

## Complement Factor H Polymorphism and Age-Related Macular Degeneration

Albert O. Edwards<sup>1\*</sup>, Robert Ritter, III<sup>1</sup>, Kenneth J. Abel<sup>2</sup>, Alisa Manning<sup>3</sup>, Carolien Panhuysen<sup>3,6</sup>, Lindsay A. Farrer<sup>3-7</sup>

<sup>1</sup>Department of Ophthalmology and McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA. <sup>2</sup>Sequenom, Inc, 3595 John Hopkins Court, San Diego, CA 92121, USA. Departments of <sup>3</sup>Medicine (Genetics Program), <sup>4</sup>Neurology, and <sup>5</sup>Genetics & Genomics, Boston University School of Medicine and Departments of <sup>6</sup>Biostatistics and <sup>7</sup>Epidemiology, Boston University School of Public Health, 715 Albany Street, Boston, MA 02118, USA.

\*To whom correspondence should be addressed. Email: albert-edwards@swbell.net

†Current address: Institute for Retina Research, 3215 Princess Lane, Dallas, TX 75229, USA.

**Age-related macular degeneration (AMD) is a common, late-onset, and complex trait with multiple risk factors. Concentrating on a region harboring a locus for AMD on 1q25-31, the *ARMD1* locus, we tested single nucleotide polymorphisms for association with AMD in two independent case-control populations. Significant association ( $P = 4.95 \times 10^{-10}$ ) was identified within the regulation of complement activation locus and centered over a Tyr402His protein polymorphism in the gene encoding complement factor H. Possession of at least one histidine at amino acid 402 increased the risk of AMD 2.7-fold and accounts for 50% of the attributable risk of AMD.**

Age-related macular degeneration (AMD) is a leading cause of blindness in older individuals (1). It is a late-onset, complex trait with hereditary, lifestyle, and medical risk factors (2). The condition typically presents in the fifth decade of life with small yellow deposits external to the outer retina and retinal pigment epithelium (RPE) called drusen. Large numbers of drusen and clinical features of damage to the RPE markedly increase the risk of complications (atrophy of the RPE and abnormal neovascularization of the outer retina), leading to severe vision loss (1).

Although the primary pathogenic mechanisms of AMD were previously unknown, there is strong evidence that genetics plays a role (3–9). The first locus for AMD (*ARMD1*) was reported in a single extended family linked to chromosome 1q25.3-31.3 (5). Because there was strong evidence for linkage to this region of chromosome 1 (fig. S1) from subsequently reported small family studies, we focused our efforts on the *ARMD1* locus (3, 4, 6, 8, 9).

We performed an allele association study on a new case-control population highly discordant for clinical phenotypes. Cases were enrolled based on ocular features (extensive drusen or pigmentary abnormalities of the macula) placing

subjects at high risk for development of the complications of AMD or the presence of those complications in one or both eyes (10). Control subjects were from the same patient population and could have no more than four small hard drusen in the central retina (macula) and no known family history of AMD. A subset of 224 cases and 134 controls meeting these criteria were selected as a discovery sample for initial genotyping (table S1). The discovery sample was enriched for AMD cases showing familial clustering of AMD and high-risk, early AMD. A second, replication sample of 176 cases and 68 controls was ascertained at the same clinic following the same protocol (table S1).

Evaluation of the reported Q5346R variation in the Fibulin 6 gene (*FIBL6*) and 23 SNPs across this gene (supporting online text) gave no evidence for allele or haplotype association in the discovery sample (11). To determine if common coding sequence variation within the *ARMD1* locus was associated with AMD, we searched publicly available databases for non-synonymous coding SNPs (nscSNPs). We identified 24 nscSNPs with known minor allele frequency of at least 10% across the 14 Mb *ARMD1* locus. Genotyping of the discovery sample gave significant allele and genotype association between AMD and nscSNPs only within the regulation of complement activation (RCA) locus in chromosome 1q31.3 (table S2). Additional evenly spaced, gene-based SNPs were evaluated across all 31 genes in clusters 1, 3, 4, and 6 at 8 kb to 25 kb density (fig. S1) and no associations with AMD were detected outside of the RCA locus (table S3). These data suggest the RCA locus contains one or more genetic variants increasing the risk of developing AMD.

