Supporting Online Material

Materials and Methods

Participant selection

From the AREDS sample, we identified 96 case subjects who, at their most recent study visit, had either uniocular choroidal neovascularization (50 cases) or geographic atrophy either central or non-central to the macula (46 cases). The fellow eye of these case subjects was required to have at least one large drusen (>125 μm in diameter), and total drusen area equivalent to a circle of at least 1061 μm in diameter. Controls were 50 individuals from the AREDS sample who had few or no drusen (< 63 μm in diameter in each eye) for the duration of their participation in AREDS.

Genotyping

We genotyped each individual using the Affymetrix GeneChip Mapping 100K Set of microarrays (*S1*). This mapping assay consisted of two chips (XbaI and HindIII) with approximately 50,000 SNPs each that were used for each individual. About 250 ng of genomic DNA was digested with two restriction enzymes XbaI and HindIII and processed according to the Affymetrix protocol (*S1*). The images were analyzed using GDAS software (Affymetrix). For the data obtained from each chip, two internal quality control measures were used: the call rate always exceeded 95% and heterozygosity on the X chromosome correctly identified the gender of the individual. To ensure that no

samples were confused, we checked that the 31 identical SNPs placed on both chips yielded the same genotype for the same individual.

To test the reproducibility of this system, we performed three experiments. First, 4 samples were processed twice with the Xba chips. Next, 2 replicates of a reference DNA positive control provided by Affymetrix were run on Xba chips alongside the samples described here. Finally, to test the accuracy of this assay, results for 3 individuals were compared with genotyping using the Affymetrix 10K SNP platform (*S2*).

To remove a SNP for which genotyping was consistently problematic, we required that it was called in at least 85% of the individuals. We then removed any SNP which was monomorphic in our data, as these SNPs are uninformative. Finally, we removed SNPs in which genotype frequencies deviated from the Hardy-Weinberg equilibrium expectation (HWD $\chi^2 > 25$, $P = 0.05$, 1df, after Bonferroni correction) as being likely to contain genotyping errors. SNPs for which no homozygotes were observed in the entire sample were also likely due to errors and removed. Taken together, 105,980 SNPs with a call rate of at least 85%, both alleles observed, at least one homozygote observed, and a HWD $\chi^2 \leq 25$ were considered for further analysis. Table S1 summarizes the quality and informativeness of the genotypes.

Statistical analysis

To test for allelic association between each SNP and disease status, we constructed a 2×2 contingency table of allele frequencies. We then computed the Pearson χ^2 value and a *P*-value based on the central χ^2 distribution under the null hypothesis of no association with 1 degree of freedom. We corrected this nominal *P*-value for multiple testing by applying the Bonferroni criterion, in which the nominal *P*-value is multiplied by 103,581 (the number of autosomal SNPs that passed our quality control checks) to produce the Bonferroni-corrected *P*-value.

We used two methods of genomic control to look for population stratification, GC and GCF (*S3*). In the first method, we took the median χ^2 value for allelic association with a number of SNPs assumed to be unassociated with the disease (null SNPs). We then divided our test statistic $\chi^2_{(1)}$ values by this median, and tested for significance using the χ^2 distribution. Alternatively, for the GCF method, we took the mean rather the median of the null χ^2 statistics, and tested significance of the quotient using the F(1,L) distribution, where L is the number of null SNPs used (*S3*). We used two different sets of unassociated SNPs: all the SNPs successfully genotyped except the two significant ones, and the set of 31 SNPs that are in common between the two chips used in the assay (see Genotyping section, above).

Odds ratios, confidence intervals, and population attributable risk were calculated as described in (*S4*). The population frequency of the alleles of interest is relatively high,

23% for and 41% for homozygous rs380390 and rs1329428, respectively. Therefore, the odds ratios necessarily calculated from the case control design study used here will over-estimate (without changing the significance levels) the equivalent relative risk estimate needed to calculate lifetime risk. A prospective cohort study design will be necessary to provide valid estimates of lifetime risk in persons who have and have not inherited the alleles.

We defined the candidate region by looking for adjacent SNPs in which all four gametes were observed (*S5*) and bounding the region there. To look at linkage disequilibrium between SNPs in the candidate region, we inferred haplotype frequencies in the region using PHASE version 2.1 (*S6, S7*). Based on the inferred haplotype frequencies across the entire region, pairwise linkage disequilibrium was calculated by first computing the two-locus haplotype frequencies implied by the overall haplotype frequencies. The measure of linkage disequilibrium, D', was then calculated using standard equations (*S8*) and plotted using the program GOLD (*S9*).

