Expansion of \textit{CTG18.1} Trinucleotide Repeat in \textit{TCF4} Is a Potent Driver of Fuchs’ Corneal Dystrophy


1The Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
2Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, United States
3Center for Human Disease Modeling, Duke University, Durham, North Carolina, United States

Purpose. To analyze the expansion of \textit{CTG18.1} allele associated with Fuchs’ corneal dystrophy (FCD) in our large cohort of late-onset FCD cases.

Methods. CTG repeats within the \textit{CTG18.1} allele were estimated by short tandem repeat (STR) and triplet primed PCR (TP-PCR) assays in our large cohort of 574 late-onset FCD cases and 354 controls and large multigeneration familial cases. The age versus severity relationships were analyzed in FCD genotypes, namely, nonexpanded (N/N), monoallelic expansion (N/X), and biallelic expansion (X/X) with N \leq 40 CTG monomers. The threshold for causality conferred by an expansion of \textit{CTG18.1} was identified by excluding the population of FCD cases who harbored an allele length equivalent to the maximum CTG monomers observed in the controls.

Results. The expanded \textit{CTG18.1} for \textit{(CTG)}_{n>40} showed a strong association (P = 1.56 \times 10^{-82}) with FCD. Importantly, we delineated the threshold of expansion to 103 CTG repeats above which the allele confers causality in 17.8% of FCD cases. Regression analyses demonstrated a significant correlation between disease severity and age in individuals who harbor either a monoallelic expansion or a biallelic expansion at \textit{(CTG)}_{n>40}. These analyses helped predict FCD in two previously unaffected individuals based on their \textit{CTG18.1} expansion genotype.

Conclusions. A monoallelic expansion of \textit{CTG18.1} contributes to increased disease severity and is causal at \textit{(CTG)}_{n>103}, whereas a biallelic expansion is sufficient to be causal for FCD at \textit{(CTG)}_{n>40}. This study highlights the largest contributory causal allele for FCD.

Keywords: Fuchs’ corneal dystrophy, \textit{TCF4}, \textit{CTG18.1}, repeat expansion, corneal endothelium

Fuchs’ corneal dystrophy (FCD; Mendelian Inheritance in Man [MIM] 136800) is characterized by a progressive decrease in visual acuity due to dysfunctional corneal endothelium. The occurrence of FCD has been estimated at 4% of the population above the age of 40. Pathological changes include thickening of Descemet membrane and formation of collagenous excrescences known as guttae. Fuchs’ corneal dystrophy accounts for a significant fraction of corneal transplantation and endothelial keratoplasty (EK) performed in the United States.

Fuchs’ corneal dystrophy is genetically heterogeneous and is inherited as an autosomal dominant trait with variable penetrance and expressivity. Linkage in large, multigenerational families has mapped four FCD loci (\textit{FCD1–4}, on chromosomes 13, 18, 5, and 9, respectively). Moreover, a combination of candidate gene studies and positional cloning has implicated mutations in \textit{SLC4A11}, \textit{TCF8}, \textit{LOXHD1}, and \textit{AGBL1} in the pathogenesis of the disease. Baratz and colleagues performed a genome-wide association scan and identified a \textit{TCF4} (chromosome 18) intronic single nucleotide polymorphism (SNP), rs613872, that was associated significantly with late-onset FCD; this association has been replicated in multiple cohorts. Recently, Wieben and colleagues reported that expansion of trinucleotide CTG repeat within \textit{TCF4}, in linkage disequilibrium to rs613872, showed a significant association with late-onset FCD. Mootha and colleagues likewise replicated this association in multiple ethnic populations.

Here, we have analyzed the CTG trinucleotide repeat expansion in \textit{TCF4} within our large cohort of both familial and sporadic cases of late-onset FCD to confirm the association and delineate the threshold of expansion for causality. We have utilized this large data set to perform regression analysis of age versus severity stratified by genotypes (no expansion, a monoallelic, or a biallelic expansion). These analyses helped predict FCD in two individuals who were previously unaffected, thereby suggesting a biallelic expansion as a possible genomic marker for FCD.

