

1 **Genome-wide association meta-analysis for early age-related** 2 **macular degeneration highlights novel loci and insights for** 3 **advanced disease**

4

5 **Authors**

6 Thomas W Winkler^{1*}, Felix Grassmann^{2,3*}, Caroline Brandl^{1,2,4}, Christina Kiel², Felix Günther^{1,5},
7 Tobias Strunz², Lorraine Weidner¹, Martina E Zimmermann¹, Christina A. Korb⁶, Alicia
8 Poplawski⁷, Alexander K Schuster⁶, Martina Müller-Nurasyid^{8,9,10}, Annette Peters^{11,12},
9 Franziska G Rauscher^{13,14}, Tobias Elze^{13,15}, Katrin Horn^{13,14}, Markus Scholz^{13,14}, Marisa
10 Cañadas-Garre¹⁶, Amy Jayne McKnight¹⁶, Nicola Quinn¹⁶, Ruth E Hogg¹⁶, Helmut Küchenhoff⁵,
11 Iris M Heid^{1§}, Klaus J Stark^{1§} and Bernhard HF Weber^{2§}

12 **Affiliations**

13 1: Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany; 2: Institute of Human Genetics,
14 University of Regensburg, Regensburg, Germany; 3: Department of Medical Epidemiology and Biostatistics, Karolinska Institutet,
15 Stockholm, Sweden; 4: Department of Ophthalmology, University Hospital Regensburg, Regensburg, Germany; 5: Statistical
16 Consulting Unit StaBLab, Department of Statistics, Ludwig-Maximilians-Universität Munich, Munich, Germany; 6: Department of
17 Ophthalmology, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany; 7: Institute for Medical
18 Biostatistics, Epidemiology and Informatics, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz,
19 Germany; 8: Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health,
20 Neuherberg, Germany; 9: Department of Internal Medicine I (Cardiology), Hospital of the Ludwig-Maximilians-University (LMU)
21 Munich, Munich, Germany; 10: Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Germany; 11: German
22 Center for Diabetes Research (DZD), Neuherberg, Germany; 12: Institute of Epidemiology, Helmholtz Zentrum München
23 Research Center for Environmental Health, Neuherberg, Germany; 13: Leipzig Research Centre for Civilization Diseases (LIFE),
24 Leipzig University, Leipzig, Germany; 14: Institute for Medical Informatics, Statistics, and Epidemiology (IMISE), Leipzig
25 University, Leipzig, Germany; 15: Schepens Eye Research Institute, Harvard Medical School, Boston, MA, United States; 16:
26 Centre for Public Health, Queen's University of Belfast, UK

27

28 * These authors contributed equally

29 § These authors jointly supervised this work

30

31 Corresponding author: iris.heid@klinik.uni-regensburg.de

32

33 **ABSTRACT**

34 **Background:** Advanced age-related macular degeneration (AMD) is a leading cause of
35 blindness. While around half of the genetic contribution to advanced AMD has been
36 uncovered, little is known about the genetic architecture of the preceding early stages of the
37 diseases.

38 **Methods:** To identify genetic factors for early AMD, we conducted a genome-wide association
39 meta-analysis with 14,034 early AMD cases and 91,214 controls from 11 sources of data
40 including data from the International AMD Genomics Consortium (IAMGDC) and the UK
41 Biobank (UKBB). We ascertained early AMD via color fundus photographs by manual grading
42 for 10 sources and by using an automated machine learning approach for >170,000 images
43 from UKBB. We searched for significant genetic loci in a genome-wide association screen
44 ($P < 5 \times 10^{-8}$) based on the meta-analysis of the 11 sources and via a candidate approach based
45 on 13 suggestive early AMD variants from Holliday et al 2013 ($P < 0.05/13$, additional 3,432
46 early AMD cases and 11,235 controls). For the novel AMD regions, we conducted in-silico
47 follow-up analysis to prioritize causal genes and pathway analyses.

48 **Results:** We identified 11 loci for early AMD, 9 novel and 2 known for early AMD. Most of
49 these 11 loci overlapped with known advanced AMD loci (near *ARMS2/HTRA1*, *CFH*, *APOE*,
50 *C2*, *C3*, *CETP*, *PVRL2*, *TNFRSF10A*, *VEGFA*), except two that were completely novel to any
51 AMD. Among the 17 genes within the two novel loci, in-silico functional annotation suggested
52 *CD46* and *TYR* as the most likely responsible genes. We found the presence or absence of
53 an early AMD effect to distinguish known pathways of advanced AMD genetics
54 (complement/lipid pathways or extracellular matrix metabolism, respectively).

55 **Conclusions:** Our data on early AMD genetics provides a resource comparable to the existing
56 data on advanced AMD genetics, which enables a joint view. Our large GWAS on early AMD
57 identified novel loci, highlighted shared and distinct genetics between early and advanced
58 AMD and provides insights into AMD etiology. The ability of early AMD effects to differentiate
59 the major pathways for advanced AMD underscores the biological relevance of a joint view on
60 early and advanced AMD genetics.

61 **KEYWORDS**

62 Genome-wide association study (GWAS); Meta-analysis; Age-related macular degeneration
63 (AMD); Early AMD; *CD46*; *TYR*; International AMD Genomics Consortium (IAMDGCC); UK
64 Biobank (UKBB); machine-learning; automated phenotyping

65

66 **BACKGROUND**

67 Age-related macular degeneration (AMD) is the leading cause of irreversible central vision
68 impairment in industrialized countries. Advanced AMD presents as geographic atrophy (GA)
69 and/or neovascular (NV) complications (1). Typically, advanced AMD is preceded by clinically
70 asymptomatic and thus often unrecognized early disease stages. Early AMD is characterised
71 by differently sized yellowish accumulations of extracellular material between Bruch's
72 membrane and retinal pigment epithelium (RPE) or between RPE and the photoreceptors
73 (drusen or subretinal drusenoid deposits, respectively). Other features of early AMD are RPE
74 abnormalities, including depigmentation or increased amount of pigment (1).

75 Early and advanced AMD can be documented by color fundus imaging of the central
76 retina and/or other multimodal imaging approaches including optical coherence tomography
77 (OCT) (1–3). While the definition of advanced AMD is reasonably homogeneous across clinical
78 and epidemiological studies, the classification of early AMD is more variable and different
79 studies traditionally apply differing classification systems (4,5).

80 Epidemiological studies show that high age is the strongest risk factor for early and
81 advanced AMD onset as well as progression (1,6–8). A robust genetic influence was shown
82 for advanced AMD (1,9,10) with 34 distinct loci at genome-wide significance in a large genome-
83 wide association study (GWAS) for advanced AMD (9). The genes underneath these advanced
84 AMD loci were found to be enriched for genes in the alternative complement pathway, HDL
85 transport, and extracellular matrix organization and assembly (9).

86 Exploring the genetics of early AMD offers the potential to understand the mechanisms
87 of early disease processes, but also for the development to advanced AMD when comparing
88 genetic effect sizes for early and advanced stages. Yet there have been few published GWAS
89 searches for early AMD. One meta-analysis on 4,089 early AMD patients and 20,453 control

90 persons reported two loci with genome-wide significance, both being well known from
91 advanced AMD, the *CFH* and the *ARMS2/HTRA1* locus (11).

92 We have thus set out to gather GWAS data for early AMD from 11 sources including
93 own study data, data from the International AMD Genomics Consortium (IAMDGC), dbGaP
94 and UK Biobank to conduct the largest GWAS meta-analysis on early AMD to date.

95

96 **METHODS**

97 **GWAS data from 11 sources**

98 We included 11 sources of data with GWAS data and color fundus photography for early AMD
99 phenotyping (**Table S1**). Our studies were primarily population-based cohort studies, where
100 the baseline survey data were used for this analysis from studies of the authors (GHS, LIFE,
101 NICOLA, KORA, AugUR) as well as for publicly available studies from dbGaP (ARIC, CHS,
102 WHI; accession numbers: phs000090.v5.p1, phs000287.v6.p1, phs000746.v2.p3). We also
103 included data from UK Biobank for participants from baseline and additional participants from
104 the follow-up survey, since the color fundus photography program had started only after the
105 main study onset (application number #33999). The studies captured an age range from 25 to
106 100 years of age (mean age from 47.5 years to 77.2 years across the 10 population-based
107 studies, AugUR with the very old individuals range from 70.3 years to 95 years). About 50% of
108 the study participants in each study were male (except for the Women's Health Initiative, WHI)
109 and all demonstrated European ancestry. For each of these cross-sectional data sets,
110 participants with at least one eye gradable for AMD (see below) and with existing GWAS data
111 were eligible for our analysis. We excluded participants with advanced AMD. We used
112 participants with ascertained early AMD as cases and participants being ascertained for not
113 having any signs of AMD as controls (n=7,363 cases, 73,358 controls across these population-
114 based studies). Case-control data were also included from IAMDGC (<http://amdgenetics.org/>).
115 The early AMD GWAS from IAMDGC is based on 24,527 individual participant data from 26
116 sources (9). This data includes 17,856 participants with no AMD and 6,671 participants with
117 early AMD (excluding the 16,144 participants with advanced AMD). The cases and controls

118 from IAMDGC were 16 to 102 years of age (mean age = 71.7y). For all of these participants,
119 DNA samples had been gathered and genotyped centrally (see below) (9).

120

121 **Genotyping and Imputation**

122 All population-based studies were genotyped, quality controlled and imputed using similar chip
123 platforms and imputation approaches (**Table S2**). As the imputation backbone, the 1000
124 Genomes Phase 1 or Phase 3 reference panel was applied (12), except GHS was imputed
125 based on the Haplotype Reference Consortium (HRC) (13) and UK Biobank was imputed
126 based on HRC and the UK10K haplotype resource (14). Details on the UK Biobank genotypic
127 resource are described elsewhere (15). For the IAMDGC case-control data, DNA samples had
128 been gathered across all participants and genotyped on an Illumina HumanCoreExome array
129 and quality controlled centrally. Genotype quality control and imputation to the 1000 Genomes
130 phase 1 version 3 reference panel (>12 million variants) were conducted centrally. Details on
131 the IAMDGC data were described in detail by Fritsche et al (9).

