Rare Complement Factor H Variant Associated With Age-Related Macular Degeneration in the Amish

Joshua D. Hoffman,¹ Jessica N. Cooke Bailey,² Laura D'Aoust,¹ William Cade,³ Juan Ayala-Haedo,³ Denise Fuzzell,² Renee Laux,² Larry D. Adams,³ Lori Reinhart-Mercer,³ Laura Caywood,³ Patrice Whitehead-Gay,³ Anita Agarwal,⁴ Gaofeng Wang,³ William K. Scott,³ Margaret A. Pericak-Vance,³ and Jonathan L. Haines^{1,2}

¹Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee, United States

²Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, United States

³John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, Florida, United States

⁴Department of Ophthalmology and Visual Sciences, Vanderbilt University, Nashville, Tennessee, United States

Correspondence: Jonathan L. Haines, Department of Epidemiology and Biostatistics, Institute for Computational Biology, 2-529 Wolstein Research Building, 2103 Cornell Road, Case Western Reserve University, Cleveland, OH 44106, USA; jlh213@case.edu.

Submitted: November 27, 2013
Accepted: May 20, 2014

Citation: Hoffman JD, Cooke Bailey JN, D'Aoust L, et al. Rare complement factor H variant associated with age-related macular degeneration in the Amish. *Invest Ophthalmol Vis Sci*. 2014;55:4455–4460. DOI:10.1167/iivs.13-13684

PURPOSE. Age-related macular degeneration is the leading cause of blindness among the adult population in the developed world. To further the understanding of this disease, we have studied the genetically isolated Amish population of Ohio and Indiana.

METHODS. Cumulative genetic risk scores were calculated using the 19 known allelic associations. Exome sequencing was performed in three members of a small Amish family with AMD who lacked the common risk alleles in complement factor H (*CFH*) and *ARMS2/HTRA1*. Follow-up genotyping and association analysis was performed in a cohort of 973 Amish individuals, including 95 with self-reported AMD.

RESULTS. The cumulative genetic risk score analysis generated a mean genetic risk score of 1.12 (95% confidence interval [CI]: 1.10, 1.13) in the Amish controls and 1.18 (95% CI: 1.13, 1.22) in the Amish cases. This mean difference in genetic risk scores is statistically significant ($P = 0.0042$). Exome sequencing identified a rare variant (P503A) in *CFH*. Association analysis in the remainder of the Amish sample revealed that the P503A variant is significantly associated with AMD ($P = 9.27 \times 10^{-13}$). Variant P503A was absent when evaluated in a cohort of 791 elderly non-Amish controls, and 1456 non-Amish cases.

CONCLUSIONS. Data from the cumulative genetic risk score analysis suggests that the variants reported by the AMDGene consortium account for a smaller genetic burden of disease in the Amish compared with the non-Amish Caucasian population. Using exome sequencing data, we identified a novel missense mutation that is shared among a densely affected nuclear Amish family and located in a gene that has been previously implicated in AMD risk.

Keywords: age-related macular degeneration, linkage analysis, rare variant, exome sequencing, risk score analysis

Age-related macular degeneration is the leading cause of blindness in individuals aged older than 65 years in the developed world.^{1,2} Age-related macular degeneration is a progressive neurodegenerative disease that results in central vision loss. Vision loss caused by AMD is generally divided into two categories. Non-neovascular or “dry” AMD is characterized by the presence of a broad range of abnormalities in the RPE layer. Clinical characteristics include the presence of drusen, hyperpigmentation or hypopigmentation of the RPE, and in later stages, geographic atrophy, which results from RPE layer cell death.³ Neovascular or “wet” AMD occurs when new blood vessels form behind the macula, leaking fluid into the space separating the Bruch’s membrane and the RPE layer.³

Age-related macular degeneration has been shown to be associated with age, race, smoking, and cardiovascular risk factors including obesity and hypertension.^{4–8} Genetic studies have identified common variants with strong associations in complement factor H (*CFH*) and *ARMS2/HTRA1*, as well as multiple loci of smaller effect, in populations of European

descent.^{9–14} A recent meta-analysis completed by the AMDGene Consortium showed that the most strongly associated variants in *CFH* and *ARMS2/HTRA1*, in combination with 17 other loci that reached genome-wide significance, account for less than 65% of the total genetic contribution to AMD.¹⁵ While this shows success in common variant analysis, there is still a gap in the total genetic variation yet to be explained for this complex disease. Some of this missing heritability is thought to lie within rare variants of large effect.^{16–19} To address the deficit in knowledge of rare variants influencing AMD, we exome sequenced individual members of a nuclear family who represent a subset of the Ohio-Indiana Amish population.