To define the smaller haplotype blocks within the 4-gamete region, we accessed the HapMap data website (http://www.hapmap.org) on December 15, 2004, and downloaded genotypes for SNPs in the region between SNPs rs10494744 and rs10484502. Genotypes for the CEU population (CEPH Utah population of northern and western European ancestry) were downloaded and visualized using Haploview version 3.0. Haplotype blocks were then defined using the method and parameters of Gabriel et al (*S10*).

Haplotypes across the narrower region defined by the HapMap block were also inferred using PHASE version 2.1. Haplotypes with an estimated frequency of at least 1% were considered for further analysis. Phylogenetic trees were built using the maximum parsimony of PHYLIP 3.62 ("dnapars" program). We then computed the odds ratio in a nested cladistic framework for the haplotypes (*S4, S11*).

Resequencing

To resequence the CFH gene, we selected 96 individuals (66 cases and 30 controls). Most of these individuals were selected either because SNP rs380390 was homozygous (representing opposite risk groups) or SNP rs10272438 was not successfully genotyped (as we planned to use the same plates to re-sequence this SNP for genotyping). Three additional individuals were randomly selected to get the total of 96 for a full plate. Primer design, PCR amplification, bi-directional sequencing of PCR products, and mutation analyses were performed by Genaissance (New Haven, CT). A list of new variants identified and primers used is provided in Tables S4 and S5, respectively.

Immunofluorescence Microscopy for CFH

Tissue preparation

Normal donor eyes were embedded in optimal cutting temperature compound (OCT; Miles Laboratory, Elkhart, IN), snap frozen, and stored at -70°C. Frozen retina sections were cut at 8 to 10 μm and placed on slides (Superfrost/Plus; Fisher Scientific, Fair

Lawn, NJ). All human eyes were obtained with the informed consent of the donors, and the research with human eyes was performed in accordance with the tenets of the Declaration of Helsinki and the Institutional Review Board (IRB).

Immunofluorescence Microscopy

For immunofluorescence labeling, frozen sections of human retina were fixed in 4% paraformaldehyde in phosphate buffer saline (PBS) for 10 minutes. The tissue sections were blocked for 30 minutes with 5% normal donkey serum (Jackson Immunoresearch, West Grove, PA), diluted in IC buffer (PBS, containing 0.2% Tween-20, 0.1% sodium azide), and incubated for 1 hour at room temperature with a goat anti-human Factor H antibody (Quidel, Santa Clara, CA) diluted 1:200 in staining buffer (IC buffer plus 1% normal donkey serum). Sections were washed repeatedly in IC buffer and incubated for 1 hour with the nuclear dye 4',6'-diamino-2-phenylindole (DAPI; 1 μg/mL) and Alexa-488 Donkey anti-goat antibodies (Molecular Probes, Eugene, OR) diluted 1:250 in staining buffer. After repeated washing with IC buffer, sections were covered in mounting medium (Gel Mount; Biomeda, Foster City, CA) and coverslipped. For the control, the same concentration of anti-human factor H antibody was preincubated for 1 hour with purified human factor H protein (Calbiochem, La Jolla, CA) at the ratio of 3 μg for 1 μl of antibodies. The pretreated antibodies were then used to stain tissue sections as just described. Specimens were analyzed on a laser scanning confocal microscope (model SP2; Leica Microsystems, Exton, PA) equipped with Nomarski optics. Immunolabeled

and negative control sections were imaged under identical scanning conditions. Images were processed with Photoshop (Adobe Systems, San Jose, CA).

Immunohistochemistry for C5b-9

Source of tissue

Post-mortem retinas from four donors were examined. Three were obtained through the Foundation Fighting Blindness (FFB) eye donor program. All of these had a clinical diagnosis of dry AMD. One pair of eyes embedded in paraffin was obtained from an 86 year old Caucasian female through the autopsy service of the Yale School of Medicine. No clinical history was available. Histologically, these retinas have multiple large or coalescing drusen with minimal RPE and photoreceptor loss consistent with a diagnosis of early AMD. Approval for research on human post mortem donor eyes was obtained from the Yale School of Medicine.

