Sequence and expression of complement factor H gene cluster variants and their roles in age-related macular degeneration risk


Published in:
Investigative Ophthalmology and Visual Science

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2016, The Authors
This is an open access article published under a Creative Commons Attribution-NonCommercial-NoDerivs License (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Sequence and Expression of Complement Factor H Gene Cluster Variants and Their Roles in Age-Related Macular Degeneration Risk

Anne E. Hughes,1 Stephen Bridgett,2 Weihua Meng,2 Mingyao Li,3 Christine A. Curcio,4 Dwight Stambolian,5 and Declan T. Bradley2

1Formerly of Centre for Public Health, Queen’s University Belfast, Belfast, United Kingdom
2Centre for Public Health, Queen’s University Belfast, Belfast, United Kingdom
3Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, United States
4Department of Ophthalmology, University of Alabama, Birmingham, Alabama, United States
5Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, United States

Correspondence: Anne E. Hughes, 169 Ballylesson Road, Belfast BT8 8JU, UK; annehughes169@btinternet.com.
Submitted: November 30, 2015
Accepted: April 10, 2016

PURPOSE. To investigate how potentially functional genetic variants are coinherited on each of four common complement factor H (CFH) and CFH-related gene haplotypes and to measure expression of these genes in eye and liver tissues.

METHODS. We sequenced the CFH region in four individuals (one homozygote for each of four common CFH region haplotypes) to identify all genetic variants. We studied associations between the haplotypes and AMD phenotypes in 2157 cases and 1150 controls. We examined RNA-seq profiles in macular and peripheral retina and retinal pigment epithelium/choroid/sclera (RCS) from eight eye donors and three liver samples.

RESULTS. The haplotypic coinheritance of potentially functional variants (including missense variants, novel splice sites, and the CFHR3–CFHR1 deletion) was described for the four common haplotypes. Expression of the short and long CFH transcripts differed markedly between the retina and liver. We found no expression of any of the five CFH-related genes in the retina or RCS, in contrast to the liver, which is the main source of the circulating proteins.

CONCLUSIONS. We identified all genetic variants on common CFH region haplotypes and described their coinheritance. Understanding their functional effects will be key to developing and stratifying AMD therapies. The small scale of our expression study prevented us from investigating the relationships between CFH region haplotypes and their expression, and it will take time and collaboration to develop epidemiologic-scale studies. However, the striking difference between systemic and ocular expression of complement regulators shown in this study suggests important implications for the development of intraocular and systemic treatments.

Keywords: complement factor H, genetic variation, gene expression, genetic association, haplotype

Ten years have passed since the association between variation in the human complement factor H (CFH) gene and risk of age-related macular degeneration (AMD) was discovered,1–3 yet this knowledge has not yet been translated into preventive measures for the aging population or treatments for patients with AMD. New therapeutic agents that modify the function of the pathway are in development and in trials, and stratified medicine approaches to the use of existing treatments may soon require ophthalmologists to choose treatments according to individual patients’ inherited genetic variants.

Dysfunction of the complement system (a central part of the innate immune system) is key to the etiology of AMD.4 The complement system comprises a series of proteins, C1 to C9, which are sequentially activated or inhibited by regulatory molecules, including complement factor H (FH), factor D (FD), and factor I (FI). The central component of the process, C3, forms enzymes that accelerate its own activation in a positive feedback loop. When in homeostatic balance, the system facilitates destruction of microbial pathogens, apoptosis, and the removal of cell debris without damage to healthy self-tissues.