The Amish are a genetically and culturally isolated founder population descended from Swiss and German Anabaptists who emigrated from Western Europe to North America in the 1700s and 1800s.²⁰ Individuals of the Amish community typically marry within the faith and observe a strict lifestyle, resulting in a community that is more genetically and environmentally homogeneous than the surrounding population. Due to the

FIGURE 1. Pedigree of nuclear family chosen for exome sequencing. *Clinical diagnoses were available for a-78, a-79, and a-80 at time of exome sequencing. Grading was carried out according to the modified AREDS scale.

intermarriage within the community and relatively recent founder event, this population may also be enriched for rare genetic variation. These factors, in addition to an extensive family record available through the Anabaptist Genealogy Database (AGDB), make the Amish a unique and valued population for genetic studies.^{21–23}

MATERIALS AND METHODS

Subjects

Amish subjects were selected from the Collaborative Aging and Memory Project (CAMP), an ongoing sample collection of Amish individuals living in the United States in Ohio and Indiana, described elsewhere.^{24,25} Construction and maintenance of the AGDB is covered under an institutional review board-approved protocol at the National Institutes of Health (Leslie Biesecker, Principal Investigator). The affection status of AMD was assigned based on self-report questionnaire responses where subjects were asked if they had ever been diagnosed with AMD by a physician. A subset of 73 participants (42 cases, 31 controls) received a follow-up clinical exam by a retinal specialist. The self-report dataset included 128 individuals with AMD, 728 individuals without AMD, and 294 with no self-report information, and all individuals were connected into a single 13-generation pedigree based on an “all common paths” query of the AGDB using pedigree and genealogy software PedHunter 2.0 (available in the public domain at <http://www.ncbi.nlm.nih.gov/CBBresearch/Schaffer/pedhunter.html>).²⁶

Non-Amish subjects were ascertained from the Duke University Eye Center, the Vanderbilt Eye Institute, and the Bascom Palmer Eye Institute at the University of Miami Miller School of Medicine. Participants were examined by a retinal specialist and graded on a severity scale derived from the Age-Related Eye Disease Study (AREDS), described elsewhere.^{27,28} Grades were given on a scale of 1 to 5 where grades 1 and 2 were assigned to controls, grade 3 represented early AMD, and grades 4 and 5 represented late AMD (geographic atrophy and choroidal neovascularization, respectively). Amish participants who received a clinical diagnosis were graded using the same

criteria. The final non-Amish dataset contained 1732 cases, 943 controls, and 310 unknown samples.

As part of ongoing AMD family studies, a single nuclear Amish family (family 1) with multiple individuals affected with AMD (three of eight siblings affected) was noted. Affected members lacked risk alleles at the Y402H locus in *CFH* and at the A69S locus in *ARMS2*. The three affected siblings (a-78, a-79, a-80; Fig. 1) were clinically evaluated to confirm disease status. The other siblings were not available for clinical examination at the time of exome sequencing, although a self-reported diagnosis was completed. Subject a-78 was diagnosed with bilateral choroidal neovascularization (grade 5); subjects a-79 and a-80 were diagnosed with large drusen in both eyes (grade 3). Details of clinical grades and age at time of diagnosis for the ascertained nuclear family are included in Figure 1. All procedures followed the tenets of the Declaration of Helsinki and were approved by the institutional review boards of the University of Miami Miller School of Medicine and Vanderbilt University. Informed consent was obtained from all research subjects involved in this study.

