Phenome-wide Mendelian randomisation analysis identifies causal factors for age-related macular degeneration.

Thomas H Julian[1,2], Johnathan Cooper-Knock[3], Stuart MacGregor[4], Hui Guo[5] Tariq Aslam[6], Eleanor Sanderson[7], Graeme Black*[1,9], Panagiotis I Sergouniotis*[1,2,8,9],

1. Division of Evolution, Infection and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK.
2. Manchester Royal Eye Hospital, Manchester University NHS Foundation Trust, Manchester, UK.
3. Department of Neuroscience, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK.
4. Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
5. Centre for Biostatistics, Division of Population Health, Health Services Research & Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK.
6. Division of Pharmacy and Optometry, Faculty of Biology, Medicine and Health, School of Health Sciences, The University of Manchester, Manchester, UK.
7. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom.
8. European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, UK.
9. Manchester Centre for Genomic Medicine, Saint Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK.

***Corresponding authors:** Panagiotis Sergouniotis & Graeme Black

ABSTRACT:

Introduction:

Age-related macular degeneration (AMD) is the leading cause of blindness in the industrialised world and is projected to affect 288 million people worldwide by 2040. Aiming to identify causal factors for this common disorder, we designed and applied a phenome-wide Mendelian randomisation (MR) approach to identify therapeutic targets and avenues for future research.

Methods:

We evaluated the effect of 4,591 exposure traits on early AMD using univariable MR. Statistically significant results were explored further using validation in an advanced AMD cohort, MR Bayesian model averaging (MR BMA) and multivariable MR.

Results:

44 traits were found to be putatively causal for early AMD in univariable analysis. MR BMA of lipid traits suggested a causal role for serum sphingomyelin (marginal inclusion probability=0.76, model averaged causal effect=0.29). Univariable MR analysis supported roles for complement and immune cell traits. Serum proteins were found to have significant relationships with AMD including S100-A5 (Odds ratio (OR)= 1.07, p = 6.80E-06), cathepsin F (OR=1.10, p = 7.16E-05) and serine palmitoyltransferase 2 (OR= 0.86, p value=1.00E-03).

Conclusions:

The results of this study support several putative causal factors in AMD and highlight avenues for future clinical research.

52 **INTRODUCTION:**

53

54 Age-related macular degeneration (AMD) is a progressive multifactorial disorder that affects individuals
55 who are ≥50 years old and is caused by the complex interplay of multiple genetic and environmental risk
56 factors. Genome-wide association studies (GWAS) have identified AMD-implicated variants in at least 69
57 loci including important risk alleles in the 1q32 and 10q26 genomic regions (corresponding to the *CFH* -
58 complement factor H – and *ARMS2/HTRA1* locus respectively) (Fritsche et al., 2016; Winkler et al., 2020).
59 Other key risk factors include age, smoking, alcohol consumption and low dietary intake of antioxidants
60 (carotenoids, zinc) (Chakravarthy et al., 2010).

61

62 AMD can be categorised according to severity (early, intermediate, or advanced) or based on the presence
63 of neovascularization (neovascular or non-neovascular). Advanced AMD results in loss of central vision,
64 often leading to severe visual impairment (Fleckenstein et al., 2021). Notably, AMD is a major cause of
65 blindness in the elderly population and represents a substantial global burden that is expected to continue
66 to grow into the future as an ageing population expands worldwide.

67

68 Mendelian randomization (MR) is a statistical approach that uses genetic variation to look for causal
69 relationships between *exposures* (such as body mass index) and medically-relevant *outcomes* (such as
70 risk of a specific disease) (Julian et al., 2021). MR is increasingly being used as it can overcome a major
71 limitation of observational studies: unmeasured confounding(Sanderson et al., 2022). To minimise issues
72 with certain types of confounding and to support causal inference statements, MR uses genetic variation
73 as an "*instrument*" (i.e. as a variable that is associated with the exposure, independent of the confounders,
74 and is not associated with the outcome, other than through the exposure). Its principles are based on
75 Mendel's laws of segregation and independent assortment, which state that offspring inherit alleles
76 randomly from their parents and randomly with respect to other locations in the genome (Julian et al., 2021).
77 As a result, genetic variants that are related to an exposure of interest can be used to proxy the part of the
78 exposure variable that is independent of possible confounding influences (e.g., from the environment and
79 other traits). It is noted that analogies have been drawn between MR and randomised controlled trials with

these two approaches considered proximal in terms of hierarchy of evidence.(Julian et al., 2021) To date, the use of MR approaches in the context of AMD has been limited, although these methods have been successfully implemented to explore the relationship between AMD and a small number of traits including lipids, thyroid function, CRP and complement factors.(Cipriani et al., 2021; Han et al., 2021, 2020b; Li et al., 2022; Zuber et al., 2020)

In this study, we have developed a systematic, broad ('phenome-wide') MR analysis methodology which serves to investigate the relationship between early AMD and several thousand exposure variables. Using this approach, we report a set of traits that are robustly associated with genetic liability to AMD.

METHODS:

Data sources

Outcome data

Two AMD phenotypes were used as outcome measures in this study. The first was early AMD. The GWAS summary statistics for this phenotype were taken from a meta-analysis by Winkler et al(Winkler et al., 2020). This study focused on populations of European ancestries and used data from the ARIC, AugUR, CHS, GHS, IAMDGC, KORA S4, LIFE-Adult NICOLA, UKBB and WHI studies (14,034 early AMD cases and 91,214 controls overall). A full description of how these studies classified participants as "early AMD" can be found in Winkler et al(Winkler et al., 2020). The second phenotype that we studied was advanced AMD. For this trait, we drew on GWAS summary statistics from a multiple trait analysis of GWAS (MTAG) study by Han et al. (Han et al., 2020a). This meta-analysis also focused on individuals of European ancestries and derived data from the IAMDGC 2013 (Fritsche et al., 2013) (17,181 cases and 60,074 controls) and 2016 (Fritsche et al., 2016) (16,144 advanced AMD cases and 17,832 controls) studies as well as the GERA study (Kvale et al., 2015) (4,017 cases and 14,984 controls). The summary statistics produced in this GWAS are primarily reflective of advanced AMD, but the GERA cohort included both advanced and intermediate AMD cases. Notably, the MTAG approach is able to leverage the high genetic correlation between the input phenotypes to detect genetic associations for advanced AMD.

108

109 **Exposure data**

110 In this study we performed a phenome-wide screen to make causal inferences on the role of an extensive
111 range of traits in early and advanced AMD. To achieve this, we used both published and unpublished
112 GWAS data from the IEU open GWAS database; these were accessible via the TwoSampleMR
113 programme in R(Hemani et al., 2018). We included all European GWAS within the database but removed
114 imaging phenotypes and eQTL data. In the early AMD analysis, we also excluded studies from the "ukb"
115 and "met-d" batches as these are studies entirely derived from UK Biobank data and therefore had
116 extensive population overlap with the early AMD GWAS (Sudlow et al., 2015). In the advanced AMD
117 analysis, "ukb" and "met-d" batches were included. The early AMD analysis was conducted on
118 30/12/2021 and a total of 10,979 traits were considered for analysis. The advanced AMD analysis was
119 conducted on 08/01/2022. On 26/01/2022 we added newly published 'finn-b' (n=2803) traits to our
120 analysis in place of outdated finn-a traits (n=1489). As it is not possible to manually inspect the degree of
121 population overlap for all traits prior to conducting the analysis, the degree of overlap for all significant
122 traits was inspected after the analysis.

123

124 **Instrument selection**

125 In this study, we used a statistically driven approach to instrumental variable selection. In general, an
126 arbitrary p-value threshold is set for the identification of appropriate single-nucleotide polymorphisms
127 (SNPs) that are used as instrumental variables (referred to thereafter as instruments). A conventional value
128 for selection of instruments is p=5E-08, but this can, in some cases, be problematic. For example, when
129 the number of instruments exceeding this threshold is small, the analysis can be underpowered or, in certain
130 cases of unbiased screens, the results can be inflated.(Boddy et al., 2022) With this in mind, we have set
131 the p-value for instrument selection for each trait to the level where >5 instruments are available for each
132 analysis. More specifically, for each trait, the analysis would first be conducted with a p-value for inclusion
133 of 5E-8 and would sequentially increase by a factor of 10 each time until >5 eligible instruments are
134 identified to a predefined maximum of p=5E-05. Final range of p values for inclusion was 5E-06 – 5E-08.

