Complement Component 3: an assessment of association with AMD and analysis of gene-gene and gene-environment interactions in a Northern Irish cohort

Gareth J. McKay,1 Shilpa Dasari,1 Christopher C. Patterson,2 Usha Chakravarthy,1 Giuliana Silvestri1

1Centre for Vision and Vascular Sciences, Queen's University of Belfast, Belfast, Northern Ireland, UK; 2Centre for Public Health, Queen's University of Belfast, Belfast, Northern Ireland, UK

Purpose: A non-synonymous single nucleotide polymorphism (SNP) in complement component 3 has been shown to increase the risk of age-related macular degeneration (AMD). We assess its effect on AMD risk in a Northern Irish sample, test for gene–gene and gene–environment interaction, and review a risk prediction model.

Methods: SNP rs2230199 was genotyped in 1,358 samples, which comprised 437 cases, 436 no-disease controls, and 485 participants randomly sampled from the Northern Ireland population. Allele frequencies were assessed in cases and controls. Logistic regression analysis was used to assess interaction and develop a risk prediction model.

Results: We report a minor allele frequency of 0.248 for rs2230199 in the population (n=485), 0.296 in cases (n=437), and 0.221 in controls (n=436; odds ratio [OR]=1.48; confidence interval [CI]: 1.19–1.85; p=0.0003). The significant association is retained following multivariate analysis with adjustment for age, smoking status, Complement Factor H (CFH), Age-Related Maculopathy Susceptibility 2 (ARMS2), Complement Component 2 (CC2), and Complement Factor B (CFB; OR=1.45; CI: 1.10–1.91; p=0.009). No evidence to support an interaction between any of the covariates within the regression model was found. The area under the receiver operator characteristic curve calculated for the fully adjusted model, including all variables, was 0.86 for late AMD.

Conclusions: Our study confirmed the association between Complement Component 3 (C3) and late-stage AMD. There was no evidence for an interaction with environmental exposures, nor did we find data to support a gene–gene effect.

Age-related macular degeneration (AMD, MIM 603075) is a degenerative retinal disease that causes progressive impairment of the central vision and is the leading cause of irreversible severe vision loss in Caucasians over the age of 50 [1]. AMD begins with drusen formation and, in some cases, degenerative changes at the level of the retinal pigment epithelium (RPE), often referred to as early AMD, that primarily affects the macula and can progress to the visually disabling late phenotypes of geographic atrophy (GA) and/or neovascular AMD. Susceptibility to AMD is complex, involving genetic, lifestyle, and environmental factors, although the specifics of the etiology remain, in part, unresolved. Advances in our understanding of the genetic basis of the disease etiology have identified risk and protective haplotypes in several genes associated with the complement pathway and chronic inflammation, such as Complement Factor H (CFH) [2-6], Complement Component 2 (CC2)/Complement Factor B (CFB) [7,8], and Complement Component 3 (C3) [9-11]. A dysfunctional complement pathway has been proposed to accentuate retinal cell damage with increased formation of drusen deposits, atrophy, and cell degeneration [12] and progression to choroidal neovascularization [5]. A non-synonymous single nucleotide polymorphism (SNP), rs2230199 (R102G), which causes an amino acid change from arginine to glycine in C3, has been most commonly associated with late-stage AMD in other studies, primarily due to the functional nature of this variant, which differentiates between the electrophoretic allotypes of C3. As high linkage disequilibrium (LD) exists between rs2230199 and rs1047286 (P314L), the latter cannot be excluded as a functional variant on the basis of a genetic signal [11].