The RCA locus spans 388 kb of genomic DNA that contains the gene encoding complement factor H (*CFH*), five genes derived from *CFH* through ancestral duplications, and the gene encoding factor 13B (Fig. 1). A total of 86 SNPs located across the RCA locus and flanking regions were

genotyped. Twenty-nine gave evidence for allele association with the majority and most significant of these, including the nscSNP rs1061170, concentrated in the *CFH* gene (table S3, GenBank accession NM\_000186).

The genotype frequency data for rs1061170 revealed that the association with AMD was largely due to an excess of CC homozygotes in cases compared to controls (Table 1). A similar pattern of association was evident with seven adjacent SNPs in *CFH* (Fig. 1 and table S3). The strength of the evidence for association diminished markedly with SNPs located immediately proximal to *CFH* and distal to its derivatives *CFHL1-CFHL5*, suggesting that the effect was due to one or more polymorphisms in the complement factor genes only (Fig. 1 and table S3).

Haplotype analyses of 34 SNPs spanning 418 kb revealed extensive linkage disequilibrium across the full length of the RCA locus (Fig. 2). The highest level of linkage disequilibrium discernable among 4 haplotype blocks was across the RCA locus. Thirteen contiguous SNPs in *CFH* (i.e., all but the first two *CFH* SNPs in Fig. 2) form a 64 kb haplotype block. A second 9 kb haplotype block contained SNPs in the proximal portion of *CFHL4*, a third 50 kb block contained SNPs in the distal portion of *CFHL4* and SNPs in *CFHL2*, and a fourth 146 kb block contained SNPs in *F13B*, *ASPM* and *FRBZI* (Fig. 2). The SNPs most significantly associated with AMD were in *CFH* or within 221 kb downstream of *CFH* (Fig. 1 and table S3).

The association between AMD and haplotypes comprising 2-5 contiguous SNPs was evaluated using the Haplo.stat software by a sliding window approach (12, 13). Additional comparisons were made using all possible haplotypes formed by pair wise combinations including at least one nscSNP within the RCA locus. The analysis did not reveal any SNP combination showing greater association with AMD than the individual SNPs. All of the AMD high-risk haplotypes including a *CFH* SNP from haplotype block 1 and a non-*CFH* SNP from other regions of the RCA locus (Fig. 1) contained the AMD risk allele from the *CFH* SNP but not necessarily the AMD risk allele from the non-*CFH* SNP. These analyses provide further evidence that the multiple signals in the RCA locus are related to a single haplotype and therefore likely caused by a single genetic effect.

To verify these findings, we genotyped 14 SNPs in the RCA locus in the replication sample. Association with AMD was observed with the 7 markers that were significant in the discovery sample, including rs1061170, but not with the 7 markers yielding negative results in the discovery sample (table S4). Notably, the genotype frequencies for rs1061170 among cases enriched for a positive family history of AMD were nearly identical to the frequencies among cases without this characteristic (Table 1), suggesting that this *CFH* protein polymorphism is a risk factor for AMD more generally.

Taking into account data from the entire sample, a conservative estimation of the relative risk for AMD conferred by having at least one C allele (i.e., having either the CC or CT genotypes) was 2.7 (95% CI = 1.9–3.9).

Complement activation has been implicated in the pathogenesis of a number of complex traits including AMD and can arise through the classical, lectin, or alternative pathways (14). All three pathways lead to the generation of a C3 convertase enzyme and subsequent activation of the immune response, the terminal pathway pore-like membrane attack complex (C5b-9), and cell lysis. The alternative complement pathway is spontaneously activated and *CFH* is an essential inhibitor preventing uncontrolled complement activation (15). Components of the terminal complement pathway and other markers of inflammation are deposited in drusen and the choroid of eyes with AMD (16, 17). Abnormal regulation of the alternative pathway of complement activation by *CFH* is consistent with these observations.

