nature genetics

Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration

Julian Maller[1,3], Sarah George[2], Shaun Purcell[1,3], Jes Fagerness[1,3], David Altshuler[1,3,4], Mark J Daly[1,3,4] & Johanna M Seddon[2,4]

Age-related macular degeneration (AMD) is a common, late-onset disease with seemingly typical complexity: recurrence ratios for siblings of an affected individual are three- to sixfold higher than in the general population, and family-based analysis has resulted in only modestly significant evidence for linkage. In a case-control study drawn from a US-based population of European descent, we have identified a previously unrecognized common, noncoding variant in *CFH*, the gene encoding complement factor H, that substantially increases the influence of this locus on AMD, and we have strongly replicated the associations of four other previously reported common alleles in three genes (*P* values ranging from 10^{-6} to 10^{-70}). Despite excellent power to detect epistasis, we observed purely additive accumulation of risk from alleles at these genes. We found no differences in association of these loci with major phenotypic categories of advanced AMD. Genotypes at these five common SNPs define a broad spectrum of interindividual disease risk and explain about half of the classical sibling risk of AMD in our study population.

We genotyped a dense set of common SNPs in numerous candidate genes, primarily in regions on chromosomes 1 and 10, on the basis of consistent positive linkage findings[1–4]. During the early phases of this study, a common coding variant in *CFH* (rs1061170; Y402H) was reported to be associated with AMD[5–8] and thus we included denser SNP coverage of this gene locus. In total, 1,536 tag SNPs were selected from HapMap Phase I (Methods) and genotyped using Illumina BeadArray at the Broad Institute/National Center for Research Resources (NCRR) Center for Genotyping and Analysis.

The study population consisted of 2,172 unrelated European-descended individuals 60 years of age or older diagnosed on the basis of ocular examination and fundus photography (1,238 affected individuals and 934 controls). Affected individuals were defined as those having advanced AMD with either geographic atrophy or neovascular disease (clinical age-related maculopathy staging system (CARMS) stages 4 and 5)[9]. Controls were individuals without AMD categorized as CARMS stage 1 (ref. 9). The mean age was 74 years for controls (54% female) and 78 years for affected individuals (55% female). We used genomic control SNPs to document the lack of artifactual elevation in the χ^2 distribution from population stratification between cases and controls.

As expected, we observed strong association to SNPs in the gene region encoding complement factor H (*CFH*). Unexpectedly, the strongest finding was the association with a previously unreported SNP (rs1410996, $P = 2.65 \times 10^{-61}$). This SNP lies within an intron of *CFH*, as do all four perfect proxies ($r^2 = 1$) of this variant on HapMap Phase II, and the functional consequence of these five SNPs is not immediately obvious from genome annotation. We also saw strong evidence for association at the previously described common coding variant Y402H ($P = 1.79 \times 10^{-59}$), with effects similar to those previously described: heterozygotes had 2.7-fold increased risk for AMD, and homozygotes 7.6-fold increased risk, as compared to the homozygous non-risk genotype.

We used stepwise logistic regression to determine whether the well-documented effect of Y402H might explain the association at the noncoding SNP (rs1410996). Even after conditioning on Y402H, rs1410996 was still strongly associated with disease risk ($P < 10^{-14}$) (**Table 1**). Similarly, conditional on rs1410996, Y402H remained associated with disease risk ($P < 10^{-9}$). Finally, although many of the other 63 SNPs genotyped in and around *CFH* showed highly significant individual associations, none of these SNPs remained significant conditional on the association at Y402H and rs1410996 together, and in addition, no other combination of two SNPs could explain the association at these loci. These results establish a second association, independent of Y402H, and indicate that Y402H and rs1410996 (or a perfect proxy) fully account for the association signal observed at the set of SNPs examined at this locus (**Supplementary Table 1** online). We cannot rule out the possibility that other variants

[1]Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge St., Boston, Massachusetts 02114, USA. [2]Epidemiology Unit, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, 243 Charles St. Boston, Massachusetts 02114, USA. [3]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [4]Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to J.M.S. (johanna_seddon@meei.harvard.edu) or M.J.D. (mjdaly@chgr.mgh.harvard.edu).