Methods

Study Participants

We recruited a large cohort of 574 FCD patients. All individuals underwent a thorough ophthalmic evaluation including slit-lamp biomicroscopy. Grading of phenotype severity was based on the scale proposed by Krachmer and colleagues, but modified as described previously. These modifications, summarized in Supplementary Table S1, reflect findings that corneal thickness and edema gradually increase throughout all stages of severity, and that subclinical disease (5–11 central
guttae) is itself a stage that correlates with FCD-associated genetic variation.21 All study participants consented to the study protocol approved by the Joint Committee on Clinical Investigation at the Johns Hopkins University School of Medicine and in accordance with the Declaration of Helsinki and Health Insurance Portability and Accountability Act regulations. Written informed consent was obtained from all individuals prior to enrollment. All participants provided a blood sample (10–20 mL) and genomic DNAs were extracted from white blood cells using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA).

Genotyping of CTG18.1 Allele
The trinucleotide repeats in intron 3 of TCF4 were determined by a previously published STR assay by Wieben et al.19 The 5′ forward primer (P1) was labeled with 6-carboxyfluorescein (FAM) (5′ CAGATGAGTTTGGTAAAGATG 3′), and unlabeled reverse primer (P2) (5′ TTGACGGCTTCTGTGTTG 3′) was used in PCR amplification carried out with Platinum PCR Supermix High Fidelity (Life Technologies, Carlsbad, CA, USA) according to the cycling conditions described by Wieben et al.19 Triplet repeat primed PCR (TP-PCR) reactions were carried out with modification to the published protocol by Mootha et al.20 Briefly, 1 μl genomic DNA was mixed with 14 μl Platinum PCR Supermix High Fidelity (Life Technologies) and 250 nM final concentrations of P1 and reverse primer (P4) (5′ TACGCATCCCCAGTTGAGCGACGCAGCCAGCAGCAGCAGCC 3′) and cycled at 95°C for 5 minutes; 95°C for 30 seconds, 58°C for 1 minute, 68°C for 4 minutes × 20 cycles; 68°C for 10 minutes. An additional 15 seconds was introduced in each of the 8th to 20th cycles. Two microliters of PCR product were used as a template for a second PCR reaction by mixing with Platinum PCR Supermix with 250 nM final concentration of 6-FAM-labeled forward primer and unlabeled reverse primer (P3) (5′ TACGCATCCCCAGTTGAGCGACGCAGCCAGCAGCAGCAGCC 3′) and cycled at 95°C for 5 minutes; 95°C for 30 seconds, 62°C for 1 minute, 68°C for 6 minutes × 25 cycles; 68°C for 10 minutes. Polymerase chain reaction products were resolved in ABI3730XL DNA Analyzer (Applied Biosystems, Inc., Foster City, CA, USA) with GeneScan 500 LIZ dye size standard mix (Life Technologies, Carlsbad, CA, USA). The results were analyzed using Gene Mapper (Applied Biosystems, Inc.). Statistical Analysis
χ² and Fisher’s exact tests were used to examine whether genotypes were associated with sex and patient case-control status. The relationship between disease severity and age was examined using a linear regression model stratified by genotype. Nonparametric Kruskal-Wallis rank testing was used to compare affection between genotypes and to compare real-time PCR data between the CTG18.1 genotypes. The mean age from different groups was compared using two-sample t-tests with unequal variances. All analyses were performed using STATA version 12.0 (Stata Statistical Software, College Station, TX, USA).

RESULTS
To understand the genetic basis of FCD, we have recruited a large cohort of both familial and sporadic cases of late-onset FCD. Our cohort of 574 cases comprised 372 females and 202 males, consistent with previous reports of increased FCD occurrence in females. The mean age and the Krachmer grading did not differ significantly between males and females (Supplementary Table S2).