Pharmaceuticals currently in clinical trials that inhibit the complement system include lampalizumab (used to inhibit FD, currently in phase III trials for the treatment of “dry” AMD or geographic atrophy in people with specific genotypes), compstatin (an inhibitor of C3, in phase II trials for treatment of AMD), and bikacibomab (an inhibitor of FB that is in phase II trials for treatment of AMD).5 Numerous studies of the impact of patients’ genetics on their response to intravitreal anti-VEGF for AMD have been carried out, yet they have primarily focused on the effect of only one polymorphism, Y402H, even though there are four distinct haplotypes of CFH with different functional variants and risk profiles. Even while potential clinical applications are being investigated there are many unanswered questions about how the complement system behaves in people with different inherited genetic variants and in the different tissues of the eye.
The major inhibitor of the system, FH, is produced in the liver and secreted into the serum, where it is highly abundant. Rare CFH mutations are associated with early-onset AMD. Relatively little is known about how variants influence pathogenesis. Factor H protein is composed of 20 short consensus repeats (SCRs), each approximately 60 amino acids in length, which share homology at specific residues. Full-length FH is encoded by 22 exons, and a shorter form, FHL-1, stops after splicing of an alternative 10th exon. Five homologous CFH-related genes (CFHR1 to CFHR5 encoding FH-R1 to FH-R5) exist in tandem on chromosome 1q in a region of strong linkage disequilibrium (Supplementary Fig. S1). Four common haplotypes of the CFH region are inherited in white Europeans, and each confers a different risk of AMD (and other diseases). A common deletion of entire CFH-related genes (CFHR3-CFHR1) has been found between sites of segmental duplication, and two distinct isoforms of FHR-1, one acidic and one basic, have been reported. Homologous sequences cause significant technical challenges for investigators trying to genotype variants and structural rearrangements.

The research questions that we sought to answer were which of the sequence variations in the CFH region might be functional and therefore may be of interest as prognostic biomarkers, predictive (of therapeutic response) biomarkers, or therapeutic targets for inhibition or synthetic imitation; and in which AMD-relevant human tissues the CFH and CFHR1-5 genes and their splicing variants are expressed (eye tissues for local production and liver for systemic production). In order to answer these questions, we undertook massively parallel genomic sequencing of the CFH region in individuals who were homozygous for each of four common haplotypes. We investigated the risk of AMD-related phenotypes from a genome-wide association study of AMD to illustrate the risk associated with each haplotype. Lastly, we studied RNA expression of CFH and CFHR gene transcripts in central and peripheral retina and retinal pigment epithelium/chorioid/scera (RCS) and in the liver. Our results will help inform the development of a personalized medicine approach to complement system therapeutics.

**Materials and Methods**

**Study Populations**

Institutional Review Board (IRB)/Ethics Committee approval was obtained, and informed consent was obtained from all participants in the original studies. The research adhered to the tenets of the Declaration of Helsinki. DNA for the sequencing experiment was from the anonymized DNA, RNA, and Serum Bank in the Centre for Public Health at Queen’s University Belfast. Patients had neovascular AMD confirmed by clinical examination, grading of digital fundal color photographs, and fluorescein angiography. No details about the clinical course or treatment were available. Research ethics approval for this bank allows analysis of DNA samples for this purpose (Office for Research Ethics Committees Northern Ireland approval reference 11/NI/0139). The in silico data analyzed for this report were from IRB-approved studies. The characteristics of participants included in these studies are summarized in Supplementary Table S1.

**Investigation of CFH and CFH-Related Gene Sequence**

To identify functional variants in the CFH gene region, we selected genomic DNA samples from individuals of white European ancestry who were homozygous for each of the four common CFH haplotypes, as described previously. Full details of the preparatory, sequencing, and bioinformatics methods have been provided (Supplementary Methods). Briefly, we quantified and fragmented genomic DNA, size selected, and indexed it for Illumina sequencing using the Truseq system. We used Nimblegen SeqCap EZ capture to target the entire CFH and CFHR genomic region. Sequencing was performed on a HiSeqation 2000 (Illumina, San Diego, CA, USA) with 100-bp paired-end reads. We aligned genomic reads to hg19 human reference sequence with the Burroughs-Wheeler Aligner (BWA) 0.5.9 aln algorithm and used SAMtools 0.1.14 for sorting, indexing, and removal of duplicate reads. We employed Genome Analysis Toolkit (GATK) to recalibrate and realign and to call polymorphisms. We predicted the effect of coding polymorphisms with PolyPhen-2 (Polymorphism Phenotyping v2).

