

Set-Based Joint Test of Interaction Between SNPs in the VEGF Pathway and Exogenous Estrogen Finds Association With Age-Related Macular Degeneration

Monique D. Courtenay,¹ William H. Cade,² Stephen G. Schwartz,³ Jaclyn L. Kovach,³ Anita Agarwal,⁴ Gaofeng Wang,² Jonathan L. Haines,^{*,5} Margaret A. Pericak-Vance,^{1,2} and William K. Scott^{1,2}

¹Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, Florida, United States

²John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, Florida, United States

³Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, Florida, United States

⁴Department of Ophthalmology, Vanderbilt University, Nashville, Tennessee, United States

⁵Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, United States

Correspondence: William K. Scott, Hussman Institute for Human Genomics, University of Miami, 1501 NW 10th Avenue, BRB 414, Miami, FL 33136, USA; bscott@med.miami.edu.

Current affiliation: *Case Western Reserve University, Department of Epidemiology and Biostatistics, Cleveland, Ohio, United States.

Submitted: April 1, 2014

Accepted: June 27, 2014

Citation: Courtenay MD, Cade WH, Schwartz SG, et al. Set-based joint test of interaction between SNPs in the VEGF pathway and exogenous estrogen finds association with age-related macular degeneration. *Invest Ophthalmol Vis Sci.* 2014;55:4873–4879. DOI:10.1167/iovs.14-14494

PURPOSE. Age-related macular degeneration (AMD) is the leading cause of irreversible visual loss in developed countries. Its etiology includes genetic and environmental factors. Although *VEGFA* variants are associated with AMD, the joint action of variants within the VEGF pathway and their interaction with nongenetic factors have not been investigated.

METHODS. Affymetrix 6.0 chipsets were used to genotype 668,238 single nucleotide polymorphisms (SNPs) in 1207 AMD cases and 686 controls. Environmental exposures were collected by questionnaire. A set-based test was conducted using the χ^2 statistic at each SNP derived from Kraft's two degree of freedom (2df) joint test. Pathway- and gene-based test statistics were calculated as the mean of all independent SNP statistics. Phenotype labels were permuted 10,000 times to generate an empirical *P* value.

RESULTS. While a main effect of the VEGF pathway was not identified, the pathway was associated with neovascular AMD in women when accounting for birth control pill (BCP) use ($P = 0.017$). Analysis of VEGF's subpathways showed that SNPs in the proliferation subpathway were associated with neovascular AMD ($P = 0.029$) when accounting for BCP use. Nominally significant genes within this subpathway were also observed. Stratification by BCP use revealed novel significant genetic effects in women who had taken BCPs.

CONCLUSIONS. These results illustrate that some AMD genetic risk factors may be revealed only when complex relationships among risk factors are considered. This shows the utility of exploring pathways of previously associated genes to find novel effects. It also demonstrates the importance of incorporating environmental exposures in tests of genetic association at the SNP, gene, or pathway level.

Keywords: age-related macular degeneration, case-control study, epidemiology, statistics, candidate genes

Age-related macular degeneration (AMD) is the leading cause of irreversible visual loss among older adults in developed countries.¹ It has a multifactorial etiology that has yet to be completely elucidated but comprises both genetic and environmental causes. Established genetic risk factors are common variants in *CFH*^{2–6} and *ARMS2/HTRA1*.^{7,8} Additional linkage analysis, candidate gene, and genome-wide association studies (GWAS) have identified numerous other potential genes and pathways that are associated with AMD.

Other nongenetic AMD risk factors include age, female sex, European ancestry, hypertension, and obesity.⁹ Smoking is the most replicated modifiable risk factor (relative risk [RR] = 2).^{10,11} Smoking may modulate the effects of variants on risk of AMD in several genes including *NOS2A*¹² and *ARMS2*.¹³ In contrast, one of the few protective factors identified for AMD is exogenous estrogen. Females who have used exogenous

estrogen in the form of birth control pills (BCPs) or hormone replacement therapy (HRT) have a reported 50% reduction in risk for the disease compared to those who have never taken exogenous estrogens.¹⁴ Recently, effect modification via HRT use and *ARMS2* variants has also been inversely associated with AMD.¹⁵

Up to 65% of the genetic contribution to AMD risk may be explained by known factors.¹⁶ A source of unexplained risk is likely to be gene-gene and gene-environment interactions, which are understudied in AMD. One way to assess the joint effects of multiple variants within a gene or biological pathway is to use a gene-set approach; such methods can have more power to detect genetic risk factors than analysis of individual single nucleotide polymorphisms (SNPs).¹⁷ These methods allow detection of genes or pathways in which multiple variants are weakly associated with disease individually, but

TABLE 1. Description of Total AMD Dataset for 1207 Cases and 686 Controls

Dataset	<i>n</i>	Mean Age* (SD)	Female (%)	Smoked (Ever %)
Total cases	1207	77.6 (7.9)	767 (63.5%)	548 (45.4%)
Total controls	686	70.9 (7.8)	380 (55.4%)	278 (40.5%)
Neovascular subset cases†	720	78.5 (7.6)	456 (63.3%)	349 (48.5%)
Neovascular subset controls†	513	70.6 (7.6)	290 (56.5%)	195 (38%)
Total dataset	1893	75.2 (8.5)	1147 (60.6%)	826 (43.6%)

* Mean age at ascertainment.