The tyrosine to histidine polymorphism (rs1061170) at amino acid 402 of *CFH* may be a primary pathogenic variation increasing the risk of developing AMD. *CFH* is composed of 20 repetitive units of 60 amino acids called short consensus repeats (SCRs). The Tyr402His polymorphism is located within SCR7, which contains the overlapping binding sites for heparin, C-reactive protein (CRP), and M-protein (18). Serum levels of CRP were elevated in AMD subjects compared to controls in one large prospective clinical trial (19). CRP activates the classic complement pathway, but reduces deposition of C5b-9 through the direct binding of *CFH* (20). Risk factors for development of complications of AMD including cigarette smoking, lack of exercise, hypertension, and obesity (2, 21) increased serum CRP levels or decreased serum *CFH* levels (22–25). Further, drusen with terminal complement deposition indistinguishable from AMD were observed in eyes from patients with a kidney disease (membranoproliferative glomerulonephritis type II) that can be caused by mutations in *CFH* (26, 27). In principle, altered binding of *CFH* to CRP or heparin on outer retinal surfaces caused by the Tyr402His substitution could affect the level of inflammation in the outer retina, thereby contributing to AMD. Although our results are consistent with the Tyr402His variant causing AMD, they do not rule out the existence of other coding or splice site variants within *CFH* that modulate risk of AMD.

More than 7 million individuals in the United States have retinal features placing them at high risk for developing vision loss from complications of AMD (28). The attributable fraction for the C allele derived from the total sample of subjects in this study is 50%, suggesting that persons either homozygous or heterozygous for histidine at amino acid 402 of *CFH* account for one-half of AMD cases. Given the rapid aging of the population, an estimated 3 million individuals

will have atrophic and exudative complications of AMD by 2020 (28). Our findings suggest new avenues for developing preventative and therapeutic strategies for AMD.

## References and Notes

1. R. Klein, B. E. Klein, S. C. Tomany, S. M. Meuer, G. H. Huang, *Ophthalmology* **109**, 1767 (2002).
2. R. Klein, T. Peto, A. Bird, M. R. Vannewkirk, *Am. J. Ophthalmol.* **137**, 486 (2004).
3. G. R. Abecasis *et al.*, *Am. J. Hum. Genet.* **74**, 482 (2004).
4. S. K. Iyengar *et al.*, *Am. J. Hum. Genet.* **74**, 20 (2004).
5. M. L. Klein *et al.*, *Arch. Ophthalmol.* **116**, 1082 (1998).
6. J. Majewski *et al.*, *Am. J. Hum. Genet.* **73**, 540 (2003).
7. J. H. Schick *et al.*, *Am. J. Hum. Genet.* **72**, 1412 (2003).
8. J. M. Seddon, S. L. Santangelo, K. Book, S. Chong, J. Cote, *Am. J. Hum. Genet.* **73**, 780 (2003).
9. D. E. Weeks *et al.*, *Am. J. Hum. Genet.* **75**, 174 (2004).
10. Materials and methods are available as supporting material on Science Online.
11. D. W. Schultz *et al.*, *Hum. Mol. Genet.* **12**, 3315 (2003).
12. D. J. Schaid, C. M. Rowland, D. E. Tines, R. M. Jacobson, G. A. Poland, *Am. J. Hum. Genet.* **70**, 425 (2002).
13. H. Zhao, R. Pfeiffer, M. H. Gail, *Pharmacogenomics* **4**, 171 (2003).
14. J. Acosta, X. Qin, J. Halperin, *Curr. Pharm. Des.* **10**, 203 (2004).
15. H. J. Muller-Eberhard, R. D. Schreiber, *Adv. Immunol.* **29**, 1 (1980).
16. R. F. Mullins, S. R. Russell, D. H. Anderson, G. S. Hageman, *FASEB J.* **14**, 835 (2000).
17. L. V. Johnson, W. P. Leitner, M. K. Staples, D. H. Anderson, *Exp. Eye. Res.* **73**, 887 (2001).
18. E. Giannakis *et al.*, *Eur. J. Immunol.* **33**, 962 (2003).
19. J. M. Seddon, G. Gensler, R. C. Milton, M. L. Klein, N. Rifai, *JAMA* **291**, 704 (2004).
20. C. Mold, H. Gewurz, T. W. Du Clos, *Immunopharmacology* **42**, 23 (1999).
21. J. M. Seddon, J. Cote, N. Davis, B. Rosner, *Arch. Ophthalmol.* **121**, 785 (2003).
22. J. Esparza-Gordillo *et al.*, *Immunogenetics* **56**, 77 (2004).
23. J. R. Greenfield *et al.*, *Circulation* **109**, 3022 (2004).
24. J. M. Backes, P. A. Howard, P. M. Moriarty, *Ann. Pharmacother.* **38**, 110 (2004).
25. M. H. Wener, P. R. Daum, G. M. McQuillan, *J. Rheumatol.* **27**, 2351 (2000).
26. R. F. Mullins, N. Aptsiauri, G. S. Hageman, *Eye* **15**, 390 (2001).
27. M. A. Dragon-Durey *et al.*, *J. Am. Soc. Nephrol.* **15**, 787 (2004).
28. D. S. Friedman *et al.*, *Arch. Ophthalmol.* **122**, 564 (2004).
29. J. C. Barrett, B. Fry, J. Maller, M. J. Daly, *Bioinformatics* **21**, 263 (2005).