Cumulative Genetic Risk Score Analysis

Genetic risk scores were calculated using the 19 variants and their effect sizes reported by the AMDGene consortium’s meta-analysis.¹⁵ Genotyping was performed using a commercial genotyping platform (Sequenom MassARRAY; Sequenom, Inc., San Diego, CA, USA). Subjects that presented with missing genotypes at any of the 19 loci were excluded from the analysis. Each variant was weighted and multiplied by the number of risk alleles present at each locus, where the SNP weight is equal to the individual SNP beta-estimate divided by the sum of the beta-estimates across all 19 loci (formula 1). Risk scores from each of the 19 loci were summed to give the cumulative genetic risk score per person. Two-sided *t*-tests assuming unequal variance were calculated in the Amish case-control group and non-Amish versus Amish case group.

$$\begin{aligned} \text{Cumulative genetic risk score} = & w_1 \text{Genotype}_1 \\ & + w_2 \text{Genotype}_2 + \dots \\ & + w_{19} \text{Genotype}_{19} \end{aligned} \quad (1)$$

TABLE 1. Genotyped Samples Utilized per Analysis Step

	Full Dataset	Linkage Scan	Exome Sequencing	Rare Variant Association	Risk Score Analysis
Amish					
Cases	128	95	3	95	88
Controls	728	1	-	653	559
Unknowns	294	-	-	225	-
Non-Amish					
Cases	1732	-	-	1456	1573
Controls	943	-	-	791	841
Unknowns	310	-	-	124	-

Sample numbers shown are postquality control.

Exome Sequencing

For all samples, DNA was extracted from whole blood by the Vanderbilt University DNA Resources core and the John P. Hussman Institute for Human Genomics (HIHG) at the University of Miami Miller School of Medicine using DNA extraction methods (PureGene; Gentra Systems, Minneapolis, MN, USA). The three clinically examined members of family 1 were selected for exome sequencing. Exome capture was performed using a commercial exome sequencing kit (Agilent SureSelect All Exon V5 kit; Agilent Life Sciences, Santa Clara, CA, USA). Exon enriched libraries were sequenced on commercial equipment (Illumina HiSeq 2000; Illumina, Inc., San Diego, CA, USA). Sequence capture and high-throughput sequencing were completed at the HIHG. Paired end reads were generated and mapped to the human reference genome (version hg19 from UCSC) using genome mapping software Burrows-Wheeler Aligner (BWA; available in the public domain at <http://bio-bwa.sourceforge.net/>).²⁹ Duplicate reads were marked using Picard tools and local realignment around insertions and deletions was performed using the genome analysis toolkit (GATK).³⁰ Realignment was performed using the base quality score recalibration walker followed by variant calling with the UnifiedGenotyper walker. Variant quality score recalibration was completed using an additional 172 sets of exomes from Amish individuals who were not members of nuclear family 1. Annotation of nucleotide variants was performed using the SeattleSeq Annotation server (available in the public domain at <http://snp.gs.washington.edu/SeattleSeqAnnotation138/>).⁴ Variant filtration was completed as follows: variants found in the 1000 Genomes Project database, NHLBI exome sequencing project exome variation server database, or dbSNP137 database were excluded. In addition, single nucleotide variants that were not missense, nonsense, splice junction, or frameshift causing mutations found within exon boundaries were excluded. Genes not known to be associated with AMD were also excluded. Variants found to be either homozygous or heterozygous and shared by all three affected siblings were retained. Variant confirmation was performed using standard forward and reverse Sanger sequencing practices on the ABI 3730xl (Applied Biosystems, Foster City, CA, USA).

Targeted Genotyping

Genotyping was performed using a commercial genotyping platform (Sequenom, Inc.). Independent Sequenom pools were genotyped and evaluated for the rare variant data, and the 19 loci used to calculate the cumulative genetic risk score. A genotyping efficiency threshold of 95% was used in both datasets to determine a valid Sequenom assay. See Table 1 for final sample sizes used in each step of the analysis.

