±132 cells/mm²; OD: p=0.0013, OS:p=0.0016). Eleven individuals were homozygous for the rs613872 risk allele. We found no association between rs613872 genotype and cell density or coefficient variation.
One of the eleven homozygous GG individuals was found to have a gutta in one sample field on specular microscopy while two of 297 TT individuals also had a gutta each in one sample field.

**Conclusions:** We were unable to detect an association between *TCF4* rs613872 genotype and the variation in corneal endothelial cell density or variation in cell morphology in a healthy young adult population.

**Financial Disclosure(s):** The authors have no proprietary or commercial interest in any materials discussed in this article.
Fuch’s Endothelial Dystrophy (FED) is a leading indication for corneal transplants in individuals living in developed countries.(1) FED is a slowly progressive, bilateral but asymmetric posterior corneal dystrophy clinically evident in the fifth and sixth decades of life(2,3) and later causes endothelial decompensation resulting in visual impairment.(4) The first sign of FED is the presence of guttae on Descemet’s membrane.(5) The relationship between corneal endothelial cell count and the variability of cell size (polymegathism) with FED is inconsistent in the literature. (6, 7) FED is a complex heterogeneous disease affecting approximately 38% of first-degree relatives of cases.(11) However, previous work suggests that there is little difference in corneal endothelial cell count between first-degree relatives of individuals with FED.(8) Approximately 70% of Fuch’s original cohort was female and more recently Krachmer found that women were 2.5 times more likely than men to be affected. (9, 10, 11) Early onset disease has been associated with mutations in the \textit{COL8A2} gene located on chromosome 1p, presumably reflecting that the α2 subunit of collagen 8 is a component of Descemet’s layer.(12) Genome-wide association studies (GWAS) is a relatively new means of investigation and has been successful in identifying single nucleotide polymorphisms (SNPs) associated with macular degeneration, pseudoexfoliation, myopia and glaucoma (amongst others).

GWAS holds great promise, particularly in the case of common complex diseases, because no prior information about underlying molecular mechanisms is required. In contrast to a candidate gene study, where typically only one particular gene locus is interrogated, in a GWAS several hundred thousand SNPs dispersed across the genome are simultaneously investigated for association with a disease. Mutations identified in \textit{KCNJ13} on chromosome 2q, \textit{SLC4A11} on chromosome 20p, and \textit{ZEB1} on chromosome 10p have also been associated with rare forms of FED.(13, 14, 15, 16) Linkage of FED to chromosome 5q, 9p, 13p and 18q has been described.(16, 17, 18,
Recently, a genome-wide association study (GWAS) identified a SNP in the transcription factor 4 gene (TCF4) (which encodes E2-2, a member of the E protein family) as being strongly associated with FED. Individuals with two copies of the high-risk allele were almost 30 times more likely to develop FED. A further study of 170 late onset FED cases and 180 age-matched controls of Northern European descent interrogated 9 SNPs spanning the TCF4 locus and identified the rs61872 G risk allele as being significantly associated with late onset FED (p=4.28 x 10^{-15}, OR = 4.2) In this study we sought to specifically investigate the role of the TCF4 gene in the normal variation in corneal endothelium in a population-based cohort of 20- to 21-year-old adults.

Methods:

The Western Australian Pregnancy Cohort (Raine) Study was established in Perth in 1989 when an unselected cohort of 2900 pregnant women was recruited at 16-18 weeks’ gestation from King Edward Memorial Hospital and surrounding practices. It is one of the largest ongoing prospective cohorts of pregnancy, childhood, adolescence, and young adulthood. The primary purpose of the Raine cohort was to investigate how events during pregnancy and at the time of birth subsequently influence health. Participants in the Raine study are currently being recruited and assessed as part of their 20- to 21-year-old follow-up, which includes each participant having specular microscopy measurements as part of a broad ophthalmic phenotyping. This is the first time that a comprehensive ophthalmic examination has been included as part of the Raine study follow up. A total of 522 individuals took part in a pilot study for the Raine Eye Health Study
(REHS). The REHS will compare ophthalmic diseases, traits and biometry with previously assembled GWAS and environmental data of the cohort.

