GENERAL ARTICLE

Family-based exome sequencing identifies rare coding variants in age-related macular degeneration

Rinki Ratnapriya[1,2,†,‡], İlhan E. Acar[3,†], Maartje J. Geerlings[3], Kari Branham[4], Alan Kwong[5], Nicole T.M. Saksens[3], Marc Pauper[3], Jordi Corominas[3], Madeline Kwicklis[1], David Zipprer[1], Margaret R. Starostik[1], Mohammad Othman[4], Beverly Yashar[4], Goncalo R. Abecasis[5], Emily Y. Chew[1], Deborah A. Ferrington[6], Carel B. Hoyng[3], Anand Swaroop[1,‡] and Anneke I. den Hollander[3,‡,*]

[1]Neurobiology, Neurodegeneration and Repair Laboratory (NNRL), National Eye Institute, Bethesda, MD 20892, USA, [2]Department of Ophthalmology, Baylor College of Medicine, Houston, TX 77030, USA, [3]Department of Ophthalmology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen 6500, The Netherlands, [4]Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI 48105, USA, [5]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA and [6]Department of Ophthalmology and Visual Neurosciences, University of Minnesota, Minneapolis, MN 55455, USA

*To whom correspondence should be addressed at: Department of Ophthalmology, Radboud University Medical Center, Philips van Leydenlaan 15, Route 409, Nijmegen 6525 EX, The Netherlands; Email: anneke.denhollander@radboudumc.nl

Abstract

Genome-wide association studies (GWAS) have identified 52 independent variants at 34 genetic loci that are associated with age-related macular degeneration (AMD), the most common cause of incurable vision loss in the elderly worldwide. However, causal genes at the majority of these loci remain unknown. In this study, we performed whole exome sequencing of 264 individuals from 63 multiplex families with AMD and analyzed the data for rare protein-altering variants in candidate target genes at AMD-associated loci. Rare coding variants were identified in the *CFH*, *PUS7*, *RXFP2*, *PHF12* and *TACC2* genes in three or more families. In addition, we detected rare coding variants in the *C9*, *SPEF2* and *BCAR1* genes, which were previously suggested as likely causative genes at respective AMD susceptibility loci. Identification of rare variants in the *CFH* and *C9* genes in our study validated previous reports of rare variants in complement pathway genes in AMD. We then extended our exome-wide analysis and identified rare protein-altering variants in 13 genes outside the AMD-GWAS loci in three or more families. Two of these genes, *SCN10A* and *KIR2DL4*, are of interest because variants in these genes also showed association with AMD in case-control cohorts, albeit not at the level of genome-wide significance. Our study presents the first large-scale, exome-wide analysis of rare variants in AMD. Further independent replications and molecular investigation of candidate target genes, reported here, would assist in gaining novel insights into mechanisms underlying AMD pathogenesis.

Introduction

The genetic basis of common diseases continues to be a subject of significant interest in human genetics research in the post-genome era. Genome-wide association studies (GWAS) have been successful in uncovering hundreds of common genetic variants associated with multiple complex traits (1). Despite representing a major advance in deciphering genetic structure of complex diseases, GWAS have faced two major setbacks. First, the associated variants often reside in non-coding regions of the genome, representing a tag SNP (single nucleotide polymorphism) for the causal variant(s). Thus, the causal gene(s) and genetic variant(s) are not immediately evident for a majority of identified loci. Second, barring few exceptions, the associated variants identified by GWAS explain only a small proportion of the disease heritability (2). These observations have led to shifts in paradigms including possible involvement of rare variants in complex traits. In fact, both common and rare variants have been recognized to contribute to the genetic architecture of multifactorial diseases (3). Nevertheless, common variants have long been central to these investigations, primarily because of the identification of millions of SNPs across the genome that could be genotyped in a high-throughput manner (4,5). The advent of next-generation sequencing technologies and reduced cost of sequencing have prompted a new wave of studies for evaluating the impact of rare and low-frequency variants in complex diseases (6).

Age-related macular degeneration (AMD; MIM, 603075) is a late-onset multifactorial, neurodegenerative disease, which affects the macular region of the retina, eventually resulting in the loss of central vision. AMD afflicts 8.7% of the worldwide population (7) with over 10 million individuals in the United States alone, posing a substantial healthcare burden (8). The clinical presentation of the disease is progressive and can be broadly divided into early, intermediate and late (advanced) stage (9). Early and intermediate stages are mostly asymptomatic and characterized by changes in the retinal pigment epithelium (RPE) and accumulation of extracellular aggregates, called drusen. Central vision loss characterizes the advanced stage, which can be further subdivided into two forms: geographic atrophy (GA) and choroidal neovascularization (CNV) (10). Therapeutic interventions are available for many patients with CNV, which represents a small proportion of all AMD cases (11).

AMD is a complex trait determined by a combination of advanced age, genetic variants and environmental factors such as smoking and nutrition (12,13). Genetic studies in AMD have contributed significantly to biological underpinnings of its pathogenesis (14,15). The most recent GWAS of 16,144 patients and 17,832 control individuals identified 52 independently associated variants at 34 loci that explain over 50% of the genomic heritability (16). These findings have highlighted the complement system, extracellular matrix remodeling and lipid metabolism in AMD pathogenesis. Rare variants are also suggested to contribute to the heritability in AMD (16,17). Targeted sequencing of candidate genes and whole exome sequencing (WES) have identified rare coding variants in the *CFH*, *C3*, *CFI*, *C9*, *CFB* and *COL8A1* genes at GWAS loci (18–24). Additional studies have established an unequivocal role of the complement system in AMD pathology (25–30). At this stage, the contribution of rare variants outside of the GWAS loci remains largely unexplored due to power issues, as case-control designs require sequencing of large numbers of samples to achieve statistical significance.

Early investigations indicating the aggregation of AMD in families strongly indicated a strong genetic contribution (31–34), and such families can serve as a reference for identifying rare variants and causal genes. Linkage analysis in AMD families had initially led to the identification of the 1q31 and 10q26 loci (35–38), which were later identified as the top two association signals in GWAS and/or by candidate gene association at the *CFH* (39–43) and *ARMS2-HTRA1* (44–46) loci, respectively. In addition, exome sequencing in large, multiplex families identified a rare variant in the *FBN2* gene (47) in the early form of macular degeneration and confirmed the occurrence of rare variants in the *CFH*, *CFI*, *C9* and *C3* genes at multiple AMD loci (21,28,29,48–52). Thus, AMD families can provide new insights into AMD by identifying additional rare variants and pointing to candidate causal genes.

In this study, we performed WES of 264 individuals from 63 multiplex families. To determine causality of genes in loci previously identified by GWAS, we analyzed these families for rare protein-altering variants in coding regions at AMD-associated loci. In addition, we examined whole exome data for rare coding variants outside the AMD-associated loci to identify new AMD candidate genes.

Results

Whole exome sequencing in AMD families

We assembled a cohort of 226 affected and 38 unaffected individuals from 63 multiplex families with a history of AMD (Table 1). A majority of the cohort (51.5%) had advanced stages of the disease (GA, CNV or GA + CNV) with a mean age of 78.2 years. Individuals with early and intermediate stages had a mean age of 69.1 years and constituted 34% of the affected individuals (Table 1). We also included 38 unaffected members (age > 55 years) from 32 families with a mean age of 72.1 years. Exome capture and sequencing was performed at two different centers [at the National Eye Institute (NEI) for the NEI families and at ErasmusMC Rotterdam for the Radboudumc families] (Fig. 1). As the WES data were captured on different platforms, we extracted variants from the shared 58.55 Mb exome regions. WES data were analyzed following the guidelines of GATK best practices (Fig. 1), resulting in the identification of 107,348 coding variants (53). We identified 4957 variants after additional, multi-tiered filtering based on quality, allele depth, minor allele frequency (MAF) in matched controls and predicted effect of the variants (Fig. 1). The variants segregating in a family were considered only if the variant was shared among 80% or more affected family members and/or segregated in all but one affected individual. We excluded the variants that were present in available unaffected member(s) of the respective family.

Survey of rare coding variants in genes within GWAS loci

We first analyzed the families for rare (MAF < 1%) coding variants within 34 AMD-associated loci (Table 2). The locus regions were defined by genome-wide significant variants and those within linkage disequilibrium (LD) ($r^2 > 0.5$), with an additional 500 kb to both sides, as described in the most recent AMD-GWAS study (16). We identified 91 rare variants in 78 genes within the LD intervals of 29 known AMD loci and initially focused on loci where rare coding variants were identified in three or more families. Variants were detected in the *CFH*, *PUS7*, *RXFP2* and *PHF12* genes in three families and in the *TACC2* gene in four families (Fig. 2A–D and F; Table 2). We also included genes that

