Pathway Analysis Integrating Genome-Wide and Functional Data Identifies \textit{PLCG2} as a Candidate Gene for Age-Related Macular Degeneration

Andrea R. Waksmunski,1–3 Michelle Grunin,2,3 Tyler G. Kinzy,3 Robert P. Igo Jr,3 Jonathan L. Haines,1–3 and Jessica N. Cooke Bailey2,3; for the International Age-Related Macular Degeneration Genomics Consortium

1Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland, Ohio, United States
2Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, Ohio, United States
3Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, United States

Correspondence: Jessica N. Cooke Bailey, Department of Population and Quantitative Health Sciences, Case Western Reserve University, 2103 Cornell Road, Suite 1313, Cleveland, OH 44106-7281, USA; jlb213@case.edu.

JLH and JNCB are joint senior authors.

See the appendix for members of the International Age-related Macular Degeneration Genomics Consortium.

Submitted: June 26, 2019
Accepted: August 14, 2019

\textbf{Purpose.} Age-related macular degeneration (AMD) is the worldwide leading cause of blindness among the elderly. Although genome-wide association studies (GWAS) have identified AMD risk variants, their roles in disease etiology are not well-characterized, and they only explain a portion of AMD heritability.

\textbf{Methods.} We performed pathway analyses using summary statistics from the International AMD Genomics Consortium’s 2016 GWAS and multiple pathway databases to identify biological pathways wherein genetic association signals for AMD may be aggregating. We determined which genes contributed most to significant pathway signals across the databases. We characterized these genes by constructing protein-protein interaction networks and performing motif analysis.

\textbf{Results.} We determined that eight genes (\textit{C2, C3, LIPC, MICA, NOTCH4, PLCG2, PPARA,} and \textit{RAD51B}) “drive” the statistical signals observed across pathways curated in the Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and Gene Ontology (GO) databases. We further refined our definition of statistical driver gene to identify \textit{PLCG2} as a candidate gene for AMD due to its significant gene-level signals (\textit{P < 0.0001}) across KEGG, Reactome, GO, and NetPath pathways.

\textbf{Conclusions.} We performed pathway analyses on the largest available collection of advanced AMD cases and controls in the world. Eight genes strongly contributed to significant pathways from the three larger databases, and one gene (\textit{PLCG2}) was central to significant pathways from all four databases. This is, to our knowledge, the first study to identify \textit{PLCG2} as a candidate gene for AMD based solely on genetic burden. Our findings reinforce the utility of integrating in silico genetic and biological pathway data to investigate the genetic architecture of AMD.

Keywords: age-related macular degeneration, pathway analysis, genome-wide association study, database, phospholipase C gamma 2

Vision loss is one of the most feared medical conditions because of its profound effect on day-to-day quality of life.\textsuperscript{1,2} Age-related macular degeneration (AMD) is the most common cause of blindness in individuals over age 60 and is responsible for almost 10% of all cases of blindness in the world.\textsuperscript{3} AMD is a late-onset disease that results from the accumulation of drusen, inflammation, and photoreceptor loss in the macular region of the eye.\textsuperscript{3} This progressive disease is categorized as either early/intermediate or advanced AMD; the latter is further subclassified as geographic atrophy (dry AMD [GA]) or choroidal neovascularization (wet AMD [CNV]).\textsuperscript{3} Early AMD is often asymptomatic and dry AMD is initially asymptomatic, but as the disease progresses, patients’ central vision begins to blur and diminish.\textsuperscript{3} Wet AMD is characterized by the growth of abnormal blood vessels in the macula, which ultimately results in severe vision loss.\textsuperscript{3}

Although both genetic and environmental factors shape AMD susceptibility, between 46% and 71% of the phenotypic variance of the disease is attributable to genetic factors.\textsuperscript{4} To understand the genetic architecture of AMD, the International Age-Related Macular Degeneration Genomics Consortium (IAMDGC) performed a large-scale genome-wide association study (GWAS) for advanced AMD cases and controls. They identified 52 independent genetic variants across 34 susceptibility loci for advanced AMD that are estimated to explain nearly two thirds of AMD heritability.\textsuperscript{5} Therefore, about one third of AMD heritability is still unexplained by the known loci. Although other studies have identified additional risk loci with modest effect for advanced AMD, more comprehensive approaches beyond GWAS must be used to find the remaining heritable variation for AMD.

