



Rare variants and loci for age-related macular degeneration in the Ohio and Indiana Amish

Andrea R. Waksmunski^{1,2,3} · Robert P. Igo Jr.³ · Yeunjoo E. Song³ · Jessica N. Cooke Bailey^{2,3} · Renee Laux³ · Denise Fuzzell³ · Sarada Fuzzell³ · Larry D. Adams⁴ · Laura Caywood⁴ · Michael Prough⁴ · Dwight Stambolian⁵ · William K. Scott⁴ · Margaret A. Pericak-Vance⁴ · Jonathan L. Haines^{1,2,3}

Received: 27 April 2019 / Accepted: 21 July 2019 / Published online: 31 July 2019
© The Author(s) 2019

Abstract

Age-related macular degeneration (AMD) is a leading cause of blindness in the world. While dozens of independent genomic variants are associated with AMD, about one-third of AMD heritability is still unexplained. To identify novel variants and loci for AMD, we analyzed Illumina HumanExome chip data from 87 Amish individuals with early or late AMD, 79 unaffected Amish individuals, and 15 related Amish individuals with unknown AMD affection status. We retained 37,428 polymorphic autosomal variants across 175 samples for association and linkage analyses. After correcting for multiple testing ($n = 37,428$), we identified four variants significantly associated with AMD: rs200437673 (*LCN9*, $p = 1.50 \times 10^{-11}$), rs151214675 (*RTEL1*, $p = 3.18 \times 10^{-8}$), rs140250387 (*DLGAPI*, $p = 4.49 \times 10^{-7}$), and rs115333865 (*CGRRF1*, $p = 1.05 \times 10^{-6}$). These variants have not been previously associated with AMD and are not in linkage disequilibrium with the 52 known AMD-associated variants reported by the International AMD Genomics Consortium based on physical distance. Genome-wide significant linkage peaks were observed on chromosomes 8q21.11–q21.13 (maximum recessive HLOD = 4.03) and 18q21.2–21.32 (maximum dominant HLOD = 3.87; maximum recessive HLOD = 4.27). These loci do not overlap with loci previously linked to AMD. Through gene ontology enrichment analysis with ClueGO in Cytoscape, we determined that several genes in the 1-HLOD support interval of the chromosome 8 locus are involved in fatty acid binding and triglyceride catabolic processes, and the 1-HLOD support interval of the linkage region on chromosome 18 is enriched in genes that participate in serine-type endopeptidase inhibitor activity and the positive regulation of epithelial to mesenchymal transition. These results nominate novel variants and loci for AMD that require further investigation.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00439-019-02050-4>) contains supplementary material, which is available to authorized users.

✉ Jonathan L. Haines
jonathan.haines@case.edu

¹ Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH, USA

² Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA

³ Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

⁴ John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

⁵ Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA, USA

Introduction

Age-related macular degeneration (AMD) is the third leading cause of vision loss in the world (Ayoub and Patel 2009). It is characterized by the deterioration of the central field of vision due to the accumulation of lipid deposits (drusen), inflammation, and neurodegeneration in the macula (Fritsche et al. 2014). AMD is a multifactorial disease with numerous influential genetic and environmental risk factors. Environmental factors constitute about 20–40% of total AMD variance and include age, smoking, hypertension, and diet (Fritsche et al. 2014; Seddon 2013). More than half of AMD heritability is explained by 52 independent common and rare genetic variants across 34 genomic loci (Fritsche et al. 2016). Therefore, about half of AMD heritability is unaccounted for by known genetic polymorphisms and may be partially resolved with the identification of rare variants (Manolio et al. 2009).

As the minor allele frequency (MAF) for a variant decreases, the sample size required to detect an association increases linearly with $1/\text{MAF}$ (Manolio et al. 2009). Consequently, traditional genome-wide-association studies (GWAS) have been largely unable to detect associations with rarer variation (typically $\text{MAF} < 0.01$) unless the sample size is in the tens of thousands. In contrast, GWAS on a sample of large families can increase the power to detect such variants through enrichment of rare variants segregating in families with many affected members (Auer and Lettre 2015). Isolated founder populations, like the Amish, can amplify the power of family-based studies through sharing of founder variants. The Amish are descendants of Swiss Anabaptists who settled in North America 200–300 years ago and maintained their culturally segregated community for generations by marrying within their faith group and adhering to a uniform, conservative lifestyle (Strauss and Puffenberger 2009). Therefore, they have lower genetic and environmental heterogeneity compared to the general population of European descent. In addition, Amish communities comprise large families with an average of seven to nine children in each nuclear family (McKusick et al. 1964). Extensive genealogical data are available for this community in the Anabaptist Genealogy Database (AGDB) making this a valuable population to identify rare genetic variation by combining the advantages of isolated populations and large pedigrees (Agarwala et al. 2003).

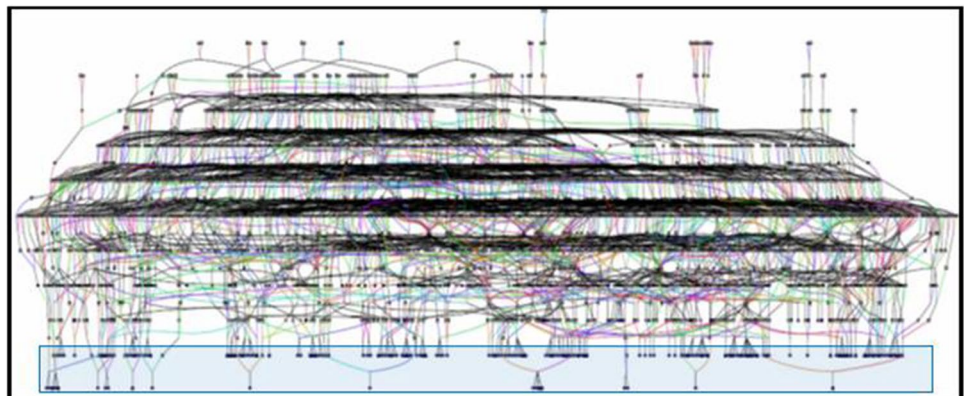
To identify novel variants underlying the pathophysiology of AMD, we performed association and linkage analyses on Illumina HumanExome chip data for 175 related Amish individuals from Ohio and Indiana. We also examined *in silico* functional annotations of the genes from our linkage regions to determine whether these genes participate in biological processes that may be relevant for AMD pathology.

Materials and methods

Study participants

The Amish individuals were identified from published Amish community directories, referral from other Amish community members, or due to close relationships to other research subjects. All study participants were at least 65 years old. The AMD affection status was defined by a self-reported diagnosis on a health questionnaire, which asked whether the individual had ever received an AMD diagnosis from a physician. We previously demonstrated that self-reported AMD statuses have positive and negative predictive values of around 90% in this population (Hoffman et al. 2014). Some Ohio Amish participants also received a clinical exam at the Wooster Eye Clinic in Wooster, Ohio where ocular coherence tomography images were obtained for both eyes. Participants' eyes were graded on a scale from 0 to 5 using the modified Clinical Age-Related Maculopathy Staging (CARMS) system (Seddon et al. 2006b; Sardell et al. 2016). Individuals with grades 0, 1, and 2 in both eyes were designated as unaffected in this study, and individuals with grades 3–5 in at least one eye were designated as affected. Grades of 3 were indicative of early/intermediate AMD, and grades of 4 or 5 were classified as advanced AMD. For this study, 181 Amish samples included 87 individuals with confirmed or self-reported AMD, 79 unaffected individuals, and 15 related Amish individuals with unknown AMD phenotype. Using the Anabaptist Genealogy Database (Agarwala et al. 2003), we constructed a 2118 person all-connecting path pedigree for the 175 Amish individuals that passed quality control and were closely related (Fig. 1). This included 86 affected with AMD, 77 unaffected individuals, and 12 relatives with an unknown AMD affection status. Informed consent was obtained from all study participants, and the study was approved by the institutional review boards at Vanderbilt University and Case Western Reserve

Fig. 1 All-connecting path pedigree of the 175 Amish individuals in this study. The pedigree was drawn using information from the Anabaptist Genealogy Database (AGDB). Circles: females. Squares: males. The genotyped individuals are within the shaded blue box



University. Study procedures were performed in accordance with the tenets of the Declaration of Helsinki.

