Human Molecular Genetics, 2014, Vol. 23, No. 15 **4001–4014**
doi:10.1093/hmg/ddu114
Advance Access published on March 14, 2014

Comprehensive analysis of gene expression in human retina and supporting tissues

Mingyao Li[1], Cheng Jia[1], Krista L. Kazmierkiewicz[2], Anita S. Bowman[2], Lifeng Tian[5], Yichuan Liu[1], Neel A. Gupta[8], Harini V. Gudiseva[2], Stephanie S. Yee[2], Mijin Kim[2], Tzvete Dentchev[3], James A. Kimble[9,10], John S. Parker[10], Jeffrey D. Messinger[10], Hakon Hakonarson[4,6,7], Christine A. Curcio[10,†] and Dwight Stambolian[2,*,†]

[1]Department of Biostatistics and Epidemiology, [2]Department of Ophthalmology, [3]Department of Dermatology and, [4]Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA, [5]Center for Applied Genomics, [6]Division of Pulmonary Medicine and [7]The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA, [8]College of Medicine, Drexel University, Philadelphia, PA 19104, USA, [9]Retina Specialists of Alabama, Birmingham, AL 35294, USA and [10]Department of Ophthalmology, University of Alabama School of Medicine, Birmingham, AL 35294, USA

Received November 12, 2013; Revised January 21, 2014; Accepted March 10, 2014

Understanding the influence of gene expression on the molecular mechanisms underpinning human phenotypic diversity is fundamental to being able to predict health outcomes and treat disease. We have carried out whole transcriptome expression analysis on a series of eight normal human postmortem eyes by RNA sequencing. Here we present data showing that ∼80% of the transcriptome is expressed in the posterior layers of the eye and that there is significant differential expression not only between the layers of the posterior part of the eye but also between locations of a tissue layer. These differences in expression also extend to alternative splicing and splicing factors. Differentially expressed genes are enriched for genes associated with psychiatric, immune and cardiovascular disorders. Enrichment categories for gene ontology included ion transport, synaptic transmission and visual and sensory perception. Lastly, allele-specific expression was found to be significant for *CFH*, *C3* and *CFB*, which are known risk genes for age-related macular degeneration. These expression differences should be useful in determining the underlying biology of associations with common diseases of the human retina, retinal pigment epithelium and choroid and in guiding the analysis of the genomic regions involved in the control of normal gene expression.

INTRODUCTION

Genome-wide association studies (GWASs) have successfully identified many single-nucleotide polymorphisms (SNPs) associated with complex diseases. Most of these associated SNPs reside within specific genomic regions that include dozens of nearby genes, falling short of proving causation. This is because ∼88% of the genetic variants currently associated with complex diseases and traits by GWAS lie within intronic or intergenic regions (1). Recent GWASs for eye diseases/traits have delivered a number of novel findings across a diverse range of diseases, including age-related macular degeneration (AMD), glaucoma and refractive error (2–4). However, despite this astonishing rate of success, the major challenge still remains to not only confirm that the genes implicated in these studies are truly the genes conferring protection from or risk of disease but also to define the functional roles these genes play in disease. Current thinking is that these nucleotide variations are likely to have causal effects by influencing gene expression rather than affecting protein function (5). A growing number of studies have shown that these SNPs influence gene expression variation in the population (6,7). Global gene expression studies—not acquiring a priori hypotheses—provide a method to investigate the pathogenesis of common disorders on a large scale (8).

*To whom correspondence should be addressed. Tel: +1 2158980305; Fax: +1 2155736728; Email: stamboli@mail.med.upenn.edu
†The last two authors should be regarded as co-corresponding authors.

The neurosensory retina is a specialized neural tissue lining the back of the eye that is responsible for processing images and sending these images to the brain. It originates as an outgrowth of the brain during ontogenesis and is thus part of the central nervous system. The cellular composition of the retina is complex with a variety of cell types. The macula has the highest overall density of neurons and is responsible for central vision. The retinal pigment epithelium and choroidal vasculature constitute the photoreceptor support system. The choroid is part of the systemic circulation, i.e. outside the blood retina barrier, and contains multiple resident and transient cells. The sclera is the protective structure of the eye and encloses the eyeball except in the region of the cornea.

Molecular phenotypes such as gene and protein expression are influenced by nearby DNA variations. Therefore, to uncover the genetic determinants regulating expression in tissues involved in chorioretinal diseases, we profiled eight human neurosensory retina and retinal pigment epithelium (RPE)/choroid/sclera (RCS) samples by RNA sequencing (RNA-Seq). We also genotyped DNA from each sample on an Omni 2.5 million SNP chip. Within the RCS sample, the RPE cells and the choroidal vasculature form the photoreceptor support system and are preferentially affected in AMD.

Previous studies on gene expression of the eye have been limited to microarrays (9–14). Our study is the first to report on the differential expression and differential alternative splicing between the macula and peripheral regions of the retina and RCS using RNA sequencing. Our comprehensive analysis of the retina and RCS gene expression profiles revealed large differences in expression between these layers of the posterior eye and a smaller but significant number of differences between the macula and periphery of each layer. Specifically, we identified 9555 differentially expressed (DE) genes between retina and RCS layers in macula and >10 149 DE genes in periphery. We identified >2051 DE genes between peripheral and macular retina (MR) and 926 between peripheral RCS (PRCS) and macular RCS (MRCS). Our differential alternative splicing (DAS) analysis also revealed differences between layers and location and varied from a low of two DAS events between MRCS and PRCS and a high of 575 DAS events between peripheral retina (PR) and PRCS. Significantly, we replicated a portion of our DE genes by QPCR and DAS events by PCR, demonstrating that our data analysis is reproducible. Finally, we have preliminary evidence demonstrating potential allele-specific expression for *CFH*, *C3* and *CFB*, three genes that have been previously reported to be associated with AMD (15). These marked differences in gene expression between the macula and periphery of each layer may offer some explanation for why some diseases that primarily affect the macula and others affect the periphery. The marked differences between retina and RCS also underscore the tremendous diversity of biological processes required to maintain good photoreceptor health.

RESULTS

Overall quality parameters of the RNA-Seq dataset

We used RNA-Seq to characterize the chorioretinal transcriptomes in a Discovery set of eight normal human eyes. For each eye, we sequenced four RNA-Seq samples and generated close to 100 million 101-bp paired-end reads per sample. We mapped the sequence reads to the reference human genome (hg19) using GSNAP (16). Our data are of high quality with 76–94% of the reads mapped to the human genome and 60–81% retained after stringent quality control filtering, among which 86–93% mapped to genes defined by RefSeq (Supplementary Material, Table S1).

Analysis of gene expression

Global analysis of gene expression

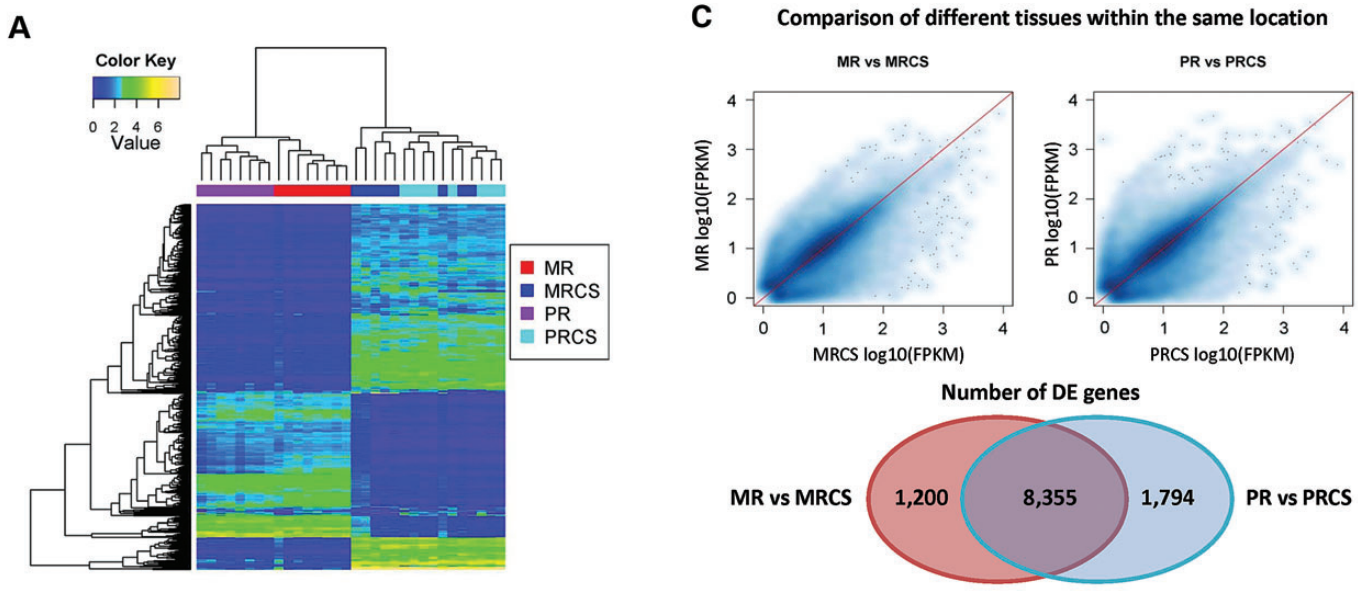
We first considered the overall gene expression in each sample. Using filtered mapped reads, we estimated the expression levels of 23 569 RefSeq protein-coding genes using the fragments per kilobase of gene per million mapped fragment (FPKM) metric (17). With coverage depth ranging from 66 to 133 million paired-end reads per sample, we detected the expression (i.e. FPKM > 0) of the majority of known protein-coding genes. The average number of expressed genes was 18 318 (78%) in MR, 17 334 (74%) in MRCS, 18 850 (80%) in PR and 19 022 (81%) in PRCS (Table 1). Although the overall numbers of expressed genes were similar in the four groups, there was an excess of moderate-to-low-abundance (FPKM < 5) genes in retina as compared with RCS (MR versus MRCS: Fisher's exact test $P = 7.8 \times 10^{-4}$; PR versus PRCS: $P = 2.0 \times 10^{-3}$). We also observed a slight excess of moderate-to-low-abundance genes in periphery as compared with macula in the RCS layer ($P = 0.029$). There was a similar trend in retina (MR versus PR: $P = 0.25$). However, since these tests were not corrected for multiple testing, the comparison of PRCS versus MRCS would not be statistically significant if multiple comparisons were considered.