Table 1. Overview of families analyzed

|	Radboudumc cohort	NEI cohort
Number of families	44	19
Total number of individuals	195	69
Males, females	73,122	30, 39
Number of individuals with advanced AMD	82	54
Mean age for advanced AMD	76.5	80.8
Number of individuals with early or intermediate AMD	82	8
Mean age for early or intermediate AMD	68.5	74.8
Number of unaffected individuals	31	7
Mean age of unaffected individuals	70.8	77.5
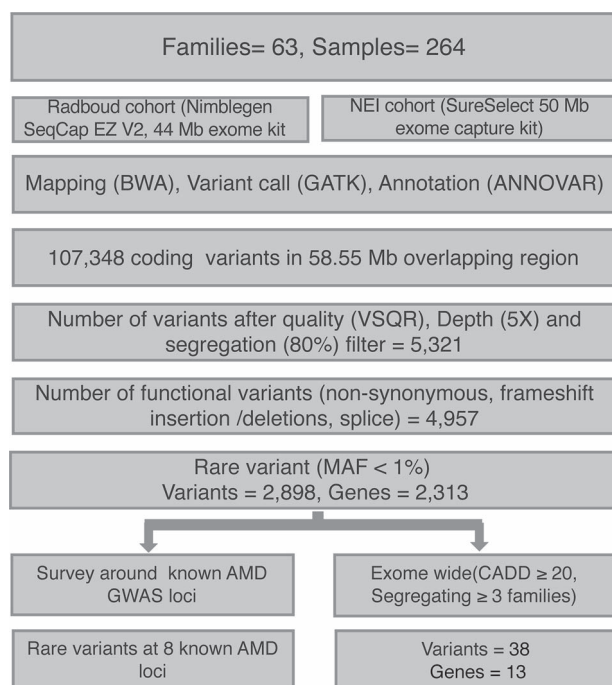
Figure 1. An overview of stepwise filtering of variants identified in multiplex AMD families using whole exome sequencing. We retained variants that were shared among 80% or more affected family members or segregated in all but one affected individual. Variants were discarded if they were present in one or more unaffected family members.

were predicted as most likely candidates using the gene priority score (GPS), as described previously (16). The scoring was based on expression of genes in AMD-relevant tissues (retina, RPE and choroid) and the presence of rare variants in AMD patients and their known biological functions. By combining our WES results with GPS score, we identified three additional likely AMD genes: *SPEF2* (PRLR-SPEF2 locus), *BCAR1* (CTRB2-CTRB1 locus) and *C9* (C9 locus) (Table 2). Two independent non-synonymous variants (R201G and R421H) in *SPEF2* were predicted to be pathogenic and identified in two families (Fig. 2G). No additional segregating rare variants were uncovered among the 16 genes within the LD region of the *PRLR-SPEF2* locus. A non-synonymous and splice variant in *BCAR1* is segregated in two families (Fig. 2H). In addition, a single family with four affected members revealed the segregation of a rare, likely pathogenic variant P167S in *C9* (Fig. 2D).

We next analyzed the expression of the eight genes where rare variants were identified in the retinal transcriptome and expression quantitative trait loci (eQTL) dataset that included both control and AMD donor samples (54) (Fig. 3). Six of these genes (*CFH*, *PUS7*, *PHF12*, *TACC2*, *SPEF2* and *BCAR1*) were expressed in the retina (counts per million, CPM > 0), while no retinal expression was detected for *RXPF2* and *C9* (Fig. 3A). Additionally, we evaluated gene expression changes during AMD progression by comparing the gene expression profile of the control retina with the profiles of the retina during early, intermediate and late stages of the disease. While we did not observe a significant difference [at false discovery rate (FDR) $\leq 10\%$] during disease progression, several candidates showed a trend of either increased or decreased expression (Fig. 3B). We note that expression of four of the candidates (*PUS7*, *PHF12*, *SPEF2*, *BCAR1*) is regulated in the retina through common eQTL variants (Fig. 3C).

Exome-wide analysis of rare pathogenic variants

We then extended our analysis in search of rare pathogenic variants in genes outside the GWAS loci. We selected genes with rare (MAF < 1%) variants that were detected in at least three families and applied an additional predictive causality filter to focus on the 1% most deleterious variants in the human genome (CADD score ≥ 20). We also removed the variants that did not pass the quality score filter in Exome Aggregation Consortium (ExAC) database. Our analyses resulted in the identification of 13 candidate genes with 38 rare variants segregating in three or more families (Table 3). As large or polymorphic genes harbor a large number of variants, we further prioritized the list based on the following criteria: (i) whether the gene was expressed in the human retina (CPM > 0) and is regulated by an eQTL, (ii) whether common variants around the candidate gene exhibited a suggestive association (P-value < 5×10^{-4}) for AMD and (iii) whether the candidate gene or a closely related gene was shown to have a function relevant to AMD pathology. Ten candidate genes demonstrated expression in the human retina with a few showing changes during AMD progression (Fig. 4A and B). Five of the genes also had an eQTL in the retina (Fig. 4C). One candidate, *SCN10A*, showed a modest association (3:38800182_C/T; P-value: 9.63×10^{-5}) in the AMD-GWAS (Fig. 4D). We identified three different rare variants in *SCN10A*, one non-sense variant (Q923X) and two non-synonymous variants (P893L and P332L) in three independent AMD families (Fig. 5A). A rare frameshift variant (S267fs) in *KIR2DL4* is segregated in three families (Fig. 5B). Based on their role in mitochondrial biogenesis, *ESRRA* and *VPS13B*, each with rare variants in three different families, could be presented as additional interesting candidates (Fig. 5C and D).

Table 2. Rare coding variants (MAF < 1%) in genes within GWAS loci identified by WES

Locus ([a] genes within LD ± 500 kb)	Gene	Family ID	Chr: Pos	rs Id	MAF_NFE	Variants	Exonic function	Protein change
Genes with rare variants in three or more families								
CFH (13)	*CFH*	AMD580	1:196621252	rs142266551	0	NM_000186: exon1:c.5G > C	Non-synonymous	R2T
	CFH	W11-2310_2	1:196646702	—	—	NM_000186: exon5:c.524G > A	Non-synonymous	R175Q
	CFH	W10-0408_1	1:196654311	—	0	NM_000186: exon7:c.908G > A	Non-synonymous	R303Q
KMT2E/SRPK2 (8)	*PUS7*	AMD930	7:105148683	rs139058270	0.006	NM_001318163: exon1:c.277 T > C	Non-synonymous	C93R
	PUS7	AMD479	7:105148683	rs139058270	0.006	NM_001318163: exon1:c.277 T > C	Non-synonymous	C93R
	PUS7	W08–0553	7:105148893	—	—	NM_001318163: exon1:c.67A > G	Non-synonymous	S23G
B3GALTL (7)	*RXFP2*	W11–4656	13:32351535	rs121918303	0.0072	NM_001166058: exon8:c.664A > C	Non-synonymous	T222P
	RXFP2	AMD580	13:32352714	rs73163317	0.0078	NM_001166058: exon9:c.779A > G	Non-synonymous	N260S
	RXFP2	W07–0199	13:32367033	rs138951290	0.0014	NM_001166058: exon15:c.1522C > G	Non-synonymous	R508G
TMEM9/VTN (38)	*PHF12*	W11-1525_1	17:27238135	rs148347485	0.0083	NM_001033561: exon10:c.2210A > G	Non-synonymous	N737S
	PHF12	AMD580	17:27251125	—	—	NM_001033561: exon4:c.517A > T	Non-synonymous	T173S
	PHF12	AMD930	17:27238135	rs148347485	0.0083	NM_001033561: exon10:c.2210A > G	Non-synonymous	N737S
ARMS2/HRTA1 (15)	*TACC2*	W09–1832	10:123844894	rs112188313	0.0068	NM_001291876: exon4:c.2879G > A	Non-synonymous	R960K
	TACC2	W10-0408_1	10:123954596	—	—	NM_001291878: exon2:c.110C > T	Non-synonymous	T37M
	TACC2	W11-1525_2	10:123844894	rs112188313	0.0068	NM_001291876: exon4:c.2879G > A	Non-synonymous	R960K
	TACC2	AMD56	10:123809983	rs202197379	—	NM_001291876: exon3:c.64G > A	Non-synonymous	A22T
Genes with highest GPS[a]								
C9 (6)	*C9*	W11–4035	5:39331894	rs34882957	0.0066	NM_001737: exon5:c.499C > T	Non-synonymous	P167S
PRLR/SPEF2 (16)	*SPEF2*	AMD930	5:35646784	rs80010329	0.0066	NM_024867: exon5:c.601A > G	Non-synonymous	R201G
	SPEF2	AMD393	5:35667268	rs139580877	0.0094	NM_024867: exon9:c.1262G > A	Non-synonymous	R421H
CTRB2/CTRB1 (20)	*BCAR1*	AMD930	16:75268977	rs61743104	0.01	NM_001170721: exon4:c.1190G > A	Non-synonymous	R397Q
	BCAR1	W11–4044	16:75286131	rs74024754	0.0029	NM_001170720: exon2:c.12 + 2 T > C	Splicing	—

[a]GPS refers to gene priority score as described in Fritsche et al. (16), which was a scoring-based method for suggesting most likely candidate at known AMD loci based on expression, presence of rare variants and known biological function that were deemed important for AMD. Minor allele frequency (MAF) was from non-Finnish Europeans from ExAC database.

Given that segregation of rare variants in small families can be detected by chance, we performed simulation analysis to test the null case of no variant effect on disease by randomly assigning alleles to subjects in our study and then calculating the probability of observing segregation in three or more families. We applied the segregation criteria similar to the criteria that were used to identify variants in the AMD families and performed 100 000 gene-level simulations for two different variant sets: one with rare variants identified at the AMD loci and the other with exome-wide, rare variants after filtering by CADD score. We observed an average gene-level type I error rate of 7.7×10^{-7} for the AMD loci analysis and 1.87×10^{-7} for the CADD score filtered exome-wide analysis. The study-level type I error was 4.6×10^{-4} (AMD loci) and 2.2×10^{-4} (exome-wide). Given the pedigree structure and variants identified in our study, it is therefore unlikely to detect segregation of rare variants in three or more families by chance.