Rather than investigating associations between single genetic variants and a phenotype, pathway analysis of GWAS data interrogates alterations in biological pathways for a trait of interest. Generally, this is done by aggregating summary statistics for these variants into genes, which are then grouped...
into pathways based on data in curated pathway databases.\textsuperscript{8} We hypothesize that applying this more comprehensive approach may help elucidate the genetic etiology of advanced AMD that has been indiscernible from GWAS. In this study, we performed in silico pathway analysis using the Pathway Analysis by Randomization Incorporating Structure (PARIS) software to identify biological pathways and processes enriched in genetic variation potentially associated with AMD in individuals of European descent. Because nomenclature, foci, and definitions vary across pathway databases,\textsuperscript{9} we utilized multiple databases to complement and validate our findings. Additionally, we sought to determine the central causal genes that “drive” the statistical signals observed for significant pathways identified by PARIS.

**METHODS**

**Study Subjects and GWAS Summary Statistics**

The participants for this study were previously ascertained by cohorts in the IAMDGC as described.\textsuperscript{5} This included 16,144 individuals with advanced AMD and 17,832 unaffected individuals. Of the advanced AMD cases, 3,235 individuals have GA only and 10,749 have CNV only. The remaining cases have both GA and CNV. All of the cases and controls used for our analyses were of European ancestry. All participants provided informed consent, and the study protocol was approved by institutional review boards as previously described.\textsuperscript{5} Data were previously collected in accordance with the tenets of the Declaration of Helsinki. The summary statistics we analyzed in this study were obtained in the 2016 GWAS performed by the IAMDGC.\textsuperscript{5} Specifically, these data include $P$ values for 445,115 directly genotyped common and rare variants from the advanced AMD case-control results. The genotypes for these variants were generated from an array (HumanCoreExome; Illumina, San Diego, CA, USA) that was designed with additional genome-wide and custom content for AMD.\textsuperscript{5}

**PARIS: Knowledge-Driven Pathway Analysis of GWAS Data**

To identify biological pathways enriched in genetic variants possibly contributing to advanced AMD risk, we performed in silico pathway analysis using the PARIS v2.4 software.\textsuperscript{10,11} PARIS uses variant summary statistics from GWAS, clusters them into features defined by the linkage disequilibrium (LD) structure of the genome based on a reference catalog of common genetic variants, and assigns significance to pathways based on permutation of the genome.\textsuperscript{10,11} In our analyses, we performed 100,000 permutations. PARIS also assigns empirical $P$ values to the genes composing a pathway based on permutation testing of features within each of the genes.\textsuperscript{10,11}

We performed PARIS using multiple pathway databases, including Kyoto Encyclopedia of Genes and Genomes (KEGG),\textsuperscript{12} Reactome,\textsuperscript{13} Gene Ontology (GO),\textsuperscript{14} and NetPath.\textsuperscript{15} KEGG, Reactome, and GO databases are extensive, curated biological pathway data repositories. NetPath is a specialized database that covers signaling pathways. Pathways with a $P$ value less than 0.0001 were prioritized for further investigation. This permutation $P$ value was calculated using the following equation: $P = (1 + b)/(1 + M)$, where $M$ is the number of permutations and $b$ is the number of randomly sampled permutation scores that are greater than the observed score. To determine if the pathway associations we observed were driven by known AMD loci, we reperformed our pathway analyses excluding variants from the 34 susceptibility loci identified by the IAMDGC (defined by the 52 genomic variants) and their proxies ($r^2 \geq 0.5$) within 500 kb.\textsuperscript{3}

**Identification of Statistical Pathway Driver Genes**

Due to disparate nomenclature and composition of pathways in the databases, we identified genes that overlapped across significant pathways within a database and across databases (regardless of pathway). This served to internally validate and complement our results. To interrogate the significant signals obtained from the pathways identified by PARIS, we queried which significant ($P < 0.0001$) genes overlapped among the significant ($P < 0.0001$) pathways within a pathway database. These genes were compared across the analyses done with each of the pathway databases (KEGG, Reactome, GO, and NetPath) to find statistical driver genes that had significant signals across three or more databases for the advanced AMD results.