Based on the estimated FPKM values in each sample, we then investigated the similarity in the global gene expression profiles among the four tissue groups using average linkage hierarchical clustering of the top 1000 most divergent genes defined by the coefficient of variation of the FPKMs (Fig. 1A). The retina samples clearly separated from the RCS samples, suggesting that contamination of retina and RCS during dissection was negligible. Moreover, there was a clear distinction between macula and periphery for retina, but the distinction between macula and periphery was less clear for RCS. These results suggest the existence of tissue-specific gene expression signatures and that regional differences were less pronounced than differences between tissues.

Table 1. Numbers of expressed genes

Group	Minimum	Maximum	Mean
MR	17 639	18 800	18 318
MRCS	16 177	18 369	17 334
PR	18 258	19 577	18 850
PRCS	18 519	20 938	19 022

A gene was considered expressed if the FPKM was >0. The minimum, maximum and mean numbers of expressed genes were calculated based on the eight samples in each eye part.



B Number of DE genes identified from each comparison

Comparison (group 1 vs group2)	Number of DE genes	Number of DE genes with higher expression level in group 1	Number of DE genes with higher expression level in group 2
MR vs MRCS	9,555	4,615	4,940
PR vs PRCS	10,149	4,918	5,231
MR vs PR	2,051	1,239	812
MRCS vs PRCS	926	653	273

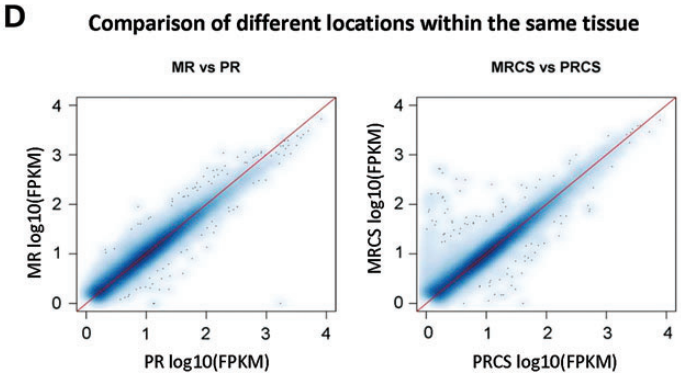


Figure 1. Gene expression. (**A**) Hierarchical clustering analysis of top 1000 most divergent genes. For each gene, we calculated its coefficient of variation (CV) based on its log-transformed FPKM values across all RNA-Seq samples. The genes were then ranked based on their CV values. The heatmap was generated by hierarchical clustering of the top 1000 genes with the largest CV values. (**B**) Number of differentially expressed (DE) genes identified from each comparison. Analysis was completed using Cuffdiff v2.1.1. A gene was considered DE if FDR-adjusted *P*-value was <0.05. (**C**) Comparison of gene expression profiles between different tissues within the same location. The Venn diagram shows the numbers of DE genes that were unique to each comparison as well as the number of DE genes that were common between the two comparisons. (**D**) Comparison of gene expression profiles between different locations within the same tissue.

Differentially expressed genes

To detail the genes expressed in each tissue sample, we carried out differential expression analysis using Cufflinks (18). Figure 1B shows the numbers of genes that were differentially expressed. Consistent with the hierarchical clustering analysis, the difference between tissues (e.g. MR versus MRCS) was larger than the difference between locations within the same tissue layer (e.g. MR versus PR). When comparing retina and RCS, we found 9555 DE genes in macula and 10 149 in periphery. In contrast, a smaller number of DE genes were identified when comparing different locations within the same tissue; for retina, the number of DE genes was 2051, whereas it was 926 for RCS. Among genes that were differentially expressed between macula and periphery, 60–70% had higher expression levels in macula (MR versus PR: $P < 2.2 \times 10^{-16}$; MRCS versus PRCS: $P < 2.2 \times 10^{-16}$). Among genes that were differentially expressed between retina and RCS, ~52% had higher expression levels in RCS (MR versus MRCS: $P = 9.2 \times 10^{-4}$; PR versus PRCS: $P = 2.0 \times 10^{-3}$). Most of the DE genes (~70%) had at least 2-fold gene expression difference, except for the comparison of PR and MR in which only ~39% had a gene expression difference of >2-fold. These results agree with the patterns shown in scatter plots of the estimated FPKMs (Fig. 1C and D).

As our differential expression analysis results were based on Cufflinks v2.1.1, we decided to utilize additional statistical programs to confirm the results. We analyzed the data using two other programs, edgeR (19) and DESeq (20), which utilize independent statistical methodologies (Supplementary Material, Fig. S1). Focusing on the DE genes identified by Cufflinks v2.1.1, we found agreement among all three programs for 8378 (88%) DE genes, when comparing MR and MRCS, 8929 (88%) for the comparison of PR and PRCS, 1195 (58%) for MR and PR and 203 (22%) for MRCS and PRCS (Supplementary Material, Fig. S1). The relatively lower percentage of agreement in the comparison of MR and PR, and the comparison of

MRCS and PRCS is due to the fact that many DE genes identified by Cufflinks v2.1.1 had a small fold change (<2) of gene expression, and these subtle changes were possibly missed by edgeR or DESeq owing to their lack of power. This was confirmed by the fact that >99% of these DE genes showed the same direction of gene expression changes across all three methods. If we required agreement with either edgeR or DESeq only, then the percentage of agreement would increase to 84% for the comparison of MR and PR and 64% for MRCS and PRCS. These comparisons suggest that our findings from Cufflinks v2.1.1 are reliable. Hereafter, we will only report results based on Cufflinks v.2.1.1.

Enrichment analysis of differentially expressed genes

To gain more insight into biological pathways that possibly explain the differences among different tissue layers of the eye, we carried out functional annotation analysis using Database for Annotation, Visualization and Integrated Discovery (DAVID) (21). For the comparison between MR and MRCS (Table 2; Supplementary Material, Table S2A), DE genes with higher expression levels in MR were significantly enriched as follows: for psychiatric disorders, chemical dependency and neurological disorders in disease category; for synaptic transmission, sensory perception of light stimulus and visual perception in gene ontology (GO); and for long-term potentiation, MAPK signaling pathway and endocytosis in Kyoto Encyclopedia of Genes and Genomes (KEGG). Genes with higher expression levels in MRCS (Supplementary Material, Table S2B) were significantly enriched as follows: for immune, cardiovascular and infection in disease category; for immune system processes, positive regulation of biological process and response to external stimulus in GO; and for ribosomes, cell adhesion molecules and focal adhesion in KEGG. Enrichment analysis for DE genes between PR and PRCS revealed a similar pattern (Table 2; Supplementary Material, Table S2C and 2D).

We also carried out functional annotation analysis for genes that were differentially expressed between macula and periphery (Table 3). For the comparison in retina, the DE genes that had higher expression levels in MR (Supplementary Material,

Table 2. DAVID enrichment analysis for DE genes between retina and RCS

Category	MR > MRCS Term	Gene count	*P*-value	MR < MRCS Term	Gene count	*P*-value
Disease class	Psych	195	2.5×10^{-19}	Immune	383	1.4×10^{-11}
	Chemdependency	57	4.3×10^{-4}	Cardiovascular	328	4.4×10^{-10}
	Neurological	140	2.5×10^{-2}	Infection	165	7.3×10^{-10}
Gene ontology	Synaptic transmission	151	1.1×10^{-23}	Immune system process	476	2.9×10^{-45}
	Sensory perception of light stimulus	119	2.1×10^{-22}	Positive regulation of biological process	802	1.1×10^{-38}
	Visual perception	119	2.1×10^{-22}	Response to external stimulus	423	4.5×10^{-36}
KEGG pathway	Long-term potentiation	35	1.4×10^{-6}	Ribosome	79	1.5×10^{-26}
	MAPK signaling pathway	84	9.7×10^{-5}	Cell adhesion molecules (CAMs)	83	2.3×10^{-10}
	Endocytosis	59	1.8×10^{-3}	Focal adhesion	113	4.5×10^{-10}
	PR > PRCS			PR < PRCS		
Disease class	Psych	190	1.6×10^{-16}	Immune	392	1.9×10^{-12}
	Chemdependency	57	6.4×10^{-4}	Infection	170	1.1×10^{-10}
	Neurological	141	3.2×10^{-2}	Cardiovascular	324	3.8×10^{-8}
Gene ontology	Cellular process	2701	7.0×10^{-26}	Immune system process	478	5.6×10^{-39}
	Synaptic transmission	154	4.9×10^{-23}	Positive regulation of biological process	819	6.1×10^{-34}
	Sensory perception of light stimulus	121	1.2×10^{-21}	Response to external stimulus	424	1.8×10^{-30}
KEGG pathway	Oxidative phosphorylation	54	6.3×10^{-6}	Ribosome	67	1.7×10^{-13}
	Alzheimer's disease	61	3.1×10^{-5}	Cell adhesion molecules (CAMs)	81	2.1×10^{-8}
	Parkinson's disease	51	2.5×10^{-5}	Focal adhesion	112	1.5×10^{-8}

P-values were Bonferroni corrected. Displayed are the top three enriched terms for each category.

Table 3. DAVID enrichment analysis for DE genes between macula and periphery

Category	MR > PR Term	Gene count	*P*-value	MR < PR Term	Gene count	*P*-value
Gene ontology	Ion transport	110	3.1×10^{-12}	Translational elongation	110	2.6×10^{-22}
	Metal ion transport	7	3.3×10^{-11}	Sensory perception of light stimulus	77	2.0×10^{-13}
	Transmission of nerve impulse	64	4.7×10^{-11}	Visual perception	64	2.0×10^{-13}
KEGG pathway	No significant enrichment			Ribosome	36	9.4×10^{-21}
	MRCS > PRCS			MRCS < PRCS		
Gene ontology	System process	167	3.2×10^{-38}	Cell adhesion	32	1.9×10^{-4}
	Muscle contraction	51	4.5×10^{-31}	Biological adhesion	28	1.0×10^{-4}
	Multicellular organismal process	289	5.2×10^{-31}	Cell development	28	1.3×10^{-3}
KEGG pathway	Dilated cardiomyopathy	21	2.9×10^{-8}	TGF-beta signaling pathway	9	1.9×10^{-2}
	Hypertrophic cardiomyopathy	19	1.9×10^{-7}			
	Calcium signaling pathway	23	4.9×10^{-5}			

P-values were Bonferroni corrected. Displayed are the top three enriched terms for each category.