† The neovascular AMD subset denotes grade 5 (cases) versus grade 1 (controls).

more strongly associated when considered together.¹⁸ Similarly, accounting for the joint effects of individual SNPs and an interacting environmental factor can increase the power to detect genetic risk factors.¹⁹ However, limited research has been conducted on the best way to incorporate such environmental factors into pathway analysis or on whether including gene-environment associations improves the power of pathway analyses.

In this study, an existing set-based pathway analysis method was modified to include genetic main effects and gene-environment interactions in a self-contained pathway test. Self-contained tests are preferred for testing the hypothesis that a specific pathway (or a small set of pathways) is associated with a phenotype,¹⁷ and the set-test implemented in PLINK (available in the public domain at <http://pngu.mgh.harvard.edu/purcell/plink/>), based on the average of individual test statistics across all variants in the set, is one of the most powerful and flexible self-contained pathway tests available.²⁰ The method used here is an extension of PLINK's set-test that uses the χ^2 statistic at each SNP obtained from the two degree of freedom (2df) joint test of genetic and environmental factors and their interaction.¹⁹ The likelihood ratio test used for the joint effects analysis simultaneously tests the association of a phenotype with the SNP minor allele and the interaction of this allele with an environmental exposure. Considering these joint test statistics instead of single SNP associations accounted for the three possible types of significant disease-SNP associations within a pathway: SNPs with strong main effects, SNPs with strong gene-environment interaction effects, and SNPs with small main and interaction effects that became significant when considered jointly. Accounting for these additional gene-environment interactions may improve the fit of the model across a pathway, increasing statistical power to detect associations with the phenotype.

Here, this method has been used to test the significance of the VEGF signaling pathway while accounting for interaction with environmental factors (smoking, BCP, and HRT). This pathway was chosen due to the discovery that variants in the *VEGF* gene have moderate effects on AMD risk^{16,21,22} and that exposure to these environmental risk factors alters VEGF expression levels.²³⁻²⁵ Despite these observations, little work has focused on finding genes or gene-environment interactions within the VEGF pathway that might influence angiogenesis and AMD development. Using an existing genome-wide association dataset of 1207 cases and 686 controls²⁶ with self-reported environmental risk factor data, this method was used to examine joint effects of smoking or estrogen exposure and variants in the VEGF pathway on AMD risk.

METHODS

Sample Ascertainment

Participant ascertainment occurred at the University of Miami's Bascom Palmer Eye Institute, the Duke University Eye Center,

and the Vanderbilt Eye Institute. Subjects were recruited by physicians, through advertisements in waiting rooms, patient newsletters, recruitment presentations, and AMD project websites. The study was conducted using protocols approved by the Institutional Review Boards at each institution and adhered to the tenets of the Declarations of Helsinki. Written informed consent was obtained from each participant.

All subjects received a comprehensive eye exam and AMD grade by a study ophthalmologist at one of the three ascertainment centers (Supplementary Table S1). Grading was conducted using color fundus photographs and the CARMS five-step grading scale,²⁷ a modified version of the Age-Related Eye Disease Study (AREDS) grading system.²⁸ This scale incorporated example slides from the Wisconsin Grading System²⁹ and used the International Classification System as a guide.³⁰ Regular exchanges of photographs were conducted with a subset of the sample to assess concordant grading across sites.³¹ Participants were assigned an AMD grade of 1 through 5 based on the more severely affected eye. Grades 1 and 2 were designated unaffected "controls" while grades 3, 4, and 5 were considered affected "cases." All included subjects were of European descent and at least 55 years old.

Detailed histories of environmental exposures including smoking history and exogenous estrogen use were collected via a self-administered questionnaire. All participants were asked to indicate if they had ever smoked at least 100 cigarettes in their lifetime. Individuals who answered "Yes" were considered "Ever Smokers" and were asked the age or year they started and if and when they had quit (Table 1). Female participants were also asked to indicate whether they had ever used birth control (including pills, shots, or implants) or HRT (including pills, creams, or patches). Participants who answered "Yes" to either of these questions were considered "Ever BCP" or "Ever HRT" users and were also prompted to provide the duration of their usage. However, information on usage was too sparse for use in this analysis (Table 2).