**Protein-Protein Interaction (PPI) Network for Statistical Pathway Driver Genes**

We searched the Search Tool for Recurring Instances of Neighbouring Genes (STRING) database\textsuperscript{16} version 10.5 for PPIs involving the proteins encoded by the genes identified as statistical driver genes. The STRING database is composed of known and predicted PPIs based on data from curated interactions databases, high-throughput lab experiments, coexpression, and text mining in the literature. We used the high confidence (0.700) minimum required interaction score to construct the protein-protein networks of interactions based on experimental data, database entries, and coexpression.

**Motif Analysis for Statistical Pathway Driver Genes**

We extracted reference genome sequences for the statistical driver genes using the UCSC Genome Table Browser.\textsuperscript{17} We included 600 nucleotides upstream from the first exon and the 5’ untranslated region (UTR) in the sequences for each gene. To identify potential sequence motifs for each of these gene sets, we utilized the Multiple Expectation Maximization (EM) for Motif Elucidation (MEME) software suite.\textsuperscript{18} Sequences were considered motifs if their lengths were between 6 and 50 nucleotides. MEME was not required to find a motif in every sequence, but motifs were required to have an E-value of 0.0001. Each motif from the gene sets was then investigated in Tomtom, which looks for transcription factors (TFs) that are associated with the motif. TF binding motifs were evaluated based on the known human TF database from JASPAR\textsuperscript{19} using HOCOMOCO.\textsuperscript{20} To validate the motifs found and to test the null hypothesis of random motifs found unrelated to the statistical driver genes, 10 permutations were run on a random gene set generator for eight genes and performed the same analyses via MEME and Tomtom. We removed motifs and TFs that appeared in both the random and actual gene sets from further analysis.

**RESULTS**

**In Silico Pathway Analysis**

We identified several biological pathways and processes from KEGG, Reactome, GO, and NetPath databases (Table 1; Supplementary Tables S1–S4) to be significantly associated with advanced AMD using PARIS. A pathway was considered significant if it had a pathway-level $P$ value less than 0.0001. The vast majority of pathways in the four databases were not
TABLE 1. Significantly Associated Pathways Across Multiple Pathway Databases for Advanced AMD

<table>
<thead>
<tr>
<th>Database</th>
<th>Count of Significant Pathways</th>
<th>Total Entries in Database</th>
<th>Proportion of Significant Pathways in Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>NetPath</td>
<td>1</td>
<td>26</td>
<td>0.038</td>
</tr>
<tr>
<td>KEGG</td>
<td>25</td>
<td>293</td>
<td>0.085</td>
</tr>
<tr>
<td>Reactome</td>
<td>50</td>
<td>1,748</td>
<td>0.029</td>
</tr>
<tr>
<td>GO</td>
<td>145</td>
<td>12,765</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Pathways were considered significant if they obtained an empirical \( P < 0.0001 \).

TABLE 2. Eight Statistical Pathway Driver Genes From Significant KEGG, Reactome, and GO Pathways

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Full Gene Name (HGNC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLCG2</td>
<td>C2</td>
<td>Complement C2</td>
</tr>
<tr>
<td>MICA</td>
<td>6</td>
<td>MHC class I polypeptide-related</td>
</tr>
<tr>
<td>NOTCH4</td>
<td>6</td>
<td>Notch receptor 4</td>
</tr>
<tr>
<td>RAD51B</td>
<td>14</td>
<td>RAD51 paralog B</td>
</tr>
<tr>
<td>LIPC</td>
<td>15</td>
<td>Lipase C, hepatic type</td>
</tr>
<tr>
<td>C3</td>
<td>19</td>
<td>Complement C3</td>
</tr>
<tr>
<td>PPARA</td>
<td>22</td>
<td>Peroxisome proliferator activated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>receptor alpha</td>
</tr>
</tbody>
</table>

The cross-database comparison of significant genes from significantly associated pathways.

FIGURE 1. Comparison of significant genes from AMD-associated KEGG, Reactome, and GO pathways identified by PARIS. Eight genes demonstrated significant signals across all three comparisons and are summarized in Table 2.