132

133 **Phenotyping**

134 Across all studies included into this analysis, early AMD and the unaffected status was
135 ascertained by color fundus photography. For participants from AugUR and LIFE, “early AMD”
136 was classified according to the Three Continent Consortium (3CC) Severity Scale (4), which
137 separates “mild early” from “moderate” and “severe early” AMD stages depending on drusen
138 size, drusen area, or the presence of pigmentary abnormalities (4). For the analysis, we
139 collapsed any of these “early” AMD stages into the definition of “early AMD”. However, the
140 3CC Severity Scale was not available for the other studies. In these, similar early AMD
141 classifications, considering drusen size or area and presence of pigmentary abnormalities,
142 were used (**Table S1**): For participants from GHS, the Rotterdam Eye Study classification was
143 applied (16). For participants from NICOLA, the Beckman Clinical Classification was utilized
144 (17). Participants from the KORA study were classified as “early AMD” based on the AREDS-
145 9 step classification scheme and we defined “early AMD” for this analysis by AREDS-9 steps

146 2-8 (18). The ascertainment of IAMDGCC study participants is described in detail elsewhere and
147 covers various classification systems (9). Of note, LIFE and NICOLA phenotyping incorporated
148 OCT information additional to the information from color fundus imaging (**Table S1**). For UK
149 Biobank participants, color fundus images were received (application number 33999); there
150 was no existing AMD classification available in UK Biobank (see below). The AMD status of a
151 person was derived based on the AMD status of the eye with the more severe AMD stage
152 (“worse eye”) when both eyes were gradable, and as the grade of the one available eye
153 otherwise. Eyes were regarded as gradable, if at least one image of the eye fulfilled defined
154 quality criteria allowing for the assessment of AMD (bright image, good color contrast, full
155 macular region captured on images). Images were excluded from AMD grading if they revealed
156 obscuring lesions (e.g. cataract) or lesions considered to be the result of a competing retinal
157 disease (such as advanced diabetic retinopathy, high myopia, trauma, congenital diseases, or
158 photocoagulation unrelated to choroidal neovascularization). Details for IAMDGCC are
159 described previously (9). Persons with gradable images for at least one eye were included in
160 this analysis. Persons with advanced AMD defined as presence of neovascularization or
161 geographic atrophy in at least one eye were excluded for the main GWAS on early AMD.

162

163 **Automated classification of early AMD in UK Biobank**

164 To obtain early AMD phenotype data for UK Biobank participants, we used a pre-trained
165 algorithm for automated AMD classification based on an ensemble of convolutional neural
166 networks (19). In the UKBB baseline data, fundus images were available for 135,500 eyes of
167 68,400 individuals with at least one image. Among the additional 38,712 images of 19,501
168 individuals in the follow-up, there were 17,198 individuals without any image from baseline.
169 For each image (eye) at baseline and follow-up, we predicted the AMD stage on the AREDS-
170 9 step severity scale using the automated AMD classification. We defined a person-specific
171 AMD stage at baseline and follow-up based on the worse eye. Eyes that were classified as
172 ungradable were treated as missing data and, if diagnosis was available for only one eye, the
173 person-specific AMD stage was based on the classification of the single eye. If we obtained

174 an automated disease classification to an AMD stage (i.e. not “ungradable” for both eyes) at
175 baseline and follow-up, we used the follow-up disease stage (and follow-up age) in the
176 association analysis. By this, we obtained an automated AMD classification for 70,349
177 individuals (2,161 advanced AMD, 3,835 early AMD, 64,353 unaffected). Individuals with
178 advanced AMD were excluded from this analysis. Finally, we yielded 57,802 unrelated
179 individuals of European ancestry with valid GWAS data that had either early AMD or were free
180 of any AMD (3,105 cases, 54,697 controls). We evaluated the performance of the automated
181 disease classification by selecting 2,013 individuals (4,026 fundus images) and manual
182 classification based on the 3CC Severity Scale. We found substantial agreement between the
183 automated and the manual classification for the four categories of “no AMD”, “early AMD”,
184 “advanced AMD” and “ungradable” (concordance=79.5%, Cohen’s kappa $\kappa=0.613$) (20).

185

186 **Study-specific association analyses**

187 Study-specific logistic regression analyses (early AMD cases versus controls, excluding
188 advanced AMD cases) were applied by study partners (in Regensburg, Leipzig, Mainz, Belfast)
189 using an additive genotype model and according to a pre-defined analysis plan. All publicly
190 available data from dbGAP (studies ARIC, CHS and WHI) and UK Biobank as well as IAMDGC
191 data was analyzed in Regensburg. All studies inferred the association of each genetic variant
192 with early AMD using a Wald test statistic as implemented in RVTESTS (21). Age and two
193 principal components (to adjust for population stratification) were included as covariates in the
194 regression models. The IAMDGC analyses were further adjusted for DNA source as done
195 previously (9). For the IAMDGC data that stemmed from 26 sources, we conducted a sensitivity
196 analysis additionally adjusting for source membership according to previous work highlighting
197 slight differences in effect estimates (22); we found the same results.

198

199 **Quality control of study-specific aggregated data**

200 GWAS summary statistics for all data sources were processed through a standardized quality-
201 control (QC) pipeline (23). This involved QC checks on file completeness, range of test

202 statistics, allele frequencies, population stratification as well as filtering on low quality data. We
203 excluded variants with low minor allele count ($MAC < 10$, calculated as $MAC = 2 * N_{eff} * MAF$, with
204 N_{eff} being the effective sample size, $N_{eff} = 4N_{Cases} * N_{Controls} / (N_{Cases} + N_{Controls})$ and MAF being the
205 minor allele frequency), low imputation quality ($rsq < 0.4$) or large standard error of the
206 estimated genetic effect ($SE > 10$). Genomic control (GC) correction was applied to each GWAS
207 result to correct for population stratification within each study (24). The estimation of the GC
208 inflation factor was based on variants outside of the 34 known advanced AMD regions
209 (excluding all variants within < 5 Mb base positions to any of the 34 known advanced AMD lead
210 variants). The GC factors ranged from 1.00 to 1.04 (**Table S2**). We transferred all variant
211 identifiers to unique variant names consisting of chromosomal, base position (hg19) and allele
212 codes in (e.g. "3:12345:A_C", allele codes in ASCII ascending order).

213

214 **Meta-analysis**

215 For signal detection and effect quantification, study-specific genetic effects were combined
216 using an inverse-variance weighted fixed effect meta-analysis method as implemented in
217 METAL (25). We performed additional quality control on meta-analysis results: We only
218 included variants for identification that were available (i) in at least two of the data sources with
219 a total effective sample size of more than 5,000 individuals ($N_{eff} > 5,000$) and (ii) for chromosome
220 and position annotation in dbSNP (hg19). A conservative second GC correction (again
221 focusing on variants outside the known advanced AMD regions) was applied to the meta-
222 analysis result, in order to correct for potential population stratification across studies (24). The
223 GC lambda factor of the meta-analysis was 1.01. We judged the variants' association at
224 genome-wide significance level ($P < 5 \times 10^{-8}$). To evaluate the robustness of any novel genome-
225 wide significant AMD locus, we performed leave-one-out (LOO) meta-analyses.

226

227 **Variant selection, locus definition, and independent replication**

228 We combined genome-wide significant variants ($P < 5.0 \times 10^{-8}$) into independent loci by using a
229 locus definition similar to what was done previously (9): the most significant variant was

230 selected genome-wide, all variants were extracted that were correlated with this lead variant
 231 ($r^2 > 0.5$, using IAMDGC controls as reference) and a further 500 kb were added to both sides.
 232 All variants overlapping the so-defined locus were assigned to the respective locus. We
 233 repeated the procedure until no further genome-wide significant variants were detected. Genes
 234 overlapping the so-defined loci were used for biological follow-up analyses (gene region
 235 defined from start to end). To identify independent secondary signals at any novel AMD locus,
 236 approximate conditional analyses were conducted based on meta-analysis summary statistics
 237 using GCTA (26).

238 For an independent replication stage for genome-wide significant lead variants
 239 identified in our GWAS meta-analysis, we utilized the data from Holliday et al. (11) where
 240 possible. Since Holliday et al. did not provide genome-wide summary statistics, this was limited
 241 to variants, for which (or for reasonable proxies) summary statistics were reported. To account
 242 for the overlap between the current meta-analysis and the Holliday et al. meta-analysis, we
 243 removed the effects of the two overlapping studies (ARIC and CHS) from the reported Holliday
 244 et al. summary statistics. For β_{Holl} , β_{ARIC} , β_{CHS} and SE_{Holl} , SE_{ARIC} , SE_{CHS} being the genetic effects
 245 and standard errors of the Holliday et al. meta-analysis (reported) and of the ARIC or CHS
 246 study (available directly or as next best proxy in our meta-analysis), we estimated the “leave-
 247 two-out” (L2O) genetic effect β_{L2O} and standard error SE_{L2O} of the Holliday et al. meta-analysis
 248 as follows:

$$249 \quad SE_{L2O} = \sqrt{\left(\frac{SE_{Holl}^2}{1 - \left(\frac{SE_{Holl}^2}{SE_{ARIC}^2} \right)} \right) \left/ \left[1 - \left(\frac{SE_{Holl}^2}{1 - \left(\frac{SE_{Holl}^2}{SE_{ARIC}^2} \right)} \right) / SE_{CHS}^2 \right]} \right.}$$

$$250 \quad \beta_{L2O} = \left(1 + \frac{SE_{L2O}^2}{SE_{CHS}^2} \right) \left[\left(1 + \frac{SE_{Holl}^2 SE_{ARIC}^2}{1 - \frac{SE_{Holl}^2}{SE_{ARIC}^2}} \right) \cdot \beta_{Holl} - \left(\frac{SE_{Holl}^2 SE_{ARIC}^2}{1 - \frac{SE_{Holl}^2}{SE_{ARIC}^2}} \right) \cdot \beta_{ARIC} \right] - \frac{SE_{L2O}^2}{SE_{CHS}^2} \cdot \beta_{CHS}$$

251 We applied a Wald test to test the corrected genetic effect for significance.

252

253 **Candidate approach**

254 Additionally to the genome-wide search in our meta-analysis of 14,034 cases and 91,214
255 controls, we adopted a candidate approach based on the 14 reported suggestive variants by
256 Holliday et al. (P-values from 8.9×10^{-6} to 1.1×10^{-6} in their meta-analysis, 4,089 cases and
257 20,453 controls). For this, we removed the overlapping studies from our meta-analysis (ARIC,
258 CHS utilizing the reported variant; yielding 13,450 cases and 84,942 controls) and judged the
259 variants' association at a Bonferroni-corrected significant association ($P < 0.05/14$).

260

261 **Gene prioritization at newly identified AMD loci**

262 To prioritize genes and variants at the newly identified AMD loci, we conducted a range of
263 statistical and functional follow-up analyses. The following criteria were used: (1) Statistical
264 evidence; we computed the Bayes Factor based posterior probability of each variant using Z-
265 scores and derived 95% credible intervals for each locus (27). The method assumes a single
266 causal signal per locus. (2) Variant effect predictor (VEP) to explore whether any of the credible
267 variants was located in a relevant regulatory gene region (28). (3) eQTL analysis: We
268 downloaded expression summary statistics for the candidate genes in retina from the EyeGEx
269 database (29) and for 44 other tissues from the GTEx database (30) (both available at
270 www.gtexportal.org/home/datasets) and evaluated whether any of the credible variants
271 showed significant effects on expression levels in the aggregated data. For each significant
272 eQTL in EyeGEx, we conducted colocalization analyses using eCAVIAR (31) to evaluate
273 whether the observed early AMD association signal colocalized with the variants' association
274 with gene expression. (4) Retinal expression: We queried the EyeIntegration database to
275 evaluate genes in the relevant loci for expression in fetal or adult retina or RPE cells (32). (5)
276 Animal model: We queried the Mouse Genome Informatics (MGI) database
277 (www.informatics.jax.org) for each gene in the relevant loci for relevant eye phenotypes in mice
278 (33). (6) Human phenotype: The Online Mendelian Inheritance in Man (OMIM)[®] database was
279 queried for human eye phenotypes (McKusick-Nathans Institute of Genetic Medicine, Johns
280 Hopkins University, Baltimore, MD, queried 07/11/2019, www.omim.org).