Association and Linkage Analysis

Variants identified in the nuclear family were evaluated for association with AMD in the full Amish dataset using the modified quasi-likelihood score statistic (MQSL; available in the public domain at <http://www.stat.uchicago.edu/~mcpeek/software/MQSL/index.html>).³¹ This score statistic is comparable with the χ^2 test with the exception that it estimates the variance on point estimates of the allele frequencies in cases and controls while taking into account the correlation between related individuals in a pedigree. This effectively allows all degrees of relationships to be included in the association analysis. Kinship and inbreeding coefficients for all possible relationships in the 13-generation pedigree were calculated using the MQSL recommended program KinInbcoef (available in the public domain at <http://www.stat.uchicago.edu/~mcpeek/software/KinInbcoef/index.html>). The MQSL analysis was performed using option 1, which allows for individuals with genotype, but no phenotype data to contribute to the analysis. An assumed disease prevalence of 10% was selected from published estimates of the prevalence of AMD in individuals aged older than 60 years in population samples.^{1,2,32} The nuclear family was genotyped as part of the full Amish dataset and was not excluded from the analysis. A subanalysis was performed in which only those cases with a clinical exam completed by a retinal specialist were analyzed.

Parametric linkage analysis was performed using the entire dataset genotyped on an array cartridge (Affymetrix Human SNP Array 6.0; Affymetrix, Inc., Santa Clara, CA, USA). The pedigree was divided into smaller, more computationally feasible pedigrees using subpedigree software PedCut (available in the public domain at <ftp://mga.bionet.nsc.ru/pedcut/>) with a bit size threshold of 24.³³ Parametric heterogeneity logarithm of the odds (HLOD) scores were calculated under affected-only dominant and recessive models, assuming incomplete penetrance in MERLIN (available in the public domain at <http://www.sph.umich.edu/csg/abecasis/merlin/index.html>).³⁴ Under the dominant model, we specified penetrance values of 0 for no copies of the disease allele and 0.0001 for one or two copies of the disease allele. Under the recessive model, we specified penetrances of 0 for zero or one copy of the disease allele and 0.0001 for two copies of the disease allele. The disease allele frequency was set to 10% and marker allele frequencies corrected for relatedness were estimated from all genotyped Amish individuals. Multipoint linkage analysis was performed on a 7-megabase region surrounding *CFH*. Linkage disequilibrium (LD) pruning for multipoint was performed using the HapMap CEPH (Centre d'Etude du Polymorphisme Humain) samples with a pairwise r^2 cutoff < 0.16 . In addition to standard multipoint linkage analysis, subanalyses specifying liability classes modeled on the effect sizes of the rare and common *CFH* alleles were examined. These included using the common SNP data derived from the

TABLE 2. Demographics for Amish and Non-Amish Datasets

	Median Age at Exam (IQR)	Percent Female
Amish		
Cases	82 (80–86)	70%
Controls	80 (76–84)	57%
Non-Amish		
Cases	78 (72–83)	63%
Controls	70 (64–76)	58%

IQR, interquartile range.

commercial chip (Affymetrix 6.0 Genome-Wide Human SNP array; Affymetrix, Inc.), assuming only *CFH* Y402H as a risk allele, assuming only *CFH* P503A as a risk allele, and assuming both *CFH* Y402H and P503A as risk alleles. Detailed methods for quality control and linkage analysis procedures can be found elsewhere.³⁵

To examine LD between the novel variant and the common Y402H variant, we extracted the most distantly related Amish subjects genotyped in the pedigree. For this analysis, we used a maximum pairwise relationship cutoff between second and third cousins, resulting 168 individuals available for an LD calculation. Linkage disequilibrium was measured using the whole genome data analysis toolset PLINK (available in the public domain at <http://pngu.mgh.harvard.edu/~purcell/plink/>).³⁶

RESULTS

Clinical Versus Self-Report of AMD

Comparing clinical diagnoses to self-report of AMD status in individuals for whom both were available, we observed positive and negative predictive values of 89% and 90%, respectively (Table 3), indicating that self-report of AMD status is a good proxy for AMD diagnosis in this population sample.

FIGURE 2. Variant filtration procedures used and the resultant number of variants after each stage. Variants passing quality control procedures are used as the initial filtration starting point. ESP refers to the NHLBI grand opportunity exome sequencing project.

TABLE 3. AMD Self-Reported Pilot Study

Clinical Exam	Self-Reported AMD, <i>n</i>	Self-Reported No AMD, <i>n</i>	Total
Diagnosed with AMD	39	3	42
No AMD	5	26	31
Total	44	29	73

Cumulative Risk Score Analysis

In our cumulative genetic risk score analysis, we observed a mean risk score of 1.12 (95% CI: 1.10, 1.13) in the Amish controls and 1.18 (95% CI: 1.13, 1.22) in the Amish cases ($P = 0.0042$). We also observed a mean risk score of 1.14 (95% CI: 1.13, 1.16) in 841 non-Amish Caucasian controls and 1.31 (95% CI: 1.30, 1.32) in 1573 non-Amish Caucasian cases. When comparing the Amish cases with the non-Amish cases, we see a significant decrease in genetic risk score ($P < 0.00001$).