Genotyping and Quality Control

Deoxyribonucleic acid from each participant was extracted from whole blood using PUREGENE methods (Gentra Systems, Minneapolis, MN, USA). Detailed genotyping and quality control methods have been described previously.²⁶ Briefly, the dataset (1216 cases and 715 controls) was genotyped on the Affymetrix 6.0 GeneChip Human Mapping 1 Million SNP array set (Affymetrix, Inc., Santa Clara, CA, USA) using established protocols. The Birdseed program in APT BirdSuite 1.8.6³² was used for genotype calling. Deoxyribonucleic acid samples from the Centre d'Etude du Polymorphisme Humain (CEPH) reference families were also represented on each 96-well plate for quality assurance. We retained 668,238 SNPs for association analysis with SNP call rate $\geq 95\%$, minor allele frequency > 0.02 , and $P > 1 \times 10^{-6}$ for tests of Hardy-Weinberg equilibrium in controls.

For sample quality control, samples with genotype call rates $\leq 95\%$ were excluded from the analysis ($n = 58$). Also,

TABLE 2. Description of the Female-Only AMD Subset Who Had a History of Exogenous Estrogen Use

Dataset	<i>n</i>	Mean Age* (SD)	HRT Use† (Ever %)	BCP Use‡ (Ever %)
Total cases	531	77.2 (8.2)	269 (51.8%)	102 (19.8%)
Total controls	242	71.5 (7.6)	162 (67.8%)	98 (41%)
Neovascular subset cases§	320	78.1 (7.6)	145 (46.5%)	54 (17.6%)
Neovascular subset controls§	180	71.3 (7.3)	121 (68.4%)	80 (44.7%)
Total dataset	773	75.4 (8.4)	431 (55.8%)	200 (25.9%)

* Mean age at ascertainment.

† 20 women were missing data on HRT use.

‡ 15 women were missing data on BCP use.

§ The neovascular AMD subset denotes grade 5 (cases) versus grade 1 (controls).

participants with a sex inconsistency between the reported sex and the heterozygous genotypes of X-linked SNPs were removed ($n = 13$). GRR³⁵ removed 22 related individuals ($IBD > 0.40$), and sample outliers identified by principal components analysis in EIGENSTRAT (available in the public domain at <http://genepath.med.harvard.edu/~reich/EIGENSTRAT.htm>) were removed ($n = 11$). The final dataset consisted of 1893 unrelated individuals (1207 cases and 686 controls, Tables 1, 2).

Statistical Analysis

To determine if the VEGF signaling pathway was associated with AMD in our dataset, PLINK's set-test function³⁴ was implemented in R (available in the public domain at <http://www.R-project.org/>) and conducted on all SNPs from the AMD GWAS within the VEGF pathway as defined by the KEGG PATHWAY database³⁵ (Supplementary Table S2) ($n = 1058$). Unlike traditional overrepresentation pathway analysis, this set-based method aggregated all SNP associations and generated a *P* value for the entire set based on the null hypothesis of no association between the pathway and AMD.¹⁷ Steps for the set-test were as follows. First, SNPs within each gene (± 20 kb) were selected and to reduce noise, variants in strong linkage disequilibrium (LD) ($R^2 > 0.8$) were removed. Next, single SNP associations were obtained from a logistic regression model adjusted for sex, age, and smoking. The set χ^2 test statistic was then calculated as the mean of all independent SNP statistics. Next, a permutation test was conducted in order to assess significance. Ten thousand copies of the dataset were generated by permuting the phenotype labels but holding the genotypes and covariate structure constant. This was performed with the BiasedUrn permutation procedure. The set *P* value was calculated as the number of times the 10,000 pathway χ^2 test statistics from the permutations exceeded the observed one from our dataset. This analysis was conducted for the overall dataset (grades 3, 4, and 5 versus grades 1 and 2) and a neovascular AMD subset of the phenotypic extremes consisting only of neovascular AMD cases (grade 5) and controls with no retinal changes (grade 1).

Next, gene-environment interactions were incorporated into the pathway set-test. The set-test was modified to accept the two degree of freedom (2df) joint test of genetic and environmental factors¹⁹ as the input statistic instead of the single SNP association statistics. To generate this statistic, a likelihood ratio test for each SNP was conducted via logistic regression between a full model with the SNP and environmental exposure main effects and the SNP-exposure interaction and a null model excluding the SNP and interaction terms. This joint effect pathway analysis was conducted for each environmental exposure of interest in the total AMD dataset (grades 3, 4, and 5 versus grades 1 and 2) and the neovascular AMD (grade 5 versus grade 1) subset. Significance was also assessed by repeating the analysis 10,000 times with the

BiasedUrn permutation procedure. For the smoking analysis, both participant age and sex were also controlled for in the logistic regression models, while participant age and smoking history were accounted for in the women-only analysis of HRT and BCPs. Stratified analyses were conducted for significant results from the joint test to see if genetic effects were restricted to particular exposure strata.