Received 1 May; accepted 1 August; published online 27 August 2006; doi:10.1038/ng1873

Table 1 Association between *CFH* variants and age-related macular degeneration

SNP	Allele	Control freq.	Affected freq	χ^2	*P* value	Conditional on rs1061170	Conditional on rs1410996
rs1061170[a]	C	0.359	0.615	264.5	1.79×10^{-59}	X	3.4×10^{-9}
rs1410996	C	0.571	0.808	272.9	2.65×10^{-61}	2.7×10^{-15}	X
rs1061170, rs1410996					3.7×10^{-64}		
	CC	0.356	0.609				
	TC	0.213	0.193				
	TT	0.428	0.194				

Two-letter allele symbols are used to indicate associated haplotypes of multiple SNPs.
[a]rs1061170 is the variant of *CFH* encoding the Y402H protein variant.

in *CFH*—particularly unexplored rare variations —might not also be associated with AMD.

These two SNPs in *CFH* are in strong linkage disequilibrium (LD) in this population, with only three of four possible haplotypes observed. Each haplotype carries a distinct level of risk for AMD: the high-risk allele at Y402H is always associated with the high-risk allele of rs1410996, but the low-risk allele of Y402H is further subdivided on the basis of genotype at rs1410996. Specifically, individuals homozygous for the high-risk haplotype are at roughly 15-fold greater risk of developing AMD than those homozygous for the low-risk haplotype. That the three haplotypes formed by these two linked SNPs carry completely distinct levels of risk suggests that it is unlikely a single undiscovered SNP (with only two alleles) explains the association. Effect modeling suggests a dominant influence of the noncoding SNP risk allele (Methods).

We also found individual SNPs under the chromosome 10 linkage peak that were highly significantly associated with risk of AMD ($P < 10^{-20}$). In addition to testing all the individual SNPs that we genotyped, we carried out multimarker haplotype tests to predict (requiring $r^2 > 0.8$) SNPs in HapMap Phase II untagged by our individual Phase I tags (**Supplementary Table 2** online). One such haplotype was markedly associated with risk of AMD ($P < 10^{-56}$), and this haplotype is included because of strong correlation with several SNPs including rs10490924 ($r^2 = 0.79$), the same SNP that has shown maximal association with risk of AMD in recently published studies[10,11]. Subsequently, we typed rs10490924 (**Table 2**) and showed that this is indeed the most associated SNP or haplotype in the region. These data unequivocally confirmed this second AMD locus, strongly implicating the A69S-encoding variant of *LOC387715* (the putative change induced by rs10490924) as causally related to AMD. Final confirmation of this as the causal variant and gene remains unproven, however: HapMap Phase II data revealed three perfect proxies for rs10490924, including one 17 kb away, intronic to the downstream gene *HTRA1*; in addition, stepwise regression suggested that some evidence of association persists in this region after conditioning on rs10490924.

Very recently, common variation in the gene locus containing the *C2* and *CFB* (also called *BF*) genes (encoding complement component 2 and complement factor B, respectively) was reported to be associated with risk of AMD[12]. We genotyped, using Sequenom (Methods), the same DNA samples for the four SNPs found in that report to be most significantly associated, and thereby strongly confirmed the association to two independent gene variants with risk of AMD. Specifically, BF:R32Q and rs547154 (C2:intron10) were both strongly associated ($P < 10^{-11}$) with AMD; these two SNPs were highly correlated ($r^2 = 0.78$) and thus represent the same association. Similarly, the correlated BF:L9H and C2:E318D variants ($r^2 = 0.8$) were each associated with risk of AMD at $P < 10^{-5}$; these two SNPs were highly correlated. Although both pairs of SNPs have minor alleles that conferred an equivalent protective effect, these effects were independent and distinct, as determined by stepwise logistic regression of each SNP pair conditional on the other (**Table 3**). With the sample size assembled here, we excluded rs547154 in favor of BF:R32Q ($P < 10^{-4}$) by stepwise regression; however, as each of BF:L9H and C2:E318D still maintained a modest effect conditional on the other, a third variant in strong LD with both variants may be the truly causal second allele at this locus.