Rare variant association in case-control cohort

We further tested whether candidate genes identified in this study exhibit rare variant associations in AMD case-control studies that employed WES or whole genome sequencing (WGS) to genotype rare variants in coding regions of the genome. We analyzed the burden of rare coding variants in the eight genes within GWAS loci (Table 2) and 13 genes identified in our exome-wide analysis (Table 3) in two different cohorts (see Methods) (24,30). A meta-analysis of rare variants in 3519 AMD subjects and 3754 controls suggested a burden of rare variants in the *C9* ($P = 0.024$) and *KIR2DL4* ($P = 0.035$) genes using the burden and variable threshold (VT) tests (Supplementary Material, Table S1).

Discussion

GWAS and deep sequencing efforts have strongly supported a role for common and rare variants in AMD susceptibility. In
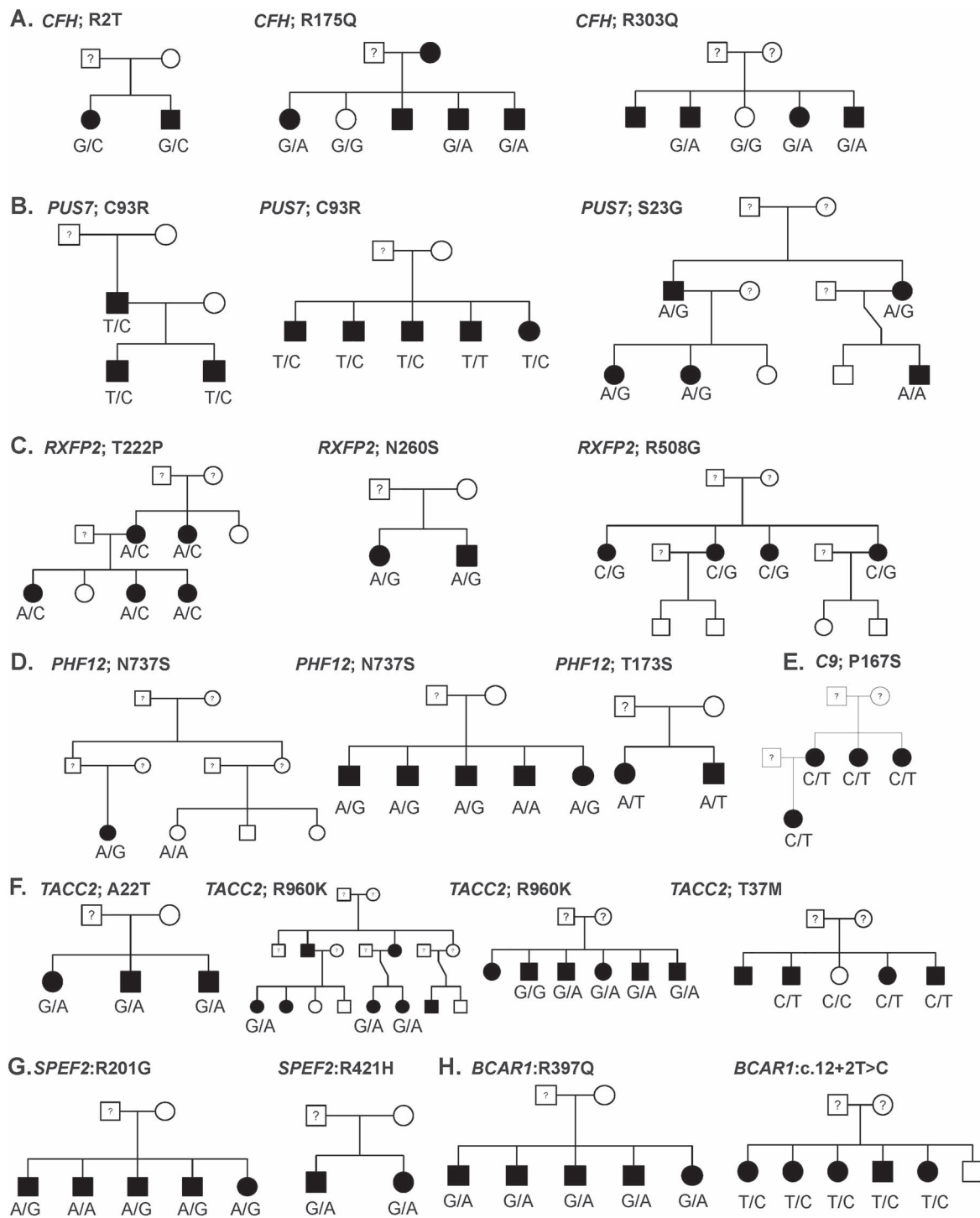
Figure 2. Segregation of rare variants in genes residing within AMD-associated loci. Filled symbols represent the individuals with AMD.

addition to conferring high risk for AMD, rare variant studies can provide important insights into the phenotypic characteristics of the carrier patients (55,56). The most recent GWAS examined both common and rare variants in a large cohort but was limited to the analyses of known rare variants that were present on the chip (16) and did not allow sequencing-based discovery of rare variants. A few sequencing studies attempted to harness the power of familial cases, but their search was restricted to known AMD loci (28,49). Here, we present the first, large-scale,

genome-wide survey of rare coding variants in multiplex AMD families that identified several novel candidate genes/variants contributing to AMD susceptibility.

Within known AMD susceptibility regions, we identified rare variants at eight loci; two of these in *CFH* and *C9* have previously been reported in case-control and familial studies (18,20). These findings further strengthen the role of the complement pathway in AMD and validate the presence of highly penetrant, functional rare variants as causal variants at these loci. Our study suggested
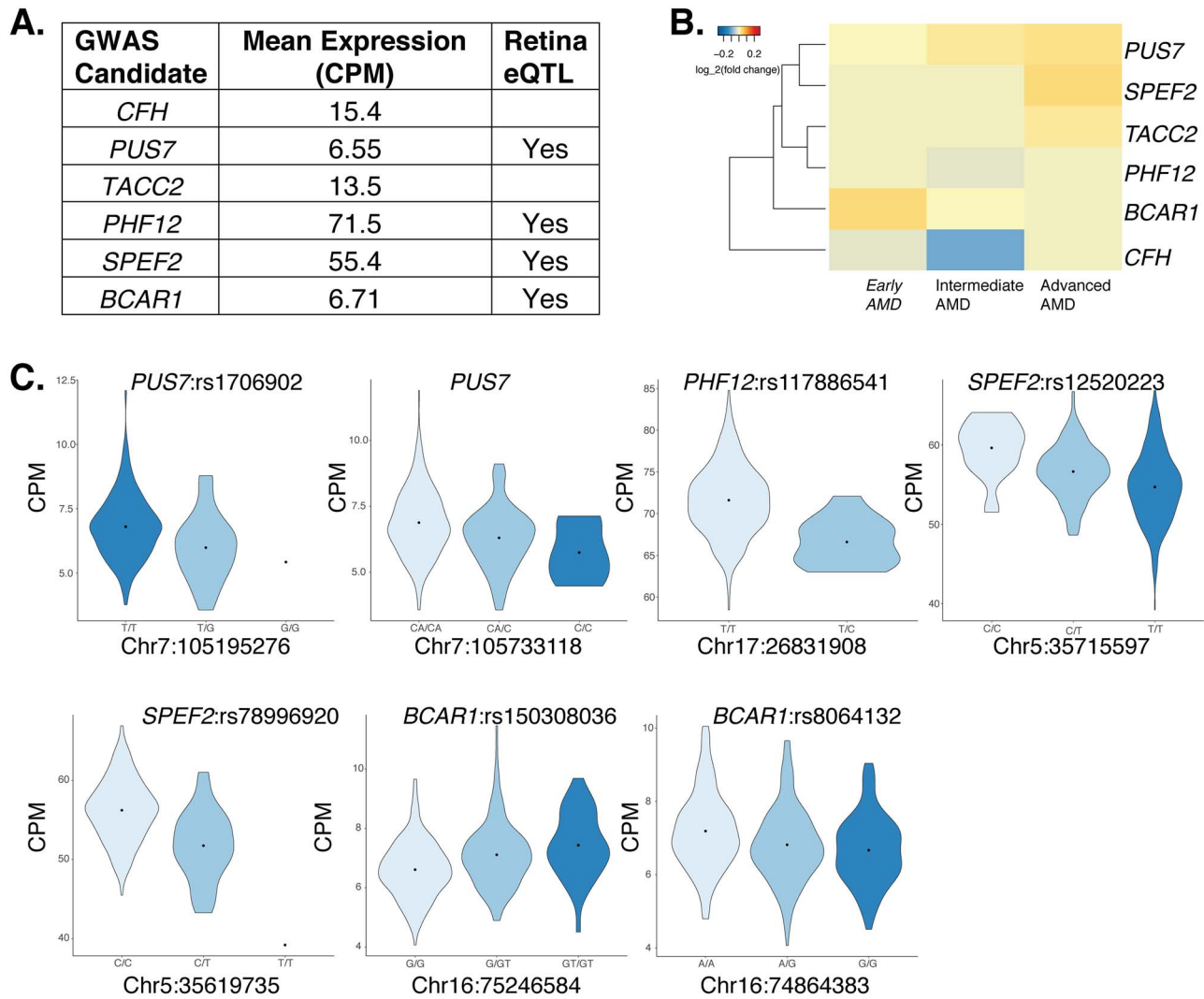
Figure 3. (A) Mean expression of candidate genes within AMD-associated loci in human donor retina. (B) Heatmap showing the fold change differences observed in genes during early-, intermediate- and late-stage AMD compared to normal retina. None of the candidate reaches statistical significance of FDR ≤ 10%. (**C**) Violin plots showing the relationship between the variant and the gene of an observed eQTL. Fewer individuals with homozygous genotypes of *PUS7* (G/G) and *SPEF2* (T/T) resulted in single point in violin plots. P-values for the eVariants: (rs170690–3.97×10^{-5}), (7:105733118–1.39×10^{-5}), (rs117886541–1.12×10^{-5}), (rs78996920–3.85×10^{-10}), (rs12520223–1.75×10^{-8}), (rs150308036–3.15×10^{-7}), (rs8064132–2.04×10^{-5}).