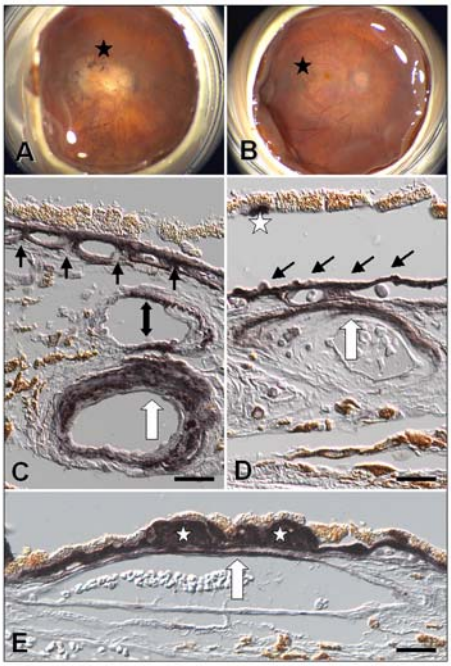
Tissue processing and histology

Upon enucleation, eyes were fixed in 4% paraformaldehyde, 0.5% glutaraldehyde in 0.1 M phosphate buffer for several days. The fixed eyes were transferred to 2% paraformaldehyde for storage. Six 0.5 cm circular punches were taken from each of the AMD donor eyes. Three of these were selected from the central retina at the junction of atrophic and more normal retina, and the remaining three from peripheral retina. Retinal plugs were embedded in paraffin and sections cut at 5μm.

Immunohistochemistry

Following deparaffinization and rehydration, antigen retrieval was performed by boiling sections in a microwave oven for 10 minutes in 10mM sodium citrate (pH 6.0). Sections were allowed to cool for 20 minutes, prior to a 5-minute endogenous peroxidase block in 5% H_2O_2. Immunohistochemistry was performed using a mouse monoclonal antibody against human activated complement C5b-9 (Quidel Corporation, San Diego, CA, catalogue #A239). Primary antibody was applied at a concentration of 1:250 in 1X PBS. Biotinylated goat anti-mouse (cat # BA9200) secondary antibody (Vector, Burlingame, CA) was used at a concentration of 1:200. Nickel enhanced diaminobenzidine (DAB; cat # SK4100; Vector) was used to visualize bound antibody. Negative controls were obtained by omission of the primary antibody. Images were obtained with a Zeiss Axioplan microscope equipped with differential interference contrast lenses and a Zeiss Axiocam digital camera.

Fig. S1. Immunohistochemistry for activated complement C5b-9. Tissues from three patients are illustrated. **(A,B)** Postmortem fundus images from patients 1 and 2, respectively. The site illustrated histologically is indicated with an asterisk. **(C)** Tissue from patient 1 is immunopositive for C5b-9 throughout Bruch's membrane and in intercapillary pillars (thin black arrows). Overlying retinal pigment epithelium is hypertrophic, and associated retina demonstrated marked photoreceptor loss (not shown). Complement deposition is also present within the elastica of a choroidal artery (double headed black arrow), as well as within the walls of a choroidal vein (white arrow). **(D)** In patient 2, C5b-9 deposition is present in Bruch's membrane, intercapillary pillars (arrows) and drusen (asterisk). The internal aspect of a choroidal vein is also immunopositive (white arrow). **(E)** Tissue from patient 3, an 86-year old with histologic evidence of early AMD. Activated complement deposition is noted throughout Bruch's membrane, in drusen (asterisks) and in the internal wall of a choroidal vein (white arrow). Bar = 20 μm (C, D); 15 μm (E)



(Figure S1)

Table S1. Summary of sample phenotypes. There are 96 cases and 50 controls.

Phenotype	Cases	Controls
Males (%)	44	54
Never smoked (%)	36	52
Formerly smoked (%)	58	48
Currently smoke (%)	5	0
Mean age (± s.d.) (years)	79±5.2	82±2.2
Age range (years)	65-89	78-87

Table S2. Quality control and informativeness statistics for genotyping. Descriptions of the quality control measures used can be found in the Materials and Methods section.

*For the most part, when SNPs do not match it is due to one of the SNPs not being called. In only 3 out of 4485 comparisons is a mismatch observed, which is equivalent to 99.93% concordance.

Per-chip data quality	
Median call rate per chip	99.1%
Minimum call rate per chip	95.6%
Chips for which gender matches	292 (100%)
Per-individual data quality	
Median call rate per individual	99.1%
Minimum call rate per individual	96.7%
Average number of matches for common SNPs between two chips	30.7/31
Minimum number of matches for common SNPs between two chips*	28/31
Reproducibility	
Xba Repeat concordance (4 replicates)	99.886%
Xba Positive control concordance (2 replicates)	99.870%
10K concordance (3 replicates)	99.767%
Call rate (per-SNP)	
Total number of SNPs	116204
SNPs with 100% call rate	81456
SNPs with call rate between 85% and 100%	33262
SNPs with call rate less than 85%	1486
Locus Polymorphism	
Number of SNPs with no polymorphism observed	8538