The REHS was conducted in accordance with the tenets of the Declaration of Helsinki, and the protocol was approved by the institutional review board at the University of Western Australia. Informed consent was obtained from all participants at the initiation of the examination session.

Of the 522 individuals, 77 participants did not have genetic data available. A subset of 445 individuals was selected for analysis based on the following criteria: singleton birth and no siblings in the study, Caucasian ethnicity, DNA sample, no congenital abnormalities and available phenotype and genotype data.

Ophthalmic measurements were obtained from a maximum of 873 eyes (17 eyes did not have adequate images) of 445 participants examined with non-contact specular microscopy (EM-3000, Tomey, Nagoya, Japan). For each eye, fifteen serial photographs of central corneal endothelium were recorded by using the auto alignment and auto shots functions of the instrument. The best quality image is automatically selected and analyzed by the pre-installed software of the instrument (Cell and Layer Analyser; Tomey). The following parameters were automatically determined using the fixed frame counting method in an area set to 100 µm²; the number of analyzed cells, cell density (CD), average cell size, standard deviation (SD), coefficient variation (CV), minimum and maximum values for cell area. No cell border correction was made. Central corneal thickness (CCT) was measured with the Tomey Specular Microscopy along with the corneal endothelial cell counts. Grading was not conducted at the time of testing and poor quality scans were later excluded and not repeated.
Study individual genotype data for rs613872 were extracted from the genome-wide Illumina 660 Quad Array. Briefly, the genotyping was performed on the Illumina BeadArray Reader at the Centre for Applied Genomics (Toronto, Ontario, Canada) using 250 nanograms of DNA. Both the left and right eye measurements were analyzed separately. Cross-sectional analyses of the eye measures were performed using multivariate linear regression adjusting for gender and rs613872 genotype coded as AA, AC or CC (i.e. co-dominant genetic model). The platform we used coded for the other strand of DNA, the SNP that was genotyped as A/C on our GWAS is presented that way in the text and tables of this manuscript but not the abstract, where we have transliterated to the HapMap database which used the other strand with T/G alleles. We have 100% concordance between our genotyped and imputed data for this SNP.

Age was investigated for association with each outcome but was non-significant over this small age range and was excluded from analysis. CD was approximately normally distributed, but CV was right skewed so a log transformation was performed prior to analysis. All analyses were performed using the statistical graphics software R version 2.7.1\textsuperscript{(23)} A post-hoc power calculation was carried out in the Quanto software package\textsuperscript{(24)} to investigate whether this study had adequate power to detect a genetic association. In this study we had more than 80% power to detect an $\beta\geq57$ cells/mm$^2$ for CD and $\beta\geq0.05$ for log(CV) under a log-additive genetic model for the SNP rs613872 at the $\alpha=0.05$ level.

Results:

Of the 445 participants who were included in the analysis, 205 (46%) were female and 240 (54%) male. The average age of the participants was 20.14 years (SD = 0.29).
The median CD and coefficient variations were determined for right and left eyes (Table 1). Regardless of the gender, the median CD was 2851 and 2850 cells/mm$^2$ in the right and left eyes, respectively. There was no significant difference in CD between the left and right eye (correlation = 0.64). There was no significant difference in variability in endothelial cell size between the left and the right eye (correlation = 0.43); however, a difference was observed between males and females with respect to variability in endothelial cell density, which appeared to be due to a difference in standard deviation rather than average size.

**Table 1** Median (interquartile range) of cell density and coefficient of variation in pilot study, including the differences seen between males and females.

(Table 1 here)

The minor allele frequency of rs613872 was 0.18 (Number in each genotype category is: AA=297, AC=137, CC=11). The SNP was in Hardy-Weinberg equilibrium (P=0.42). and had 100% call rate. Association analysis between the CD and rs613872 genotype in the left and right eye is shown in Table 2. Association analysis between the cell size variability and rs613872 genotype in the left and right eye is displayed in Table 3.

**Table 2**

Table 2 presents the effect size and statistical significance of each risk genotype of rs613872 in comparison to the major homozygote (AA) for the genetic association analysis of both cell density and coefficient of variation.
In the HapMap A is annotated by the other strand T and C as G.