the causal genes at other AMD loci. *PUS7* encodes a key RNA-modifying enzyme. *RXFP2* encodes a receptor for glycoprotein hormones that contains a unique low-density lipoprotein type A (LDLa) module (57); though this gene is not expressed in the adult human retina (54), its homolog in zebrafish is transcribed in the developing brain and retina (58). *PHF12* encodes a member of the PHD zinc finger family of proteins involved in regulation of ribosomal biogenesis and senescence (59) and could contribute to AMD pathology through activation of senescence. Although present in only two families, the findings of rare variants in *SPEF2* and *BCAR1* are also noteworthy as both of these candidates were highlighted as most likely candidate at the AMD-GWAS loci based on a gene priority score (16). *SPEF2* is highly expressed in the human retina and shows a slight increase in advanced stages of AMD. SPEF2 is a cilia-related protein and its absence causes male infertility and primary ciliary dyskinesia (60). This protein is shown to interact with IFT-related protein, IFT20 (61) that is required for opsin trafficking and photoreceptor outer segment development (62). *BCAR1* encodes a Src family kinase substrate

that is shown to be involved in early retinal development, and a mouse model of *Bcar1* exhibits dramatic disruption of the ganglion cell layer (63). Recently, *BCAR1* was also identified as a putative causal gene at the *CTRB2/CTRB1* locus by co-localization of GWAS and single-cell eQTL data (64).

Our survey of candidate genes harboring rare variants outside AMD loci was aimed at identifying potential novel candidate genes underlying AMD pathology. Such candidates could have been missed in GWAS and other association studies if the gene confers the risk to AMD solely through rare variants, or if the effect size of the common risk variant is very low, warranting larger numbers of cases and controls to detect such associations. We identified 13 candidate genes that harbor rare variants in more than three families and show, by simulation studies, that the probability of these variants segregating by chance is very low. Thus, these genes represent reasonable candidates for AMD susceptibility. We propose *SCN10A* and *KIR2DL4* as the most interesting AMD candidates. Common variants at the *SCN10A* gene have been associated with AMD in a recent AMD-GWAS

Table 3. Candidates with rare coding variants (MAF < 1%) segregating in three or more families identified in exome-wide analysis

Gene	Chr: Pos	rs Id	No. of family	MAF_NFE	Exonic function	Nucleotide change	Amino acid change	CADD_phred
CTDSP2	chr12: 58220816	rs76940645	11	0	Non-synonymous	NM_005730: exon4:c.317 T > C	I106T	25.3
TTN	chr2: 179395813	rs55865284	7	0.0001	Non-synonymous	NM_003319: exon186:c.78334G > A	V26112M	22.1
	chr2: 179482937	rs72677232		0.0026	Non-synonymous	NM_003319: exon80:c.20053G > A	V6685I	21.6
	chr2: 179482994	rs72677231		0.0035	Non-synonymous	NM_003319: exon80:c.19996C > T	R6666C	22.2
	chr2: 179486037	rs72677225		0.0065	Non-synonymous	NM_003319: exon74:c.18213G > T	K6071N	22.3
	chr2: 179486223	rs17354992		0.0079	Non-synonymous	NM_003319: exon73:c.18133G > A	D6045N	23.2
	chr2: 179486345	rs114331773		0.0015	Non-synonymous	NM_003319: exon73:c.18011A > T	E6004V	22.8
	chr2: 179537200	rs202014478		0.006	Non-synonymous	NM_133378: exon149:c.30961G > A	V10321I	21.5
	chr2: 179549407	rs72650031		0.0077	Non-synonymous	NM_133378: exon128:c.28892C > T	P9631L	20.1
	chr2: 179582913	rs72648981		0.003	Non-synonymous	NM_133378: exon83:c.21088G > A	E7030K	22.7
	chr2: 179585312	rs17452588		0.008	Non-synonymous	NM_133378: exon77:c.19445C > T	S6482L	22.5
	chr2: 179588622	rs201394117		0.0007	Non-synonymous	NM_133378: exon70:c.17632G > A	A5878T	23.6
KIR3DL3	chr19: 55237616	rs143765860	4	NA	Non-synonymous	NM_153443: exon3:c.168C > A	N56K	NA
	chr19: 55246731	rs602444		NA	Non-synonymous	NM_153443: exon6:c.961C > T	H321Y	NA
DCHS2	chr4: 155156598	rs149548848	4	0.002	Non-synonymous	NM_017639: exon25:c.7841C > T	P2614L	24.9
	chr4: 155411721	rs199621086		0.0042	Non-synonymous	NM_001142552: exon1:c.787C > T	R263W	25.9
	chr4: 155411948	rs184619033		0.0074	Non-synonymous	NM_001142552: exon1:c.560G > T	R187L	24.2
DNAH14	chr1: 225340410	rs17578819	4	0.0094	Non-synonymous	NM_001373: exon32:c.4970G > A	G1657E	21.8
	chr1: 225490915	rs140066130		0.0042	Non-synonymous	NM_001373: exon55:c.8410C > T	R2804C	34
	chr1: 225528175	rs184094753		0.0034	Stopgain	NM_001373: exon67:c.10171G > T	E3391X	58
SCN10A	chr3: 38768123	—	3	0	Stopgain	NM_001293307: exon15:c.2767C > T	Q923X	24.9
	chr3: 38768212	rs138413438		0.0014	Non-synonymous	NM_001293307: exon15:c.2678C > T	P893L	24.7
	chr3: 38798606	—		0	Non-synonymous	NM_001293306: exon8:c.995C > T	P332L	33
ESRRA	chr11: 64083290	rs150848359	3	0.0098	Non-synonymous	NM_001282450: exon7:c.1124G > A	R375Q	22.4
	chr11: 64083293	rs201971362		0.0098	Non-synonymous	NM_001282450: exon7:c.1127G > T	R376L	22.1
ADPRHL1	chr13: 114077172	rs145187729	3	0.0084	Non-synonymous	NM_138430: exon7:c.1030G > A	A344T	20.5
	chr13: 114107571	rs138029763		0.0076	Non-synonymous	NM_138430: exon1:c.182 T > C	M61T	26.2
KIR2DL4	chr19: 55324674	rs11371265	3	NA	frameshift_insertion	NM_001080772: exon6:c.802dupA	S267fs	NA
PLB1	chr2: 28805359	rs144737372	3	0.0013	Non-synonymous	NM_001170585: exon24:c.1687A > G	R563G	22.3
	chr2: 28854972	rs74701215		NA	Non-synonymous	NM_001170585: exon54:c.3934C > A	P1312T	28.1
TAF1C	chr16: 84212875	rs199976567	3	0.0037	Non-synonymous	NM_001243158: exon11:c.1286C > T	S429L	31
	chr16: 84213651	-		NA	Non-synonymous	NM_001243158: exon10:c.604C > A	Q202K	23.3
	chr16: 84215010	-		0	Non-synonymous	NM_001243158: exon7:c.170G > A	R57H	34
UNC80	chr2: 210683829	rs200473652	3	0.0019	Non-synonymous	NM_032504: exon12:c.1806G > C	Q602H	23.1
	chr2: 210707144	rs78912192		0.0003	Non-synonymous	NM_032504: exon21:c.3434A > C	E1145A	27.9
	chr2: 210860221	rs199783352		0.0005	Non-synonymous	NM_182587: exon63:c.9607G > A	E3203K	23.4
VPS13B	chr8: 100147957	rs143205296	3	0.0005	Non-synonymous	NM_015243: exon11:c.1559A > G	H520R	24.8
	chr8: 100443885	rs61753722		0.0091	Non-synonymous	NM_017890: exon22:c.3203C > T	T1068I	24.2

A.