Type 1 Error and Power Simulations

To evaluate the validity of this set-based approach, GenomeSIMLA³⁶ was used to simulate 1000 null case-control datasets with sample sizes and prevalence of BCP use and smoking based on the females in the project data ($n = 514$ cases and 239 controls, BCP prevalence = 27%, smoking prevalence = 51%). GenomeSIMLA generated genomes with markers, allele frequencies, and linkage disequilibrium patterns based on SNPs from the Affymetrix 500K SNP array. The null datasets were generated with environmental effects (BCP $RR_c = 0.5$ and smoking $RR_c = 2.0$) but no significant genetic effects or gene-environment interactions. Under the null hypothesis of no association with AMD, the VEGF signaling pathway ($n = 444$ SNPs) was tested with the pathway-based 2df joint test incorporating BCP use and controlling for smoking. The type 1 error was estimated as the number of null replicates with significant pathway associations and compared to the expected 5% error rate.

Next we tested the power of this method to detect a genetic effect under a range of possible gene-environment interactions with an environmental factor modeled on the known effects of smoking in our dataset. In the first simulation, five independent SNPs within the pathway were chosen to have a genetic main effect ($RR_g = 1.2$). This served as the comparison group to assess any gain in power when accounting for environment and gene-environment interactions. Next, main genetic and environmental effects were kept constant ($RR_g = 1.2$ for SNPs and smoking $RR_c = 2.0$), while the number and magnitude of SNPs with gene-environment interactions were varied. Using QUANTO,³⁷ we calculated that individual SNPs in a GWAS of this size would need a main effect of $RR > 1.69$ to be detected and implicate a pathway containing that SNP ($\alpha = 5 \times 10^{-8}$). Therefore, we explored a range of main and moderate interaction effects that fell below this threshold. For each condition, 1000 datasets were generated to mimic our entire smoking dataset ($n = 894$ cases and 525 controls, smoking prevalence = 51%), and power was estimated as the proportion of replicates with significant pathway associations ($\alpha = 0.05$). Summing χ^2 statistics within a set (rather than averaging them) to obtain a cumulative statistic is used for gene-based methods like SKAT,³⁸ VEGAS,³⁹ and RASSOC.^{40,41} Therefore, we also evaluated the performance of our set-based 2df joint test when χ^2 statistics were summed instead of averaged.

TABLE 3. Pathway Set-Test Results for Main Effects Within the VEGF Signaling Pathway (894 SNPs) for 1207 Cases and 686 Controls Who Had a History of Smoking and/or Exogenous Estrogen Use

Individuals With Data On:	Total AMD Dataset*	Neovascular AMD Subset†
	P Value	P Value
Smoking	0.27	0.083
HRT‡	0.59	0.12
BCPs‡	0.70	0.2

* The total AMD dataset denotes grades 3, 4, and 5 (cases) versus grades 1 and 2 (controls).

† The neovascular AMD subset denotes grade 5 (cases) versus grade 1 (controls).

‡ Dataset includes women only.

RESULTS

The KEGG PATHWAY database version of the VEGF signaling pathway was used to determine the pathway's genes ($n = 67$ genes, 894 SNPs) and the specific functional arms, or subpathways (Supplementary Table S2). Three subpathways were identified: proliferation ($n = 35$ genes, 624 SNPs), migration ($n = 21$ genes, 181 SNPs), and survival/permeability ($n = 18$ genes, 164 SNPs). Although cell survival and permeability are distinct physiological functions, there was significant overlap between the two gene networks that necessitated their combination for analysis. Migration and survival/permeability also shared several genes at the onset of both subpathways ($n = 11$). VEGFA, its receptor KDR, and receptor-associated proteins SH2D2A and SHC2 were included in the total VEGF pathway analysis but not in individual subpathways.

First we evaluated the association between the VEGF signaling pathway and total AMD or neovascular AMD by conducting a pathway analysis of genetic effects only. The pathway-based set-test of SNP main effects was not significant at the Bonferroni corrected P value ≤ 0.017 after correcting for three exposure datasets (Table 3). Next, we incorporated environmental exposure into our analysis. The total VEGF signaling pathway was not associated with AMD or neovascular AMD when we conducted the pathway-based 2df joint test for gene-environment interaction with HRT use or smoking. However, there was a significant pathway association ($P = 0.017$) for neovascular AMD in women after taking into account BCP use (Table 4). Stratification by BCP use found this association restricted to women who had taken BCPs ($P = 0.0096$).