Genotyping and quality control

A total of 192 samples, which included 181 unique Amish samples, 5 Amish duplicates, and 6 Coriell HapMap trio controls of European descent, were genotyped using the Illumina HumanExome BeadChip v1.1 at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) DNA Resources Core. Genotype data were called and clustered with Illumina's GenomeStudio (Illumina). The accuracy of genotype calling using GenomeStudio is improved by increasing the sample size (Grove et al. 2013). Therefore, an additional 2280 non-Amish samples that were genotyped by VANTAGE with the same protocols and at the same time as the Amish samples were imported into GenomeStudio and clustered with the Amish samples. The data from these non-Amish samples were not used in the subsequent association and linkage analyses. All samples demonstrated call rates greater than 95% and were used for extensive quality control in GenomeStudio as summarized in Supplemental Fig. 1 and Supplemental Table 1 (Guo et al. 2014). Briefly, we excluded variants for low call rate, GenTrain score for cluster quality, cluster separation, and intensity values (Supplemental Table 1). Hemizygous variants were evaluated as well as variants with Mendelian and replication errors (Supplemental Table 1). Rare variant calling was performed manually with GenomeStudio's GenCall algorithm and with zCall, a separate rare variant calling algorithm (Goldstein et al. 2012). zCall is utilized as a post-processing step for calling variants with no genotypes assigned after initial variant calling with GenomeStudio's GenCall algorithm. It uses the intensity profile of the homozygous cluster of the common allele to determine the location of the other homozygous and heterozygous genotype clusters (Goldstein et al. 2012). All variants with at least four new heterozygous calls in zCall were manually reviewed in GenomeStudio.

A total of 146,571 variants and 2466 samples, including 181 Amish individuals, passed quality control in GenomeStudio and were evaluated for sex mismatch, race mismatch, and heterozygosity outliers using PLINK v1.07 (Purcell et al. 2007). Sex mismatched samples were evaluated using heterozygosity rates for the X chromosome and call rates for the Y chromosome (Supplemental Fig. 2). One unresolvable sex mismatch in our Amish samples was identified and removed from the study. An XXY Amish male and an XO Amish female were identified and retained. Ancestry mismatch was assessed using the EIGENSTRAT software package and about 3000 ancestry-informative markers (AIMs) from the exome chip (Price et al. 2006). Principal component analysis (PCA) demonstrated that the combined sample set is multi-ethnic

and that the Amish samples clustered together with other Europeans as expected (Supplemental Fig. 3). None of our Amish samples were in the extremes of the heterozygosity distribution calculated by PLINK v1.07 (Purcell et al. 2007), so none of them were excluded as outliers (Supplemental Fig. 4). The heterozygosity rate can indicate problematic, potentially contaminated DNA samples (Guo et al. 2014). Of the 242,901 variants genotyped, 37,428 were polymorphic in the Amish sample, passed extensive quality control, and served as input for our association and linkage analyses. Variants were considered polymorphic if they had a non-zero minor allele frequency calculated by the ROBust Association-Detection Test for Related Individuals with Population Substructure (ROADTRIPS) software v2.0. We omitted five Amish individuals from our association and linkage analyses due to excessively distant relatedness in the pedigree (Supplemental Fig. 5). Of the remaining 175 Amish individuals, 86 are affected, 77 are unaffected, and 12 have an unknown AMD status.

Association analysis

The samples in this study are highly interrelated (Fig. 1). To account for these relationships and population structure, we performed association testing using the ROADTRIPS software v2.0 (Thornton and McPeck 2010). This program determines associations between genotypes and binary traits while adjusting for the relatedness and population structure of the samples. It calculates and uses an empirical covariance matrix based on the genotypes from the genotype data to account for population structure. ROADTRIPS determines three different test statistics for association: the RM, R_{χ^2} , and RW tests (Thornton and McPeck 2010). For each of these tests, p values are calculated based on a Chi squared distribution with 1 df asymptotic null distribution (Thornton and McPeck 2010). We used the RM test statistic because it allows inclusion of both unaffected individuals and individuals with unknown phenotype in the analysis and is more powerful than the other two when samples are related and pedigree structure is known (Thornton and McPeck 2010). We used kinship coefficients for the 175 Amish individuals calculated by KinInbcoef to account for the pedigree structure in our data (Bourgain 2003). The software program determines the weight of individuals with unknown phenotype in the analysis by a user-defined population prevalence for the trait and the phenotypes of their genotyped relatives (Thornton and McPeck 2010). We set the population prevalence of AMD to 0.13 based on published estimates of its prevalence in individuals of European descent of an age typical of this pedigree (Shalev et al. 2011; Smith et al. 2001).

Linkage analysis

We performed multipoint linkage analyses on a panel of 5668 autosomal variants chosen to be informative for stretches of identity by descent on the exome chip (Abecasis Lab 2013). The 175 Amish individuals in this study are connected in a multigenerational pedigree of 2118 Amish individuals (Fig. 1). Since analyzing very large, complex pedigrees is computationally intractable, we used PedCut to partition the all-connecting path pedigree into densely affected sub-pedigrees with a maximum bit size of 24 (Liu et al. 2008). Parametric heterogeneity LOD (HLOD) scores were determined with the Multi-Point Engine for Rapid Likelihood Inference (MERLIN) software tool (Abecasis et al. 2002) under affected-only dominant and recessive models. The penetrance values for the dominant model were 0, 0.0001, or 0.0001 for 0, 1, or 2 copies of the disease allele, respectively. Under the recessive model, penetrance values were defined as 0, 0, or 0.0001 for 0, 1, or 2 copies of the disease allele, respectively. We set disease allele frequencies to 0.01 and 0.10 for each of these models. We performed multipoint linkage analyses across all autosomes with HLOD scores evaluated at each marker and at every 1 centimorgan (cM; Haldane) with a bit size threshold of 24. For chromosomes with genome-wide significant multipoint HLOD scores, we repartitioned the all-connecting path pedigree into sub-pedigrees with maximum bit sizes of 23 and 25 to test the robustness of our findings to changes in sub-pedigree construction. The model parameters for these analyses were consistent with those used in the initial multipoint analyses.

For chromosome 1, we performed a multipoint linkage analysis with liability classes for carriers of the *CFH* Y402H (rs1061170) (Edwards et al. 2005; Hageman et al. 2005; Haines et al. 2005; Klein et al. 2005) and P503A (rs570523689) variants (Hoffman et al. 2014). *CFH* Y402H failed quality control in our dataset. Therefore, we identified a surrogate variant (rs1329424, $r^2 = 1.0$) for *CFH* Y402H using the LDmatrix module of LDlink to compare linkage disequilibrium statistics among *CFH* variants that had passed quality control in our study and the Y402H variant in the CEU population from Phase 3 (Version 5) of the 1000 Genome Project (Machiela and Chanock 2015). We identified *CFH* P503A carriers by performing customized TaqMan genotyping assays of the variant in our Amish cohort (unpublished data). Briefly, P503A genotyping was carried out using a Custom TaqMan SNP Genotyping Assay (Thermo Fisher) to interrogate the presence of a C (P503) or G (A503) at position 1507 in the *CFH* transcript. Assays were carried out per the manufacturer's instructions using TaqMan Genotyping Master Mix (Thermo Fisher) and 10 ng of genomic DNA per reaction. Reactions were run on a QuantStudio 7 PCR machine, and genotypes were called by

the QuantStudio analysis software. In our conditional linkage analysis, penetrance values were 0.13, 0.312, and 0.624 for 0, 1, and 2 copies of the disease allele, respectively, for carriers of only the surrogate for Y402H. We chose these values based on the population prevalence of AMD in the Caucasian population (13%) (Shalev et al. 2011; Smith et al. 2001) and the expected prevalence based on the published odds ratio (OR) for Y402H, which is 2.4 per copy of the risk allele in individuals of European descent (Haines et al. 2005; Sofat et al. 2012). For carriers of only the P503A variant or both variants, penetrance values were 0.13, 0.6, and 0.6 for 0, 1, and 2 copies of the disease allele, respectively. The penetrance value for the P503A risk allele was derived from the number of affected P503A carriers we identified in a previous study (Hoffman et al. 2014). For non-carriers of both variants, penetrance values were consistent with those used for the multipoint linkage analyses of the other autosomes.