Statistical Driver Genes Among Advanced AMD-Associated Pathways

Because pathway structure and terminology vary across databases, we determined which genes were significantly contributing to the overall pathway signals detected by PARIS. We compared the significant genes in significant pathways from KEGG, Reactome, and GO (Fig. 1; Table 2) and identified eight such genes. Upon removing variants from our analyses that fell within the 34 known AMD susceptibility loci as defined in Supplementary Table S5 in the IAMDGC GWAS,5 we found that two genes (PPARA and PLCG2) remained statistical driver genes across associated pathways from KEGG, Reactome, and GO.

To identify evidence of PPI for the proteins encoded by the eight statistical driver genes in our analyses (C2, C3, LIPC, MICA, NOTCH4, PPARA, PLCG2, and RAD51B), we queried the STRING database. Each of these proteins have multiple binding interaction partners identified through functional studies or in silico predictions (Fig. 2). When considering no more than 50 interaction partners for each of the eight proteins, we found three distinct clusters of PPIs (Fig. 2). One cluster connects MICA, PLCG2, LIPC, C2, C3, and other immune-related proteins (Fig. 2A); another connects NOTCH4, PPARA, and other signaling proteins (Fig. 2B); and the third contains RAD51B and other DNA repair proteins (Fig. 2C).

Using the MEME software suite, we identified sequence motifs with known TF binding sites near the eight statistical driver gene sequences from the UCSC Genome Table Browser.17 Five motifs were present for most of the statistical driver genes and contain binding sites for TFs (Table 3). Only one sequence motif ([GCA][AC][CT][AG][AT][G][CA][T][A]A[AT][A][G][AT][CA][T][G][G][GA][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][TG][T...
Susceptibility loci described earlier. This suggests that modest effects aggregating in these pathways may contribute to the missing heritability of AMD. Although the Wnt pathway from NetPath was no longer significant in our sensitivity analysis, the Wnt signaling pathway from GO remained associated with AMD. This results from the difference in the pathway definitions. These pathways are nearly identical in size ($n = 45$ and 41 genes for NetPath and GO, respectively); however, only two genes overlap between them (PLCG2 and FZD4).

Furthermore, the Wnt signaling pathway in KEGG ($n = 140$ genes) and the signaling by Wnt pathway in Reactome ($n = 294$ genes) only achieved pathway-level $P$ values of 0.032 and 0.037 in our analyses, respectively. These pathway definition differences further justify our use of multiple curated databases in our analyses to uncover AMD-associated pathways and genes driving their statistical significance.

Due to varying nomenclature for pathways across databases and as a way of internal validation, we focused on eight statistical driver genes ($C2$, $C3$, $LIPC$, $MICA$, $NOTCH4$, $PPARA$, $PLCG2$, and $RAD51B$) that were consistently significant across GO, Reactome, and KEGG pathways. $PPARA$ and $PLCG2$ were not previously identified as a part of the 34 IAMDGC loci associated with AMD risk. The strongest single-marker $P$ values observed in $PLCG2$ and $PPARA$ were $2.05 \times 10^{-19}$ and $3.10 \times 10^{-9}$.
respectively, and do not meet the classical GWAS significance levels. In our sensitivity analysis, PPARA and PLCG2 remained statistical driver genes in pathways from KEGG, Reactome, and GO, suggesting that pathway analysis can identify novel AMD genes. Additionally, the aggregation of nominally significant independent variants in PLCG2 suggests that the gene-wide significance of PLCG2 is greater than that of the individual variants and emphasizes the power of pathway analysis for identifying gene-wide signals rather than single-variant associations.

DNA motif analysis identified five sequence motifs adjacent to the eight statistical driver genes in their promoter regions. These motifs represent sites of known TF binding and suggest that the expression of these genes may be controlled by similar mechanisms. One motif ([GCA][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]) was adjacent to the start positions of all eight statistical driver genes and contains known binding sites of several TFs (Table 3). Functional studies are required to confirm these in silico findings and elucidate the transcriptional mechanisms of these statistical driver genes in the context of AMD.

