A Hybrid CFHR3-1 Gene Causes Familial C3 Glomerulopathy

Talat H. Malik,* Peter J. Lavin,†‡ Elena Goicoechea de Jorge,* Katherine A. Vernon,* Kirsten L. Rose,* Mitali P. Patel,* Marcel de Leeuw,§ John J. Neary,¶ Peter J. Conlon,‖ Michelle P. Winn,† and Matthew C. Pickering*

*Centre for Complement and Inflammation Research, Imperial College, London, United Kingdom; †Department of Medicine, Duke University Medical Center, Durham, North Carolina; §Beckman Coulter Genomics, Grenoble, France; ‡Department of Renal Medicine, Royal Infirmary Edinburgh, Edinburgh, United Kingdom; and ‖Department of Nephrology, Beaumont Hospital, Dublin, Ireland

ABSTRACT

Controlled activation of the complement system, a key component of innate immunity, enables destruction of pathogens with minimal damage to host tissue. Complement factor H (CFH), which inhibits complement activation, and five CFH-related proteins (CFHR1–5) compose a family of structurally related molecules. Combined deletion of CFHR3 and CFHR1 is common and confers a protective effect in IgA nephropathy. Here, we report an autosomal dominant complement-mediated GN associated with abnormal increases in copy number across the CFHR3 and CFHR1 loci. In addition to normal copies of these genes, affected individuals carry a unique hybrid CFHR3–1 gene. In addition to identifying an association between these genetic observations and complement-mediated kidney disease, these results provide insight into the protective role of the combined deletion of CFHR3 and CFHR1 in IgA nephropathy.


Polymorphic variation within the CFH-CFHR gene locus influences susceptibility to diverse pathologies including age-related macular degeneration (AMD), meningococcal sepsis, and thrombotic and inflammatory kidney disease such as atypical hemolytic uremic syndrome (aHUS) and C3 glomerulopathy, respectively. Recently, a common copy number polymorphism that results in the deletion of both CFHR3 and CFHR1 genes (ΔCFHR3–1) has been associated with protection against the most common form of GN worldwide, IgA nephropathy, a pathology associated with complement deposition within the kidney. Although CFH1 and CFHR3 have complement regulatory activity in vitro, both have been considered biologically redundant because their complete absence in healthy populations is common. For example, the ΔCFHR3–1 homozygous deletion frequency is 4.7% in European-American and 16% in African-American populations. In fact, ΔCFHR3–1 is actually associated with disease protection. ΔCFHR3–1 is protective in both AMD and IgA nephropathy, pathologies associated with complement deposition in affected tissues. On the basis of a protective role for the ΔCFHR3–1 in IgA nephropathy, we developed the hypothesis that CFHR1 and CFHR3 impair complement processing within the kidney. This hypothesis would predict that an increase in CFHR1 and CFHR3 copy number would enhance susceptibility to complement-mediated kidney injury. Here, we report a novel CFHR3–1 hybrid gene located on an allele that also contained intact copies of the CFHR1 and CFHR3 genes. The hybrid gene was detected exclusively in eight affected individuals among a large kindred with familial complement-mediated GN.

