Cigarette Smoking Strongly Modifies the Association of *LOC387715* and Age-Related Macular Degeneration

Silke Schmidt,[1,*] Michael A. Hauser,[1,*] William K. Scott,[1] Eric A. Postel,[2] Anita Agarwal,[3] Paul Gallins,[1] Frank Wong,[1] Yu Sarah Chen,[1] Kylee Spencer,[4] Nathalie Schnetz-Boutaud,[4] Jonathan L. Haines,[4] and Margaret A. Pericak-Vance[1]

[1]Center for Human Genetics and [2]Duke Eye Center, Duke University Medical Center, Durham, NC; and [3]Department of Ophthalmology and [4]Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN

We used iterative association mapping to identify a susceptibility gene for age-related macular degeneration (AMD) on chromosome 10q26, which is one of the most consistently implicated linkage regions for this disorder. We employed linkage analysis methods, followed by family-based and case-control association analyses, using two independent data sets. To identify statistically the most likely AMD-susceptibility allele, we used the Genotype-IBD Sharing Test (GIST) and conditional haplotype analysis. To incorporate the two most important known AMD risk factors—smoking and the Y402H variant of the complement factor H gene (*CFH*)—we used logistic regression modeling to test for gene-gene and gene-environment interactions in the case-control data set and used the ordered-subset analysis to account for genetic linkage heterogeneity in the family-based data set. Our results strongly implicate a coding change (Ala69Ser) in the *LOC387715* gene as the second major identified AMD-susceptibility allele, confirming earlier suggestions. This variant's effect on AMD is statistically independent of *CFH* and is of similar magnitude to the effect of Y402H. The overall effect is driven primarily by a strong association in smokers, since we observed significant evidence for a statistical interaction between the *LOC387715* variant and a history of cigarette smoking. This gene-environment interaction is supported by statistically independent family-based and case-control analysis methods. We estimate that *CFH*, *LOC387715*, and cigarette smoking together explain 61% of the population-attributable risk (PAR) of AMD. The adjusted PAR percentage estimates are 20% for smoking, 36% for *LOC387715*, and 43% for *CFH*. We demonstrate, for the first time, that a genetic susceptibility coupled with a modifiable lifestyle factor such as cigarette smoking confers a significantly higher risk of AMD than either factor alone.

Age-related macular degeneration (AMD) is a common complex disorder that affects the central region of the retina (macula) and is the leading cause of legal blindness in white Americans aged >65 years. The prevalence of AMD and its significant morbidity will rise sharply as the population ages. AMD is a clinically heterogeneous disorder with a poorly understood etiology. Population-based longitudinal studies[1,2,3] have established that the presence of extracellular protein and/or lipid deposits (drusen) between the basal lamina of the retinal pigment epithelium (RPE) and the inner layer of the Bruch's membrane is associated with an increased risk of progressing to an advanced form of AMD, either geographic atrophy (dry AMD) or choroidal neovascularization (wet AMD). The presence of large and indistinct (soft) drusen, coupled with RPE abnormalities, is considered to be an early form of the disorder and is often referred to as "age-related maculopathy."

Epidemiologically, AMD is a complex disorder, with contributions of environmental factors and genetic susceptibility.[4] Many environmental and lifestyle factors have been postulated, but the strongest nongenetic risk factor for AMD is clearly cigarette smoking.[5,6] Much progress has been made recently in identifying and characterizing the genetic basis of AMD. In a remarkable example of the convergence of methods for disease-gene discovery, multiple independent research efforts identified the Y402H variant in *CFH* (the complement factor H [MIM 134370]) on chromosome 1q32 as the first major AMD-susceptibility allele.[7–12] Whereas one of the studies was able to pinpoint *CFH* on the basis of a whole-genome association study,[9] most focused on the 1q32 region because it had been implicated consistently by several whole-genome linkage scans. A second genomic region with similarly consistent linkage evidence is chromosome 10q26, which was identified as the single

Received November 28, 2005; accepted for publication March 2, 2006; electronically published March 20, 2006.

Address for correspondence and reprints: Dr. Silke Schmidt, Center for Human Genetics, Duke University Medical Center, Durham, NC 27710. E-mail: silke.schmidt@duke.edu

The novel SNPs discovered in this study were submitted to the National Center for Biotechnology Information's SNP database (dbSNP) and will have "rs" numbers when dbSNP build 126 is released.

* These two authors contributed equally to this work.

Am. J. Hum. Genet. 2006;78:852–864. 0002-9297/2006/7805-0011$15.00

Table 1

Demographic and Clinical Characteristics of the Study Population

Characteristic	Family Data Set	Independent Case-Control Data Set	
		Cases	Controls
No. of multiplex families	140	…	…
No. of affected sibling pairs[a]	169	…	…
No. of other affected relative pairs[a]	37	…	…
No. of singleton families	60	…	…
No. of discordant sibling pairs[b]	158	…	…
No. of individuals	526	610	259
No. (%) with macular findings:			
Grade 1[c]	85 (16.2)	…	193 (74.5)
Grade 2[d]	50 (9.5)	…	66 (25.5)
Grade 3[e]	109 (20.7)	140 (23.0)	…
Grade 4[f]	61 (11.6)	77 (12.6)	…
Grade 5[g]	221 (42.0)	393 (64.4)	…
No. (%) female	348 (66.2)	396 (64.9)	148 (57.1)
Mean (±SD) age at examination (years)	72.6 ± 9.9	76.8 ± 7.7	66.7 ± 8.1

[a] In multiplex families.

[b] In multiplex and singleton families.

[c] No drusen or nonextensive small (<63 μm) drusen without RPE abnormalities.

[d] Extensive small drusen or nonextensive intermediate (≥63 μm and <125 μm) drusen and/or RPE hyper- or hypopigmentation.

[e] Extensive intermediate drusen or any large (≥125 μm), soft drusen, including drusenoid RPE detachment.

[f] Geographic atrophy (area of RPE atrophy with sharp margins, usually visible choroidal vessels, at least 175 μm in diameter).

[g] Exudative AMD, including nondrusenoid RPE detachment, choroidal neovascularization, subretinal hemorrhage or fibrosis, or photocoagulation scar consistent with treatment of AMD.

most-promising region by a recent meta-analysis of published linkage screens.[13]

Two recent studies have suggested specific AMD-susceptibility genes that are located on chromosome 10q26. One study used a combination of family-based and case-control analyses to implicate *PLEKHA1* (pleckstrin homology domain–containing, family A, member 1 [MIM 607772]) and the predicted *LOC387715* gene.[14] However, the association signals for SNPs in these two genes were statistically indistinguishable. The other study used two independent case-control data sets and concluded that the T allele of SNP *rs10490924* in *LOC387715,* a coding change (Ala69Ser) in exon 1 of this poorly characterized gene, was the most likely AMD-susceptibility allele.[15] Both studies reported that the chromosome 10q26 variant confers an AMD risk similar in magnitude to that conferred by the Y402H variant in *CFH*. Here, we describe highly significant association of SNPs in *LOC387715* with AMD. In our data, only SNPs in this gene, including *rs10490924,* explained the strong linkage and association signal in this region. Given a previous report of an effect of cigarette smoking on the linkage findings in the 10q26 region,[16] we tested whether smoking modifies this association. Our data suggest that variant genotypes at *rs10490924* confer a substantially larger AMD risk to cigarette smokers than to nonsmokers. This observation is supported by traditional case-control modeling; by ordered-subset linkage analysis (OSA), incorporating pack-years of cigarette smoking as a covariate; and by family-based association analysis, using a more homogeneous set of families as defined by OSA.

Subjects and Methods

Study Population

As part of an ongoing, large-scale study of genetic and environmental risk factors for AMD, we ascertained patients with AMD, their affected and unaffected family members, and a group of unrelated controls of similar age and ethnic background at two sites in the southeastern United States: Duke University Eye Center (DUEC) and Vanderbilt University Medical Center (VUMC). By use of stereoscopic color fundus photographs, all enrolled individuals were assigned (by E.A.P. and A.A.) one of five different grades of macular findings, as described elsewhere[17,18] and summarized in table 1. Our AMD classification is a modification of the Age-Related Eye Disease Study grading system, with the use of Wisconsin grading-system example slides[19] and the International Classification System[20] as guides. The more severely affected eye was used to classify each individual. Unrelated controls were enrolled

via (i) study advertisement in DUEC- and VUMC-specific newsletters; (ii) recruitment presentations by study coordinators at local retirement communities, the residents of which were likely to obtain health care at DUEC or VUMC; and (iii) AMD-related seminars for the general public sponsored by DUEC or VUMC ophthalmology clinics. Spouses of patients with AMD were also asked to participate as controls. All cases and controls included in this study were white and at least 55 years old. The study protocol was approved by the institutional review boards of the Duke University Medical Center and VUMC, the research adhered to the tenets of the Declaration of Helsinki, and informed consent was obtained from all study participants. Blood samples were collected, and genomic DNA was extracted from whole blood by using the PureGene system (Gentra Systems) on an Autopure LS.