135 **Proxies**

136 Where an exposure instrument was not present in the outcome dataset, we sought to identify a suitable
137 proxy(Hartwig et al., 2016). In the early AMD analysis, we were able to achieve this by using the
138 TwoSampleMR software with a linkage disequilibrium R^2 value of 0.9 or above for identification of
139 SNPs(Purcell et al., 2007). For the advanced AMD phenotype we used data which was not derived from
140 the TwoSampleMR database and therefore the ensembl server was used to derive proxies(Cunningham
141 et al., 2022; Hemani et al., 2018).
142
143 **Clumping**
144 SNPs were clumped using a linkage disequilibrium R^2 value of 0.001 and a genetic distance cut-off of
145 10,000 kilo-bases.
146
147 **Harmonisation**
148 The effects of instruments on outcomes and exposures were harmonised in order to ensure that the beta
149 values (i.e. the regression analysis estimates of effect size) were signed with respect to the same alleles
150 (Hartwig et al., 2016). For palindromic alleles (i.e., alleles that are the same on the forward as on the reverse
151 strand), those with minor allele frequency (MAF) > 0.42 were omitted from the analysis in order to reduce
152 the risk of errors due to strand issues.
153
154 **Removal of pleiotropic genetic variants and outliers**
155 In addition to using a range of robust measures of scale (which quantify the statistical dispersion in a sample
156 of numerical data while resisting outliers) and quality control measures (as detailed below), we endeavoured
157 to remove pleiotropic instruments and outliers from the analysis. We operated this using a statistical
158 approach in which instruments which were more significant for the outcome than the exposure were
159 removed(Hemani et al., 2017). Radial MR, a simulation-based approach that detects outlying instruments
160 was also utilised(Bowden et al., 2018).
161
162 **Causal inference**

MR estimation was primarily performed using a multiplicative random effects (MRE) inverse variance weighted (IVW) analysis. MRE IVW was selected over a fixed effects (FE) approach as it allows inclusion of heterogeneous instruments, certain to occur within the breadth of this screen(Burgess et al., 2019). A range of robust measures of scale were utilised to ensure the accuracy of the results and to cover for a range of violations of the MR assumptions(Burgess et al., 2019). The robust measures selected were the weighted median(Bowden et al., 2016a), Egger(Burgess and Thompson, 2017a), weighted mode(Hartwig et al., 2017) and radial MR with modified second order weights(Bowden et al., 2018).

Further quality control

Instrument strength was determined using the F statistic (which tests the association between the instruments and the exposure) (Burgess et al., 2011). F statistics were calculated against the final included set of instruments. A mean F statistic >10 was considered sufficiently strong.

The Cochran's Q test was performed for each analysis. Cochran's Q is a measure of heterogeneity among causal estimates and serves as an indicator of the presence of horizontal pleiotropy (which occurs when an instrument exhibits effects on the outcome through pathways other than the exposure)(Bowden and Holmes, 2019). It is important to note that a heterogeneous instrument is not necessarily invalid, but rather calls for a primary assessment with an MRE IVW rather than an FE approach; this has been conducted as standard throughout our analysis.

The MR-Egger intercept test was used to detect horizontal pleiotropy. When this occurs, the Egger regression is robust to horizontal pleiotropy under the assumption that that pleiotropy is uncorrelated with the association between the SNP and the exposure(Burgess and Thompson, 2017b). In this study, assuming there is not demonstrable horizontal pleiotropy we do not utilise the Egger regression to determine causal effect given its low power with a small number of SNPs (Bowden et al., 2015).

189 The I^2 statistic was calculated as a measure of heterogeneity between variant specific causal estimates. A
190 I^2 <0.9 indicates that Egger is more likely to be biased towards the null through violation of the `NO
191 Measurement Error' (NOME) assumption (Bowden et al., 2016b) .
192
193 Leave-one-out (LOO) cross-validation was performed for every analysis to determine if any particular SNP
194 was driving the significance of the causal estimates.
195
196 **Management of duplicate traits/GWAS**
197 As the GWAS database used contains multiple different GWAS studies for the certain traits, some
198 exposures were analysed on multiple occasions. Where this occurred, we considered the largest sample
199 size study to be the primary analysis. Where there were duplicate studies in the same population, we used
200 the study with the largest F statistic.
201
202 **Identification of significant results**
203 Before considering an MR result to be significant, the results of a range of causal inference and quality
204 control tests should be considered. Notably, it is not necessary for a study to find significance in all
205 measures to determine a true causal relationship. MR is a low power study type and as such an overly
206 conservative approach to multiple testing can be excessive(Burgess et al., 2019). However, in the context
207 of the present study we have chosen to consider the results of our early AMD phenome-wide screen as
208 significant only if they remained: significant after false discovery rate correction (FDR) in the MRE IVW;
209 nominally significant in weighted mode and weighted median; and nominally significant throughout the
210 leave-one-out analysis (IVW-MRE)(Benjamini and Hochberg, 1995). This conservative approach was
211 selected as a large number of phenotypes was considered and because we wanted to focus on high
212 confidence signals.
213
214 Where we identified causal traits for early AMD, we have additionally explored their relationship with
215 advanced AMD so as to ascertain whether these traits are also implicated in more severe disease. These
216 two AMD classifications are phenotypically distinct but are generally part of the same disease spectrum.

217 For this reason, where traits failed to replicate as causal factors in the advanced AMD dataset this would
218 not necessarily imply that they are not causal for early AMD. However, significance in both AMD traits
219 provides support for the detected causal links and evidence that a factor plays a role across the spectrum
220 of disease.
221
222 **Multivariable mendelian randomisation**
223 Multivariable MR was performed in circumstances where we aimed to estimate the effect of >1 closely
224 related or potentially confounding exposure traits.(Sanderson et al., 2019) P-values for the inclusion of
225 instruments for the exposures of interest were optimised to obtain sufficiently high (>10) conditional F-
226 statistics for reliable analysis (Sanderson et al., 2021). With this in mind, selection for exposures began at
227 p>5E-08 and where a trait's instruments had a conditional F-statistic <10 the p-value for selection was
228 reduced in an automated manner by factor of 10 until an F-statistic >10 was obtained in the final analysis.
229 We utilised the same clumping procedure as in our univariable MR analysis. Adjusted Cochran's Q statistics
230 were calculated, with a p-value of <0.05 indicating significant heterogeneity. Where the Cochran's Q
231 statistic indicated heterogeneity, a Q-statistic minimisation procedure was used to evaluate the causal
232 relationship, with testing both assuming high levels (0.9) and low levels (0.1) of phenotypic correlation
233 (Sanderson et al., 2021).
234
235 **Two-sample multivariable Mendelian randomisation approach based on Bayesian model averaging**
236 **(MR-BMA)**
237
238 Multivariable MR can be used to obtain effect estimates for a few (potentially related) traits, however it
239 cannot be directly applied when there are many traits to be considered. MR-BMA is a Bayesian approach
240 developed by Zuber et al *(Zuber et al., 2020)*. Their approach searches over sets of potential risk factors to
241 determine which are most likely to be causal risk factors.
242
243 Notably, Zuber *et al*. have previously performed an excellent analysis which considered the role of lipids
244 against an older AMD GWAS. This study served as the proof of the MR-BMA method and also

245 demonstrated that several lipid traits have causal roles in AMD. However, this analysis has two potential
246 limitations. First, it downweighed fatty acid traits through limiting composite traits for SNP identification to
247 HDL, LDL and triglycerides. Second, numerous lipids with a potential role in AMD were not included in the
248 analysis. For these reasons, we chose to conduct a more inclusive analysis with a slightly altered approach.
249 The following study design modifications were made compared to Zuber et al.(Zuber et al., 2020):
250 (1) Fatty acids were included as a composite trait (utilising GWAS data for serum fatty acids derived from
251 Nightingale Health 2020 listed in TwoSampleMR (Hemani et al., 2018)).
252 (2) All lipid and fatty acid measures in the Kettunen *et al* GWAS were considered as potential causal traits
253 (n=102 traits)(Kettunen et al., 2016);
254 (3) We utilised a more recent, early AMD GWAS(Winkler et al., 2020);
255 (4) In general, multivariable MR (of any sort) cannot produce reliable results where the included traits are
256 ≥ 0.99 correlated with respect to the included instruments. For this reason, where two traits were highly
257 correlated, one was removed at random rather than manually selecting traits in a manner which risks
258 selection bias.
259
260 Additionally, we performed MR-BMA for immune cell and complement phenotypes. In this analysis
261 instruments were obtained at genome wide significance (5E-08) for every included exposure in the model
262 given that composite traits were not available and/or applicable. For the immune cell MR-BMA, all immune
263 traits in the Orru et al. GWAS present in TwoSampleMR were used as exposures(Orrù et al., 2020). In the
264 complement analysis, we used all complement traits available in both Sun et al. and Suhre et al. with the
265 exception of complement subfractions (Suhre et al., 2017; Sun et al., 2018).
266
267 For our MR-BMA analysis, the prior probability was set to 0.1 and the prior variance was set to 0.25. We
268 operated our stochastic search with 10,000 iterations and calculated empirical p-values with 100,000
269 permutations.
270
271 **Presentation of effect sizes:**

We aim to present effect sizes to enable readers to appraise the impact of putative risk factors. In univariable MR, these are presented as OR per 1SD change in the exposure for continuous traits and beta values for binary exposure traits. We present this way because odds ratios are uninformative for binary exposure traits(Burgess and Labrecque, 2018). Similarly, beta values are not presented by all MR authorities where an exposure variable is binary in nature and it's important to recognise that these are indicative only. Multivariable MR effect sizes are presented as beta estimates irrespective of the nature of the exposure variable, given that the role of Multivariable MR within this study is purely to identify confounders.

MR-BMA effect sizes are presented in the form of MACE. The MACE is a conservative estimate of the direct causal effect of an exposure on an outcome averaged across models. It should be noted that the primary function of MR-BMA is to highlight the probable causal trait amongst a number of candidate causal risk factors, and the authors of this study warn that the causal effects should be used to interpret direction of effect but not necessarily as an absolute guide to magnitude (Zuber et al., 2020).

Software

R version 4.1.0.

TwoSampleMR version 0.5.6.

RadialMR version 1.0.