Beyond the complement pathway, the Age Related Maculopathy Susceptibility 2 (ARMS2) locus at chromosome 10q26 has been implicated as the other major genetic contributor to the AMD disease process [13-16]. Although discussion continues regarding the source of the genetic effect observed at this locus due to high LD within the region, data have been reported supporting mitochondrial involvement through interaction with the translocase of outer mitochondrial membrane proteins and co-localization to the mitochondrial-rich ellipsoid region of the photoreceptors [17,18]. Mitochondrial dysfunction and resultant increased oxidative stress have been implicated as a component of AMD disease etiology, with reports of increased mitochondrial damage in the neural retina and RPE with aging, particularly in AMD. Variation within the mitochondrial genome at...
NADH ubiquinone oxidoreductase chain 2 (MTND2 A4917G) has been reported to confer increased risk associated with AMD [19,20] and as such would provide an excellent biologic candidate of functional significance, although subsequent replication and validation in independent cohorts is necessary. Here, we assess the effect of the C3 functional variant rs2230199 independent of CFH, ARMS2, CC2, CFB, age, and smoking status in a Northern Irish cohort; we check for potential gene–gene or gene–environment interactions, and we use the area under the receiver operator characteristic (ROC) curve to quantify the predictive ability of a logistic regression model, incorporating both genetic and environmental risk factors.

**METHODS**

**Study participants:** Patients with end-stage AMD (n=437; average age=77.6 years; 39.1% male), either GA or neovascular in at least one eye were recruited opportunistically between June 2002 and September 2006 from ophthalmology clinics in the Royal Victoria Hospital, Belfast, UK, which is the regional referral center for Northern Ireland. DNA was available from an age-matched control group (n=436; average age=74.9 years; 38.5% male) recruited through random sampling of older adults (aged 65 years and above) resident in Northern Ireland and who consented to stereoscopic digital fundus photography and DNA analysis to ascertain disease status. Fundi of control participants were either free of any drusen or had fewer than five hard drusen of diameter less than 63 µm and absence of focal pigmentary irregularities, such as hyperpigmentation or hypopigmentation, with image grading undertaken by trained reading center staff using the Wisconsin Age-related Maculopathy Grading System [21]. All participants were from Northern Ireland and described themselves as of European descent. An additional control sample (n=485) obtained through random sampling of the population from across Northern Ireland in 2006 was used to assess allele frequencies only but had not been assessed for AMD disease status.

**Clinical assessment of cases:** Participants recruited from ophthalmology clinics had a confirmed diagnosis of late AMD made following clinical examination and fluorescein angiography [21]. In 32 participants, the diagnosis was advanced atrophic AMD, i.e., geographic atrophy, and 405 participants were diagnosed with exudative AMD. These late-stage manifestations were present in at least one eye of each participant. Cases with macular pathology due to other primary causes that mimic neovascular AMD, such as myopic degeneration, adult vitelliform, central serous retinopathy, diabetic retinopathy, or idiopathic macular telangiectasia, were excluded. Those with lens opacities sufficient to obscure retinal detail were also excluded. A history of cigarette smoking was ascertained at the time of examination, and the questionnaire used had been previously validated in the Whitehall study [22], with participants classified as never smokers, ex-smokers or current smokers.

Informed written consent was obtained from all subjects involved. The study was approved by the Research Ethics Committee of the Queen's University of Belfast and the Office for Research Ethics Committee, Lisburn, Northern Ireland and conformed to the tenets of the Declaration of Helsinki.

**Genotyping of AMD susceptibility loci:** DNA was extracted from peripheral blood leukocytes or frozen buffy coat samples using standard protocols in accordance with the manufacturers’ instructions (Promega Wizard® Genomic DNA Extraction Kit, Southampton, UK). Genotyping was undertaken by multiplex PCR and primer extension methodology using standard manufacturer’s protocols (ABI Snapshot, ABI, Warrington, UK) using custom-designed primer and probe sequences (Table 1) to ascertain the genetic data for rs1061170, rs800292, rs419137, rs6677604, rs2284664, and rs3753396 in CFH; rs10490924 in ARMS2; rs9332739 in CC2; rs641153 in CFB; and rs2230199 in C3.