Information about the smoking history of study participants was obtained from a self-administered questionnaire that was formatted to maximize readability for individuals with low vision. However, if participants indicated that they could not complete the form, a project coordinator offered to assist the participants in filling out the questionnaire. Regular cigarette smoking was assessed by two questions: (1) "Have you smoked at least 100 cigarettes in your lifetime?" and (2) "Did you ever smoke cigarettes at least once per week?". Individuals answering "yes" to both questions were asked the average number of cigarettes they smoked per day, the year that they started smoking, whether they had quit smoking, and, if so, what year. This information was used to calculate pack-years of smoking as (cigarettes per day × years smoked)/20 cigarettes per pack. The most general measurement of smoking history was constructed as a binary "ever or never" variable based on a participant's response to question (1) above.

The study population for the analysis presented here included 810 unrelated patients with early (grade 3) or advanced (grades 4 and 5) AMD. Of these, 200 had at least one sampled (affected or unaffected) relative and thus contributed to the family-based association analysis. The remaining 610 patients with AMD who did not have sampled relatives and 259 unrelated controls without AMD (grades 1 and 2) made up an independent case-control data set. Demographic and clinical information for these individuals and their relatives (1,395 individuals total) is shown in table 1.

Genotyping and Linkage and Association Analysis

Previous work by our group[21] and by others[16,22–24] suggested the presence of an AMD-susceptibility locus on chromosome 10q26, with the linkage peak centered at ~122 Mb. To narrow the region most likely to harbor an AMD-susceptibility allele, we genotyped SNPs in the 112–132-Mb interval, which extended 10 Mb on either side of the reported linkage peak. We started with a density of ~1 SNP per 1 Mb and filled in the 117–127-Mb region immediately surrounding the 122-Mb peak with a higher density of 1 SNP per 140 kb on average. All SNPs were selected using SNPSelector software[25] to have approximately equal spacing and minor-allele frequency (MAF) $\geqslant$5%. Genotyping was performed with the TaqMan allelic discrimination assay, by use of either Assays-On-Demand or Assays-By-Design products (Applied Biosystems). For quality-control purposes, two CEPH standards were included in each 96-well plate, and samples from six individuals were duplicated across all plates, with the laboratory technicians blinded to their identities. Analysis required matching quality-controlled genotypes within and across plates and at least 95% genotyping efficiency. The Y402H variant in *CFH* was genotyped by sequencing, as described elsewhere.[7]

After the first round of genotyping and statistical analysis, we applied iterative association mapping[26] to select another set of SNPs in the peak region, defined approximately as the 1-LOD support interval surrounding the peak multipoint LOD score (123.8–126.6 Mb). Our final SNP density was an average of 1 SNP per 33 kb, for a total of 84 SNPs in the 2.8-Mb support interval. Outside this interval, 101 SNPs were genotyped in the 112–132-Mb region, for a total of 185 SNPs. In addition to using SNPSelector,[25] SNPs were identified through resequencing of *LOC387715* and *CUZD1* (CUB and zona pellucida–like domains 1 [HGNC accession number 17937]) in 48–72 unrelated, affected and unaffected individuals. We chose to sequence only *LOC387715* and *CUZD1,* because they were the only genes that harbored SNPs with statistically significant association signals when the false-discovery rate (FDR) was used to correct for multiple testing and controlled at a level of 5%. Individuals were selected for resequencing on the basis of homozygosity at SNP *rs10490924* in *LOC387715* and SNP *rs1891110* in *CUZD1*. The reason for this selection was that individuals homozygous for the risk-associated variant would be most likely to carry the risk allele at a different intragenic sequence variant. Thus, if the observed association were due to linkage disequilibrium (LD) with an untyped causal variant, this should maximize the probability of identifying this variant.

The genotype data were analyzed with MERLIN[27] (for MERLIN software, see Center for Statistical Genetics Web site) to calculate nonparametric two-point and multipoint LOD scores under the exponential model, denoted as LOD*.[28] Allele frequencies were estimated from all genotyped individuals. Parametric affecteds-only heterogeneity LOD scores (HLODs) under a dominant (disease-allele frequency 0.01) or recessive (disease-allele frequency 0.2) model were also computed with MERLIN. To avoid an inflation of linkage evidence due to intermarker LD,[29] we used recently described methods for estimating haplotype frequencies of SNP clusters in high pairwise LD, using a threshold of $r^2 = 0.16$ to define these clusters.[30] The LD pattern in the region of interest was analyzed with the Haploview program,[31] with the generated genotypes from unrelated patients with AMD as the input. Association analysis was applied to all 185 SNPs, by use of the family-based association in the presence of linkage (APL) test[32] (for APL software, see Center for Human Genetics Web site) and standard logistic regression analysis for case-control comparisons with adjustment for age and sex (SAS version 8.02 [SAS Institute]). An additive coding scheme was used, with the SNP model covariate taking on the values −1, 0, and 1 for genotypes 1/1, 1/2, and 2/2, respectively, where 2 is the minor allele in controls. As described above, we divided our total sample into cases contributing to the APL analysis (affected individuals with at least one sampled relative; $n = 200$ families) and an independent sample of cases without sampled relatives ($n = 610$) who were compared with 259 unrelated controls.

As mentioned above, we used the FDR to correct for multiple

testing.[33] Only SNPs that were significant with an FDR of 5% were analyzed with the Genotype-IBD Sharing Test (GIST) method[34] (for GIST software, see Center for Human Genetics Research Web site) to examine which of the most strongly associated SNPs best explained the linkage evidence in the region. We also used the COCAPHASE module of the UNPHASED software package[35] (see UNPHASED page at the Medical Research Council Biostatisics Unit Web site) to perform conditional haplotype analysis. This analysis tested whether conditioning on the risk allele at a particular SNP accounted for the association signal in the region. If the association signal in the region was driven by a single SNP, conditioning on its effect was expected to remove all evidence of association for the remaining SNPs.

Interaction Analysis

We conducted additional analyses to incorporate effects of the two most important known AMD risk factors—smoking and the *CFH* gene. First, we fit a series of logistic regression models to the combined case-control data set (including probands from the family data set) to identify the model that best described (1) the joint effects of *CFH* and *LOC387715* and (2) the joint effects of smoking and *LOC387715*. We followed a recently proposed modeling strategy[36] in which the best-fitting model was derived on the basis of Akaike's information criterion (AIC). The AIC compares different models with a log-likelihood ratio test that is penalized for the number of model parameters to identify the most parsimonious model that adequately fits the data. For each genotype, two model terms were tested: one coding for additive effects at the first, second, or both loci (ADD1, ADD2, and ADDBOTH), using the coding described above, and the other coding for dominance effects (DOM1, DOM2, and DOMBOTH). For example, the ADD1 model contained only model term x_1, coded as -1 for genotype GG at *rs10490924,* as 0 for genotype GT, and as 1 for genotype TT. The ADD2 model included only model term x_2, coded as -1 for genotype TT at Y402H, as 0 for genotype TC, and as 1 for genotype CC. Both models were nested within the larger ADDBOTH model, which included both x_1 and x_2 model terms. This model was again nested within the larger DOMBOTH model, which included x_1, x_2, and two additional terms: z_1, coded as 0.5 for genotype GT at *rs10490924* and as -0.5 for genotypes GG and TT, and z_2, coded analogously for genotypes at Y402H. To test for deviation from joint additive and joint dominance effects (on the logarithmic scale), three additional models (ADDINT, ADDDOM, and DOMINT) were fit. The ADDINT model included the product term $x_1 \times x_2$; the ADDOM model included the product terms $x_1 \times x_2$, $x_1 \times z_2$, and $x_2 \times z_1$; and the DOMINT model included all the above product terms plus the term $z_1 \times z_2$. The model hierarchy was similar for testing the joint effects of *LOC387715* and smoking, except that smoking was coded as 0 for never-smokers and 1 for ever-smokers, and thus only two higher-order models (ADD_SMOKE_INT and DOM_SMOKE_INT) were fit. Since all models are nested within the largest model considered, their respective log-likelihood ratio statistics can be formally compared using the AIC measure. Models for which the AIC differed by <2 units were considered statistically indistinguishable,[36] and the model with the fewest parameters was chosen as the best fitting and most parsimonious model. For example, when the ADDINT model did not provide a substantially better fit than the ADDBOTH model, this was interpreted as lack of evidence for statistical interaction between the two factors. Thus, they each had independent main effects that were multiplicative (additive on the logarithmic scale) such that the best estimate of the odds