The clinical description of this family has previously been reported and the clinical progress of the affected individuals is described in Table 1. The kindred contains eight individuals with the condition over three generations (Figure 1). Renal biopsies were performed in six affected individuals. In the remaining two affected individuals, disease positivity...
was deduced from abnormal urinalysis (Figure 1). Renal biopsy showed a distinct pattern of glomerular inflammation termed membranoproliferative GN type III (MPGN type III). Evidence of glomerular immunostaining for C3 in the absence of Igs was present. These features indicated Ig-independent complement activation through the complement alternate pathway and are the defining features of C3 glomerulopathy, a term that includes MPGN type III, C3 GN, and dense deposit disease.3 AMD-like lesions may develop in C3 glomerulopathy but ophthalmologic assessment in individuals 105, 212, 213, and 214 did not reveal any evidence of ocular drusen. C3 glomerulopathy has been associated with genetic and acquired defects in complement regulation.3 No evidence of anti-CFH autoantibodies or C3 nephritic factors, acquired causes of C3 glomerulopathy, were detected in the affected individuals and plasma C3 and CFH levels were normal (Table 2). Previous linkage data indicated that the causative defect was within a 22-cM region at chromosome 1q31–32 (maximum logarithm [base 10] of odds score = 3.86), which included the CFH-CFHR locus.13 Using Sanger sequencing, we did not identify any mutations in the CFH or the CFHR genes in the index case (pedigree number 105; Figure 1, data not shown). We next examined copy number variation (CNV) across the CFH-CFHR locus using multiplex ligation-dependent probe assay and TaqMan copy number assays (Supplemental Figure 1, A and B). Combined data from these approaches demonstrated the presence of three copies of two genomic segments within the CFHR locus. These segments included exons 1, 2, and 3 of CFHR3 and exons 3, 4, 5, and 6 of CFHR1 (Figure 2A). Abnormal rearrangement within the CFH-CFHR gene cluster can also impair the surface recognition domains of CFH and predispose individuals to familial aHUS.14,15 Importantly, in affected individuals, there was no evidence of structural mutations affecting the CFH gene, serum CFH levels were normal, and anti-CFH antibodies were not detected.

We therefore hypothesized that a complex rearrangement that resulted in a CFHR3-CFHR1 hybrid gene in addition to the normal diploid copies of both CFHR3 and CFHR1 genes had occurred, and that this was inherited in an autosomal dominant fashion in affected individuals. The region of interest (NCBI37/hg19: Chr1: 196621008–196978800) containing the CFHR locus was successfully targeted using the Agilent SureSelect target enrichment system. Deep sequencing with coverage of >1000× was performed in one affected individual (pedigree number 214) and one related unaffected individual (pedigree number 104). Two partial gene duplication events were identified by comparing sequence coverage discrepancies between individuals; the first starting at least 10 kb upstream of CFHR3’s transcription start site and terminating in intron 3 of CFHR3, the second starting in intron 1 of CFHR1 and terminating at least 13 kb downstream of CFHR1 (Supplemental

Table 1. Updated clinical summaries of affected individuals

<table>
<thead>
<tr>
<th>Pedigree No.</th>
<th>Presentation</th>
<th>End Stage Renal Failure</th>
<th>Renal Transplantation</th>
<th>Ocular Drusen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age (yr)</td>
<td>Proteinuria (g/24 h)</td>
<td>Hematuria</td>
<td>Biopsy Diagnosis</td>
</tr>
<tr>
<td>103</td>
<td>25</td>
<td>&gt; 3</td>
<td>+++</td>
<td>MPGN</td>
</tr>
<tr>
<td>105</td>
<td>51</td>
<td>3.06</td>
<td>++</td>
<td>MPGN</td>
</tr>
<tr>
<td>107</td>
<td>51</td>
<td>+++a</td>
<td>+++</td>
<td>NA</td>
</tr>
<tr>
<td>207</td>
<td>22</td>
<td>+++a</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>212</td>
<td>28</td>
<td>3.77</td>
<td>+</td>
<td>MPGN</td>
</tr>
<tr>
<td>213</td>
<td>4</td>
<td>5.0</td>
<td>Negative</td>
<td>MPGN</td>
</tr>
<tr>
<td>214</td>
<td>21</td>
<td>3.13</td>
<td>+</td>
<td>MPGN</td>
</tr>
<tr>
<td>223</td>
<td>16</td>
<td>0.3</td>
<td>+++</td>
<td>MPGN</td>
</tr>
</tbody>
</table>

MPGN, membranoproliferative GN; NA, not assessed.

aQualitative proteinuria.