Association of CTG18.1 With FCD
To determine the contribution of the CTG18.1 expansion in our cohort, we genotyped all 574 late-onset FCD cases representing one affected individual per family, as well as 354 ethnically matched control subjects consisting of individuals above the age of 60 without clinical signs of FCD by utilizing an established PCR-based short tandem repeat (STR) assay.19,20 We were unable to detect two alleles in 18% of controls (64/354) and 10.3% of cases (59/574), suggesting that either these individuals may have the same number of CTG monomers in both alleles (homozygous for CTG18.1 allele) or the length of the expanded allele might lie beyond the detectable limits of the STR assay (Supplementary Fig. S1). To examine these possibilities, we used a previously published TP-PCR assay with minor modifications (see Methods; representative resolution of PCR products for each genotype is shown in Supplementary Fig. S1). The TP-PCR assay confirmed that 31/574 FCD cases (5.4%) and 59/354 (16.6%) control individuals were homozygous for the CTG18.1 allele.
The two previous studies by Wieben and colleagues19 and Mootha and colleagues,20 which reported the association of the CTG18.1 expansion with FCD, considered CTG units below 40 and 50, respectively, as unexpanded. We, therefore, designated any allele with 40 or more CTG repeats as expanded (X) and less than 40 CTG repeats as not expanded (N). Based on these criteria, we found one expanded allele (N/X) in 59.8% (343/574) and two expanded alleles (X/X) in 2.3% (13/574) of FCD cases. In sharp contrast, only 3.67% of controls harbored one expanded allele; none carried two expanded alleles (P = 1.52 × 10⁻⁸²; Fig. 1A; Table). CTG18.1 Expansion Contributes to Disease Severity
Since the affection grade of our cohort ranged from 0.5 (trace) to 6 (severe phenotype that required transplantation), we next examined the genotyping data to determine if the 38% of cases who were N/N had a lower affection grade (modified Krachmer grading). The median modified Krachmer grading scores were 2 (N/N, n = 218), 3 (N/X, n = 343), and 4 (X/X, n = 13) for the respective genotypes, suggesting that the expansion of CTG18.1 contributes to a higher disease severity (P < 7.87 × 10⁻⁵, Kruskal-Wallis rank test). We also considered age as a factor that may result in lower affection; we found no significant difference between the average age of the different genotypic groups (Supplementary Table S3, P < 0.32; t-test, as age is normally distributed for all the groups). Since the female FCD population is almost two times higher in our cohort, we also tested if the expanded allele is more prevalent in females. Among the 372 females with FCD, 55.4% were N/X and 1.6% were X/X, whereas among 202 males, 67.8% were N/X and 3.5% were X/X. These data suggested that, within our FCD cohort, the expanded allele is more prevalent in males (57% females, 71.3% males; P = 4.82 × 10⁻⁴ by χ² test).

Next, we investigated whether the expanded allele alters any possible age versus severity profiles stratified by genotypes. We did not observe a relationship between age and disease severity for the FCD cases in the N/N category using a linear regression model, with a slope (average severity increase per year) of −0.008 (P = 0.362); however, the slopes for the N/X and X/X groups of FCD cases were 0.019 (P = 0.021) and 0.094 (P = 0.008), respectively, suggesting a significant
correlation between age and disease progression in FCD cases with the expanded allele (Fig. 2).

**Delineating the Threshold of Expansion for Causality**

We utilized the above data to determine the number of CTG repeats in CTG18.1 above which an expansion might be deterministic for FCD. We found that 3.7% of unaffected controls harbored a heterozygous expanded CTG18.1 allele at (CTG)>40, indicating that this number of repeats does not constitute causality for FCD. Therefore, we increased the threshold of the CTG to the largest repeat size (103 repeats) present in the controls (1/354) (Fig. 1B) to determine any causality conferred by this locus. The elimination of the presence of expansion above 103 CTG repeats in controls accounted for 17.8% of FCD cases who harbored the expanded allele sufficient to be causal (\( P = 1.5 \times 10^{-23} \)), suggesting that the FCD cases who harbor the CTG18.1 causal allele show a significantly increased progression (Supplementary Fig. S2).

In order to understand if the expanded alleles segregated with FCD, we analyzed the CTG repeats within the CTG18.1 allele in two large multigenerational families previously recruited for the FCD study (Fig. 3). We were particularly intrigued by the observation of two individuals (a 34-year-old female and a 56-year-old male) within these families with CTG18.1 biallelic expansion for (CTG)>40 and negative for FCD at the time of enrollment (Fig. 3). We therefore hypothesized that these individuals developed the disease over the several years that had elapsed since their initial examination. A follow-up clinical examination revealed FCD corresponding to modified Krachmer grading 1 for the female (34 years) and 3 for the male (56 years). This observation, together with the absence of biallelic expansion of CTG18.1 for (CTG)>40 from controls, suggests that this genotype is sufficient to cause FCD.