Table 2

SNPs Identified in *LOC387715* Sequencing of Individuals Homozygous for *rs1891110* or *rs10490924* Variant, by AMD Grade

	Minor Allele (MAF) in Homozygotes for			
	rs1891110 Variant		*rs10490924* Variant	
Variant[a]	Grade 1	Grade 5	Grade 1	Grade 5
rs10490923	A (.11)	A (.14)	ND	ND
rs2736911	T (.188)	T (.159)	ND	ND
rs10490924	T (.188)	T (.523)	ND	ND
rs17623531	T (.11)	T (.077)	ND	ND
C124204957T (ss51855974)	T (1)	T (.04)	ND	ND
G124204966T (ss51855975)	T (.060)	T (.538)	T (1)	T (1)
A124205188G (ss51855976)	G (.189)	G (.523)	G (.974)	G (.983)
T124205201C (ss51855977)	C (.189)	C (.523)	C (.974)	C (.983)
rs3750848	G (.189)	G (.523)	G (.974)	G (.983)
rs3750846	C (.189)	C (.523)	C (1)	C (1)
rs2736912	T (.167)	T (.182)	ND	ND
rs3750847	T (.189)	T (.523)	T (1)	T (1)
rs10664316	AT (.66)	AT (.94)	AT (1)	AT (1)
T124206387C (ss51855978)	C (.15)	C (.16)	ND	ND
rs7088128	G (.15)	G (.16)	ND	ND

Note.—ND = not determined.

[a] Physical location according to NCBI build 35.

ratio (OR) for being exposed to both factors was the product of the two main-effect ORs.

Our second approach for incorporating AMD-associated covariates was motivated by earlier reports that the 10q26 linkage evidence resulted primarily from families with heavy smokers.[16] As in the previous study, we used an OSA[37] (for OSA software, see Center for Human Genetics Web site) with the family-average of smoking pack-years as a covariate. To avoid an undue influence of zero pack-years values on family averages, pack-years were coded as missing for nonsmokers. With the high-to-low ordering of family-averaged pack-years, OSA tested whether a subset of families with heavy smokers provided significantly greater linkage evidence than did the reference data set, which, in this case, was restricted to families for whom nonmissing covariate values could be computed. Thus, the baseline LOD score was computed for families in which there was at least one affected smoker with pack-years information.

Results

Linkage and Association Analysis

Resequencing of *LOC387715* and *CUZD1* identified 21 known and 23 novel SNPs (tables 2 and 3). Sequencing primers and conditions are available (from M.A.H.) on request. Of these 44 SNPs, 19 were genotyped in our entire data set. Genotypes for all SNPs analyzed here were in Hardy-Weinberg equilibrium in unrelated controls ($P > .01$). We observed high LD ($D' > 0.9$) across a 60-kb region that included a frequent coding SNP in exon 12 of *PLEKHA1* (*rs1045216*), three coding SNPs in *LOC387715* (*rs10490923, rs2736911,* and *rs10490924*), and several additional noncoding SNPs in *PLEKHA1* and *LOC387715,* replicating earlier observations.[15] Notably, the adjacent downstream gene *PRSS11,* also known as "*HTRA1*" (HtrA serine peptidase 1 [MIM 602194]), was not included in this 60-kb region (fig. 1).

In the family-based linkage analysis, a peak multipoint LOD score was obtained at 124.7 Mb (HLOD 3.0 under affecteds-only dominant model; nonparametric LOD* 2.6) (fig. 2). SNP *rs10664316* in *LOC387715* (124.2 Mb) gave a maximum nonparametric two-point LOD score of 3.2. In the case-control analysis, four highly correlated SNPs in *LOC387715,* including the frequent and previously implicated coding change *rs10490924* in exon 1,[15] were very strongly associated with AMD, with logistic regression *P* values on the order of 10^{-8} (table 4). The MAF of these highly correlated SNPs was ~41.7% in cases, very similar to that reported by Rivera et al.,[15] and ~25.8% in controls, somewhat higher than the 19.6% reported by Rivera et al. Within the 60-kb LD block and in the entire 20-Mb interval we screened, association signals of this order of magnitude were observed only for this set of highly correlated SNPs. In particular, the coding SNP in exon 12 of *PLEKHA1* (*rs1045216*) showed substantially weaker evidence for association, both in terms of magnitude and statistical significance (MAF_{cases} 28.2%; $MAF_{controls}$ 36.8%; OR 0.6; $P = .02$). Unlike in the previous reports, we detected a second region of association 400 kb distal to *LOC387715* that included several SNPs in *CUZD1* and an even more distal SNP in *FAM24A* (family with sequence similarity 24, member A [HGNC accession number 23470]). These SNPs, which were in LD with each other but were not in LD with the associated SNPs in *LOC387715* (fig. 1), showed independent evidence for association with AMD risk, although at much lower statistical significance (MAF_{cases} ~55%; $MAF_{controls}$ ~48%; *P* range .0002–.0058). Case-control association results for all 185 SNPs are shown in table 5.

Table 3

SNPs Identified in *CUZD1* Sequencing of Individuals Homozygous for *rs1891110* Variant, by AMD Grade

	Minor Allele (MAF) in Homozygotes for *rs1891110* Variant		
Variant[a]	Grade 1	Grade 3	Grade 5
rs7908196	A (.184)	ND	A (.095)
rs11248321	A (.065)	A (.023)	A (.044)
A124585820G (*ss51855956*)	G (.022)	G (0)	G (0)
C124586887T (*ss51855957*)	T (0)	T (0)	T (.021)
A124586911G (*ss51855958*)	G (.022)	T (0)	G (0)
C124587151T (*ss51855959*)	T (0)	T (0)	T (.021)
A124590492G (*ss51855960*)	G (0)	G (0)	G (.024)
C124595656T (*ss51855961*)	T (.136)	T (0)	T (.136)
C124599157T (*ss51855962*)	T (0)	T (0)	T (.022)
C124599185T (*ss51855963*)	T (0)	T (0)	T (.022)
A124601941T (*ss51855964*)	T (0)	T (0)	T (.022)
G124608497A (*ss51855965*)	A (0)	A (0)	A (.021)
A124612240G (*ss51855966*)	G (0)	G (0)	G (.021)
A124612340G (*ss51855967*)	G (.043)	G (.063)	G (.042)
G124612380A (*ss51855968*)	A (0)	A (0)	A (.021)
T124612649C (*ss51855969*)	C (0)	C (0)	C (.021)
rs4638251	A (.184)	A (.341)	A (.357)
T124628837C (*ss51855970*)	C (0)	C (.045)	C (.024)
C124629013T (*ss51855971*)	T (0)	T (0)	T (.024)
A124629083G (*ss51855972*)	G (0)	G (0)	G (.024)
rs2950355	T (.09)	T (.022)	T (.087)
rs2950354	T (.048)	T (0)	T (0)
rs9423288	T (.048)	T (0)	T (0)
rs10902838	A (.095)	A (0)	A (0)
rs11248323	T (.048)	T (0)	T (0)
T124647361A (*ss51855973*)	A (0)	A (0)	A (.022)
rs11248329	C (.043)	C (.022)	C (.043)
rs4403744	G (.087)	G (.022)	G (.114)
rs1891113	T (.0)	T (.023)	T (0)

Note.—ND = not determined.

[a] Physical location according to NCBI build 35.