MVMR version 0.3

MR BMA code was sourced from https://github.com/verena-zuber/demo_AMD.

RESULTS:

Overview

Focusing on early AMD, we applied univariable MR analysis to a broad range of traits. Following careful quality control, we sought to replicate significant results in an advanced AMD dataset and conducted further analyses using multivariable MR and MR-BMA.

300

301 Overall, 4,591 traits were eligible for analysis. Among these, 44 were found to be putatively causal for early
302 AMD **(Table 1 and Supplementary Data 1 & 2)**. Most of these causal traits were serum lipoprotein
303 concentration and compositional measures (n=29). Other significant traits identified included immune cell
304 phenotypes (n=5), serum proteins (n=6) and disease phenotypes (n=4).

305

Trait Name	**Odds Ratio**	**FDR Adjusted IVW P Value**	**MRE IVW Beta**
Rheumatoid arthritis	NA	1.51E-06	0.08
Unswitched memory B cell %B cell	1.11	5.40E-05	0.10
CD62L- Dendritic Cell %Dendritic Cell	0.95	3.32E-02	-0.05
Effector Memory CD8+ T cell Absolute Count	1.08	1.79E-04	0.07
CD25 on IgD+ CD38- naive B cell	0.93	1.25E-04	-0.07
CD80 on plasmacytoid Dendritic Cell	0.96	1.63E-02	-0.04
Total cholesterol in IDL	0.80	6.48E-08	-0.22
Free cholesterol in IDL	0.80	5.51E-08	-0.22
Total lipids in IDL	0.79	3.35E-09	-0.23
Concentration of IDL particles	0.80	9.83E-09	-0.23
Phospholipids in IDL	0.79	1.86E-09	-0.23
Triglycerides in IDL	0.84	2.40E-07	-0.18
Total cholesterol in large LDL	0.83	9.85E-08	-0.19
Cholesterol esters in large VLDL	0.83	1.41E-07	-0.19
Free cholesterol in large LDL	0.83	7.07E-07	-0.18
Total lipids in large LDL	0.83	1.04E-07	-0.19
Concentration of large LDL particles	0.83	2.40E-07	-0.19
Phospholipids in large LDL	0.83	1.33E-06	-0.18
Cholesterol esters in large VLDL	0.82	6.55E-03	-0.20
18:2, linoleic acid (LA)	0.80	2.93E-07	-0.23
Total cholesterol in LDL	0.82	4.46E-08	-0.19
Total cholesterol in medium LDL	0.82	3.64E-09	-0.20
Cholesterol esters in medium LDL	0.82	5.93E-09	-0.20

Total lipids in medium LDL	0.81	3.35E-09	-0.21
Concentration of medium LDL particles	0.81	3.64E-09	-0.21
Phospholipids in medium LDL	0.81	1.86E-09	-0.21
Total cholesterol in small LDL	0.81	9.85E-08	-0.21
Total lipids in small LDL	0.81	2.65E-07	-0.21
Total cholesterol in small VLDL	0.85	2.54E-02	-0.16
Serum total cholesterol	0.77	3.67E-08	-0.26
Total phosphoglycerides	0.81	3.93E-03	-0.21
Triglycerides in very large HDL	0.87	4.37E-04	-0.14
Total lipids in very small VLDL	0.84	1.97E-03	-0.18
Concentration of very small VLDL particles	0.84	9.04E-04	-0.18
Phospholipids in very small VLDL	0.83	3.63E-03	-0.19
Interferon alpha-10	1.14	1.83E-02	0.13
Protein S100-A5	1.07	6.94E-04	0.07
Serine palmitoyltransferase 2	0.86	3.17E-02	-0.15
CD59 glycoprotein	1.10	1.77E-03	0.09
Complement factor H-related protein 5	1.09	9.30E-05	0.09
Cathepsin F	1.10	4.44E-03	0.10
Benign neoplasm: Skin, unspecified	NA	8.23E-03	0.07
Psychiatric diseases	NA	8.54E-05	-0.43
Mytonic disorders	NA	2.32E-02	0.00

Table 1.

This table contains data from analysis in an early AMD cohort. Only traits passing stringent quality control described in the methods are listed. FDR = false discovery rate. IVW = inverse variance weighted. MRE = multiplicative random effects. % = as a proportion of. NA = not applicable, exposure traits of binary nature do not produce accurate odds ratios and as such beta can be used to infer direction of effect but not necessarily magnitude.

Lipoprotein metabolism is linked to AMD risk

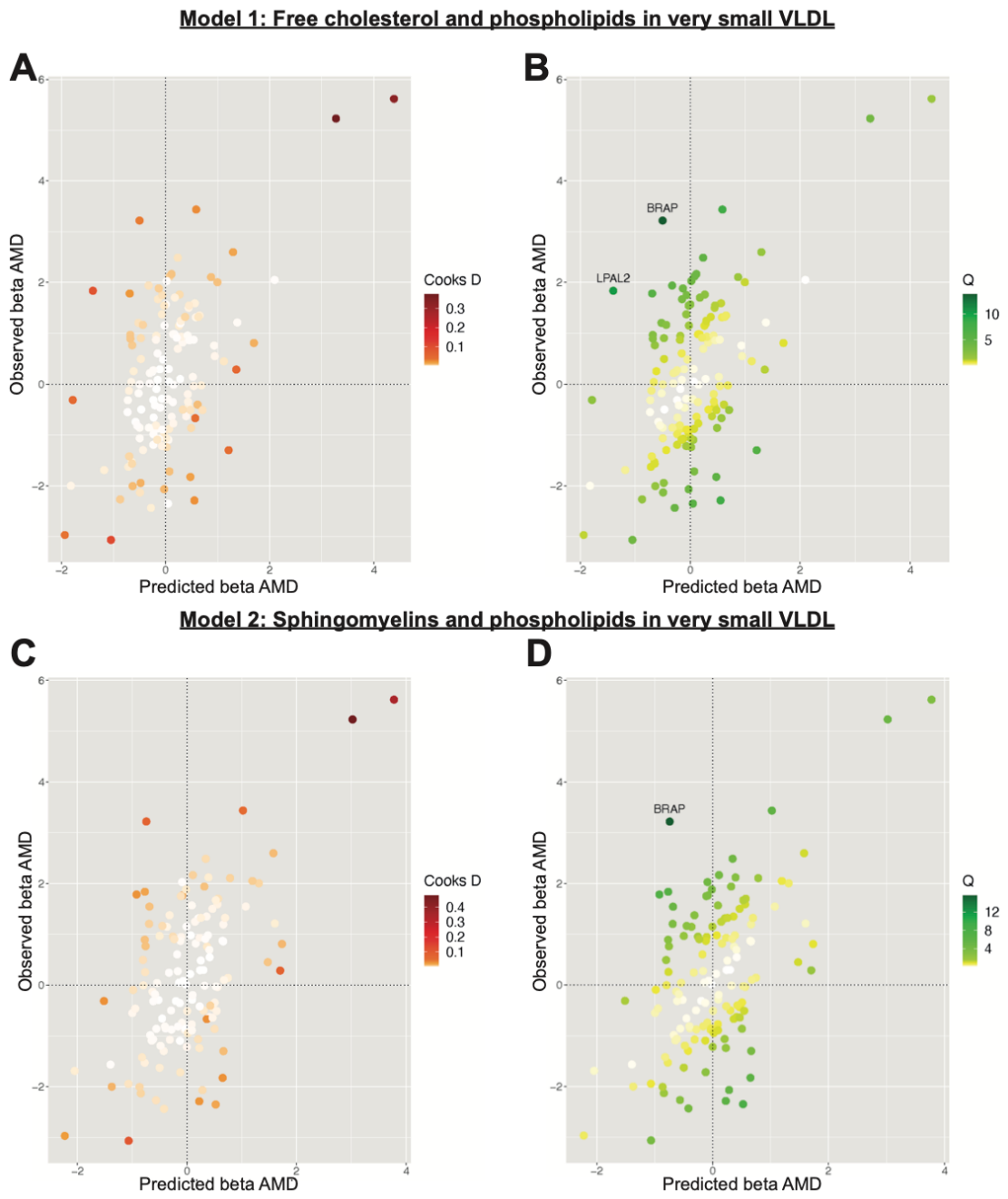
315 Univariable MR demonstrated significant causal relationships for 28 serum lipoprotein measures and 1
316 serum fatty acid concentration (18:2 linoleic acid) in early AMD (**Table 1)**. These relationships were also
317 strongly supported by the results of an advanced AMD analysis (Kettunen et al., 2016) **(Supplementary**
318 **data 3).**
319
320 Serum metabolites are highly correlated traits, and the instruments for serum lipoprotein and fatty acid
321 measures in our study were demonstrably correlated **(Supplementary Figure 1)**. MR-BMA was used in
322 order to discern which traits were driving the causal relationships. In our initial analysis, two genetic variants
323 (rs11065987 [*BRAP*] and rs10455872 [*LPAL2*]) were found to be outliers in terms of Q statistic (Q statistic
324 13.84-15.71 and 6.79-10.46 respectively across models) **(Figures 1A-D)**. These SNPs were therefore
325 omitted and the analysis was re-run. In the subsequent analysis, no outlier SNPs were identified. The top
326 10 models ranked according to their posterior probability are presented in **Supplementary Data 4** and the
327 top 10 individual risk factors with respect to their marginal inclusion probability (MIP; defined as the sum of
328 the posterior probabilities over all the models where the risk factor is present) are shown in **Table 2**. The
329 MIPs of all traits included in our model are plotted in **Figure 2.** The top four traits with respect to their MIP
330 are serum sphingomyelins (MIP=0.76, model averaged causal effects (MACE) =0.29), triglycerides in IDL
331 (MIP=0.32, MACE= -0.16), free cholesterol (MIP=0.20, MACE=0.07) and phospholipids in very small VLDL
332 (MIP=0.63, MACE= -0.31). It is important to note that whilst MR-BMA functions to select the likely causal
333 risk factor amongst a set of candidate causal traits, it is not possible to achieve this with certainty. As such,
334 a definitive statement of causality for individual lipid traits cannot be obtained within the MR causal inference
335 framework.
336