281

282 **Phenome-wide association study for newly identified AMD loci**

283 We used 82 other traits and queried reported genome-wide significant ($P < 5.0 \times 10^{-8}$) lead
284 variants and proxies ($r^2 > 0.5$) for any of these traits for overlap with genes underneath our novel
285 loci as done previously (34). For this, we used GWAS summary results that were previously
286 aggregated from GWAS catalogue (35), GWAS central (36) and literature search.

287 For the novel early AMD lead variants, we further evaluated their association with 118
288 non-binary and 660 binary traits from the UK Biobank (37). The Phenome-wide association
289 study (PheWAS) web browser “GeneATLAS” (www.geneatlas.roslin.ed.ac.uk) was used for
290 the UK Biobank lookup. For each variant, association P values were corrected for the testing
291 of multiple traits by the Benjamini-Hochberg false-discovery-rate (FDR) method (38).

292

293 **Interaction analyses**

294 For the novel early AMD effects and for the 34 known advanced AMD lead variants (9), we
295 investigated whether age modulated early AMD effects by analyzing variant x age interaction
296 in seven data sources for which we had individual participant data available in Regensburg
297 (ARIC, CHS, WHI, IAMDGC, UKBB, AugUR and KORA). For each source, we applied logistic
298 regression and included a variant x AGE interaction term in the model (in addition to the
299 covariates used in the main analysis). We conducted meta-analysis across the seven sources
300 to obtain pooled variant x age interaction effects and applied a Wald test to test for significant
301 interaction (at a Bonferroni-corrected alpha-level). For the novel early AMD effects, we further
302 investigated whether age modulated advanced AMD effects by evaluating publically available
303 data from IAMDGC (39). Finally, we investigated whether a novel early AMD lead variant
304 modulated any of the effects of the 34 known AMD variants on advanced AMD (9). We used
305 the IAMDGC data and applied one logistic regression model for each pair of known advanced
306 AMD variants and novel early AMD variants including the two respective variants and their
307 interaction (and the same other covariates as before).

308

309 **Comparison of genetic effects on early and advanced stage AMD**

310 We estimated the genetic correlation between early and advanced AMD by utilizing the LDSC
311 tool (40) with the GWAS summary statistics for early and advanced AMD (from the current
312 meta-analysis and the IAMDGC (9), respectively). We used pre-calculated LD scores for
313 European ancestry
314 (https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2). We further
315 compared genetic effect sizes between early and advanced AMD for the novel early AMD lead
316 variants and for the 34 known advanced AMD lead variants (9). For this, we queried the novel
317 early AMD lead variants in the IAMDGC GWAS for advanced AMD (9) and (vice-versa) queried
318 the 34 known advanced AMD lead variants (9) in the early AMD meta-analysis results. We
319 compared effect sizes in a scatter plot and clustered the lead variants by their nominal
320 significant association on advanced and/or early AMD. We classify different types of loci in a
321 similar fashion as done previously for adiposity trait genetics (41): (1) “advanced-and-early”
322 AMD loci ($P_{\text{early}} < 0.05$, $P_{\text{adv}} < 0.05$), (2) “advanced-only” AMD loci ($P_{\text{early}} \geq 0.05$, $P_{\text{adv}} < 0.05$), (3)
323 “early-only” AMD loci ($P_{\text{early}} < 0.05$, $P_{\text{adv}} \geq 0.05$).

324

325 **Pathway analysis**

326 To evaluate whether “advanced-and-early” AMD loci versus “advanced-only” AMD loci
327 distinguished the major known pathways for advanced AMD, we performed pathway
328 enrichment analysis separately for these two classes. We used the genes in the gene
329 prioritization for all advanced AMD loci as previously described (9), derived the gene
330 prioritization score and selected the best scored gene in each locus (two genes in the case of
331 ties). We then separated the gene list according to the class of the respective locus, and
332 performed pathway enrichment analysis via Enrichr (42) with default settings searching
333 Reactome’s cell signaling pathway database 2016 (n=1,530 pathways). P-values were
334 corrected for multiple testing with Benjamini-Hochberg procedure (38).

335

336 **RESULTS**

337 **Eight genome-wide significant loci from a GWAS on early AMD**

338 We conducted a meta-analysis of genotyped and imputed data from 11 sources (14,034 early
339 AMD cases, 91,214 controls, for study-specific genotyping, analysis and quality control
340 **(Tables S1-S2)**). For all participants, early AMD or control status (i.e. no early nor late AMD)
341 was ascertained via color fundus photographs **(Table S1)**. This included automated machine-
342 learning based AMD classification of UK Biobank fundus images (application number 33999;
343 56,699 individuals from baseline, 13,650 additional individuals from follow-up) (15,19) .

344 Based on logistic regression association analysis in each of the 11 data sets meta-
345 analyzed via fixed effect model, we identified eight distinct loci with genome-wide significance
346 ($P=1.3 \times 10^{-116}$ to 4.7×10^{-8} , **Figure 1, Table 1**; “locus” defined by the lead variant and proxies,
347 $r^2 \geq 0.5$, +/-500 kb). Six of these loci were novel for early AMD; two loci had been identified for
348 early AMD previously (11).

349 Most of these loci overlap with known loci for advanced AMD (9) (*CFH*, *ARMS2/HTRA1*,
350 *C2*, *C3*, *CETP*, *VEGFA*, *TNFRSF10A*), except one which has not been identified in early or
351 advanced AMD GWAS before ($P=4.7 \times 10^{-8}$, lead variant rs4844620, near *CD46*, **Figure S1**).
352 This novel locus is additionally supported by independent data from Holliday et al. (11) (3,432
353 cases and 11,235 controls after removing the overlap to our meta-analysis: $P=7.0 \times 10^{-5}$).

354 This novel locus on chromosome 1 is >10 million base positions distant from the closest
355 known advanced AMD locus (around *CFH*) and the lead variant is uncorrelated with any of the
356 eight independent *CFH* locus variants known for advanced AMD ($r^2 < 0.01$). This locus showed
357 no second signal (GCTA (26) conditional $P \geq 5.0 \times 10^{-8}$, **Figure S1**). Since early AMD
358 ascertainment and classification was heterogeneous across the 11 data sources of our meta-
359 analysis, we conducted sensitivity analysis leaving out one data set at a time and found the
360 *CD46* locus association to be robust, except for a slightly larger early AMD effect when
361 excluding IAMDG data **(Figure S2)**. Taken together, we identified six novel loci for early AMD
362 including a novel locus for any AMD with genome-wide significance near *CD46*.

363

364 **Three significant loci from a candidate-based approach of 14 variants**

365 Subsequently, we applied a candidate-based approach by investigating the lead variants of
366 the 13 loci reported as suggestive by the previous GWAS for early AMD (4,089 early AMD
367 cases , 20,453 controls; reported P between 1.1×10^{-6} and 8.9×10^{-6}) (11); one further variant
368 reported as suggestive was in the *CD46* locus that we have identified with genome-wide
369 significance in our data and for which we have already utilized parts of the reported Holliday
370 et al. data for independent replication (see above). We re-analyzed our data excluding the
371 overlap with the previous GWAS (i.e. excluding ARIC, CHS study, yielding 13,450 early AMD
372 cases and 84,942 controls). We found three of the 13 variants as significantly associated at a
373 Bonferroni-corrected level ($P < 0.05/13 = 0.0038$) depicting three novel loci for early AMD (**Table**
374 **2**). This included two variants in known advanced AMD loci (*APOE/TOMMS*; *PVRL2*) and one
375 completely novel for any AMD near *TYR* (rs621313, $P = 6.8 \times 10^{-4}$, **Figure S3**). There was no
376 second signal within this locus (conditional $P > 0.0036$, **Figure S3**). The observed effect was
377 robust upon exclusion of any of the 11 data sets (**Figure S4**). Altogether, we have thus
378 identified 9 novel loci for early AMD (6 from GWAS, 3 from candidate-based approach)
379 including two completely novel for any AMD (1 from GWAS near *CD46*, 1 from candidate-
380 based approach near *TYR*).

381

382 **Gene prioritization at the two novel loci**

383 To prioritize variants and genes at the *CD46* and *TYR* locus, we conducted *in silico* follow-up
384 analyses for all variants and overlapping genes for each of these two loci (4,451 or 5,729
385 variants, 10 or 7 genes, respectively). We found several interesting aspects (**Table 3**): (1)
386 When prioritizing variants according to their statistical evidence for being the driver variant by
387 computing 95% credible sets of variants (27), we found 23 and 294 credible set variants for
388 the *CD46* and *TYR* locus, respectively (**Table S3**). (2) Using the Variant Effect Predictor (28),
389 we assessed overlap of credible set variants with functional regulatory regions and found
390 variants influencing the transcript and/or the protein for four genes (**Table S4**): variants causing
391 an alternative splice form for *CD46*, a nonsense-mediated mRNA decay (NMD) for *CR1L*, a
392 missense variant for *TYR* (rs1042602, $r^2 = 0.56$ to the lead variant rs621313), and NMD variants

393 for *NOX4*. (3) We investigated credible set variants for being an expression quantitative trait
394 locus (eQTL) for any of the 17 genes in retina (Eye Genotype Expression database, EyeGEx
395 (29)) or in 44 other tissues (Genotype-Tissue Expression database, GTEx (30)). For the *CD46*
396 locus, we observed significant association of the lead variant and additional 16 credible set
397 variants on *CD46* expression in retina (FDR<5%, **Table S5**); the early AMD risk increasing
398 alleles of all 17 variants were associated with elevated *CD46* expression. Importantly, we
399 observed the expression signal to colocalize with the early AMD association signal using
400 eCAVIAR (31) (3 variants with colocalization posterior probability CLPP>0.01, **Table S6**,
401 **Figure S5-S6**). We also found credible variants to be associated with *CD46* expression in 15
402 other tissues from GTEx, including four brain tissues (FDR<0.05, **Table S7**). Among the
403 credible set variants in the two loci, we found no further eQTL for any of the other genes. When
404 extending beyond the credible set, we found one further *CD46* locus variant as eQTL for *CD55*,
405 but without colocalization (**Table S6, Figure S5-S6**). These findings support the idea that the
406 credible set captures the essential signal. (4) We queried the 17 genes overlapping the two
407 loci for expression in eye tissue and cells in EyeIntegratation summary data (43). We found five
408 and three genes, respectively, expressed in adult retina and adult RPE cells (*CD46, PLXNA2*,
409 *CR1, CD34, CD55; TYR, GRM5, NOX4; Figure S7-S8*). (5) When querying the 17 genes in
410 the Mouse Genome Informatics, MGI (33) or Online Mendelian Inheritance in Man, OMIM®,
411 database, for eye phenotypes in mice or humans, we identified relevant eye phenotypes for
412 five genes in mice (*CD46, CR1, CR1L, PLXNA2; TYR; Table S8*) and for one gene in human
413 (*TYR; Table S9*).