Table S3A) were significantly enriched for ion transport, transmission of nerve impulse and synaptic transmission in GO; genes that had higher expression levels in periphery (Supplementary Material, Table S3B) were significantly enriched for translational elongation, sensory perception of light stimulus and visual perception in GO, and for ribosomes in KEGG. The overall pattern was different for the comparison in the RCS layer (Table 3; Supplementary Material, Table S3C and 3D). DE genes that had higher expression levels in MRCS were significantly enriched for system processes, muscle contraction and multicellular organismal process in GO and for dilated cardiomyopathy, hypertrophic cardiomyopathy and calcium muscle contraction in KEGG; genes that had higher expression levels in PRCS were enriched for biological adhesion, cell adhesion and cell development in GO.

Overlap with GWAS signals on eye diseases

Next, we sought to examine whether eye disease susceptibility genes were expressed in these tissues. We assembled a list of 91 unique genes implicated in recent GWAS on AMD, glaucoma and myopia based on NHGRI GWAS catalog (http://www.genome.gov/26525384, last accessed date on October 10, 2013). Among the assembled genes, 80 were expressed (mean FPKM values >0) in at least one eye part, although some were expressed at very low levels (Supplementary Material, Table S4). Most of these genes were differentially expressed between retina and RCS; 57 (71%) were in the comparison of MR and MRCS and 61 (76%) in the comparison of PR and PRCS. For example, *CFH*, a strong genetic risk factor for AMD, had much higher gene expression levels in RCS than in retina (FPKM was ~200 in RCS but was only ~5 in retina). With respect to the genes associated with AMD, 22 genes were differentially expressed between retina and RCS macula and periphery, and 11 did not demonstrate any differential expression. We then focused on 14 genes that have been well-studied in genetics and tissue-level studies. Five genes in the alternative complement pathway ($C3>CFH>>CFI=CFB>C2$) were expressed at much higher levels in RCS versus retina, with regional differential expression for *CFI* only (MR > PR). Three genes in cholesterol transport pathways ($APOE>>CETP>>LIPC$) were also expressed at higher levels in RCS versus retina, with differences as to regional effects (*APOE*, periphery > macula for both retina and RCS; *CETP*, similar in macula and periphery; *LIPC*, higher in periphery only). *VEGFA* was unusual in that it was expressed at much higher levels in MR than MRCS, without evidence of regional differential expression. *ARMS2* showed no differential expression whereas *HTRA1*, which is in strong linkage disequilibrium with the *ARMS2* locus, was interestingly higher in MR and MRCS versus PR and PRCS, respectively. *SLC16A8*, encoding a monocarboxylate transporter called MCT3, and *CD46*, a membrane-bound regulator of complement, are both exclusively expressed on the basolateral RPE and exhibit stereotypical stage-specific mis-localization and decline in expression level in geographic atrophy of AMD (22). *TIMP3* protein is abundant in aging Bruch's membrane and in sub-RPE drusen and is known to be expressed by RPE, yet in our samples, it was almost exclusively expressed by retina (23,24). The fact that many AMD genes are expressed in retina and RCS, although at very different levels, underscores the continuing need for precise localization of gene products in validation studies because both the subretinal and sub-RPE compartments each have specific and distinctive AMD pathology (subretinal drusenoid deposit versus drusen; type 3 neovascularization versus type 1 neovascularization) (25,26).

Since we also generated DNA genotypes using Illumina Omni 2.5 SNP chip, we examined whether GWAS SNPs showed evidence of allele-specific expression (ASE) in our RNA-Seq data. ASE quantifies transcription from both paternal and maternal haplotypes using individuals heterozygous at tested SNPs. We extracted 120 SNPs associated with eye diseases at *P*-values of $<5\times10^{-8}$ based on the NHGRI GWAS catalog. Among these, 10 are exonic and are in the Omni 2.5 SNP chip. For each of these 10 SNPs, we extracted RNA-Seq reads for those subjects that were heterozygous in the DNA and tested ASE using the chi-squared statistic. We found evidence of ASE for three SNPs, rs1061147 in *CFH*, rs2230199 in *C3* and rs641153 in *CFB* (Table 4), all previously reported in AMD GWAS (15). Both rs2230199 and rs641153 are non-synonymous and had been annotated as functional based on ENCODE data (27). rs2230199 (Arg80Gly) is located within a regulatory region identified by DNase-Seq from multiple cell types by ENCODE. This SNP is also located within a protein-binding region based on CHIP-Seq analyses. rs641153 (Glu32Arg) is located within a protein-binding region for TCF4, and this region showed evidence of histone modifications. However, rs1061147 is synonymous and had not been found to be functional based on ENCODE. The fact that this SNP showed evidence of ASE in the eye suggests that it may associate with AMD by regulating the gene expression of *CFH*.

Table 4. Evidence of ASE at GWAS-associated SNPs for eye diseases

SNP	Position	Gene	Reference allele	Alternative allele	Subject	Sample	ASE ratio	*P*-value
rs2230199	chr19:6718387	C3	G	C	11–1516	PR	0.75	2.54E-02
					11–1516	PRCS	0.73	3.66E-09
					11–1516	MRCS	0.65	3.75E-05
					11–2043	MR	0.60	6.33E-05
rs1061147	chr1:196654324	CFH	A	C	11–1516	PRCS	0.46	4.42E-02
					11–1624	PRCS	0.47	4.77E-02
					11–1624	MRCS	0.41	6.89E-08
rs641153	chr6:31914180	CFB	G	A	11–1614	MRCS	0.69	4.11E-02
					11–1833	PRCS	0.66	2.93E-04

ASE ratio was calculated as the fraction of the alternative allele among all alleles at the SNP. All three SNPs were associated with AMD.

Biological replication of differential expression
A total of 204 genes detected to be DE by Cufflinks from the Discovery set were chosen for biological replication based on their biological relevance and significance of the *P*-values. The Replication set included eight normal eyes, but two MR samples were eliminated owing to poor RNA quality. Paired *t*-test was used to determine *P*-value for differential expression. Percentages of DE genes replicated were as follows: MR versus MRCS (61% replicated), PR versus PRCS (83% replicated), MR versus PR (59% replicated) and MRCS versus PRCS (29% replicated).

Analysis of alternative splicing

Global analysis of alternative splicing
RNA-Seq's specificity for closely related transcript isoforms allowed us to go beyond gene expression level and identify isoform differences owing to pre-mRNA alternative splicing (AS). Using MATS (28), we identified a large number of AS events, particularly in retina (Fig. 2A). The average numbers of AS events were 15 560 in MR and 15 290 in PR, whereas the corresponding numbers were 11 760 in MRCS and 10 050 in PRCS, respectively. These results suggest that many genes were alternatively spliced and the AS patterns were different between the retina and RCS tissue layers.

For insight into AS patterns in different chorioretinal tissues and locations, we carried out hierarchical clustering analysis based on the top 150 most divergent AS events defined by the coefficient of variation of the exon-inclusion levels (Fig. 2B). Interestingly, the retina samples clearly separated from the RCS samples, although only 150 AS events were used in the clustering. However, within each tissue, the distinction between macula and periphery was less clear with retina showing a slightly better separation of macula and periphery. These results suggest the existence of tissue-specific AS signatures, but the difference between locations within the tissue layer was subtle.

Differential alternative splicing
Next, we examined whether an AS event showed different degrees of splicing between different chorioretinal tissues of the eye by testing whether the difference in exon-inclusion levels exceeded 5% for any cassette exon of a gene. Figure 2C shows the numbers of differential AS (DAS) events. When comparing retina and RCS, we found 431 DAS events in macula and 575 in periphery. In contrast, we detected a much smaller number of DAS events between MR and PR (39 events) and between MRCS and PRCS (2 events). These results were consistent with the patterns revealed by the hierarchical clustering analysis.

For DAS events identified from the comparison of retina and RCS, we further examined their exon-inclusion levels (Fig. 2D and E). In macula, more DAS events appeared to have higher exon-inclusion levels in retina than in RCS (MR versus MRCS: $P = 0.027$). There was a similar trend in periphery, but the difference was not significant (PR versus PRCS: $P = 0.13$).

We recognize that the analysis of DAS is challenging because detection of AS events relies on the number of junction reads. To assess the reliability of our findings, we classified DAS events into three tiers according to the degrees of confidence. Tier 1 includes high-confidence events that have been annotated by RefSeq, i.e. both the inclusion and the exclusion isoforms are present in RefSeq's annotation. Tier 2 includes moderate-confidence events for which the corresponding genes are alternatively spliced according to uniProt. Tier 3 includes the remaining events. Based on these classifications, we confirmed that $>80\%$ of the DAS events were in tier 1 for comparisons between MR and MRCS, MR and PR and PR and PRCS. Of the two DAS events identified for the comparison between MRCS and PRCS, one was in tier 1 and the other was in tier 2 (Fig. 2F). These results suggest that our findings on DAS were reliable.

Splicing factors are RNA-binding proteins that play key roles in coordination of spliceosome functions and regulation of AS. We compiled a list of 60 well-studied splicing factors and examined their expression in all samples. Thirty-seven splicing factor genes were expressed in chorioretinal tissues: 12 being DE between MR and MRCS, 16 between PR and PRCS and 11 DE between retina and RCS in both regions. As expected, the number of DE genes was much smaller when comparing different regions; only three genes were differentially expressed between MR and PR, and none were differentially expressed between MRCS and PRCS. These findings are consistent with the number of DAS events in that the number of DAS events identified in each comparison perfectly correlated with the number of splicing factors that were differentially expressed, suggesting that the differences in DAS events were likely due to the expression difference of related splicing factors.

Enrichment analysis of genes with differential alternative splicing
Next, we carried out functional annotation analysis for genes showing DAS when comparing retina and RCS. For both macula and periphery, genes were significantly enriched for actin filament-based process, cytoskeleton organization and anatomical structure development in GO, and for tight junction in KEGG. We did not perform functional annotation analysis for the comparison between macular and periphery owing to the small number of genes showing DAS.