Exome Sequencing and Linkage Analysis

Exome sequencing was performed on the three individuals in family 1. We generated on average 4.6 million reads per sample. Of those reads, 75% were on target with a depth of coverage of $\times 10$ or higher. After extensive quality control, we examined these data for rare variants in known AMD genes that might explain AMD in this family. After variant filtration procedures, we identified a single nonsynonymous mutation in *CFH* that predicts a proline to alanine amino acid change at position 503 (P503A; Fig. 2).

Case control analysis using self-reported affection status in the Amish sample population identified a significant association of AMD with P503A ($P = 9.27 \times 10^{-13}$). Results were consistent in the subanalysis of subjects with clinically confirmed AMD ($P = 5.21 \times 10^{-7}$). Out of the Amish samples that were not part of the original nuclear family and that were genotyped for the P503A variant, we observed 15 additional carriers of the P503A variant with eight subjects affected, five unaffected, and two subjects of unknown case control status. When evaluating the variant in our non-Amish Caucasian dataset, we did not observe the risk allele in either the 791 controls or the 1456 cases. Multipoint linkage analysis carried out on chromosome 1 identified an HLOD peak of 5.12 on chromosome 1 spanning the *CFH* gene. After incorporating a liability class for carriers of the *CFH* P503A rare variant, we observe a maximum LOD score of 4.53 in this region; by including the common Y402H variant, we observe a maximum LOD score of 3.72, when we include liability classes for carriers of both the common and rare *CFH* variants into the model, we observe a maximum HLOD score of 3.28. Upon examination of LD between the Y402H variant and the P503A variant, we observe an r^2 value of 0.002.

DISCUSSION

We determined that the genetic burden of known AMD loci is substantially lower in the Amish than in a general European population sample. Given that AMD is at least as frequent in the Amish as in other European ancestry populations (Stambolian D, unpublished observations, [2012]) and that the Amish generally do not smoke (the strongest known AMD environmental risk factor outside of age), our data supports the hypothesis that other genetic loci are segregating in the Amish.

We identified a densely affected nuclear Amish family in which affected siblings a-78, a-79, and a-80 do not carry the

Y402H or A69S risk variants in *CFH* and *ARMS2*, respectively, loci that account for the majority of the genetic risk of AMD in populations of European descent.⁷ The absence of these risk alleles in affected members lead us to hypothesize that other rare variants of large effect may be contributing to AMD in this family. Using exome sequencing data, we identified a novel missense mutation that is shared among the affected siblings and located in the *CFH* gene. This mutation is a cytosine to guanine transversion resulting in the substitution of an alanine for a proline at amino acid position 503 (P503A). Proline P503A has a genomic evolutionary rate profiling score of 3.64, indicating strong conservation across mammalian species. PolyPhen2, which predicts the possible impact of amino acid substitutions on the structure and function of human proteins, indicates that this variant is probably damaging.

Complement factor H is a regulator of the alternative complement cascade and associations to variation in *CFH* have consistently been replicated in AMD linkage and association analyses.^{9–12,14,15} Complement factor H inhibits activation of complement component 3 (C3) to C3a and C3b, and in addition, direct inactivation of C3b. Previous work has shown binding sites for C3b in short consensus repeat (SCR) domains 1 through 4, 6 through 10, and 16 through 20 of *CFH*.^{37,38} The P503A variant is located within SCR domain 8, and thus may affect C3b binding affinity. Expanding the analysis of this variant to the full Amish dataset, we observe a total of 19 carriers including the four carriers in the nuclear Amish family, 11 reported as having AMD, six reported not having AMD, and two with an unknown affection status. Of the 19 self-reported carriers, six had their diagnosis confirmed by a retinal specialist.