The three VEGF subpathways were analyzed separately in women with neovascular AMD to see if joint genetic-BCP effects were clustered within a specific functional arm of the pathway. The VEGF subpathway responsible for proliferation was also associated with neovascular AMD. Again, stratification by BCP use showed that the association was only in women who had ever taken BCPs (Table 4).

Finally, gene-based 2df joint tests were conducted for genes within the proliferation subpathway for the BCP dataset to see which genes may be driving the association (Supplementary Table S3). One gene, *MAPK3*, did not have any SNPs that were included on the Affymetrix 6.0 chipset and could not be analyzed. Six genes within the proliferation subpathway were nominally associated ($P < 0.05$) with AMD for women with a history of BCP use (Table 4): *NEATC1*, *NEATC2*, *NEATC4*, *PLCG1*, *PLCG2*, and *PPP3R1*. All but *PPP3R1* were associated with AMD only in BCP users, consistent with the previous pathway and subpathway results. *NEATC1* and *PPP3R1* were

significant in both the SNP main effect and the 2df joint set-tests, suggesting that their associations may be driven by strong main effects. However, four genes on different chromosomes, *NEATC2*, *NEATC4*, *PLCG1*, and *PLCG2*, were significant only for the 2df joint set-tests and would not have been detected as associated with AMD without incorporating the environmental interaction.

Simulations generated a type I error rate of 4.3%, which is slightly under the expected frequency of 5%. These results establish that our test is valid under the null hypothesis for the types of genetic and environmental effects studied here. For power simulations, there was between 12% and 40% power to detect a pathway association in our dataset when genetic main effects were restricted to two to five independent SNPs (Supplementary Table S4). This was consistent with what we expected since we also found no main effects in our analysis. Next, for four or five independent SNPs, approximately 1% of the total pathway size, there was 80% to 94% power to detect an associated pathway after accounting for a modest gene-environment interaction with all SNPs ($RR_{ge} = 1.3$). This highlights the strength of the set-based 2df joint test and most likely explains the new association we observed in the VEGF signaling pathway, in which independent SNPs in several genes demonstrated significant interactions with BCP use. For lower numbers of interacting SNPs, the required gene-environment interaction to detect a pathway effect remained high at $RR_{ge} = 2.4$ to 3.9. This indicated that the lack of significant results for HRT or smoking in our dataset may reflect either a true lack of significant interaction across the pathways, or low numbers of interacting SNPs that did not reach this high threshold. There were no differences in power for averaging versus summing the χ^2 statistics in our dataset (data not shown).

DISCUSSION

Variants with large main effects can be readily detected with standard single-locus analysis, and these methods have identified variants that explain up to 65% of the genetic effect underlying AMD.¹⁶ The remaining genetic factors may be genes or pathways containing multiple variants with small effect, or variants that interact with the environment (or each other) to modulate disease. Such effects are detectable through methods that consider the joint action of variants within a gene or across a pathway and the interaction of those variants with nongenetic factors. This approach successfully identified the association of multiple genes within the VEGF pathway with AMD, but only when considering history of BCP use in women. This illustrates that some potential AMD biological mechanisms may be revealed only when complex relationships among risk factors are considered.

To perform this pathway-based gene-environment interaction analysis, we extended an existing gene-based test to incorporate a 2df joint test of genetic main effect and gene-environment interaction as the input test statistic. Unlike standard set-based methods that consider main effects only, the joint effects analysis simultaneously tested the effects of a genotype and its interaction with an environmental exposure. This method does not test gene-environment interaction directly but instead considers the joint effect of genetic variants and environmental exposures, summarized across the gene or pathway. This approach maintained the correct type I error rate and had adequate power to detect a pathway effect in our dataset.

For our analysis of the VEGF signaling pathway we found novel pathway and gene associations with neovascular AMD, after accounting for BCP use, that were not detected in the same tests of genetic main effects alone. Although set-based

TABLE 4. Two Degree of Freedom Joint Test Results for the Pathway- and Gene-Based Analysis for the VEGF Signaling Pathway for 514 Female Cases and 239 Female Controls Who Had a History of Smoking and Birth Control Pill Use