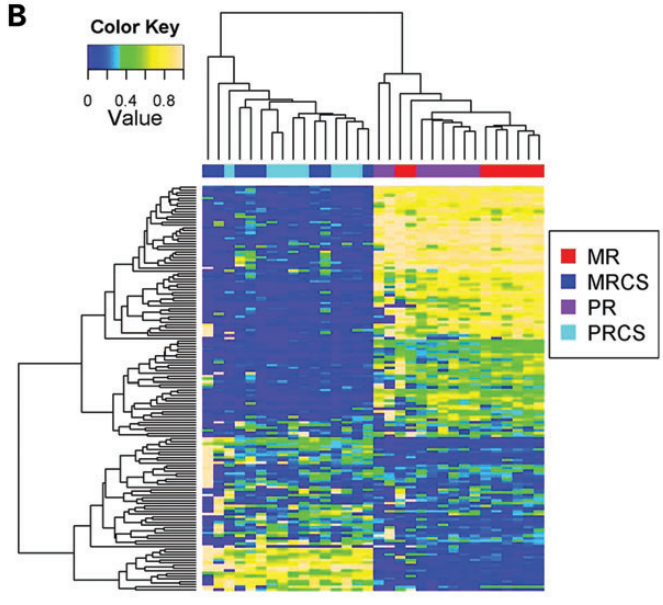
Overlap with GWAS signals on eye diseases
For genes showing DAS, we further examined whether they had been implicated in recent GWAS on AMD (15), glaucoma (29) and myopia (30) based on the NHGRI GWAS catalog. Among the 91 unique GWAS genes we examined, three genes—*CD46*, *DDR1* and *RREB1*—showed DAS between PR and PRCS. Interestingly, all three were identified from GWAS on AMD (15), and all had higher expression levels in RCS than in retina. Of these genes, *CD46* (also called membrane cofactor protein or MCP) is notably the only membrane-bound regulator of complement that is expressed on the basolateral RPE (22,31). *DDR1* is a hypoxia-induced gene involved regulation of cell growth and is expressed in lamina cribrosa cells in the optic nerve (32). We did not find overlap with GWAS signals for genes showing DAS identified from other comparisons.

Biological replication of differential alternative splicing
Five genes showing DAS by MATS were chosen for replication in a Replication set of an additional eight normal postmortem eyes. All five genes were replicated in all eight eyes demonstrating the robustness of our analysis. Figure 3 shows gel images for two representative eyes.

A

Number of AS events

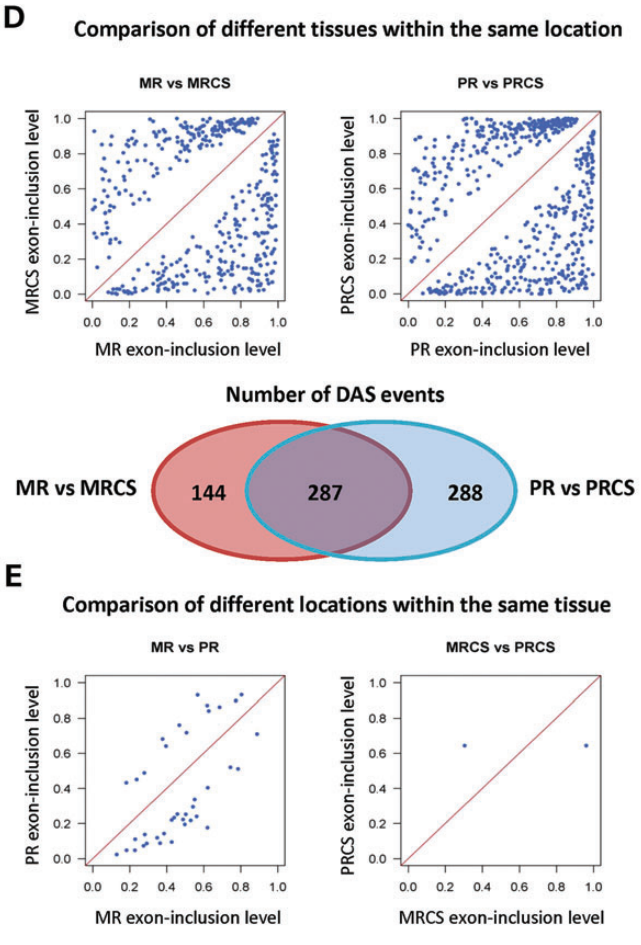
Group	Minimum	Maximum	Mean
MR	12,830	17,080	15,560
MRCS	9,744	17,850	11,760
PR	12,560	16,960	15,290
PRCS	9,033	12,840	10,050



C

Number of DAS events identified from each comparison

Comparison	Number of DAS events	Number of DAS events with higher exon-inclusion level in group 1	Number of DAS events with higher exon-inclusion level in group 2
MR vs MRCS	431	239	192
PR vs PRCS	575	306	269
MR vs PR	39	26	13
MRCS vs PRCS	2	1	1



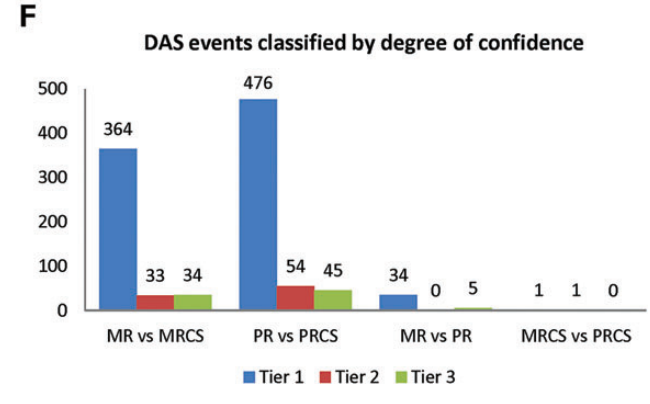


Figure 2. Alternative splicing (AS). (**A**) Numbers of AS events. An exon was considered alternatively spliced if its exon-inclusion level was >0 but <1. The minimum, maximum and mean numbers of AS events were calculated based on the eight samples in each eye part. (**B**) Hierarchical clustering analysis of top 150 most divergent alternative splicing events. For each AS event, we calculated its CV based on its exon-inclusion levels across all RNA-Seq samples. The AS events were then ranked based on their CV values. The heatmap was generated by hierarchical clustering of the top 150 AS events with the largest CV values. (**C**) Number of DAS genes identified from each comparison. Analysis was completed using MATS. An exon was considered DAS if FDR-adjusted *P*-value was <0.05. (**D**) Comparison of exon-inclusion levels between different tissues within the same location. The Venn diagram shows the numbers of DAS events that were unique to each comparison as well as the number of DE genes that were common between the two comparisons. (**E**) Comparison of exon-inclusion levels between different locations within the same tissue. (**F**) Classification of DAS events by degree of confidence. Tier 1 includes high-confidence events that have been annotated by RefSeq; tier 2 includes moderate-confidence events for each the corresponding genes are alternatively spliced according to uniProt; tier 3 includes the remaining events.

DISCUSSION

GWASs have been used to correlate genetic variants with disease in a hypothesis-generating manner. Linking these genetic variants with causation has been the next major milestone, with tissue specificity, low resolution of DNA genotypes and technical challenges of assaying molecular traits being the limiting factors to date. Previous studies have examined ocular

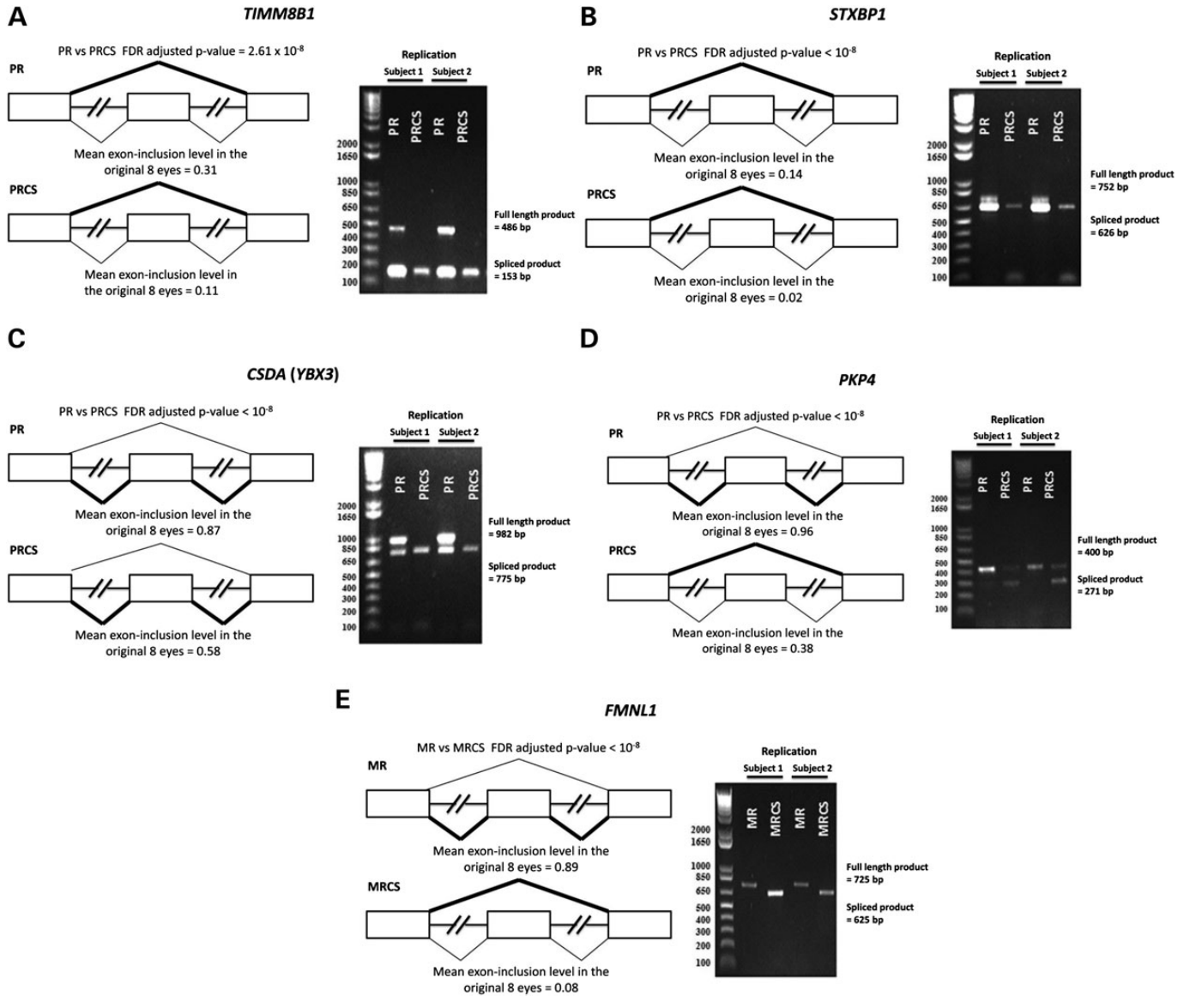


Figure 3. RT-PCR biological replication of differential alternative splicing (DAS) events. (**A**) Replication of *TIMM8B1* DAS event from PR versus PRCS comparison. (**B**) Replication of *STXBP1* DAS event from PR versus PRCS comparison. (**C**) Replication of *CSDA* (*YBX3*) DAS event from PR versus PRCS comparison. (**D**) Replication of *PKP4* DAS event from PR versus PRCS comparison. (**E**) Replication of *FMNL1* DAS event from MR versus MRCS comparison.