Using HapMap data to evaluate the evolutionary history of these alleles (**Supplementary Fig. 1** online), we observed that these rare, protective variants arose on two distinct haplotype backgrounds, suggesting it is highly unlikely that a single variant exists that could fully capture the association to both of these SNPs. Thus, we strongly confirmed the association of at least two independent common variants at the *C2-CFB* locus with risk of AMD. (Because of the low frequency and nearly identical risk of these protective alleles, we considered individual risk at *C2-CFB* in subsequent modeling in two categories: high-risk (individuals carrying neither protective variant) and low-risk (individuals carrying one or two of the protective variants) (**Table 3**; **Fig. 1**)).

We then evaluated the role of epistasis among the five common variants at these three loci (*CFH*, *LOC387715* and *C2-CFB*). Although our analysis had excellent power to detect epistasis that contributes substantially to risk (consider that the risk alleles are common and the *P* values for association in this sample range from 10^{-6} to 10^{-70}), we observed no statistically significant non-additive interactions. Specifically, a model in which the risk alleles at the three genes act completely independently (individual risks are multiplied to generate a combined risk profile) provided a better fit to the observed data than the same model with the inclusion of interlocus interaction terms. Model selection was done using Akaike's Information Criteria (AIC) and complete details of each model attempted are provided in Methods. Thus, although we cannot rule

Table 2 Association between *LOC387715* variants and age-related macular degeneration

SNP	Allele	Control freq.	Affected freq.	χ^2	Nominal *P* value
rs10490924	T	0.194	0.455	315.075	1.71×10^{-70}

Table 3 Association between *CFB* and *C2* variants and age-related macular degeneration

SNP	Allele	Control freq.	Affected freq.	χ^2	Nominal *P* value	*P* value conditional on rs9332739	*P* value conditional on rs641153
rs9332739[a]	C	0.054	0.026	23.6	1.10×10^{-6}	X	9.61×10^{-7}
rs547154	A	0.097	0.054	27.8	1.30×10^{-7}	6.09×10^{-8}	0.0508
rs4151667	A	0.050	0.029	13.8	0.0002	0.0397	6.8×10^{-5}
rs641153[b]	T	0.102	0.045	51.9	5.5×10^{-13}	2.99×10^{-13}	X

[a]rs9332739 is the variant of *C2* encoding the E318D protein variant, highly correlated to rs4151667 (BF:L9H). [b]rs641153 is the variant of *BF* encoding the R32Q protein variant, highly correlated to rs547154.

out the existence of small amounts of non-additivity, the observed data are consistent with a model in which there is none.

We tested whether or not these SNPs were specifically associated with phenotypic class (geographic atrophy (dry) versus neovascular disease (wet)). Despite very good power to detect association, none of these risk alleles showed significant differences in association with these phenotypic subclassifications of advanced AMD (all $P > 0.05$). We also did not see any meaningful differences in association by gender.

Finally, our finding of a new common, noncoding SNP at *CFH*, replication of common variation at *LOC387715* and the first unequivocal confirmation of association at *C2-CFB* allowed us to estimate a model of risk for AMD based on these five validated common variants. In contrast to the modest elevation in overall risk to siblings (three- to sixfold[13–15]), the predictive value of specific genotype combinations was notable. For example, we define baseline risk as that observed for individuals carrying the lowest-risk genotypes at all three loci (approximately 2% of the population sampled). As compared to this low-risk group, we estimated approximately 10% of the population to have 40-fold greater risk and 1% (high-risk homozygotes at all three loci) to have a more than 250-fold greater risk (**Fig. 1**, **Supplementary Tables 3** and **4** online). These risk estimates were dependent on the specific inclusion and diagnostic criteria used in this study; however, we note that the risk estimates at all three loci were similar to those from previous studies. Expressed another way, these genotypes apparently identify individuals whose lifetime risk of AMD ranges from less than 1% to more than 50%; however, longitudinal studies are needed to define the true risk attributable to these loci and the ways in which these might interact with the known environmental and lifestyle risk factors[16].