**Investigation of the Effect of CFH Region Haplotypes on AMD Phenotypes**

To investigate the effect of polymorphisms on the differential risk on effect of neovascular AMD compared to drusen only (as a proxy for progression), we conducted a genome-wide case–control study and candidate gene studies using the Michigan, Mayo, AREDS, Pennsylvania (MMAP; dbGAP accession code 10000169; v1.p1) AMD study consisting of 2157 cases of neovascular AMD, geographic atrophy, or drusen and 1150 healthy controls of white European ancestry (Supplementary Table S1). The candidate gene studies aimed to investigate the roles of the four CFH haplotypes and a representative polymorphism from each of the other known major AMD loci (CFB, C3, and HTRA1) on the differential risk of neovascular AMD compared to drusen. The risks of drusen and neovascular AMD compared to disease-free controls are shown to illustrate the effects associated with each haplotype.

Full details of quality control and analyses are shown in the Supplementary Methods. Briefly, genotyping was conducted on Illumina Human570 microarray chips. Final analyses included 867 individuals with neovascular AMD, 519 with drusen, and 1115 healthy controls. A Q-Q plot is shown in Supplementary Figure S2. We used additive model univariate binary logistic regression in PLINK v1.07. Statistical significance was accepted at two-sided P < 0.05. After Bonferroni correction for multiple testing in the genome-wide association studies. We did not apply correction to the candidate polymorphisms because their selection was based on prior hypotheses due to their known association with the AMD phenotype. We phased haplotypes of CFH in PLINK and used SNP Annotation and Proxy Search (Broad Institute) to identify proxy single nucleotide polymorphisms (SNPs).
Sequence and Expression of CFH Gene Variants

Investigation of FH, FHL-1, and FHR-1 to FHR-5 Expression in Liver

To investigate the relative expression of FH, FHL-1, and FHR-1 to FHR-5 in liver, we accessed liver RNA-seq reads from three individuals (no details of ancestry were available) from the EBI Illumina body map and EBI Expression Atlas and aligned them to the human reference hg19 chromosome 1 using STAR aligner (version 2.4.0f1). This method was chosen as it allowed calculation of an appropriately weighted average proportion (taking into account the total number of reads from each individual) with an overall estimate of uncertainty (confidence interval [CI]) for each tissue.

Investigation of Risk of AMD Phenotypes

The risks of drusen and neovascular AMD associated with each CFH haplotype were calculated to illustrate the effects of each CFH haplotype (Table 2). We compared genome-wide variation in individuals with neovascular AMD to those with drusen and found that two SNPs were significantly associated with differential risk (a proxy for a “progression” phenotype) after Bonferroni correction for multiple testing: rs932275 and rs2248799, both in HTRA1 (Supplementary Table S5). We identified several SNPs with a suggestive P value < 5×10⁻², many of which were in biologically plausible genes (Supplementary Table S3). We also compared individuals with neovascular AMD to those with drusen for previously reported major AMD loci (Supplementary Table S4), but the only association identified from among these candidates was the same one that was identified at the HTRA1 locus in the genome-wide analysis. The complement system gene variants were not significantly different between people with drusen and people with neovascular AMD in this cohort.

Retinal Expression of CFH and CFH-Related Genes

We found no expression of any CFH-related genes (CFHR1-5) in any of the eye tissues studied. Expression of both long and short CFH transcripts was 30- to 50-fold higher in the RCS than in peripheral or macular retina (Table 3). Within the retina, the full-length FH transcripts were considerably more abundant than FHL-1 transcripts in the macula compared to the periphery (Table 3). The RCS showed increased ratio of FH expression in peripheral or macular retina (Table 3). Within the retina, the full-length FH transcripts were considerably more abundant than FHL-1 transcripts in the macula compared to the periphery (Table 3). The RCS showed increased ratio of FH to FHL-1 transcripts in the periphery compared to the macula in seven of eight eyes (Supplementary Table S5).

Expression of CFH and CFH-Related Genes in Liver and Confirmation of the Effect of rs4085749 T on CFHR2 mRNA

Relative to total CFH transcription in the liver, CFHR1 was expressed approximately 2-fold more strongly, CFHR2 at a
reduce expression of \( CFH \) gene variants, the odds ratio for AMD progression, odds ratio for NV AMD vs. drusen, and odds ratio for drusen vs. normal-eye controls (95% CI) are shown in Table 2.