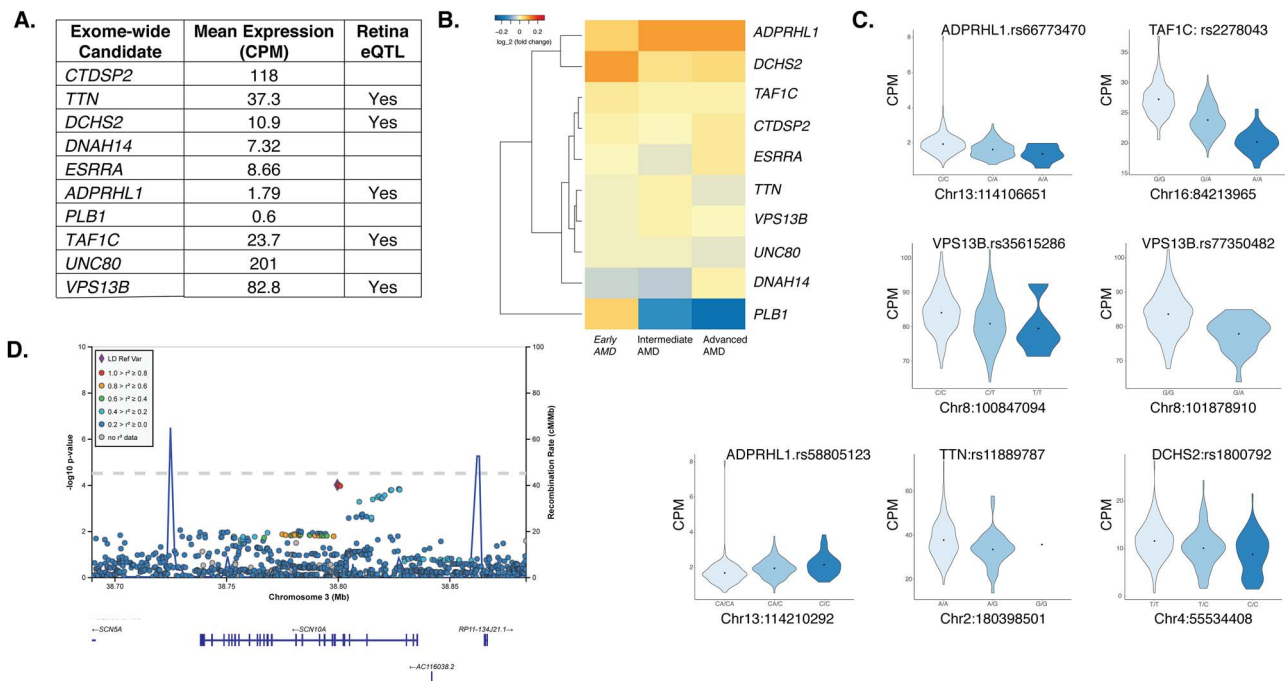| Exome-wide Candidate	Mean Expression (CPM)	Retina eQTL
CTDSP2	118	
TTN	37.3	Yes
DCHS2	10.9	Yes
DNAH14	7.32	
ESRRA	8.66	
ADPRHL1	1.79	Yes
PLB1	0.6	
TAF1C	23.7	Yes
UNC80	201	
VPS13B	82.8	Yes

Figure 4. (**A**) Mean expression of the candidate genes with rare variants (MAF < 1%) in three or more families that are located outside of AMD-associated loci. (**B**) Heatmap showing the fold change differences observed in genes during early-, intermediate- and late-stage AMD compared to normal retina. (**C**) Violin plots showing the relationship between the variant and the gene of an eQTL. Fewer individuals with homozygous genotypes of *TTN* (G/G) resulted in single point in the violin plot. *P*-values for the eVariants: (rs66773470–2.66×10^{-6}), (rs2278043–5.92×10^{-60}), (rs35615286–4×10^{-5}), (rs58805123–2.50×10^{-5}), (rs11889787–4.11×10^{-5}), (rs1800792–1.56×10^{-5}). (**D**) LocusZoom plot (80) showing the association signals around *SCN10A* in most recent GWAS analysis in AMD (16).

(16) though not at the level of genome-wide significance, and a suggestive burden of rare variants in *KIR2DL4* has been detected in a meta-analysis of WGS data in a case-control cohort composed of 3519 cases and 3754 controls (65). SCN10A is a voltage-gated sodium channel expressed in starburst amacrine cells and a subset of retinal ganglion cells (66), which is likely the reason for not observing its expression in the human retina. Gain-of-function mutations in this gene cause axonal degeneration leading to painful neuropathy (67). Several retinal diseases are caused by ion channel mutations including inherited macular degeneration (68). *KIR2DL4*, a member of the human killer cell Ig-like receptor (KIR) family, is expressed primarily in natural killer (NK) cells and plays an important role in innate immunity (69). NK cells interact with HLA class ligand through their KIR receptors and a certain combination of HLA-C, and KIR gene variants have been associated with AMD (70). *ESRRA*, an estrogen-related receptor is a key regulator of energy homeostasis and mitochondrial function. *ESRRA*-null mice show altered regulation of enzymes involved in lipid, eicosanoid and steroid synthesis (71). Dysfunction in lipid metabolism has been linked to AMD pathogenesis (72), making *ESRRA* an attractive candidate for further investigation. *VPS13B* is a transmembrane protein with a function in vesicle-mediated transport. Mutations in this gene cause Cohen syndrome, whose clinical features include non-progressive psychomotor retardation, microcephaly and retinal dystrophy (73).

Our study demonstrates that WES in extended families could prove valuable for identifying rare coding risk variants in potentially novel AMD genes. Similar results could be obtained from traditional case-control studies; however, that would require study designs involving much larger cohorts. A further incorporation of expression data from a large cohort of normal and AMD donors (54) also uncovered new insights. These candidates

do not reach statistical threshold (FDR ≤ 10%) in the differential expression analysis because of the clinical and genetic heterogeneity among AMD and normal individuals and limited power of the study to detect such differences. Investigation of larger cohorts as well as other AMD-relevant tissues such as RPE and choroid will be useful as additional line of evidence in validating these candidates as AMD-causing genes.

In conclusion, our family-based exome sequencing studies identified rare coding variants for novel candidate genes at eight known GWAS loci and 13 additional candidates outside GWAS regions. Further independent replications and molecular investigations of the candidate genes and variants, reported here, could provide novel mechanisms and pathways underlying AMD pathogenesis.

Materials and Methods

Study samples

The study population consisted of 264 samples from 63 multiplex AMD families is collected at two different centers. The Radboudumc cohort (Fig. 1, Table 1) comprised of 44 families (195 subjects). All patients gave written consent and the local ethics committees on research involving human subjects approved the study. Family ascertainment and disease classification for NEI cohort have been described in detail elsewhere (16,37). Briefly, samples were ascertained from the clinical practice at the Kellogg Eye Center. Fundus photographs, fluorescein angiograms and eye examinations records were obtained for all probands and family members and were updated every 1–2 years. The recruitment and research protocols were reviewed and approved by the University of Michigan institutional review board, and informed consent was obtained from all study participants.
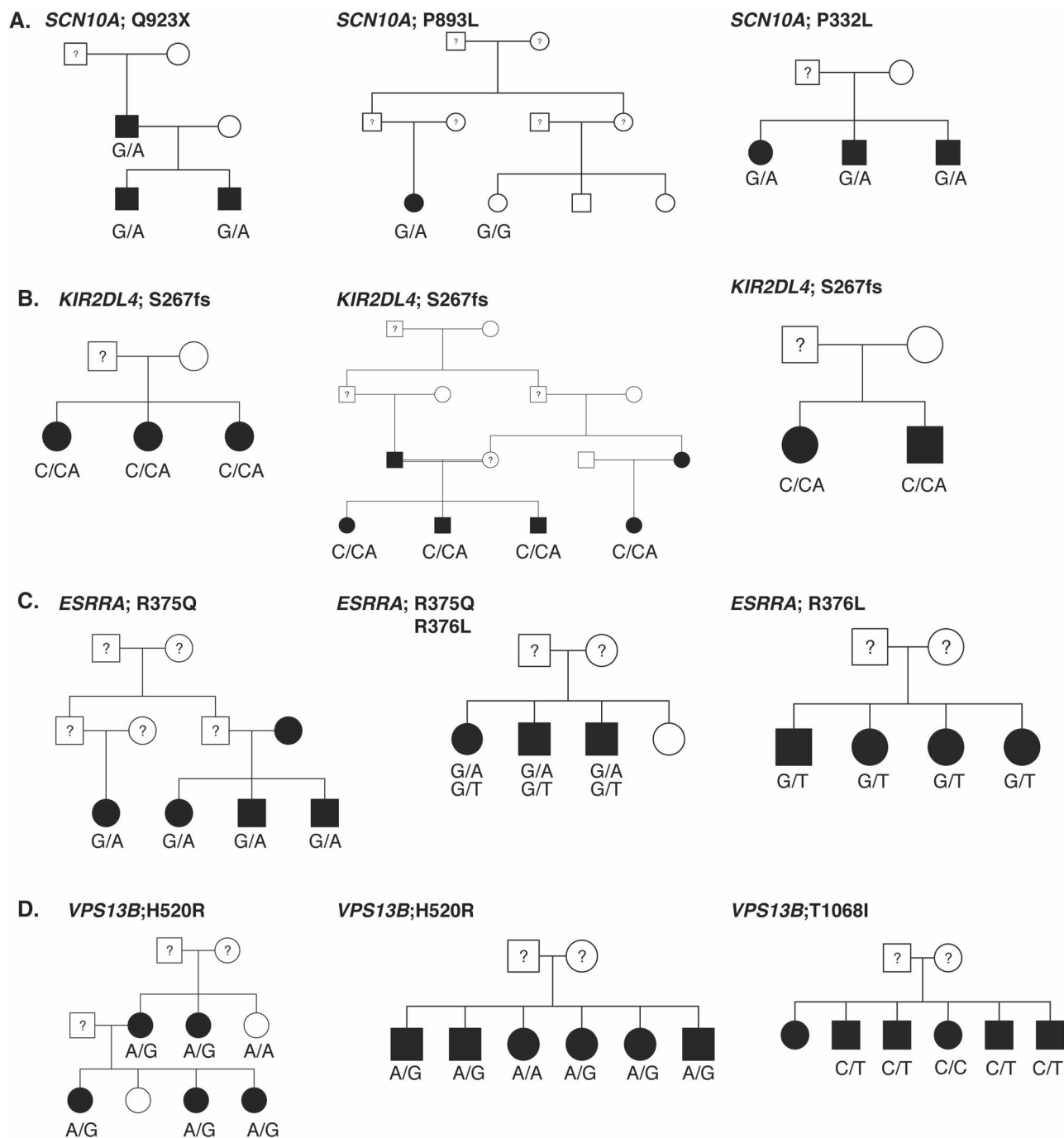
Figure 5. Segregation of the rare variants in *SCN10A* (**A**), *KIR2DL4* (**B**), *ESRRA* (**C**) and *VPS13B* (**D**) in AMD families.