Figure 1. Pedigree with familial C3 glomerulopathy. Affected individuals indicate those with C3 glomerulopathy confirmed on renal biopsy. Probable affected indicates those with significant proteinuria (>300 mg/24 h or at least 3+ on urinalysis) or hematuria (at least 3+ on two occasions). Unaffected individuals are healthy and do not have abnormal urinalysis.
Figure 1C). Analysis of the sequencing which compared atypical mapping of paired-end reads between these individuals revealed two genomic breakpoints (NCBI37/hg19:chr1:196754062 at CFHR3 intron 3 and NCBI37/hg19: chr1: 196790079 at CFHR1 intron 1) present only in the affected individual (Supplemental Figures 2 and 3).

To confirm the presence of these genomic breakpoints, primers located within the duplicated regions were used to generate a 6-kb amplicon. As expected, it was possible to generate this amplicon using genomic DNA from affected individuals only. Amplicon sequencing confirmed the genomic breakpoints identified using the paired-end mapping analysis. These data indicated a CFHR3-CFHR1 hybrid gene comprising exons 1, 2, and 3 from CFHR3 and exons 2, 3, 4, 5, and 6 from CFHR1 (Figure 2A). This hybrid gene was present on a CFH-CFHR allele that contained intact copies of the CFHR1 and CFHR3 genes. We next designed a genomic PCR that would amplify this breakpoint (757-bp amplicon) in addition to normal genomic sequence (366-bp amplicon). The abnormal amplicon was present in all affected members of the kindred but not in unaffected individuals (Supplemental Figure 4) or 450 Caucasian control samples (data not shown).

Analysis of the predicted hybrid protein cDNA sequence indicated that splicing from CFHR3 exon 3 to CFHR1 exon 2 would generate a novel protein containing 454 amino acids. The protein would comprise leader peptide sequence and the initial two protein domains, termed short consensus repeat domains, of CFHR3 (derived from CFHR3 exons 1, 2, and 3) and all five short consensus repeat domains of CFHR1 (derived from CFHR1 exons 2, 3, 4, 5, and 6; Figure 2A).

Using a polyclonal anti-human CFH antibody we were able to demonstrate the presence of two aberrant protein bands of similar molecular weight in affected but not unaffected sera (Figure 2B). Two CFHR1 protein bands of similar molecular mass (37-kD CFHR1α and 43-kD CFHR1-β) are seen in sera and we postulated that the two aberrant protein bands represented differentially glycosylated isoforms of the CFHR3–1 hybrid protein (Figure 2B). Using heparin-purified sera from affected individuals, we were able to isolate and analyze the aberrant proteins by peptide mass fingerprinting (Supplemental Figure 5). This analysis identified seven peptides within the predicted hybrid protein sequence (Supplemental Figure 5), including a unique peptide sequence that spanned the junction between CFHR3 protein sequence and CFHR1 protein sequence (VTTFCDFPK) (Supplemental Figure 5).

It is important to note that linkage to the disease locus was reported in this family a decade ago and exon sequencing of the CFH–CFHR locus did not detect the structural variation. Paired-end mapping analysis and copy number assays across the locus successfully identified the hybrid CFHR3–1 gene. Hence, target enrichment followed by paired-end mapping analysis is an efficient method of detecting rearrangements and CNV in linkage peaks and this approach can be applied on a genome-wide level.

The CFHR3–1 hybrid gene is a unique cause of C3 glomerulopathy. The association is robust because we were able to demonstrate complete segregation of disease in a large kindred.