In silico functional analysis of genes from 1-HLOD support intervals of linkage regions

To determine whether genes present in the 1-HLOD support intervals (Supplemental Tables 2–4) of our linkage regions on chromosomes 8 and 18 are functionally related to one another, we extracted gene boundaries from Ensembl (human genome build 37) and performed gene ontology (GO) enrichment analyses with the ClueGO v2.5.3 plug-in (Bindea et al. 2009) of Cytoscape v3.7.1 (<https://cytoscape.org/>). Specifically, terms from the following classifications in GO were considered: Immune System Process, Biological Process, Molecular Function, and Cellular Component. We included all evidence codes (experimental, non-experimental, author statements from publication, and curator statements) in the analysis. We calculated p values using right-sided hypergeometric tests and the Benjamini–Hochberg correction for multiple testing. We chose medium network specificity to identify representative pathways for our genes of interest. This level of specificity examines GO levels 3–8, requires at least three genes per GO term, and ensures that mapped genes represent at least 4% of the total associated genes. We also used a Kappa score of 0.4 for our GO term network connectivity threshold. GO terms were iteratively compared, merged into functional groups, and visualized in Cytoscape.

Results

Association analysis

Of the 37,428 polymorphic autosomal variants that passed quality control, four variants met the Bonferroni correction threshold ($p < 1.34 \times 10^{-6}$): rs200437673 (chromosome 9,

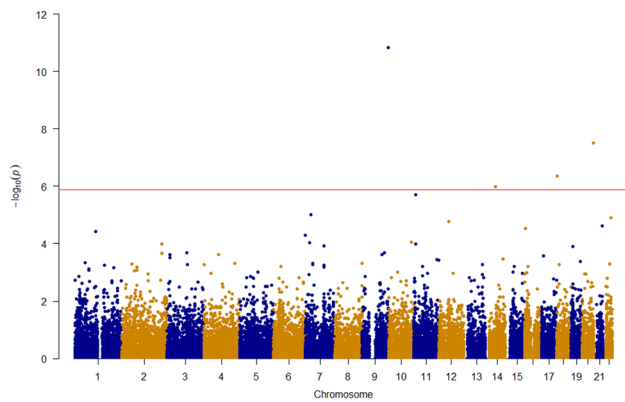


Fig. 2 Manhattan plot of p values obtained from association testing using ROADTRIPS. The red line denotes the Bonferroni correction threshold for multiple testing ($p < 1.34 \times 10^{-6}$). Four variants passed this threshold and are summarized in Table 1

LCN9, $p = 1.50 \times 10^{-11}$, rs151214675 (chromosome 20, *RTEL1*, $p = 3.18 \times 10^{-8}$, rs140250387 (chromosome 18, *DLGAPI*, $p = 4.49 \times 10^{-7}$), and rs115333865 (chromosome 14, *CGRRF1*, $p = 1.05 \times 10^{-6}$) (Fig. 2; Table 1). p values were not corrected for inflation because the genomic control

factor was 1.05 (Supplemental Fig. 6). These variants demonstrate novel associations with AMD and do not map to the AMD susceptibility loci identified by the International AMD Genomics Consortium (IAMDGC) based on physical distance (Fritsche et al. 2016). While these variants were identified in the Amish, we found that they are rare to low-frequency ($MAF < 0.01$) in outbred populations of European descent including the IAMDGC dataset and the gnomAD database (Table 2).

Linkage analysis

Multipoint linkage analyses under autosomal dominant and recessive models were performed to identify genomic loci linked to AMD in 16 Amish sub-pedigrees (Supplemental Figs. 7, 8; Supplemental Tables 5, 6). Under the recessive model with a disease allele frequency of 0.10, we identified genome-wide significant signals ($HLOD > 3.6$) on chromosome 8 (Table 3; Fig. 3). We also identified significant signals on chromosome 18 under the dominant and recessive models (Table 3; Fig. 4). These signals were robust to varying sub-pedigree structure (Supplemental Figs. 9, 10). Although we initially observed genome-wide significant

Table 1 AMD-associated variants identified with ROADTRIPS testing of Amish families

rsID	Chr.	Position	Alleles	Gene(s)	Consequence	Allele count			P
						Affected	Unaffected	Unknowns	
rs200437673	9	138,555,237	A/G	<i>LCN9</i>	Missense	5	0	1	1.50×10^{-11}
rs151214675	20	62,293,235	G/A	<i>RTEL1</i>	Missense	6	1	0	3.18×10^{-8}
				<i>RTEL1-TNFRSF6B</i>	Non-coding Transcript				
rs140250387	18	3,534,424	A/G	<i>DLGAPI</i>	Synonymous	3	1	0	4.49×10^{-7}
rs115333865	14	55,004,449	G/A	<i>CGRRF1</i>	Missense	10	2	1	1.05×10^{-6}

p values were obtained from the RM test

Variant positions are given for build 37 (hg19) of the human genome

AMD, age-related macular degeneration; ROADTRIPS, ROBust Association-Detection Test for Related Individuals with Population Substructure; Chr., chromosome; *LCN9*, lipocalin 9; *RTEL1*, regulator of telomere elongation helicase 1; *RTEL1-TNFRSF6B*, *RTEL1-TNFRSF6B* readthrough (NMD candidate); *DLGAPI*, DLG-associated protein 1; *CGRRF1*, cell growth regulator with ring finger domain 1

Table 2 Allele frequencies for the AMD-associated variants identified with ROADTRIPS testing of Amish families in outbred populations (IAMDGC and gnomAD release 2.1)

Variant		Allele frequency		
rsID	Chromosome	IAMDGC		gnomAD
		Advanced AMD cases	Controls	Non-Finnish Europeans
rs200437673	9	0.004118	0.0043691	0.004341
rs151214675	20	0.00024779	0.00016824	0.0001936
rs140250387	18	0.001146	0.0012338	0.001611
rs115333865	14	0.019112	0.019993	0.01879

AMD, age-related macular degeneration; ROADTRIPS, ROBust Association-Detection Test for Related Individuals with Population Substructure; IAMDGC, International Age-related Macular Degeneration Genomics Consortium; gnomAD, Genome Aggregation Database

Table 3 Significant linkage loci identified from model-based multipoint linkage analyses of Amish families with disease allele frequency of 0.10

Model	Chromosome	Peak HLOD score	Region with HLOD > 3.6 (cM)	1-HLOD support interval (cM)
Recessive	8	4.03	97.28–99.61	96.745–105.65
Dominant	18	3.87	75.82–85.41	70.88–87.23
Recessive	18	4.27	74.39–78.51	74.025–78.87 81.31–82.18 83.46–84.80

The peak HLOD score is the maximum HLOD score obtained for the designated chromosome

HLOD, heterogeneity LOD; cM, centimorgans

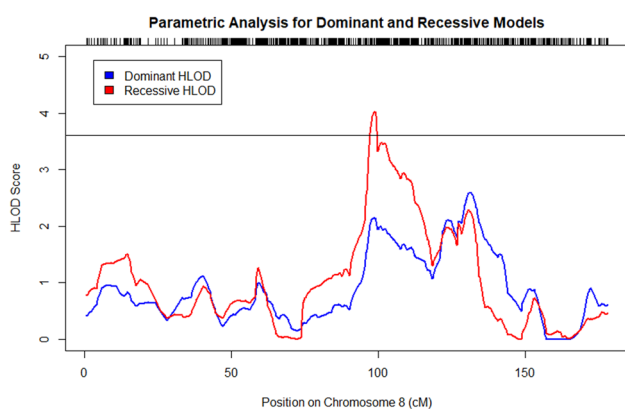


Fig. 3 HLOD scores obtained from multipoint linkage analysis in MERLIN under dominant and recessive models on chromosome 8. The black line denotes genome-wide significance (HLOD Score > 3.6). A genome-wide significant HLOD score of 4.03 was observed under the recessive model. Tick marks along the upper *x*-axis correspond to the marker positions

signals (HLOD > 3.6) on chromosome 2 under the dominant and recessive models (Supplemental Table 7; Supplemental Fig. 11) and on chromosome 15 under the recessive model (Supplemental Table 7; Supplemental Fig. 12), these signals greatly reduced when we re-performed our analyses with different sub-pedigree structures (Supplemental Figs. 13, 14).