Fundus findings in each eye were classified on the basis of a standardized set of diagnostic criteria established by the International Age-Related Maculopathy Epidemiological Study (74). The Declaration of Helsinki principles was followed for all procedures. The NEI cohort (Fig. 1, Table 1) consisted of 19 families (69 subjects) that were collected at the University of Michigan. All patients signed informed consent, and the Institutional Review Boards of the University of Michigan and the National Eye Institute approved the study.

Whole exome sequencing

Genomic DNA was extracted from the peripheral blood using standard methods. Genomic DNA samples were quantified using the Promega QuantiFluor® dsDNA system (Promega, Madison WI, USA), according to the manufacturer's instructions. The NEI families were subjected to exome sequencing using standard library preparation protocol with Agilent SureSelect Human All exon 50 Mb kit (Agilent Technologies, Santa Clara, CA), following the manufacturer's instructions (75). Captured libraries were

amplified and converted to clusters using Illumina Cluster Station, and single-end 101 bp sequencing was performed on Illumina GAIIx (Illumina, Inc., San Diego, CA). The Nimblegen Seq-Cap EZ Exome v2.0 44 Mb kit (Roche Nimblegen, Inc., Madison, WI) was used for performing WES in the Radboudumc families. Illumina HiSeq2000 sequencer was used to perform paired-end sequencing, using Illumina TruSeq V3 chemistry (Illumina, Inc., San Diego, CA).

Primary bioinformatics analysis

FastQC (available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to confirm quality of sequencing, after which adapter indexes were removed using Trimmomatic (76). Mapping was performed on the hg19/GRCh37 human reference genome sequence build using BWA (77). Aligned reads were processed to mark duplicates using Picard (http://picard.sourceforge.net). The Genome Analysis Toolkit (GATK) recommendations for best practices were applied for variant calling, local realignment, base quality recalibration and variant recalibration (78). The annotation of variants was performed with ANNOVAR (79). Overlapping regions between the capture kits were identified using the mergeBed tools from the bedtools package (80). Alternate allele frequencies of the variants in non-Finnish European (NFE) populations were obtained from the Exome Aggregation Consortium, which constituted data from 33,370 individuals.

Variant filtering and prioritization

We removed the variants that did not pass the variant quality score recalibration (VQSR) filter. We further retained variants based on read depth ($\geq 5\times$), and at least 20% of the reads were attributed to the alternate allele. We applied a segregation filter where we retained variants that were shared among 80% or more affected family members or segregated in all but one affected individual. Variants were discarded if they were present in one or more unaffected family members. We generated a list of rare variants with MAF $< 1\%$ and low-frequency variants based on allele frequency data from ExAC non-Finnish Europeans (81). The AMD locus regions were defined by the genome-wide significant variants and the variants within linkage disequilibrium (LD) ($r^2 > 0.5$), adding a further 500 kb to both sides, as described in the most recent AMD-GWAS analysis (16). Finally, the variant-level data were collapsed into gene-level data by combining all variants observed in each gene across different families. For exome-wide analysis, we applied additional filters of predictive causality (CADD ≥ 20) and retained variants that passed variant quality score recalibration in the ExAC database. Common AMD-associated variants from the most recent GWAS (16) were visualized using LocusZoom (82).

Simulation analysis of rare variant segregation

Simulation analysis was performed to assess type I errors. We included all the variants that passed quality control, MAF and functional filters. We did not include the segregation filter in order to access the probability of genes/variants, segregating in three or more families. We started with our list of filtered variants with allele counts from the AMD loci analysis (2862 variant) and the exome-wide analysis (1299 variants) as our input. For each variant, we simulated the null case by randomly distributing its minor alleles across the 264 subjects in the study, with at most one allele per subject. Then, we applied the same segregation criteria we used in our real data analysis. Briefly, if

any minor allele was assigned to a control subject, the variant was discarded; otherwise, we evaluated the segregation pattern of each family and counted properly segregating families as described above. We repeated this for all variants within each gene and then identified all segregating families. We performed 100,000 simulations for each gene in two different scenarios: (1) 2862 variants in 600 genes within the known AMD Loci and (2) 1299 variants with a CADD score ≥ 20 in 1176 genes exome-wide. The study-level type I error was calculated by multiplying the gene-level type I errors with the number of genes tested.

Expression and eQTL analysis of control and AMD retina

The transcriptome data of post-mortem human donor retinas from 453 donors at different stages of AMD and controls were analyzed for candidate gene prioritization. The donor retina that were graded for normal and disease status of AMD using Minnesota Grading System (MGS), with criteria similar to the Age-Related Eye Disease Study (AREDS). We analyzed 105 MGS1 (normal), 175 MGS2 (early AMD), 112 MGS3 (intermediate AMD) and 61 MGS4 (advanced AMD) after initial RNA-seq quality control described elsewhere. The mean age of donors was 80 years (range 55–107). The methods for RNA sequencing, gene expression quantitation, differential gene expression and eQTL analysis are described in detail elsewhere (54). Briefly, RSEM expected counts from 453 samples and 18,053 genes passing the expression-level filter (≥ 1 CPM in $\geq 10\%$ of samples) were transformed into TMM-normalized CPM and then converted into $\log_2$ CPM with an offset of 1. Batch effects were estimated and corrected using the supervised surrogate variable analysis (supervised sva) method within the bioconductor sva package in R. The differential expression between MGS 1 controls and each disease stage was assessed using the limma package in R. Fold changes between these comparisons were presented as heatmaps and violin plots using in-house scripts in R. The eQTL analysis included 17,389 genes and 8,924,684 genotyped and imputed common variants. The mapping of *cis*-eQTLs [as defined by SNP-gene combination within ± 1 Mb of the transcriptional start site (TSS) of each gene] was performed using QTLtools to identify genetic variants (eVariants) that control expression of genes (eGenes) at FDR ≤ 0.05.

Rare variant association analysis in AMD case-control studies

We analyzed two different cohorts for this analysis. The first cohort was analyzed using whole genome sequencing data in 4787 subjects consisting of 2394 cases and 2393 controls (30,81). All the cases had advanced AMD (GA or CNV). The second cohort was analyzed using whole exome sequencing in 1125 cases and 1361 controls (24). RAREMETAL was used to perform the meta-analysis of gene-based tests for rare variants (83).

Supplementary material

Supplementary material is available at *HMG* online.

Acknowledgments

References

1. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. *et al.* (2014) The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.*, **42**, D1001–D1006.

2. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747–753.

3. Zwick, M.E., Cutler, D.J. and Chakravarti, A. (2000) Patterns of genetic variation in Mendelian and complex traits. *Annu. Rev. Genomics Hum. Genet.*, **1**, 387–407.

4. International HapMap Consortium (2003) The international HapMap project. *Nature*, **426**, 789–796.

5. Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A. *et al.* (2015) A global reference for human genetic variation. *Nature*, **526**, 68–74.

6. Bomba, L., Walter, K. and Soranzo, N. (2017) The impact of rare and low-frequency genetic variants in common disease. *Genome Biol.*, **18**, 77.

7. Wong, W.L., Su, X., Li, X., Cheung, C.M., Klein, R., Cheng, C.Y. and Wong, T.Y. (2014) Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob. Health*, **2**, e106–e116.

8. Friedman, D.S., O'Colmain, B.J., Munoz, B., Tomany, S.C., McCarty, C., de Jong, P.T., Nemesure, B., Mitchell, P. and Kempen, J. (2004) Prevalence of age-related macular degeneration in the United States. *Arch. Ophthalmol.*, **122**, 564–572.

9. Ferris, F.L., 3rd, Wilkinson, C.P., Bird, A., Chakravarthy, U., Chew, E., Csaky, K. and Sadda, S.R. (2013) Clinical classification of age-related macular degeneration. *Ophthalmology*, **120**, 844–851.

10. Ratnapriya, R. and Chew, E.Y. (2013) Age-related macular degeneration-clinical review and genetics update. *Clin. Genet.*, **84**, 160–166.

11. Holz, F.G., Schmitz-Valckenberg, S. and Fleckenstein, M. (2014) Recent developments in the treatment of age-related macular degeneration. *J. Clin. Invest.*, **124**, 1430–1438.

12. Age-Related Eye Disease Study Research Group (2000) Risk factors associated with age-related macular degeneration. A case-control study in the age-related eye disease study: age-related eye disease study report number 3. *Ophthalmology*, **107**, 2224–2232.

13. Swaroop, A., Chew, E.Y., Rickman, C.B. and Abecasis, G.R. (2009) Unraveling a multifactorial late-onset disease: from genetic susceptibility to disease mechanisms for age-related macular degeneration. *Annu. Rev. Genomics Hum. Genet.*, **10**, 19–43.

14. Fritsche, L.G., Fariss, R.N., Stambolian, D., Abecasis, G.R., Curcio, C.A. and Swaroop, A. (2014) Age-related macular degeneration: genetics and biology coming together. *Annu. Rev. Genomics Hum. Genet.*, **15**, 151–171.

15. Grassmann, F., Ach, T., Brandl, C., Heid, I.M. and Weber, B.H.F. (2015) What does genetics tell us about age-related macular degeneration? *Annu. Rev. Vis. Sci.*, **1**, 73–96.

16. Fritsche, L.G., Igl, W., Bailey, J.N., Grassmann, F., Sengupta, S., Bragg-Gresham, J.L., Burdon, K.P., Hebbring, S.J., Wen, C., Gorski, M. *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.