gene expression by RT-PCR and microarrays but have not yet attempted the entire transcriptome of human neurosensory retina and photoreceptor support system by RNA-Seq. In the present study, we conducted a comprehensive RNA-Seq analysis (16,33) of MR and PR and macular and PRCS from eight phenotypically well-characterized normal adult eyes. The molecular signatures that we obtained for macula and periphery will be very useful biomarkers for characterizing different stages of chorioretinal disease. Our results for RNA-Seq are to our knowledge novel owing to the comparison of MR and PR and for the same regional comparison of the RCS.

RNA-Seq findings from multicellular tissues should be interpreted in light of the cellular composition of the tissue samples. Human neurosensory retina is dominated by densely packed neurons and includes significant minority populations of radial glia (Müller) cells, astrocytes, microglia and vascular endothelial cells and pericytes. RCS samples, in contrast, include multifunctional RPE, multicellular choroid (with vascular and lymphatic endothelia, fibroblasts, melanocytes, mast cells, autonomic neuronal ganglia and resident and transient cells of monocyte lineage) and structural sclera (with sparse fibroblasts and myofibroblasts) (34,35). To effectively study and understand biological phenomena at the molecular level, variability in gene expression must be taken into account. This variability in gene expression can exist for two reasons. First, the tissue may be dominated by a large number of different cell populations in differing amounts, which can result in a dilution of the signal from less-populated cells below the lower detection limit for RNA-Seq. Second, cell populations can average out individual cellular co-expression patterns, making it difficult to

determine whether two RNAs always occur within the same cell. We hypothesized that our expression from the retina was not sensitive enough to detect gene expression from less abundant retinal cells. We expected neurosensory retina RNA expression to be dominated by the transcriptome of the numerically dominant photoreceptor population (>100 M cells), yet we found robust RNA-Seq signal for genes expressed by less populous neuronal and glial cell classes (36). The cell classes include *POU4F2* (ganglion cells) (37), *SNAP25* (horizontal cells) (38), *CALB2* (amacrine cells) (39), *RCVRN* (bipolar cells) (39) and *GLUL* (Müller cells) (40). FPKM values for all five genes varied from a low of 1.0 for *POU4F2* to a high of 4148 for *RCVRN*, suggesting that we had excellent representation of the retina's less numerous cells. Next, we asked whether signal for specific neuronal types was also represented in the RNA-Seq data. We found high RNA expression for *CBP5* (39) (specific for bipolar cells DB3, DB4, RB; 258 FPKM) and *SYT2* (39) (specific for amacrine cells A8; 23 FPKM) among others, suggesting that our utilization of whole retina RNA was sensitive enough to detect RNA expression from less common retinal cells. We did not interrogate the RCS samples with the same questions because cell populations in those tissues are not as well defined quantitatively as they are in neurosensory retina, yet we would expect on the basis of this analysis that all RCS cell types have been adequately represented as well.

The largest difference in expression in these eight unremarkable adult human eyes occurs between the neurosensory retina and the RCS. In total, 9555 genes in macula, and 10 149 genes in periphery were differentially expressed between layers, and these were evenly distributed between those that were higher in retina and those that were higher in RCS in both locations. Of the genes enriched in the retina as compared with RCS and those enriched in MR as compared with PR, a number of genes were found to be associated with psychiatric disorders following DAVID analysis (41). We examined the top 25 most significant DE genes for each pairwise comparison (MR versus MRCS, PR versus PRCS and MR versus PR). This yielded a list of 50 unique genes. Following an expansive literature review, we found that 45 of them were implicated in one or more psychiatric disorders, including schizophrenia ($n = 24$), Alzheimer's disease ($n = 8$), autism ($n = 8$), depression ($n = 6$) and bipolar disorder ($n = 6$) (Supplementary Material, Table S5). Of these 50 genes, mutations in eight are known to have a specific eye phenotype in humans. These include *DDC*, *GAD1*, *GAD2*, *MAPT*, *PAX6*, *PPP2R2B*, *SYP* and *TFAP2B*. Common eye phenotypes seen include oculomotor apraxia, ptosis, nystagmus and ophthalmoplegia. One gene of particular interest is *AQP4*, aquaporin 4. Enhanced *AQP4* expression has been found in the brains of patients with Alzheimer's disease (42). An eye phenotype in humans has not been documented, but $Aqp4^{tm1Ask}/Aqp4^{tm1Ask}$ mice display significantly reduced electroretinogram (ERG) b-wave potentials (43). Patients with Alzheimer's disease are known to display attenuated pattern- and multifocal-electroretinogram responses as well as decreases in retinal nerve fiber layer thickness as determined by optical coherence tomography (OCT) (44–47). Together, these findings suggest that ERG b-wave response and OCT may be useful tools to identify individuals at risk for developing Alzheimer's disease.

We further explored tissue-specific variation in chorioretinal gene expression by examining the top 25 most significant DE genes between retina and RCS for both macula and periphery (Supplementary Material, Table S6). Many DE genes preferentially expressed in RCS fall into the cardiometabolic category. They are a diverse group, expressing in muscle, vessel walls and blood cells, and they are well explored in systemic vascular diseases. The top genes expressed at higher levels in RCS than retina in both macula and periphery include *BDKRB2*, *CYP4B1*, *DES*, *HBB*, *JPH2* and *SAA1*. Of these, *BDKRB2*, *DES* and *HBB* have been investigated with regard to the retinal circulation in the context of diabetic retinopathy and remain to be explored in choroid. SAA1 encodes an apolipoprotein and acute phase responder found in isolated RPE-capped drusen (48). The top genes expressed at higher levels in RCS than retina for macula only (Supplementary Material, Table S6) notably include *COL1A1*, *COL1A2* and *COL3A1*, which encode structural collagens, consistent with the fact that choroid and sclera are both at their thickest under the macula and thus contain many fibroblasts and myofibroblasts (27,49). Polymorphisms of COL genes are associated with systemic hypermobility of joints and disorders of vascular disintegrity. Other highly expressed macula RCS genes of interest for roles in AMD initiation and/or progression are *CR1* (found in RPE, BrM and blood leukocytes), *IL6* (an acute mediator of inflammatory responses found in aqueous of AMD eye), *THBS1* (an anti-angiostatic protein reduced in AMD BrM) and *VCAM1*. The top genes expressed at higher levels in RCS than retina in periphery only include many cellular receptors (*ADRA1A*, *ADRA2C*, *AGTR1*, *CD36*, *NPR1* and *NR1H4*). RCS expression of these genes is apparently unmasked by thinner neurosensory retina in this region. As the support system for the photoreceptors, the RCS exhibits diverse gene expression in service of maintaining the health of these cells for good visual function. An important next step will be assigning these expression patterns to the multiple RCS cell types.

As an external validation of genes important in cardiovascular disease, we compared our chorioretinal RNA-Seq expression data to microarray gene expression data for chorioretinal cholesterol pathways published by Zheng *et al*.(50), who analyzed neurosensory retina and RPE of six human donor eyes using a commercially available microarray. This comparison is of interest because esterified cholesterol and phosphatidylcholine are the major volumetric components of drusen (48), AMD's pathognomonic extracellular lesions located on the inner surface of the choroid, and unesterified cholesterol is present in subretinal drusenoid deposit (reticular pseudodrusen), a lesion located between the photoreceptors and the RPE in many AMD eyes (51,52). Zheng *et al*. (50) reported expression levels for 84 genes in these two tissues normalized against 6 housekeeping genes. For these 84 genes, we ranked expression levels within our MR, MRCS, PR and PRCS groups, from highest to lowest and compared their ranks to the Zheng dataset (50) (Supplementary Material, Fig. S2). Two observations can be made. First, all microarray genes were also detected by RNA-Seq in all four tissue sample groups. Second, RNA-Seq expression levels were strongly positively correlated with microarray results for these genes. We note that agreement was higher for MR and PR (0.67) than for MRCS and PRCS (0.51–0.54). These findings are plausible because both studies assayed retina, whereas Zheng *et al*. (50) assayed RPE only versus RCS for RNA-Seq. We can conclude that many major genes in cholesterol and lipoprotein

pathways are expressed and likely functional in retina, RPE and choroid. This comparison comprised an external validation of our RNA-Seq results with regard to a focused set of genes. These corroborative findings also emphasize that pathogenic models based on GWAS-identified SNPs in genes well known from plasma HDL metabolism and expressed in liver, intestine and adipose tissue should incorporate the intraocular expression of these genes.