In our multipoint linkage analysis under the dominant model with a disease allele frequency of 0.10, we observed a region on chromosome 1 that was nearly genome-wide significant (maximum HLOD = 3.50 at 234.60 cM; Supplemental Table 5). The peak of this region is about 16 MB from the complement factor H (*CFH*) gene, which has been repeatedly associated with AMD risk in the general population (Fritsche et al. 2016). We performed a conditional linkage analysis on the markers on chromosome 1 with liability classes for carriers of two AMD risk variants in *CFH* (Y402H and P503A) to determine whether these markers

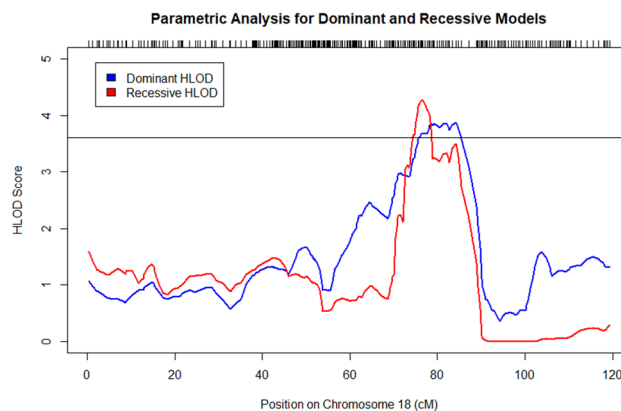


Fig. 4 HLOD scores obtained from multipoint linkage analysis in MERLIN under dominant and recessive models on chromosome 18. The black line denotes genome-wide significance (HLOD Score > 3.6). The maximum dominant HLOD score was 3.87. The maximum recessive HLOD score was 4.27. Tick marks along the upper *x*-axis correspond to the marker positions

were driving the peak linkage signal on chromosome 1. Of the 86 affected individuals in the linkage analysis, 8 carry at least one copy of the P503A variant, and 63 have at least 1 copy of the surrogate for Y402H. Three of the P503A carriers are also heterozygous for the surrogate of Y402H. Our conditional linkage analysis on chromosome 1 using distinct classes for carriers of the surrogate for *CFH* Y402H and/or the *CFH* P503A variant demonstrated that the *CFH* locus was contributing to the peak signal we observed in the unconditioned analysis (Fig. 5).

In silico functional analysis of genes from 1-HLOD support intervals of linkage regions

We performed GO enrichment analyses using right-sided hypergeometric tests on the genes from the 1-HLOD support intervals our linkage regions to identify any functional relationships through GO terms. *p* values were corrected for multiple testing using the Benjamini–Hochberg method. The region we identified with significant evidence of linkage on chromosome 8q21.11–q21.13 includes 19 genes (Supplemental Table 2). The 1-HLOD support interval of this peak contains 80 unique genes (Supplemental Table 2), including those that are involved in fatty acid binding and triglyceride catabolic processes (Fig. 6a; Supplemental Table 8). The region we found with significant evidence of linkage on chromosome 18q21.2–21.32 under the dominant model includes 47 genes (Supplemental Table 3). The 1-HLOD support interval of this peak contains 102 unique genes (Supplemental Table 3), including those that are implicated in serine-type endopeptidase inhibitor activity and positive regulation of epithelial to mesenchymal transition (Fig. 6b; Supplemental Table 9).

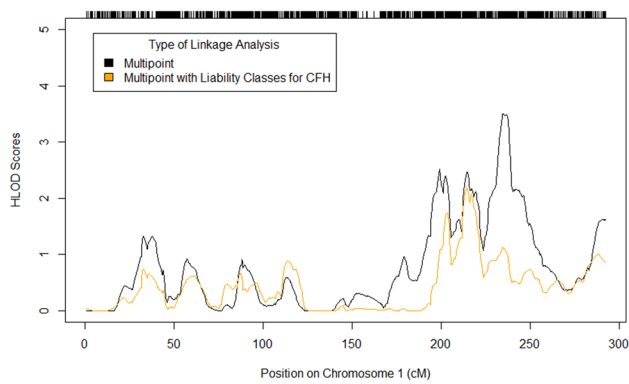


Fig. 5 Conditional linkage analysis on chromosome 1 taking into account the Y402H and P503A carrier statuses. The black line denotes the HLOD scores from the multipoint linkage analysis performed without liability classes for the *CFH* variants. The orange line represents the distribution of HLOD scores from the conditional linkage analysis. The peak HLOD score in the unconditioned analysis was 3.503 at 234.6 cM (212,619,339 bp, build 37). The peak HLOD score in the conditioned analysis was 2.1836 at 214.47 cM (197,070,697 bp, build 37). The *CFH* gene boundaries are 196,621,008–196,716,634 bp (Ensembl, build 37)

The significant linkage region we observed on chromosome 18q21.2–21.31 under the recessive model includes 19 genes (Supplemental Table 4). The 1-HLOD support interval of this peak contains 46 unique genes (Supplemental Table 4), which were not overrepresented in any particular GO terms in our enrichment analysis. When we analyzed all the genes from our 1-HLOD support intervals on chromosome 8 and 18 together in ClueGO, the same GO terms from our chromosome-specific analyses were identified.

Discussion

We performed association and linkage analyses on Amish families from Ohio and Indiana to uncover novel variants and loci for AMD. Using ROADTRIPS and kinship coefficients derived from the relationships detailed in the all-connecting path pedigree, we identified four novel variants associated with AMD in the Amish. Since we analyzed exome chip data, we recognize that these variants may not be the functional variants underlying these loci and that we might be observing their association signals because they are in linkage disequilibrium with the true functional variants for AMD. These four variants are independent and physically distant from the 52 AMD-associated variants identified by the IAMDGC (Fritsche et al. 2016). Association signals were not detected in the 5 directly genotyped variants in the *LCN9* gene on chromosome 9 or the 27 directly genotyped variants in the *RTEL1/RTEL1-TNFRSF6B* locus on chromosome 20 from the IAMDGC GWAS for advanced AMD. However, a few nominal signals ($p < 0.05$) were found in the IAMDGC loci for *DLGAP1* ($n = 3$ out of 144) on chromosome 18 and *CGRRF1* ($n = 1$ out of 6) on chromosome 14 (Fritsche et al. 2016). These three *DLGAP1* variants are not in linkage disequilibrium (LD) with one another ($r^2 < 0.2$) but are within 500 KB of the variant we identified in the Amish. The *CGRRF1* variant that achieved a p value less than 0.05 in the IAMDGC GWAS and the variant we identified with ROADTRIPS in the Amish are 500 bp apart but are not in LD ($r^2 = 0.0001$ in CEU population). The variant we found on chromosome 9 (rs200437673) is about 36.6 MB, 65.1 MB, 61.9 MB, and 30.9 MB away from the lead variants identified by the IAMDGC on this chromosome (rs1626340, rs71507014, rs10781182, and rs2740488,