	No. of Genes (SNPs)	Total AMD Dataset*	Neovascular AMD Subset†		
		2df Joint Test P Value, 514 Cases/239 Controls	2df Joint Test P Value, 307 Cases/179 Controls	Ever BCP Users P Value, 102 Cases/98 Controls	Never BCP Users P Value, 412 Cases/144 Controls
Total VEGF signaling pathway					
Total VEGF pathway	67 (894)	0.31	0.017	0.0096	0.67
VEGF subpathways					
Proliferation	35 (624)	0.25	0.029	0.010	0.79
Migration	21 (181)	0.77	0.31	-	-
Surv/perm	18 (164)	0.86	0.50	-	-
Significant proliferation genes‡					
<i>NEATC1</i> chr18:77160326-77289323	1 (16)	0.15	0.0023	0.0097	0.078
<i>NEATC2</i> chr20:50003494-50159258	1 (45)	0.068	0.015	0.038	0.079
<i>NEATC4</i> chr14:24836145-24848810	1 (15)	0.19	0.039	0.016	0.41
<i>PLCG1</i> chr20:39766161-39804357	1 (5)	0.026	0.0094	0.0056	0.24
<i>PLCG2</i> chr16:81812899-81991899	1 (82)	0.39	0.036	0.016	0.67
<i>PPP3R1</i> chr2:68405989-68479651	1 (4)	0.21	0.041	0.27	0.032

Bold typeface indicates a significant association ($P < 0.05$).

* The total AMD dataset denotes grades 3, 4, and 5 (cases) versus grades 1 and 2 (controls).

† The neovascular AMD subset denotes grade 5 (cases) versus grade 1 (controls).

‡ Significant genes had $P < 0.05$ in either the original set-test or 2df joint test.

tests cannot tell us the direction of the associations, stratification by environmental exposure found these genetic effects only in women who had a history of BCP use. Further analysis showed that the total pathway association seemed to be driven by six genes within the proliferation subpathway. This subpathway was the largest of the three analyzed, and it also contained several nested pathways, including the initiation of the mitogen-activated protein kinase and calcium signaling pathways. All six genes were directly connected at the beginning of the proliferation subpathway, near the beginning of the calcium signaling pathway. They were also in the gene network that leads to the metabolism of proinflammatory arachidonic acid; this is notable because inflammation is an important reported feature of AMD pathogenesis.

In addition, several genes identified in this study have been implicated in ocular diseases or are connected to other AMD effectors. The results found within the proliferation subpathway mirrored known biological interactions for several genes. Ethinyl estradiol, the most common synthetic estrogen found in BCPs, induces expression of *PLA2G12B*,⁴² *PLCG2*,⁴³ and *PPP3R1*.⁴² Known estrogen receptor modulators, tamoxifen and raloxifene, can also alter the expression of *NEATC1*^{44,45} and *PPP3R1*.⁴² *PLCG1* has an established interaction with Ginsenoside Rg1,⁴⁶ a drug known for exerting estrogen-like activity. Also, *PLCG2* interacts with dehydroepiandrosterone (DHEA),⁴⁷ a chemical precursor that is converted to estrogen, and exemestane,⁴⁸ an aromatase inhibitor that blocks estrogen synthesis. In contrast to these findings, estrogens (including BCP use) have previously been inversely associated with AMD in women.¹⁴ However, this overall effect ignores potential interactions with genes. Estrogen may still increase AMD risk in women with some genetic backgrounds. Reconciling the conflicting results may require biological studies to explain how genes within this subpathway interact with estrogen to influence proliferation and modulate AMD risk.

Furthermore, the VEGF pathway's *PLA2* and *PLCG* genes that are present in the eye are activated by light,⁴⁹ and excessive sunlight exposure has also been implicated in AMD

pathogenesis. Other genes within the PLA2 family, notably sPLA2-IIa, have been implicated in ocular surface physiology and inflammation.^{50,51} How these genes work together to influence AMD is unknown and will require molecular and functional studies.

Nevertheless, anti-VEGF drugs that target VEGFA and its receptor are the main treatments for combating the destructive angiogenesis characterizing neovascular AMD. However, not all patients respond similarly to this treatment. In the future, pathway-based approaches targeting multiple genes might lead to more effective therapies for attenuating this pathway in AMD. For example, a *PLCG1* modulator, c-Cbl, has already been found to inhibit angiogenesis from VEGF stimulation.⁵² The identification of associated gene networks within this study will allow more focused fine-mapping of the complex interactions that jointly influence the risk of developing AMD; the results bear examination in additional large datasets to evaluate the reproducibility of this effect. A limitation of this study is that we were able to test in only one dataset, so results may have been generated by chance; but if reproducible, these results point toward new genes for biological studies and possible treatment targets.

This study illustrates the utility of exploring pathways of previously associated genes to find novel gene associations. While the *VEGFA* gene itself was previously associated with increased risk of AMD, this study implicates the VEGF signaling pathway more broadly, and suggests that this effect may be restricted only to those women who used BCPs. These results also demonstrate the importance of incorporating environmental exposures in tests of genetic association in complex diseases, at the SNP, gene, or pathway level. Exploring more complex models of gene-environment interactions may find associated variants that account for remaining heritability and more closely align with disease etiology.