GIST Analysis

Seven SNPs with *P* values ≤.0019 in the case-control analysis were statistically significant by the FDR criterion ((0.05 × 7)/185 = 0.0019). These seven SNPs were

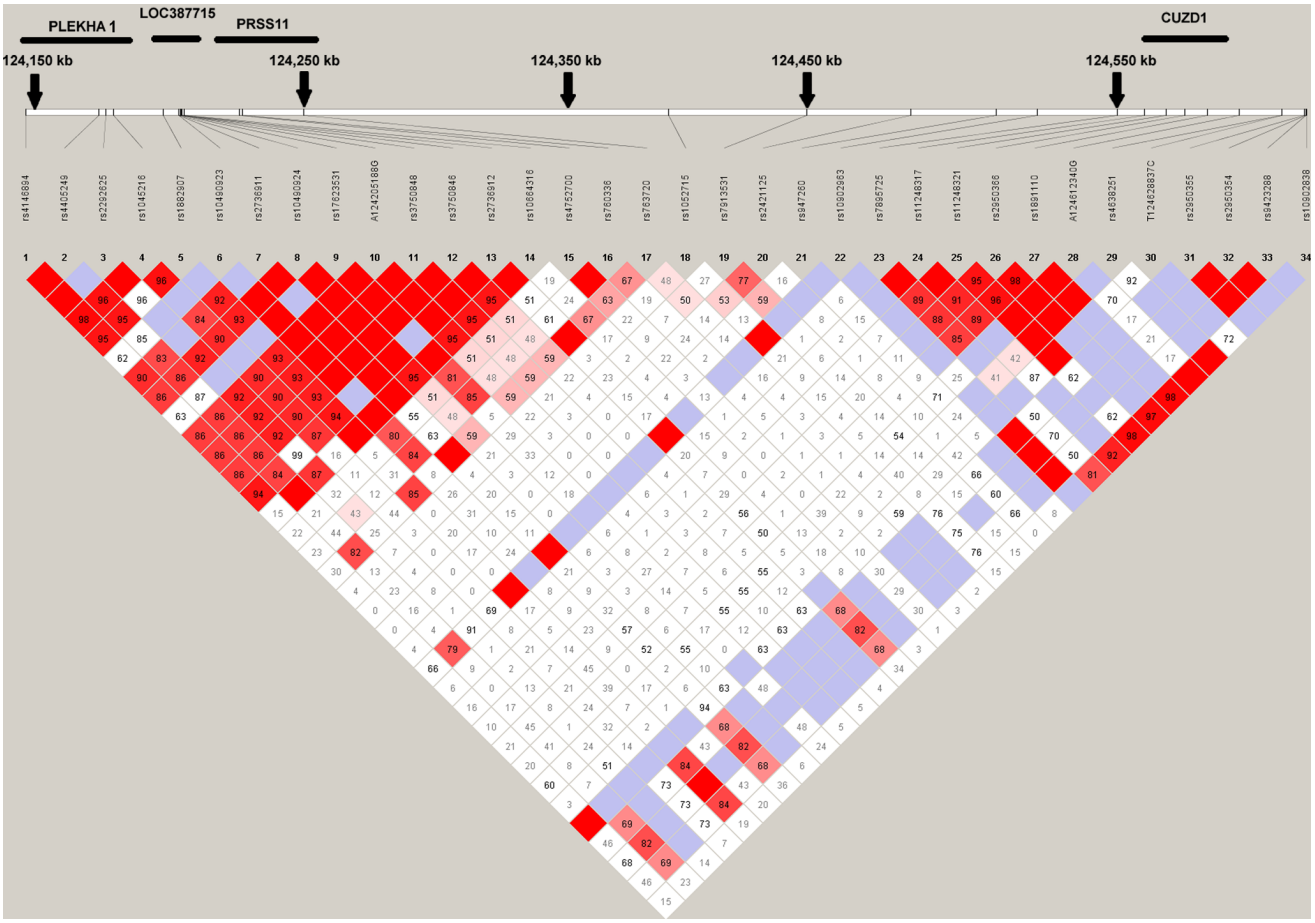

Figure 1 LD pattern in 492.8-kb region from *PLEKHA1* to *CUZD1*. The relative physical position of each SNP is given in the upper diagram, and the pairwise D' between all SNPs is given below each SNP combination. Red-shaded squares indicate $D' > 0.80$. When $D' = 1.0$, no number is given inside the square, and the result is either significant (*red*) or nonsignificant (*blue*) on the basis of the Haploview default definition.[44]

analyzed with GIST to test whether they explained the linkage signal in the region. Under the additive weighting scheme suggested by the case-control analysis,[34] only the four SNPs in *LOC387715* were significant in the GIST analysis (table 4). This suggests that *LOC387715* alone is responsible for the 10q26 linkage evidence.

Conditional Haplotype Analysis

With the combined case-control data set, we used conditional haplotype modeling to identify the statistically most likely AMD-susceptibility variant from among all the SNPs with strong evidence of association. We tested each SNP in table 4, conditioning on the risk allele of the most strongly associated SNP in *CUZD1, FAM24A,* and *LOC387715*. Conditioning on the risk allele at *rs1891110* in *CUZD1, rs10490924* was strongly associated ($P = 7.6 \times 10^{-5}$), whereas none of the other SNPs were significant ($P > .05$). Conditioning on the risk allele at *rs2293435* in *FAM24A, rs10490924* was strongly associated ($P = 7.1 \times 10^{-05}$), whereas none of the other SNPs were significant ($P > .05$). Only conditioning on the risk allele at *rs10490924* fully explained the association signal in the region, such that none of the other SNPs showed any evidence for association ($P > .6$). Thus, this analysis also strongly implicates the T allele at *LOC387715* as a major AMD-susceptibility allele, consistent with the study by Rivera et al.[15]

Interaction Analysis

The results of the AIC modeling strategy (table 6) suggested that the joint action of the Y402H and the *rs10490924* variants was best described by independent multiplicative effects, without statistically significant evidence of dominance effects or epistatic interaction. Previous data presented by our group[38] and by others[39] suggested that the joint action of Y402H and cigarette smoking was also best described by independent multiplicative effects. In contrast, we found strong evidence

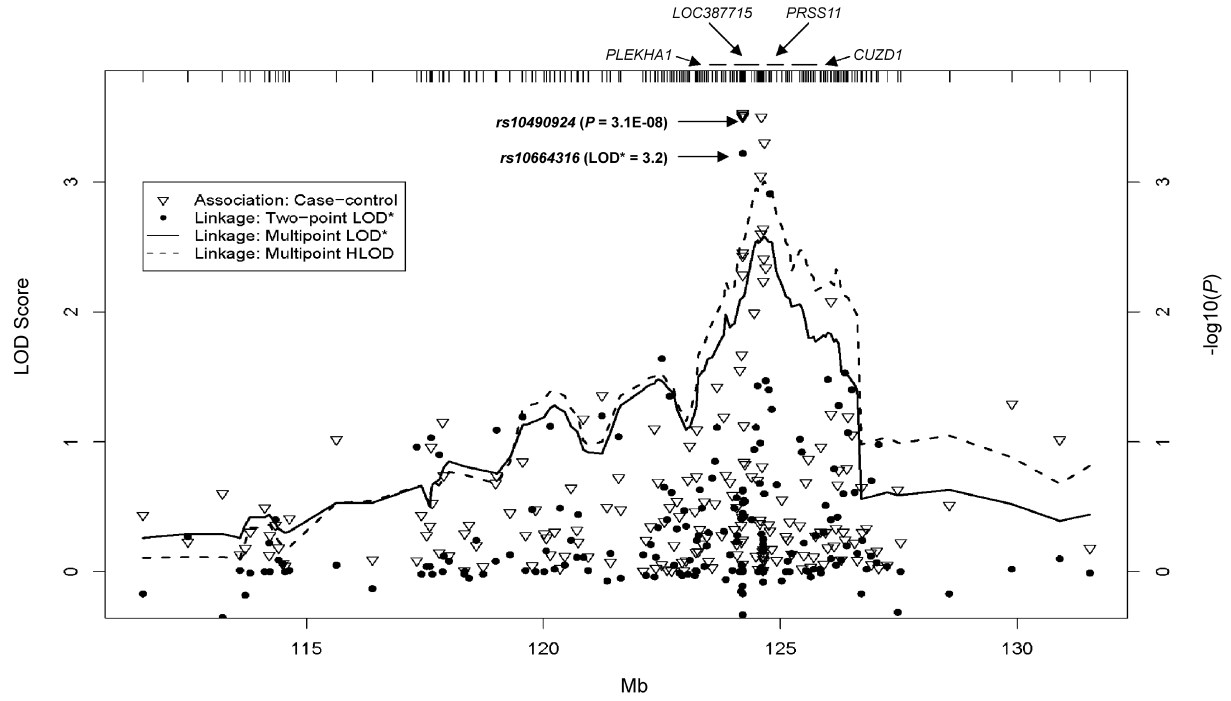

Figure 2 Results of linkage and association analysis. *Left* Y-*axis*, Two-point and multipoint LOD scores. *Right* Y-*axis*, Log_{10}-transformed *P* values from logistic regression of case-control data set, with additive coding (as described in the text) and adjusted for age and sex. For exact *P* values in the 112–132-Mb region that are $<10^{-3}$, see table 4.