Table 3. Sequence Motifs With TF Binding Sites Near Statistical Driver Genes

<table>
<thead>
<tr>
<th>Motif Consensus Sequence</th>
<th>TF</th>
<th>P Value</th>
<th>Statistical Driver Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>KLF5</td>
<td>0.0095</td>
<td>C2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>KLF12</td>
<td>0.011</td>
<td>LIPC</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>THAI</td>
<td>0.012</td>
<td>MICA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZNF63</td>
<td>0.013</td>
<td>NOTCH4</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>IRF2</td>
<td>0.013</td>
<td>PPARA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>NFIA</td>
<td>0.017</td>
<td>RAD51B</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZNF449</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ELF2</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZBTB6</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>RARG</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>PTP1</td>
<td>0.0052</td>
<td>C2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>SOX5</td>
<td>0.0093</td>
<td>LIPC</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>AIRE</td>
<td>0.010</td>
<td>MICA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>CEBPE</td>
<td>0.011</td>
<td>NOTCH4</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>MAFB</td>
<td>0.0089</td>
<td>C2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>MAFF</td>
<td>0.010</td>
<td>LIPC</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>HTF4</td>
<td>0.011</td>
<td>MICA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>MAIK</td>
<td>0.012</td>
<td>NOTCH4</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>FOXA2</td>
<td>0.012</td>
<td>PPARA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>TFE2</td>
<td>0.014</td>
<td>RAD51B</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>BACH2</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>TFAP4</td>
<td>0.0047</td>
<td>C2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZNF322</td>
<td>0.0062</td>
<td>LIPC</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZNF41</td>
<td>0.011</td>
<td>MICA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>CRX</td>
<td>0.013</td>
<td>NOTCH4</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZIC3</td>
<td>0.015</td>
<td>PPARA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>NKX21</td>
<td>0.020</td>
<td>RAD51B</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>GLI3</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>HEN1</td>
<td>0.0025</td>
<td>C2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ZSC31</td>
<td>0.0029</td>
<td>C3</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>PKNX1</td>
<td>0.0034</td>
<td>LIPC</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>NKX21</td>
<td>0.0037</td>
<td>MICA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>PBRX3</td>
<td>0.0066</td>
<td>NOTCH4</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>TYY1</td>
<td>0.010</td>
<td>PPARA</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>NR2C1</td>
<td>0.011</td>
<td>PLCG2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>VDR</td>
<td>0.014</td>
<td>PLCG2</td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>CREB1</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>RFX2</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>ATF1</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>G[TG][AC][CT][AT][GA][AT][TA][CG][TA][CT][CT][TG][CG]</td>
<td>CEBPE</td>
<td>0.022</td>
<td></td>
</tr>
</tbody>
</table>

For each motif, we identified TFs associated with the motif sequence using ToToMot. The P value represents the strength of the match between the sequence motif identified adjacent to the statistical driver genes and the curated sequences of the TF binding motifs in the HOCOMOCO database.

10^-5, respectively, and do not meet the classical GWAS significance levels. In our sensitivity analysis, PPARA and PLCG2 remained statistical driver genes in pathways from KEGG, Reactome, and GO, suggesting that pathway analysis can identify novel AMD genes. Additionally, the aggregation of nominally significant independent variants in PLCG2 suggests that the gene-wide significance of PLCG2 is greater than that of the individual variants and emphasizes the power of pathway analysis for identifying gene-wide signals rather than single-variant associations.

One gene, PLCG2, was central to multiple pathways in all four databases and remained significant after our sensitivity analysis. PLCG2 encodes a signaling enzyme (phospholipase C gamma 2, PLCG2) that utilizes calcium to catalyze the hydrolysis of PIP2 into second messengers IP3 and DAG.21 These molecules initiate intracellular calcium flux and activate protein kinase C, respectively.21 The enzymatic activity of PLCG2 results from tyrosine phosphorylation performed by growth factor receptors, immune receptors, and G protein-coupled receptors as well as the activity of lipid-derived second messengers in the cell.21 This enzyme is highly expressed in cells of hematopoietic origin and is responsible...
for regulating immune responses and platelet adhesion and spreading.22–26