**Table 2.** \( CFH \) Region Haplotype Frequencies and Associated Risk of Neovascular AMD and Drusen (Additive Model Univariate Logistic Regression With Each Haplotype Compared to All Others); the Reported Risk of Invasive Meningococcal Disease and Atypical Hemolytic Uremic Syndrome Shown to Provide Context

<table>
<thead>
<tr>
<th>Gene Variants</th>
<th>Haplotype A</th>
<th>Haplotype B</th>
<th>Haplotype C</th>
<th>Haplotype D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency in normal-eye controls, ( n = 1115 )</td>
<td>0.36</td>
<td>0.17</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>Frequency in individuals with drusen, ( n = 519 )</td>
<td>0.59</td>
<td>0.15</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>Frequency in individuals with NV AMD, ( n = 867 )</td>
<td>0.60</td>
<td>0.15</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>Odds ratio for drusen vs. normal-eye controls (95% CI)</td>
<td>2.59 (2.21–3.04)</td>
<td>0.89 (0.73–1.08)</td>
<td>0.44 (0.35–0.55)</td>
<td>0.47 (0.37–0.59)</td>
</tr>
<tr>
<td>Odds ratio for NV AMD vs. normal-eye controls (95% CI)</td>
<td>2.75 (2.39–3.16)</td>
<td>0.90 (0.75–1.07)</td>
<td>0.41 (0.34–0.49)</td>
<td>0.41 (0.34–0.50)</td>
</tr>
<tr>
<td>Odds ratio for AMD progression, NV AMD vs. drusen (95% CI)</td>
<td>1.05 (0.89–1.24)</td>
<td>1.02 (0.82–1.26)</td>
<td>0.95 (0.74–1.21)</td>
<td>0.88 (0.68–1.14)</td>
</tr>
<tr>
<td>Meningococcal disease risk(^{13})</td>
<td>Neutral</td>
<td>Reduced</td>
<td>Neutral</td>
<td>Increased in homozygosity</td>
</tr>
<tr>
<td>Atypical hemolytic uremic syndrome risk(^{8,15})</td>
<td>-</td>
<td>-</td>
<td>Increased in homozygosity</td>
<td>Neutral</td>
</tr>
</tbody>
</table>

*NV*, neovascular.

**DISCUSSION**

The central role of factor H in innate immunity and the effects of its genetic variation on risk of AMD, atypical hemolytic uremic syndrome (aHUS),\(^8\) systemic lupus erythematosus,\(^36\) rheumatoid arthritis,\(^39\) and invasive meningococcal infection\(^9\) make it intriguing, and it has yielded many surprises. Our study is the first to investigate the extended \( CFH \) region in individuals who are homozygous for the common European haplotypes. The results of this study should provide a clear foundation for future studies of FH function and aid the development of personalized approaches based on patients’ own complement system genetics.

Haplotype A confers greatest risk of AMD and carries the \( CFHR3 \) region were found on haplotype B, \( CFH \) 936D and \( CFHR3 \) 241S, two variants in the 3’ untranslated region (UTR) of \( CFHR3 \) (rs402372 and rs390837), and \( CFHR2 \) 72Y are carried on haplotype B. This haplotype also produces the acidic isoform of FH-1.

Haplotype C protects against AMD.\(^12\) \( CFHR1 \) 157Y, 159V, and 175Q, which have arisen by conversion from \( CFH \), code for the basic isoform of FHR-1 and FH. The central role of \( CFHR1 \) in innate immunity and the effects of its genetic variation on risk of AMD, atypical hemolytic uremic syndrome (aHUS),\(^8\) systemic lupus erythematosus,\(^38\) rheumatoid arthritis,\(^38\) and invasive meningococcal infection\(^9\)