of statistical interaction between smoking and genotypes at *rs10490924*. The ADD_SMOKE_INT model provided a significantly better fit to the data, by 5.2 AIC units, compared with the model without this term (table 6). A significant product term with positive regression coefficient for smoking and *rs10490924* (additive coding) indicated more than multiplicative joint effects ($P = .007$). A case-only analysis of *rs10490924* and pack-years of smoking (as a continuous variable) also supported the presence of gene-environment interaction ($P = .05$, adjusted for age and sex). The relative frequency of TT genotypes in affected individuals increased almost linearly with increasing pack-years of smoking, with a corresponding decrease in GG genotype frequencies (fig. 3*A*). This pattern was strikingly similar to results for simulated data when the disease status was generated with a logistic regression model that included a gene-environment interaction term.[40] Genotype frequencies at *rs10490924* were not related to pack-years of smoking in our control sample (fig. 3*B*), confirming that the result for cases was due to gene-environment interaction rather than population correlation of the two factors.

The results of our model-fitting strategy suggested that

Table 4

SNPs in the 112–132-Mb Region with $P \leq .0019$ in the Case-Control Association Analysis

SNP (Risk Allele)	Gene	Type	MAF_{cases}	$MAF_{controls}$	OR_{het} (95% CI)	OR_{hom} (95% CI)	*P*	GIST *P*
rs10490924 (T)	*LOC387715*	Ala69Ser	.410	.259	1.65 (1.12–2.43)	5.73 (3.07–10.71)	3.13×10^{-8}	***.05***
A124205188G (*ss51855976*) (G)	*LOC387715*	Intronic	.412	.257	1.70 (1.14–2.53)	6.04 (3.22–11.33)	1.34×10^{-8}	***.05***
rs3750848 (G)	*LOC387715*	Intronic	.412	.257	1.66 (1.12–2.47)	5.93 (3.16–11.14)	2.38×10^{-8}	***.05***
rs3750846 (G)	*LOC387715*	Intronic	.409	.264	1.52 (1.02–2.24)	5.54 (2.96–10.38)	1.44×10^{-7}	***.05***
rs11248321 (A)	*CUZD1*	Intronic	.558	.490	1.40 (.89–2.22)	2.37 (1.42–3.96)	.0009	.41
rs1891110 (G)	*CUZD1/FAM24B*	5′ UTR (*CUZD1*); Leu2Pro (*FAM24B*)	.568	.489	1.55 (.90–2.68)	3.24 (1.73–6.08)	.0002	.51
rs2293435 (G)	*FAM24A*	5′ UTR	.537	.466	1.51 (.94–2.41)	2.65 (1.53–4.61)	.0005	.26

NOTE.—ORs were adjusted for age and sex and estimated separately for heterozygous (OR_{het}) and homozygous (OR_{hom}) carriers of the minor allele. *P* values are from additive coding of SNP covariate, as described in the text. Significant *P* values for GIST are in bold italics.

Table 5

Case-Control Association Results for 185 SNPs in the 112–132-Mb Region

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

the most parsimonious model for estimating joint ORs for all combinations of smoking history and the risk variants in *CFH* and *LOC387715* was one that included main effect terms for smoking, genotypes GT and TT at *rs10490924,* genotypes TC and CC at Y402H, and two interaction (product) terms: one for smoking and the GT genotype and one for smoking and the TT genotype. The nonsmoker/GG/TT combination was used as the referent group for all ORs (table 7). For nonsmokers, the combination of the TC genotype at Y402H and the GT genotype at *LOC387715* conferred a 1.8-fold increase in AMD risk, and the combination of the CC genotype at Y402H and the TT genotype at *LOC387715* conferred a 10-fold increase in risk. In the absence of either susceptible genotype, the effect of cigarette smoking was not significantly different from 1.0. The effect of Y402H was similar for smokers and nonsmokers, with the most common genotype GG at *rs10490924,* with an OR of ~1.4 for the TC genotype and ~4.5 for the CC genotype. However, in the presence of the susceptible genotype at *rs10490924,* the effect of cigarette smoking was two- to three-times stronger than under a simple multiplicative model. For example, the OR for individuals with the TC genotype at Y402H and the TT genotype at *LOC387715* increased from 3.2 for nonsmokers to 10.8 for smokers, and the OR for individuals with the CC genotype at Y402H and the TT genotype at *LOC387715* increased from 10.2 for nonsmokers to 34.5 for smokers (table 7). Averaged across Y402H genotypes, the presence of the susceptibility allele in *LOC387715* did not confer a significantly increased risk of AMD to nonsmokers (for the GT genotype, OR 1.2, 95% CI 0.6–2.2, $P = 0.59$; for the TT genotype, OR 2.1, 95% CI 0.8–5.1, $P = .12$). However, in smokers, the GT genotype increased the risk 2.7-fold (95% CI 1.5–4.9; $P = .001$), and the TT genotype increased the risk 8.2-fold (95% CI 3.5–19.2; $P < .0001$). Table 8 shows the exposure frequencies for smoking and the four susceptible genotypes at *CFH* and *LOC387715* in controls, as well as the estimated population-attributable risk percentage (PAR%) for each factor, with adjustment for the other two factors.[41] In our data set, smoking accounted for ~20% of AMD, *rs10490924* for ~36%, and Y402H for ~43%, consistent with our previous report.[7] The summary PAR% for all three factors was 61.0%, which is less than the sum of the adjusted PAR% estimates because exposures were not mutually exclusive.

Family-Based Gene-Environment Interaction Analysis

The highly significant association of AMD with *rs10490924* that was observed in the initial case-control analysis was not replicated in the family-based analysis with the APL test. This could be because of the smaller size of our family-based data set or because of between-family heterogeneity. To test the latter possibility, we applied OSA to our multiplex family data set, using the average pack-years of smoking in affected individuals as the OSA covariate (ordered from high to low). OSA indicated that the majority of linkage evidence in the 10q26 region was contributed by only 40 families, with an average of ⩾44 pack-years of smoking (fig. 4). The difference in nonparametric LOD scores between the 90 multiplex families with sufficient information to calculate average smoking pack-years and the 40 families with heavy smokers was significant ($P = .048$) on the basis of 10,000 runs of the OSA permutation test.[37] When the APL analysis was repeated using only multiplex and singleton families that met the "heavy smoking" criterion in affected individuals (family average of ⩾44 pack-years of smoking; 46 families total), the results confirmed the case-control association analysis; the APL *P* value for *rs10490924* and *rs3750848* in *LOC387715* was .02. Three SNPs in other genes also had *P* values

Table 6

Results of Fitting Two-Factor Models by Logistic Regression, Adjusted for Age and Sex

Factor 2 and Model	AIC	AIC Difference[a]
Y402H (*rs1061170*):		
MEAN	936.8	262.4
ADD1	852.0	177.6
ADD2	719.2	44.8
ADDBOTH	675.3	.9[b]
DOM1	851.5	177.1
DOM2	719.1	44.7
DOMBOTH	674.4	0[c]
ADDINT	677.2	2.8
ADDDOM	677.5	3.1
DOMINT	678.5	4.1
Smoking (ever vs. never):		
MEAN	936.8	288.0
ADD	852.0	203.2
SMOKE	708.5	59.7
ADD_SMOKE	654.0	5.2
DOM	851.5	202.7
ADD_SMOKE_INT	648.8	0[c]
DOM_SMOKE_INT	652.4	3.6

NOTE.—Factor 1 is *rs10490924*. Detailed model definitions are given in the "Subjects and Methods" section.

[a] AIC difference is the difference from the AIC of the best-fitting model.

[b] Most parsimonious model.

[c] Best fit.