Multipoint linkage analysis carried out across chromosome 1 shows an HLOD score above 5 within the region harboring the *CFH* rare P503A and common Y402H variants. When including liability classes to account for these variants, we see a minimal reduction in the HLOD score, indicating that these variants may be only partially contributing to the observed linkage signal. Subsequent to the exome sequencing, three more members of the original nuclear family used to identify the P503A variant received a clinical diagnosis. Of these three subjects that originally self-reported as not having AMD, one individual, a-81, was seen by a retinal specialist and was reclassified as a case. The self-report age of exam (AOE) of this individual was 66 and the clinical AOE was 74, suggesting that this individual may have progressed to a nonsevere case during the 8-year time-span between the self-report and clinical exam. Genotyping results of a-81 showed that this person was a noncarrier of the P503A variant, but did carry one copy of the Y402H risk allele (Fig. 1). Age of exam is a concern with diagnosing AMD due to the variable age of onset, where no control can definitively be classified without some chance of progression to AMD. Our Amish study sample, although mainly classified through self-report status, is well represented, with late age controls matched to cases, an important factor in phenotype assignment for AMD (Table 2). Of the other two unaffected siblings, a-83 was a carrier of both the P503A variant and A69S, and presented with small drusen at the time of exam. This reaffirms the need for further molecular characterization of this variant and its role in the AMD pathway.

Using the extensive genealogical data available in the AGDB, we observed that the carriers of this variant can be traced back four generations to a shared common ancestor and 15 of the 19 *CFH* rare variant carriers reside within one county in Ohio. We failed to observe the *CFH* P503A variant in 2247 non-Amish individuals (1456 non-Amish cases and 791 non-Amish controls) or in publicly available databases. This suggests that

this *CFH* variant may have become enriched in the Ohio-Indiana Amish populations due to a recent founder event.

Acknowledgments

We thank the community members and family participants for agreeing to participate in this ongoing study. Some of the samples used in this study were collected while WKS and MAP-V were faculty members at Duke University. The authors recognize the efforts of the late Charles E. Jackson for his contributions to the overall Amish projects. We also acknowledge work for this study that was performed using the Vanderbilt Center for Human Genetics Research Core facilities: the Genetic Studies Ascertainment Core, the DNA Resources Core, and the Computational Genomics Core.

Supported by the National Institutes of Health Grants AG019085 (JLH, MAP-V), AG019726 (WKS), EY21453-2 (JNCB), and AG044089 (JDH).

Disclosure: **J.D. Hoffman**, None; **J.N. Cooke Bailey**, None; **L. D'Aoust**, None; **W. Cade**, None; **J. Ayala-Haedo**, None; **D. Fuzzell**, None; **R. Laux**, None; **L.D. Adams**, None; **L. Reinhart-Mercer**, None; **L. Caywood**, None; **P. Whitehead-Gay**, None; **A. Agarwal**, None; **G. Wang**, None; **W.K. Scott**, None; **M.A. Pericak-Vance**, None; **J.L. Haines**, None

References

1. Friedman DS, O'Colmain BJ, Munoz B, et al. Prevalence of age-related macular degeneration in the United States. *Arch Ophthalmol*. 2004;122:564–572.
2. Klein R, Chou CF, Klein BE, Zhang X, Meuer SM, Saaddine JB. Prevalence of age-related macular degeneration in the US population. *Arch Ophthalmol*. 2011;129:75–80.
3. Haug D. *Retinal Imaging*. Philadelphia, PA: Elsevier; 2006; 145–163.
4. Exome Variant Server. *NHLBI GO Exome Sequencing Project (ESP)*. Available at: <http://evs.gs.washington.edu/EVS/>. Accessed February 25, 2013.
5. Klein R, Klein BE, Knudtson MD, Meuer SM, Swift M, Gangnon RE. Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study. *Ophthalmology*. 2007;114:253–262.
6. Schaumberg DA, Christen WG, Hankinson SE, Glynn RJ. Body mass index and the incidence of visually significant age-related maculopathy in men. *Arch Ophthalmol*. 2001;119:1259–1265.
7. 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.
8. 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.
9. 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.
10. Jakobsdottir J, Conley YP, Weeks DE, Mah TS, Ferrell RE, Gorin MB. Susceptibility genes for age-related maculopathy on chromosome 10q26. *Am J Hum Genet*. 2005;77:389–407.
11. Kenealy SJ, Schmidt S, Agarwal A, et al. Linkage analysis for age-related macular degeneration supports a gene on chromosome 10q26. *Mol Vis*. 2004;10:57–61.
12. Majewski J, Schultz DW, Weleber RG, et al. Age-related macular degeneration—a genome scan in extended families. *Am J Hum Genet*. 2003;73:540–550.
13. Maller J, George S, Purcell S, et al. Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration. *Nat Genet*. 2006;38:1055–1059.