414 While it is debatable how to prioritize evidence for a gene's probability to be causal,
415 one approach is to count any of the following characteristics for each of the 17 genes (Gene
416 Prioritization Score, GPS **Table 3**): any credible set variant is (i) protein-coding, (ii) involved in
417 NMD, (iii) affecting splice function, (iv) an eQTL for this gene in retina (EyeGEx) or in any other
418 tissue (GTEx), or/and the gene (v) is expressed in retina or RPE, (vi) linked to eye phenotype
419 in mouse or (vii) human. This approach offered *CD46* and *TYR* as the highest scored gene in
420 the respective locus (GPS=4 for each; **Table 3**).

421

422 **Phenome-wide association search for the two novel loci**

423 Co-association of variants in the two novel loci for early AMD with other traits and diseases
424 may provide insights into shared disease mechanisms. We queried different data sets on
425 numerous phenotypes by a gene-based and by a locus-based view.

426 For the gene-based view, we focused on 82 traits and evaluated reported genome-wide
427 significant ($P < 5.0 \times 10^{-8}$) lead variants (and proxies, $r^2 > 0.5$) for overlap with any of the 17 gene
428 regions (**Table S10**). For the *CD46* locus, we found significant association corrected for
429 multiple testing (false-discovery rate, $FDR < 5\%$) for schizophrenia (in *CD46* and *CR1L*) and
430 for Alzheimer's disease (in *CR1*, **Table S11**). For the *TYR* locus, we found significant
431 associations for eye color, skin pigmentation and skin cancer (in *GRM5* and *TYR*, **Tables S11**).

432 For the locus-based view, we conducted a phenome-wide association study
433 (PheWAS): we evaluated whether the two lead variants were associated with any of the 778
434 traits in UK Biobank using GeneAtlas ($n=452,264$, age-adjusted estimates; **Table S12**) (37).
435 For the *CD46* lead variant, we identified 27 significant trait associations ($FDR < 5\%$), including
436 four with particularly strong evidence ($P < 5.0 \times 10^{-8}$; white blood cell, neutrophil, monocyte count
437 and plateletcrit); the early AMD risk increasing allele (G, frequency=79%) was associated
438 consistently with increased blood cell counts. We did not find a significant association of the
439 *CD46* lead variant with schizophrenia in UK Biobank ($FDR > 5\%$; Alzheimer's disease not
440 available). For the *TYR* lead variant (rs621313, G allele associated with increased early AMD
441 risk, frequency=48%), we identified 20 significant trait associations including Melanoma
442 ($FDR < 5\%$, G allele associated with increased Melanoma risk) and two with particularly strong
443 evidence for skin color and ease of skin tanning ($P < 5.0 \times 10^{-8}$, G allele associated with brighter
444 skin color and increased ease of skin tanning).

445

446 **Advanced AMD association and interaction analyses for the two novel loci**

447 Next, we investigated whether the early AMD loci *CD46* or *TYR* were associated with advanced
448 AMD. We thus queried the two lead variants for early AMD (rs4844620 and rs621313,

449 respectively) for their advanced AMD association in the IAMDGC data (**Table S13**). We
450 observed nominally significant directionally consistent effects for advanced AMD ($OR_{adv}=1.05$,
451 95% confidence interval, $CI=[1.01, 1.09]$ and $1.03 [1.00, 1.07]$, $P_{adv}=0.02$ and 0.05 , respectively)
452 that were slightly smaller compared to early AMD effects ($OR_{early}=1.10 [1.06, 1.14]$ and 1.05
453 $[1.02, 1.08]$, $P_{early}=4.7 \times 10^{-8}$ and 6.8×10^{-4}).

454 When exploring variant x age interaction for early AMD (in a subset of our meta-analysis
455 of 10,890 early AMD cases and 54,697 controls) or for advanced AMD (IAMDGC data (39))
456 for the two novel locus lead variants, we found no statistically significant interaction at a
457 Bonferroni-corrected level for early or advanced AMD ($P_{G \times AGE} > 0.05/2 = 0.025$, **Table S14-S15**).

458 We were interested in whether one of the two novel lead variants showed interaction
459 with any of the 34 known advanced AMD variants for association with advanced AMD
460 (IAMDGC data). We found no significant interaction ($P_{G \times G} > 0.05/34/2$), **Table S16**), which
461 suggests that the known advanced AMD effects are not modulated by the two novel early AMD
462 variants.

463

464 **Dissecting advanced AMD genetics into shared and distinct genetics for early AMD**

465 We were interested in whether we could learn about advanced AMD genetics from a joint view
466 of advanced and early AMD genetic effects. First, when computing genetic correlation of
467 advanced AMD genetics with early AMD genetics, we found a substantial correlation of 78.8%
468 (based on LD-score regression). Second, we contrasted advanced AMD effect sizes (IAMDGC
469 data (9)) with early AMD effect sizes (our meta-analysis,) for the 34 known advanced AMD
470 lead variants (**Figure 2, Table S13**). We found two classes of variants: (1) 25 variants showed
471 nominally significant effects on early AMD ($P < 0.05$; “advanced-and-early-AMD loci”), all
472 directionally consistent and all smaller for early vs. advanced AMD ($OR_{early}=1.04-1.47$; OR_{adv}
473 $=1.10-2.81$); (2) nine variants had no nominally significant effect on early AMD ($P \geq 0.05$;
474 “advanced-only AMD loci”). We did not find any variant with early AMD effects into the opposite
475 direction as the advanced AMD effects. Also, we did not find any variant-age interaction on
476 early AMD (**Table S15**).

477 We observed that complement genes *CFH*, *CFI*, *C3*, *C9*, and *C2* were all included in
478 the 25 advanced-and-early-AMD loci. We were thus interested in whether advanced-and-early-
479 AMD loci suggested different pathways compared to advanced-only-AMD loci. For this, we
480 utilized the GPS from our previous work on advanced AMD (9) to select the best-supported
481 genes in each of these loci (**Table S17**). We applied Reactome pathway analyses via Enrichr
482 (42) twice: (i) for the 35 genes in the 25 advanced-and-early-AMD loci and (ii) for the nine
483 genes in eight advanced-only-AMD loci (no gene in the “narrow” locus definition of the *RORB*
484 locus). This revealed significant enrichment (corrected $P < 0.05$) for genes from “complement
485 system” and “lipoprotein metabolism” in the 25 advanced-and-early-AMD loci and enrichment
486 for genes in the pathways “extracellular matrix organization” and “assembly of collagen fibrils”
487 in the 8 advanced-only-AMD loci (**Table 4**). This suggested that the early AMD effect of
488 advanced AMD variants distinguished the major known pathways for advanced AMD.

489

490 **DISCUSSION**

491 Based on the largest genome-wide meta-analysis for early AMD to date encompassing
492 ~14,000 cases and ~91,000 controls, all color fundus photography confirmed, we identified 11
493 loci for early AMD: 9 loci highlighted here for the first time with significant association for early
494 AMD and two previously identified (11). Nine of the 11 loci overlapped with known loci for
495 advanced AMD (9) and two had not been detected by GWAS for early or advanced AMD so
496 far. Our post-GWAS approach highlighted *CD46* and *TYR* as compelling candidate genes in
497 the two loci. Our joint view on early and advanced AMD genetics allowed us to differentiate
498 between shared and distinct genetics for these two disease stages, which the pathway
499 analyses suggested to be biologically relevant.

500 The locus around *CD46*, had not been identified with genome-wide significance by the
501 previous largest GWAS for advanced AMD (9) (16,144 advanced AMD cases, 17,832 controls)
502 or early AMD (11) (4,089 early AMD cases, 20,453 controls). Our meta-analysis was more
503 than three times larger than the previous early AMD GWAS (effective sample size 48,651
504 compared to a 13,631(11)) and had a larger power to detect an “any AMD” effect with genome-

505 wide significance than the previous advanced AMD GWAS (e.g. for OR=1.10, allele frequency
506 30%: power=92% compared to 61%, respectively). The *TYR* locus had not been genome-wide
507 significant in any previous GWAS on advanced or early AMD; it was significant for early AMD
508 at a Bonferroni-corrected level in our candidate approach (11).

509 Prioritization of genes underneath association signals is a known challenge, but highly
510 relevant for selecting promising candidates for functional follow-up. Our systematic approach,
511 scrutinizing all genes underneath our two newly identified loci, highlighted *CD46* and *TYR* as
512 the most supported genes. *CD46* is an immediate compelling candidate as a part of the
513 complement system (44). Complement activation in retina is thought to have a causal role for
514 AMD (45,46). Importantly, we found our *CD46* GWAS signal to colocalize with *CD46*
515 expression with the early AMD risk increasing allele (rs4844620 G) increasing *CD46*
516 expression in retinal cells. On the one hand, this contrasts the presumption that a higher *CD46*
517 expression in eye tissue should protect from AMD, based on previous *CD46* expression data
518 (47) and a documented AMD risk increasing effect for increased complement inhibition (48).
519 On the other hand, *CD46* had also been found to have pathogenic receptor properties for
520 human viral and bacterial pathogens (e.g. measles virus) (49) and is known to down-modulate
521 adaptive T helper type 1 cells (50). Furthermore, a GWAS on neutralizing antibody response
522 to measles vaccine had identified two intronic *CD46* variants (rs2724384, rs2724374) (51). In
523 our data, these two variants were in the 95% credible set for the *CD46* locus, highly correlated
524 with our lead variant rs4844620 ($r^2 \geq 0.95$), and the major alleles (rs2724374 T, rs2724384 A)
525 increased early AMD risk. Interestingly, the rs2724374 G was shown for *CD46* exon skipping
526 resulting in a shorter *CD46* isoform with a potential role in pathogen binding (51). Based on
527 this, one may hypothesize that the observed *CD46* signal in early AMD is related to pathogenic
528 receptor properties rather than complement inactivation.

529 At the second locus, *TYR* appears as the best supported gene by our systematic
530 scoring. This locus and gene was already discussed by Holliday and colleagues (11), although
531 they did not have statistically significant association with early AMD in their data. *TYR* is
532 important for melanin production and *TYR* variants in human were associated with skin, eye

533 and hair color (52–54). Melanin has a protective function in RPE against oxidative stress from
534 UV radiation (55) and RPE pigmentation alterations are linked to AMD (56). In contrast to skin,
535 melanin pigment in RPE was shown to be synthesized prenatally and stored in melanosomes
536 throughout life, while tyrosinase activity in adult RPE remained controversial (57,58). Our study
537 provides additional insight: our query of EyeInteGration data (32) found *TYR* expressed in adult
538 human RPE, which may influence melanin production and its protective function in retina.
539 Therefore, variants in or near *TYR* may represent risk factors beyond fetal melanin production.
540 While we did not identify any *cis* effect between variants and *TYR* expression, one of our
541 credible set variants in *TYR* (rs1042602) is a missense variant. Interestingly, this variant is a
542 GWAS lead variant not only for skin color (53), but also for macular thickness in UK Biobank
543 (59); the allele associated with thicker retina showed increased early AMD risk in our data.
544 Since thicker RPE/Bruch's membrane complex was associated with increased early AMD risk
545 in the AugUR study (60), this would be in line with our early AMD risk increasing allele being
546 linked to a process of increased accumulation of drusenoid debris in the RPE/Bruch's
547 membrane complex. Although *CD46* and *TYR* are the most supported genes in the two loci,
548 we cannot rule-out the relevance of other genes in the loci.