Rank	Risk factor (Trait ID)	Marginal Inclusion	Average Effect	P value	FDR p value
1	Sphingomyelins (met-c-935)	0.76	0.30	2.40E-04	5.76E-03
2	Phospholipids in very small VLDL (met-c-955)	0.63	-0.31	1.00E-05	7.20E-04

3	Triglycerides in IDL (met-c-872)	0.32	-0.16	2.10E-04	5.76E-03
4	Free cholesterol (met-c-858)	0.20	0.07	2.83E-03	5.09E-02
5	Omega-3 fatty acids (met-c-855)	0.07	0.01	3.70E-02	2.34E-01
6	Free cholesterol in very large HDL (met-c-944)	0.07	0.01	2.74E-02	2.34E-01
7	Total lipids in very small VLDL (met-c-953)	0.06	-0.02	2.08E-02	2.34E-01
8	Cholesterol esters in medium VLDL (met-c-910)	0.05	0.01	3.66E-02	2.34E-01
9	Cholesterol esters in very large HDL (met-c-943)	0.05	0.01	4.87E-02	2.34E-01
10	Ratio of bisallylic groups to double bonds (met-c-844)	0.05	0.01	4.41E-02	2.34E-01

Table 2:

The lead causal traits identified by MR BMA of lipids in early AMD ranked according to their marginal inclusion probability (MIP).



342

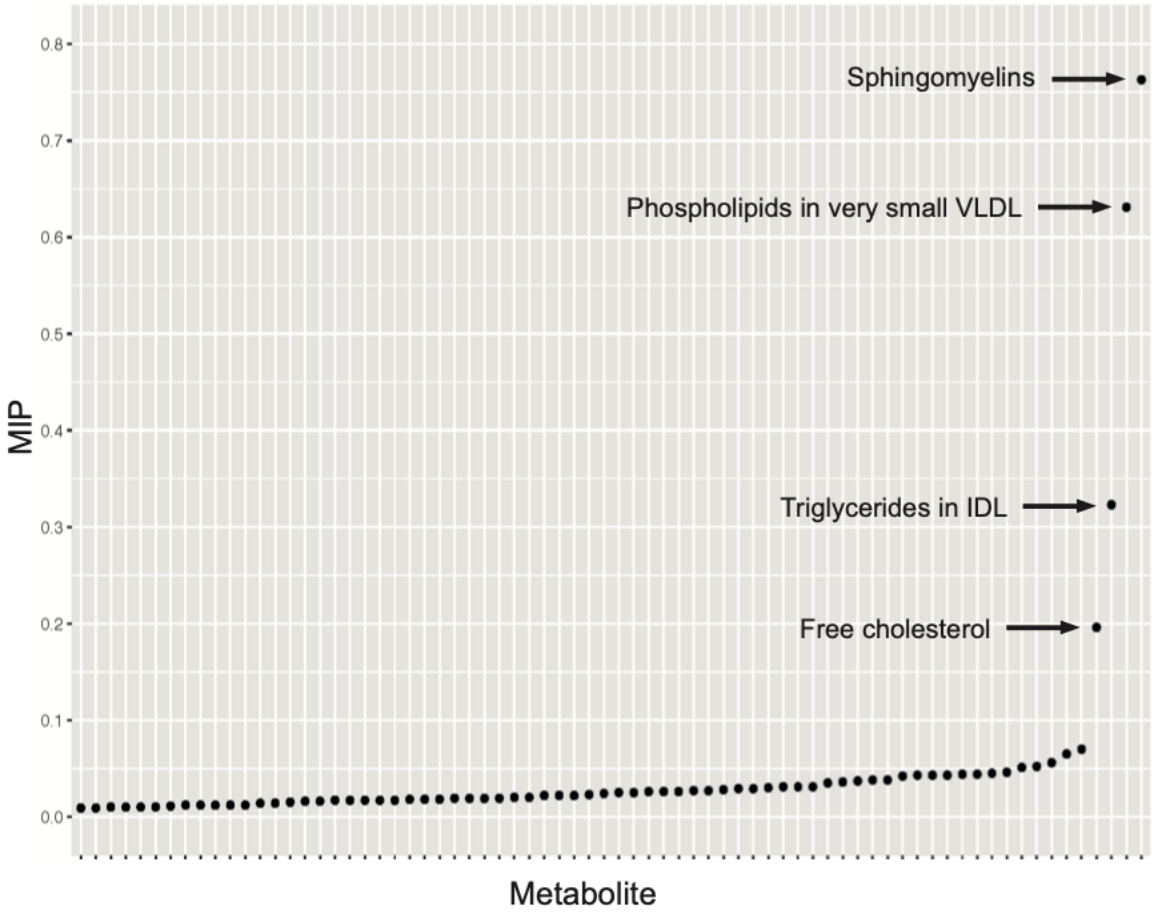
343 **Figure 1A-D:**

344 **Plots generated for the top-ranking models in the first run of MR-BMA of lipid traits and early**

345 **AMD. Plots 2A and 2C are for Cook's distance and 2B and 2D of Cochran's Q. Outlier instruments**

346 **are annotated.**

347



348

349 **Figure 2:**

350 **A graph detailing the results of MR-BMA. Lipids are ranked according to their MIP. The four likely**

351 **true-causal traits are identified.**

352

353 Serine palmitoyltransferase 2, an enzyme which catalyses the first committed step in sphingolipid

354 biosynthesis, was robustly associated with early AMD. The relevant odds ratio highlighted that genetic

355 liability to increasing serum enzyme levels is protective of AMD (**Table 1, Supplementary Data 1**, OR 0.86,

356 p = 1.00E-03). It is of interest that increasing serum levels of the associated enzyme serine

357 palmitoyltransferase 1 also appeared to be protective of early AMD in most measures (**Supplementary**

358 **Data 2**, OR 0.94, IVW p =2.00E-03, FDR IVW p= 0.05, weighted median p = 0.05, significant throughout

359 leave-one-out analysis). Serine palmitoyltransferase enzymes were not significantly related to risk of
360 advanced AMD **(Supplementary Data 3)**.
361
362 **Reinforcing a causal role for complement**
363 Increasing serum levels of the complement proteins CD59 glycoprotein (OR 1.10, IVW p= 2.04E-05) and
364 Complement factor H-related protein 5 (OR 1.10, IVW p = 7.70E-07) were found to be associated with
365 genetic liability to early AMD and passed multiple testing correction. Whilst Complement C4, Complement
366 factor B and Complement factor I were also related to early AMD risk at nominal level of significance, they
367 did not exceed the conservative criteria for causal inference imposed in this study with respect to robust
368 measures **(Supplementary Data 2)**. In subsequent advanced AMD analyses (**Supplementary Data 3**),
369 CD59 glycoprotein remained related to disease risk (OR 1.07, IVW p= 0.04, weighted median p=9.00E-04,
370 MR-Egger p=2.00E-03, weighted mode p =7.00E-03). Complement factor I reached a nominal level of
371 significance in the IVW measure (OR 0.95, p=0.05) but did not remain significant in robust tests. No other
372 complement measures nominally significant for early AMD were significantly related to advanced AMD.
373
374 Complement traits are correlated with one another, and as such MR BMA was attempted. In the initial
375 analysis, the relationship between complement proteins was strongly suggested to be driven by
376 complement factor H (Posterior probability=0.83, MIP=1.00, MIP=-0.68). However, rs2274700 (*CFHR3*,
377 Cook's distance 26.26-55.24 & Cochran's Q 0.87-2.12 across models) and rs10824796 (*MBL2*, Cook's
378 distance 0.13-9.93 & Cochran's Q 0.87-22.12 across models) were identified as outlier instruments and
379 therefore warranted removal from subsequent analysis. rs2274700, a genetic variant known to be
380 associated with AMD risk (Liao et al., 2016), represented the only instrument strongly associated with
381 complement factor H in the MR BMA analysis and as such its removal precluded an informative high-
382 throughput analysis of the complement cascade. Whilst it is clear that there is a causal relationship between
383 complement and AMD, we cannot comment with precision about the specific complement-related molecule
384 driving this.
385
386 **Other immune traits have mixed protective and causal effects in AMD**