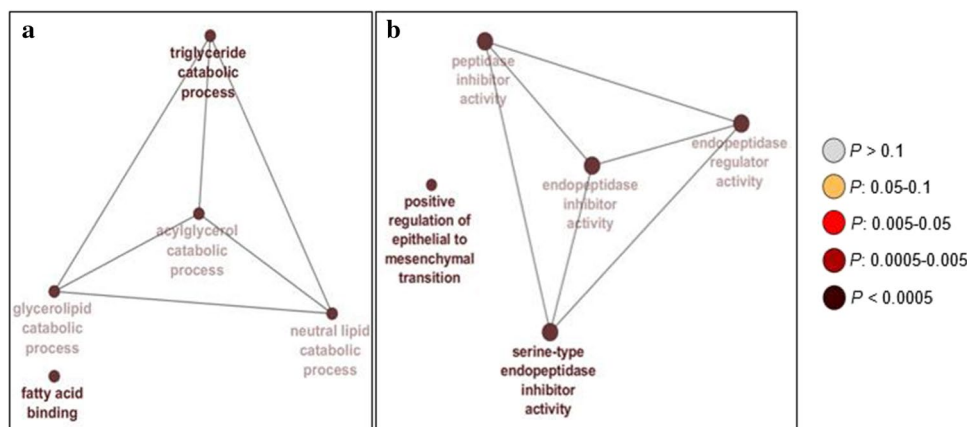


Fig. 6 Gene Ontology (GO) Enrichment Networks for Genes from 1-HLOD Support Intervals from **a** chromosome 8 and **b** chromosome 18. The size of the nodes in each network illustrates the number of mapped genes to the depicted GO term, and the color of the node

shows the significance of the term in each enrichment analysis. The leading term of each GO group is depicted in bold and represents the most significant ontology from the group. Full descriptions of these terms are available in Supplemental Tables 8 and 9

respectively) (Fritsche et al. 2016). The variant we identified on chromosome 20 (rs151214675) is about 5.6 and 17.7 MB away from the lead variants from the IAMDC GWAS on chromosome 20 (rs201459901 and rs142450006, respectively) (Fritsche et al. 2016). The IAMDC did not identify any AMD-associated variants on chromosome 18 (Fritsche et al. 2016). The lead variant in the AMD locus identified by the IAMDC on chromosome 14 (rs61985136) (Fritsche et al. 2016) is almost 14 MB away from the variant we identified in the Amish (rs115333865). The four variants we identified in the Amish are rare to low-frequency in the IAMDC data and the gnomAD database. While allele frequencies for these variants are fairly similar in advanced AMD cases and controls from the IAMDC, these variants were significantly associated with AMD in our Amish cohort. It was previously demonstrated that the Amish population has a lower genetic burden of known AMD variants (Hoffman et al. 2014); therefore, this may suggest that the Amish have different etiology for AMD than the general population of European descent.

Of the AMD-associated variants that we identified in this study, only the *RTEL1* variant (rs151214675) is cataloged in the ClinVar database, but it is considered a variant of unknown significance (<https://www.ncbi.nlm.nih.gov/clinvar/RCV000504185/>). This variant also maps to the naturally occurring *RTEL1-TNFRSF6B* read-through transcript of this locus, which is a candidate for nonsense-mediated mRNA decay and unlikely to yield a protein product. *RTEL1* encodes a regulator of a telomere elongation helicase, which maintains telomere length and genomic stability (Barber et al. 2008; Codd et al. 2013; Ding et al. 2004). Telomere length has been hypothesized as a marker of aging because telomeres shorten with age and the presence of short telomeres directs the cell to enter senescence (von Zglinicki and Martin-Ruiz 2005). Oxidative stress can also contribute to the reduction of telomere length in cells and has been characterized as a contributing factor to AMD pathophysiology (Shaw et al. 2016; von Zglinicki and Martin-Ruiz 2005). Lipocalins constitute a family of extracellular proteins that are responsible for transporting small lipids such as fatty acids, retinoids, and steroids (Suzuki et al. 2004). There are no *LCN9* variants documented in the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>). However, another lipocalin family member (lipocalin-2, *LCN2*) may modulate inflammation in retinal degeneration by promoting cell survival responses and regulating the production of inflammatory proteins (Parmar et al. 2018). Additionally, tear lipocalins constitute a large group of lipid-binding proteins in tears and may serve as potential biomarkers for diabetic retinopathy and Alzheimer's disease (Kallo et al. 2016; Wang et al. 2017). *DLGAP1* encodes a guanylate kinase-associated protein involved in protein–protein interactions with scaffold proteins in the post-synaptic density of excitatory

synapses (Feng and Zhang 2009; Li et al. 2016). *DLGAP1* also is one of nine genes located in a chromosomal region (myopia-2, *MYP2*) that demonstrated significant genetic linkage with autosomal dominant high myopia (Scavell et al. 2005; Young et al. 1998). The proteins encoded by *DLGAP1* and lipocalin genes (*LCN1* and *LCN2*) were also found to be expressed in the human choroid-retinal pigment epithelial complex (Skeie and Mahajan 2014). The variant we identified in *DLGAP1* is about 3 MB away from an AMD variant (rs9973159) identified in a genome-wide association study accounting for age-stratified effects in the IAMDC data (Winkler et al. 2018). *CGRRF1* encodes a cell growth regulator that has been associated with eye morphology in humans (Lee et al. 2017). While the protein product of this gene is not well-characterized, it has proposed as a modulator of Evi, which is a transmembrane protein involved in Wnt protein secretion (Glaeser et al. 2018). Canonical Wnt signaling has been implicated in retinal inflammation and may have a role in AMD pathology (Zhou et al. 2010). Additional studies will be required to elucidate the roles these genes might have in AMD etiology.

In our linkage screens in the Amish, we identified a novel susceptibility locus for AMD on chromosome 8q21.11–q21.13 under our recessive model. This region does not overlap with AMD loci previously identified with genome-wide linkage screens, which occur on chromosomes 8p21 and 8q11.2 (Seddon et al. 2003). In their recent GWAS, the IAMDC identified a susceptibility locus for AMD on chromosome 8p21.3 with the most significant signals coming from rs13278062 and rs79037040 in the *TNFRSF10A/LOC389641* and *TNFRSF10A* genes (Fritsche 2013, 2016). In our linkage analyses of chromosome 8, we observed a maximum HLOD score of 0.67 in this gene region under the recessive model with disease allele frequency of 0.10. The strongest single variant *p* value observed by the IAMDC in our linkage region was 6.54×10^{-4} , which does not reach classical GWAS significance ($p < 5 \times 10^{-8}$). Common variants associated with optic nerve degeneration in glaucoma have also been identified in 8q22, which is near our linkage region (Wiggs et al. 2012).

Based on our GO enrichment analysis with ClueGO, the genes from the 1-HLOD support interval of our significant linkage region on chromosome 8q21.11–q21.13 have functional annotations related to triglyceride catabolic processes and fatty acid binding. Components of lipid metabolism have been previously implicated in the genetic etiology and pathophysiology of AMD. Lipids comprise about 40% of the composition of drusen, which are a hallmark of AMD (Wang et al. 2010). In their most recent GWAS, the IAMDC determined that several lipid pathways from the Reactome and GO pathway databases were enriched for genes from the 34 AMD susceptibility loci they identified (Fritsche et al. 2016). Known AMD-associated genes (*APOE* and *LIPC*)

(Fritsche et al. 2016) are also described as members of the triglyceride, neutral lipid, acylglycerol, and glycerolipid catabolic processes (GO:0019433, GO:0046461, GO:0046464, and GO:0046503, respectively), which were enriched in genes from our linkage region. Additionally, although consuming high amounts of saturated and monounsaturated fats has been associated with AMD risk (Delcourt et al. 2007), dietary intake of omega-3-fatty acids has been attributed to reducing the risk of AMD (Delcourt et al. 2007; SanGiovanni and Chew 2005; SanGiovanni et al. 2008; Seddon et al. 2006a). There is not a consistent association between triglyceride levels and AMD risk. Some studies observed lower triglyceride levels in patients with early AMD (Klein et al. 2003), the choroidal neovascularization subtype of advanced AMD (Merle et al. 2014), or any type of AMD (Paun et al. 2015; Roh et al. 2008; Semba et al. 2014). However, other studies found higher triglyceride levels were associated with AMD status (Munch et al. 2013; Nowak et al. 2005). Recently, levels of triglycerides were associated with a decreased risk for AMD and smaller drusen area (Colijn et al. 2018). Further functional studies would need to be performed to definitively implicate the role of fatty acid binding and triglyceride levels in AMD pathology.