549 It is a strength of the current study that early AMD and control status was ascertained
550 by color fundus photography, not relying on health record data. However, the early AMD
551 classification in our GWAS is heterogeneous across the 11 data sets: two studies incorporated
552 information from optical coherence tomography (NICOLA, GHS), the UK Biobank classification
553 was derived by a machine-learning algorithm (19), and the IAMDGCC data was multi-site with
554 different classification approaches (9). The uncertainty in early AMD classification and the
555 substantial effort required for any manual AMD classification are likely reasons for the sparsity
556 of early AMD GWAS so far. Our sensitivity analysis showed that our findings did not depend
557 on one or the other data source or classification approach. Our data on early AMD genetics is
558 comparable in size to the existing data on advanced AMD genetics from IAMDGCC (summary
559 statistics at <http://amdgenetics.org/>) and thus provides an important resource (summary
560 statistics at <http://genepi-regensburg.de>) to enable a joint view.

561 By this joint view, we were able to differentiate the 34 loci known for advanced AMD
562 into 25 “advanced-and-early-AMD loci” and nine “advanced-AMD-only loci”. Pathway
563 enrichment analyses conducted separately for these two groups effectively discriminated the
564 major known pathways for advanced AMD genetics (9): complement complex and lipid
565 metabolism for “advanced-and-early-AMD” loci; extracellular matrix metabolism for “advanced-
566 AMD-only” loci. The two novel loci around *CD46* and *TYR* fit to the definition of “advanced-
567 and-early-AMD” loci and the *CD46* being part of the complement system supports the above
568 stated pathway pattern. The larger effect size for early compared to advanced AMD for the two
569 novel loci may – in part – be winner’s curse.

570 How do our observations relate to potential etiological models? (1) For a genetic variant
571 capturing an underlying mechanism that triggers both early and advanced AMD, we would
572 expect the variant to show association with early and advanced AMD (compared to “healthy”)
573 with directionally consistent effects (**Figure 3**, Model 1). This would be in line with our observed
574 associations for the 27 “advanced-and-early-AMD” loci (25 known advanced AMD loci, 2 novel
575 loci). This would also suggest that mechanisms of complement system or lipid metabolism
576 trigger both early and advanced disease. (2) For a mechanism that triggers advanced AMD no
577 matter whether the person is “healthy” or has early AMD, we would anticipate a variant effect
578 for advanced AMD, but not for early AMD (**Figure 3**, Model 2). This would be in line with our
579 observed associations for the nine “advanced-AMD-only” loci. This would also suggest that
580 mechanisms of extracellular matrix metabolism trigger advanced AMD rather than early AMD.
581 Of note, these include the *MMP9* locus, which is thought to trigger vascularization and wet
582 AMD (9). (3) Another mechanism is conferred by variants that are purely responsible for
583 progression from early to advanced AMD, but do not increase advanced AMD risk for “healthy”
584 individuals. In such a scenario, the advanced AMD risk increasing allele would be under-
585 represented among persons with early AMD (**Figure 3**, Model 3), particularly at older age, and
586 it would be associated with decreased risk of early AMD (compared to “healthy”). None of the
587 identified variants showed this pattern overall or for older age in the variant x age interaction
588 analyses. (4) For a mechanism that triggers early AMD, but has no impact on the progression

589 from early to advanced AMD, we would have an effect on early AMD, but no effect on advanced
590 AMD (**Figure 3**, Model 4). We did not find such a variant.

591 Our data and joint view on effects for both disease stages support two of the four
592 etiological models. One may hypothesize that the unsupported models are non-existing or
593 unlikely. There are limitations to consider: (1) To reduce complexity, we adopted an isolated
594 view per variant with some accounting for interaction, but ignoring more complex networks. (2)
595 Early AMD effects were estimated predominantly in population-based studies, while advanced
596 AMD effects were from a case-control design. (3) The cut-off of “nominal significance” for
597 separating variants into “advanced-and-early” or “advanced-only” loci is arbitrary and larger
598 data might give rise to re-classification. Still, the power to detect effects for early AMD in our
599 meta-analysis was similar to the power in the advanced AMD data from IAMDGC (for OR=1.05,
600 allele frequency 30%, nominal significance: power=94% or 83%, respectively). (4) An improved
601 disentangling of genetic effects for the two chronologically linked disease stages will be an
602 important subject of further research, requiring large-scale population-based studies with long-
603 term follow-up and the estimation of transition probabilities.

604

605 **CONCLUSIONS**

606 In summary, our large GWAS on early AMD identified novel loci, highlighted shared and distinct
607 genetics between early and advanced AMD and provides insights into AMD etiology. The
608 ability of early AMD effects to differentiate the major pathways for advanced AMD underscores
609 the biological relevance of a joint view on early and advanced AMD genetics.

610

611 **REFERENCES**

- 612 1. Lim, L. S., Mitchell, P., Seddon, J. M., et al. (2012) Age-related macular degeneration.
613 *Lancet*, **379**, 1728–1738.
- 614 2. Garrity, S. T., Sarraf, D., Freund, K. B., et al. (2018) Multimodal Imaging of
615 Nonneovascular Age-Related Macular Degeneration. *Invest. Ophthalmol. Vis. Sci.*, **59**,
616 AMD48–AMD64.
- 617 3. Forte, R., Querques, G., Querques, L., et al. (2012) Multimodal imaging of dry age-
618 related macular degeneration. *Acta Ophthalmol.*, **90**, 281–287.

- 619 4. Klein, R., Meuer, S. M., Myers, C. E., et al. (2014) Harmonizing the classification of age-
620 related macular degeneration in the three-continent AMD consortium. *Ophthalmic*
621 *Epidemiol*, **21**, 14–23.
- 622 5. Brandl, C., Zimmermann, M. E., Günther, F., et al. (2018) On the impact of different
623 approaches to classify age-related macular degeneration: Results from the German
624 AugUR study. *Sci. Rep.*, **8**, 1–10.
- 625 6. AREDS (2000) Risk factors associated with age-related macular degeneration. A case-
626 control study in the age-related eye disease study: age-related eye disease study report
627 number 3. Age-Related Eye Disease Study Research Group. *Ophthalmology*, **107**,
628 2224–2232.
- 629 7. Smith, W., Assink, J., Klein, R., et al. (2001) Risk factors for age-related macular
630 degeneration: Pooled findings from three continents. *Ophthalmology*, **108**, 697–704.
- 631 8. Yonekawa, Y., Miller, J. and Kim, I. (2015) Age-Related Macular Degeneration:
632 Advances in Management and Diagnosis. *J. Clin. Med.*, **4**, 343–359.
- 633 9. Fritsche, L. G., Igl, W., Bailey, J. N. C., et al. (2016) A large genome-wide association
634 study of age-related macular degeneration highlights contributions of rare and common
635 variants. *Nat. Genet.*, **48**, 134–143.
- 636 10. Grassmann, F., Fritsche, L. G., Keilhauer, C. N., et al. (2012) Modelling the genetic risk
637 in age-related macular degeneration. *PLoS One*, **7**.
- 638 11. Holliday, E. G., Smith, A. V., Cornes, B. K., et al. (2013) Insights into the Genetic
639 Architecture of Early Stage Age-Related Macular Degeneration: A Genome-Wide
640 Association Study Meta-Analysis. *PLoS One*, **8**, e53830.
- 641 12. Auton, A., Abecasis, G. R., Altshuler, D. M., et al. (2015) A global reference for human
642 genetic variation. *Nature*, **526**, 68–74.
- 643 13. McCarthy, S., Das, S., Kretzschmar, W., et al. (2016) A reference panel of 64,976
644 haplotypes for genotype imputation. *Nat. Genet.*, **48**, 1279–1283.
- 645 14. Walter, K., Min, J. L., Huang, J., et al. (2015) The UK10K project identifies rare variants
646 in health and disease. *Nature*, **526**, 82–89.
- 647 15. Bycroft, C., Freeman, C., Petkova, D., et al. (2018) The UK Biobank resource with deep
648 phenotyping and genomic data. *Nature*, **562**, 203–209.
- 649 16. Korb, C. A., Kottler, U. B., Wolfram, C., et al. (2014) Prevalence of age-related macular
650 degeneration in a large European cohort: Results from the population-based Gutenberg
651 Health Study. *Graefe's Arch. Clin. Exp. Ophthalmol.*, **252**, 1403–1411.
- 652 17. Ferris, F. L., Wilkinson, C. P., Bird, A., et al. (2013) Clinical Classification of Age-related
653 Macular Degeneration. *Ophthalmology*, **120**, 844–851.
- 654 18. Brandl, C., Breinlich, V., Stark, K. J., et al. (2016) Features of Age-Related Macular
655 Degeneration in the General Adults and Their Dependency on Age, Sex, and Smoking:
656 Results from the German KORA Study. *PLoS One*, **11**, e0167181.
- 657 19. Grassmann, F., Mengelkamp, J., Brandl, C., et al. (2018) A Deep Learning Algorithm
658 for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related
659 Macular Degeneration from Color Fundus Photography. *Ophthalmology*, **125**, 1410–
660 1420.
- 661 20. Landis, J. R. and Koch, G. G. (1977) The measurement of observer agreement for
662 categorical data. *Biometrics*, **33**, 159–74.
- 663 21. Zhan, X., Hu, Y., Li, B., et al. (2016) RVTESTS: An efficient and comprehensive tool for
664 rare variant association analysis using sequence data. *Bioinformatics*, **32**, 1423–1426.
- 665 22. Gorski, M., Günther, F., Winkler, T. W., et al. (2019) On the differences between mega-