Intermediate risks were observed for most combinations of common alleles, documenting the classic model of quantitative genetics predicted by Fisher[17], in which genotypes at a handful of loci can combine to create a nearly continuous gradient of risk in a population. A corollary of this is that, although genotype can identify a subset of individuals with extreme risk or protection (those who inherit a preponderance of risk or protective alleles), the majority of the population inherits some of each category, resulting in a lifetime risk of AMD that is only modestly different from the overall population average. We note, however, that other genetic factors most likely have yet to be discovered; these, in combination with environmental variables, will further stratify individual risk more precisely.

We estimate that these five common variants explain roughly half the excess risk of AMD to siblings of affected individuals. Specifically, the relative risk to siblings (λ_s) attributable to these loci are ≈ 1.4 for the two-SNP combination at *CFH*, ≈ 1.45 for rs10490924 and 1.05 for *C2-CFB*. Given the independence of the effects, common variants at these three loci thus show λ_s of slightly more than 2, a major fraction of the overall λ_s of 3–6.

In summary, we identify a previously unknown, independent and noncoding common allele at *CFH* associated with AMD. We also provide strong additional statistical support for the effects of two well-described common variants on the risk of AMD (Y402H at *CFH* and A69S at *LOC387715*), and we confirm for the first time the association of two independent alleles at the *C2-CFB* locus. Although these results may not be representative of other diseases, these common variants offer empirical data that provide input into some widely debated points in the field of human genetics: (i) that there can exist common alleles of substantial effect on common disease (in AMD, these explain at least half of all risk to siblings); (ii) that these can be found outside of 'candidate genes' and outside of coding regions; (iii) that for such alleles, association offers much greater power and reproducibility than linkage (association results much stronger than similarly sized linkage studies); (iv) that even for late-onset diseases with partial heritability, common genotypes can strongly influence individual risk; and (v) that there need not be epistasis among, or phenotypic sub-stratification by, genes of substantial population effect.

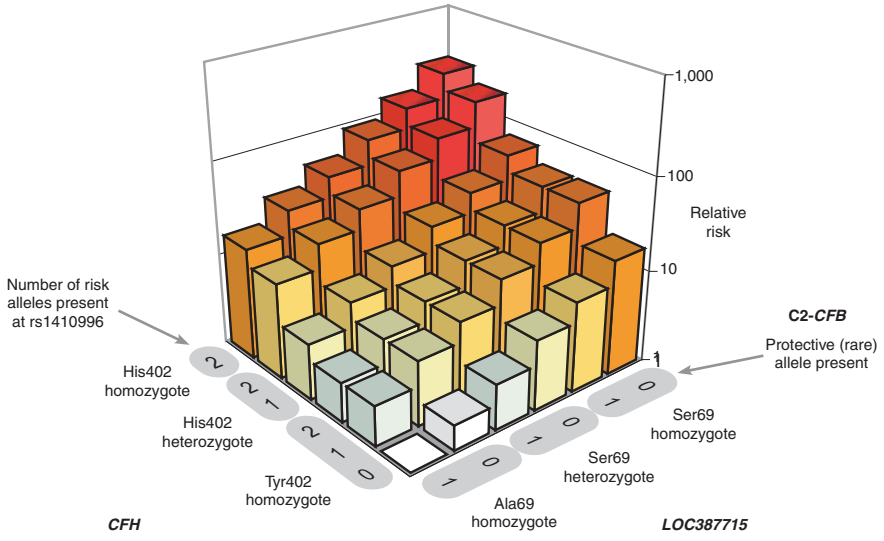

Figure 1 Relative risk plotted as a function of the genetic load of the five variants that influence risk of AMD. Two variants are in the *CFH* gene on chromosome 1: Y402H and rs1410996. Another common variant (A69S) is in hypothetical gene *LOC387715* on chromosome 10. Two relatively rare variants are observed in the *C2* and *BF* genes on chromosome 6. We find no evidence for interaction between any of these variants, suggesting an independent mode of action. (See also **Supplementary Table 4**.)