Figure is common among healthy individuals and protective in IgA nephropathy and AMD, suggested that the CFHR3–1 protein causes C3 glomerulopathy through a dominant mechanism. However, further studies are needed to demonstrate causation in this family. In a previous study, we identified a heterozygous internal duplication within CFHR5 in familial C3 glomerulopathy,7 whether this represented a loss of function or dominant negative change remained uncertain because the consequences of complete CFHR5 deficiency, unlike complete deficiency of CFHR1 and CFHR3, were unknown. We speculate that the CFHR3–1 and CFHR5 mutant proteins mediate abnormal accumulation of C3 within the kidney through a common dominant mechanism most likely interfering with the C3 regulation by CFH and CFHR5. It may be the case that normal CFHR3 and CFHR1 proteins also interfere with this process, albeit to a lesser degree; consequently, the absence of these proteins, afforded by the ΔCFHR3–1 polymorphism, results in amelioration of complement-mediated renal injury in IgA nephropathy.

Table 2. Complement profile

<table>
<thead>
<tr>
<th>Pedigree No.</th>
<th>Normal range</th>
<th>Status</th>
<th>C3 0.7–1.7 g/L</th>
<th>C4 0.16–0.54 g/L</th>
<th>CH100 50%–150% NHP</th>
<th>AP100 50%–150% NHP</th>
<th>CFH% NHP</th>
<th>CFI% NHP</th>
<th>C3NeF n/a</th>
<th>Anti-CFH AEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>Affected</td>
<td>0.93</td>
<td>0.39</td>
<td>158</td>
<td>109</td>
<td>102</td>
<td>152</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>105</td>
<td>Affected</td>
<td>0.84</td>
<td>0.29</td>
<td>155</td>
<td>39</td>
<td>77</td>
<td>87</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>212</td>
<td>Affected</td>
<td>1.16</td>
<td>0.54</td>
<td>126</td>
<td>108</td>
<td>119</td>
<td>123</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>213</td>
<td>Affected</td>
<td>1.12</td>
<td>0.40</td>
<td>172</td>
<td>32</td>
<td>84</td>
<td>96</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>214</td>
<td>Affected</td>
<td>0.79</td>
<td>0.26</td>
<td>145</td>
<td>64</td>
<td>86</td>
<td>103</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>215</td>
<td>Unaffected</td>
<td>1.12</td>
<td>0.23</td>
<td>155</td>
<td>74</td>
<td>77</td>
<td>96</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

NHP, normal human pool; C3NeF, C3 nephritic factor; AEU, arbitrary ELISA units; CFI, complement factor I; ND, not detected.

*No validated reference range exists for CFH and CFI and results are thus expressed as the percentage of NHP. C3NeF was detected using a commercial assay (The Binding Site, UK).
Figure 2. The CFHR3–1 hybrid gene structure, genomic breakpoint, and protein. (A) Schematic representation of the CFHR3–1 hybrid gene and protein. Red arrows indicate sites of probes giving three copy signals, whereas blue arrows depict probes giving two copy signals using copy number assays. Exons and protein domains are numbered. The intergenic distance between CFHR3 and CFHR1 is 24 kb. The genomic breakpoint (4959 bp into CFHR3 intron 3) was identified using paired-end mapping analysis of next-generation sequence and long-range genomic PCR and amplicon sequencing. The breakpoint sequence contained a CAG triplet not identified in the introns of either gene (red arrow). (B) Serum Western blot using polyclonal anti-CFH antibody demonstrated the presence of two additional protein bands of similar molecular mass only in sera from affected individuals. The two bands represent differentially glycosylated isoforms (designated α and β) comparable with that seen with CFHR1. CFHL-1, CFH-like protein 1, a protein derived from alternative splicing of the CFH gene that is visible on this gel only in the absence of CFHR1 protein (lane 2).
CONCISE METHODS

Sequencing
Exon sequencing of *CFH* and *CFHR1–5* genes was performed in two affected individuals (numbers 105 and 213). Amplicons that targeted the exons and at least 50 bp of contiguous intronic sequence were amplified by PCR using primers designed using Primer3 software. Sequencing reactions were performed using BigDye Terminator 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Both PCR and sequencing reaction products were purified using Edge Biosystems Gel Filtration plates (Edge Biosystems, Gaithersburg, MD). The sequencing was run on a 3730 DNA Analyzer and analyzed using Sequencer DNA sequencing analysis software (Gene Codes Corp, Ann Arbor, MI).