Another striking finding is that >2000 genes were differentially expressed between MR and PR and >900 genes between MRCS and PRCS. These genes were slightly tilted toward those that were more highly expressed in RCS versus retina. The differential expression between macula and periphery is attributable in part to differences in the cellular composition of these regions. For retina, rod photoreceptors are the most abundant cell type (100 million) (53) with cones more numerous in macula than they are in the periphery but still a minority (<10% of total in macula). The cone-dominated fovea, <0.8 mm in diameter, contains 32 000 cones (53,54). The ratio of macula photoreceptor densities to peripheral photoreceptor densities is 1.9 for cones (because cones dominate a small area) and 2.7 for rods (53,55). A feature unique to macula is the Henle fiber layer, which contains processes of cone and rod photoreceptors transmitting graded potentials from the light-capturing inner segment to synaptic terminals, interleaved by parallel processes of Müller cells. Individual Henle fibers can be up to 600 μm long, and collectively the Henle fiber component of the anatomical outer plexiform layer (Fig. 4) occupies 14% of macular volume at its thickest portion (56–58). The macula contains half (∼750 000) of all brain-projecting ganglion cells of the human retina (59); these cells are dominated by midget cells required for high-acuity vision (60). Mean ganglion cell density is 100-fold greater in macula than in periphery. For the choroid, differential cellular composition may be attributable to blood cells retained within the more highly vascularized macula choroid (61) and numerous melanocytes that impart to the macula its greater pigmentation relative to periphery in this mature age group (62). Therefore, our differential expression results between macula and periphery are expected based on the anatomical differences between these regions and will likely be borne out by regional differences in choroid that are being explored anew through high-resolution imaging.

In conclusion, the transcriptome data presented here is the first report showing differential expression between regions within the retina and RCS by RNA-Seq. The magnitude of these differences is striking and expands our catalog of the gene expression variation that exists in the normal eye. The data underscore the need to not only fully understand the expression differences between layers but also between regions of a layer. Overall biological interpretation of regulatory effects—much like in the case of complex traits—is tissue- and region-dependent, highlighting the value of multiple tissue expression datasets. Gene expression is a critical phenotype that reveals biological properties and allows us to probe cellular functions. Combining statistical methods with relevant sample collections of tissues and cell types from well-phenotyped individuals enables the integrated treatment of biological and epidemiological information in an iterative way. This provides us with the highest possible resolution and will reveal the real causes for disease predisposition.

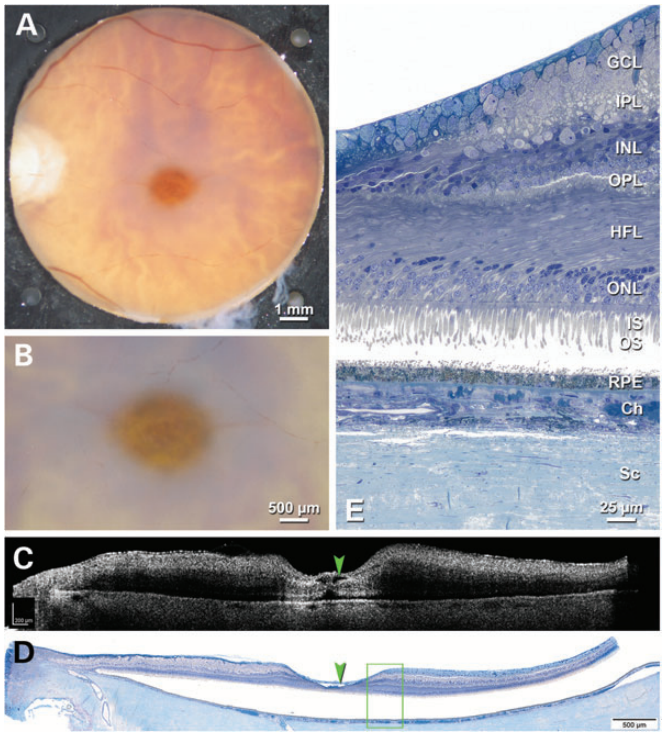


Figure 4. *Ex vivo* imaging and histology of eye contralateral to that used for RNA-Seq. (**A** and **B**) Color image of excised preserved macula (A) and fovea (B) from a 79-year-old Caucasian male. Optic nerve is at the left edge of the punch. (**C**) *Ex vivo* spectral domain optical coherence tomography of excised preserved macula, with neurosensory retina at the top and RPE–choroid–sclera at the bottom. The optic nerve is at the left, and the fovea is the dip in the middle. This fovea has postmortem cystic change (arrowheads). The macula is otherwise unremarkable. (**D**) 0.8-μm-thick epoxy section of tissue post-fixed with the OTAP method, stained with toluidine blue and scanned. Detachment of the neurosensory retina from the RPE is common in postmortem specimens, even if microscopically well-preserved like this one. Green frame indicates area shown at higher magnification in **E**. (E) Chorioretinal layers on the foveal slope. This area of the retina is cone-dominated, as evidence by the few darkly stained rod nuclei that contrast with numerous cone nuclei in the ONL. Layers: GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; HFL, Henle fiber layer; IS, inner segments of photoreceptors; OS, outer segments of photoreceptors; RPE, retinal pigment epithelium; Ch, choroid; Sc, sclera. Detached retina was digitally reapposed to RPE for illustrative purposes (Eye 2011001R is shown as a left eye by UAB lab convention. As AEB 11-1515-P, it is contralateral to 11-1516-P analyzed by RNA-Seq.).

This view has the potential to also lead to higher confidence candidates in the absence of direct functional support for any one gene, as might be the case for inconclusive GWAS findings where the SNPs identified have no known functional role. Generating even larger-scale expression datasets may provide a path to more rapidly elucidating not only the genetic basis of eye disease but also the impact of gene expression on molecular networks that in turn induce variations in disease associated traits.

MATERIALS AND METHODS

Eye collection

This study conformed to Institutional Review Board regulations for use of human tissues at University of Alabama (UAB) and at

University of Pennsylvania (Penn). Our study utilized two sets of eight pairs of eyes from non-diabetic Caucasian donors at a death-to-preservation interval of <6 h. This interval was chosen because RPE RNA quality is adequate during this time frame and begins to decline thereafter (63). A Discovery set of eyes consisted of five males and three females [73.9 year $\pm$ 12.5 year (mean $\pm$ standard deviation)]. A Replication set also consisted of five males and three females (85.9 $\pm$ 3.0 year) with somewhat older ages owing to an adjustment of the tissue collection protocol to maximize the number of AMD eyes collected for separately reported studies. Ocular health histories were not available. Eyes were opened anteriorly by Alabama Eye Bank recovery personnel using an 18-mm-diameter trephine, followed by a radially oriented snip to the iris margin to facilitate penetration of preservatives into the fundus. Preservatives used were RNAlater (Qiagen, Valencia, CA, USA) for the left eye and 2% glutaraldehyde and 1% paraformaldehyde in 0.1 M phosphate buffer for the right eye, both at 4°C. Left eyes were shipped on wet ice via overnight courier to Penn where they were processed upon arrival.

Maculopathy status of right eyes was assessed at UAB by a three-component protocol. Eyes underwent multimodal *ex vivo* imaging of excised 8-mm-diameter macular punches using digital color photography and spectral domain optical coherence tomography volume scans (SD-OCT; Spectralis, Heidelberg Engineering) with a custom tissue holder. They underwent internal globe examination using a dissecting scope (Nikon SMZ-U) with oblique trans- and epi-illumination in consultation with an MD medical retina specialist (J.A.K.). Finally, eyes were submitted for histopathology using macula-wide high-resolution sections. Macular punches including retina, RPE, choroid and sclera were post-fixed in osmium tannic acid paraphenylenediamine to accentuate neutral lipid-rich lesions associated with AMD (52,58). Sections of 0.8 μm in thickness through rod-free foveola and rod-dominant perifovea at 2000 μm superior to the fovea were stained with toluidine blue, examined and annotated (digital sections are available for viewing at http://projectmacula.cis.uab.edu/?page_id=541). These locations were chosen to capture distinctive features of macular photoreceptor inhomogeneity and because distinctive pathologies associated with AMD (severe RPE change with extracellular lesions, with or without neovascularization) are located in each of these sites (52). Discovery and Replication sets of eyes used in this study had only age-appropriate chorioretinal changes, by all three examinations (Fig. 4).

Left eyes were examined at Penn prior to preparation for RNA-Seq by internal examination and photography with a stereo dissecting microscope. This procedure was adequate to detect between-eye differences in chorioretinal disease status, as dictated by a long-term objective of analyzing AMD eyes, in studies to be reported separately. The fellow eye design was adopted owing to the strong correlation between eyes of the same individual to manifest an AMD phenotype (64–67). We do realize that key phenotypic features of age-related chorioretinal disease, i.e. abundance of AMD-specific lesions, are best recognized in high-resolution histopathologic preparations, but these methods are incompatible with using the same maculas for RNA-Seq.

RNA-Seq library preparation and sequencing

RNA for the Discovery set of eight eyes was extracted using the AllPrep DNA/RNA Mini Kit (Qiagen). For each eye, we dissected macula and periphery from retina and RCS resulting in four samples per eye. We denote these samples as MR, MRCS (macular RPE/choroid/sclera), PR (periphery retina) and PRCS. The tissues were isolated at the macula and periphery using a circular 10-mm biopsy punch followed by a second biopsy punch of 8 mm in the center of the 10-mm punch after retina removal to minimize contamination of underlying tissues with retina. Extracted RNA samples underwent quality control assessment using R6K Screen Tape on a 2200 Tape Station (Agilent, Santa Clara, CA, USA) and were quantified using Qubit 2.0 Fluorimeter from Life Technologies (Grand Island, NY). All RNA samples selected for sequencing had an RNA integrity number of ≥ 8. Strand-specific RNA library was prepared from 100 ng total RNA using the Encore Complete RNA-Seq library kit (Nugen Technologies, Inc., San Carlos, CA, USA) according to the manufacturer's protocol. Briefly, non-rRNA-enriched cDNA was prepared and converted to libraries with adaptors and reagents provided in the kit. The library quality was assessed by qPCR using Kappa Library Quant kit (Kappa Biosystems, Woburn, MA, USA). RNA sequencing was performed at the Center for Applied Genomics at the Children's Hospital of Philadelphia per standard protocols. The prepared libraries were clustered and then sequenced using HiSeq 2000 sequencer (Illumina, Inc., San Diego, CA, USA) with four RNA-seq libraries per lane (2×101-bp paired-end reads).