From our multipoint linkage analyses with dominant and recessive models of inheritance, we identified novel AMD loci on chromosome 18q21.2–21.32 and 18q21.2–21.31, respectively. These regions do not overlap with the AMD-associated variant (rs140250387) we identified in our association test. Although both analyses interrogate the genetic etiology of the trait, linkage analyses identify chromosomal segments co-segregating with traits in families (Morton 1955); whereas, association analyses compare allele frequencies between individuals with and without the trait of interest (Hirschhorn and Daly 2005). Therefore, their results can be independent of one another. The IAMDGC did not identify any AMD susceptibility loci on chromosome 18 in either of their recent GWAS, and another study using IAMDGC data identified an AMD gene \times age interaction on chromosome 18p, which is independent of our linkage peak (Fritsche 2013, 2016; Winkler et al. 2018). The most notable single-marker p value observed in the IAMDGC GWAS was 1.27×10^{-4} , which occurred for a variant that falls within our linkage regions under both dominant and recessive models. Previous GWAS have found suggestive AMD variants on chromosomes 18q22.1 (Naj et al. 2013) and 18q23 (Sobrin et al. 2012), which are located outside of the 1-HLOD support interval of our linkage region. The C allele of the 18q22.1 variant (rs17073641) has a strong protective effect in never smokers but increases the risk of AMD in smokers (Naj et al. 2013). The 18q23 variant (rs1789110), which occurs near the myelin basic protein (*MBP*) gene, demonstrated suggestive evidence of association with geographic atrophy (Sobrin et al. 2012).

The 1-HLOD support interval of our significant linkage region on chromosome 18q21.2–21.32 contains genes that have functional annotations in GO processes such as positive regulation of epithelial to mesenchymal transition (EMT) and serine-type endopeptidase inhibitor activity. Another known AMD risk gene, *TGFBR1*, (Fritsche 2013, 2016) has been previously described as a part of the positive regulation of EMT (GO:0010718). The EMT of retinal pigment epithelial (RPE) cells is considered one of several biological processes that are responsible for the formation of subretinal fibrosis in the macula following choroidal neovascularization in advanced AMD (Ishikawa et al. 2016a). EMT has also been proposed as a mechanism employed by RPE cells to survive in the stressful macular microenvironment during dry AMD progression (Ghosh et al. 2018). Therefore, it has been suggested that therapeutically targeting RPE cells in EMT may help treat patients with advanced AMD and subretinal fibrosis (Ghosh et al. 2018; Ishikawa et al. 2016b; Kobayashi et al. 2019). Several of the serine proteinase inhibitor (serpin) family members from the 1-HLOD support interval have also been described in the context of AMD-related pathologies. The proteins encoded by *SERPINB5* (maspin) and *SERPINB13* (headpin) are known for their anti-angiogenic properties, which may have implications for the choroidal neovascularization subtype of advanced AMD (Pescosolido et al. 2014; Shellenberger et al. 2005; Zhang et al. 2000). A genetic variant near *SERPINB2* was identified as a risk factor for AMD in smokers and a protective factor in nonsmokers in gene-environment interaction analyses (Naj et al. 2013). Another gene from our 1-HLOD support interval, *TCF4*, has repeatedly been associated with Fuchs endothelial dystrophy, which is characterized by the deterioration of the corneal epithelium (Baratz et al. 2010; Kuot et al. 2012; Li et al. 2011; Riazuddin et al. 2011; Thalamuthu et al. 2011). *C3* and *TIMP3* are known AMD-associated genes identified by the IAMDGC that are also described in the following GO terms from our analyses: endopeptidase regulator activity (GO:0061135), peptidase inhibitor activity (GO:0030414), and endopeptidase inhibitor activity (GO:0004866). Additionally, the only subtype-specific variant identified for the CNV subtype of advanced AMD was located near the *MMP9* gene on chromosome 20, which is a member of the endopeptidase family that participates in extracellular matrix degradation (Fritsche et al. 2016).

Although the major peak in our initial linkage analysis on chromosome 1 was diminished in our conditional linkage analysis, a suggestive linkage peak (HLOD \sim 2) remained around 213–217 cM. This genomic region corresponds to a locus of approximately 5 MB that contains multiple genes including *CFH* and *CFHR1-5*. These genes are located within the regulation of complement activation (RCA) gene cluster on chromosome 1 and encode members of the factor H/CFHR family of proteins (Male et al. 2000). Polymorphisms

and deletions in *CFH* and *CFHRs* have been previously associated with AMD (Edwards et al. 2005; Fritsche et al. 2016; Hageman et al. 2005; Haines et al. 2005; Klein et al. 2005; Lores-Motta et al. 2018; Martinez-Barricarte et al. 2012), including rare variants in *CFH* (Hoffman et al. 2014; Raychaudhuri et al. 2011) and a protective deletion of *CFHR1* and *CFHR3* (Fritsche et al. 2010; Hageman et al. 2006; Hughes et al. 2006). More detailed mapping of this region in our cohort will be necessary to determine which polymorphisms in these gene(s) are contributing to our linkage signal.

In this study, we performed association and linkage screens in Amish families from Ohio and Indiana. We identified four novel rare variants associated with increased AMD risk in the Amish and novel susceptibility loci on chromosomes 8q21.11–q21.13 (maximum recessive HLOD = 4.03) and 18q21 (maximum dominant HLOD = 3.87; maximum recessive HLOD = 4.27). These findings suggest novel genetic factors for AMD in the Amish and demonstrate the benefits of analyzing the genetic architecture of a complex trait, like AMD, in closely related individuals from an isolated population. While family-based studies do not require sample sizes as large as traditional GWAS, the sample size for this study ($n = 175$) limited its power to detect novel AMD variants and loci. This could be remedied with the ascertainment of additional members of many Amish communities. Additionally, more genetic data could be obtained from whole exome sequencing or dense genotyping for a panel such as the Multi-Ethnic Genotyping Array (MEGA) of about 2 million exonic markers for GWAS and ancestry-specific variants. Additional studies are needed to replicate their associations and validate their roles in the development or progression of AMD in outbred populations. Analyses could also be performed for the subtypes or endophenotypes of AMD to identify subtype-specific trends or associations with other binary and quantitative traits that may be pertinent for AMD etiology. Such studies may help resolve the missing heritability of AMD and provide novel insights to AMD pathophysiology.

Acknowledgements We thank the Amish families for allowing us into their communities and for participating in our study. We also acknowledge the contributions of the Anabaptist Genealogy Database and Swiss Anabaptist Genealogy Association. The authors acknowledge the contributions of the Vanderbilt Technologies for Advanced Genomics (VANTAGE) DNA Resources Core. This work was supported by the National Institutes of Health [EY023164]. J.N.C.B. was supported in part by a PhRMA Informatics Postdoctoral Fellowship. This work made use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use,

distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abecasis Lab (2013) Exome chip design wiki site. https://genome.sph.umich.edu/wiki/Exome_Chip_Design. Accessed 9 May 2017
- 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. <https://doi.org/10.1038/ng786>
- Agarwala R, Biesecker LG, Schaffer AA (2003) Anabaptist genealogy database. *Am J Med Genet C Semin Med Genet* 121C:32–37. <https://doi.org/10.1002/ajmg.c.20004>
- Auer PL, Lettre G (2015) Rare variant association studies: considerations, challenges and opportunities. *Genome Med* 7:16. <https://doi.org/10.1186/s13073-015-0138-2>
- Ayoub T, Patel N (2009) Age-related macular degeneration. *J R Soc Med* 102:56–61. <https://doi.org/10.1258/jrsm.2009.080298>
- Baratz KH et al (2010) E2-2 protein and Fuchs's corneal dystrophy. *N Engl J Med* 363:1016–1024. <https://doi.org/10.1056/NEJMoal007064>
- Barber LJ et al (2008) RTEL1 maintains genomic stability by suppressing homologous recombination. *Cell* 135:261–271. <https://doi.org/10.1016/j.cell.2008.08.016>
- Bindea G et al (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25:1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>
- Bourgain C (2003) KinInbcoef: calculation of kinship and inbreeding coefficients
- Codd V et al (2013) Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet* 45:422–427. <https://doi.org/10.1038/ng.2528>
- Colijn JM et al (2018) Increased high-density lipoprotein levels associated with age-related macular degeneration: evidence from the EYE-RISK and European Eye Epidemiology Consortia. *Ophthalmology*. <https://doi.org/10.1016/j.ophtha.2018.09.045>
- Delcourt C, Carriere I, Cristol JP, Lacroux A, Gerber M (2007) Dietary fat and the risk of age-related maculopathy: the POLANUT study. *Eur J Clin Nutr* 61:1341–1344. <https://doi.org/10.1038/sj.ejcn.1602685>
- Ding H et al (2004) Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* 117:873–886. <https://doi.org/10.1016/j.cell.2004.05.026>
- Edwards AO, Ritter R 3rd, Abel KJ, Manning A, Panhuysen C, Farrer LA (2005) Complement factor H polymorphism and age-related macular degeneration. *Science* 308:421–424. <https://doi.org/10.1126/science.1110189>
- Feng W, Zhang M (2009) Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci* 10:87–99. <https://doi.org/10.1038/nrn2540>
- Fritsche LG et al (2010) An imbalance of human complement regulatory proteins CFHR1, CFHR3 and factor H influences risk for age-related macular degeneration (AMD). *Hum Mol Genet* 19:4694–4704. <https://doi.org/10.1093/hmg/ddq399>
- Fritsche LG et al (2013) Seven new loci associated with age-related macular degeneration. *Nat Genet* 45:433–439, 439e431–432. <https://doi.org/10.1038/ng.2578>
- Fritsche LG, Fariss RN, Stambolian D, Abecasis GR, Curcio CA, Swaroop A (2014) Age-related macular degeneration: genetics and biology coming together. *Annu Rev Genom Hum Genet* 15:151–171. <https://doi.org/10.1146/annurev-genom-090413-025610>