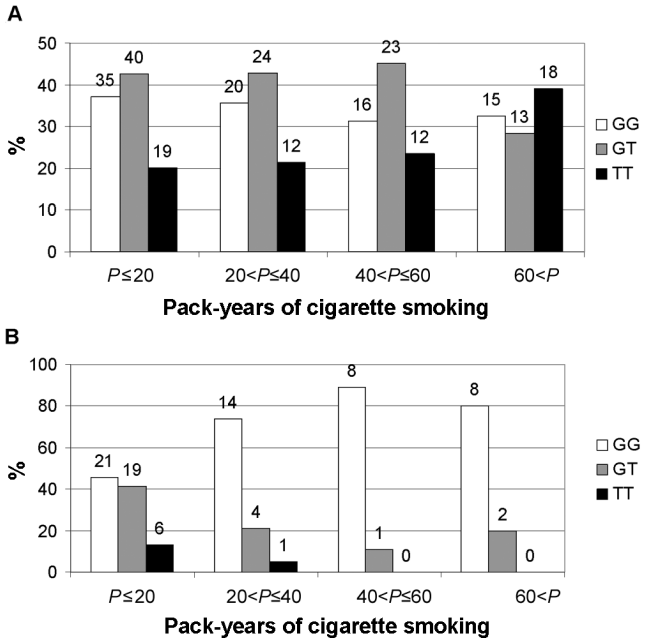

Figure 3 *A,* Genotype frequencies at *rs10490924* in unrelated patients with AMD, by pack-years of cigarette smoking. The number of individuals is shown above each bar. Of 314 reported smokers in this group, 67 did not provide information on smoking duration and/or the average number of packs smoked per day. *B,* Genotype frequencies at *rs10490924* in unrelated controls without AMD, by pack-years of cigarette smoking. The number of individuals is shown above each bar. Of 105 reported smokers in this group, 21 did not provide information on smoking duration and/or average number of packs smoked per day.

of .02: *rs760336* in *PRSS11* (adjacent to *LOC387715*), *rs1052715* in *DMBT1* (deleted in malignant brain tumors 1 [MIM 601969]), and *hcv2917031* in *GPR26* (G protein-coupled receptor 26 [MIM 604847]). Neither SNP had a case-control association $P < .05$ in the overall analysis.

Clinical Subgroup Analysis

It is of great clinical interest to determine whether the *LOC387715* association is present for both geographic atrophy (grade 4) and neovascular AMD (grade 5). Our data set had limited statistical power for the AMD subtype comparison because it included a much smaller number of patients with geographic atrophy than of patients with choroidal neovascularization (table 1), and because smoking-history information was not available for all study participants. Therefore, we did not calculate *rs10490924* genotype frequencies across AMD grades separately for smokers and nonsmokers.

Table 9 shows that the frequency of the T allele was higher in patients with choroidal neovascularization (47.6%) than in patients with geographic atrophy (39.0%), but the clinical interpretation of this finding awaits replication in an independent data set that includes large numbers of patients with each of these AMD subtypes.

Discussion

Using iterative high-density SNP association mapping, we identified a coding change in *LOC387715,* at SNP *rs10490924,* to be the most likely major AMD-susceptibility allele on chromosome 10q26, the second identified to date. We also, for the first time, presented statistical evidence of gene-environment interaction for this variant, suggesting that a genetic susceptibility coupled with a modifiable lifestyle factor, such as cigarette smoking, confers a significantly higher risk of AMD than either factor alone. Genotype frequencies at *rs10490924* were strongly correlated with pack-years of smoking in patients with AMD, consistent with heterogeneity analysis of the genetic linkage data. It is striking that we have observed evidence for gene-environment interaction in two different data sets, using two statistically independent approaches. However, the presence of statistical interaction does not prove biological interaction, and much work remains to be done to identify the molecular mechanism underlying the increased AMD risk.

Our data did not support the previously reported association of AMD with the *GRK5/RGS10* region[14] at ~121 Mb, because the four SNPs (*hcv1809962,*

Table 7

Estimated Joint ORs and 95% CIs for Smoking History (Ever vs. Never), Genotype at *rs10490924* in *LOC387715,* and Genotype at Y402H in *CFH*

rs10490924 Genotype	OR (95% CI) for Nonsmoker and Y402H Genotype			OR (95% CI) for Smoker and Y402H Genotype		
	TT	TC	CC	TT	TC	CC
GG	1.0 (Ref)	1.42 (.87–2.33)	4.56 (2.55–8.18)	.87 (.49–1.57)	1.24 (.57–2.72)	3.99 (1.75–9.12)
GT	1.24 (.65–2.37)	1.76 (.78–3.99)	5.66 (2.36–13.60)	2.48 (1.33–4.59)	3.52 (1.55–8.01)	11.30 (4.75–26.90)
TT	2.24 (.85–5.88)	3.18 (1.08–9.39)	10.21 (3.27–31.94)	7.56 (3.15–18.14)	10.75 (3.92–29.49)	34.51 (11.87–100.32)

NOTE.—Estimated from 511 patients with AMD (grades 3–5) and 208 controls (grades 1–2) with complete data on the three factors of interest and adjusted for age and sex. Ref = referent.

Table 8

Exposure Frequencies for Smoking and Susceptible Genotypes at *LOC387715* and *CFH* in 208 Controls (Grades 1–2) and PAR% for Each Exposure, Adjusted for the Other Two Factors[41]

Exposure	Frequency in Controls (%)	Adjusted PAR%
Cigarette smoking (ever)	49.8	20.0
rs10490924:		
GT	34.6	36.3
TT	9.0	
CFH Y402H:		
TC	53.9	42.6
CC	16.0	

NOTE.—Summary PAR% for all three factors is 61.0%, which is less than the sum of the adjusted PARs because exposures are not mutually exclusive.

rs871196, *rs1537576*, and *rs1467813*) that we genotyped in this region did not demonstrate significant association ($P > .05$). The GIST and conditional haplotype analyses suggested that only *rs10490924*, and the surrounding *LOC387715* SNPs in high LD with it, explained the linkage and association signals in this region. Neither analysis supported SNPs in the nearby *PLEKHA1* and *PRSS11* genes as being responsible for the linkage or association evidence. Consistent with these results, the most-significant single-SNP associations, the highest ORs, and the highest nonparametric two-point LOD score of 3.2 were contributed by SNPs in *LOC387715*. Although we did not resequence the nearby *PLEKHA1* and *PRSS11* genes, we genotyped the vast majority of SNPs examined by the earlier studies in our data set, including all known nonsynonymous coding SNPs spanning the region between and including these two genes. Several SNPs in *CUZD1*, which is not in LD with the *PLEKHA1*/*LOC387715* LD block, gave substantial association signals with logistic regression (smallest $P = .0002$), but allele frequency differences in cases and controls were much less pronounced for these SNPs (MAF_{cases} ~55%; $MAF_{controls}$ ~48%) compared with the SNPs in *LOC387715* (MAF_{cases} ~41%; $MAF_{controls}$ ~26%). In addition, the GIST method and the conditional haplotype analysis suggested that these SNPs did not explain the linkage and association signals in this region.