30. We thank the McDermott Sequencing Center at UTSW for genotyping assistance and T. Hyatt for technical advice and assistance. Supported by the National Eye Institute (EY014467).

## Supporting Online Material

[www.sciencemage.org/cgi/content/full/1110189/DC1](http://www.sciencemage.org/cgi/content/full/1110189/DC1)

Materials and Methods

Fig. S1

Tables S1 to S4

References

25 January 2005; accepted 22 February 2005

Published online 10 March 2005; 10.1126/science.1110189

Include this information when citing this paper.

**Fig. 1.** The regulation of complement activation (RCA) locus located within chromosome 1q31.3 includes the gene for complement factor H (*CFH*), 5 related genes derived from *CFH* through ancestral duplications, and the gene for factor 13B (*F13B*). A megabase (Mb) scale of this region is provided at the top of the figure. SNPs genotyped across the RCA locus are shown along the bottom of the figure. The negative natural logarithm of the significance of allele association to AMD for each SNP is given in the graph (10). The 0.05 significance level is shown by the dotted line. Values greater than 15 on the y-axis correspond to *P*-values less than  $10^{-7}$ .

**Fig. 2.** Pairwise linkage disequilibrium (LD) among 38 SNPs in and surrounding the RCA locus, spanning the region including the 5' upstream region of *CFH* (represented by the SNP rs3753394) and the proximal portion of *FRBZ1* (represented by the SNP rs10732295), a distance of 516 kb. LD was measured by the  $D'$  statistic using data from all subjects in the discovery sample (10). A  $D'$  value of 1 indicates complete linkage disequilibrium between two markers and a  $D'$  value of 0 indicates complete linkage equilibrium. Using the Haploview version 3.0 software, we could divide the RCA locus into four haplotype blocks (29).  $D'$  values for all, or nearly all, pairwise comparisons within a haplotype block are at least 0.8 (black squares). The haplotype block structure is similar to that obtained by the HapMap Project for this region in Caucasians with the notable difference that blocks 2 and 3 in our structure comprise one block in the HapMap report (10). The lone SNP (rs379370) between blocks 2 and 3 is a nscSNP in the *CFHL4* gene (table S2) showing moderate LD (gray squares), modest LD (light gray squares), or low LD (white squares) with the SNPs in block 3.

**Table 1.** Association between the Try402His polymorphism (rs1061170) in *CFH* and AMD. The C allele codes for histidine. The genotype association compares CC to CT and TT.

| Sample      | Allele distribution |          | Allele Association<br>( <i>P</i> -value) | Genotype distribution |          | Genotype Association<br>( <i>P</i> -value) |
|-------------|---------------------|----------|--|-----------------------|----------|--|
|             | Cases               | Controls |  | Cases                 | Controls |  |
| Discovery   | C                   | 0.553    | 3.68 x 10 <sup>-8</sup>                  | CC                    | 0.320    | 7.67 x 10 <sup>-7</sup>                    |
|             | T                   | 0.447    |  | TC                    | 0.467    |  |
|             |                     |          |  | TT                    | 0.213    |  |
| Replication | C                   | 0.544    | 0.0039                                   | CC                    | 0.306    | 0.0135                                     |
|             | T                   | 0.456    |  | TC                    | 0.477    |  |
|             |                     |          |  | TT                    | 0.218    |  |
| Total       | C                   | 0.549    | 4.95 x 10 <sup>-10</sup>                 | CC                    | 0.314    | 1.4 x 10 <sup>-8</sup>                     |
|             | T                   | 0.451    |  | TC                    | 0.471    |  |
|             |                     |          |  | TT                    | 0.215    |  |

## 1q31.3