No significant association was detected between either CD or cell size variability and rs613872 genotype in this sample of 445 individuals. An analysis of the other endothelial cell counts available, including number of cells, average and SD of cell size, minimum and maximum cell size, demonstrated no obvious associations with the rs613872 genotype, except for in the minimum cell size variable, which indicated an increased minimum cell size with the more copies of the risk (C) allele.

Association analysis using an additive genetic model was conducted for each SNP within an 800 kilobase of chromosome 18 that spans the TCF4 gene and beyond from the imputed data. Figure 1 indicates that SNPs across the TCF4 region were not associated with CD and cell size variability in this sample after a Bonferonni adjustment for the multiple tests conducted (P<7x10^{-5}).

(Figure 1 here)
Association results of endothelial cell density and size variability for single nucleotide polymorphisms (SNPs) at the Chromosome 18q21 TCF4 locus. SNPs are plotted according to chromosomal location (x-axis) and the level of association for each SNP is displayed as the $-\log_{10} p$ value, with stronger associations being displayed higher on the y-axis. The colour scheme indicates linkage disequilibrium between rs613872 and other SNPs in the region, where warmer colours represent stronger correlation between the relative SNPs. The corresponding position of the TCF4 gene is depicted in the ideogram at the bottom of each plot. Panels display the right and left eyes separately.

One of the homozygous carriers of the risk allele showed a single endothelial gutta in the measurement zone of the left eye. The participants had specular microscopy that was not graded at the time of examination and follow up of abnormal results was not possible. No other carrier had any abnormality detected. Two of the non-carriers had a single endothelial gutta each in their sample field.

As Fuchs endothelial dystrophy is associated with corneal edema we also compared CCT in the cohort. CCT measurements were available on 436 right eyes (236 male, 200 female) and 437 left eyes (238 male, 199 female). Mean (SD) for CTT measurement total set genders combined: Right: 534.7µm (34.6µm), Left: 527.9µm (34.7µm); genders separated: Male Right: 533.6µm (35.8µm), Female Right: 535.9µm (33.3µm), Male Left: 526.6µm (36.5µm), Female Left: 529.5µm (32.4µm). There was no significant difference between male and females; however, there was an unexplained difference between right and left eyes (p=0.004).
Discussion:

Identification of the TCF4 locus has enabled detection of those at high risk for developing FED. Through following people who harbour FED causing TCF4 mutations closely we hope to recognize early signs of the disease and to identify other potential risk factors that exacerbate progression in those at increased genetic risk.

This study reveals that guttae, one of the early signs of FED, were not common among those young adults carrying the high risk alleles for the disease (one in 22 eyes). The youngest affected individual with a COL8A2 mutation in other genetic studies of FED was pre-symptomatic in his third decade with 1–2 mm of grouped central corneal guttae. In addition there was no difference in corneal endothelial CD or polymegathism when compared to wild type controls.

Follow-up studies of at-risk individuals in cohorts of young adults such as this will be of value in understanding the onset of visual impairment from FED. The finding in this study of an increased variability of cell size (greater distribution) in females compared with males may explain the higher prevalence of FED dystrophy in women. Further genetic investigation of endothelial cell counts with GWAS may identify a sex-limited gene associated with this and FED. The G allele of rs613872 was not previously found to be more prevalent in females with FED. Although the G allele of rs613872 was associated with increased risk for FED, it was not associated with severity of disease. Limitations of this study include the relatively small number of individuals involved. It is hoped that we will be able to expand the study to a full 1,500 individuals with available genetic data in the coming years. Only one specular microscopy reading was taken and this was not graded at the time of participant assessment. Moreover, further corneal examination to confirm other pathology and other signs of FED was not conducted.
In conclusion identification of risk alleles in the TCF4 gene by Baratz and colleagues\textsuperscript{20} allows pre-symptomatic detection of individuals at higher risk of FED.

We did not identify any significant abnormality in endothelial cell counts in homozygous carriers aged 20-21 years and guttae were not common in high-risk individuals of this age (one of 11 homozygous carriers compared with two of 297 non-carriers).