Number of SNPs with minor allele frequency < 0.01	3604
Number of SNPs with only heterozygotes observed	19
Number of polymorphic SNPs with no heterozygotes observed	71
Hardy-Weinberg Equilibrium	
Number of SNPs significantly out of equilibrium	231

Table S3. Haplotypes in the haplotype block that harbors the putative disease variant. Haplotype frequencies are estimated using the program PHASE (*S6*, *S7*). The SNPs used to construct the haplotypes are the SNPs from the mapping microarrays found in the 41kb haplotype block defined by the HapMap data. Frequencies are the estimated frequency of each haplotype in the combined case and control population. The two SNPs that show association in the initial analysis are indicated in boldface.

Name	rs2019727	rs10489456	rs3753396	**rs380390**	rs2284664	**rs1329428**	Frequency
N1	A	C	T	C	C	G	0.59
N1	A	C	T	G	C	G	0.0068
N3	A	C	T	G	T	A	0.12
N4	A	T	C	G	C	G	0.15
N5	T	C	T	G	C	A	0.12
N6	T	C	T	G	C	G	0.0071

Table S4. New polymorphisms identified through resequencing. Location of each polymorphism refers to the position on GenBank accession AL049744.8. MAF is minor allele frequency.

Region	Position	Change	Type	MAF	AA Change	rs Number
promoter	120992	A/G	noncoding	0.005263		
promoter	120865	A/G	noncoding	0.010526		
promoter	120546	C/T	noncoding	0.242105		rs3753394
promoter	120410	T/C	noncoding	0.005263		
promoter	120294	A/G	noncoding	0.005263		
intron 1	99391	C/T	noncoding	0.117021		rs511397
exon 2	99242	T/G	nonsynonomous	0.005319	Ser 58 Ala	
exon 2	99230	G/A	nonsynonomous	0.117021	Val 62 Ile	rs800292
intron 2	99114	G/A	noncoding	0.005319		
intron 3	98283	T/C	noncoding	0.005263		
intron 3	98188	T/G	noncoding	0.005263		
exon 4	96315	G/A	nonsynonomous	0.005263	Arg 127 His	
exon 7	87139	A/C	synonomous	0.415789		rs1061147
intron 7	83059	T/C	noncoding	0.005263		
intron 7	82966	G/T	noncoding	0.410526		rs482934
intron 7	82957	A/G	noncoding	0.005263		
exon 9	82232	C/A	nonsynonomous	0.005208	Gln 400 Lys	
exon 9	82226	C/T	nonsynonomous	0.414894	His 402 Tyr	rs1061170
intron 9	58652	T/C	noncoding	0.005319		
exon 10	58516	G/A	synonomous	0.22043		rs2274700

intron 10	58319	A/G	noncoding	0.005319		rs203678
intron 10	58260	C/G	noncoding	0.005319		
intron 10	56838	G/T	noncoding	0.367021		rs203674
exon 12	47084	G/A	nonsynonomous	0.005263	Val 609 Ile	
intron 12	46992	T/G	noncoding	0.005208		
exon 13	45721	A/G	synonomous	0.143617		rs3753396
exon 15	43875	A/G	synonomous	0.005376		
intron 15	40549	A/G	noncoding	0.215054		rs7514261
intron 15	40445	C/T	noncoding	0.021277		
intron 15	40412	G/C	noncoding	0.365591		rs380390
intron 15	40335	G/C	noncoding	0.005319		rs380060
intron 15	40179	C/T	noncoding	0.215054		rs7540032
intron 15	35577	T/G	noncoding	0.005208		rs435628
intron 15	35537	C/A	noncoding	0.357895		rs375046
intron 16	35263	C/T	noncoding	0.005263		rs428060
exon 17	34821	C/T	synonomous	0.026316		
exon 17	34786	G/T	nonsynonomous	0.005263	Ser 890 Ile	rs515299
intron 17	31825	A/C	noncoding	0.005319		
exon 18	31689	G/T	nonsynonomous	0.154255	Glu 936 Asp	rs1065489
intron 18	30673	T/G	noncoding	0.005556		rs385892
intron 18	30547	T/C	noncoding	0.111702		rs16840522
intron 18	30546	A/G	noncoding	0.005319		rs385543
exon 19	30396	G/T	nonsynonomous	0.005319	Val 1007 Leu	rs534399
intron 19	28886	T/C	noncoding	0.154255		rs513699
exon 20	28877	C/T	synonomous	0.154255		rs513729

exon 20	28867	A/T	nonsynonomous	0.015957	Asn 1050 Tyr
intron 20	28592	A/G	noncoding	0.012987	
intron 20	26589	G/C	noncoding	0.005618	
exon 22	25219	C/A	nonsynonomous	0.005556	Pro 1166 Gln
exon 22	25088	C/T	nonsynonomous	0.005618	Arg 1210 Cys

Table S5. Primer sequences used in resequencing.