387 Our screen identified other serum immune traits with potential causal roles in early AMD, with several of
388 these related to dendritic cell populations. These traits were Unswitched memory B cell as a proportion of
389 B cell (OR=1.11, p=4E-07), CD62L- Dendritic Cell as a proportion of Dendritic Cells (OR=0.95, p=1.22E-
390 03), Effector Memory CD8+ T cell Absolute Count (OR=0.95, p=1.56E-06), CD80 on plasmacytoid
391 Dendritic Cell (OR=0.96, p=4.20E-04) and interferon alpha 10 (OR=1.14, p=5.38E-04). Notably, numerous
392 immune traits were also found to be significantly associated with advanced AMD (**Supplementary Data 3)**.
393 As with lipids, these traits are highly correlated and it is not possible to confidently address which immune
394 cell trait is truly causal in a univariable analysis due to potential overlap in the genetic instruments for each
395 trait. MR BMA using the Orru et al. GWAS for the above immune traits was not possible due to too few
396 genetic variants strongly instrumenting across the exposure variables leading to a failure to construct a
397 meaningful model (Orrù et al., 2020). For this reason, we are unable to determine which of these specific
398 immune traits are driving the casual relationship between immune cell traits and AMD, though dendritic cell
399 traits dominate the univariable analysis.
400
401 **Disease phenotypes are related to AMD risk:**
402 There are four disease phenotypes/groups which were found to be significant at a level which satisfy the
403 criteria for causal inference defined in this study. These traits are rheumatoid arthritis (beta= 0.08, IVW
404 p=9.86E-09), psychiatric diseases (beta=-0.43,IVW p=6.88E-07), 'benign neoplasm: skin, unspecified'
405 (beta= 0.07 , IVW p= 1.76E-04) and myotonic disorders (beta=-0.003,IVW p= 7.49E-04). None of these
406 disorders were significant in advanced AMD. The very broad and non-specific definitions encapsulated in
407 the GWAS of benign skin neoplasms, psychiatric disease and myotonic disorder make it challenging to
408 apply further analyses within the scope of this study, but these may be interesting phenotypes for
409 exploration in future studies.
410
411 In univariable MR, rheumatoid arthritis was identified to be causally related to AMD as detailed above.
412 Multivariable MR with effector memory CD8+ T cell absolute count as a second exposure variable (utilising
413 a Q statistic minimisation procedure) demonstrated that the causal effect of rheumatoid arthritis was either
414 mediated by immune cells or underpinned by correlated pleiotropy (rheumatoid arthritis effect size: 0.03,

95% CI: -0.007-0.055; effector memory CD8+ T cell absolute count effect size 0.07, 95% CI: 0.001-0.19). This finding persists in models assuming both very high (0.9) and low (0.1) phenotypic correlation.

With respect to the relationship between AMD and psychiatric disease, no reverse causation was identified (p=0.31). More specific psychiatric traits which could be analysed were not significantly related to genetic liability to early AMD **(Supplementary Data 2)**. We also explored an extensive range of psychiatric disorders in the advanced AMD dataset and found no significant relationships **(Supplementary Data 3)**. Due to the broad nature of this trait, and negative results for all disorders explored, it is not possible to discern what the driving force behind the significant result for psychiatric disease. It appears probable that psychiatric disease is a false positive result given that it lacks the supporting evidence of biological plausibility, validation in an advanced AMD dataset, or supportive evidence through significance of other similar traits in the early AMD analysis.

Serum cathepsin F and S100 proteins have a causal role in AMD

In addition to the four previously mentioned serum proteins causally associated with early AMD (**Table 1**, Interferon alpha-10, CD59 glycoprotein, complement factor H-related protein and serine palmitoyltransferase 2) a further two serum proteins were identified to have causal effects on AMD. These are S100-A5 (OR 1.07, IVW p = 6.80E-06) and Cathepsin F (OR= 1.10, IVW p = 7.16E-05).

Serum cathepsin F (OR= 1.10, IVW p = 7.16E-05) was shown to be causally related to early AMD risk. However, no other serum cathepsin group proteins produced significant results in early AMD and no cathepsin group proteins were robustly significant in the advanced AMD analysis (although cathepsin S and G have significant IVW values of p=1.71E-05 and p=0.001 respectively, **Supplementary data 3**).

Serum protein S100-A5 was demonstrated to be causal for AMD (OR 1.07, IVW p = 6.80E-06). Additionally, protein S100-A13 was significantly related to early AMD risk in both IVW (OR=1.07, p=1.22E-05) and robust measures (weighted median p=3.31E-05, weighted mode p=9.00E-04 , Egger p= 1.00E-03) but did not remain significant throughout the leave-one-out analysis, suggesting the relationship for this particular

443 protein could be driven by a small number of influential variants. Whilst protein S100-A2 was nominally
444 significant in the IVW (p=0.04), it did not pass any robust measures. No S100 group proteins were
445 significantly related to advanced AMD risk **(Supplementary Data 3)**.
446
447 **DISCUSSION:**
448 In this study we used MR to advance our understanding of AMD pathogenesis. Our findings indicate
449 protective effects for serum VLDL and IDL compositional and concentration traits in AMD. Causal effects
450 for serum free cholesterol and sphingolipid metabolism were also highlighted. Notably, our results suggest
451 that increasing serum levels of sphingomyelins are causally linked to AMD and we report protective
452 relationships for specific enzymes implicated in sphingolipid biosynthesis. It is known that sphingolipids play
453 key roles in retinal physiology, and there is emerging evidence that some sphingolipids may play a role in
454 AMD pathogenesis.(Simon et al., 2021) Further, our findings are in keeping with those of a meta-analysis
455 of metabolomic studies which described pathway enrichment analysis suggesting a role for sphingolipid
456 metabolism in AMD (Hou et al., 2020). Whilst sphingomyelin itself has not previously been described as a
457 causative factor for AMD, other substances involved in sphingolipid metabolism (including ceramide, which
458 can serve as both a substrate for sphingomyelin synthesis or a product of sphingomyelin metabolism) have
459 been suggested as causal factors through regulation of retinal cell death, inflammation and
460 neovascularization (Simon et al., 2021). It is known that the balance of sphingolipids have crucial roles in
461 determining cell fates, and therefore it is interesting to note the contrasting effects of serine
462 palmitoyltransferase and sphingomyelin in this study(Taniguchi and Okazaki, 2020). Our study therefore
463 adds weight to the assertion that sphingolipids play a key role in AMD pathophysiology. Further study of
464 sphingolipid metabolism in AMD subjects is expected to provide important insights and to highlight potential
465 treatment targets.
466
467 Our MR findings support the well-documented assertion that complement plays a key role in AMD
468 pathogenesis.(Clark and Bishop, 2018; Jha et al., 2007; Sivaprasad and Chong, 2006). Whilst MR-BMA
469 was not feasible, univariable analysis supported roles for complement factor H related protein 5 and CD59
470 glycoprotein. A causal signal was obtained for Complement factor H-related protein 5, a finding in keeping

471 with recent studies highlighting the role of this and other factor-H-related proteins (FHR-1 to FHR4; all
472 involved in the regulation of complement factor C3b turnover) in AMD(Cipriani et al., 2021). An intriguing
473 observation was the apparently causal role of increasing serum levels of the complement-related CD59
474 glycoprotein (an inhibitor of membrane attack complex formation) in both early and advanced AMD. This
475 finding appears to be at odds with previous studies that investigated the impact of high levels of CD59 in a
476 model of AMD and found a potential therapeutic benefit (Cashman et al., 2011). The relationship between
477 systemic and local complement activation is poorly understood and studies have produced conflicting
478 results with respect to the roles of each in AMD(Clark and Bishop, 2018; Jha et al., 2007). It can therefore
479 be speculated that the superficially inconsistent results between our study and previous work on CD59
480 reflects the complicated relationship between serum and ocular CD59 expression.
481
482 It is interesting to note that a connection between complement and lipid accumulation has been proposed
483 in AMD. This includes evidence that dysregulation of the complement system results in lipid deposition both
484 systemically and in the retina, and metabolomic evidence that decreased VLDL and increased HDL are
485 associated with increased complement activation independent of AMD status(Armento et al., 2021). Given
486 our results, the relationship between VLDL and complement is an area warranting further study.
487
488 Cathepsins are a very diverse family of lysosomal proteases which play an important part in cellular
489 homeostasis by participating in antigen processing and degrade chemokines and proteases(Yadati et al.,
490 2020). Cathepsins are susceptible to age-related alterations, and have been linked to modulation of pro-
491 inflammatory signalling pathways. Additionally, cathepsins D and S have a direct role to play in the retina
492 in the degradation of photoreceptor outer segments. It has also been shown that homozygosity for variant
493 B cystatin C (an inhibitor of cysteine proteases) causes reduced secretion of mature cystatin C and is
494 associated with increased susceptibility to neovascular AMD(Turk et al., 2012; Zurdel et al., 2002). For
495 these reasons, it has been suggested that dysregulation of cathepsin activity may be a factor in AMD
496 pathophysiology. Here, we show data that suggests serum cathepsin F, a ubiquitously expressed cysteine
497 cathepsin, has a causal relationship with early AMD. The roles of specific cathepsins in AMD is an area
498 which has not yet been mechanistically explored.

499

500 S100 proteins act as intracellular regulators and extracellular signalling proteins. Intracellularly, S100
501 proteins have a wide range of roles including regulating proliferation, differentiation, apoptosis, calcium
502 homeostasis, metabolism and inflammation (Donato et al., 2013). Extracellular S100 proteins have
503 important roles in immunity and inflammation. Of the S100 protein family, serum S100-A5 is the only protein
504 to be robustly identified as a causal factor for AMD in our study. Whilst S100 proteins have received little
505 study in AMD thus far, the binding of S100B to RAGE and the subsequent increase in VEGF have previously
506 been shown to be linked to the development of AMD(Ma et al., 2007; Xia et al., 2017). S100-A5 is a protein
507 which has received very little study and as such its biological functions are mostly uncharacterised, but it is
508 known that it is involved in inflammation via the activation of RAGE(Wheeler and Harms, 2017). Whilst it
509 is outside the scope of this study to assert a mechanistic hypothesis for this finding, it seems likely that the
510 role of S100-A5 would be linked to the inflammatory processes.