Although \( CFH \) was expressed in the retina and in the RCS of normal aged eyes, we found no evidence of transcription of any of the \( CFH \)-related genes (\( CFHR1 \)–5) in these tissues. This finding contrasts with reports by Bennis et al.\(^{45}\) and Booij et al.\(^{46,47}\) of a microarray experiment that measured transcriptome expression of six human eyes, which reported \( CFHR1 \) transcription at approximately one-fourth the level of \( CFH \) (and lower levels of \( CFHR4 \), \( CFHR5 \), and \( CFHR2 \), but not \( CFHR3 \)). They found no \( CFHR1 \) expression in mouse eye tissues, which is consistent with an earlier report by Luo et al.\(^{48}\) in which RT-PCR indicated no expression of \( CFHR1 \) in mouse retina or RPE/choroidal tissue. The conflicting results indicate either that our method did not detect \( CFHR \) mRNA where it was present, or that the Bennis et al.\(^{45}\) and Booij et al.\(^{46,47}\) experiments falsely detected \( CFHR \) mRNA. If \( CFHR \) genes are expressed in these tissues, the most likely explanation for their apparent absence in our experiment would be differences in the tissue processing between the two studies. Alternatively, the Bennis et al.\(^{45}\) and Booij et al.\(^{46,47}\) custom microarrays may have
suffered cross-reactivity from FH and FHL-1 transcripts (Supplementary Fig. S1), leading to the apparent presence of CFHR genes. Absent expression would not necessarily indicate that CFHR genes have no role in these tissues, as it is likely that they could diffuse across Bruch’s membrane, as FHL-1 does (though FH does not).\textsuperscript{49} Further studies are needed to assess CFHR expression (if any) in AMD eyes and the effect of inflammation. CFH transcripts for FH and FHL-1 showed relatively low expression in the macular and peripheral retina and much greater expression in the RSC, suggesting reduced protection in the retina and greater protection in at least one of the RCS tissues. Our results suggested a greater proportion of FH compared to FHL-1 in the macular retina compared to the peripheral retina. This may be due to regulation by microRNAs. Both transcripts appear to be heavily regulated by microRNAs; however, target sites for mir146a and mir155 are adjacent on the 3′ UTR of CFH and absent from the 3′ UTR of the short (FHL-1) transcript. These are among the microRNAs with greatest influence on immune signalling pathways and immune homeostasis.\textsuperscript{50}

In contrast to the eye, we found that all CFH-related transcripts were present in liver RNA. Both CFH transcripts, CFHR1, CFHR2, and CFHR3 were much more highly expressed than CFHR4 or CFHR5. Our analysis of liver RNA-seq data suggested that haplotype C may be associated with either considerable reduction in FH or increase in FHL-1 mRNA transcript levels compared to the other haplotypes. Unexpectedly, there seems to be poor correlation between levels of transcripts in the liver and the reported concentrations of their translated products in the blood plasma. FHL-1 is reported at only 2% to 8% of FH levels in plasma, despite having transcript levels in the liver greater than that of CFH mRNA.\textsuperscript{6} It may be important to understand the stability of FH, FHL-1, and their degradation products. Measurement of FH and related proteins is difficult because of the potential for future development of personalized treatments based on these mechanisms.

**Acknowledgments**

We thank the participants in the National Eye Institute Study of Age-Related Macular Degeneration (NEI-AMD) and the NEI-AMD Research Group for their valuable contribution to this research. The AMD dataset used for the analyses described in this paper was obtained from the NEI-AMD database, through dbGaP accession no. phs000182.v1.p1. Funding support for NEI-AMD was provided by the National Eye Institute.

Supported by Guide Dogs for the Blind Association Grant OR2010-03c (AEH), National Eye Institute Grant R01EY023164, and the Arnold and Mabel Beckman Initiative for Macular Research (CAC and DS).

Disclosure: A.E. Hughes, None; S. Bridgett, None; W. Meng, None; M. Li, None; C.A. Curcio, Genentech (C), Novartis (C), Merck (C), Janssen Cell Therapy (C), IOVS (S); D. Stambolian, None; D.T. Bradley, None

**References**


26. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses.


30. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads.


34. E-MTAB-513 - RNA-Seq of human individual tissues and mixture of 16 tissues (Illumina Body Map).

35. Zhao J, Wu H, Khosravi M, et al. Association of genetic variants of LOC387715 near TIMP3 and high-density lipoprotein-associated loci with increased susceptibility to age-related macular degeneration.


40. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses.

41. Tortajada A, Montes T, Martinez-Barricarte R, Morgan BP, Harris CL, de Cordoba SR. The disease-protective complement factor (Y402H) and FHL-1 in age-related macular degeneration.