The PLCG2 protein interacts with several members (HCK, LYN, PIK3R1, and SYK) of the microglia pathogen phagocytosis pathway in humans.27 Its interaction partners also play roles in oxidative stress, angiogenesis, and platelet activation. BLNK and BTK are central to facilitating B-cell apoptosis following oxidative stress.28,29 Exposure to oxidative stress activates EGFR, which promotes retinal epithelial cell health and survival through EGFR/Akt, PI3K, and ERK/MAPK signaling pathways.30,31 EGFR downstream signaling also contributes to retinal pigment epithelial cell proliferation and migration in wound healing.32,33 PIK3R1 is a regulatory subunit of PI3K in the PI3K/Akt/mTOR pathway, which is a possible target for treating ocular neovascularization.34 PI3K and Tec protein kinases regulate platelet activation,35 and signaling cascades from LCP2 (also called SLP-76) and SYK are responsible for separating blood and lymphatic vasculatures in the human body.36 These interactions and processes, coupled with PLCG2’s role in the VEGF pathway,37–38 could be pertinent for understanding the role of PLCG2 and its interaction partners in the choroidal neovascularization subtype of advanced AMD. In the CNV-only case-control GWAS performed by the IAMDGC, no PLCG2 variants were genome-wide significant; however, 13 variants were nominally associated with CNV ($P < 0.05$).5 Of the 65 PLCG2 variants analyzed by PARIS, 31 exhibited lower $P$ values in the CNV-specific IAMDGC GWAS than in the combined advanced AMD IAMDGC GWAS. Heterozygous gain-of-function mutations in PLCG2 result in constitutive phospholipase activity and PLCG2-associated antibody deficiency and immune dysregulation, which is characterized by immunodeficiency and autoimmunity.39 This gene was recently identified as a candidate gene for rheumatoid arthritis (RA) due to its overexpression in RA patients compared to controls.40 Genetic risk scores for RA are associated with increased AMD risk,14 and individuals with RA are at a higher risk of developing AMD.15 PLCG2 is also highly expressed in microglia43 and has been previously implicated in the genetic etiology of late-onset Alzheimer’s disease.

**Figure 3.** Identification of **PLCG2** as a candidate gene for advanced AMD. A comparison of the significant genes from significant KEGG, GO, NetPath, and Reactome pathways in our PARIS pathway analysis converged on one gene (**PLCG2**), which encodes a protein that is common to several pathways.

<table>
<thead>
<tr>
<th>KEGG</th>
<th>GO</th>
<th>NetPath</th>
<th>Reactome</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fc epsilon RI</td>
<td>• B cell receptor</td>
<td>• Wnt</td>
<td></td>
</tr>
<tr>
<td>signaling pathway</td>
<td>signaling pathway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Natural killer cell</td>
<td>• Blood coagulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mediated cytoxicity</td>
<td>• Cytosol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Metabolic pathways</td>
<td>• Extracellular vesicular</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• MicroRNAs in cancer</td>
<td>exosome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Epstein-Barr virus</td>
<td>• Innate immune response</td>
<td></td>
<td></td>
</tr>
<tr>
<td>infection</td>
<td>• Inositol phosphate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Epithelial cell</td>
<td>metabolic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>signaling in</td>
<td>• Protein binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>• Wnt signaling pathway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>infection</td>
<td></td>
<td></td>
<td>• PLC-mediated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hydrolysis of PIP2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Metabolism</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Inositol phosphate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>metabolicism</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Immune System</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Innate Immune System</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Adaptive Immune System</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Dectin-2 family</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Hemostasis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Platelet activation, signaling and aggregation</td>
</tr>
</tbody>
</table>

**Pathway Analysis Implicates PLCG2 in AMD**

IOVS | September 2019 | Vol. 60 | No. 12 | 4046

Downloaded from tvst.arvojournals.org on 10/01/2019
Specifically, GWAS identified a protective effect for a rare variant in the coding region of \textit{PLCG2} on LOAD.\textsuperscript{44,45} This variant is considered hypermorphic because the mutant enzyme experiences a small increase in enzymatic activity compared to wild-type enzyme, which would imply that mildly activating \textit{PLCG2} could be a therapeutic intervention for LOAD.\textsuperscript{45} Functional studies would need to be performed to determine if \textit{PLCG2}'s enzymatic activity could be modulated by a similar mechanism in patients with AMD.