511

512 For the most part, the present study has captured the known AMD risk factors. A notable omission however
513 is smoking, an accepted AMD risk factor. In this study, cigarettes smoked per day was associated with early
514 AMD at a nominal level of significance (IVW p=0.009) but was not significant with FDR or in robust
515 measures. This finding is most probably reflective of the underpowered nature of MR. Given the hypothesis
516 free nature of this study and the low power of MR, we have not commented on negative results as evidence
517 of a null relationship between environmental exposures and AMD.

518

519 In summary, this study has identified previously undescribed causal factors for AMD in addition to
520 reinforcing several previously suggested AMD risk factors. Future research will follow-up these causal
521 factors from a mechanistic perspective in order to open up avenues for therapeutic intervention.

522

523 **Ethical approval**

524 This study utilised publicly available data and no additional ethical approval is required.

525

526 **AUTHOR CONTRIBUTIONS:**

THJ, JCK, PIS, GB, HG and ES were involved in study conception and design. THJ and JCK performed the statistical analysis. PIS, GB, GH, TA, SM and ES reviewed the methodology and statistical analysis. All authors aided in interpretation of the results. All authors contributed to manuscript drafting and editing. All authors approved the final manuscript. All authors meet the ICMJE authorship criteria.

ACKNOWLEDGEMENTS:

This project was funded by the NIHR and the Wellcome Institutional Strategic Support Fund for Causal Inference. SM acknowledges Program Grant (1150144) and Senior Research Fellowship funding from the Australian National Health and Medical Research Council (NHMRC).

COMPETING INTERESTS:

TA is involved on advisory boards and has received grants and speaker fees from Allergan, Novartis, Bayer, Roche, Bausch and Lomb, Heidelberg, Topcon and Canon.

The remaining authors have no declarations.

DATA AVAILABILITY:

All data associated with this study are provided within the supplementary data.

546 **REFERENCES:**

547 Armento A, Ueffing M, Clark SJ. 2021. The complement system in age-related macular degeneration. *Cell*
548 *Mol Life Sci* **78**:4487–4505.

549 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to
550 multiple testing. *J R Stat Soc* **57**:289–300.

551 Boddy S, Islam M, Moll T, Kurz J, Burrows D, McGown A, Bhargava A, Julian TH, Harvey C, Marshall
552 JNG, Hall BPC, Allen SP, Kenna KP, Sanderson E, Zhang S, Ramesh T, Snyder MP, Shaw PJ,
553 McDermott C, Cooper-Knock J. 2022. Unbiased metabolome screen leads to personalized medicine
554 strategy for amyotrophic lateral sclerosis. *Brain Commun* **4**. doi:10.1093/braincomms/fcac069

555 Bowden J, Davey Smith G, Burgess S. 2015. Mendelian randomization with invalid instruments: effect
556 estimation and bias detection through Egger regression. *Int J Epidemiol* **44**:512–525.

557 Bowden J, Davey Smith G, Haycock PC, Burgess S. 2016a. Consistent Estimation in Mendelian
558 Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*
559 **40**:304–314.

560 Bowden J, Del Greco M. F, Minelli C, Smith GD, Sheehan NA, Thompson JR. 2016b. Assessing the
561 suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger
562 regression: the role of the I2 statistic. *International Journal of Epidemiology*. doi:10.1093/ije/dyw220

563 Bowden J, Holmes MV. 2019. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods*
564 **10**:486–496.

565 Bowden J, Spiller W, Del Greco M F, Sheehan N, Thompson J, Minelli C, Davey Smith G. 2018.
566 Improving the visualization, interpretation and analysis of two-sample summary data Mendelian
567 randomization via the Radial plot and Radial regression. *Int J Epidemiol* **47**:1264–1278.

568 Burgess S, Labrecque JA. 2018. Mendelian randomization with a binary exposure variable: interpretation
569 and presentation of causal estimates. *European Journal of Epidemiology*. doi:10.1007/s10654-018-
570 0424-6

571 Burgess S, Smith GD, Davies NM, Dudbridge F, Gill D, Maria Glymour M, Hartwig FP, Holmes MV,
572 Minelli C, Relton CL, Theodoratou E. 2019. Guidelines for performing Mendelian randomization
573 investigations. *Wellcome Open Research* **4**. doi:10.12688/wellcomeopenres.15555.2

574 Burgess S, Thompson SG. 2017a. Interpreting findings from Mendelian randomization using the MR-
575 Egger method. *Eur J Epidemiol* **32**:377–389.
576 Burgess S, Thompson SG. 2017b. Erratum to: Interpreting findings from Mendelian randomization using
577 the MR-Egger method. *Eur J Epidemiol* **32**:391–392.
578 Burgess S, Thompson SG, CRP CHD Genetics Collaboration. 2011. Avoiding bias from weak instruments
579 in Mendelian randomization studies. *Int J Epidemiol* **40**:755–764.
580 Cashman SM, Ramo K, Kumar-Singh R. 2011. A Non Membrane-Targeted Human Soluble CD59
581 Attenuates Choroidal Neovascularization in a Model of Age Related Macular Degeneration. *PLoS*
582 *ONE*. doi:10.1371/journal.pone.0019078
583 Chakravarthy U, Wong TY, Fletcher A, Piault E, Evans C, Zlateva G, Buggage R, Pleil A, Mitchell P.
584 2010. Clinical risk factors for age-related macular degeneration: a systematic review and meta-
585 analysis. *BMC Ophthalmology*. doi:10.1186/1471-2415-10-31
586 Cipriani V, Tierney A, Griffiths JR, Zuber V, Sergouniotis PI, Yates JRW, Moore AT, Bishop PN, Clark SJ,
587 Unwin RD. 2021. Beyond factor H: The impact of genetic-risk variants for age-related macular
588 degeneration on circulating factor-H-like 1 and factor-H-related protein concentrations. *Am J Hum*
589 *Genet* **108**:1385–1400.
590 Clark SJ, Bishop PN. 2018. The eye as a complement dysregulation hotspot. *Semin Immunopathol*
591 **40**:65–74.
592 Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov
593 AG, Barnes I, Bennett R, Berry A, Bhai J, Bignell A, Billis K, Boddu S, Brooks L, Charkhchi M,
594 Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, Donaldson S, El Houdaigui B, El Naboulsi T,
595 Fatima R, Giron CG, Genez T, Martinez JG, Guijarro-Clarke C, Gymer A, Hardy M, Hollis Z, Hourlier
596 T, Hunt T, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Marugán JC, Mohanan S,
597 Mushtaq A, Naven M, Ogeh DN, Parker A, Parton A, Perry M, Piližota I, Prosovetskaia I, Sakthivel
598 MP, Salam AIA, Schmitt BM, Schuilenburg H, Sheppard D, Pérez-Silva JG, Stark W, Steed E,
599 Sutinen K, Sukumaran R, Sumathipala D, Suner M-M, Szpak M, Thormann A, Tricomi FF, Urbina-
600 Gómez D, Veidenberg A, Walsh TA, Walts B, Willhoft N, Winterbottom A, Wass E, Chakiachvili M,
601 Flint B, Frankish A, Giorgetti S, Haggerty L, Hunt SE, IIsley GR, Loveland JE, Martin FJ, Moore B,