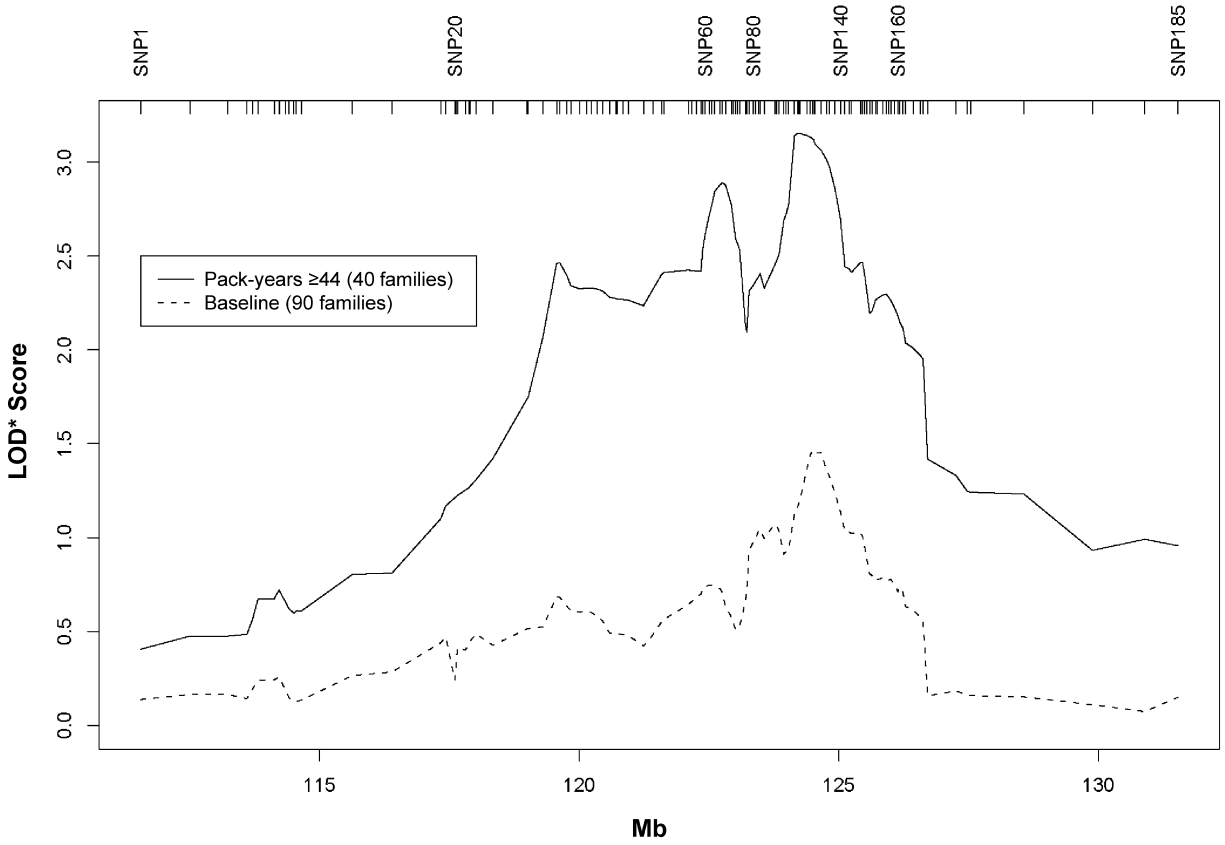

Figure 4 OSA of 90 multiplex AMD-affected families with information on pack-years of cigarette smoking. *Dashed line*, Multipoint LOD* in 90 families. *Solid line*, Multipoint LOD* in 40 families with ⩾44 pack-years, averaged across family members with AMD.

Table 9

MAF and Genotype Frequency at *rs10490924,* by AMD Grade

AMD Grade	MAF	Frequency of *rs10490924* Genotype (No. of Individuals with Genotype)		
		GG	GT	TT
1	.275	.520 (140)	.409 (110)	.071 (19)
2	.307	.479 (57)	.429 (51)	.092 (11)
3	.318	.460 (115)	.444 (111)	.096 (24)
4	.390	.404 (55)	.412 (56)	.184 (25)
5	.476	.295 (177)	.458 (275)	.247 (148)

Note.—Estimated by combining family-based and case-control data sets, including related individuals.

The limitations of any retrospective epidemiologic study apply to our findings, including the potential for recall bias of past exposures. The validity of the summary PAR% estimates depends on the extent to which our case-control data set is representative of a population-based sample of patients with AMD and controls. Since our data set was used to identify the susceptibility variant in *LOC387715*, it is possible that its effect size, and thus its PAR%, was overestimated.[42,43] Independent large population-based studies, ideally collected in a prospective fashion, are needed to confirm the statistical interaction between smoking and *rs10490924* in contributing to AMD and its clinical subtypes and to refine estimates of their individual and joint PAR%s. Because of the high prevalence of smoking, its interaction with genotypes at *LOC387715* induces a strong marginal effect, which explains why our study and others[14,15] detected significant association signals even when smoking-history information was not included in the analysis. The frequencies of the T allele in smokers (40.4%) and nonsmokers (41.4%) with AMD were very similar in the study by Rivera et al.,[15] but were quite different in our study (45.7% in smokers; 38.4% in nonsmokers). The reason for this difference is not clear, but the two samples are also quite different with respect to the reported proportion of smokers (in Rivera et al., 25.2% of the 842 patients with information on smoking history; in our study, 60.3% of 521). The controls in our data set have a lower mean age at examination, and thus some of them may develop AMD in the future. This may explain why the estimated allele frequency of *rs10490924* in our data set (25.8%) is slightly higher than that reported by Rivera et al. (19.6%). It suggests that the inclusion of younger controls in our study is more likely to have created a bias toward the null, rather than an overestimated effect of this SNP on the risk of AMD.

There is currently no biological explanation for the mechanism by which *LOC387715* may increase the risk of AMD. It is not clear whether this statistical association provides further support for the role of the innate immunity system that was highlighted by the recent discovery of *CFH*. *LOC387715* is a two-exon gene that encodes a protein of 107 aa, whose only homologue is a chimpanzee gene with 97% protein identity. No significant matches were found with any known protein motifs. ESTs have been recovered from the placenta and the testis, and this gene has recently been reported to be weakly expressed in the retina.[15]

In summary, we have replicated and refined previous reports implicating a coding change in *LOC387715* as the second major AMD-susceptibility allele. The effect of *rs10490924* appears to be completely independent of the Y402H variant in *CFH*. The joint effect of these two susceptibility genes is consistent with a multiplicative model. Previous data reported by our group[38] and by others[39] suggested that the joint effects of *CFH* and smoking are also consistent with a multiplicative model. In contrast, the effect of *rs10490924* appears to be strongly modified by cigarette smoking. Although the marginal effect of this SNP was strong enough to be detected without incorporation of smoking-history information, an effect modification of a genetic susceptibility by a lifestyle factor such as smoking has important implications for the clinical interpretation of this finding. Our data suggest that the T allele at *rs10490924* may only moderately increase the AMD risk in nonsmokers and likely exerts its strongest effect on heavy smokers. This has the potential to reduce the impact of an AMD-susceptibility allele on the aging population through public health efforts such as smoking-prevention and smoking-cessation programs. Our replication of the 10q26 linkage heterogeneity due to smoking and the consistency of results from multiple statistically independent approaches for assessing gene-environment interaction reported here are unusual in genetic studies of complex human diseases and provide substantial support for our findings. The obvious next task is to examine the function of *LOC387715* in AMD-relevant tissue at the molecular level, particularly in light of the role it may play in smoking-associated pathways leading to this devastating disorder.

Acknowledgments

This study was supported by grants EY12118 (to M.A.P.-V. and J.L.H.) and EY015216 (to S.S.) from the National Eye Institute, National Institutes of Health (NIH), and by grant AG11268 from the National Institute on Aging, NIH (to Dr. Harvey Cohen). It was also supported in part by a General Clinical Research Center award (RR 00095) to Vanderbilt University. We express our appreciation to all the participants and their relatives who generously participated in the study. We thank Dr. Monica de la Paz, Jennifer Caldwell, Molly Klein, Ruth Domurath, Katie Haynes, Valerie Mitchell, Maureen Shaw, and Jason Galloway for participant ascertainment and data management. We also thank the following clinics and clinicians for referring individuals to the study: Southern Ret-

ina (Dr. Charles Harris), Vitreo-Retinal Surgeons (Dr. Michael E. Duan and Dr. Christopher J. Devine), Georgia Retina, and The Retina Group of Washington.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Center for Human Genetics, http://wwwchg.duhs.duke.edu/research/software.html (for OSA and APL software)
Center for Human Genetics Research, http://chgr.mc.vanderbilt.edu/mam/ (for GIST software)
Center for Statistical Genetics, http://www.sph.umich.edu/csg/abecasis/Merlin/ (for MERLIN software)
dbSNP, http://www.ncbi.nlm.nih.gov/SNP/
HUGO Gene Nomenclature Committee (HGNC), http://www.gene.ucl.ac.uk/nomenclature/ (for *CUZD1* [accession number 17937] and *FAM24A* [accession number 23470])
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *CFH, PLEKHA1, HTRA1, DMBT1,* and *GPR26*)
UNPHASED software, http://www.mrc-bsu.cam.ac.uk/personal/frank/software/unphased/