**TABLE.** Distribution of the CTG18.1 Expansion Genotype in Controls and FCD Cases

<table>
<thead>
<tr>
<th></th>
<th>CTG in Both Alleles &lt; 40</th>
<th>CTG in 1 Allele &gt; 40</th>
<th>CTG in Both Alleles &gt; 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls, n = 354</td>
<td>341, 96.3%</td>
<td>13, 3.7%</td>
<td>0, 0%</td>
</tr>
<tr>
<td>FCD, n = 574</td>
<td>218, 38%</td>
<td>343, 59.8%</td>
<td>13, 2.3%</td>
</tr>
<tr>
<td>Females, n = 372</td>
<td>160</td>
<td>206</td>
<td>6</td>
</tr>
<tr>
<td>Males, n = 202</td>
<td>58</td>
<td>137</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CTG in Both Alleles &lt; 103</th>
<th>CTG in 1 Allele &gt; 103</th>
<th>CTG in Both Alleles &gt; 103</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls, n = 354</td>
<td>354, 100%</td>
<td>0, 0%</td>
<td>0, 0%</td>
</tr>
<tr>
<td>FCD, n = 574</td>
<td>390, 82.2%</td>
<td>102, 17.8%</td>
<td>0, 0%</td>
</tr>
<tr>
<td>Females, n = 372</td>
<td>316</td>
<td>56</td>
<td>46</td>
</tr>
<tr>
<td>Males, n = 202</td>
<td>156</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( P = 1.114 \times 10^{-41} \) for (CTG)>40 and \( 1.5 \times 10^{-23} \) for (CTG)>103 by Fisher's exact test.
In this study, we have analyzed the distribution of the expanded CTG18.1 allele in our large cohort of 574 late-onset FCD cases, 354 controls, and two multigenerational families. Our results support the contributory role of the expanded repeat and a correlation between the size of the repeat (and zygosity) and disease severity. These analyses have also helped predict the outcome of a biallelic expansion of CTG18.1 in individuals who were unaffected previously.

Trinucleotide repeat (TNR) expansions cause multiple human genetic disorders, such as Huntington disease, fragile-X syndrome, and myotonic dystrophy. Noncoding repeat expansions have been linked to myotonic dystrophy type 2 ([CCTG]n in intron 1 of CNBP/ZNF9), Friedreich’s ataxia ([GAA]n in intron 1 of FXN), and spinocerebellar ataxia ([ATTCT]n in intron 9 of ATX10). Tissue-specific expansion of the (CCTG) repeat in ZNF9 in some individuals with myotonic dystrophy type 2 suggests that somatic instability associated with these expansions can occur. Analysis of the CTG repeat expansion in our families does not suggest expansion of the CTG18.1 allele between generations, although we do not exclude the possibility of somatic instability that may play a role during development.

The identification of intronic repeats expanded within ZNP9, FXN, and ATX10 estimates the number of repeats as disease causing between 75 and 11,000 units of CCTG within ZNP9, 70 and 1000 units of GAA within FXN, and 400 and 4500 units of ATTCT within ATX10; the “normal” range of repeats in all three genes is 5 to 30 monomers. Moreover, intergeneration expansion of repeats in myotonic dystrophy results in anticipation. In our analyses, we do not see a large intergeneration expansion of CTG18.1 allele, although a relatively smaller increase in the range of 10 to 30 repeats cannot be determined due to limitations of the STR and TP-PCR assays. Moreover, we have not observed earlier onset of the disease in our pedigrees with successive generations.