Alignment of RNA-Seq reads

The RNA-Seq data were aligned to the hg19 reference genome using GSNAP (16) with default options. In order to eliminate mapping errors and reduce potential mapping ambiguity owing to homologous sequences, several filtering steps were applied. Specifically, we required the mapping quality score of ≥ 30 for each read, reads from the same pair were mapped to the same chromosome with expected orientations and the mapping distance between members of the read pair was $<500\,000$ bp. Quality control analysis of the aligned data was performed using program RNA-SeQC (33). All subsequent analyses were based on filtered alignment files.

Analysis of differential gene expression

Transcripts were assembled using Cufflinks software (18). For each gene, we compared the expression levels between chorioretinal samples, including MR versus MRCS, PR versus PRCS, PR versus MR and PRCS versus MRCS. The first two comparisons allow us to identify genes that are differentially expressed between different tissues, whereas the last two comparisons allow us to identify genes that are differentially expressed between different locations within the same tissue layer. A gene was considered differentially expressed if the false discovery rate (FDR)-adjusted P-value is <0.05.

Analysis of differential alternative splicing

The availability of RNA-Seq data on the same subject in different tissues and locations provides an opportunity to identify differential alternative splicing (AS) events. To identify such events, we used MATS (28). This software implements a Bayesian approach that detects differential AS under two conditions by examining whether the difference in the exon-inclusion levels between two samples exceeds a given user-defined threshold (0.05 in our analysis). An AS event was declared if the inclusion level of an exon is between 0 and 1. A differential AS (DAS) event was declared if the FDR-adjusted *P*-value is <0.05.

Functional annotation

To identify overrepresented functional categories among genes that are differentially expressed or differentially spliced, we carried out annotation analysis using the DAVID (41). Differentially expressed genes or genes showing DAS were used as input gene list. We looked for enrichment for genetic association with disease class, biological processes in GO and KEGG pathways. Multiple testing was adjusted using the Benjamini–Hochberg approach, and enrichment was declared if the adjusted *P*-value is <0.05.

Biological replication of differential gene expression results via Nanostring

Using the Nanostring nCounter Analysis System (Nanostring Technologies, Seattle, WA, USA), gene expression analysis was conducted on six new phenotypic normal maculas and eight new phenotypic normal peripheral tissues using a custom-designed code set containing 204 genes as previously described (68). These genes were chosen based on biological relevance and having evidence of DE based on Cufflinks analysis (fold change >2 and an FDR-adjusted $P < 0.05$). Each reaction contained 100 ng of total RNA in a 5 ul aliquot, plus reporter and capture probes. Analysis and normalization of the raw Nanostring data was conducted using nSolver Analysis Software v1.1 (Nanostring Technologies). Raw counts were normalized to internal levels of 14 reference genes: *ADAM10, APP*, *APH1A*, *FHIT*, *SNCA*, *GSTP1*, *NFE212*, *DIRAS2*, *SIRT2*, *CYC1*, *GIGYF2*, *CTIF*, *YWHAZ* and *RPL22.* The background was subtracted from the data using the average of the negative controls (included in each Nanostring Codeset). Paired *t*-test was used to assess differential expression. A gene was considered replicated if the *P*-value is <0.05.

Biological replication of differential alternative splicing results via PCR

Replication of our DAS for five genes, *TIMM8B1*, *CSDA*, *PKP4*, *STXBP1* and *FMNL1*, was performed in a Replication set of eight additional eyes. To simplify the replication analysis, genes were chosen from the MATS results that had only two isoforms. Using the genomic coordinates included in the MATS output, the UCSC genome browser was used to visually localize the exon's location in relation to the flanking exons shared by both isoforms. A portion of the sequence from the exons flanking the differentially spliced exon was taken from sequence data at NCBI and used to create the primers using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/). To ensure that these primers were target-specific, Primer-BLAST was used (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primers are listed in Supplementary Material, Table S7. RNA from the eight replication eyes was extracted using the AllPrep DNA/RNA Mini prep Kit (Qiagen). Extracted RNA samples underwent quality control assessment using R6K ScreenTape on a 2200 Tape station (Agilent) and quantified using Qubit 2.0 Fluorometer (Life Technologies, Foster City, CA, USA). In total, 20 ng of RNA was used to make cDNA using random hexamers for priming and Superscript III first strand synthesis system (Life Technologies). The cDNA was prepared according to the manufacturer's protocol. The cDNA was amplified using GoTaq Green Master Mix (Promega, Madison, WI, USA) and gene-specific primers. A touchdown PCR method with 68–55°C temperature gradient was used to amplify the different transcripts. The PCR products were visualized on an Agarose 1.5% gel using ethidium bromide dye.

DNA genotyping and quality control

DNA for the eight Discovery eyes was extracted using AllPrep DNA/RNA Mini Kit. Simultaneous purification of genomic DNA and total RNA from same sample was achieved using this kit. Lysate from each tissue was passed through an AllPrep DNA spin column to selectively isolate DNA and then through RNeasy MinElute spin column to selectively isolate RNA. The extracted genomic DNA quality and quantity was measured using NanoDrop 8000 (Thermo Fisher, DE). All eight DNA samples had a 260/280 ratio of $\geq$ 1.7. All eyes were genotyped at the Center for Inherited Disease Research using the Illumina HumanOmni2.5-4v1_B SNP array.

We performed quality control analysis to ensure the quality of the DNA genotypes. First, we required all samples to have genotype call rate of $>97.5\%$ and no excess or deficient heterozygosity (inbreeding coefficient $|F| < 0.1$). Next, we performed quality control measures to exclude unreliable SNPs. We eliminated SNPs with genotype call rate of $<98\%$ or if there was significant departure from Hardy–Weinberg equilibrium ($P < 0.0001$).

Analysis of allele-specific expression

To test for allele-specific expression (ASE), we considered exonic SNPs reaching GWAS significance ($P < 5 \times 10^{-8}$) for eye-related diseases based on NHGRI GWAS catalog (http://www.genome.gov/26525384). We identified subjects that are heterozygous at the test SNPs. We counted the number of RNA-Seq reads carrying each allele and tested ASE using a one degree of freedom χ^2 goodness-of-fit test for equal frequencies of both alleles.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

The Arnold and Mabel Beckman Initiative for Macular Research.