Although \textit{PLCG2} has not been previously associated with AMD in a case-control GWAS, variants in this gene were associated with AMD when accounting for birth control pill usage in women with CNV.\textsuperscript{46} These associations were undetectable when gene-environment interactions between \textit{PLCG2} variants and exogenous estrogen exposure were not considered.\textsuperscript{46} Other interaction studies have identified \textit{PLCG2} variants as genetic modifiers of previously identified associations among menopausal hormone therapy, mammographic density, and breast cancer risk, which could suggest sex-specific effects of genetic variants in this gene for disease risk.\textsuperscript{47,48}

While our study provides in silico evidence for the roles of these statistical driver genes and pathways in AMD, it does not biologically confirm them. Functional studies are required to determine causality for these genes and pathways in patients with AMD. Knowledge-driven pathway analyses are subject to the quality and coverage of the knowledge in a given database. We attempted to circumvent this limitation by utilizing multiple databases in our analyses and integrating our results. The GWAS data used in this study were generated from individuals of European descent. Consequently, these findings may not be applicable to non-European populations. The IAMDGC GWAS dataset is considered the largest available dataset for advanced AMD cases and controls in the world. We are unaware of any comparable datasets available for replication.
Acknowledgments

The authors thank the study participants for their contribution to this study. This work made use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University.

Supported by Grants 1X01HG006934-01, R01 EY022310, the National Institutes of Health National Center for Advancing Translational Science (NCATS) Grant KL2TR004440 (JNCB), and the VSTP Training Grant T32EY001757-18 (MG).

Disclosure: A.R. Waksmanusk, None; M. Grunin, None; T.G. Kinzy, None; R.P. Igo Jr, None; J.L. Haines, None; J.N. Cooke, None.

References

Pathway Analysis Implicates PLCG2 in AMD

10Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, United States
11Department of Ophthalmology, Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
12Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
13Department of Neuroscience-Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
14Institute of Genetic Medicine:Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
15Institut de la Vision, Université Pierre et Marie Curie, Paris, France
16Hôpital Intercommunal de Créteil, Hôpital Henri Mondor-Université Paris Est Créteil, France
17University of Bonn, Department of Ophthalmology, Bonn, Germany
18Louis Stokes Cleveland VA Medical Center, Cleveland, Ohio, United States
19Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, Wisconsin, United States
20Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States
21Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, United States
22John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, Florida, United States
23Centre for Vision Research, Department of Ophthalmology and Westmead Millennium Institute for Medical Research, University of Sydney, Sydney, Australia
24Department of Ophthalmology Columbia University, New York, New York, United States
25UCL Institute of Ophthalmology, University College London, London, United Kingdom
26Moorfields Eye Hospital, London, United Kingdom
27Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee, United States
28University Hospital of Cologne, Department of Ophthalmology, Cologne, Germany
29Department of Ophthalmology, Radboud University Medical Centre, Nijmegen, the Netherlands
30Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye and Ear Hospital, East Melbourne, Victoria, Australia
31South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Brownsville, Texas, United States
32Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, United States
33MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland, United Kingdom
34Department of Ophthalmology, Erasmus Medical Center, Rotterdam, the Netherlands
35Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands
36Quantitative Biomedical Research Center, Department of Clinical Science, University of Texas Southwestern Medical Center, Dallas, Texas, United States
37Center for the Genetics of Host Defense, University of Texas Southwestern Medical Center, Dallas, Texas, United States
38Neurobiology Neurodegeneration & Repair Laboratory (N-NRL), National Eye Institute, National Institutes of Health, Bethesda, Maryland, United States
39Department of Ophthalmology and Visual Sciences, University of Michigan, Kellogg Eye Center, Ann Arbor, Michigan, United States
40Department of Ophthalmology, Flinders Medical Centre, Flinders University, Adelaide, South Australia, Australia
41Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western Australia, Perth, Western Australia, Australia
42Sichuan Provincial Key Laboratory for Human Disease Gene Study, Hospital of the University of Electronic Science and Technology of China and Sichuan Provincial People’s Hospital, Chengdu, China
43Sichuan Translational Medicine Hospital, Chinese Academy of Sciences, Chengdu, China
44Molecular Medicine Research Center, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Sichuan, China
45EyeCentre, Stuttgart, Germany
46University Eye Clinic, Ludwig-Maximilians-University Munich, Munich, Germany
47Department of Ophthalmology, University Hospital Regensburg, Regensburg, Germany
48Department of Ophthalmology and Visual Sciences, University of Utah, Salt Lake City, Utah, United States
49Department of Medicine (Biomedical Genetics), Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States
50Department of Ophthalmology, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States
51Department of Neurology, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States
52Department of Epidemiology, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States
53Department of Biostatistics, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, United States
54Department of Ophthalmology, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Republic of Korea
55Centre for Experimental Medicine, Queen’s University, Belfast, United Kingdom
56Department of Ophthalmology, University of Thessaly, School of Medicine, Larissa, Greece
57Department of Ophthalmology, Seoul National University Bundang Hospital, Seongnam, Republic of Korea
58Department of Ophthalmology, Weill Cornell Medical College, New York, New York, United States
59Scheie Eye Institute, Department of Ophthalmology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States
60Department of Biostatistics and Epidemiology University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States
61Department of Ophthalmology, The University of Alabama at Birmingham, Birmingham, Alabama, United States
62Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France
63UPMC Univ Paris 06, UMR_S 968, Institut de la Vision, Department of Genetics, Paris, France
64Le Centre National de la Recherche Scientifique (CNRS), UMR 7210, Paris, France
65Centre Hospitalier National d’Ophthalmologie des Quinze-Vingts, INSERM-DHOS CIC 503, Paris, France
66Fondation Ophthalmologique Adolphe de Rothschild, Paris, France
Pathway Analysis Implicates \( \text{PLCG2} \) in AMD