- 666 and meta-imputation and analysis exemplified on the genetics of age-related macular
667 degeneration. *Genet. Epidemiol.*, **43**, 559–576.
- 668 23. Winkler, T. W., Day, F. R., Croteau-Chonka, D. C., et al. (2014) Quality control and
669 conduct of genome-wide association meta-analyses. *Nat. Protoc.*, **9**, 1192–212.
- 670 24. Devlin, A. B., Roeder, K. and Devlin, B. (2013) Genomic Control for Association. **55**,
671 997–1004.
- 672 25. Willer, C. J., Li, Y. and Abecasis, G. R. (2010) METAL: Fast and efficient meta-analysis
673 of genomewide association scans. *Bioinformatics*, **26**, 2190–2191.
- 674 26. Yang, J., Ferreira, T., Morris, A. P., et al. (2012) Conditional and joint multiple-SNP
675 analysis of GWAS summary statistics identifies additional variants influencing complex
676 traits. *Nat. Genet.*, **44**, 369–375.
- 677 27. Kichaev, G., Yang, W. Y., Lindstrom, S., et al. (2014) Integrating Functional Data to
678 Prioritize Causal Variants in Statistical Fine-Mapping Studies. *PLoS Genet.*, **10**.
- 679 28. McLaren, W., Gil, L., Hunt, S. E., et al. (2016) The Ensembl Variant Effect Predictor.
680 *Genome Biol.*, **17**, 122.
- 681 29. Ratnapriya, R., Sosina, O. A., Starostik, M. R., et al. (2019) Retinal transcriptome and
682 eQTL analyses identify genes associated with age-related macular degeneration. *Nat.*
683 *Genet.*, **51**, 606–610.
- 684 30. Aguet, F., Brown, A. A., Castel, S. E., et al. (2017) Genetic effects on gene expression
685 across human tissues. *Nature*, **550**, 204–213.
- 686 31. Hormozdiari, F., van de Bunt, M., Segrè, A. V., et al. (2016) Colocalization of GWAS
687 and eQTL Signals Detects Target Genes. *Am. J. Hum. Genet.*, **99**, 1245–1260.
- 688 32. Bryan, J. M., Fufa, T. D., Bharti, K., et al. (2018) Identifying core biological processes
689 distinguishing human eye tissues with precise systems-level gene expression analyses
690 and weighted correlation networks. *Hum. Mol. Genet.*, **27**, 3325–3339.
- 691 33. Bult, C. J., Blake, J. A., Smith, C. L., et al. (2019) Mouse Genome Database (MGD)
692 2019. *Nucleic Acids Res.*, **47**, D801–D806.
- 693 34. Grassmann, F., Kiel, C., Zimmermann, M. E., et al. (2017) Genetic pleiotropy between
694 age-related macular degeneration and 16 complex diseases and traits. *Genome Med.*,
695 **9**, 1–13.
- 696 35. MacArthur, J., Bowler, E., Cerezo, M., et al. (2017) The new NHGRI-EBI Catalog of
697 published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.*, **45**,
698 D896–D901.
- 699 36. Beck, T., Hastings, R. K., Gollapudi, S., et al. (2014) GWAS Central: A comprehensive
700 resource for the comparison and interrogation of genome-wide association studies. *Eur.*
701 *J. Hum. Genet.*, **22**, 949–952.
- 702 37. Canela-Xandri, O., Rawlik, K. and Tenesa, A. (2018) An atlas of genetic associations in
703 UK Biobank. *Nat. Genet.*, **50**, 1593–1599.
- 704 38. Benjamini, Yoav ; Hochberg, Y. (1995) Controlling the False Discovery Rate - a Practical
705 and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society
706 Series B-Methodological 1995.pdf. *J. R. Stat. Soc. Ser. B*.
- 707 39. Winkler, T. W., Brandl, C., Grassmann, F., et al. (2018) Investigating the modulation of
708 genetic effects on late AMD by age and sex: Lessons learned and two additional loci.
709 *PLoS One*, **13**, 1–21.
- 710 40. Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., et al. (2015) LD score regression
711 distinguishes confounding from polygenicity in genome-wide association studies. *Nat.*
712 *Genet.*, **47**, 291–295.

- 713 41. Winkler, T. W., Günther, F., Höllerer, S., et al. (2018) A joint view on genetic variants for
714 adiposity differentiates subtypes with distinct metabolic implications. *Nat. Commun.*, **9**.
- 715 42. Kuleshov, M. V, Jones, M. R., Rouillard, A. D., et al. (2016) Enrichr: a comprehensive
716 gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, W90-7.
- 717 43. Bryan, J. M., Fufa, T. D., Bharti, K., et al. (2018) Identifying core biological processes
718 distinguishing human eye tissues with precise systems-level gene expression analyses
719 and weighted correlation networks. *Hum. Mol. Genet.*, **27**, 3325–3339.
- 720 44. Fabregat, A., Jupe, S., Matthews, L., et al. (2018) The Reactome Pathway
721 Knowledgebase. *Nucleic Acids Res.*, **46**, D649–D655.
- 722 45. Hageman, G. S., Luthert, P. J., Victor Chong, N. H., et al. (2001) An integrated
723 hypothesis that considers drusen as biomarkers of immune-mediated processes at the
724 RPE-Bruch's membrane interface in aging and age-related macular degeneration. *Prog*
725 *Retin Eye Res*, **20**, 705–732.
- 726 46. Johnson, L. V, Leitner, W. P., Staples, M. K., et al. (2001) Complement activation and
727 inflammatory processes in Drusen formation and age related macular degeneration. *Exp*
728 *Eye Res*, **73**, 887–896.
- 729 47. Vogt, S. D., Curcio, C. A., Wang, L., et al. (2011) Retinal pigment epithelial expression
730 of complement regulator CD46 is altered early in the course of geographic atrophy. *Exp*
731 *Eye Res*, **93**, 413–423.
- 732 48. Seya, T. and Atkinson, J. P. (1989) Functional properties of membrane cofactor protein
733 of complement. *Biochem. J.*, **264**, 581–588.
- 734 49. Cattaneo, R. (2004) Four Viruses, Two Bacteria, and One Receptor: Membrane
735 Cofactor Protein (CD46) as Pathogens' Magnet. *J. Virol.*, **78**, 4385–4388.
- 736 50. Cardone, J., Le Friec, G. and Kemper, C. (2011) CD46 in innate and adaptive immunity:
737 an update. *Clin Exp Immunol*, **164**, 301–311.
- 738 51. Haralambieva, I. H., Ovsyannikova, I. G., Kennedy, R. B., et al. (2017) Genome-wide
739 associations of CD46 and IFI44L genetic variants with neutralizing antibody response
740 to measles vaccine. *Hum Genet*, **136**, 421–435.
- 741 52. Lewis, R. A. (1993) Oculocutaneous Albinism Type 1. *Oculocutaneous Albinism Type*
742 *1*; University of Washington, Seattle, (1993) .
- 743 53. Galván-Femenía, I., Obón-Santacana, M., Piñeyro, D., et al. (2018) Multitrait genome
744 association analysis identifies new susceptibility genes for human anthropometric
745 variation in the GCAT cohort. *J. Med. Genet.*, 765–778.
- 746 54. Stokowski, R. P., Pant, P. V. K., Dadd, T., et al. (2007) A Genomewide Association
747 Study of Skin Pigmentation in a South Asian Population. *Am. J. Hum. Genet.*, **81**, 1119–
748 1132.
- 749 55. Hu, D.-N., Simon, J. D. and Sarna, T. (2008) Role of Ocular Melanin in Ophthalmic
750 Physiology and Pathology. *Photochem. Photobiol.*, **84**, 639–644.
- 751 56. Schraermeyer, U. and Heimann, K. (1999) Current Understanding on the Role of Retinal
752 Pigment Epithelium and its Pigmentation. *Pigment Cell Res.*, **12**, 219–236.
- 753 57. Julien, S., Kociok, N., Kreppel, F., et al. (2007) Tyrosinase biosynthesis and trafficking
754 in adult human retinal pigment epithelial cells. *Graefe's Arch. Clin. Exp. Ophthalmol.*,
755 **245**, 1495–1505.
- 756 58. SCHRAERMEYER, U. (1993) Does Melanin Turnover Occur in the Eyes of Adult
757 Vertebrates? *Pigment Cell Res.*, **6**, 193–204.
- 758 59. Gao, X. R., Huang, H. and Kim, H. (2019) Genome-wide association analyses identify
759 139 loci associated with macular thickness in the UK Biobank cohort. *Hum. Mol. Genet.*,

760 **28**, 1162–1172.

761 60. Brandl, C., Brückmayer, C., Günther, F., et al. (2019) Retinal Layer Thicknesses in
762 Early Age-Related Macular Degeneration: Results From the German AugUR Study.
763 *Invest. Ophthalmol. Vis. Sci.*, **60**, 1581–1594.

764

765 TABLES

766 Table 1. Genome-wide search for early AMD association.

Rs identifier	chr:pos [hg19]	EA	OA	EAF	logOR	SE	OR	P	N cases	N controls	Known advanced AMD locus (Fritsche et al. ⁹)	Locus name
<i>Novel early AMD loci:</i>												
rs4844620	1:207980901	g	a	0.79	0.095	0.017	1.10	4.7E-08	14,031	91,179	no	CD46
rs547154	6:31910938	g	t	0.91	0.218	0.025	1.24	1.3E-18	14,027	91,137	yes	C2
rs943080	6:43826627	t	c	0.51	0.080	0.015	1.08	4.7E-08	13,220	85,747	yes	VEGFA
rs13278062	8:23082971	t	g	0.52	0.080	0.014	1.08	2.0E-08	13,644	85,908	yes	TNFRSF10A
rs5817082	16:56997349	c	ca	0.26	0.108	0.017	1.11	1.0E-10	12,599	81,863	yes	CETP
rs11569415	19:6716279	a	g	0.21	0.116	0.018	1.12	1.7E-10	13,115	83,117	yes	C3
<i>Known early AMD loci:</i>												
rs4658046	1:196670757	c	t	0.39	0.321	0.014	1.38	2.9E-114	14,034	91,201	yes	CFH
rs3750847	10:124215421	t	c	0.22	0.384	0.017	1.47	1.3E-116	14,025	91,171	yes	ARMS2/HTRA1

EA = effect allele, OA = other allele, EAF = effect allele frequency, logOR = log odds ratio, SE = standard error of logOR; OR = odds ratio, P = double GC corrected early association P value from the meta-analysis

767

768 The table shows the eight genome-wide significant ($P < 5 \times 10^{-8}$) lead variants from the early AMD meta-analysis.

769

770 **Table 2. Candidate approach to search for early AMD association.**

Rs identifier	chr:pos [hg19]	Locus (Holliday et al. ¹¹)	Holliday et al. (11)					this meta-analysis (excluding ARIC, CHS)				Known advanced AMD locus (Fritsche et al. ⁹)	
			EA	OA	EAF	OR [CI 95%]	P	EAF	OR [CI 95%]	P	Neff		
<i>Novel loci for early AMD (P < 0.05/13)</i>													
rs621313	11:88913663	TYR	A	G	0.51	0.87 [0.83;0.92]	3.5E-06	0.52	0.95 [0.93;0.98]	6.8E-04	41,661	no	
rs6857	19:45392254	PVRL2	T	C	0.15	0.81 [0.74;0.88]	1.4E-06	0.16	0.92 [0.89;0.96]	7.1E-05	38,938	yes	
rs2075650	19:45395619	APOE/TOMM40	A	G	0.86	1.23 [1.13;1.34]	1.1E-06	0.86	1.08 [1.04;1.13]	2.6E-04	40,405	yes	
<i>No significant association (P ≥ 0.05/13):</i>													
rs16851585	1:177568799	-	C	G	0.92	0.77 [0.69;0.86]	5.0E-06	0.89	1.04 [0.99;1.09]	0.099	42,119	no	
rs6721654	2:121301911	GLI2,INHBB	T	C	0.08	1.26 [1.14;1.4]	6.5E-06	0.08	1.01 [0.96;1.07]	0.65	41,718	no	
rs17586843	4:116924184	-	T	C	0.78	1.18 [1.1;1.27]	1.5E-06	0.78	1.02 [0.98;1.05]	0.31	42,119	no	
rs7750345	6:106260128	-	A	G	0.75	1.16 [1.09;1.24]	6.8E-06	0.74	1.02 [0.98;1.05]	0.32	42,119	no	
rs2049622	7:42176282	GLI3	A	G	0.49	0.87 [0.83;0.93]	8.9E-06	0.52	0.99 [0.96;1.02]	0.41	42,060	no	
rs11986011	8:127332657	FAM84B	T	C	0.02	2.5 [1.68;3.71]	5.0E-06	-	-	-	-	no	
rs6480975	10:54574996	MBL2	C	G	0.84	1.21 [1.12;1.32]	2.8E-06	0.85	0.99 [0.95;1.03]	0.63	40,651	no	
rs4293143	11:82821382	PCF11,RAB30	T	G	0.69	0.85 [0.79;0.91]	7.8E-06	0.70	0.99 [0.96;1.02]	0.50	42,119	no	
rs9646096	13:38065446	POSTN,TRPC4	A	C	0.95	0.74 [0.65;0.84]	6.0E-06	0.96	0.99 [0.92;1.06]	0.76	39,782	no	
rs10406174	19:3944240	ITGB1BP3,DAPK3	A	G	0.11	1.24 [1.13;1.36]	5.6E-06	0.11	1.00 [0.93;1.07]	0.92	17,936	yes	