Only by thoroughly testing common genetic variation in other large clinical cohorts will the generality (or lack thereof) of these observations become clear.

METHODS

Subjects. The methods used in this study conformed to the tenets of the Declaration of Helsinki and received approval from the institutional review board at the Massachusetts Eye and Ear Infirmary, Boston. Informed consent forms were signed by all participants. Some methods have been described in detail previously[1]. For these analyses, unrelated European-descended individuals with extremely discordant phenotypes were included. Affected individuals were defined as those having advanced age-related macular degeneration (AMD) with either geographic atrophy or neovascular disease as determined by fundus photography and ocular examination (clinical age-related maculopathy grading system (CARMS) stages 4 and 5)[9,18]. Controls were unrelated to the affected individuals, were 60 years of age or older, and were defined as individuals without macular degeneration, and without early or intermediate disease, categorized as stage 1, on the basis of fundus photography and ocular examination[9]. Subjects were derived from ongoing AMD study protocols with similar procedures including the Progression of AMD Study[19,20], AMD Registry Study[21], Family Study of AMD[1], The US Twin Study of AMD[22] and the Age-Related Eye Disease Study (AREDS)[18]. A subset of the AREDS samples have been used in previous studies of *CFH*:Y402H[5,16]. Estimates of risk were equivalent in AREDS and non-AREDS samples, so this subset of AREDS samples, previously unexplored during the design of this study for all other loci, were included.

Gene selection. We selected the *CFH* gene region on the basis of previous reports establishing association to the Y402H variant of this gene[5–8]. We chose to look further 3′ into the region because of the high linkage disequilibrium across that area of the genome. We did this in order to positively rule out any other markers in the area that could also affect the association results of the previous studies, given the high degree of linkage disequilibrium in that region. We genotyped 1,118 SNPs across 19.9 Mb on the distal end of chromosome 10 where we and others had found linkage in previous studies and where a recent meta-analysis[23] had found the strongest linkage signal. We also chose to genotype SNPs in individual genes that had previous positive associations with AMD[24]. We selected several other candidate genes on the basis of animal study data and expression levels in the eye and other functional pathway involvement in genes thought to be related to AMD.

SNP selection. We genotyped 1,536 tag SNPs within our selected candidate genes and under the previously reported linkage peak on chromosome 10 extending 19.9 Mb. These tag SNPs were selected from the HapMap Project database using the Tagger tool (http://www.broad.mit.edu/mpg/tagger/) and are based on the HapMap genotypes for the CEPH (Centre d'Etude du Polymorphisme Humain) population of 30 US trios with northern and western European ancestry. We selected SNPs with a minor allele frequency of $>10\%$, and we set a minimum threshold value of 0.8 for the r^2 parameter, which represents the multivariate coefficient of determination for all alleles that are to be captured. Thus, the SNPs that were selected should have been highly representative of the genetic variance within each region of interest because they were direct proxies of other SNPs in those areas, or the SNPs were part of a multimarker haplotype made up of other selected SNPs that were themselves in strong LD.

Genotyping. Genotyping was carried out on an Illumina system at The Broad Institute Center for Genotyping and Analysis (TBICGA), which was subsidized through a grant from the National Center for Research Resources. Several duplicates were added to each 96-well sample plate for quality assurance and quality control validation of inter-plate discordance, and we placed an extra 179 duplicates into our sample set in order to test for experiment-wide discordance. The data completion rate was 98.4%. The process for Illumina genotyping can be found on their website (http://www.illumina.com/technology/tech_overview.ilmn).