602 Mudge JM, Muffato M, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Dyer S, Harrison PW,
603 Howe KL, Yates AD, Zerbino DR, Flicek P. 2022. Ensembl 2022. *Nucleic Acids Res* **50**:D988–D995.
604 Donato R, Cannon BR, Sorci G, Riuzzi F, Hsu K, Weber DJ, Geczy CL. 2013. Functions of S100 proteins.
605 *Curr Mol Med* **13**:24–57.
606 Fleckenstein M, Keenan TDL, Guymer RH, Chakravarthy U, Schmitz-Valckenberg S, Klaver CC, Wong
607 WT, Chew EY. 2021. Age-related macular degeneration. *Nat Rev Dis Primers* **7**:31.
608 Fritsche LG, Chen W, Schu M, Yaspan BL, Yu Y, Thorleifsson G, Zack DJ, Arakawa S, Cipriani V, Ripke
609 S, Igo RP Jr, Buitendijk GHS, Sim X, Weeks DE, Guymer RH, Merriam JE, Francis PJ, Hannum G,
610 Agarwal A, Armbrecht AM, Audo I, Aung T, Barile GR, Benchaboune M, Bird AC, Bishop PN,
611 Branham KE, Brooks M, Brucker AJ, Cade WH, Cain MS, Campochiaro PA, Chan C-C, Cheng C-Y,
612 Chew EY, Chin KA, Chowers I, Clayton DG, Cojocaru R, Conley YP, Cornes BK, Daly MJ, Dhillon B,
613 Edwards AO, Evangelou E, Fagerness J, Ferreyra HA, Friedman JS, Geirsdottir A, George RJ,
614 Gieger C, Gupta N, Hagstrom SA, Harding SP, Haritoglou C, Heckenlively JR, Holz FG, Hughes G,
615 Ioannidis JPA, Ishibashi T, Joseph P, Jun G, Kamatani Y, Katsanis N, N Keilhauer C, Khan JC, Kim
616 IK, Kiyohara Y, Klein BEK, Klein R, Kovach JL, Kozak I, Lee CJ, Lee KE, Lichtner P, Lotery AJ,
617 Meitinger T, Mitchell P, Mohand-Saïd S, Moore AT, Morgan DJ, Morrison MA, Myers CE, Naj AC,
618 Nakamura Y, Okada Y, Orlin A, Ortube MC, Othman MI, Pappas C, Park KH, Pauer GJT, Peachey
619 NS, Poch O, Priya RR, Reynolds R, Richardson AJ, Ripp R, Rudolph G, Ryu E, Sahel J-A,
620 Schaumberg DA, Scholl HPN, Schwartz SG, Scott WK, Shahid H, Sigurdsson H, Silvestri G,
621 Sivakumaran TA, Smith RT, Sobrin L, Souied EH, Stambolian DE, Stefansson H, Sturgill-Short GM,
622 Takahashi A, Tosakulwong N, Truitt BJ, Tsironi EE, Uitterlinden AG, van Duijn CM, Vijaya L,
623 Vingerling JR, Vithana EN, Webster AR, Wichmann H-E, Winkler TW, Wong TY, Wright AF, Zelenika
624 D, Zhang M, Zhao L, Zhang K, Klein ML, Hageman GS, Lathrop GM, Stefansson K, Allikmets R,
625 Baird PN, Gorin MB, Wang JJ, Klaver CCW, Seddon JM, Pericak-Vance MA, Iyengar SK, Yates
626 JRW, Swaroop A, Weber BHF, Kubo M, Deangelis MM, Léveillard T, Thorsteinsdottir U, Haines JL,
627 Farrer LA, Heid IM, Abecasis GR, AMD Gene Consortium. 2013. Seven new loci associated with
628 age-related macular degeneration. *Nat Genet* **45**:433–9, 439e1–2.
629 Fritsche LG, Igl W, Bailey JNC, Grassmann F, Sengupta S, Bragg-Gresham JL, Burdon KP, Hebbring SJ,

630 Wen C, Gorski M, Kim IK, Cho D, Zack D, Souied E, Scholl HPN, Bala E, Lee KE, Hunter DJ, Sardell
631 RJ, Mitchell P, Merriam JE, Cipriani V, Hoffman JD, Schick T, Lechanteur YTE, Guymer RH,
632 Johnson MP, Jiang Y, Stanton CM, Buitendijk GHS, Zhan X, Kwong AM, Boleda A, Brooks M, Gieser
633 L, Ratnapriya R, Branham KE, Foerster JR, Heckenlively JR, Othman MI, Vote BJ, Liang HH,
634 Souzeau E, McAllister IL, Isaacs T, Hall J, Lake S, Mackey DA, Constable IJ, Craig JE, Kitchner TE,
635 Yang Z, Su Z, Luo H, Chen D, Ouyang H, Flagg K, Lin D, Mao G, Ferreyra H, Stark K, von Strachwitz
636 CN, Wolf A, Brandl C, Rudolph G, Olden M, Morrison MA, Morgan DJ, Schu M, Ahn J, Silvestri G,
637 Tsironi EE, Park KH, Farrer LA, Orlin A, Brucker A, Li M, Curcio CA, Mohand-Saïd S, Sahel J-A,
638 Audo I, Benchaboune M, Cree AJ, Rennie CA, Goverdhan SV, Grunin M, Hagbi-Levi S,
639 Campochiaro P, Katsanis N, Holz FG, Blond F, Blanché H, Deleuze J-F, Igo RP Jr, Truitt B, Peachey
640 NS, Meuer SM, Myers CE, Moore EL, Klein R, Hauser MA, Postel EA, Courtenay MD, Schwartz SG,
641 Kovach JL, Scott WK, Liew G, Tan AG, Gopinath B, Merriam JC, Smith RT, Khan JC, Shahid H,
642 Moore AT, McGrath JA, Laux R, Brantley MA Jr, Agarwal A, Ersoy L, Caramoy A, Langmann T,
643 Saksens NTM, de Jong EK, Hoyng CB, Cain MS, Richardson AJ, Martin TM, Blangero J, Weeks DE,
644 Dhillon B, van Duijn CM, Doheny KF, Romm J, Klaver CCW, Hayward C, Gorin MB, Klein ML, Baird
645 PN, den Hollander AI, Fauser S, Yates JRW, Allikmets R, Wang JJ, Schaumberg DA, Klein BEK,
646 Hagstrom SA, Chowers I, Lotery AJ, Léveillard T, Zhang K, Brilliant MH, Hewitt AW, Swaroop A,
647 Chew EY, Pericak-Vance MA, DeAngelis M, Stambolian D, Haines JL, Iyengar SK, Weber BHF,
648 Abecasis GR, Heid IM. 2016. A large genome-wide association study of age-related macular
649 degeneration highlights contributions of rare and common variants. *Nat Genet* **48**:134–143.
650 Han X, Gharahkhani P, Mitchell P, Liew G, Hewitt AW, MacGregor S. 2020a. Genome-wide meta-
651 analysis identifies novel loci associated with age-related macular degeneration. *J Hum Genet*
652 **65**:657–665.
653 Han X, Ong J-S, An J, Hewitt AW, Gharahkhani P, MacGregor S. 2020b. Using Mendelian randomization
654 to evaluate the causal relationship between serum C-reactive protein levels and age-related macular
655 degeneration. *Eur J Epidemiol* **35**:139–146.
656 Han X, Ong J-S, Hewitt AW, Gharahkhani P, MacGregor S. 2021. The effects of eight serum lipid
657 biomarkers on age-related macular degeneration risk: a Mendelian randomization study. *Int J*

658 *Epidemiol* **50**:325–336.
659 Hartwig FP, Davey Smith G, Bowden J. 2017. Robust inference in summary data Mendelian
660 randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* **46**:1985–1998.
661 Hartwig FP, Davies NM, Hemani G, Davey Smith G. 2016. Two-sample Mendelian randomization:
662 avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J*
663 *Epidemiol* **45**:1717–1726.
664 Hemani G, Tilling K, Davey Smith G. 2017. Orienting the causal relationship between imprecisely
665 measured traits using GWAS summary data. *PLoS Genet* **13**:e1007081.
666 Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J,
667 Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM,
668 Davey Smith G, Gaunt TR, Haycock PC. 2018. The MR-Base platform supports systematic causal
669 inference across the human phenome. *Elife* **7**. doi:10.7554/eLife.34408
670 Hou X-W, Wang Y, Pan C-W. 2020. Metabolomics in Age-Related Macular Degeneration: A Systematic
671 Review. *Invest Ophthalmol Vis Sci* **61**:13.
672 Jha P, Bora PS, Bora NS. 2007. The role of complement system in ocular diseases including uveitis and
673 macular degeneration. *Mol Immunol* **44**:3901–3908.
674 Julian TH, Boddy S, Islam M, Kurz J, Whittaker KJ, Moll T, Harvey C, Zhang S, Snyder MP, McDermott C,
675 Cooper-Knock J, Shaw PJ. 2021. A review of Mendelian randomization in amyotrophic lateral
676 sclerosis. *Brain*. doi:10.1093/brain/awab420
677 Kettunen J, Demirkan A, Würtz P, Draisma HHM, Haller T, Rawal R, Vaarhorst A, Kangas AJ, Lyytikäinen
678 L-P, Pirinen M, Pool R, Sarin A-P, Soininen P, Tukiainen T, Wang Q, Tiainen M, Tynkkynen T, Amin
679 N, Zeller T, Beekman M, Deelen J, van Dijk KW, Esko T, Hottenga J-J, van Leeuwen EM, Lehtimäki
680 T, Mihailov E, Rose RJ, de Craen AJM, Gieger C, Kähönen M, Perola M, Blankenberg S, Savolainen
681 MJ, Verhoeven A, Viikari J, Willemsen G, Boomsma DI, van Duijn CM, Eriksson J, Jula A, Järvelin M-
682 R, Kaprio J, Metspalu A, Raitakari O, Salomaa V, Slagboom PE, Waldenberger M, Ripatti S, Ala-
683 Korpela M. 2016. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel
684 systemic effects of LPA. *Nat Commun* **7**:11122.
685 Kvale MN, Hesselson S, Hoffmann TJ, Cao Y, Chan D, Connell S, Croen LA, Dispensa BP, Eshragh J,