Acknowledgments

The authors thank all of the study participants. We also thank James Hicks for his editorial assistance.

Supported by National Institutes of Health Grant R01-EY012118 and partially supported by resources from National Institutes of Health Center Grant P30-EY014801 and an unrestricted grant to the University of Miami from Research to Prevent Blindness, New York, New York.

Disclosure: **M.D. Courtenay**, None; **W.H. Cade**, None; **S.G. Schwartz**, None; **J.L. Kovach**, None; **A. Agarwal**, None; **G. Wang**, None; **J.L. Haines**, None; **M.A. Pericak-Vance**, None; **W.K. Scott**, None

References

- Resnikoff S, Pascolini D, Etya'ale D, et al. Global data on visual impairment in the year 2002. *Bull World Health Organ*. 2004; 82:844-851.
- Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308:385-389.
- Haines JL, Hauser MA, Schmidt S, et al. Complement factor H variant increases the risk of age-related macular degeneration. *Science*. 2005;308:419-421.
- Hageman GS, Anderson DH, Johnson LV, 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 U S A*. 2005;102:7227-7232.
- Edwards AO, Ritter R III, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. *Science*. 2005;308:421-424.
- Zarepari S, Branham KE, Li M, et al. Strong association of the Y402H variant in complement factor H at 1q32 with susceptibility to age-related macular degeneration. *Am J Hum Genet*. 2005;77:149-153.
- Rivera A, Fisher SA, Fritsche LG, 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*. 2005; 14:3227-3236.
- Yang Z, Camp NJ, Sun H, et al. A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. *Science*. 2006;314:992-993.
- Klein R, Peto T, Bird A, Vannewkirk MR. The epidemiology of age-related macular degeneration. *Am J Ophthalmol*. 2004; 137:486-495.
- Vingerling JR, Hofman A, Grobbee DE, de Jong PT. Age-related macular degeneration and smoking. The Rotterdam Study. *Arch Ophthalmol*. 1996;114:1193-1196.
- Smith W, Assink J, Klein R, et al. Risk factors for age-related macular degeneration: pooled findings from three continents. *Ophthalmology*. 2001;108:697-704.
- Ayala-Haedo JA, Gallins PJ, Whitehead PL, et al. Analysis of single nucleotide polymorphisms in the NOS2A gene and interaction with smoking in age-related macular degeneration. *Ann Hum Genet*. 2010;74:195-201.
- Schmidt S, Hauser MA, Scott WK, et al. Cigarette smoking strongly modifies the association of LOC387715 and age-related macular degeneration. *Am J Hum Genet*. 2006;78:852-864.
- Snow KK, Cote J, Yang W, Davis NJ, Seddon JM. Association between reproductive and hormonal factors and age-related maculopathy in postmenopausal women. *Am J Ophthalmol*. 2002;134:842-848.
- Edwards DR, Gallins P, Polk M, et al. Inverse association of female hormone replacement therapy with age-related macular degeneration and interactions with ARMS2 polymorphisms. *Invest Ophthalmol Vis Sci*. 2010;51:1873-1879.
- Fritsche LG, Chen W, Schu M, et al. Seven new loci associated with age-related macular degeneration. *Nat Genet*. 2013;45: 433-439, 439e1-439e2.
- Holmans P. Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. *Adv Genet*. 2010;72:141-179.
- Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet*. 2010;11: 843-854.
- Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered*. 2007;63:111-119.
- Gui H, Li M, Sham PC, Cherny SS. Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's Disease dataset. *BMC Res Notes*. 2011;4:386.
- Haines JL, Schnetz-Boutaud N, Schmidt S, et al. Functional candidate genes in age-related macular degeneration: significant association with VEGF, VLDLR, and LRP6. *Invest Ophthalmol Vis Sci*. 2006;47:329-335.
- Yu Y, Bhangale TR, Fagerness J, et al. Common variants near FRK/COL10A1 and VEGFA are associated with advanced age-related macular degeneration. *Hum Mol Genet*. 2011;20: 3699-3709.
- Conklin BS, Zhao W, Zhong DS, Chen C. Nicotine and cotinine up-regulate vascular endothelial growth factor expression in endothelial cells. *Am J Pathol*. 2002;160:413-418.
- Cullinan-Bove K, Koos RD. Vascular endothelial growth factor/vascular permeability factor expression in the rat uterus: rapid stimulation by estrogen correlates with estrogen-induced increases in uterine capillary permeability and growth. *Endocrinology*. 1993;133:829-837.
- Kwekel JC, Burgoon LD, Burt JW, Harkema JR, Zacharewski TR. A cross-species analysis of the rodent uterotrophic program: elucidation of conserved responses and targets of estrogen signaling. *Physiol Genomics*. 2005;23:327-342.
- Naj AC, Scott WK, Courtenay MD, et al. Genetic factors in nonsmokers with age-related macular degeneration revealed through genome-wide gene-environment interaction analysis. *Ann Hum Genet*. 2013;77:215-231.
- Seddon JM, Sharma S, Adelman RA. Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology*. 2006;113:260-266.
- The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. *Control Clin Trials*. 1999;20:573-600.
- Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology*. 1991;98:1128-1134.
- Bird AC, Bressler NM, Bressler SB, et al. An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. *Surv Ophthalmol*. 1995;39: 367-374.
- De La Paz MA, Guy VK, Abou-Donia S, et al. Analysis of the Stargardt disease gene (ABCR) in age-related macular degeneration. *Ophthalmology*. 1999;106:1531-1536.
- Korn JM, Kuruvilla FG, McCarroll SA, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet*. 2008;40:1253-1260.
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics*. 2001;17:742-743.
- Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27-30.
- Edwards T, Bush W, Turner S, et al. Generating linkage disequilibrium patterns in data simulations using genomeSIM-