Institute of Ophthalmology, University College of London, London, United Kingdom
68 Académie des Sciences–Institut de France, Paris, France
69 Department of Molecular Genetics, Institute of Ophthalmology, London, United Kingdom
70 Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, United Kingdom
71 University Hospital Southampton, Southampton, United Kingdom
72 Department of Ophthalmology, Hadassah Hebrew University Medical Center, Jerusalem, Israel
73 Center for Human Disease Modeling, Duke University, Durham, North Carolina, United States
74 Department of Cell Biology, Duke University, Durham, North Carolina, United States
75 Department of Pediatrics, Duke University, Durham, North Carolina, United States
76 CEPH Fondation Jean Dausset, Paris, France
77 Commissariat à l’Énergie Atomique et aux Énergies Alternatives (CEA), Institut de Génomique–Centre National de Génotypage, Evry Cédex, France
78 Cole Eye Institute, Cleveland Clinic, Cleveland, Ohio, United States
79 Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina, United States
80 Department of Medicine, Duke University Medical Center, Durham, North Carolina, United States
81 Duke Molecular Physiology Institute, Duke University Medical Center, Durham, North Carolina, United States
82 Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Naples, Florida, United States
83 Department of Ophthalmology, NYU School of Medicine, New York, New York, United States
84 Department of Ophthalmology, Royal Perth Hospital, Perth, Western Australia, Australia
85 Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom
86 Department of Ophthalmology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom
87 Department of Ophthalmology UCSF Medical School, San Francisco, California, United States
88 Department of Ophthalmology and Visual Sciences, Vanderbilt University, Nashville, Tennessee, United States
89 Casey Eye Institute, Oregon Health & Science University, Portland, Oregon, United States
90 Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, United States
91 School of Clinical Sciences University of Edinburgh, Edinburgh, Scotland, United Kingdom
92 Center for Inherited Disease Research (CIDR) Institute of Genetic Medicine–Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
93 Department of Ophthalmology, David Geffen School of Medicine, UCLA, Stein Eye Institute, Los Angeles, California, United States
94 Department of Human Genetics, David Geffen School of Medicine, UCLA, Los Angeles, California, United States
95 Department of Human Genetics, Radboud University Medical Centre, Nijmegen, the Netherlands
96 Department of Pathology & Cell Biology, Columbia University, New York, New York, United States
97 Center for Translational Medicine, Moran Eye Center, University of Utah School of Medicine, Salt Lake City, Utah, United States
98 Division of Preventive Medicine, Brigham & Women’s Hospital, Harvard Medical School, Boston, Massachusetts, United States
99 Division of Epidemiology and Clinical Applications, Clinical Trials Branch, National Eye Institute, National Institutes of Health, Bethesda, Maryland, United States
100 Institute for Computational Biology, Case Western Reserve University School of Medicine, Cleveland, Ohio, United States