**Statistical analysis:** All SNPs were verified, validated, and assessed for Hardy–Weinberg equilibrium (HWE) separately in cases and controls, using a $\chi^2$ goodness-of-fit test. Allele frequencies for the SNPs being tested were compared in cases and controls by the Pearson’s $\chi^2$ test of association. The assessment of associations of genetic markers with AMD risk and interactions between genetic markers and between genetic markers and smoking were obtained using likelihood ratio $\chi^2$ tests in a logistic regression model. Age was included in the model as a continuous variable. Tests for trend for the number of risk alleles for each genetic variant (0, 1, or 2) were calculated. Initially only rs2230199 was fitted in the model and additional covariates were then added to adjust for their possible confounding effect. CFH haplotypes were entered into the logistic regression model as a count (0, 1, or 2) depending on the number of copies of the haplotype an individual carried. The lowest risk haplotype was omitted from the model to avoid collinearity problems so that odds ratios (ORs) were measured relative to homozygote carriers of this haplotype. For our final risk analysis, the number of copies carried for rs10490924, rs9332739, and rs641153 were coded 0, 1, or 2. Tests for interactions between each of the genotypes and between genotypes and smoking were used to check for effect modification. The fitted logistic model was used to order individuals in terms of their predicted AMD risk dependent on their rs2230199, CFH, rs10490924, rs9332739, and rs641153 genotypes and their smoking status.

A value for the area under the ROC curve, also known as the C statistic, was evaluated by using only phenotypes from either end of the severity spectrum, i.e., those with symptoms of end-stage disease only and those with no features of early AMD (namely pigmentary irregularities or soft drusen deposits less than 63 µm). A method to compare the area under two ROC curves derived from paired measurements was used.
RESULTS

Univariate analysis: Genomic SNP rs2230199 was assessed by a \( \chi^2 \) test and was found to be in HWE in both cases and controls. ORs, confidence intervals (CI), and associated p values are shown in Table 2. A minor allele frequency of 0.248 was recorded for rs2230199 in the general population (n=485), and a frequency of 0.296 in cases (n=437) and 0.221 in controls (n=436; OR=1.48; CI: 1.19–1.85; p=0.0003), similar to reports from previous studies [9-11].

Multivariate analysis: The significant effect associated with rs2230199 was maintained after adjustment for the effects of age, cigarette smoking status, and carriage of the adverse alleles of CFH, ARMS2, and CC2/CFB (Table 3: OR=1.45; CI=1.09–1.91; p=0.009).

Receiver operator characteristic curves: An ROC curve for cases versus controls with the covariates CFH, ARMS2, CC2, CFB, C3, smoking status, and age fitted in the regression model is presented in Figure 1. The area under the ROC curve that was generated from the model using all covariates was 0.860 (95% CI: 0.834, 0.885; p<0.001), which showed a non-significant (p=0.22) reduction to 0.857 (95% CI: 0.831, 0.883; p<0.001) upon removal of C3.

Potential gene–gene and gene–environment interactions: Potential gene–gene and gene–environment interactions were assessed using logistic regression for all the combinations of C3 and the other parameters recorded within the full model, such as CFH, ARMS2, CC2, CFB, smoking status, and age. There was no evidence to support any interaction between the parameters measured.
DISCUSSION

The polymorphic variation measured for the non-synonymous SNP rs2230199 in C3 shows significant association with late AMD in the Northern Irish population and supports the previously noted dose effect with an increased risk of 1.55 for heterozygotes (CI: 1.15–2.08; p=0.003) and 1.91 for homozygotes (CI: 1.11–3.30; p=0.013) [9-11].

The OR per allele copy decreases slightly from the univariate value calculated following adjustment through regression analysis for the following covariates: CFH, ARMS2, CC2, CFB, age, and smoking (OR=1.45; CI: 1.10–1.91; p=0.009). The minor allele frequency of rs2230199 ascertained from the general population sample (n=485) approximates at 0.25.