- LA. In: Marchiori E, Moore JH, Rajapakse JC, eds. *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*. Berlin, Heidelberg: Springer; 2008:24-35.
37. Gauderman W, Morrison JM. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. 2006. Available at: <http://hydra.usc.edu/gxe>. Accessed October 31, 2011.
 38. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*. 2011;89:82-93.
 39. Liu JZ, McRae AF, Nyholt DR, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet*. 2010; 87:139-145.
 40. Kinnamon DD, Hershberger RE, Martin ER. Reconsidering association testing methods using single-variant test statistics as alternatives to pooling tests for sequence data with rare variants. *PLoS One*. 2012;7:e30238.
 41. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet*. 2009;5: e1000384.
 42. Fong CJ, Burgoon LD, Williams KJ, Forgacs AL, Zacharewski TR. Comparative temporal and dose-dependent morphological and transcriptional uterine effects elicited by tamoxifen and ethynylestradiol in immature, ovariectomized mice. *BMC Genomics*. 2007;8:151.
 43. Boverhof DR, Burgoon LD, Williams KJ, Zacharewski TR. Inhibition of estrogen-mediated uterine gene expression responses by dioxin. *Mol Pharmacol*. 2008;73:82-93.
 44. Ball LJ, Levy N, Zhao X, et al. Cell type- and estrogen receptor-subtype specific regulation of selective estrogen receptor modulator regulatory elements. *Mol Cell Endocrinol*. 2009; 299:204-211.
 45. Bebo BF Jr, Dehghani B, Foster S, Kurniawan A, Lopez FJ, Sherman LS. Treatment with selective estrogen receptor modulators regulates myelin specific T-cells and suppresses experimental autoimmune encephalomyelitis. *Glia*. 2009;57: 777-790.
 46. Zong Y, Ai QL, Zhong LM, et al. Ginsenoside Rg1 attenuates lipopolysaccharide-induced inflammatory responses via the phospholipase C-gamma1 signaling pathway in murine BV-2 microglial cells. *Curr Med Chem*. 2012;19:770-779.
 47. Werle-Schneider G, Wolfelschneider A, von Brevern MC, et al. Gene expression profiles in rat liver slices exposed to hepatocarcinogenic enzyme inducers, peroxisome proliferators, and 17alpha-ethinylestradiol. *Int J Toxicol*. 2006;25:379-395.
 48. Miki Y, Suzuki T, Hatori M, et al. Effects of aromatase inhibitors on human osteoblast and osteoblast-like cells: a possible androgenic bone protective effects induced by exemestane. *Bone*. 2007;40:876-887.
 49. Kashiwagi T, Meyer-Rochow VB, Nishimura K, Eguchi E. Light activation of phospholipase A2 in the photoreceptor of the crayfish (*Procambarus clarkii*). *Acta Neurobiol Exp (Wars)*. 2000;60:9-16.
 50. Wei Y, Epstein SP, Fukuoka S, Birmingham NP, Li XM, Asbell PA. sPLA2-IIa amplifies ocular surface inflammation in the experimental dry eye (DE) BALB/c mouse model. *Invest Ophthalmol Vis Sci*. 2011;52:4780-4788.
 51. Wei Y, Pinhas A, Liu Y, Epstein S, Wang J, Asbell P. Isoforms of secretory group two phospholipase A (sPLA2) in mouse ocular surface epithelia and lacrimal glands. *Invest Ophthalmol Vis Sci*. 2012;53:2845-2855.
 52. Meyer RD, Husain D, Rahimi N. c-Cbl inhibits angiogenesis and tumor growth by suppressing activation of PLCgamma1. *Oncogene*. 2011;30:2198-2206.