17. Yang, H.J., Ratnapriya, R., Cogliati, T., Kim, J.W. and Swaroop, A. (2015) Vision from next generation sequencing: multi-dimensional genome-wide analysis for producing gene regulatory networks underlying retinal development, aging and disease. *Prog. Retin. Eye Res.*, **46**, 1–30.

18. Raychaudhuri, S., Iartchouk, O., Chin, K., Tan, P.L., Tai, A.K., Ripke, S., Gowrisankar, S., Vemuri, S., Montgomery, K., Yu, Y. *et al.* (2011) A rare penetrant mutation in CFH confers high risk of age-related macular degeneration. *Nat. Genet.*, **43**, 1232–1236.

19. Helgason, H., Sulem, P., Duvvari, M.R., Luo, H., Thorleifsson, G., Stefansson, H., Jonsdottir, I., Masson, G., Gudbjartsson, D.F., Walters, G.B. *et al.* (2013) A rare nonsynonymous sequence variant in C3 is associated with high risk of age-related macular degeneration. *Nat. Genet.*, **45**, 1371–1374.

20. Seddon, J.M., Yu, Y., Miller, E.C., Reynolds, R., Tan, P.L., Gowrisankar, S., Goldstein, J.I., Triebwasser, M., Anderson, H.E., Zerbib, J. *et al.* (2013) Rare variants in CFI, C3 and C9 are associated with high risk of advanced age-related macular degeneration. *Nat. Genet.*, **45**, 1366–1370.

21. Zhan, X., Larson, D.E., Wang, C., Koboldt, D.C., Sergeev, Y.V., Fulton, R.S., Fulton, L.L., Fronick, C.C., Branham, K.E., Bragg-Gresham, J. *et al.* (2013) Identification of a rare coding variant in complement 3 associated with age-related macular degeneration. *Nat. Genet.*, **45**, 1375–1379.

22. van de Ven, J.P., Nilsson, S.C., Tan, P.L., Buitendijk, G.H., Ristau, T., Mohlin, F.C., Nabuurs, S.B., Schoenmaker-Koller, F.E., Smailhodzic, D., Campochiaro, P.A. *et al.* (2013) A functional variant in the CFI gene confers a high risk of age-related macular degeneration. *Nat. Genet.*, **45**, 813–817.

23. Momozawa, Y., Akiyama, M., Kamatani, Y., Arakawa, S., Yasuda, M., Yoshida, S., Oshima, Y., Mori, R., Tanaka, K., Mori, K. *et al.* (2016) Low-frequency coding variants in CETP and CFB are associated with susceptibility of exudative age-related macular degeneration in the Japanese population. *Hum. Mol. Genet.*, **25**, 5027–5034.

24. Corominas, J., Colijn, J.M., Geerlings, M.J., Pauper, M., Bakker, B., Amin, N., Lores Motta, L., Kersten, E., Garanto, A., Verlouw, J.A.M. *et al.* (2018) Whole-exome sequencing in age-related macular degeneration identifies rare variants in COL8A1, a component of Bruch's membrane. *Ophthalmology*, **125**, 1433–1443.

25. Anderson, D.H., Radeke, M.J., Gallo, N.B., Chapin, E.A., Johnson, P.T., Curletti, C.R., Hancox, L.S., Hu, J., Ebright, J.N., Malek, G. *et al.* (2010) The pivotal role of the complement system in aging and age-related macular degeneration: hypothesis re-visited. *Prog. Retin. Eye Res.*, **29**, 95–112.

26. Stanton, C.M., Yates, J.R., den Hollander, A.I., Seddon, J.M., Swaroop, A., Stambolian, D., Fauser, S., Hoyng, C., Yu, Y., Atsuhiro, K. *et al.* (2011) Complement factor D in age-related macular degeneration. *Invest. Opthamol. Vis. Sci.*, **52**, 8828–8834.

27. Duvvari, M.R., Paun, C.C., Buitendijk, G.H., Saksens, N.T., Volokhina, E.B., Ristau, T., Schoenmaker-Koller, F.E., van de

Ven, J.P., Groenewoud, J.M., van den Heuvel, L.P. *et al.* (2014) Analysis of rare variants in the C3 gene in patients with age-related macular degeneration. *PloS One*, **9**, e94165.

28. Yu, Y., Triebwasser, M.P., Wong, E.K., Schramm, E.C., Thomas, B., Reynolds, R., Mardis, E.R., Atkinson, J.P., Daly, M., Raychaudhuri, S. *et al.* (2014) Whole-exome sequencing identifies rare, functional CFH variants in families with macular degeneration. *Hum. Mol. Genet.*, **23**, 5283–5293.

29. Saksens, N.T., Geerlings, M.J., Bakker, B., Schick, T., Daha, M.R., Fauser, S., Boon, C.J., de Jong, E.K., Hoyng, C.B. and den Hollander, A.I. (2016) Rare genetic variants associated with development of age-related macular degeneration. *JAMA Ophthalmol.*, **134**, 287–293.

30. Pietraszkiewicz, A., van Asten, F., Kwong, A., Ratnapriya, R., Abecasis, G., Swaroop, A. and Chew, E.Y. (2018) Association of rare predicted loss-of-function variants in cellular pathways with sub-phenotypes in age-related macular degeneration. *Ophthalmology*, **125**, 398–406.

31. Heiba, I.M., Elston, R.C., Klein, B.E. and Klein, R. (1994) Sibling correlations and segregation analysis of age-related maculopathy: the beaver dam eye study. *Genet. Epidemiol.*, **11**, 51–67.

32. Klaver, C.C., Wolfs, R.C., Assink, J.J., van Duijn, C.M., Hofman, A. and de Jong, P.T. (1998) Genetic risk of age-related maculopathy: population-based familial aggregation study. *Arch. Ophthalmol.*, **116**, 1646–1651.

33. Seddon, J.M., Ajani, U.A. and Mitchell, B.D. (1997) Familial aggregation of age-related maculopathy. *Am. J. Ophthalmol.*, **123**, 199–206.

34. Klein, M.L., Schultz, D.W., Edwards, A., Matise, T.C., Rust, K., Berselli, C.B., Trzupek, K., Weleber, R.G., Ott, J., Wirtz, M.K. *et al.* (1998) Age-related macular degeneration. Clinical features in a large family and linkage to chromosome 1q. *Arch. Ophthalmol.*, **116**, 1082–1088.

35. Weeks, D.E., Conley, Y.P., Mah, T.S., Paul, T.O., Morse, L., Ngo-Chang, J., Dailey, J.P., Ferrell, R.E. and Gorin, M.B. (2000) A full genome scan for age-related maculopathy. *Hum. Mol. Genet.*, **9**, 1329–1349.

36. Majewski, J., Schultz, D.W., Weleber, R.G., Schain, M.B., Edwards, A.O., Matise, T.C., Acott, T.S., Ott, J. and Klein, M.L. (2003) Age-related macular degeneration–a genome scan in extended families. *Am. J. Hum. Genet.*, **73**, 540–550.

37. Abecasis, G.R., Yashar, B.M., Zhao, Y., Ghiasvand, N.M., Zareparsi, S., Branham, K.E., Reddick, A.C., Trager, E.H., Yoshida, S., Bahling, J. *et al.* (2004) Age-related macular degeneration: a high-resolution genome scan for susceptibility loci in a population enriched for late-stage disease. *Am. J. Hum. Genet.*, **74**, 482–494.

38. Fisher, S.A., Abecasis, G.R., Yashar, B.M., Zareparsi, S., Swaroop, A., Iyengar, S.K., Klein, B.E., Klein, R., Lee, K.E., Majewski, J. *et al.* (2005) Meta-analysis of genome scans of age-related macular degeneration. *Hum. Mol. Genet.*, **14**, 2257–2264.

39. Edwards, A.O., Ritter, R., 3rd, Abel, K.J., Manning, A., Panhuysen, C. and Farrer, L.A. (2005) Complement factor H polymorphism and age-related macular degeneration. *Science*, **308**, 421–424.

40. Hageman, G.S., Anderson, D.H., Johnson, L.V., Hancox, L.S., Taiber, A.J., Hardisty, L.I., Hageman, J.L., Stockman, H.A., Borchardt, J.D., Gehrs, K.M. *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. U. S. A.*, **102**, 7227–7232.

41. Haines, J.L., Hauser, M.A., Schmidt, S., Scott, W.K., Olson, L.M., Gallins, P., Spencer, K.L., Kwan, S.Y., Noureddine, M., Gilbert, J.R. *et al.* (2005) Complement factor H variant increases the risk of age-related macular degeneration. *Science*, **308**, 419–421.

42. Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J.Y., Sackler, R.S., Haynes, C., Henning, A.K., SanGiovanni, J.P., Mane, S.M., Mayne, S.T. *et al.* (2005) Complement factor H polymorphism in age-related macular degeneration. *Science*, **308**, 385–389.

43. Zareparsi, S., Branham, K.E., Li, M., Shah, S., Klein, R.J., Ott, J., Hoh, J., Abecasis, G.R. and Swaroop, A. (2005) Strong association of the Y402H variant in complement factor H at 1q32 with susceptibility to age-related macular degeneration. *Am. J. Hum. Genet.*, **77**, 149–153.

44. Jakobsdottir, J., Conley, Y.P., Weeks, D.E., Mah, T.S., Ferrell, R.E. and Gorin, M.B. (2005) Susceptibility genes for age-related maculopathy on chromosome 10q26. *Am. J. Hum. Genet.*, **77**, 389–407.