Considerable LD exists across this region, with the majority of investigations reporting the association observed at this locus on the basis of the genotype at rs2230199 [9-11]. With evidence supporting a biologic functional effect through the formation of two common allotypes of C3 (C3F [fast] and C3S [slow]), primarily due to variation in the electrophoretic mobility, the rationale for targeting this SNP for association studies is strong [24]. However, Park et al. [11] presented data supporting extensive LD within the region and suggested that until functional data are produced supporting rs2230199 as the causal variant in AMD disease etiology, it is important to consider the potential role of the non-synonymous SNP rs1047286 (P314L), which is in high LD with rs2230199, or indeed the potential role of any additional and as yet undetermined variants that may also be in complete LD.

Inclusion of rs2230199 as a variable has helped refine our AMD disease prediction model, providing an area under the receiver operator characteristic (ROC) curve for prediction of age-related macular degeneration. The area under the ROC curve was 0.860 (95% confidence interval [CI]: 0.834–0.885; p<0.001) for cases versus controls with the covariates CFH, ARMS2, CC2, CFB, age, and smoking fitted in the regression model (solid line). The area under the ROC curve showed a non-significant reduction (dotted line; p=0.22) to 0.857 (95% CI: 0.831, 0.883; p<0.001) upon removal of C3. The dashed line represents a reference area of 0.50.

### Table 3. Estimates of AMD Risk from a Logistic Regression Model with Terms for Smoking, CFH Haplotype [6], Genotype at ARMS2 A69S (rs10490924) and the Risk Associated with the CC2/CFB Region [8].

| Regression term | OR     | 95% C.I. | Significance |
|-----------------|--------|----------|--------------|
| Ex-smoker versus never smoker | 2.34   | 1.59     | 3.5          | <0.001 |
| Current smoker versus never smoker | 5.04   | 2.96     | 8.58         | <0.001 |
| Number of CFH haplotype 1 | 3.66   | 2.24     | 5.98         | <0.001 |
| Number of CFH haplotype 2 | 4.5    | 2.85     | 7.09         | <0.001 |
| Number of CFH haplotype 3 | 2.95   | 1.8      | 4.81         | <0.001 |
| Number of CFH haplotype 4 | 1.68   | 1.01     | 2.81         | 0.047  |
| ARMS2 A69S heterozygous | 3.64   | 2.47     | 5.36         | <0.001 |
| ARMS2 A69S homozygous | 25.3   | 13.6     | 47.2         | <0.001 |
| rs641153 (CFB) | 0.3    | 0.18     | 0.5          | <0.001 |
| rs9332739 (CC2) | 0.55   | 0.3      | 1.01         | 0.056  |
| rs2230199 (C3) | 1.45   | 1.1      | 1.91         | 0.009  |
| Age/year | 1.1    | 1.08     | 1.13         | <0.001 |