The range of tri- and tetranucleotide repeats that confers pathogenicity mentioned above varies between loci. The previously published association of the expansion of CTG18.1 with FCD regarded 40 to 50 repeats as “normal.” Our data from ~600 FCD cases enabled us to identify the threshold of CTG repeat expansion above which the expansion is causal. This notion is further strengthened by the lack of any causal variants identified within the FCD2 linkage or the rs613872 linkage disequilibrium. Moreover, the absence of biallelic expansion for (CTG)n > 40 among any of our sporadic or familial controls suggests that the causal allele might be completely penetrant and essentially mendelize the phenotype. The delineation of the threshold for causality helped us predict the outcome of the specific genotype in two cases of biallelic CTG18.1 expansion. In these cases, both of whom were previously determined to be unaffected but identified with an expansion for (CTG)n > 40 in both alleles, follow-up examination after several years revealed FCD in both. Our analyses have determined the length of the expansion of CTG repeats sufficient to drive the phenotype in 17.8% and significantly associated with 44.4% of the FCD cases in our cohort. Nevertheless, the ability of this model to test the genotype–phenotype correlation needs further refinement, although the

**Figure 2.** CTG repeat expansion leads to increased disease severity with respect to age. Comparison of the age versus severity profiles of all FCD cases stratified by genotypes N/N (A), N/X (B), or the X/X (C), based on the expansion of the CTG18.1 allele at N ≤ 40 < X. The lines within the graphs represent the predicted trend using a linear regression model. The equation in all the graphs represents the linear regression model in the form y = a + bx. The slope of the line is b and a, the intercept. (D) The number of observations (n), Pearson’s coefficient (r), probability (P), and the slope representing change in average disease severity per year are shown for three genotypes as shown in the top row.
there is little doubt on the association signal. Additionally, to better understand the effect on severity, future evaluation with objective techniques such as retroillumination photography analysis, which have been utilized in FCD to determine genotype–phenotype correlation, may provide more detailed resolution of FCD severity relative to a 1 through 6 grading scale, which may be limited in its interrater reliability. 29,30 Moreover, the interrater reliability associated with subjective clinical grading by corneal specialists is modest, and a quantitative and reproducible assessment would be helpful. 31

The TCF4 transcript undergoes multiple alternative splicing and results in the expression of as many as 21 different isoforms known to date. 32 The role of these isoforms has not been completely elucidated, and how many of these alternatively spliced products are indeed translated to a functional protein remains unknown. We hypothesize that the expansion of CTG18.1 allele results in aberrant splicing or alters the translation of the protein. TCF4 has one activation domain near the N-terminus (AD1), a nuclear localization signal, another activation domain (AD2), and a basic helix-loop-helix domain toward the C-terminus. 33,34 The presence of AD1 at the N-terminus of some of the TCF4 isoforms likely imparts an isoform-specific regulatory function. In that context, it will be important to determine whether the interactome of this alternative isoform is perturbed in FCD-affected corneal endothelial tissue. Ultimately, understanding how expansion of this allele impairs specificity of the phenotype will be of interest not only for understanding the pathomechanism of FCD, but also with respect to the broader effect of triplet expansion on transcriptional fidelity.

**FIGURE 3.** Follow-up examination of two previously unaffected individuals with a biallelic CTG18.1 expansion demonstrated FCD. (A, B) The number of (CTG) repeats in all the individuals enrolled in the FCD study from families 1 and 2 was calculated using a combination of STR and TP-PCR assays and is represented below each individual. The shaded black males and females represent individuals with FCD. The individuals in black boxes were followed up for an evaluation based on the hypothesis that a biallelic expansion for (CTG)n>40 is causal. Reexamination with slit-lamp biomicroscopy revealed FCD in these individuals. Genotype represented with “X” suggests that the expansion is larger than the detectable limits of STR and TP-PCR assays.
Acknowledgments

The authors thank all the participants in the study, including the families, for being supportive toward our research.

Supported by grants R01EY016835, K12EY015025, and P30EY01765 from the National Eye Institute, National Institutes of Health, Bethesda, Maryland; the Pakula Family Research Fund; and the Faller Family LLC. The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure: S. Vasanth, None; A.O. Eghrari, None; B.C. Gapsis, None; J. Wang, None; N.F. Haller, None; W.J. Stark, None; N. Katsanis, None; S.A. Riazuddin, None; J.D. Gottsch, None.

References

34. Massari ME, Jennings PA, Murre C. The AD1 transactivation domain of E2A contains a highly conserved helix which is required for its activity in both Saccharomyces cerevisiae and mammalian cells. Mol Cell Biol. 1996;16:121–129.