- Fritsche LG et al (2016) A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat Genet* 48:134–143. <https://doi.org/10.1038/ng.3448>
- Ghosh S et al (2018) A role for betaA3/A1-crystallin in type 2 EMT of RPE cells occurring in dry age-related macular degeneration. *Invest Ophthalmol Vis Sci* 59:AMD104–AMD113. <https://doi.org/10.1167/iovs.18-24132>
- Glaeser K et al (2018) ERAD-dependent control of the Wnt secretory factor Evi. *EMBO J*. <https://doi.org/10.15252/embj.201797311>
- Goldstein JI et al (2012) zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* 28:2543–2545. <https://doi.org/10.1093/bioinformatics/bts479>
- Grove ML et al (2013) Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One* 8:e68095. <https://doi.org/10.1371/journal.pone.0068095>
- Guo Y et al (2014) Illumina human exome genotyping array clustering and quality control. *Nat Protoc* 9:2643–2662. <https://doi.org/10.1038/nprot.2014.174>
- Hageman GS 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. <https://doi.org/10.1073/pnas.0501536102>
- Hageman GS et al (2006) Extended haplotypes in the complement factor H (CFH) and CFH-related (CFHR) family of genes protect against age-related macular degeneration: characterization, ethnic distribution and evolutionary implications. *Ann Med* 38:592–604. <https://doi.org/10.1080/07853890601097030>
- Haines JL et al (2005) Complement factor H variant increases the risk of age-related macular degeneration. *Science* 308:419–421. <https://doi.org/10.1126/science.1110359>
- Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6:95–108. <https://doi.org/10.1038/nrg1521>
- Hoffman JD et al (2014) Rare complement factor H variant associated with age-related macular degeneration in the Amish. *Invest Ophthalmol Vis Sci* 55:4455–4460. <https://doi.org/10.1167/iovs.13-13684>
- Hughes AE, Orr N, Esfandiary H, Diaz-Torres M, Goodship T, Chakravarthy U (2006) A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat Genet* 38:1173–1177. <https://doi.org/10.1038/ng1890>
- Illumina GenomeStudio Software. <http://www.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html>. Accessed 22 Jan 2016
- Ishikawa K, Kannan R, Hinton DR (2016a) Molecular mechanisms of subretinal fibrosis in age-related macular degeneration. *Exp Eye Res* 142:19–25. <https://doi.org/10.1016/j.exer.2015.03.009>
- Ishikawa K, Sreekumar PG, Spee C, Nazari H, Zhu D, Kannan R, Hinton DR (2016b) AlphaB-crystallin regulates subretinal fibrosis by modulation of epithelial–mesenchymal transition. *Am J Pathol* 186:859–873. <https://doi.org/10.1016/j.ajpat.2015.11.014>
- Kallo G, Emri M, Varga Z, Ujhelyi B, Tozser J, Csutak A, Csosz E (2016) Changes in the chemical barrier composition of tears in Alzheimer's disease reveal potential tear diagnostic biomarkers. *PLoS One* 11:e0158000. <https://doi.org/10.1371/journal.pone.0158000>
- Klein R, Klein BE, Marino EK, Kuller LH, Furberg C, Burke GL, Hubbard LD (2003) Early age-related maculopathy in the cardiovascular health study. *Ophthalmology* 110:25–33
- Klein RJ et al (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–389. <https://doi.org/10.1126/science.1109557>
- Kobayashi M, Tokuda K, Kobayashi Y, Yamashiro C, Uchi SH, Hatano M, Kimura K (2019) Suppression of epithelial–mesenchymal transition in retinal pigment epithelial cells by an MRTF—an inhibitor. *Invest Ophthalmol Vis Sci* 60:528–537. <https://doi.org/10.1167/iovs.18-25678>
- Kuot A et al (2012) Association of TCF4 and CLU polymorphisms with Fuchs' endothelial dystrophy and implication of CLU and TGFBI proteins in the disease process. *Eur J Hum Genet* 20:632–638. <https://doi.org/10.1038/ejhg.2011.248>
- Lee MK et al (2017) Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. *PLoS One* 12:e0176566. <https://doi.org/10.1371/journal.pone.0176566>
- Li YJ et al (2011) Replication of TCF4 through association and linkage studies in late-onset Fuchs endothelial corneal dystrophy. *PLoS One* 6:e18044. <https://doi.org/10.1371/journal.pone.0018044>
- Li J, Wilkinson B, Clementel VA, Hou J, O'Dell TJ, Coba MP (2016) Long-term potentiation modulates synaptic phosphorylation networks and reshapes the structure of the postsynaptic interactome. *Sci Signal* 9:rs8. <https://doi.org/10.1126/scisignal.aaf6716>
- Liu F, Kirichenko A, Axenovich TI, van Duijn CM, Aulchenko YS (2008) An approach for cutting large and complex pedigrees for linkage analysis. *Eur J Hum Genet* 16:854–860. <https://doi.org/10.1038/ejhg.2008.24>
- Lores-Motta L et al (2018) Genome-wide association study reveals variants in CFH and CFHR4 associated with systemic complement activation: implications in age-related macular degeneration. *Ophthalmology* 125:1064–1074. <https://doi.org/10.1016/j.ophtha.2017.12.023>
- Machiela MJ, Chanock SJ (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31:3555–3557. <https://doi.org/10.1093/bioinformatics/btv402>
- Male DA, Ormsby RJ, Ranganathan S, Giannakis E, Gordon DL (2000) Complement factor H: sequence analysis of 221 kb of human genomic DNA containing the entire fH, fHR-1 and fHR-3 genes. *Mol Immunol* 37:41–52
- Manolio TA et al (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753. <https://doi.org/10.1038/nature08494>
- Martinez-Barricarte R et al (2012) Relevance of complement factor H-related 1 (CFHR1) genotypes in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 53:1087–1094. <https://doi.org/10.1167/iovs.11-8709>
- McKusick VA, Hostetler JA, Egeland JA (1964) Genetic studies of the Amish, background and potentialities. *Bull Johns Hopkins Hosp* 115:203–222
- Merle BM, Benlian P, Puche N, Bassols A, Delcourt C, Souied EH, Nutritional AMDTSG (2014) Circulating omega-3 Fatty acids and neovascular age-related macular degeneration. *Invest Ophthalmol Vis Sci* 55:2010–2019. <https://doi.org/10.1167/iovs.14-13916>
- Morton NE (1955) Sequential tests for the detection of linkage. *Am J Hum Genet* 7:277–318
- Munch IC, Linneberg A, Larsen M (2013) Precursors of age-related macular degeneration: associations with physical activity, obesity, and serum lipids in the inter99 eye study. *Invest Ophthalmol Vis Sci* 54:3932–3940. <https://doi.org/10.1167/iovs.12-10785>
- Naj AC et al (2013) Genetic factors in nonsmokers with age-related macular degeneration revealed through genome-wide gene-environment interaction analysis. *Ann Hum Genet* 77:215–231. <https://doi.org/10.1111/ahg.12011>
- Nowak M et al (2005) Changes in lipid metabolism in women with age-related macular degeneration. *Clin Exp Med* 4:183–187. <https://doi.org/10.1007/s10238-004-0054-z>
- Parmar T, Parmar VM, Perusek L, Georges A, Takahashi M, Crabb JW, Maeda A (2018) Lipocalin 2 plays an important role in regulating inflammation in retinal degeneration. *J Immunol* 200:3128–3141. <https://doi.org/10.4049/jimmunol.1701573>