Risks associated with smoking are relative to never smokers. CFH haplotypes (1–4; haplotypes 1 and 2 increase risk and are in complete LD with the C risk allele at Y402H while 3 and 4 are non-risk and in complete LD with the T allele at Y402H) are measured against haplotype 5 (the most protective haplotype, which is in complete LD with the deletion at CFHR-3 ) [6]. ARMS2 is measured on the heterozygosity or homozygosity of the A69S risk allele (rs10490924 ). rs641153 (CFB), rs9332739 (CC2) and rs2230199 (C3) were coded as 0, 1 or 2 depending on the number of copies of the minor allele present.
ROC curve of 0.860 (Figure 1). Seddon et al. [25] reported an ROC curve of 0.831 in a longitudinal study of progression to advanced AMD, using a model incorporating information on the genetic loci rs1061170 and rs1410996 in CFH; rs19490924 in ARMS2; rs9332739 in CC2; rs641153 in CFB; and rs2230199 in C3 in addition to the response to therapeutic supplementation. Our ROC area was comparable to that of Seddon et al. [25], despite the cross-sectional nature of our study. The similarity of the values determined from this study to that reported by Seddon et al. [25] may be due to the homogeneity of our sample in terms of the late-stage disease phenotype, which could have contributed to an improved ROC area. Also, we used an informative representation of the CFH locus identifying five haplotypes [6] as opposed to the binary format of the Y402H locus, which is used in most investigations. Although Seddon et al. [25] used two SNPs, offering improved power over Y402H alone, we contend that there is merit in using the haplotypic structure of this locus, which improved the ROC statistic [6]. This haplotype is based upon several SNPs from this region, which enable the differentiation of five haplotypes, including the deletion at the CFHR1–3 locus. It was interesting to note that addition of rs2230199 within our regression model did not significantly improve the area under the ROC curve, with only a moderate increase from 0.857 to 0.860. Evaluation of risk prediction models requires model performance measures, such as the popular ROC metric [23]. Its probabilistic interpretation is based upon the ability of the models’ predicted probabilities to correctly discriminate between randomly selected diseased and nondiseased subjects. However, it is known that the area under the ROC curve often shows only small and nonsignificant improvements, even for a predictor variable that makes a significant independent contribution in a logistic regression analysis. More recently, alternative methods for evaluating the usefulness of a new marker have been proposed, such as event-specific reclassification tables and integrated (average) sensitivity, although these methodologies are more useful in cohort than case-control studies [26].

In conclusion, the C3 locus has a small but significant role in the etiology of the disease process of AMD in comparison to the other established genetic risk factors. Smoking and the genetic risk factors were found to exert an independent effect in this cohort study. While this conclusion supports the findings of other independent investigations, reports have been published supporting gene–gene or gene–environment interactions in AMD. Perhaps it would be appropriate to exercise caution in defining interaction on the basis of single studies but implementation of replication in independent data sets in a manner similar to that required for purely genetic association studies would provide a more robust measure.

Our findings add to the literature on calculating AMD risk and could, in time, facilitate earlier practical intervention for those at risk, with either lifestyle modification strategies or preclinical therapeutic intervention when these become available. These models will also enable us to identify predicted high-risk individuals who are aged and have not progressed to a severe disease phenotype or, conversely, individuals with low predicted risk and disease, for inclusion within future studies for elucidating novel risk factors that contribute to the AMD disease process.

**ACKNOWLEDGMENTS**

The authors are indebted to those who have participated in this study. We thank Evelyn Moore and Vittorio Silvestri for technical assistance. This research was funded by The Guide Dogs for the Blind Association UK (2008–5a); Research and Development Office, Northern Ireland Health and Personal Social Services (RRG 4.5); EVI-GENORET, an integrated project funded through the European Union Research Project (FP6).

**REFERENCES**

1. Centers for Disease Control and Prevention (CDC). Prevalence of visual impairment and selected eye diseases among persons aged >/= 50 years with and without diabetes—United States, 2002. MMWR Morb Mortal Wkly Rep 2004; 53:1069-71. [PMID: 15549022]
2. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. Science 2005; 308:385-9. [PMID: 15761122]
3. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Nouredine M, Gilbert JR, Schnetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA. Complement factor H variant increases the risk of age-related macular degeneration. Science 2005; 308:419-21. [PMID: 15761120]
4. Edwards AO, Ritter R III, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. Science 2005; 308:421-4. [PMID: 15761121]
5. Hageman GS, Anderson DH, Johnson LV, Hancox LS, Taiber AJ, Hardisty LI, Hageman JL, Stockman HA, Borchardt JD, Gehrs KM, Smith RJ, Silvestri G, Russell SR, Klaver CC, Barbazzoleti I, Chang S, Yannuzzi LA, Barile GR, Merriam JC, Smith RT, Olsh AK, Bergeron J, Zernant J, Merriam JE, Gold B, Dean M, Allikmets R. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. Proc Natl Acad Sci USA 2005; 102:7227-32. [PMID: 15870199]
6. Hughes AE, Orr N, Esfandiary H, az-Torres M, Goodship T, Chakravarty U. A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. Nat Genet 2006; 38:1173-7. [PMID: 16998489]
7. Gold B, Merriam JE, Zernant J, Hancox LS, Taiber AJ, Gehrs K, Cramer K, Neel J, Bergeron J, Barile GR, Smith RT. AMD Genetics Clinical Study Group, Hageman GS, Dean M,
Allikmets R. Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. Nat Genet 2006; 38:458-62. [PMID: 16518403]