Multiplex Ligation-Dependent Probe Amplification
Probes used targeted *CFH* exons 1, 2, 3, 4, 6, 11, 12, 13, 18, and 23; *CFHR1* exons 1, 2, 3, 4, and 6; *CFHR1* exons 1, 3, 5, and 6; *CFHR2* exons 1, 2, 3, and 4; and *CFHR5* exons 1, 2, and 3 (MLPA P236-A1; MRC Holland, www.mlp.com). Genomic DNA was processed according to the manufacturer’s instructions. Assay readings were validated using control samples that included individuals with known heterozygous or homozygous polymorphic deletion of the *CFHR1* and *CFHR3* genes. Data were depicted as mean ± SD.

CNV with Real-Time PCR
TaqMan probes (Applied Biosystems) were used to further dissect CNV within the *CFHR3* and *CFHR1* genes. Genomic DNA was processed according to manufacturer’s instructions. Assay readings were normalized to control samples and values represent mean ± SD. All probes were validated using genomic DNA from controls with either heterozygous or homozygous polymorphic deletion of the *CFHR1* and *CFHR3* genes.

In-Solution Target Enrichment of the *CFHR* Locus and Next-Generation Sequencing
Extracted DNA from affected individual 214 and unaffected individual 104 were controlled using fluorescent dosage with picogreen as well as Nanodrop analysis. Illumina libraries were constructed according to the manufacturer’s SureSelect protocol, which involves replacement of one of the Illumina paired-end adapters with the SureSelect capture adapter and capture PCR. Library construction was carried out using the automated SPRI-TE system (Beckman Coulter Genomics). Libraries were controlled on the Agilent Bioanalyzer. In parallel, pairs were designed on the *CFHR* locus defined as chr1:196621008–196978804 using the online Agilent eArray design tool. Default design parameters were used to generate a total of 2179 baits of 120-bp length, covering 150 kb of the 358-kb region of interest, essentially avoiding nonunique locations such as long interspersed nuclear element repeats and long terminal repeats with the default limited 20-bp allowance of bait overlap with nonunique sequence. This represents around 4% of bait space provided for custom bait libraries by the manufacturer. Expected actual extent covered by captured fragments is in general several tens of kilobases higher because the method also captures fragments having only partial overlaps with the 120-bp reads. The design was visually checked on the UCSC genome browser (University of California, Santa Cruz, California). Bait sets were ordered from the manufacturer; upon reception, recommended quantities of baits were mixed with denatured Illumina libraries and hybridization was carried out over 24 hours. After capture, noncaptured DNA was washed away, the hybridized molecules were eluted, and postcapture PCR was performed integrating the Illumina 3’ index adapters so as to tag the libraries with DNA barcodes. Library quality control and dosage was carried out so as to pool an equivalent number of molecules from the two libraries for sequencing in a single lane. Paired-end 2×50-bp sequencing was performed on the Illumina Genome Analyzer IIx, generating 22M read pairs for the affected and 18M read pairs for the unaffected individuals.

Breakpoint Analyses
Sequences were mapped on the human genome by the manufacturer’s Eland program, comprised in the Illumina CASAVA 1.8 software suite. The mapped reads were realigned using the SRMA 0.15 assembler. Resulting sequence coverage was examined using Integrative Genomics Viewer software. Subsequently, atypical pairs were isolated using a BAMtools 0.9 custom filter, specifying insert sizes >1 kb. These atypical pairs were provided as input to SVDetect 0.7. SVDetect was instructed to compare the two libraries so as to isolate locations with atypical read pairs present in the affected sequence set and absent from the unaffected sequence set. Finally, Circos 0.55 was configured to project the output from SVDetect onto a circular *CFHR* locus representation (Supplemental Figure 3A). The output from SVDetect was also uploaded onto the Santa Cruz UCSC genome browser so as to compare putative breakpoint locations against known transcripts (Supplemental Figure 3B).