- Paun CC et al (2015) Genetic variants and systemic complement activation levels are associated with serum lipoprotein levels in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 56:7766–7773. <https://doi.org/10.1167/iovs.15-17035>
- Pescosolido N, Barbato A, Pascarella A, Giannotti R, Genzano M, Nebbio M (2014) Role of protease-inhibitors in ocular diseases. *Molecules* 19:20557–20569. <https://doi.org/10.3390/molecules191220557>
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909. <https://doi.org/10.1038/ng1847>
- Purcell S et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575. <https://doi.org/10.1086/519795>
- Raychaudhuri S et al (2011) A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat Genet* 43:1232–1236. <https://doi.org/10.1038/ng.976>
- Riazuddin SA, McGlumphy EJ, Yeo WS, Wang J, Katsanis N, Gottsch JD (2011) Replication of the TCF4 intronic variant in late-onset Fuchs corneal dystrophy and evidence of independence from the FCD2 locus. *Invest Ophthalmol Vis Sci* 52:2825–2829. <https://doi.org/10.1167/iovs.10-6497>
- Roh MI, Kim JH, Byeon SH, Koh HJ, Lee SC, Kwon OW (2008) Estimated prevalence and risk factor for age-related maculopathy. *Yonsei Med J* 49:931–941. <https://doi.org/10.3349/yjm.2008.49.6.931>
- SanGiovanni JP, Chew EY (2005) The role of omega-3 long-chain polyunsaturated fatty acids in health and disease of the retina. *Prog Retin Eye Res* 24:87–138. <https://doi.org/10.1016/j.preteyeres.2004.06.002>
- SanGiovanni JP et al (2008) The relationship of dietary omega-3 long-chain polyunsaturated fatty acid intake with incident age-related macular degeneration—AREDS report no. 23. *Arch Ophthalmol* 126:1274–1279. <https://doi.org/10.1001/archophth.126.9.1274>
- Sardell RJ et al (2016) Heritability of choroidal thickness in the Amish. *Ophthalmology* 123:2537–2544. <https://doi.org/10.1016/j.ophtha.2016.09.001>
- Scavallo GS Jr, Paluru PC, Zhou J, White PS, Rappaport EF, Young TL (2005) Genomic structure and organization of the high grade Myopia-2 locus (MYP2) critical region: mutation screening of 9 positional candidate genes. *Mol Vis* 11:97–110
- Seddon JM (2013) Genetic and environmental underpinnings to age-related ocular diseases. *Invest Ophthalmol Vis Sci* 54:ORSF28–ORSF30. <https://doi.org/10.1167/iovs.13-13234>
- Seddon JM, Santangelo SL, Book K, Chong S, Cote J (2003) A genome-wide scan for age-related macular degeneration provides evidence for linkage to several chromosomal regions. *Am J Hum Genet* 73:780–790. <https://doi.org/10.1086/378505>
- Seddon JM, George S, Rosner B (2006a) Cigarette smoking, fish consumption, omega-3 fatty acid intake, and associations with age-related macular degeneration: the US Twin Study of Age-Related Macular Degeneration. *Arch Ophthalmol* 124:995–1001. <https://doi.org/10.1001/archophth.124.7.995>
- Seddon JM, Sharma S, Adelman RA (2006b) Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology* 113:260–266. <https://doi.org/10.1016/j.ophtha.2005.11.001>
- Semba RD et al (2014) Serum carboxymethyllysine, an advanced glycation end product, and age-related macular degeneration: the Age, Gene/Environment Susceptibility-Reykjavik Study. *JAMA Ophthalmol* 132:464–470. <https://doi.org/10.1001/jamaophthalmol.2013.7664>
- Shalev V, Srur M, Goldshtein I, Kokia E, Chodick G (2011) Statin use and the risk of age related macular degeneration in a large health organization in Israel. *Ophthalmol Epidemiol* 18:83–90. <https://doi.org/10.3109/09286586.2011.560746>
- Shaw PX, Stiles T, Douglas C, Ho D, Fan W, Du H, Xiao X (2016) Oxidative stress, innate immunity, and age-related macular degeneration. *AIMS Mol Sci* 3:196–221. <https://doi.org/10.3934/molsci.2016.2.196>
- Shellenberger TD et al (2005) Headpin: a serpin with endogenous and exogenous suppression of angiogenesis. *Cancer Res* 65:11501–11509. <https://doi.org/10.1158/0008-5472.CAN-05-2262>
- Skeie JM, Mahajan VB (2014) Proteomic landscape of the human choroid-retinal pigment epithelial complex. *JAMA Ophthalmol* 132:1271–1281. <https://doi.org/10.1001/jamaophthamol.2014.2065>
- Smith W et al (2001) Risk factors for age-related macular degeneration: pooled findings from three continents. *Ophthalmology* 108:697–704
- Sobrin L et al (2012) Heritability and genome-wide association study to assess genetic differences between advanced age-related macular degeneration subtypes. *Ophthalmology* 119:1874–1885. <https://doi.org/10.1016/j.ophtha.2012.03.014>
- Sofat R et al (2012) Complement factor H genetic variant and age-related macular degeneration: effect size, modifiers and relationship to disease subtype. *Int J Epidemiol* 41:250–262. <https://doi.org/10.1093/ije/dyr204>
- Strauss KA, Puffenberger EG (2009) Genetics, medicine, and the Plain people. *Annu Rev Genom Hum Genet* 10:513–536. <https://doi.org/10.1146/annurev-genom-082908-150040>
- Suzuki K, Lareyre JJ, Sanchez D, Gutierrez G, Araki Y, Matusik RJ, Orgebin-Crist MC (2004) Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2. *Gene* 339:49–59. <https://doi.org/10.1016/j.gene.2004.06.027>
- Thalamuthu A et al (2011) Association of TCF4 gene polymorphisms with Fuchs' corneal dystrophy in the Chinese. *Invest Ophthalmol Vis Sci* 52:5573–5578. <https://doi.org/10.1167/iovs.11-7568>
- Thornton T, McPeck MS (2010) ROADTRIPS: case-control association testing with partially or completely unknown population and pedigree structure. *Am J Hum Genet* 86:172–184. <https://doi.org/10.1016/j.ajhg.2010.01.001>
- von Zglinicki T, Martin-Ruiz CM (2005) Telomeres as biomarkers for ageing and age-related diseases. *Curr Mol Med* 5:197–203
- Wang L, Clark ME, Crossman DK, Kojima K, Messinger JD, Mobley JA, Curcio CA (2010) Abundant lipid and protein components of drusen. *PLoS One* 5:e10329. <https://doi.org/10.1371/journal.pone.0010329>
- Wang JC, Ku HY, Chen TS, Chuang HS (2017) Detection of low-abundance biomarker lipocalin 1 for diabetic retinopathy using optoelectrokinetic bead-based immunosensing. *Biosens Bioelectron* 89:701–709. <https://doi.org/10.1016/j.bios.2016.11.014>
- Wiggs JL et al (2012) Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. *PLoS Genet* 8:e1002654. <https://doi.org/10.1371/journal.pgen.1002654>
- Winkler TW et al (2018) Investigating the modulation of genetic effects on late AMD by age and sex: lessons learned and two additional loci. *PLoS One* 13:e0194321. <https://doi.org/10.1371/journal.pone.0194321>
- Young TL et al (1998) Evidence that a locus for familial high myopia maps to chromosome 18p. *Am J Hum Genet* 63:109–119. <https://doi.org/10.1086/301907>
- Zhang M, Volpert O, Shi YH, Bouck N (2000) Maspin is an angiogenesis inhibitor. *Nat Med* 6:196–199. <https://doi.org/10.1038/72303>
- Zhou T, Hu Y, Chen Y, Zhou KK, Zhang B, Gao G, Ma JX (2010) The pathogenic role of the canonical Wnt pathway in age-related macular degeneration. *Invest Ophthalmol Vis Sci* 51:4371–4379. <https://doi.org/10.1167/iovs.09-4278>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.