8. McKay GJ, Silvestri G, Patterson CC, Hogg RE, Chakravarthy U, Hughes AE. Further assessment of the Complement Component 2 and Factor B region associated with Age-related Macular Degeneration. Invest Ophthalmol Vis Sci 2009; 50:533-9. [PMID: 18806297]

9. Yates JR, Sepp T, Matharu BK, Khan JC, Thurlby DA, Shahid H, Clayton DG, Hayward C, Morgan J, Wright AF, Armbrrecht AM, Dhillon B, Deary IJ, Redmond E, Bird AC, Moore AT. Genetic Factors in AMD Study Group. Complement C3 variant and the risk of age-related macular degeneration. N Engl J Med 2007; 357:553-61. [PMID: 17634448]

10. Maller JB, Fagerness JA, Reynolds RC, Neale BM, Daly MJ, Seddon JM. Variation in complement factor 3 is associated with risk of age-related macular degeneration. Nat Genet 2007; 39:1200-1. [PMID: 17767156]

11. Anderson DH, Mullins RF, Hageman GS, Johnson LV. A role for local inflammation in the formation of drusen in the aging eye. Am J Ophthalmol 2002; 134:411-31. [PMID: 12208254]

12. Jakobsdottir J, Conley YP, Weeks DE, Mah TS, Ferrell RE, Gorin MB. Susceptibility genes for age-related maculopathy on chromosome 10q26. Am J Hum Genet 2005; 77:389-407. [PMID: 16080115]

13. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, Meitinger T, Weber BH. Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum Mol Genet 2005; 14:3227-36. [PMID: 16174643]

14. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; 44:837-45. [PMID: 3203132]

15. Botto M, Fong KY, So AK, Koch C, Walport MJ. Molecular basis of polymorphisms of human complement component C3. J Exp Med 1990; 172:1011-7. [PMID: 1976733]

16. Yang Z, Camp NJ, Sun H, Tong Z, Gibbs D, Cameron DJ, Chen H, Zhao Y, Pearson E, Li X, Chien J, Dewan A, Harmon J, Bernstein PS, Shridhar V, Zabriskie NA, Hoh J, Howes K, Zhang K. A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. Science 2006; 314:992-3. [PMID: 17053109]

17. Kanda A, Chen W, Othman M, Branham KE, Brooks M, Khanna R, He S, Lyons R, Abecasis GR, Swaroop A. A variant of mitochondrial protein LOC387715/ARMS2, not HTRA1, is strongly associated with age-related macular degeneration. Proc Natl Acad Sci USA 2007; 104:16227-32. [PMID: 17884985]

18. Fritsche LG, Loenhardt T, Janssen A, Fisher SA, Rivera A, Keilhauer CN, Weber BH. Age-related macular degeneration is associated with an unstable ARMS2 (LOC387715) mRNA. Nat Genet 2008; 40:892-6. [PMID: 18511946]

19. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. Ophthalmology 1991; 98:1128-34. [PMID: 1634535]

20. Botto M, Fong KY, So AK, Koch C, Walport MJ. Molecular basis of polymorphisms of human complement component C3. J Exp Med 1990; 172:1011-7. [PMID: 1976733]

21. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. Ophthalmology 1991; 98:1128-34. [PMID: 1634535]

22. Reid DD, Hamilton PJ, McCartney P, Rose G, Jarrett RJ, Keen H. Smoking and other risk factors for coronary heart disease. Lancet 1976; 2:979-84. [PMID: 62262]

23. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008; 27:157-72. [PMID: 17569110]