Genotyping at the *C2-CFB* locus[12] was carried out using Sequenom technology at The Broad Institute Center for Genotyping and Analysis. More information on this technology can be found on their website (http://www.sequenom.com/applications/hme_assay.php).

Statistical analysis. To further evaluate potential population stratification effects in this sample, we also carried out an analysis that was based on the average identity-by-state (IBS) for each pair of individuals, using the program PLINK (http://pngu.mgh.harvard.edu/purcell/plink/). We then applied multidimensional scaling to this IBS distance matrix, to extract two main axes of variation that can be correlated with phenotype; next, we applied hierarchical complete linkage clustering. Neither methods detected any evidence for stratification.

We used the software package *whap* (http://pngu.mgh.harvard.edu/purcell/whap/) to carry out the conditional analyses. *whap* uses a weighted maximum-likelihood approach to allow for potential ambiguity in statistically inferred haplotypes. As well as basic omnibus and haplotype-specific tests, it is possible to construct likelihood ratio tests between two nested models, which allows for conditional analysis of SNP effects independent of their haplotypic background. The method implemented in *whap* has been previously described[25] and was recently found to perform well in a comprehensive simulation study comparing different haplotype-based methods for case-control association analysis[26]. For example, as shown in **Supplementary Table 5** online, comparing the alternate with the first null model provides a test of independent effects of the first SNP, that is, controlling for the second SNP; comparing the alternate model to the second null model provides a similar test for the second SNP. (That haplotypes CT and TT both have the β_4 indicates that they are constrained to have the same estimated coefficient.)

Genotypic and allelic odds ratios were computed directly as P(affected individual|genotype)/P(affected individual|not-genotype); odds ratios for haplotypes were computed by *whap*. Sibling risk estimates were derived algebraically from the odds ratios following ref. 27.

Genomic control. The method of genomic control implemented in PLINK follows ref. 28 by simply dividing the median χ^2 association statistic by 0.456 to estimate an inflation factor. We used 48 unlinked SNPs (not including previously reported associations at *CFH* and *LOC387715*), and the median χ^2 value was 0.436. Hence, we did not use a correction factor.

Multilocus modeling. We used logistic regression (implemented in the glm() function of the R statistical analysis package; http://www.r-project.org/) to carry out model selection for the three loci. For each locus individually, we established the most parsimonious best-fit model on the basis of likelihood ratio test statistics considering a series of nested submodels, as described below. The three models were then combined into a single model to provide a test for interlocus interaction, or epistasis (**Supplementary Table 6** online).

We considered the additive effects of each haplotype (coded 0, 1 and 2 for the number of copies of each haplotype) and extra terms for dominance effects (coded 1 for possessing exactly 1 copy of a haplotype, and otherwise 0). In each model, one haplotype effect was designated as the baseline and therefore removed from the model; although this step is necessary for identifiability, it does not alter the results.

For *CFH*, the additive effect of the TT haplotype was used as the baseline haplotype. Compared to the full model A1, which allows for dominance effects for all haplotype combinations, in the A2 model we dropped all dominance terms. This resulted in a significant reduction in fit ($\chi^2_3 = 9.9$; $P = 0.019$), suggesting that dominance effects play a role at this locus. On the basis of the parameter coefficients in A1, we next fit a model that allowed only TC to show dominance effects; this model also equates the dominance terms $D_{TC}D_{TT}$ and $D_{TC}D_{CC}$, which implies that the effect of having exactly one copy of TC is the same whether it is paired with the TT or CC haplotype. (Note: equating coefficients for $D_{TC}D_{TT}$ and $D_{TC}D_{CC}$ is equivalent to just entering D_{TC} alone into the model.) Compared to A1, A3 did not show any significant reduction in fit ($\chi^2_2 = 1.6$; $P = 0.449$). Finally, we asked whether the TC haplotype showed complete or incomplete dominance: model A4 enters a recoded variable {0,1,1} for having 0,1 or 2 copies of the TC haplotype (DD_{TC}) (note: this model is still nested in models A1–A3.) Compared to A3, model A4 does not show any significant deterioration in fit, suggesting that a model in which the TC haplotype is completely dominant (having one copy has the same effect as having two copies) is a better description of the data. Model A4 has the lowest AIC of the models tested, suggesting it is the best-fitting model. In summary, these analyses suggest the best-fitting model for the *CFH* locus is, relative to the

TT haplotype, an additive effect of CC and a dominant effect of TC; this model was therefore used in the multilocus modeling.