PCR
We designed a genomic PCR that would generate a single 366-bp amplicon in individuals with at least one copy of the *CFHR3* and *CFHR1* genes using the following primers: forward, 5’-CTCTGTGACTTCTCATCTG3’; and reverse, 5’-AGAAAAACCCATCTCATGTGC3’, which targets GCACATGAGTGTTTTCTT in *CFHR3* intron 3–4. In the presence of the *CFHR3*–1 hybrid gene, an additional 757-bp amplicon is generated using the same forward primer: 5’-CTCTGTGACTTCTCATCTG3’ and the reverse primer 5’-GTTGGGTATGGCCGTA3’, which targets TTGAGGGATAGGCCACCC in the forward strand of *CFHR1* intron 1–2. PCR was performed using Qiagen HotStar Taq DNA polymerase with 10 ng of genomic DNA according to the manufacturer’s protocol. Both reactions were multiplexed in the same reaction. Affected individuals had both a 366-bp and a 757-bp amplicon. In contrast, only the 366-bp control amplicon was seen in all unaffected individuals examined.

Western Blots
Western blots for CFHR proteins were performed using heparin-purified sera from healthy controls and an affected individual (Supplemental Figure 3A). Samples were enriched for CFHR proteins using heparin affinity chromatography and immunoprecipitation. Samples were incubated with heparin-coated sepharose beads (GE Healthcare). After centrifugation and washing, bound proteins were eluted with 500 mM NaCl, 20 mM TRIS-HCL, pH 7.5. Eluted proteins were immunoprecipitated with a mouse monoclonal anti-human CFHR1 antibody (a gift from Dr. Claire Harris, Cardiff University, Wales), which recognizes an epitope common to CFHR1, CFHR2, and CFHR3.
CFHR5, and run on standard 12% SDS PAGE under nonreducing conditions and probed with a goat polyclonal anti-human CFH antibody (Quidel).

**Peptide Mass Fingerprinting**

After electrophoresis, a replicate gel was stained with Coomassie blue and the aberrant bands were manually excised and the gel fragment subjected to peptide mass fingerprinting. Protein samples were reduced and alkylated with iodoacetamide and then digested with trypsin. Peptide mass determination was performed using a Bruker Autoflex III MALDI TOF/TOF instrument by Alphalyse (www.alphalyse.com).

**ACKNOWLEDGMENTS**

The authors thank the patients and their families.

M.C.P. is a Wellcome Trust Senior Fellow in Clinical Science (WT082291MA) and T.H.M., E.G.d.J., and K.L.R. are funded by this fellowship. Additional support was obtained from the National Institute for Health Research Biomedical Research Centre Funding Scheme. K.A.V. is a Kidney Research UK Clinical Fellow (TF8/2009).

**DISCLOSURES**

None.

**REFERENCES**


This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2012020166/-/dcSupplemental.
Supplemental figure 1. **Copy number assays using genomic DNA from affected and unaffected individuals.**

(A) Multiplex ligation-dependent probe assay (MLPA).

Probes used targeted complement factor H (*CFH*) exons 1, 2, 3, 4, 6, 11, 12, 13, 18, 23; complement factor H-related 3 (*CFHR3*) exons 1, 2, 3, 4, 6; complement factor H-related 1 (*CFHR1*) exons 1, 3, 5, 6; complement factor H-related 2 (*CFHR2*) exons 1, 2, 3, 4; and complement factor H-related 5 (*CFHR5*) exons 1, 2 and 3 (MLPA P236-A1, MRC Holland, [www.mlpa.com](http://www.mlpa.com)). Genomic DNA was processed according to manufacturer’s instructions. Assay readings were validated using control samples which included individuals with known heterozygous or homozygous polymorphic deletion of the *CFHR1* and *CFHR3* genes. Data depict mean +/- standard deviation for affected individuals (n=7, red squares, pedigree numbers: 105, 107, 207, 212, 213, 214 and 223) and unaffected individuals (n=11, blue circles, pedigree numbers: 101, 104, 106, 108, 109, 110, 114, 206, 220, 221, 222). In all affected individuals 3 copies of *CFHR3* exons 1, 2 and 3 and *CFHR1* exons 3, 5 and 6 are present. In contrast in all unaffected individuals two copies of these exons are present.