REFERENCES

1. Freedman, M.L., Monteiro, A.N., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D. *et al.* (2011) Principles for the post-GWAS functional characterization of cancer risk loci. *Nat. Genet.*, **43**, 513–518.
2. Stambolian, D. (2013) Genetic susceptibility and mechanisms for refractive error. *Clin. Genet.*, **84**, 102–108.
3. Cooke Bailey, J.N., Sobrin, L., Pericak-Vance, M.A., Haines, J.L., Hammond, C.J. and Wiggs, J.L. (2013) Advances in the genomics of common eye diseases. *Hum. Mol. Genet.*, **22**, R59–R65.
4. 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.
5. Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Dolan, M.E. and Cox, N.J. (2010) Trait associated SNPs are more likely to be eQTLS: Annotation to enhance discovery from GWAS. *PLoS Genet.*, **6**, e1000888.
6. Goring, H.H., Curran, J.E., Johnson, M.P., Dyer, T.D., Charlesworth, J., Cole, S.A., Jowett, J.B., Abraham, L.J., Rainwater, D.L., Comuzzie, A.G. *et al.* (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat. Genet.*, **39**, 1208–1216.
7. Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S. and Cheung, V.G. (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature*, **430**, 743–747.
8. Twine, N.A., Janitz, K., Wilkins, M.R. and Janitz, M. (2011) Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's disease. *PLoS One*, **6**, e16266.
9. Cai, H., Fields, M.A., Hoshino, R. and Priore, L.V. (2012) Effects of aging and anatomic location on gene expression in human retina. *Front. Aging Neurosci.*, **4**, 8.
10. Hornan, D.M., Peirson, S.N., Hardcastle, A.J., Molday, R.S., Cheetham, M.E. and Webster, A.R. (2007) Novel retinal and cone photoreceptor transcripts revealed by human macular expression profiling. *Invest. Ophthalmol. Vis. Sci.*, **48**, 5388–5396.
11. Sharon, D., Blackshaw, S., Cepko, C.L. and Dryja, T.P. (2002) Profile of the genes expressed in the human peripheral retina, macula, and retinal pigment epithelium determined through serial analysis of gene expression (SAGE). *Proc. Natl. Acad. Sci. USA*, **99**, 315–320.
12. Wagner, A.H., Anand, V.N., Wang, W.H., Chatterton, J.E., Sun, D., Shepard, A.R., Jacobson, N., Pang, I.H., Deluca, A.P., Casavant, T.L. *et al.* (2013) Exon-level expression profiling of ocular tissues. *Exp. Eye Res.*, **111**, 105–111.
13. Yoshida, S., Yashar, B.M., Hiriyanna, S. and Swaroop, A. (2002) Microarray analysis of gene expression in the aging human retina. *Invest. Ophthalmol. Vis. Sci.*, **43**, 2554–2560.
14. Bowes Rickman, C., Ebright, J.N., Zavodni, Z.J., Yu, L., Wang, T., Daiger, S.P., Wistow, G., Boon, K. and Hauser, M.A. (2006) Defining the human macula transcriptome and candidate retinal disease genes using EyeSAGE. *Invest. Ophthalmol. Vis. Sci.*, **47**, 2305–2316.
15. Fritsche, L.G., Chen, W., Schu, M., Yaspan, B.L., Yu, Y., Thorleifsson, G., Zack, D.J., Arakawa, S., Cipriani, V., Ripke, S. *et al.* (2013) Seven new loci associated with age-related macular degeneration. *Nat. Genet.*, **45**, 433–439. 439e1-2.
16. Wu, T.D. and Watanabe, C.K. (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*, **21**, 1859–1875.
17. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **28**, 511–515.
18. Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L. and Pachter, L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.*, **31**, 46–53.
19. Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
20. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106-2010-11-10-r106. Epub 2010 Oct 27.
21. Dennis, G. Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C. and Lempicki, R.A. (2003) DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.*, **4**, P3.
22. Vogt, S.D., Curcio, C.A., Wang, L., Li, C.M., McGwin, G. Jr., Medeiros, N.E., Philp, N.J., Kimble, J.A. and Read, R.W. (2011) Retinal pigment epithelial expression of complement regulator CD46 is altered early in the course of geographic atrophy. *Exp. Eye Res.*, **93**, 413–423.
23. Kamei, M. and Hollyfield, J.G. (1999) TIMP-3 in bruch's membrane: changes during aging and in age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.*, **40**, 2367–2375.
24. Ruiz, A., Brett, P. and Bok, D. (1996) TIMP-3 is expressed in the human retinal pigment epithelium. *Biochem. Biophys. Res. Commun.*, **226**, 467–474.
25. Zweifel, S.A., Spaide, R.F., Curcio, C.A., Malek, G. and Imamura, Y. (2010) Reticular pseudodrusen are subretinal drusenoid deposits. *Ophthalmology*, **117**, 303–312. e1.
26. Freund, K.B., Zweifel, S.A. and Engelbert, M. (2010) Do we need a new classification for choroidal neovascularization in age-related macular degeneration? *Retina*, **30**, 1333–1349.
27. Schaub, M.A., Boyle, A.P., Kundaje, A., Batzoglou, S. and Snyder, M. (2012) Linking disease associations with regulatory information in the human genome. *Genome Res.*, **22**, 1748–1759.
28. Shen, S., Park, J.W., Huang, J., Dittmar, K.A., Lu, Z.X., Zhou, Q., Carstens, R.P. and Xing, Y. (2012) MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-seq data. *Nucleic Acids Res.*, **40**, e61.
29. Wiggs, J.L., Yaspan, B.L., Hauser, M.A., Kang, J.H., Allingham, R.R., Olson, L.M., Abdrabou, W., Fan, B.J., Wang, D.Y., Brodeur, W. *et al.* (2012) Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. *PLoS Genet.*, **8**, e1002654.
30. Verhoeven, V.J., Hysi, P.G., Wojciechowski, R., Fan, Q., Guggenheim, J.A., Hohn, R., MacGregor, S., Hewitt, A.W., Nag, A., Cheng, C.Y. *et al.* (2013) Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat. Genet.*, **45**, 314–318.
31. Vogt, S.D., Barnum, S.R., Curcio, C.A. and Read, R.W. (2006) Distribution of complement anaphylatoxin receptors and membrane-bound regulators in normal human retina. *Exp. Eye Res.*, **83**, 834–840.
32. Kirwan, R.P., Felice, L., Clark, A.F., O'Brien, C.J. and Leonard, M.O. (2012) Hypoxia regulated gene transcription in human optic nerve lamina cribrosa cells in culture. *Invest. Ophthalmol. Vis. Sci.*, **53**, 2243–2255.
33. DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W. and Getz, G. (2012) RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*, **28**, 1530–1532.
34. Lutty, G.A., Merges, C., Threlkeld, A.B., Crone, S. and McLeod, D.S. (1993) Heterogeneity in localization of isoforms of TGF-beta in human retina, vitreous, and choroid. *Invest. Ophthalmol. Vis. Sci.*, **34**, 477–487.
35. Nickla, D.L. and Wallman, J. (2010) The multifunctional choroid. *Prog. Retin. Eye Res.*, **29**, 144–168.
36. Rodieck, R.W. (1999) *The First Steps in Seeing*. Sinauer Associates, Sunderland, MA.
37. Li, C.M., Presley, J.B., Zhang, X., Dashti, N., Chung, B.H., Medeiros, N.E., Guidry, C. and Curcio, C.A. (2005) Retina expresses microsomal triglyceride transfer protein: Implications for age-related maculopathy. *J. Lipid Res.*, **46**, 628–640.
38. Hirano, A.A., Brandstatter, J.H., Morgans, C.W. and Brecha, N.C. (2011) SNAP25 expression in mammalian retinal horizontal cells. *J. Compar. Neurol.*, **519**, 972–988.
39. Neumann, S. and Haverkamp, S. (2013) Characterization of small-field bistratified amacrine cells in macaque retina labeled by antibodies against synaptotagmin-2. *J. Compar. Neurol.*, **521**, 709–724.
40. Nishikawa, S. and Tamai, M. (2001) Müller cells in the human foveal region. *Curr Eye Res*, **22**, 34–41.
41. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
42. Chung, T.S. and Lung, F.W. (2012) Different impacts of aquaporin 4 and MAOA allele variation among olanzapine, risperidone, and paliperidone in schizophrenia. *J. Clin. Psychopharmacol.*, **32**, 394–397.
43. Li, J., Patil, R.V. and Verkman, A.S. (2002) Mildly abnormal retinal function in transgenic mice without muller cell aquaporin-4 water channels. *Invest. Ophthalmol. Vis. Sci.*, **43**, 573–579.

44. Berisha, F., Feke, G.T., Trempe, C.L., McMeel, J.W. and Schepens, C.L. (2007) Retinal abnormalities in early Alzheimer's disease. *Invest. Ophthalmol. Vis. Sci.*, **48**, 2285–2289.
45. Katz, B., Rimmer, S., Iragui, V. and Katzman, R. (1989) Abnormal pattern electroretinogram in Alzheimer's disease: evidence for retinal ganglion cell degeneration? *Ann. Neurol.*, **26**, 221–225.
46. Katz, B. and Rimmer, S. (1989) Ophthalmologic manifestations of Alzheimer's disease. *Surv. Ophthalmol.*, **34**, 31–43.
47. Moschos, M.M., Markopoulos, I., Chatziralli, I., Rouvas, A., Papageorgiou, S.G., Ladas, I. and Vassilopoulos, D. (2012) Structural and functional impairment of the retina and optic nerve in Alzheimer's disease. *Curr. Alzheimer Res.*, **9**, 782–788.
48. Wang, L., Clark, M.E., Crossman, D.K., Kojima, K., Messinger, J.D., Mobley, J.A. and Curcio, C.A. (2010) Abundant lipid and protein components of drusen. *PLoS One*, **5**, e10329.
49. Mrejen, S. and Spaide, R.F. (2013) Optical coherence tomography: Imaging of the choroid and beyond. *Surv. Ophthalmol.*, **58**, 387–429.
50. Zheng, W., Reem, R.E., Omarova, S., Huang, S., DiPatre, P.L., Charvet, C.D., Curcio, C.A. and Pikuleva, I.A. (2012) Spatial distribution of the pathways of cholesterol homeostasis in human retina. *PLoS One*, **7**, e37926.
51. Oak, A.S.W., Messinger, J.D. and Curcio, C.A. (2014) Subretinal drusenoid deposits: further characterization by lipid histochemistry. *Retina*. Feb. 28. PMID: 24589874.
52. Curcio, C.A., Messinger, J.D., Sloan, K.R., McGwin, G., Medeiros, N.E. and Spaide, R.F. (2013) Subretinal drusenoid deposits in non-neovascular age-related macular degeneration: morphology, prevalence, topography, and biogenesis model. *Retina*, **33**, 265–276.
53. Curcio, C.A., Sloan, K.R., Kalina, R.E. and Hendrickson, A.E. (1990) Human photoreceptor topography. *J. Comp. Neurol.*, **292**, 497–523.
54. Curcio, C.A. (2001) Photoreceptor topography in ageing and age-related maculopathy. *Eye (Lond)*, **15**, 376–383.
55. Curcio, C.A., Millican, C.L., Bailey, T. and Kruth, H.S. (2001) Accumulation of cholesterol with age in human bruch's membrane. *Invest. Ophthalmol. Vis. Sci.*, **42**, 265–274.
56. Polyak, S.L. (1941) *The Retina*. The University of Chicago Press, Chicago.
57. Drasdo, N., Millican, C.L., Katholi, C.R. and Curcio, C.A. (2007) The length of henle fibers in the human retina and a model of ganglion receptive field density in the visual field. *Vision Res.*, **47**, 2901–2911.
58. Curcio, C.A., Messinger, J.D., Sloan, K.R., Mitra, A., McGwin, G. and Spaide, R.F. (2011) Human chorioretinal layer thicknesses measured in macula-wide, high-resolution histologic sections. *Invest. Ophthalmol. Vis. Sci.*, **52**, 3943–3954.
59. Curcio, C.A. and Allen, K.A. (1990) Topography of ganglion cells in human retina. *J. Comp. Neurol.*, **300**, 5–25.
60. Dacey, D.M. (1993) The mosaic of midget ganglion cells in the human retina. *J. Neurosci.*, **13**, 5334–5355.
61. Spraul, C.W., Lang, G.E. and Grossniklaus, H.E. (1996) Morphometric analysis of the choroid, bruch's membrane, and retinal pigment epithelium in eyes with age-related macular degeneration. *Invest. Ophthalmol. Vis. Sci.*, **37**, 2724–2735.
62. Weiter, J.J., Delori, F.C., Wing, G.L. and Fitch, K.A. (1986) Retinal pigment epithelial lipofuscin and melanin and choroidal melanin in human eyes. *Invest. Ophthalmol. Vis. Sci.*, **27**, 145–152.
63. Malik, K.J., Chen, C.D. and Olsen, T.W. (2003) Stability of RNA from the retina and retinal pigment epithelium in a porcine model simulating human eye bank conditions. *Invest. Ophthalmol. Vis. Sci.*, **44**, 2730–2735.
64. Barondes, M., Pauleikhoff, D., Chisholm, I.C., Minassian, D. and Bird, A.C. (1990) Bilaterality of drusen. *Br. J. Ophthalmol.*, **74**, 180–182.
65. Bellmann, C., Jorzik, J., Spital, G., Unnebrink, K., Pauleikhoff, D. and Holz, F.G. (2002) Symmetry of bilateral lesions in geographic atrophy in patients with age-related macular degeneration. *Arch. Ophthalmol.*, **120**, 579–584.
66. Fleckenstein, M., Adrion, C., Schmitz-Valckenberg, S., Gobel, A.P., Bindewald-Wittich, A., Scholl, H.P., Mansmann, U. and Holz, F.G. and FAM Study Group. (2010) Concordance of disease progression in bilateral geographic atrophy due to AMD. *Invest. Ophthalmol. Vis. Sci.*, **51**, 637–642.
67. Wang, J.J., Mitchell, P., Smith, W. and Cumming, R.G. (1998) Bilateral involvement by age related maculopathy lesions in a population. *Br. J. Ophthalmol.*, **82**, 743–747.
68. Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T. *et al.* (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.*, **26**, 317–325.