For *LOC387715*, we designated the additive effects of the ACA haplotype as the baseline effect. On the basis of the full model, B1, we dropped all terms except for an additive effect of the AGA haplotype (model B2). This reduced model was not significantly different from B1 as determined by a likelihood ratio test ($\chi^2_8 = 10.7$; $P = 0.219$). We therefore used *B2* as the best-fit model for *LOC387715*—that is, a purely additive effect of the AGA haplotype.

For *C2-CFB* we did not carry out formal model selection, as the protective haplotypes are very rare in their homozygous states and so any test would be drastically underpowered. Instead, we selected a model for this locus in which the common GCTC haplotype shows a recessive effect. As well as consideration of genotype frequencies, this is based on the *whap* analysis that indicates that the two rare haplotypes are functionally indistinguishable. That is, parameters for the additive effects of GATT and CCAC are not significantly different from each other ($\chi^2_1 = 0.04$; $P = 0.833$). This simplifies the model to a single recessive effect of GCTC (or, described alternately, a single dominant effect ascribable to carriage of the rare allele at either associated SNP).

A multilocus model (M1) was constructed to include all the terms from the selected models for the three loci and all pairwise interlocus interactions. As a whole, there was no compelling evidence for any epistasis from this analysis. One term ($DD_{TC}\ R_{GCTC}$) was significant at $P = 0.036$, although this was determined without controlling for the multiple testing inherent in testing five epistatic components. More importantly, the likelihood ratio test contrasting models M1 and M2 showed no evidence of epistasis ($\chi^2_5 = 5.4$; $P = 0.369$).