\textbf{Acknowledgements}

The authors are grateful to the Raine Study participants, their families and the Raine Study research staff for cohort coordination and data collection. Startup funds for the Raine Eye Health Study pilot were provided by the University of Western Australia and the Australian Foundation for the Prevention of Blindness. The authors gratefully acknowledge the NH\&MRC for their long term contribution to funding the study over the last 20 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, The Telethon Institute for Child Health Research and Women and Infants Research Foundation. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (a National Health and Medical Research Council of Australia National Enabling Facility). The authors also acknowledge the support of the National Health and Medical Research Council of Australia (Grant ID 572613, ID 403981 and ID 003209) and the Canadian Institutes of Health Research (Grant ID MOP-82893).


Table 1: Median (interquartile range) of cell density and coefficient of variation in pilot study, including the differences seen between males and females

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total (n=445)</th>
<th>Males (n=240)</th>
<th>Females (n=205)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell Density (cells/mm²)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>2851 (2692, 3006)</td>
<td>2846 (2688, 2998)</td>
<td>2864 (2715, 3008)</td>
<td>0.502</td>
</tr>
<tr>
<td>Left</td>
<td>2850 (2706, 2974)</td>
<td>2843 (2707, 2962)</td>
<td>2858 (2704, 2985)</td>
<td>0.722</td>
</tr>
<tr>
<td><strong>Coefficient Variation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>37.00 (33.00, 42.00)</td>
<td>36.00 (32.00, 40.00)</td>
<td>37.50 (34.00, 44.00)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Left</td>
<td>36.00 (33.00, 41.00)</td>
<td>35.00 (32.00, 40.00)</td>
<td>38.00 (34.00, 42.00)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Average Cell size (µm²)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>350.5 (333.0,371.2)</td>
<td>351.0 (333.2,372.0)</td>
<td>349.0 (332.2,368.0)</td>
<td>0.473</td>
</tr>
<tr>
<td>Left</td>
<td>351.0 (336.0,369.0)</td>
<td>351.5 (337.0,369.0)</td>
<td>350.0 (334.5,369.5)</td>
<td>0.709</td>
</tr>
<tr>
<td><strong>SD Cell size (µm²)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>127.0 (114.0,148.2)</td>
<td>124.0 (113.0,142.8)</td>
<td>132.0 (116.2,156.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>Left</td>
<td>127.0 (114.0,146.0)</td>
<td>124.0 (112.8,140.2)</td>
<td>134.0 (115.5,150.5)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

* T-test was used to test for a difference in means of normally distributed outcomes (cell density) and wilcoxon rank sum test was used to test for a difference in medians for skewed outcomes (coefficient of variation, average and SD)

Table 2: Results of genetic association analysis between cell density, coefficient of variation and rs613872 genotype

<table>
<thead>
<tr>
<th></th>
<th>Left Eye</th>
<th></th>
<th>Right Eye</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell Density</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Effect (SE)</td>
<td>P-value</td>
<td>Effect (SE)</td>
<td>P-value</td>
</tr>
<tr>
<td>rs613872 heterozygote (AC)</td>
<td>-21.34 (24.77)</td>
<td>0.389</td>
<td>8.58 (25.87)</td>
<td>0.740</td>
</tr>
<tr>
<td>rs613872 risk homozygote (CC)</td>
<td>-61.06 (72.52)</td>
<td>0.400</td>
<td>-116.78 (76.07)</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>Loge(coefficient of variation)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Effect (SE)</td>
<td>P-value</td>
<td>Effect (SE)</td>
<td>P-value</td>
</tr>
<tr>
<td>rs613872 heterozygote (AC)</td>
<td>0.018 (0.019)</td>
<td>0.329</td>
<td>0.025 (0.019)</td>
<td>0.193</td>
</tr>
<tr>
<td>rs613872 risk homozygote (CC)</td>
<td>-0.028 (0.054)</td>
<td>0.603</td>
<td>-0.049 (0.059)</td>
<td>0.405</td>
</tr>
</tbody>
</table>