45. Rivera, A., Fisher, S.A., Fritsche, L.G., Keilhauer, C.N., Lichtner, P., Meitinger, T. and Weber, B.H. (2005) Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. *Hum. Mol. Genet.*, **14**, 3227–3236.

46. Kanda, A., Chen, W., Othman, M., Branham, K.E., Brooks, M., Khanna, R., He, S., Lyons, R., Abecasis, G.R. and Swaroop, A. (2007) A variant of mitochondrial protein LOC387715/ARMS2, not HTRA1, is strongly associated with age-related macular degeneration. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 16227–16232.

47. Ratnapriya, R., Zhan, X., Fariss, R.N., Branham, K.E., Zipprer, D., Chakarova, C.F., Sergeev, Y.V., Campos, M.M., Othman, M., Friedman, J.S. *et al.* (2014) Rare and common variants in extracellular matrix gene Fibrillin 2 (FBN2) are associated with macular degeneration. *Hum. Mol. Genet.*, **23**, 5827–5837.

48. Geerlings, M.J., de Jong, E.K. and den Hollander, A.I. (2017) The complement system in age-related macular degeneration: a review of rare genetic variants and implications for personalized treatment. *Mol. Immunol.*, **84**, 65–76.

49. Wagner, E.K., Raychaudhuri, S., Villalonga, M.B., Java, A., Triebwasser, M.P., Daly, M.J., Atkinson, J.P. and Seddon, J.M. (2016) Mapping rare, deleterious mutations in factor H: association with early onset, drusen burden, and lower antigenic levels in familial AMD. *Sci. Rep.*, **6**, 31531.

50. Duvvari, M.R., van de Ven, J.P., Geerlings, M.J., Saksens, N.T., Bakker, B., Henkes, A., Neveling, K., del Rosario, M., Westra, D., van den Heuvel, L.P. *et al.* (2016) Whole exome sequencing in patients with the cuticular drusen subtype of age-related macular degeneration. *PloS One*, **11**, e0152047.

51. Pras, E., Kristal, D., Shoshany, N., Volodarsky, D., Vulih, I., Celniker, G., Isakov, O., Shomron, N. and Pras, E. (2015) Rare genetic variants in Tunisian Jewish patients suffering from age-related macular degeneration. *J. Med. Genet.*, **52**, 484–492.

52. Hoffman, J.D., Cooke Bailey, J.N., D'Aoust, L., Cade, W., Ayala-Haedo, J., Fuzzell, D., Laux, R., Adams, L.D., Reinhart-Mercer, L., Caywood, L. *et al.* (2014) Rare complement factor H variant associated with age-related macular degeneration in the Amish. *Invest. Ophthalmol. Vis. Sci.*, **55**, 4455–4460.

53. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. *et al.* (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, **20**, 1297–1303.

54. Ratnapriya, R., Sosina, O.A., Starostik, M.R., Kwicklis, M., Kapphahn, R.J., Fritsche, L.G., Walton, A., Arvanitis, M., Gieser, L., Pietraszkiewicz, A. *et al.* (2019) Retinal transcriptome and

eQTL analyses identify genes associated with age-related macular degeneration. *Nat. Genet.*, **51**, 606–610.

55. Ferrara, D. and Seddon, J.M. (2015) Phenotypic characterization of complement factor H R1210C rare genetic variant in age-related macular degeneration. *JAMA Ophthalmol.*, **133**, 785–791.

56. Kersten, E., Geerlings, M.J., den Hollander, A.I., de Jong, E.K., Fauser, S., Peto, T. and Hoyng, C.B. (2017) Phenotype characteristics of patients with age-related macular degeneration carrying a rare variant in the complement factor H gene. *JAMA Ophthalmol.*, **135**, 1037–1044.

57. Bathgate, R.A., Halls, M.L., van der Westhuizen, E.T., Callander, G.E., Kocan, M. and Summers, R.J. (2013) Relaxin family peptides and their receptors. *Physiol. Rev.*, **93**, 405–480.

58. Donizetti, A., Fiengo, M., Del Gaudio, R., Iazzetti, G., Pariante, P., Minucci, S. and Aniello, F. (2015) Expression pattern of zebrafish rxfp2 homologue genes during embryonic development. *J. Exp. Zool. B. Mol. Dev. Evol.*, **324**, 605–613.

59. Graveline, R., Marcinkiewicz, K., Choi, S., Paquet, M., Wurst, W., Floss, T. and David, G. (2017) The chromatin-associated Phf12 protein maintains nucleolar integrity and prevents premature cellular senescence. *Mol. Cell. Biol.*, **37**, e00522–e00516.

60. Sironen, A., Kotaja, N., Mulhern, H., Wyatt, T.A., Sisson, J.H., Pavlik, J.A., Miiluniemi, M., Fleming, M.D. and Lee, L. (2011) Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia. *Biol. Reprod.*, **85**, 690–701.

61. Sironen, A., Hansen, J., Thomsen, B., Andersson, M., Vilkki, J., Toppari, J. and Kotaja, N. (2010) Expression of SPEF2 during mouse spermatogenesis and identification of IFT20 as an interacting protein. *Biol. Reprod.*, **82**, 580–590.

62. Keady, B.T., Le, Y.Z. and Pazour, G.J. (2011) IFT20 is required for opsin trafficking and photoreceptor outer segment development. *Mol. Biol. Cell*, **22**, 921–930.

63. Riccomagno, M.M., Sun, L.O., Brady, C.M., Alexandropoulos, K., Seo, S., Kurokawa, M. and Kolodkin, A.L. (2014) Cas adaptor proteins organize the retinal ganglion cell layer downstream of integrin signaling. *Neuron*, **81**, 779–786.

64. Orozco, L.D., Chen, H.H., Cox, C., Katschke, K.J., Jr., Arceo, R., Espiritu, C., Caplazi, P., Nghiem, S.S., Chen, Y.J., Modrusan, Z. et al. (2020) Integration of eQTL and a single-cell atlas in the human eye identifies causal genes for age-related macular degeneration. *Cell Rep.*, **30**, 1246–1259.

65. Kwong, , A.Z.X., Fritsche, L.G., Bragg-Gresham, J., Branham, K.E., Othman, M., Boleda, A., Gieser, L., Ratnapriya, R., Stambolian, D., Chew, E.Y., Swaroop, A. and Abecasis, G. (2014) Whole-genome sequencing study of ∼6,000 samples for age-related macular degeneration. (Abstract #385) Presented at the 64th Annual Meeting of The Americal Society of Human Genetics. San Diego, California.

66. O'Brien, B.J., Caldwell, J.H., Ehring, G.R., Bumsted O'Brien, K.M., Luo, S. and Levinson, S.R. (2008) Tetrodotoxin-resistant voltage-gated sodium channels Na(v)1.8 and Na(v)1.9 are expressed in the retina. *J. Comp. Neurol.*, **508**, 940–951.

67. Faber, C.G., Lauria, G., Merkies, I.S., Cheng, X., Han, C., Ahn, H.S., Persson, A.K., Hoeijmakers, J.G., Gerrits, M.M., Pierro, T. et al. (2012) Gain-of-function Nav1.8 mutations in painful neuropathy. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 19444–19449.

68. Giblin, J.P., Comes, N., Strauss, O. and Gasull, X. (2016) Ion channels in the eye: involvement in ocular pathologies. *Adv. Protein Chem. Struct. Biol.*, **104**, 157–231.

69. Bashirova, A.A., Martin, M.P., McVicar, D.W. and Carrington, M. (2006) The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. *Annu. Rev. Genomics Hum. Genet.*, **7**, 277–300.

70. Goverdhan, S.V., Khakoo, S.I., Gaston, H., Chen, X. and Lotery, A.J. (2008) Age-related macular degeneration is associated with the HLA-Cw∗0701 genotype and the natural killer cell receptor AA haplotype. *Invest. Opthalmol. Vis. Sci.*, **49**, 5077–5082.

71. Luo, J., Sladek, R., Carrier, J., Bader, J.A., Richard, D. and Giguere, V. (2003) Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor alpha. *Mol. Cell. Biol.*, **23**, 7947–7956.

72. van Leeuwen, E.M., Emri, E., Merle, B.M.J., Colijn, J.M., Kersten, E., Cougnard-Gregoire, A., Dammeier, S., Meester-Smoor, M., Pool, F.M., de Jong, E.K. et al. (2018) A new perspective on lipid research in age-related macular degeneration. *Prog. Retin. Eye Res.*, **67**, 56–86.

73. Kolehmainen, J., Black, G.C., Saarinen, A., Chandler, K., Clayton-Smith, J., Traskelin, A.L., Perveen, R., Kivitie-Kallio, S., Norio, R., Warburg, M. et al. (2003) Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am. J. Hum. Genet.*, **72**, 1359–1369.

74. Bird, A.C., Bressler, N.M., Bressler, S.B., Chisholm, I.H., Coscas, G., Davis, M.D., de Jong, P.T., Klaver, C.C., Klein, B.E., Klein, R. et al. (1995) An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. *Surv. Ophthalmol.*, **39**, 367–374.

75. Priya, R.R., Chew, E.Y. and Swaroop, A. (2012) Genetic studies of age-related macular degeneration: lessons, challenges, and opportunities for disease management. *Ophthalmology*, **119**, 2526–2536.

76. Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

77. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, **25**, 1754–1760.

78. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M. et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.*, **43**, 491–498.

79. Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.*, **38**, e164.

80. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

81. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**, 285–291.

82. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336–2337.

83. Feng, S., Liu, D., Zhan, X., Wing, M.K. and Abecasis, G.R. (2014) RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics*, **30**, 2828–2829.