771 EA = effect allele, OA = other allele, EAF = effect allele frequency, OR = odds ratio, CI = confidence interval, P = P value from Holliday et al or Double GC corrected early association P value from this meta-analysis, Neff =

772 effective sample size

773 The table shows results for the 13 lead variants reported as suggestive for early AMD by Holliday et al. (11) (effective sample size = 13,631) for their
774 association with early AMD in our data set (P<0.05/13= 0.0038, tested at Bonferroni-corrected significance level, effective sample size up to 42,119),
775 excluding the variant in the CD46 locus that we have already identified with genome-wide significance (see Table 1). The ARIC and CHS studies
776 were excluded from our meta-analysis data to avoid overlap with the data by Holliday et al. (11).

777 **Table 3. Summary of *in silico* follow-up and gene prioritization score (GPS).**

Locus	Candidate gene	Chr	Pos-Start	Pos-End	Number of variants in 95% credible set	GPS	Annotation for variants in 95% credible set				Biology of the gene		
							Protein Altering	NMD	Altered splicing	eQTL *	Expressed in Eye tissue §	MGI Mouse eye phenotype	OMIM Human eye phenotype
CD46	<i>CD46</i>	1	207925382	207968861	11	4	0	0	1	1	1	1	0
CD46	<i>CR1L</i>	1	207818457	207897036	1	1	0	0	0	0	0	1	0
CD46	<i>PLXNA2</i>	1	208195587	208417665	0	2	0	0	0	0	1	1	0
CD46	<i>CR1</i>	1	207669472	207815110	0	2	0	0	0	0	1	1	0
CD46	<i>LOC148696</i>	1	207991723	207995941	1	0	0	0	0	NA	NA	0	0
CD46	<i>CD34</i>	1	208059882	208084683	0	1	0	0	0	0	1	0	0
CD46	<i>CD55</i>	1	207494816	207534311	0	1	0	0	0	0	1	0	0
CD46	<i>CR2</i>	1	207627644	207663240	0	0	0	0	0	0	0	0	0
CD46	<i>MIR29B2</i>	1	207975787	207975868	0	0	0	0	0	NA	NA	0	0
CD46	<i>MIR29C</i>	1	207975196	207975284	0	0	0	0	0	NA	NA	0	0
TYR	<i>TYR</i>	11	88911039	89028927	39	4	1	0	0	0	1	1	1
TYR	<i>NOX4</i>	11	89057521	89322779	1	3	0	1	1	0	1	0	0
TYR	<i>GRM5</i>	11	88237743	88796846	109	1	0	0	0	0	1	0	0
TYR	<i>FOLH1B</i>	11	89392464	89431886	0	0	0	0	0	0	0	0	0
TYR	<i>GRM5-AS1</i>	11	88237743	88257222	0	0	0	0	0	0	NA	0	0
TYR	<i>TRIM49</i>	11	89530822	89541743	0	0	0	0	0	0	NA	0	0
TYR	<i>TRIM77</i>	11	89443466	89451040	0	0	0	0	0	0	NA	0	0

* Variants in 95% credible set are a local expression quantitative trait locus for this gene in retina (EyeGEx) or any tissue included in the GTEx database (cis for genes in locus); § Expression in Eye Integration data; NMD = nonsense-mediated mRNA decay; OMIM = Online Mendelian Inheritance in Man (<https://www.omim.org/>); NA = data not available; The gene start and end positions were extracted from the hg19 gene range list from <http://www.cog-genomics.org/plink/1.9/resources>.

778

779 The table summarizes statistical and functional evidence for 10 and seven candidate genes of the novel early AMD loci on chromosome 1 and
 780 chromosome 11, respectively. Detailed results on the individual statistical and functional analyses are shown in **Tables S3-S9**. For the GPS, the
 781 sum of cell entries for “annotation” and “biology” was computed per row.

782 **Table 4. Enriched pathways.**

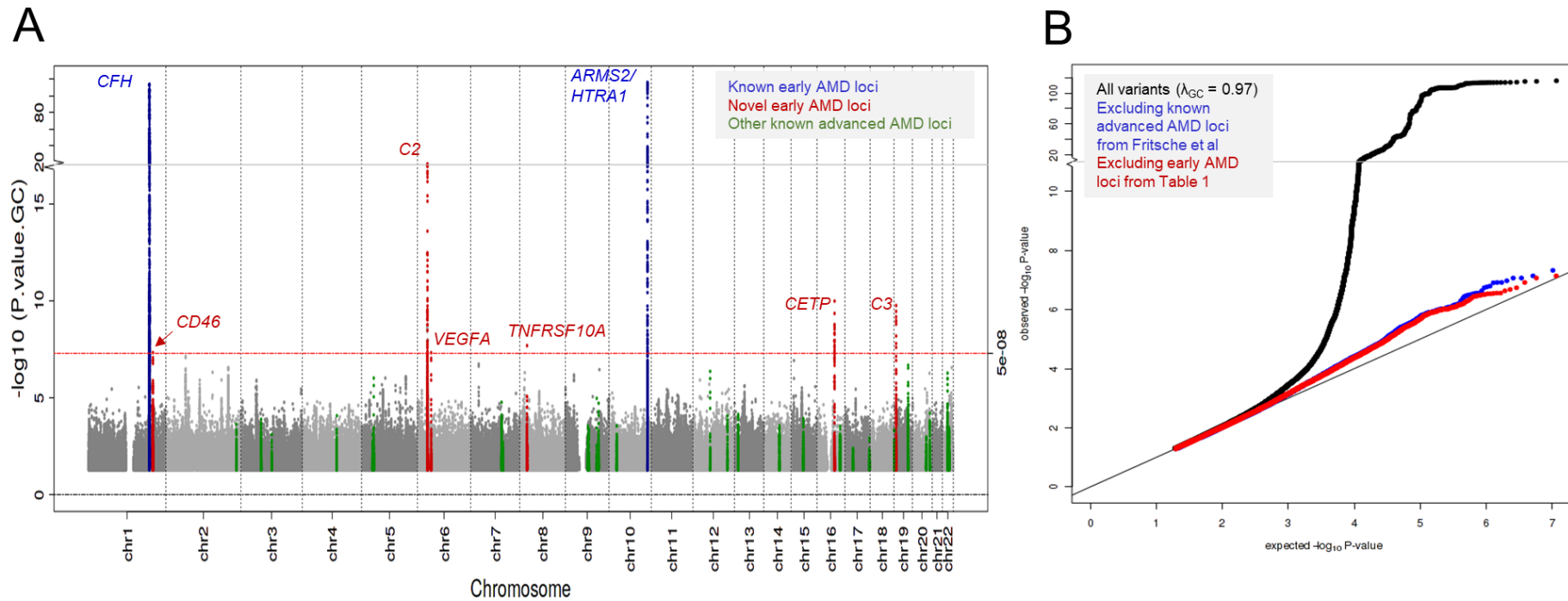
Gene group	Reactome pathway	#Genes in gene set	#AMD loci in gene set	Raw P	Corrected P	Genes contributing to enrichment
Effects on early and advanced AMD	Regulation of Complement cascade (R-HSA-977606)	26	5	7.8×10^{-10}	1.2×10^{-6}	<i>C3;CFH;C9;CFI;CFB</i>
	Lipoprotein metabolism (R-HSA-174824)	34	4	3.5×10^{-7}	1.8×10^{-4}	<i>ABCA1;CETP;LIPC;APOE</i>
	Complement cascade (R-HSA-166658)	80	5	2.7×10^{-7}	2.0×10^{-4}	<i>C3;CFH;C9;CFI;CFB</i>
	HDL-mediated lipid transport (R-HSA-194223)	19	3	4.7×10^{-6}	1.8×10^{-3}	<i>ABCA1;CETP;APOE</i>
	Lipid digestion, mobilization, and transport (R-HSA-73923)	71	4	7.0×10^{-6}	2.2×10^{-3}	<i>ABCA1;CETP;LIPC;APOE</i>
Activation of C3 and C5 (R-HSA-174577)	6	2	4.4×10^{-5}	0.011	<i>C3;CFB</i>	
no effects on early AMD	Assembly of collagen fibrils and other multimeric structures (R-HSA-2022090)	54	3	1.5×10^{-6}	2.4×10^{-3}	<i>COL15A1;COL8A1;MMP9</i>
	Collagen formation (R-HSA-1474290)	85	3	6.1×10^{-6}	3.1×10^{-3}	<i>COL15A1;COL8A1;MMP9</i>
	Extracellular matrix organization (R-HSA-1474244)	283	4	4.7×10^{-6}	3.6×10^{-3}	<i>VTN;COL15A1;COL8A1;MMP9</i>

783

784 The table shows enriched pathways for highest prioritized genes (from Fritsche et al. 2016 without modifications) in the 25 late AMD loci with early
785 AMD effects (35 genes) versus the 8 loci with no effect on early AMD (9 genes). Pathways with significant corrected P-value ($P_{\text{corr}} < 0.05$) for each
786 gene group from EnrichR querying human Reactome database 2016 are shown.