Region	Forward primer sequence	Reverse primer sequence
promoter	AGAATCGTGGTCTCTGTGTGTGG	AGCAGCTGGTGATATCCTCTGG
promoter	TCAAATGAGAGTGAGCCAGTTGC	CTGTTCACAACGTCCAGTTCTCC
exon 1	GTGGGAGTGCAGTGAGAATTGG	AACTCAACAATGTCAAAAGCC
exon 2	GATAGACCTGTGACTGTCTAGGC	GGCAATAGTGATATAATTCAGGC
exon 3	ACCTCAGCCTCCCAAAGTGC	TGCATACTGTTTTCCCACTCTCC
exon 4	AAGGAGGAGGAGAAGGAGGAAGG	CAGGCTGCATTCGTTTTTGG
exon 5	CCACTCCCATAGAAAAGAATCAGG	ACTTCTTTGCACCAGTCTCTTCC
exon 6	GATAAATCATTTATTAAGCGG	GAACCTTGAACACAGAAAATGC
exon 7	GGATGACTTTGGAGAAGAAGG	TATGAGTTTCGGCAACTTCG
exon 8	TCATCTTCATTAACAAAGACC	AGATCTATTTTGGTCACTTTGC
exon 9	CTTTGTTAGTAACTTTAGTTCG	TTATACACAGTTGAAAAACC
exon 10	GGCAACTCTGAGCTTATTTTCC	AGAGTAGGAAAAGCCTGAATGG
exon 11	CATAGATTATTTTTGTACGG	CAAAACTCCCTTCTTTTCCC
exon 12	ATCTGATGCCCCTCTGTATGACC	ATTCAGTACTCAATACATGTCC
exon 13	CACCATTCTTGATTGTTTAGG	GAATCTCCATAGTAATAAGG
exon 14	CAATGTGTTGATGGAGAGTGG	ATTGAATTATAAGCAATATGC
exon 15	CATTTCAGCGACAGAATACAGG	GTGTGTGTGTGTGTGTGTGC
intron 15	AAGGCAGGAAAGTGTCCTTATGC	GTCAAATTACTGAAAATCACC
exon 16	AACTGTTACACAGCTGAAAAG	GTGGTGATTGATTAATGTGC
exon 17	GGTGGAGGAATATATCTTTGC	ATAGAATAGATTCAATCATGC
exon 18	CGATAGACAGACAGACACCAGAAGG	CAGCTATAATTTCCCACAGCAGTCC
exon 19	GTGTAATCTCAATTGCTACGGCTACC	CAAGTAGCTGGGACTTCAGATGC
exon 20	TAGTTTCATGTCTTTTCCTC	GAATTTTAAGCACCATCAGTC

exon 21	CCAGGACTCATTTCTTTCACC	CTTTCTGACAGAAATATTTGG
exon 22	TGATGTTTCTACATAGTTGG	GGAGTAAAACAATACATAAAAAATG

Supplementary Reference List

S1. H. Matsuzaki *et al.*, *Nature Methods* **1**, 109 (2004).
S2. H. Matsuzaki *et al.*, *Genome Res.* **14**, 414 (2004).
S3. B. Devlin, S. A. Bacanu, K. Roeder, *Nature Genet.* **36**, 1129 (2004).
S4. P. Armitage, G. Berry, J. N. S. Matthews, in *Statistcal Methods in Medical Research*. (Blackwell Scientific Publications, Oxford, 2002) pp. 671-685.
S5. R. R. Hudson, N. L. Kaplan, *Genetics* **111**, 147 (1985).
S6. M. Stephens, P. Donnelly, *Am. J. Hum. Genet.* **73**, 1162 (2003).
S7. M. Stephens, N. J. Smith, P. Donnelly, *Am. J. Hum. Genet.* **68**, 978 (2001).
S8. B. Devlin, N. Risch, *Genomics* **29**, 311 (1995).
S9. G. R. Abecasis, W. O. Cookson, *Bioinformatics* **16**, 182 (2000).
S10. S. B. Gabriel *et al.*, *Science* **296**, 2225 (2002).
S11. A. R. Templeton, E. Boerwinkle, C. F. Sing, *Genetics* **117**, 343 (1987).