14. 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.
15. Fritsche LG, Chen W, Schu M, et al. Seven new loci associated with age-related macular degeneration. *Nat Genet.* 2013;45:433-439.
16. Helgason H, Sulem P, Duvvari MR, et al. A rare nonsynonymous sequence variant in C3 is associated with high risk of age-related macular degeneration. *Nat Genet.* 2013;45:1371-1374.
17. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461:747-753.
18. Zhan X, Larson DE, Wang C, et al. Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nat Genet.* 2013;45:1375-1379.
19. Raychaudhuri S, Iartchouk O, Chin K, et al. A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat Genet.* 2011;43:1232-1236.
20. Hostetler J. *Amish Society*. 4th ed. Baltimore, MD: Johns Hopkins University Press; 1993.
21. Agarwala R, Biesecker LG, Schaffer AA. Anabaptist genealogy database. *Am J Med Genet C Semin Med Genet.* 2003;121C:32-37.
22. Agarwala R, Schaffer AA, Tomlin JF. Towards a complete North American Anabaptist Genealogy II: analysis of inbreeding. *Hum Biol.* 2001;73:533-545.
23. Agarwala R, Biesecker LG, Tomlin JF, Schaffer AA. Towards a complete North American Anabaptist genealogy: a systematic approach to merging partially overlapping genealogy resources. *Am J Med Genet.* 1999;86:156-161.
24. Hahs DW, McCauley JL, Crunk AE, et al. A genome-wide linkage analysis of dementia in the Amish. *Am J Med Genet B Neuropsychiatr Genet.* 2006;141B:160-166.
25. McCauley JL, Hahs DW, Jiang L, et al. Combinatorial mismatch scan (CMS) for loci associated with dementia in the Amish. *BMC Med Genet.* 2006;7:19.
26. Lee WJ, Pollin TI, O'Connell JR, Agarwala R, Schaffer AA. PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County. *BMC Med Genet.* 2010;11:68.
27. Schmidt S, Saunders AM, De La Paz MA, et al. Association of the apolipoprotein E gene with age-related macular degeneration: possible effect modification by family history, age, and gender. *Mol Vis.* 2000;6:287-293.
28. Seddon JM, Ajani UA, Mitchell BD. Familial aggregation of age-related maculopathy. *Am J Ophthalmol.* 1997;123:199-206.
29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754-1760.
30. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491-498.
31. Thornton T, McPeck MS. Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *Am J Hum Genet.* 2007;81:321-337.
32. Mitchell P, Smith W, Attebo K, Wang JJ. Prevalence of age-related maculopathy in Australia. The Blue Mountains Eye Study. *Ophthalmology.* 1995;102:1450-1460.
33. Liu F, Kirichenko A, Axenovich TI, van Duijn CM, Aulchenko YS. An approach for cutting large and complex pedigrees for linkage analysis. *Eur J Hum Genet.* 2008;16:854-860.
34. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet.* 2002;30:97-101.
35. Cummings AC, Jiang L, Velez Edwards DR, et al. Genome-wide association and linkage study in the Amish detects a novel candidate late-onset Alzheimer disease gene. *Ann Hum Genet.* 2012;76:342-351.
36. 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.
37. Pangburn MK, Pangburn KL, Koistinen V, Meri S, Sharma AK. Molecular mechanisms of target recognition in an innate immune system: interactions among factor H, C3b, and target in the alternative pathway of human complement. *J Immunol.* 2000;164:4742-4751.
38. Sharma AK, Pangburn MK. Identification of three physically and functionally distinct binding sites for C3b in human complement factor H by deletion mutagenesis. *Proc Natl Acad Sci U S A.* 1996;93:10996-11001.