To confirm the absence of epistasis, we also tested an 'allelic epistasis' model, in which we simply entered allelic effects for the four haplotypes (TC and CC at *CFH*, GCTC at *LOC387715* and AGA at *C2-CFB*) and all interlocus pairwise epistatic terms, instead of relying on the best-fit models derived from the single-locus modeling. This alternate approach also did not yield any evidence for epistasis (results not shown).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank the study participants, their families and numerous ophthalmologists throughout the country who participated in this study. We particularly thank D. Mirel and the NCRR Broad Institute Center for Genotyping and Analysis for expert design and execution of the SNP genotyping reported herein, and P. de Bakker for comments and **Figure 1** graphics. We also thank AREDS participants and investigators and the EMMES Corporation for their work on the AREDS Genetic Repository. This research was supported by EY11309 from the US National Institutes of Health; the Foundation Fighting Blindness; Massachusetts Lions Research Fund, Inc.; the Epidemiology Unit AMD Genetics Research Fund, Massachusetts Eye and Ear Infirmary; and the Broad Institute Center for Genotyping and Analysis, supported by grant U54 RR020278 from the NCRR.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Seddon, J.M. *et al.* A genomewide scan for age-related macular degenerations providences evidence for linkage to several chromosomal regions. *Am. J. Hum. Genet.* **73**, 780–790 (2003).
2. Weeks, D.E. *et al.* Age-related maculopathy: an expanded genome-wide scan with evidence of susceptibility loci within the 1q31 and 17q25 regions. *Am. J. Ophthalmol.* **132**, 682–692 (2001).
3. Majewski, J. *et al.* Age-related macular degeneration—a genome scan in extended families. *Am. J. Hum. Genet.* **73**, 540–550 (2003).
4. Iyengar, S.K. *et al.* Dissection of genomewide-scan data in extended families reveals a major locus and oligogenic susceptibility for age-related macular degeneration. *Am. J. Hum. Genet.* **74**, 20–39 (2004).
5. Klein, R.J. *et al.* Complement factor H polymorphism in age-related macular degeneration. *Science* **308**, 385–389 (2005).
6. Edwards, A.O. *et al.* Complement factor H polymorphism and age-related macular degeneration. *Science* **308**, 421–424 (2005).
7. Hageman, G.S. *et al.* A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc. Natl. Acad. Sci. USA* **102**, 7227–7232 (2005).
8. Haines, J.L. *et al.* Complement Factor H variant increases the risk of age-related macular degeneration. *Science* **308**, 419–421 (2005).
9. Seddon, J.M., Sharma, S. & Adelman, R.A. Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology* **113**, 260–266 (2006).
10. Rivera, A. *et al.* 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.* **14**, 3227–3236 (2005).
11. Jakobsdottir, J. *et al.* Susceptibility genes for age-related maculopathy on chromosome 10q26. *Am. J. Hum. Genet.* **77**, 389–407 (2005).
12. Gold, B. *et al.* Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. *Nat. Genet.* **38**, 458–462 (2006).
13. Seddon, J.M., Ajani, U.A. & Mitchell, B.D. Familial aggregation of age-related maculopathy. *Am. J. Ophthalmol.* **123**, 199–206 (1997).
14. Heiba, I.M., Elston, R.C., Klein, B.E. & Klein, R. Sibling correlations and segregation analysis of age-related maculopathy: the Beaver Dam Eye Study. *Genet. Epidemiol.* **11**, 51–67 (1994).
15. Klaver, C.C. *et al.* Genetic risk of age-related maculopathy. Population-based familial aggregation study. *Arch. Ophthalmol.* **116**, 1646–1651 (1997).
16. Seddon, J., George, S., Rosner, B. & Klein, M. CFH gene variant, Y402H, and smoking, body mass index, environmental associations with advanced age-related macular degeneration. *Hum. Hered.* **61**, 157–165 (2006).
17. Fisher, R.A. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinburgh.* **52**, 399–433 (1918).
18. Age-Related Eye Disease Study Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C & E, beta carotene, and zinc for age-related macular degeneration and vision loss. *Arch. Ophthalmol.* **119**, 1417–1436 (2001).
19. Seddon, J.M., Cote, J., Davis, N. & Rosner, B. Progression of age-related macular degeneration: Association with body mass index, waist circumference and waist-hip ratio. *Arch. Ophthalmol.* **121**, 785–792 (2003).
20. Seddon, J.M., Cote, J., Davis, N. & Rosner, B. Progression of age-related macular degeneration: association with dietary fat, transunsaturated fat, nuts, and fish intake. *Arch. Ophthalmol.* **121**, 1728–1737 (2003).
21. Snow, K.K. *et al.* Association between reproductive factors and age-related maculopathy in post-menopausal women. *Am. J. Ophthalmol.* **134**, 842–848 (2002).
22. Seddon, J.M., Cote, J., Page, W.F., Aggen, S. & Neale, M. The US twin study of age-related macular degeneration: relative roles of genetic and environmental influences. *Arch. Ophthalmol.* **123**, 321–327 (2005).
23. Fisher, S.A. *et al.* Meta-analysis of genome scans of age-related macular degeneration. *Hum. Mol. Genet.* **14**, 2257–2264 (2005).
24. Haddad, S., Chen, C., Santangelo, S.L. & Seddon, J.M. The genetics of age-related macular degeneration: a review of progress to date. *Surv. Ophthalmol.* **51**, 316–363 (2006).
25. Sham, P.C. *et al.* Haplotype association analysis of discrete and continuous traits using mixture of regression models. *Behav. Genet.* **34**, 207–214 (2004).
26. Cordell, H.J. Estimation and testing of genotype and haplotype effects in case-control studies: comparison of weighted regression and multiple imputation procedures. *Genet. Epidemiol.* **30**, 259–275 (2006).
27. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517 (1996).
28. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).