References

1. Klaver CCW, Assink JJM, van Leeuwen R, Wolfs RCW, Vingerling JR, Stijnen T, Hofman A, de Jong PTVM (2001) Incidence and progression rates of age-related maculopathy: the Rotterdam Study. Invest Ophthalmol Vis Sci 42:2237–2241
2. van Leeuwen R, Klaver CC, Vingerling JR, Hofman A, de Jong PT (2003) The risk and natural course of age-related maculopathy: follow-up at 6 1/2 years in the Rotterdam Study. Arch Ophthalmol 121:519–526
3. Klein R, Klein BE, Tomany SC, Cruickshanks KJ (2003) The association of cardiovascular disease with the long-term incidence of age-related maculopathy: the Beaver Dam Eye Study. Ophthalmology 110:1273–1280
4. Klein R, Peto T, Bird A, Vannewkirk MR (2004) The epidemiology of age-related macular degeneration. Am J Ophthalmol 137:486–495
5. Smith W, Assink J, Klein R, Mitchell P, Klaver CC, Klein BE, Hofman A, Jensen S, Wang JJ, de Jong PT (2001) Risk factors for age-related macular degeneration: pooled findings from three continents. Ophthalmology 108:697–704
6. Khan JC, Thurlby DA, Shahid H, Clayton DG, Yates JR, Bradley M, Moore AT, Bird AC (2006) Smoking and age related macular degeneration: the number of pack years of cigarette smoking is a major determinant of risk for both geographic atrophy and choroidal neovascularisation. Br J Ophthalmol 90:75–80
7. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Noureddine M, Gilbert JR, Schnetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA (2005) Complement factor H variant increases the risk of age-related macular degeneration. Science 308:419–421
8. Hageman GS, Anderson DH, Johnson LV, Hancox LS, Taiber AJ, Hardisty LI, Hageman JL, et al (2005) 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
9. 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 (2005) Complement factor H polymorphism in age-related macular degeneration. Science 308:385–389
10. Edwards AO, Ritter R III, Abel KJ, Manning A, Panhuysen C, Farrer LA (2005) Complement factor H polymorphism and age-related macular degeneration. Science 308:421–424
11. Zareparsi S, Branham KE, Li M, Shah S, Klein RJ, Ott J, Hoh J, Abecasis GR, Swaroop A (2005) Strong association of the Y402H variant in complement factor H at 1q32 with susceptibility to age-related macular degeneration. Am J Hum Genet 77:149–153
12. Conley YP, Thalamuthu A, Jakobsdottir J, Weeks DE, Mah T, Ferrell RE, Gorin MB (2005) Candidate gene analysis suggests a role for fatty acid biosynthesis and regulation of the complement system in the etiology of age-related maculopathy. Hum Mol Genet 14:1991–2002
13. Fisher SA, Abecasis GR, Yashar BM, Zareparsi S, Swaroop A, Iyengar SK, Klein BE, Klein R, Lee KE, Majewski J, Schultz DW, Klein ML, Seddon JM, Santangelo SL, Weeks DE, Conley YP, Mah TS, Schmidt S, Haines JL, Pericak-Vance MA, Gorin MB, Schulz HL, Pardi F, Lewis CM, Weber BH (2005) Meta-analysis of genome scans of age-related macular degeneration. Hum Mol Genet 14:2257–2264
14. Jakobsdottir J, Conley YP, Weeks DE, Mah TS, Ferrell RE, Gorin MB (2005) Susceptibility genes for age-related maculopathy on chromosome 10q26. Am J Hum Genet 77:389–407
15. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, Meitinger T, Weber BH (2005) 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
16. Weeks DE, Conley YP, Tsai HJ, Mah TS, Schmidt S, Postel EA, Agarwal A, Haines JL, Pericak-Vance MA, Rosenfeld PJ, Paul TO, Eller AW, Morse LS, Dailey JP, Ferrell RE, Gorin MB (2004) Age-related maculopathy: a genomewide scan with continued evidence of susceptibility loci within the 1q31, 10q26, and 17q25 regions. Am J Hum Genet 75:174–189
17. Schmidt S, Saunders AM, De La Paz MA, Postel EA, Heinis RM, Agarwal A, Scott WK, Gilbert JR, McDowell JG, Bazyk A, Gass JD, Haines JL, Pericak-Vance MA (2000) Association of the apolipoprotein E gene with age-related macular degeneration: possible effect modification by family history, age, and gender. Mol Vis 6:287–293
18. Seddon JM, Ajani UA, Mitchell BD (1997) Familial aggregation of age-related maculopathy. Am J Ophthalmol 123:199–206
19. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L (1991) The Wisconsin age-related maculopathy grading system. Ophthalmology 98:1128–1134
20. Bird AC, Bressler NM, Bressler SB, Chisholm IH, Coscas G, Davis MD, de Jong PT, Klaver CC, Klein BE, Klein R (1995) An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. Surv Ophthalmol 39:367–374
21. Kenealy SJ, Schmidt S, Agarwal A, Postel EA, De La Paz MA, Pericak-Vance MA, Haines JL (2004) Linkage analysis for age-related macular degeneration supports a gene on chromosome 10q26. Mol Vis 10:57–61
22. Majewski J, Schultz DW, Weleber RG, Schain MB, Edwards AO, Matise TC, Acott TS, Ott J, Klein ML (2003) Age-related macular degeneration—a genome scan in extended families. Am J Hum Genet 73:540–550
23. Seddon JM, Santangelo SL, Book K, Chong S, Cote J (2003) A genomewide scan for age-related macular degeneration provides evidence for linkage to several chromosomal regions. Am J Hum Genet 73:780–790
24. Iyengar SK, Song D, Klein BE, Klein R, Schick JH, Humphrey J, Millard C, Liptak R, Russo K, Jun G, Lee KE, Fijal B, Elston RC

(2004) 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
25. Xu H, Gregory SG, Hauser ER, Stenger JE, Pericak-Vance MA, Vance JM, Zuchner S, Hauser MA (2005) SNPselector: a web tool for selecting SNPs for genetic association studies. Bioinformatics 21:4181–4186
26. Oliveira SA, Li YJ, Noureddine MA, Zuchner S, Qin X, Pericak-Vance MA, Vance JM (2005) Identification of risk and age-at-onset genes on chromosome 1p in Parkinson disease. Am J Hum Genet 77:252–264
27. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101
28. Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179–1188
29. Boyles AL, Scott WK, Martin ER, Schmidt S, Li YJ, Ashley-Koch A, Bass MP, Schmidt M, Pericak-Vance MA, Speer MC, Hauser ER (2005) Linkage disequilibrium inflates type I error rates in multipoint linkage analysis when parental genotypes are missing. Hum Hered 59:220–227
30. Abecasis GR, Wigginton JE (2005) Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. Am J Hum Genet 77:754–767
31. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265
32. Martin ER, Bass MP, Hauser ER, Kaplan NL (2003) Accounting for linkage in family-based tests of association with missing parental genotypes. Am J Hum Genet 73:1016–1026
33. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 57:289–300
34. Li C, Scott LJ, Boehnke M (2004) Assessing whether an allele can account in part for a linkage signal: the Genotype-IBD Sharing Test (GIST). Am J Hum Genet 74:418–431
35. Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25:115–121
36. North BV, Curtis D, Sham PC (2005) Application of logistic regression to case-control association studies involving two causative loci. Hum Hered 59:79–87
37. Hauser ER, Watanabe RM, Duren WL, Bass MP, Langefeld CD, Boehnke M (2004) Ordered subset analysis in genetic linkage mapping of complex traits. Genet Epidemiol 27:53–63
38. Scott WK, Schmidt S, Hauser MA, Gallins P, Kwan S, Olson LM, Schnetz-Boutaud N, Spencer KL, Gilbert JR, Agarwal A, Postel EA, Haines JL, Pericak-Vance MA (2005) Interaction of CFH T1277C polymorphism and cigarette smoking in age-related macular degeneration. Paper presented at the American Society of Human Genetics 55th Annual Meeting, Salt Lake City, October 25–29
39. Sepp T, Khan JC, Thurlby DA, Shahid H, Clayton DG, Moore AT, Bird AC, Yates JR (2006) Complement factor H variant Y402H is a major risk determinant for geographic atrophy and choroidal neovascularization in smokers and nonsmokers. Invest Ophthalmol Vis Sci 47:536–540
40. Schmidt M, Hauser E, Martin E, Schmidt S (2005) Extension of the SIMLA package for generating pedigrees with complex inheritance patterns: environmental covariates, gene-gene and gene-environment interaction. Stat Appl Genet Mol Biol 4 (http://www.bepress.com/sagmb/vol4/iss1/art15/) (electronically published June 6, 2005; accessed March 1, 2006)
41. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C (1985) Estimating the population attributable risk for multiple risk factors using case-control data. Am J Epidemiol 122:904–914
42. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33:177–182
43. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet 29:306–309
44. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296:2225–2229