787 **FIGURES**

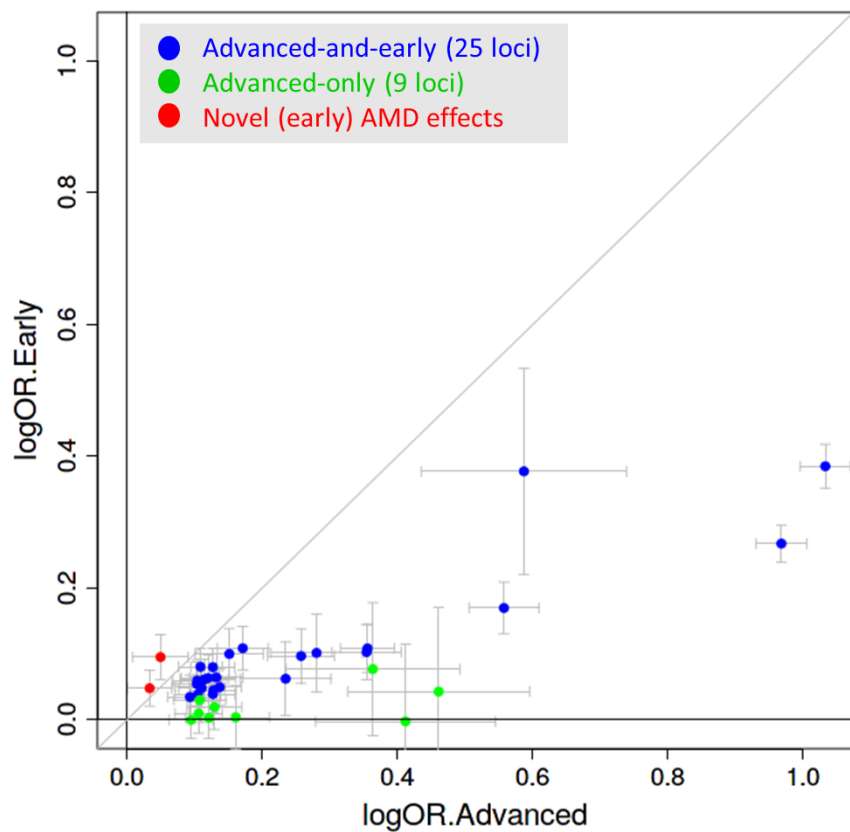
788 **Fig 1. Early AMD meta-analysis.** Shown are the association P-values of the meta-analysis for early AMD by their position on the genome (A,
789 Manhattan plot) as well as their distribution (B, QQ plot). In A, color indicates whether the locus was previously identified by Holliday et al (11) (blue),
790 novel for early AMD (red), or among the other advanced AMD loci identified by Fritsche et al (9) (green).



791

792

793 **Fig 2. Advanced vs early AMD effect sizes.** Shown are advanced AMD effect sizes contrasted to early AMD effect sizes (effect sizes as log odds
794 ratios) for the 34 known advanced AMD variants (9) (blue or green for $P_{\text{early}} < 0.05$ or $P_{\text{early}} \geq 0.05$, respectively) and for the two novel (early) AMD
795 variants (red, near *CD46*, *TYR*). Detailed results are shown in **Table S13**.



796

797 **Fig 3. Etiological models and the respective expected association of a variant with early and advanced AMD.**

		<u>Expected direction of effect in association data</u>	
		Early AMD (vs. healthy)	Advanced AMD (vs. healthy)
Model 1: „Trigger for early and advanced AMD“	<pre> graph LR healthy --> early early --> advanced </pre>	+	+
Model 2: „Trigger for advanced AMD“	<pre> graph LR healthy --> advanced early --> advanced healthy -.-> advanced </pre>	0	+
Model 3: „Trigger for progression“	<pre> graph LR healthy early --> advanced </pre>	-	+
Model 4: „Trigger for early AMD“	<pre> graph LR healthy --> early advanced </pre>	+	0

798

799 **LIST OF ABBREVIATIONS**

800 3CC: Three Continent Consortium

801 AMD: Age-related macular degeneration

802 ARIC: The Atherosclerosis Risk in Communities Study

803 AugUR: Age-related diseases: Understanding Genetic and non-genetic influences - a study at
804 the University of Regensburg

805 CHS: Cardiovascular Health Study

806 eQTL: Expression quantitative trait locus

807 FDR: False-discovery-rate

808 GCTA: Genome-wide Complex Trait Analysis

809 GC: Genomic control

810 GHS: Gutenberg Health Study

811 GPS: Gene Prioritization Score

812 GWAS: Genome-wide association study

813 HRC: Haplotype Reference Consortium

814 IAMDGC: International AMD Genomics Consortium

815 KORA: KOoperative Gesundheitsforschung in der Region Augsburg

816 LIFE: Leipzig Research Centre for Civilization Based Diseases - LIFE Adult population-based
817 study, city of Leipzig, Germany

818 LOO: leave-one-out

819 MAC: minor allele count

- 820 MGI: Mouse Genome Informatics
- 821 NICOLA: Northern Ireland Cohort for Longitudinal Study of Ageing
- 822 NMD: nonsense-mediated mRNA decay
- 823 OCT: optical coherence tomography
- 824 OMIM: Online Mendelian Inheritance in Man
- 825 QC: quality control
- 826 RPE: retinal pigment epithelium
- 827 UKBB: UK Biobank
- 828 VEP: Variant effect predictor
- 829 WHI: Women's Health Initiative

830

831 **DECLARATIONS**

832 **Ethics approval and consent to participate**

833 The Institutional Review Board (IRB) of the University of Utah was the umbrella IRB for all
834 other studies contributing data to the International Age-related Macular Degeneration
835 Genomics Consortium (IAMDGC), except for the Beaver Dam Eye Study (BDES). The
836 University of Utah approved and certified each individual study ethic committee's conduct for
837 the data used in this study. Data provided by BDES was approved by the IRB of the University
838 of Wisconsin. Local ethics approval for data access to the studies deposited in dbGAP (WHI,
839 ARIC and CHS) was granted by the IRB of the University of Regensburg. For all other studies,
840 study participants obtained informed consent and local ethics committees approved the study
841 protocols.

842

843

844 **Consent for publication**

845 Not applicable.

846 **Availability of data and materials**

847 The genome-wide meta-analysis summary statistics are available for download from
848 <http://genepi-regensburg.de>.

849 **Competing interests**

850 M.S. receives funding from Pfizer Inc. for a project not related to this research. Retinal grading
851 of the NICOLA study was supported by Novartis (for R.E.H) and Bayer (for Usha Chakravarthy,
852 not a co-author). A.K.S. received financial and research support by Heidelberg Engineering,
853 Novartis, Bayer Vital and Allergen without a link to the content of this work. I.M.H. received
854 funding from Roche Diagnostics for a project not related to this research. None of the other
855 authors have any conflicts of interest.

856 **Funding**

857 The International AMD Genomics Consortium (IAMDGC) is supported by a grant from NIH
858 (R01 EY022310). Genotyping was supported by a contract (HHSN268201200008I) to the
859 Center for Inherited Disease Research (<http://amdgenetics.org/>). In-depth analyses to estimate
860 genetic effects in the IAMDGC data was supported by DFG HE 3690/5-1 to I.M.H. The AugUR
861 study was supported by grants from the German Federal Ministry of Education and Research
862 (BMBF 01ER1206, BMBF 01ER1507 to I.M.H.) and the University of Regensburg. The KORA
863 study was initiated and financed by the Helmholtz Zentrum München – German Research
864 Center for Environmental Health, which is funded by the German Federal Ministry of Education
865 and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was
866 supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-
867 Universität, as part of LMUinnovativ. This publication is supported by the Leipzig Research
868 Centre for Civilization Diseases (LIFE), an organizational unit affiliated to the Medical Faculty
869 of Leipzig University. LIFE is funded by means of the European Union, by the European

870 Regional Development Fund (ERDF) and by funds of the Free State of Saxony within the
871 framework of the excellence initiative (project numbers: 713-241202, 14505/2470,
872 14575/2470). Franziska G. Rauscher (F.G.R.) holds a grant from the German Federal Ministry
873 of Education and Research: i:DSem - Integrative data semantics in systems medicine
874 (031L0026). Tobias Elze (T.E.) is funded by the Lions Foundation, Grimshaw-Gudewicz
875 Foundation, Research to Prevent Blindness, BrightFocus Foundation, Alice Adler Fellowship,
876 NEI R21EY030142, NEI R21EY030631, NEI R01EY030575, and NEI Core Grant
877 P30EYE003790. The Gutenberg Health Study is funded through the government of Rhineland-
878 Palatinate („Stiftung Rheinland-Pfalz für Innovation“, contract AZ 961-386261/733), the
879 research programs “Wissen schafft Zukunft” and “Center for Translational Vascular Biology
880 (CTVB)” of the Johannes Gutenberg-University of Mainz, and its contract with Boehringer
881 Ingelheim and PHILIPS Medical Systems, including an unrestricted grant for the Gutenberg
882 Health Study. Alexander K Schuster (A.K.S.) holds the professorship for ophthalmic healthcare
883 research endowed by „Stiftung Auge“ and financed by „Deutsche Ophthalmologische
884 Gesellschaft“ and „Berufsverband der Augenärzte Deutschland e.V.“. The Atherosclerosis Risk
885 in Communities study has been funded in whole or in part with Federal funds from the National
886 Heart, Lung, and Blood Institute, National Institute of Health, Department of Health and Human
887 Services, under contract numbers (HHSN268201700001I, HHSN268201700002I,
888 HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I). The research
889 reported in this article was supported by contract numbers N01-HC-85079, N01-HC-85080,
890 N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-
891 85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133,
892 N01-HC-85239 and HHSN268201200036C; grant numbers U01 HL080295 from the National
893 Heart, Lung, and Blood Institute and R01 AG-023629 from the National Institute on Aging, with
894 additional contribution from the National Institute of Neurological Disorders and Stroke. The
895 WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of
896 Health, U.S. Department of Health and Human Services through contracts N01WH22110,
897 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-

898 32, and 44221. This study has been conducted using the UK Biobank resource under
899 Application Number 33999. The UK Biobank was established by the Wellcome Trust medical
900 charity, Medical Research Council, Department of Health, Scottish Government and the
901 Northwest Regional Development Agency. It has also had funding from the Welsh Assembly
902 Government, British Heart Foundation and Diabetes UK. The analyses were supported by
903 German Research Foundation (DFG HE-3690/5-1 to I.M.H.) and by the National Institutes of
904 Health (NIH R01 EY RES 511967 to I.M.H.). Felix Grassmann (F.G.) was a Leopoldina
905 Postdoctoral Fellow (Grant No. LPDS 2018-06) funded by the Academy of Sciences
906 Leopoldina. The position of Tobias Strunz (T.S.) is financed by the Helmut-Ecker-Foundation
907 (# 05/17 to B.H.F.W.) and of Christina Kiel (C.K.) by a grant from the German Research
908 Foundation to F.G. and B.H.F.W. (GR 5065/1-1).

909

910 **Acknowledgments**

911 Extended acknowledgments are shown in **Table S1**.

912

913 **Authors' contributions**

914 TWW, FGr, CB, CK, FGü, IMH, KJS and BHFw designed the study and wrote the manuscript.
915 TWW and FGr conducted the meta-analysis. FGü applied the automated grading of UK
916 Biobank fundus images. TWW and CK conducted the PheWAS. TWW, LW, MEZ and KJS
917 conducted analysis for the gene priority scoring. KS conducted the pathway analyses. CAK,
918 AP and AKS contributed data from the GHS study. MMN and AP contributed data from the
919 KORA study. FGR, TE, KH, and MS contributed data from the LIFE-Adult study. MCG, AJMK,
920 NQ and REH contributed data from the NICOLA study. IMH, KJS and BHFw supervised the
921 study. All authors read and approved the final manuscript.