(B) TaqMan® copy number assays

TaqMan® probes ([www.appliedbiosystems.com](http://www.appliedbiosystems.com)) were used to further dissect copy number variation (CNV) within the *CFHR3* and *CFHR1* genes. Genomic DNA was processed according to manufacturer’s instructions. Assay readings were normalised to control samples and values represent mean +/- standard deviation. All probes were validated using genomic DNA from controls with either heterozygous or homozygous polymorphic deletion of the *CFHR1* and *CFHR3* genes (data not shown). The schematic of the *CFHR3* and *CFHR1* genes depicts the TaqMan® probe locations. Exons are indicated by vertical numbered bars. Probes P2, P3, P5 and P6 were custom designed and supplied by Applied Biosystems. Three copies were detected using probes P1, P2 and P5-7 in every affected individual from the pedigree. In contrast, two copies were detected with these probes in all unaffected individuals examined (n=21, pedigree numbers 1, 101, 203, 206, 104, 115, 110, 215, 106, 217, 218, 114, 220, 221, 222, 108, 112, 224, 225, 226 and 109). For clarity we illustrate probe data from three unaffected (pedigree numbers: 101, 104, 222) and three affected (pedigree numbers: 105, 107, 212) individuals.

(C) *CFHR3-CFHR1* hybrid gene: copy number variation as revealed by sequence coverage discrepancies
Coverage discrepancies between individual 214 (IL1142.tdf, affected, top) and individual 104 (IL1160.tdf, unaffected, bottom) suggest an additional copy of two genomic regions. Although bait density and capture efficiency varies importantly over the extent of the region of interest, the resulting sequence coverage is in general very reproducible from sample to sample. The coverage envelope, depicted at linear scale by the Integrative Genome Viewer, suggests two partial gene duplication events, the first starting at least 10kb upstream of the CFHR3 transcription start site and terminating in intron 3 of CFHR3, the second starting in intron 1 of CFHR1 and terminating at least 13kb downstream of CFHR1.
Supplemental figure 2. **Breakpoint analysis using paired-end reads and atypical mapping.**

Breakpoint analysis was performed with SV detect. Individual 214 (affected) is represented in the upper panel and individual 104 (unaffected) is in the lower panel. Atypical mapping occurred when the upstream end of the paired end sequence mapped >1kb away from where it was predicted to map. Paired end reads with atypical mapping are represented by red lines in the lower two panels. The green rectangle shows atypical mapping which only occurs in individual 214 and represents the breakpoint at CFHR3 intron 3 (NCBI37/hg19:chr1:196754062).
Supplemental figure 3. **CFHR3-CFHR1 hybrid gene: fusion between CFHR3 intron 3 and CFHR1 Intron 1.**

(A) This Circos diagram represents the fusion breakpoint. The heavy red line represents 97 paired-end reads with atypical mapping and spanning CFHR3 intron 3 to CFHR1 intron 1. The numerous thin grey lines in the diagram represent paired-end reads with atypical mapping with much lower numbers of atypical reads (range 3-12). This is likely due to the high number of Long Interspersed Elements (LINES) and repetitive elements in this region and have not been confirmed by alternative methods.