686 Finn A, Gollub J, Iribarren C, Jorgenson E, Kushi LH, Lao R, Lu Y, Ludwig D, Mathauda GK, McGuire
687 WB, Mei G, Miles S, Mittman M, Patil M, Quesenberry CP Jr, Ranatunga D, Rowell S, Sadler M,
688 Sakoda LC, Shapero M, Shen L, Shenoy T, Smethurst D, Somkin CP, Van Den Eeden SK, Walter L,
689 Wan E, Webster T, Whitmer RA, Wong S, Zau C, Zhan Y, Schaefer C, Kwok P-Y, Risch N. 2015.
690 Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology
691 Research on Adult Health and Aging (GERA) Cohort. *Genetics* **200**:1051–1060.
692 Liao X, Lan C-J, Cheuk I-W-Y, Tan Q-Q. 2016. Four complement factor H gene polymorphisms in
693 association with AMD: A meta-analysis. *Arch Gerontol Geriatr* **64**:123–129.
694 Li X, Li H, Cheng J, Wang M, Zhong Y, Shi G, Yu A-Y. 2022. Causal Associations of Thyroid Function
695 and Age-Related Macular Degeneration: A Two-Sample Mendelian Randomization Study. *Am J*
696 *Ophthalmol* **239**:108–114.
697 Ma W, Lee SE, Guo J, Qu W, Hudson BI, Schmidt AM, Barile GR. 2007. RAGE ligand upregulation of
698 VEGF secretion in ARPE-19 cells. *Invest Ophthalmol Vis Sci* **48**:1355–1361.
699 Orrù V, Steri M, Sidore C, Marongiu M, Serra V, Olla S, Sole G, Lai S, Dei M, Mulas A, Virdis F, Piras
700 MG, Lobina M, Marongiu M, Pitzalis M, Deidda F, Loizedda A, Onano S, Zoledziewska M, Sawcer S,
701 Devoto M, Gorospe M, Abecasis GR, Floris M, Pala M, Schlessinger D, Fiorillo E, Cucca F. 2020.
702 Complex genetic signatures in immune cells underlie autoimmunity and inform therapy. *Nat Genet*
703 **52**:1036–1045.
704 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW,
705 Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based
706 linkage analyses. *Am J Hum Genet* **81**:559–575.
707 Sanderson E, Davey Smith G, Windmeijer F, Bowden J. 2019. An examination of multivariable Mendelian
708 randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol* **48**:713–
709 727.
710 Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, Palmer T, Schooling CM,
711 Wallace C, Zhao Q, Davey Smith G. 2022. Mendelian randomization. *Nat Rev Methods Primers* **2**.
712 doi:10.1038/s43586-021-00092-5
713 Sanderson E, Spiller W, Bowden J. 2021. Testing and correcting for weak and pleiotropic instruments in

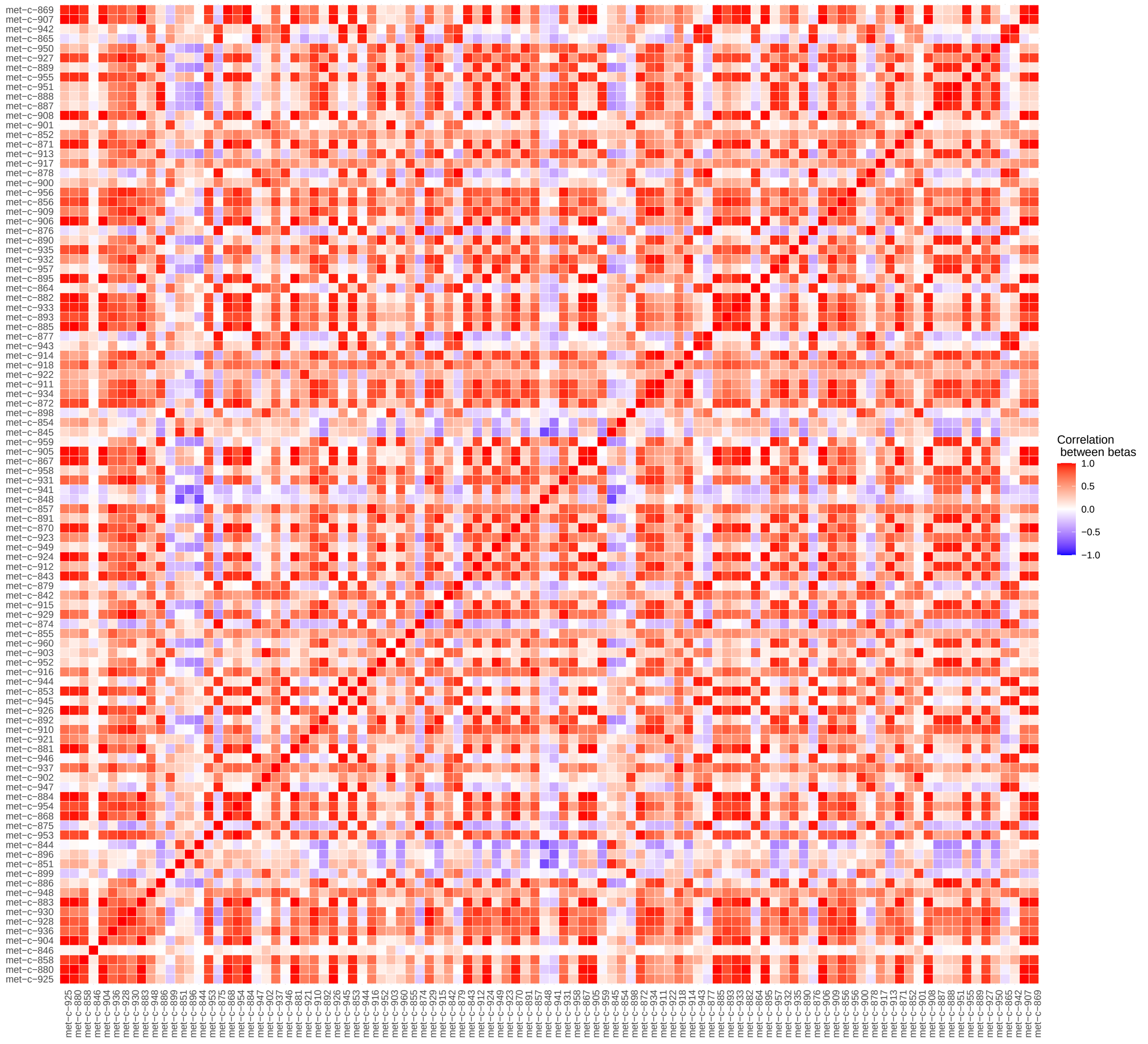
714 two-sample multivariable Mendelian randomization. *Stat Med* **40**:5434–5452.
715 Simon MV, Basu SK, Qaladize B, Grambergs R, Rotstein NP, Mandal N. 2021. Sphingolipids as critical
716 players in retinal physiology and pathology. *J Lipid Res* **62**:100037.
717 Sivaprasad S, Chong NV. 2006. The complement system and age-related macular degeneration. *Eye*
718 **20**:867–872.
719 Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M,
720 Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. 2015. UK
721 biobank: an open access resource for identifying the causes of a wide range of complex diseases of
722 middle and old age. *PLoS Med* **12**:e1001779.
723 Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, Sarwath H, Thareja G, Wahl A, DeLisle
724 RK, Gold L, Pezer M, Lauc G, El-Din Selim MA, Mook-Kanamori DO, Al-Dous EK, Mohamoud YA,
725 Malek J, Strauch K, Grallert H, Peters A, Kastenmüller G, Gieger C, Graumann J. 2017. Connecting
726 genetic risk to disease end points through the human blood plasma proteome. *Nat Commun*
727 **8**:14357.
728 Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E,
729 Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N,
730 Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouwehand WH, Todd JA,
731 Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. 2018.
732 Genomic atlas of the human plasma proteome. *Nature* **558**:73–79.
733 Taniguchi M, Okazaki T. 2020. Ceramide/sphingomyelin rheostat regulated by sphingomyelin synthases
734 and chronic diseases in Murine models. *J Lipid Atheroscler* **9**:380–405.
735 Turk V, Stoka V, Vasiljeva O, Renko M, Sun T, Turk B, Turk D. 2012. Cysteine cathepsins: from structure,
736 function and regulation to new frontiers. *Biochim Biophys Acta* **1824**:68–88.
737 Wheeler LC, Harms MJ. 2017. Human S100A5 binds Ca2+ and Cu2+ independently. *BMC Biophys* **10**:8.
738 Winkler TW, Grassmann F, Brandl C, Kiel C, Günther F, Strunz T, Weidner L, Zimmermann ME, Korb CA,
739 Poplawski A, Schuster AK, Müller-Nurasyid M, Peters A, Rauscher FG, Elze T, Horn K, Scholz M,
740 Cañadas-Garre M, McKnight AJ, Quinn N, Hogg RE, Küchenhoff H, Heid IM, Stark KJ, Weber BHF.
741 2020. Genome-wide association meta-analysis for early age-related macular degeneration highlights

742 novel loci and insights for advanced disease. *BMC Med Genomics* **13**:120.
743 Xia C, Braunstein Z, Toomey AC, Zhong J, Rao X. 2017. S100 Proteins As an Important Regulator of
744 Macrophage Inflammation. *Front Immunol* **8**:1908.
745 Yadati T, Houben T, Bitorina A, Shiri-Sverdlov R. 2020. The Ins and Outs of Cathepsins: Physiological
746 Function and Role in Disease Management. *Cells* **9**. doi:10.3390/cells9071679
747 Zuber V, Colijn JM, Klaver C, Burgess S. 2020. Selecting likely causal risk factors from high-throughput
748 experiments using multivariable Mendelian randomization. *Nat Commun* **11**:29.
749 Zurdel J, Finckh U, Menzer G, Nitsch RM, Richard G. 2002. CST3 genotype associated with exudative
750 age related macular degeneration. *Br J Ophthalmol* **86**:214–219.

751

752

753


Correlation between betas
1.0
0.5
0.0
−0.5
−1.0