(B) Fusion breakpoint stands out on the UCSC browser. Sequences from genomic fragments that were caught by baits which were designed to map CFHR1 actually map to CFHR3 at their 5’ end where no baits were designed. This represents an approximately 35kb deletion.
Affected individuals in the pedigree: 103, 105, 107, 207, 212, 213, 214, 223

Marker lane
2000, 1200, 800, 400 and 200bp

CFHR3/CFHR1 hybrid
control

CFHR3/CFHR1 hybrid
control

CFHR3/CFHR1 hybrid
control

Malik et al, Supplemental figure 4
Supplemental figure 4. Genomic polymerase chain reaction (PCR) identifying the CFHR3/CFHR1 breakpoint within the pedigree.

We designed a genomic PCR which would generate a single 366bp amplicon in individuals with at least one copy of the CFHR3 and CFHR1 genes. In the presence of the CFHR3-1 hybrid gene an additional 757bp amplicon is generated. Primers used: 5’CCTTGTTGACTTTCCATCTCG3’, 5’GGTGCTTATGCCTGCAA3’ and 5’AGGAAACCATCTCATGTGC3’. In all affected members of the pedigree (n=8, pedigree numbers: 103, 105, 107, 207, 212, 213, 214, 223) the 757bp amplicon is present. In contrast in all unaffected members examined only the 366bp control amplicon is seen.
A

B

CFHR3

MLLINVLTLWVSCANGQVKPCDFPDIKHGGLFHENMRRPYFPVAVGKYYSYYCDE
HETPSGSYWDYIHCTQNGWSAPAVPCLRKYCFPYLENGYNQNYGRK
ACHPGYGLPKAQTTVTCTEKGWSTPTPRCIRVYT

CFHR1

TFCDFPKINHGLYDEEKKPFSQVPTGEVFYYSCEYNFVSPSKSFWTRITCTEEGW
STPKCLRLCFFPVENGHSESGQTLEGDTVQICTGVQLGNENNISCVERGW
STPPKCRSTDSCVNPPTVQNAHILSRUMSKYPSGERVERYECRSPYEMFGDEEVM
CLNGWTEPPQCKDSTGKCQPPIIDNGDITSFPAPSSVAPASSVEYQCNLYQLEG
NKRTCRNGQWSEPPKCLHCVISREIMENYIALRWTAQKLYLRTGESAEFVCKR
GYRLSSRSHTLRTCCWDGKLEYPTCAKR

C

<table>
<thead>
<tr>
<th>CODON</th>
<th>ATC</th>
<th>CGT</th>
<th>GTC</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMINOACID</td>
<td>I</td>
<td>R</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CFHR3 exon 3</td>
<td>CA</td>
<td>ACA</td>
<td>TTT</td>
<td>TGT</td>
</tr>
<tr>
<td>CFHR1 exon 2</td>
<td>CA</td>
<td>ACA</td>
<td>TTT</td>
<td>TGT</td>
</tr>
</tbody>
</table>

= T  T  F  C

GA = R in CFHR3 exon 4

Malik et al, Supplemental figure 5
Supplemental figure 5. **Proteomic analysis of the CFHR3-1 hybrid gene product.**

(A) Western blot for complement factor H-related (CFHR) proteins using heparin-purified sera from healthy control and an affected individual. In control sera (lane 2) the two isoforms of CFHR1 are evident. Using sera from an affected individual (lane 1), two additional bands are seen running above CFHR1-β (red arrows). After electrophoresis a replicate gel was stained with Coomassie Blue and the aberrant bands were manually excised and the gel fragment subjected to peptide mass fingerprinting.

(B) and (C) Protein identification by mass spectrophotometry peptide mapping and sequencing analysis.

Seven peptides matched the predicted CFHR3-1 protein sequence (underlined in bold). Importantly, one peptide (VTTFCDFPK) spanned the amino acid junction between CFHR3 and CFHR1. This peptide included the unique amino acid (the initial T in the sequence VTTFCDFPK) generated by the abnormal splicing between CFHR3 exon 3 and CFHR1 exon 2 (C).